Investigating diet-microbe-host interaction in preterm infants at risk of necrotising enterocolitis

Andrea Chiara Masi

Thesis submitted for the degree of Doctor of Philosophy

Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University

Submitted March 2023

Abstract

Necrotising enterocolitis (NEC) is a devastating intestinal disease primarily affecting preterm infants born <32 weeks gestation. The underlying mechanisms are poorly understood, but mothers own breast milk (MOM) and infant gut microbiome play a role. In MOM, human milk oligosaccharides (HMOs) influence the infant health in multiple ways, including shaping infant gut colonisation by potentially beneficial bacteria (mainly bifidobacteria). In the infant gut, a lack of colonisation with bifidobacteria, coupled with a community rich in Pseudomonadota, have been associated with NEC development.

In this thesis, the interaction between HMO profiles and infant gut microbiome in NEC was investigated. To further understand the potential mechanism of HMOs in promoting infant health, basic microbiology and an experimental preterm intestinal-derived organoid model were employed. It was hypothesised that the HMO and bacterial profiles in MOM given to NEC infants would differ from healthy infants, and that these differences would be associated with the infant gut microbiome.

The results confirmed and expanded upon previous work, finding that the concentration of a single HMO, disialyllacto-N-tetraose (DSLNT), was significantly lower in MOM received by NEC infants compared to controls. No difference in the microbiome of MOM was observed between mothers whose infants developed NEC or not, while infants receiving low MOM DSLNT were associated with reduced transition into preterm gut community types dominated by *Bifidobacterium* spp. Numerous bifidobacteria isolated from preterm infants showed growth *in vitro* on selected HMOs, however, only a *B. bifidum* isolate could metabolise DSLNT. Finally, the interaction between selected HMOs and preterm intestinal-derived organoids from NEC and non-NEC tissue was investigated, finding differences in specific genes involved in cell differentiation and proliferation.

This work provides new insights into the interaction between MOM, infant gut microbiome and the intestine in preterm infants with association to NEC development.

Impact of COVID-19

As the COVID-19 pandemic started during the first year of my PhD project, I have stated the impact the pandemic has had on my research.

Due to the pandemic, access to the laboratory was limited for over 6 months, and when the access to the laboratory was allowed again, a limited number of people could work in each lab space, reducing the amount of time everyone could use to do wet lab work. This PhD project was heavily laboratory based, and due the reduced accessibility to the lab facilities, the aims of the study had to be adjusted. As part of my project, I investigated the interaction between human milk oligosaccharides (HMOs) and preterm intestinal organoids. Performance of such experiments was postponed due to order delays for most of the lab consumables required, including Transwell inserts which could not be delivered for over 6 months. Moreover, the effects of HMOs were investigated in combination with pro-inflammatory stimuli, to study the anti-inflammatory properties of these sugars. Due to the mentioned delays, the pro-inflammatory stimuli could not be optimised, and thus we decided to mimic published studies. One of the major aims of this project was to investigate the role of HMOs in necrotising enterocolitis onset. The samples were analysed by our collaborator Prof Lars Bode (University of California San Diego) prior to the pandemic. Such work led to one of my first-author papers published in *Gut* and presented in chapter 3. Additional breast milk samples were sent to Prof Bode to expand on the published work and to address some of the limitations of the study. Due to a covid-related backlog experienced by Prof Bode, the additional samples could not be analysed and included in this thesis.

During the period in which I could not access the laboratory, I was able to analyse readily available RNA-sequencing data from organoids experiments previously performed by Dr Stewart. Such opportunity allowed me to learn how to analyse whole transcriptome data, which I subsequently applied to data acquired from other organoid experiments. Moreover, such analysis was later published as a Letter in *Gut.* Despite this clear impact of COVID-19, I am confident I have achieved the main aims I had set for my project.

Acknowledgements

I would like to start by thanking my great supervisors, Dr. Chris Stewart, Dr. Chris Lamb and Prof. John Kirby. I will always be grateful for their support and guidance through this journey, which has made me grow as a scientist but also as a person. A particular thanks goes to my primary supervisor Chris S, it has been a pleasure being your first PhD student, and I really appreciate your support inside and outside of the lab.

I need to acknowledge the important role played by Dr. Janet Berrington and Prof. Nick Embleton, neonatal consultants who make our work in collaboration with the Newcastle upon Tyne NICU possible. Thanks to them and to Dr. Claire Grainger for instructing me with the details of working directly with these fragile babies day to day, putting in perspective the importance of the research we have done and will continue to carry out together.

A big thanks goes to the extended Stewart and Lamb group, who has given me a fantastic support through the past three and a half years, in particular during the COVID-19 pandemic which kept me away from my family for months. And thanks to Lauren Beck for her continuous guidance through the world of bioinformatics. I am also really grateful to Dr Greg Young for his help with samples hunting in the numerous freezers.

Thanks to all our collaborators who made this project possible and successful. I need to thank Prof. Lars Bode and his team for the collaboration on human milk oligosaccharides profiling in our fragile preterm cohort. I would like also to thank him for having hosted during my visit at the University of California San Diego.

I need to acknowledge the support provided by various groups at Baylor College of Medicine, Houston. They have helped me and the wider group in many projects with their bioinformatic skills and their vast knowledge of intestinal organoids. A special thanks to Dr. Sashi Ramani for hosting me in her lab and for organising a very fruitful time for me to spend at BCM, and to all the lab groups who made me feel welcome and shared their knowledge with me.

I would like to acknowledge Glycom/DSM for providing human milk oligosaccharides for me to test on organoids and for our microbiology work. Thanks also to all the families who have trusted us and donated the biological material we use for our research, which would not be possible otherwise.

I want also to thank Katie, Jenny and Chong with whom I have shared this journey with since the start. And to all the old and new friends who have supported me in these past years.

I want to give a special acknowledgement to my parents. Grazie per avermi sempre supportata, amata e per aver sempre creduto in me negli anni. Non sarei chi sono oggi senza di voi.

At last, I would like to thank my partner Marcus, who has encouraged me through the last few months of this journey. Thanks to him for believing in me and supporting me especially through the hard times.

Publications and presentations

Key publications

The following publications have formed the foundations for part of the introduction and of chapter results included in this thesis.

* Co-first author

- **Masi AC**, Stewart CJ. (2022). Untangling human milk oligosaccharides and infant gut microbiome. iScience. 10.1016/j.isci.2021.103542
- **Masi AC***, Fofanova TY*, Lamb CA, Auchtung JM, Britton RA, Estes MK, Ramani S, Cockell SJ, Coxhead J, Embleton ND, Berrington JE, Petrosino JF, Stewart CJ. (2022). Distinct gene expression profiles between human preterm-derived and adult-derived intestinal organoids exposed to Enterococcus faecalis: a pilot study. Gut 71, 2141-2143. 10.1136/gutjnl-2021-326552
- **Masi AC**, Embleton ND, Lamb CA, Young G, Granger CL, Najera J, Smith DP, Hoffman KL, Petrosino JF, Bode L, Berrington JE, Stewart CJ. (2020). Human milk oligosaccharide DSLNT and gut microbiome in preterm infants predicts necrotising enterocolitis. Gut 70, 999. 10.1136/gutjnl-2020-322457
- **Masi AC**, Stewart CJ. (2019). The role of the preterm intestinal microbiome in sepsis and necrotising enterocolitis. Early Hum Dev 138, 104854. 10.1016/j.earlhumdev.2019.104854

Additional publications

The following list reports publications I have contributed to during my PhD.

- Beck LC, **Masi AC**, Young G, Vatanen T, Lamb CA, Smith R, Coxhead J, Butler A, Marsland B, Embleton N, Berrington J, Stewart CJ. (2022). Strain-specific impacts of probiotics are a significant driver of gut microbiome development in very preterm infants. Nat Microbiol 7, 1525-1535. 10.1038/s41564-022-01213-w
- Beck LC*, Granger CL*, **Masi AC***, Stewart CJ. (2021). Use of omic technologies in early life gastrointestinal health and disease: from bench to bedside Expert Review of Proteomics 18, 247-259. 10.1080/14789450.2021.1922278
- Granger CL, Lamb CA, Embleton ND, Beck LC, **Masi AC**, Palmer JM, Stewart CJ, Berrington JE. (2021). Secretory immunoglobulin A in preterm infants: determination of normal values in breast milk and stool. Pediatr Res. 10.1038/s41390-021-01930-8
- **Masi AC***, Oppong YEA*, Haugk B, Lamb CA, Sharp L, Shaw JM, Stewart CJ, Oppong KW. (2021). Endoscopic ultrasound (EUS)-guided fine needle biopsy

(FNB) formalin fixed paraffin-embedded (FFPE) pancreatic tissue samples are a potential resource for microbiota analysis. Gut 70, 999. 10.1136/gutjnl-2020-322457

- **Masi AC***, Koo S*, Lamb CA, Hull MK, Sharp L, Nelson A, Hampton JS, Rees CJ, Stewart CJ. (2021). Using faecal immunochemical test (FIT) undertaken in a national screening programme for large-scale gut microbiota analysis. Gut 70, 429. 10.1136/gutjnl-2020-321594

List of oral presentations

- Human milk oligosaccharide DSLNT and gut microbiome, but not breast milk microbiome, in preterm infants predicts necrotising enterocolitis, 2022 IMGC Hybrid Symposium, UC Davis, California
- Human milk oligosaccharides (HMOs) and infant gut microbiome, 2022 ISRHML Conference, Panama, Panama
- Human milk oligosaccharide DSLNT and gut microbiome in preterm infants predicts necrotising enterocolitis, 2021 NEC Society Virtual Sessions
- Human milk oligosaccharide DSLNT and gut microbiome in preterm infants predicts necrotising enterocolitis, 2021 IHMC 20
- Human milk oligosaccharide DSLNT and gut microbiome in preterm infants predicts necrotising enterocolitis, 2021 The Origins and Benefits of Biologically Active Components in Human Milk Conference
- Distinct gene expression profiles between preterm- and adult-derived intestinal organoids exposed to Enterococcus faecalis: A pilot study, 2021 North East PostGraduate Conference. Newcastle Upon Tyne, UK
- Human milk oligosaccharide DSLNT and gut microbiome in preterm infants predicts necrotising enterocolitis, 2020 North East PostGraduate Conference. Newcastle Upon Tyne, UK

Table of contents

Table of figures

Table of tables

Abbreviations

2’FL – 2’-fucosyllactose

3’SL – 3’-sialyllactose

3FL – 3-fucosyllactose

6’SL – 6’-sialyllactose

ABC – ATP-binding cassette

ADDD – added drop by drop till dissolved

ATIMA – Agile Toolkit for Incisive Microbial Analysis

AUC – area under the curve

BHI – brain heart infusion

BLASTn – nucleotide basic local alignment search tool

BMI – body mass index

BP – biological process

BSM – bifidus selective medium

CCS – complete chelating solution

cDNA – complementary DNA

CFU – colony-forming unit

CI – confidence interval

CM – conditioned media

CMGF – complete media growth factor

CoNS – coagulative-negative Staphylococci

CPSE – CHROMID® CPS® Elite

CTRL – control

DEG – differentially expressed gene

DFLac – difucosyllactose

DFLNH – difucosyllacto-N-hexaose

DFLNT – difucosyllacto-N-tetraose

DMEM – Regular Dulbecco's Modified Eagle Medium

DMM – Dirichlet Multinomial Mixtures

DOL – day of life

DSLNH – disialyllacto-N-hexaose

DSLNT – disialyllacto-N-tetraose

EDTA – ethylenediaminetetraacetic acid

ELISA – enzyme-linked immunosorbent-assay

EOS – early-onset sepsis

FAA – fastidious anaerobe agar

FBS – foetal bovine serum

FC – fold change

FDR – false discovery rate

FDSLNH – fucosyl-disialyllacto-N-hexaose

FFPE – formalin-fixed paraffin-embedded

FIP – focal intestinal perforation

FLNH – fucosyllacto-N-hexaose

FOS – fructo-oligosaccharides

Fuc – fucose

GA – gestational age

Gal – galactose

GalNAc – N-acetylgalactosamine

GALT – gut-associated lymphoid tissue

Glc – glucose

GlcNAc – N-acetylglucosamine

GNB – galacto-N-biose

GO – Gene Ontology

GOS – galacto-oligosaccharides

H&E - hemotoxylin and eosin

HIO – human intestine-derived organoid

HMO – human milk oligosaccharide

HPLC – high-performance liquid chromatography

ICS – incomplete chelating solution

IgA – immunoglobulin A

IL – interleukin

KEGG - Kyoto Encyclopedia of Genes and Genomes

LacNAc – N-acetyllactosamine

LCMS – liquid chromatography mass spectrometry

LDFT – lactodifucotetraose

LNB – Lacto-N-biose

LNDFH – lacto-N-difucohexaose

LNFP – lacto-N-fucopentaose

LNH – lacto-N-hexaose

LNnT – lacto-N-neotetraose

LNT – lacto-N-tetraose

LNTri – lacto-N-triose

LOS – late onset sepsis

LPS – lipopolysaccharide

LST – sialyl-lacto-N-tetraose

MALDI-TOF – Matrix-Assisted Laser Desorption/Ionisation-Time Of Flight

MC – mast cells

MCCV – Monte-Carlo cross validation

MDA – Mean decrease accuracy

mMRS – modified De Man, Rogosa and Sharpe

MOM – mother's own milk

MRS – De Man, Rogosa and Sharpe

MS – mass spectrometry

MT – metallothionein

NCBI – National Center for Biotechnology Information

NEC – necrotising enterocolitis

Neu5Ac – N-acetylneuraminic acid

NICU – neonatal intensive care unit

NMR – nuclear magnetic resonance

OACC – organoid-anaerobe co-culture

OD – optical density

OPLS-DA - Orthogonal Partial Least Squares - Discriminant Analysis

OR – odds ratio

ORA – over-representation analysis

OTU– operational taxonomic unit

PBS – phosphate-buffered saline

PCA – principal component analysis

PCoA – principal coordinates analysis

PCR – polymerase chain reaction

PERMANOVA – permutational multivariate analysis of variance

PFA – para formaldehyde

PGCT – preterm gut community type

PIO – preterm intestine-derived organoid

PLS-DA – Partial least squares – discriminant analysis

PMA – post-menstrual age

PUL – polysaccharide utilisation loci

ROC – receiver operating characteristic

SBP – solute binding protein

SCFA – short-chain fatty acid

SDWDC – Stir to dissolve in water to desired concentration

SERVIS – Supporting Enhanced Research in Vulnerable Infants

SLC – solute carrier family

SVM – support vector machine

TC – tissue culture

TEER – trans-epithelial electrical resistance

TLR4 – toll-like receptor 4

TNF – tumor necrosis factor

TOS – TOS-propionate agar medium

VFS – vacuum-filtered sterilisation

VLBW – very-low birthweight infants

WHO – World Health Organization

YPD – Yeast Extract Peptone Dextrose Agar

ZF – zinc finger

Chapter 1. Introduction

1.1 The infant gut microbiome

1.1.1 The human microbiome

The collection of microorganisms inhabiting the human body is referred to as the microbiome and is composed of bacteria, fungi, archaea, viruses, and bacteriophages. Various body sites are colonised by specific microbial communities, which can provide protection from potential pathogenic organisms and participate in host homeostasis (Pflughoeft and Versalovic, 2012). The gastrointestinal tract harbours the densest microbial community in the human body, with up to 10^{14} bacteria (Sender et al., 2016), and represents a vast interface where the immune system interacts with the outside world. Correspondingly, the intestinal tract hosts the largest number of immune cells, collectively known as gut-associated lymphoid tissue (GALT), accounting for almost 70% of the entire immune system (Jung et al., 2010). The role of the gut microbiome in human health and disease has long been established. The first use of faecal transplantation dates back to 1958, when four patients with severe pseudomembranous colitis were successfully treated using faecal enemas (Eiseman et al., 1958). The gut microbiome participates in human health by digesting food and fibres, producing essential vitamins and amino acids, and by competing with pathobionts, thus helping to prevent potential infections (Jandhyala et al., 2015) (Figure 1.1).

A balanced gut-microbiome interaction is essential for host health, and it is shaped from early life (Rautava, 2016, Rodriguez et al., 2015). For instance, it has been demonstrated in germ-free mice that a lack in microbial colonisation leads to impaired antibody production and development of gut-associated lymphoid tissue, low number of Peyer's patches and mesenteric lymph nodes, and altered microvilli formation by intestinal epithelial cells (Round and Mazmanian, 2009). Indeed, the gut microbiome in early life is important for immune training and metabolic programming of the infant, and a status of dysbiosis in this period has been associated to disease later in life (Rautava et al., 2012, Rodriguez et al., 2015). Early *Bifidobacterium* colonisation may exert a central role in shaping the human health status, with short and long-term effects. Decreased bifidobacteria colonisation together with lower microbial diversity early in life have been linked with higher incidence of various diseases, such as necrotising enterocolitis (NEC) (Stewart et al., 2016) and late-onset sepsis (LOS) (Stewart et al., 2017b) in infants, as well as obesity (Kalliomäki et al., 2008) and atopy (Kalliomäki et al., 2001, Fujimura et al., 2016) later in life.

1.1.2 Techniques to study the microbiome

To study and characterise the human microbiome, different methods can be applied, and we can divide them into two main groups: culture dependent and culture independent techniques, with the latter corresponding mainly to DNA sequencing methods (Figure 1.1). Each technique has its own benefits and drawbacks, and an integrative use of both approaches is needed to deeply study the microbiome ecology.

1.1.3 Culturing techniques

Microbiome studies in the past usually relied exclusively on direct culturing of microorganisms, with identification based on biochemical tests and more recently the use of Sanger or capillary sequencing. This approach requires the generation of pure cultures of living organisms which can be subsequently characterised and used for experimentation, such as in mechanistic studies of microbiome-host interaction.

Historically, an estimated 80% of intestinal bacteria was thought to be unculturable (Eckburg et al., 2005). However, advances in the past 10 years in so-called 'culturomics' have vastly improved the number and throughput of culture-based experiments, with a study from 2019 estimating that 67% of intestinal bacteria can be currently cultured (Diakite et al., 2019). Culturomics involves the utilisation of several culture conditions characterised by multiple enrichment and inoculation media, environmental conditions (temperature, oxygen levels, etc), and supplementation with different components (such as antibiotics, phages, amino acids) to enrich or deplete for specific species (Lagier et al., 2018, Alou et al., 2020) (Figure 1.1a). This step usually yields to hundreds or thousands of colonies which are rapidly identified using mass spectrometry (MS) or 16S rRNA gene sequencing when MS is inconclusive (Lagier et al., 2018).

Lagier *et al.* (2012) were the first to use a culturomics approach by coupling 212 culture conditions with identification of microorganisms through Matrix-Assisted Laser Desorption/Ionisation-Time Of Flight (MALDI-TOF) MS and 16S rRNA gene sequencing (Lagier et al., 2012). Such effort led to the identification of 174 species which had not been reported before as part of the human gut microbiome, of which 31 were new species and genera (Lagier et al., 2012). Further work from the group expanded the collection of isolated species from the gut microbiome and showed the importance of culturing (the uncultured species) to classify sequences not previously assigned in sequencing studies (Lagier et al., 2016). Over the past decade, more groups have applied such techniques to expand the collection of cultured

microorganisms, but also to optimise and decrease the number of conditions necessary for culturomics studies (Diakite et al., 2019, Forster et al., 2019, Diakite et al., 2020).

1.1.4 Sequencing techniques

Owing to the relatively low-throughput and inability to detect a large proportion of the bacteria within a given sample by utilising microbial culturing, over the past decade application of culture independent approaches has increased, with the advent of next generation sequencing techniques. Indeed, sequencing methodologies allow to detect the vast majority of microorganisms belonging to a microbial ecosystem, even though in certain cases they might fail to detect species isolated by culturing (Stewart et al., 2012).

Targeted 16S rRNA gene sequencing and metagenomics are the two principal culture-independent techniques adopted to date (Figure 1.1b). The first method consists in the targeted sequencing of the 16S rRNA gene, while sequences of random fragments are obtained with the metagenomic method. 16S rRNA gene sequencing is cost effective but is limited to the detection at the genus level. On the other hand, metagenomics is more expensive and is more computationally demanding, but since it is not based on sequencing a specific amplicon universal for bacteria, metagenomics facilitates the detection of all microorganisms (i.e., not only bacteria), provides resolution to species/strain level, and detects genes encoded by the microbiome (e.g., antibiotic resistance genes) (Ranjan et al., 2016). However, 16S rRNA gene sequencing maintain high value for studying samples in which low abundance of bacterial DNA and/or an abundance of host DNA is present, owing to the amplification step. Since sequencing methods allow researchers to deeply investigate microbial ecosystem composition but cannot give information about viability of microorganisms detected and relies on databases containing bacterial genomes known to date, coupling culture dependent and independent methods remains the desirable approach.

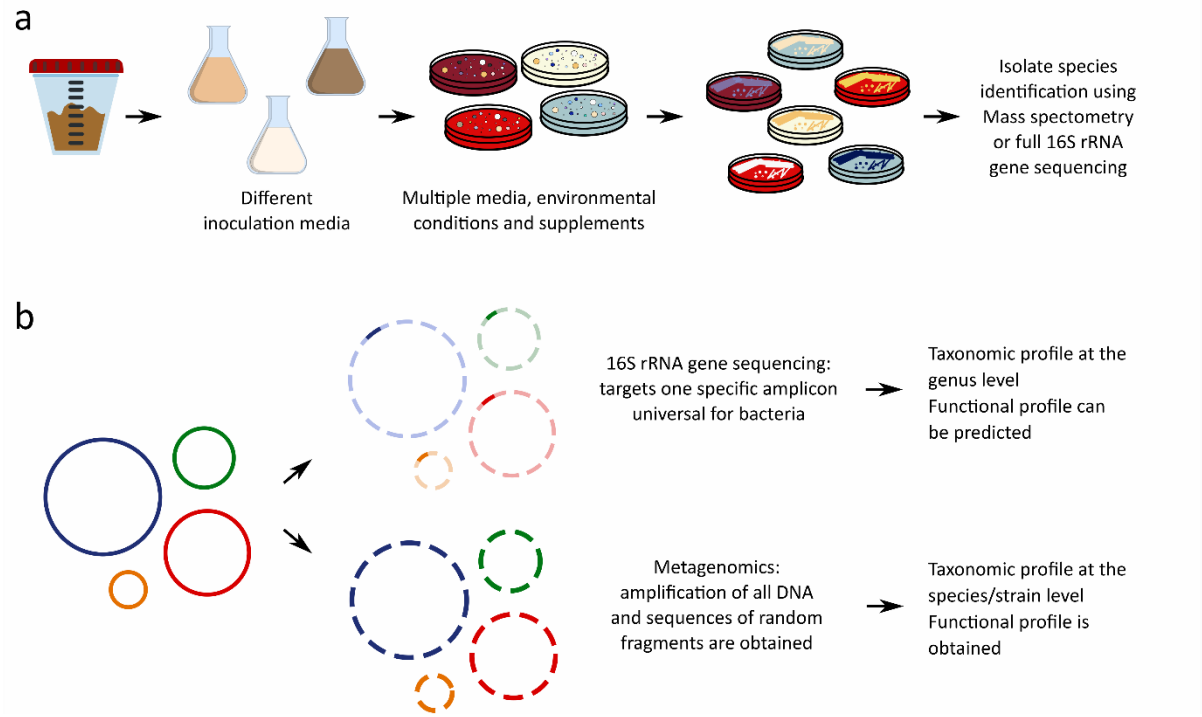

Figure 1.1. Schematic representation of techniques to study the microbiome composition. (a) Summary of steps involved in culturomics approach. (b) Sequencing techniques used to study microbial communities, 16S rRNA gene sequencing and metagenomic sequencing.

1.1.5 The term infant gut microbiome composition

Microbial colonisation begins at birth, when the neonate is exposed to viable bacteria and other microorganisms for the first time. Although some reports over the past years have hinted towards in utero detection of bacterial DNA within the placenta and amniotic fluid, and thus potentially the foetus, recent evaluation of published data suggests that such DNA fragments are not genuine and are likely contaminants (Kennedy et al., 2023). The infant gut is initially colonised by aerobic and facultative anaerobic bacteria (e.g., *Escherichia* and *Staphylococcus* spp.), followed by establishment of a more anaerobic community (e.g., *Bacteroides*, *Bifidobacterium* and *Clostridium* spp.) (Rodriguez et al., 2015, Palmer et al., 2007, Reyman et al., 2019). The determinants of succession are not completely understood and likely represent a combination of location seeding different pioneering and succession species, which reduce oxygen levels over the initial weeks, supporting strict anaerobes (Ferretti et al., 2018). The maturation of this microbial niche is characterised by an increase in bacterial richness and diversity, and an adult-like microbiome is achieved around 2–5 years of age, with a preponderance of Bacteroidota and Bacillota (Palmer et al., 2007, Rodriguez et al., 2015, Stewart et al., 2018).

The microbiome undergoes dynamic development that is influenced by a wide range of factors. Overall, mode of delivery (vaginal vs. C-section), use of antibiotics, either during gestation or after the delivery, and diet (breast milk vs. formula), play a major role (Rautava et al., 2012, Rodriguez et al., 2015). Term infants delivered by C-section showed an impaired colonisation by *Bacteroides*, suggesting the maternal gut to be an important source for infant seeding during the delivery (Stewart et al., 2018, Wampach et al., 2018, Shao et al., 2019, Reyman et al., 2019). Moreover, vaginal delivery associated with higher *Bifidobacterium* spp. abundance and lower abundance of pathogenic species, such as *Enterococcus* and *Klebsiella* spp. (Reyman et al., 2019), with alterations remaining during infancy, thus potentially having long-lasting effects (Shao et al., 2019). Mother to infant transmission of bacteria continues after delivery and maternal intestinal strains show the highest persistence in the infant gut (Ferretti et al., 2018). Antibiotic administration causes a decrease in the microbiome diversity, which can lead to an over-growth of antibiotic resistant microorganisms and potential pathogens (Palmer et al., 2007, Rodriguez et al., 2015).

Infant feeding has been found to be the most important contributor to microbiome development in term infants (Stewart et al., 2018). Breastfed infants harbour a microbiome dominated by bifidobacteria, while formula feeding is linked to a more diverse microbial composition containing lower levels of potentially beneficial bacteria, including *Lactobacillus* and *Bifidobacterium* spp. (Rautava et al., 2012, Rodriguez et al., 2015, Stewart et al., 2018). Defining a healthy infant gut microbiome is challenging, but microbial communities rich in *Bifidobacterium* spp. have been associated with positive outcomes and lower risk of various pathologies (Fujimura et al., 2016, Kalliomäki et al., 2008, Stewart et al., 2017b). Colonisation of the infant gut by this genus is known to be influenced by type of feeding, largely because of the prebiotic effect exerted by human milk oligosaccharides (HMOs) (Berger et al., 2020, Lawson et al., 2019, Vatanen et al., 2018), which will be discussed in later sections. Other variables influencing the development of the microbiome comprise environment, geographical location, and exposure to animals and siblings (Rodriguez et al., 2015, Stewart et al., 2018), as well as preterm delivery (Jia et al., 2020).

1.1.6 Preterm birth

Infants born <37 weeks of gestation are defined premature by the World Health Organization (WHO) (1977). More than one in ten babies was born premature in 2020 worldwide, and preterm birth rates varied by region, ranging from 4% up to 16% (Lawn et al., 2023, WHO, 2023). Multiple subpopulations can be defined based on the gestational age (GA) at birth.

Preterm birth is classified as late preterm (34 to <37 weeks of GA), moderate preterm (32 to <34 weeks of GA), very preterm (28 to <32 weeks of GA) and extremely preterm (<28 weeks of GA) (Natarajan and Shankaran, 2016, Patel, 2016). In this study we will focus on very and extremely preterm infants.

1.1.7 Gut microbiome development in preterm infants

The gut microbiome establishment in preterm infants is usually different from infants born at term. Unlike term infants, preterm babies born <32 weeks gestation are underdeveloped and immature, both anatomically and immunologically. In addition, preterm infants are more likely to be born by caesarean section, receive a greater number of antibiotics and antifungals, cannot be directly breastfed, are less likely to receive adequate breastmilk, and are housed within clean incubators with limited exposure to environmental microbes. The disruption of mother-infant contact can promote the colonisation of the infant gut by microorganisms found in the neonatal intensive care unit (NICU) environment (Brooks et al., 2017). Several NICU surfaces, such as infant incubators, soap dispensers, thermometers and sink drain, are colonised by potential pathogens (Brooks et al., 2014). Combined, these factors have profound and potentially detrimental influences on the developing preterm gut microbiome (Rodriguez et al., 2015, Brooks et al., 2017, Brooks et al., 2014, Barrett et al., 2013, Rahman et al., 2019, Jia et al., 2020).

The preterm gut microbiome is characterised by a reduced microbial diversity and is usually dominated by Pseudomonadota and Bacillota, with low levels of Actinomycetota and Bacteroidota, especially in the first weeks of life (Barrett et al., 2013). Indeed, compared to term infants, the preterm population shows a delayed colonisation by *Bifidobacterium* spp. while their gut microbiome remains dominated by pathobionts such ad *Escherichia* and *Klebsiella* (Jia et al., 2020). Moreover, when looking at longitudinal development of the microbial community, the preterm microbiome appeared to go through many shifts in the bacterial composition (La Rosa et al., 2014), instability that might represent a risk factor for health outcomes (Stewart et al., 2016). Probiotic mixtures containing *Bifidobacterium* spp. and *Lactobacillus* spp. are often administered to preterm infants due to their supposed beneficial effects. Samara *et al.* (2022) recently reported probiotic use to be linked to an acceleration towards the establishment of a more mature and stable microbial community in the preterm gut (Samara et al., 2022). A similar finding has been reported by Beck *et al.* (2022), who found probiotics administration as being associated to establishment of bifidobacteria rich communities, which are found in oldest infants, suggesting these to be more mature microbial

communities (Beck et al., 2022). Moreover, effect of different probiotic products could be evaluated, and different strains of the same probiotic species were demonstrated to individually influence the establishment of the preterm microbiome and showed different persistence potential after probiotic administration was stopped (Beck et al., 2022). This suggests that selecting the right probiotic strain might be pivotal to result in the health benefit desired.

1.2 Human breast milk

Human milk is a complex biofluid, which not only provides nourishment to the new-born, but also helps the immature body cope with foreign environmental stimuli and the microorganisms it first encounters (Ballard and Morrow, 2013). Breast milk is composed of a mix of nutrients, prebiotics, viable microorganisms, and antimicrobial and immunomodulatory molecules. Human breast milk composition is variable, and is affected by multiple factors, such as maternal diet, body mass index (BMI) and maternal age (Gila-Diaz et al., 2019). Breast milk varies also depending on lactation period, and its composition is dynamic and adapts to the infant needs (Gila-Diaz et al., 2019, Bode et al., 2020).

1.2.1 Breast milk in preterm infants

Breast milk feeding is beneficial to health of both infant and mother, and in preterm babies it is associated with lower incidence of various diseases, including NEC and LOS (Bode et al., 2020). However, due to the immature digestive and immune systems, preterm infants are initially parenteral fed and enteral feeding is introduced slowly and with caution. Furthermore, mothers who deliver prematurely are often unable to produce sufficient breast milk, so alternative nutritive sources must be utilised, such as bovine-milk based formulas, human-milk fortifier and pasteurised donor human milk (Shulhan et al., 2017).

Preterm breast milk is different in composition when compared to term milk (Jie et al., 2018, Li et al., 2019, Ingvordsen Lindahl et al., 2019, Xavier et al., 2018). True protein and fat concentration are higher in preterm milk during the first few days, while lactose appears to be unaltered (Gidrewicz and Fenton, 2014). Due to the paucity of breast milk, mothers are encouraged to express when milk is available; it is subsequently frozen for storage, and thawed when the baby needs to be fed. This management of preterm breast milk might impact on infant health, as freeze-thaw cycles can alter milk composition as observed for secretory immunoglobulin A (IgA) concentration (Granger et al., 2021b). Nonetheless, while fresh breast milk would be optimum for infant nutrition, consideration must be given to milk availability and the practicalities of caring for preterm infants in the NICU environment. When mother's

own milk (MOM) is not available, donor human milk is the best alternative for neonatal feeding. However, donor human milk is provided by term mothers in a late stage of lactation, when protein content decreases, and it undergoes pasteurisation, which might deactivate some important bioactive components (Shulhan et al., 2017). Thus, identifying specific factors which drive major protective effects in preterm health is of vital importance, since they might be utilised to improve infant nutrition.

1.2.2 Human milk microbiome

Human breast milk carries around 10^6 bacteria per ml, typically ranging from 10^5 up to 10^8 (Boix-Amoros et al., 2016). Breast milk microbiota seems to be initially seeded through the entero-mammary pathway starting from the late stages of pregnancy. According to this theory, bacteria would be transported by immune cells from the gut through to the mammary glands, where they are released and colonise this body site (Moossavi and Azad, 2019). After birth, the infant's mouth represents an additional source shaping breast milk microbiome, as milk can flow back into the breast during suction (Moossavi and Azad, 2019). This last mechanism might be crucial for the development of a "healthy" breast milk microbiome. A study including 393 breastfeeding dyads reported that indirect breastfeeding (e.g., use of breast milk pumps) was associated with higher abundance of potential opportunistic pathogens possibly associated with breast-pumps and less frequent presence of bifidobacteria (Moossavi et al., 2019b). As mentioned above, preterm babies cannot be directly breastfed during the first weeks of life, and use of breast-pumps might enhance the onset of a gut microbiome abundant in potential pathogens.

Multiple factors might shape the breast milk microbiome, including BMI, parity, and mode of delivery (Moossavi et al., 2019b, Asbury et al., 2020). HMOs composition has also been linked to the breast milk microbiome, although different findings are reported by the various research groups (Moossavi et al., 2019a, Ramani et al., 2018, Hunt et al., 2012). Studies focusing on the characterisation of breast milk microbiome in moms who delivered at term reported again different findings, likely mirroring the influence of location and diet on this microbial niche. *Staphylococcus*, *Streptococcus* and *Pseudomonas* are the bacterial genera commonly shared between mothers and cohorts (Boix-Amoros et al., 2016, Moossavi et al., 2019b, Hunt et al., 2011, Urbaniak et al., 2016, Biagi et al., 2018, Lackey et al., 2019). *Bifidobacterium* and *Lactobacillus* spp., which are notable beneficial bacteria, are also frequently found in this biological fluid in lower relative abundances (Boix-Amoros et al., 2016, Moossavi et al., 2019b, Urbaniak et al., 2016, Biagi et al., 2018, Lackey et al., 2019, Khodayar-Pardo et al., 2014). In

a study from 2016, Urbaniak *et al*. found that breast milk microbiome composition was not affected by time of delivery in their cohort (Urbaniak et al., 2016), however, this study relied on 16S rRNA gene sequencing which provides bacterial relative and not absolute abundances. A study from Khodayar-Pardo *et al*. (2014) found preterm delivery to affect total bifidobacteria count, with lower loads when compared to term milk, which might reduce transmission chances of this beneficial gut commensal to the infant (Khodayar-Pardo et al., 2014). Recently, Asbury *et al.* (2020) published the largest preterm breast milk microbiome cohort to date, which included weekly sampling from 86 mothers of infants born <1250g up over the first 8 weeks post-partum (Asbury et al., 2020). Similar to term cohorts, the most abundant genera included *Staphylococcus, Acinetobacter* and *Pseudomonas*. Despite many factors being associated with breast milk microbiome changes over time, including delivery mode and mother's antibiotics exposure, each mother showed an overall individualised MOM microbial community (Asbury et al., 2020). The lower abundance of the oral coloniser *Streptococcus* observed in preterm MOM compared to term cohorts might be a consequence of lack of direct breastfeeding (i.e., skin-oral contact), since its relative abundance has been reported in preterm milk after initiation of infant's direct suction from the breast (Biagi et al., 2018, Asbury et al., 2020). The lack of initial direct breastfeeding might also be associated with the relatively high abundance of *Stenotrophomonas* and Pseudomonadaceae observed in the preterm MOM, microorganisms which have been correlated with breast pump use by Moossavi *et al.* (2019) (Moossavi et al., 2019b). Such findings underline the altered exposure to microorganisms encountered by infants born prematurely.

1.2.3 Human milk oligosaccharides: structures and profiles

HMOs are a family of structurally complex unconjugated glycans characteristic of human milk, which are involved in the modulation of epithelium (He et al., 2016, Lane et al., 2013, Natividad et al., 2020, Wu et al., 2019), immune system (Bode et al., 2004, Eiwegger et al., 2004), and microbiome (Lawson et al., 2019, Borewicz et al., 2020). HMOs represent the third most abundant solid component in human milk after lactose and lipids, with a concentration of 9–24 g/L, which usually exceeds the quantity of total proteins (Bode, 2012, Zuurveld et al., 2020). More than 200 structurally diverse HMOs have been reported, although 20–25 of them are expressed in appreciable quantities and account for >95% of total HMOs.

HMOs derive from the arrangement of five monosaccharides: glucose (Glc), galactose (Gal), N-acetylglucosamine (GlcNAc), fucose (Fuc), and sialic acid (N-acetylneuraminic acid (Neu5Ac)) (Figure 1.2a). GlcNAc and Gal can form two different disaccharides: lacto-N-biose

(LNB) in case of β1-3 linkage, N-acetyllactosamine (LacNAc) with β1-4 linkage (Figure 1.2b). Every HMO is composed by a lactose molecule at the reducing end and a variable number of LNB (type 1 chain) and LacNAc (type 2 chain) (Figure 1.2c). The lactose molecule is initially elongated with one of the two disaccharides through β1-3 or β1-6 linkage. Although the addition of LNB terminates the chain and is not affected by additional modifications, LacNAc can be further elongated with the introduction of supplementary disaccharides. Branched molecules are formed when the disaccharide is added with a β1-6 linkage. The resulting HMOs can then be modified by the addition of Fuc (α1-2, α1-3, and α1-4 linkage) and/or Neu5Ac (α2-3, α2-6 linkage) (Bode, 2012, Zuurveld et al., 2020).

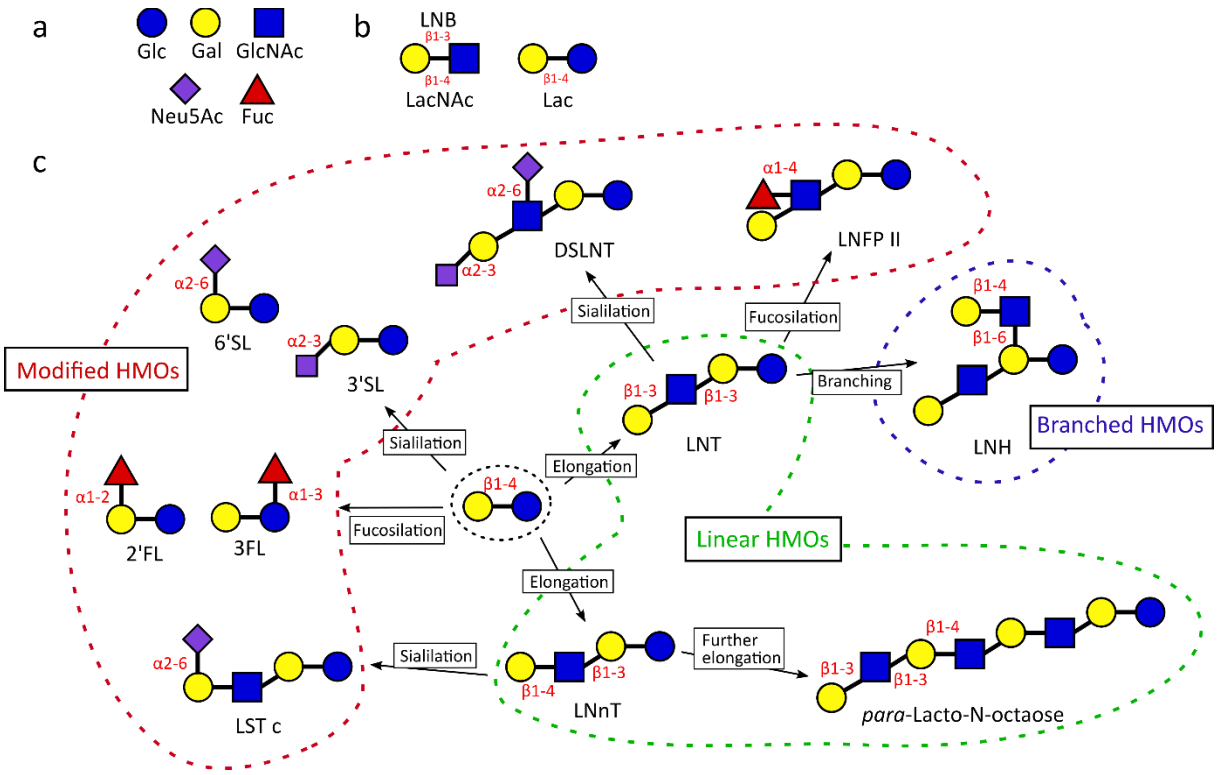

Figure 1.2. Human milk oligosaccharides structures. (a) Monosaccharides and (a) disaccharides composing HMOs. (c) Examples of type of linkages and modifications characterising HMOs. Glc, glucose; Gal, galactose; GlcNAc, N-acetylglucosamine; Neu5Ac, N-acetylneuraminic acid; Fuc, fucose; LNB, lacto-N-biose; LacNAc, N-acetyllactosamine; Lac, lactose; HMOs, human milk oligosaccharides; 2'FL; 2'-fucosyllactose; 3FL, 3-fucosyllactose; LNnT, lacto-N-neotetraose; 3'SL, 3'-sialyllactose; 6'SL, 6'-sialyllactose; LNT, lacto-N-tetraose; LNFP II, lacto-N-fucopentaose II; LST c, sialyl-LNT c; LNH, lacto-N-hexaose; DSLNT, disialyllacto-N-tetraose.

Although the more abundant HMOs are expressed in comparable amounts between mothers, each woman will produce only a subset of the possible structures, leading to an oligosaccharide composition specific to each mother (Bode, 2012). The synthesis of the HMOs relies on many

glycosyl-transferases specific for the various monomers, however, the HMO composition depends on the genetic profile of the mother determining the activity of two fucosyl-transferases. Secretor women express the Se gene encoding for an active α1-2-fucosyltransferase FUT2, and Lewis positive women carry a Le gene encoding for an active α1-3-fucosyltransferase FUT3. Secretor mothers will be characterised by a breast milk with high concentration of α1-2-fucosylated HMOs (e.g., 2'-fucosyllactose – 2'FL), while the presence of an active Le gene is associated with enrichment in α1-4-fucosylated HMOs (e.g., lacto-N-fucopentaose (LNFP) II – LNFP II). Le-negative and Se-negative women, thus women not expressing active enzymes encoded by these two gene loci, will still have low concentrations of α1-2 and α1-4-fucosylated HMOs, suggesting that other fucosyl-transferases might be involved, even though with low activity (Bode, 2012, Zuurveld et al., 2020). Interestingly, Secretor status does not only affect fucosylated HMOs, but also the concentration of non-fucosylated HMOs has been shown to be influenced by this genetic factor (Azad et al., 2018).

Additional factors which influence the HMO composition include maternal age, ethnicity, parity, obesity, country, lactation stage and diet (McGuire et al., 2017, Azad et al., 2018, Neville et al., 2021, Seferovic et al., 2020). HMO composition might also be shaped by preterm delivery, however, contrasting findings have been published (Kunz et al., 2017, Nakhla et al., 1999, Austin et al., 2019, De Leoz et al., 2012). Wang *et al.* (2001) found a greater concentration of sialic acid HMOs in preterm milk compared to the term counterpart (Wang et al., 2001). In the largest longitudinal study to date involving 27 and 34 mothers of preterm and term infants respectively, Austin *et al.* (2019) found that disialyllacto-N-tetraose (DSLNT), 3'-sialyllactose (3'SL) and sialyl-lacto-N-tetraose (LST) b were higher in the preterm milk between 2 to 8 weeks post-partum, while LSTc and 6SL were higher in the term milk during the first month (Austin et al., 2019). Moreover, most HMOs decreased over-time in both groups, with the exception of DSLNT, lacto-N-tetraose (LNT) and lacto-N-difucohexaose I (LNDFH I) which peaked after the first week (Austin et al., 2019). Similarly, in the biggest term cohort (427 mothers), most HMOs decreased in concentration over the course of lactation, except for DSLNT (Azad et al., 2018).

1.2.4 Human milk oligosaccharides activities

HMOs are considered prebiotics because, being indigestible to the infant, they reach the gut intact where they promote the growth of potentially beneficial bacteria (Engfer et al., 2000). They are mainly digested by *Bifidobacterium* spp. but can also be used by other bacteria thanks

to the HMOs resemblance to mucins (Yu et al., 2013). Indeed, species belonging to the genera *Bacteroides, Roseburia, Eubacterium* and *Akkermansia* have been reported to utilise HMOs through genes involved in mucin degradation (Marcobal et al., 2011, Pichler et al., 2020, Luna et al., 2022). HMOs protect the infant gut from potential pathogen colonisation, not only promoting beneficial bacteria colonisation to occupy available niches, but also by directly acting as antiadhesive antimicrobials (Bode, 2012, Zuurveld et al., 2020) (Figure 1.3). Indeed, they are able to coat pathogens, preventing their adhesion to epithelial surfaces and thus reducing risk of infection, as reported for *Escherichia coli* (Wang et al., 2020) and *Campylobacter jejuni* (Ruiz-Palacios et al., 2003). Such action as decoy receptors is made possible by their resemblance to cell surface glycans which are used by pathogenic species during the infection process (Zuurveld et al., 2020). However, it is notable that specific HMOs are also able to increase infectivity of specific neonatal rotavirus strains, but not adult or bovine strains, highlighting the need for further work (Ramani et al., 2018).

HMOs can directly interact with and modulate the immune and gastrointestinal systems (Zuurveld et al., 2020) (Figure 1.3). At the level of intestinal tissue, HMOs are able to regulate the development and maturation of the gut. As reported by Lane *et al.* (2013), exposure of HT29 cells to HMOs impacted the gene expression profile, with modulation of cytokines, chemokines, and cell surface receptors (Lane et al., 2013). In another study, T84 epithelial cell line exposed to HMOs produced lower levels of proinflammatory mediators, whereas cytokines involved in tissue repair and homeostasis were increased (He et al., 2016). Intestinal barrier function can also be modulated by HMOs, with reduced gut permeability observed after treatment with these glycans (Natividad et al., 2020, Wu et al., 2019).

Sialylated HMOs have been reported to influence lymphocyte maturation and modulate a low-level immunity by shaping the immune system towards a regulatory type Th1 response (Eiwegger et al., 2004). Acidic HMOs were also able to modulate and lower monocyte, lymphocyte, and neutrophil adhesion to endothelial cells, where excessive leukocyte infiltration can cause major tissue injury in inflammatory diseases (Bode et al., 2004). Beyond the gut, it has been further reported that 1% of HMOs can be absorbed through the gut into the systemic circulation, giving rise to the possible impacts on modulating systemic immunity. To show this, mothers were given 13C-labelled galactose, and the 13C-labelled HMOs which resulted by this supplementation (Rudloff et al., 2006), could also be found in the urine of their breast-fed infants, suggesting that these sugars can be partially absorbed (Obermeier et al., 1999, Rudloff et al., 2012).

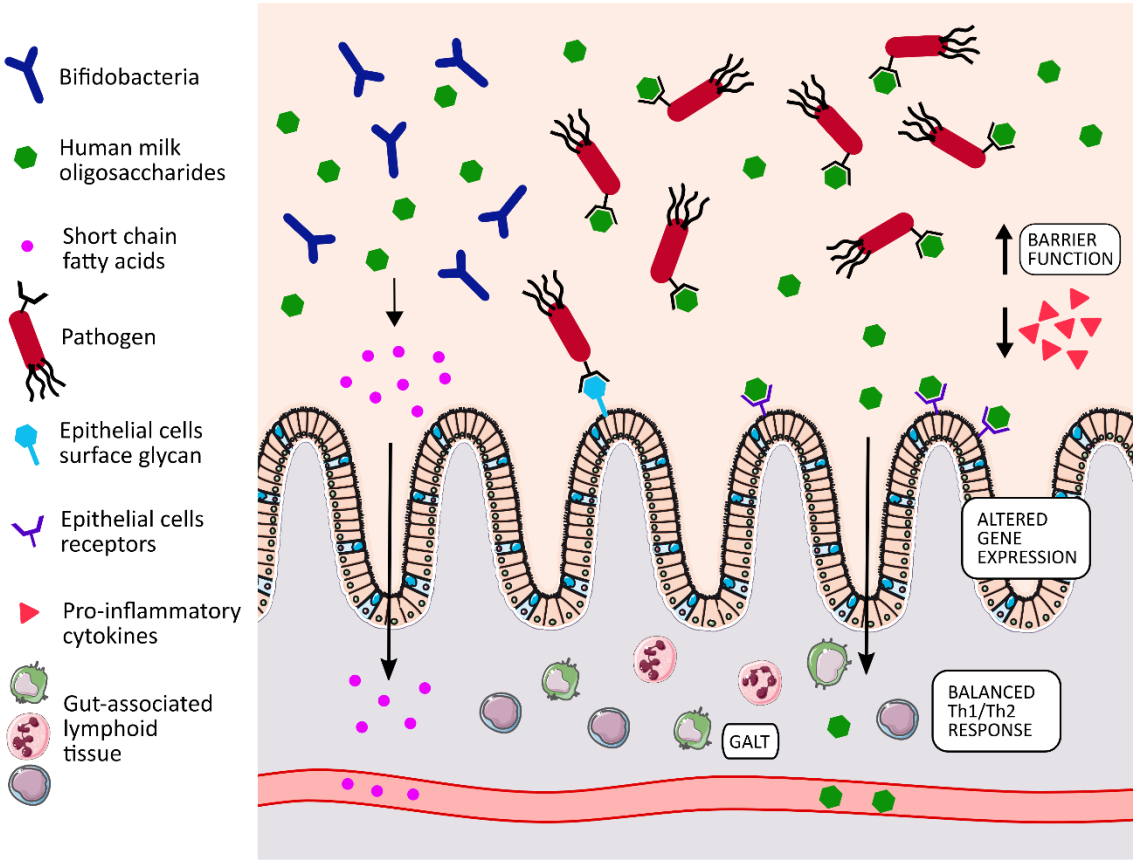

Figure 1.3. Mechanisms by which human milk oligosaccharides shape infant health.

1.2.5 Human milk oligosaccharides and the infant gut microbiome

The microbial shaping effect exerted by breast milk feeding is driven through various mechanisms, including by directly providing potential colonisers from the breast milk microbiome, immune factors (e.g., secretory IgA, antimicrobial peptides, and proteins), and HMOs (Granger et al., 2021a). Breast milk feeding shapes the infant gut microbiome towards a gut community rich in *Bifidobacterium* spp. The first observation of the enrichment in bifidobacteria in the stool of breast milk fed infants compared to formula fed infants dates back to 1900, through the work of Tissier at the Pasteur Institute (Tissier, 1900). The central role of HMOs in explaining breast milk's bifidogenic-effect was then discovered in the mid-1900s by György and colleagues (György et al., 1954, Gyorgy et al., 1954b, Gyorgy et al., 1954a, Gauhe et al., 1954). They performed subsequent experiments including growth of *Lactobacillus bifidus*, then renamed *Bifidobacterium bifidum*, in media supplemented with breast milk and other various different supplements, as well as additional experiments, to isolate the milk components having the bifidogenic effect (Tissier, 1900). Recent advances in DNA sequencing technologies have further shown the correlation between breast milk ingestion and bifidobacteria colonisation, with multiple studies characterising early life gut microbiome establishment (Berger et al., 2020, Borewicz et al., 2020, Lawson et al., 2019, Vatanen et al.,

2018). However, the first studies elucidating the mechanisms and genes involved in HMOs utilisation by species belonging to this genus are relatively recent, dating to early 2000 (Katayama et al., 2004, Kitaoka et al., 2005, Møller et al., 2001, Nishimoto and Kitaoka, 2007a, Sela et al., 2008), and will be discussed in later sections.

There are many ways through which HMOs shape the infant gut microbiome composition. As mentioned above, bifidobacteria have an advantage in colonising the gut of breast milk fed infants as they are the main HMOs utilisers. They are thus able to fill this microbial niche, and their metabolism produces an acidic environment which in return inhibits the growth of potential pathogens. Moreover, HMOs can increase the adhesion of microorganisms to the intestinal epithelium, including *Bifidobacterium* (Wickramasinghe et al., 2015, Chichlowski et al., 2012) and *Lactobacillus* spp. (Kong et al., 2021), thus favouring their colonisation. Correlation between singular HMOs and specific bacteria (e.g., *Staphylococcus* and *Streptococcus*) have also been reported (Borewicz et al., 2020), however the nature of some of these correlations is not clear yet. For instance, it has been shown *in vitro* that HMOs can inhibit the growth of group B *Streptococcus* (Lin et al., 2017, Andreas et al., 2016), and an association between its colonisation of the infant gut and the HMO profile was also found (Andreas et al., 2016). A different example is instead represented by an enteropathogenic *E. coli*, which in the presence of HMOs showed decreased epithelial cells invasion *in vitro* and lower intestinal colonisation of a neonatal mice model (Manthey et al., 2014). These differences were mediated through HMOs coating of the bacterium and not by affecting its growth (Manthey et al., 2014, Lin et al., 2014). Thus, the infant gut microbiome modulation exerted by HMOs is the results of multiple different mechanisms of effect, and many of the correlations between specific structures and bacteria have yet to be elucidated.

1.2.6 Techniques to study HMOs composition

Interest in HMOs in health and disease has gained increased interest, in part, owing to advances in techniques applied for their analysis (Smilowitz et al., 2014, van Leeuwen, 2019). HMOs are challenging molecules to measure due to their complexity, for instance they can be highly branched, can present with multiple structural isomers and they lack intrinsic chromophores (Mantovani et al., 2016). Various methods have been applied to date, including Nuclear magnetic resonance (NMR), high-performance liquid chromatography (HPLC), and MS. NMR has been widely used to elucidate HMOs structures, however, branching information is difficult to be assigned through NMR data and high quantities of purified HMOs are needed. Despite

not being ideal for studying human samples, NMR coupled with MS has valuable role in determining HMOs structures (Mantovani et al., 2016).

HPLC coupled with online fluorescence detection (HPLC-FL) and liquid chromatography mass spectrometry (LCMS) are frequently used in HMO studies, where relative or absolute abundances of HMOs can be determined with these methods (van Leeuwen, 2019, Smilowitz et al., 2014). These chromatography techniques are indeed used to separate and identify the different HMO structures (Mantovani et al., 2016). However, due to the lack of intrinsic chromophore group in HMOs, derivatisation of these glycans is usually performed in order to increase sensitivity of the techniques, allowing also to improve the resolution of separation by introducing additional charges (Mantovani et al., 2016). However, identification of the different HMO structures remains challenging due to the oligosaccharide complexity of milk samples, and limitations are imposed by the restricted availability of commercial HMOs standards. Coupling chromatography with MS is often performed to fully characterise complex HMOs mixtures, since further structural information can then be included (van Leeuwen, 2019).

1.2.7 Human milk oligosaccharides enzymatic degradation

For HMO utilisation, the single monosaccharides need to be released, and the breakdown of each linkage requires specific enzymes that exist in microbes. To be able to utilise the core structure, the modifications are often removed first (Ashida et al., 2009). The enzymes responsible for fucose release are named fucosidases, and two types are present: 1,2-α-L-fucosidase acting on α-1,2 linkage and 1,3-1,4-α-L-fucosidase acting on fucose added with α-1,3 and α-1,4 linkages. The first enzyme will act mainly on 2'FL, LNFP I (Katayama et al., 2004), and the second on 3-fucosyllactose (3FL), LNFP II, and LNFP III (Ashida et al., 2009). Sialidases on the other hand are enzymes responsible for the liberation of Neu5Ac from the core structure by acting on the α-2,3 and α-2,6 linkages (Kiyohara et al., 2011). Bacteria can further metabolise Fuc and Neu5AC released from the core structures (Brigham et al., 2009, Bunesova et al., 2016, Salli et al., 2021).

After modifications are removed, the core structure can be digested. Two enzymes can target Gal: β-1,3-galactosidase which acts mainly on LNT, but also on β-1,4 linkage of Lac, LNB and lacto-N-neotetraose (LNnT) (James et al., 2016, Yoshida et al., 2011); β-1,4-galactosidases act on the β-1,4 linkage found in type 2 chains and Lac (James et al., 2016, Miwa et al., 2010). Further disassembling the HMOs structures, the bond between GlcNAc and Gal is cleaved by β-N-acetylglucosaminidases. Various enzymes of this type have been reported in literature, each having their own preferences for the specific HMOs targeted and some of which being

able to act on β-1,3 and β-1,6 linkages equally, whereas others having a preference on specific linkage type (Honda et al., 2013, James et al., 2016, Miwa et al., 2010). In certain cases, the first and outer Gal residue needs to be removed, before the enzyme can free the GlcNAc (Garrido et al., 2012). Finally, the enzyme Lacto-N-biosidase can act on LNT generating LNB and Lac, which can be further metabolised (Sakurama et al., 2013, Wada et al., 2008).

1.2.8 Methodologies to identify HMO utilisation genes

Generation of genomic libraries has often been applied for the discovery and characterisation of genes implied in HMOs degradation (Katayama et al., 2004, Møller et al., 2001, Sakurama et al., 2013) (Figure 1.4a). This approach consists of various steps: initially the whole genome of the bacterium of interest is randomly fragmented; the DNA fragments obtained are inserted in an appropriate vector, and the collection of vectors obtained is transfected in a suitable bacterial host (usually *E. coli*) (Clark and Pazdernik, 2013). The collection of clones obtained is the genomic library, which can subsequently be screened for the target function by observation of the phenotype of interest (Clark and Pazdernik, 2013). For instance, to determine novel genes involved in fucose utilisation, Katayama *et al*. (2004) lysed each transformed colony, and the cell content was incubated with 2'FL (Katayama et al., 2004). The reaction mixture was then analysed using a thin-layer chromatography to identify the colonies harbouring the function desired (Katayama et al., 2004). After selecting the clones showing the acquired phenotype, retrieval of the gene sequence can be performed by sequencing the fragment inserted in the vector (Katayama et al., 2004).

Transcriptome profiling, utilising RNA-sequencing (Garrido et al., 2016) or microarrays (James et al., 2016), has also been used (Figure 1.4b). The bacterial species are initially grown in the specific substrate of interest, the RNA is subsequently extracted, converted to complementary DNA, and the transcription profile defined. To identify induced genes associated with the substrate metabolism, a known and well characterised condition is used as control (James et al., 2016, Garrido et al., 2016). For instance, Garrido *et al*. (2016) grew *B. longum* in lactose and performed transcriptome analysis to identify HMO utilisation genes in this species (Garrido et al., 2016). Genes with significantly elevated expression by exposure to HMOs or related structures can then be further characterised in subsequent experiments, such as purification and characterisation of the protein encoded (James et al., 2016, Garrido et al., 2016). The putative role exerted by the identified genes can also be determined *in silico* by comparing the gene sequence and/or protein sequence with known genes and proteins stored in specific databases (James et al., 2016). This process is also central to the whole genome sequencing approach

(Ashida et al., 2009, LoCascio et al., 2010, Miwa et al., 2010, Sela et al., 2008). In the first step, the bacterial genome sequence is annotated, which consists in identification of genes, promoters, rRNA genes, untranslated regions, and pseudogenes (Beckloff et al., 2012). The function of the genes is then identified by comparing their sequence identity towards genes of known function stored in publicly available databases. Owing to declining sequencing costs, the number of genomes deposited in the dedicated databases is increasing, allowing the study of the conservation of HMO utilisation genes and gene clusters in bacteria families, genera, species, and strains (Kitaoka et al., 2005).

The genomic library screening and transcriptome profiling approaches allow the identification of novel genes, as they do not rely on characterised genes reported in the databases. On the contrary, while less laborious, the whole genome sequencing method can only find homologous of known HMO utilisation genes in uncharacterised bacterial species and strains. Thus, complementing the different techniques is central to expanding the collection of the well characterised genes and investigating their distribution and conservation in bacteria.

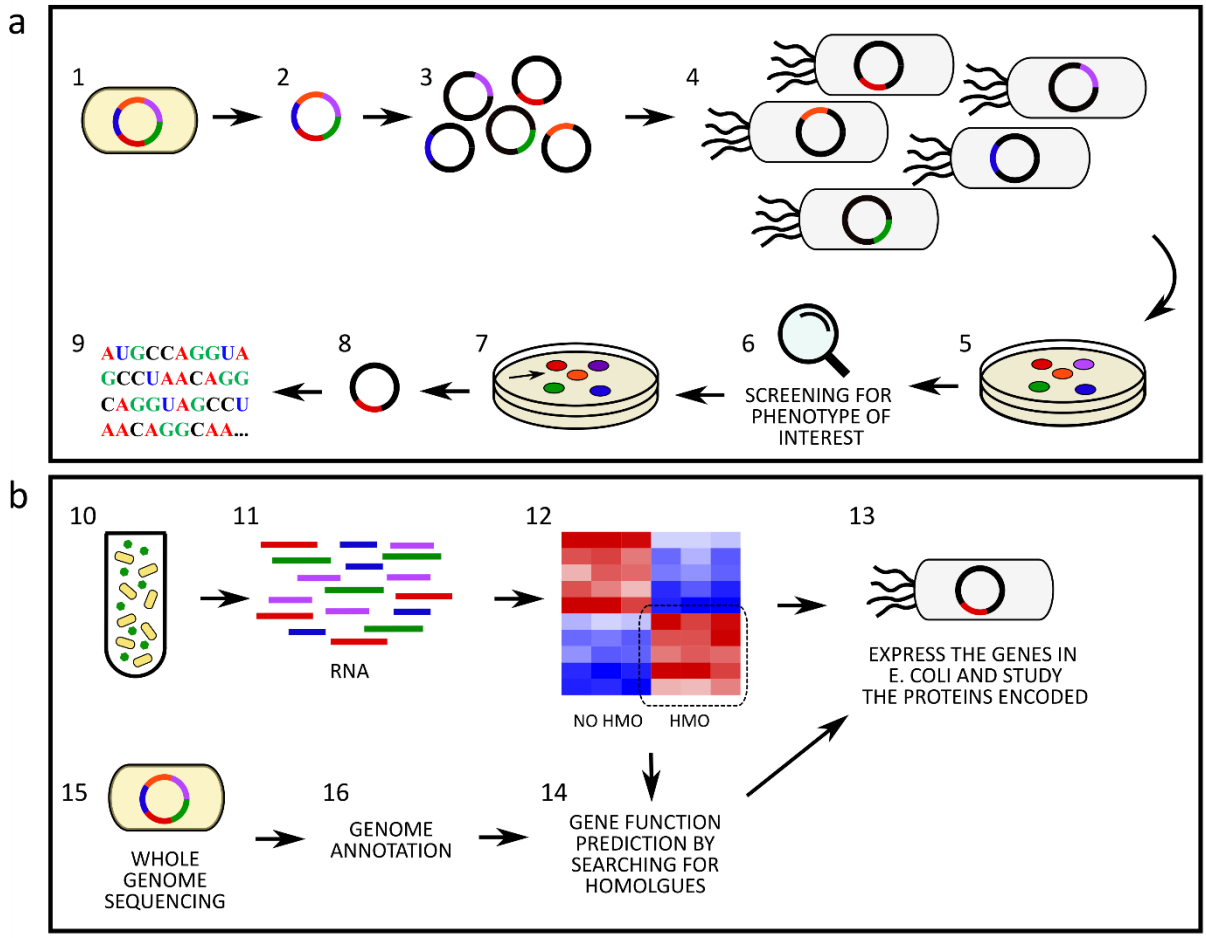

Figure 1.4. Methodologies for the identification of human milk oligosaccharides utilisation genes. (a) Approach involving genomic library formation. The bacterium of interest (1) is lysed, and the genome is retrieved (2). The genome is randomly fragmented, and the fragments are

inserted in a cloning vector (3), which are then transfected in the host bacterium (4), creating the genome library (5). The clones composing the genome library are then screened for the identification of the phenotype of interest (6) and the colony harbouring such a function (7) is retrieved. The vector is reisolated from the identified clone (8) and the gene sequence is determined. (b) Approaches relying on sequencing methods. The bacterium of interest is grown in media containing HMOs as the sole carbon source (10), the RNA is extracted (11) and the transcriptome profile determined. The function of the genes upregulated in media containing HMOs vs the reference sugar is subsequently determined by inserting it in a vector and expressing it in a host bacterium (13), or by predicting its function searching for homologues (14). Alternatively, the sequencing of the bacterial genome is performed (15), the genome is annotated (16) and the gene function is predicted by searching for homologues (14). The gene function can be further studied by expressing the gene in a host bacterium, as described above (13).

1.2.9 Bifidobacterium

Bifidobacteria are anaerobic Gram-positive bacteria belonging to the Actinomycetota phylum and are generally considered to be beneficial to the human body through various mechanisms including production of short-chain fatty acids (SCFAs) (Alessandri et al., 2019). The ability of *Bifidobacterium* spp. to colonise the human gut differ by the host stage of life, with *B. breve*, *B. bifidum* and *B. longum* subsp. *infantis* being found mainly in breastfed infants, whereas older subjects usually carry *B. adolescentis* and *B. catenulatum*. Only *B. longum* subsp. *longum* has been reported to colonise the human gut throughout life (Alessandri et al., 2019). Such differences can be explained, in part, by the bacteria's capability to utilise HMOs, which can greatly vary between different species, subspecies, and strains.

Two different strategies for HMOs digestion have been reported: *B. bifidum* is equipped with extracellular glycosidases (Ashida et al., 2009, Katayama et al., 2004, Møller et al., 2001, Wada et al., 2008), which therefore act on the HMO linkages outside of the cell and release the monosaccharides and disaccharides (LNB and Lac) in the surrounding environment, which can either be left for the growth of other bacteria or transported internally to be metabolised inside the cell (Asakuma et al., 2011, Gotoh et al., 2018, Kiyohara et al., 2011, Miwa et al., 2010, Nishimoto and Kitaoka, 2007b, Suzuki et al., 2008) (Figure 1.5a). On the contrary, *B. longum* subsp. *infantis*, *B. breve*, and *B. longum* tend to internalise the oligosaccharides through ATP-binding cassette (ABC) transporters and digest the sugar structure internally (Asakuma et al., 2011, Garrido et al., 2016, Møller et al., 2001, Sakanaka et al., 2019b, Sela et al., 2008) (Figure 1.5b-d). Extracellular glycosidases have also been found in *B. longum*, but with a lower frequency (Sakanaka et al., 2019a, Sakurama et al., 2013). Likely because of the strategy adopted, *B. bifidum* has been demonstrated to cross-feed other *Bifidobacterium* spp. not

equipped for HMO degradation, but able to utilise the released degradants, as reported both *in vitro* (Gotoh et al., 2018) and *in vivo* from infant gut microbiome studies (Tannock et al., 2013). *B. bifidum* and *B. infantis* strains are the most frequent utilisers of HMOs, and many different structures, modified and non-modified (i.e., structures carrying or not Fuc and/or Neu5Ac), can be digested by these two (sub)species (Garrido et al., 2016, Katayama et al., 2004, LoCascio et al., 2007, LoCascio et al., 2010, Sela et al., 2008, Sela et al., 2012). On the contrary, *B. longum* and *B. breve* can utilise only LNT, LNB, and LNnT, whereas utilisation of modified HMOs has been described only in a few strains (Asakuma et al., 2011, Thongaram et al., 2017). *B. infantis* is reported to be the most efficient utiliser of HMOs and is able to consume up to 64% of total pooled HMOs, compared to a utilisation between 23% and 43% displayed by other species (LoCascio et al., 2007, LoCascio et al., 2010). Sela *et al.* (2008) were the first to report a vast HMO-utilisation cluster in *B. infantis* comprising 30 genes, some of which are likely subjected to a communal transcriptional regulation (Sela et al., 2008). This cluster comprises 4 glycosidases (a fucosidase, a sialidase, a β-N-acetylglucosaminidase, and a β-galactosidase), 2 ABC transport permeases and associated ATPase, and 7 solute binding proteins (SBPs) predicted to bind oligosaccharides (Sela et al., 2008). Other genes implicated in HMO utilisation are also found in other positions in the genome. Moreover, a total of 21 copies of family 1 SBPs were found in this subspecies, compared to 10 and 11 found in *B. longum* and *B. adolescentis*, and 6 of the SBPs in the cluster show evolutionary divergence compared to other family 1 SBPs (Sela et al., 2008). These characteristics of *B. infantis* guarantee its potential to utilise many different HMOs applying a strategy of internal hydrolysis and compete in the infant gut. It has to be noted that not all *B. infantis* strains are equipped with the full set of genes reported in strain JCM 1260 (LoCascio et al., 2010), and indeed showed lower capacity of growth in HMOs compared to other *B. infantis* strains (Locascio et al., 2009). Notably, not all genes involved in HMO degradation have been identified to date. Lawson *et al*. (2019) isolated bifidobacterial strains that were able to grow on HMOs, but lacking known genes and clusters for their digestion, suggesting the presence of uncharacterised HMO utilisation genes (Lawson et al., 2019). This underlines the necessity of further studies, such as the recent work expanding the understanding of genes responsible for HMO utilisation in commercial *B. infantis* strains (Duar et al., 2020).

Table 1.1. Genes characterised in papers cited in section 1.2.9. Taken from Masi and Stewart (2022) (Masi and Stewart, 2022).

	Bifidobacterium strain	Protein/enzyme type	Abbreviation or Gene locus	Enzymatic activity	Preferred HMO substrates	Reference
1	*B. bifidum* JCM1254	1,2-α-L-fucosidase	AfcA	Extracellular	2'FL, LNFP I, limited activity on 3FL and LNFP V	(Katayama et al., 2004)
2	*B. bifidum* JCM1254	1,3-1,4-α-L-fucosidase	AfcB	Extracellular	3FL, LNFP II, LNFP III	(Ashida et al., 2009)
3	*B. bifidum* JCM1254	Exo-α-sialidase	SiaBb2	Extracellular	3'SL, DSLNT, 6'SL	(Kiyohara et al., 2011)
4	*B. bifidum* JCM1254	β-galactosidase	BbgIII	Extracellular	LacNAc, LNnT, LNH	(Miwa et al., 2010)
5	*B. bifidum* JCM1254	Lacto-N-biosidase	LnbB	Extracellular	LNT, LNH	(Wada et al., 2008)
6	*B. bifidum* JCM1254	β-N-Acetylglucosaminidase	BbhI	Extracellular	LNTri	(Miwa et al., 2010)
7	*B. bifidum* JCM1254	GNB/LNB phosphorylase	LnpA1	Intracellular	LNB/GNB	(Nishimoto and Kitaoka, 2007b)
8	*B. bifidum* JCM1254	GNB/LNB phosphorylase	LnpA2	Intracellular	LNB/GNB	(Nishimoto and Kitaoka, 2007b)
9	*B. longum* subsp. *infantis* ATCC 15697	Transporter SBP	FL1-BP	-	2'FL	(Sakanaka et al., 2019b)
10	*B. longum* subsp. *infantis* ATCC 15697	Transporter SBP	FL2-BP	-	2'FL, 3FL, LDFT, LNFP I	(Sakanaka et al., 2019b)
11	*B. longum* subsp. *infantis* ATCC 15697	α-L-fucosidase	AfcA	Intracellular	LNFP I, 2'FL, 3FL	(Sela et al., 2012)
12	*B. longum* subsp. *infantis* ATCC 15697	1,3-1,4-α-L-fucosidase	AfcB	Intracellular	LNFP III, 3FL	(Sela et al., 2012)
13	*B. longum* subsp. *infantis* ATCC 15697	1,3-1,4-α-L-fucosidase	Blon_0248	Intracellular	LNFP III	(Sela et al., 2012)
14	*B. longum* subsp. *infantis* ATCC 15697	α-L-fucosidase	Blon_0426	Intracellular	LNFP III	(Sela et al., 2012)
15	*B. longum* subsp. *infantis* ATCC 15697	β-galactosidase	Bga2A	Intracellular	Lac, LacNAc, LNnT	(Yoshida et al., 2011)
16	*B. longum* subsp. *infantis* ATCC 15697	β-N-Acetylglucosaminidase	Blon_0459	Intracellular	LNT, LNH, LNTri	(Garrido et al., 2012)
17	*B. longum* subsp. *infantis* ATCC 15697	LNT β-1,3-Galactosidase	Bga42A	Intracellular	LNT, LNB	(Yoshida et al., 2011)
18	*B. longum* subsp. *infantis* ATCC 15697	β-N-Acetylglucosaminidase	Blon_0732	Intracellular	LNT, LNH, LNTri	(Garrido et al., 2012)
19	*B. longum* subsp. *infantis* ATCC 15697	β-N-Acetylglucosaminidase	Blon_2355	Intracellular	LNT, LNH, LNTri	(Garrido et al., 2012)
20	*B. breve* UCC2003	Transporter SBP	NahS	-	LNnT	(James et al., 2016)

	Bifidobacterium strain	Protein/enzyme type	Abbreviation or Gene locus	Enzymatic activity	Preferred HMO substrates	Reference
21	*B. breve* UCC2003	β-galactosidase	LntA	Intracellular	LNT, LNnT, Lac	(James et al., 2016)
22	*B. breve* UCC2003	β-galactosidase	LacZ2	Intracellular	LNnT, Lac	(James et al., 2016)
23	*B. breve* UCC2003	β-galactosidase	LacZ6	Intracellular	LNnT, Lac	(James et al., 2016)
24	*B. breve* UCC2003	β-N-Acetylglucosaminidase	NahA	Intracellular	LNTri	(James et al., 2016)
25	*B. breve* UCC2003	GNB/LNB phosphorylase	LnbP	Intracellular	LNB	(James et al., 2016)
26	*B. longum* subsp. *longum* JCM1217	Lacto-N-biosidase	LnbX	Extracellular	LNT, LNH, LNFP I, LST a	(Sakurama et al., 2013)
27	*B. longum* subsp. *longum* JCM1217	Chaperon for LnbX	LnbY	Extracellular	-	(Sakurama et al., 2013)
28	*B. longum* subsp. *longum* JCM1217	Transporter SBP	GL-BP	-	LNB, GNB	(Suzuki et al., 2008)
29	*B. longum* subsp. *longum* JCM1217	β-N-Acetylglucosaminidase	BLLJ_1391	Intracellular	LNTri	(Honda et al., 2013)
30	*B. longum* subsp. *longum* JCM1217	N-acetylhexosamine 1-kinase	NahK	Intracellular	GlcNAc/GalNAc	(Kitaoka et al., 2005)
31	*B. longum* subsp. *longum* JCM1217	GNB/LNB phosphorylase	LnpA	Intracellular	LNB/GNB	(Kitaoka et al., 2005)

HMO, human milk oligosaccharide; 2'FL, 2'-fucosyllactose; LNFP I, lacto-N-fucopentaose I; LNFP II, lacto-N-fucopentaose II; LNFP III, lacto-N-fucopentaose III; LNFP V, lacto-N-fucopentaose V; 3FL, 3-fucosyllactose; 3'SL, 3'-sialyllactose; DSLNT, disialyllacto-N-tetraose; 6'SL, 6'-sialyllactose; LacNAc, N-acetyllactosamine; LNnT, lacto-N-neotetraose; LNH, lacto-N-hexaose; LNT, lacto-N-tetraose; LNTri, ; LNB, lacto-N-biose; GNB, galacto-N-biose; LDFT, lactodifucotetraose; Lac, lactose; LST a, sialyllacto-N-tetraose a; GlcNAc, N-acetylglucosamine ; GalNAc, N-acetylgalactosamine.

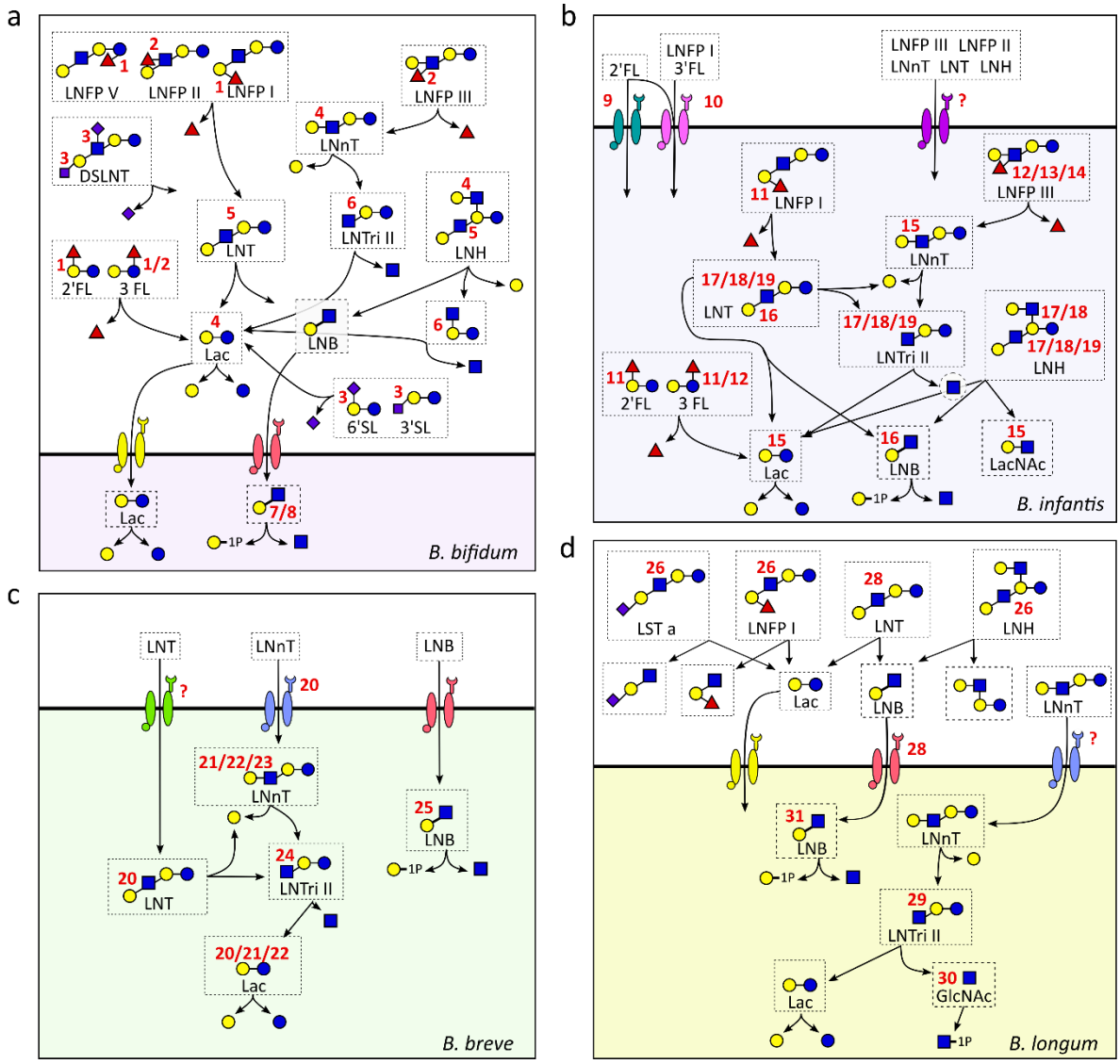

Figure 1.5. Schematic representation of the main human milk oligosaccharides metabolised by *B. bifidum* (a), *B. longum* subsp. *infantis* (b), *B. breve* (c), and *B. longum* subsp. *longum* (d). Proteins responsible for cleavage of the various bonds or transport of the oligosaccharides are represented by the red numbers; numbers refer to the gene list reported in Table 1.1. Where no protein harbouring the specific activity has been identified, a ‘‘?’’ has been added. In other instances where no red symbol is reported, the reaction is mediated by homologues of known genes. 2’FL, 2’- fucosyllactose; 3FL, 3-fucosyllactose; LNnT, lacto-N-neotetraose; 3’SL, 3’-sialyllactose; 6’SL, 6’-sialyllactose; LNT, lacto- N-tetraose; LNFP I, lacto-N-fucopentaose I; LNFP II, lacto-N-fucopentaose II; LNFP III, lacto-N-fucopentaose III; LNFP V, lacto-N-fucopentaose V; LST a, sialyllacto-N-tetraose a; LNH, lacto-N-hexaose; DSLNT, disialyllacto-N-tetraose; Lac, lactose; LNB, lacto-N-biose; LNTri II, lacto-N-triose II; LacNAc, N-acetyllactosamine; GlcNAc, N-acetylglucosamine. Taken from Masi and Stewart (2022) (Masi and Stewart, 2022).

1.2.10 Bacteroides

Bacterial species belonging to the *Bacteroides* genus are frequent colonisers of the term new-born and adult intestine (Shao et al., 2019, Stewart et al., 2018). Vaginal delivery seems to be

the key to colonisation of the infant gut by this genus and studies suggest the vertical transfer of *Bacteroides* spp. from the mother's gut microbiome during the delivery (Shao et al., 2019). *Bacteroides* contain specialised genes clustered in genomic regions referred to as polysaccharide utilisation loci (PUL) that provide capacity to break down various complex polysaccharides derived from the diet (Flint et al., 2012) and composing the fungal cell wall as well as host-derived glycans (Temple et al., 2017). Indeed, they can hydrolyse the intestinal mucus, which, being composed of Neu5Ac, Fuc, GlcNAc, Gal, is structurally similar to HMOs, placing *Bacteroides* in a favourable position for utilising milk glycans (Marcobal et al., 2011). Multiple studies have investigated the ability of *Bacteroides* spp. to utilise HMOs, albeit to a lesser extent compared to bifidobacteria. Strains of *B. fragilis*, *B. vulgatus*, *B. thetaiotaomicron*, and *B. caccae* were able to grow on pooled HMOs (Marcobal et al., 2010, Marcobal et al., 2011). *B. fragilis* grew better and consumed a higher proportion of HMOs when compared to *B. vulgatus*. The two *Bacteroides* spp. also differed in their preference for which HMOs were utilised, with *B. fragilis* preferring non-modified HMOs compared to fucosylated glycans, whereas no difference in HMOs preference was observed for *B. vulgatus*. Involving the same strains, Yu *et al*. (2013) reported *B. vulgatus* and *B. fragilis* could utilise 2'FL, 3FL, lactodifucotetraose (LDFT), and 6'-sialyllactose (6'SL); *B. vulgatus* could also metabolise 3'SL (Yu et al., 2013). Growth of a different strain of *B. thetaiotaomicron* was also investigated, and all HMOs mentioned with the exception of LDFT were degraded by this bacterium (Yu et al., 2013). Other different strains of *B. vulgatus*, *B. fragilis*, and *B. thetaiotaomicron* could grow on 2'FL, 3FL, difucosyllactose (DFLac), and Fuc (Salli et al., 2021). This latter *B. thetaiotaomicron* was demonstrated to partially utilise also LNT, LNnT, LNFP I, LNFP II, LNFP III, and LNFP V (Chia et al., 2020).

The genes involved in HMO utilisation in *Bacteroides* spp. were elucidated by Marcobal *et al*. (2011) (Marcobal et al., 2011). Both *B. fragilis* and *B. thetaiotaomicron* utilised PULs loci involved in mucin degradation, even though the genes activated in the two species were different, suggesting distinct strategies applied in different species. Although *B. thetaiotaomicron* could efficiently grow on HMOs, it was readily outcompeted by *B. infantis* in an *in vivo* mouse model fed with LNnT (Marcobal et al., 2011). These data suggest that, even though *Bacteroides* spp. are able to opportunistically hydrolyse HMOs thanks to their resemblance to mucin structure, bifidobacteria have evolved genes for the selective utilisation of glycan structures exclusive of human milk. This would explain why bifidobacteria can colonise the infant gut at higher relative abundances compared to other HMO utilising bacteria.

1.2.11 Lactobacillus and other gut commensals

Bifidobacterium and *Bacteroides* are not the only colonisers of the infant gut microbiome, and other genera frequently found in this niche include *Lactobacillus*, *Escherichia*, *Klebsiella*, *Enterococcus*, *Staphylococcus*, and *Clostridium* (Shao et al., 2019, Stewart et al., 2018). Utilisation of HMOs from these gut commensals has been reported, even though it is sporadic and often strain-specific (Hoeflinger et al., 2015, Marcobal et al., 2010, Salli et al., 2021, Schwab and Gänzle, 2011, Thongaram et al., 2017, Yu et al., 2013). In this regard, Lactobacilli are the most studied as they are often found in the infant developing intestinal microbiome even if at lower persistence and abundance compared to bifidobacteria (Shao et al., 2019, Stewart et al., 2018). Lactobacilli are also often included in infant probiotic formulations together with bifidobacteria and for this reason have received more attention compared to other gut colonisers.

Many infant gut commensals, considered non-HMO utilisers, have been tested for growth on these glycans showing weak or absent utilisation. Frequent HMOs used for growth curve experiments include pooled HMOs, 2'FL, 3FL, and LNnT. Moderate or little growth on pooled or singular HMOs was reported for strains of *Clostridium perfringens* (Marcobal et al., 2010, Salli et al., 2021), *Escherichia coli* (Salli et al., 2021, Marcobal et al., 2010), *Enterococcus faecalis* (Yu et al., 2013), *Enterobacter cloacae* (Yu et al., 2013), *Lactobacillus acidophilus* (Thongaram et al., 2017), *Lacticaseibacillus casei* (Salli et al., 2021), *Limosilactobacillus fermentum* (Salli et al., 2021), *Lactiplantibacillus plantarum* (Salli et al., 2021, Thongaram et al., 2017), *Lacticaseibacillus rhamnosus* (Salli et al., 2021, Yu et al., 2013), *Ligilactobacillus salivarius* (Salli et al., 2021), *Limosilactobacillus reuteri (*Thongaram et al., 2017), *Staphylococcus epidermidis* (Yu et al., 2013*)*, and *Streptococcus thermophilus* (Marcobal et al., 2010, Yu et al., 2013). Strains which cannot utilise HMOs, can often metabolise the building blocks composing their structure, including Fuc, GlcNAc, Neu5Ac, and other monosaccharides and disaccharides which can be released when these sugars are digested (Hoeflinger et al., 2015, Salli et al., 2021, Schwab and Gänzle, 2011, Thongaram et al., 2017). Thus, they might profit from *Bifidobacterium* and *Bacteroides* spp. breaking down the HMOs available and releasing the fragments in the environment. Many homologous of HMO utilisation genes have been identified when mining bacterial genomes deposited in the databases (Katayama et al., 2004, Kiyohara et al., 2011, Miwa et al., 2010, Sakurama et al., 2013); however, the enzymatic activity of such genes on HMOs is largely unexplored. Notably, the presence of genes for sugar degradation alone does not guarantee the ability to grow in the presence of HMOs. For instance, *L. casei* BL23 is equipped with an α-fucosidase which can act on 2'FL, but it is unable to grow

on this HMO (Rodríguez-Díaz et al., 2011). This inability to grow is likely because of the internal localisation of the enzyme and lack of a trans-membrane transporter for 2'FL, further highlighting the necessity to test *in vitro* the actual capability of the bacteria to grow on HMOs.

1.2.12 HMOs metabolism by-products

One of the mechanisms involved in the health-shaping effect exerted by the gut microbiota is through the interaction between microbial metabolites and the host. Some of these potentially beneficial metabolites are produced through the fermentation of diet indigestible carbohydrates in the intestine, including HMOs in the infant population (Stewart, 2021). SCFAs are among the most studied microbial metabolites, with particular attention given to acetate, butyrate, and propionate. SCFAs represent an energy source for colonocytes and can influence host physiology and immune system (Plaza-Díaz et al., 2018). The mechanisms mediating SCFAs beneficial effects include lowering the pH of the intestinal environment, thus preventing the growth of pathobionts (Alessandri et al., 2019, Zuurveld et al., 2020). SCFAs can also directly influence the intestinal health through increase in mucin production and enhancement of the barrier function at the level of intestinal cells, coupled with modulation of the immune system by promoting a GALT population enriched with regulatory T cells (Zuurveld et al., 2020, Alessandri et al., 2019). Partial absorption and release of SCFAs in the systemic circulation leads to a widespread effect of SCFAs in the human body, including modulation of glucose homeostasis, lipid metabolism, and appetite regulation (Morrison and Preston, 2016). Moreover, decreased colonisation by SCFA producing bacteria in early life has been associated with Type 1 diabetes (Vatanen et al., 2018) and allergy (Abrahamsson et al., 2012) onset later in life.

Intestinal SCFA composition evolves during the first months of life, likely mirroring the changes occurring in the gut microbial community. Although acetate is produced by most intestinal bacteria and is indeed the most abundant SCFA in the gut, butyrate and propionate can be produced only by a subpopulation of the gut colonisers (Koh et al., 2016, Morrison and Preston, 2016). Consequence of these differences, acetate can be found in detectable quantities since early life, whereas propionate and butyrate concentrations increase over the first year of life, in parallel with the colonisation of bacteria involved in their production (Appert et al., 2020, Nilsen et al., 2020, Tsukuda et al., 2021). Other less studied SCFAs associated molecules, lactate and succinate, and the SCFA formate, were instead higher in the first months of life and decreased until 1 year of age (Tsukuda et al., 2021). Metabolism of fucosylated HMOs by bifidobacteria has been reported to be central to formate production. Bifidobacteria are also

able to produce acetate, while they lack the pathway for butyrate and propionate synthesis (Louis and Flint, 2017, Rivière et al., 2016). However, bifidobacterial metabolism of Fuc is able to produce butyrate and propionate precursors which might be utilised by other gut commensals to produce these SCFAs (Bunesova et al., 2016). Work from Frese *et al*. (2017) supports the correlation between HMOs, *B. infantis* supplementation and SCFAs production (Frese et al., 2017). Breastfed infants supplemented with a commercial *B. infantis* strain showed increased quantities of lactate, acetate, butyrate, and formate, but not propionate, in the stool compared to non-supplemented breastfed infants (Frese et al., 2017). Similarly, in preterm infants supplemented with *B. bifidum* and *L. acidophilus*, an increase in stool acetate and lactate was found compared to non-supplemented preterm infants (Alcon-Giner et al., 2020).

1.2.13 Human milk oligosaccharides synthesis and supplementation

The many beneficial effects exerted by human milk through HMOs and other bioactive molecules are continuously being elucidated. However, for various reasons not all infants receive their mom's milk and providing the best nutritive formulations to these infants are of great importance. Owing to the current difficulties in synthetising HMOs at industrial levels (Zeuner et al., 2019), infant formulas are often supplemented with prebiotic molecules mimicking the human milk glycans, such as galacto-oligosaccharides and fructo-oligosaccharides (GOS and FOS, respectively) (Zuurveld et al., 2020). Like HMOs, GOS and FOS are indigestible to the infant, reach the infant's intestine intact, promote the growth of *Bifidobacterium*, can influence the immune system, and bind pathobionts (Zuurveld et al., 2020). However, they are not naturally present in human milk and thus past and current research aims to widen the range of HMOs available to supplement infant formulas to better mirror the human milk formulation.

Great advances have been, and continue to be, made in our understanding of HMOs synthesis, but only a few HMOs can currently be produced in quantities sufficient for their supplementation in infant formulas (Zeuner et al., 2019). The first randomised multicentre trial testing HMOs supplementation was performed in the USA using 2'FL. This trial did not show an impact on infant growth rate compared to non-supplemented formulas and breast milk fed infants (Marriage et al., 2015). Notably, 2'FL was well tolerated, and no adverse events were reported. In a subset of the cohort, plasma cytokines concentrations were also measured and infants receiving formula supplemented with 2'FL displayed lower pro-inflammatory cytokine and tumour necrosis factor (TNF) α concentrations in the plasma when compared to non-supplemented formula, whereas no difference was observed with respect to breastfed infants

(Goehring et al., 2016). In a different randomised trial conducted in Belgium and Italy, supplementation with 2'FL and LNnT again showed HMO supplementation was well-tolerated and did not alter the infant growth (Puccio et al., 2017). Administration of these HMOs also shifted the developing gut microbiome toward a composition more similar to breastfed infants, and decreased medication use was reported compared to non-supplemented formula fed infants (Berger et al., 2020). Very recently, a European trial involving multiple centres across different countries, reported the outcome of supplementation of 5 HMOs – 2'FL, 3FL, LNT, 3'SL, and 6'SL (Parschat et al., 2021). Overall, the 5-HMOs blend was well-tolerated by term infants enrolled and no difference in growth rates were observed when compared to non-supplemented formula (Parschat et al., 2021).

1.3 Diseases associated with preterm birth

Preterm birth is the leading cause of mortality in children under 5 years of age (Wolke et al., 2019). Death causes in very preterm infants arise from consequences of premature birth. Indeed, preterm infants are at high risk of developing many diseases, including NEC, LOS and focal intestinal perforation (FIP) which will be discussed in the next paragraphs.

1.3.1 Necrotising enterocolitis

NEC is a life-threatening intestinal disease which affects 5-10% of preterm infants born <32 weeks gestation (Embleton et al., 2017). The mortality can reach 30% and an important morbidity might affect the survivors, causing neurodevelopmental delays and short bowel syndrome when surgery is necessary (Pammi et al., 2017). NEC is characterised by severe inflammation, pneumatosis intestinalis and intestinal necrosis, and the most common symptoms comprise feeding intolerance, abdominal distention and bloody stool (Neu and Walker, 2011). Due to the absence of any specific symptoms, NEC remains hard to diagnose and likely represents the final stage of different pathologies, which might require different therapeutic strategies (Neu and Walker, 2011). Although no accurate diagnostic benchmarks are available yet, Bell Staging criteria proposed by Dr Bell in 1978 to describe NEC severity with some modifications are still in use in some countries, although this is now rarely used in the UK (Bell et al., 1978, Gregory et al., 2011). Stage I describes cases of suspected NEC, and symptoms include lethargy, apnoea and temperature instability. Stage II is the subsequent step, and includes proven NEC cases, with signs of pneumatosis intestinalis. The last stage, Bell III or advanced NEC, is characterised by most of the symptoms associated to the pathology, including intestinal bleeding or septic shock (Gregory et al., 2011).

In case of suspected NEC, feed is usually slowed or completely stopped, and replaced by intravenous hyperalimentation. When NEC is diagnosed, diseased infants can be treated with a medical approach, which includes broad-spectrum antibiotics in addition to changes in feeding mode. For more severe cases, surgery procedures might be needed, including resection of the diseased bowel or ostomy (Gregory et al., 2011, Neu and Walker, 2011). Around 20-40% of NEC cases are estimated to need surgical intervention, with a mortality rate of 50%, and many morbidities in the survivors, including short bowel syndrome (Gregory et al., 2011).

1.3.2 Preterm gut microbiome in NEC

The role of microbiome in NEC onset is an active area of research. Although NEC pathophysiology is still poorly understood, low bacterial diversity, prolonged use of antibiotics, and immaturity of the intestine and immune system are among the factors associated with disease development (Berrington et al., 2013) (Figure 1.6). The importance of the microbiome in NEC has been demonstrated circumstantially in germ free mice being unable to develop the disease and NEC occurring after the initial colonisation of the intestine (Pammi et al., 2017, Neu and Walker, 2011). Despite these observations, no specific microorganism has been consistently associated with NEC and many organisms which appear to be implicated in disease development can also be found to dominate the gut microbiome in heathy infants (Stewart et al., 2012, Stewart et al., 2013b, Sim et al., 2015, Abdulkadir et al., 2016). On the other hand, recent work has shown that preterm infants with high microbial diversity and an abundance of *Bifidobacterium* spp. in the gut may be protected from NEC onset, but whether this is cause or effect has not been delineated (Stewart et al., 2016).

One of the first studies sequencing the stool microbiome in preterm infants suggested a low microbial diversity coupled with the predominance of single genus belonging to Pseudomonadota phylum could be the key difference leading to NEC onset (Wang et al., 2009). Indeed, the higher consensus in the literature is observed at phylum level, with an overrepresentation of Pseudomonadota in NEC infants. Larger cohort studies employing 16S rRNA gene sequencing have corroborated that Pseudomonadota is higher in NEC infants prior to disease diagnosis, including a study by Warner *et al*., which included 2492 stool samples from 122 infants, of whom 28 developed NEC (Warner et al., 2016). They found higher proportions of Gammaproteobacteria (Pseudomonadota phylum) and lower proportions of Clostridia (Bacillota phylum) and Negativicutes (Bacillota phylum) in infants who developed the disease compared to the healthy counterpart (Warner et al., 2016). The difference was mostly attributable to cases of late onset NEC occurring after day 30 of life, which is after the

typical median age of NEC onset (day 20-30 of life). In another longitudinal study, Stewart *et al.* (2016) found that high microbiome instability and lack of microbial diversity were associated to NEC onset. In accordance with literature, no specific microbial composition was associated to disease in this cohort, but a diverse microbiome rich in *Bifidobacterium* spp. was found only in healthy babies (Stewart et al., 2016). A large meta-analysis of publicly available 16S rRNA gene sequencing of stool samples also found higher Pseudomonadota in NEC, but no significant difference in the bacterial genera (Pammi et al., 2017). However, it is notable that 48% of the samples included in this meta-analysis were obtained solely from Warner *et al*. study (Warner et al., 2016). Supporting case/control cohorts, studies comparing the gut microbiome of twins discording for NEC onset have also shown a less diverse microbiome and an increased abundance of Pseudomonadota in the twin who developed NEC (Wang et al., 2009, Hourigan et al., 2016). In addition to stool-based studies, a study which employed formalin fixed paraffin embedded (FFPE) tissue from the site of disease also found higher relative abundance of Pseudomonadota, lower Bacillota and Bacteroidota, and lower microbial diversity in NEC infants compared to spontaneous isolated perforation controls (Stewart et al., 2019).

It is not clear if a higher relative abundance of Pseudomonadota is a cause or a consequence of NEC. Hyperresponsiveness of the immune system mediated by toll-like receptor 4 (TLR4) induction has been proposed as one of the potential pathways inducing NEC (Leaphart et al., 2007). Indeed, TLR4 is over-expressed in the premature infant gut, and it is involved in maintaining the balance between injury and repair of the intestine (Neu and Walker, 2011). In a mouse model, the developing intestine harboured a high expression level of TLR4, which decreased at term (Gribar et al., 2009). Bacterial species within the Pseudomonadota phylum are recognised by the host through the interaction of lipopolysaccharide (LPS) with these specific receptors. Thus, an increase in the abundance of Pseudomonadota could trigger an excessive immune response in the preterm gut, leading to intestinal injury and necrosis. Moreover, LPS together with other stimuli, such as hypoxia and jejunal distention, can induce higher expression of TLR4 in rats (Luo et al., 2012, Chan et al., 2009). LPS exposure alone could also trigger TLR4 overexpression, although conflicting results are reported, requiring further investigation (Chan et al., 2009, Guo et al., 2013).

16S rRNA gene sequencing has been the most utilised sequencing technique to study association between microbiome and NEC onset. Reflecting the complexity of the disease, studies implying different cohorts usually report conflicting results, and application of metagenome analysis could profoundly improve our understanding of the role played by intestinal microorganisms. A recent study analysed longitudinal faecal samples (n = 1163) from

34 NEC babies and 126 matched healthy infants through metagenomic sequencing (Olm et al., 2019). The study strengthens the idea that no single causative microorganism associates with disease development and reported also that no differences were observed in microbiome composition. However, they found higher bacterial replication rates in samples collected pre-NEC, which might trigger preterm immune response and inflammation. In particular, the genus *Klebsiella* and its secondary metabolites interacting with TLR4 (LPS, fimbriae) were associated with faecal samples prior to NEC onset, corroborating the hypothesis of TLR4 as key in disease onset (Olm et al., 2019).

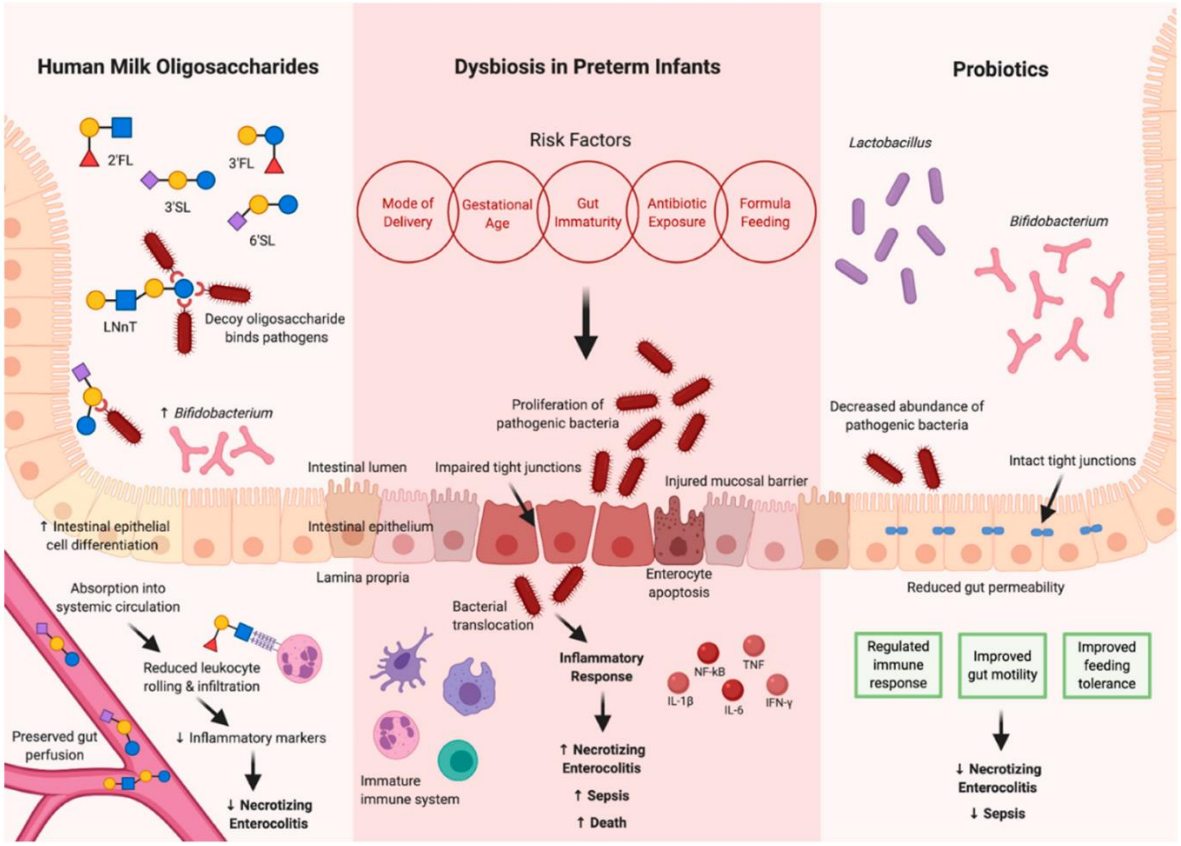

Figure 1.6. Impact of human milk oligosaccharides, gut dysbiosis and probiotics on the preterm intestinal barrier function. Taken from Nolan *et al.* (2020) (Nolan et al., 2020).

1.3.3 Breast milk role in necrotising enterocolitis

Breast milk feeding is included among the factors associated with protection against NEC development in preterm infants. Indeed, breast milk feeding brings to a 6 to 10-fold reduction of NEC development in a dose related manner, either when used exclusively or in conjunction with other types of feeding (Meinzen-Derr et al., 2009). The mechanisms by which breast milk exerts its defensive action in NEC are still unknown and likely numerous. Exclusively breast milk fed infants can still develop NEC, leading to the hypothesis that variation in human milk composition might be linked to disease onset.

HMOs are thought to play a major role, as they are absent in infant formulas, are associated with colonisation by potentially beneficial bacteria, modulate epithelial cells' gene expression and might regulate the immune system's inflammatory response. One specific HMO, disialyllacto-N-tetraose (DSLNT), is attracting attention since it prevents NEC onset in a neonatal rat model (Jantscher-Krenn et al., 2012). DSLNT is an acidic HMO derived by the elongation of the lactose molecule with LNB through β1-3 linkage; LNB is further modified with addition of two Neu5Ac on Gal and GlcNAc through α2-3 and α2-6 linkages respectively. Jantscher-Krenn *et al*. (2012) reported that DSLNT alone could explain the protective effect driven by HMOs in a neonatal rat model, and this action was strictly dependent on its molecular structure (Jantscher-Krenn et al., 2012). Later work from the same research group found that DSLNT concentration was lower in human breast milk of mothers whose infants developed NEC, when compared to healthy mother-infant pairs, with DSLNT deficiency being higher in more severe NEC cases respect to moderate ones (Autran et al., 2018). The concentration of no other HMO was altered. Interestingly, similar results were also reported by Hassinger's group (2020) (Hassinger et al., 2020). In this cohort, breast milk from mothers whose babies developed NEC showed DSLNT z-score below the age specific mean, even though no difference in DSLNT concentration between diseased and healthy babies was found, potentially due to small patient numbers (48 mothers, 5 NEC babies) (Hassinger et al., 2020). Low HMOs diversity has also been linked to NEC affected babies in an independent study (Wejryd et al., 2018), although other groups did not report any such association (Autran et al., 2018, Hassinger et al., 2020). Despite studies reporting differences in HMO profile in NEC infants, the potential implications of the breast milk microbiome have been poorly studied to date. For instance, Stewart *et al*. (2013) investigated the correlation between twins microbiome and their mom's breast milk microbiota (Stewart et al., 2013a). This work included twins discordant for NEC development but lacked power to perform a direct investigation of MOM microbiome in NEC development. Thus, further studies are needed.

1.3.4 Late onset sepsis

Neonatal sepsis is an infectious condition characterised by the presence of bacteria, fungi or viruses in the bloodstream. In neonatal patients, this pathology can be classified as early-onset sepsis (EOS) or LOS, depending on the age of the infant at the time of clinical manifestation (Shane et al., 2017). EOS occurs during the first 72 h of life and is likely caused by direct transmission of pathogens from the mother to the infant, before or during delivery (Shane et al., 2017). Instead, LOS comprises sepsis episodes occurring after 72 h of life and the source of

harmful bacteria is represented by hospital personal and environment, or the microbiome (Sharma et al., 2018). It has been estimated that ~21% of very low birth weight (VLBW) infants (i.e, born weighing <1500g) develop at least one episode of LOS confirmed by blood culture, with high rate of mortality and severe morbidity, including poor neurodevelopmental outcomes (Camacho-Gonzalez et al., 2013). Neonatal sepsis is hard to diagnose due to lack of specific clinical manifestations. Frequent symptoms include fever, pneumonia, apnoea, lethargy, poor feeding and bleeding manifestation, but all of them are common to other neonatal conditions making sepsis difficult to recognise (Camacho-Gonzalez et al., 2013, Sharma et al., 2018). The "gold standard" for diagnosis is the isolation of the causative microorganism from blood culture. However, false negatives may occur, and clinical signs are not correlated with a positive blood culture result, underlining the importance of seeking novel and effective biomarkers (Camacho-Gonzalez et al., 2013, Berrington et al., 2013).

The skin and gut epithelium provide protective barriers, but preterm neonates are underdeveloped and prone to leaky gut, increasing the risk for translocation of microorganisms in the blood stream (Shaw et al., 2015). Along with an immature immune system and immature intestine, other risk factors for LOS comprise a status of dysbiosis, prolonged use of central catheters, delayed initiation of enteral feeding and increased duration of ventilator support (Stoll et al., 2002). Instead, as observed with NEC, breast milk feeding associates with a reduction in risk to develop LOS (Patel et al., 2013). The main microorganisms causing LOS in preterm infants are Gram-negative enteric bacteria including *Klebsiella* spp., *Pseudomonas* spp., *Escherichia coli* and Gram-positive bacteria including *Streptococcus* spp., *Enterococcus* spp., and coagulative-negative Staphylococci (CoNS), with the latter causing the highest rate of neonatal sepsis episodes (Shaw et al., 2015, Mai et al., 2013). Overall, an aberrant gut microbiome development characterised by low bacterial diversity, higher Pseudomonadota and Bacillota, and delayed colonisation by obligate anaerobic species has been linked to LOS onset (Shaw et al., 2015, Mai et al., 2013, Taft et al., 2015, Stewart et al., 2017b). In addition, some anaerobic bacteria (e.g., bifidobacteria) have been associated with increased intestinal barrier function (Maier et al., 2014, Ling et al., 2016). A longitudinal analysis involving regular stool sampling from birth also demonstrated that in most cases of LOS the causative organism was abundant in the gut at the time of diagnosis, often reducing following antibiotic treatment (Stewart et al., 2017b). While LOS was traditionally thought to be a consequence of microbial translocation from the skin, new evidence suggest that this group of microorganisms might also be of intestinal origin (Stewart et al., 2017b, Shaw et al., 2015, Mai et al., 2013, Taft et al., 2015).

1.3.5 Focal intestinal perforation

Together with NEC, FIP is a common form of intestinal injury occurring in premature infants (Gordon, 2009). NEC and FIP present with similar characteristics and symptoms and misdiagnosis is common (Berrington and Embleton, 2021), however they are two distinct pathologies. While NEC starts in the small or large bowel but then usually affects other organs too, FIP is restricted to a limited area surrounded by healthy tissue and does not present ischemic damage (Gordon, 2009). Moreover, FIP tends to occur earlier in life than NEC, even though there is a window of time where an overlap is observed (Berrington and Embleton, 2021). Due to the lower mortality and morbidity of FIP compared to NEC, FIP is less studied and, to the best of my knowledge, the potential implication of HMOs and the gut microbiome on disease onset have not been investigated yet. A microbiome study utilising FFPE tissue from NEC and FIP patients suggests that differences in the microbial community at site of disease exist between the two pathologies, with Pseudomonadota being high in NEC and Bacteroidota being higher in FIP (Stewart et al., 2019).

1.4 Human intestinal organoids

1.4.1 Models to study the microbiome-host interaction in preterm infants

Since the early 2000s, the focus on preterm microbiome has increased exponentially, including studies on its link with NEC. Microbiome-host interaction studies are crucial for the characterisation of microorganisms. This type of experiment is typically performed involving animal models or tumour-derived cell line cultures (e.g., Caco-2 cells), but they cannot effectively model the preterm human intestine. Animal models allow investigation of microorganisms on an experimental model comprising all the tissues and organ systems of a body, for instance circulating immune system and crosstalk between cells and organs. However, animal models cannot completely recapitulate human prematurity and each animal species harbours a distinct enteric microbiome, making it difficult to translate findings into humans. Various animal models have been developed over the years to study NEC, including mice and piglet models. Pups are usually delivered prematurely by C-section and NEC-like injury is induced applying stress conditions, which vary among different models developed (Sulistyo et al., 2018). Epithelial-cell lines such as Caco-2 cells provide a simple model composed by a single cell type, intestinal epithelial cells, and are usually derived by tumour cells, thus they cannot reproduce intestine prematurity and carry alterations in the gene and protein expression when compared to healthy intestinal cells. To overcome these potential limitations, the use of

human intestine-derived organoids (HIOs) is gaining popularity and could improve our understanding of microbiome-host crosstalk.

1.4.2 Intestinal organoids: 3D and 2D form

With the first reported long-term intestinal organoid culture dating to 2009, HIOs are generated from resected intestinal tissue by stimulation of Lgr5+ stem cells found in the crypts (Sato et al., 2009). The epithelium in the small intestine can be divided into two areas: the villus, which is composed by both secretory and absorptive cells and protrudes into the lumen, and the crypts, where the stem and Paneth cells reside (Leushacke and Barker, 2014) (Figure 1.7a). The villus is composed mainly of enterocytes, which connect the luminal compartment to the portal circulation by secretion and absorbance of various substances. Surrounded by enterocytes, we also find enteroendocrine cells, secreting neuroendocrine factors, and goblet cells, responsible of mucus production and secretion. In the crypt, Paneth cells secrete anti-microbial factors, while stem cells proliferate, differentiate and migrate to renovate the villus compartment (Blutt et al., 2018).

Owing to their multipotent nature, stem cell expressing Lgr5 gene are able to generate all cell types composing the intestinal epithelium (Sato et al., 2009). *In vitro*, these Lgr5+ cells can self-assemble forming HIOs, when supported by an extracellular matrix surrogate (e.g., Matrigel) and stimulated with specific growth factors promoting stem cell proliferation (e.g., Wnt-3a, R-spondin and Noggin). If this mix of factors is removed, stem cells will stop proliferating and will start to differentiate in the cell types composing the villus (Blutt et al., 2018). HIOs remain morphologically representative of the segments of intestine they were taken from (e.g., ileum is distinct from colon), are composed of all the major cell types of the intestinal epithelium (enterocyte, Paneth, goblet, enteroendocrine, and stem), and are physiologically active (e.g., secrete mucus) (Blutt et al., 2018). Middendorp *et al.* (2013) investigated gene expression profile of organoids derived from various regions of the small intestine (Middendorp et al., 2014). Initially, location-specific genes were determined by mRNA sequencing of crypts and villi isolated by duodenum, jejunum and ileum. Subsequently, organoids were generated from crypts isolated from each intestinal compartment and their gene-expression profile was determined by mRNA sequencing. By comparing expression profile of each region-specific organoid with corresponding location-specific gene profile, organoids demonstrated the ability to retain part of the location-specific genes, from a minimum of 49% up to 87% (Middendorp et al., 2014).

HIOs represent an innovative *ex vivo* model which allows to investigate the response of the intestine to microorganisms. However, because HIOs self-polarise the villi protrude towards the centre of the organoid (Figure 1.7a), which can therefore be identified as the luminal compartment, where the interaction with the microbiome takes place *in vivo*. This introduces challenges in co-culturing with bacteria, and to effectively recapitulate the *in vivo* scenario techniques such as micro-injection are required to add microorganisms to the "lumen" of the intestinal organoid. Overcoming this, an alternative approach to using 3D HIOs involves dissociating 3D structures into 2D monolayers of cells derived from HIOs, allowing the interaction between microorganisms and intestinal cells to be studied. HIOs are disrupted to single cells which are then seeded onto Transwells and, once confluent, are differentiated by removing stemness factors from the media (Figure 1.7b). The monolayer of cells obtained will then recapitulate the villus composition, with the presence of enterocytes, goblet cells, enteroendocrine cells, Paneth cells, and others. Moreover, each HIO line retains the patient genetic information, allowing for the interaction between gut microbes and human epithelial cells from multiple individuals (i.e., not limited to a single cancerous cell line such as Caco-2) to be mechanistically investigated, thus taking into account possible patient genetic contributions.

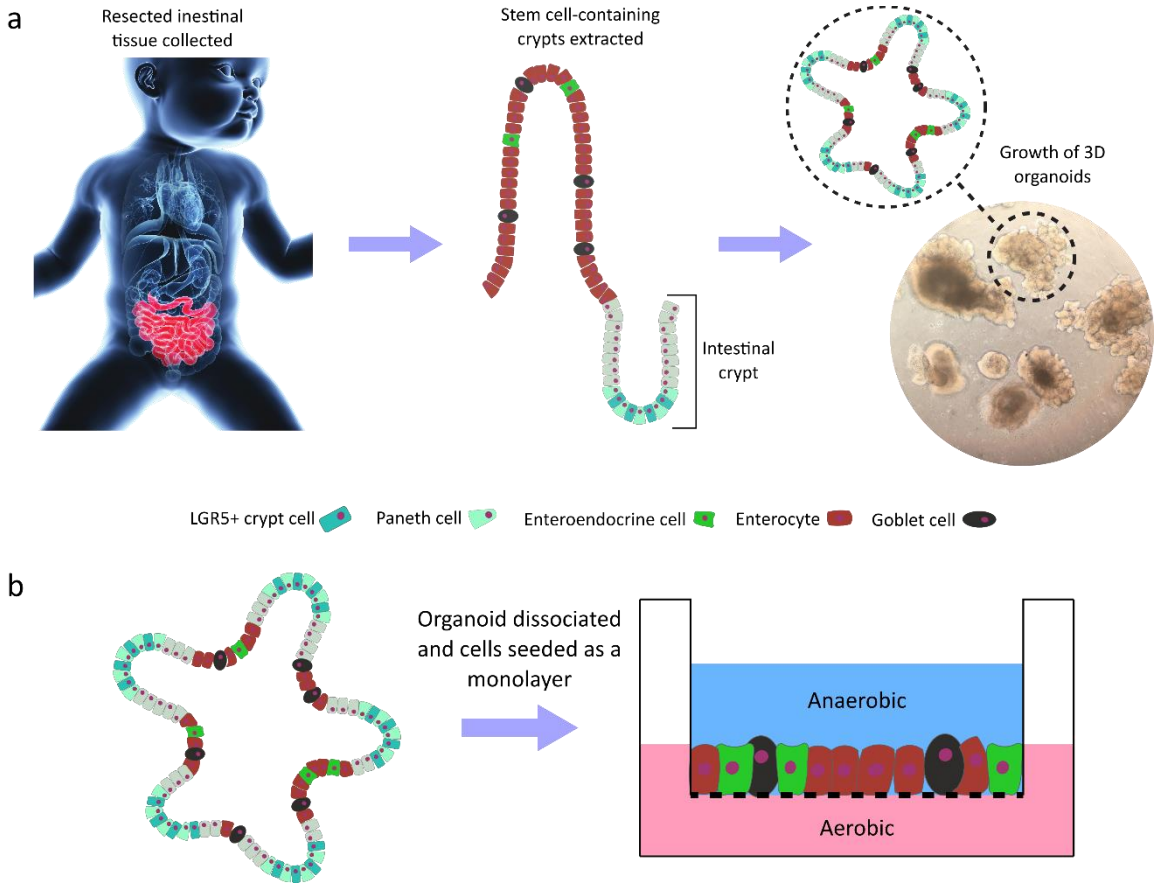

Figure 1.7. Representation of intestinal epithelium structure in-vivo, of ex-vivo HIOs and of organoid derived monolayers. (a) Small intestine epithelium structure. Lgr5+ stem cells

can be found in the crypt-compartment. The crypt compartment can be isolated and used to establish intestinal organoids: a villus-like composition faces the internal compartment of the organoid, while crypt-like structures point outwards. (b) Schematic representation of the establishment of intestinal organoid derived monolayers onto Transwell inserts. Image taken from Chapman and Stewart (2023) (Chapman and Stewart, 2023).

1.4.3 Bacteria-organoid co-culture systems

Intestinal cells *in vivo* are exposed to an oxygen level ranging from 1 to 11% (situation referred to as physiological hypoxia or "physioxia"), while incubation conditions normally used for cell cultures (i.e., 5% CO_2) present an estimated oxygen saturation of 17%. These high oxygen levels alter the cell gene expression, making the classical incubator environment not well representative of physiology of the intestine (Fofanova et al., 2019). A step forward in this field has been made with the development of a co-culture system (organoid-anaerobe co-culture system, OACC system) which allows the characterisation of anaerobic microorganisms belonging to the gut microbiome (Fofanova et al., 2019). In the OACC system, HIO cells are supplied of physiological concentration of oxygen (5.6%) from the basolateral side of the Transwell (Figure 1.8). The oxygen provided is entirely consumed by the monolayer, thus maintaining the apical compartment anaerobic and suitable for intestinal microorganisms to grow (e.g., *Bifidobacterium* spp.) (Fofanova et al., 2019, Jalili-Firoozinezhad et al., 2019). Host-microbial interaction can be characterised by applying multiple assays, e.g., measuring epithelial integrity, bacterial translocation, gene expression levels, monolayer histology, and production of metabolites and proteins.

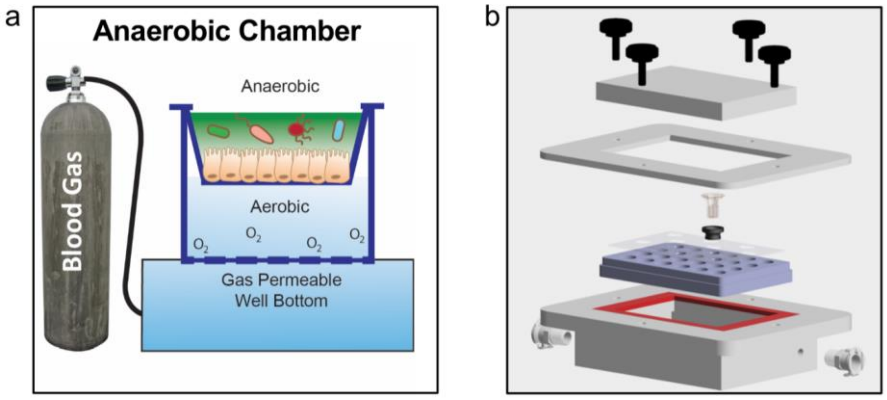

Figure 1.8. Representation of the organoids-anaerobe co-culture system. (a) Schematic representation of the system at the single well level. (b) Transwells are seeded with intestinal organoids monolayers and sealed onto a 24-well gas permeable plate thanks to special gaskets and double-sided sticky tape. The system is maintained in an anaerobic chamber, oxygen is provided through the base of the system and the apical compartment can stay anaerobic. Taken from Fofanova *et al.* (2019) (Fofanova et al., 2019).

Additional systems built to study host-microbiome interaction in the gut have been developed to date, each with its own strengths and drawbacks. The HoxBan system couples microbes

grown in solid agar and interacting with intestinal cells grown on coverslips and placed upside-down on the agar (Sadaghian Sadabad et al., 2015). No specialist equipment is needed, and it is a low-cost approach, however, no customisation of the system is possible, and bacteria are not free to move in the media. Setups similar to the OACC system described above have also been reported, however, no manipulation of the oxygen level supplemented to the intestinal cells is possible (Ulluwishewa et al., 2015, Sasaki et al., 2020). While the systems described so far use intestinal cell monolayers, more complex systems have also been developed to incorporate measurement of oxygen and trans-epithelial electrical resistance (TEER) (Shah et al., 2016), or to better mimic the villus like structure (Lanik et al., 2020). However, they do not allow direct interaction between bacteria and host cells and mucus is incorporated artificially (Shah et al., 2016), they cannot maintain the luminal compartment anaerobic (Lanik et al., 2020), and they are custom built and not easily scalable (Shah et al., 2016, Lanik et al., 2020). For these reasons, the OACC system represents a valid approach to study the interaction between the preterm intestine and gut bacteria in a physiologically relevant oxygen environment.

1.5 Aims

The aims of this study included addressing various aspects of diet-microbiome-host interaction in preterm infants at risk of necrotising enterocolitis. The overarching aims were:

- Identify specific HMOs or microbial species, from both infant intestine and breast milk microbiome, associated with NEC or health status.
- Investigate the potential of bacteria isolated from preterm infants to metabolise specific HMOs.
- Characterise how the preterm intestinal epithelium responds to direct HMOs exposure using preterm intestinal organoids.

Chapter 2. Materials and Methods

2.1 Cohort description

2.1.1 Ethics and samples collection

The preterm infants (born at <32 weeks gestation) involved in this study were born in or transferred to a single large tertiary level NICU in Newcastle upon Tyne, UK, recruited to the Supporting Enhanced Research in Vulnerable Infants (SERVIS) study (REC10/H0908/39) with written parental consent covering data and sample collection. A NICU classified as level 3 treats babies who are the sickest as it can provide the whole range of neonatal care. It also offers its services to babies and families who are recommended by the neonatal network in addition to the local community.

Stool samples were collected from the nappy by the nurses and stored in sterile universals. Breast milk samples were either collected fresh after expression or at the end of the infant's feeding from the tube used for the process, which can last up to 24 hours. Where the sample was collected after feeding, the day of the MOM sample reflects the day the infant received the milk and does not necessarily reflect the day milk was expressed due to standard practice that often involves milk storage. After collection for research, samples were stored in sterile tubes. All samples were stored in the NICU at -20°C between one week up to six months, before being transferred and stored at -80°C until the sample was analysed. The samples included were collected from babies born between December 2014 and January 2021.

2.1.2 Population description and clinical data availability

Diagnoses were made using an extensive combination of clinical, x-ray and histological findings and blindly agreed by two neonatal clinicians, Dr Janet E Berrington and Dr Nicholas D Embleton. Clinical variables including gestational age, birthweight and sex were available together with detailed information on type of feeding and antibiotics administration. For breast milk feeding, formula feeding, breast milk fortifier and probiotics administration, day of life of start and end were recorded. Detailed information on day of start and end and type of antibiotics administered was also available.

Standard clinical protocols recommended the routine use of supplemental probiotics when more than 30ml/kg/day of MOM was tolerated for at least 1-2 days. The number of infants receiving probiotics varied between chapters and the relevant information relating to probiotic use is

described in each chapter where appropriate. The probiotics administered were either Labinic (*L. acidophilus, B. infantis* and *B. bifidum*) or Infloran (*L. acidophilus* and *B. bifidum*).

2.2 Human milk oligosaccharides analysis

2.2.1 Human milk oligosaccharides analysis

The absolute quantification for the 19 most abundant HMOs was determined by HPLC following derivatisation as per the protocol described by Bode *et al*. (Bode et al., 2012). Raffinose was added to every sample before analysis to work as internal standard. Lipids, proteins, lactose, salts, and peptides were removed by stepwise solid-phase extraction. HMOs were labelled by adding the fluorescent tag 2-aminobenzamide to the reducing end and subsequently analysed by HPLC on an amide-80 column. The HMOs quantified account for >95% of total HMOs and included: 2'FL, 3FL, LNnT, 3'SL, DFlac, 6'SL, LNT, LNFP I, LNFP II, LNFP III, LSTb, LSTc, difucosyl-LNT (DFLNT), lacto-N-hexaose (LNH), DSLNT, fucosyllacto-N-hexaose (FLNH), difucosyllacto-N-hexaose (DFLNH), fucosyldisialyllacto-N-hexaose (FDSLNH) and disialyllalcto-N-hexaose (DSLNH). Maternal secretor (presence of an active FUT2 gene) status was determined by presence or near-absence of 2'FL in the breast milk analysed.

2.2.2 Statistical analysis

Statistical analysis of cross-sectional HMO profiles was performed using MetaboAnalyst V.3.0.28. Orthogonal Partial Least Squares - Discriminant Analysis (OPLS-DA) and Partial Least Squares - Discriminant Analysis (PLS-DA), for 2 or more group comparison respectively, were performed on HMO data normalised by logarithmic transformation and 2000 random permutations were used to test the significance of group separation. HMO Shannon diversity was calculated using "vegan" (version 2.5-6) package (Jari Oksanen, 2019) in R (version 3.6.3). Wilcoxon rank-sum test or Kruskal Wallis test were used for variables comparison between two or more groups, respectively. P values were adjusted for multiple comparisons by applying the Benjamini & Hochberg correction (Benjamini and Hochberg, 1995). Variables with >2 groups deemed significant with Kruskal-Wallis underwent Dunn's post-hoc test to determine P values specific to each group comparisons and resulting P values were adjusted applying Bonferroni method (Dunn, 1961).

To test the potential role of individual HMOs as biomarker for disease development, univariate receiver operating characteristic (ROC) curve analysis was performed, and an optimal cut-off was defined by the closest point of the curve to the top-left corner. Multivariate ROC curves

were also generated using linear Support Vector Machine (SVM) classification method coupled with Monte-Carlo cross validation (MCCV) to test the classification performance obtained by using model iterations containing increasing numbers of features: 2, 3, 5, 7, 10, or 19 HMOs. In each MCCV step, two thirds of the dataset were used to determine feature importance and classification model performance was evaluated with the remaining third of the samples which was left out.

Correlation between clinical variables and individual HMOs was tested by performing a multivariate adjusted linear model in R (version 3.6.3). HMO concentrations were normalised by log-transformation prior to analysis and P values were adjusted applying the Benjamini & Hochberg correction (Benjamini and Hochberg, 1995). For all analysis, P values <0.05 were considered significant.

2.3 Generation of pre-existing metagenomic dataset from stool

At the time of starting my PhD, the Stewart Lab had recently acquired metagenomic data from preterm infant stool samples through a collaboration with Astarte Medical. Where infants with stool metagenome overlapped with the breast milk samples from my work, I was able to use this dataset for my analysis. This included 644 stool samples from 34 controls and 14 infants with NEC. DNA was extracted from infant stool samples for downstream analysis using the DNeasy PowerSoil Kit (QIAGEN). Extraction was performed on ~0.1g of stool following the manufacturer's instructions. The bead beating step was performed by vortexing the samples at maximum speed for 20 mins using a Vortex Adapter tube holder and the subsequent steps were performed according to the protocol. The DNA was eluted in 100 µl of C6 solution and the samples were stored at -80°C. In addition to stool samples, extraction was performed on a positive (Zymo Microbial Community Standard) and negative (no sample at all) controls. A negative control was extracted in every batch of 24 samples.

Library prep was performed using the Nextera DNA Flex Kit. Sequencing was performed on the HiSeq X Ten (Illumina) with a target read depth of 10M reads per sample with a read length of 150bp paired end reads. Raw fastq files were quality trimmed and Illumina adapters removed using bbduk (BBMap version 38.69) (Truong et al., 2015). Trimming parameters included kmer length of 19, allowing one mismatch, and a minimum Phred score of 20. Post-trimming, reads with a minimum average Phred <17 and length <50 bp were discarded. Host contamination reads were identified by mapping trimmed fastq files to a combined database containing the hg38 reference human genome and PhiX (standard Illumina spike in) using bbmap (BBMap version 37.58) (Truong et al., 2015) with kmer length of 15, bloom filter enabled, and fast

search settings. Host reads were subsequently removed, and the remaining processed fastq files were mapped against the MetaPhlan2 marker gene database (mpa_v20_m200) using bbmap with the bloom filter enabled and fast search settings (Truong et al., 2015). Finally, the metaphlan.py script was used to generate kingdom specific taxonomic profiles.

2.4 DNA extraction and 16S rRNA gene sequencing of mother's own milk

DNA was extracted from 130 MOM samples for downstream analysis using the DNeasy PowerLyzer PowerSoil Kit (QIAGEN) following the manufacturer's instructions with some modifications. 400 µl of MOM sample were added to the PowerBead tube and mixed with 400 µl of PowerBead solution and 60 µl of solution C1. Where 400 µl of samples were not available, a sample volume as low as 150 µl was used. The bead beating step was performed by vortexing the samples at maximum speed for 20 mins using a Vortex Adapter tube holder. The tubes were subsequently centrifuged at 10.000 g for 1 minute and the supernatant moved to a clean tube. Solutions C2 and C3 were mixed in 1:1 proportion, and 200 µl of the mix were added before a unique 5 mins incubation at 4 °C. The subsequent steps were performed as per the protocol instructions. At last, 60 µl of Solution C6 were added for the elution step; the samples were left 5 min at room temperature and then centrifuged at 10.000 g for 1 min, the column discarded, and the samples stored at -80 °C. A negative control was extracted in every batch of 24 samples. Polymerase chain reaction (PCR) was used to amplify the V4 region of the 16S rRNA gene using the barcoded Illumina adapter-containing primers 515F and 806R (Caporaso et al., 2012). Sequencing was performed on the MiSeq platform (Illumina), with a target read depth of 10k and a read length of 250 bp paired end. Raw fastq files were demultiplexed using the Illumina 'blc2fastq' software, followed by quality trimming and Illumina adapters and PhiX reads removal using bbduk (BBMap version 38.82) (Truong et al., 2015). Reads with a Phred quality score below 15 and length below 100 bp after trimming were discarded. Reads are then merged using bbmerge (BBMap version 38.82) (Truong et al., 2015), with subsequent further filtering using 'vsearch' (Rognes et al., 2016) setting the maximum expected error of 0.05, maximum length of 254 bp and minimum length of 252 bp. Using the UPARSE algorithm (Edgar, 2013), sequences were clustered into Operational Taxonomic Units (OTUs) applying a similarity cut-off of 97% and using a stepwise approach. USEARCH and UCHIME (Edgar et al., 2011) were then used to remove chimeras, and USEARCH (Edgar, 2010) was used to determine taxonomies by mapping the OTUs against a version of the SILVA Database (Quast et al., 2013) containing only the 16S V4 region. Abundances were then recovered by mapping the demultiplexed reads to the OTUs file and were then used for subsequent analysis.

2.5 Statistical analysis

2.5.1 Stool metagenomic data

A total of 10,015,821,590 mapped reads (median 14,426,827 reads per sample) were obtained from metagenomic sequencing of the 644 preterm infant stool samples. The lowest sample contained 152,718 mapped reads. The cross-sectional cohort of stool samples collected from NEC infants before diagnosis and matched controls was analysed using MicrobiomeAnalyst (Chong et al., 2020, Dhariwal et al., 2017). Alpha diversity analysis was performed based on observed species (richness) and Shannon diversity, and beta-diversity was performed using Bray-Curtis principal coordinate analysis and differences between groups performed using permutational multivariate analysis of variance (PERMANOVA). MetagenomeSeq was used to assess differential abundance at the phyla and species level. This approach utilises both cumulative sum scaling normalisation and zero-inflated Gaussian distribution mixture or zero-inflated Log-Normal mixture model. Dirichlet Multinomial Mixtures (DMM) clusters samples on the basis of microbial community structure (Holmes et al., 2012) and was used to determine the preterm gut community types (PGCTs) from all samples, as performed previously (Stewart et al., 2017a, Stewart et al., 2018). The appropriate number of clusters was determined based on the lowest Laplace approximation score (Holmes et al., 2012). Five PGCTs were found to be optimal, and these were ordered 1-5 based on the average day of life (DOL) of samples within that PGCT, where PGCT-1 contained on average the samples collected from the earliest DOL and PGCT-5 contained on average the samples collected from the oldest DOL. To avoid introducing biases of repeated measures, analysis was performed at specific time windows, including only a single sample per infant in each time point. In cases where an infant had more than one sample within a given time window, the chosen sample reflected the PGCT which was most common among an infant's samples within the given time window. The ratios of each PGCT were compared by chi-square test.

2.5.2 Mom's own milk 16s rRNA gene sequencing data

A total of 3,898,316 mapped reads were obtained from the 16S rRNA gene sequencing of 130 MOM DNA samples. Each sample was rarefied to 1231 reads, and a total of 111 samples were included in the final analysis (median 22,560 reads per sample). The data analysis was performed using the Agile Toolkit for Incisive Microbial Analysis (ATIMA; https://atima.research.bcm.edu/). In accordance with the stool metagenomic data, alpha diversity analysis was performed based on observed species (richness) and Shannon diversity

(richness and evenness). Beta-diversity was performed using Bray-Curtis principal coordinate analysis, with statistical significance between groups determined using PERMANOVA. Mann-Whitney test was performed to assess differential abundance at phylum and genus level, and p-values were adjusted using the false discovery rate (FDR) algorithm.

2.5.3 Integration of human milk oligosaccharide profile data, stool metagenomes and mom's own milk 16S rRNA gene sequencing data

Random Forest algorithm on MetaboAnalyst V.3.0.28 was used for comparing the performance of classification models built using cross-sectional HMO profile data (nmol/ml), cross-sectional pre-NEC and matched control metagenomic data (count), and both HMO and metagenome datasets combined. The contribution given by each variable was evaluated through the Mean Decrease Accuracy (MDA) value, which indicates how much the accuracy of the model decreases when that variable is removed. The higher the MDA value, the more important that feature is. Variables associated to a negative MDA value were removed and a new model was built on the subset of variables. This step of feature filtering was performed until the model with best classification performance was obtained.

The association of various clinical variables with the HMO, breast milk microbiome and infant gut microbiome profiles was tested by applying the function "adonis" of "vegan" (version 2.5-6) package (Jari Oksanen, 2019) in R (version 3.6.3). Bray-Curtis dissimilarity was used for calculating the dissimilarity matrix, and 1000 permutations were applied. Each test was performed stepwise and P values were adjusted using Benjamini & Hochberg (Benjamini and Hochberg, 1995). Clinical factors tested were delivery mode, gestational age, birthweight, sex, maternal secretor status, infant antibiotic administration (i.e., receiving antibiotics yes/no at time of sample), infant probiotic administration (i.e., receiving probiotic at time of sample, before, after, or never), diet (i.e., combinations of expressed breast milk and formula), DOL and post-menstrual age (PMA) of sample, and disease status.

To understand if any correlation between MOM HMO profile, MOM microbiome and infant gut microbiome, could be found, spearman correlation analysis was performed between the 3 possible pairs of comparison. The analysis was performed in R (version 3.6.3) using the "psych" package (version 2.2.5) (Ravelle, 2016), and p-values were adjusted using the FDR algorithm.

2.6 Microbiology techniques

2.6.1 Bacterial isolation

MOM and stool samples were cultured on agar-based media to isolate and identify viable anaerobic and/or aerobic bacteria, with particular interest in the isolation of *Bifidobacterium* spp. in the stool and the 10 most abundant genera present in the breast milk microbiome. Isolation of fungi from MOM samples was also attempted in aerobic conditions.

100 mg of stool sample were homogenised in 1ml of Phosphate-buffered saline (PBS). 100 µl of homogenised stool sample or 100 µl of neat breast milk, and 100 µl of their 10^{-2} and 10^{-4} dilutions, were cultured on several culture media, listed in Table 2.1. After plating, the samples were left incubating 72/96h. Colonies with diverse morphology were sub-cultured at least twice to obtain pure single colonies, which were then grown in broth media. After incubation overnight, bacterial broths were mixed with glycerol, 25% final concentration, for storage at -80°C. In case of isolation of aerobic microorganisms, every step was performed on the bench next to the flame, to keep the environment sterile. Isolation of bacteria was performed at 37°C, while 30°C were used for fungi (Mini Incubator, Labnet International Inc.) In case of isolation of anaerobic bacteria, every step was performed in anaerobic atmosphere [~60ppm O_2, 2.5% H_2] at 37°C in a Coy Type B Anaerobic Chamber and agar plates, broth media and liquid reagents were left in the anaerobic chamber overnight before usage.

Isolates identification was performed by either full-length 16S rRNA gene sequencing or by MALDI-TOF of single fresh colonies. MALDI-TOF was performed in collaboration with Dr John Perry, Freeman Hospital, Newcastle upon-Tyne.

For full 16S rRNA gene sequencing, bacterial DNA was extracted using DNeasy PowerLyzer PowerSoil Kit (QIAGEN), following manufacturer's instructions with minor edits. Solutions C2 and C3 were mixed in 1:1 proportion, and 200 µl of the mix were added before a unique 5 mins incubation at 4 °C. Next steps were performed with no deviation from the protocol, and DNA was resuspended in 70 µl of C6 solution. Full-length 16S rRNA gene was amplified through PCR (27F 5'-AGAGTTTGATCCTGGCTCAG3'; 1492R 5'-GGTTACCTTGTTACGACTT-3'; Dream-Taq™ Hot Start PCR Master Mix (Thermo Scientific)), including in the reaction a negative control, to check for eventual contamination. The success of the amplification and the lack of contamination were checked by running the samples for 30 min at 200 mA in agarose gel 1 %. The PCR product was cleaned using ExoSAP-IT™ (Applied Biosystems) following the manufacturer's instructions. The cleaned PCR product was then sequenced by Eurofins Genomics using the 27F primer and identity was

determined using nucleotide basic local alignment search tool (BLASTn) in National Center for Biotechnology Information (NCBI) website and selecting the rRNA/ITS databases and using the blastn algorithm.

Table 2.1. Description of the microbiological media used for bacteria and fungi isolation. All media were sterilised using a Prestige Medical Classic Media Autoclave.

Media	Abbreviation	Manufacturer	Target	Aerobic or anaerobic culturing	Notes	Corresponding broth used
deMan, Rogosa, Sharpe	MRS	Sigma-Aldrich	Bacteria	Anaerobic	Prepared according to the manufacturer's instructions. After autoclave, all MRS media was suppelemented with 50 mg/l of L-cysteine-HCl (Sigma-Aldrich). 50 mg/l mupirocin (PanReac AppliChem) were also added to MRS+.	MRS
Bifidus Selective	BSM	Millipore	Bacteria	Anaerobic	Prepared according to the manufacturer's	BSM
TOS-propionate agar medium	TOS	Sigma-Aldrich	Bacteria	Anaerobic	Prepared according to the manufacturer's instructions	BHI
Brain Heart Infusion	BHI	Sigma-Aldrich	Bacteria	Aerobic	Prepared according to the manufacturer's instructions	BHI
Fastidious Anaerobe Agar	FAA	EO Labs	Bacteria	Anaerobic	Prepared according to the manufacturer's instructions	BHI + 5% horse blood (Thermo Scientific)
CHROMID® CPS® Elite	CPSE	Biomerieux	Bacteria	Aerobic	Prepared according to the manufacturer's instructions	BHI
MacConkey Agar	MacConkey	OXOID	Bacteria	Aerobic	Prepared according to the manufacturer's instructions	BHI
Yeast Extract Peptone Dextrose Agar	YPD	Prepared in house	Fungi	Aerobic	1% Yeast extract (Sigma-Aldrich), 2% D-glucose (AnalaR), 2% Bacteriological peptone (OXOID) and Agar (Sigma-Aldrich) were mixed in distilled water. After autoclave, 25/25 U/ml of Penicillin/Streptomycin were added.	YPD
Sabouraud dextrose agar	Sabouraud	OXOID	Fungi	Aerobic	Prepared according to the manufacturer's instructions	Sabouraud

2.6.2 Optimisation of growth curve conditions

A preliminary step of media optimisation was carried out prior to testing bacterial isolates for their ability to utilise HMOs. The goal was to find a singular medium that could support the growth of various bacterial species and isolates when the carbon source was present while also allowing for minimal growth in its absence. Four bacterial media were tested: minimal medium, rich medium as reported by Zabel et al. (2020) (Zabel et al., 2020), modified De Man, Rogosa and Sharpe (mMRS - MRS without carbon source) and ZMB1 as reported by Zhang et al. (2009) (Zhang et al., 2009). The composition and preparation protocol for each media are reported in Table 2.2, 2.3, 2.4 and 2.5. As described in Table 2.2, minimal medium was prepared following two different protocols, with L-cysteine HCl being added either before or after the autoclaving step. All media were tested with and without glucose, and various glucose concentrations were used: 1%, 1.5% and 2%.

Bacteria isolates tested were grown overnight in their preferred medium: *Escherichia coli*, *Enterococcus faecium*, *Enterococcus faecalis*, *Klebsiella pneumoniae*, *Klebsiella oxytoca* and *Lacticaseibacillus rhamnosus* were grown in brain hearth infusion (BHI) medium, *Bifidobacterium longum* and *Bifidobacterium breve* isolates in MRS supplemented with filter-sterilised L-cysteine HCl (0.05% w/v). Growth curves were measured in a Cerillo Stratus plate reader for 24h. Bacterial overnight growth was mixed with the tested media in a 1:10 ratio and 200 µl and 100 µl of final volume were tested.

Where the bacteria were washed in PBS, the overnight growth was centrifuged at 5000 G for 5 mins at 4°C, and the pellet was resuspended in the same amount of anaerobic PBS.

Table 2.2. Recipe for Minimal medium preparation.

Reagent	**1l of media**	**Preparation 1**	**Preparation 2**
NH4 SO4	1 g	All components except for histidine-hematin solution were mixed in a 2x concentration and autoclaved for 20 minutes. After autoclave, the 2x medium was mixed with defined volumes of sterile dH_2O and/or sterile 4% glucose solution to reach the desired final glucose concentration. Histidine-hematin solution was then added.	All components except for histidine-hematin solution and L-cysteine HCl were mixed in a 2x concentration and autoclaved for 20 minutes. After autoclave, the 2x medium was mixed with defined volumes of sterile dH_2O and/or sterile 4% glucose solution to reach the desired final glucose concentration. Histidine-hematin solution and filter-sterilised L-cysteine HCl were then added.
Na2CO3	1 g		
L-cysteine HCl	0.5 g		
1 M KPO4 pH 7.2	100 ml		
Vitamin K solution, 1mg/ml	1 ml		
FeSO4, 0.4 mg/ml	10 ml		
Resazurin, 0.25 mg/ml	4 ml		
Vitamin B12, 0.01 mg/ml	0.5 ml		
Mineral Salts for defined medium	50 ml		
Histidine-hematin solution (12 mg hematin dissolved in 10 ml of 0.2 M histidine pH 8, then filter-sterilise using 0.2 µm filter)	1 ml		

Table 2.3. Recipe for Rich medium preparation, as reported by Zabel et al. (2020) (Zabel et al., 2020).

Reagent	1l of media	Preparation
Bacteriological peptone	10 g	All components except for L-cysteine HCl were mixed in a 2x concentration. pH was adjusted at 6.4 and the solution autoclaved for 20 minutes. After autoclave, the 2x medium was mixed with defined volumes of sterile dH_2O and/or sterile 4% glucose solution to reach the desired final glucose concentration. Then, filter-sterilised L-cysteine HCl was added to the medium.
Yest extract	2 g	
NaCl	5 g	
Diammonium citrate	2 g	
Magnesium sulphate	0.2 g	
Dipotassium hydrogen phosphate	2 g	
L-cysteine HCl	0.5 g	

Table 2.4. Recipe for modified MRS supplemented with 0.05% L-cysteine HCl.

Reagent	1l Medium	Preparation
Bacteriological peptone	10 g	All components were mixed in a 2x concentration and autoclaved for 20 minutes. After autoclave, the 2x medium was mixed with defined volumes of sterile dH_2O and/or sterile 4% glucose solution to reach the desired final glucose concentration. Medium was then supplemented with 0.05% w/v L-cysteine HCl.
Beef extract	10 g	
Yeast extract	4 g	
Sodium acetate trihydrate	5 g	
Tween 80	1 g	
Dipotassium hydrogen phosphate	2 g	
Triammonium citrate	2 g	
Magnesium sulphate heptahydrate	0.2 g	
Manganese sulfate tetrahydrate	0.05 g	

Table 2.5. Recipe for ZMB1 media. Table adapted from Zhang et al. (2009) (Zhang et al., 2009). All components were mixed, and defined volumes of sterile dH_2O and/or 250 mg/ml of glucose were added to reach the desired glucose concentration.

Reagent	Procedures for preparing solutions	Stock solution (mg/ml)	Final concentration (mg/ml)
L-Histidine (His)	Stir to dissolve in water to desired concentration (SDWDC) with magnetic bars; Vacuum-filtered sterilization (VFS, 0.22 µm pore size); precipitate at 4 ºC.	40	0.17
L-Isoleucine (Ile)	SDWDC; VFS.	20	0.24
L-Leucine (Leu)	Stir; 10 N NaOH added drop by drop till dissolved (ADDD); VFS.	25	1
L-Methionine (Met)	SDWDC; VFS.	15	0.06
L-Valine (Val)	SDWDC; VFS.	17.5	0.7
L-Arginine (Arg)	SDWDC; VFS.	20	0.72
Inositol (Ino)	SDWDC; VFS.	1	0.002

Reagent	Procedures for preparing solutions	Stock solution (mg/ml)	Final concentration (mg/ml)
KH_2PO_4	SDWDC; autoclave.	150	3.1
K_2HPO_4	SDWDC; autoclave.	500	6.4
L-Glutamic acid (Glu)	Stir; 10 N NaOH ADDD; VFS.	15	0.6
L-Phenylalanine (Phe)	SDWDC; VFS.	10	0.4
L-Proline (Pro)	SDWDC; VFS.	17.5	0.7
L-Asparagine (Asn)	Stir; 10 N NaOH ADDD; VFS; precipitate at 4 ºC; heat sensitive.	12.5	0.5
L-Aspartic acid (Asp)	Stir; 10 N NaOH ADDD; VFS.	1.25	0.05
L-Glutamine (Glx)	SDWDC; VFS; heat sensitive.	15	0.6
L-Serine (Ser)	SDWDC; VFS.	50	0.5
L-Threonine (Thr)	SDWDC; VFS.	12.5	0.5
L-Cysteine (Cys) HCl	SDWDC; VFS; heat sensitive; fresh preparation preferred; reducing agent.	50	0.2
L-Alanine (Ala)	SDWDC; VFS.	10	0.4
Glycine (Gly)	SDWDC; VFS.	7.5	0.3
L-Lysine (Lys) HCl	SDWDC; VFS.	12.5	0.5
L-Tryptophan (Trp)	Stir; 10 N NaOH ADDD; VFS; heat sensitive.	5	0.2
L-Tyrosine (Tyr)	Stir; 10 N NaOH ADDD; VFS.	7.5	0.3
Biotin	Stir and heat; 10 N NaOH ADDD; VFS.	0.5	0.006
Calcium pantothenate	SDWDC; only stable at pH 5 to 7.	0.1	0.0012
Niacin	SDWDC; VFS.	0.075	0.0009
Pyridoxal HCl	SDWDC; VFS; heat sensitive.	0.4	0.0048
Riboflavin	SDWDC; VFS; heat sensitive; unstable at alkaline solution.	0.075	0.0009
$MgSO_4.7H_2O$	Prepared seperately from phosphates; SDWDC; autoclave.	50	1
$FeSO_4.7H_2O$	SDWDC; VFS; unstable when exposed to air; fresh preparation preferred.	0.4	0.004
$ZnSO_4.7H_2O$	SDWDC; VFS.	2	0.005
Folic acid	dissolve in NaOH; VFS; light sensitive; stable at alkaline solution.	0.075	0.00056
p-Aminobenzoic acid	SDWDC; VFS.	0.1	0.000056
Thiamine HCl	SDWDC; VFS; heat sensitive; unstable if pH > 5.	1	0.00056
Potassium acetate	SDWDC; autoclave.	200	0.9
lipoic acid	SDWDC; VFS.	0.1	0.001
Tween 80	Heat while stirring, then add water to dissolve it when it is hot (but not too hot); autoclave.	50	0.5
Adenine	Heat while stirring; VFS.	1.5	0.011
Guanine	Stir, 10 N H2SO4 ADDD (around pH 1), then 10 N NaOH ADDD to raise pH to be 2.75 (for stability); VFS.	0.75	0.0056
Uracil	Stir; 10 N NaOH ADDD; VFS.	3	0.023
Xanthine	Stir; 10 N NaOH ADDD; VFS.	0.5	0.0038
MOPS	SDWDC; VFS.	500	15
Tricine	SDWDC; VFS.	80	1.5
$(NH_4)_6Mo_7O_{24} \cdot 4H_2O$	SDWDC; VFS.	0.025	0.00019

Reagent	Procedures for preparing solutions	Stock solution (mg/ml)	Final concentration (mg/ml)
$MnSO_4.4\ H_2O$	SDWDC; VFS.	0.05	0.00038
$CaCl_2.2\ H_2O$	SDWDC; autoclave.	5	0.04
$CoCl_2.6\ H_2O$	SDWDC; VFS.	0.025	0.00019
$CuSO_4.5\ H_2O$	SDWDC; VFS.	0.025	0.00019
H_3BO_3	SDWDC; VFS.	0.1	0.00075
K_2SO_4	SDWDC; VFS.	3	0.023
KI	SDWDC; VFS; light sensitive.	0.3	0.00011
EDTA	Stir; 10 N NaOH ADDD; autoclave.	1	0.0075
NTA	Stir; 10 N NaOH ADDD (pH 6.5); VFS.	1	0.0075
Glutathione (GSH)	SDWDC; VFS; reducing agent.	50	0.015
$(NH_4)_2SO_4$	SDWDC; autoclave.	50	1
NaCl	SDWDC; autoclave.	300	3

SDWDC, stir to dissolve in water to desired concentration; VFS, vacuum-filtered sterilization; ADDD, added drop by drop till dissolved.

2.6.3 Bacterial isolates growth curves

The 17 isolates listed in Table 2.6 were tested. All isolates were grown overnight in BHI, with the exception for bifidobacteria which were grown in MRS supplemented with L-cysteine HCl (0.05% w/v). ZMB1 without glucose was prepared, and various carbon sources were tested in 1% concentration: no carbon source, glucose, lactose, 2'FL, DSLNT, LNT, LNnT, LNFP I, 6SL. The overnight growth for each isolate was centrifuged at 5000 G for 5 mins at 4°C, and the pellet was resuspended in the same amount of anaerobic PBS. 20 µl of the bacteria in PBS was spiked in 180 µl of media, and the growth was measured in a Cerillo Stratus plate reader for 150h. All isolates were tested in triplicate, and wells with media only were included in each plate to check for eventual contamination.

Table 2.6. List of isolates tested for growth on HMOs. Information on the isolate number, species, sample they were isolated form and sample type are reported.

Isolate	Species	Sample	Sample type
PP1	*Bifidobacterium animalis*	Proprems	Commercial probiotic
CS27	*Bifidobacterium bifidum*	NA	Stool
AM76	*Bifidobacterium breve*	H001_10	Stool
BM6-3	*Bifidobacterium breve*	NA	Breast milk
VK68	*Bifidobacterium breve*	S009_106	Stool
LB1	*Bifidobacterium longum* subspecies *infantis*	Labinic	Commercial probiotic
AM7	*Bifidobacterium longum*	B013_1	Stool
VK53	*Bifidobacterium longum*	C001_18	Stool
JC10	*Enterobacter cloacae*	38	Stool
AM3	*Enterococcus faecalis*	B013_1	Stool
AM9	*Enterococcus faecium*	B013_1	Stool
AM114	*Escherichia coli*	G012_34	Stool
VK66	*Klebsiella oxytoca*	S009_106	Stool
MR1	*Klebsiella pneumoniae*	1238	Stool
MR5	*Lacticaseibacillus rhamnosus*	5657	Stool
MR3	*Staphylococcus aureus*	1235	Stool
AM72	*Staphylococcus epidermidis*	H001_10	Stool

2.7 Human intestinal-derived organoids experiments

2.7.1 Intestinal organoids media preparation

Media for organoids propagation was prepared in-house and included the preparation of media conditioned with L-Wnt3A, Noggin and R-spondin.

2.7.2 L-Wnt3A conditioned media

L-Wnt3A cells (ATCC CRL-2647) were thawed from liquid nitrogen by placing the vial in a 37°C incubator for 2 minutes, and cells were propagated overnight in Regular Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% heat-inactivated Foetal Bovine Serum (FBS). The subsequent day, media was supplemented with 400 µg/ml of G418 (Invitrogen), and cells were propagated in T75 Flasks. When flasks reached confluence, 4 ml of PBS pre-warmed at 37°C were used to wash the flask, 1.5 ml of 0.25% Trypsin (Corning)

were then added to the flask and cells were incubated for 5 minutes at 37 °C. After incubation, 8.5 ml of DMEM supplemented with 10% FBS and 400 µg/ml G418 were added, and cells were used to seed new flasks. One confluent flask was used to seed ten new T75 flasks and cells were cultured in 10 ml of media. Once enough L-Wnt3A cells were obtained, one confluent T75 flask was used to seed one T875 super flask. 110 ml of Advanced DMEM/F12 supplemented with 1x Glutamax (Gibco) and 10% FBS were added to the T785 flask and pre-warmed at 37°C. Cells were washed with 4 ml of PBS, 1.5 ml of trypsin were added, the flask was incubated at 37°C for 5 minutes, after which 3.5 ml of Advanced DMEM/F12 supplemented with 1x Glutamax and 10% FBS were added. L-Wnt3A cells were then moved to the T875 multi-flask (Falcon), and media was added to reach a final volume of 120 ml of Advanced DMEM/F12 supplemented with 1x Glutamax and 10% FBS. The first batch of L-Wnt3A conditioned media was harvested after 3 days and 120 ml of new Advanced DMEM/F12 supplemented with 1x Gllutamax and 10% FBS were added to the flask, which was incubated an additional 2 days before harvesting the second batch of conditioned media. Right after harvesting each batch of conditioned media, all media was centrifuged at 1000 G for 10 minutes at 4°C to remove the cell content, followed by filtration using a 0.22 µm filter unit. L-Wnt3A conditioned media was stored at -80°C until usage.

2.7.3 R-spondin conditioned media

293T-HA-RspoI-Fc cells (kindly donated by Calvin Kuo, Palo Alto, CA) were thawed from liquid nitrogen by placing the vial in a 37°C incubator for 2 minutes, and cells were propagated overnight in Regular DMEM supplemented with 10% FBS. The subsequent day, media was supplemented with 300 µg/ml Zeocin (Invitrogen), and cells were propagated in T75 Flasks. When cells reached 80% confluence, they were passaged as described for the L-Wnt3A cell line, with the variation that 0.05% trypsin was used, and the passage ratio was one to five T75 flasks. Once enough 293T-HA-RspoI-Fc cells were obtained, five 80% confluent T75 flask were used to seed one T875 super flask. 230 ml of Advanced DMEM/F12 supplemented with 1x Glutamax and 10% FBS were added to the T785 flask and pre-warmed at 37°C. Cells were washed with 4 ml of PBS, 1.5 ml of 0.05% trypsin were added to each flask and incubated at 37°C for 5 minutes, after which 3.5 ml of Regular DMEM supplemented with 10% FBS were added. 293T-HA-RspoI-Fc cells from each of the 5 T75 flasks were combined in a 50 ml Falcon tube, centrifuged 4 minutes at 200 G and resuspended in 10 ml of Advanced DMEM/F12 supplemented with 1x Glutamax and 10% FBS. The cells were then moved to the T785 superflask reaching a final volume of 240 ml of Advanced DMEM/F12 supplemented with 1x

Glutamax and 10% FBS. The conditioned media was harvested after 6 days, centrifuged at 1000 G for 10 minutes at 4°C to remove the cell content, followed by filtration using a 0.22 µm filter unit. The 293T-HA-RspoI-Fc (RSpondin) conditioned media was stored at -80°C until usage.

2.7.4 Noggin conditioned media

Noggin-producing cells (kindly donated by G. R. van den Brink, Amsterdam, The Netherlands) conditioned media was prepared following the same method described for 293T-HA-RspoI-Fc cells with some variations. The antibiotic used for Noggin-producing cell line expansion was 10 µg/ml Puromycin (Invitrogen), 0.08 % trypsin was used for cell passage and after the last cell passage before T875 multi-flask seed, Noggin-producing cells were cultured in Regular DMEM supplemented with 10% FBS without the addition of Puromycin. Noggin conditioned media was stored at -80°C until usage.

2.7.5 Intestinal organoids extraction and propagation

Human intestinal-derived organoid lines were generated from preterm and adult ileal resected tissue following the protocol described by Stewart *et al*. (2020) (Stewart et al., 2020). The samples used were regulated under the ethics for the Supporting Enhanced Research in Vulnerable Infants (SERVIS) study (REC10/H0908/39). Briefly, ~5 mm^2 of tissue were minced and thoroughly washed with complete chelating (CCS) solution (recipe in Table 2.7) supplemented with 2.5 µg/ml Amphotericin b (Gibco) and 100 µg/ml Penicillin/Streptomycin (Gibco). Intestinal crypts were subsequently extracted by incubating the tissue twice in 3 ml CCS with increasing concentrations of Ethylenediaminetetraacetic acid (EDTA; 0.03 M and 0.04 M the first and second incubation, respectively) on orbital shaker at 300 rpm, for 30 minutes, at 4 °C. The supernatant containing the crypts was collected after each incubation and centrifuged 5 minutes, 4 °C, 1000 rpm. The isolated crypts were resuspended in 70 µl Matrigel (Corning) and two dots were dispensed in three wells of a 24 well tissue culture plate. High Wnt media prepared in house (recipe in Table 2.8) was supplemented with 10 µM Y-27632 (Sigma Aldrich) and 500 µl were added to each well. The media was changed every 2-3 days. 2.5 µg/ml Amphotericin b and 100 µg/ml Penicillin/Streptomycin were added to the media for the first 3 passages.

Organoids were passaged by mechanical disruption using a 25 gauge needle every 5-7 days. Briefly, media was removed from the wells and 500 µl of cold (4°C) Complete Media Growth Factor negative (CMGF-; recipe in Table 2.9) was added to the first well to disrupt the Matrigel

and resuspend the organoids. The 500 µl of CMGF- containing the resuspended organoids were passed from one well to the next one, thus collecting the organoids from all the wells seeded. The resuspended organoids were then moved to a 15 ml Falcon tube, and an additional 500 µl of CMGF- were used to wash the wells a second time and collect any organoids left. A 25-gauge needle was then used to mechanically disrupt the organoids, which were subsequently centrifuged 5 minutes, 1000 rpm at 4°C. Organoids were passaged in ratio 1:2 and resuspended in an amount of Matrigel sufficient to seed 30 µl in each well of a 24 well tissue culture plate. As described above, two dots of Matrigel containing the organoids were dispensed in each well, and 500 µl of High Wnt media supplemented with 10 µM Y-27632 were added. Media was changed every 2-3 days.

Table 2.7. Recipe for complete chelating solution preparation. The solution was sterilised using a 0.22 µm filter unit.

Reagent	Manufacturer	Quantity
MilliQ H2O	Sigma-Aldrich	50 ml
Na_2HPO_4-$2H_2O$	Sigma-Aldrich	0.5 g
KH_2PO_4	Sigma-Aldrich	0.54 g
NaCl	Sigma-Aldrich	2.8 g
KCl	Sigma-Aldrich	60 mg
Sucrose	Sigma-Aldrich	7.5 g
D-Sorbitol	Sigma-Aldrich	2 g
DL-Dithiothreitol solution 1M	Sigma-Aldrich	26 µl

Table 2.8. Recipe for High Wnt medium.

Reagent	Manufacturer	Quantity
CMGF+ (recipe in Table 2.10)	Produced in-house	250 ml
Wnt3A conditioned medium	Produced in-house	250 ml

CMGF +, Complete Media Growth Factor positive.

Table 2.9. Recipe Complete Media Growth Factor negative (CMGF-).

Reagent	Manufacturer	Quantity
Advanced DMEM/F12	Gibco	500 ml
Glutamax 100x	Gibco	5 ml
HEPES 1M	Gibco	5 ml

DMEM, Dulbecco's Modified Eagle Medium.

Table 2.10. Recipe Complete Media Growth Factor positive (CMGF+).

Reagent	Manufacturer	Quantity
CMGF- (recipe in Table 2.9)	Produced in-house	78 ml
L-Wnt3A conditioned medium	Produced in-house	250 ml
R-spondin conditioned medium	Produced in-house	100 ml
Noggin conditioned medium	Produced in-house	50 ml
B27 (50x)	Gibco	10 ml
N2 (100x)	Gibco	5 ml
Nicotinamide 1M	Sigma-Aldrich	5 ml
N-Acetylcysteine 500 mM	Sigma-Aldrich	1 ml
A-83-01 500 µM	Sigma-Aldrich	500 µl
[Leu15]-Gastrin I 10 µM	Sigma-Aldrich	50 µl
SB202190 30 mM	Sigma-Aldrich	157 µl
House recombinant EGF 50 µg/ml	Sigma-Aldrich	50 µl

CMGF -, Complete Media Growth Factor negative; DMEM, Dulbecco's Modified Eagle Medium.

2.7.6 Intestinal organoid derived monolayers formation

Monolayers of tissue derived organoids were generated for the culture in the intestinal OACC model. Transwell inserts (6.5 mm, 0.4 µm pore size; Corning) were coated with 100 µl of Matrigel diluted 1:40 in PBS and incubated at 37 °C during the following passages. Organoids were washed in 0.5 mM EDTA. Single cell suspensions were obtained by incubating organoids with 1 ml 0.05% Trypsin – 0.5 mM EDTA for 5 minutes at 37 °C, and subsequently passed through a 40 µm cell strainer. After removing excess PBS, Transwell inserts were seeded with $5x10^5$ cells suspended in 200 µl of CMGF+ (recipe in Table 2.10) media supplemented with 10 µM Y-27632 (Sigma Aldrich), and 650 µl of media were added to the basolateral side. TEER was measured every day, and differentiation media was used after TEERs reached ~300 Ohm. The media was renewed every other day, and monolayers were cultured in the OACC system after 4 days of differentiation.

2.7.7 Organoid monolayers conditions optimisation

Prior to generating organoid derived monolayers for the experiments described in sections 2.7.8 and 2.7.9, an optimisation of the protocol was performed. In the first optimisation step two different cell counts were used for seeding 0.4 µm pore size Transwell inserts: $5x10^5$ and $8x10^5$

cells, each tested in duplicate. In the second step two different pore size Transwell inserts were tested: $5x10^5$ cells were used to seed 0.4 µm and 3 µm pore size Transwell inserts, each pore size was tested in triplicate.

The OACC model was assembled as described by Fofanova *et al.* (2019) (Fofanova et al., 2019). 650 µl of differentiation media were added to a 24-well gas-permeable plate and Transwell inserts containing the differentiated monolayers were placed into appropriate gaskets and subsequently sealed using double-sided adhesive tape onto the plate. 200 µl of fresh differentiation media were added to the apical compartment, before moving the console to the anaerobic chamber. A gas mix containing 5% O_2, 5% CO_2, 90% N_2, was supplied to the basolateral media by diffusion through the base of the gas permeable plate. Culture was stopped after 24 hours; the TEER values were then measured before fixing the monolayers in 4% paraformaldehyde (PFA) for 1 hour. Monolayer slides were stained using hemotoxylin and eosin (H&E) by the Molecular Pathology Node (Royal Victoria Infirmary, Newcastle).

2.7.8 Organoid monolayers exposure to Enterococcus faecalis

The OACC model was assembled as described by Fofanova *et al.* (2019) (Fofanova et al., 2019). Briefly, 650 µl of differentiation media were added to a 24-well gas-permeable plate and Transwell inserts containing the differentiated monolayers were placed into appropriate gaskets and subsequently sealed using double-sided adhesive tape onto the plate. The console was moved to the anaerobic chamber and a gas mix containing 5% O_2, 5% CO_2, 90% N_2, was supplied to the basolateral media by diffusion through the base of the gas permeable plate. 200 µl of mix 1:1 of differentiation media and bioreactor media (Auchtung et al., 2020) were added to the apical side, after equilibrating the system for one hour. Where monolayers exposed to *E. faecalis*, $5x10^5$ bacterial cells/ml were added to the apical media. The co-culture was stopped after 8 hours or 24 hours. The monolayers were washed with PBS and RNA was immediately extracted using the QIAGEN RNeasy Mini Kit following the manufacturer's instructions.

2.7.9 Organoid monolayers exposure to human milk oligosaccharides

The OACC model was assembled as described in the paragraph above, with some modifications. Briefly, 650 µl of differentiation media supplemented with 1 ng/ml of interleukin (IL) 1 beta (IL-1β (Sigma-Aldrich) were added to a 24-well gas-permeable plate. Transwell inserts carrying differentiated organoids were sealed onto the plate and the console was moved to the anaerobic chamber. The apical stimulus comprised 200 µl of differentiation media supplemented with 1000 ng/ml of LPS (Sigma-Aldrich) and the different HMOs were added to

the apical compartment of the monolayers. HMOs were tested in the concentration they were present in MOM for the healthy infants included in the HMO profile cohort (analysed in chapter 3); details on the conditions tested can be found in Table 2.11. All conditions were tested in triplicate and the co-culture was stopped after 24 hours. TEERs were measured before and after the co-culture, and the apical and basolateral media were collected after 24h for subsequent downstream analysis. The Transwell membrane carrying the organoid monolayer was cut two halves: one used to extract RNA, one fixed in 4% (w/v) PFA for 1 hour. Triplicates for the same condition were pooled together for RNA extraction and PFA fixation. RNA was extracted using the QIAGEN RNeasy Mini Kit following the manufacturer's instructions.

Table 2.11. Exposure conditions tested in experiments described in section 2.4.8.

Condition	Apical	Basolateral
Control	1000 ng/ml LPS	1 ng/ml IL-1β
DSLNT	1000 ng/ml LPS + 0.5 mg/ml DSLNT	1 ng/ml IL-1β
2'FL	1000 ng/ml LPS + 2 mg/ml 2'FL	1 ng/ml IL-1β
6SL	1000 ng/ml LPS + 0.75 mg/ml 6SL	1 ng/ml IL-1β
LNnT	1000 ng/ml LPS + 0.25 mg/ml LNnT	1 ng/ml IL-1β
LNT	1000 ng/ml LPS + 1 mg/ml LNT	1 ng/ml IL-1β
LNFP I	1000 ng/ml LPS + 1 mg/ml LNFP I	1 ng/ml IL-1β
HMOs Mix	1000 ng/ml LPS + 0.5 mg/ml DSLNT + 2 mg/ml 2'FL + 0.75 mg/ml 6SL + 0.25 mg/ml LNnT + 1 mg/ml LNT + 1 mg/ml LNFP I	1 ng/ml IL-1β

LPS, lipopolysaccharide; IL-1β, interleukin 1 beta; DSLNT, disialyllacto-N-tetraose; 2'FL, 2'-fucosyllactose; 6'SL, 6'-sialyllactose; LNnT, lacto-N-neotetraose; LNT, lacto-N-tetraose; LNFP I, lacto-N-fucopentaose I.

2.7.10 RNA sequencing and statistical analysis

The monolayers were washed with PBS where exposed to *E. faecalis.* For all conditions, RNA was immediately extracted using the QIAGEN RNeasy Mini Kit following the manufacturer's instructions. Total RNA-Seq libraries were prepared for differential gene expression analysis using the SMART-Seq® Stranded Kit (Takara) following manufacturer's instructions. The pooled library of samples was sequenced on a NovaSeq 6000 using a SP v1.0 100 cycle sequencing kit (Illumina) providing ~20 million 100 bp single reads per sample. Quality control of raw reads was performed with 'FASTX Toolkit' (version 0.0.14) with Phred quality score threshold set to 28 and minimum read length of 50bp. Transcript abundance was estimated using 'salmon' (v1.1.0) (Patro et al., 2017) with the GRCh38 transcriptome from GenCode (release 33). Transcript-level counts produced by Salmon were aggregated at the gene level by

'tximport' (version 1.14.2) (Soneson et al., 2015). The "DESeq2" package (version 1.26.0) (Love et al., 2014) in the R (version 3.6.3) environment was used for counts normalisation and gene expression comparison between groups. Genes were considered significantly differentially expressed if associated to an adjusted P value < 0.05 and Fold Change (FC) > 2. PLS-DA for two group comparison was performed running "ropls" package (version 1.18.8) (Thevenot et al., 2015) in R. Normalised counts transformed with "rlogTransformation" function of the "DESeq2" package were used and significance of group separation was determined running 2000 permutations. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) Biological Processes (BPs) over-representation analysis (ORA) was performed using "GOstats" package (version 2.52.0) in R environment (Falcon and Gentleman, 2007) or WebGestalt online tool and P values were adjusted for multiple comparisons applying the FDR algorithm.

All data from the *E. faecalis* experiment is available at the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/), Accession: GSE161953.

2.7.11 Enzyme-linked immunosorbent-assay

IL-8 was measured from apical and basolateral media harvested after the co-culture using the DuoSet Enzyme-linked immunosorbent-assay (ELISA) Kit (R&D Systems) following the manufacturer's instructions with some modifications. The volumes used at each step were halved, except for the addition of Substrate Solution and Stop Solution which were added in the volume suggested. Apical media was tested as a 1:2 dilution, while basolateral media was use as neat, and technical triplicates were performed. The absorbance at 450 nm was measured in a microplate reader. A logistic curve fit was used in GraphPad Prism™ software (GraphPad, San Diego, California, USA) to create a standard curve, and samples concentrations were determined using interpolated data. Wilcoxon rank test followed by FDR adjustment were used to determine the p-values for group comparisons.

Chapter 3. Human milk oligosaccharide profile and preterm infant gut microbiome

3.1 Introduction

Breast milk feeding has been associated with up to 10-fold decrease in risk of developing NEC in preterm infants (Meinzen-Derr et al., 2009). This beneficial outcome is probably the result of the cumulative protective effect exerted by multiple components present in this biofluid, and HMOs are likely one of these. Recent work has begun to elucidate the potential contribution of HMOs to preterm infant health. In a neonatal rat model, DSLNT, a non-fucosylated but double-sialylated HMO, significantly reduced NEC development and improved NEC-associated mortality rate (Jantscher-Krenn et al., 2012). An association of lower DSLNT concentration in MOM and subsequent higher risk of NEC onset in the infant has since been observed in preterm human studies (Autran et al., 2018, Van Niekerk et al., 2014). To date, these studies have included very small numbers of infants with NEC (between 4 and 8), with a broad range of NEC phenotypes.

NEC development has been repeatedly linked to the infant gut microbiome development and composition (Stewart et al., 2016, Sim et al., 2015). Despite it has been shown that MOM feeding influences the term infant gut microbiome establishment through HMOs among other factors, this relationship has not been investigated to date in preterm infants, either in healthy cohorts or in association with NEC onset. In particular, HMOs shape the term infant gut microbiome by favouring the establishment of a bifidobacteria rich bacterial community (Lawson et al., 2019, Vatanen et al., 2018), which in preterm cohorts might help preventing NEC development (Stewart et al., 2016).

The first aim of this study was to validate the results just discussed in a larger cohort and expand the findings with infant stool gut microbiome analysis. As described in chapter 2, MOM samples from NEC and matched control infants were selected, and sent to Prof Lars Bode (University California San Diego, San Diego), who applied HPLC to determine the absolute concentration of the 19 most abundant HMOs present in human breast milk. Furthermore, longitudinal infant stool metagenomic data was available for a sub-cohort of the study through a collaboration with Astarte Medical. The MOM HMO profile and infant metagenomic data were analysed by performing a variety of tests including ordination plots, statistical group comparisons to evaluate differences for specific variables, and application of univariate and multivariate models, and random forest classification, to explore the potential use of HMO profile and/or infant gut metagenomic profile for the prediction of NEC onset.

3.2 Human milk oligosaccharide profile in necrotising enterocolitis

HMO profiles were analysed for 37 control and 33 NEC infants. NEC and control groups were matched by various confounders, and no significant difference was observed, except for sex (Table 3.1). A higher percentage of male infants was present in the NEC group compared to the control group (p=0.016); however, this has previously been reported in the literature with male preterm babies having a higher risk of developing NEC (van Westering-Kroon et al., 2021).

Table 3.1. Demographics of the analytical cohort with human milk oligosaccharide profile data. Differences between groups were tested applying Chi-square test and Wilcoxon rank test where applicable.

	Control	NEC	P value
Number of patients	37	33	-
Secretors	25 (68%)	20 (61%)	0.544
Male	14 (38%)	22 (67%)	0.016
Vaginal delivery	25 (68%)	17 (52%)	0.171
Gestational age	25 [24; 26]	25 [24; 27]	0.881
Birthweight	670 [585; 830]	670 [600; 840]	1.000
Probiotics ever	37 (100%)	31 (94%)	0.855
MOM only	3 (8%)	6 (18%)	0.468
MOM + Formula	11 (30%)	12 (37%)	
MOM + BMF	10 (27%)	7 (21%)	
MOM + Formula + BMF	13 (35%)	8 (24%)	
DOL breast milk sample	18 [12; 31]	18 [13; 34]	0.636
DOL disease onset	-	19 [14; 35]	-
NEC surgical	-	16	-

NEC, necrotising enterocolitis; MOM, mother's own breast milk; BMF, breast milk fortifier; DOL, day of life.

Since one of the major factors influencing HMO profiles is secretor status, the first step was to look at the differences between samples based on this variable. MOM samples clustered according to maternal secretor status and 16 of the 19 HMOs measured were present in significantly different concentrations, secretor mothers had a higher total HMO concentration, a higher HMO Shannon diversity and a significantly higher concentration of overall HMO-

bound fucose (Figure 3.1). Thus, where relevant, the data was stratified and adjusted for maternal secretor status in subsequent analyses.

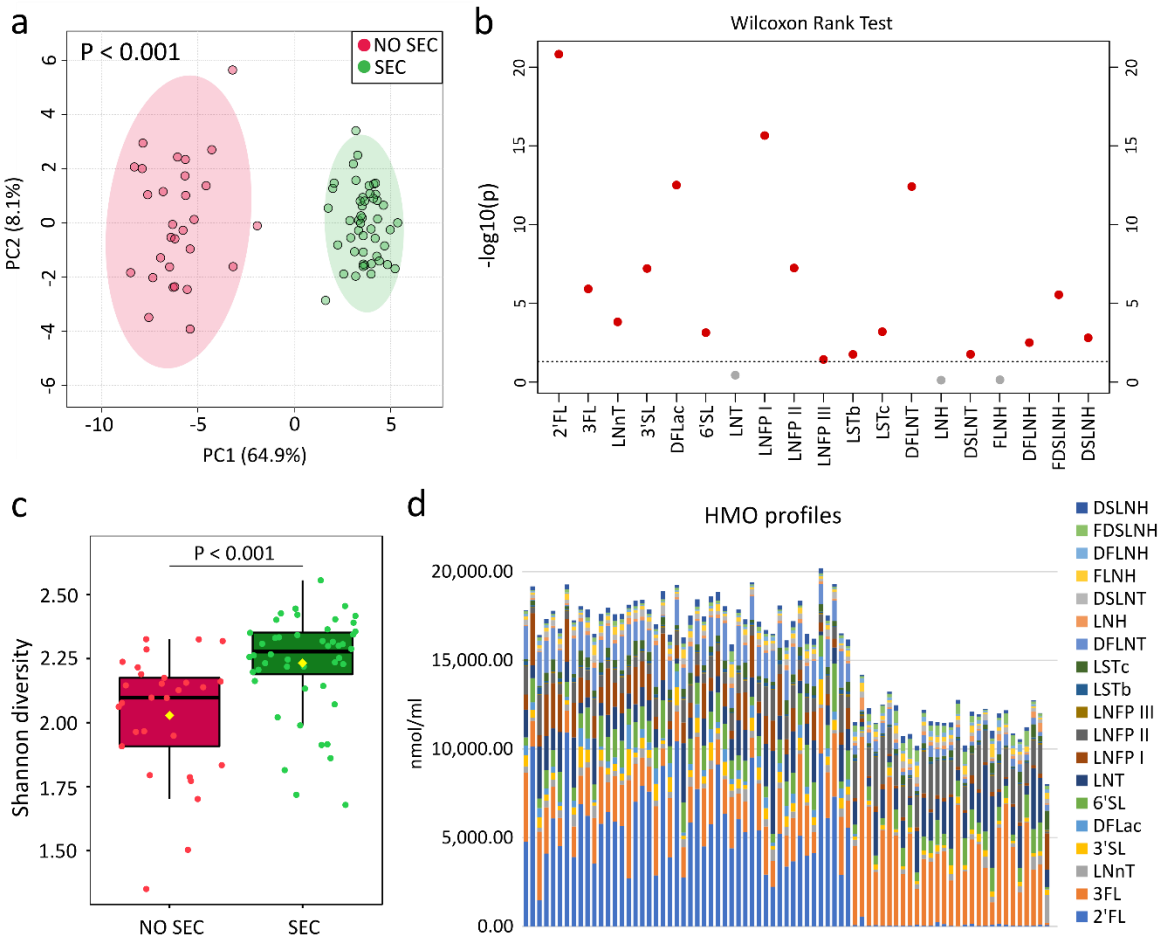

Figure 3.1. Comparison of human milk oligosaccharide (HMO) profiles by maternal secretor status. All infants were included (n = 70). (a) Principal component analysis (PCA) showing the clustering of HMO profiles based on secretor status. (b) Visual representation of P values for comparison of individual HMOs between secretor and non-secretor groups. 16 of the 19 HMOs were different between the two groups. Wilcoxon rank test was applied, and P values were adjusted using FDR algorithm. (c) HMO Shannon diversity was higher in breast milk from secretor mothers compared to non-secretors. Wilcoxon rank sum test was applied. (d) Stacked bar plot of HMOs concentrations describing HMO profile of each breast milk sample analysed. 2'FL used for identifying secretor status is almost absent in non-secretor breast milks, and present in relatively high concentration in samples from secretor mothers.

The next step was to investigate if any difference was present in the HMO profiles depending on disease status. HMO profiles showed significant separation of NEC and control infants (2000 permutations, $p<0.001$; Figure 3.2a), and this was consistent when secretor and non-secretor samples were analysed separately (both 2000 permutations, $p<0.001$; Figure 3.2b,c). Individually, of the 19 HMOs quantified in this study, only DSLNT was significantly different between NEC and controls, with a lower concentration in infants with NEC (adjusted (adj.)

p<0.001; Figure 3.2d) independently of secretor status (Figure 3.2e,f). No significant associations were found in the Shannon diversity of HMOs between NEC and matched controls for the full cohort, or when stratified by maternal secretor status (all p>0.05; Figure 3.2g,h,i).

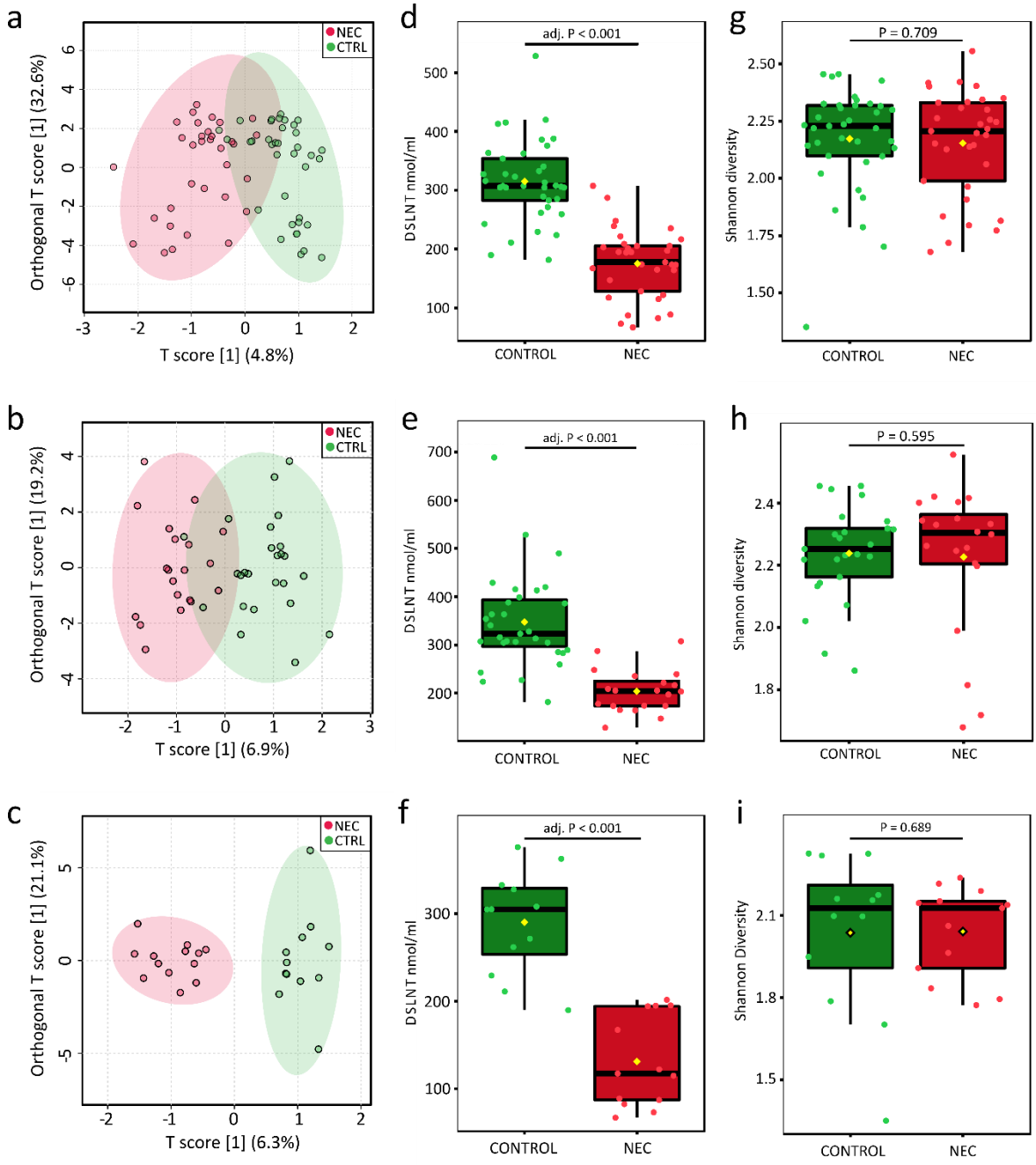

Figure 3.2. Analysis of HMO profiles and DSLNT concentration in NEC and controls. Orthogonal partial least squares discriminant analysis of maternal HMO profiles fed to infants diagnosed with NEC and controls from full cohort (a), secretors (b) and non-secretors (c). P values were calculated based on 2000 permutations. Box plot showing the concentration of DSLNT between NEC and controls in the full cohort (d), secretors (e) and non-secretors (f). Group comparison was performed applying Wilcoxon rank sum test and P values adjusted using FDR. Box plot showing the Shannon diversity between NEC and controls in the full cohort (g), secretors (h) and non-secretors (i). Group comparison was performed applying Wilcoxon rank

sum test. adj, adjusted; CTRL, control; DSLNT, disialyllacto-N-tetraose; NEC, necrotising enterocolitis.

Given that lower DSLNT was associated with NEC independent of secretor status, the utility of this HMO as a biomarker for NEC development was explored. Univariate ROC curve analysis determined that 241 nmol/ml (or 310.93 μg/ml) was the optimal DSLNT concentration in MOM for distinguishing NEC and control infants (Figure 3.3a). At this threshold, the area under the curve (AUC) was 0.947, with a sensitivity of 0.9 and a specificity of 0.9, correctly identifying 91% of infants with NEC (below threshold) and 86% of control healthy infants (above threshold).

To test if integration of additional HMOs could improve the classification performance, multivariate ROC curves built on increasing number of HMOs were performed (Figure 3.3b). Inclusion of two HMOs (the minimum in multivariate analysis) resulted in the optimal performance, with DSLNT being selected as a discriminatory feature in 100% of permutations (Figure 3.3c). 3FL and LNnT were the second and third most selected features, with a selection frequency of around 30%, being more abundant in cases of NEC. However, the integration of any additional HMOs to DSLNT in the multivariate model resulted in similar performance compared with the univariate model using DSLNT only (AUCs of 0.946 and 0.947, respectively).

To validate the 241 nmol/ml threshold defined in the current study in an independent cohort, we analysed data from Autran *et al.* (Autran et al., 2018) which contained 8 NEC and 40 matched control infants. Since this study included temporal sampling before disease, nearest milk sample to NEC onset were selected for each infant and control samples were matched by sample DOL and included only DSLNT concentration. Using a DSLNT threshold of 241 nmol/ml, the MOM sample for 100% (8/8) infants with NEC fell under the threshold, while 60% (24/40) control samples had a DSLNT concentration above 241 nmol/ml (Figure 3.3d).

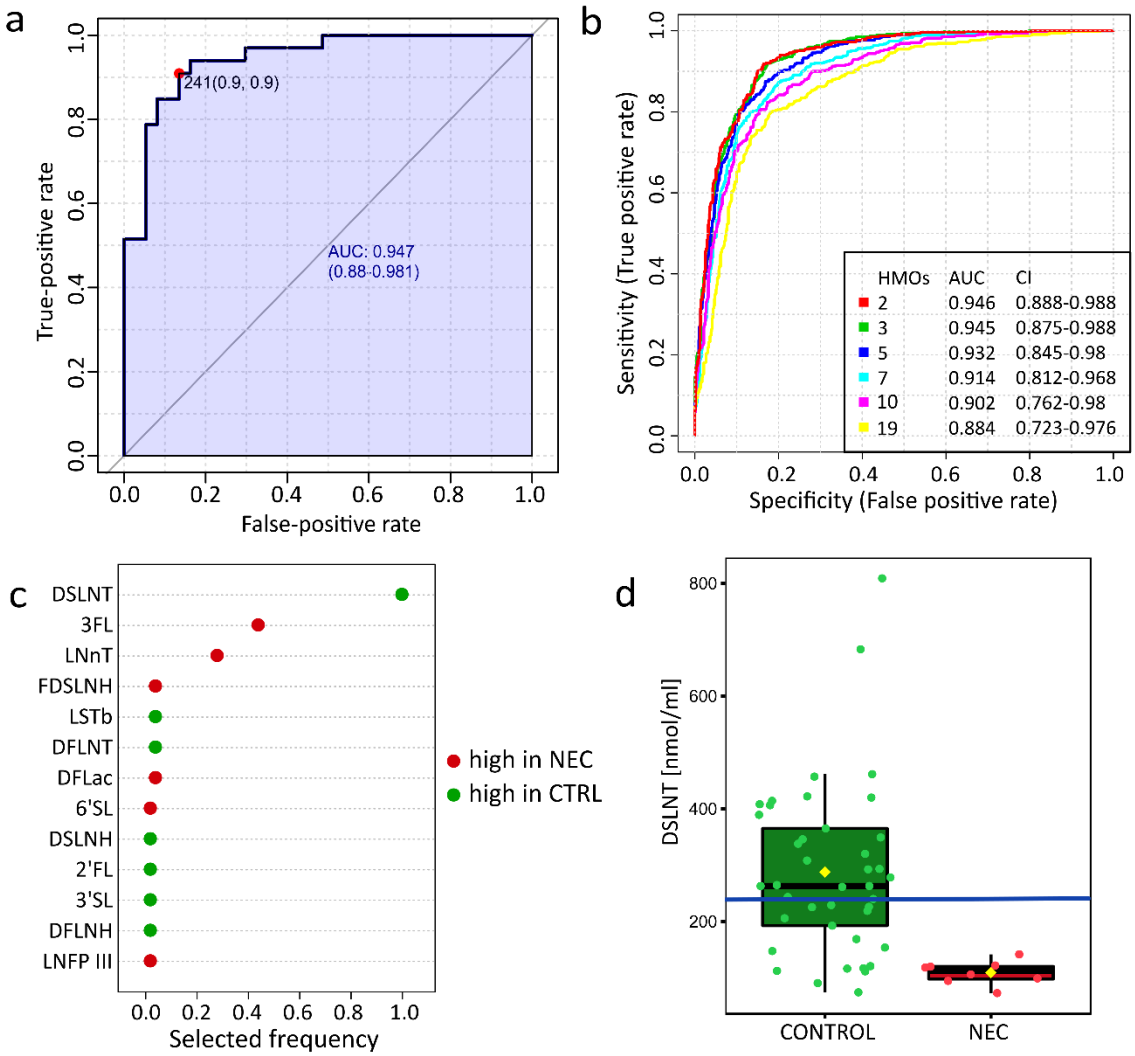

Figure 3.3. Human milk oligosaccharide (HMO) profiles were predictive of necrotising enterocolitis (NEC) status. (a) Univariate receiver operating characteristic (ROC) curve generated on DSLNT concentration identified 241 nmol/ml as the best threshold for NEC prediction. The performance of the classification is defined by the area under the curve (AUC), specificity (false positive rate) and sensitivity (false negative rate). (b) ROC curves generated using linear support vector machine (SVM) classification of HMO profiles between NEC and control groups. Increasing numbers of HMOs were included in the model and performance was described by the AUC. Two HMOs model gave the optimal performance. (c) Feature importance for the two HMOs model. Disialyllacto-N-tetraose (DSLNT) was selected as discriminatory feature in 100% of the permutations. (d) Box plot showing the concentration of disialyllacto-N-tetraose (DSLNT) between NEC and controls from the validation dataset, Autran et al. (2018). Blue line represents the 241 nmol/ml threshold. 2'FL, 2'-fucosyllactose; 3FL, 3-fucosyllactose; 6'SL, 6'-sialyllactose; adj, adjusted; AUC, area under the curve; CI, confidence interval; CTRL, control; DFLac, difucosyllactose; DFLNH, difucosyllacto-N-hexaose; DFLNT, difucosyllacto-N-tetraose; DSLNH, disialyllacto-N-hexaose; DSLNT, disialyllacto-N-tetraose; FDSLNH, fucodisialyllacto-N-hexaose; HMO, human milk oligosaccharide, 3'SL, 3'-sialyllactose; LNFP, lacto-N-fucopentaose; LNnT, lacto-N-neotetraose; NEC, necrotising enterocolitis.

Infants with NEC were then stratified by disease severity, and a comparison between medically managed NEC (NEC-M), where infants did not undergo surgery or die from NEC (i.e., had less severe disease) and infants with NEC that underwent surgery (NEC-S) was performed. NEC-M and NEC-S clustered together and were distinct from matched controls (2000 permutations, $p<0.001$; Figure 3.4a). Two HMOs were found to be significantly different, with DSLNT lower in MOM in both NEC-M (adj. $p<0.001$) and NEC-S (adj. $p<0.001$) compared with controls (Figure 3.4b). In addition, LNnT in MOM was significantly lower in NEC-S in comparison to both NEC-M (adj. $p=0.002$) and matched controls (adj. $p=0.042$) (Figure 3.4c). Subsequently, the potential association between DSLNT and LNnT concentrations and clinical variables was investigated by applying an adjusted linear model. DSLNT was negatively correlated to both disease types, with coefficients equal to -0.60 for NEC-M (adj. $p<0.001$) and -0.67 for NEC-S (adj. $p<0.001$) (Figure 3.4d). However, LNnT was not associated with disease type following adjusted linear modelling (both adj. $p>0.05$). DSLNT and LNnT were both significantly higher in secretor mothers (adj. $p=0.008$ and adj. $p<0.001$, respectively). DSLNT in MOM also positively correlated to gestational age (adj. $p=0.008$) and negatively to birth weight (adj. $p=0.008$). Neither HMO correlated to sex, delivery mode, postmenstrual age or DOL of the MOM sample (Figure 3.4d).

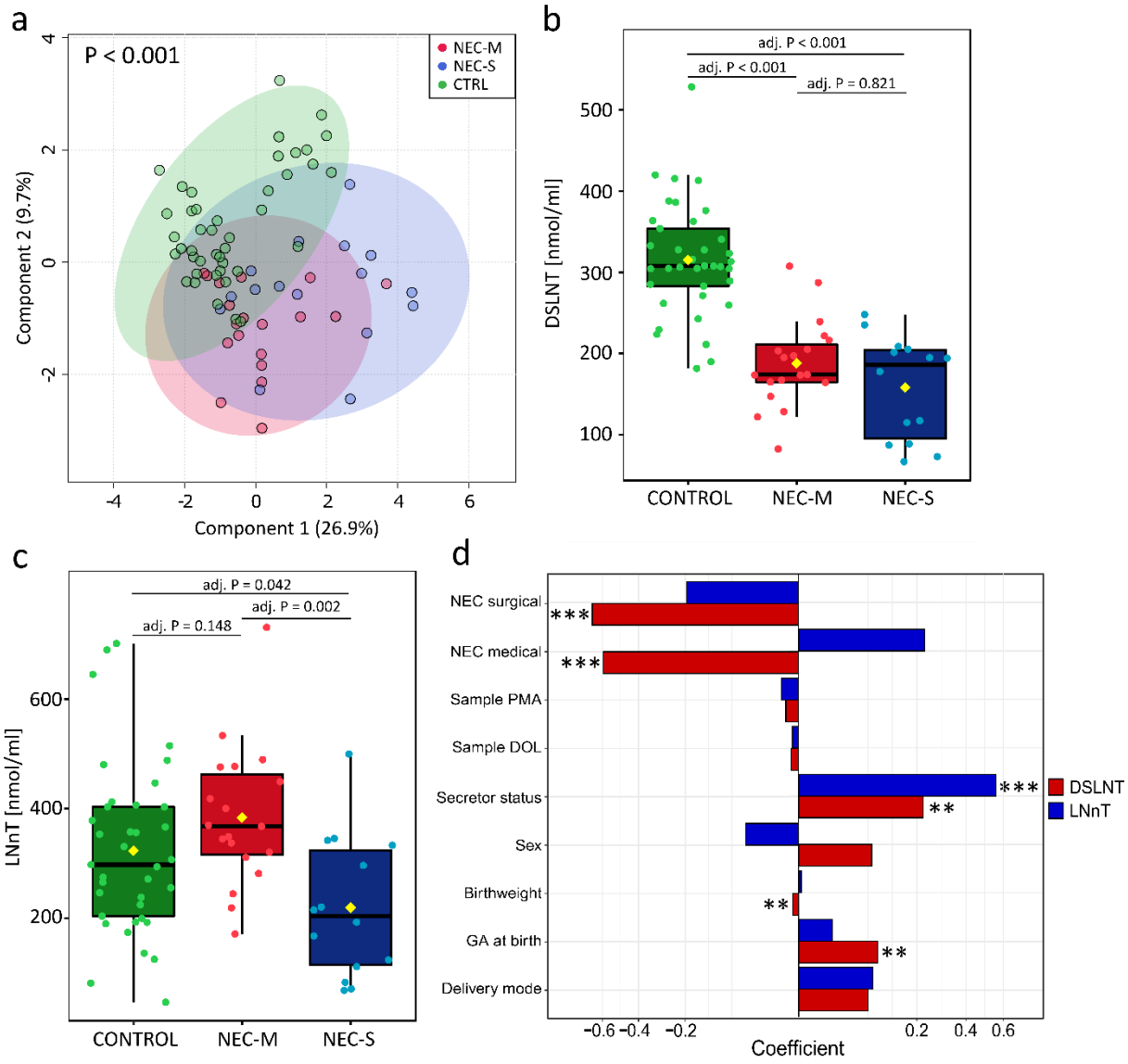

Figure 3.4. Analysis of HMO profiles with stratification of NEC-M and NEC-S. (a) Partial least squares discriminant analysis of HMO profiles from control, NEC-M and NEC-S infants. NEC-M and NEC-S clustered together and separately from controls ($p<0.001$). P values were calculated based on 2000 permutations. Box plots of (b) DSLNT and (c) LNnT concentration between control, NEC-M and NEC-S infants. Kruskal-Wallis followed by Dunn's test using Bonferroni adjustment was applied. (d) Adjusted linear regression model for DSLNT and LNnT including potential clinical confounders. P values were corrected by false discovery rate (FDR). Significant variables are indicated by asterisks: *** denotes FDR $p<0.001$; ** denotes FDR $p<0.01$. CTRL, control; DOL, day of life; DSLNT, disialyllacto-N-tetraose; GA, gestational age; HMO, human milk oligosaccharide; LNnT, lacto-N-neotetraose; NEC, necrotising enterocolitis; NEC-M, medically managed NEC; NEC-S, infants with NEC that underwent surgery; PMA postmenstrual age.

3.3 Human milk oligosaccharide in necrotising enterocolitis: comparison with focal intestinal perforation

HMO profiling data was also available for 7 infants diagnosed with FIP. NEC and FIP infants were matched to control infants independently, however, when combining all samples significant differences were found in clinical variables between the groups (Table 3.2). Gestational age and birthweight were significantly higher in the FIP group compared to both control (p=0.008 and p=0.01, respectively) and NEC (p=0.021 and p=0.016, respectively) infants. Vaginal delivery frequency was less frequent in FIP infants compared to the control group (p=0.009), but not compared to the NEC group.

Table 3.2. Demographics of the analytical cohort with human milk oligosaccharide profile data, including FIP. Differences between groups were tested applying Chi-square test and Dunn's post-hoc test where applicable.

	Control	NEC	FIP	NEC-CTRL	FIP-CTRL	FIP-NEC
Number of patients	37	33	7	-	-	-
Secretors	25 (68%)	20 (61%)	3 (43%)	0.544	0.213	0.388
Male	14 (38%)	22 (67%)	3 (43%)	0.016	0.803	0.237
Vaginal delivery	25 (68%)	17 (52%)	1 (14%)	0.171	0.009	0.072
Gestational age	25 [24; 26]	25 [24; 27]	30 [27.5; 30]	0.881	0.008	0.021
Birth weight	670 [585; 830]	670 [600; 840]	1400 [1035; 1420]	1.000	0.010	0.016
Probiotics ever	37 (100%)	31 (94%)	7 (100%)	0.129	-	0.504
MOM only	3 (8%)	6 (18%)	0 (0%)			
MOM+formula	11 (30%)	12 (37%)	3 (43%)	0.468	0.714	0.531
MOM+BMF	10 (27%)	7 (21%)	1 (14%)			
MOM+formula+BMF	13 (35%)	8 (24%)	3 (43%)			
DOL MOM sample	18 [12; 31]	18 [13; 34]	30 [16; 33.5]	0.636	0.713	1.000
DOL disease onset	-	19 [14; 35]	9 [5.5; 16.5]	-	-	-
NEC surgical	-	16		-	-	-

NEC, necrotising enterocolitis; FIP, focal intestinal perforation; MOM, mother's own breast milk; BMF, breast milk fortifier; DOL, day of life

No difference in the overall HMO profiles was observed for FIP vs. control, and FIP vs. NEC comparisons in the PLSDA analysis (Figure 3.5a). Consistent with earlier findings, only DSLNT concentration was significantly different between the groups compared (adj. p<0.001). Subsequent Dunn's post-hoc test revealed that NEC samples had lower concentrations of DSLNT compared to both control (adj. p<0.001) and FIP (adj. p=0.046) groups, while no difference was found between control and FIP samples (Figure 3.5b). Moreover, when using the 241 nmol/ml threshold defined previously to separate NEC and control infants, 5/7 (71%) of FIP infants were correctly classified as non-NEC infants. Interestingly, the 2 misclassified FIP infants had initially been diagnosed with NEC and then re-classified to FIP after histology.

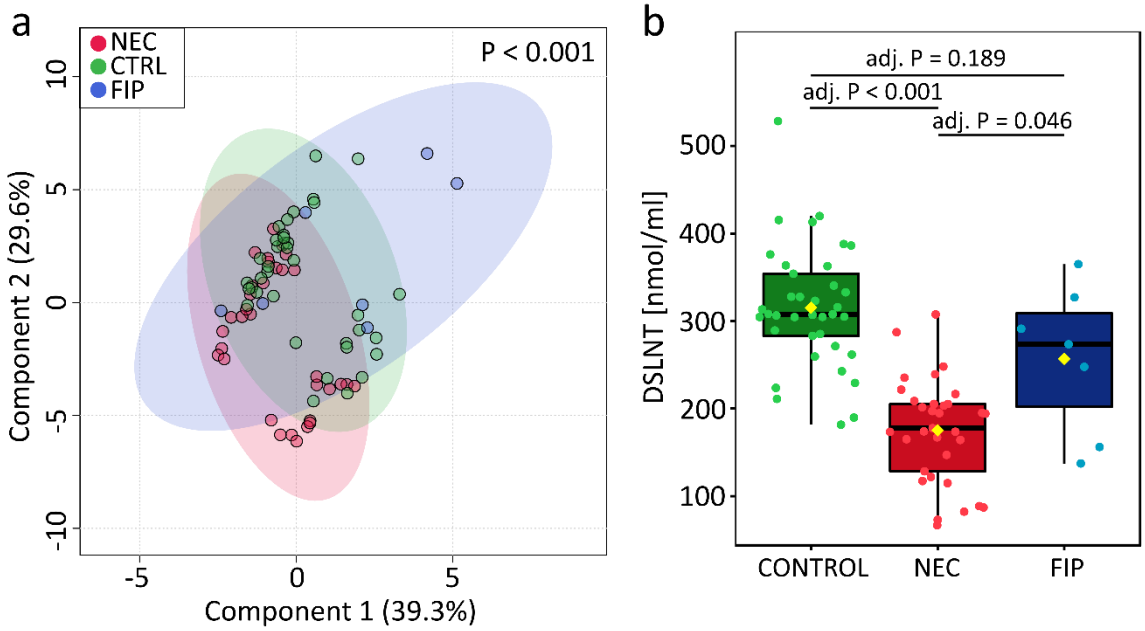

Figure 3.5. Disialyllacto-N-tetraose (DSLNT) is lower in necrotising enterocolitis (NEC) infants compared to control and focal intestinal perforation (FIP) groups. (a) Partial least squares discriminant analysis (PLS-DA) of log-normalised HMO profiles from NEC, FIP and control infants. NEC and control groups harboured a different HMO profile (P < 0.001), while no difference was found between FIP-NEC and FIP-control. P values were calculated based on 2000 permutations. (b) DSLNT concentration in maternal milk was lower in NEC infants compared to both FIP and control group. P values were generated applying Dunn's post-hoc test with Bonferroni adjustment. NEC, necrotising enterocolitis; FIP, focal intestinal perforation; DSLNT, disialyllacto-N-tetraose; adj., adjusted.

3.4 Infant gut microbiome in necrotising enterocolitis

HMOs are known to shape the infant gut microbiome, and to advance our work on HMOs in NEC, we further explored the preterm intestinal microbiome development where metagenomic data was available through a collaboration with Astarte Medical. This included 644 stool samples from 34 controls and 14 infants with NEC which again differed for the percentage of

male infants in each group (Table 3.3), as reported for the full cohort with HMO profile data available (Table 3.1).

Table 3.3. Sub-cohort of infants with longitudinal metagenome data from stool samples. Differences between groups were tested applying Chi-square test and Wilcoxon rank test where applicable.

	Control	NEC	P value
Number of patients	34	14	-
Number of stool samples	449	195	-
Secretors	23 (68%)	10 (71%)	0.797
Male	12 (35%)	11 (79%)	0.006
Vaginal delivery	24 (71%)	7 (50%)	0.175
Gestational age	25 [24; 26]	25 [24; 26.75]	0.908
Birthweight	645 [586.3; 747.5]	670 [562.5; 735]	0.447
Probiotics ever	34 (100%)	13 (93%)	0.871
MOM only	2 (6%)	1 (7%)	0.315
MOM + Formula	10 (29%)	8 (57%)	
MOM + BMF	9 (27%)	2 (14%)	
MOM + Formula + BMF	13 (38%)	3 (22%)	
DOL NEC onset	-	28 [13; 51]	-
NEC surgical	-	4	-

NEC, necrotising enterocolitis; MOM, mother's own breast milk; BMF, breast milk fortifier; DOL, day of life

To overcome challenges of repeated measures and to compare results with existing published work, the initial analysis included one stool sample per infant closest to NEC onset (median of 3 days before NEC) and a corresponding control sample matched by DOL (Figure 3.6). This cross-sectional analysis showed infants with NEC had significantly lower richness (p=0.027) but comparable Shannon diversity (p=0.443, Figure 3.7a). Bray-Curtis Principal Coordinates Analysis (PCoA) showed no significant difference between the bacterial profiles of NEC and controls (PERMANOVA p=0.182, Figure 3.7b). Analysis at the phylum level showed significantly lower relative abundance of Actinomycetota (adj. p=0.034) and higher relative abundance of Pseudomonadota (adj. p=0.034) in infants with NEC (Figure 3.7c). Correspondingly, at the species level, infants with NEC had lower relative abundance of

B. longum (adj. p=0.012) and higher relative abundance of *Enterobacter cloacae* (adj. p=0.012) compared with controls (Figure 3.7d).

Figure 3.6. Sampling schematic for the entire cohort. Only samples collected in the first 100 days of life are shown. Shapes represent sample timing in relation to the diagnosis of disease; control (n = 37) and necrotising enterocolitis (NEC; n = 33). Colours indicate if the sample on that day of life was a maternal breast milk human milk oligosaccharide (HMO) or infants stool metagenome, or if both sample/data types were generated from each sample collected on that day.

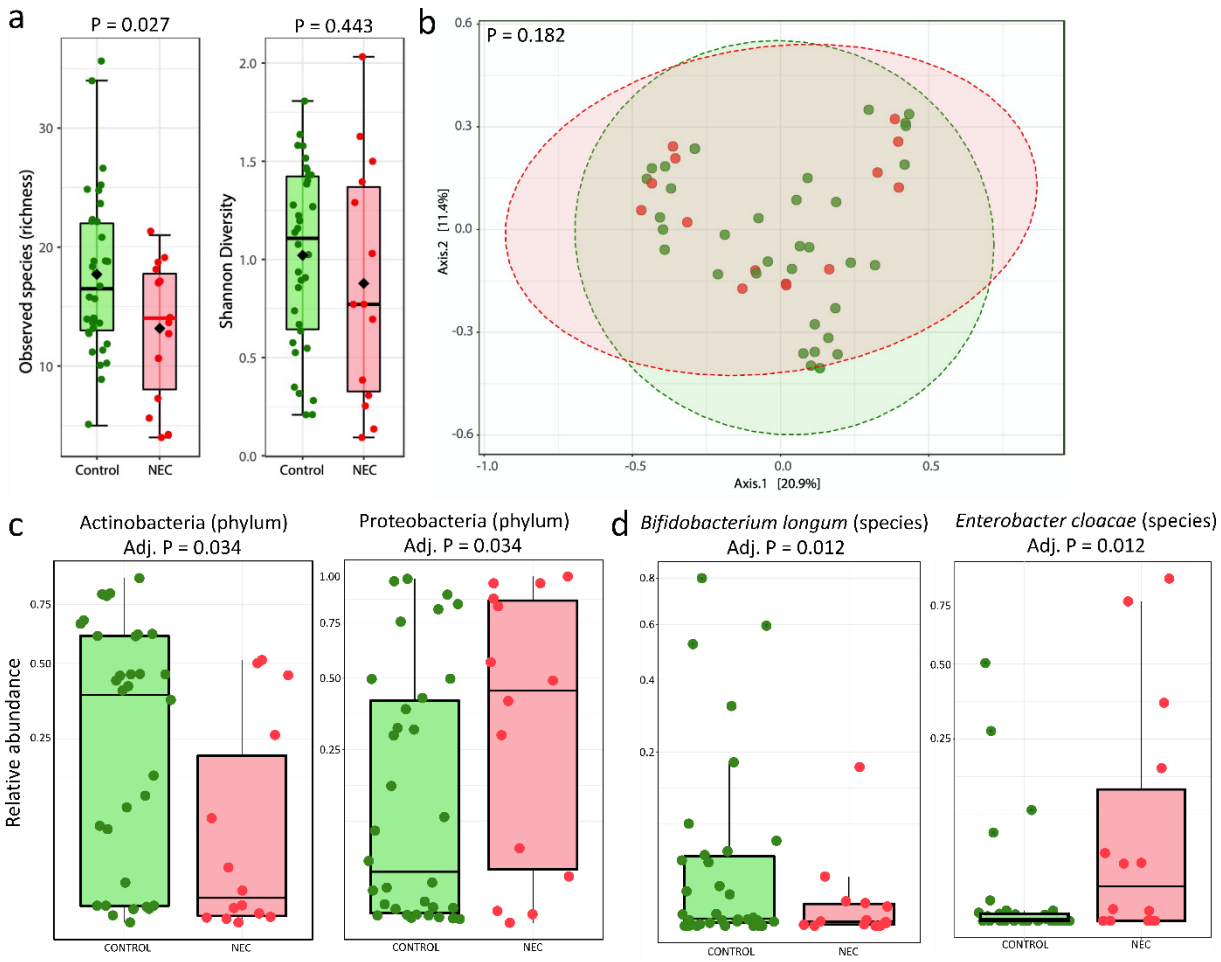

Figure 3.7. Cross-sectional analysis of preterm stool metagenome profiles between NEC and matched controls. Analysis includes the sample closest to NEC onset (median of 3 days prior to NEC) and a corresponding control sample matched by day of life. (a) Alpha diversity based on observed species (richness) and Shannon diversity. (b) Bray-Curtis principal coordinate analysis. (c) Box plots showing the relative abundance of significant phyla. (d) Box plots showing the relative abundance of significant species. NEC, necrotising enterocolitis. The figure adopts the old taxonomy names, and here is a list of the new taxonomy: Proteobacteria is now Pseudomonadota; Actinobacteria is now Actinomycetota.

3.5 Infant gut microbiome in necrotising enterocolitis and its correlation with disialyllacto-N-tetraose

DMM clustering was used to determine PGCTs using species-level data, and five PCGTs were deemed optimal (Figure 3.8a). PGCT-1 was characterised by high relative abundance of *Staphylococcus* spp. and *Enterococcus faecalis*, PGCT-2 had high *Escherichia* spp; PGCT-3 had high *Klebsiella* spp; and PGCT-4 and PGCT-5 had high *Bifidobacterium* spp, with *B. breve* notably high in PGCT-5. Using the PGCT clusters, the temporal transition of an infant's gut microbiome over the first 60 days of life was analysed by defining distinct time points and including only one sample per infant at each time point. Based on the distribution of samples across all time points and all clusters, the temporal transition of the microbiome over the first

60 days of life was significantly different in infants in receipt of MOM below the DSLNT threshold of 241 nmol/ml compared with infants above the DSLNT threshold ($\chi 2$ test, $p<0.001$; Figure 3.8b).

The PGCTs were named according to the average age of samples within that cluster, where PGCT-1 contained on average the earliest samples and PGCT-5 on average the latest samples. The number of samples from all time points in only PGCT-1 and PGCT-5 were compared to investigate associations between the MOM DSLNT threshold and gut microbiome development from the typically younger to the typically older PGCTs. Infants receiving MOM with DSLNT level below 241 nmol/ml had significantly more samples remaining within PGCT-1 throughout all time points (78% in PCGT-1 vs 22% in PGCT-5, $\chi 2$ test $p<0.001$), whereas infants receiving MOM with DSLNT above this threshold transitioned from PGCT-1 to PGCT-5 as demonstrated by a similar number of samples in each PGCT across all time points (48% in PCGT-1 vs 52% in PGCT-5, $\chi 2$ test $p=0.717$) (Figure 3.8b). In addition to comparing samples from all time points, samples from the final time point only were next compared (i.e., DOL 50–60). After correcting for uneven frequency of sampling between groups, at the final time point, infants receiving MOM above the DSLNT threshold were twice as likely to be in PGCT-5 (3/11 samples below vs 12/22 samples above DSLNT threshold; odds ratio (OR) 3.20, 95% confidence interval (CI) 0.6657 to 15.3819), which was characterised by high relative abundance of *Bifidobacterium* (Figure 3.8b).

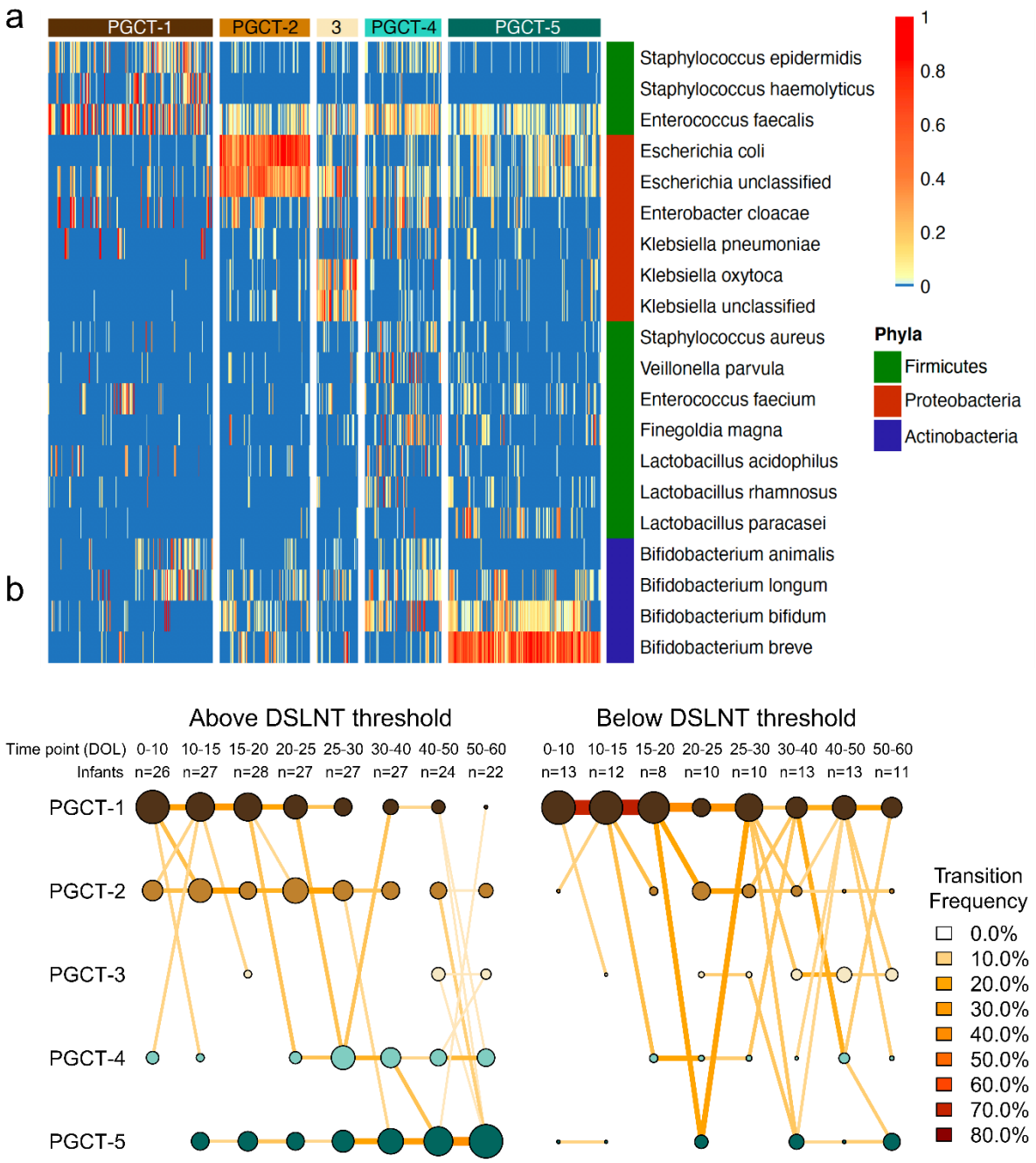

Figure 3.8. Analysis of PGCTs by infants receiving maternal milk above or below the 241 nmol/ml DSLNT threshold. The entire dataset of 644 samples formed five distinct clusters based on lowest Laplace approximation following Dirichlet multinomial clustering. (a) Heatmap showing the relative abundance of dominant bacterial species within each PGCT cluster. The phyla for each species are also shown. (b) Transition model showing the progression of samples through each PGCT, from day of life 0–60 across eight distinct time points. Plots are separated based on whether the concentration of DSLNT in maternal milk was above or below the 241 nmol/ml threshold. Nodes and edges are sized based on the total counts. Nodes are coloured according to Dirichlet Multinomial Mixtures (DMM) cluster number and edges are coloured by the transition frequency. Transitions with less than 5% frequency are not shown. DSLNT, disialyllacto-N-tetraose; PGCT, preterm gut community type. The figure adopts the old taxonomy names, and here is a list of the new taxonomy: *Lactobacillus rhamnosus* is now *Lacticaseibacillus rhamnosus: Lactobacillus casei* is now *Lacticaseibacillus*

casei: Firmicutes is now Bacillota; Proteobacteria is now Pseudomonadota; Actinobacteria is now Actinomycetota.

The performance of random forest classification models built on the cross-sectional subset of HMO profile data, metagenomic sequencing data and the two datasets combined to classify an infant as NEC or healthy was investigated, given that all this information is available before onset of disease and could therefore function as a risk stratification system in clinical practice. The HMO profile alone had a classification error of 0.146, with 21% (3/14) NEC and 12% (4/34) control infants misclassified. DSLNT had the greatest contribution to classification with MDA of 0.11. Other HMOs contributing to classification accuracy included LNH (MDA=0.012) and DFLNH (MDA=0.011), which were non-significantly higher in infants with NEC. Random forest generated using the metagenomic sequencing data was characterised by a classification error of 0.229, with 43% (6/14) NEC and 15% (5/34) control infants misclassified. *E. cloacae* was the most important feature guiding the classification (MDA=0.036), with higher relative abundance in infants with NEC, followed by *B. bifidum* (MDA=0.024) and *B. longum* (MDA=0.013), which had higher relative abundance in control infants. Combining HMO and metagenome datasets slightly improved the performance compared with using HMOs alone, with 21% (3/14) of infants with NEC and (9%) 3/34 controls misclassified. In this combined model, DSLNT was enriched in controls, and DSLNH and the relative abundance of *Escherichia* unclassified were higher in infants with NEC (Figure 3.9).

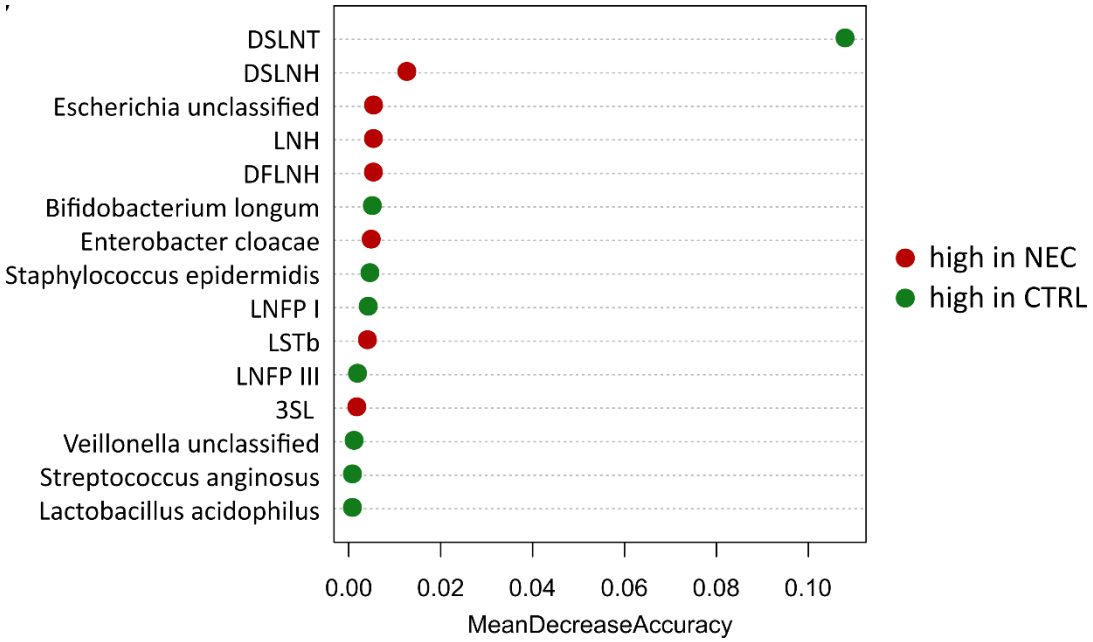

Figure 3.9. Modelling of cross-sectional HMO and infant stool metagenomic profiles using random forest. Feature importance from combined HMO and metagenome random forest classification model. Mean decrease accuracy value defines the contribution given by a certain feature to classification process. CTRL, control; NEC, necrotising enterocolitis; DSLNT,

disialyllacto-N-tetraose; DSLNH, disialyllacto-N-hexaose; DFLNH, difucosyllacto-N-hexaose; LNFP I, lacto-N-fucopentaose I; LSTb, sialyllacto-N-tetraose b; LNFP III lacto-N-fucopentaose III; 3'SL, 3'-sialyllactose.

3.6 Discussion

HMO profiling of MOM and metagenomic sequencing of preterm stool were used to investigate their correlation with NEC. The main factor influencing the HMO profile was the mother's secretor status, with differences in 16/19 HMOs measured observed between secretor and non-secretor MOM. However, one single HMO, DSLNT, was lower in MOM fed to infants who developed NEC independently of secretor status, confirming the previous smaller studies (Van Niekerk et al., 2014, Autran et al., 2018). The univariate ROC curve model identified 241 nmol/ml as the optimal DSLNT concentration to separate NEC and control infants with a 0.9 specificity and sensitivity. An additional HMO, LNnT, was found lower in infants who developed a more severe NEC compared to the less severe NEC manifestation, suggesting its potential contribution to disease protection. An altered gut microbiome development was also found in infants receiving low DSLNT, with a decreased progression toward the PGCT abundant in *Bifidobacterium* spp. and typical of older infants. Moreover, a lower relative abundance in *B. longum* was observed in NEC infants prior to disease onset. These results strengthen the idea that bifidobacteria might exert a protective role against NEC development.
Multiple factors might influence MOM composition and infant gut microbiome, thus infants with disease were matched by various confounders to control infants. Variables used for matching included gestational age, sample DOL and delivery mode. NEC group harboured higher percentage of male infants compared to the control group, however, this has been reported before in other studies (van Westering-Kroon et al., 2021). Instead, FIP group was characterised by higher gestational age and birthweight compared to both NEC and control groups, and less frequency of vaginal birth compared to the control group. Moreover, FIP occurs very early in life and the median DOL of onset in our cohort was day 9, while the median DOL of the MOM sample analysed was day 30. The potential role of MOM in FIP protection was thus hard to investigate. These factors must need to be taken into account when interpreting the data and no effective investigation of MOM protection against FIP could be performed. Despite these downfalls, comparison of the two diseases might still give interesting hypotheses which will need to be tested in a dedicated study.
The HMO results from the current study build on previous findings in humans, showing reduced DSLNT in MOM received by infants developing NEC (Van Niekerk et al., 2014, Autran et al., 2018). This is also supported by rodent studies where total and individual HMOs including

2'FL and DSLNT have shown a protective effect against NEC development (Jantscher-Krenn et al., 2012, Good et al., 2016, Autran et al., 2016). However, 2'FL and mixtures of HMOs (one of which included DSLNT) did not show any protection in NEC piglet models (Rasmussen et al., 2017, Cilieborg et al., 2016). Importantly from a clinical perspective, in rats, the protection provided by pooled HMOs could be reproduced with DSLNT alone, with specific dependence on its precise structure since closely related sialyllacto-N-tetraose (identical in structure to DSLNT but lacking one sialic acid residue) did not provide protection, suggesting a highly structure-specific mechanism (Jantscher-Krenn et al., 2012). The reduction in DSLNT was independent of maternal secretor status, which was the only factor in this dataset and was significantly associated to HMO profile overall, but also to this HMO singularly. Our findings further extend the evidence for the specificity of DSLNT in the NEC pathway, as MOM fed to FIP infants showed higher DSLNT concentration to the NEC group, while no difference was observed comparing FIP to the control group. Using a 241 nmol/ml threshold of DSLNT from a single MOM sample correctly identified 91% of infants with NEC (below DSLNT threshold), 86% of control healthy infants (above DSLNT threshold) and 71% of FIP infants. Of the three infants who developed NEC who received MOM with DSLNT above the threshold, two had not received MOM in the 3 weeks prior to disease onset, and the remaining infant had a DSLNT concentration of 248 nmol/ml.

Within the validation dataset (Autran et al., 2018), 100% of infants with NEC were correctly classified, but only 60% of controls. Making a robust diagnosis of NEC is difficult, and it is possible that the specific threshold value of DSLNT we identified will have a different predictive value in other populations, or where other criteria are used to determine the presence of disease. Our study contains a large number of cases coded clinically as NEC independently validated by blinded review. In addition, our cohort is more homogenous (predominantly white Caucasian) and the concentration of DSLNT was less variable (current study IQR 184–321 nmol/ml vs Autran *et al.* IQR 122–346 nmol/ml) despite using the same analytical platform. Given HMO composition and DSLNT concentrations may be influenced by genetic factors, geographical location, ethnicity (McGuire et al., 2017) and seasonality (Azad et al., 2018), differential thresholds may improve diagnostic performance in other settings. Taken together, this external validation and potential variation in DSLNT concentration by maternal factors underscore the need for large multicentre studies to both refine a universal or stratified threshold for DSLNT concentration in predicting NEC and potentially prospectively identifying milk samples that may benefit from supplementation with synthetically produced DSLNT.

Information on the NEC severity was also available in this study. In contrast to the study from Autran *et al.* (2018), which found Bell stage 1 NEC having higher DSLNT concentration compared to Bell stages 2 and 3, this study found no differences in DSLNT concentration depending on

disease severity. However, Autran *et al.* (2018) classified NEC cases depending on Bell stages, for which Bell stage 2 and 3 identifies infants with definite NEC, while Bell stage 1 identifies infants with suspected NEC. In this cohort instead, all disease infants were diagnosed with definite NEC and disease severity was defined depending on treatment. Beyond DSLNT, LNnT was found to be lower in MOM given to NEC infants who required surgery compared to NEC cases managed medically. However, correlation between NEC and LNnT was lost when the model was adjusted for other clinical variables, potentially as a consequence of the low numbers in each NEC group after stratification for severity. Overall, this suggests DSLNT to be the major factor protecting against NEC development, but also that other HMOs might participate in determining disease severity.

In addition to HMO profiles, our extensive longitudinal stool metagenomic analysis represents one of the largest datasets to date. This extends the Stewart lab previous work (Stewart et al., 2016, Stewart et al., 2017a, Stewart et al., 2017b), where DMM was used to facilitate analysis of temporal microbiome development and which was used here to integrate the HMO DSLNT threshold of 241 nmol/ml with infant gut microbiome profiles. We observed a difference in microbiome development between DSLNT groups, with infants receiving MOM with lower DSLNT tending to have delayed progression into the PGCT typically expected in older infants (i.e., PGCT-5). This supports the theory that concentrations of specific HMOs in MOM are associated with differences in gut microbiome development. On the contrary, transition into PGCT-5 was twice as likely in infants receiving MOM with DSLNT above the threshold, which was characterised by high relative abundance of *Bifidobacterium* spp. *Bifidobacterium* has previously been linked to health in preterm infants, (Stewart et al., 2016, Stewart et al., 2017b, Torrazza et al., 2013) and our current findings in pre-NEC samples further support the association of reduced *Bifidobacterium* spp., specifically *B. longum*, as a risk factor for NEC. In addition, our species-level metagenome data advanced previous associations of Enterobacteriaceae with NEC (Sim et al., 2015, Morrow et al., 2013, Mai et al., 2011), showing *E. cloacae* relative abundance was higher before NEC.

Random forest analysis confirmed the capability of HMO profiles to identify infants who developed NEC and slightly outperformed metagenome profiles by correctly classifying three more NEC cases and one more control. Combining HMO and metagenome data before disease accurately classified 87.5% of infants as healthy or having NEC, with DSLNT and the bacterial species identified as important in the random forest analysis being comparable to the unsupervised analysis in the current study and in previous studies. Further work is needed to determine if DSLNT functions via modulation of the microbiome or by acting directly on the host, such as acting in a structure-specific receptor-mediated way to alter immune functioning and to reduce inflammation leading to necrosis. In the event of the latter, a microbial community with less use of DSLNT could provide an advantage to reducing NEC risk.

3.1 Study limitations and conclusions

This study found various novel results, but some limitations should be considered. First, the cross-sectional HMO profiling data precluded assessment of changes within mothers over time and how this may relate to NEC development. The milk sample was selected based on the day of infant feeding, and the actual expression of milk may have occurred several days earlier, which may be important clinically. Moreover, milk samples were collected from the feeding syringe after the feed, which can last up to 24 hours, and thus the HMO profile assessed might have been different to the original milk sample. Further studies will need to include HMO profile comparison between fresh and frozen MOM and the same sample collected from the infant feeding tube to be able to better interpret the data discussed in this thesis. Moreover, current published data suggest that the concentration of HMOs, including DSLNT, is relatively stable over time (Austin et al., 2019), but validation in longitudinal preterm cohorts involving multiple NICUs is needed. The amount of MOM an infant receives and tolerates each day is variable, and DSLNT exposure is dependent on both concentration and volume. Although this study identifies DSLNT concentration alone may be useful from both a diagnostic and therapeutic perspective, further studies should consider the volume of milk received in addition to concentration.

FIP infants were included in the analysis as their HMO profile data was available. However, major differences were present in clinical variables between FIP infants and NEC and control groups, including gestational age and birthweight. While no difference in sample DOL were present between the groups, the median DOL of FIP onset was day 9, while the median DOL of the MOM samples analysed was day 30. Thus, it could not be evaluated if a potential difference in HMO profile was present in MOM given to FIP infants prior to disease development. This thesis aimed to investigate HMOs in NEC in comparison to healthy infants, and a study directed to understanding the role of HMOs in FIP infants is warranted.

Inclusion of metagenome data was opportunistic based on available data, and cost prohibited sequencing all infants in the cohort. As such, the classification accuracy of the model might be impacted by the reduced sample size in comparison to the full HMO cohort, necessitating the need for follow-up analyses in larger cohorts. Despite this, the sample size of 644, including 195 samples from 14 preterm infants who developed NEC, makes this dataset one of the largest published to date. Furthermore, the gene relative abundance data warrant further investigation, in combination with other experimental approaches, to help inform the HMO use capacity of different strains.

Taken together, the current findings and recent work highlighting the ability of *Bifidobacterium* spp. to use HMOs is strain specific (Lawson et al., 2019, Gotoh et al., 2018) underscore the need for further research to better understand the complexity of human milk and other nutritional exposures, including the use of supplements such as prebiotics and probiotics in preterm infants. In addition to therapeutics, the classifiers may provide a basis for the development of biomarkers predicting NEC risk. While additional work is needed, the addition of microbial biomarkers may allow for the most accurate predictions and could inform NEC risk for infants where MOM (and thus HMO information) is not available.

In summary, HMO profiling of MOM coupled to metagenomic sequencing of preterm stool showed that the concentration of a single HMO, DSLNT, was lower in milk received by infants who developed NEC. The lower concentration of DSLNT was associated with altered microbiome development, specifically a reduced progression toward PGCT-4 and PGCT-5 typically found in the older infants and abundant in *Bifidobacterium* spp. These results suggest MOM HMO profiling may provide potential targets for biomarker development and disease risk stratification. They may also guide focused donor milk use (e.g., prioritise high DSLNT for preterm infants) and novel avenues for supplements that may prevent this life-threatening disease.

Chapter 4. Mom's own milk microbiome in necrotising enterocolitis and its correlation with human milk oligosaccharide profile and infant gut microbiome

4.1 Introduction

MOM microbiome, HMO profiles, and infant gut microbiome, have all been linked to one another in various studies. In term infants, MOM microbiome seems to be a potential provider of bacteria which can colonise the infant gut (Solis et al., 2010, Bogaert et al., 2023), and HMOs shape the neonatal gut bacterial community acting as prebiotics and favouring *Bifidobacterium* spp. colonisation (Masi and Stewart, 2022). Moreover, HMOs might also modulate the MOM microbial community and vice versa, but the relationship between these two breast milk components requires further study (Moossavi et al., 2019a, Ramani et al., 2018, Hunt et al., 2012). Notably, most studies investigating these connections have been performed in term infants, which are likely distinct to their premature counterparts. Preterm infants receive multiple antibiotics at birth, they are more likely to be born by C-section, are kept in 'sterile' incubators, they cannot be directly breastfed, and mother-infant contact is disrupted, resulting in an altered gut microbiome establishment compared to term infants.

To date, studies focused on the role of MOM microbiome in NEC are still lacking. Stewart *et al.* (2013) investigated the relationship between preterm twins' microbiome and MOM microbiome, including twin pairs discordant for NEC development (Stewart et al., 2013a). To the best of my knowledge, this study is the only one to date including MOM microbiome from NEC infants, however, no direct investigation of MOM microbiome correlation with NEC was performed. Additionally, the potential relationship between MOM microbiome, HMO profiles, and preterm infant gut microbiome have not been investigated yet. Previous work, including the results presented in chapter 3, have reported an alteration in the HMO profile of MOM fed to NEC infants, but whether this difference is correlated with the MOM microbiome has not been explored (Autran et al., 2018). Thus, in this chapter we investigated the MOM microbiome in an expanded cohort of NEC and matched control infants, which included the babies from the cohort described in chapter 3 (same sample was used where possible, or a different sample was included if available) and additional new NEC and control infants. Cross-sectional MOM microbiome was determined using amplicon-based 16S rRNA gene sequencing, owing to the low biomass nature of this sample type. The potential association of MOM microbiome with NEC development was investigated for all available samples. In a subset of infants for whom

HMO profile, MOM microbiome and gut metagenome were available, an integration of the three datasets was performed applying correlation analysis.

4.2 Mom's own milk microbiome in necrotising enterocolitis

A total of 126 MOM samples were sequenced from 72 control and 54 NEC infants, with the number of reads ranging from 7 to 138,913. Samples were rarefied at 1,231 reads, retaining 62 control and 48 NEC samples. 7 kit negative controls were also sequenced, obtaining between 1 and 17 reads and so subsequently lost during rarefaction, suggesting that no contamination happened during extraction or sequencing. Information on the cohort which was analysed is reported in Table 4.1. No significant difference was found between NEC and control groups for the variables listed.

Table 4.1. Demographics of the analytical cohort with cross-sectional 16S rRNA gene sequencing of MOM samples. Differences between groups were tested applying Chi-square test and Wilcoxon Rank test where applicable.

	Control	**NEC**	**P value**
Number of patients	62	48	-
Male	30 (48%)	32 (65%)	0.074663
Vaginal delivery	36 (58%)	27 (55%)	0.754404
Gestational age	26 [24; 27]	25 [24; 27]	0.5923
Birthweight	800 [640; 900]	713 [591.5; 855]	0.1339
DOL breast milk sample	19 [14.25; 28.75]	21 [14; 27.5]	0.6395
DOL disease onset	-	16 [12; 29]	-
HMO profile available	36	30	

NEC, necrotising enterocolitis; HMO, human milk oligosaccharide; DOL, day of life.

The dataset was analysed at phylum and genus level, the latter being the deepest level that can be confidently and accurately assigned using V4 16S rRNA gene sequencing. The samples carried a median number of 10 OTUs (IQR 6-15) and a median Shannon diversity of 0.63 (IQR 0.27-1.18) (Figure 4.1a). At the phylum level, Bacillota dominated the preterm MOM samples, followed by Pseudomonadota and Actinomycetota (Figure 4.1b). At the genus level the 10 most abundant bacteria in rank order were *Staphylococcus*, *Acinetobacter*, *Enterobacter*,

Pseudomonas, *Enterococcus*, *Corynebacterium*, *Cutibacterium*, *Finegoldia*, *Streptococcus* and *Bifidobacterium* (Figure 4.1c). The existence and viability of 9 of the top 10 most abundant genera was confirmed using culturing techniques and will be discussed in chapter 5.

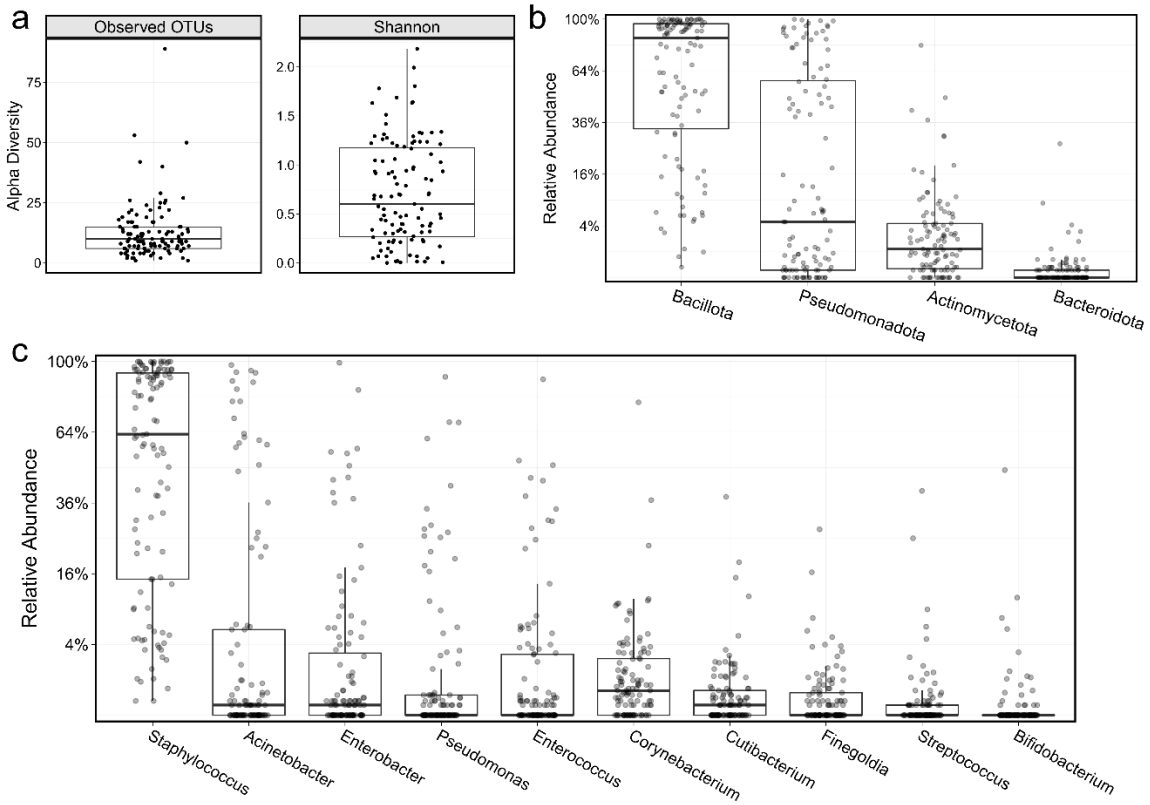

Figure 4.1. Overview of the preterm MOM 16S rRNA gene profile. (a) Box plots showing the alpha diversity based on observed OTUs (richness) and Shannon diversity. Box plots showing the relative abundance of phyla (b) and of the top 10 most abundant genera (c). OTU, operational taxonomic unit.

The variation in MOM microbiome profile over the first 50 days of life was investigated. At the alpha diversity level, while the number of OTUs remained stable over time, an increase in Shannon diversity was observed at early DOL with a peak reached at day ~15, after which a decrease was observed with lowest value found at DOL ~30 after which the diversity increased again (Figure 4.2a). At the phylum level, Bacillota relative abundance stayed stable over time (Figure 4.2b). Pseudomonadota increased in relative abundance over the first two weeks of life while Actinomycetota decreased rapidly during the same time period, after which the relative abundances for both phyla remained stable (Figure 4.2b). Further analysing the composition at the genus level, *Staphylococcus* relative abundance slightly decreased over time (Figure 4.2c). *Acinetobacter* slightly increased up to ~10 DOL and remained stable until day ~30 after which its relative abundance decreased. *Enterobacter*, *Pseudomonas* and *Enterococcus* relative abundances showed a similar trend, with an increase during the first ~15-18 days of life, reaching the peak around the day for which the highest Shannon diversity was observed, before

decreasing in their relative abundances (Figure 4.2c). After day ~25 *Pseudomonas* relative abundance remained stable, while an increase in *Enterobacter* and *Enterococcus* was found (Figure 4.2c).

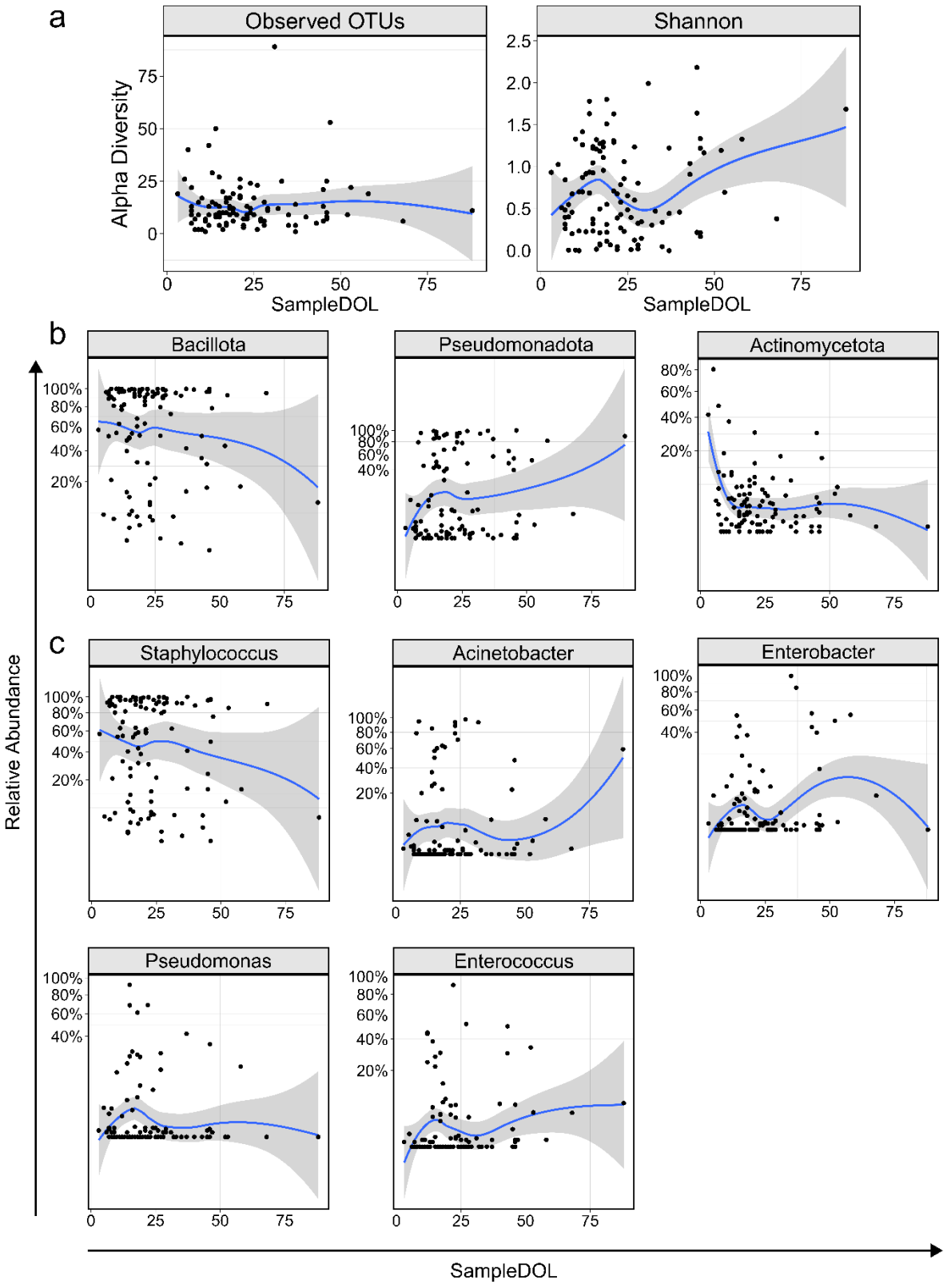

Figure 4.2. Descriptive overview of the preterm MOM 16S rRNA gene profile in the first 50 days of life. Local regression fits (95% CIs shaded in grey) over time for alpha diversity expressed as observed OTUs (richness) and Shannon diversity (a), phyla (b), and top 5 most abundant genera in MOM (c). OTU, operational taxonomic unit; DOL, day of life.

We then sought to determine if a difference in MOM microbiome was present between NEC and control infants. First, the alpha diversity was analysed finding no significant difference in either richness (median NEC 9.5 (IQR 6-15) vs median control 10 (IQR 7-16.5; P = 0.7) or Shannon diversity (median NEC 0.67 (IQR 0.22-1.13) vs median control 0.56 (IQR 0.31-1.19); P = 0.81) (Figure 4.3a). We next explored beta-diversity in MOM using weighted Bray-Curtis dissimilarity, finding samples from NEC and control infants had no separation (P value = 0.937; Figure 4.3b). Taxonomic profile further supported that the MOM microbiome was comparable between NEC and control infants, finding no difference in the relative abundance of any phylum (all adj. P >0.05; Figure 4.1c) or genus (all adj. P >0.05; Figure 4.1d).

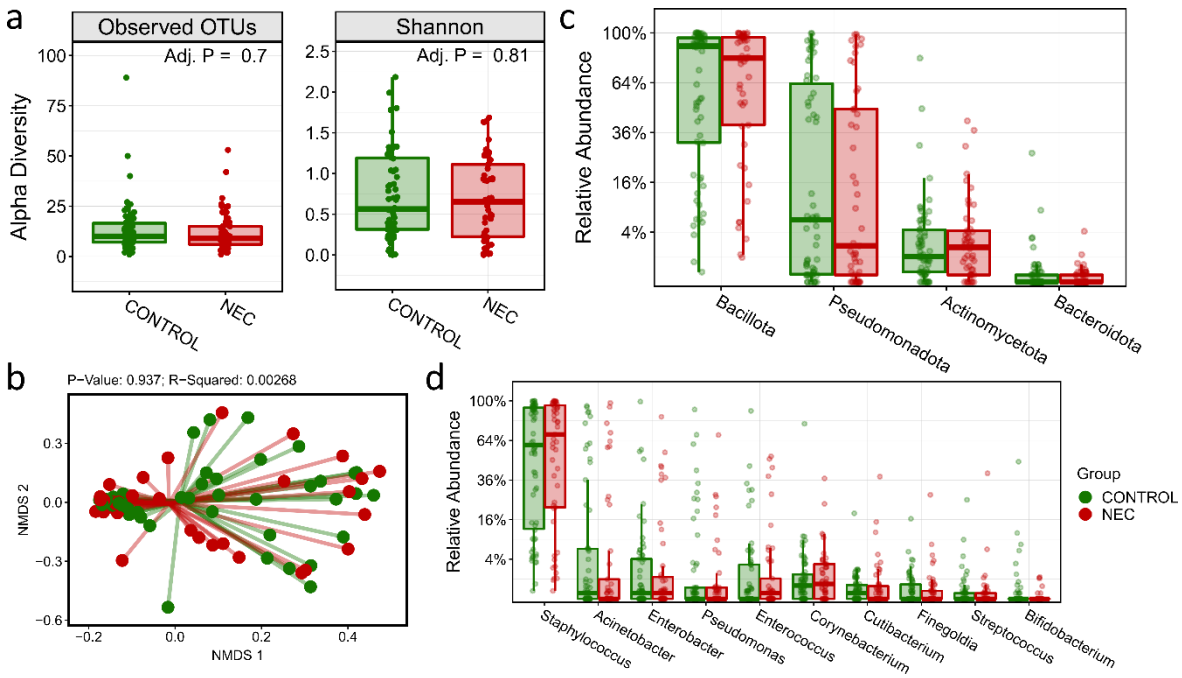

Figure 4.3. Cross-sectional analysis of preterm MOM 16S rRNA gene profile between NEC and matched controls. (a) Box plots showing the alpha diversity based on observed OTUs (richness) and Shannon diversity. P values were calculated by applying the Mann-Whitney test and adjusted using the FDR algorithm. (b) NMDS plot of beta-diversity measure applying weighted Bray-Curtis dissimilarity. Box plots showing the relative abundance of phyla (c) and of the top 10 most abundant genera (d). P values were calculated by applying the Mann-Whitney test and adjusted using the FDR algorithm. NEC, necrotising enterocolitis; OTU, operational taxonomic unit.

4.3 Integration of human milk oligosaccharides profile, infant gut microbiome and mom's own milk microbiome

The next step was integrating the cross-sectional breast milk microbiome, HMO profiling data, and metagenomic dataset. Firstly, the correlation between clinical factors and the different datasets was investigated using 'adonis'. For MOM microbiome, all infants who had also HMO profile data were included, while from the other two datasets only the subset of the cohort with both HMO profile and stool metagenomic sequencing was included. No covariate was found to be significantly associated with the MOM microbiome profile, but probiotic use in the NICU (adj. P=0.799, R2=0.036), infant sex (adj. P=0.0799, R2=0.020), diagnosis of NEC (adj. P=0.799, R2=0.018), and DOL MOM was received (adj. P=0.799, R2=0.018) explained the most variance (Figure 4.4). Combining all 11 co-variates analysed explained only 15% of the overall variance of MOM microbiome data. Within HMO profiling data, secretor status alone explained 56% of the variance (adj. p<0.001), but no other factor was significantly associated with this data (Figure 4.4). In contrast, the infant gut bacterial profiles were significantly associated with both PMA (R2=0.07, adj. p=0.006) and DOL (R2=0.07, adj. p=0.006), as well as receipt of antibiotics at the time of sampling (R2=0.06, adj. p=0.006) and receipt of probiotics (R2=0.12, adj. p=0.006), but not maternal secretor status (R2=0.02, adj. p=0.58; Figure 4.4). Together, these findings highlight that MOM HMO and infant gut bacterial profiles are influenced by non-overlapping factors related to early life in preterm infants. MOM was not associated with any of the infant-based variables.

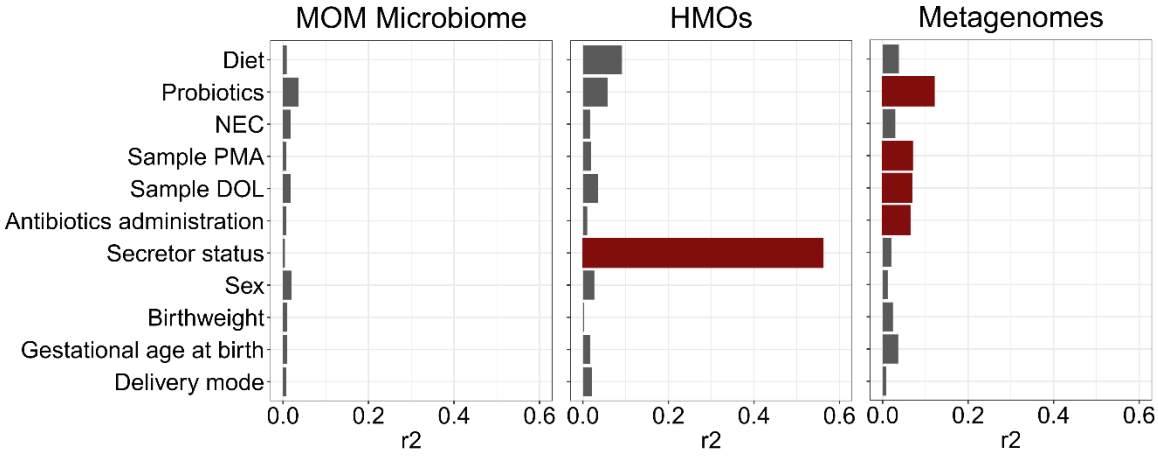

Figure 4.4. Modelling of cross-sectional MOM microbiome, HMO profile and infant stool metagenomic profile using 'adonis'. Horizontal bar plots showing the variance (r2) in MOM microbiome, maternal HMO and infant stool metagenomic profiles explained by clinical covariates as modelled by adonis. For MOM microbiome, all infants who had also HMO profile data were included. For HMO profile and infant gut metagenomes, only infants with a sample in both datasets were included. From the longitudinal stool metagenome dataset, the stool

sample per infant closest to NEC onset (median of 3 days before NEC) and a corresponding control sample matched by DOL were chosen and included in this analysis, and represent the same samples used in chapter 3, section 3.4. Sample DOL is specific to each sample. Variables with an FDR P < 0.05 are shown in red. MOM, mom's own milk; HMOs, human milk oligosaccharides; DOL, day of life; PMA, post-menstrual age.

It has been reported that the HMO profile, MOM microbiome and infant gut microbiome might be interconnected. To understand if any correlation between the different profiles was present, Spearman correlation analysis was performed between MOM samples and the nearest available infant gut microbiome samples based on DOL MOM was fed to the infant and stool DOL.

The correlation between MOM microbiome and infant gut microbiome was investigated to understand if the MOM could directly seed the infant gut, and the top 10 most abundant genera in the breast milk microbiome were included from both datasets. Genus level was used since 16S rRNA gene sequencing cannot give information at the species/strain level. In order of most abundant, the genera included were *Staphylococcus*, *Acinetobacter*, *Enterobacter*, *Pseudomonas*, *Enterococcus*, *Corynebacterium*, *Cutibacterium*, *Finegoldia*, *Streptococcus* and *Bifidobacterium*. Significant positive correlations were found between *Acinetobacter* in the infant gut microbiome and *Staphylococcus* and *Acinetobacter* in the MOM microbiome (R = -0.48, adj. p<0.01 and R = 0.49, adj. p<0.01, respectively; Figure 4.5a,b). A significant negative correlation between *Acinetobacter* and *Staphylococcus* within the MOM microbiome was found (R = -0.52, adj. p<0.001, Figure 4.5c), suggesting that the negative correlation observed between the MOM *Staphylococcus* and the infant gut *Acinetobacter* might be a reflection of the MOM microbiome composition where *Staphylococcus* and *Acinetobacter* are negatively correlated.

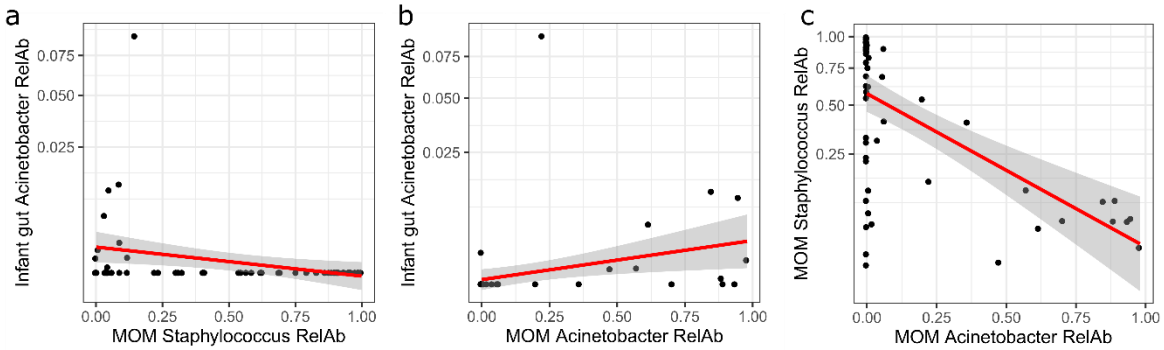

Figure 4.5. Correlation plots between infant gut relative abundance (RelAb) of Acinetobacter, MOM Acinetobacter RelAb and MOM Staphylococcus RelAb. Plot of MOM *Staphylococcus* RelAb and infant *Acinetobacter* RelAb (a), MOM *Acinetobacter* RelAb and infant *Acinetobacter* RelAb (b) and MOM *Acinetobacter* RelAb and MOM *Staphylococcus* RelAb (c). R and P values were calculated by Spearman correlation analysis and P values were adjusted using the FDR algorithm. MOM, mother's own milk; RelAb, relative abundance.

Spearman correlation analysis was also performed between HMO profiles vs. MOM microbiome and HMO profiles vs. infant gut microbiomes. For both analyses, whole dataset and the top 10 genera specific to each microbiome dataset were tested but no significant correlation was found between any HMO and any genus in both microbiome profiles.

4.4 Discussion

MOM feeding is protective against NEC onset and multiple factors might contribute to this protection. MOM might shape the infant gut microbiome not only by providing HMOs, but also by delivering potential gut colonisers from the mom to the infant. To the best of my knowledge, this is the first study investigating MOM microbiome in NEC, and integrating MOM microbiome, HMO profiles, and infant gut microbiome in a cohort of preterm infants. No statistically significant differences were found in MOM microbiome between infants who went on to develop NEC and matched controls at any level of analysis. Contrary to what has been reported for term infants, showing HMO profiles are linked to the developing gut microbiome (Borewicz et al., 2020, Pace et al., 2021), no correlation between HMO profile and preterm infant gut microbiome was found. On the contrary, there was a significant correlation in *Acinetobacter* relative abundances in MOM microbiome and infant gut microbiome.

16S rRNA gene sequencing was applied to study the MOM microbiome owing to the samples being low biomass. Thus, they did not harbour sufficient quantity of microbial DNA for metagenomic sequencing. As observed in previous term and preterm studies, *Staphylococcus* was the most abundant genera found in MOM samples. This is not surprising considering that *Staphylococcus* is a skin coloniser and translocation of the bacterium from the breast skin to the mammary gland or directly in the MOM upon expression could be expected. Compared to what reported in the biggest preterm MOM microbiome study to date, the MOM microbiome in our cohort showed a much lower richness and Shannon diversity (Asbury et al., 2020). As both studies used an OTU approach with 97% similarity threshold, this cannot be explained by the bioinformatic approach used. In addition, the timeframe of sampling was comparable (50 vs 56 days). Despite this difference in alpha diversity, similar to what found by Asbury and colleagues (2020), the top 10 most abundant genera included *Acinetobacter, Pseudomonas, Corynebacterium, Finegoldia* and *Streptococcus (Asbury et al., 2020)*. In contrast to the published data, the other most abundant genera found in this cohort were *Enterococcus*, *Enterobacter*, *Cutibacterium* and *Bifidobacterium*. The maternal gut microbiome, which is thought to be the main seeding source to the MOM bacterial community, is shaped by the mother's diet which might vary upon geographical location, potentially explaining the

differences observed between the two cohorts. Probiotics were used in the NICU this cohort was recruited from, contrary to the published study (Asbury et al., 2020), and further strain-typing work will allow to understand if the bifidobacteria found in MOM samples were a result of cross-contamination due to the probiotics strains colonising the NICU environment. Indeed, multiple studies involving cohorts from probiotic-utilising NICUs have reported that probiotic strains can be found in preterm infants also in absence of direct probiotic administration (Beck et al., 2022), and thus a similar cross-contamination could also be postulated for MOM samples handled in the NICU.

As reported in chapter 3, a difference in DSLNT concentration was found in MOM of NEC compared to control infants. Taking into account the prebiotic effect exerted by HMOs, the MOM microbiome from an extended cohort was also investigated to understand if the HMO profile alteration could be linked to changes in the milk microbial community. Despite variations in HMO profile in MOM received by preterm infants who developed NEC, no differences in MOM microbiome were observed depending on disease status. Thus, the lower DSLNT concentration in NEC might not be a consequence of altered MOM microbiome, or vice versa, contrary to what observed for the infant gut microbiome. Spearman correlation analysis also showed no link between each singular HMO and MOM genera. Published studies have found potential correlations between specific HMOs and milk microbiome, however, the results vary, and no consistent correlation has been identified (Moossavi et al., 2019a, Hunt et al., 2012, Ramani et al., 2018). Notably, while this study was focused on preterm infants, the papers cited included term cohorts.

In contrast to MOM microbiome, the impact of HMOs on the term infant gut microbial community have been widely studied, with the most consistent correlation reporting that HMOs consumption shapes a gut microbiome rich in bifidobacteria when compared to infants receiving formula only (Berger et al., 2020, Borewicz et al., 2020, Vatanen et al., 2018, Lawson et al., 2019). All preterm infants included in this study received some breast milk, preventing analysis exploring the potential microbiome shaping effect exerted by receipt of HMOs compared to never receiving HMOs. It needs also to be noted that preterm cohorts are different from term infants in two pivotal aspects. First of all, preterm infants are routinely treated with antibiotics at birth and are also more likely to receive antibiotics during the first weeks of life due to their immaturity and high risk of infections. Bifidobacteria are known to be susceptible to most antibiotics, potentially explaining the delayed bifidobacteria abundance in the preterm gut which might be affected by this routine practice (Jia et al., 2020). Secondly, 97% of the preterm infants included in this study were supplemented with probiotics containing *Bifidobacterium* spp., which could have covered any potential correlation between HMO

profiles and bacterial species belonging to this genus. Moreover, not much is known about the relationship between specific HMO structures and specific bacterial species in the infant gut microbiome, likely due to the strain variability reported for the capability to metabolise HMOs. It might then not be surprising that no correlation between specific HMOs and microbiome at genus level could be identified.

Previous studies found that MOM-harboured bacterial strains could be found in the term infant gut, suggesting a direct seeding of the mom through this biofluid (Solis et al., 2010). In this study, a positive correlation was found between *Acinetobacter* genus in the MOM and in the infant gut microbiome. A negative correlation was found between infant gut *Acinetobacter* and MOM *Staphylococcus*, likely mirroring the negative correlation observed between these two genera in the MOM microbiome. While metagenomic sequencing can give information at the strain level and was used for stool samples, 16S rRNA gene sequencing was applied to study the MOM samples. Due to the lack of strain information for MOM bacteria, it could not be verified if there had been a direct passage of *Acinetobacter* from the MOM to the preterm gut. While *Acinetobacter* could not be isolated from MOM samples, isolates for the other 9/10 top genera could be obtained, and taken together with the correlation analysis, the results suggest that MOM could potentially directly seed the preterm gut. *Acinetobacter* has been repeatedly found to be one of the abundant genera in breast milks from both preterm (Biagi et al., 2018, Asbury et al., 2020) and term cohorts (Boix-Amoros et al., 2016). *Acinetobacter* spp. are Gram-negative bacteria which harbour multi-drug resistance genes and preterm infants are at increased risk of developing sepsis caused by species belonging to this genus (Zarrilli et al., 2018, Shete et al., 2009). Owing to this antibiotic resistance, *Acinetobacter* might be advantaged in being transmitted from MOM to preterm infants who are often subjected to multiple antibiotic treatments, potentially explaining the correlation between its relative abundance in MOM and stool samples.

Adonis analysis showed no significant correlation between clinical variables tested and MOM microbiome. Only secretor status was associated with the HMO profiles and the infant gut microbiome was associated with antibiotics and probiotics administration, as well as infant age (i.e., DOL and PMA). Notably, while granular metadata was available on infant variables, no information on maternal specific variables were available, which may influence the MOM microbiome composition (e.g., BMI, age, diet) (Gila-Diaz et al., 2019). The infant metadata included information such as day of start and stop for various treatments such as antibiotics, probiotics, receipt of expressed breast milk, formula feeding, and breast milk fortifier usage. The infant factors associated with the infant stool metagenomes included probiotics and antibiotics administration in line with what reported by Beck *et al.* (2022), who analysed

longitudinal stool metagenomes profiles from very preterm infants not affected by intestinal diseases and who were cared for in the same NICU as the infants included in this study (Beck et al., 2022). Delivery mode was not found to be associated with infant gut microbiome in this cohort, as it has been reported in other preterm cohorts (Beck et al., 2022, Stewart et al., 2017a) and contrary to what observed in term infants (Stewart et al., 2018, Wampach et al., 2018, Shao et al., 2019, Reyman et al., 2019). Better understanding what are the factors shaping the HMO and MOM microbiome profiles could however be useful to potentially shape this biofluid composition in order to benefit the preterm health.

4.1 Study limitations and conclusions

This novel study integrated MOM HMOs, MOM microbiome, and infant gut microbiome in a preterm cohort, and investigated MOM microbiome in relation to NEC disease. There are some limitations that should be considered when interpretating the results. First, one single MOM sample per mother-infant pair was analysed preventing the investigation of the MOM microbiome development over time. Only relative abundance data was available for the microbiome datasets and using absolute abundance data could give new insights on the role of MOM microbiome in NEC and on the potential correlation between HMOs, MOM microbiome and infant gut microbiome. 97% of preterm infants included in this study were supplemented with probiotics which include *Bifidobacterium* spp., and this might have masked any potential correlation present between diet-microbiome interaction in the cohort. To better study the correlation between HMOs and intestinal microbiome in a preterm population, comparison of exclusively formula-fed and breastmilk-fed infants would be needed. Moreover, having probiotics-free cohort would allow analysis of the 'natural' microbiome development, and obtaining information on absolute bacteria abundances would aid an additional level of information. No strain characterisation for the MOM microbiome dataset was available, and it could not be verified if the *Acinetobacter* correlation between MOM and infant gut was the result of a direct transmission of the bacteria. Finally, no data on mothers' characteristics was available, apart from secretor status, and the correlation between MOM composition and clinical variables could not be properly investigated.

While HMO profiling might have future applications in NEC prediction and risk stratification, the same cannot be said for the MOM microbiome. This lack of association between NEC and MOM microbiome might have positive implications for the use of donor human milk which is pasteurised before usage resulting in the killing of the live microbial component. However, there might be a direct seeding of the infant gut microbiome through the MOM microbiome.

Further work investigating direct microbial transmission through MOM and understanding what shapes the MOM microbial profile, as well as the HMO composition, will be useful to better understand preterm infant microbiome development and its link with short- and long-term health.

Chapter 5. Bacterial isolation from preterm mom's own milk and stool samples and their growth on human milk oligosaccharides

5.1 Introduction

Bacterial strain variability has many consequences, including influencing the outcome of the interaction between the bacterium and the host and its potential to utilise certain nutrients (like HMOs) (Garrido et al., 2016, LoCascio et al., 2010, Marcobal et al., 2010). Diverse population, such as preterm or term infants, might harbour bacterial strains with a different pathogenic and metabolic capability. To better characterise the preterm population, bacteria were isolated from preterm stool and MOM samples. Aims for this step included isolating bacteria composing the MOM microbiome, to investigate their viability in this sample type, and isolation of bifidobacteria strains from both stool and MOM, because of their potential role in NEC protection and HMO utilisation. Stool samples cultured were selected from babies who were cared for prior to probiotics usage in our NICU at the Newcastle Upon Tyne RVI, and from hospitals which did not utilise probiotics (from the MAGPIE study) (Embleton et al., 2021), to enhance the chance to isolate preterm-associated bifidobacteria strains and not probiotic ones. MOM samples were only selected from mother-infant pairs treated in our NICU where probiotics are used. Fungi isolation from MOM was also attempted, as these microorganisms have been reported to colonise breast milk in term cohorts (Boix-Amorós et al., 2017). After a first step of optimisation, selected isolates were tested for their potential to utilise specific HMOs, including DSLNT. HMOs were kindly donated by Glycom/DSM.

5.2 Isolation of bacteria from preterm mom's own milk and preterm stool

A total of 13 preterm breast milk samples and 30 preterm stool samples were cultured as described in chapter 2, section 2.6.1. 191 and 281 isolates were obtained from MOM and stool, respectively. Of these, 171 and 128 isolates from MOM and preterm stool, respectively, were further analysed and the species identified. Since we did not have strain level information and limited capacity to test isolates of interest, in cases where a species was isolated multiple times from the same sample, we considered all the isolates of that species to be the identical strain. Making the assumption that only one strain per species could be isolated from each sample, 64 and 52 'unique' isolates were obtained from MOM and stool samples, respectively. A full list of the bacterial species obtained from MOM and infant stool samples are listed in Tables 5.1 and 5.2.

In MOM samples, *Staphylococcus epidermidis* and *Cutibacterium acnes* were both isolated from 8 samples (62%), followed by *Enterococcus faecalis* present in 5 samples (39%), and *B. breve*, *E. cloacae* and *Staphylococcus lugdunensis* were obtained from 3 samples (23%) (Table 5.1). Other *Bifidobacterium* spp. isolated included *B. animalis* (2), *B. longum* (1) and *B. bifidum* (1). Of note, probiotics containing *B. longum, B. bifidum,* and *B. animalis* (contaminant as described here (Beck et al., 2022)) were in use in the NICU during the time some of the samples were obtained and so it is likely some of the *Bifidobacterium* originated from those commercial probiotic products. Except for *Acinetobacter,* at least one isolate was obtained for all the top 10 genera composing the MOM microbiome, suggesting that these genera are present alive in the breast milk. Notably, culture conditions were not specific for *Acinetobacter* and so not culturing this organism does not mean it was not present and viable. Despite the utilisation of media targeted to bacteria isolation, *Candida parapsilosis*, a fungal species, was also isolated. Targeted fungal isolation from 6 MOM samples was also attempted by utilising media designed for these microorganisms, however, no fungi grew, and no additional samples were cultured on fungal media and in aerobic conditions.

In preterm infant stool samples, *E. faecalis* was isolated most frequently (10 samples, 33%) followed by *S. epidermidis* (6 samples, 20%), *Cutibacterium avidum* (5, 16%), *B. breve* (5, 17%) and *C. perfringens* (4, 13%) (Table 5.2). *B. longum* and *B. animalis* were also isolated from stool.

Table 5.1. MOM culture collection. Number of unique identified species isolated from each sample and media they could be isolated from.

Species	**Number**	**Agar media**
Anaerococcus senegalensis	1	FAA
Atlantibacter hermannii	1	YPD + P/S
Bifidobacterium animalis	2	BSM, TOS, MRS+, MRS-
Bifidobacterium bifidum	1	MRS-
Bifidobacterium breve	3	BSM, MRS-, MRS+, TOS
Bifidobacterium longum	1	MRS+
Candida parapsilosis	1	BHI, BHI + 5% blood, CPSE
Corynebacterium kroppenstedtii	1	BHI
Corynebacterium pyruviciproducens	1	BHI + 5% blood
Cutibacterium/Propionibacterium sp.	2	MRS-, FAA
Cutibacterium acnes	8	FAA, BSM, MRS+, MRS-
Cutibacterium avidum	1	BSM
Enterobacter bugandensis	1	CPSE

Species	Number	Agar media
Enterobacter cloacae complex	3	BSM, BHI + 5% blood, MacConkey, BHI, CPSE
Enterococcus faecalis	5	CPSE, BHI, BSM, TOS, MRS-, FAA, MRS+
Escherichia coli	1	CPSE
Finegoldia magna	1	FAA
Klebsiella grimontii	1	YPD + P/S
Lactobacillus gasseri	1	MRS-
Pantoea septica	2	MacConkey
Pseudomonas lactis	1	YPD + P/S
Pseudomonas plecoglossicida	1	YPD + P/S
Schaalia radingae	1	FAA
Sphingomonas paucimobilis	1	YPD + P/S
Staphylococcus aureus	2	CPSE, BSM
Staphylococcus capitis	1	BHI
Staphylococcus condimenti	1	CPSE
Staphylococcus epidermidis	8	CPSE, BHI, BHI + 5% blood, TOS, MRS-, FAA, BSM, MacConkey
Staphylococcus haemolyticus	1	CPSE
Staphylococcus hominis	2	CPSE, BHI, MacConkey, BHI + 5% blood
Staphylococcus lugdunensis	3	BHI, BHI + 5% blood, TOS, MacConkey, FAA
Staphylococcus warneri	1	BSM, BHI
Stenotrophomonas maltophilia	1	YPD + P/S
Stenotrophomonas rhizophila	1	YPD + P/S
Streptococcus anginosis	1	BHI

BHI, brain heart infusion; MRS, De Man, Rogosa and Sharpe; MRS+, MRS supplemented with L-cysteine and mupirocin; MRS-, MRS supplemented with L-cysteine; TOS, Transgalctosylated Oligiosaccharide; BSM, bifidus selective medium; FAA, fastidious anaerobe agar; YPD, yeast extract peptone dextrose; P/S, penicillin and streptomycin supplement.

Table 5.2. Stool culture collection. Number of unique identified species isolated from each sample and media they could be isolated from.

Species	Number of isolates	Agar media
Bifidobacterium animalis	2	MRS+, BSM, TOS, FAA
Bifidobacterium breve	5	MRS-, TOS, MRS+, BSM
Bifidobacterium longum	3	TOS, MRS+, MRS-
Clostridium baratii	3	TOS, MRS+, MRS-
Clostridium perfringens	4	TOS, MRS+, MRS-
Clostridium saudiense	1	FAA
Cutibacterium avidum	5	MRS-, MRS+, BSM

Species	Number of isolates	Agar media
Enterococcus faecalis	10	BSM, TOS, MRS-, MRS+
Enterococcus faecium	3	BSM, TOS, MRS-, MRS+
Escherichia coli	3	TOS, MRS-
Klebsiella oxytoca	2	MRS+, MRS-
Klebsiella variicola	1	TOS
Staphylococcus epidermidis	6	MRS-, TOS, BSM
Staphylococcus haemolyticus	3	MRS-, TOS, BSM
Staphylococcus warneri	1	BSM

BHI, brain heart infusion; MRS, De Man, Rogosa and Sharpe; MRS+, MRS supplemented with L-cysteine and mupirocin; MRS-, MRS supplemented with L-cysteine; TOS, transgalctosylated oligiosaccharide; BSM, bifidus Selective Medium; FAA, fastidious anaerobe agar; YPD, yeast extract peptone dextrose; P/S, penicillin and streptomycin supplement.

5.3 Optimisation of bacterial medium for growth curves

As described in the previous section, multiple different bacterial species could be cultured from infant stool and MOM. Before being able to investigate the ability of the different species to utilise HMOs using growth curves, a first step of optimisation of the medium used for the experiments was performed. The aim was to identify a single medium which could sustain the growth of the different isolates in the presence of a carbon source leading at the same time to minimal growth in its absence.

E. coli AM114, *K. oxytoca* VK66, *E. faecalis* AM3, *B. breve* AM76 and *B. longum* AM7 were used for the optimisation stage. Firstly, their growth was tested in Minimal medium prepared by autoclaving all components together, Rich medium and mMRS with and without 2% glucose. No media was able to sustain the growth of all the species tested while preventing the growth in absence of the carbon source (Figure 5.1a-e). As Minimal medium satisfied the conditions for *K. oxytoca* VK66, *E. coli* AM114 and *E. faecalis* AM3 (Figure 5.1a-c), but not for the bifidobacteria tested (Figure 5.1d,e), Minimal medium was prepared with a different protocol. This medium contains L-cysteine HCl which is known to be important for bifidobacteria growth, and which structure can be affected by the autoclaving process. Minimal medium was thus prepared following the previous protocol but adding filter-sterilised L-cysteine HCl after the autoclave cycle as described in chapter 2, Table 2.2. This new version of the Minimal medium sustained the growth of all bacteria tested in the first step, including bifidobacteria (Figure 5.1f,g, Figure 5.2), while no or very low growth was observed in glucose absence.

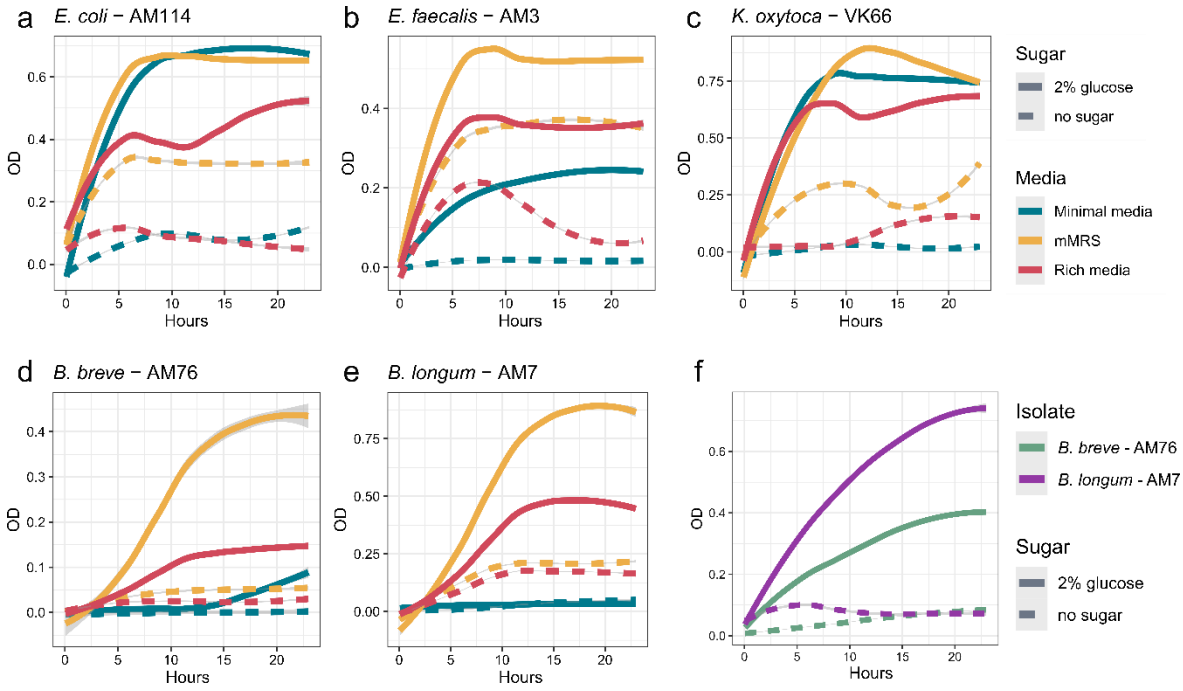

Figure 5.1. 24 hours growth curves of bacteria isolated from preterm stool in different media. The media used included Minimal medium, modified MRS (mMRS) and Rich medium. Their composition and preparation are described in chapter 2, section 2.3.2. The volume used for the growth was of 200 µl. Growth curve of *Escherichia coli* AM114 (a), *Enterococcus faecalis* AM3 (b), *Klebsiella oxytoca* VK66 (c), *Bifidobacterium breve* AM76 (d) and *Bifidobacterium longum* AM7 (e). Plots in panels a-e are coloured by medium used and presence or absence of glucose is represented by line type. Minimal medium was prepared according to the first ptorocol. (f) 24 hours growth curve of *Bifidobacterium breve* AM76 and *Bifidobacterium longum* AM7 in Minimal medium prepared according to protocol two with filter-sterilised L-cysteine HCl added after the autoclave sterilisation process. Plots in panel f is coloured by isolate tested and presence or absence of glucose is represented by line type. For each isolate tested, isolate number and species are reported. OD, optical density.

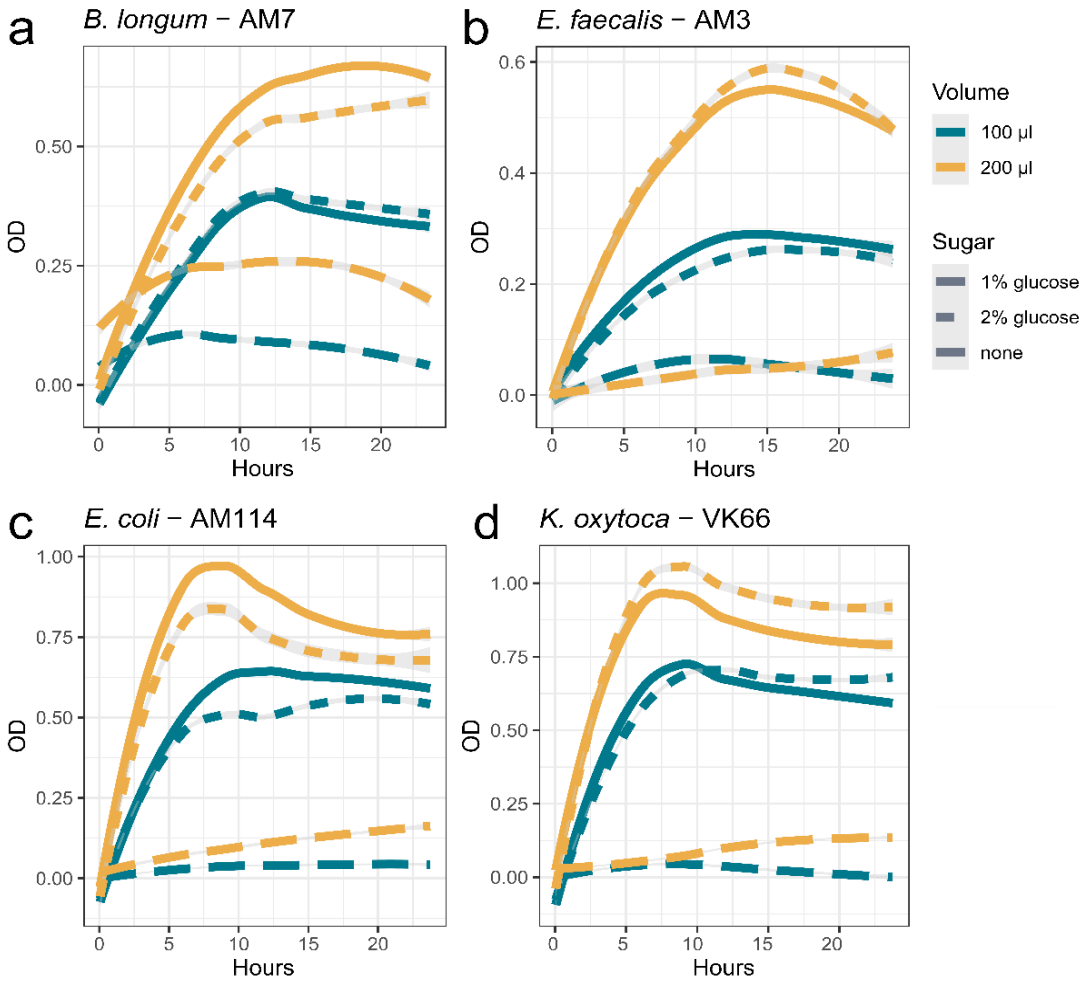

Figure 5.2. 24 hours growth curves of bacteria isolated from preterm stool in Minimal medium tested with different glucose concentrations and growth volumes. Minimal medium prepared by adding filter-sterilised L-cysteine HCl after autoclave was used. Growth curve of *Bifidobacterium longum* AM7 (a), *Enterococcus faecalis* AM3 (b), *Escherichia coli* AM114 (c), and *Klebsiella oxytoca* VK66 (d). Glucose was tested in 1% and 2% w/v concentration, and growth volumes of 100 µl and 200 µl were used. Plots are coloured by growth volume, and the line type indicates the glucose concentration. For each isolate tested, isolate number and species are reported. OD, optical density.

Synthesised HMOs are expensive and the quantity available in our laboratory was limited. To be able to test as many bacterial isolates as possible, two different concentrations of glucose and of growth volume were tested. Growth of the species analysed was comparable in 1% and 2% glucose concentrations (Figure 5.2). On the contrary, the optical density (OD) measured was lower when growth was measured in 100 µl volume compared to 200 µl, resulting in a better separation between presence and absence of glucose in the higher volume tested. We therefore opted to use 1% sugar concentration in 200 µl of growth volume in the subsequent experiments.

Before moving onto performing the growth curves with the whole sets of isolates of interest, an additional medium was tested. ZMB1 is a complex defined medium which contains 57

different components, and which was developed to allow the growth of many different bacterial species. At this stage, the isolates tested included the ones reported in Figures 5.1 and 5.2, and new additional isolates. All isolates analysed could grow in ZMB1, even though growth in absence of carbon source was observed for some isolates, including *E. faecalis* AM3, *S. aureus* MR3, *B. longum* AM7 and *B. longum* VK53 (Figure 5.3a,b). Since there might be carry over of nutrients not metabolised in the overnight growth media, isolates were tested after resuspending the growth in PBS before being spiked in the growth plate, as described in chapter 2, section 2.6.2. Resuspending the overnight culture in PBS before performing the growth curve assay resulted in a lower OD in absence of carbon source for the isolates tested (Figure 5.3c-f). Removal of spent medium and resuspension in PBS was then included the growth curve preparation protocol.

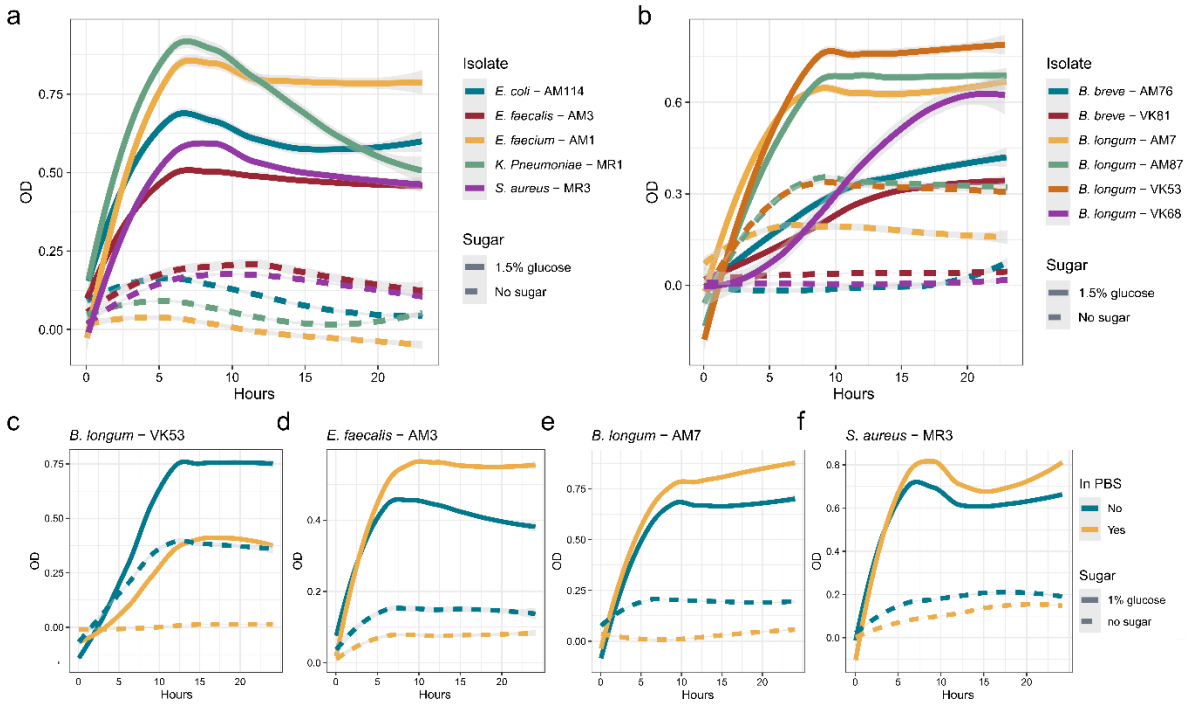

Figure 5.3. Optimisation of growth curves using ZMB1 medium. Non-bifidobacteria isolates (a) and bifidobacteria isolates (b) not resuspended in PBS and grown for 24 hours in ZMB1 with or without glucose. Each isolate was plotted with a different colour, and sugar supplementation was indicated by line type. 24 hour growth curve in ZMB1 with and without glucose of *B.* longum VK53 (c), *Enterococcus faecalis* AM3 (d), *B. longum* AM7 (e) and *S. aureus* MR3 (f) resuspended or not in PBS. Isolates were plotted by colouring depending on presence or absence of PBS resuspension step, while sugar supplementation was indicated by line type. For each isolate tested, isolate number and species are reported. OD, optical density.

5.4 Growth curves of selected isolates on human milk oligosaccharides

A total of 17 different bacterial isolates spanning pathobionts and potential probiotics species (e.g., *Bifidobacterium*) were tested for growth on HMOs using ZMB1 with and without a carbon source.

Lactose, the most abundant sugar in human milk, was also tested. All bacteria reached the plateau in their growth. For each isolate, the maximum OD reached in medium without a carbon source (i.e., the blank) was removed from the maximum OD reached in supplemented medium, and this value was used in the plots. A clustering based on HMOs utilisation could be observed both in Principal Component Analysis (PCA) and heatmap plots of the data (Figure 5.4a,b). The probiotic strain *B. infantis* was the main HMO utiliser and clustered distinctly in the PCA plot (Figure 5.4a). *B. infantis* could reach ODs above 0.8 on all carbon sources except for DSLNT, which it could not metabolise (Figure 5.4b). *B. bifidum* also clustered distinct from the other bifidobacteria (Figure 5.4a). It could metabolise most carbon sources with exception of 2'FL; however, it showed much lower maximum OD reached, with growth in LNFP I showing the highest OD of 0.46 (Figure 5.4b). All other bifidobacteria isolates, except for *B. animalis*, could utilise LNT, and 3/5 could also metabolise LNnT, but not the other HMOs (Figure 5.4b). These isolates clustered together in the PCA but separately from *B. bifidum*, underlining the differences in HMOs utilisation capacity, while they clustered together with *B. bifidum* in the heatmap. The last cluster observed was composed of isolates which could only utilise glucose and in some cases lactose, but not HMOs (Figure 5.4a,b). This cluster included the probiotic *B. animalis* and other preterm gut pathobionts not belonging to the *Bifidobacterium* genus.

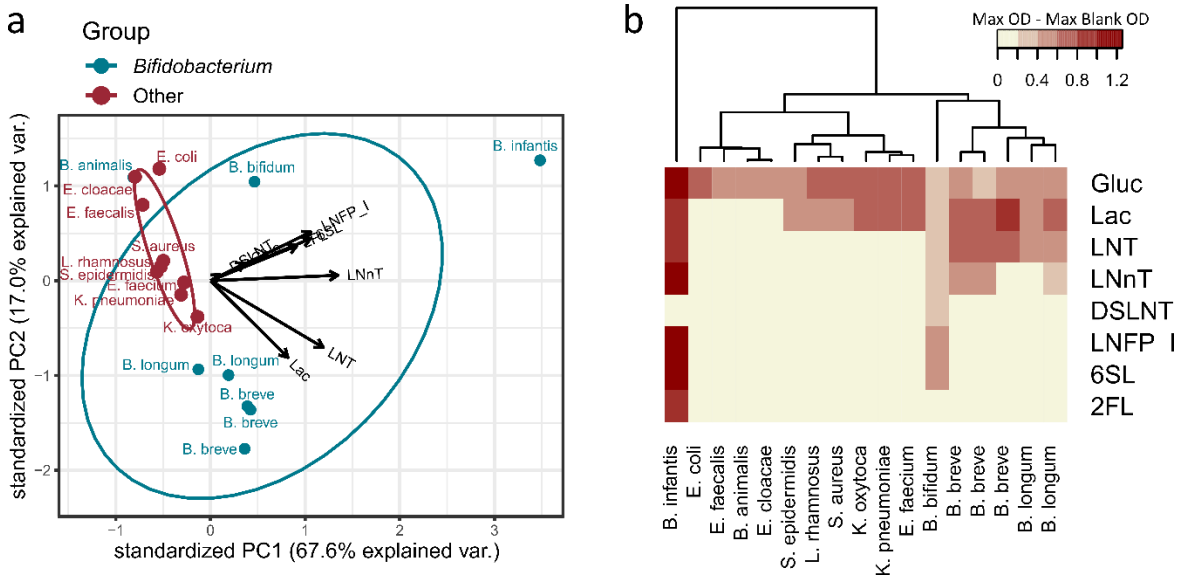

Figure 5.4. Growth curve data of isolates grown on HMOs. PCA plot (a) and heatmap (b) based on the optical density (OD) reached by each isolate on the various carbon sources. The isolate's maximum OD in medium without carbon source was subtracted to the maximum OD reached by the isolate in each condition. Gluc, glucose; Lac, lactose; LNT, lacto-N-tetraose; LNnT, lacto-N-neotetraose; DSLNT, disialyllacto-N-tetraose; LNFP, lacto-N-fucopentaose; 6'SL, 6'-sialyllactose; 2'FL, 2'-fucosyllactose; OD, optical density.

5.5 Discussion

Strain variation in bacteria is known to define the fitness of the microorganism depending on the environment characteristics, and preterm stool and MOM samples were cultured to obtain preterm-related strains. This work's focus was directed to understand the potential of preterm bacteria, in particular *Bifidobacterium* spp., to utilise DSLNT and other relevant HMOs, considering the alteration in MOM HMOs and infant gut microbiome reported in NEC (see chapter 3, (Autran et al., 2018)). *Bifidobacterium* spp. were confirmed to be able to utilise some of the HMOs tested. *B. infantis* could use 5/6 HMOs included, with DSLNT being the only sugar it could not metabolise and reached the highest OD on all sugars compared to the other bacterial isolates. *B. bifidum* could also grow on 5/6 HMOs utilised, 2'FL was the only structure it could not metabolise, and it was the only isolate which could grow on DSLNT, despite showing a lower median OD compared to *B. infantis.*

Cultured stool samples were selected from infants who did not receive any probiotics in Newcastle upon Tyne NICU or who were treated in a probiotic-free NICUs to avoid potential isolation of probiotic strains. MOM samples were instead selected from our NICU, thus cross-contamination with probiotics strains might have happened. Bifidobacteria were not the only target, and the isolation process aimed also at obtaining the top 10 genera present in MOM samples according to the 16S rRNA gene sequencing. Indeed, sequencing techniques do not give information on bacteria viability and isolation was attempted to investigate the presence of viable bacteria in this sample type. *Bifidobacterium* spp. could be isolated from preterm MOM and stool samples, and 9/10 most abundant genera in MOM microbiome data could be isolated from milk samples. Only *Acinetobacter* could not be cultured from the group of top 10 genera in MOM samples. This was potentially the consequence of storage settings or not adopting the optimal media and conditions for the culturing process, which have likely affected the isolation of other genera present in the samples cultured.

Term studies have reported the presence of fungi in breast milk (Boix-Amorós et al., 2017). To expand our work, MOM samples were cultured for the isolation of these microorganisms by utilising media optimised for fungi culturing and which are supplemented with antibiotics to prevent bacteria growth (see chapter 2, section 2.3.1). Six MOM samples were cultured on these media; however, no fungi could be isolated. Due to paucity of MOM samples and the volume of sample needed for culturing, fungal isolation was not attempted from additional samples. Interestingly, only one fungal species, *Candida parapsilosis*, was obtained from a sample which was cultured for bacteria isolation. This suggests that preterm MOM might harbour a fungal microbiome, but potentially less abundant and readily culturable than in term cohorts.

The results obtained from the growth curve experiments largely confirmed what has been found from other groups. *B. infantis* and *B. bifidum* could utilise most HMOs, except for DSLNT and 2'FL, respectively. *B. breve* and *B. longum* strains could all utilise LNT, and most could use LNnT, confirming that a strain variability is present in isolates belonging to the same species (Asakuma et al., 2011, Thongaram et al., 2017). Finally, other gut commensals could not grow on HMOs, and some of them could only grow on glucose, without being able to metabolise lactose, which is the most abundant sugar present in MOM. No whole genome sequencing was available for the isolates tested in the experiment, so the potential of the isolates to use HMOs based on gene content could not be inferred and did not guide isolate selection. Proper comparison of growth between different strains tested was made hard by the different maximum ODs reached by each isolate. For instance, *B. infantis* appears to be more efficient in HMO utilisation as it is reaching ODs superior to 0.8 in all sugar (except DSLNT), while the maximum OD reached by *B. bifidum* was 0.46 OD. While these data suggest *B. infantis* can metabolise HMOs more efficiently, *B. bifidum* reached the highest ODs on most HMOs when normalised for growth on both lactose and glucose. Thus, the differences might be due to varied potential of growth of the isolates in the specific medium utilised, and a normalisation factor would be needed. Potential avenues of normalisation could include counting the number of colony-forming units (CFUs) at the start and end of growth; however, this will require further optimisation.

Recent studies reported how *Bifidobacterium* spp. promote the formation of bifidobacteria consortia by cross-feeding bacteria belonging to the same genus (Lawson et al., 2019, Tannock et al., 2013). In our small cohort, only a single unique *Bifidobacterium* spp. was isolated from individual samples. Nonetheless, it is possible other species existed but were not identified and further work is needed to determine if the same species from within a sample are also the same strain. This prevented the investigation of the potential cross-feeding between bifidobacteria co-existing in the same microbial community. Notably, studies investigating bifidobacteria consortia have typically involved full-term breastfed infants who harbour a gut microbiome dominated by bifidobacteria (Lawson et al., 2019, Stewart et al., 2018). On the contrary, due to high antibiotics administration rates and lower breastmilk intake, preterm infants show a delayed bifidobacteria colonisation (Jia et al., 2020), making less likely the establishment of multi-*Bifidobacterium* spp. ecosystems.

To the best of my knowledge, this was the first time synthetised DSLNT was used in isolation to test bacterial growth. Previous studies only tested its consumption when part of an HMO pool (Marcobal et al., 2010). An enzyme from *B. bifidum* has been found to be able to catalyse the release of the sialic acids from DSLNT with the production of LNT, however, it was tested

in its purified version and the ability of this species to grow on DSLNT was not investigated (Kiyohara et al., 2011). In the experiments performed as part of this thesis, we showed that *B. bifidum* was able to metabolise and grow on DSLNT, and was the only isolate tested able to use this sugar. *B. bifidum* metabolises HMOs through extracellular glycosidases, and can subsequently internalise and utilise lactose, LNB and some of the monosaccharides which are released (Masi and Stewart, 2022). Some studies have also hypothesised that the HMO-degradants produced by *B. bifidum* extracellular enzymes could also be digested by other bifidobacteria favouring the formation of a community rich in this genus (Gotoh et al., 2018). This hypothesis was generated based on *in vitro* (Gotoh et al., 2018) and *in vivo* studies (Tannock et al., 2013), and it could be a possible mechanism for NEC protection exerted by DSLNT. Tannock *et al.* (2013) reported that breast-fed infants harbouring a *B. bifidum* relative abundance of at least 10% in their gut microbiome were associated to the highest abundance of total *Bifidobacteriaceae* (Tannock et al., 2013). Using *in vitro* experiments, Gotoh and colleagues found that *B. longum* could only grow on HMOs when co-cultured with four *B. bifidum* strains which were tested separately (Gotoh et al., 2018). Moreover, faecal microbial communities grown in HMOs and supplemented with the *B. bifidum* strains tested showed an enhanced expansion of bifidobacteria which was reduced when *B. bifidum* strains were not added (Gotoh et al., 2018). Despite what shown by Gotoh *et al.*, the degradants produced by the growth of *B. bifidum* on HMOs might also give access to other gut commensals to nutrients, also favouring the growth of potential pathogens. Further *in vitro* experiments in complex communities are needed to verify if the cross-feeding of *B. bifidum* is exclusive to other bifidobacteria or if it is extended to other gut commensals. Also, studies in non-probiotic preterm cohorts might elucidate the potential role of DSLNT and *B. bifidum* in favouring the establishment of a bifidobacteria rich community and its protective role against NEC development. While no correlation between DSLNT and *B. bifidum* or other bifidobacteria was found in the cohort study, infants receiving high DSLNT transitioned to bifidobacteria-rich gut communities compared to infants receiving low DSLNT.

B. infantis is one of the most frequent *Bifidobacterium* (sub)species used in probiotics products for infant supplementation. Indeed, *B. infantis* is specialised in utilising HMOs and can thrive in the infant gut, with colonisation by this species associated with increased SCFAs production (Frese et al., 2017, Alcon-Giner et al., 2020) and decreased inflammation in the infant (Henrick et al., 2021). However, HMOs might benefit the infant in further ways other than acting as prebiotics, like directly modulating the gastrointestinal and immune systems (Masi and Stewart, 2022). The long-term effects of *B. infantis* supplementation and its full consumption of HMOs have yet to be elucidated. Postulating that DSLNT might protect infant from NEC through a

microbiome-independent mechanism, this lack of DSLNT utilisation by *B. infantis* might be of relevance when optimising probiotics for reducing disease risk in preterm infants.

5.6 Study limitations and conclusions

There are some limitations to the experiments presented in this chapter. Samples used for the isolation of microorganisms were not stored in the optimal conditions to favour microbial culturing. Samples cultured had been previously used for other analyses, such as microbial DNA extraction for microbiome sequencing, so they had been subjected to multiple freeze-thaw cycles, potentially hindering the quantity of viable bacteria. Moreover, samples were exposed to an aerobic environment, which could affect anaerobic bacteria survival, and were not stored in cryoprotectant solutions to preserve bacteria viability. The use of additional microbiological media and varied culturing conditions might have allowed the isolation of additional species. However, the scope of the isolation process was to obtain multiple bacterial species for diet-microbiome-host interaction experiments, and not characterisation of the bacterial viable fraction of preterm stool and MOM microbiome.

Isolates obtained were not characterised at the strain level, preventing the determination if multiple strains were co-existing in the same sample. MOM samples used for culturing were collected from a probiotic-using NICU and the bifidobacteria isolates obtained could be derived from the probiotic products. A further limitation is that, due to paucity in HMOs quantities, only one isolate was tested for each of the non-bifidobacteria species. As strain-variability has been found for HMO utilisation, including additional isolates will be important for future studies.

In summary, bifidobacteria could be isolated from preterm MOM and stool samples, and the top ten genera composing the MOM microbiome could be isolated from breast milk proving their viability in this sample type, with the exception of *Acinetobacter*. Growth curve data confirmed *Bifidobacterium* spp. were the main HMOs utiliser, and none of the other gut commensals tested could utilise these complex sugars. *B. infantis* and *B. bifidum* could metabolise most HMOs, with exception of DSLNT and 2’FL, respectively. *B. bifidum* was the only isolate tested which could utilise DSLNT.

Chapter 6 Exposure of preterm and adult derived ileal organoids to bacteria and human milk oligosaccharides

6.1 Introduction

HMO profiles are altered in MOM fed to infants who develop NEC, with DSLNT being present in a lower concentration compared to control infants, as described in chapter 3. The mechanism behind DSLNT protective effect against NEC is yet to be identified, and it might act through multiple avenues. Indeed, DSLNT might shape the infant gut microbiome as discussed in the previous chapters. However, HMOs have also been shown to interact with and modulate the intestinal epithelium (Lane et al., 2013, He et al., 2016, Natividad et al., 2020, Wu et al., 2019). Such potentially contrasting mechanisms raise the conundrum that if HMOs act directly on the host epithelium, colonisation by HMO-utilising bacteria (typically considered probiotic) could be disadvantageous.

In this chapter, the direct response of the premature intestine to HMOs was investigated using preterm ileum-derived organoids (PIOs). Two datasets were analysed: the first one was readily available when I started my PhD and I led the analysis while unable to access the lab during the covid pandemic; the second was generated during my PhD once lab space was reopened. The first experiment exposed organoids to *E. faecalis* and in the second experiment organoids were exposed to HMOs, IL-1β and LPS. In both experiments, PIOs and adult ileum-derived organoids (AIOs) were used. The main output analysed was RNA-sequencing to investigate the transcriptional response of the organoids to the applied stimuli. Firstly, the transcriptome profiles of PIOs and AIOs were compared to understand if any differences were present and if the "prematurity" of the preterm cohort was retained. Subsequently, NEC and non-NEC PIOs were compared, and the impact of HMOs on PIOs and AIOs were investigated.

6.2 Optimisation of transwell pore size and cell number for formation of organoid monolayers

Monolayers of cells were produced using PIOs cells. The protocol was optimised in two steps. In the first part, monolayers were produced by seeding 0.4 μm pore size transwells with $8x10^5$ or $5x10^5$ cells. The two conditions were compared after 4 days of differentiation. Cell confluence was assessed visually using microscopy directly above the transwell and by TEER measurement (Figure 6.1a), while monolayer morphology and cell polarisation were assessed by H&E staining (Figure 6.1b). Monolayers could properly establish in both conditions;

however, higher TEER measures were observed in monolayers formed by $5x10^5$ cells seeding, suggesting higher barrier integrity. Moreover, higher cell concentration resulted in increased portions characterised by cells growing on each other. $5x10^5$ cells seeding was used for following experiments since resulting monolayers reached confluence and correctly polarised. Of note, $5x10^5$ cells is the amount suggested from Fofanova et al. (2019) for monolayer formation (Fofanova et al., 2019). After determining that $5x10^5$ cells was optimal, we proceeded testing transwells of different pore sizes and assembling the OACC system. All monolayers survived in the system, as it is demonstrated by TEER measures which remained above 1000 Ω (Figure 6.1c). Transwells of 3 µm pore size showed organoid cells in the pores and translocation of cellular content to the bottom side of the transwell (Figure 6.1d). This was not observed in 0.4 µm pore size transwells which also reported higher TEER measures. Transwell inserts of 0.4 µm pore size were used in the subsequent experiments.

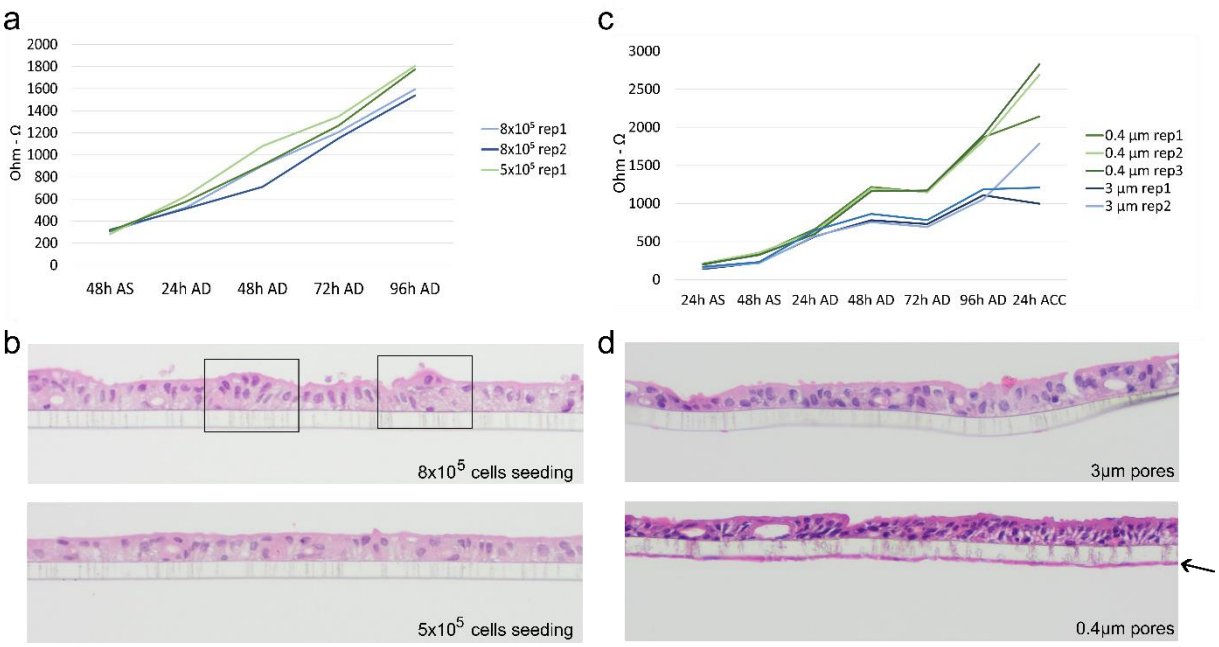

Figure 6.1. TEER measures and microscope pictures of monolayers derived from preterm intestinal organoids. TEER measures (a) and H&E staining (b) of monolayers originated from seeding $8x10^5$ and $5x10^5$ cells; black boxes indicate examples of regions in which cells grew on each other. TEER measures (c) and H&E staining (d) of monolayers originated on 0.4 µm and 3 µm pore size transwells; $5x10^5$ cells were used for seeding. Rep, replicate; AS, after seeding; AD, after Differentiation media was added; ACC, after co-culture system was assembled.

6.3 Impact of exposure time on transcriptome profile of preterm organoids exposure to *Enterococcus faecalis*

At the start of my PhD, RNA-sequencing data from organoids experiments was readily available and I performed the analysis, which is included in this thesis. The experiments comprised four PIO and five AIO lines which were co-cultured with *E. faecalis* (see methods;

chapter 2, section 2.7.8). A control condition without the bacterium was also included. AIOs were co-cultured for 24 hours, while PIOs were co-cultured for 8 hours and 24 hours.

At first, the influence of culture time on the transcriptome profile was investigated within PIOs by comparing monolayers maintained in the OACC model for 8 hours and 24 hours. Transcriptome profiles were significantly different depending on time of culture (2000 permutations, pR2Y=0.006, pQ2 < 0.001; Figure 6.2a). A total of 598 differentially expressed genes (DEGs) were identified when comparing the two time points without separating for exposure condition (all adj. p<0.05), of which 420 (70%) were downregulated (Table 6.1, Figure 6.2b). The DEGs included genes associated to cell cycle, such as M-phase inducer phosphatase 3 (*CDC25C*) and G1 to S phase transition 2 (*GSPT2*), and immune response, including IL-6 receptor (*IL6R*) and IL-18 binding protein (*IL18BP*). All the genes just mentioned were upregulated at 24 hours, except for *GSPT2* which was upregulated at 8 hours (Figure 6.2b). Additional DEGs which were differentially expressed included 14 solute carrier family (SLC) genes, involved in metabolites and molecules transport across the membranes; 19 genes for zinc finger (ZF) proteins which might be involved in transcriptional regulation; 5 metallothionein (MT) genes which encode for proteins binding heavy metals; hypoxia inducible lipid droplet associated (*HILPDA*) gene which was downregulated in the 8 hours culture compared to 24 hours. Next, we compared the transcriptome profile between 8 hours and 24 hours co-culture after dividing by exposure condition (*E. faecalis* or media only). Comparing between 8 hours and 24 hours within media only cultures showed 112 DEGs (77 downregulated; 69%), while comparing between the two time points following bacterial co-culture showed 595 DEGs (408 downregulated, 69%) (Table 6.1, Figure 6.3a). 3 SLC and 1 ZF genes were differentially expressed in the media only comparison, while 21 SLC and 17 ZF genes were significant in the bacteria exposure comparison. ORA was performed on the DEGs, showing no gene set was significantly over-represented.

Table 6.1. Number of differentially expressed genes and proportion of down and up regulated genes in the comparisons listed.

	Total DEGs	Downregulated	Upregulated
8h vs 24h Both conditions together	598	420 (70%)	178 (30%)
8h vs 24h Media only	112	77 (69%)	35 (31%)
8h vs 24h *E. faecalis*	595	408 (69%)	187 (31%)
E. faecalis vs Media 8h	43	24 (56%)	19 (44%)
E. faecalis vs Media 24h	843	282 (35%)	561 (65%)

DEG, differentially expressed gene.

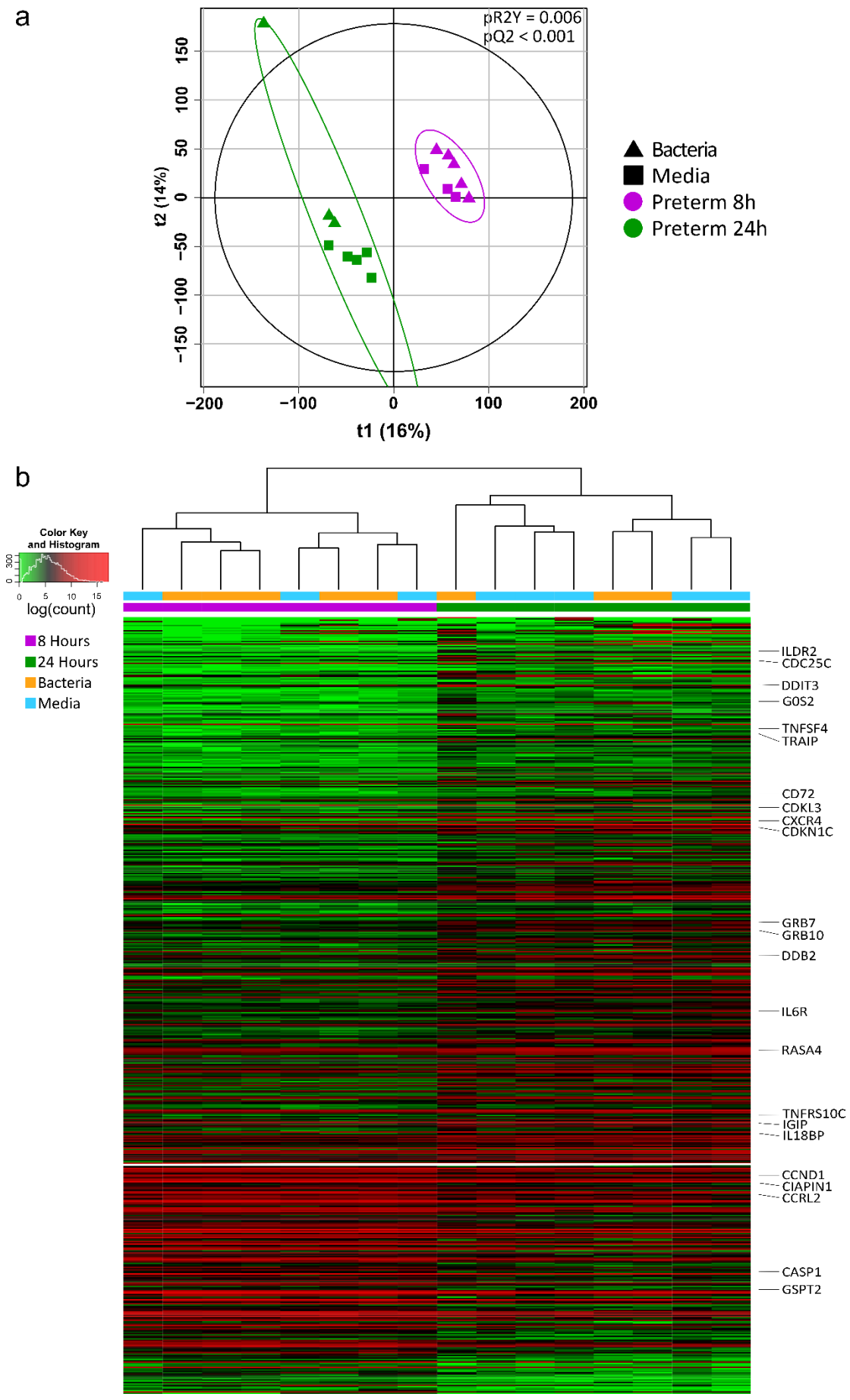

Figure 6.2. Comparison of preterm lines cultured for 8 hours and 24 hours. Bacterial treatment was with *Enterococcus faecalis*. (a) Partial least squares discriminant analysis of

preterm monolayers cultured for 8 hours and 24 hours. Plot coloured by time point and shaped by exposure (media only or bacteria). P values were calculated based on 2000 permutations. (b) Heatmap of genes differentially expressed between preterm organoids cultured for 8 and 24 hours (all adj. p<0.05).

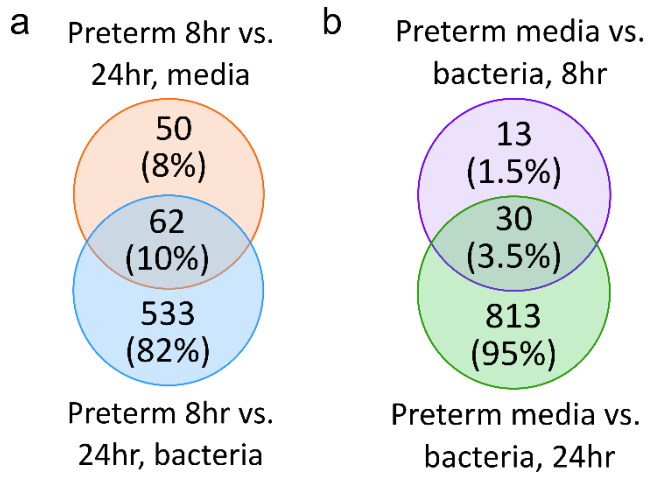

Figure 6.3. Venn-diagrams showing the number of shared genes differentially expressed between various comparisons. Bacterial treatment was with *Enterococcus faecalis*. (a) Venn diagram of significantly differentially expressed genes associated to comparison of preterm organoids cultured for 8 and 24 hours separating by exposure condition (all adj. p<0.05). (b) Venn diagram of significantly differentially expressed genes between preterm organoids exposed to media and bacteria by time point (8 and 24 hours) (all adj. p<0.05).

Next, the impact of exposure (media only or bacteria) within each time point in PIOs was analysed. At 8 hours, 43 DEGs were found between media only and bacteria of which 24 were downregulated (56%), whereas at 24 hours 843 DEGs were found of which 282 were downregulated (35%) (Table 6.1). Of these DEGs, 30 were consistently up or down regulated in both exposure times except for one gene which was upregulated at 24 hours and downregulated at 8 hours, but for which no functional characterisation was available (Figure 6.3b). A higher number of genes related to immune response and cell death were significantly differentially expressed in the comparison of 24 hours conditions. These DEGs included death associated protein kinase 3 (*DAPK3*), DNA damage inducible transcript 3 (*DDIT3*), 15 heat shock protein genes, IL-10 receptor subunit beta (*IL10RB*), IL-1 receptor associated kinase 2 (*IRAK2*), which were all upregulated in bacteria versus media comparison at the 24 hours timepoint.

6.4 Comparison of transcriptome profile of preterm ileal-derived organoids between tissue culture incubator compared to the organoid-anaerobe co-culture system

Data on the transcriptome profiles of monolayers exposed to the control condition (IL-1β and LPS) in a tissue culture (TC) incubator (5% CO_2) were also collected for 4 PIO lines (1 NEC,

3 controls), allowing comparison of PIOs kept in a TC incubator (estimated 17% oxygen) vs the OACC system (5.6% oxygen). As observed for the previous comparison, samples clustered mainly by line and not by exposure condition (Figure 6.4a). A total of 353 genes were differentially expressed (all adj. $p<0.05$), of which 149 (42%) were downregulated in the TC incubator compared to the OACC system (Figure 6.4b). DEGs included 3 ATP binding cassette (ABC) transporters and 11 SLC genes, suggesting alterations in metabolites and molecules transport across the membranes. Interestingly, genes from same family of proteins were differentially expressed, including 18 ZF proteins which might be involved in transcriptional regulation and 10 MT genes which encode for proteins binding heavy metals.

ORA was subsequently performed on the DEGs. A total of 5 GO BPs and one KEGG pathway were significantly over-represented (all adj. $p<0.05$; Figure 6.4c). The KEGG pathway and BPs included 9 of the 11 metallothionein genes. All the metallothionein genes were downregulated in the TC incubator (Figure 6.4b).

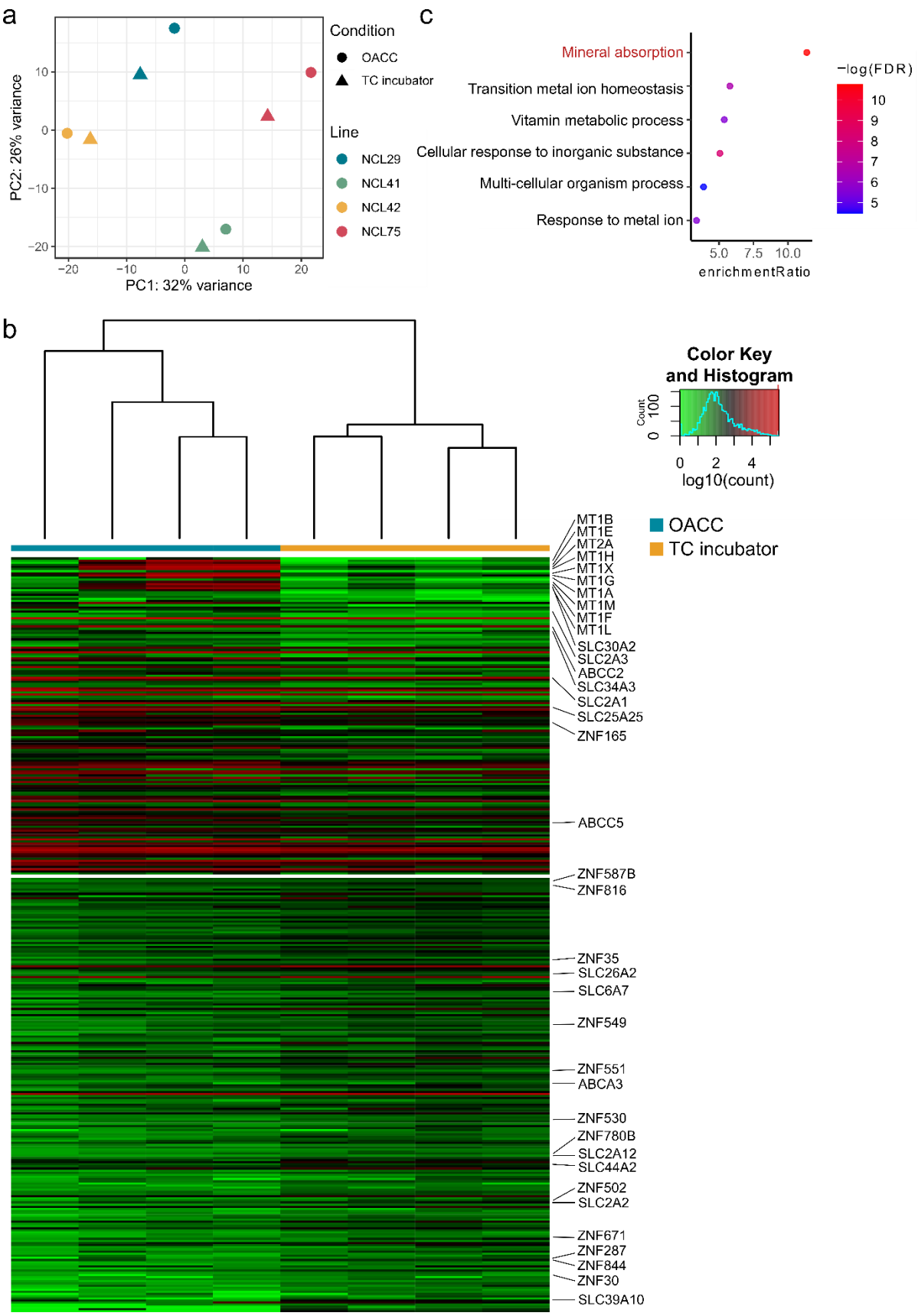

Figure 6.4. Comparison of transcriptome profile from preterm organoids cultured in the OACC system and TC incubator. All organoids were exposed to IL-1β apically and LPS basolaterally, and all plots show comparison of preterm intestinal organoids cultured in the OACC system with 5.6% oxygen supplemented basolaterally and organoids cultured in a TC incubator. (a) Principal component analysis plot of the transcriptome profiles. Plot coloured by

organoid line and shaped by incubation condition. (b) Heatmap normalised counts for the differentially expressed genes. All adj. p<0.05 and absolute log2 fold change>0. (c) KEGG (in red) and GO Biological Processes (black) significantly over-represented in comparison of PIOs cultured in the OACC system or TC incubator. All adj. p<0.05.

IL-8 is a known marker of stress and response to an inflammatory stimulus and IL-8 concentration was analysed in apical and basolateral media using ELISA. Higher production and secretion of IL-8 was found in the apical compartment from PIOs in the TC incubator compared to the OACC system (p=0.033, Figure 6.5a). A similar trend was observed in the basolateral compartment, but this did not reach statistical significance (p=0.486, Figure 6.5b).

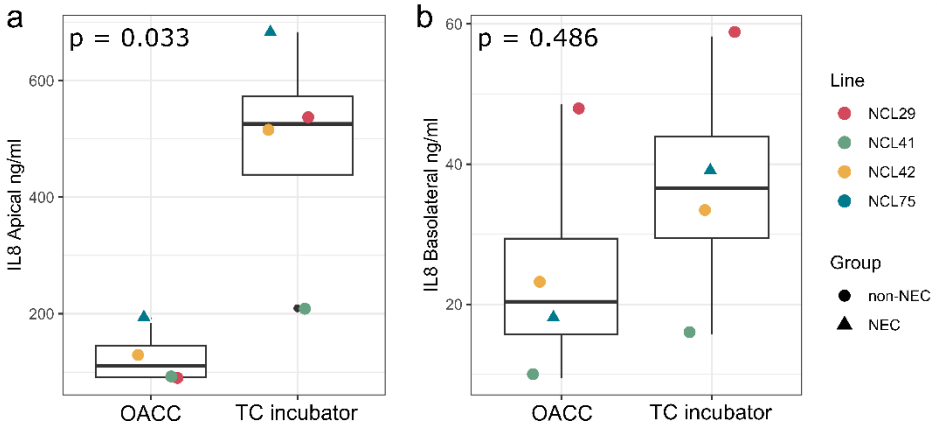

Figure 6.5. Preterm organoids cultured in the TC incubator produced and released higher concentrations of IL-8 in the apical compartment compared to culture in the OACC system. All organoids were exposed to IL-1β apically and LPS basolaterally. Apical (a) and basolateral (b) IL-8 concentrations comparison for preterm organoids incubated in the OACC system or TC incubator. The plot is coloured by organoid line and shaped by group (NEC or non-NEC). The data is shown as the median between triplicate conditions for each line. IL-8 concentration was measured through ELISA assay, and P values were calculated by applying the Wilcoxon rank test. NEC, necrotising enterocolitis; IL-8, interleukin 8.

6.5 Comparison of preterm and adult derived ileal organoids

6.5.1 Preterm and adult ileal organoids exposed to bacteria

PIO and AIO monolayers following 24 hours of culture in the OACC system and exposed to *E. faecalis* were compared. The transcriptome profiles were significantly different in PIOs compared to AIOs lines (2000 permutations, pR2Y=0.047, pQ2=0.007; Figure 6.6a). A total of 490 DEGs were found when comparing all PIO and AIO monolayers (all adj. p<0.05), of which 252 were downregulated (51%) (Table 6.2, Figure 6.6b). These included immune response associated genes such as eight human leucocyte antigen (HLA) system genes, *IL7*, IL-1 receptor antagonist (*IL1RN*) and *IL10RB*. KEGG pathways significantly over-represented included

'Antigen processing and presentation' and 'Intestinal immune network for IgA production', as well as various metabolic processes (all adj. $p<0.05$, Figure 6.7a). Metabolic processes were also over-represented when the analysis was run selecting the GO BPs database (Figure 6.7b).

Table 6.2. Number of differentially expressed genes and proportion of down and up regulated genes between preterm and adult organoids exposed to the conditions listed.

	Total DEGs	Downregulated	Upregulated
Media only exposure and *E. faecalis* exposure together	490	252 (51%)	238 (49%)
Media exposure	137	82 (60%)	55 (40%)
E. faecalis exposure	1204	490 (41%)	714 (59%)

DEG, differentially expressed gene

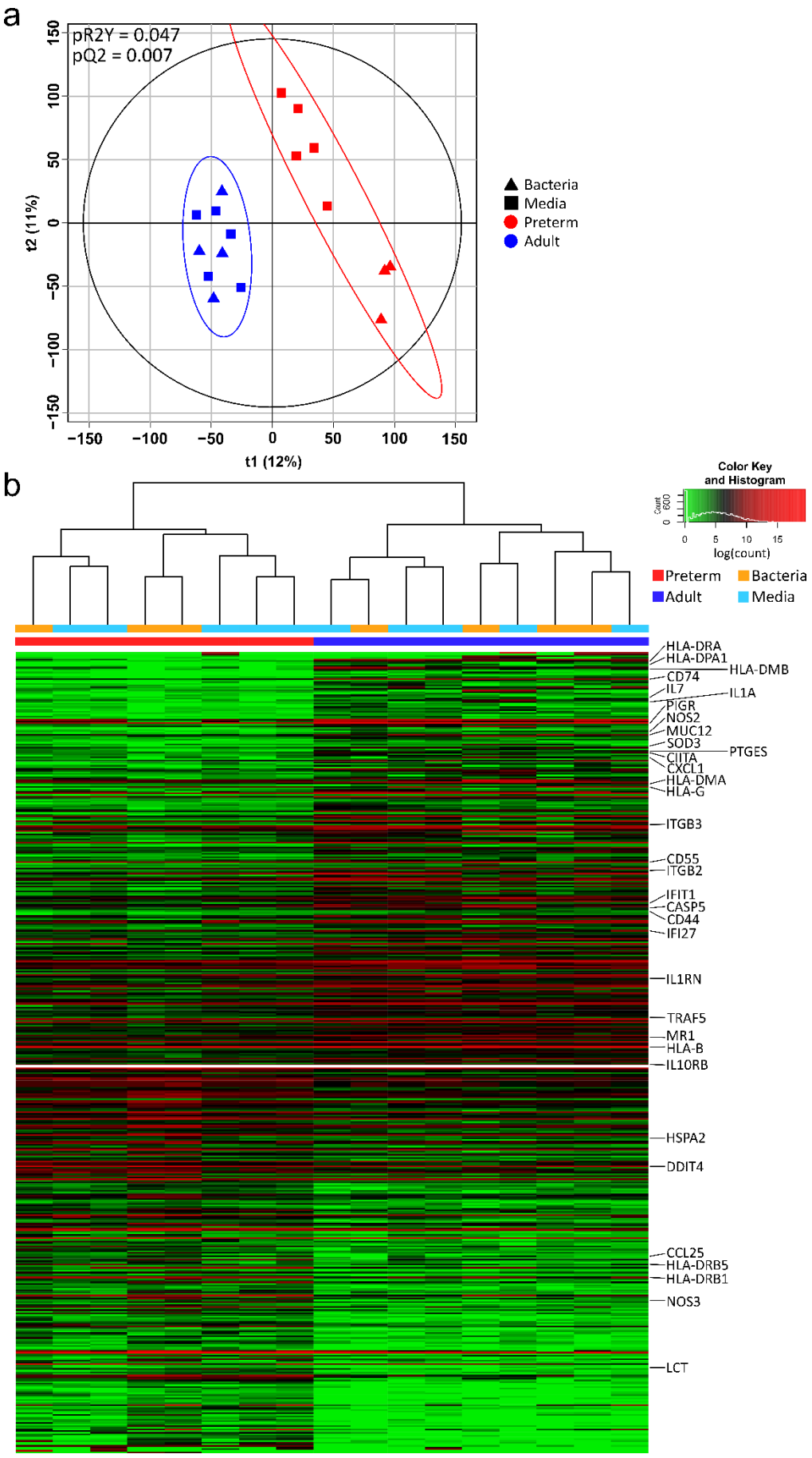

Figure 6.6. Comparison of preterm and adult ileal organoid lines following 24 hours of culture. Bacterial treatment was with *Enterococcus faecalis*. (a) Partial least squares

discriminant analysis of preterm and adult intestinal organoid lines. Plot coloured by line and shaped by exposure (media only or bacteria). P values were calculated based on 2000 permutations. (b) Heatmap of genes differentially expressed between preterm and adult organoids (all adj. p<0.05).

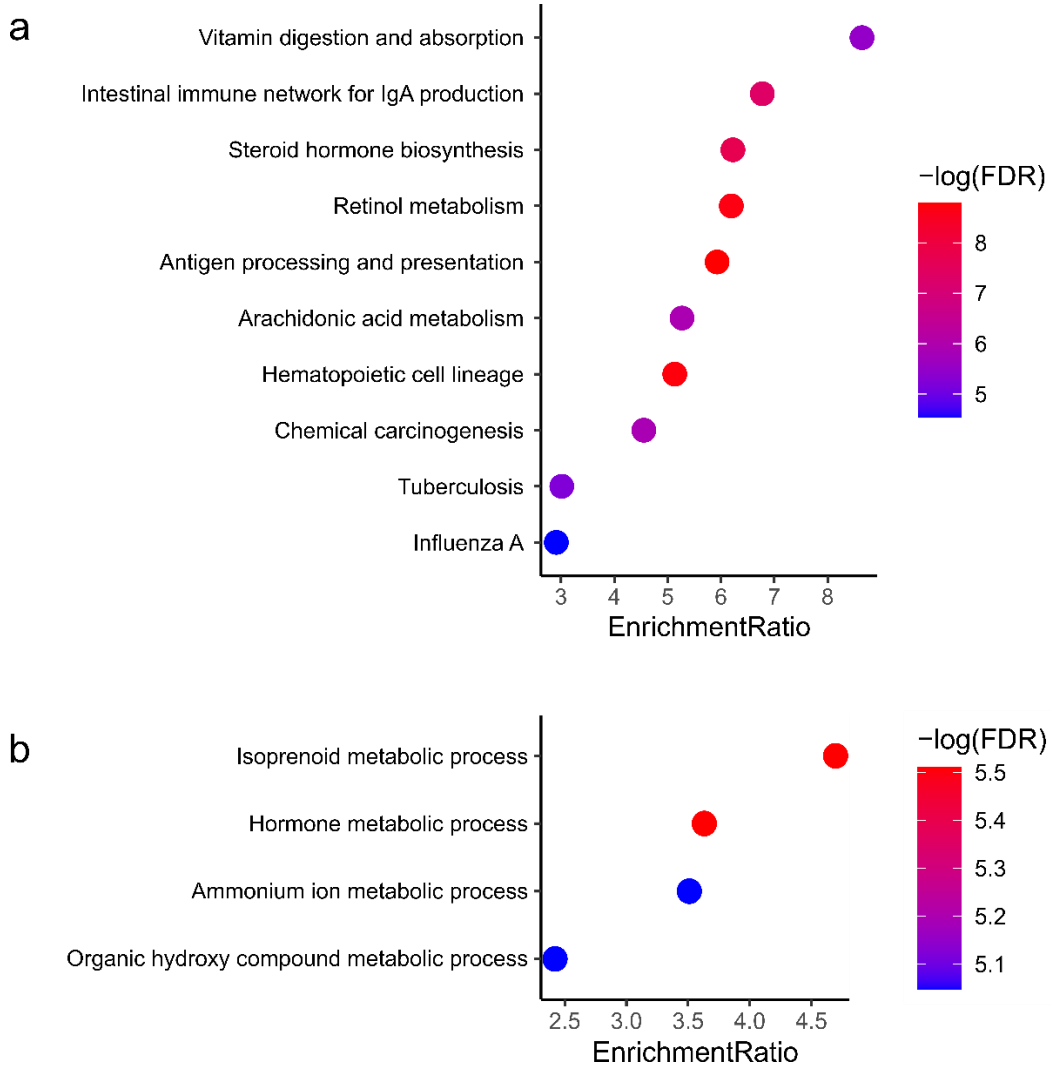

Figure 6.7. Over-representation analysis for the comparison of preterm and adult ileal organoid lines following 24 hours of culture. Bacterial treatment was with *Enterococcus faecalis*. Over-representation analysis for KEGG (a) and GO Biological processes (b) on DEGs from the preterm and adult comparison including all exposure conditions together.

To further explore the response to bacteria within PIOs compared with AIO lines, the two exposure conditions (media only or bacteria) were next analysed separately. While 137 DEGS were found between PIO and AIO monolayers exposed to media only (82 downregulated, 60%), the presence of bacteria increased the number of DEGs to 1204 (490 downregulated, 41%) (Table 6.2). The two comparisons shared 65 common DEGs which were consistently up or down regulated (Figure 6.8a,b), including genes associated with antigen presentation (*CD74*, class II major histocompatibility complex trans-activator (*CIITA*) and five HLA genes), lactase (*LCT*), polymeric immunoglobulin receptor (*PIGR*) and insulin-like growth factor 2 mRNA

binding proteins (*IGF2BP1*, *IGF2BP3*). This suggests that PIOs respond differently to specific bacteria (*E. faecalis*) when compared with ileal AIOs, however, there are some underlying differences which are independent of bacterial exposure.

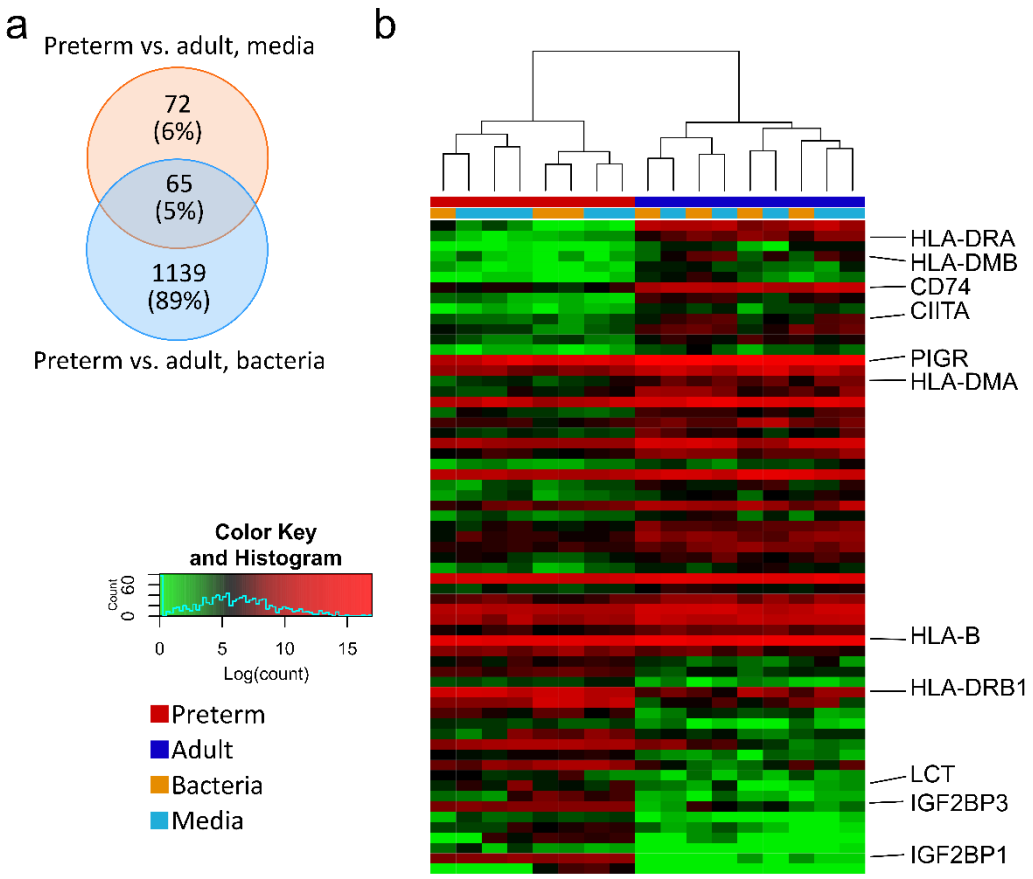

Figure 6.8. Genes differentially expressed between preterm and adult ileal organoid lines independently of exposure condition. (a) Venn diagram of significantly differentially expressed genes between preterm and adult lines within media only or bacterial exposure conditions (all adj. p<0.05). (b) Heatmap of the 65 genes differentially expressed shared between comparisons of preterm and adult organoids separated by exposure condition (all adj. p<0.05).

6.5.2 Preterm and adult ileal organoids exposed to human milk oligosaccharides and comparison with the above dataset

To validate and expand the results from *E. faecalis* exposure, a new experiment was performed exposing six PIO and three AIO lines to HMOs (see methods chapter 2, section 2.7.9). In accordance with the previous experiment, this data showed a clear separation of PIOs and AIOs based on the gene expression profile, with further clustering within each group based on the organoid line (Figure 6.9a). The 'adonis' test was applied to understand how the different variables impact the transcriptome profile of ileal organoids (Figure 6.9b). Organoid line explained 77% of the transcriptome profile variance (adj. p<0.001), followed by age (preterm

or adult) with 26% of variance explained (adj. p<0.001). Exposure condition (IL-1β and LPS with and without the HMOs tested) explained 6% of the variance but was not significant (adj. p=0.9967).

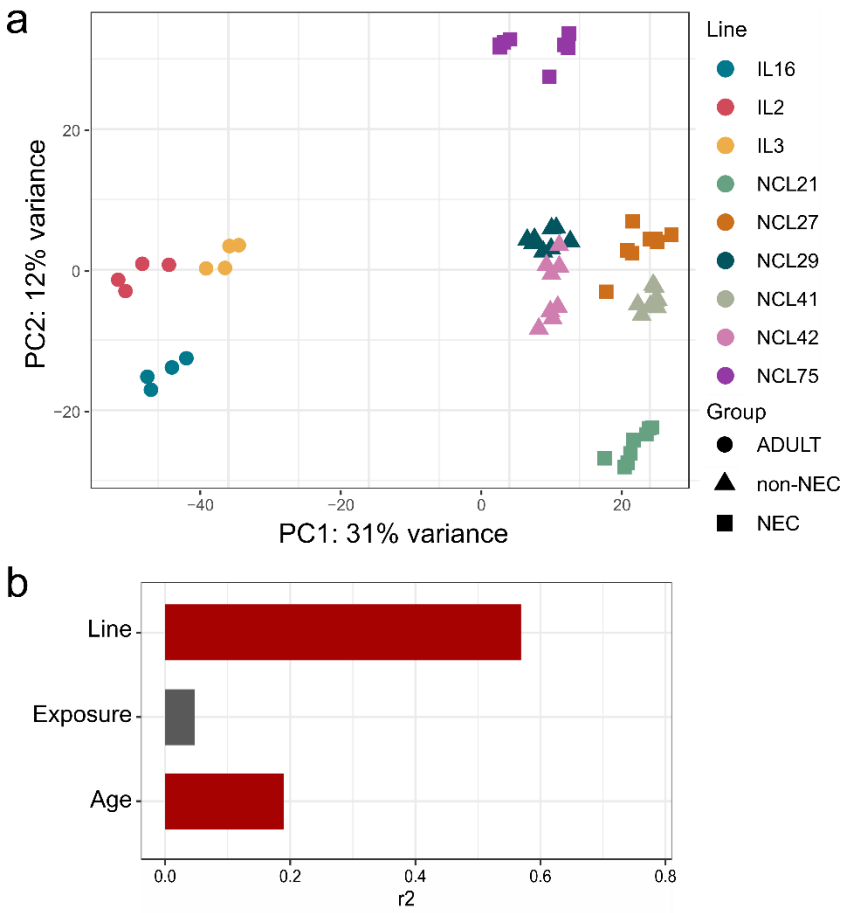

Figure 6.9. Modelling of preterm and adult organoids transcriptome profile using PCA and 'adonis'. (a) PCA of preterm and adult intestinal organoid lines exposed to HMOs. Plot coloured by line and shaped by group (adult, NEC or non-NEC). (b) Horizontal bar plots showing the variance (r2) in organoids transcriptome profile by organoid line, exposure condition (IL-1β and LPS with and without the HMOs tested) and age (adult or preterm) as modelled by 'adonis'. P values were adjusted applying the FDR algorithm. The red bar indicates significant adj. P value < 0.05. NEC, necrotising enterocolitis.

The transcription profile of PIOs and AIOs was then compared after separating by exposure conditions. A total of 813 DEGs were shared between the different exposure conditions (LPS and IL-1β with or without various HMOs), and of these, 129 genes were shared with the dataset described in the previous section and were consistently up or down regulated in every exposure condition (Figure 6.10). Genes shared between the two datasets include apolipoprotein (APO) encoding genes (*APOA4*, *APOE*, *APOL4*), immune response genes (C-C motif chemokine ligand 25 (*CCL25*), C-X-C motif chemokine ligand 1 (*CXCL1)*, interferon alpha inducible protein 27 (*IFI27*), *IL1A*, *IL1RN*, *PIGR*, nitric oxide synthase 3 (*NOS3*), peptidase inhibitor 3 (*PI3*)), *IGF2BP1*, *IGF2BP3*, and *LCT*.

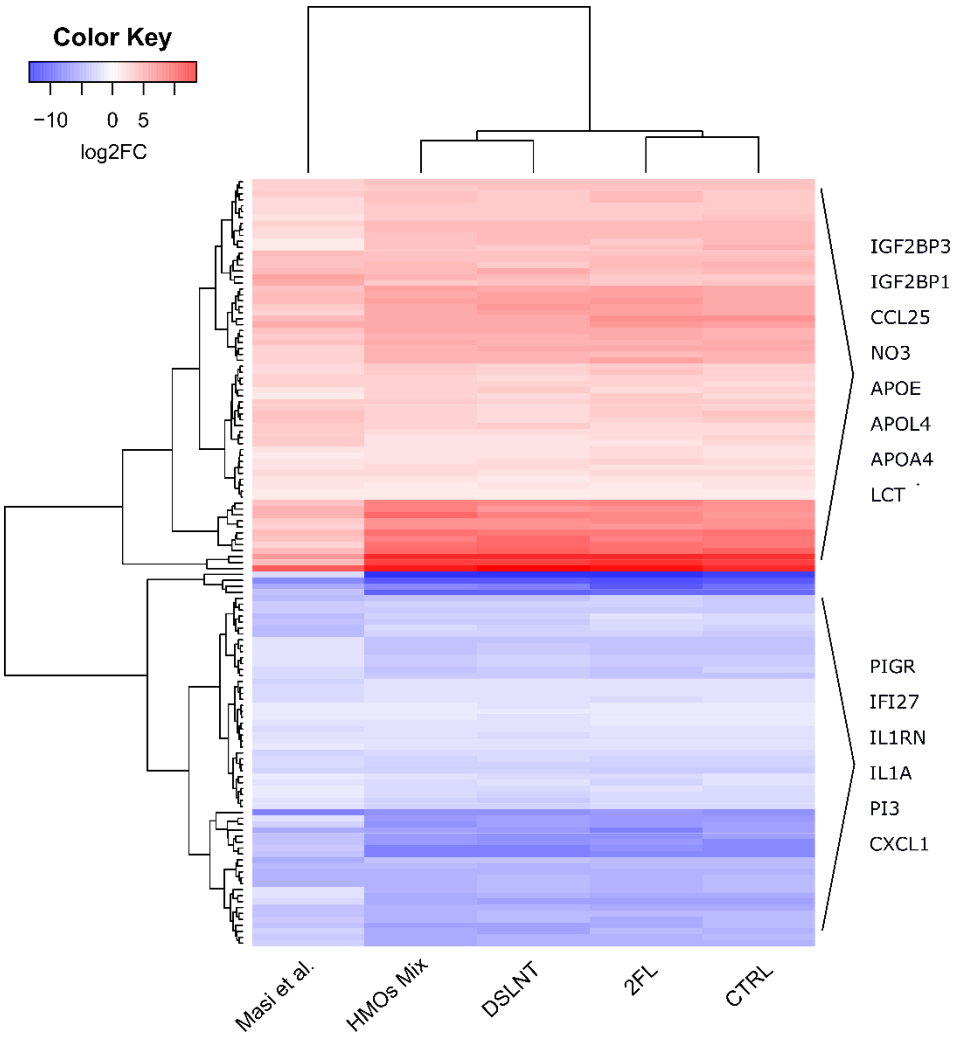

Figure 6.10. Genes differentially expressed between preterm and adult ileal organoid lines and shared between two different datasets. Heatmap of log2 fold changes for the differentially expressed genes when comparing preterm and adult organoids and shared between the published dataset (Masi et al., 2022) and discussed in the previous section, and the HMO dataset analysed after separating by exposure condition (all adj. $p<0.05$). In red are the genes upregulated in PIOs, in blue are the genes which are downregulated in PIOs. Log2FC, log2 fold change; HMO, human milk oligosaccharide; 2'FL, 2'-fucosyllactose; DSLNT, disialyllacto-N-tetraose; CTRL, control.

To further investigate the differences between PIOs and AIOs, the focus was put on the control condition that contained no HMOs (exposure to LPS and IL-1β only). This stratified analysis was performed to specifically explore the differences in response to inflammatory stimuli. A total of 1345 DEGs were found between preterm and adult lines, of which 712 (53%) were downregulated in PIOs. DEGs included intestinal alkaline phosphatase (*ALPI*, upregulated in PIOs) and lysozyme (*LYZ*, downregulated in PIOs), two enzymes key in the intestine homeostasis involved in LPS detoxification and an antimicrobial agent, respectively. Some pro-apoptotic genes were found to be downregulated in PIOs compared to AIOs (BCL2 interacting killer (*BIK*), BCL2 family apoptosis regulator (*BOK*), BH3 interacting domain death agonist

(*BID*)). Moreover, differences in expression of genes belonging to common families were observed between the two PIOs and AIOs. These included differences in mucin genes (*MUC1*, *MUC4*, *MUC5B*, *MUC12*, *MUC20*) and claudin genes (*CLDN1*, *CLDN2*, *CLDN6*, *CLDN9*, *CLDN18*), suggesting differences in barrier function and mucus production, which are key to protection against pathogens. Galectin-encoding genes were found upregulated in PIOs (*LGALS14*, *LGALS7*, *LGALS2*, *LGALS7B*), and are implicated in defence against pathogens, cell differentiation and immune regulation. Genes from the cytochrome 450 family were also differentially expressed between PIOs and AIOs, with *CYP2B6*, *CYP2C8* and *CYP2C9* being upregulated in preterm lines while *CYP4F2*, *CYP4F3* and *CYP4F12* were upregulated in adult organoids.

ORA on the DEGs returned 28 significant GO BPs and 4 significant KEGG pathways. The top 5 significant BPs and the KEGG pathways are reported in Figure 6.11. Both databases returned processes involved in fatty acids metabolism as being different in the two types of organoids.

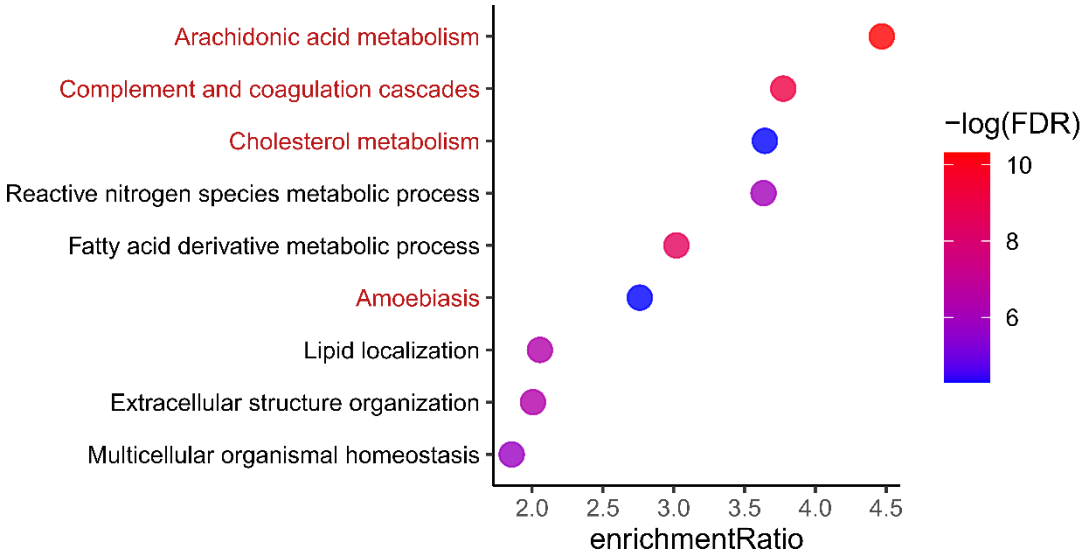

Figure 6.11. Over-representation analysis performed on differentially expressed genes from the comparison of preterm and adult organoids exposed to the control condition. KEGG (in red) and top 5 GO Biological Processes (in black) significantly over-represented in comparison of preterm and adult organoids exposed to the control condition composed by IL-1β basolaterally and LPS apically. All P values were adjusted using the FDR algorithm and deemed significant when adj. $p<0.05$.

IL8 production levels were also measured by ELISA. Significantly lower production and secretion of IL-8 in the apical media was observed in preterm organoids compared to adult lines ($p=0.024$, Figure 6.12a). A similar trend was present in the basolateral media, but this did not reach statistical significance ($p=0.095$, Figure 6.12b). At the gene expression level, significantly lower C-X-C motif chemokine ligand 8 (*CXCL8*; i.e., IL-8 gene) counts were present in preterm samples compared to adult lines, mirroring the IL-8 apical concentration results (lof2FC=-2.38,

adj. p=0.047; Figure 6.12c). At the level of barrier function and integrity, no differences in TEER were found (p=0.262; Figure 6.12d).

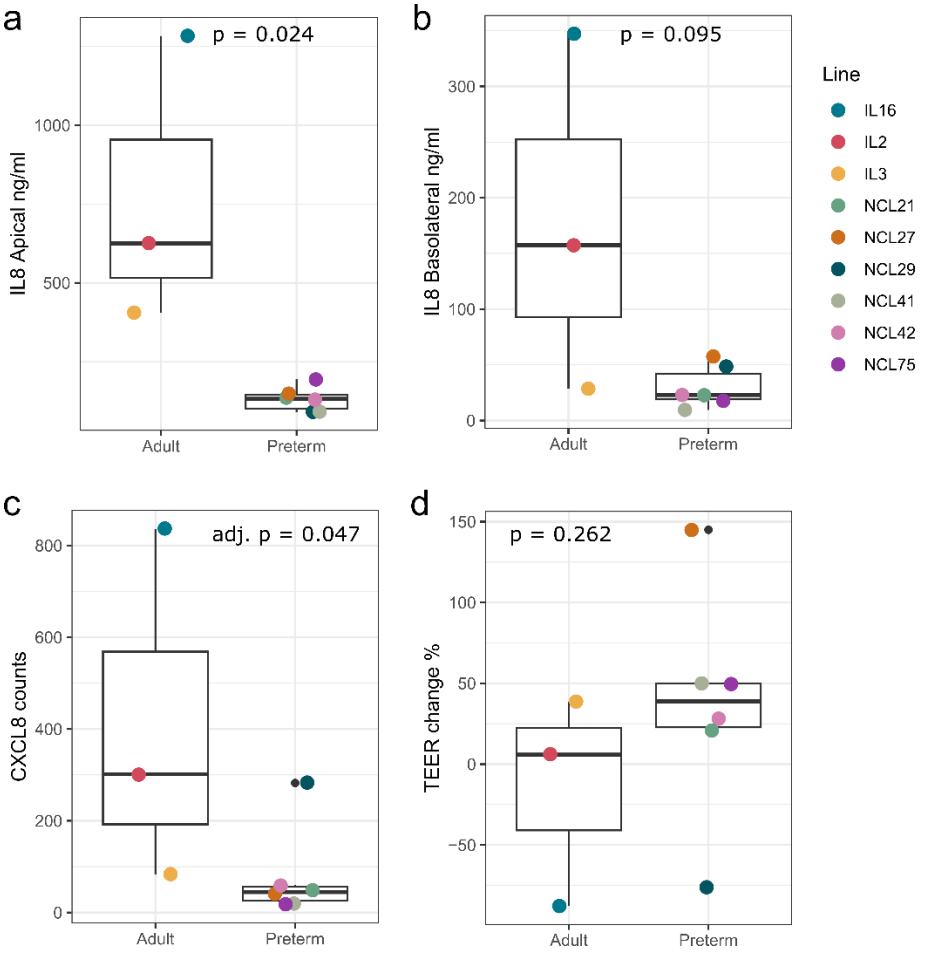

Figure 6.12. Adult organoids produced and released higher concentrations of IL-8 in the apical compartment. All organoids were exposed to IL-1β basolaterallyy and LPS apically. Apical (a) and basolateral (b) IL-8 concentrations comparison for preterm and adult organoids. The data is shown as the median between triplicate conditions for each line. IL-8 concentration was measured through ELISA assay, and P values were calculated by applying the Wilcoxon rank test. (c) *CXCL8* normalised counts for preterm and adult organoids. Log2 fold change and adjusted p-values were calculated using the DSeq2 package in R environment. (d) Boxplot showing Trans-epithelial electrical resistance (TEER) after 24 hours of co-culture shown as the relative change compared to the TEER at the start of the co-culture. Wilcoxon rank test was applied to calculate the P value. IL-8, interleukin 8; *CXCL8*, C-X-C motif chemokine ligand 8.

6.6 Preterm ileal organoids exposure to selected human milk oligosaccharides

6.6.1 Comparison of ileal organoids derived from infant affected by necrotising enterocolitis and control infants

Due to our interest in better understanding what the mechanisms leading to NEC are, before analysing the impact of HMOs on the preterm intestine, the first step was to understand if any

underlying differences were present in the transcriptome profile of organoids derived from NEC infants compared to infants who had surgery for non-NEC related conditions. 3 NEC and 3 FIP lines were then included in our experiments. FIP lines will from now on be called non-NEC lines.

As previously observed, overall transcriptomic profiles clustered by line, with no clear separation depending on NEC diagnosis or HMOs exposure being found (Figure 6.13).

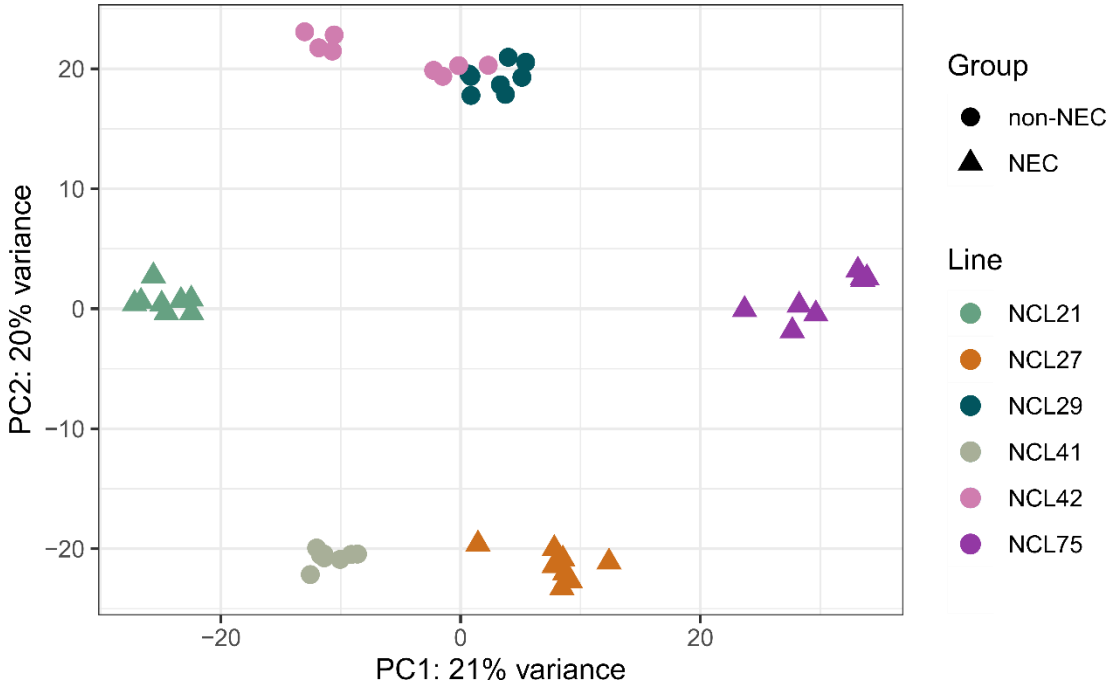

Figure 6.13. Principal component analysis plot of preterm organoids transcriptome profile. (a) Principal component analysis plot of the transcriptome profiles of NEC and non-NEC organoids. All exposure conditions were included, and the plot coloured by organoid line and shaped by group, NEC or non-NEC. NEC, necrotising enterocolitis.

We next sought to analyse the impact of specific HMOs between lines from NEC infants compared to non-NEC controls. Samples were firstly separated by condition exposure (LPS and IL-1β with or without various HMOs), and NEC versus non-NEC comparison was performed. A total of 6 DEGs were shared between all the exposure conditions and were thus differentially expressed independently of addition of the various HMOs. All these 6 DEGs were downregulated in NEC compared to non-NEC lines and included major histocompatibility complex class I H (*HLA-H*) which is a pseudogene, the Family With Sequence Similarity 43 Member A (*FAM43A*), for which no functional information is available on the Uniprot database, and the long intergenic non-protein coding RNA 2535 (*LINC02535*), which is an RNA gene. The other 3 genes encoded for protein WWC3, which is involved in the negative regulation of organ growth, cell migration and regulation of hippo signalling, bone morphogenetic protein 8A (BMP8A), which among other functions is involved in cell

differentiation and activation of the SMAD cascade, and elafin (PI3), which inhibits tissue damage mediated by neutrophil produced elastase.

The difference in response between NEC and non-NEC lines was then analysed separately for each specific HMO. The highest number of DEGs between NEC and non-NEC organoids was observed when organoids were exposed to DSLNT (140) and LNnT (120), followed by 2'FL (84), HMOs Mix (i.e., all HMOs were added to the apical side together with LPS) (81), CTRL (62) and LNT (62), 6SL (33), LNFP I (14) (Table 6.3, Figure 6.14a). The number of genes which were significant exclusively to each condition were also reported. LNnT showed the highest number of exclusive genes (56, 47%), followed by DSLNT (50, 36%) (Table 6.3, Figure 6.14b). DEGs exclusively altered in LNnT exposure included genes associated with regulation of apoptosis (uveal autoantigen with coiled-coil domains and ankyrin repeats, *UACA*), cell cycle (sushi domain containing 2, *SUSD2*) and biosynthesis of O-linked oligosaccharides like mucins (polypeptide N-acetylgalactosaminyltransferase 4 and 13 (*GALNT4*, *GALNT13*)). Similarly, DSLNT exposure specific DEGs included genes associated to cell apoptosis and proliferation (TNF receptor superfamily member 12A (*TNFRSF12A*), *LGALS1*, bradykinin receptor B2 (*BDKRB2*), BCL2 binding component 3 (*BBC3*), hyaluronoglucosaminidase 3 (*HYAL3*), polo like kinase 3 (*PLK3*)), glucose (pyruvate dehydrogenase kinase 4 (*PDK4*), C1q and TNF related 12 (*C1QTNF12*), gastric inhibitory polypeptide (*GIP*)) and lipid (*APOE*, fatty acid binding protein 3 (*FABP3*)) regulation.

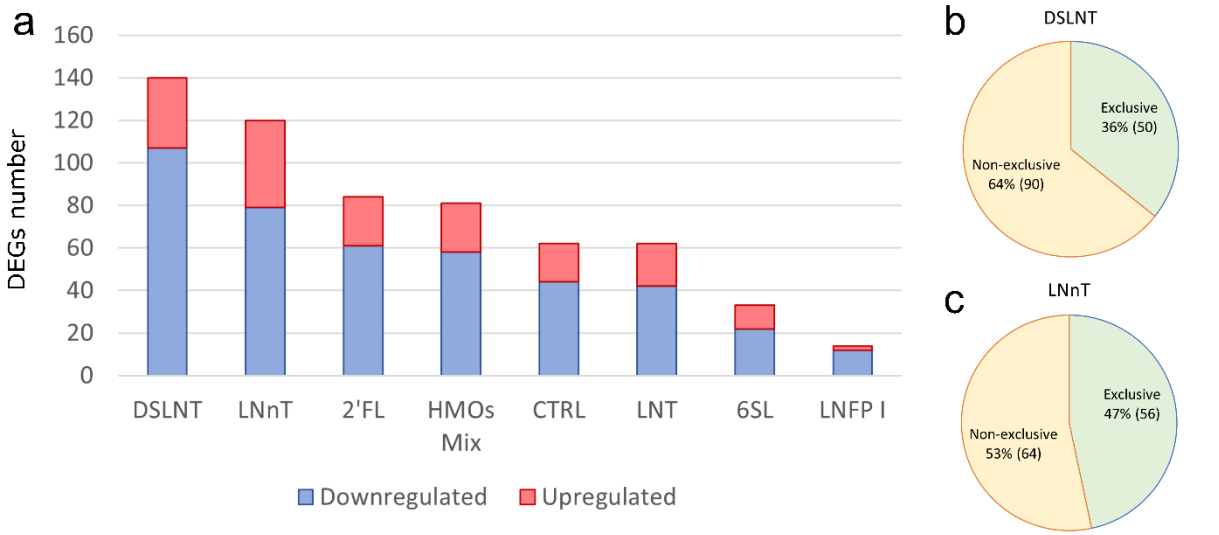

Figure 6.14. Visual representation of numbers of DEGs in NEC and non-NEC comparison after separating by HMO exposure. (a) Stacked bar plot showing the number of up and down regulated genes by comparing NEC vs non-NEC organoids separating each HMO exposure. Pie chart representing number of DEGs exclusive or not to (b) DSLNT and (c) LNnT. DEG, differentially expressed gene; DSLNT, disialyllacto-N-tetraose; LNnT, lacto-N-neotetraose.

Table 6.3. Number of differentially expressed genes and proportion of down and up regulated genes in the comparison between NEC and non-NEC organoids exposed to the conditions listed. 'Exclusive' column refers to the number of genes which were significant exclusively in that specific exposure condition.

	Total DEGs	Downregulated	Upregulated	Exclusive
DSLNT	140	107 (76%)	33 (24%)	50 (36%)
LNnT	120	79 (66%)	41 (34%)	56 (47%)
2'FL	84	61 (73%)	23 (27%)	10 (12%)
HMOs Mix	81	58 (72%)	23 (28%)	8 (10%)
CTRL	62	44 (71%)	18 (29%)	8 (13%)
LNT	62	42 (68%)	20 (32%)	13 (21%)
6'SL	33	22 (67%)	11 (33%)	5 (15%)
LNFP I	14	12 (86%)	2 (14%)	1 (7%)

DEG, differentially expressed gene; DSLNT, disialyllacto-N-tetraose; LNnT, lacto-N-neotetraose; 2'FL, 2'-fucosyllactose; HMOs, human milk oligosaccharides; 6'SL, 6'-sialyllactose; LNFP, lacto-N-fucopentaose; LNT, lacto-N-tetraose.

Next, ORA was performed on the whole set of DEGs for each condition. Significant results were obtained when testing for GO BPs for DSLNT, LNFP I and HMOs Mix, with 8 BPs significant for DSLNT and 1 BP significant for both HMOs Mix and LNFP I (Figure 6.15). For the DSLNT condition, over-represented GO BPs included "response to carbohydrate", "isoprenoid metabolic process" and "response to nutrient level".

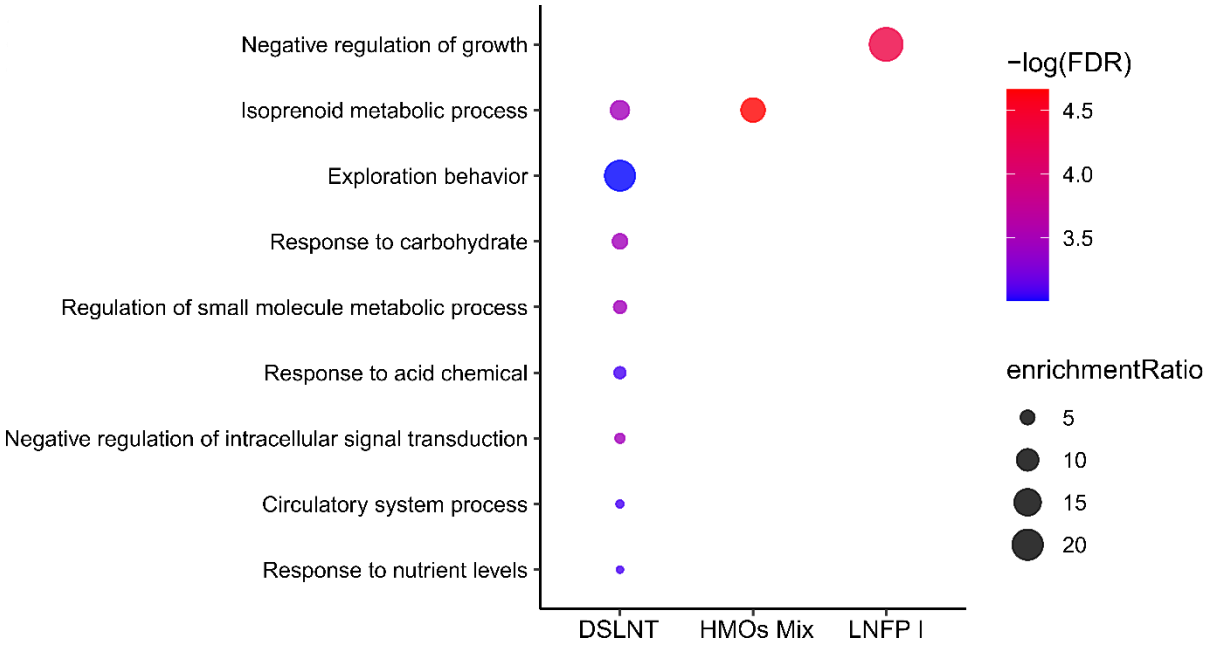

Figure 6.15. Over-representation analysis performed on differentially expressed genes from the comparison of NEC and non-NEC organoids in various exposure conditions. GO

Biological Processes significantly over-represented in comparison of NEC and non-NEC organoids separated by exposure condition. All adj. $p<0.05$. DSLNT, disialyllacto-N-tetraose; HMOs, human milk oligosaccharides; LNFP I, lacto-N-fucopentaose I; FDR, false discovery rate.

No differences were observed between NEC and non-NEC organoids for IL-8 production in either apical or basolateral compartment for any exposure condition, or for TEER change after co-culture relative to the starting point (all adj. $p >0.05$; Figure 6.16)

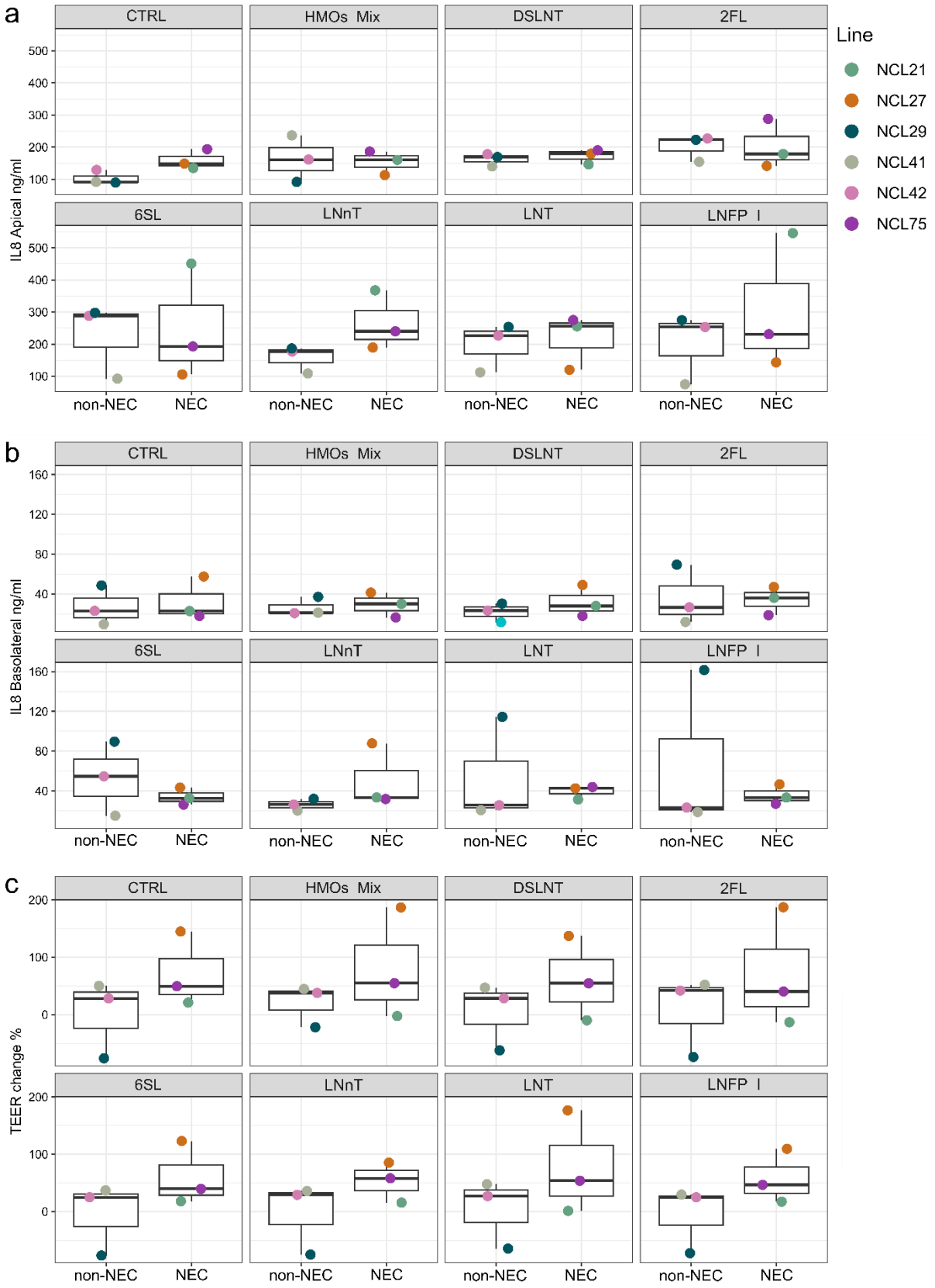

Figure 6.16. No differences in IL-8 production or trans-epithelial electrical resistance were observed between NEC and non-NEC organoids in any exposure condition. Apical (a) and basolateral (b) IL-8 concentrations comparison between NEC and non-NEC organoids exposed to various HMOs. IL-8 concentration was measured through ELISA assay. (c) Boxplot

showing trans-epithelial electrical resistance (TEER) after 24 hours of co-culture shown as the relative change compared to the TEER at the start of the co-culture. All plots are coloured by organoid line. The data is shown as the median between triplicate conditions for each line. P values were calculated by applying the Wilcoxon rank test and adjusted using FDR algorithm. All adj. P > 0.05. IL-8, interleukin 8; TEER, trans-epithelial electrical resistance; CTRL, control; DSLNT, disialyllacto-N-tetraose; LNnT, lacto-N-neotetraose; 2'FL, 2'-fucosyllactose; HMOs, human milk oligosaccharides; 6'SL, 6'-sialyllactose; LNFP, lacto-N-fucopentaose; LNT, lacto-N-tetraose; NEC, necrotising enterocolitis.

6.6.2 Comparison between human milk oligosaccharide exposure and control condition

The previous section analysed the differences *between* NEC vs non-NEC organoid lines. Since a different transcriptome profile was found when comparing NEC and non-NEC organoids, the impact of HMOs exposure compared to the control condition (exposure to LPS and IL-1β without HMOs) were investigated separately *within* NEC and *within* non-NEC lines. Overall, NEC organoids showed a higher number of DEGs for all comparisons, except for the exposure to LNFP I (Table 6.4). The highest response in non-NEC organoids was observed following exposure to LNFP I (17 DEGs) and LNnT (15), while in NEC organoids a higher number of DEGs was found when exposing to LNT (51 DEGs) and LNnT (36 DEGs).

Table 6.4. Number of differentially expressed genes and proportion of down and up regulated genes in the comparison between HMOs listed and control condition after separating NEC and non-NEC organoids.

	Non-NEC organoids			**NEC organoids**		
	Total DEGs	Downregulated	Upregulated	Total DEGs	Downregulated	Upregulated
DSLNT	10	7 (70%)	3 (30%)	13	4 (31%)	9 (69%)
LNnT	15	4 (27%)	11 (73%)	36	21 (58%)	15 (42%)
2'FL	1	0 (0%)	1 (100%)	5	0 (0%)	5 (100%)
HMOs Mix	4	3 (75%)	1 (25%)	14	6 (43%)	8 (57%)
6S'L	3	0 (0%)	3 (100%)	14	12 (86%)	2 (14%)
LNFP I	17	2 (12%)	15 (88%)	14	7 (50%)	7 (50%)
LNT	7	3 (43%)	4 (57%)	51	28 (55%)	23 (45%)

NEC, necrotising enterocolitis; DEG, differentially expressed gene; DSLNT, disialyllacto-N-tetraose; LNnT, lacto-N-neotetraose; 2'FL, 2'-fucosyllactose; HMOs, human milk oligosaccharides; 6'SL, 6'-sialyllactose; LNFP, lacto-N-fucopentaose; LNT, lacto-N-tetraose.

The dataset was then manually screened to search for potentially interesting genes. In various conditions, DEGs included pseudogenes and novel uncharacterised proteins. In non-NEC

organoids, genes deemed of interest included a solute transporter, *SLC28A4*, which was upregulated in 6'SL and LNFP I exposures. In the LNFP I condition, also 6 MT genes were significantly upregulated, whilst 2 of these genes were also upregulated in the LNnT exposure. LNFP I and LNnT conditions shared the upregulation of aldo-keto reductase family (AKRC) member C1 (*AKR1C1*) and C2 (*AKR1C2*) which encode for two cytosolic aldo-keto reductases. In NEC organoids, potential genes of interest included serpin family F member 2 (*SERPINF2*) which was downregulated upon exposure to LNFP I and LNT, and which is a protease inhibitor which targets include plasmin, trypsin and chymotrypsin. Organoids exposure to LNFP I and LNnT showed increased expression of Wnt family member 7a (*WNT7a*) which is involved in cell proliferation and differentiation. Specific to LNnT, a downregulation of F-box protein 10 (F*BXO10*) and carcinoembryonic antigen related cell adhesion molecule 20 (*CEACAM20*) was found, genes which are respectively involved in apoptosis regulation and enhancement of production of *CXCL8*/IL8. In LNnT and LNT, delta like canonical Notch ligand 1 (*DLL1*) was upregulated compared to the control exposure, a gene important for intestinal stem cells homeostasis through the Notch signalling pathway. Finally, DEGs of interest specific to LNT exposure included downregulation of chromogranin A (*CHGA*), and upregulation of *LYZ* and *CEACAM7*, which have roles in the response to intestinal bacteria, and downregulation of *SLC43A2* and *SLC7A5* which are involved in amino acids transport.

No significant differences in IL-8 production and secretion in apical or basolateral media, or in TEER changes relative to starting point, were found in either NEC or non-NEC organoids when comparing the various exposure conditions (adj. $p > 0.05$; Figure 6.17).

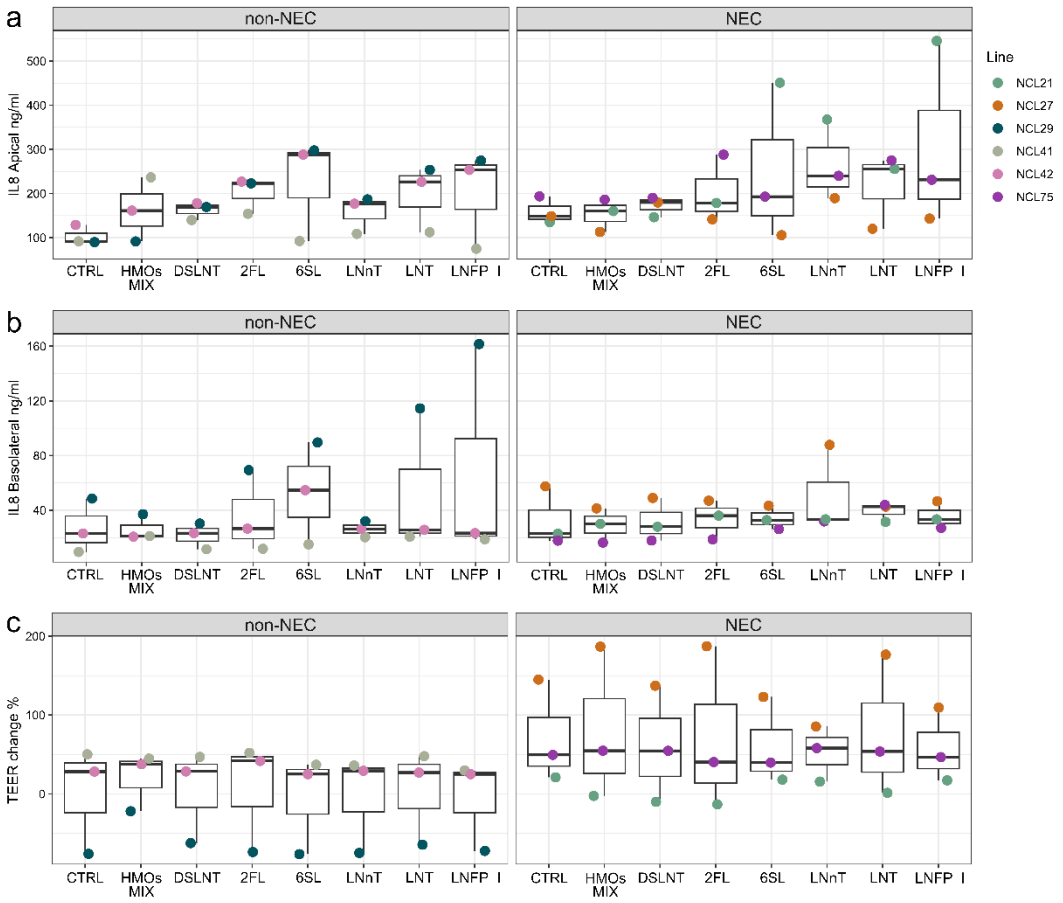

Figure 6.17. No differences in IL-8 production or trans-epithelial electrical resistance changes were observed between HMOs exposure compared to no-HMO control in either NEC or non-NEC organoids lines. Apical (a) and basolateral (b) IL-8 concentrations comparison between exposure conditions after separating NEC and non-NEC organoids. IL-8 concentration was measured through ELISA assay. (d) Boxplot showing trans-epithelial electrical resistance (TEER) after 24 hours of co-culture shown as the relative change compared to the TEER at the start of the co-culture. All plots are coloured by organoid line. The data is shown as the median between triplicate conditions for each line. P values were calculated by applying the Wilcoxon rank test of each HMO condition compared to the control reference and adjusted using FDR algorithm. All adj. P > 0.05. IL-8, interleukin 8; NEC, necrotising enterocolitis; DSLNT, disialyllacto-N-tetraose; LNnT, lacto-N-neotetraose; 2'FL, 2'-fucosyllactose; HMOs, human milk oligosaccharides; 6'SL, 6'-sialyllactose; LNFP, lacto-N-fucopentaose; LNT, lacto-N-tetraose; CTRL, control; TEER, trans-epithelial electrical resistance.

6.7 Adult ileal organoids exposure to selected human milk oligosaccharides

Moving beyond preterm infant organoids, we next analysed the impact of HMOs on adult organoids. For these experiments, only the HMOs DSLNT, 2'FL and HMOs Mix were tested, as these deemed to be of highest interest. A similar number of DEGs was found to all exposures in comparison to the control condition (IL-1β and LPS, but no HMOs), with 22 DEGs in DSLNT exposure and 20 DEGs when organoids were exposed to 2'FL and HMOs Mix (Table 6.5). These genes included pseudogenes and genes encoding for novel proteins for which the

function has not been identified yet. Among the characterised genes, no genes of interest were identified.

Next, DEGs from the same exposure condition for adult, NEC and non-NEC organoids were compared to determine if any common genes were significant following exposure to a specific HMO, independently of age and disease status. Only 1 and 2 DEGs were found to be significant in both adult and NEC organoids when exposed to DSLNT and HMOs Mix, respectively. The DEG common in DSLNT exposure is consistently upregulated in adult and NEC organoids and it encodes for an uncharacterised protein. Of the 2 DEGs significant in HMOs Mix exposure, one is consistently downregulated in both organoid types and the other is upregulated in NEC organoids and downregulated in adult organoids. However, both these genes are pseudogenes. No differences in IL-8 concentration were found in both apical and basolateral media between the various exposure conditions (Figure 6.18).

Table 6.5. Number of differentially expressed genes and proportion of down and up regulated genes in the comparison of adult organoids exposed to the listed conditions and the control exposure.

	Total DEGs	Downregulated	Upregulated	Exclusive
DSLNT	22	11 (50%)	11 (50%)	5 (23%)
2'FL	20	6 (30%)	14 (70%)	5 (25%)
HMOs Mix	20	12 (60%)	23 (40%)	7 (35%)

DEG, differentially expressed gene; DSLNT, disialyllacto-N-tetraose; 2'FL, 2'-fucosyllactose; HMOs, human milk oligosaccharides.

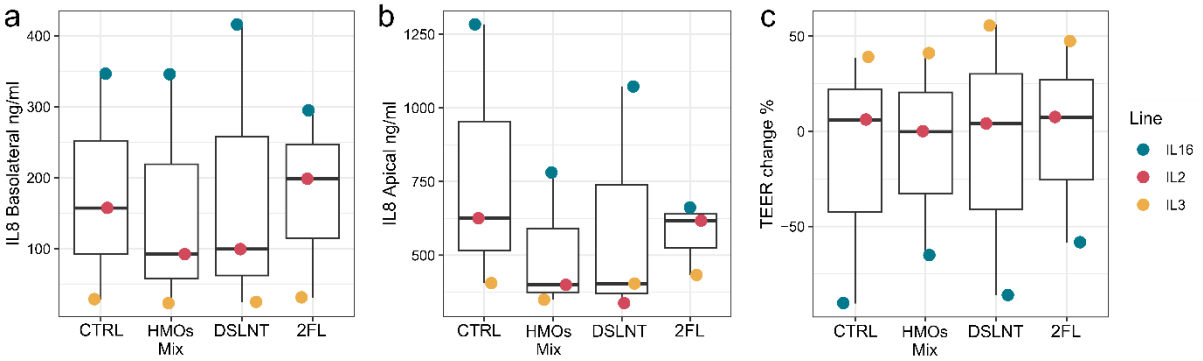

Figure 6.18. No differences in IL-8 production or TEER were observed between control condition and HMOs exposure in adult organoids. Apical (a) and basolateral (b) IL-8 concentrations comparison between exposure conditions in adult organoids. IL-8 concentration was measured through ELISA assay. (c) Boxplot showing trans-epithelial electrical resistance (TEER) after 24 hours of co-culture shown as the relative change compared to the TEER at the start of the co-culture. All plots are coloured by organoid line. The data is shown as the median between triplicate conditions for each line. P values were calculated by applying the Wilcoxon rank test and adjusted using FDR algorithm. All adj. p>0.05. TEER, trans-epithelial electrical

resistance; IL-8, interleukin 8; CTRL, control; HMOs, human milk oligosaccharides; DSLNT, disialyllacto-N-tetraose; 2'FL, 2'-fucosyllactose.

6.8 Discussion

Necrotising enterocolitis is a disease which affects the intestine of preterm infants, and human intestinal organoids are a powerful model that allow for mechanistic investigation of the interaction between the human intestine and various components including HMOs and microbes. Using a novel organoid-anaerobe co-culture system, the impact of culture time, differences between preterm and adult derived intestinal organoids, and impact of HMOs were investigated using RNA-sequencing.

HIOs derived from preterm infants were used as a model to study the potential relation between HMOs and the NEC development. However, the first step was to evaluate how various factors impact the HIOs transcriptome profile. At the start of my PhD, RNA-sequencing data was available from preterm and adult HIOs exposed to *E. faecalis*, an abundant gut microbe that is known to cause neonatal sepsis. The information gained from these experiments helped to inform the experiments performed as part of this thesis, where preterm and adult HIOs were exposed to HMOs. Since HMOs are reported to have a direct anti-inflammatory effect on intestinal cells (Lane et al., 2013, He et al., 2016, Holscher et al., 2017), HIOs were also exposed to IL-1β (basolateral compartment) and LPS (apical compartment) to induce an inflammatory stimulus which was hypothesised would be counteracted by HMOs. These two pro-inflammatory stimuli were chosen to recapitulate NEC, and due to COVID 19 delays, their concentration was chosen based on published studies (Meng et al., 2020, Ling et al., 2016, Zheng et al., 2020). IL-1β has been previously used by a world-leading group working on NEC development mechanisms to mimic disease inflammation *in vitro* (Zheng et al., 2020, Meng et al., 2020). LPS is the component of Gram-negative bacteria that is recognised by TLR4, which is reported to be upregulated in the intestine of preterm infants, and LPS has been previously utilised as stimulus in NEC research (Ling et al., 2016, Senger et al., 2018). Higher relative abundance of Gram-negative bacteria is positively associated to NEC, and this could be caused by a higher presence of LPS in the preterm intestine which could trigger TLR4 response.

The first condition analysed was the impact of exposure time on the transcriptome profile of preterm HIOs. A difference in transcriptome profile was observed in preterm organoids co-cultured for 8 hours compared to 24-hour time point. Importantly, this difference was enhanced when HIOs were exposed to bacteria, which resulted in a higher number of genes differentially expressed, including genes related to immune response and apoptosis that are relevant when studying bacterial exposure. Considering the transcription of genes dynamically changes

depending on the stimuli received by cells, the choice of time of exposure might be of vital importance to observe a response at the RNA level. Various genes and pathways will need different times to be activated and show a variation in expression upon exposure to the desired stimulus. Based on the results discussed, 24 hours co-culture of HIOs and components of interest was chosen for subsequent experiments, to try and obtain a wider gene response.

Transcriptome data on impact of oxygen concentration on the transcriptome profile of preterm HIOs exposed to IL-1β and LPS was produced. Indeed, PIOs were kept either in a TC incubator, characterised by estimated 17% of oxygen or in the OACC system which provides 5.6% oxygen to the organoids. Fofanova *et al.* (2019) were the developers of the OACC system, in which 5.6% oxygen is supplemented to the HIO monolayers to mimic the low oxygen levels observed in the gut at the lumen interface (Fofanova et al., 2019). Among the results reported, gene expression data acquired through a PCR array was included, with focus on tight junction and immune-response related genes. Comparison of the results showed that none of the genes reported as being differentially expressed by Fofanova and colleagues were significant in this dataset, potentially as a consequence of using two different techniques to investigate gene expression. Moreover, this study used preterm ileal organoids, while their experiments involved adult jejunal organoids, which might have underlying differences in their transcriptome profile and response to hypoxia. Indeed, gene-expression differences have been reported depending on the intestinal region the HIOs were generated from (Middendorp et al., 2014). The exposure conditions were also different, as no stimulus was used by Fofanova *et al.* (2019), while in the current study preterm HIOs were exposed to inflammatory stimuli IL-1β and LPS. Finally, despite not finding a down-regulation in *CXCL8* in the OACC system compared to exposure in TC incubator, a significantly lower concentration of IL-8 was found in the apical compartment in this study, in agreement with the expression profile reported by Fofanova *et al.* (2019) (Fofanova et al., 2019).

In this study, differences found when comparing TC incubator and OACC system included ABC transporters and SLC proteins, which as a consequence might alter the transport of molecules through the cellular membrane. Interestingly, 9/19 MT genes present in the human genome were upregulated in the OACC system compared to the TC incubator. Previous studies reported correlation between hypoxia and upregulation of MT genes, however, they were in carcinogenic non-intestinal cells, including prostate cancer cells (Yamasaki et al., 2007) and bladder carcinoma cells (Tsui et al., 2019). Up-regulation of MTs was also found to protect against hypoxia-induced damage in pancreatic beta cells (Li et al., 2004) and rat kidneys (Kojima et al., 2009). On the contrary, in the gut epithelium, an inverse correlation between hypoxia (and hypoxia inducible factor 1 alpha (*HIF1A*)) and MT genes has been reported

(Devisscher et al., 2011). *HIF1A* is the master regulator of hypoxia, and the expression of this gene was not significantly altered by oxygen availability in this study. Notably, the adjusted P value for *HIF1A* was < 0.05, but the log2 fold change was only -0.56 (significance for log2 fold change was set at 1, absolute value). Despite not passing our threshold for statistical significance, a similar inverse correlation between MTs and *HIF1A* was observed, even if it was opposed to what could be expected since *HIF1A,* which is induced during hypoxia, was downregulated. Notably, the oxygen conditions applied in the OACC system aim to reflect the low levels of oxygen epithelial cells are exposed to *in vivo,* especially at the interface with the anoxic lumen. Thus, the levels used in the system might not be perceived as hypoxia by the cells, explaining why a downregulation of *HIF1A* and upregulation of MTs were observed.

Using HIOs as a model has many strengths, including the potential to derive organoids from multiple patients each having a different genetic composition, they retain the intestinal region-specific gene profile and morphology, and age-specific organoids can be generated. Even though region specificity in transcriptome profile has been previously investigated (Middendorp et al., 2014, Kayisoglu et al., 2021, Kraiczy et al., 2019), there are not many studies exploring the differences depending on patient age especially for premature cohorts. Differences in HIOs transcription profiles based on gestational age have relied on murine embryo-derived HIOs which were compared to adult ones (Kayisoglu et al., 2021). Kayisoglu and colleagues investigated the expression of innate immune system genes in various regions of murine embryo- and adult-derived HIOs. Despite some of the genes involved in immune response were location-specific and expressed at a similar level in adult and embryo organoids, there were some differences in expression profile depending on the developmental stage (Kayisoglu et al., 2021). However, there is a lack of exposure to the environment in murine embryo-derived organoids, which is not representative of preterm infants who develop NEC.

Owing to the two datasets analysed and discussed in this thesis, a comparison of preterm and adult ileal organoids could be evaluated. A clear separation between PIOs and AIOs transcriptome profile was observed. Even though DEGs unique to specific exposures, including bacteria and HMOs, were found, there were also genes which were differentially expressed in a consistent direction between the various exposures and between the two different datasets. Interestingly, an upregulation of *LCT* in preterm compared to adult organoids was observed. In humans the highest expression in lactase is observed at birth and during milk feeding and declines after weaning (Deng et al., 2015). Polymeric IgA receptor (*PIGR*) is involved in the transport of IgAs across the intestinal epithelium and was downregulated in preterm organoids, potentially reflecting the lack of endogenous IgA production in preterm infants in the first weeks of life (Gopalakrishna et al., 2019). The IGF2BP family of proteins is key during

development, and in particular IGF2BP1 and 3 are expressed mainly in early life and present at low levels in adults (Ramesh-Kumar and Guil, 2022), and were upregulated in PIOs compared to AIOs. These results suggest that some of the age related differences in gene expression observed *in vivo* in the human intestine are maintained in the tissue derived organoids from the different life stages. This underscores that preterm derived organoids might be a better model for the study of prematurity related diseases such as NEC.

A further analysis of differences between PIOs and AIOs was performed on organoids exposed to IL-1β and LPS. Interestingly, PIOs showed a higher expression of the intestinal alkaline phosphatase (*ALPI*), which is involved in detoxification of LPS. This was contrary to what has been reported in a rat pup model, which showed lower *ALPI* expression compared to adult rats (Whitehouse et al., 2010). On the contrary, lysozyme (*LYZ*) was downregulated in PIOs compared to AIOs, which might be related to the lower number of Paneth cells and to their immaturity observed in the neonatal intestine compared to adult intestine (Demers-Mathieu, 2022). LYZ is an enzyme participating in the innate immunity as ALPI, but has activity on peptidoglycan and has bactericidal activity on both Gram-positive and Gram-negative bacteria. Interestingly, *ALPI* decreased expression (Biesterveld et al., 2015) and higher expression of *LYZ* (Diez et al., 2022) have been associated with NEC development. Mucins and claudins exert a pivotal role in gut health, the first by forming a mucus layer which protects the intestinal cells from potential invading pathogens and also through cell signalling (Grondin et al., 2020), while claudins are involved in tight junction formation and thus barrier integrity (Lu et al., 2013). Gastrointestinal immaturity in preterm infants increases the risk of a leaky gut and alteration in expression of specific members of mucin and claudin family have been associated with NEC (Pan et al., 2021, Liu et al., 2021, Clark et al., 2006, Buonpane et al., 2020, Ravisankar et al., 2018). In this study, four mucin and five claudin genes were differentially expressed between preterm and adult organoids. Finally, galectin encoding genes were upregulated in PIOs. These are receptors which are involved in the regulation of immune and inflammatory responses, and they are also among the receptors which can interact with HMOs (Rousseaux et al., 2021), thus of particular interest in NEC and infant research.

Understanding what the drivers of the gene expression profile in HIOs are is key to be able to fully interpret the results obtained. The variance explained by various factors on HIOs transcriptome profile was investigated in organoids exposed to HMOs. Organoid line (i.e., the patient from which the line was derived) had the highest impact on overall transcriptome profiles, followed by the age of the patient the HIOs were derived from (i.e., preterm or adult) Exposure condition (i.e., IL-1β and LPS with and without the various HMOs tested) was not significantly associated with organoid transcriptome profiles. This is partly in accordance with

the study by Criss *et al.* (2021) that investigated factors influencing transcriptome profiles in HIOs by integrating multiple datasets generated as part of separate projects testing different treatments, which included calcitriol, bacteria, LMD4 bacterial media, and norovirus and rotavirus strains (Criss et al., 2021). It was found that the highest variance was explained by impact of format (2D vs 3D HIOs), substrate (Matrigel or collagen), differentiation status (differentiated or non-differentiated), project/treatment and intestinal region (Criss et al., 2021, 2nd et al., 2021). However, in their dataset it was also found that, while line partly explained the transcriptome profile variance, no impact of treatment (which was specific to each project included) was observed, in agreement with the results reported in this thesis. That patient-to-patient variability was the significant factors associated with HIOs transcriptome further highlights the advantage of using patient derived organoid lines over carcinogenic lines, which are unlikely to capture true biological complexity and individuality. This could also limit generalisation of the results and hinder translation.

Multiple studies to date have investigated the differences in gene expression profiles directly in NEC compared to non-NEC tissues in preterm infants (Chan et al., 2014, Tremblay et al., 2021, Tremblay et al., 2016). However, no comparison at the organoid level has been made. Interestingly, six genes were downregulated in NEC compared to control HIOs in a consistent manner in all eight conditions tested. These included *WWC3* which has been demonstrated to control cell differentiation, proliferation and apoptosis through the hippo signalling in vascular smooth muscle cells (Chen and Liu, 2018). WWC proteins might also act through the Wnt/β-catenin signalling pathway in glioma cells (Wang et al., 2018), pathway that is important in intestinal epithelial cells differentiation and development and in crypt formation (Kandasamy et al., 2014) and which might be central to NEC disease. Peptidase inhibitor 3 (*PI3*, elafin) was also downregulated in NEC organoids. Elafin has anti-inflammatory properties thanks to its anti-peptidase activity against neutrophil derived elastase and proteinase 3 (Motta et al., 2011). Moreover, it has been demonstrated to have anti-bacterial and anti-fungal activity *in vitro* against various species including the Gram-negative *P. aeruginosa* (Simpson et al., 1999) and *K. pneumoniae* (Baranger et al., 2008). The role of neutrophils in NEC development is yet to be elucidated. While likely necessary for infection clearance, excessive neutrophil activity might cause tissue damage (Christensen et al., 2015, Musemeche et al., 1991). In an IBD mouse model, elafin expression was associated with protection against colitis (Motta et al., 2011). No association between elafin and NEC development has been reported to date, however, due to the similarities between NEC and IBD, this gene/protein might play an anti-inflammatory role in this multifactorial disease.

Subsequently, the impact of HMOs exposure was investigated. No clear differences in response to the various HMOs was observed in the comparison between NEC and non-NEC HIOs. The biggest difference was observed when organoids were exposed to LNnT and DSLNT. This might be of interest since these two HMOs were found to be the only HMOs significantly different between MOM samples fed to NEC and matched controls (chapter 3). LNnT was lower in the most severe cases of NEC, and DSLNT was in lower in NEC independently of disease severity. Galectin 1 (*LGALS1*) was exclusively differentially expressed (downregulated) in NEC compared to non-NEC organoids exposed to DSLNT. Importantly, *LGALS1* is involved in modulating the viability of enterocytes in the mouse, and might thus influence integrity of the intestinal epithelium and villus (Muglia et al., 2011).

IGFBP3 was found to be downregulated in NEC compared to non-NEC organoids when exposed to DSLNT, 2'FL and the HMOs Mix. DEGs also included components of the isoprenoid/retinoid metabolic process and glucose regulation. IGFBP3 is a member of the family of proteins binding IGF1 and it is the most abundant circulating IGFBP in the blood (Varma Shrivastav et al., 2020). Its role is to stabilise IGF1 in the blood so that it can reach the target tissue, also defining its action and availability (Varma Shrivastav et al., 2020). IGF1 is the main hormone contributing to the foetus growth in utero and its levels drop below corresponding in-utero concentrations after preterm birth (Hellström et al., 2016). Among its functions, IGF1 regulates glucose uptake (Hellström et al., 2016) and in the intestine promotes cell proliferation and differentiation and has an apoptosis inhibitory effect (Baregamian et al., 2006, Hunninghake et al., 2010, Jeschke et al., 2007). Impact of preterm supplementation with recombinant human (rh) IGF1/binding protein-3 (rhIGF1/BP3) is currently being tested in clinical trials for the protection against the development of complications of prematurity, with focus on retinopathy of prematurity (Ley et al., 2019). The potential outcome on preterm gut health and NEC has not been fully investigated yet. In a preterm pig model of NEC, supplementation of rhIGF1/BP3 reduced disease severity and incidence of severe NEC lesions in the colon, while in the small intestine it was associated with increased villi length and brush border enzymes activity (Holgersen et al., 2020). Moreover, IGFBP3 can also interact with lactoferrin which is thought to be protective against NEC development, and with genes involved in retinol metabolism (Varma Shrivastav et al., 2020), for which alterations were observed in the NEC and non-NEC organoids comparison in response to DSLNT. Together, changes in expression of IGFBP3 and of retinol metabolism and glucose regulation genes in NEC organoids in response to exposure to various HMOs, including DSLNT, might be related to the mechanisms leading to disease onset. However, the role of these genes and pathways is yet to be determined and further studies are needed.

A major outstanding question regarding the potential mechanisms of HMO in improving health relates to their ability to act directly on the host (and not only as a prebiotic). We sought to test the potential anti-inflammatory properties of HMOs by exposing HIOs to immune and bacterial inflammatory stimuli (IL-1β and LPS) with and without HMOs. Compared to the control condition which was represented by media supplemented with IL-1β and LPS, no clear response was found at the transcriptome profile level and no differences were present in IL-8 production in any of the organoid groups investigated when HMOs were added. In NEC organoids, LNnT exposure caused upregulation of specific genes *WNT7a* and *DLL1*, which are both involved in cell proliferation and differentiation, and a downregulation of *FBXO10* which is involved in the apoptotic process. Moreover, there was a downregulation of *CEACAM20* which is involved in the enhancement of *CXCL8*/IL-8 production, although this was not reflected in *CXCL8* expression or IL-8 concentration in the apical and basolateral media. These results contrast data from a small intestine foetal cell line that found LNnT reduced IL-8 production (Cheng et al., 2021).

Many studies have investigated the response of human intestinal cells to HMOs, either as a pool or as single HMOs, finding that these sugars have an impact on barrier function (Natividad et al., 2020), expression of immune genes (Huang et al., 2021, Sodhi et al., 2020, Lane et al., 2013) and goblet cell markers (Wu et al., 2022, Figueroa-Lozano et al., 2021). However, most studies utilised a single cell line, including transformed colonic cells (Wu et al., 2022, Lane et al., 2013, Figueroa-Lozano et al., 2021), while in this study ileum sections were used. In studies that utilised multiple cell-lines, the ratio between the cell-types was experimentally defined (Natividad et al., 2020), which cannot be controlled for in HIOs derived monolayers. Indeed, differentiated intestinal organoid derived monolayers contain most epithelial cell types including enterocytes, goblet cells and enteroendocrine cells, and information on their proportion in the samples used was not available. Thus, differences observed at the gene expression level might be a consequence of the different composition of the various organoid monolayers used. Using single cell RNA-seq instead of bulk-RNA would help to disentangle this, but such a technique is expensive and outside the scope of the current work.

The mechanism behind DSLNT protection from NEC is not yet understood. *In vitro,* sialylated HMOs have been demonstrated to prevent neutrophil rolling and adhesion to endothelial cells by interacting with selectins. However, DSLNT does not have the structure recognised by selectins as it lacks the fucose, and therefore might be exerting other functions. A novel study correlated DSLNT protection against NEC intestinal damage with mast cells (MC) activity (Zhang et al., 2021). MCs are residential immune cells residing at the mucosal border involved in innate and adaptive immunity and in maintaining gut homeostasis (Albert-Bayo et al., 2019).

DSLNT reduced MC degranulation, prevented MC accumulation and damage in the rat NEC model intestine and protected neonatal ileum segments from damage caused by MC chymase, an enzyme produced by these cells and which has roles in tissue remodelling (Zhang et al., 2021). These results were further confirmed in the Caco-2 colonic cell line, in which DSLNT restored *ZO-1*, *FAK* and *P38* expression upon MC chymase treatment, with subsequent effects on cell integrity and cell cycle (Lian et al., 2022). The protective effect of DSLNT against intestinal epithelium damage was observed upon treatment with MC chymase in absence of entire mast cells. This suggests that DSLNT either blocks MC chymase, or directly interact with the intestinal epithelium to modulate its response to MC chymase. Pre-incubating MC chymase with DSLNT before investigating its effect on intestinal cells, might help understanding the mechanism behind the putative protective effect exerted by this HMO.

6.9 Study limitations and conclusions

Despite the novelty of some of the organoid work performed in this chapter, some limitations should be considered. In the first dataset discussed HIOs were only exposed directly to bacteria. Including bacterial supernatant and non-live bacteria cells would also help understanding if the response of HIOs to the bacterium is a consequence of the direct interaction of the microorganism with the intestinal cells, or if it is due to the effect exerted by bacterial-produced metabolites. The comparison of intestinal organoids derived from preterm and adult patients in the second dataset, and from NEC and non-NEC preterm infants, was only performed on organoids exposed to some stimuli (HMOs and/or IL-1β and LPS). Including non-stimulated (i.e., basic organoid media only) organoids in future studies might be beneficial to understand the differences observed between the various organoid types and stimuli. Despite finding differences in NEC and non-NEC organoids, it could not be determined if these alterations were present before disease status and might be linked to its development, or if they are a consequence of the conditions the intestine was exposed to as a result of disease onset.

Bulk RNA-sequencing was used to study the gene expression profile, while single-cell RNA-sequencing would have aid deeper understanding of the results obtained. HIOs are composed by multiple intestinal cell types and the differences observed between PIOs and AIOs or NEC and non-NEC organoids might be a consequence of differences in abundances of the distinct cell types which have a diversified gene expression profile. Single-cell RNA sequencing data could have also informed on the response of each specific cell type to the stimuli added. Utilising various organoid lines is one of the strengths of this study, as the effect of the stimulus can be investigated on patient specific lines with different genetic and epigenetic backgrounds,

which is more reflective of the human population. However, to account for the variability in responses and gene expression for each patient line, the study would have benefited from having additional lines in each group.

The exposure of intestinal organoids to HMOs might benefit from further optimisation. No test to assess the purity of the HMO products used was performed. HMOs are often produced in engineered bacteria, such as *E. coli*, thus LPS or other bacteria metabolites might be present in the HMO solution and affect the results. HMOs together with an inflammatory stimulus were added contemporaneously, and thus the intestinal cells might have had too many stimuli to adapt to, rendering the output of the experiment difficult to interpret. In this study, the HMOs were tested as a single dose corresponding to the average concentration in the breast milk from mothers whose infants did not develop any intestinal disease. Testing multiple doses with increasing HMOs concentrations will be useful to understand if a dose-response can be observed in HIOs.

In summary, differences in intestinal organoids transcriptome profiles were found based on various factors. Co-culture time affects the gene expression profile with a longer incubation resulting in an enhanced response when preterm intestinal organoids were exposed to *E. faecalis*. The main factor determining transcriptome profile was organoid/patient line. However, age (preterm or adult) and disease status (NEC or non-NEC) also significantly influenced the transcriptome profile of intestinal organoids. Differences in immune-response and metabolism-related genes were observed, and some of these differences were independent of exposure condition. This suggests that the preterm intestine's prematurity is partly maintained in tissue-derived intestinal organoids, and preterm organoids might represent a more relevant model to study diet-microbe-host interaction in this population and in relation to relevant diseases such as NEC. NEC and non-NEC organoids were also different in their transcriptome profile, and some of the alterations were independent of exposure conditions. Further studies are needed to better understand the interaction between HMOs and preterm intestinal organoids.

Chapter 7. General discussion

7.1 Introduction

The aim of this study was to investigate the interaction between MOM feeding, the infant gut microbiome and preterm infant intestinal health, with a focus on its role in NEC. The most abundant HMOs, MOM microbiome and infant gut microbiome were compared between infants who developed NEC and infants who remained healthy. Since it has been reported that each of these components can influence the composition of the other, the datasets were also integrated to characterise the interaction between them. Microbiological techniques were applied to isolate and characterise preterm bacterial isolates. In particular, the potential of bacterial species relevant to the preterm gut were tested for their capacity to metabolise various HMOs. Finally, preterm derived intestinal organoids were used as a model to study the interaction between HMOs and the preterm infant gut. A novel co-culture system, the organoid-anaerobe co-culture system, was used and the response of the organoids investigated through RNA sequencing and ELISA for IL-8 production.

7.2 General discussion

Receipt of human breast milk and early life gut microbiome development are intrinsically linked, and both influence the risk of NEC in preterm infants. This study represents the largest analysis of HMOs and breast milk microbiome in NEC and the first to integrate HMO, MOM microbiome and preterm gut metagenomic data. DSLNT was found to be present in significantly lower concentrations in MOM fed to infants diagnosed with NEC, independently of NEC severity. Furthermore, lower DSLNT concentrations in MOM were associated with reduced transition into preterm gut community types typically observed in older infants and high in *Bifidobacterium* spp. relative abundance. The HMO profiles of infants with FIP were similar to controls, suggesting a clear clinical categorisation of FIP infants from NEC. Conversely, no differences in MOM microbiome were found in relation to NEC onset. No correlation between the HMOs and MOM microbiome and infant gut microbiome was observed, while a significant positive correlation between the relative abundance of *Acinetobacter* in MOM and in the infant gut was found, suggesting a potential direct seeding of the preterm gut through breast milk feeding. Growth curves of selected isolates on HMOs generally confirmed what has been reported in literature, with *B. infantis* and *B. bifidum* being the main (sub)species able to metabolise HMOs (Garrido et al., 2016, Katayama et al., 2004, LoCascio et al., 2010), followed by the other bifidobacteria which could grow on LNT and

LNnT (Asakuma et al., 2011, Thongaram et al., 2017, Lawson et al., 2019). Other infant gut commensals could not metabolise any of the HMOs tested and, in some cases, could not utilise free lactose. This is the first study to the best of my knowledge reporting the growth of *B. bifidum* on DSLNT, although how this might link to the association of higher DLSNT reducing NEC risk remains unknown. Finally, preterm intestinal organoids derived from NEC infants compared to non-NEC infants showed underlying differences in their gene expression profile and response to HMOs directly (i.e., without the presence of the microbiome), including alterations in *PI3* and *IGFBP3* expression.

HMOs are known for their beneficial effects on infants through multiple mechanisms, but the primary aspects relate to microbiome modulation by acting as prebiotics. Indeed, only a subset of the infant intestine colonisers can metabolise these complex sugars, and bifidobacteria are the main bacteria equipped with the genes to break down HMOs. Reduced DSLNT concentration in MOM received by infants developing NEC was found, and it correlated with a delayed progression into a gut microbial community typically expected in older infants and rich in bifidobacteria. *B. bifidum*, which has been associated to the establishment of a bifidobacteria rich microbial community in infants (Tannock et al., 2013), was the only bacterium tested in the growth curve experiments which could utilise DSLNT. *B. bifidum* is equipped with external glycosidases which release HMOs components from these saccharide structures, degradants which then can be internalised and metabolised by *B. bifidum* and which can also be utilised by other bifidobacteria, and potentially other species, to grow. All infants included in the metagenomic analysis, except for one NEC infant, received probiotic products which contained bifidobacteria species. However, infants who received low DSLNT showed impaired transition into a bifidobacteria rich microbial community. This might be correlated with the cross-feeding showed by *B. bifidum* species; however, this hypothesis needs to be tested.

No alteration in MOM microbiome was found despite the differences in HMO profile in NEC compared to healthy controls. These results suggest that the bacterial community in MOM does not have a role in NEC development, but it could provide potential colonisers of the preterm gut. Considering *B. bifidum* is the only bacterial species to date, which was found to metabolise DSLNT, it might not be surprising. Moreover, the environment within breast milk might not be optimal for bacterial growth, which would be essential to prevent HMOs metabolism in the mammary gland, so that they can be ingested by the infant and exert their multiple beneficial effects in the infant gut. While not directly associated with NEC risk, MOM microbiome might still play a role in direct microbial seeding and potentially influence preterm infant health. For instance, although *Acinetobacter* has not been previously linked to NEC, correlation of this

genus between MOM and infant gut microbiome suggests that there might be a direct seeding of the preterm gut through breast milk microbial community. Thus, investigating how the MOM microbiome is shaped might be of importance in preterm health modulation.

Multiple studies, including this thesis, have reported associations between preterm gut microbiome and HMO profile and NEC. Intestinal organoids represent a valid model to investigate the impact that the intestinal microorganisms and HMOs have on the preterm intestinal cells, helping to elucidate the role that these components might have in NEC development. Intestinal organoids are composed by all the major cell types found in the intestine, retain the specificity of the region they were derived from at the gene expression profile, and might also retain age specific characteristics. Comparison of adult and preterm derived intestinal organoids showed that differences in gene expression exist depending on life course stage. Differentially expressed genes included components of immune response and gut barrier function, suggesting that using preterm derived organoids might partially retain the intestinal prematurity of the preterm population. Given the patient population of interest for this thesis, preterm derived intestinal organoids were employed to investigate the effects of HMOs exposure on preterm intestine transcription. Differences between NEC and non-NEC organoids were also studied and DEGs included genes involved in cell differentiation and response to bacteria. NEC organoids showed a lower expression of *PI3* independently of exposure condition. *PI3* encodes for elafin, a protein which has anti-peptidase activity against neutrophil derived elastase and proteinase 3 as well as anti-bacterial properties against pathobionts, including Gram-negative bacteria. This alteration might be important in NEC onset, since intestinal damage might be caused by neutrophils and Gram-negative bacteria relative abundance is increased in NEC as reported in this thesis and in published studies (Wang et al., 2009, Warner et al., 2016).

The potential protective mechanism exerted by DSLNT in preventing NEC has not been elucidated yet. As discussed above, it might be linked to modulation of the gut microbial community composition, however the beneficial effects might be at multiple levels. New studies suggest the role of DSLNT in protecting the human intestine from damage provoked by mast cells activation (Zhang et al., 2021, Lian et al., 2022). Preterm ileal resected tissue pre-treated with DSLNT showed reduced MC chymase caused damage, and MC accumulation and intestinal damage were also reduced in a rat NEC model upon DSLNT supplementation (Zhang et al., 2021). DSLNT could also directly alter the gene expression profile of Caco-2 cells by stimulating cell cycle and proliferation (Lian et al., 2022). When comparing NEC and non-NEC organoids, some genes involved in cell apoptosis and proliferation were found to be specific to DSLNT exposure. The cell proliferation mediated by DSLNT in preterm intestine might also

be correlated with modulation of *IGFBP3* gene expression, which was found to be significantly different in the NEC versus non-NEC comparison upon exposure to DSLNT, 2'FL and the mix of HMOs. Despite no definitive conclusions on the effect of DSLNT exposure on preterm intestinal cells, published results on mast cells and results discussed in this thesis, suggest that DSLNT might have NEC-protective effects by directly interacting with the gut epithelium.

Findings reported in this thesis confirmed that DSLNT concentration in MOM is associated with NEC protection and suggested potential mechanisms for this beneficial effect which does not seem to be related to the MOM microbiome. These mechanisms might include modulation of the infant gut microbial community establishment through *B. bifidum* cross-feeding of other bifidobacteria, and direct modulation of the preterm intestinal epithelium gene expression. Moreover, underlying mechanisms in gene expression profile might be present in the preterm intestine of NEC and non-NEC infants, and these differences might be important to understanding the mechanisms to leading NEC onset.

7.3 Future work

The results presented in this thesis can lay the foundation for future research to expand on the findings and address the key limitations listed in each chapter's discussion.

Breast milk samples analysed in this thesis were salvaged directly from the infant feeding tube after the feed, which can take up to 24 hours. This is necessary to ensure the infant receives optimal nutrition, but this might cause an alteration to the HMO profile that was measured. The stability of various breast milk components at room temperature has been investigated, however not much is known about HMOs. Future studies aiming at understanding the impact on HMOs of MOM handling and storage in the NICU are needed. Furthermore, according to recently published data, the concentration of HMOs, including DSLNT, is comparatively stable throughout time, but this needs to be confirmed in long-term preterm cohorts. To address these questions a new set of MOM samples were sent to our collaborator for HMO analysis, however, the data is not available yet. The samples sent included MOM from nine mothers either frozen as fresh or after having been collected from the infant's feeding tube, and longitudinal fresh-frozen MOM samples from ten mothers to investigate the composition changes in HMOs over the first 50 days of the infant's life.

Reduced DSLNT concentration was found to be specific to NEC onset, and not involved in FIP disease. However, the groups were not properly matched, and the results might be an artifact of the clinical variable differences. To validate these results and understand the potential use of DSLNT measurement to predict NEC onset, it would be useful to study its specificity to NEC

disease by investigating the HMO profile in different preterm diseases, including FIP and LOS, in a dedicated study. Notably, FIP tends to occur earlier in life compared to NEC making the establishment of cohorts of matched infants difficult.

Growth curves experiments performed as part of this thesis will help to generate hypotheses on the link between HMOs and the preterm gut microbiome. However, further studies should aim at testing multiple preterm isolates for the various bacterial species. Indeed, only one isolate was included for the non-bifidobacteria gut commensals. The results shown as part of this thesis are in concordance with published studies, suggesting that non-bifidobacteria species tend to not be equipped with genes for HMO metabolism and high strain variability has been reported. Nonetheless, further investigating HMO utilisation capability of isolates obtained from preterm samples would be of interest.

Over the past 10 years multiple studies have aimed at identifying bacterial genes involved in HMO utilisation. Novel genes have yet to be identified since some bifidobacteria strains not carrying any of the known HMO utilisation genes have been shown to be able to grow on these complex sugars. Bacteria tested in the HMO growth curve presented and additional isolates will be sent for whole-genome sequencing, and bacteria which could grow on HMOs will be subjected to RNA-sequencing to help identify the genes involved in HMOs degradation. Metabolomics on the growth supernatant will also be performed to identify metabolites produced as a consequence of HMO metabolism.

Coupled with published studies, the results of this thesis suggest that a potential mechanism behind DSLNT protection might be through *B. bifidum* and its postulated cross-feeding of HMO degradants to other bifidobacteria. Experiments aimed at understanding the evolution of a complex gut microbial community composed by *B. bifidum,* other *Bifidobacterium* spp. and non-bifidobacteria gut commensals, such as *E. coli* or *Klebsiella* spp., in the presence of HMOs is needed. Conditioned media produced by *B. bifidum* and *B. infantis* metabolism on HMOs will be tested to evaluate its effect on other bacterial isolates. For instance, it will be investigated if the metabolites produced by these two bifidobacteria in the presence of HMOs are able to block or enhance the growth of other bifidobacteria and non-bifidobacteria species. Growth in presence of HMOs of a mock microbial community composed by various gut commensals will also be evaluated by monitoring the microbial growth using quantitative PCR and primers specific to the species composing the mock community. Such experiments could inform if there are mechanisms through which *B. bifidum* could preferentially cross-feed HMO degradants to other bifidobacteria, or if any bacterium present in the community would benefit from such release of HMO components, including potential pathogenic bacteria.

Further optimisation of the preterm intestinal organoids model to study NEC will be needed. Optimisation will include testing the pre-exposure of organoids to HMOs prior to the inflammatory challenge. Organoid monolayers could be exposed to HMOs prior or during differentiation, or after

the differentiation has happened. Alternative pro-inflammatory stimuli might also be tested, such as flagellin or TNFα. Future studies will aim to investigate the response of preterm organoid to HMOs, bacteria and to mix of these two. Preterm organoids will represent a valuable model to understand the interaction between HMOs, microbiome and host.

Despite these limitations and proposed future work, this thesis involved the largest cohort to date investigating the relationship between HMO composition and NEC development and included one of the most extensive longitudinal stool metagenomic analyses of preterm infants. The role of MOM microbiome in NEC was investigated for the first time, and it was also the first study integrating HMO composition, infants stool metagenome and MOM microbiome data in a preterm cohort. Finally, it was one of the first studies comparing the transcriptome profile of preterm and adult intestinal organoids, as well as NEC and non-NEC intestinal organoids. The results presented in this study will help advance the understanding of the mechanisms leading to NEC onset and stimulate the generation of new hypotheses to test.

Taken together, the results show that preterm tissue derived ileal organoids behave differently to adult ileal organoids when exposed to various stimuli. Reproducibility between experiments was observed, with genes being differentially expressed in a consistent manner in two separate datasets. Differences at the transcriptome level were also found between NEC and non-NEC organoids, which responded differently to HMOs exposures. This highlights the importance of using preterm organoids under physiologically relevant oxygen conditions to model host-microbiome interactions and to study the underlying mechanisms involved in the pathogenesis of NEC.

References

1977. WHO: recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths. Modifications recommended by FIGO as amended October 14, 1976. *Acta Obstet Gynecol Scand,* 56**,** 247-53.

2ND, Z. K. C., BHASIN, N., RIENZI, S. C. D., RAJAN, A., DEANS-FIELDER, K., SWAMINATHAN, G., KAMYABI, N., ZENG, X.-L., DODDAPANENI, H., MENON, V. K., CHAKRAVARTI, D., ESTRELLA, C., YU, X., PATIL, K., PETROSINO, J. F., FLEET, J. C., VERZI, M. P., CHRISTAKOS, S., HELMRATH, M. A., ARIMURA, S., DEPINHO, R. A., BRITTON, R. A., MARESSO, A. W., GRANDE-ALLEN, K. J., BLUTT, S. E., CRAWFORD, S. E., ESTES, M. K., RAMANI, S. & SHROYER, N. F. 2021. Drivers of transcriptional variance in human intestinal epithelial organoids. *Physiological Genomics,* 53**,** 486-508.

ABDULKADIR, B., NELSON, A., SKEATH, T., MARRS, E. C., PERRY, J. D., CUMMINGS, S. P., EMBLETON, N. D., BERRINGTON, J. E. & STEWART, C. J. 2016. Stool bacterial load in preterm infants with necrotising enterocolitis. *Early Hum Dev,* 95**,** 1-2.

ABRAHAMSSON, T. R., JAKOBSSON, H. E., ANDERSSON, A. F., BJÖRKSTÉN, B., ENGSTRAND, L. & JENMALM, M. C. 2012. Low diversity of the gut microbiota in infants with atopic eczema. *J Allergy Clin Immunol,* 129**,** 434-40, 440.e1-2.

ALBERT-BAYO, M., PARACUELLOS, I., GONZÁLEZ-CASTRO, A. M., RODRÍGUEZ-URRUTIA, A., RODRÍGUEZ-LAGUNAS, M. J., ALONSO-COTONER, C., SANTOS, J. & VICARIO, M. 2019. Intestinal Mucosal Mast Cells: Key Modulators of Barrier Function and Homeostasis. *Cells* [Online], 8.

ALCON-GINER, C., DALBY, M. J., CAIM, S., KETSKEMETY, J., SHAW, A., SIM, K., LAWSON, M. A. E., KIU, R., LECLAIRE, C., CHALKLEN, L., KUJAWSKA, M., MITRA, S., FARDUS-REID, F., BELTEKI, G., MCCOLL, K., SWANN, J. R., KROLL, J. S., CLARKE, P. & HALL, L. J. 2020. Microbiota Supplementation with Bifidobacterium and Lactobacillus Modifies the Preterm Infant Gut Microbiota and Metabolome: An Observational Study. *Cell Rep Med,* 1**,** 100077.

ALESSANDRI, G., OSSIPRANDI, M. C., MACSHARRY, J., VAN SINDEREN, D. & VENTURA, M. 2019. Bifidobacterial Dialogue With Its Human Host and Consequent Modulation of the Immune System. *Frontiers in Immunology,* 10.

ALOU, M. T., NAUD, S., KHELAIFIA, S., BONNET, M., LAGIER, J.-C. & RAOULT, D. 2020. State of the Art in the Culture of the Human Microbiota: New Interests and Strategies. *Clinical Microbiology Reviews,* 34**,** e00129-19.

ANDREAS, N. J., AL-KHALIDI, A., JAITEH, M., CLARKE, E., HYDE, M. J., MODI, N., HOLMES, E., KAMPMANN, B. & MEHRING LE DOARE, K. 2016. Role of human milk oligosaccharides in Group B Streptococcus colonisation. *Clin Transl Immunology,* 5**,** e99.

APPERT, O., GARCIA, A. R., FREI, R., RODUIT, C., CONSTANCIAS, F., NEUZIL-BUNESOVA, V., FERSTL, R., ZHANG, J., AKDIS, C., LAUENER, R., LACROIX, C. & SCHWAB, C. 2020. Initial butyrate producers during infant gut microbiota development are endospore formers. *Environmental Microbiology,* 22**,** 3909-3921.

ASAKUMA, S., HATAKEYAMA, E., URASHIMA, T., YOSHIDA, E., KATAYAMA, T., YAMAMOTO, K., KUMAGAI, H., ASHIDA, H., HIROSE, J. & KITAOKA, M. 2011. Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria. *J Biol Chem,* 286**,** 34583-92.

ASBURY, M. R., BUTCHER, J., COPELAND, J. K., UNGER, S., BANDO, N., COMELLI, E. M., FORTE, V., KISS, A., LEMAY-NEDJELSKI, L., SHERMAN, P. M., STINTZI, A., TOMLINSON, C., WANG, P. W. & O'CONNOR, D. L. 2020. Mothers of Preterm Infants Have Individualized Breast Milk Microbiota that Changes Temporally Based on Maternal Characteristics. *Cell Host & Microbe*.

ASHIDA, H., MIYAKE, A., KIYOHARA, M., WADA, J., YOSHIDA, E., KUMAGAI, H., KATAYAMA, T. & YAMAMOTO, K. 2009. Two distinct alpha-L-fucosidases from Bifidobacterium bifidum are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates. *Glycobiology,* 19**,** 1010-7.

AUCHTUNG, J. M., PREISNER, E. C., COLLINS, J., LERMA, A. I. & BRITTON, R. A. 2020. Identification of Simplified Microbial Communities That Inhibit Clostridioides difficile Infection through Dilution/Extinction. *mSphere,* 5.

AUSTIN, S., DE CASTRO, C. A., SPRENGER, N., BINIA, A., AFFOLTER, M., GARCIA-RODENAS, C. L., BEAUPORT, L., TOLSA, J. F. & FISCHER FUMEAUX, C. J. 2019. Human Milk Oligosaccharides in the Milk of Mothers Delivering Term versus Preterm Infants. *Nutrients,* 11.

AUTRAN, C. A., KELLMAN, B. P., KIM, J. H., ASZTALOS, E., BLOOD, A. B., SPENCE, E. C. H., PATEL, A. L., HOU, J., LEWIS, N. E. & BODE, L. 2018. Human milk oligosaccharide composition predicts risk of necrotising enterocolitis in preterm infants. *Gut,* 67**,** 1064-1070.

AUTRAN, C. A., SCHOTERMAN, M. H., JANTSCHER-KRENN, E., KAMERLING, J. P. & BODE, L. 2016. Sialylated galacto-oligosaccharides and 2'-fucosyllactose reduce necrotising enterocolitis in neonatal rats. *Br J Nutr,* 116**,** 294-9.

AZAD, M. B., ROBERTSON, B., ATAKORA, F., BECKER, A. B., SUBBARAO, P., MORAES, T. J., MANDHANE, P. J., TURVEY, S. E., LEFEBVRE, D. L., SEARS, M. R. & BODE, L. 2018. Human Milk Oligosaccharide Concentrations Are Associated with Multiple Fixed and Modifiable Maternal Characteristics, Environmental Factors, and Feeding Practices. *J Nutr,* 148**,** 1733-1742.

BALLARD, O. & MORROW, A. L. 2013. Human milk composition: nutrients and bioactive factors. *Pediatric clinics of North America,* 60**,** 49-74.

BARANGER, K., ZANI, M. L., CHANDENIER, J., DALLET-CHOISY, S. & MOREAU, T. 2008. The antibacterial and antifungal properties of trappin-2 (pre-elafin) do not depend on its protease inhibitory function. *Febs j,* 275**,** 2008-20.

BAREGAMIAN, N., SONG, J., JESCHKE, M. G., EVERS, B. M. & CHUNG, D. H. 2006. IGF-1 protects intestinal epithelial cells from oxidative stress-induced apoptosis. *J Surg Res,* 136**,** 31-7.

BARRETT, E., KERR, C., MURPHY, K., O'SULLIVAN, O., RYAN, C. A., DEMPSEY, E. M., MURPHY, B. P., O'TOOLE, P. W., COTTER, P. D., FITZGERALD, G. F., ROSS, R. P. & STANTON, C. 2013. The individual-specific and diverse nature of the preterm infant microbiota. *Arch Dis Child Fetal Neonatal Ed,* 98**,** F334-40.

BECK, L. C., MASI, A. C., YOUNG, G. R., VATANEN, T., LAMB, C. A., SMITH, R., COXHEAD, J., BUTLER, A., MARSLAND, B. J., EMBLETON, N. D., BERRINGTON, J. E. & STEWART, C. J. 2022. Strain-specific impacts of probiotics are a significant driver of gut microbiome development in very preterm infants. *Nature Microbiology,* 7**,** 1525-1535.

BECKLOFF, N., STARKENBURG, S., FREITAS, T. & CHAIN, P. 2012. Bacterial genome annotation. *Methods Mol Biol,* 881**,** 471-503.

BELL, M. J., TERNBERG, J. L., FEIGIN, R. D., KEATING, J. P., MARSHALL, R., BARTON, L. & BROTHERTON, T. 1978. Neonatal necrotizing enterocolitis. Therapeutic decisions based upon clinical staging. *Ann Surg,* 187**,** 1-7.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological),* 57**,** 289-300.

BERGER, B., PORTA, N., FOATA, F., GRATHWOHL, D., DELLEY, M., MOINE, D., CHARPAGNE, A., SIEGWALD, L., DESCOMBES, P., ALLIET, P., PUCCIO, G.,

STEENHOUT, P., MERCENIER, A. & SPRENGER, N. 2020. Linking Human Milk Oligosaccharides, Infant Fecal Community Types, and Later Risk To Require Antibiotics. *mBio,* 11.

BERRINGTON, J. & EMBLETON, N. D. 2021. Discriminating necrotising enterocolitis and focal intestinal perforation. *Archives of Disease in Childhood - Fetal and Neonatal Edition,* fetalneonatal-2020-321429.

BERRINGTON, J. E., STEWART, C. J., EMBLETON, N. D. & CUMMINGS, S. P. 2013. Gut microbiota in preterm infants: assessment and relevance to health and disease. *Archives of Disease in Childhood - Fetal and Neonatal Edition,* 98**,** F286-F290.

BIAGI, E., ACETI, A., QUERCIA, S., BEGHETTI, I., RAMPELLI, S., TURRONI, S., SOVERINI, M., ZAMBRINI, A. V., FALDELLA, G., CANDELA, M., CORVAGLIA, L. & BRIGIDI, P. 2018. Microbial Community Dynamics in Mother's Milk and Infant's Mouth and Gut in Moderately Preterm Infants. *Front Microbiol,* 9**,** 2512.

BIESTERVELD, B. E., KOEHLER, S. M., HEINZERLING, N. P., RENTEA, R. M., FREDRICH, K., WELAK, S. R. & GOURLAY, D. M. 2015. Intestinal alkaline phosphatase to treat necrotizing enterocolitis. *J Surg Res,* 196**,** 235-40.

BLUTT, S. E., CRAWFORD, S. E., RAMANI, S., ZOU, W. Y. & ESTES, M. K. 2018. Engineered Human Gastrointestinal Cultures to Study the Microbiome and Infectious Diseases. *Cell Mol Gastroenterol Hepatol,* 5**,** 241-251.

BODE, L. 2012. Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology,* 22**,** 1147-62.

BODE, L., KUHN, L., KIM, H. Y., HSIAO, L., NISSAN, C., SINKALA, M., KANKASA, C., MWIYA, M., THEA, D. M. & ALDROVANDI, G. M. 2012. Human milk oligosaccharide concentration and risk of postnatal transmission of HIV through breastfeeding. *Am J Clin Nutr,* 96**,** 831-9.

BODE, L., KUNZ, C., MUHLY-REINHOLZ, M., MAYER, K., SEEGER, W. & RUDLOFF, S. 2004. Inhibition of monocyte, lymphocyte, and neutrophil adhesion to endothelial cells by human milk oligosaccharides. *Thromb Haemost,* 92**,** 1402-10.

BODE, L., RAMAN, A. S., MURCH, S. H., ROLLINS, N. C. & GORDON, J. I. 2020. Understanding the mother-breastmilk-infant "triad". *Science,* 367**,** 1070-1072.

BOGAERT, D., VAN BEVEREN, G. J., DE KOFF, E. M., LUSARRETA PARGA, P., BALCAZAR LOPEZ, C. E., KOPPENSTEINER, L., CLERC, M., HASRAT, R., ARP, K., CHU, M. L. J. N., DE GROOT, P. C. M., SANDERS, E. A. M., VAN HOUTEN, M. A. & DE STEENHUIJSEN PITERS, W. A. A. 2023. Mother-to-infant microbiota

transmission and infant microbiota development across multiple body sites. *Cell Host & Microbe,* 31**,** 447-460.e6.

BOIX-AMOROS, A., COLLADO, M. C. & MIRA, A. 2016. Relationship between Milk Microbiota, Bacterial Load, Macronutrients, and Human Cells during Lactation. *Front Microbiol,* 7**,** 492.

BOIX-AMORÓS, A., MARTINEZ-COSTA, C., QUEROL, A., COLLADO, M. C. & MIRA, A. 2017. Multiple Approaches Detect the Presence of Fungi in Human Breastmilk Samples from Healthy Mothers. *Sci Rep,* 7**,** 13016.

BOREWICZ, K., GU, F., SACCENTI, E., HECHLER, C., BEIJERS, R., DE WEERTH, C., VAN LEEUWEN, S. S., SCHOLS, H. A. & SMIDT, H. 2020. The association between breastmilk oligosaccharides and faecal microbiota in healthy breastfed infants at two, six, and twelve weeks of age. *Sci Rep,* 10**,** 4270.

BRIGHAM, C., CAUGHLAN, R., GALLEGOS, R., DALLAS, M. B., GODOY, V. G. & MALAMY, M. H. 2009. Sialic acid (N-acetyl neuraminic acid) utilization by Bacteroides fragilis requires a novel N-acetyl mannosamine epimerase. *J Bacteriol,* 191**,** 3629-38.

BROOKS, B., FIREK, B. A., MILLER, C. S., SHARON, I., THOMAS, B. C., BAKER, R., MOROWITZ, M. J. & BANFIELD, J. F. 2014. Microbes in the neonatal intensive care unit resemble those found in the gut of premature infants. *Microbiome,* 2**,** 1.

BROOKS, B., OLM, M. R., FIREK, B. A., BAKER, R., THOMAS, B. C., MOROWITZ, M. J. & BANFIELD, J. F. 2017. Strain-resolved analysis of hospital rooms and infants reveals overlap between the human and room microbiome. *Nature Communications,* 8**,** 1814.

BUNESOVA, V., LACROIX, C. & SCHWAB, C. 2016. Fucosyllactose and L-fucose utilization of infant Bifidobacterium longum and Bifidobacterium kashiwanohense. *BMC Microbiology,* 16**,** 248.

BUONPANE, C., ARES, G., YUAN, C., SCHLEGEL, C., LIEBE, H. & HUNTER, C. J. 2020. Experimental Modeling of Necrotizing Enterocolitis in Human Infant Intestinal Enteroids. *Journal of Investigative Surgery***,** 1-8.

CAMACHO-GONZALEZ, A., SPEARMAN, P. W. & STOLL, B. J. 2013. Neonatal Infectious Diseases: Evaluation of Neonatal Sepsis. *Pediatric Clinics of North America,* 60**,** 367-389.

CAPORASO, J. G., LAUBER, C. L., WALTERS, W. A., BERG-LYONS, D., HUNTLEY, J., FIERER, N., OWENS, S. M., BETLEY, J., FRASER, L., BAUER, M., GORMLEY, N., GILBERT, J. A., SMITH, G. & KNIGHT, R. 2012. Ultra-high-throughput microbial

community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal,* 6**,** 1621-1624.

CHAN, K. L., WONG, K. F. & LUK, J. M. 2009. Role of LPS/CD14/TLR4-mediated inflammation in necrotizing enterocolitis: pathogenesis and therapeutic implications. *World J Gastroenterol,* 15**,** 4745-52.

CHAN, K. Y., LEUNG, K. T., TAM, Y. H., LAM, H. S., CHEUNG, H. M., MA, T. P., LEE, K. H., TO, K. F., LI, K. & NG, P. C. 2014. Genome-wide expression profiles of necrotizing enterocolitis versus spontaneous intestinal perforation in human intestinal tissues: dysregulation of functional pathways. *Ann Surg,* 260**,** 1128-37.

CHAPMAN, J. A. & STEWART, C. J. 2023. Methodological challenges in neonatal microbiome research. *Gut Microbes,* 15**,** 2183687.

CHEN, B. & LIU, G. 2018. WWC3 inhibits intimal proliferation following vascular injury via the Hippo signaling pathway. *Mol Med Rep,* 17**,** 5175-5183.

CHENG, L., KONG, C., WANG, W., GROENEVELD, A., NAUTA, A., GROVES, M. R., KIEWIET, M. B. G. & DE VOS, P. 2021. The Human Milk Oligosaccharides 3-FL, LNnT, and LDFT Attenuate TNF-alpha Induced Inflammation in Fetal Intestinal Epithelial Cells In Vitro Through Shedding or Interacting with TNF Receptor 1. *Mol Nutr Food Res***,** e2000425.

CHIA, L. W., MANK, M., BLIJENBERG, B., AALVINK, S., BONGERS, R. S., STAHL, B., KNOL, J. & BELZER, C. 2020. Bacteroides thetaiotaomicron Fosters the Growth of Butyrate-Producing Anaerostipes caccae in the Presence of Lactose and Total Human Milk Carbohydrates. *Microorganisms,* 8.

CHICHLOWSKI, M., DE LARTIGUE, G., GERMAN, J. B., RAYBOULD, H. E. & MILLS, D. A. 2012. Bifidobacteria isolated from infants and cultured on human milk oligosaccharides affect intestinal epithelial function. *J Pediatr Gastroenterol Nutr,* 55**,** 321-7.

CHONG, J., LIU, P., ZHOU, G. & XIA, J. 2020. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. *Nat Protoc,* 15**,** 799-821.

CHRISTENSEN, R. D., YODER, B. A., BAER, V. L., SNOW, G. L. & BUTLER, A. 2015. Early-Onset Neutropenia in Small-for-Gestational-Age Infants. *Pediatrics,* 136**,** e1259-67.

CILIEBORG, M. S., BERING, S. B., OSTERGAARD, M. V., JENSEN, M. L., KRYCH, L., NEWBURG, D. S. & SANGILD, P. T. 2016. Minimal short-term effect of dietary 2'-fucosyllactose on bacterial colonisation, intestinal function and necrotising enterocolitis in preterm pigs. *Br J Nutr,* 116**,** 834-41.

CLARK, D. P. & PAZDERNIK, N. J. 2013. Chapter 7 - Cloning Genes for Analysis. *In:* CLARK, D. P. & PAZDERNIK, N. J. (eds.) *Molecular Biology (Second Edition).* Boston: Academic Press.

CLARK, J. A., DOELLE, S. M., HALPERN, M. D., SAUNDERS, T. A., HOLUBEC, H., DVORAK, K., BOITANO, S. A. & DVORAK, B. 2006. Intestinal barrier failure during experimental necrotizing enterocolitis: protective effect of EGF treatment. *American Journal of Physiology-Gastrointestinal and Liver Physiology,* 291**,** G938-G949.

CRISS, Z. K., 2ND, BHASIN, N., DI RIENZI, S. C., RAJAN, A., DEANS-FIELDER, K., SWAMINATHAN, G., KAMYABI, N., ZENG, X. L., DODDAPANENI, H., MENON, V. K., CHAKRAVARTI, D., ESTRELLA, C., YU, X., PATIL, K., PETROSINO, J. F., FLEET, J. C., VERZI, M. P., CHRISTAKOS, S., HELMRATH, M. A., ARIMURA, S., DEPINHO, R. A., BRITTON, R. A., MARESSO, A. W., GRANDE-ALLEN, K. J., BLUTT, S. E., CRAWFORD, S. E., ESTES, M. K., RAMANI, S. & SHROYER, N. F. 2021. Drivers of transcriptional variance in human intestinal epithelial organoids. *Physiol Genomics,* 53**,** 486-508.

DE LEOZ, M. L., GAERLAN, S. C., STRUM, J. S., DIMAPASOC, L. M., MIRMIRAN, M., TANCREDI, D. J., SMILOWITZ, J. T., KALANETRA, K. M., MILLS, D. A., GERMAN, J. B., LEBRILLA, C. B. & UNDERWOOD, M. A. 2012. Lacto-N-tetraose, fucosylation, and secretor status are highly variable in human milk oligosaccharides from women delivering preterm. *J Proteome Res,* 11**,** 4662-72.

DEMERS-MATHIEU, V. 2022. The immature intestinal epithelial cells in preterm infants play a role in the necrotizing enterocolitis pathogenesis: A review. *Health Sciences Review,* 4**,** 100033.

DENG, Y., MISSELWITZ, B., DAI, N. & FOX, M. 2015. Lactose Intolerance in Adults: Biological Mechanism and Dietary Management. *Nutrients* [Online], 7.

DEVISSCHER, L., HINDRYCKX, P., OLIEVIER, K., PEETERS, H., DE VOS, M. & LAUKENS, D. 2011. Inverse correlation between metallothioneins and hypoxia-inducible factor 1 alpha in colonocytes and experimental colitis. *Biochemical and Biophysical Research Communications,* 416**,** 307-312.

DHARIWAL, A., CHONG, J., HABIB, S., KING, I. L., AGELLON, L. B. & XIA, J. 2017. MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res,* 45**,** W180-W188.

DIAKITE, A., DUBOURG, G., DIONE, N., AFOUDA, P., BELLALI, S., NGOM, II, VALLES, C., MILLION, M., LEVASSEUR, A., CADORET, F., LAGIER, J. C. &

RAOULT, D. 2019. Extensive culturomics of 8 healthy samples enhances metagenomics efficiency. *PLoS One,* 14**,** e0223543.

DIAKITE, A., DUBOURG, G., DIONE, N., AFOUDA, P., BELLALI, S., NGOM, I. I., VALLES, C., TALL, M. L., LAGIER, J.-C. & RAOULT, D. 2020. Optimization and standardization of the culturomics technique for human microbiome exploration. *Scientific Reports,* 10**,** 9674.

DIEZ, S., RENNER, M., BAHLINGER, V., HARTMANN, A., BESENDÖRFER, M. & MÜLLER, H. 2022. Increased expression of OLFM4 and lysozyme during necrotizing enterocolitis in neonates: an observational research study. *BMC Pediatr,* 22**,** 192.

DUAR, R. M., CASABURI, G., MITCHELL, R. D., SCOFIELD, L. N. C., ORTEGA RAMIREZ, C. A., BARILE, D., HENRICK, B. M. & FRESE, S. A. 2020. Comparative Genome Analysis of Bifidobacterium longum subsp. infantis Strains Reveals Variation in Human Milk Oligosaccharide Utilization Genes among Commercial Probiotics. *Nutrients,* 12**,** 3247.

DUNN, O. J. 1961. Multiple Comparisons Among Means. *Journal of the American Statistical Association,* 56**,** 52-64.

ECKBURG, P. B., BIK, E. M., BERNSTEIN, C. N., PURDOM, E., DETHLEFSEN, L., SARGENT, M., GILL, S. R., NELSON, K. E. & RELMAN, D. A. 2005. Diversity of the human intestinal microbial flora. *Science,* 308**,** 1635-8.

EDGAR, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics,* 26**,** 2460-1.

EDGAR, R. C. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods,* 10**,** 996-8.

EDGAR, R. C., HAAS, B. J., CLEMENTE, J. C., QUINCE, C. & KNIGHT, R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics,* 27**,** 2194-200.

EISEMAN, B., SILEN, W., BASCOM, G. S. & KAUVAR, A. J. 1958. Fecal enema as an adjunct in the treatment of pseudomembranous enterocolitis. *Surgery,* 44**,** 854-9.

EIWEGGER, T., STAHL, B., SCHMITT, J., BOEHM, G., GERSTMAYR, M., PICHLER, J., DEHLINK, E., LOIBICHLER, C., URBANEK, R. & SZÉPFALUSI, Z. 2004. Human Milk–Derived Oligosaccharides and Plant-Derived Oligosaccharides Stimulate Cytokine Production of Cord Blood T-Cells In Vitro. *Pediatric Research,* 56**,** 536-540.

EMBLETON, N., BERRINGTON, J., CUMMINGS, S., DORLING, J., EWER, A., FRAU, A., JUSZCZAK, E., KIRBY, J., LAMB, C., LANYON, C., LETT, L., MCGUIRE, W., PROBERT, C., RUSHTON, S., SHIRLEY, M., STEWART, C. & YOUNG, G. R. 2021. Efficacy and Mechanism Evaluation. *Lactoferrin impact on gut microbiota in preterm*

infants with late-onset sepsis or necrotising enterocolitis: the MAGPIE mechanisms of action study. Southampton (UK): NIHR Journals Library

EMBLETON, N. D., BERRINGTON, J. E., DORLING, J., EWER, A. K., JUSZCZAK, E., KIRBY, J. A., LAMB, C. A., LANYON, C. V., MCGUIRE, W., PROBERT, C. S., RUSHTON, S. P., SHIRLEY, M. D., STEWART, C. J. & CUMMINGS, S. P. 2017. Mechanisms Affecting the Gut of Preterm Infants in Enteral Feeding Trials. *Front Nutr,* 4**,** 14.

ENGFER, M. B., STAHL, B., FINKE, B., SAWATZKI, G. & DANIEL, H. 2000. Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract. *The American Journal of Clinical Nutrition,* 71**,** 1589-1596.

FALCON, S. & GENTLEMAN, R. 2007. Using GOstats to test gene lists for GO term association. *Bioinformatics,* 23**,** 257-8.

FERRETTI, P., PASOLLI, E., TETT, A., ASNICAR, F., GORFER, V., FEDI, S., ARMANINI, F., TRUONG, D. T., MANARA, S., ZOLFO, M., BEGHINI, F., BERTORELLI, R., DE SANCTIS, V., BARILETTI, I., CANTO, R., CLEMENTI, R., COLOGNA, M., CRIFÒ, T., CUSUMANO, G., GOTTARDI, S., INNAMORATI, C., MASÈ, C., POSTAI, D., SAVOI, D., DURANTI, S., LUGLI, G. A., MANCABELLI, L., TURRONI, F., FERRARIO, C., MILANI, C., MANGIFESTA, M., ANZALONE, R., VIAPPIANI, A., YASSOUR, M., VLAMAKIS, H., XAVIER, R., COLLADO, C. M., KOREN, O., TATEO, S., SOFFIATI, M., PEDROTTI, A., VENTURA, M., HUTTENHOWER, C., BORK, P. & SEGATA, N. 2018. Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome. *Cell Host & Microbe,* 24**,** 133-145.e5.

FIGUEROA-LOZANO, S., AKKERMAN, R., BEUKEMA, M., VAN LEEUWEN, S. S., DIJKHUIZEN, L. & DE VOS, P. 2021. 2′-Fucosyllactose impacts the expression of mucus-related genes in goblet cells and maintains barrier function of gut epithelial cells. *Journal of Functional Foods,* 85**,** 104630.

FLINT, H. J., SCOTT, K. P., DUNCAN, S. H., LOUIS, P. & FORANO, E. 2012. Microbial degradation of complex carbohydrates in the gut. *Gut microbes,* 3**,** 289-306.

FOFANOVA, T. Y., STEWART, C., AUCHTUNG, J. M., WILSON, R. L., BRITTON, R. A., GRANDE-ALLEN, K. J., ESTES, M. K. & PETROSINO, J. F. 2019. A novel human enteroid-anaerobe co-culture system to study microbial-host interaction under physiological hypoxia. *bioRxiv***,** 555755.

FORSTER, S. C., KUMAR, N., ANONYE, B. O., ALMEIDA, A., VICIANI, E., STARES, M. D., DUNN, M., MKANDAWIRE, T. T., ZHU, A., SHAO, Y., PIKE, L. J., LOUIE, T., BROWNE, H. P., MITCHELL, A. L., NEVILLE, B. A., FINN, R. D. & LAWLEY, T. D. 2019. A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nature Biotechnology,* 37**,** 186-192.

FRESE, S. A., HUTTON, A. A., CONTRERAS, L. N., SHAW, C. A., PALUMBO, M. C., CASABURI, G., XU, G., DAVIS, J. C. C., LEBRILLA, C. B., HENRICK, B. M., FREEMAN, S. L., BARILE, D., GERMAN, J. B., MILLS, D. A., SMILOWITZ, J. T., UNDERWOOD, M. A. & KRAJMALNIK-BROWN, R. 2017. Persistence of Supplemented Bifidobacterium longum subsp. infantis EVC001 in Breastfed Infants. *mSphere,* 2**,** e00501-17.

FUJIMURA, K. E., SITARIK, A. R., HAVSTAD, S., LIN, D. L., LEVAN, S., FADROSH, D., PANZER, A. R., LAMERE, B., RACKAITYTE, E., LUKACS, N. W., WEGIENKA, G., BOUSHEY, H. A., OWNBY, D. R., ZORATTI, E. M., LEVIN, A. M., JOHNSON, C. C. & LYNCH, S. V. 2016. Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nature Medicine,* 22**,** 1187-1191.

GARRIDO, D., RUIZ-MOYANO, S., KIRMIZ, N., DAVIS, J. C., TOTTEN, S. M., LEMAY, D. G., UGALDE, J. A., GERMAN, J. B., LEBRILLA, C. B. & MILLS, D. A. 2016. A novel gene cluster allows preferential utilization of fucosylated milk oligosaccharides in Bifidobacterium longum subsp. longum SC596. *Scientific Reports,* 6**,** 35045.

GARRIDO, D., RUIZ-MOYANO, S. & MILLS, D. A. 2012. Release and utilization of N-acetyl-d-glucosamine from human milk oligosaccharides by Bifidobacterium longum subsp. infantis. *Anaerobe,* 18**,** 430-435.

GAUHE, A., GYÖRGY, P., HOOVER, J. R. E., KUHN, R., ROSE, C. S., RUELIUS, H. W. & ZILLIKEN, F. 1954. Bifidus factor. IV. Preparations obtained from human milk. *Archives of Biochemistry and Biophysics,* 48**,** 214-224.

GIDREWICZ, D. A. & FENTON, T. R. 2014. A systematic review and meta-analysis of the nutrient content of preterm and term breast milk. *BMC Pediatr,* 14**,** 216.

GILA-DIAZ, A., ARRIBAS, S. M., ALGARA, A., MARTIN-CABREJAS, M. A., LOPEZ DE PABLO, A. L., SAENZ DE PIPAON, M. & RAMIRO-CORTIJO, D. 2019. A Review of Bioactive Factors in Human Breastmilk: A Focus on Prematurity. *Nutrients,* 11.

GOEHRING, K. C., MARRIAGE, B. J., OLIVER, J. S., WILDER, J. A., BARRETT, E. G. & BUCK, R. H. 2016. Similar to Those Who Are Breastfed, Infants Fed a Formula Containing 2′-Fucosyllactose Have Lower Inflammatory Cytokines in a Randomized Controlled Trial. *The Journal of Nutrition,* 146**,** 2559-2566.

GOOD, M., SODHI, C. P., YAMAGUCHI, Y., JIA, H., LU, P., FULTON, W. B., MARTIN, L. Y., PRINDLE, T., NINO, D. F., ZHOU, Q., MA, C., OZOLEK, J. A., BUCK, R. H., GOEHRING, K. C. & HACKAM, D. J. 2016. The human milk oligosaccharide 2'-fucosyllactose attenuates the severity of experimental necrotising enterocolitis by enhancing mesenteric perfusion in the neonatal intestine. *Br J Nutr,* 116**,** 1175-1187.

GOPALAKRISHNA, K. P., MACADANGDANG, B. R., ROGERS, M. B., TOMETICH, J. T., FIREK, B. A., BAKER, R., JI, J., BURR, A. H. P., MA, C., GOOD, M., MOROWITZ, M. J. & HAND, T. W. 2019. Maternal IgA protects against the development of necrotizing enterocolitis in preterm infants. *Nat Med,* 25**,** 1110-1115.

GORDON, P. V. 2009. Understanding intestinal vulnerability to perforation in the extremely low birth weight infant. *Pediatr Res,* 65**,** 138-44.

GOTOH, A., KATOH, T., SAKANAKA, M., LING, Y., YAMADA, C., ASAKUMA, S., URASHIMA, T., TOMABECHI, Y., KATAYAMA-IKEGAMI, A., KURIHARA, S., YAMAMOTO, K., HARATA, G., HE, F., HIROSE, J., KITAOKA, M., OKUDA, S. & KATAYAMA, T. 2018. Sharing of human milk oligosaccharides degradants within bifidobacterial communities in faecal cultures supplemented with Bifidobacterium bifidum. *Sci Rep,* 8**,** 13958.

GRANGER, C. L., EMBLETON, N. D., PALMER, J. M., LAMB, C. A., BERRINGTON, J. E. & STEWART, C. J. 2021a. Maternal breastmilk, infant gut microbiome and the impact on preterm infant health. *Acta Paediatr,* 110**,** 450-457.

GRANGER, C. L., LAMB, C. A., EMBLETON, N. D., BECK, L. C., MASI, A. C., PALMER, J. M., STEWART, C. J. & BERRINGTON, J. E. 2021b. Secretory immunoglobulin A in preterm infants: determination of normal values in breast milk and stool. *Pediatric Research***,** 1-8.

GREGORY, K. E., DEFORGE, C. E., NATALE, K. M., PHILLIPS, M. & VAN MARTER, L. J. 2011. Necrotizing enterocolitis in the premature infant: neonatal nursing assessment, disease pathogenesis, and clinical presentation. *Adv Neonatal Care,* 11**,** 155-64; quiz 165-6.

GRIBAR, S. C., SODHI, C. P., RICHARDSON, W. M., ANAND, R. J., GITTES, G. K., BRANCA, M. F., JAKUB, A., SHI, X. H., SHAH, S., OZOLEK, J. A. & HACKAM, D. J. 2009. Reciprocal expression and signaling of TLR4 and TLR9 in the pathogenesis and treatment of necrotizing enterocolitis. *J Immunol,* 182**,** 636-46.

GRONDIN, J. A., KWON, Y. H., FAR, P. M., HAQ, S. & KHAN, W. I. 2020. Mucins in Intestinal Mucosal Defense and Inflammation: Learning From Clinical and Experimental Studies. *Front Immunol,* 11**,** 2054.

GUO, S., AL-SADI, R., SAID, H. M. & MA, T. Y. 2013. Lipopolysaccharide causes an increase in intestinal tight junction permeability in vitro and in vivo by inducing enterocyte membrane expression and localization of TLR-4 and CD14. *Am J Pathol,* 182**,** 375-87.

GYORGY, P., HOOVER, J. R., KUHN, R. & ROSE, C. S. 1954a. Bifidus factor. III. The rate of dialysis. *Arch Biochem Biophys,* 48**,** 209-13.

GYORGY, P., KUHN, R., ROSE, C. S. & ZILLIKEN, F. 1954b. Bifidus factor. II. Its occurrence in milk from different species and in other natural products. *Arch Biochem Biophys,* 48**,** 202-8.

GYÖRGY, P., NORRIS, R. F. & ROSE, C. S. 1954. Bifidus factor. I. A variant of Lactobacillus bifidus requiring a special growth factor. *Archives of Biochemistry and Biophysics,* 48**,** 193-201.

HASSINGER, D., CLAUSEN, D. M., NITKA, S., HERDT, A. & GRIFFIN, I. 2020. Analysis of Disialyllacto-N-Tetraose (DSLNT) Content in Milk From Mothers of Preterm Infants. *J Hum Lact***,** 890334420904041.

HE, Y., LIU, S., KLING, D. E., LEONE, S., LAWLOR, N. T., HUANG, Y., FEINBERG, S. B., HILL, D. R. & NEWBURG, D. S. 2016. The human milk oligosaccharide 2′-fucosyllactose modulates CD14 expression in human enterocytes, thereby attenuating LPS-induced inflammation. *Gut,* 65**,** 33-46.

HELLSTRÖM, A., LEY, D., HANSEN-PUPP, I., HALLBERG, B., LÖFQVIST, C., VAN MARTER, L., VAN WEISSENBRUCH, M., RAMENGHI, L. A., BEARDSALL, K., DUNGER, D., HÅRD, A. L. & SMITH, L. E. 2016. Insulin-like growth factor 1 has multisystem effects on foetal and preterm infant development. *Acta Paediatr,* 105**,** 576-86.

HENRICK, B. M., RODRIGUEZ, L., LAKSHMIKANTH, T., POU, C., HENCKEL, E., ARZOOMAND, A., OLIN, A., WANG, J., MIKES, J., TAN, Z., CHEN, Y., EHRLICH, A. M., BERNHARDSSON, A. K., MUGABO, C. H., AMBROSIANI, Y., GUSTAFSSON, A., CHEW, S., BROWN, H. K., PRAMBS, J., BOHLIN, K.,

MITCHELL, R. D., UNDERWOOD, M. A., SMILOWITZ, J. T., GERMAN, J. B., FRESE, S. A. & BRODIN, P. 2021. Bifidobacteria-mediated immune system imprinting early in life. *Cell,* 184**,** 3884-3898.e11.

HOEFLINGER, J. L., DAVIS, S. R., CHOW, J. & MILLER, M. J. 2015. In vitro impact of human milk oligosaccharides on Enterobacteriaceae growth. *J Agric Food Chem,* 63**,** 3295-302.

HOLGERSEN, K., GAO, X., NARAYANAN, R., GAUR, T., CAREY, G., BARTON, N., PAN, X., MUK, T., THYMANN, T. & SANGILD, P. T. 2020. Supplemental Insulin-Like Growth Factor-1 and Necrotizing Enterocolitis in Preterm Pigs. *Front Pediatr,* 8**,** 602047.

HOLMES, I., HARRIS, K. & QUINCE, C. 2012. Dirichlet multinomial mixtures: generative models for microbial metagenomics. *PLoS One,* 7**,** e30126.

HOLSCHER, H. D., BODE, L. & TAPPENDEN, K. A. 2017. Human Milk Oligosaccharides Influence Intestinal Epithelial Cell Maturation In Vitro. *J Pediatr Gastroenterol Nutr,* 64**,** 296-301.

HONDA, Y., NISHIMOTO, M., KATAYAMA, T. & KITAOKA, M. 2013. Characterization of the Cytosolic β-N-Acetylglucosaminidase from Bifidobacterium longum subsp. longum. *Journal of Applied Glycoscience,* 60**,** 141-146.

HOURIGAN, S. K., TA, A., WONG, W. S., CLEMENCY, N. C., PROVENZANO, M. G., BAVEJA, R., IYER, R., KLEIN, E. & NIEDERHUBER, J. E. 2016. The Microbiome in Necrotizing Enterocolitis: A Case Report in Twins and Minireview. *Clin Ther,* 38**,** 747-53.

HUANG, Z., LI, Y., LUO, Y. & GUO, H. 2021. Human milk oligosaccharides 3′-sialyllactose and 6′-sialyllactose protect intestine against necrotizing enterocolitis damage induced by hypoxia. *Journal of Functional Foods,* 86**,** 104708.

HUNNINGHAKE, G. W., DOERSCHUG, K. C., NYMON, A. B., SCHMIDT, G. A., MEYERHOLZ, D. K. & ASHARE, A. 2010. Insulin-like growth factor-1 levels contribute to the development of bacterial translocation in sepsis. *Am J Respir Crit Care Med,* 182**,** 517-25.

HUNT, K. M., FOSTER, J. A., FORNEY, L. J., SCHUTTE, U. M., BECK, D. L., ABDO, Z., FOX, L. K., WILLIAMS, J. E., MCGUIRE, M. K. & MCGUIRE, M. A. 2011. Characterization of the diversity and temporal stability of bacterial communities in human milk. *PLoS One,* 6**,** e21313.

HUNT, K. M., PREUSS, J., NISSAN, C., DAVLIN, C. A., WILLIAMS, J. E., SHAFII, B., RICHARDSON, A. D., MCGUIRE, M. K., BODE, L. & MCGUIRE, M. A. 2012.

Human milk oligosaccharides promote the growth of staphylococci. *Appl Environ Microbiol,* 78**,** 4763-70.

INGVORDSEN LINDAHL, I. E., ARTEGOITIA, V. M., DOWNEY, E., O'MAHONY, J. A., O'SHEA, C. A., RYAN, C. A., KELLY, A. L., BERTRAM, H. C. & SUNDEKILDE, U. K. 2019. Quantification of Human Milk Phospholipids: the Effect of Gestational and Lactational Age on Phospholipid Composition. *Nutrients,* 11.

JALILI-FIROOZINEZHAD, S., GAZZANIGA, F. S., CALAMARI, E. L., CAMACHO, D. M., FADEL, C. W., BEIN, A., SWENOR, B., NESTOR, B., CRONCE, M. J., TOVAGLIERI, A., LEVY, O., GREGORY, K. E., BREAULT, D. T., CABRAL, J. M. S., KASPER, D. L., NOVAK, R. & INGBER, D. E. 2019. A complex human gut microbiome cultured in an anaerobic intestine-on-a-chip. *Nature Biomedical Engineering,* 3**,** 520-531.

JAMES, K., MOTHERWAY, M. O., BOTTACINI, F. & VAN SINDEREN, D. 2016. Bifidobacterium breve UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways. *Sci Rep,* 6**,** 38560.

JANDHYALA, S. M., TALUKDAR, R., SUBRAMANYAM, C., VUYYURU, H., SASIKALA, M. & NAGESHWAR REDDY, D. 2015. Role of the normal gut microbiota. *World journal of gastroenterology,* 21**,** 8787-8803.

JANTSCHER-KRENN, E., ZHEREBTSOV, M., NISSAN, C., GOTH, K., GUNER, Y. S., NAIDU, N., CHOUDHURY, B., GRISHIN, A. V., FORD, H. R. & BODE, L. 2012. The human milk oligosaccharide disialyllacto-N-tetraose prevents necrotising enterocolitis in neonatal rats. *Gut,* 61**,** 1417-25.

JARI OKSANEN, F. G. B., MICHAEL FRIENDLY, ROELAND KINDT, PIERRE LEGENDRE, DAN MCGLINN, PETER R. MINCHIN, R. B. O'HARA, GAVIN L. SIMPSON, PETER SOLYMOS, M. HENRY H. STEVENS, EDUARD SZOECS, HELENE WAGNER 2019. vegan: Community Ecology Package. R package version 2.5-6. . *https://CRAN.R-project.org/package=vegan*.

JESCHKE, M. G., BOLDER, U., CHUNG, D. H., PRZKORA, R., MUELLER, U., THOMPSON, J. C., WOLF, S. E. & HERNDON, D. N. 2007. Gut mucosal homeostasis and cellular mediators after severe thermal trauma and the effect of insulin-like growth factor-I in combination with insulin-like growth factor binding protein-3. *Endocrinology,* 148**,** 354-62.

JIA, J., XUN, P., WANG, X., HE, K., TANG, Q., ZHANG, T., WANG, Y., TANG, W., LU, L., YAN, W., WANG, W., HU, T. & CAI, W. 2020. Impact of Postnatal Antibiotics

and Parenteral Nutrition on the Gut Microbiota in Preterm Infants During Early Life. *JPEN J Parenter Enteral Nutr,* 44**,** 639-654.

JIE, L., QI, C., SUN, J., YU, R., WANG, X., KORMA, S. A., XIANG, J., JIN, Q., AKOH, C. C., XIAO, H. & WANG, X. 2018. The impact of lactation and gestational age on the composition of branched-chain fatty acids in human breast milk. *Food Funct,* 9**,** 1747-1754.

JUNG, C., HUGOT, J.-P. & BARREAU, F. 2010. Peyer's Patches: The Immune Sensors of the Intestine. *International Journal of Inflammation,* 2010**,** 823710.

KALLIOMÄKI, M., COLLADO, M. C., SALMINEN, S. & ISOLAURI, E. 2008. Early differences in fecal microbiota composition in children may predict overweight. *Am J Clin Nutr,* 87**,** 534-8.

KALLIOMÄKI, M., KIRJAVAINEN, P., EEROLA, E., KERO, P., SALMINEN, S. & ISOLAURI, E. 2001. Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. *J Allergy Clin Immunol,* 107**,** 129-34.

KANDASAMY, J., HUDA, S., AMBALAVANAN, N. & JILLING, T. 2014. Inflammatory signals that regulate intestinal epithelial renewal, differentiation, migration and cell death: Implications for necrotizing enterocolitis. *Pathophysiology,* 21**,** 67-80.

KATAYAMA, T., SAKUMA, A., KIMURA, T., MAKIMURA, Y., HIRATAKE, J., SAKATA, K., YAMANOI, T., KUMAGAI, H. & YAMAMOTO, K. 2004. Molecular cloning and characterization of Bifidobacterium bifidum 1,2-alpha-L-fucosidase (AfcA), a novel inverting glycosidase (glycoside hydrolase family 95). *J Bacteriol,* 186**,** 4885-93.

KAYISOGLU, O., WEISS, F., NIKLAS, C., PIEROTTI, I., POMPAIAH, M., WALLASCHEK, N., GERMER, C.-T., WIEGERING, A. & BARTFELD, S. 2021. Location-specific cell identity rather than exposure to GI microbiota defines many innate immune signalling cascades in the gut epithelium. *Gut,* 70**,** 687-697.

KENNEDY, K. M., DE GOFFAU, M. C., PEREZ-MUÑOZ, M. E., ARRIETA, M.-C., BÄCKHED, F., BORK, P., BRAUN, T., BUSHMAN, F. D., DORE, J., DE VOS, W. M., EARL, A. M., EISEN, J. A., ELOVITZ, M. A., GANAL-VONARBURG, S. C., GÄNZLE, M. G., GARRETT, W. S., HALL, L. J., HORNEF, M. W., HUTTENHOWER, C., KONNIKOVA, L., LEBEER, S., MACPHERSON, A. J., MASSEY, R. C., MCHARDY, A. C., KOREN, O., LAWLEY, T. D., LEY, R. E., O'MAHONY, L., O'TOOLE, P. W., PAMER, E. G., PARKHILL, J., RAES, J., RATTEI, T., SALONEN, A., SEGAL, E., SEGATA, N., SHANAHAN, F., SLOBODA, D. M., SMITH, G. C. S., SOKOL, H., SPECTOR, T. D., SURETTE, M.

G., TANNOCK, G. W., WALKER, A. W., YASSOUR, M. & WALTER, J. 2023. Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies. *Nature,* 613**,** 639-649.

KHODAYAR-PARDO, P., MIRA-PASCUAL, L., COLLADO, M. C. & MARTINEZ-COSTA, C. 2014. Impact of lactation stage, gestational age and mode of delivery on breast milk microbiota. *J Perinatol,* 34**,** 599-605.

KITAOKA, M., TIAN, J. & NISHIMOTO, M. 2005. Novel putative galactose operon involving lacto-N-biose phosphorylase in Bifidobacterium longum. *Appl Environ Microbiol,* 71**,** 3158-62.

KIYOHARA, M., TANIGAWA, K., CHAIWANGSRI, T., KATAYAMA, T., ASHIDA, H. & YAMAMOTO, K. 2011. An exo-alpha-sialidase from bifidobacteria involved in the degradation of sialyloligosaccharides in human milk and intestinal glycoconjugates. *Glycobiology,* 21**,** 437-47.

KOH, A., DE VADDER, F., KOVATCHEVA-DATCHARY, P. & BÄCKHED, F. 2016. From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell,* 165**,** 1332-1345.

KOJIMA, I., TANAKA, T., INAGI, R., NISHI, H., ABURATANI, H., KATO, H., MIYATA, T., FUJITA, T. & NANGAKU, M. 2009. Metallothionein is upregulated by hypoxia and stabilizes hypoxia-inducible factor in the kidney. *Kidney International,* 75**,** 268-277.

KONG, C., AKKERMAN, R., KLOSTERMANN, C. E., BEUKEMA, M., OERLEMANS, M. M. P., SCHOLS, H. A. & DE VOS, P. 2021. Distinct fermentation of human milk oligosaccharides 3-FL and LNT2 and GOS/inulin by infant gut microbiota and impact on adhesion of Lactobacillus plantarum WCFS1 to gut epithelial cells. *Food Funct,* 12**,** 12513-12525.

KRAICZY, J., NAYAK, K. M., HOWELL, K. J., ROSS, A., FORBESTER, J., SALVESTRINI, C., MUSTATA, R., PERKINS, S., ANDERSSON-ROLF, A., LEENEN, E., LIEBERT, A., VALLIER, L., ROSENSTIEL, P. C., STEGLE, O., DOUGAN, G., HEUSCHKEL, R., KOO, B. K. & ZILBAUER, M. 2019. DNA methylation defines regional identity of human intestinal epithelial organoids and undergoes dynamic changes during development. *Gut,* 68**,** 49-61.

KUNZ, C., MEYER, C., COLLADO, M. C., GEIGER, L., GARCIA-MANTRANA, I., BERTUA-RIOS, B., MARTINEZ-COSTA, C., BORSCH, C. & RUDLOFF, S. 2017. Influence of Gestational Age, Secretor, and Lewis Blood Group Status on the Oligosaccharide Content of Human Milk. *J Pediatr Gastroenterol Nutr,* 64**,** 789-798.

LA ROSA, P. S., WARNER, B. B., ZHOU, Y., WEINSTOCK, G. M., SODERGREN, E., HALL-MOORE, C. M., STEVENS, H. J., BENNETT, W. E., SHAIKH, N. & LINNEMAN, L. A. 2014. Patterned progression of bacterial populations in the premature infant gut. *Proceedings of the National Academy of Sciences,* 111**,** 12522-12527.

LACKEY, K. A., WILLIAMS, J. E., MEEHAN, C. L., ZACHEK, J. A., BENDA, E. D., PRICE, W. J., FOSTER, J. A., SELLEN, D. W., KAMAU-MBUTHIA, E. W., KAMUNDIA, E. W., MBUGUA, S., MOORE, S. E., PRENTICE, A. M., K, D. G., KVIST, L. J., OTOO, G. E., GARCIA-CARRAL, C., JIMENEZ, E., RUIZ, L., RODRIGUEZ, J. M., PAREJA, R. G., BODE, L., MCGUIRE, M. A. & MCGUIRE, M. K. 2019. What's Normal? Microbiomes in Human Milk and Infant Feces Are Related to Each Other but Vary Geographically: The INSPIRE Study. *Front Nutr,* 6**,** 45.

LAGIER, J. C., ARMOUGOM, F., MILLION, M., HUGON, P., PAGNIER, I., ROBERT, C., BITTAR, F., FOURNOUS, G., GIMENEZ, G., MARANINCHI, M., TRAPE, J. F., KOONIN, E. V., LA SCOLA, B. & RAOULT, D. 2012. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clinical Microbiology and Infection,* 18**,** 1185-1193.

LAGIER, J. C., DUBOURG, G., MILLION, M., CADORET, F., BILEN, M., FENOLLAR, F., LEVASSEUR, A., ROLAIN, J. M., FOURNIER, P. E. & RAOULT, D. 2018. Culturing the human microbiota and culturomics. *Nat Rev Microbiol,* 16**,** 540-550.

LAGIER, J. C., KHELAIFIA, S., ALOU, M. T., NDONGO, S., DIONE, N., HUGON, P., CAPUTO, A., CADORET, F., TRAORE, S. I., SECK, E. H., DUBOURG, G., DURAND, G., MOUREMBOU, G., GUILHOT, E., TOGO, A., BELLALI, S., BACHAR, D., CASSIR, N., BITTAR, F., DELERCE, J., MAILHE, M., RICABONI, D., BILEN, M., DANGUI NIEKO, N. P., DIA BADIANE, N. M., VALLES, C., MOUELHI, D., DIOP, K., MILLION, M., MUSSO, D., ABRAHÃO, J., AZHAR, E. I., BIBI, F., YASIR, M., DIALLO, A., SOKHNA, C., DJOSSOU, F., VITTON, V., ROBERT, C., ROLAIN, J. M., LA SCOLA, B., FOURNIER, P. E., LEVASSEUR, A. & RAOULT, D. 2016. Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat Microbiol,* 1**,** 16203.

LANE, J. A., O'CALLAGHAN, J., CARRINGTON, S. D. & HICKEY, R. M. 2013. Transcriptional response of HT-29 intestinal epithelial cells to human and bovine milk oligosaccharides. *Br J Nutr,* 110**,** 2127-37.

LANIK, W. E., LUKE, C. J., NOLAN, L. S., GONG, Q., RIMER, J. M., GALE, S. E., LUC, R., BIDANI, S. S., SIBBALD, C. A., LEWIS, A. N., MIHI, B., AGRAWAL, P.,

GOREE, M., MAESTAS, M., HU, E., PETERS, D. G. & GOOD, M. 2020. Microfluidic Device Facilitates Novel In Vitro Modeling of Human Neonatal Necrotizing Enterocolitis-on-a-Chip. *bioRxiv***,** 2020.11.29.402735.

LAWN, J. E., OHUMA, E. O., BRADLEY, E., IDUETA, L. S., HAZEL, E., OKWARAJI, Y. B., ERCHICK, D. J., YARGAWA, J., KATZ, J., LEE, A. C. C., DIAZ, M., SALASIBEW, M., REQUEJO, J., HAYASHI, C., MOLLER, A.-B., BORGHI, E., BLACK, R. E., BLENCOWE, H., ASHORN, P., BLACK, R. E., LAWN, J. E., ASHORN, U., KLEIN, N., HOFMEYR, G. J., TEMMERMAN, M., ASKARI, S., OHUMA, E. O., MOLLER, A.-B., BRADLEY, E., CHAKWERA, S., HUSSAIN-ALKHATEEB, L., LEWIN, A., OKWARAJI, Y. B., RETNO MAHANANI, W., WHITE JOHANSSON, E., LAVIN, T., ESTEVEZ FERNANDEZ, D., GATICA DOMÍNGUEZ, G., DE COSTA, A., CRESSWELL, J. A., KRASEVEC, J., LAWN, J. E., BLENCOWE, H., REQUEJO, J., MORAN, A. C., PINGRAY, V., CORMICK, G., GIBBONS, L., BELIZAN, J., GUEVEL, C., WARRILOW, K., GORDON, A., FLENADY, V., SEXTON, J., LAWFORD, H., PAIXAO, E. S., ROCHA FALCÃO, I., LIMA BARRETO, M., LISONKOVA, S., WEN, Q., MARDONES, F., CAULIER-CISTERNA, R., ACUÑA, J., VELEBIL, P., JIROVA, J., HORVÁTH-PUHÓ, E., TOFT SØRENSEN, H., SAKKEUS, L., ABULADZE, L., GISSLER, M., MORADI-LAKEH, M., HEIDARZADEH, M., KHALILI, N., A. YUNIS, K., AL BIZRI, A., NAKAD, P., DEVI KARALASINGAM, S., R JEGANATHAN, J. R., BINTI BAHARUM, N., SUÁREZ-IDUETA, L., BARRANCO FLORES, A., GONZALEZ ROLDAN, J. F., LOPEZ ALVAREZ, S., VAN DIJK, A. E., BROEDERS, L., HUICHO, L., QUEZADA PINEDO, H. G., CAJACHAGUA-TORRES, K. N., CARRILLO-LARCO, R. M., TARAZONA MEZA, C. E., GUZMAN-VILCA, W. C., OLUKADE, T. O., ALI, H. A., ALYAFEI, F., ALQUBAISI, M., ALTURK, M. R., KIM, H. Y., CHO, G. J., RAZAZ, N., SÖDERLING, J., et al. 2023. Small babies, big risks: global estimates of prevalence and mortality for vulnerable newborns to accelerate change and improve counting. *The Lancet,* 401**,** 1707-1719.

LAWSON, M. A. E., O'NEILL, I. J., KUJAWSKA, M., GOWRINADH JAVVADI, S., WIJEYESEKERA, A., FLEGG, Z., CHALKLEN, L. & HALL, L. J. 2019. Breast milk-derived human milk oligosaccharides promote Bifidobacterium interactions within a single ecosystem. *ISME J.*

LEAPHART, C. L., CAVALLO, J., GRIBAR, S. C., CETIN, S., LI, J., BRANCA, M. F., DUBOWSKI, T. D., SODHI, C. P. & HACKAM, D. J. 2007. A critical role for TLR4

in the pathogenesis of necrotizing enterocolitis by modulating intestinal injury and repair. *J Immunol,* 179**,** 4808-20.

LEUSHACKE, M. & BARKER, N. 2014. Ex vivo culture of the intestinal epithelium: strategies and applications. *Gut,* 63**,** 1345-54.

LEY, D., HALLBERG, B., HANSEN-PUPP, I., DANI, C., RAMENGHI, L. A., MARLOW, N., BEARDSALL, K., BHATTI, F., DUNGER, D., HIGGINSON, J. D., MAHAVEER, A., MEZU-NDUBUISI, O. J., REYNOLDS, P., GIANNANTONIO, C., VAN WEISSENBRUCH, M., BARTON, N., TOCOIAN, A., HAMDANI, M., JOCHIM, E., MANGILI, A., CHUNG, J. K., TURNER, M. A., SMITH, L. E. H. & HELLSTRÖM, A. 2019. rhIGF-1/rhIGFBP-3 in Preterm Infants: A Phase 2 Randomized Controlled Trial. *J Pediatr,* 206**,** 56-65.e8.

LI, S., ZHANG, L., ZHOU, Q., JIANG, S., YANG, Y. & CAO, Y. 2019. Characterization of Stem Cells and Immune Cells in Preterm and Term Mother's Milk. *J Hum Lact,* 35**,** 528-534.

LI, X., CHEN, H. & EPSTEIN, P. N. 2004. Metallothionein protects islets from hypoxia and extends islet graft survival by scavenging most kinds of reactive oxygen species. *J Biol Chem,* 279**,** 765-71.

LIAN, X., ZHANG, W., HE-YANG, J. & ZHOU, X. 2022. Human milk oligosaccharide disialyllacto-n-tetraose protects human intestinal epithelium integrity and permeability against mast cell chymase-induced disruption by stabilizing ZO-1/FAK/P38 pathway of intestinal epithelial cell. *Immunopharmacol Immunotoxicol***,** 1-10.

LIN, A. E., AUTRAN, C. A., ESPANOLA, S. D., BODE, L. & NIZET, V. 2014. Human Milk Oligosaccharides Protect Bladder Epithelial Cells Against Uropathogenic Escherichia coli Invasion and Cytotoxicity. *The Journal of Infectious Diseases,* 209**,** 389-398.

LIN, A. E., AUTRAN, C. A., SZYSZKA, A., ESCAJADILLO, T., HUANG, M., GODULA, K., PRUDDEN, A. R., BOONS, G. J., LEWIS, A. L., DORAN, K. S., NIZET, V. & BODE, L. 2017. Human milk oligosaccharides inhibit growth of group B Streptococcus. *J Biol Chem,* 292**,** 11243-11249.

LING, X., LINGLONG, P., WEIXIA, D. & HONG, W. 2016. Protective Effects of Bifidobacterium on Intestinal Barrier Function in LPS-Induced Enterocyte Barrier Injury of Caco-2 Monolayers and in a Rat NEC Model. *PLoS One,* 11**,** e0161635.

LIU, D., XU, Y., FENG, J., YU, J., HUANG, J. & LI, Z. 2021. Mucins and Tight Junctions are Severely Altered in Necrotizing Enterocolitis Neonates. *Am J Perinatol,* 38**,** 1174-1180.

LOCASCIO, R. G., DESAI, P., SELA, D. A., WEIMER, B. & MILLS, D. A. 2010. Broad conservation of milk utilization genes in Bifidobacterium longum subsp. infantis as revealed by comparative genomic hybridization. *Appl Environ Microbiol,* 76**,** 7373-81.

LOCASCIO, R. G., NINONUEVO, M. R., FREEMAN, S. L., SELA, D. A., GRIMM, R., LEBRILLA, C. B., MILLS, D. A. & GERMAN, J. B. 2007. Glycoprofiling of Bifidobacterial Consumption of Human Milk Oligosaccharides Demonstrates Strain Specific, Preferential Consumption of Small Chain Glycans Secreted in Early Human Lactation. *Journal of Agricultural and Food Chemistry,* 55**,** 8914-8919.

LOCASCIO, R. G., NIÑONUEVO, M. R., KRONEWITTER, S. R., FREEMAN, S. L., GERMAN, J. B., LEBRILLA, C. B. & MILLS, D. A. 2009. A versatile and scalable strategy for glycoprofiling bifidobacterial consumption of human milk oligosaccharides. *Microb Biotechnol,* 2**,** 333-42.

LOUIS, P. & FLINT, H. J. 2017. Formation of propionate and butyrate by the human colonic microbiota. *Environ Microbiol,* 19**,** 29-41.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.

LU, Z., DING, L., LU, Q. & CHEN, Y. H. 2013. Claudins in intestines: Distribution and functional significance in health and diseases. *Tissue Barriers,* 1**,** e24978.

LUNA, E., PARKAR, S. G., KIRMIZ, N., HARTEL, S., HEARN, E., HOSSINE, M., KURDIAN, A., MENDOZA, C., ORR, K., PADILLA, L., RAMIREZ, K., SALCEDO, P., SERRANO, E., CHOUDHURY, B., PAULCHAKRABARTI, M., PARKER, C. T., HUYNH, S., COOPER, K. & FLORES, G. E. 2022. Utilization Efficiency of Human Milk Oligosaccharides by Human-Associated Akkermansia Is Strain Dependent. *Appl Environ Microbiol,* 88**,** e0148721.

LUO, H., GUO, P. & ZHOU, Q. 2012. Role of TLR4/NF-κB in damage to intestinal mucosa barrier function and bacterial translocation in rats exposed to hypoxia. *PLoS One,* 7**,** e46291.

MAI, V., TORRAZZA, R. M., UKHANOVA, M., WANG, X., SUN, Y., LI, N., SHUSTER, J., SHARMA, R., HUDAK, M. L. & NEU, J. 2013. Distortions in Development of Intestinal Microbiota Associated with Late Onset Sepsis in Preterm Infants. *PLOS ONE,* 8**,** e52876.

MAI, V., YOUNG, C. M., UKHANOVA, M., WANG, X., SUN, Y., CASELLA, G., THERIAQUE, D., LI, N., SHARMA, R., HUDAK, M. & NEU, J. 2011. Fecal microbiota in premature infants prior to necrotizing enterocolitis. *PLoS One,* 6**,** e20647.

MAIER, E., ANDERSON, R. C. & ROY, N. C. 2014. Understanding how commensal obligate anaerobic bacteria regulate immune functions in the large intestine. *Nutrients,* 7**,** 45-73.

MANTHEY, C. F., AUTRAN, C. A., ECKMANN, L. & BODE, L. 2014. Human milk oligosaccharides protect against enteropathogenic Escherichia coli attachment in vitro and EPEC colonization in suckling mice. *J Pediatr Gastroenterol Nutr,* 58**,** 165-8.

MANTOVANI, V., GALEOTTI, F., MACCARI, F. & VOLPI, N. 2016. Recent advances on separation and characterization of human milk oligosaccharides. *Electrophoresis,* 37**,** 1514-24.

MARCOBAL, A., BARBOZA, M., FROEHLICH, J. W., BLOCK, D. E., GERMAN, J. B., LEBRILLA, C. B. & MILLS, D. A. 2010. Consumption of human milk oligosaccharides by gut-related microbes. *J Agric Food Chem,* 58**,** 5334-40.

MARCOBAL, A., BARBOZA, M., SONNENBURG, E. D., PUDLO, N., MARTENS, E. C., DESAI, P., LEBRILLA, C. B., WEIMER, B. C., MILLS, D. A., GERMAN, J. B. & SONNENBURG, J. L. 2011. Bacteroides in the infant gut consume milk oligosaccharides via mucus-utilization pathways. *Cell Host Microbe,* 10**,** 507-14.

MARRIAGE, B. J., BUCK, R. H., GOEHRING, K. C., OLIVER, J. S. & WILLIAMS, J. A. 2015. Infants Fed a Lower Calorie Formula With 2'FL Show Growth and 2'FL Uptake Like Breast-Fed Infants. *Journal of pediatric gastroenterology and nutrition,* 61**,** 649-658.

MASI, A. C., FOFANOVA, T. Y., LAMB, C. A., AUCHTUNG, J. M., BRITTON, R. A., ESTES, M. K., RAMANI, S., COCKELL, S. J., COXHEAD, J., EMBLETON, N. D., BERRINGTON, J. E., PETROSINO, J. F. & STEWART, C. J. 2022. Distinct gene expression profiles between human preterm-derived and adult-derived intestinal organoids exposed to Enterococcus faecalis: a pilot study. *Gut,* 71**,** 2141-2143.

MASI, A. C. & STEWART, C. J. 2022. Untangling human milk oligosaccharides and infant gut microbiome. *Iscience,* 25**,** 103542.

MCGUIRE, M. K., MEEHAN, C. L., MCGUIRE, M. A., WILLIAMS, J. E., FOSTER, J., SELLEN, D. W., KAMAU-MBUTHIA, E. W., KAMUNDIA, E. W., MBUGUA, S., MOORE, S. E., PRENTICE, A. M., KVIST, L. J., OTOO, G. E., BROOKER, S. L., PRICE, W. J., SHAFII, B., PLACEK, C., LACKEY, K. A., ROBERTSON, B., MANZANO, S., RUIZ, L., RODRIGUEZ, J. M., PAREJA, R. G. & BODE, L. 2017. What's normal? Oligosaccharide concentrations and profiles in milk produced by healthy women vary geographically. *Am J Clin Nutr,* 105**,** 1086-1100.

MEINZEN-DERR, J., POINDEXTER, B., WRAGE, L., MORROW, A. L., STOLL, B. & DONOVAN, E. F. 2009. Role of human milk in extremely low birth weight infants' risk of necrotizing enterocolitis or death. *J Perinatol,* 29**,** 57-62.

MENG, D., SOMMELLA, E., SALVIATI, E., CAMPIGLIA, P., GANGULI, K., DJEBALI, K., ZHU, W. & WALKER, W. A. 2020. Indole-3-lactic acid, a metabolite of tryptophan, secreted by Bifidobacterium longum subspecies infantis is anti-inflammatory in the immature intestine. *Pediatr Res,* 88**,** 209-217.

MIDDENDORP, S., SCHNEEBERGER, K., WIEGERINCK, C. L., MOKRY, M., AKKERMAN, R. D., VAN WIJNGAARDEN, S., CLEVERS, H. & NIEUWENHUIS, E. E. 2014. Adult stem cells in the small intestine are intrinsically programmed with their location-specific function. *Stem Cells,* 32**,** 1083-91.

MIWA, M., HORIMOTO, T., KIYOHARA, M., KATAYAMA, T., KITAOKA, M., ASHIDA, H. & YAMAMOTO, K. 2010. Cooperation of β-galactosidase and β-N-acetylhexosaminidase from bifidobacteria in assimilation of human milk oligosaccharides with type 2 structure. *Glycobiology,* 20**,** 1402-9.

MØLLER, P. L., JØRGENSEN, F., HANSEN, O. C., MADSEN, S. M. & STOUGAARD, P. 2001. Intra- and extracellular beta-galactosidases from Bifidobacterium bifidum and B. infantis: molecular cloning, heterologous expression, and comparative characterization. *Applied and environmental microbiology,* 67**,** 2276-2283.

MOOSSAVI, S., ATAKORA, F., MILIKU, K., SEPEHRI, S., ROBERTSON, B., DUAN, Q. L., BECKER, A. B., MANDHANE, P. J., TURVEY, S. E., MORAES, T. J., LEFEBVRE, D. L., SEARS, M. R., SUBBARAO, P., FIELD, C. J., BODE, L., KHAFIPOUR, E. & AZAD, M. B. 2019a. Integrated Analysis of Human Milk Microbiota With Oligosaccharides and Fatty Acids in the CHILD Cohort. *Front Nutr,* 6**,** 58.

MOOSSAVI, S. & AZAD, M. B. 2019. Origins of human milk microbiota: new evidence and arising questions. *Gut Microbes***,** 1-10.

MOOSSAVI, S., SEPEHRI, S., ROBERTSON, B., BODE, L., GORUK, S., FIELD, C. J., LIX, L. M., DE SOUZA, R. J., BECKER, A. B., MANDHANE, P. J., TURVEY, S. E., SUBBARAO, P., MORAES, T. J., LEFEBVRE, D. L., SEARS, M. R., KHAFIPOUR, E. & AZAD, M. B. 2019b. Composition and Variation of the Human Milk Microbiota Are Influenced by Maternal and Early-Life Factors. *Cell Host Microbe,* 25**,** 324-335 e4.

MORRISON, D. J. & PRESTON, T. 2016. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut Microbes,* 7**,** 189-200.

MORROW, A. L., LAGOMARCINO, A. J., SCHIBLER, K. R., TAFT, D. H., YU, Z., WANG, B., ALTAYE, M., WAGNER, M., GEVERS, D., WARD, D. V., KENNEDY, M. A., HUTTENHOWER, C. & NEWBURG, D. S. 2013. Early microbial and metabolomic signatures predict later onset of necrotizing enterocolitis in preterm infants. *Microbiome,* 1**,** 13.

MOTTA, J. P., MAGNE, L., DESCAMPS, D., ROLLAND, C., SQUARZONI–DALE, C., ROUSSET, P., MARTIN, L., CENAC, N., BALLOY, V., HUERRE, M., FRÖHLICH, L. F., JENNE, D., WARTELLE, J., BELAAOUAJ, A., MAS, E., VINEL, J. P., ALRIC, L., CHIGNARD, M., VERGNOLLE, N. & SALLENAVE, J. M. 2011. Modifying the Protease, Antiprotease Pattern by Elafin Overexpression Protects Mice From Colitis. *Gastroenterology,* 140**,** 1272-1282.

MUGLIA, C., MERCER, N., TOSCANO, M. A., SCHATTNER, M., POZNER, R., CERLIANI, J. P., GOBBI, R. P., RABINOVICH, G. A. & DOCENA, G. H. 2011. The glycan-binding protein galectin-1 controls survival of epithelial cells along the crypt-villus axis of small intestine. *Cell Death & Disease,* 2**,** e163-e163.

MUSEMECHE, C., CAPLAN, M., HSUEH, W., SUN, X. & KELLY, A. 1991. Experimental necrotizing enterocolitis: The role of polymorphonuclear neutrophils. *Journal of Pediatric Surgery,* 26**,** 1047-1050.

NAKHLA, T., FU, D., ZOPF, D., BRODSKY, N. L. & HURT, H. 1999. Neutral oligosaccharide content of preterm human milk. *Br J Nutr,* 82**,** 361-7.

NATARAJAN, G. & SHANKARAN, S. 2016. Short- and Long-Term Outcomes of Moderate and Late Preterm Infants. *Am J Perinatol,* 33**,** 305-17.

NATIVIDAD, J. M., RYTZ, A., KEDDANI, S., BERGONZELLI, G. & GARCIA-RODENAS, C. L. 2020. Blends of Human Milk Oligosaccharides Confer Intestinal Epithelial Barrier Protection in Vitro. *Nutrients,* 12.

NEU, J. & WALKER, W. A. 2011. Necrotizing enterocolitis. *N Engl J Med,* 364**,** 255-64.

NEVILLE, J., PAWLAK, R., CHANG, M., FURST, A., BODE, L. & PERRIN, M. T. 2021. A Cross-Sectional Assessment of Human Milk Oligosaccharide Composition of Vegan, Vegetarian, and Nonvegetarian Mothers. *Breastfeeding Medicine,* 17**,** 210-217.

NILSEN, M., MADELEN SAUNDERS, C., LEENA ANGELL, I., ARNTZEN, M. Ø., LØDRUP CARLSEN, K. C., CARLSEN, K.-H., HAUGEN, G., HELDAL HAGEN, L., CARLSEN, M. H., HEDLIN, G., MONCEYRON JONASSEN, C., NORDLUND, B., MARIA REHBINDER, E., SKJERVEN, H. O., SNIPEN, L., CATHRINE STAFF, A., VETTUKATTIL, R. & RUDI, K. 2020. Butyrate Levels in the Transition from an

Infant- to an Adult-Like Gut Microbiota Correlate with Bacterial Networks Associated with Eubacterium Rectale and Ruminococcus Gnavus. *Genes,* 11**,** 1245.

NISHIMOTO, M. & KITAOKA, M. 2007a. Identification of N-acetylhexosamine 1-kinase in the complete lacto-N-biose I/galacto-N-biose metabolic pathway in Bifidobacterium longum. *Applied and environmental microbiology,* 73**,** 6444-6449.

NISHIMOTO, M. & KITAOKA, M. 2007b. Identification of the Putative Proton Donor Residue of Lacto-N-biose Phosphorylase (EC 2.4.1.211). *Bioscience, Biotechnology, and Biochemistry,* 71**,** 1587-1591.

NOLAN, L. S., RIMER, J. M. & GOOD, M. 2020. The Role of Human Milk Oligosaccharides and Probiotics on the Neonatal Microbiome and Risk of Necrotizing Enterocolitis: A Narrative Review. *Nutrients,* 12.

OBERMEIER, S., RUDLOFF, S., POHLENTZ, G., LENTZE, M. J. & KUNZ, C. 1999. Secretion of 13C-labelled oligosaccharides into human milk and infant's urine after an oral [13C]galactose load. *Isotopes Environ Health Stud,* 35**,** 119-25.

OLM, M. R., BHATTACHARYA, N., CRITS-CHRISTOPH, A., FIREK, B. A., BAKER, R., SONG, Y. S., MOROWITZ, M. J. & BANFIELD, J. F. 2019. Necrotizing enterocolitis is preceded by increased gut bacterial replication, Klebsiella, and fimbriae-encoding bacteria. *Sci Adv,* 5**,** eaax5727.

PACE, R. M., WILLIAMS, J. E., ROBERTSON, B., LACKEY, K. A., MEEHAN, C. L., PRICE, W. J., FOSTER, J. A., SELLEN, D. W., KAMAU-MBUTHIA, E. W., KAMUNDIA, E. W., MBUGUA, S., MOORE, S. E., PRENTICE, A. M., KITA, D. G., KVIST, L. J., OTOO, G. E., RUIZ, L., RODRÍGUEZ, J. M., PAREJA, R. G., MCGUIRE, M. A., BODE, L. & MCGUIRE, M. K. 2021. Variation in Human Milk Composition Is Related to Differences in Milk and Infant Fecal Microbial Communities. *Microorganisms,* 9**,** 1153.

PALMER, C., BIK, E. M., DIGIULIO, D. B., RELMAN, D. A. & BROWN, P. O. 2007. Development of the human infant intestinal microbiota. *PLoS Biol,* 5**,** e177.

PAMMI, M., COPE, J., TARR, P. I., WARNER, B. B., MORROW, A. L., MAI, V., GREGORY, K. E., KROLL, J. S., MCMURTRY, V., FERRIS, M. J., ENGSTRAND, L., LILJA, H. E., HOLLISTER, E. B., VERSALOVIC, J. & NEU, J. 2017. Intestinal dysbiosis in preterm infants preceding necrotizing enterocolitis: a systematic review and meta-analysis. *Microbiome,* 5**,** 31.

PAN, H. X., ZHANG, C. S., LIN, C. H., CHEN, M. M., ZHANG, X. Z. & YU, N. 2021. Mucin 1 and interleukin-11 protein expression and inflammatory reactions in the intestinal

mucosa of necrotizing enterocolitis children after surgery. *World J Clin Cases,* 9**,** 7372-7380.

PARSCHAT, K., MELSAETHER, C., JÄPELT, K. R. & JENNEWEIN, S. 2021. Clinical Evaluation of 16-Week Supplementation with 5HMO-Mix in Healthy-Term Human Infants to Determine Tolerability, Safety, and Effect on Growth. *Nutrients,* 13**,** 2871.

PATEL, A. L., JOHNSON, T. J., ENGSTROM, J. L., FOGG, L. F., JEGIER, B. J., BIGGER, H. R. & MEIER, P. P. 2013. Impact of early human milk on sepsis and health-care costs in very low birth weight infants. *Journal of Perinatology,* 33**,** 514-519.

PATEL, R. M. 2016. Short- and Long-Term Outcomes for Extremely Preterm Infants. *Am J Perinatol,* 33**,** 318-28.

PATRO, R., DUGGAL, G., LOVE, M. I., IRIZARRY, R. A. & KINGSFORD, C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods,* 14**,** 417-419.

PFLUGHOEFT, K. J. & VERSALOVIC, J. 2012. Human microbiome in health and disease. *Annu Rev Pathol,* 7**,** 99-122.

PICHLER, M. J., YAMADA, C., SHUOKER, B., ALVAREZ-SILVA, C., GOTOH, A., LETH, M. L., SCHOOF, E., KATOH, T., SAKANAKA, M., KATAYAMA, T., JIN, C., KARLSSON, N. G., ARUMUGAM, M., FUSHINOBU, S. & ABOU HACHEM, M. 2020. Butyrate producing colonic Clostridiales metabolise human milk oligosaccharides and cross feed on mucin via conserved pathways. *Nat Commun,* 11**,** 3285.

PLAZA-DÍAZ, J., FONTANA, L. & GIL, A. 2018. Human Milk Oligosaccharides and Immune System Development. *Nutrients,* 10**,** 1038.

PUCCIO, G., ALLIET, P., CAJOZZO, C., JANSSENS, E., CORSELLO, G., SPRENGER, N., WERNIMONT, S., EGLI, D., GOSONIU, L. & STEENHOUT, P. 2017. Effects of Infant Formula With Human Milk Oligosaccharides on Growth and Morbidity: A Randomized Multicenter Trial. *J Pediatr Gastroenterol Nutr,* 64**,** 624-631.

QUAST, C., PRUESSE, E., YILMAZ, P., GERKEN, J., SCHWEER, T., YARZA, P., PEPLIES, J. & GLÖCKNER, F. O. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res,* 41**,** D590-6.

RAHMAN, S. F., OLM, M. R., MOROWITZ, M. J. & BANFIELD, J. F. 2019. Functional potential of bacterial strains in the premature infant gut microbiome is associated with gestational age. *bioRxiv***,** 530139.

RAMANI, S., STEWART, C. J., LAUCIRICA, D. R., AJAMI, N. J., ROBERTSON, B., AUTRAN, C. A., SHINGE, D., RANI, S., ANANDAN, S., HU, L., FERREON, J. C.,

KURUVILLA, K. A., PETROSINO, J. F., VENKATARAM PRASAD, B. V., BODE, L., KANG, G. & ESTES, M. K. 2018. Human milk oligosaccharides, milk microbiome and infant gut microbiome modulate neonatal rotavirus infection. *Nat Commun,* 9**,** 5010.

RAMESH-KUMAR, D. & GUIL, S. 2022. The IGF2BP family of RNA binding proteins links epitranscriptomics to cancer. *Seminars in Cancer Biology,* 86**,** 18-31.

RANJAN, R., RANI, A., METWALLY, A., MCGEE, H. S. & PERKINS, D. L. 2016. Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochem Biophys Res Commun,* 469**,** 967-77.

RASMUSSEN, S. O., MARTIN, L., OSTERGAARD, M. V., RUDLOFF, S., ROGGENBUCK, M., NGUYEN, D. N., SANGILD, P. T. & BERING, S. B. 2017. Human milk oligosaccharide effects on intestinal function and inflammation after preterm birth in pigs. *J Nutr Biochem,* 40**,** 141-154.

RAUTAVA, S. 2016. Early microbial contact, the breast milk microbiome and child health. *Journal of Developmental Origins of Health and Disease,* 7**,** 5-14.

RAUTAVA, S., LUOTO, R., SALMINEN, S. & ISOLAURI, E. 2012. Microbial contact during pregnancy, intestinal colonization and human disease. *Nat Rev Gastroenterol Hepatol,* 9**,** 565-76.

RAVELLE, W. 2016. psych: Procedures for Personality and Psychological Research. *http://CRAN.R-project.org/package=psych*.

RAVISANKAR, S., TATUM, R., GARG, P. M., HERCO, M., SHEKHAWAT, P. S. & CHEN, Y.-H. 2018. Necrotizing enterocolitis leads to disruption of tight junctions and increase in gut permeability in a mouse model. *BMC Pediatrics,* 18**,** 372.

REYMAN, M., VAN HOUTEN, M. A., VAN BAARLE, D., BOSCH, A., MAN, W. H., CHU, M., ARP, K., WATSON, R. L., SANDERS, E. A. M., FUENTES, S. & BOGAERT, D. 2019. Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. *Nat Commun,* 10**,** 4997.

RIVIÈRE, A., SELAK, M., LANTIN, D., LEROY, F. & DE VUYST, L. 2016. Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. *Frontiers in Microbiology,* 7.

RODRÍGUEZ-DÍAZ, J., MONEDERO, V. & YEBRA, M. J. 2011. Utilization of natural fucosylated oligosaccharides by three novel alpha-L-fucosidases from a probiotic Lactobacillus casei strain. *Appl Environ Microbiol,* 77**,** 703-5.

RODRIGUEZ, J. M., MURPHY, K., STANTON, C., ROSS, R. P., KOBER, O. I., JUGE, N., AVERSHINA, E., RUDI, K., NARBAD, A., JENMALM, M. C., MARCHESI, J. R. &

COLLADO, M. C. 2015. The composition of the gut microbiota throughout life, with an emphasis on early life. *Microb Ecol Health Dis,* 26**,** 26050.

ROGNES, T., FLOURI, T., NICHOLS, B., QUINCE, C. & MAHÉ, F. 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ,* 4**,** e2584.

ROUND, J. L. & MAZMANIAN, S. K. 2009. The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol,* 9**,** 313-23.

ROUSSEAUX, A., BROSSEAU, C., LE GALL, S., PILOQUET, H., BARBAROT, S. & BODINIER, M. 2021. Human Milk Oligosaccharides: Their Effects on the Host and Their Potential as Therapeutic Agents. *Front Immunol,* 12**,** 680911.

RUDLOFF, S., OBERMEIER, S., BORSCH, C., POHLENTZ, G., HARTMANN, R., BRÖSICKE, H., LENTZE, M. J. & KUNZ, C. 2006. Incorporation of orally applied (13)C-galactose into milk lactose and oligosaccharides. *Glycobiology,* 16**,** 477-87.

RUDLOFF, S., POHLENTZ, G., BORSCH, C., LENTZE, M. J. & KUNZ, C. 2012. Urinary excretion of in vivo ^{13}C-labelled milk oligosaccharides in breastfed infants. *Br J Nutr,* 107**,** 957-63.

RUIZ-PALACIOS, G. M., CERVANTES, L. E., RAMOS, P., CHAVEZ-MUNGUIA, B. & NEWBURG, D. S. 2003. Campylobacter jejuni binds intestinal H(O) antigen (Fuc alpha 1, 2Gal beta 1, 4GlcNAc), and fucosyloligosaccharides of human milk inhibit its binding and infection. *J Biol Chem,* 278**,** 14112-20.

SADAGHIAN SADABAD, M., VON MARTELS, J. Z. H., KHAN, M. T., BLOKZIJL, T., PAGLIA, G., DIJKSTRA, G., HARMSEN, H. J. M. & FABER, K. N. 2015. A simple coculture system shows mutualism between anaerobic faecalibacteria and epithelial Caco-2 cells. *Scientific Reports,* 5**,** 17906.

SAKANAKA, M., GOTOH, A., YOSHIDA, K., ODAMAKI, T., KOGUCHI, H., XIAO, J. Z., KITAOKA, M. & KATAYAMA, T. 2019a. Varied Pathways of Infant Gut-Associated Bifidobacterium to Assimilate Human Milk Oligosaccharides: Prevalence of the Gene Set and Its Correlation with Bifidobacteria-Rich Microbiota Formation. *Nutrients,* 12.

SAKANAKA, M., HANSEN, M. E., GOTOH, A., KATOH, T., YOSHIDA, K., ODAMAKI, T., YACHI, H., SUGIYAMA, Y., KURIHARA, S., HIROSE, J., URASHIMA, T., XIAO, J.-Z., KITAOKA, M., FUKIYA, S., YOKOTA, A., LO LEGGIO, L., ABOU HACHEM, M. & KATAYAMA, T. 2019b. Evolutionary adaptation in fucosyllactose uptake systems supports bifidobacteria-infant symbiosis. *Science Advances,* 5**,** eaaw7696.

SAKURAMA, H., KIYOHARA, M., WADA, J., HONDA, Y., YAMAGUCHI, M., FUKIYA, S., YOKOTA, A., ASHIDA, H., KUMAGAI, H., KITAOKA, M., YAMAMOTO, K.

& KATAYAMA, T. 2013. Lacto-N-biosidase encoded by a novel gene of Bifidobacterium longum subspecies longum shows unique substrate specificity and requires a designated chaperone for its active expression. *J Biol Chem,* 288**,** 25194-25206.

SALLI, K., HIRVONEN, J., SIITONEN, J., AHONEN, I., ANGLENIUS, H. & MAUKONEN, J. 2021. Selective Utilization of the Human Milk Oligosaccharides 2'-Fucosyllactose, 3-Fucosyllactose, and Difucosyllactose by Various Probiotic and Pathogenic Bacteria. *J Agric Food Chem,* 69**,** 170-182.

SAMARA, J., MOOSSAVI, S., ALSHAIKH, B., ORTEGA, V. A., PETTERSEN, V. K., FERDOUS, T., HOOPS, S. L., SORAISHAM, A., VAYALUMKAL, J., DERSCH-MILLS, D., GERBER, J. S., MUKHOPADHYAY, S., PUOPOLO, K., TOMPKINS, T. A., KNIGHTS, D., WALTER, J., AMIN, H. & ARRIETA, M.-C. 2022. Supplementation with a probiotic mixture accelerates gut microbiome maturation and reduces intestinal inflammation in extremely preterm infants. *Cell Host & Microbe,* 30**,** 696-711.e5.

SASAKI, N., MIYAMOTO, K., MASLOWSKI, K. M., OHNO, H., KANAI, T. & SATO, T. 2020. Development of a Scalable Coculture System for Gut Anaerobes and Human Colon Epithelium. *Gastroenterology,* 159**,** 388-390.e5.

SATO, T., VRIES, R. G., SNIPPERT, H. J., VAN DE WETERING, M., BARKER, N., STANGE, D. E., VAN ES, J. H., ABO, A., KUJALA, P., PETERS, P. J. & CLEVERS, H. 2009. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature,* 459**,** 262-5.

SCHWAB, C. & GÄNZLE, M. 2011. Lactic acid bacteria fermentation of human milk oligosaccharide components, human milk oligosaccharides and galactooligosaccharides. *FEMS Microbiology Letters,* 315**,** 141-148.

SEFEROVIC, M. D., MOHAMMAD, M., PACE, R. M., ENGEVIK, M., VERSALOVIC, J., BODE, L., HAYMOND, M. & AAGAARD, K. M. 2020. Maternal diet alters human milk oligosaccharide composition with implications for the milk metagenome. *Sci Rep,* 10**,** 22092.

SELA, D. A., CHAPMAN, J., ADEUYA, A., KIM, J. H., CHEN, F., WHITEHEAD, T. R., LAPIDUS, A., ROKHSAR, D. S., LEBRILLA, C. B., GERMAN, J. B., PRICE, N. P., RICHARDSON, P. M. & MILLS, D. A. 2008. The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome. *Proceedings of the National Academy of Sciences,* 105**,** 18964-18969.

SELA, D. A., GARRIDO, D., LERNO, L., WU, S., TAN, K., EOM, H. J., JOACHIMIAK, A., LEBRILLA, C. B. & MILLS, D. A. 2012. Bifidobacterium longum subsp. infantis ATCC 15697 α-fucosidases are active on fucosylated human milk oligosaccharides. *Appl Environ Microbiol,* 78**,** 795-803.

SENDER, R., FUCHS, S. & MILO, R. 2016. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol,* 14**,** e1002533.

SENGER, S., INGANO, L., FREIRE, R., ANSELMO, A., ZHU, W., SADREYEV, R., WALKER, W. A. & FASANO, A. 2018. Human Fetal-Derived Enterospheres Provide Insights on Intestinal Development and a Novel Model to Study Necrotizing Enterocolitis (NEC). *Cell Mol Gastroenterol Hepatol,* 5**,** 549-568.

SHAH, P., FRITZ, J. V., GLAAB, E., DESAI, M. S., GREENHALGH, K., FRACHET, A., NIEGOWSKA, M., ESTES, M., JÄGER, C., SEGUIN-DEVAUX, C., ZENHAUSERN, F. & WILMES, P. 2016. A microfluidics-based in vitro model of the gastrointestinal human–microbe interface. *Nature Communications,* 7**,** 11535.

SHANE, A. L., SÁNCHEZ, P. J. & STOLL, B. J. 2017. Neonatal sepsis. *The Lancet,* 390**,** 1770-1780.

SHAO, Y., FORSTER, S. C., TSALIKI, E., VERVIER, K., STRANG, A., SIMPSON, N., KUMAR, N., STARES, M. D., RODGER, A., BROCKLEHURST, P., FIELD, N. & LAWLEY, T. D. 2019. Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. *Nature,* 574**,** 117-121.

SHARMA, D., FARAHBAKHSH, N., SHASTRI, S. & SHARMA, P. 2018. Biomarkers for diagnosis of neonatal sepsis: a literature review. *The Journal of Maternal-Fetal & Neonatal Medicine,* 31**,** 1646-1659.

SHAW, A. G., SIM, K., RANDELL, P., COX, M. J., MCCLURE, Z. E., LI, M.-S., DONALDSON, H., LANGFORD, P. R., COOKSON, W. O. C. M., MOFFATT, M. F. & KROLL, J. S. 2015. Late-Onset Bloodstream Infection and Perturbed Maturation of the Gastrointestinal Microbiota in Premature Infants. *PLOS ONE,* 10**,** e0132923.

SHETE, V. B., GHADAGE, D. P., MULEY, V. A. & BHORE, A. V. 2009. Acinetobacter septicemia in neonates admitted to intensive care units. *J Lab Physicians,* 1**,** 73-6.

SHULHAN, J., DICKEN, B., HARTLING, L. & LARSEN, B. M. 2017. Current Knowledge of Necrotizing Enterocolitis in Preterm Infants and the Impact of Different Types of Enteral Nutrition Products. *Adv Nutr,* 8**,** 80-91.

SIM, K., SHAW, A. G., RANDELL, P., COX, M. J., MCCLURE, Z. E., LI, M. S., HADDAD, M., LANGFORD, P. R., COOKSON, W. O., MOFFATT, M. F. & KROLL, J. S. 2015.

Dysbiosis anticipating necrotizing enterocolitis in very premature infants. *Clin Infect Dis,* 60**,** 389-97.

SIMPSON, A. J., MAXWELL, A. I., GOVAN, J. R., HASLETT, C. & SALLENAVE, J. M. 1999. Elafin (elastase-specific inhibitor) has anti-microbial activity against gram-positive and gram-negative respiratory pathogens. *FEBS Lett,* 452**,** 309-13.

SMILOWITZ, J. T., LEBRILLA, C. B., MILLS, D. A., GERMAN, J. B. & FREEMAN, S. L. 2014. Breast milk oligosaccharides: structure-function relationships in the neonate. *Annu Rev Nutr,* 34**,** 143-69.

SODHI, C. P., WIPF, P., YAMAGUCHI, Y., FULTON, W. B., KOVLER, M., NINO, D. F., ZHOU, Q., BANFIELD, E., WERTS, A. D., LADD, M. R., BUCK, R. H., GOEHRING, K. C., PRINDLE, T., JR., WANG, S., JIA, H., LU, P. & HACKAM, D. J. 2020. The human milk oligosaccharides 2'-fucosyllactose and 6'-sialyllactose protect against the development of necrotizing enterocolitis by inhibiting toll-like receptor 4 signaling. *Pediatr Res*.

SOLIS, G., DE LOS REYES-GAVILAN, C. G., FERNANDEZ, N., MARGOLLES, A. & GUEIMONDE, M. 2010. Establishment and development of lactic acid bacteria and bifidobacteria microbiota in breast-milk and the infant gut. *Anaerobe,* 16**,** 307-10.

SONESON, C., LOVE, M. I. & ROBINSON, M. D. 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res,* 4**,** 1521.

STEWART, C. J. 2021. Breastfeeding promotes bifidobacterial immunomodulatory metabolites. *Nature Microbiology,* 6**,** 1335-1336.

STEWART, C. J., AJAMI, N. J., O'BRIEN, J. L., HUTCHINSON, D. S., SMITH, D. P., WONG, M. C., ROSS, M. C., LLOYD, R. E., DODDAPANENI, H., METCALF, G. A., MUZNY, D., GIBBS, R. A., VATANEN, T., HUTTENHOWER, C., XAVIER, R. J., REWERS, M., HAGOPIAN, W., TOPPARI, J., ZIEGLER, A. G., SHE, J. X., AKOLKAR, B., LERNMARK, A., HYOTY, H., VEHIK, K., KRISCHER, J. P. & PETROSINO, J. F. 2018. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature,* 562**,** 583-588.

STEWART, C. J., EMBLETON, N. D., CLEMENTS, E., LUNA, P. N., SMITH, D. P., FOFANOVA, T. Y., NELSON, A., TAYLOR, G., ORR, C. H., PETROSINO, J. F., BERRINGTON, J. E. & CUMMINGS, S. P. 2017a. Cesarean or Vaginal Birth Does Not Impact the Longitudinal Development of the Gut Microbiome in a Cohort of Exclusively Preterm Infants. *Front Microbiol,* 8**,** 1008.

STEWART, C. J., EMBLETON, N. D., MARRS, E. C., SMITH, D. P., NELSON, A., ABDULKADIR, B., SKEATH, T., PETROSINO, J. F., PERRY, J. D., BERRINGTON,

J. E. & CUMMINGS, S. P. 2016. Temporal bacterial and metabolic development of the preterm gut reveals specific signatures in health and disease. *Microbiome,* 4**,** 67.

STEWART, C. J., EMBLETON, N. D., MARRS, E. C. L., SMITH, D. P., FOFANOVA, T., NELSON, A., SKEATH, T., PERRY, J. D., PETROSINO, J. F., BERRINGTON, J. E. & CUMMINGS, S. P. 2017b. Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. *Microbiome,* 5**,** 75.

STEWART, C. J., ESTES, M. K. & RAMANI, S. 2020. Establishing Human Intestinal Enteroid/Organoid Lines from Preterm Infant and Adult Tissue. *Methods Mol Biol,* 2121**,** 185-198.

STEWART, C. J., FATEMIZADEH, R., PARSONS, P., LAMB, C. A., SHADY, D. A., PETROSINO, J. F. & HAIR, A. B. 2019. Using formalin fixed paraffin embedded tissue to characterize the preterm gut microbiota in necrotising enterocolitis and spontaneous isolated perforation using marginal and diseased tissue. *BMC Microbiol,* 19**,** 52.

STEWART, C. J., MARRS, E. C., MAGORRIAN, S., NELSON, A., LANYON, C., PERRY, J. D., EMBLETON, N. D., CUMMINGS, S. P. & BERRINGTON, J. E. 2012. The preterm gut microbiota: changes associated with necrotizing enterocolitis and infection. *Acta Paediatr,* 101**,** 1121-7.

STEWART, C. J., MARRS, E. C. L., NELSON, A., LANYON, C., PERRY, J. D., EMBLETON, N. D., CUMMINGS, S. P. & BERRINGTON, J. E. 2013a. Development of the Preterm Gut Microbiome in Twins at Risk of Necrotising Enterocolitis and Sepsis. *PLOS ONE,* 8**,** e73465.

STEWART, C. J., NELSON, A., SCRIBBINS, D., MARRS, E. C., LANYON, C., PERRY, J. D., EMBLETON, N. D., CUMMINGS, S. P. & BERRINGTON, J. E. 2013b. Bacterial and fungal viability in the preterm gut: NEC and sepsis. *Arch Dis Child Fetal Neonatal Ed,* 98**,** F298-303.

STOLL, B. J., HANSEN, N., FANAROFF, A. A., WRIGHT, L. L., CARLO, W. A., EHRENKRANZ, R. A., LEMONS, J. A., DONOVAN, E. F., STARK, A. R., TYSON, J. E., OH., W., BAUER, C. R., KORONES, S. B., SHANKARAN, S., LAPTOOK, A. R., STEVENSON, D. K., PAPILE, L.-A. & POOLE, W. K. 2002. Late-Onset Sepsis in Very Low Birth Weight Neonates: The Experience of the NICHD Neonatal Research Network. *Pediatrics,* 110**,** 285-291.

SULISTYO, A., RAHMAN, A., BIOUSS, G., ANTOUNIANS, L. & ZANI, A. 2018. Animal models of necrotizing enterocolitis: review of the literature and state of the art. *Innov Surg Sci,* 3**,** 87-92.

SUZUKI, R., WADA, J., KATAYAMA, T., FUSHINOBU, S., WAKAGI, T., SHOUN, H., SUGIMOTO, H., TANAKA, A., KUMAGAI, H., ASHIDA, H., KITAOKA, M. & YAMAMOTO, K. 2008. Structural and Thermodynamic Analyses of Solute-binding Protein from Bifidobacterium longum Specific for Core 1 Disaccharide and Lacto-N-biose I*. *Journal of Biological Chemistry,* 283**,** 13165-13173.

TAFT, D. H., AMBALAVANAN, N., SCHIBLER, K. R., YU, Z., NEWBURG, D. S., DESHMUKH, H., WARD, D. V. & MORROW, A. L. 2015. Center Variation in Intestinal Microbiota Prior to Late-Onset Sepsis in Preterm Infants. *PLOS ONE,* 10**,** e0130604.

TANNOCK, G. W., LAWLEY, B., MUNRO, K., GOWRI PATHMANATHAN, S., ZHOU, S. J., MAKRIDES, M., GIBSON, R. A., SULLIVAN, T., PROSSER, C. G., LOWRY, D. & HODGKINSON, A. J. 2013. Comparison of the compositions of the stool microbiotas of infants fed goat milk formula, cow milk-based formula, or breast milk. *Appl Environ Microbiol,* 79**,** 3040-8.

TEMPLE, M. J., CUSKIN, F., BASLÉ, A., HICKEY, N., SPECIALE, G., WILLIAMS, S. J., GILBERT, H. J. & LOWE, E. C. 2017. A Bacteroidetes locus dedicated to fungal 1,6-β-glucan degradation: Unique substrate conformation drives specificity of the key endo-1,6-β-glucanase. *J Biol Chem,* 292**,** 10639-10650.

THEVENOT, E. A., ROUX, A., XU, Y., EZAN, E. & JUNOT, C. 2015. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J Proteome Res,* 14**,** 3322-35.

THONGARAM, T., HOEFLINGER, J. L., CHOW, J. & MILLER, M. J. 2017. Human milk oligosaccharide consumption by probiotic and human-associated bifidobacteria and lactobacilli. *J Dairy Sci,* 100**,** 7825-7833.

TISSIER, H. 1900. *Recherches sur la flore intestinale des nourrissons : (état normal et pathologique).*

TORRAZZA, R. M., UKHANOVA, M., WANG, X., SHARMA, R., HUDAK, M. L., NEU, J. & MAI, V. 2013. Intestinal microbial ecology and environmental factors affecting necrotizing enterocolitis. *PLoS One,* 8**,** e83304.

TREMBLAY, É., FERRETTI, E., BABAKISSA, C., BURGHARDT, K. M., LEVY, E. & BEAULIEU, J.-F. 2021. IL-17-related signature genes linked to human necrotizing enterocolitis. *BMC Research Notes,* 14**,** 82.

TREMBLAY, É., THIBAULT, M.-P., FERRETTI, E., BABAKISSA, C., BERTELLE, V., BETTOLLI, M., BURGHARDT, K. M., COLOMBANI, J.-F., GRYNSPAN, D.,

LEVY, E., LU, P., MAYER, S., MÉNARD, D., MOUTERDE, O., RENES, I. B., SEIDMAN, E. G. & BEAULIEU, J.-F. 2016. Gene expression profiling in necrotizing enterocolitis reveals pathways common to those reported in Crohn's disease. *BMC Medical Genomics,* 9**,** 6.

TRUONG, D. T., FRANZOSA, E. A., TICKLE, T. L., SCHOLZ, M., WEINGART, G., PASOLLI, E., TETT, A., HUTTENHOWER, C. & SEGATA, N. 2015. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods,* 12**,** 902-3.

TSUI, K.-H., HOU, C.-P., CHANG, K.-S., LIN, Y.-H., FENG, T.-H., CHEN, C.-C., SHIN, Y.-S. & JUANG, H.-H. 2019. Metallothionein 3 Is a Hypoxia-Upregulated Oncogene Enhancing Cell Invasion and Tumorigenesis in Human Bladder Carcinoma Cells. *International journal of molecular sciences,* 20**,** 980.

TSUKUDA, N., YAHAGI, K., HARA, T., WATANABE, Y., MATSUMOTO, H., MORI, H., HIGASHI, K., TSUJI, H., MATSUMOTO, S., KUROKAWA, K. & MATSUKI, T. 2021. Key bacterial taxa and metabolic pathways affecting gut short-chain fatty acid profiles in early life. *The ISME Journal,* 15**,** 2574-2590.

ULLUWISHEWA, D., ANDERSON, R. C., YOUNG, W., MCNABB, W. C., VAN BAARLEN, P., MOUGHAN, P. J., WELLS, J. M. & ROY, N. C. 2015. Live Faecalibacterium prausnitzii in an apical anaerobic model of the intestinal epithelial barrier. *Cell Microbiol,* 17**,** 226-40.

URBANIAK, C., ANGELINI, M., GLOOR, G. B. & REID, G. 2016. Human milk microbiota profiles in relation to birthing method, gestation and infant gender. *Microbiome,* 4**,** 1.

VAN LEEUWEN, S. S. 2019. Challenges and Pitfalls in Human Milk Oligosaccharide Analysis. *Nutrients,* 11.

VAN NIEKERK, E., AUTRAN, C. A., NEL, D. G., KIRSTEN, G. F., BLAAUW, R. & BODE, L. 2014. Human milk oligosaccharides differ between HIV-infected and HIV-uninfected mothers and are related to necrotizing enterocolitis incidence in their preterm very-low-birth-weight infants. *J Nutr,* 144**,** 1227-33.

VAN WESTERING-KROON, E., HUIZING, M. J., VILLAMOR-MARTÍNEZ, E. & VILLAMOR, E. 2021. Male Disadvantage in Oxidative Stress-Associated Complications of Prematurity: A Systematic Review, Meta-Analysis and Meta-Regression. *Antioxidants (Basel),* 10.

VARMA SHRIVASTAV, S., BHARDWAJ, A., PATHAK, K. A. & SHRIVASTAV, A. 2020. Insulin-Like Growth Factor Binding Protein-3 (IGFBP-3): Unraveling the Role in Mediating IGF-Independent Effects Within the Cell. *Front Cell Dev Biol,* 8**,** 286.

VATANEN, T., FRANZOSA, E. A., SCHWAGER, R., TRIPATHI, S., ARTHUR, T. D., VEHIK, K., LERNMARK, A., HAGOPIAN, W. A., REWERS, M. J., SHE, J. X., TOPPARI, J., ZIEGLER, A. G., AKOLKAR, B., KRISCHER, J. P., STEWART, C. J., AJAMI, N. J., PETROSINO, J. F., GEVERS, D., LAHDESMAKI, H., VLAMAKIS, H., HUTTENHOWER, C. & XAVIER, R. J. 2018. The human gut microbiome in early-onset type 1 diabetes from the TEDDY study. *Nature,* 562**,** 589-594.

WADA, J., ANDO, T., KIYOHARA, M., ASHIDA, H., KITAOKA, M., YAMAGUCHI, M., KUMAGAI, H., KATAYAMA, T. & YAMAMOTO, K. 2008. Bifidobacterium bifidum lacto-N-biosidase, a critical enzyme for the degradation of human milk oligosaccharides with a type 1 structure. *Applied and environmental microbiology,* 74**,** 3996-4004.

WAMPACH, L., HEINTZ-BUSCHART, A., FRITZ, J. V., RAMIRO-GARCIA, J., HABIER, J., HEROLD, M., NARAYANASAMY, S., KAYSEN, A., HOGAN, A. H., BINDL, L., BOTTU, J., HALDER, R., SJÖQVIST, C., MAY, P., ANDERSSON, A. F., DE BEAUFORT, C. & WILMES, P. 2018. Birth mode is associated with earliest strain-conferred gut microbiome functions and immunostimulatory potential. *Nature Communications,* 9**,** 5091.

WANG, B., BRAND-MILLER, J., MCVEAGH, P. & PETOCZ, P. 2001. Concentration and distribution of sialic acid in human milk and infant formulas. *Am J Clin Nutr,* 74**,** 510-5.

WANG, Y., HOENIG, J. D., MALIN, K. J., QAMAR, S., PETROF, E. O., SUN, J., ANTONOPOULOS, D. A., CHANG, E. B. & CLAUD, E. C. 2009. 16S rRNA gene-based analysis of fecal microbiota from preterm infants with and without necrotizing enterocolitis. *ISME J,* 3**,** 944-54.

WANG, Y., JIANG, M., YAO, Y. & CAI, Z. 2018. WWC3 Inhibits Glioma Cell Proliferation Through Suppressing the Wnt/β-Catenin Signaling Pathway. *DNA Cell Biol,* 37**,** 31-37.

WANG, Y., ZOU, Y., WANG, J., MA, H., ZHANG, B. & WANG, S. 2020. The Protective Effects of 2'-Fucosyllactose against E. Coli O157 Infection Are Mediated by the Regulation of Gut Microbiota and the Inhibition of Pathogen Adhesion. *Nutrients,* 12**,** 1284.

WARNER, B. B., DEYCH, E., ZHOU, Y., HALL-MOORE, C., WEINSTOCK, G. M., SODERGREN, E., SHAIKH, N., HOFFMANN, J. A., LINNEMAN, L. A., HAMVAS, A., KHANNA, G., ROUGGLY-NICKLESS, L. C., NDAO, I. M., SHANDS, B. A., ESCOBEDO, M., SULLIVAN, J. E., RADMACHER, P. G., SHANNON, W. D. &

TARR, P. I. 2016. Gut bacteria dysbiosis and necrotising enterocolitis in very low birthweight infants: a prospective case-control study. *Lancet,* 387**,** 1928-36.

WEJRYD, E., MARTI, M., MARCHINI, G., WERME, A., JONSSON, B., LANDBERG, E. & ABRAHAMSSON, T. R. 2018. Low Diversity of Human Milk Oligosaccharides is Associated with Necrotising Enterocolitis in Extremely Low Birth Weight Infants. *Nutrients,* 10.

WHITEHOUSE, J. S., RIGGLE, K. M., PURPI, D. P., MAYER, A. N., PRITCHARD, K. A., JR., OLDHAM, K. T. & GOURLAY, D. M. 2010. The protective role of intestinal alkaline phosphatase in necrotizing enterocolitis. *J Surg Res,* 163**,** 79-85.

WHO. 2023. *https://www.who.int/news-room/fact-sheets/detail/preterm-birth* [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/preterm-birth [Accessed].

WICKRAMASINGHE, S., PACHECO, A. R., LEMAY, D. G. & MILLS, D. A. 2015. Bifidobacteria grown on human milk oligosaccharides downregulate the expression of inflammation-related genes in Caco-2 cells. *BMC Microbiol,* 15**,** 172.

WOLKE, D., JOHNSON, S. & MENDONÇA, M. 2019. The Life Course Consequences of Very Preterm Birth. *Annual Review of Developmental Psychology,* 1**,** 69-92.

WU, R. Y., LI, B., HORNE, R. G., AHMED, A., LEE, D., ROBINSON, S. C., ZHU, H., CADETE, M., ALGANABI, M., FILLER, R., JOHNSON-HENRY, K. C., DELGADO-OLGUIN, P., PIERRO, A. & SHERMAN, P. M. 2022. Structure–Function Relationships of Human Milk Oligosaccharides on the Intestinal Epithelial Transcriptome in Caco-2 Cells and a Murine Model of Necrotizing Enterocolitis. *Molecular Nutrition & Food Research,* 66**,** 2100893.

WU, R. Y., LI, B., KOIKE, Y., MAATTANEN, P., MIYAKE, H., CADETE, M., JOHNSON-HENRY, K. C., BOTTS, S. R., LEE, C., ABRAHAMSSON, T. R., LANDBERG, E., PIERRO, A. & SHERMAN, P. M. 2019. Human Milk Oligosaccharides Increase Mucin Expression in Experimental Necrotizing Enterocolitis. *Mol Nutr Food Res,* 63**,** e1800658.

XAVIER, A. A. O., DIAZ-SALIDO, E., ARENILLA-VELEZ, I., AGUAYO-MALDONADO, J., GARRIDO-FERNANDEZ, J., FONTECHA, J., SANCHEZ-GARCIA, A. & PEREZ-GALVEZ, A. 2018. Carotenoid Content in Human Colostrum is Associated to Preterm/Full-Term Birth Condition. *Nutrients,* 10.

YAMASAKI, M., NOMURA, T., SATO, F. & MIMATA, H. 2007. Metallothionein is up-regulated under hypoxia and promotes the survival of human prostate cancer cells. *Oncol Rep,* 18**,** 1145-53.

YOSHIDA, E., SAKURAMA, H., KIYOHARA, M., NAKAJIMA, M., KITAOKA, M., ASHIDA, H., HIROSE, J., KATAYAMA, T., YAMAMOTO, K. & KUMAGAI, H. 2011. Bifidobacterium longum subsp. infantis uses two different β-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides. *Glycobiology,* 22**,** 361-368.

YU, Z. T., CHEN, C. & NEWBURG, D. S. 2013. Utilization of major fucosylated and sialylated human milk oligosaccharides by isolated human gut microbes. *Glycobiology,* 23**,** 1281-92.

ZABEL, B. E., GERDES, S., EVANS, K. C., NEDVECK, D., SINGLES, S. K., VOLK, B. & BUDINOFF, C. 2020. Strain-specific strategies of 2'-fucosyllactose, 3-fucosyllactose, and difucosyllactose assimilation by Bifidobacterium longum subsp. infantis Bi-26 and ATCC 15697. *Sci Rep,* 10**,** 15919.

ZARRILLI, R., BAGATTINI, M., ESPOSITO, E. P. & TRIASSI, M. 2018. Acinetobacter Infections in Neonates. *Current Infectious Disease Reports,* 20**,** 48.

ZEUNER, B., TEZE, D., MUSCHIOL, J. & MEYER, A. S. 2019. Synthesis of Human Milk Oligosaccharides: Protein Engineering Strategies for Improved Enzymatic Transglycosylation. *Molecules,* 24.

ZHANG, G., MILLS, D. A. & BLOCK, D. E. 2009. Development of Chemically Defined Media Supporting High-Cell-Density Growth of Lactococci, Enterococci, and Streptococci. *Applied and Environmental Microbiology,* 75**,** 1080-1087.

ZHANG, W., HE-YANG, J., ZHUANG, W., LIU, J. & ZHOU, X. 2021. Causative role of mast cell and mast cell-regulatory function of disialyllacto-N-tetraose in necrotizing enterocolitis. *International Immunopharmacology,* 96**,** 107597.

ZHENG, N., GAO, Y., ZHU, W., MENG, D. & WALKER, W. A. 2020. Short chain fatty acids produced by colonizing intestinal commensal bacterial interaction with expressed breast milk are anti-inflammatory in human immature enterocytes. *PLoS One,* 15**,** e0229283.

ZUURVELD, M., VAN WITZENBURG, N. P., GARSSEN, J., FOLKERTS, G., STAHL, B., VAN'T LAND, B. & WILLEMSEN, L. E. M. 2020. Immunomodulation by Human Milk Oligosaccharides: The Potential Role in Prevention of Allergic Diseases. *Frontiers in immunology,* 11**,** 801-801.