PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
SCHOOL OF ENGINEERING

MOLECULAR STRATEGIES AND INTERACTIONS BETWEEN REPRESENTATIVE BACTERIA OF THE HUMAN GUT MICROBIOTA DURING THE UTILIZATION OF GLYCOMACROPEPTIDE

KEVIN JOSÉ GONZÁLEZ MORELO

Thesis submitted to the Office of Graduate Studies in partial fulfillment of the requirements for the Degree of Doctor in Engineering Sciences

Advisor:
DANIEL GARRIDO CORTÉS

Santiago de Chile, January, 2024

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
SCHOOL OF ENGINEERING

MOLECULAR STRATEGIES AND INTERACTIONS BETWEEN REPRESENTATIVE BACTERIA OF THE HUMAN GUT MICROBIOTA DURING THE UTILIZATION OF GLYCOMACROPEPTIDE

KEVIN JOSÉ GONZÁLEZ MORELO

Members of the Committee:

DANIEL GARRIDO

CÉSAR RAMÍREZ

FRANCO PEDRESCHI

JUAN UGALDE

ERNESTO PEREZ-RUEDA

IGNACIO LIRA

Thesis submitted to the Office of Graduate Studies in partial fulfillment of the requirements for the Degree of Doctor in Engineering Sciences

Santiago de Chile, January, 2024

To my parents, sisters, my wife Angye, and future generations of my entire family.

ACKNOWLEGMENTS

Firstly, I would like to express my sincere gratitude for the unwavering support and encouragement provided by God, my parents, my sisters Gabriela, Angie, and my wife Angye in writing this thesis.

In addition, I would like to extend my appreciation to Professor Daniel Garrido Cortes for his trust in me and for his guidance during the development of my doctoral research, laboratory training, and further teaching experience.

I would like to express my gratitude to my colleagues in the laboratory, including Marco Vega, Kineret Serebrinsky, Felipe Melis, Belén Hirmas, Alexis Torres, Tomas Pröschle, Aline Ovalle, and Romina Diaz, for their unwavering support throughout my doctoral studies.

I would like to express my gratitude to Professors Ernesto Perez-Rueda and Edgardo Galan from the Universidad Autónoma de Mexico (UNAM) for their mentorship during my internship in their research group.

Last but not least, I would like to express my gratitude to the Colombian research group IDAA integrated by Professor Diofanor Acevedo, Piedad M. Montero, Diego F. Tirado, Jose Torres, Katherine Paternina, Luis Gallo, and Edinson Castillo for their advice and knowledge.

Additionally, I extend my appreciation to the Vicerrectoría de Investigación Scholarship (VRI) of the Pontificia Universidad Católica de Chile, the ANID National Doctorate Scholarship 2019, the Fondecyt N° 1190074, and Fondecyt N°1230764 for funding my doctoral research.

LIST OF PAPERS

The thesis is based on the following papers, referred to in the text by the respective chapters:

CHAPTER I: *O-LINKED* GLYCAN UTILIZATION BY HUMAN GUT MICROBES: González-Morelo, K. J., Vega-Sagardía, M., & Garrido, D. (2020). **Molecular insights into *O*-linked glycan utilization by gut microbes.** Frontiers in Microbiology, 11, 591568.

CHAPTER II: STRUCTURE OF CO-EXPRESSION NETWORKS OF *BIFIDOBACTERIUM* SPECIES IN RESPONSE TO HUMAN MILK OLIGOSACCHARIDES. González-Morelo, K. J., Galán-Vásquez, E., Melis, F., Pérez-Rueda, E., & Garrido, D. (2023). Frontiers in Molecular Biosciences**,** 10, 1040721.

CHAPTER III: CROSS-FEEDING INTERACTIONS OF GUT MICROBES MEDIATED FROM *O-LINKED* GLYCANS CASEIN GLYCOMACROPEPTIDE. González-Morelo, K. J., Garrido, D. (2023). Manuscript submitted for publication.

PROCEEDINGS

Parts of this work have been presented at international congresses under the following references:

González-Morelo, K. J., Garrido, D.
Utilization of glycomacropeptide-derived O-glycans by representative members of the gut microbiota.
In: XLII Congreso Digital SOMICH 2020 (Oral Presentation)
December 2020.

González-Morelo, K. J., Garrido, D.
Metabolic interactions between different members of the human gut microbiota using glycomacropeptide,
In: ASM Microbe 2022, (Poster Presentation)
Washington D.C., EEUU., 2022.

González-Morelo, K. J., Galán-Vásquez, E., Melis, F., Pérez-Rueda, E., & Garrido, D.
Structure of co-expression networks of Bifidobacterium species in response to human milk oligosaccharides,
In: XLIV Congreso Chileno de Microbiología (Poster Presentation)
La Serena, Chile, November 2022.

González-Morelo, K. J., Garrido, D.
Metabolic interactions and gene expression between members of gut microbiota using glycomacropeptide,
In: ASM Microbe 2023 (Oral and Poster Presentation)
Houston, Texas, EEUU., 2023.

FIGURE INDEX

SUPPLEMENTARY FIGURE INDEX

TABLE INDEX

SUPPLEMENTARY TABLE INDEX

Page

LIST OF CONTENT

ABBREVIATIONS

ABC transport: ATP-binding cassette transport

Asn: Asparagine.

CCF: Commensal Colonization Factor.

cDNA: Complementary DNA.

Endo-α-GalNAcase: endo-α-*N*-Acetylgalactosaminidase.

FL: Fucosyllactose.

FOS: Fructooligosaccharides.

Fuc: Fucose.

Gal: Galactose

GalNAc: *N*-Acetylgalactosamine.

GF: Germ-free.

GH: Glycosyl hydrolase.

Glc: Glucosamine

GlcNAc: *N*-Acetylglucosamine.

GMP: Glycomacropeptide.

Gn: *N*-Acetylgalactosamine.

GNB: Galacto-*N*-biose.

GOS: Galactooligosaccharides.

HMM: Hidden Markov model.

HMO: Human Milk Oligosaccharides.

HPLC: high-performance liquid chromatography

IBD: Inflammatory Bowel Disease.

IL: Interleukins.

KEEG: Kyoto Encyclopedia of Genes and Genomes.

Lac: Lactose

LacNAc: *N*-acetyllactosamine.

LNB: Lacto-*N*-biose.

LNnT: lacto-*N*-neotetraose

LNT: lacto-*N*-tetraose.

Man: Mannose.

MRS: Man, Rogosa and Sharpe medium.

MUC: mucin.

mZMB: Modified ZMB culture.

NeuAc: Sialic acid/*N*-Acetylneuraminic acid.

NTC: Negative Treatment Control.

OD: Optical density

PUL: Polysaccharide Utilization System.

RCM: Reinforced clostridial medium.

RNA-seq: RNA sequencing.

SCFA: Short chain fatty acids.

Ser: Serine.

SL: sialyllactose.

Sup: Supernatant.

Sus: Starch Utilization System.

T: Transferase.

TF: Transcriptional Factor.

Thr: threonine.

TNF: Tumor necrosis factor.

TOM: Topological Overlap Measure.

UC: Ulcerative Colitis.

VNTR: Variable Number of Tandem Repeats.

WGCNA: Weighted gene co-expression network analysis.

ABSTRACT

The human gut microbiota plays an essential role in metabolizing complex dietary compounds that host enzymes cannot degrade. The prebiotic role of milk glycans, such as free human milk oligosaccharides (HMO), currently included in milk formulas to stimulate a healthy gut microbiota, has recently been validated. In addition, *O*-glycans are similarly complex oligosaccharides commonly found in intestinal mucins and milk proteins. While large extensions of *O*-glycans protect the epithelium cells, some gut microbes have been described to utilize them as carbon sources for colonization. *O*-glycans have been proposed as emerging prebiotics due to their remarkable similarity to host-associated glycans and HMO. Glycromacropeptide (GMP) is an *O*-glycosylated peptide obtained from whey during cheese manufacture. GMP could be considered a simple model of *O*-glycans to analyze the molecular mechanisms involved in metabolic interactions between gut microbes. This doctoral work aimed to determine the molecular strategies and microbial interactions while utilizing GMP as a carbon source. For this reason, this work pursues three specific objectives: i) To study co-expression networks in human milk oligosaccharides in *Bifidobacterium* species, ii) To study the mechanism of GMP *O*-glycan utilization by bacteria representative of the adult intestinal microbiome, and iii) To determine the metabolic cross-feeding interactions between representative bacteria of the human gut microbiome using GMP as a source of *O*-glycans.

Chapter I critically discusses the different molecular strategies of *O*-glycan consumption by prominent bacteria of the gut microbiota currently available in the literature. Certain gut microbes can access, release, and consume *O*-linked glycans as a carbon source. Among these, *Bifidobacterium bifidum* and *Bacteroides thetaiotaomicron* are prominent *O*-linked glycan utilizers. Their consumption strategies include specialized *α*-fucosidases, α-sialidases, and endo-*α*-*N*-

Acetylgalactosaminidases that release galacto-*N*-biose (GNB) from peptides backbones. *O*-linked glycan utilization by certain gut microbes represents an important niche that allows them to predominate and modulate host responses such as inflammation.

Chapter II focuses on the gene regulation of HMO to observe possible differences in gene expression architecture among Bifidobacterium species, as these oligosaccharides are chemically related to *O*-linked glycans. This chapter highlights *in silico* results from the co-expression networks obtained through WGCNA. This powerful systems biology methodology allows a better understanding of these oligosaccharides' consumption and gene expression. RNA-seq data obtained from Geo Datasets were obtained for *Bifidobacterium longum* subsp. *infantis, Bifidobacterium bifidum and Bifidobacterium longum* subsp. *longum*. Between 10 and 20 co-expressing modules were obtained for each dataset. HMO-associated genes appeared in the modules with more genes for *B. infantis* and *B. bifidum*, in contrast with *B. longum*. Hub genes were identified in each module and generally participated in conserved essential processes. Certain modules were differentially enriched with LacI-like transcription factors, and others with specific metabolic pathways, such as the biosynthesis of secondary metabolites. The three *Bifidobacterium* transcriptomes showed distinct regulation patterns for HMO utilization. HMO-associated genes in *B. infantis* co-expressed in two modules according to their participation in galactose or *N*-Acetylglucosamine utilization. Instead, *B. bifidum* showed a less structured co-expression of genes participating in HMO utilization. Finally, this category of genes in *B. longum* clustered in a small module, indicating a lack of co-expression with main cell processes, and suggesting a recent acquisition.

Chapter III highlights complex strategies for utilizing *O*-glycans in GMP consumption among gut microbes based on cross-feeding and competition. Individual cultures of representative bacteria allowed the identification of the major GMP-degraders. Unidirectional assays identified galacto-*N*-biose, galactose, *N*-

Acetylgalactosamine, and sialic acid as by-products, providing a perspective on microbial interactions during GMP fermentation. Bidirectional assays demonstrated cross-feeding activity, competition between gut microbes, and promotion of butyrate from the fatty acids derived from GMP. *O*-glycan-specific enzyme expression was identified for *B. infantis* ATCC 15697 and *B. bifidum* JCM 1254 during GMP cross-feeding consumption. This study highlights complex strategies for utilizing *o*-glycans in GMP consumption among gut microbes based on cross-feeding and competition.

KEYWORDS: Gut microbiota, glycomacropeptide, *O*-glycans.

RESUMEN

La microbiota intestinal humana desempeña un papel esencial en el metabolismo de compuestos complejos de la dieta que las enzimas del huésped no pueden degradar,. El rol prebiótico de los glicanos de la leche tales como los oligosacáridos de la leche materna (HMO), actualmente incluidos en fórmulas lácteas para estimular una microbiota intestinal saludable ha sido recientemente validado. Asimismo, los *O*-glicanos son oligosacáridos similarmente complejos que se encuentran habitualmente en las mucinas intestinales y proteínas lácteas. Mientras las grandes extensiones de *O*-glicanos protegen las células del epitelio, se ha descrito que algunos microbios intestinales los utilizan como fuentes de carbono para su colonización. Los *O*-glicanos se han propuesto como prebióticos emergentes debido a su notable similitud con los glicanos asociados al huésped y a los HMO. El glicomacropéptido (GMP) es un péptido *O*-glicosilado que se obtiene del suero durante la fabricación del queso. El GMP podría considerarse un modelo simple de *O*-glicanos para analizar los mecanismos moleculares implicados en las interacciones metabólicas entre los microbios intestinales. Este trabajo doctoral tiene como objetivo determinar las estrategias moleculares y las interacciones microbianas al utilizar GMP como fuente de carbono. Para ello, este trabajo persigue tres objetivos específicos: i) Estudiar las redes de co-expresión en oligosacáridos de leche humana en especies de *Bifidobacterium*, ii) Estudiar el mecanismo de utilización de *O*-glicanos de GMP por bacterias representativas de la microbiota intestinal adulta, y iii) Determinar las interacciones metabólicas de alimentación cruzada entre bacterias representativas de la microbiota intestinal humana utilizando GMP como fuente de *O*-glicanos. El Capítulo I discute críticamente las diferentes estrategias moleculares del consumo de *O*-glicanos por bacterias prominentes de la microbiota intestinal. Ciertos microbios intestinales pueden acceder, liberar y consumir *O*-glicanos como fuente de carbono. Entre estas, *Bifidobacterium bifidum*, y *Bacteroides thetaiotaomicron* son prominentes utilizadores de *O*-glicanos. Sus estrategias de consumo incluyen *α*-fucosidasas especializadas y α-sialidasas, además de endo-*α*-*N*-Acetylgalactosaminidasas que liberan galacto-*N*-biosa (GNB) de estructuras peptídicas. La utilización de *O*-glicanos por ciertos microbios intestinales representa un importante nicho que les permite predominar y modular las respuestas del huésped tales como la inflamación.

El Capitulo II está enfocado en la regulación génica de los HMO con la finalidad de observar diferencias en la arquitectura de expresión de genes entre especies de *Bifidobacterium*, siendo estos oligosacáridos químicamente relacionados a *O*-glicanos. Este capítulo destaca resultados *in silico* de redes de co-expresión obtenidas mediante WGCNA, una poderosa metodología de biología de sistemas que permite un mejor entendimiento del consumo de estos oligosacáridos y su expresión génica. Se obtuvieron datos de secuenciación de ARN de la Geo Datasets para *Bifidobacterium longum* subsp. *infantis*, *Bifidobacterium bifidum* y *Bifidobacterium longum* subsp. *longum*. Se obtuvieron entre 10 y 20 módulos de co-expresión para cada conjunto de datos. Los genes asociados a HMO aparecieron en los módulos con más genes para *B. infantis* y *B. bifidum*, en contraste con *B. longum*. En cada módulo se identificaron genes centrales, que en general participaban en procesos esenciales conservados. Ciertos módulos estaban diferencialmente enriquecidos con factores de transcripción tipo LacI, y otros con ciertas vías metabólicas como la biosíntesis de metabolitos secundarios. Los tres transcriptomas de *Bifidobacterium* mostraron patrones de regulación distintos para la utilización de HMO. Los genes asociados a HMO en *B. infantis* se co-expresaron en dos módulos según su participación en la utilización de galactosa o *N*-Acetilglucosamina. En cambio, *B. bifidum* mostró una co-expresión menos estructurada de los genes que participan en la utilización de HMO. Por último, esta categoría de genes en *B. longum* se agrupó en un pequeño módulo, lo que indica una falta de co-expresión con los principales procesos celulares, y sugiere una adquisición reciente.

El Capitulo III destaca las complejas estrategias para la utilización de O-glicanos en el consumo de GMP entre microbios intestinales, basado en comportamientos de alimentación cruzada y competencia. Los cultivos individuales de bacterias representativas permitieron identificar los principales degradadores de GMP. Los ensayos unidireccionales identificaron galacto-*N*-biosa, galactosa, *N*-Acetilgalactosamina y ácido siálico como subproductos, proporcionando una perspectiva sobre las interacciones microbianas durante la fermentación de GMP. Los ensayos bidireccionales demostraron la actividad de alimentación cruzada y la competencia entre los microbios intestinales, además de la promoción del butirato a partir de los ácidos grasos derivados del uso de GMP. Se identificó la expresión de enzimas especificas a *O*-glicanos para *B. infantis* ATCC 15697 y *B. Bifidum* JCM 1254 durante el

consumo cruzado de GMP. Este estudio pone de relieve estrategias complejas de utilización de *O*-glicanos en el consumo de GMP entre microbios intestinales, basadas en la alimentación cruzada y la competencia.

KEYWORDS: Microbiota intestinal, glicomacropéptido, *O*-glicanos.

1. INTRODUCTION

1.1 HUMAN GUT MICROBIOTA

The human gut microbiota is an extensive assemblage of microorganisms that colonize the human gastrointestinal tract (Ruan et al., 2020). This ecological consortium covers a spectrum of life forms from the prokaryotic domain and, to a lesser extent, includes fungi, parasites, and archaea (Berg et al., 2020). Viruses also constitute integral components of this intricate ecosystem (Lecuit & Eloit, 2017). The composite genetic and functional profile of these microbial entities collectively is referred to as the gut microbiome (Hills et al., 2019). From a broader perspective, it is conceivable to categorize individuals based on their enterotypes (Arumugam et al., 2011; Bresser et al., 2022), which represent distinct gut microbial compositions commonly observed within specific human populations.

Within a single individual, the aggregated gut microbial species have been demonstrated to possess a genetic repertoire of approximately 3.3 million genes (Zhu et al., 2010). When juxtaposed against the 23,000 genes encoded within the human genome, this genetic reservoir underscores these microbial inhabitants' magnitude and potential impact on human health (Al Bander et al., 2020). Notably, it is of consequence to recognize that the human body harbors nearly equivalent quantities of bacterial cells as human cells (Dekaboruah et al., 2020).

Under specific circumstances, alterations occur within the communities comprising the gut microbiota, encompassing their composition and bacterial diversity changes. These perturbations are recognized as dysbiosis, a term elucidated by Bernstein & Forbes (2017). Dysbiosis can be instigated by factors such as antibiotic consumption or adopting a low-fiber diet, which diminishes microbial diversity. Studies have demonstrated that alterations in the intestinal microenvironment, brought about by variations in metabolite profiles or shifts in pH levels (Sanders et al., 2021), lead to an escalation in the prevalence of opportunistic pathogens. Consequently, this intricate consortium of bacteria constituting the gut microbiome plays a pivotal role in shaping the host's state of health and may contribute to developing specific diseases (Nagpal et al., 2016).

1.2 CONVENTIONAL PREBIOTICS

The structure and activities of the human gut microbiota can be modulated by prebiotics (Peredo-Lovillo et al., 2020). A recent elucidation of the prebiotic concept aligns with substrates selectively utilized by host microorganisms, conferring health advantages upon the host (Gibson et al., 2017). The concept includes three essential parts: a substance, a physiologically beneficial effect, and a microbiota-mediated mechanism. According to the concept and literature, certain polysaccharides and oligosaccharides with a beneficial function on the host have been recognized as prebiotics (Davani-Davari et al., 2019). Among the benefits of consuming prebiotics are a reduced presence of pathogenic microorganisms (Gibson et al., 2005), stimulation of immune system activity (Shokryazdan et al., 2017), and a lower incidence of allergies (Brosseau et al., 2019).

These beneficial effects can, in part, be attributed to the capacity of prebiotics to be fermented by gut microbes (Rivière et al., 2016). The prebiotic effect may be enhanced by the series of metabolic interactions by degradation and fermentation products that may take place between gut microbes (Scott et al., 2013). These metabolic interactions can be through cross-feeding, defined as the degradation by-product of one species being consumed by another, and the opposite case refers to competitiveness for the substrate. These degradation by-products include sugars and short-chain fatty acids. (SCFA) (Saa et al., 2022).

Under their selective fermentability, prebiotics induce distinct alterations in the composition and activity of the gastrointestinal microbiota, with subsequent responses in host health. In a broader context, the repertoire of substances recognized as conventional prebiotics encompasses, but is not exhaustively restricted to, inulin, fructooligosaccharides, and galactooligosaccharides (Rawi et al., 2020).

1.2.1 Human Milk Oligosaccharides (HMO)

Human milk oligosaccharides (HMO) are also considered prebiotics because of their resistance to acidic pH and enzymatic degradation, as well as their ability to reach the large intestine intact to support microbial growth (Okburan & Kızıler, 2023). HMO constitute a group of bioactive carbohydrates typically composed of 3 to 10 monosaccharide units or sugar residues (Figure 1) (Bode, 2012). Following lactose and lipids, they represent the third most abundant solid component in human milk (Goehring et al., 2014). HMO concentrations

typically range from 5 to 23 g/L on average, with the highest levels, approximately 20–30 g/L, found in colostrum and 10–15 g/L in mature milk (Marx et al., 2014).

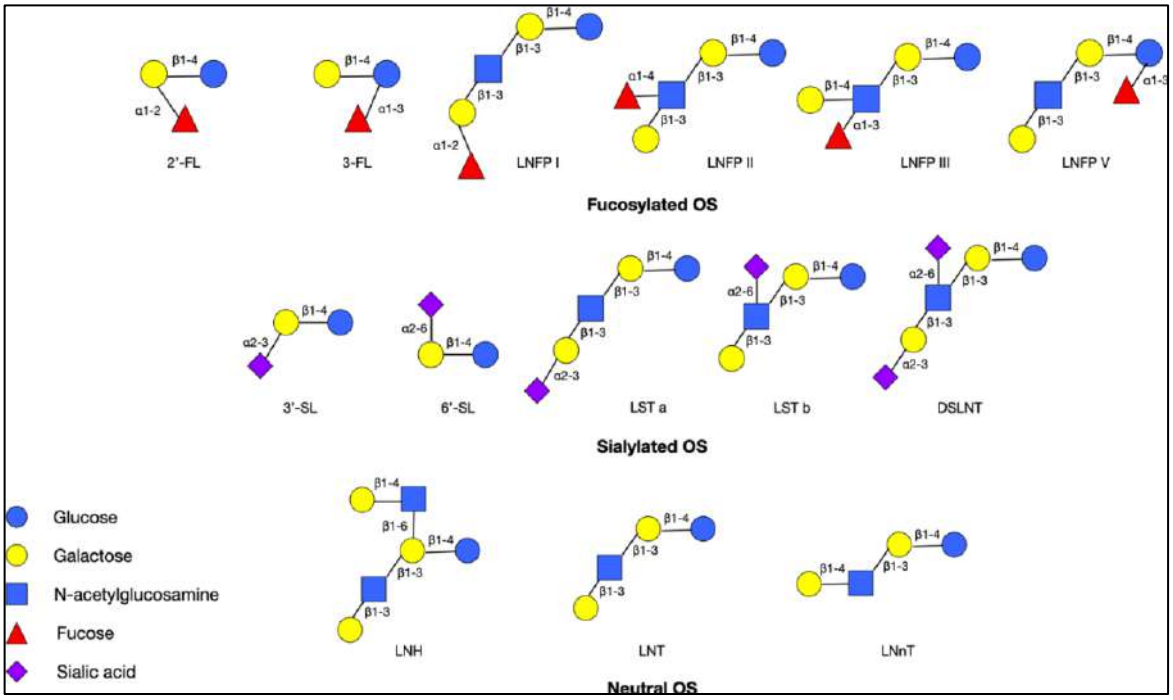

Figure 1- Schematization of main HMO. Source: (Zhang et al., 2021).

Five fundamental monosaccharides, namely D-glucose (Glu), D-galactose (Gal), *N*-Acetylglucosamine (GlcNAc), *L*-fucose (Fuc), and *N*-Acetylneuraminic acid (NeuAc) constitute the core building blocks of all HMO (Oliveira et al., 2015). With few exceptions, nearly all HMO share a common core structure featuring lactose (Galβ1-4Glu) at the reducing end, extended by the addition of lacto-*N*-biose units (Gal *β*1-3GlcNAc, LNB, type 1) or lactosamine units (Gal *β*1-4GlcNAc, LacNAc, type 2) through glycosyltransferases (Auer et al., 2021). The core structure can undergo further modifications such as fucosylation, sialylation through the addition of fucose in *α*-(1–2), *α*-(1–3), and *α*-(1–4) linkages, sialic acids in *α*-(2–3) or *α*-(2–6) linkages (Mendis & Jackson, 2022). These modification processes not only introduce structural variations in HMO but also categorize them into two groups: neutral oligosaccharides, which include fucosylated carbohydrates and are more abundant in human milk, and acid oligosaccharides, containing either sialic acids (Thurl et al., 2017).

Despite their structural complexity and diversity, HMO have been recognized as growth factors for "*bifidus flora"* in breast-fed infants (Borewicz et al., 2019). Considering the mutualistic interactions encompassed by the host and the breastfed newborn in the first years of life, the infant gut microbiota is highly colonized by *Bifidobacterium* genus members as microbes adapted to HMO for utilization. For example, *Bifidobacterium longum* subsp. *infantis* (*B. infantis*) is a prototypical bifidobacterial species that internalizes intact HMO through ABC transporters, broken down into monomers for intracellular metabolization (Garrido, Barile, et al., 2012). Into its genome, *B. infantis* ATCC 15697 is characterized by at least four HMO clusters, one devoted to type 1 HMO (LNB/GNB cluster), another for fucosylated and type 2 HMO consumption (HMO Cluster I), and two clusters targeting FL (Dedon et al., 2020; Sela et al., 2008; Zabel et al., 2020). In contrast, *B. bifidum* species produce extracellular enzymes that degrade the oligosaccharides and internalize the monomers (Kitaoka, 2012a). The subspecies *B. longum* subsp. *longum* (*B. longum*) can grow on various HMO structures, with a greater preference for fucosylated HMO. *B. longum* species are thought to be adapted to the adult gut microbiota, but some isolates also contain HMO clusters, allowing type 1 and FL consumption (Sakanaka et al., 2019).

Although it is well known that the World Health Organization (WHO) recommends breastfeeding babies for the first six months of life to provide health benefits to infants, breastfeeding and supplementation using formula milk with these molecules may not always be possible (Martin et al., 2016). Lately, alternatives such as fructooligosaccharides, galactooligosaccharides, and synthetic HMO, including 2FL and LNnT as commercially available for addition to infant formula, as well as glycans from bovine milk glycoconjugates are currently being studied (Thomson et al., 2018b).

1.3 GLYCANS FROM MILK GLYCOCONJUGATES

In addition to free oligosaccharides such as those found in HMO, glycoconjugates are another type of structure formed by glycosylation and composed of glycans (glycone) covalently attached to a non-carbohydrate moiety (aglycone). Within the glycoconjugates classes are glycoproteins, proteins that have glycans attached within their amino acidic

structure (Wandall et al., 2021). Among the glycoconjugates, *N*-(Asn)-oligosaccharides (*N*-glycans) and *O*-(Ser/Thr)-oligosaccharide (*O*-glycans) are found (Varki, 2017).

N-glycans are complex oligosaccharides that possess a standard pentasaccharide core sequence, present in all *N*-glycans: Manα1-6(Manα1-3)Manβ1-4GlcNacβ1-4GlcNacβ1-Asn-X-Ser/Thr. These structures can be found in whey, specifically in lactoferrins and immunoglobulins, and in smaller proportions, lactalbumin, lactoperoxidase, and antitrypsin (Guan et al., 2022).. These *N*-glycans can be categorized into three distinct types: 1) oligomannose, characterized by the addition of mannose residues to the core; 2) complex *N*-glycans, featuring antennas initiated by *N*-Acetylglucosamine appended to the core, with sialic acid often serving as the terminal residue (Zhang et al., 2023); and 3) hybrid *N*-glycans, where mannose residues are exclusively attached to the Manα1-6 arm of the core, and one or two GlcNAc units are added to the Manα1-3 arm (Hirata & Kizuka, 2021). Bacterial endo-*β*-*N*-Acetylglucosaminidases are enzymes that cleave the *N*,*N*′-diacetyl chitobiose moiety characteristic of all *N*-glycans (Mészáros et al., 2021).

For their part, *O*-glycans are another complex oligosaccharides characterized by an *N*-Acetylgalactosamine (GalNAc) residue linked to Ser/Thr as a common core (Figure 2). The elongation of *O*-linked glycans and their attachment to secreted proteins begins in the Golgi apparatus. A polypeptide-*N*-Acetyl-galactosaminyltransferase (ppGalNAcT) catalyzes the addition of a GalNAc*α*1 to a Ser/Thr available as a glycosylation site (Varki & Lowe, 2009). Several glycosyltransferases act in conjunction to attach sugar residues in single *O*-glycans. The Core 1 *O*-glycan (Gal*β*1-3GalNAcα1-Ser/Thr) is the first synthesized, and then the Core 2 unit [Galβ1-3(GlcNAcβ1,6)GalNAcα1-Ser/Thr; Itano, (2019)].

In milk proteins, *O*-linked glycans are a significant component of immunoglobulins and *κ*-caseins (Magnelli et al., 2011). These *O*-glycosylated proteins have a higher resistance to proteolysis due to their structural complexity, which also provides the physical function of gelation (Boutrou et al., 2008; Kesimer & Sheehan, 2012). For utilization as a carbon source, bacterial endo-*α*-*N*-Acetylgalactosaminidases are endoglycosidases that have been involved in releasing intact *O*-glycans from Ser/Thr linkage.

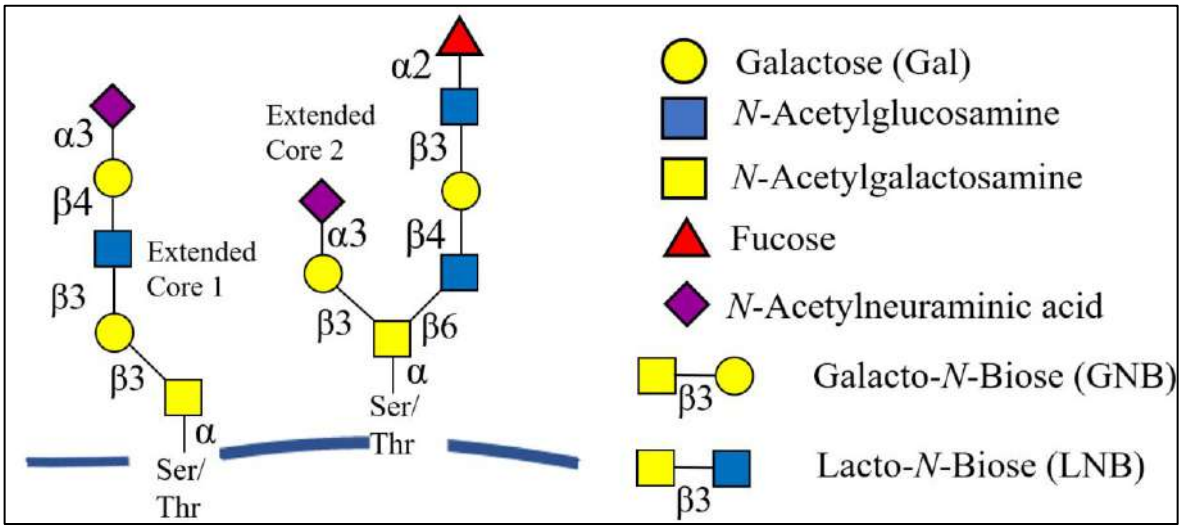

Figure 2- Representative examples of complex *O*-glycans with different types of cores.

1.4 GMP AS A SOURCE OF *O*-LINKED GLYCANS

During cheese manufacturing, milk coagulation is one of the most important stages in fat and protein concentration. The proteolytic enzyme chymosin is added to the milk at pH 6.6, and coagulation is initiated (Kethireddipalli & Hill, 2015). Chymosin hydrolyzes the κ-casein protein at the specific Phe105-Met106 bond, splitting the macromolecule into para-kappa-casein fractions, and the glycomacropeptides (GMP) are transferred to the resulting liquid fraction called whey (Hidalgo et al., 2010). After coagulation, GMP is incorporated as a component of liquid whey (20–25% of the protein moiety) resulting from κ-casein treatment (Keel et al., 2022). Interestingly, GMP is a hydrophilic, negatively charged, 64 amino acid *O*-glycosylated peptide with five different types of *O*-linked cores (Figure 3**)** (González-Morelo et al., 2020a). The molecular weight of GMP ranges from 7 to 11 kDa (Córdova-Dávalos et al., 2019a), depending on the level of glycosylation and isolation method (Li and Mine 2004). Due to its complexity, it probably reaches lower gut sections and interacts with the gut microbiota.

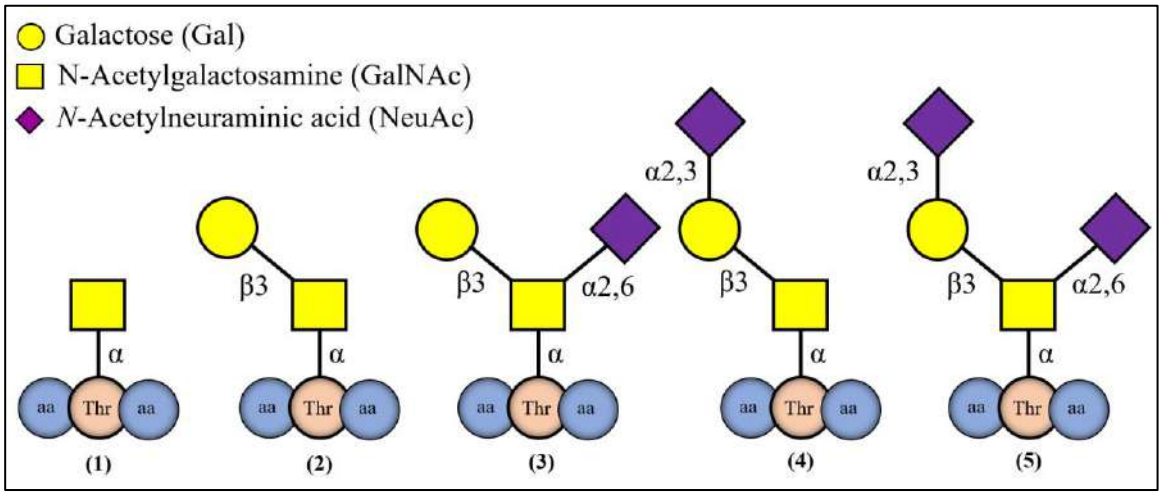

Figure 3- Structure of *O*-liked glycans from GMP.

Reported studies have shown that *B. infantis* ATCC 15697 is able to grow using the glycosidic fraction of GMP (O'Riordan et al., 2018). Its bacterial growth was proportional to the concentration of GMP and was stimulated only by the oligosaccharide moiety of the glycopeptide. Several molecular adaptations for accessing and consuming *O*-glycans have been described in gut microorganisms, revealing the utilization of intra- and extracellular enzymes (Morozumi et al., 2023).

Among some species of *B. infantis*, *B. longum*, and *B. bifidum*, conserved bacterial endo-*α*-*N*-Acetylgalactosaminidases (EC 3.2.1.97) have been discovered that catalyze the hydrolysis reaction of the *O*-glycosidic bond between the α-GalNAc of the reducing end of the oligosaccharide chain and Ser/Thr as the glycosylation site (Katoh et al., 2020). Some bifidobacteria can produce sialidases that break down GMP, separating NeuAc and allowing the action of other glycosyl hydrolases and a GNB transporter to import them. However, little is known about the enzymes and molecular mechanisms involved in this process of *O*-glycan utilization (Wang, 2020).

1.5 IMPLICATIONS OF GLYCAN PROCESSING IN THE HUMAN GUT ENVIRONMENT

A remarkable number of reports have shown that the intestinal mucus is a modulator of human health, functioning as a defensive barrier against physical and biological attacks (Breugelmans et al., 2022). Mucin can contain up to 80% *O*-glycans and serve as a source of nutrients for some commensal microbes and pathogens. The main component of secreted colonic mucus is a network of cross-linked, high molecular-weight Muc2 glycoproteins (MUC2) (Liu et al., 2020). It has been reported that in situations of dietary polysaccharide deprivation, it promoted specific bacterial populations and their enzymes that target host mucin glycans as an energy source, leading to erosion of the colonic mucosal layer and increased susceptibility to enteric pathogens (La Rosa et al., 2022).

On the other hand, milk glycans provide a defense against pathogens and can block their adhesion (Liu & Newburg, 2013). Many pathogens adhere to the intestinal epithelial surfaces, and milk glycans have epitopes identical to the intestinal surfaces, as they are generated from similar enzymes (Peterson et al., 2013). Consequently, pathogens bind to milk glycans instead of host cell ligands. Another possible mechanism of glycan protection from pathogens is the competitive binding of glycans with host cell surface receptors (Morrow et al., 2005). HMO have also been shown to alter the intracellular fate of pathogens. For example, pretreatment with HMO reduced UPEC-mediated activation of key signaling proteins such as MAP kinase and the master regulator NF-κB, thereby protecting the bladder epithelial cells from the pro-inflammatory effects of bacterial infection (Lin et al., 2014). Due to the relevance in maintaining the host health and providing a nutrient niche for the microbiome, there is a great interest in understanding how different microbes degrade and consume *O*-glycans, which are considered as emerging prebiotics, as well as analyzing the regulation of similar and better-known carbon sources such as HMO. In the present study, GMP, a glycopeptide with sialylated and neutral *O*-glycans, was used as a study model to determine the metabolic interactions between representative intestinal microbes.

1.1 HYPOTHESIS AND OBJECTIVES

O-glycans are important structures commonly found in the intestine. They can be endogenously in mucins or supplied by the diet in glycoproteins. GMP is a simple model glycopeptide with *O*-glycans. Although it has been considered a prebiotic in some studies as well as HMO, its influence on the gut microbiota, the strategies associated with its consumption, and its impact on metabolic interactions, have not been studied.

The main hypothesis of this doctoral thesis is that representative bacteria of the human gut microbiota deploy distinct transcriptional modules and interact metabolically during the utilization of milk-derived HMO or GMP O-glycans.

General Objective

The general objective of this thesis was to determine the molecular strategies and metabolic interactions of representative bacteria of the human gut microbiota in the utilization of GMP as a source of *O*-glycans.

The specific objectives are:

1. To study co-expression networks in human milk oligosaccharides in *Bifidobacterium* species.
2. To study the mechanism of GMP *O*-glycan utilization by bacteria representative of the adult intestinal microbiome.
3. To determine the metabolic *cross-feeding* interactions between representative bacteria of the human gut microbiome using GMP as a source of *O*-glycans.

1.2 OUTLINE OF THIS THESIS

A general outline of this thesis is summarized in Figure 4, which considers three chapters.

The first chapter is a critical review of the distinct molecular mechanisms of consumption of *O*-linked GalNAc glycans by prominent gut microbes of reports currently available in the literature, especially from mucin and casein glycomacropeptide (GMP), highlighting the potential of these structures as emerging prebiotics.

The second chapter mainly focuses on the HMO, oligosaccharides chemically similar to *O*-linked glycans from GMP. Before delving into *O*-glycans, this chapter highlights the distinct co-expression architectures in bifidobacterial genomes during HMO consumption. It contributes to understanding gene regulation and co-expression in these gut microbiome species.

The third chapter mainly focuses on GMP as a carbon source. This chapter study highlights complex strategies for utilizing *O*-glycans in GMP consumption among gut microbes based on cross-feeding and competition.

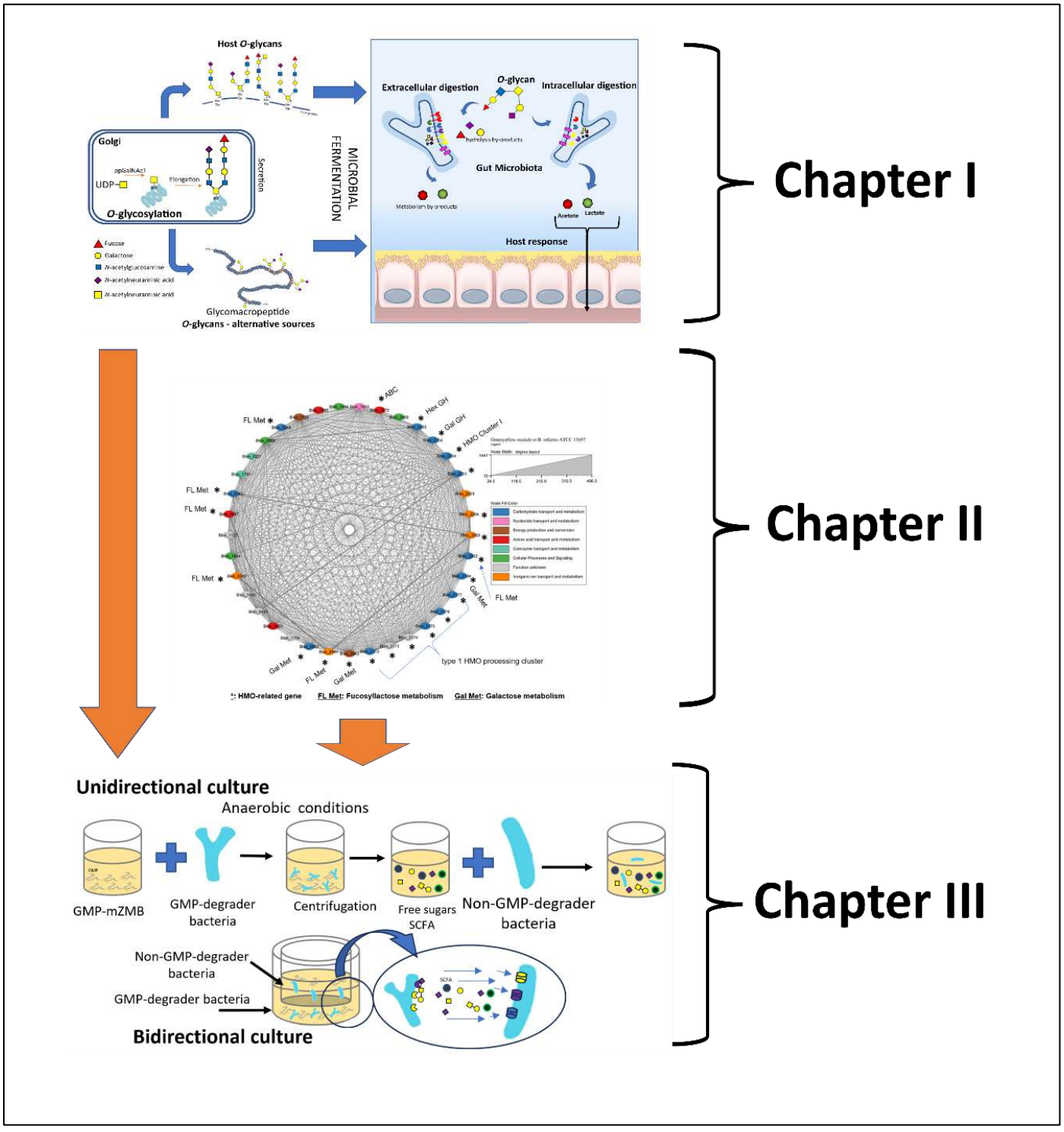

Figure 4- Overview of the studies compromising in this thesis.

2 CHAPTER I: *O*-LINKED GLYCAN UTILIZATION BY HUMAN GUT MICROBES

2.1 INTRODUCTION

There is a great interest regarding the impact and modulation of the gut microbiota through our diet. Among several dietary interventions, consumption of fibers and prebiotics has mainly been considered as positive for our health (Sawicki et al., 2017). A recent definition of prebiotic corresponds to a substrate that is selectively utilized by host microorganisms conferring a health benefit (Gibson et al., 2017). These benefits include a reduced microbial load of pathogens (Gibson et al., 2005), stimulation of the immune system (Shokryazdan et al., 2017), and lower allergy rates (Brosseau et al., 2019). These effects are attributed in part to the ability of prebiotics to be fermented by healthy microorganisms and stimulate the production of certain short-chain fatty acids (SCFA), especially butyrate (Rivière et al., 2016).

Traditionally, carbohydrates such as inulin and fructo-oligosaccharides (FOS) have been studied as prebiotics (BeMiller, 2019; Vandeputte et al., 2017; Wilson & Whelan, 2017). In general, these plant-derived prebiotics have a simple structure containing one monosaccharide and repeats of one linkage. In contrast, host-derived glycans such as human milk oligosaccharides (HMO) and those found in glycoproteins, are structurally more complex and thought to be more suitable and selective toward beneficial members of the gut microbiota. HMO are complex glycans that promote the growth and activity of beneficial gut microbes such as infant-gut associated bifidobacteria, among other effects (Thompson et al., 2018). Several advances have been made to synthesize HMO at the industrial level, and they are currently being added as functional ingredients to infant formula (Puccio et al., 2017).

Besides, some human and bovine milk proteins possess similar glycans to HMO in their structures. Glycosylation is a post-translational modification where oligosaccharides are covalently bound to asparagine (*N*-glycans), or serine or threonine (Ser/Thr; *O*-glycans; Vliegenthart, 2017). These glycans serve as signaling molecules for secretion and other cellular processes, providing increased resistance to proteolysis (Baum & Cobb, 2016).

These glycans are being proposed as emerging prebiotics due to their similarity to host-derived glycans compared to plant-derived prebiotics and their enrichment of dominant and health-promoting gut microbes, such as *Bacteroides* spp., *Bifidobacterium* spp., and *Akkermansia muciniphila* (Bergstrom & Xia, 2013; Kirmiz et al., 2018). Several molecular adaptations for accessing and consuming *N*- and *O*-glycans have been described in gut microorganisms, indicating that the gut microbiota is quite adapted for metabolizing these oligosaccharides. Importantly, some of these bioactive glycans could also be found in dairy streams, warranting a wide availability for potential prebiotic use for the food industry.

N-linked glycans are complex oligosaccharides that possess *N*-Acetylglucosamine (GlcNAc) as a common core, conjugated with an additional GlcNAc and three mannose (Man) residues, forming a $Man_3GlcNAc_2$ motif found in all *N*-glycans. *N*-glycans could be further modified by extensive mannosylation (high-mannose *N*-glycans) or by lactosamine chains (Galβ1-4GlcNAc; complex *N*-glycans). Terminal fucose (Fuc) and *N*-Acetylneuraminic acid (NeuAc) modifications are commonly added to complex *N*-glycans (Stanley et al., 2017). *N*-glycans are characteristic of immunoglobulins and lactoferrin in milk (Davis et al., 2016; Parc et al., 2015). *N*-glycans are an example of host-derived oligosaccharides that can be used by beneficial microbes. *Bifidobacterium longum* subsp. *infantis* ATCC 15697 is a dominant beneficial infant gut microbe that has been shown to access *N*-glycans in vitro by a specialized endo-β-*N*-Acetylglucosaminidase (Garrido, Nwosu, et al., 2012b). It also shows vigorous growth in vitro using *N*-glycans (Karav et al., 2016). The released *N*-glycans from milk glycoproteins such as lactoferrin and immunoglobulins have been well characterized (Parc et al., 2015). These results, while promising, are yet to be tested in vivo for claiming any prebiotic effect.

Less attention has been paid to *O*-linked glycans. These are characterized by an *N*-Acetylgalactosamine (GalNAc) residue linked to Ser/Thr as a common core. GalNAc is usually bound to Galβ1-3, forming galacto-*N*-biose (GNB) as a building block in Core 1 and Core 2, which could be further extended forming larger chains similar to *N*-glycans. *O*-linked glycans are a significant component of mucins (MUCs; Figure 5), also found in immunoglobulins and κ-caseins (Magnelli et al., 2011). *O*-glycosylated proteins have a higher resistance to proteolysis (Boutrou et al., 2008; Kesimer & Sheehan, 2012), probably reaching lower sections of the gut and interacting the gut microbiota. Therefore, it is not

surprising that gut microbes have evolved strategies to utilize *O*-linked glycans as a carbon source.

Prominent gut microbes such as *A. muciniphila*, *Bifidobacterium bifidum*, and *Bacteroides thetaiotaomicron* are representative species capable of *O*-linked glycan consumption. However, the mechanisms involved and the consequences of this process for the host are not fully known. This chapter aims to summarize and discuss current research regarding the structures of *O*-linked glycans, their potential prebiotic activity, and consumption by gut microorganisms.

2.2 STRUCTURES OF *O*-LINKED GLYCANS

The elongation of *O*-linked glycans and their attachment to secreted proteins begins in the Golgi apparatus (Brockhausen & Stanley, 2017). A polypeptide-*N*-Acetyl-galactosaminyltransferase (ppGalNAcT) catalyzes the addition of a GalNAcα1 to a Ser/Thr available as a glycosylation site (Varki & Lowe, 2009). Several glycosyltransferases act in conjunction to attach sugar residues in single *O*-glycans. The Core 1 *O*-glycan (Galβ1-3GalNAcα1-Ser/Thr) is the first synthesized, and then the Core 2 unit [Galβ1-3(GlcNAcβ1,6)GalNAcα1-Ser/Thr; Itano, 2019].

Here, we will focus on two model glycoproteins containing *O*-glycans: MUC and glycomacropeptide (GMP). The wide availability of mucins from animal sources makes them reference glycoproteins for the study of their contribution to host processes and their interaction with the gut microbiota. Similarly, GMP is a glycopeptide derived from cheesemaking, available in large quantities. It contains neutral and acidic *O*-glycans, which are shorter and simpler compared to mucin *O*-glycans. Both types of glycans have been shown to interact with the gut microbiota and could be considered promising prebiotics considering their stimulation of beneficial gut microbes.

2.3.1 *O*-linked glycans from mucin

Mucins are highly glycosylated (up to 90%), high-molecular-weight (200 kDa–200 MDa), and large (Rg 10–300 nm) extracellular glycoproteins that serve as a dense barrier between the intestinal lumen and epithelium (Kesimer & Sheehan, 2012). They are generally found in epithelial tissues in the gastrointestinal tract and certain secretions. They provide a

crucial role in forming a physicochemical barrier against the luminal compartment through their gel-forming properties (Bansil & Turner, 2018). Interestingly, mucin serves as a scaffold for the attachment and colonization by certain microorganisms (Ringot-Destrez et al., 2017). Two different mucus layers have been identified: an outer mucus layer (containing 106 microbial cells/g) and a tight inner mucus layer (105 microbial cells/g; Atuma et al., 2001). Mucins can be found in different parts of the human body: kidneys and bladder, respiratory and urinary tracts, among others (Atuma et al., 2001). They are produced by goblet cells, specialized secretory cells found in the epithelial layer (Lamacchia et al., 2018).

There are at least 20 different MUC genes, whose products are classified into two families: secreted and transmembrane mucins (Corfield, 2015; Dhanisha et al., 2018). Secreted mucins can be small and soluble, or large gel-forming mucins. The latter are able to cross-link, forming complex networks and contributing to the viscosity of mucins. MUC5AC is among the most studied mucin being secreted by epithelial cells in the stomach, while MUC6 is secreted in the deep gastric glands in the stomach and ileum (Magalhães et al., 2016). MUC2 is the most abundant secreted mucin in the small intestine and the colon (Arike et al., 2017; Corfield, 2000). Transmembrane mucins participate primarily in cellular adhesion, while secreted mucins are responsible for viscoelasticity for both the inner and outer mucus layers (Demouveaux et al., 2018; Lillehoj et al., 2013). All these differences between secreted and membrane-bound mucins contribute to the dynamic properties of the mucus layer across the GI tract.

Mucins could also be found in secretions and milk. MUC1 is a transmembrane mucin expressed on the apical surface of most epithelial cells (Ross et al., 2015). This type of mucin can be found in breast milk, which is transferred to the milk fat globule membrane upon its secretion from the cells. MUC15 is also a mucin-type isolated from bovine milk fat globule membranes (Bernard et al., 2018). The function of milk mucins has not been extensively studied, but it is suggested to be mostly structural (Donovan, 2019).

Eight cores of *O*-glycan structures have been identified in mucins, four of them being the most predominant (Core 1–4; Figure 5). These *O*-glycans are found in the region known as the variable number of tandem repeats (VNTR; Arike & Hansson, 2016). The VNTR section is rich in Ser/Thr, which can make up to 80% of the weight of the mucin. In MUC2, VNTRs are accompanied by two cysteine-rich regions at their ends and a D domain involved

in mucin polymerization (Arike & Hansson, 2016). Depending on the tissue, MUC5AC and MUC6 could be glycosylated with predominant Core 1–2 structures (Jin et al., 2017). The *O*-linked glycans of colonic MUC2 are predominantly Core 1–3 structures (Bergstrom et al., 2017).

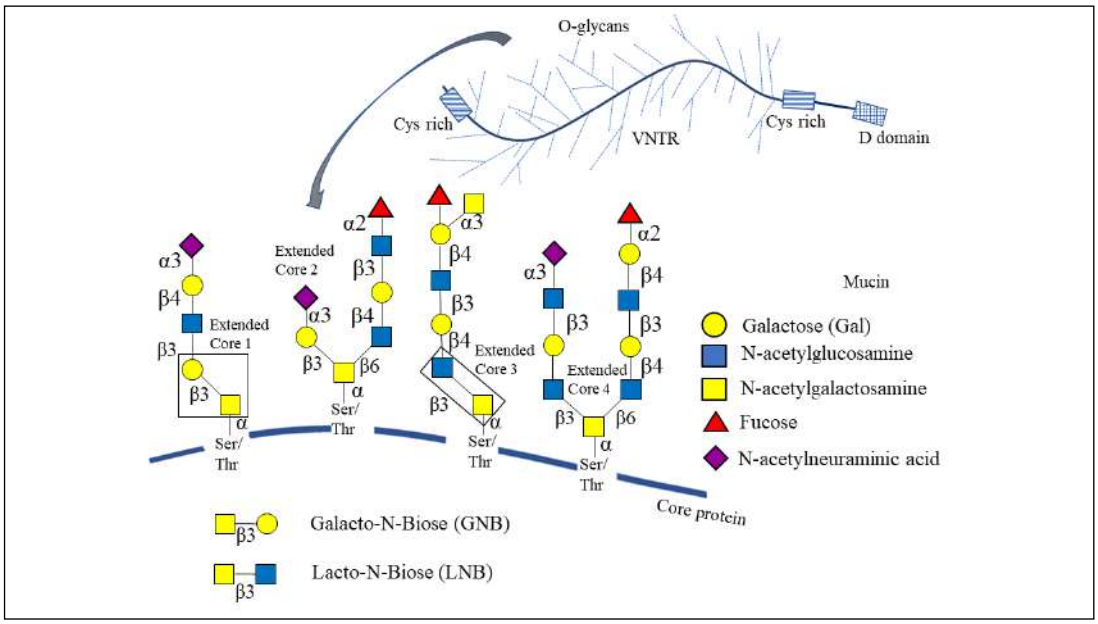

Figure 5- Schematic representation of *O*-glycan cores linked to mucin.

Mucin *O*-glycans can be either branched or linear in structure, depending on their Core (e.g., Core 1 glycans are linear, and Core 2 glycans are branched; Li & Chai, 2019; Podolsky, 1985. *O*-glycans have been found to contain up to 20 residues and may include blood group determinants of ABO, Lewis groups, and glycan epitopes such as the linear antigen i (Galβ1-4GlcNAcβ1-3Gal; Vliegenthart, 2017). Fuc and NeuAc are monosaccharides decorating *O*-glycans in terminal α-linkages, and usually mucins contain sulfate groups. The latter two confer a negative charge on the mucin structure, crucial for selective mucin permeability and its rheological properties (De Weirdt & Van de Wiele, 2015). An acidic gradient, based on increasing amounts of NeuAc in mucins, has been shown in the gastrointestinal tract (Robbe et al., 2003).

Several studies have shown how alterations in mucin *O*-glycosylation patterns participate in disease. In cystic fibrosis patients, the respiratory epithelium shows higher degradation rates for MUC5B and MUC5AC, in addition to reduced sulfation, higher

sialylation, and lower fucosylation (Schulz et al., 2007). This disease is characterized by chronic pulmonary infection and severe inflammation, and the changes mentioned above could be used as biomarkers. In the stomach, a broad diversity of *O*-glycans has been observed, but specific epitopes, such as Lewis b and α1-4GlcNAc, were found to promote adhesion of Helicobacter pylori to the mucus layer (Jin et al., 2017; Rossez et al., 2012). This pathogen is a direct cause of gastrointestinal ulcers and gastric cancer.

Similarly to mucins, glycoproteins in cancer cells show alterations in their *O*-glycosylation profiles. Truncated *O*-glycans such as a GalNAc residue with no further glycosylation (GalNAcα-Ser/Thr; Tn antigen) is considered a tumor marker given its presence is abnormal in glycoproteins (Itzkowitz et al., 1989). Sialylated Tn is also a common feature of cancer cells. An increase in core fucosylation is also observed in these cells. Consequently, these changes interfere with cell-cell adhesion processes, promoting tumor cell invasion (Pinho & Reis, 2015).

Importantly, *O*-linked glycans from human gastric mucin are structurally similar to those from the porcine stomach mucin. The oligosaccharides in both glycoproteins are mainly extended with Core 1 and Core 2 structures. The availability of porcine stomach mucin has facilitated the study of mucin function and its interactions with gut microorganisms. However, oligosaccharides in pig mucins have unique characteristics, like that they could contain extended Core 3 and Core 4 glycans (Padra et al., 2018). Pig gastric mucins are highly sulfated, usually terminating with galactose residues and have lower sialylation (Quintana-Hayashi et al., 2018).

Mucin glycosylation is in part mediated by gut microbiota. Using germ-free (GF) mice as control, it has been shown that the presence of certain members of gut microbiota is critical for the expression of glycosyltransferases participating in mucin *O*-glycosylation for example, ppGalNAcT, Core 1 β1,3-Galactosyltransferase (C1GALT1), and Fucosyltransferase (FUT2; Johansson et al., 2015). MUC2 *O*-linked glycans from colonic tissues of conventionally raised animals were more sialylated, fucosylated, and longer than GF mice (Arike et al., 2017). MUC2 from GF mice showed a reduced abundance of β1,6-*N*-Acetylglucosaminyltransferases, enzymes responsible for the formation of Core 2 and Core 4 *O*-glycans. These GF animals tended to produce shorter *O*-glycans (Arike et al.,

2017). This result suggests that the gut microbiota modulates the *O*-glycosylation patterns of mucins, which in turn influence the composition the gut microbiota.

Mucin utilization by the gut microbiota is important for host health. It is known that animals that are fed mucin show higher fecal butyrate levels. This compound is an anti-inflammatory SCFA considered positive for health and significantly decreased in inflammatory diseases such as Ulcerative Colitis (UC; Chen et al., 2018; Yamada et al., 2019). In UC subjects, mucin *O*-glycan abundance correlates negatively with butyrate production in feces, showing a reduced utilization of *O*-glycans by the gut microbiota (Yamada et al., 2019). This study presented important in vivo evidence linking the foraging of host-derived glycans, the action of the gut microbiota, and inflammatory bowel diseases (IBDs). In a different context, mucin *O*-glycan consumption by gut microbes could be considered unfavorable for the host. Animals deprived of fiber in their diets show the promotion of gut inflammation and colitis. This was explained partly by their gut microbiota turning into endogenous mucin resources as the last carbon source available. Accessing the mucin layer by mucin degraders permits the colonization by *Citrobacter rodentium*, a mucosal pathogen of mice sharing several pathogenic features with human gastrointestinal pathogens such as enteropathogenic *Escherichia coli* (Collins et al., 2014; Desai et al., 2016).

Mucins and their *O*-glycans are essential in the gut barrier function. Modifications of their expression, organization, or glycosylation are likely to prevent this effect, as observed in IBD. The altered balance between pro-inflammatory cytokines (TNF, IL-1b, IL-8, and IL-17), anti-inflammatory cytokines (IL-4 and IL-13), and immuno-regulatory cytokines described in IBD is likely to modify mucin expression and glycosylation (Q. Guan & Zhang, 2017). Furthermore, altered glycosylation and sulfation of colonic mucins in IBD subjects could alter the protective role of the colonic mucus barrier (Groux-Degroote et al., 2020).

2.3.2 *O*-linked glycans from GMP

Glycomacropeptide is a constituent of whey (20–25% of the protein moiety), derived from κ-casein after chymosin treatment during cheesemaking (Manso & López-Fandiño, 2004; Sunds et al., 2019). GMP is a hydrophilic, negatively charged, 64 amino acid

glycopeptide. The molecular weight of GMP ranges from 6.7 to 8 kDa (Córdova-Dávalos et al., 2019a). GMP is available in large quantities in dairy streams (Rojas & Torres, 2013).

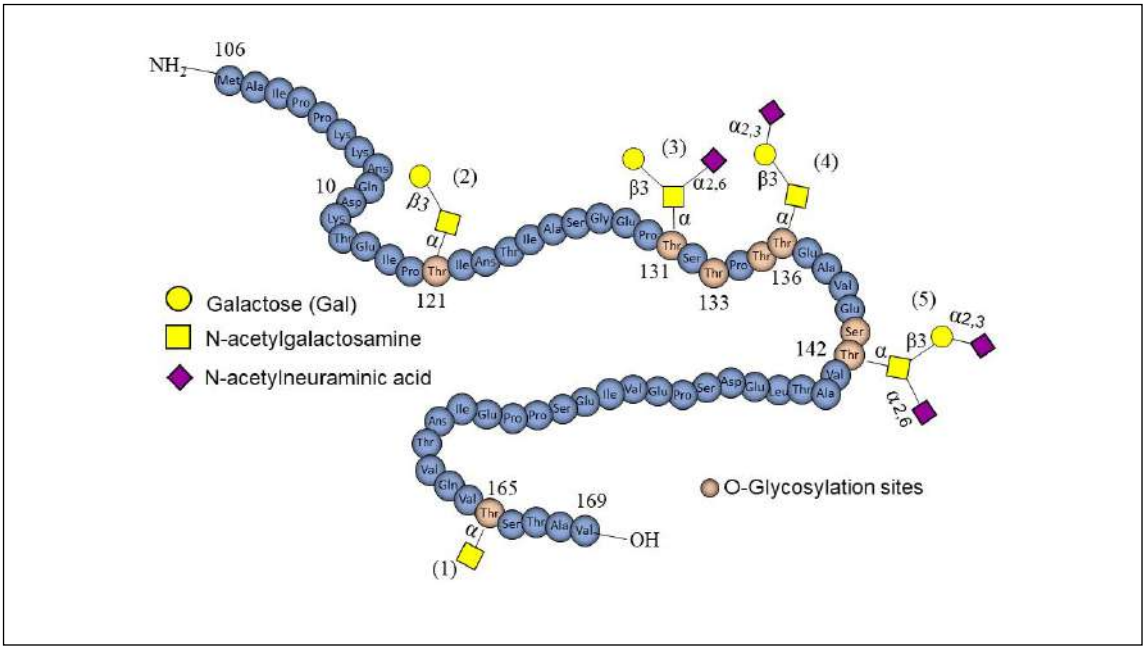

Figure 6- Schematic representation of *O*-linked glycans found in casein glycomacropeptide (GMP).

The glycan portion is simpler compared to mucin oligosaccharides, containing only NeuAc, GalNAc, and Gal. GMP has five distinct *O*-linked glycans bound to Ser/Thr (Figure 6), corresponding to monosaccharides (0.8%), disaccharides (6.3%), trisaccharides (36.5%), and tetrasaccharides (56%; Saito & Itoh, 1992). The disaccharide GNB (Figure 6) is a building block found in GMP (Fiat & Jollès, 1989). Different *O*-glycosylation sites have been proposed in GMP, as determined by 2D gel electrophoresis and tandem mass spectrometry (Figure 2; Neelima et al., 2013). Threonine in positions 121, 131, 133, 136, and 142 appear to be used as glycosylation sites, and positions 165, 135, and 141 have been proposed as potential sites (Thomä-Worringer et al., 2006). As expected, the functional and adhesion properties of GMP will depend on the position where *O*-glycans are attached, especially the sialylated. NeuAc is an acid sugar easily recognizable by mucin-degrading bacteria (Neelima et al., 2013).

2.4 UTILIZATION OF MUCIN *O*-GLYCANS BY GUT MICROBES

While certain gut bacteria use mucin glycan as attachment sites, a few others go beyond and additionally deploy enzymatic activities to use mucin *O*-glycans as a nutrient (Tailford et al., 2015). Interestingly, this enzymatic capability is distributed in members across the four most dominant phyla in the gut microbiota (Bacteroidota, Actinomycetota, Bacillota, and Verrucomicrobia), suggesting it is an important trait for the gut microbiota.

Mucin *O*-glycan degradation by intestinal microorganisms requires an extensive array of glycosyl hydrolases (GHs), as expected from the complex structures, compositions, and diverse linkages found (Tailford et al., 2015). Most mucin-degrading bacteria base their strategies on exo-acting enzymes, including sialidases (GH33), fucosidases (GH29 and GH95), exo- and endo-β-*N*-Acetylglucosaminidases (GH84 and GH85), β-galactosidases (GH2, GH20, and GH42), among others. Their activities imply sequential degradation. Other enzymes required for mucin utilization are α-*N*-Acetylglucosaminidases (GH89) and α-*N*-Acetylgalactosaminidases (GH101, GH129), which cleave a monosaccharide-peptide linkage after glycosidase activity. One potential exception to this exo-acting mechanism is a GH16 endoactive *O*-glycanase with mucin breakdown activities (Crouch et al., 2020). These activities release mono or disaccharides, which could be imported inside the bacterial cell or serve as cross-feeding metabolites (Turroni et al., 2018). Several enzymes related to mucin deglycosylation have been described, but the full mechanisms by which gut microbes utilize these glycans are unknown.

2.4.1 Utilization of *O*-Linked glycans by *Bacteroidota* spp

The Bacteroides genus is predominant in the adult gut microbiota, being generally considered as commensals. Genomic and functional studies in animals show that these species show a preference for utilizing complex carbohydrates (e.g., HMO), rather than simple carbohydrates (Wexler & Goodman, 2017). Bacteroides devote a large part of their genomes to polysaccharide utilization loci (PULs), which correlates with their broad diverse polysaccharide utilization. These gene clusters encode multiple extracellular GHs enzymes, oligosaccharide transporters, and binding proteins.

Early works studied the molecular system of complex glycan degradation in starch utilization in these species (Anderson & Salyers, 1989). *Bacteroides thetaiotaomicron* VPI-5482 is a model bacterium for mucin *O*-glycan utilization. The microorganism shows a remarkable growth in porcine mucin glycans (Martens et al., 2008). These oligosaccharides can induce the expression of at least 16 PULs in *B. thetaiotaomicron*, aiding in identifying loci involved in *O*-glycan utilization. These genes encode for putative glycolytic enzymes, including an *α-L*-fucosidase, endo-*β*-GlcNAcase, endo-*β*-galactosidase, *α*-GalNAcase, in addition to proteases, neuraminidase, and a sulfatase. The enzymes are suggested to play an orchestrated degradation of *O*-linked glycans (Luis et al., 2018; Sicard et al., 2017). The relevance of these genetic determinants has been confirmed in vivo experiments in GF mice, where mutants for these loci show reduced gut colonization.

These enzymes are part of PULs also known as Starch utilization system (Sus), an operon of eight genes (SusRABCDEFG). Their products are predicted to locate in the periplasm or the outer membrane (Figure 7). Carbohydrates are metabolized in response to the SusR signal, which triggers the expression of hydrolyzing proteins. The processed oligosaccharides are transported by a TonB protein and are metabolized inside the cells (Cockburn & Koropatkin, 2016). SusD homologs possess a binding site for *O*-glycans (Tailford et al., 2015) and have been implicated in the utilization of *O*-linked mucin glycans. Interestingly, the transcriptional response mounted to utilize *O*-glycans is similar to that required for HMO consumption (Marcobal et al., 2011).

Four PULs in *B. thetaiotaomicron* were found to be induced in HMO but not in the presence of mucin glycans, indicating that this microorganism can respond to building blocks in HMO not found in mucin. Regardless, the deletion of these PULs did not affect the ability of *B. thetaiotaomicron* to consume HMO. Apparently, for this strain *O*-glycans are not a premium substrate since several mono or simple carbohydrates repressed these genes (Pudlo et al., 2015). This was not observed for other Bacteroides species, indicating the complex regulation of glycan utilization and diversity on preference for carbon sources even at the species level (Pudlo et al., 2015).

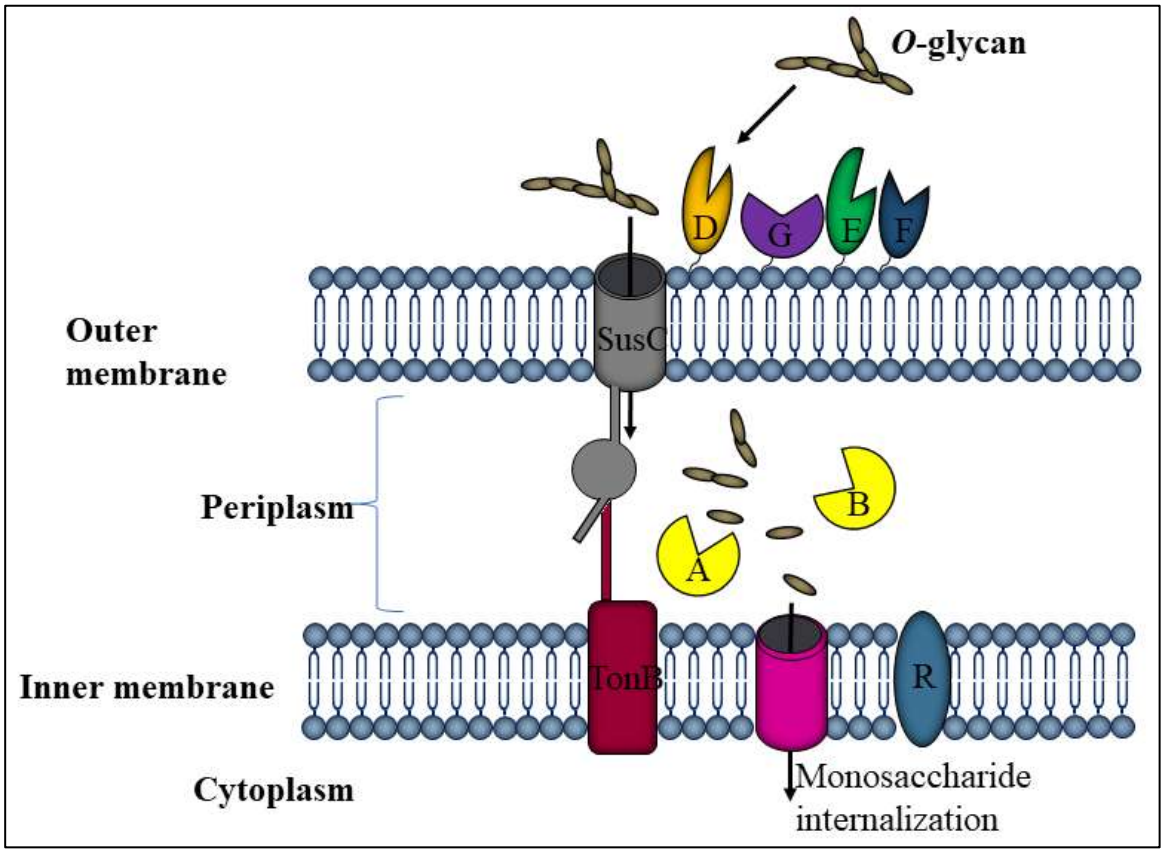

Figure 7- Representation of *O*-linked glycan utilization by *Bacteroides thetaiotaomicron*.

Monocolonization of GF mice with *B. thetaiotaomicron* modulates cellular responses in mucin-producing goblet cells, increasing their differentiation and synthesis of mucins, changing their glycosylation patterns with a higher NeuAc content (Wrzosek et al., 2013). In the presence of *Faecalibacterium prausnitzii*, a keystone microbe and butyrate producer, these cellular responses were attenuated, probably maintaining epithelial homeostasis between cell lines.

Bacteroides fragilis is a commensal species, but certain strains could be enterotoxigenic causing serious diseases (L. Chung et al., 2018; Purcell et al., 2017). This species also has a PUL involved in the consumption of host *O*-glycans. This PUL is known as the commensal colonization factor (CCF; Lee et al., 2013). The genes within the CCF are homologous to *B. thetaiotaomicron* PULs, which are up-regulated possibly producing extracellular enzymes to sense and mediate *O*-glycan processing. These enzymes allow the commensal colonization of mucus, specifically in the crypts of the colon. Some of these have been characterized biochemically (Praharaj et al., 2018; Yamamoto et al., 2018).

2.4.2 Utilization of *O*-linked glycans by *Bacillota* spp.

This phylum contains a broad diversity of genera, including commensal Clostridium, Eubacterium, and Lactobacillus species. In general, most Firmicutes in the gut microbiota prefer the assimilation of smaller rather than complex carbohydrates (Cockburn & Koropatkin, 2016; Ravcheev & Thiele, 2017). Several *Lactobacillus* species are endowed with mucin binding proteins (van Tassell & Miller, 2011), but they apparently lack any mucin utilization mechanism. However, within the Firmicutes a few members have acquired the ability to gain access to *O*-linked glycans, deploying several extracellular GHs.

A few pathogens in this group have evolved mechanisms for O-glycan utilization. Clostridium perfringens, an opportunistic pathogenic bacterium, can consume N-glycans and *O*-glycans from intestinal mucin by releasing extracellular glycosidases such as sialidases (Koutsioulis et al., 2008; Pluvinage et al., 2019). *Enterococcus faecalis* strains possess genes encoding endo-*β*-*N*-Acetylglucosaminidase and endo-*α*-*N*-Acetylgalactosaminidase (endo-*α*-*N*-GalNAcase) that release *N*-glycans. However, this pathobiont appears not to be able to utilize mucin *O*-linked glycans (Morio et al., 2019; Roberts et al., 2000). These microorganisms are not usually dominant in the gut microbiota, probably due to the barrier effect mounted by this community.

Commensal Firmicutes such as *Ruminococcus torques* and *Ruminococcus gnavus*, belonging to the *Lachnospiraceae* family, can also access the *O*-linked glycans from mucin, especially MUC2. *Ruminococcus torques* is endowed with an *α*-sialidase (GH33), *α*-fucosidase (GH29), *β*-galactosidase, *β*-*N*-Acetylgalactosaminidase, *β*-*N*-Acetylglucosaminidase, sialate *O*-acetylesterase, and glycosulfatase activities. Mucin deglycosylation is carried out from a mixture of exo and endo activities (Crost et al., 2016). *Ruminococcus gnavus* contains a similar enzymatic arsenal, but its strategy involves the chemical modification of NeuAc by a trans-sialidase, releasing 2,7-anhydro-Neu5Ac (Bell et al., 2019). Interestingly, this modification allows this microorganism to use this carbohydrate, making it inaccessible for others gut microbes (Bell et al., 2019). This strategy has proven essential for proper colonization of *R. gnavus* in mice.

Ruminococcus gnavus is part of the healthy gut microbiota but appears to be increased in IBDs (Hall et al., 2017). Similar to *B. thetaiotaomicron*, *R. gnavus* has the ability to modulate host cellular responses, especially genes encoding MUC2 and certain

glycosyltransferases (Graziani et al., 2016). However, both *R. gnavus* and *R. torques* are increased several-fold in the intestinal mucosa of UC and Crohn's disease subjects, suggesting that their mucolytic activities contribute to disease progression (Png et al., 2010).

Sheridan et al., (2016) used in vitro growth assays and comparative genomics to identify genes involved in general carbohydrate metabolism in 11 *Roseburia* spp. and *Eubacterium rectale* strains. These bacteria have been suggested next-generation probiotics (Lordan et al., 2019), considered desirable for the host due to the production of butyrate. Two *Roseburia inulinivorans* strains were found to contain a Gram-positive PUL (gpPUL) with putative mucin glycan degradation genes such as a desulfatase, four glycosidases, and an ATP-binding cassettes (ABC) transporter. While this evidence suggests certain butyrate-producers could use mucin *O*-glycans as a carbon source, none of the strains of this study was able to grow using type 2 or type 3 porcine-derived mucin. Further experiments are required to clarify the ability of these microorganisms to target mucin as carbon source and the role of the gpPUL.

2.4.3 Utilization of *O*-linked glycan by *Bifidobacterium*

A few members of the phylum Actinomycetota can utilize glycans found in mucin, especially *Bifidobacterium* (Katoh et al., 2017; Turroni et al., 2010). This genus contains mostly commensal or beneficial microorganisms, dominant in the infant's gut.

Milani et al., (2016) showed that *B. bifidum* PRL2010 has a set of chromosomal loci that allows both *N*-glycans and *O*-glycan utilization. Seventy-seven genes, including GHs, glycosyltransferases, and glycosyl esterases were identified. *Bifidobacterium bifidum* contains a complete set of extracellular GHs, including exo-*α*-sialidases, 1,2-*α*-*L*-fucosidase, 1,3/4-*α*-*L*-fucosidase, *N*-Acetyl-*β*-hexosaminidases, and four *β*-galactosidases (Figure 8). Several of these enzymes have been biochemically characterized, showing that they are membrane-bound and essential for mucin *O*-glycan assimilation (Shimada et al., 2015).

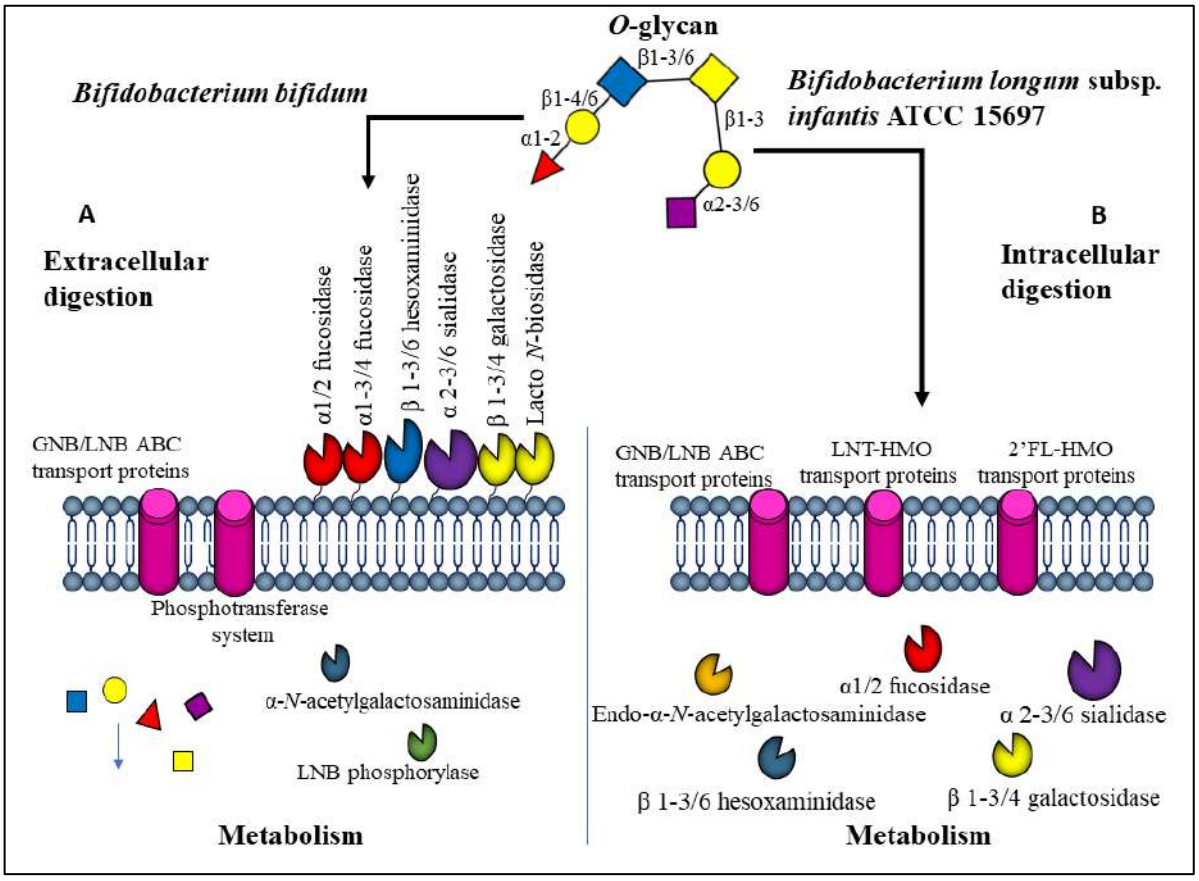

Figure 8- Representation of *O*-linked glycan utilization by *Bifidobacterium*: (A) *Bifidobacterium bifidum*, (B) *Bifidobacterium infantis*.

The consumption strategy of *B. bifidum* is mostly extracellular, where all these enzymes contain transmembrane domains and participate in mucin deglycosylation rendering GNB as a result. Apparently, *B. bifidum* does not utilize Fuc or NeuAc from complex glycans (Garrido et al., 2015c). However, sialidase activity in this bacterium appears to promote bacterial adhesion to the epithelium (Nishiyama et al., 2017). This consumption mechanism allows the resulting mono and disaccharides to participate in cross-feeding with other microorganisms, for example, with *Bifidobacterium breve* and *Eubacterium hallii* (Bunesova et al., 2018; Egan et al., 2014).

The Lacto-*N*-Biose (LNB)/GNB cluster present in most bifidobacteria allows these microorganisms to consume Lacto-*N*-Tetraose from HMO and GNB resulting from *O*-glycan degradation. This cluster is present in several infant and adult-associated bifidobacteria. It contains an ABC transporter specific for these two disaccharides and intracellular enzymes for LNB/GNB utilization in the bifid shunt (Thompson et al., 2018).

A few enzymes in this genus have been characterized by their ability to target mucin glycans. Earlier works described an endo-*α*-*N*-GalNAcase in *B. longum* subsp. *longum* JCM 1217 (EngBF), which is a conserved extracellular GH101 enzyme in this subspecies, being highly specific for Core 1 *O*-glycans (Fujita et al., 2005). An *α*-acetylgalactosaminidase (NagBb) from *B. bifidum* JCM 1254 is a GH101 enzyme that hydrolyzes *O*-glycans from Core 1 and Core 3 from mucin MUC2 (Kiyohara et al., 2012). The enzyme shows a high affinity for the Tn antigen (GalNAc*α*-Ser/Thr). Finally, a GH89 from the same microorganism shows affinity for GlcNAc*α*1-4Gal, an epitope commonly found in terminal positions in mucin glycans (Shimada et al., 2015). All these enzymes have great potential as a tool in glycobiology studies, and the obtention of prebiotics from complex mucin sources. Finally, evidence regarding the utilization of mucin *O*-glycans by *Bifidobacterium* has primarily been studied in vitro, and further studies are required to understand the implications of these strategies for the host. *B. longum*, as well as inulin, show protective effects against mucus deterioration caused by a mucin-eroding gut microbiota (Schroeder et al., 2018).

2.4.4 Utilization of *O*-linked glycans by Verrucomicrobia

In recent years, *A. muciniphila* has received considerable attention for its associations with health. Its abundance inversely correlates with the onset of IBDs and metabolic disorders (Dao et al., 2016; Lopez-Siles et al., 2018; Yassour et al., 2018), and it has been suggested to play an anti-inflammatory role (Cani & de Vos, 2017). It has also been proposed as a candidate next-generation probiotic, which in addition to the modulation of host responses possesses several technological properties compatible with scale-up processes (Cani & de Vos, 2017). Recently, the administration of pasteurized doses of *A. muciniphila* to overweight individuals improved certain inflammation and metabolic markers, being well-tolerated and safe (Depommier et al., 2019).

Akkermansia muciniphila is a Gram-negative bacterium that specializes in using mucins as nitrogen and carbon sources. This contrasts with other generalist gut microbes such as *Bacteroides* spp., which can target multiple sources for growth. Its specificity for mucin also presents a remarkable example of microbial adaptations to host conditions. *A.*

muciniphila is the only member of the phylum *Verrucomicrobiota* cultured and found in the gut microbiota, representing the 3% of intestinal bacteria detected in adult feces (Everard et al., 2014). Despite being considered a strict anaerobic microbe, recent studies have determined that it colonizes the mucus layer near epithelial cells, characterized by microaerophilic levels of oxygen (Ouwerkerk et al., 2016). *A. muciniphila* cannot synthesize Thr; however, it obtains this amino acid from mucin, which is one of the most abundant amino acids in its structure (van der Ark et al., 2018).

This microorganism is underrepresented in IBD, especially UC patients. Yamada et al. (2019) reported that fermentation of mucin O-glycans is associated with the production of butyrate in healthy subjects. However, subjects with UC show altered mucin *O*-glycosylation patterns, and the abundance of A. muciniphila was inversely proportional to the markers of inflammation (Earley et al., 2019; Yamada et al., 2019). This suggests that the *O*-glycosylation patterns present in mucins influence the anti-inflammatory activity of certain commensal microorganisms.

Studies show that the mucolytic activity and metabolite production by *A. muciniphila* promotes beneficial microbial networks (Belzer et al., 2017). Mucin degradation by gut microorganisms might be considered a pathogenic trait because it reduces the protective layer of the complex mucus layer. Intriguingly, in IL10-genetically deficient mice A. muciniphila contributes to colitis (Seregin et al., 2017). However, these degrading microorganisms may stimulate mucin accumulation, renewal, and thickening of the mucosal layer. In addition, the complexity of the mucin *O*-glycans makes these structures protective and inaccessible to most bacteria, except those that have developed consumption mechanisms such as *A. muciniphila* and *B. thetaiotaomicron*.

The genome analysis of *A. muciniphila* ATCC BAA-835, isolated from the human gastrointestinal tract, predicted the presence of a large number of mucinases (van Passel et al., 2011). This is a general term referring to GHs or lyases with combined activity against mucin. It was also reported that 26% of its proteome contains a signal peptide site, indicating a preference for extracellular degradation of macromolecules. Consequently, the putative strategy employed by this microorganism is to secrete an extensive array of extracellular proteins that hydrolyze *N*-glycans and *O*-glycans in simple sugars, some of which are later internalized (Shin et al., 2019). Regarding its enzymatic machinery, *A. muciniphila* is

predicted to contain *α*/*β*-D-galactosidase, *α*-*L*-fucosidase, *α*/*β*-*N*-Acetylgalactosaminidase, two *α*-*N*-Acetyl-glucosaminidases, neuraminidase, and a sulfatase. Few studies have confirmed the activities of some of these enzymes (Meng et al., 2020; Shin et al., 2019; M. Wang et al., 2018), including sulfatases and proteases targeting MUC2. Major end-metabolites produced by this bacterium are propionate, acetate, and sulfate (Ottman et al., 2017).

2.5 UTILIZATION OF GMP-DERIVED GLYCANS BY GUT MICROBES

Glycomacropeptide is a highly sialylated glycopeptide resulting from the cheesemaking process. It is an abundant source of *O*-glycans. However, not much attention has been paid to its potential applications for promoting a healthy gut microbiota. Compared to mucin, most studies have focused on the utilization of GMP by probiotic microorganisms in vitro, sometimes with contradictory results (Córdova-Dávalos et al., 2019b). Early work from Idota et al., (1994) showed that certain *Bifidobacterium*, including *B. bifidum*, *B. breve*, *B. longum*, and *B. adolescentis* could grow on GMP with OD_{600nm} values ranging from 0.7 to 2.70. Although purified GMP contains small amounts of lactose, which could contribute to the above results, the latter two microorganisms do not have the sialidases required to hydrolyze the GMP *O*-glycans (O'Callaghan & van Sinderen, 2016).

B. infantis is a dominant bacterium in the gut microbiota of newborns. It has been well characterized by its ability to fully utilize several HMO as a carbon source, with protective effects on the infant (Henrick et al., 2019; Thompson et al., 2018). The utilization strategy for this microorganism relies on ABC for internalization of intact glycans, subsequently degraded by intracellular GHs (Figure 4B; Zúñiga et al., 2018). *B. infantis* shows a preference for host glycans containing within its genome 16 GH, including several *α*-fucosidases, *β*-galactosidases, *β*-hexosaminidases, and *α*-sialidases (Thompson et al., 2018). This microorganism has also evolved determinants and mechanisms for the vigorous utilization of milk-derived *N*-glycans (Garrido, Nwosu, et al., 2012c; Karav et al., 2016).

Considering the structural similarity between HMO, *N*-glycans, and *O*-glycans, it could be expected that this microorganism also targets mucins as a growth substrate. Interestingly, *B. infantis* is not able to access the *O*-glycans of these glycoproteins (Kim et

al., 2013; Turroni et al., 2018). Most glycosidases in *B. infantis* are intracellular, and it lacks homologs to *α-N*-Acetylgalactosaminidases found in *B. bifidum* and *B. longum*.

Strikingly, *B. infantis* has been shown to grow vigorously using GMP (O'Riordan et al., 2018). A 20.6 ± 3.6% increase in OD600nm was observed in the mid-exponential phase of *B. longum* ssp. *infantis* cultures supplemented with GMP, indicating a growth-promoting effect. In vitro, it has been shown that *B. infantis* growth is proportional to GMP concentration in the culture media. GMP periodate treatment (GMP-P) inactivates the glycan portion of the peptide chain, and growth was compared with native GMP. GMP-P resulted in a substantially lower (5.5%) increase in growth compared with full GMP in the mid-exponential phase. Under these conditions, *B. infantis* was not able to grow using GMP-P, indicating that the GMP *O*-glycan portion is essential in *B. infantis* growth.

Transcriptomic analysis of the GMP utilization revealed the induction of two intracellular GH25 enzymes, a family related to bacterial lysozymes, and a solute binding protein from an ABC transporter (O'Riordan et al., 2018). These results provide new insights regarding the adaptations of this probiotic microorganism for a milk-derived substrate. However, how this infant-gut associated bacterium can use GMP *O*-glycans as the sole carbon source is not clear, especially considering it cannot access mucin *O*-glycans. It is possible that GMP, a smaller and structurally and physicochemical simpler molecule, is easier to access. Another possibility is the participation of proteases, which could facilitate the import of GMP-derived *O*-glycans into the bacterial cell for intracellular processing.

2.6 CONCLUSION

O-linked glycans are complex oligosaccharides decorating host glycoproteins. They are produced endogenously in mucins or could be found in secretions such as milk. These glycans are accessed, released, and consumed by individual members of the gut microbiota, especially beneficial gut microbes such as *Bifidobacterium*, *Bacteroides*, *Akkermansia*, and stimulate the growth of next-generation probiotics such as *Roseburia* and other butyrate-producing bacteria.

O-glycans serve as signaling molecules for cell secretion and provide greater resistance to proteolysis (Boutrou et al., 2008). Under conditions of dietary fiber depletion, endogenous mucin degradation occurs associated with microbial activity, promoting inflammation, and intestinal disease. *O*-glycans derived from milk glycopeptides such as in GMP are an attractive opportunity to use them in foods as emerging prebiotics, promoting a healthy gut microbiota. Unfortunately, there are still several limitations to this goal, and no human studies have evaluated its prebiotic effect. While, we have advanced in understanding some of the molecular mechanisms involved in O-linked glycan utilization in single microorganisms, the impact of this utilization in metabolic interactions and networks has been evaluated only in a small number of studies. These gaps have been hindered in part because we lack enzymes or other tools to recover full *O*-glycans from dairy streams or mucin sources. In contrast, certain enzymes have been described for the recovery of *N*-glycans from dairy byproducts. This offers an opportunity for identifying novel enzymes from gut microbes or improving the activity of current endo-*α*-*N*-Acetylgalactosaminidases.

Moreover, the consequences for the host of the microbial utilization of *O*-glycans, or glycoproteins containing *O*-glycans, have not been evaluated and remains a critical question. Whether gut microbes accessing the mucus layer and releasing *O*-glycans is beneficial for the host is still unclear. Furthermore, the study of *O*-glycans utilization by the gut microbiota, and how these microbes shape mucin glycosylation patterns, should be further studied in the context of intestinal inflammation and IBDs, as well as cancer progression.

3 CHAPTER II: STRUCTURE OF CO-EXPRESSION NETWORKS OF *BIFIDOBACTERIUM* SPECIES IN RESPONSE TO HUMAN MILK OLIGOSACCHARIDES

3.1 INTRODUCTION

The human gut microbiota is a community of anaerobic microorganisms that plays an important role in the metabolization of complex carbohydrates that are not degradable by host enzymes (Flint et al., 2012). Their presence directly influences gastrointestinal physiology. Consumption of fiber and prebiotics has been considered positive for our health. These benefits include a reduced load of pathogens (Gibson et al., 2005; Markowiak & Śliżewska, 2017), stimulation of the immune system (Shokryazdan et al., 2017), lower allergy rates (Brosseau et al., 2019), and production of short-chain fatty acids as metabolites resulting from their degradation (Tsukuda et al., 2021). For this reason, a great interest in this research has arisen, considering the contribution of the intestinal microbiota to our wellbeing (Couto et al., 2020).

Free human milk oligosaccharides (HMO) are the third most abundant component in human milk after lactose and lipids. They are structurally complex glycans composed of different monomer units that act as prebiotics (Castro et al., 2022). HMO can be classified into three groups: 1) HMO decorated with fucose or *N*-Acetylneuraminic acid linked to lactose as a common core, producing neutral or acidic HMO such as 2′- or 3-fucosyllactose (FL), and 3′- or 6′-sialyllactose (SL); 2) type 1 HMO, characterized by lacto-*N*-biose (Galβ1-3GLcNAc) repeats attached to a lactose core, rendering molecules such as lacto-*N*-tetraose (LNT); and 3) type 2 HMO, composed of *N*-Acetyllactosamine units (LacNAc; Galβ1-4GlcNAc) attached to a lactose core, forming molecules such as lacto-*N*-neotetraose (LNnT) (Thomson et al., 2018a).

The *Bifidobacterium* genus is the most dominant in the infant intestinal microbiota, stimulated by HMO in the first years of life (Turroni et al., 2022). Most infant bifidobacteria are well known for their adaptations to the infant gut, displaying several mechanisms for utilization of HMO (B. Zhang et al., 2022). These include the presence in their genomes of several ATP-binding cassette (ABC) and major facilitator superfamily (MFS) transporters,

glycolytic enzymes targeting different linkages in HMO, and feeder pathways deriving HMO molecules to central metabolism (Sakanaka et al., 2020). How bifidobacteria orchestrate molecular responses to HMO using transcriptional factors has been little studied.

RNA-sequencing (RNA-seq) is a powerful high-throughput technology that provides insights into differential gene expression and allows the identification of co-expressed genes in a particular condition (Ozsolak & Milos, 2011). Using RNA-seq data provided from multiple samples, network analysis has been used as an approach to study biological systems. This analysis models the interaction of real biological networks and can be intuitively understood by users (Chung et al., 2021; DiLeo et al., 2011; Kukurba & Montgomery, 2015). In the context of this analysis, a network has a set of nodes represented by genes and a set of edges, indicating significant co-expression relationships. In these networks, there are highly connected nodes (hubs) and a large number of nodes with a small number of connections. Both maintain the structural properties of real networks, such as scale-free topology (Galán-Vásquez & Perez-Rueda, 2019; Stuart et al., 2003).

Weighted gene co-expression network analysis (WGCNA) is a practical methodology for network reconstruction that considers the co-expression patterns between two genes, and the overlapping of neighbor genes (Langfelder & Horvath, 2008). To do this, clusters of co-expressed molecules known as modules are constructed, reflecting different groups (DiLeo et al., 2011; Zhang & Horvath, 2005). The data are included in an adjacency matrix, in which the linkage intensity between genes is defined. Also, a soft threshold parameter is used, which is essential in reconstructing the network. Then, a topological overlap measure (TOM) is implemented as a proximity measure of genes in network modules that combine the adjacency of two genes and the intensity of their connections with neighboring genes. Gene co-expression networks have been used to predict functions of unknown genes, possible relationships with diseases (Liao et al., 2020; Rezaei et al., 2022), or how microorganisms behave in response to intra- or extracellular signals (Duran-Pinedo et al., 2011). Transcriptional factors (TFs) and metabolic enzymes have been found and considered as conserved processes between organisms focused on gene regulation and metabolism (Jha et al., 2020). Due to its relevance in the genome and microbial adaptations, it is essential to evaluate how gene regulation patterns are present to explain the biological activity in each module.

Some studies have demonstrated that the complexity in the structure of HMO drives species of the Bifidobacterium genus to adapt their gene expression for their molecular utilization (Alessandri et al., 2021). These changes allow the production of enzymes necessary for processing these compounds. For example, *B. longum* subsp. *infantis* (*B. infantis*) can internalize intact HMO through ABC transporters, which are broken down into monomers for intracellular metabolization (Masi & Stewart, 2022). In contrast, *B. bifidum* produces extracellular enzymes that degrade the oligosaccharides and internalize the monomers (Kitaoka, 2012b). The subspecies *B. longum* subsp. Longum (*B. longum*) can grow on various HMO structures, with a greater preference to fucosylated HMO. HMO-utilization genes are usually contained in discrete clusters, controlled by predicted TFs, and scattered across the genome (Duar et al., 2020). *B. infantis* ATCC 15697 is characterized by at least four HMO clusters, one devoted to type 1 HMO (LNB/GNB cluster), another for fucosylated and type 2 HMO consumption (HMO Cluster I), and two cluster targeting FL (Dedon et al., 2020; D. A. Sela et al., 2008; Zabel et al., 2020). *B. longum* species are thought to be adapted to the adult gut microbiota, but some isolates also contain HMO clusters allowing type 1 and FL consumption (Sakanaka et al., 2019).

Although molecular strategies for HMO utilization have been well described recently, it is unknown how transcriptional responses to these molecules are orchestrated at genome levels. Furthermore, despite the fact that all these organisms belong to the Bifidobacterium genus, they have unique HMO utilization patterns, so the regulation of these responses might differ. Therefore, in this work, gene co-expression networks based on WGCNA were constructed and studied for three representative species of the Bifidobacterium genus during HMO utilization. Also, co-expressed modules are identified and evaluated with regulatory proteins and metabolic maps to identify unexplored co-expression patterns.

3.2 METHODS

3.2.1 Datasets

The gene expression dataset was obtained from NCBI Geo DataSets (https://www.ncbi.nlm.nih.gov/gds). The dataset includes 18 RNA-seq samples of transcriptome response for *B. infantis* ATCC 15697 (GSE58773) and *B. longum* SC596

(GSE87697), excepting *B. bifidum* SC555 (GSE59053) with 20 samples (mucin as carbon source included) to pooled and individual HMO (Garrido et al., 2015c, 2016). Among evaluated individual substrates were LNT, LNnT, 2FL, 3FL, and 6SL. Pooled HMO were evaluated at the early (OD600 nm = 0.25), middle (OD600 nm = 0.5–0.7), and late time (OD600 nm = 0.9–1.1) points of growth (Supplementary Table S1).

GoodSamplesgenes function on the WGCNA R package was used to inspect the dataset results for missing values. Genes and samples classified as "good genes" and "good samples" were respectively conserved (Langfelder & Horvath, 2008). A schematic representation of all analysis procedures is included in Figure 9.

3.2.2 Construction of co-expression networks

Gene co-expression networks were created through WGCNA in R from normalized samples Figure S1. This package was useful for clustering samples, module distribution, and determining topological properties (Langfelder & Horvath, 2008). Power (β) values were calculated per organism using pickSoftThereshold function (Table 1). Scale-free topology properties of biological networks were added for this purpose. Then, an adjacency matrix was constructed for each bacterial strain using correlation networks, where negative correlations in genes were considered unconnected.

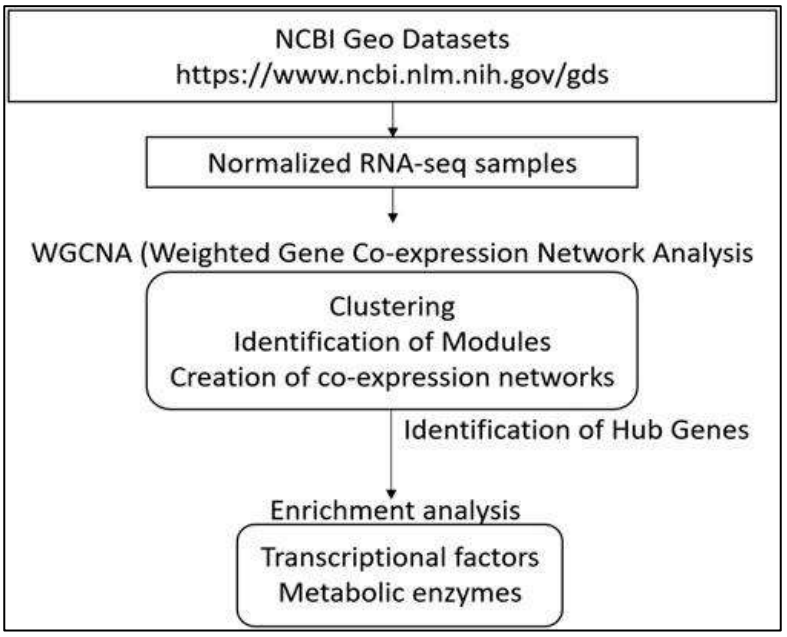

Figure 9- Overview for conducting Weighted Correlation Network Analyses (WGCNA) of different strains of *Bifidobacterium* genus using several HMO.

After, the adjacency matrix was converted into a TOM matrix to minimize noise effects and spurious associations. Higher TOM values allowed the identification of gene modules. Therefore, signed correlation networks were used, pairwise biweight midcorrelation coefficients and β values. Clustered genes were put into modules with analog expression patterns using the average linkage hierarchical clustering algorithm (flashClust function). The cutreeDynamic function cut the dendrogram branches and generated the gene modules. It used 1-TOM as a distance or dissimilarity matrix with a minimum module size equal to 20. Finally, modules with highly correlated eigengenes were merged based on a minimum height of 0.20 (mergeCloseModules function). Each module was differentiated by a specific color, where grey was set for uncorrelated and discarded genes (Horvath, 2011). The rest of the modules were renamed with a number.

Modules were exported using the exportNetworkToCytoscape function to analyze hubs about modules of interest. The 100 most highly correlated genes were chosen for each module. The hub genes were those most highly connected nodes within the module, therefore, the degree of connectivity for each node (K) was calculated, which is defined as the number of edges adjacent to each node (Junker & Schreiber, 2008). For smaller modules, genes were filtered to the top 50% according to the threshold value of the correlation. KEEG ortholog categories were assigned to each gene and the legend was created using Legend Creator App on Cytoscape. A general description of data information of genomes was included in Table 1 and all scripts in Supplementary Section.

Table 1. Overview of dataset and co-expression modules in this study.

Genome	**Sample**	**Genes in WGCNA**	**Size Max/Min genes per module**	***β***	**Grey Module**	**Coverage**	**# Modules without grey**	**# Genes delete in GoodSample Genes**
B. longum infantis **ATCC 15697**	18	2577	516/27	12	0	2577	20	0
B. bifidum **SC555**	20	1894	532/23	12	2	1892	10	0
B. longum **subsp. *longum* SC596**	18	2251	700/66	16	125	2219	11	7

3.2.3 Distribution of TFs and enzymes in modules

Each *Bifidobacterium* genome gene was associated with its Enzyme Commission (EC) number from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa & Goto, 2000). Each enzyme with an E.C. number was related to its respective metabolic map. Similarly, for TFs, we used the compendium of TFs predicted by Flores-Bautista et al., (2020), assigned from the hidden Markov model (HMM) profiles, and orthologous comparisons. The abundance and distribution of each dataset were determined by calculating an incidence rate and heatmap for each genome.

3.2.4 Enrichment analysis

Enrichment analysis was necessary to evaluate the functional relationship between obtained modules, TFs, and enzymes through a hypergeometric test. Statistical significance at a *p*-value of <0.05 was set. The enrichment analysis was performed using in house scripts in Python language (https://www.python.org/).

3.3 RESULTS

3.3.1 Construction of gene co-expression networks

A diagram of the analyses performed in this study is shown in Figure **9**. All RNA-req samples for each *Bifidobacterium* strain were obtained from NCBI Geo Datasets (Garrido et al., 2015b, 2016). For WGCNA analysis, log10 normalized read counts for all samples were taken from NCBI Geo Datasets (Supplementary Figure S1). RNA-seq datasets (GSE58773, GSE87697, GSE59053) were evaluated by sample clustering according to the Euclidean distance between different samples observed for each bacterium (Figure 10). No outliers were detected in the clusters; therefore, 56 samples were used to construct a hierarchical clustering tree (Supplementary Figure S2).

WGCNA identified 20 modules for *B. infantis*, 10 for *B. bifidum* SC555, and 11 for *B. longum* SC596 (Figure 11). The soft-threshold power was adjusted to 12, 12, and 16 for *B. infantis* ATCC 15697, *B. bifidum* SC555, and *B. longum* SC596, respectively. These values were selected to define the adjacency matrix based on the criterion of approximate scale-free topology (Supplementary Figure S3), with a minimum module size of 20, and 0.20 cut

height for merging of modules, which means that the modules whose eigengenes are correlated above 0.80 must be merged (Supplementary Figure S4).

Dominant modules such as Blue and Greenyellow in *B. infantis*, included sugar transport proteins and carbohydrate metabolism genes (Supplementary Table S2; Supplementary Table S3). For *B. bifidum* SC555, the Cyan module was the most extensive module exceeding 500 genes (Figure 11). This module contained diverse categories of biological functions, such as genes encoding sugar transporter proteins, transcriptional regulators, ribosomal proteins, MFS transporters, ATP-binding proteins, and central metabolism genes. However, this module did not contain any HMO-related gene. All modules in *B. bifidum* contained regulatory and metabolic pathway genes. The Darkturquoise and Purple modules included the highest number of HMO genes, with 13 and 12, respectively (Supplementary Table S4; Supplementary Table S5). Finally, in *B. longum* SC596 the Cyan module was the largest (Figure 11). Darkgreen and Green modules in this genome contained carbohydrate metabolism genes specifically for HMO, not being dominant compared to other modules (Supplementary Table S6; Supplementary Table S7).

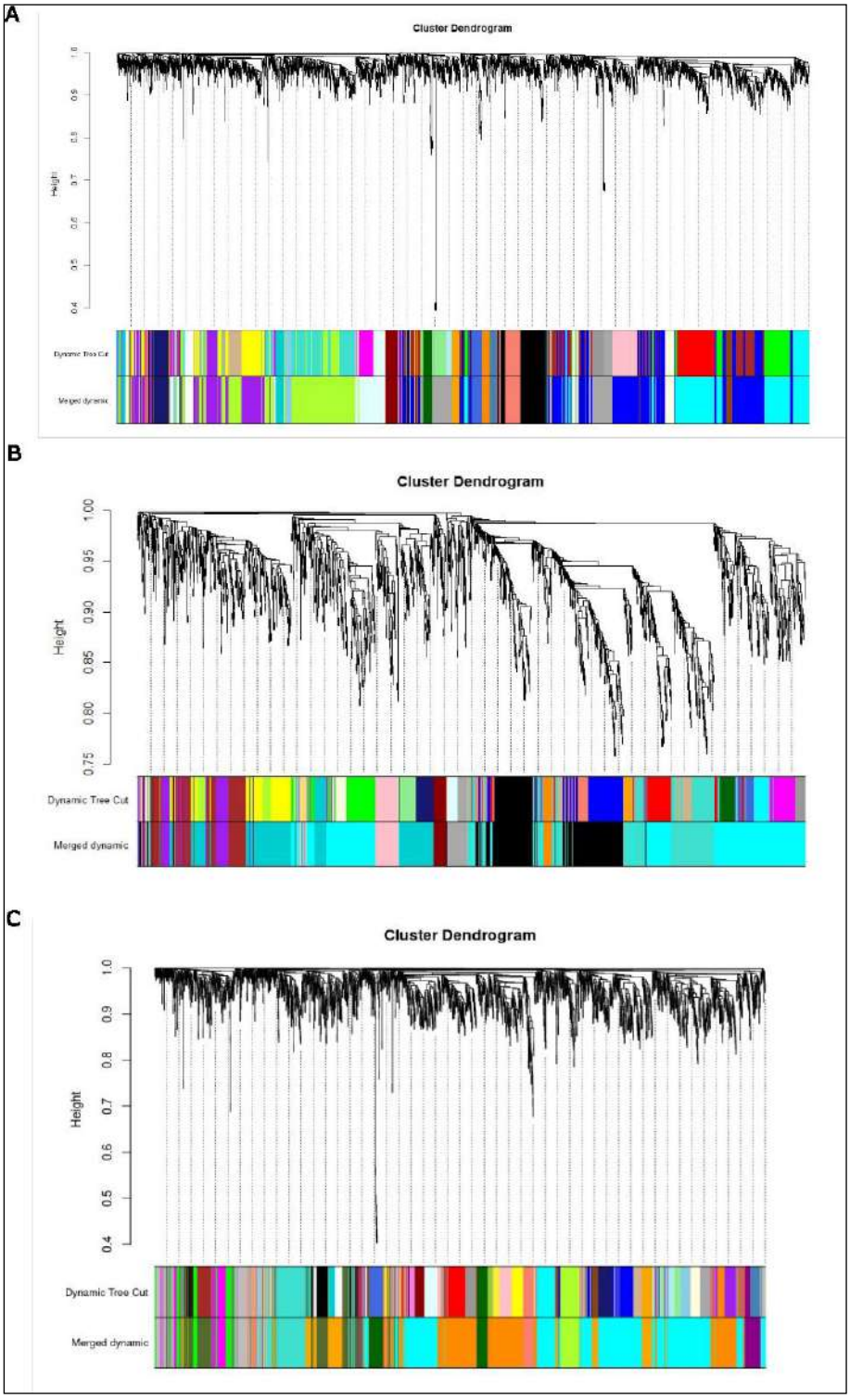

Figure 10- Gene clustering, with dissimilarity based on topological overlap (TOM), with the corresponding module colors indicated by the color row (Merged dynamic).
(A) *B. longum* subsp. *infantis* ATCC 15697. **(B)** *B. bifidum* SC555, and **(C)** *B. longum* subsp. *longum* SC596.

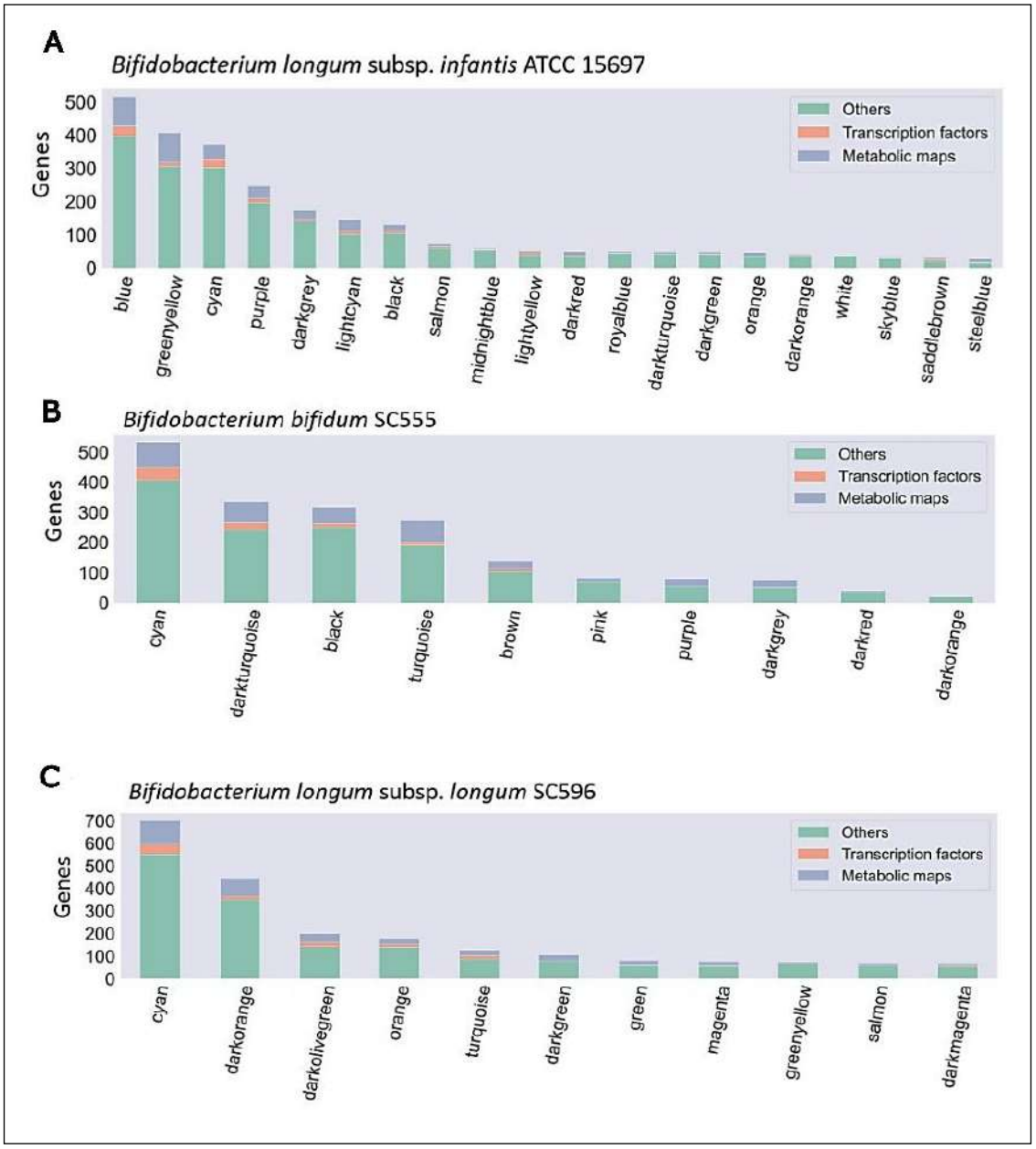

Figure 11- Bacteria co-expression modules in (A) *B. infantis* ATCC 15697, (B) *B. bifidum* SC55, and (C) *B. longum* subsp. *longum* SC596.

3.3.2 Identification of hub genes

Later, the top 100 most highly correlated genes were chosen for each module in order to identify hub genes. These are determined considering the most highly connected node within the module, calculating the degree of connectivity for each node. For smaller modules, genes were filtered to the top 50% according to the correlation threshold value. Hub genes are shown in Supplementary

Table S8 – S10. They contained diverse functions, from a hypothetical protein with a transmembrane domain, ABC permeases not related to HMO utilization, metabolic enzymes, and a transposase in *B. infantis* (Supplementary

Table S8).

Hub genes in *B. bifidum* modules appeared to be related to protein synthesis (Supplementary Table S9) and in *B. longum* were more varied and included proteases, cell division and transport proteins (Supplementary Table S10). These highly connected genes could be important for bifidobacterial physiology or HMO metabolism, and their actual role in these processes could be validated using directed mutagenesis.

3.3.3 Enrichment analysis for TFs and metabolic pathways

Gene regulation and metabolism are among the most conserved processes among microorganisms. They are characterized by DNA-binding regulatory proteins and enzymes involved in metabolic processes (Downs, 2003; Schmid, 2018). The distributions between modules were mapped to determine similar co-expression patterns between metabolism and gene regulation processes. Consequently, TFs from Hidden Markov model (HMM) profiles and enzymes from the KEGG database were associated with the genes of each *Bifidobacterium* strain.

For all three bifidobacterial genomes, enzymes involved in metabolic processes and TFs were found in almost all co-expression network modules (Figure 11). However, some modules contained a significant enrichment (-log10 (*p*-value) > 1.5) in TFs or metabolic pathways across the three genomes (Figure 12). Enriched modules with HMO-associated genes differed from those enriched with TFs (Figure 12).

The most enriched modules with TFs contained on average 15.48% of the predicted genes with this function in *B. infantis* modules, an 8.08% average for *B. bifidum* SC555, and 10.65% average for *B. longum* SC596. Meanwhile, the modules enriched with metabolic enzymes contained on average 27.47% genes predicted to be related to metabolism in *B. infantis*, 29.57% for *B. bifidum* SC555, and 26.20% for *B. longum* SC596. Analyzing the enriched functions identified, there was at least one module with a high percentage of TFs

and enzymes in each bacterium. This case led us to evaluate whether the richer modules prefer a particular TF family or metabolic maps.

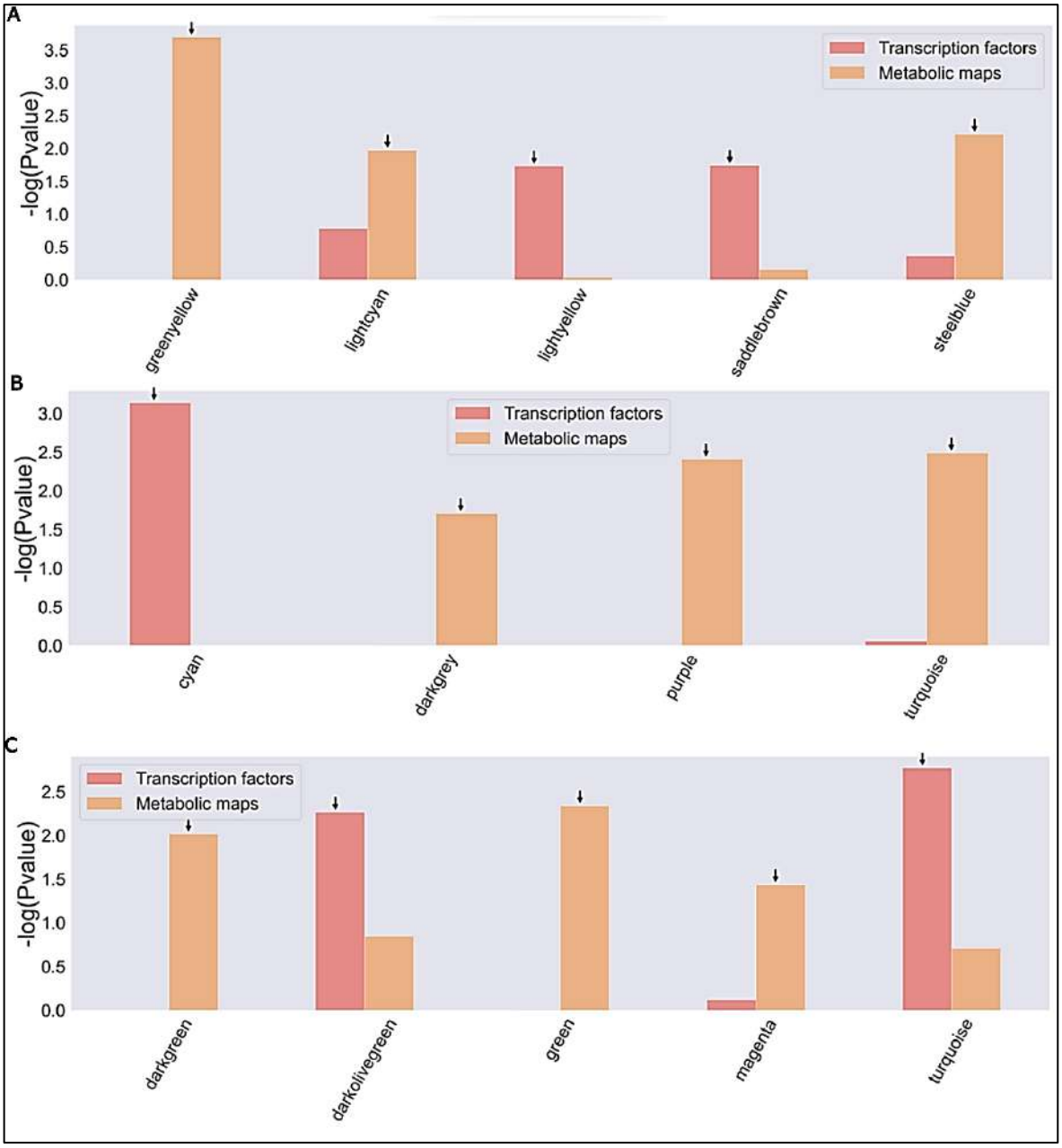

Figure 12- Enrichment of TFs and metabolic maps for Bifidobacterium strains. (A) *B. longum* subsp. *infantis* ATCC 15697, *B. bifidum* SC555, (C) *B. longum* subsp. *longum* SC596.

3.3.4 TFs and bifidobacterial metabolism

Each TF family of the enriched modules in Figure 12 was classified using the Pfam database and the z-score of frequencies of the clustered families (Sun et al., 2017). This clustering was performed by determining the Euclidean distance measure and Ward's method for linkage analysis (Dvorak et al., 2017). Overall, the TF with the highest frequency was the LacI family (PF00356) in each enriched module (Figure 13).

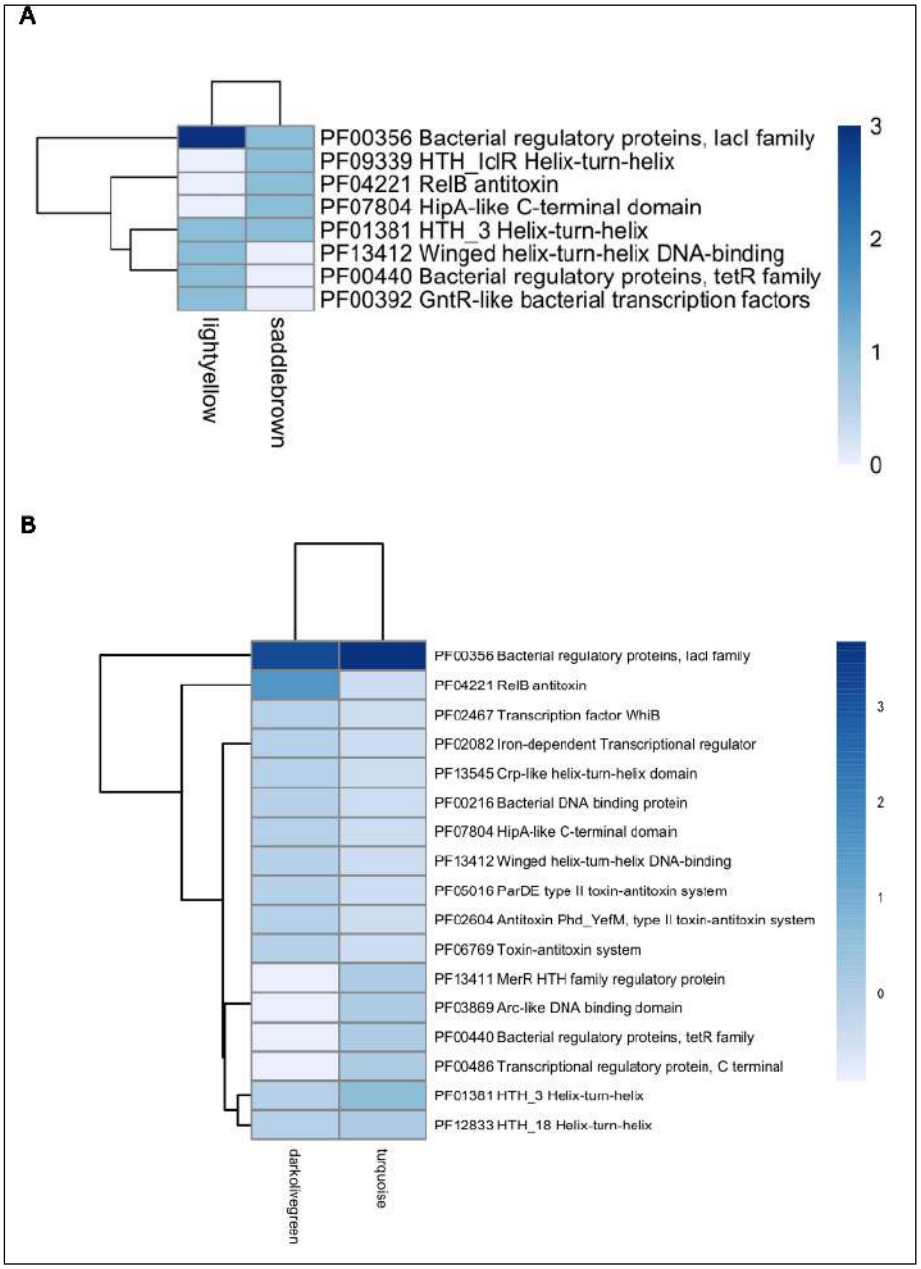

Figure 13- Heat map of TFs abundance for *Bifidobacterium* species. (A) *B. longum* subsp. *infantis* ATCC 15697, (B) *B. longum* subsp. *longum* SC596.

This family is usually associated with regulation of carbohydrate metabolism. Other TFs were identified among evaluated *Bifidobacterium* genomes, such as TetR family (PF00440). TetR was present in the Lightyellow module in *B. infantis*, Cyan module in *B. bifidum* SC555 (Table 2), and Turquoise module in *B. longum* SC596. GntR family (PF00392) was also identified in both *B. infantis* and *B. bifidum* SC555, but not in *B. longum* SC596. Similarly, metabolic enzymes were classified according to the KEGG pathway database (Kanehisa et al., 2019). z-scores of the frequency of each metabolic map were clustered together, similar to those performed for transcriptional factors (Figure 14). Modules significantly enriched were characterized by an overabundance of general metabolic pathways, followed by biosynthesis of amino acids, cofactors, and secondary metabolites. This trend was common for the three transcriptomes studied. Biosynthesis of secondary metabolites was the only category differentially enriched among modules (Figure 14).

3.3.5 Metabolism and regulation for HMO consumption

Later, modules were manually analyzed, and those containing HMO-metabolizing genes were used to create co-expression networks (Figure 15-17). This analysis considered the neighboring genes of each node with the highest correlation. Selected modules were Greenyellow and Blue in *B. infantis*, Darkturquoise and Purple in *B. bifidum*, and Green and Darkgreen in *B. longum* (Figure 15-17). The Greenyellow module network included 22 HMO-related genes (Figure 15A; Supplementary Table S2). These genes include transporters and enzymes for FL metabolism (Blon_0247, Blon_0248, Blon_0341 - Blon_0343, Blon_2202 - Blon_2204), which showed an important degree of co-expression. The Greenyellow network also included the GNB/LNB and type 1 HMO processing cluster (Blon_2172-Blon_2177), and galactose metabolism enzymes (Blon_2062-Blon_2063, Blon_2184) (Supplementary Table S1). Therefore, this module appears to orchestrate metabolic responses to HMO-derived galactose and fucose. Interestingly, a few genes of the complete HMO cluster I (D. A. Sela et al., 2008) appeared in this module (Blon_2331, permease; Blon_2334, β-galactosidase; Blon_2354, SBP; Blon_2355, hexosaminidase). This suggests that the HMO cluster I does not behave as one single transcriptional unit.

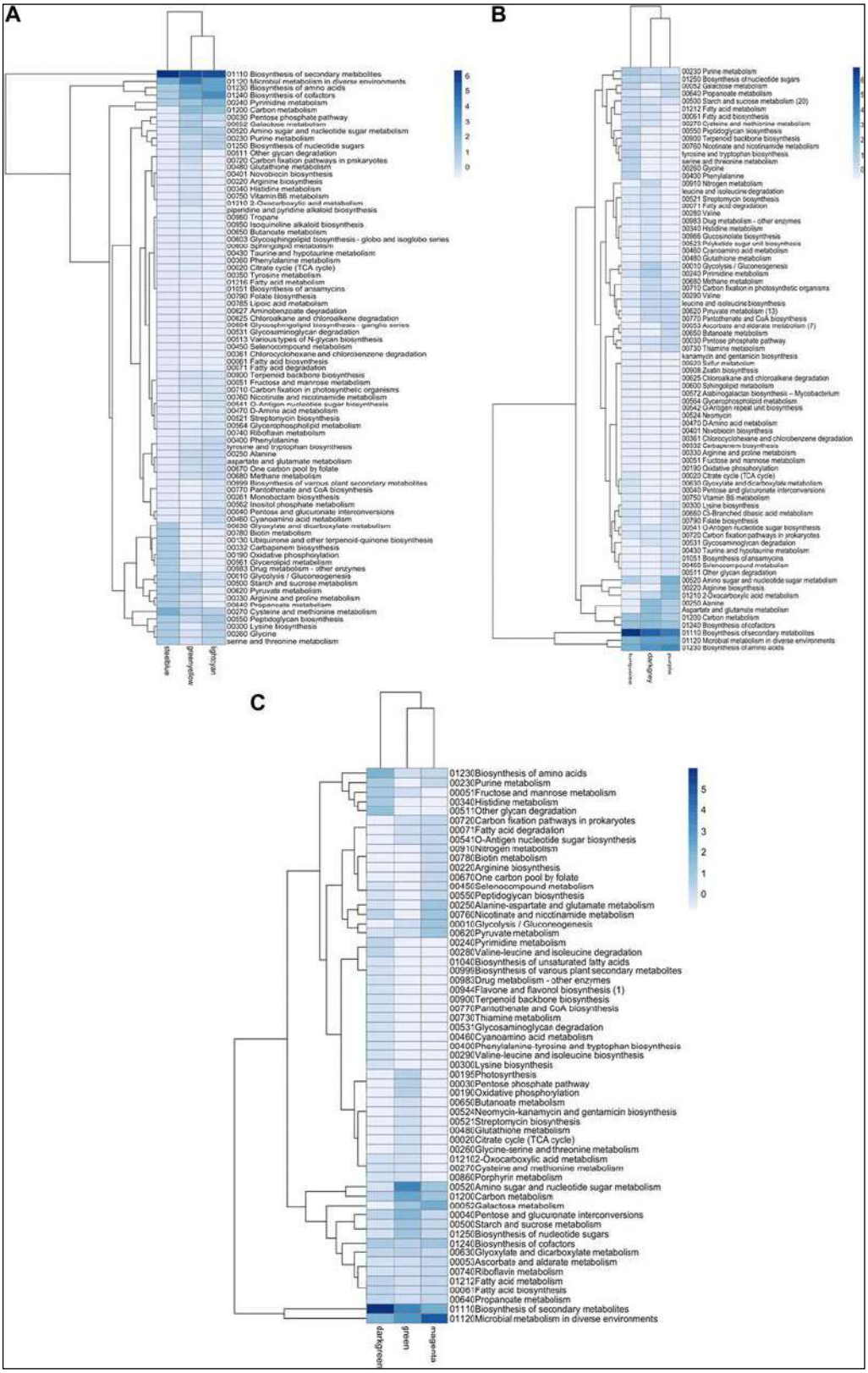

Figure 14- Metabolic map abundance in *Bifidobacterium* species. (A) *B. infantis* ATCC 15697, (B) *B. bifidum* SC555, (C) *B. longum* subsp. *longum* SC596.

Table 2. TF families found in Cyan as an Enriched module in *B. bifidum*.

Name	**Pfam family**	**Frequency**
MerR family regulatory protein	PF00376	1.0
Bacterial regulatory proteins, tetR family	PF00440	2.0
MerR HTH family regulatory protein	PF131411	2.0
GntR-like bacterial transcription factors	PF00392	2.0
MarR family	PF01047	1.0
Transcription factor WhiB	PF02467	1.0
Bacterial regulatory proteins, lacI family	PF00356	5.0
RelB antitoxin	PF04221	3.0
Transcriptional regulatory protein, C terminal	PF00486	3.0
LuxR-type DNA-binding HTH domain	PF00196	3.0
PspC domain	PF04024	1.0
Cold-shock domain	PF00313	1.0
Repressor lexA	PF01726	1.0
Helix-turn-helix	PF09339 PF01381 PF13936	1.0 2.0 2.0
PucR C-terminal helix-turn-helix domain	PF01402	1.0
Ribbon-helix-helix protein, copG family	PF13412	1.0
Winged helix-turn-helix DNA-binding	PF01418	1.0

The Blue module for *B. infantis* contained 19 genes related to HMO, being the second with the highest number of genes of this category (Figure 15B; Supplementary Table S3). This module showed a higher degree of connectivity between nodes compared to the Greenyellow. The remaining genes of the HMO cluster I were included in this module, containing functions such as ABC transporters, fucose metabolism and HMO-glycolytic enzymes. These genes displayed a high co-expression **(**Figure 15B). Outside this cluster, the module also contained a *β-N*-Acetylhexosaminidase (Blon_0732), TFs and enzymes for GlcNAc metabolism (Blon_0879, Blon_0881), and a *β*-galactosidase (Blon_2016).

Blon_0732 showed high co-expression with two ribosomal proteins and an MFS porter. Other single HMO-utilization genes were scattered in several modules (Supplementary Table S**3**).

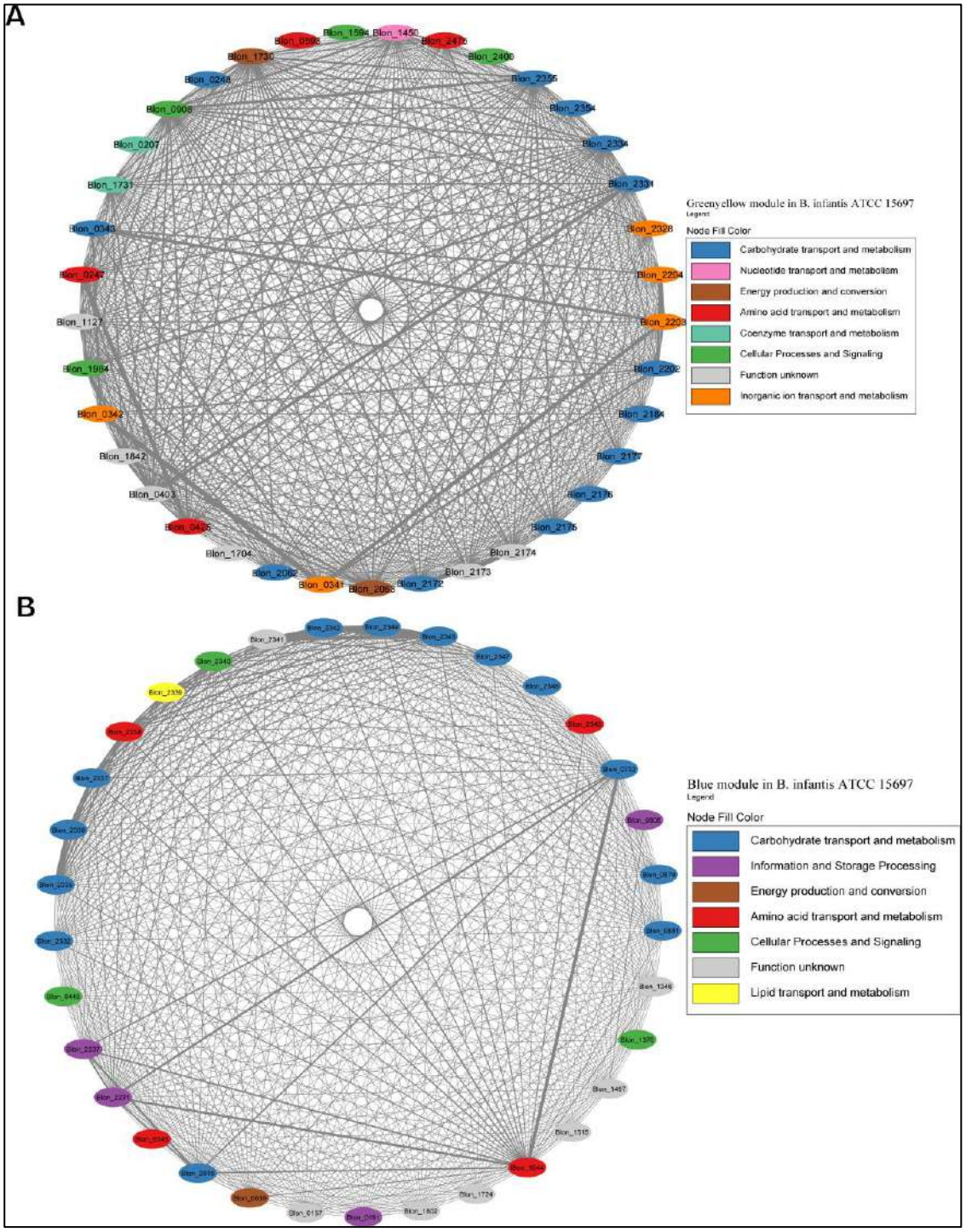

Figure 15- HMO-related co-expression networks in *B. infantis*. (A) Greenyellow and (B) Blue module in for *B. infantis*.

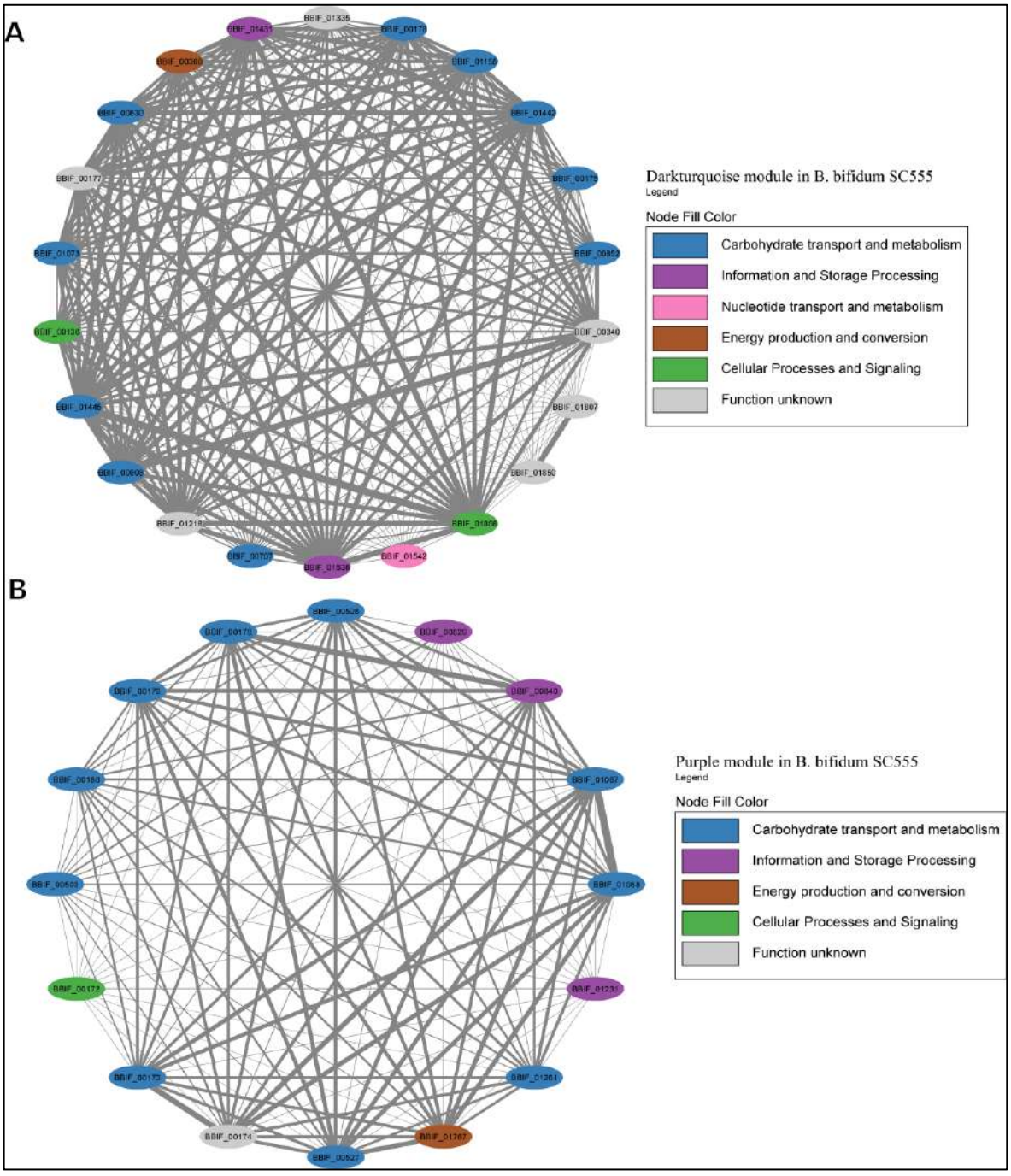

Figure 16- HMO-related co-expression networks in *B. bifidum*. (A) Darkturquoise and (B) Purple module for *B. bifidum* SC555.

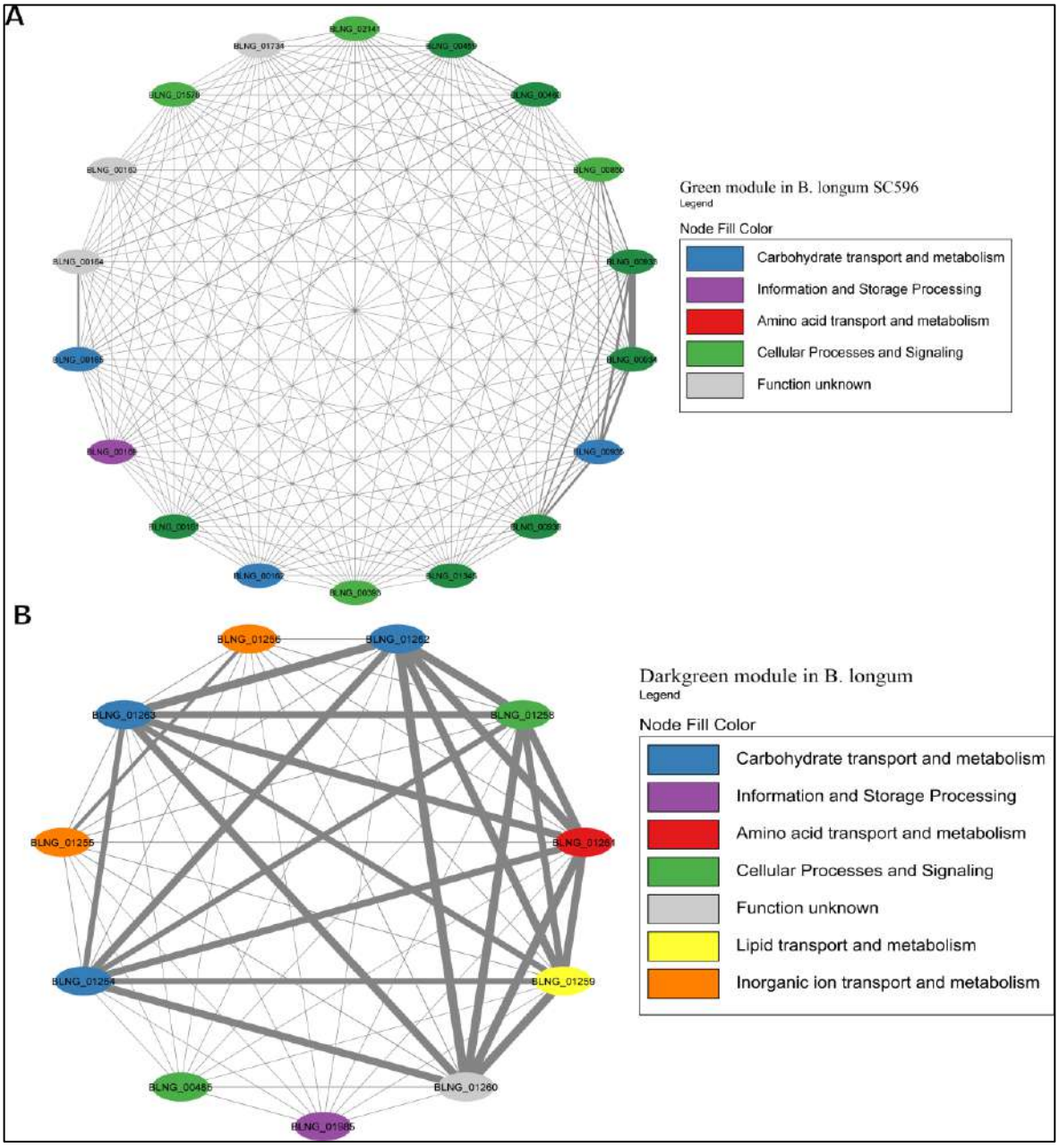

Figure 17- HMO-related co-expression networks in *B. longum*. (A) Green and (B) Darkgreen modules for *B. longum* SC596.

For *B. bifidum*, the highest number of genes related to HMO consumption was included in the Darkturquoise module, with 13 genes (Figure 16A). Modules in *B. bifidum* showed a less structured organization compared to *B. infantis*. The Darkturquoise module contained glycolytic enzymes, TFs for GlcNAc metabolism, and LNB processing enzymes (Figure 16A; Supplementary Table S4). The Purple module in *B. bifidum* contained 12 HMO-utilization genes related to galactose metabolism, transport, GlcNAc metabolism and an α-L-fucosidase (BBIF_01261). These genes showed a high degree of connectivity (Figure 16B; Supplementary Table S5). The Turquoise module contained five genes for galactose metabolism (BBIF_00368, BBIF_00550, BBIF_00871). Other HMO-utilization genes were found dispersed in other modules. Within the network, a total of nine genes unrelated to HMO utilization such as hypothetical proteins, a cell wall biosynthesis protein, and a transcriptional regulator, among others, were identified.

The Green module contained the highest number of genes in *B. longum*, with 12 HMO-related genes (Figure 17A; Supplementary Table S6). This module was only the seventh with the higher number of genes in *B. longum* (Figure 10C), suggesting HMO responses are rather modular in this species and do not co-express with other conserved processes. The Green module contained HMO transport (BLNG_00161, BLNG_00162, BLNG_00933-BLNG_00936, BLNG_01345) and galactose metabolism genes (BLNG_0163-BLNG_00165, BLNG_00459, BLNG_00460). The Darkgreen was only the sixth module with the most genes in *B. longum* (Figure 11) and contained nine genes participating in FL transport (BLNG_01255, BLNG_01256), fucose metabolism (BLNG_01258-BLNG_01262), and α-L-fucosidases (BLNG_01263, BLNG_01264) (Figure 17B; Supplementary Table S7). This modularity and lack of association with other cellular processes suggests a recent acquisition of this gene cluster. Finally, other genes in *B. longum* appeared as pairs or single in other modules.

3.4 DISCUSSIONS

Bifidobacteria are important members of the gut microbiota in both infants and adults. HMO are considered the primary substrates for the abundance of these species in the infant gut (Walsh et al., 2020). Considering the predicted transcriptional responses to *Bifidobacterium* species mediated by the structural complexity of HMO, likewise in order to have a more comprehensive and global understanding of the gene associations involved, gene co-expression networks were constructed for three representative species such as *B. infantis, B. bifidum,* and *B. longum*.

Co-expression networks provide a simpler way to analyze genes that are correlated between biological processes that could be candidates, e.g., for the study of diseases (van Dam et al., 2017). WGCNA analysis has been successfully used as a system biology method for describing expression correlation patterns among genes across RNA samples corresponding to pooled and individual types of HMO generating modules. By clustering the expression data, *B. infantis* possessed a higher number of modules compared to *B. bifidum* and *B. longum* (Figure 11). The number of modules in a co-expression network can be explained by multiple reasons, including physiological aspects such as adaptations to environmental circumstances and resource management between species (Duran-Pinedo et al., 2011). The variety of modules in *B. infantis* could indicate a more complex or diverse expression response compared to *B. bifidum* and *B. longum* (Duar et al., 2020; Horvath, 2011; Zabel et al., 2020). Interestingly, some modules have been identified that group the majority of genes focused on metabolizing HMO.

WGCNA modules may represent independent units responsible for certain biological functions (Jia et al., 2020). This observation was similar to detecting biomarkers linked to gut microbiota (Vernocchi et al., 2020), and biofilm formation genes from the bacterial community (Chen & Ma, 2021; Liu et al., 2021). HMO metabolization could also be a multifactorial process that includes other genes concerned with other biological functions, as observed a co-expression study for *B. longum* FGSZY16M3 (Z. Liu et al., 2021). More experiments are needed to confirm the function of some unknown function genes and the relevance of identified hub genes in these networks.

Among TF regulatory processes found in enriched modules, LacI family (PF00356) was with a high frequency in each genome. LacI family transcriptional regulators are a group

of allosteric DNA-binding regulators with conserved amino acid sequences (Lewis, 2005; Ravcheev et al., 2014). Most of the characterized LacI family transcriptional regulators sense sugar effectors and regulate carbohydrate utilization genes (Tsevelkhoroloo et al., 2021). Likewise, TetR family (PF00440) is known to be involved in multidrug resistance, and presumably controls the glucoside and galactoside utilization pathways (De Bruyn et al., 2013; Ramos et al., 2005). In the *Bifidobacterium* genus, TetR family has within its regulators the BgrT genes are co-localized with genes encoding various *β*-glucoside or *β*-galactoside hydrolases (e.g., bglB, bgaB, bglX, bglY), and *β*-glucoside or *β*-galactoside transporters of the MFS and ABC families. Specifically, BgrT1 regulon has been identified with the *β*-glucosidase gene bglX in *B. infantis*. For *B. longum* NCC2705, *β*-glucosidase genes bglX2 and bglX3 also have been identified with the BgrT2 regulon (Khoroshkin et al., 2016). In the case of *B. bifidum*, no regulator genes or regulons have yet been identified.

Concerning to HMO utilization, *B. infantis* is a dominant strain of the infant gut microbiota that can efficiently consume several classes of structures (Lawson et al., 2020; Ojima et al., 2022). HMO-related genes appeared to be sorted by functionality and activity of HMO clusters in the key modules. In the case of *B. bifidum* SC555, the organization of HMO-related genes can be explained by considering the consumption mechanism. *B. bifidum* prefers short HMO, using enzymes to release fucose and sialic acid decorations on the oligos, which it ultimately does not use. This preference may also be directed to other carbon sources (Garrido et al., 2015). It should be highlighted that the HMO cluster I is found specifically in *B. infantis*, while fucose clusters are found in some *B. longum*, B. breve and *B. bifidum* strains. The LNT utilization cluster (Blon_2171—Blon_2177) is present in most infant-associated bifidobacteria (Zúñiga et al., 2018b), and in *B. infantis* clustered with 2FL and galactose pathways.

A few *B. longum* strains can utilize certain HMO, especially LNT and 2FL (Díaz et al., 2021). However, this subspecies is adapted to the adult gut microbiome. *B. longum* SC596 is a strain isolated from an infant. *B. longum* key modules appeared to be closer to the HMO utilization genes, considering that Green module contained HMO transport and galactose metabolism genes (Figure 13A), and Darkgreen module was addressed to genes of FL metabolism and transport, including additionally fucosidases (Figure 17B). Previous transcriptomic analyses have revealed the affinity of *B. longum* to consume fucosylated

HMO allowing the strain to be more selective to HMO consumption compared to *B. infantis*, being an evolved strain (Garrido et al., 2016). This observation points at the Darkgreen as a single fucose metabolism module.

Regarding hub genes among three evaluated genomes (Supplementary Tables S8–S10), hypothetical, transport, and ribosomal proteins were identified as the most connected genes among modules. Hypothetical proteins could take a role as glycosyl hydrolase in the utilization of HMO. However, future analysis will allow us to determine their particular functional role (Galán-Vásquez & Perez-Rueda, 2019). Likewise, inside smaller modules, some hub genes with 3, 4, 5, 6, and 7 degrees of connections were identified, considered the most important in the co-expression network.

This study illustrates how HMO genetic responses in *Bifidobacterium* are coordinated according to the constituent monosaccharide: galactose, GlcNAc, and fucose responses usually appear in distinct modules and clusters. This correlates with monosaccharides potential role in activating or repressing TFs and triggering the expression of cognate clusters. This organization in the modules has also been previously observed, in the transcriptional response in denitrifying bacteria on carbon nanotubes, whereby WGCNA they have obtained specific modules corresponding to the activity of various types of carbon nanotube structures (Zheng et al., 2018).

Despite the extensive analysis of this study, there are still some limitations, such as low reports of WGCNA studies in bifidobacteria and the deficiency in gene annotation of novel isolates bacteria. More datasets from *Bifidobacterium* species under different conditions are required to refine their regulatory networks. In some cases, the most important hub and non-HMO-related genes encodes for hypothetical proteins, allowing future analysis to determine their functional role. Also, experimental validation of co-expression networks and incorporation of different species in utilizing HMO could further provide a widely understanding. These experiments could be combined with mutational analysis of hub or essential genes. The results obtained by WGCNA in the Bifidobacterium strains tested with different HMO provide evidence of expression patterns identified with genes unrelated to HMO metabolism.

3.5 CONCLUSIONS

In this work, we identified and analyzed modules considered metabolically and regulatory relevant in a set of infant bifidobacteria, using a weighted gene co-expression analysis method (WGCNA) in the context of HMO utilization. From this analysis, we identified some modules enriched with TFs and metabolic enzymes. In the case of regulation, we identified TFs from the LacI, TetR, RelB, and HTH_3, GntR families, which are related to sugar utilization and biological processes, such as biosynthetic processes, and cellular metabolic processes. Our approach also identified genes involved in similar metabolic or regulatory functions. Among modules for *Bifidobacterium* strains, ABC and superfamily MFS transport proteins, transcriptional regulators such as TetR were identified as hubs genes because of their high correlation with other genes. The networks generated allowed us to identify co-expressed genes involved in responses to HMO consumption. Substantial differences were found in the structure of modules and regulation across these three *Bifidobacterium* species. In summary, this analysis allowed us to determine that, despite the diversity of experimental information available for each organism, these mechanisms are similar in all organisms, which will allow us to address new experimental results, such as the use of gene expression data in metagenomic studies.

4 CHAPTER III: CROSS-FEEDING INTERACTIONS OF GUT MICROBES MEDIATED FROM *O*-LINKED GLYCANS CASEIN GLYCOMACROPEPTIDE.

4.1 INTRODUCTION

Human gut microbiota (HGM) is a complex community of microorganisms that includes mainly bacteria, yeast, viruses, and some eukaryotes (Passos & Moraes-Filho, 2017). HGM consists of more than 1500 species distributed in more than 50 different phyla (Conz et al., 2023; Hou et al., 2022). *Bacteroidota* and *Bacillota,* followed by *Pseudomonadota*, *Fusobacteriota*, *Mycoplasmatota*, *Actinomycetota*, and *Verrucomicrobiota,* are the most dominant phyla, making up to 90% of the total microbial population in humans (Sekirov et al., 2010). The gut microbiota has many significant functions in the human body, vital to digestion and metabolism (Vernocchi et al., 2020). Among several factors that can change the gut microbiota composition and function, diet is one of the most influential (Hasan & Yang, 2019). Currently, there is a great interest in the impact and modulation of the gut microbiota by the daily diet.

In addition to inulin and fructo-oligosaccharides (FOS), which are carbohydrates commonly studied as prebiotics (BeMiller, 2019; Parhi et al., 2021), there are other carbon sources attached in some proteins that could play a prebiotic role. These protein-glycan bonds are caused by glycosylation, a post-translational modification process where oligosaccharides are covalently bound to asparagine (*N*-glycans) (La Rosa et al., 2022), or serine or threonine (Ser/Thr; *O*-glycans) (Zhang et al., 2021). These complex glycans serve as signaling molecules for secretion and other cellular processes, increasing proteolysis resistance (Baum & Cobb, 2016; Goettig, 2016). These glycans are being proposed as emerging prebiotics due to their similarity to host-derived glycans compared to plant-derived prebiotics and their enrichment of dominant and health-promoting gut microbes, such as *Bacteroides* spp., *Bifidobacterium* spp. (Cunningham et al., 2021; Kirmiz et al., 2018).

N-linked glycans possess *N*-acetylglucosamine (GlcNAc) as a common core, conjugated with an additional GlcNAc and three mannose (Man) residues, forming a $Man_3GlcNAc_2$ motif found in all *N*-glycans. *N*-glycans are an example of host-derived

oligosaccharides that can be used by beneficial microbes. *Bifidobacterium longum* subsp. *infantis* ATCC 15697 (*B. infantis* ATCC 15697) is a dominant beneficial infant gut microbe that has been shown to access *N*-glycans *in vitro* by a specialized endo-*β*-*N*-Acetylglucosaminidase (Garrido, Nwosu, et al., 2012a). In contrast, less attention has been paid to *O*-linked glycans. These oligosaccharides are characterized by an *N*-Acetylgalactosamine (GalNAc) residue linked to Ser/Thr as a common core. GalNAc is usually bound to Gal*β*1-3, forming galacto-*N*-biose (GNB) as a building block in Core 1 and Core 2, which could be further extended, forming larger chains similar to *N*-glycans. *O*-linked glycans are a significant component of mucins and *κ*-caseins (Magnelli et al., 2011).

Glycomacropeptide (GMP) is a constituent of whey (20–25% of the protein moiety) derived from *κ*-casein after chymosin treatment during cheesemaking (Keel et al., 2022). Interestingly, GMP is a hydrophilic, negatively charged, 64 amino acid *O*-glycopeptide with five different types of *O*-linked cores: 1) GalNAc-*α*-Ser/Thr; 2) Core 1 Gal*β*1-3GalNAc-*α*-Ser/Thr; 3) Core 2 Gal*β*1-3(NeuAc-*α*2-6) GalNAc-*α*-Ser/Thr; 4) Core 3 NeuAc-*α*2-3Gal*β*1-3GalNAc-*α*-Ser/Thr; 5) Core 4 NeuAc-*α*2-3Gal*β*1-3GalNAc(NeuAc-*α*2-6)-*α*-Ser/Thr (González-Morelo et al., 2020b). The molecular weight of GMP ranges from 7 to 11 kDa (Córdova-Dávalos et al., 2019b), depending on the level of glycosylation and isolation method (E. W. Y. Li & Mine, 2004). *O*-glycosylated proteins such as GMP have a higher resistance to proteolysis (Amore et al., 2017), probably reaching lower sections of the gut and interacting the gut microbiota.

On the other hand, bacteria can form trophic networks of metabolic interactions from degradation and fermentation products (Gralka et al., 2020; Saa et al., 2022). These dominant metabolic interactions found in the gut microbiome may be by competition (Ding et al., 2021), which practically refers to one species being able to efficiently consume a resource required by a second species. Likewise, cooperative cross-feeding between two species may also occur (D'Souza et al., 2018). In this last one, prebiotics ingested from the diet can be degraded by certain intestinal microorganisms (primary fiber degraders). At the same time, by-products can be utilized by other members of the community (secondary degraders) that could not have accessed them prior to the degradation of complex substrates (Vega-Sagardía, Cabezón, et al., 2023; Vega-Sagardía, Delgado, et al., 2023). These metabolic interactions are a key part of the assembly and functioning of the gut microbiome influenced by dietary

compounds. Differences in abundances may reflect both direct associations with the gut environment and indirect effects due to microbiome interactions. Released monosaccharides, disaccharides, and SCFA (acetate, lactate, butyrate, propionate, and succinate) are fermentation by-products allow cross-feeding during metabolic interactions (C. T. Peterson et al., 2022; Saa et al., 2022).

Previous studies have shown that *B. infantis* ATCC 15697 is able to grow using the glycosidic fraction of GMP (O' Riordan et al., 2018). Several molecular adaptations for accessing and consuming *N*- and *O*-glycans have been described in gut microorganisms, revealing the utilization of intra- and extracellular enzymes (Morozumi et al., 2023). Considering the high abundance of *O*-glycans in host secretions, it is not surprising that gut microbes have evolved strategies to utilize *O*-linked glycans from GMP as a carbon source. However, little is known regarding the enzymes and molecular mechanisms involved in this process. Moreover, how *O*-glycans promote ecological interactions of cross-feeding or competition is not known. In order to gain a broader understanding of how members of gut microbiota use *O*-linked glycans, GMP could be used as a model for the study of metabolic interactions between microorganisms. In this study, we determined the molecular strategies and metabolic interactions of the members of the representative members of HGM during GMP utilization as a carbon source.

4.2 METHODS

4.2.1 Bacterial strains and culture media

Collection strains used in this study were acquired from BEI Resources, the ATCC, or the UC Davis Culture Collection. Isolate strains were obtained from Chilean subjects and stored at -80 °C in the Microbiology and Systems Laboratory of the Pontifical Catholic University of Chile (Table 3). GMP with 1% of lactose was obtained from Agropur Ingredients (Eden Prairie, MN).

First, microbes were cultured in their preferred culture media described in Table 3. *Clostridium* medium (RCM; Becton, Dickinson, Franklin Lakes, NJ) was supplemented with 0.5 g/L of L-cysteine (Loba Chemie, India). De Man, Rogosa, and Sharpe medium (MRS; Becton, Dickinson, Franklin Lakes, NJ) was supplemented with 0.5 g/L of *L*-cysteine. All

incubations were performed at 37 °C for 24 h to 48 h in an anaerobic jar (Anaerocult; Merck, Darmstadt, Germany) with anaerobic packs (GasPak EM; Becton, Dickinson, Franklin Lakes, NJ).

Table 3. Description of the strains used in the study.

Phylum	Strains	Abbreviation	Culture media	Source
Actinomycetota	*Bifidobacterium longum* subsp. Infantis ATCC 15697	Bi	MRS – cys	UC Davis collection
	Bifidobacterium longum subsp. *Longum* M12	M12	MRS – cys	Chilean isolate*
	Bifidobacterium breve I1	I1	MRS – cys	Chilean isolate*
	Bifidobacterium bifidum JCM 1254	JCM 1254	MRS – cys	UC Davis Collection
Bacteroidota	*Bacteroides thetaiotaomicron* VPI-5482	Bt	RCM	UC Davis Collection
	Bacteroides fragilis J4	J4	RCM	Chilean isolate*
Bacillota	*Lachnoclostridium symbiosum* WAL-14673	Ls	RCM	BEI resources
	Ruminococcus gnavus CC55_001C	Rg	RCM	BEI resources
	Streptococcus thermophilus LM5	LM5	RCM	Chilean isolate*
	Streptococcus thermophilus LM8	LM8	RCM	Chilean isolate*
	Clostridium perfringens ATCC 13124	Cp	RCM	UC Davis collection
	Enterococcus faecalis H1	H1	RCM	Chilean isolate*

*The isolated strains were obtained from Chilean subjects (Thomson et al., 2019).

4.2.2 Individual assays of strains using GMP

Microorganisms listed in Table 3 were inoculated in optimized formulation mZMB (Medina et al., 2017), with a fixed Ph of 5.5 and supplemented with GMP (10mg/mL) as a carbon source. GMP concentration was chosen from preliminary experiments (data not provided). GMP used as a substrate was previously filtered using 0.22-μm filters (Jet Biofil, China). mZMB with/without lactose (1 g/mL) were used as a positive and negative control, respectively. Microbes were individually grown in 96-well plates with a 200 μL mZMB, 6% inoculum, and 30 μL mineral oil. Experiments were conducted under anaerobic conditions with nitrogen at 37 °C for 48 h using the Bactron-300 Anaerobic Chamber (Sheldon Manufacturing, Cornelius, OR, USA). Optical density (OD_{620nm}) was measured through a Tecan F50 Robotic (Infinite®, Männedorf, Switzerland) inside the Bactron 300. The OD obtained for each strain grown on the different substrates was compared with the OD

obtained in the absence of a sugar source. The level of growth was classified as follows: negative (maximum $OD_{620nm} < 0.250$); low (OD_{620nm} from 0.250 to 0.500); moderate (OD_{620nm} from 0.500 to 0.800); high ($OD_{620nm} > 0.800$).

4.2.3 Unidirectional and Bidirectional bacterial assays

GMP-degraders and non-degrader were determined according to 4.2.2. For unidirectional experiments, GMP-degrader bacteria were anaerobically incubated in GMP-mZMB cultures at 37 °C for 24 h. Subsequently, GMP-mZMB cultures were centrifugated at 4000 x *g* for 2 min, pellets were discarded, and supernatants were recovered and filtrated for use. Then, 6% inoculum of each strain was incubated with 200 µL mZMB spent supernatants and 30 µL mineral oil in a 96-well plate. Lactose- mZMB supernatant was used as a positive control and only mZMB as a negative control. For unidirectional assays were carried out for 48 h with anaerobic conditions.

For bidirectional assays, 24-Transwell tissue culture plates with 0.1 µm semi-permeable membrane insert were used as a co-culture system (Supplementary Figure S5) (Corning Costar, Cambridge, MA). Bacterial combinations for inoculation were made per duplicate (Supplementary Table S11). Prior to the experiment, the cell density of the cultures used for apical (AW), and lower (LW) wells was equalized to 1.0 OD_{620nm}. Strains were seeded with 6 % v/v inoculum in both the apical and lower. Bidirectional assays were performed under anaerobic conditions at 37 °C. OD_{620nm} was measured at 0, 24, and 48 h by resuspending the content of each well or by transferring 200 µL to a new 96-well plate. OD_{620nm} was measured through a Tecan M200 Robotic (Infinite®, Männedorf, Switzerland) inside the Bactron 300. The mean of each basal medium without bacteria was subtracted for the final OD values of both cultures according to the carbon source. A paired Student's *t*-test assuming equal variances was performed on the maximum OD value for each condition (Hirmas et al., 2022). Schematic representation of both types of assays can be found in Figure S5.

4.2.4 Thin Layer chromatography (TLC)

Cultures from the above experiments were centrifugated at 12000 x *g* for 2 min. Pellets were discarded and supernatants were stored at -20 °C for use. For visualizing sugars in supernatants, 2 µL for each was placed in an Aluminum silica gel plate (Merck, Darmstadt, Germany). Standard sugars such as glucose, galactose, lactose, galacto-*N*-biose, and GalNAc were included at 1mg/mL. 1-butanol: acetic acid: water (2:1:1 ratio, by volume) was used as a developing solvent. After drying, plates were sprayed with 0.5% α-napthol and 5% H_2SO_4 in 95 % ethanol. Plates were dried and revealed at 150 °C for 10 min (Garrido et al., 2012).

4.2.5 Sialic acid quantification

The sialic acid content in the individual and bidirectional assays was quantified enzymatically (Sigma MAK-314). Measurements were performed on the samples using the fluorimetric procedure with λ_{ex} = 555/λ_{em} = 585 nm using Synergy H1 (Plate Reader; BioTek, USA). A hydrolyzation treatment was necessary to determine the total sialic acid content. Free sialic acid content in GMP was quantified by discarding the previous pre-treatment.

4.2.6 Quantification of SCFA by HPLC

Acetic, lactic, propionic, and butyric acid were quantified at the end of each bidirectional assays by high-performance liquid chromatography (HPLC) according to a protocol previously reported (Gutiérrez & Garrido, 2019; Mendoza et al., 2017). An ion exchange column of organic acids and carbohydrates, Aminex HPX-87H (Bio-Rad, Hercules, CA), was used in a Lachrom L-700 HPLC system (Hitachi, Japan) at 35°C with a flow rate of 0.45 mL/min of 5 Mm H_2SO_4. As standard, concentrations from 30 g/L to 0.155 g/L of each acid were used.

4.2.7 RNA extraction and reverse transcription

RNA was extracted from the pellets of the samples analyzed by HPLC. Total RNA was extracted E.Z.N.A.® Bacterial RNA Kit (Omega Biotek, Norcross, GA, USA). Purity

was determined through a 260/280 absorbance ratio in a Tecan Infinite M200 Pro plate reader (Tecan, Austria). After extraction, samples were immediately treated with DNase I (New England BioLabs) using the manufacturer's protocol for 15 min. RNA was converted into cDNA with the SensiFast™ cDNA Synthesis Kit (Bioline, UK), using the manufacturer's protocol and random primers and negative control with free nuclease water (Sigma-Aldrich) and an RNase block control for two of each bacterial sample. The final quality of cDNA was assessed by measuring the 260/280 absorbance ratio in a Tecan Infinite M200 Pro plate reader, and absorbance curves from 230 nm to 300 nm were obtained to rule out chemical contamination. All cDNA samples were stored at −20°C until use.

4.2.8 Quantitative PCR and gene expression

Primers for calculating changes in gene expression were designed using Primer-BLAST for select genes (Supplementary Table S12). Four genes (Blon_2348: Sialidase-1, Blon_1831: GH25 hydrolase, Blon_0646: sialidase-1, Blon_2177: Multisugar transport protein) were chosen for *B. infantis* ATCC 15697, and three genes (Ga0128330_11411, Ga0128330_10636: exo-*α*-sialidase, Ga0128330_104513: endo-*α*-*N*-GalNAcase, Ga0128330_100248: ABC Transport protein) for *B. bifidum* JCM 1254 were chosen. These genes were chosen considering their predicted activity in *O*-linked glycans.

Quantitative PCRs (qPCRs) were prepared using the SensiFAST SYBR No-ROX kit (Bioline, UK) under the manufacturer instructions and 2,2 µl of each cDNA. Amplification was performed in an AriaMx real-time PCR system (Agilent Technologies), using 96-well optical plates MicroAmp Fast Optical (ThermoFisher, USA). Reactions were carried out in a two-step cycling format, with an initial cycle at 95°C for 2 min, and 40 cycles at 95°C for 5 s (denaturation), 60 °C for 30 s (annealing/extension). All conditions had two biological replicates with three technical replicates. A standard curve was included by loading 10-fold dilutions of genomic DNA, and a negative control was included for each pair of primers. Threshold cycle (C_T) values and efficiency of the reactions were calculated using the Agilent Aria 1.7 software. Changes in gene expression were calculated using the efficiency-corrected method (Pfaffl, 2001). We separated the effect of the carbon source from the culture with the other bacteria, so the monoculture was used as the basal condition for the

cocultures in each carbon source, and the monoculture in mZMB was used as the basal condition for the monocultures. The efficiency of each PCR was obtained from the standard curve and considered equal to all samples for that gene. Technical replicates were averaged, but the ratio was calculated independently for biological replicates and then averaged. The ANOVA test was used to evaluate global significance under all conditions and paired *t* test between each pair of conditions when the ANOVA *P* value was smaller than 0.1.

4.3 RESULTS

4.3.1 Individual growth of intestinal microbiota strains with GMP

To investigate the GMP consumption by different strains of the HGM, we anaerobically cultured 12 bacterial strains in GMP-mZMB (10 mg/mL) during 48 h at 37 °C Figure 18). These strains belong to the Actinomycetota phylum (*B. infantis* ATCC 15697, *B. longum* M12, *B. bifidum* JCM 1254, *B. breve* I1), Bacteroidota (*B. thetaiotaomicron* VPI-5482, *B. fragilis* J4), and Bacillota (*Lachnoclostridium symbiosum* WAL-14673, *Clostridium perfringens* ATCC 13124, *Ruminococcus gnavus* CC55_001C, *Streptococcus thermophilus* LM5, *Enterococcus faecalis* H1).

B. infantis ATCC 15697 and *B. bifidum* JCM 1254 exhibited high growth on GMP with maximum OD_{620nm} values of 0.932 ± 0.04 and 0.9638 ± 0.1, respectively (Figure 18). Other strains such as *B. thetaiotaomicron* VPI-5482, *B. fragilis* J4, *C. perfringens* ATCC 13124, *E. faecalis* H1, and *R.* gnavus CC55_001C also displayed a high growth with an average OD_{620nm} values exceeding 0.8 on GMP. In contrast, *B. breve* I1 and *B. longum* M12 did not growth on this substrate. Similarly, *S. thermophilus* LM5, S. *thermophilus* LM8, and *L. symbiosum* WAL-14673 showed no grow with GMP. From the above, it can be noted that *B. infantis*, *B. bifidum*, *B. thetaiotaomicron* and *C. perfringens* were selected as GMP degraders for the unidirectional assays.

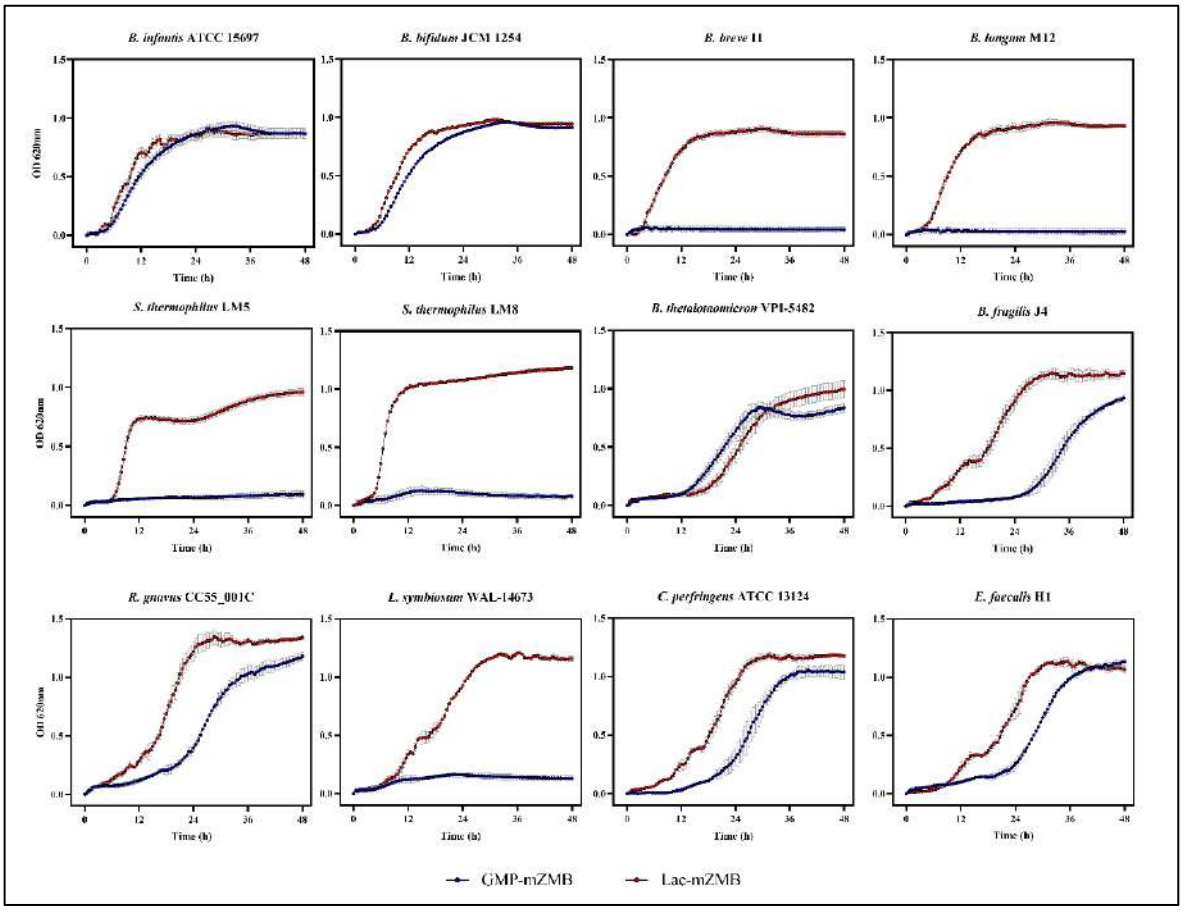

Figure 18- Growth of bacterial strains using GMP as a carbon source.
mZMB with/without lactose (1 g/mL) were used as positive and negative experimental controls, respectively. Microorganisms were individually grown in 96-well plates with 200 µL GMP-mZMB, 6% inoculum, and 30 µL mineral oil. Experiments were carried out under anaerobic conditions.

Regarding TLC assays of the supernatants from individual cultures (Supplementary Figure 19), a GNB band was visualized considering the same height with the standard. For *B. infantis* ATCC 15697 and *B. bifidum* JCM 1254, the GNB was present in the 12 h-point growth, and then disappeared in 24 h- and 48 h-point growth. The same observation was made for *C. perfringens* ATCC 13124. For *B. thetaiotaomicron* VPI-5482 and *B. fragilis* J4, the GNB band was slightly identified in 12h- and 24h- point growth. These results suggest that GNB is released from GMP (intra or extracellularly), but it is imported and consumed by these microorganisms. No spent free sugars were detected in the supernatants of *B. breve* I1, *B. longum* M12, *S. thermophilus* LM5/LM8, *L. symbiosum* WAL-14673, and *E. faecalis* H1.

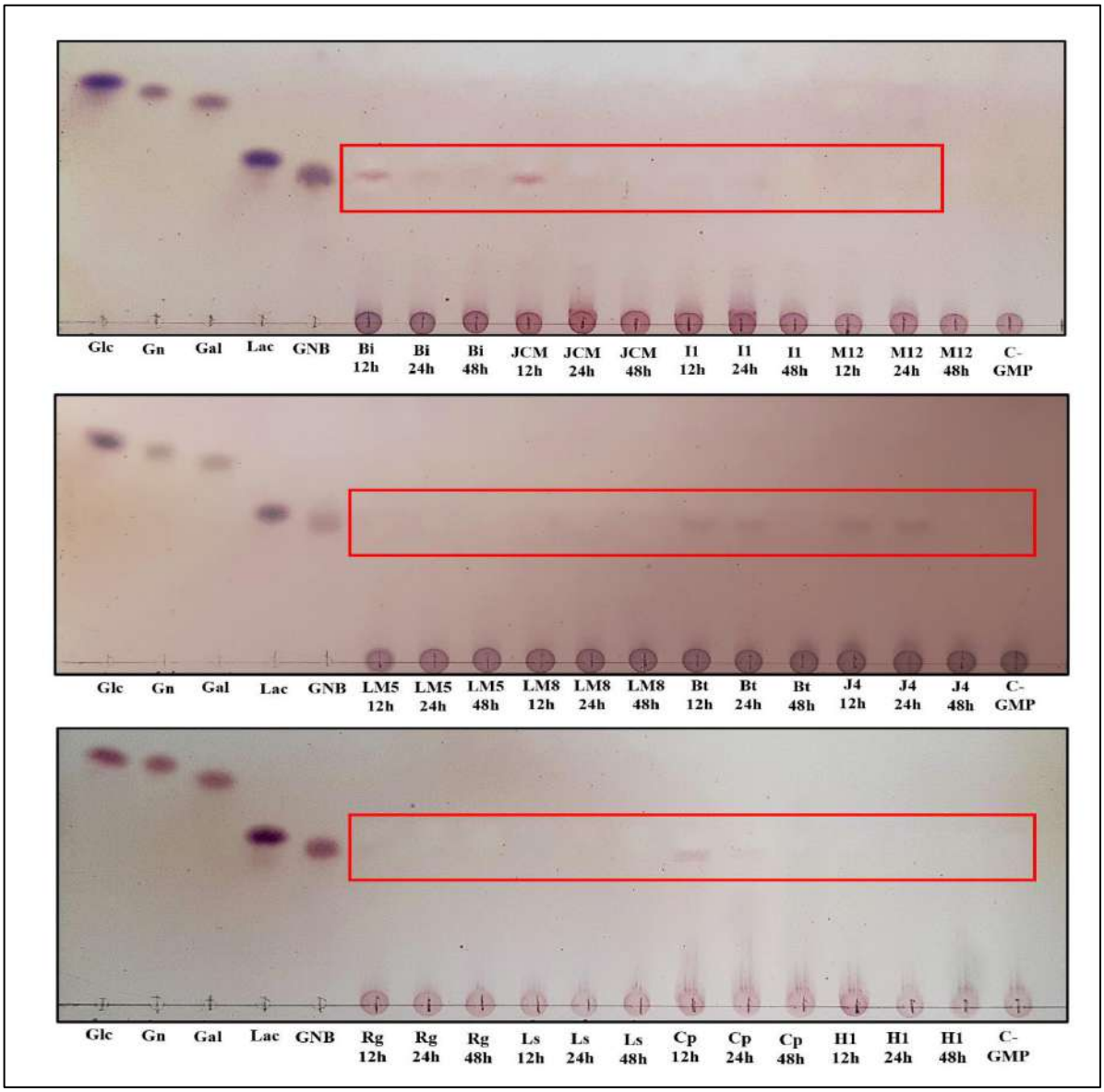

Figure 19- Thin Layer Chromatographic of individual point growth of strains using GMP as a carbon source over time.

Supernatants of individual cultures were recovered after centrifugation at 4000 x *g* for 2 min. Standard sugar such as glucose (Glc), *N*-Acetylgalactosamine (Gn), galactose (Gal), lactose (Lac), and galacto-*N*-biose (GNB) were used at 1 mg/mL. Bi: *B. infantis* ATCC 15697; JCM: *B. bifidum* JCM 1254; I1: *B. breve* I1; M12: *B. longum* M12; LM5: *S. thermophilus* LM5; LM8: *S. thermophilus* LM8; Bt: *B. thetaiotaomicron* VPI-5482; J4: *B. fragilis* J4; Rg: *R. gnavus* CC55_001C; Ls: *L. symbiosum* WAL-14673; Cp: *C. perfringens* ATCC 13124; H1: *E. faecalis* H1.

Total sialic acid was quantified in these strains supernatants at 12, 24, and 48 h (Figure 20). Sialic acid showed an average of 300 µM ± 10.64 and free sialic acid 20 µM ± 6.18 for GMP-mZMB. *B. infantis* ATCC 15697 showed a significative but small consumption of sialic acid after 48 h. Because *B. infantis* ATCC 15697 has no extracellular sialidases (Sela et al., 2011), these results indicate that *B. infantis* ATCC 15697 use primarily free sialic acid from GMP. In contrast *B. bifidum* JCM 1254 showed no reduction of sialic acid in GMP-mZMB (Figure 20), which indicates it does not consume this monosaccharide, while it still can consume GNB. Likewise, sialic acid concentration appeared to increase in *B. thetaiotaomicron* VPI-5482 supernatant over time, but *B. fragilis* J4, *R. gnavus* CC55_001C and *C. perfringens* ATCC 13124 appeared to consume part of sialic acid derived from GMP. In summary, these results indicate which bacteria were able to degrade GMP and which were not degraders, as well as visualizing the consumption of free sugars, the strains able to release into the extracellular medium and metabolize sialic acid and their own.

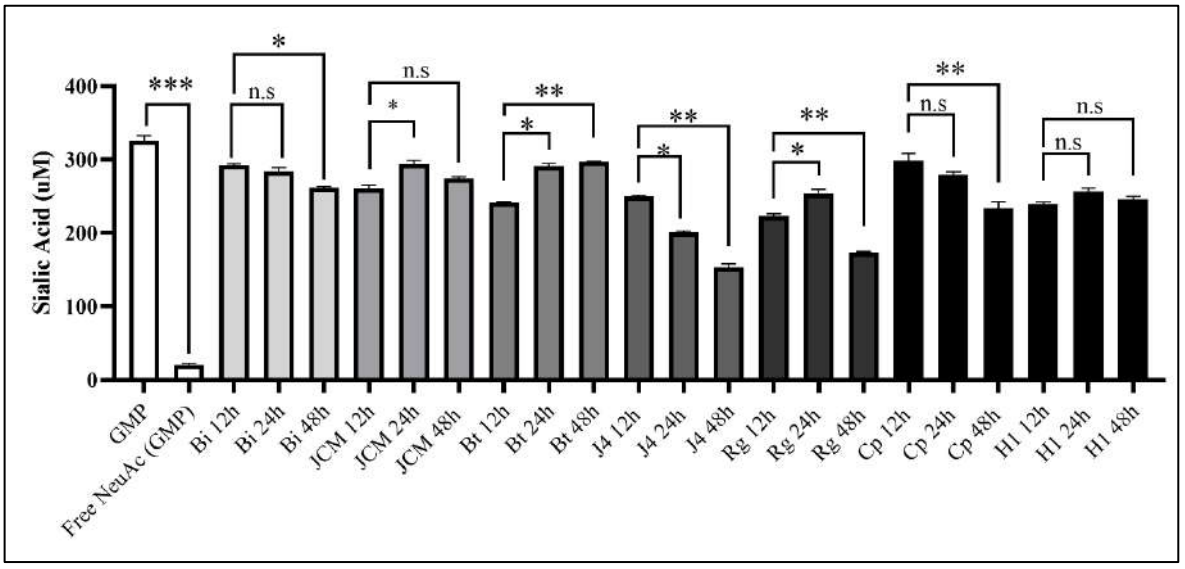

Figure 20- Quantification of sialic acid of GMP from studied bacterial strains.

The sialic acid concentration of samples was quantified using the fluorometric procedure, which is 10-fold more sensitive than the colorimetric assay. Sialic acid standards at 0, 30, 60, and 100 µM were made. Measurements were made at λex = 555/λ_{em} = 585 nm using Synergy H1.

4.3.2 Unidirectional and Bidirectional experiments

Considering the behavior of individual cultures, *B. infantis* ATCC 15697, *B. bifidum* JCM 1254, *B. thetaiotaomicron* VPI-5482, and *C. perfringens* ATCC 13124 were selected as GMP-degraders for unidirectional assays. After re-growth using these GMP-degrader bacteria for 24 h, the remaining supernatant was used to evaluate growth in the remaining bacterial strains.

a) Unidirectional cultures

For this type of bacterial culture, strains with an OD_{620nm} greater than 0.2 were considered stimulated in growth. From spent GMP-mZMB by *B. infantis* ATCC 15697, *B.* breve I1, *B. longum* M12, *S. thermophilus* LM5, *S. thermophilus* LM8, *R. gnavus* CC55_001C, *L. symbiosum* WAL-14673 were stimulated in their growth. *C. perfringens* ATCC 13124 and *E. faecalis* H1 could not growth in this spent media. On the other hand, *B. infantis* ATCC 15697, and *S. thermophilus* LM8 could not grow with *B. bifidum* JCM 1254 supernatant (Figure 21). *B. thetaiotaomicron* VPI-5482 had a maximum near to 0.2 OD_{620nm}. In the case of spent GMP-mZMB by *B. thetaiotaomicron*, it was observed that bifidobacteria strains could not grow.

However, *C. perfringens* ATCC 13124 exhibited a moderate growth rate with an OD_{620nm} maximum value of 0.53 ± 0.03. Likewise, *R. gnavus* CC55_001C, *L. symbiosum* WAL-14673, and *E. faecalis* H1 were stimulated in their growth. With the spent medium GMP-mZMB by *C. perfringens* ATCC 13124, the bifidobacteria could not grow. In contrast, *B. thetaiotaomicron* VPI-5482, *B. fragilis* J4, *S. thermophilus* LM5, S. *thermophilus* LM8, *L. symbiosum* WAL-14673, *R. gnavus* CC55_001C, and *E. faecalis* H1 were stimulated in their growth. Of the four supernatants of GMP degraders, *B. thetaiotaomicron* VPI-5482 was the most beneficial in being able to grow on all spent media. TLC assays from unidirectional experiments showed that galacto-*N*-biose and galactose disappeared over time in the supernatants of non-GMP-degrading bacteria, suggesting they participate in cross-feeding (Supplementary Figure S6).

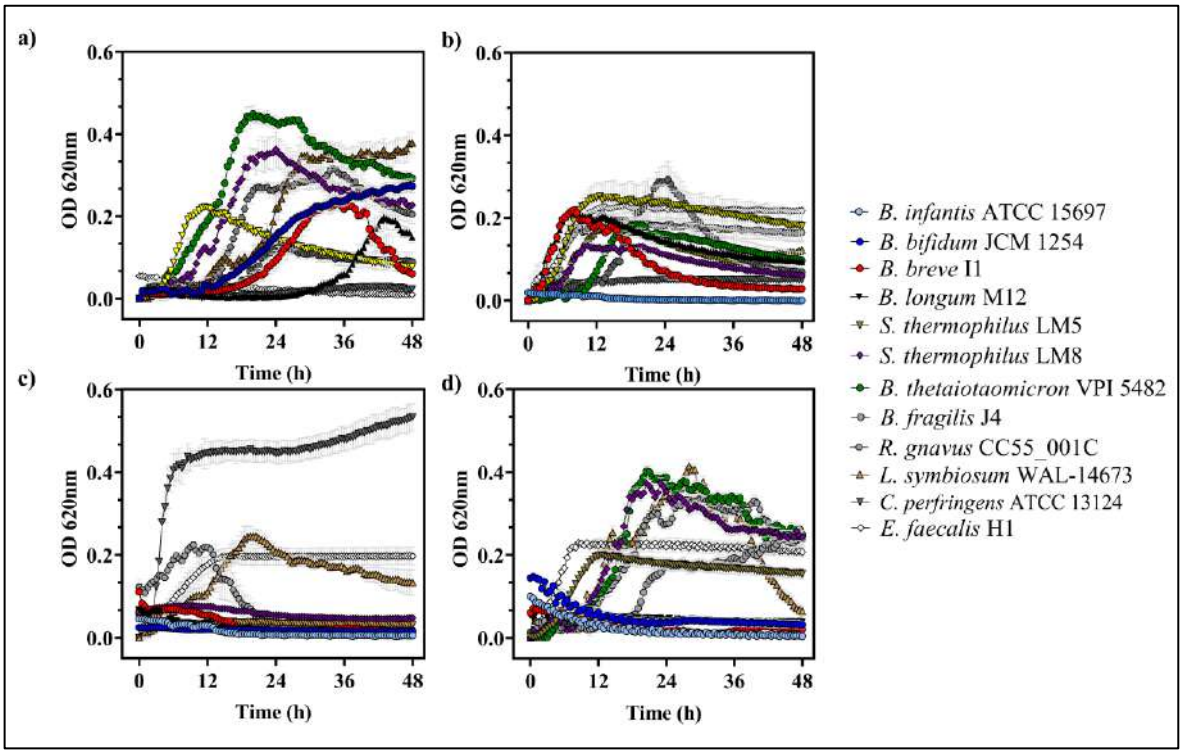

Figure 21- Unidirectional assays using a) *B. infantis* ATCC 15697, b) *B. bifidum* JCM 1254, c) *B. thetaiotaomicron* VPI-5482, and d) *C. perfringens* ATCC 13124 as GMP-degrading bacteria supernatants.

The above GMP-degrading bacteria were incubated for 24 h, then centrifuged, the pellets were discarded, and the supernatants were recovered and filtered for use as medium for the other study strains.

a) Bidirectional assays

For the bidirectional experiments (Figure 22), the pairs were selected considering the most marked cooperative or competitive behavior observed in the unidirectional assays (Supplementary Table S11). Bidirectional experiments were characterized with TLC assays, determination of sialic acid content, and quantification of SCFA such as acetate, lactate, and butyrate (Supplementary Figure S6-Figure S8). Bidirectional assays showed significant growth of non-GMP degrading bacteria when grown simultaneously with GMP-degrading bacteria (Figure 22). Most of the bidirectional assays showed cross-feeding interactions (Figure 22A to Figure 22G), and one case showed substrate competition (Figure 22H). From non-GMP-degrader strains (Figure 22), an increase in biomass was evidenced in *B. breve* I1

(ΔOD = 0.240), for *B. longum* M12 (ΔOD = 0.160), and *L. symbiosum* WAL-14673 (ΔOD = 0.366). *B. bifidum* JCM 1254 was mostly unaffected by the presence of these microorganisms. In the presence of *B. infantis* ATCC 15697, an increase was observed for *S. thermophilus* LM5 (ΔOD = 0.160), *B. longum* M12 (ΔOD = 0.240), and *B. breve* I1 (ΔOD = 0.310). GNB and sialic acid are likely mediators in these cross-feeding interactions (Supplementary Figure S8). In the bidirectional culture of *B. bifidum* JCM 1254 and *R. gnavus* CC55_001C, both bacteria showed a reduced growth in co-culture compared to monocultures, suggesting a negative interaction (Figure 22H). Sialic acid concentrations from bidirectional assays were determined at 24 and 48 h. For example, *B. bifidum* JCM 1254 released sialic acid in monoculture, and *B. breve* did not (Figure 23). These concentrations were reduced in *B. breve* compartment, suggesting sialic acid is used for growth and explains its higher growth in co-culture. Similar observations were made for the biculture of *B. bifidum* JCM 1254 with *L. symbiosum* WAL-14673 and *B. longum* M12, and *B. infantis* ATCC 15697 with *S. thermophilus* LM5, *B. longum* M12 and *B. breve* I1.

Finally, we observed a high production of acetate and lactate in most bidirectional assays, with a maximum in *B. bifidum-B. breve* co-culture. Notably, an increase in butyrate by *L. symbiosum* WAL-14673 during co-culture with *B. bifidum* JCM 1254 was observed, which correlates with the removal of lactate in *L. symbiosum* co-culture (Supplementary Figure S7). In summary, these results indicate the simultaneous growth of non-GMP-degrading strains, the use of GNB and sialic acid as available sugar as a degradation product and the production of butyrate by cross feeding between particular cases of strains that can metabolize these SCFA.

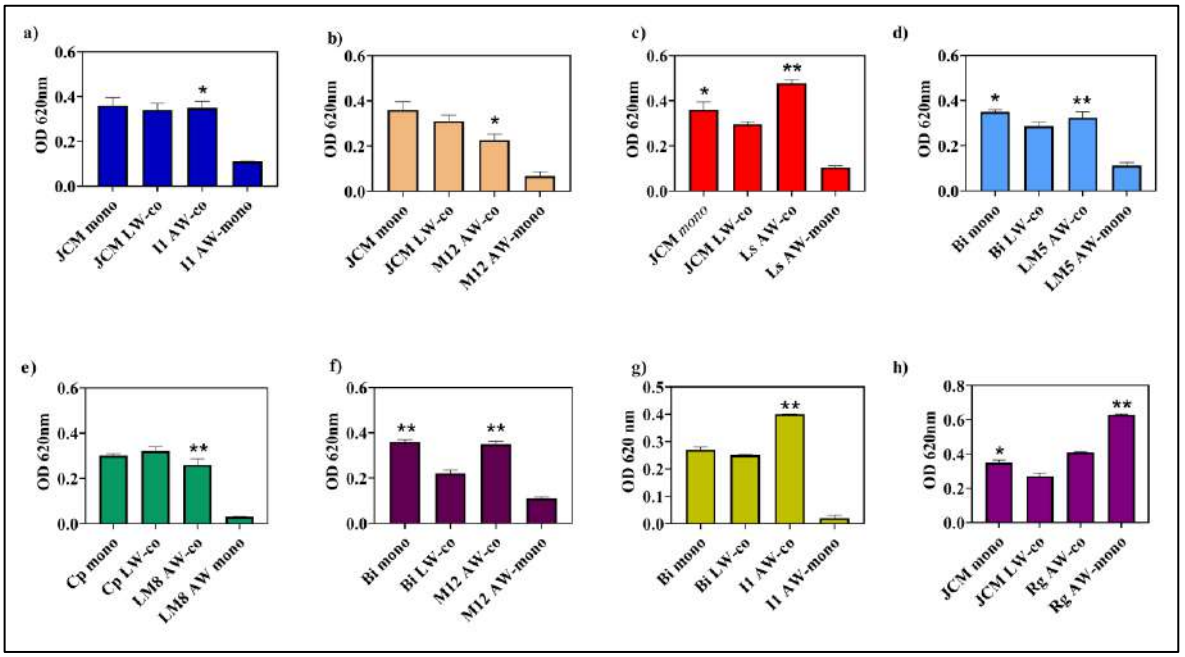

Figure 22- Bidirectional culture using GMP as a carbon source.

Maximum OD values were expressed with the mean and standard deviation. Mono: Monoculture; LW-co: Lower well of the co-culture treatment; AW-co: Apical well of the co-culture treatment. a) *B. bifidum* JCM 1254 vs *B. breve* I1; b) *B. bifidum* JCM 1254 vs *B. longum* M12; c) *B. bifidum* JCM 1254 vs *L. symbiosum* WAL-14673; d) *B. infantis* ATCC 15697 vs *S. thermophilus* LM5; e) *C. perfringens* ATCC 13124 vs *S. thermophilus* LM8; f) *B. infantis* ATCC 15697 vs *B. longum* M12; g) *B. infantis* ATCC 15697 vs *B. breve* I1; h) *B. bifidum* JCM 1254 vs *R. gnavus*. The monoculture was compared with the biculture in each strain. * $p < 0.05$; **: $p < 0.01$.

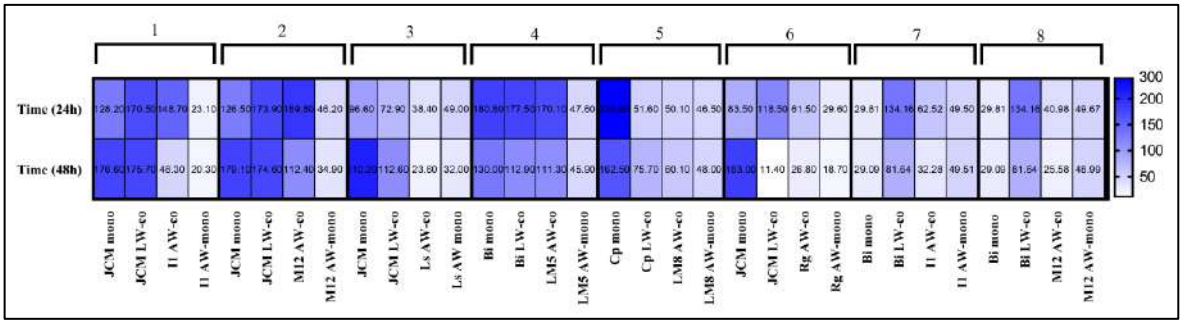

Figure 23- Quantification of sialic acid from individual cultures of studied strains using GMP as a carbon source.

The sialic acid concentration of samples was quantified using the fluorometric procedure, which is 10-fold more sensitive than the colorimetric assay. Sialic acid standards at 0, 30, 60, and 100 µM were made. Measurements were made at $\lambda ex = 555/\lambda_{em} = 585$ nm using Synergy H1.

4.3.3 Relative gene quantification in bidirectional cultures using GMP as a carbon source

We finally studied changes in gene expression of candidate genes participating in GMP consumption in *B. infantis* ATCC 15697 and *B. bifidum* JCM 1254 during the bidirectional assay (Figure 24). These strains were selected considering their background with chemically similar compounds such as HMO to GMP to utilize sugar residues.

Their co-culture revealed small changes in their $OD_{620nm,}$ with an increase in *B. infantis* ATCC 15697 in co-culture with *B. bifidum* JCM 1254. Statistically significant differences in gene expression patterns were found in their monoculture and co-cultures (Figure 24). For *B. infantis* ATCC 15697, the presence of *B. bifidum* JCM 1254 and during the consumption of GMP, we observed the induction associated to GMP of a GH25 hydrolase (Blon_1831), and two sialidases (Blon_0646, Blon_2348). In *B. bifidum* JCM 1254, there was a 6-fold and 45-fold overexpression of endo-*α*-*N*-GalNAcase and exo-*α*-sialidase in GMP relative to lactose and compared to co-culture with *B. infantis* ATCC 15697. These results suggest that co-culture conditions increase the expression of important genes related to GMP utilization in *B. infantis* ATCC 15697 and *B. bifidum* JCM 1254, compared to single culture.

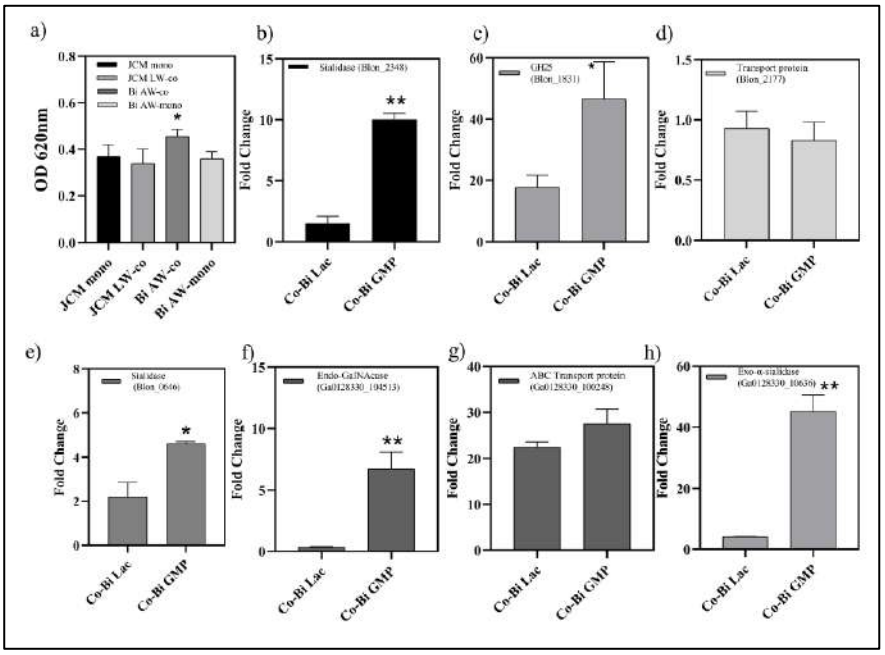

Figure 24- Relative gene expression in the co-culture assay between *B. infantis* and *B. bifidum* JCM 1254.

Quantitative PCRs (qPCRs) were prepared using the SensiFAST SYBR No-ROX kit (Bioline, UK) under the manufacturer's instructions and 2,2 µL of each cDNA. Amplification was performed in an AriaMx real-time PCR system (Agilent Technologies), using 96-well optical plates MicroAmp Fast Optical (ThermoFisher, USA). *: $p < 0.05$; **: $p < 0.01$.

4.4 DISCUSSIONS

Casein glycomacropeptide (GMP) is a simple glycopeptide with neutral and sialylated *O*-glycans corresponding to monosaccharides (0.8%), disaccharides (6.3%), trisaccharides (36.5%), and tetrasaccharides (56%) (Saito & Itoh, 1992). Considering their similarity to the *O*-glycans of intestinal mucin and other structures such as HMO (Pruss et al., 2021), the impact of glycomacropeptide as a carbon source in representative bacteria of the HGM, their metabolic interactions and relative gene expression was studied.

In first instance, individual cultures have allowed to identify which strains were able to grow with GMP as the only carbon source. Single cultures were carried out to evaluate 12 strains belonging to a Chilean isolated and collection group (Table 3). *B. infantis* ATCC 15697 and *B. bifidum* JCM 1254 could grow in GMP-mZMB, similarly reported in other studies (Morozumi et al., 2023; O'Riordan et al., 2018). In the present study, *B. infantis* ATCC 15697 could grow with GMP without incorporating glucose or other sugars, confirming its use. This could be explained by the fact that the modified ZMB medium has in its preparation the inclusion of amino acids and vitamins that could be used for stability and growth within the metabolization of GMP (Medina et al., 2017).

The GMP consumption by *B. bifidum* JCM 1254 is explained by the fact that its genome contains a complete set of genes encoding to extracellular glycosyl hydrolases, including exo-*α*-sialidases, galactosidases, endo-*α*-*N*-GalNAcase, essential for the breakdown of GMP *O*-glycans (Katoh et al., 2017; Shimada et al., 2015). These by-products are released into the medium and for the convenience of other strains, demonstrated by TLC assays and similar to those reported in other studies (Katoh et al., 2020). Isolates strains such as *B. breve* I1 and *B. longum* M12, *S. thermophilus* LM5/LM8, and *L. symbiosum* WAL-14673 could not grow with GMP, which is dependent on the enzymatic machinery, and oligosaccharide transporters available in the strains to breakdown GMP *O*-glycans (Takihara & Okuda, 2023).

For *Bacteroidota* strains, *B. thetaiotaomicron* VPI-5482 and *B. fragilis* J4 could grow with GMP. The *B. thetaiotaomicron* VPI-5482 enzymatic activity has been reported, including sialidase cluster, galactosidase, and constituent enzymes of the SUS system in mucin *O*-glycans (Cheng et al., 2022). For *B. fragilis*, this PUL involved in *O*-glycans consumption is homologous with *B. thetaiotaomicron* VPI-5482 (Fletcher et al., 2009; Luis & Hansson, 2023).

For *Bacillota* strains, in some cases their activity has been observed as pathological behavior for the benefit of growth and survival during interactions with human gut mucin (Y. Wang, 2020). *R. gnavus* is an early colonizer of the infant gut but persists in adults (Crost et al., 2023). Likewise, *R. gnavus* is part of the healthy gut microbiota of humans. However, when it exhibits disproportionate representation, it may play a pathogenic role in inflammatory gut diseases (Zhai et al., 2023). In GMP, *R. gnavus* CC55_001C demonstrated growth and the ability to use GMP *O*-glycans, similar to intestinal mucin. This highlights its enzymatic arsenal for mucin deglycosylation reported in other studies, including sialidase, sialic acid metabolization enzymes, and endo-β-galactosidases (Berkhout et al., 2022; Crost et al., 2016). The ability of *C. perfringens* ATCC 13124 to grow with this glycopeptide lies primarily in the release of sialic acid as a terminal residue that decorates *O*-glycans. In the case of GMP, *C. perfringens* ATCC 13124 can release sialic acid through the action of its three GH33 sialidases (nanH: CPF_0985, nanJ: CPF_0532, nanI: CPF_0721) (MacMillan et al., 2019) and partially metabolize. Subsequently, sialidases expose the GMP *O*-glycans, permitting the GH101 EngCP activity by cleaving the *O*-glycosidic α-linkage for their use. In contrast, little ability to degrade sialylated *O*-glycans in mucin is reported for *E. faecalis* (Raev et al., 2023), the absence of sialidase activity in this strain suggests its preference for neutral *O*-glycans with the action of GH101 EngEF and Nag31 for GMP (Cheung et al., 2020; Koutsioulis et al., 2008b; Miyazaki et al., 2022).

The Unidirectional assays may represent the first insights into microbial interactions between the studied strains considering GMP-degraders. Surprisingly, strains such as *B. breve* I1, *B. longum* M12, and *L. symbiosum* WAL-14673 were stimulated in their growth in *B. bifidum* JCM 1254 and *B. infantis* ATCC 15697 supernatants. Growth stimulation for *B. breve* I1 and *L. symbiosum* WAL-14673 could be explained mainly by using sialic acid present extracellularly in the medium (Yokoi et al., 2022). Into its genome, *B. breve* I1 and

L. symbiosum have the complete metabolic pathway to metabolize sialic acid intracellularly, which is conserved among other strains (Egan et al., 2015). In contrast, *B. longum* spp. Such as M12, does not have this advantage and only utilizes simple sugars as by-products of the GMP breakdown. This behavior can be explained because the *B. longum* strains have a high preference for plant-derived oligosaccharides compared to *O*-glycans and for HMO such as Lacto-*N*-tetraose and Lacto-*N*-neotetraose (González-Morelo et al., 2023; Thompson et al., 2018). In the same way, it has been reported that *B. longum* M12 prefers fucosylated oligosaccharides similar to *B. longum* SC596, thanks to it contains a cluster dedicated to the degradation of this type of substrate (Díaz et al., 2021). On the other hand, *C. perfringens* ATCC 13124 could not grow in supernatants of the two GMP-consuming bifid strains (Figure 21A and Figure 21B). Reported *in vitro* studies on the interaction between bifidobacteria and *C. perfringens* are often confounded by various internal and external factors. For example, bifidobacteria can be affected by the toxin production of *C. perfringens*, while mild oxygenation of the medium can affect *C. perfringens* to a greater extent than bifidobacteria (Shaw et al., 2020). Nevertheless, antagonistic activity against enteropathogens has been reported in both *Lactobacillus* and *Bifidobacterium* strains. This could be due to the release of bacteriocins present in the medium that inhibits its growth (Martinez et al., 2013).

Bidirectional assays allow a wider visualization of microbial interactions simultaneously between degraders and non-degraders of carbon sources (Dahal et al., 2023). In the case of GMP, the findings obtained from the unidirectional assays were considered to establish several co-cultures for analysis (Figure 21). Cross-feeding interactions were found in most of the studied pairs of bidirectional assays (Figure 22A - Figure 22G), except for the last co-culture between *B. bifidum* JCM 1254 and *R. gnavus* CC55_001C, which was considered a competitive behavior. *B. bifidum* JCM 1254 was considered one of the critical strains among GMP-degraders, possessing the necessary enzymatic machinery (galactosidase, endo-GalAcase, and sialidase) to deglycosylate GMP and release sugar by-products such as GNB, Gal, GalNAc, and sialic acid into the extracellular medium. Strains such as *B. breve* I1, *B. longum* M12, and *S. thermophilus* LM5/LM8 were significantly growth stimulated and confirmed by the difference between monocultures and co-cultures samples. This class of bacterial interactions with sialic acid is similar to those observed with

sugars such as fucose, which is also a terminal sugar in the fucosylated oligosaccharide structures of HMO, involving enzymatic requirements such as fucosidases and metabolic pathways for its utilization (Thompson et al., 2018). Similarly, GNB is related to interactions with Lacto-*N*-biose (LNB), which is a comparable disaccharide (Asakuma et al., 2011). *B. infantis* ATCC 15697 was the second most important in stimulating the growth of non-GMP degraders observed in the OD difference between monocultures and co-cultures. In the case of competitive behavior (Figure 22H), although *B. bifidum* JCM 1254 does not utilize sialic acid from complex glycans, *R. gnavus* chemically modifies the free sialic acid to form 2,7-anhydro-Neu5Ac, making the sugars with high enzyme specificity and less accessible to other intestinal microbes (Bell et al., 2019).

TLC assays of the bidirectional assays ratify the consumption of oligosaccharides released into the medium concerning time, among those GNB, Gal, and GalNAc (Supplementary Figure S8). Due to the difficulty of visualizing sialic acid through TLC assays, sialic acid was quantified in the eight subsequent bidirectional assays. Statistically significant differences were found in comparing the lower co-cultures (LW-co) and the apical co-cultures (AW-co) wells. Cross-feeding behavior conferring free sialic acid is observed in co-cultures such as *B. bifidum* JCM 1254 with *B. breve* I1, and *L. symbiosum* WAL-14673.

While bacteria can use their enzymatic machinery to cleave and release simpler substrates for utilization (Wardman et al., 2022), SCFAs were among the other degradation by-products generated and analyzed between the two-way culture assays. Bifidobacterial species such as *B. infantis* ATCC 15697, *B. bifidum* JCM 1254, and *B. longum* M12 can produce acetate and lactate. The results of this study show butyrate production by *L. symbiosum* WAL-14673 which correlated with lactate depletion in co-culture , similarly occurred on other substrates such as inulin and xylan (Hirmas et al., 2022; A. J. M. Martin et al., 2023).

An additional bidirectional assay between *B. infantis* ATCC 15697 and *B. bifidum* JCM 1254 was carried out to analyze the expression of genes related to *O*-glycan consumption. Cross-feeding behavior was evidenced in the OD_{620nm} values between *B. infantis* ATCC 15697 and *B. bifidum* JCM 1254, comparing monoculture and co-culture treatments. Compare with the lactose treatment in co-culture, relative gene expression assays

confirm the utilization of *O*-glycans by *B. bifidum* JCM 1254 with the use of exo-alpha sialidases, endo-*α*-*N*-GalNAcases, and the utilization of an ABC transporter protein for sugar import (Morozumi et al., 2023). The results demonstrate cross-feeding activity due to the presence of sialylated O-glycans induced sialidase encoding of *B. bifidum* JCM 1254, furthermore this suggests for further utilization with the expression of its Endo-GalAcase to release it from the amino acid structure, similar to other studies (Morozumi et al., 2023). In comparison, *B. infantis* ATCC 15697 has the expression of Blon_1831, a GH25 glycosyl hydrolase that could facilitate cell division (O'Riordan et al., 2018). Also, Blon_2348 and Blon_0636 are overexpressed intracellularly, overexpressing the sialic acid present in the *O*-glycans of GMP, which could suggest an import of sialylated disaccharides into the cell.

Despite the extensive analysis of this study, the limitations of this study lie in incorporating transcriptomics, proteomics, and comparisons with synthetic *O*-glycans due to their low yield and excessive cost in all analyses. However, individual, unidirectional, and bidirectional assays and complementary analyses of the by-products of GMP consumption such as liquid and thin layer chromatography and gene expression, allowed us to express a broader perspective of the interactions occurring in GMP consumption.

4.5 CONCLUSIONS

GMP can be considered a simple model of *O*-glycosylated structure for studying metabolic interactions between bacteria of the human gut microbiota. Within representative bacteria of the gut microbiota, non-GMP-degrading strains could be identified. Unidirectional and bidirectional cultures confirmed the altruistic activity of some microbiota strains in benefiting non-GMP-degrading strains by consuming simpler by-products. Galacto-*N*-biose, galactose, *N*-acetylgalactosamine, and sialic acid were among the substrates resulting from GMP consumption offered to non-GMP degrading bacteria, being dependent on their metabolic capacity of utilization. The fermentation of GMP promotes the release of its *O*-glycans, and also allows the production of fatty acid derivatives essential for the production of butyrate as an anti-inflammatory molecule. These results are expected to serve as a foundation for developing clinical and in vivo assays for next-generation prebiotic use.

5. CONCLUSIONS AND FUTURE PERSPECTIVES

This doctoral dissertation elucidates the intricate molecular mechanisms underlying the utilization of GMP *O*-glycans by representative gut microbes. The initial segment of this research involves a comprehensive literature review, in which the most recent advances in understanding the various molecular mechanisms employed by host organisms for *O*-glycan consumption are discussed, with particular attention to the alternative source of glycoprotein-derived *O*-glycans. Prominent bacteria of the human gut microbiota demonstrate prospects for adapting to changes in more complex substrates such as *O*-glycans, which will depend on the production of specific enzymes and metabolic pathways that facilitate their utilization.

Likewise, differences in gene expression patterns in chemically similar structures to *O*-glycans, such as HMO, mainly highlight the creation of networks for metabolizing specific substrates as a carbon source. In particular, *B. infantis*, a prototypical bifidobacterial species of the infant gut microbiota, organizes its genes into two modules according to their involvement in utilizing galactose or *N*-Acetylglucosamine. In contrast, *B. bifidum* shows a less structured co-expression of genes involved in HMO utilization. *B. longum* shows a distinct clustering of these genes within a small module, suggesting a lack of co-expression with essential cellular processes and implying a relatively recent acquisition of these genetic elements. Despite variations in experimental data among these organisms, the fundamental mechanisms compared to *O*-glycan utilization appear to be conserved, providing a basis for future exploration, including integrating gene expression data into metagenomic studies.

As for the work developed in the final section of the thesis, GMP can be considered a simple model *O*-glycosylated peptide in addition to the possibility of being used as a prebiotic. Due to the complexity of GMP *O*-glycans, they induce the formation of cross-feeding or competitive interactions with GMP-degrading microbes and those not adapted to the consumption of this substrate. Studied strains such as *B. infantis*, *B. bifidum*, *B. thetaiotaomicron*, *B. fragilis*, *R. gnavus*, *C. perfringens*, and *E. faecalis* are capable of utilizing GMP *O*-glycans. This observation has repercussions on the production of metabolites resulting from *O*-glycan degradation during GMP fermentation, such as sugar release and short-chain fatty acids useful for butyrate production, as a molecule of interest.

B. bifidum, a GMP degrading strain, can behave as an altruistic strain as substrate donors to non-consuming strains of complex compounds. Despite the extensive analysis of the present doctoral work, the limitations of the present work lie in the acquisition of more complex substrates to be used as standards, the incorporation of transcriptomics, proteomics, and comparisons with synthetic *O*-glycans due to their low throughput and high cost in all analyses. However, individual, unidirectional, and bidirectional assays and complementary analyses of the by-products of GMP consumption, such as liquid and thin-layer chromatography and gene expression, allowed us to express a broader perspective of the interactions that occur in GMP consumption.

The results of this doctoral thesis are expected to provide an opening for future research, including *in vivo* studies, dynamic and static *in vitro* digestion experiments, and clinical investigations, thus establishing GMP as an emerging prebiotic, especially in light of its potential to counteract opportunistic colonization by strains detrimental to host health.

REFERENCES

Al Bander, Z., Nitert, M. D., Mousa, A., & Naderpoor, N. (2020). The gut microbiota and inflammation: an overview. *International Journal of Environmental Research and Public Health*, *17*(20), 7618. https://doi.org/10.3390/ijerph17207618

Alessandri, G., van Sinderen, D., & Ventura, M. (2021). The genus *Bifidobacterium*: from genomics to functionality of an important component of the mammalian gut microbiota. *Computational and Structural Biotechnology Journal*, *19*(1), 1472–1487. https://doi.org/10.1016/j.csbj.2021.03.006

Amore, A., Knott, B. C., Supekar, N. T., Shajahan, A., Azadi, P., Zhao, P., Wells, L., Linger, J. G., Hobdey, S. E., Vander Wall, T. A., Shollenberger, T., Yarbrough, J. M., Tan, Z., Crowley, M. F., Himmel, M. E., Decker, S. R., Beckham, G. T., & Taylor, L. E. (2017). Distinct roles of *N*- and *O*-glycans in cellulase activity and stability. *Proceedings of the National Academy of Sciences*, *114*(52), 13667–13672. https://doi.org/10.1073/pnas.1714249114

Anderson, K. L., & Salyers, A. A. (1989). Biochemical evidence that starch breakdown by *Bacteroides thetaiotaomicron* involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes. *Journal of Bacteriology*, *171*(2), 3192–3198.

Arike, L., & Hansson, G. C. (2016). The densely *O*-glycosylated MUC2 mucin protects the intestine and provides food for the commensal bacteria. *Journal of Molecular Biology*, *428*(16), 3221–3229. https://doi.org/10.1016/j.jmb.2016.02.010

Arike, L., Holmén-Larsson, J., & Hansson, G. C. (2017). Intestinal Muc2 mucin *O*-glycosylation is affected by microbiota and regulated by differential expression of glycosyltranferases. *Glycobiology*, *27*(4), 318–328. https://doi.org/10.1093/glycob/cww134

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., Fernandes, G. R., Tap, J., Bruls, T., Batto, J.-M., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L., Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., … Bork, P. (2011). Enterotypes of the human gut microbiome. *Nature*, *473*(7346), 174–180. https://doi.org/10.1038/nature09944

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. *Journal of Biological Chemistry*, *286*(40), 34583–34592. https://doi.org/10.1074/jbc.M111.248138

Atuma, C., Strugala, V., Allen, A., & Holm, L. (2001). The adherent gastrointestinal mucus gel layer: thickness and physical state in vivo. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *280*(5), G922–G929. https://doi.org/10.1152/ajpgi.2001.280.5.G922

Auer, F., Jarvas, G., & Guttman, A. (2021). Recent advances in the analysis of human milk oligosaccharides by liquid phase separation methods. *Journal of Chromatography B*, *1162*, 122497. https://doi.org/10.1016/j.jchromb.2020.122497

Bansil, R., & Turner, B. S. (2018). The biology of mucus: composition, synthesis and organization. *Advanced Drug Delivery Reviews*, *124*(1), 3–15. https://doi.org/10.1016/j.addr.2017.09.023

Baum, L. G., & Cobb, B. A. (2016). The direct and indirect effects of glycans on immune function. *Glycobiology*, *27*(310), 619–624.

Bell, A., Brunt, J., Crost, E., Vaux, L., Nepravishta, R., Owen, C. D., Latousakis, D., Xiao, A., Li, W., Chen, X., Walsh, M. A., Claesen, J., Angulo, J., Thomas, G. H., & Juge, N. (2019). Elucidation of a sialic acid metabolism pathway in mucus-foraging *Ruminococcus gnavus* unravels mechanisms of bacterial adaptation to the gut. *Nature Microbiology*, *4*(12), 2393–2404. https://doi.org/10.1038/s41564-019-0590-7

Belzer, C., Chia, L. W., Aalvink, S., Chamlagain, B., Piironen, V., Knol, J., & de Vos, W. M. (2017). Microbial metabolic networks at the mucus layer lead to diet-independent butyrate and vitamin B12 production by intestinal symbionts. *MBio*, *8*(5), e00770-17. https://doi.org/10.1128/mBio.00770-17

BeMiller, J. N. (2019). Oligosaccharides. In J. N. BeMiller (Ed.), *Carbohydrate Chemistry for Food Scientists* (3th ed., pp. 49–74). Woodhead Publishing and AACC International Press (Elsevier). https://doi.org/10.1016/B978-0-12-812069-9.00003-0

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C. C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G. H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J. A., Maguin, E., Mauchline, T., McClure, R., … Schloter, M. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome*, *8*(1), 103. https://doi.org/10.1186/s40168-020-00875-0

Bergstrom, K., Fu, J., Johansson, M. E. V, Liu, X., Gao, N., Wu, Q., Song, J., McDaniel, J. M., McGee, S., Chen, W., Braun, J., Hansson, G. C., & Xia, L. (2017). Core 1– and 3–derived *O*-glycans collectively maintain the colonic mucus barrier and protect against spontaneous colitis in mice. *Mucosal Immunology*, *10*(1), 91–103. https://doi.org/10.1038/mi.2016.45

Bergstrom, K., & Xia, L. (2013). Mucin-type *O*-glycans and their roles in intestinal homeostasis. *Glycobiology*, *23*(9), 1026–1037. https://doi.org/10.1093/glycob/cwt045

Berkhout, M. D., Plugge, C. M., & Belzer, C. (2022). How microbial glycosyl hydrolase activity in the gut mucosa initiates microbial cross-feeding. *Glycobiology*, *32*(3), 182–200. https://doi.org/10.1093/glycob/cwab105

Bernard, L., Bonnet, M., Delavaud, C., Delosière, M., Ferlay, A., & Fougère, H. (2018). Milk fat globule in ruminant: major and minor compounds, nutritional regulation and differences among species. *European Journal of Lipid Science and Technology*, *120*(5), 1700039.

Bernstein, C. N., & Forbes, J. D. (2017). Gut microbiome in inflammatory bowel disease and other chronic immune-mediated inflammatory diseases. *Inflammatory Intestinal Diseases*, *2*(2), 116–123. https://doi.org/10.1159/000481401

Bode, L. (2012). Human milk oligosaccharides: Every baby needs a sugar mama. *Glycobiology*, *22*(9), 1147–1162. https://doi.org/10.1093/glycob/cws074

Borewicz, K., Suarez-Diez, M., Hechler, C., Beijers, R., de Weerth, C., Arts, I., Penders, J., Thijs, C., Nauta, A., Lindner, C., Van Leusen, E., Vaughan, E. E., & Smidt, H. (2019). The effect of prebiotic fortified infant formulas on microbiota composition and dynamics in early life. *Scientific Reports*, *9*(1), 2434. https://doi.org/10.1038/s41598-018-38268-x

Boutrou, R., Jardin, J., Blais, A., Tomé, D., & Léonil, J. (2008). Glycosylations of κ-casein-derived caseinomacropeptide reduce its accessibility to endo- but not exointestinal brush border membrane peptidases. *Journal of Agricultural and Food Chemistry*, *56*(17), 8166–8173. https://doi.org/10.1021/jf801140d

Bresser, L. R. F., de Goffau, M. C., Levin, E., & Nieuwdorp, M. (2022). Gut Microbiota in Nutrition and Health with a Special Focus on Specific Bacterial Clusters. *Cells*, *11*(19), 3091. https://doi.org/10.3390/cells11193091

Breugelmans, T., Oosterlinck, B., Arras, W., Ceuleers, H., De Man, J., Hold, G. L., De Winter, B. Y., & Smet, A. (2022). The role of mucins in gastrointestinal barrier function during health and disease. *The Lancet Gastroenterology & Hepatology*, *7*(5), 455–471. https://doi.org/10.1016/S2468-1253(21)00431-3

Brockhausen, I., & Stanley, P. (2017). O-GalNAc glycans. In A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, & P. H. Seeberger (Eds.), *Essentials of Glycobiology* (3rd ed.). Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/glycobiology.3e.010

Brosseau, C., Selle, A., Palmer, D. J., Prescott, S. L., Barbarot, S., & Bodinier, M. (2019). Prebiotics: mechanisms and preventive effects in allergy. *Nutrients*, *11*(8), 1841. https://doi.org/10.3390/nu11081841

Bunesova, V., Lacroix, C., & Schwab, C. (2018). Mucin cross-feeding of infant Bifidobacteria and *Eubacterium hallii*. *Microbial Ecology*, *75*(1), 228–238. https://doi.org/10.1007/s00248-017-1037-4

Cani, P. D., & de Vos, W. M. (2017). Next-generation beneficial microbes: the case of *Akkermansia muciniphila*. *Frontiers in Microbiology*, *8*, 1765. https://doi.org/10.3389/fmicb.2017.01765

Castro, I., García-Carral, C., Furst, A., Khwajazada, S., García, J., Arroyo, R., Ruiz, L., Rodríguez, J. M., Bode, L., & Fernández, L. (2022). Interactions between human milk oligosaccharides, microbiota and immune factors in milk of women with and without mastitis. *Scientific Reports*, *12*(1), 1367. https://doi.org/10.1038/s41598-022-05250-7

Chen, G., Ran, X., Li, B., Li, Y., He, D., Huang, B., Fu, S., Liu, J., & Wang, W. (2018). Sodium butyrate inhibits inflammation and maintains epithelium barrier integrity in a TNBS-induced inflammatory bowel disease mice model. *EBioMedicine*, *30*, 317–325. https://doi.org/10.1016/j.ebiom.2018.03.030

Chen, X., & Ma, J. (2021). Weighted gene co-expression network analysis (WGCNA) to explore genes responsive to Streptococcus oralis biofilm and immune infiltration analysis in human gingival fibroblasts cells. *Bioengineered*, *12*(1), 1054–1065.

https://doi.org/10.1080/21655979.2021.1902697

Cheng, J., Hu, J., Geng, F., & Nie, S. (2022). Bacteroides utilization for dietary polysaccharides and their beneficial effects on gut health. *Food Science and Human Wellness*, *11*(5), 1101–1110. https://doi.org/10.1016/j.fshw.2022.04.002

Cheung, J. K., Adams, V., D'Souza, D., James, M., Day, C. J., Jennings, M. P., Lyras, D., & Rood, J. I. (2020). The EngCP endo-*α*-*N*-acetylgalactosaminidase is a virulence factor involved in *Clostridium perfringens* gas gangrene infections. *International Journal of Medical Microbiology*, *310*(2), 151398. https://doi.org/10.1016/j.ijmm.2020.151398

Chung, L., Thiele Orberg, E., Geis, A. L., Chan, J. L., Fu, K., DeStefano Shields, C. E., Dejea, C. M., Fathi, P., Chen, J., Finard, B. B., Tam, A. J., McAllister, F., Fan, H., Wu, X., Ganguly, S., Lebid, A., Metz, P., Van Meerbeke, S. W., Huso, D. L., … Housseau, F. (2018). *Bacteroides fragilis* toxin coordinates a pro-carcinogenic inflammatory cascade via targeting of colonic epithelial cells. *Cell Host and Microbe*, *23*(2), 203–214. https://doi.org/10.1016/j.chom.2018.01.007

Chung, M., Bruno, V. M., Rasko, D. A., Cuomo, C. A., Muñoz, J. F., Livny, J., Shetty, A. C., Mahurkar, A., & Dunning Hotopp, J. C. (2021). Best practices on the differential expression analysis of multi-species RNA-seq. *Genome Biology*, *22*(1), 121. https://doi.org/10.1186/s13059-021-02337-8

Cockburn, D. W., & Koropatkin, N. M. (2016). Polysaccharide degradation by the intestinal microbiota and its influence on human health and disease. *Journal of Molecular Biology*, *428*(16), 3230–3252. https://doi.org/10.1016/j.jmb.2016.06.021

Collins, J. W., Keeney, K. M., Crepin, V. F., Rathinam, V. A. K., Fitzgerald, K. A., Finlay, B. B., & Frankel, G. (2014). *Citrobacter rodentium*: infection, inflammation and the microbiota. *Nature Reviews Microbiology*, *12*(9), 612–623. https://doi.org/10.1038/nrmicro3315

Conz, A., Salmona, M., & Diomede, L. (2023). Effect of non-nutritive sweeteners on the gut microbiota. *Nutrients*, *15*(8), 1869. https://doi.org/10.3390/nu15081869

Córdova-Dávalos, L., Jiménez, M., & Salinas, E. (2019a). Glycomacropeptide bioactivity and health: a review highlighting action mechanisms and signaling pathways. *Nutrients*, *11*(3), 598. https://doi.org/10.3390/nu11030598

Córdova-Dávalos, L., Jiménez, M., & Salinas, E. (2019b). Glycomacropeptide bioactivity and health: A review highlighting action mechanisms and signaling pathways. *Nutrients*, *11*(3), 598. https://doi.org/10.3390/nu11030598

Corfield, A. P. (2000). Mucins and mucosal protection in the gastrointestinal tract: new prospects for mucins in the pathology of gastrointestinal disease. *Gut*, *47*(4), 589–594. https://doi.org/10.1136/gut.47.4.589

Corfield, A. P. (2015). Mucins: A biologically relevant glycan barrier in mucosal protection. In *Biochimica et Biophysica Acta - General Subjects* (Vol. 1850, Issue 1, pp. 236–252). Elsevier B.V. https://doi.org/10.1016/j.bbagen.2014.05.003

Couto, M. R., Gonçalves, P., Magro, F., & Martel, F. (2020). Microbiota-derived butyrate

regulates intestinal inflammation: Focus on inflammatory bowel disease. *Pharmacological Research*, *159*, 104947. https://doi.org/10.1016/j.phrs.2020.104947

Crost, E. H., Coletto, E., Bell, A., & Juge, N. (2023). *Ruminococcus gnavus*: friend or foe for human health. *FEMS Microbiology Reviews*, *47*(2). https://doi.org/10.1093/femsre/fuad014

Crost, E. H., Tailford, L. E., Monestier, M., Swarbreck, D., Henrissat, B., Crossman, L. C., & Juge, N. (2016). The mucin-degradation strategy of *Ruminococcus gnavus* : The importance of intramolecular trans-sialidases. *Gut Microbes*, *7*(4), 302–312. https://doi.org/10.1080/19490976.2016.1186334

Crouch, L. I., Liberato, M. V, Urbanowicz, P. A., Baslé, A., Lamb, C. A., Stewart, C. J., Cooke, K., Doona, M., Needham, S., Brady, R. R., Berrington, J. E., Madunic, K., Wuhrer, M., Chater, P., Pearson, J. P., Glowacki, R., Martens, E. C., Zhang, F., Linhardt, R. J., … Bolam, D. N. (2020). Prominent members of the human gut microbiota express endo-acting *O*-glycanases to initiate mucin breakdown. *Nature Communications*, *11*(1), 1–13. https://doi.org/10.1038/s41467-020-17847-5

Cunningham, M., Azcarate-Peril, M. A., Barnard, A., Benoit, V., Grimaldi, R., Guyonnet, D., Holscher, H. D., Hunter, K., Manurung, S., Obis, D., Petrova, M. I., Steinert, R. E., Swanson, K. S., van Sinderen, D., Vulevic, J., & Gibson, G. R. (2021). Shaping the future of probiotics and prebiotics. *Trends in Microbiology*, *29*(8), 667–685. https://doi.org/10.1016/j.tim.2021.01.003

D'Souza, G., Shitut, S., Preussger, D., Yousif, G., Waschina, S., & Kost, C. (2018). Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Natural Product Reports*, *35*(5), 455–488. https://doi.org/10.1039/C8NP00009C

Dahal, S., Hurst, G. B., Chourey, K., Engle, N. L., Burdick, L. H., Morrell-Falvey, J. L., Tschaplinski, T. J., Doktycz, M. J., & Pelletier, D. A. (2023). Mechanism for utilization of the populus-derived metabolite salicin by a *Pseudomonas—Rahnella* co-Culture. *Metabolites*, *13*(2), 140. https://doi.org/10.3390/metabo13020140

Dao, M. C., Everard, A., Aron-Wisnewsky, J., Sokolovska, N., Prifti, E., Verger, E. O., Kayser, B. D., Levenez, F., Chilloux, J., Hoyles, L., Dumas, M. E., Rizkalla, S. W., Doré, J., Cani, P. D., Clément, K., Le Mouhaër, S., Cotillard, A., Kennedy, S. P., Pons, N., … Lepage, P. (2016). *Akkermansia muciniphila* and improved metabolic health during a dietary intervention in obesity: relationship with gut microbiome richness and ecology. *Gut*, *65*(3), 426–436. https://doi.org/10.1136/gutjnl-2014-308778

Davani-Davari, D., Negahdaripour, M., Karimzadeh, I., Seifan, M., Mohkam, M., Masoumi, S., Berenjian, A., & Ghasemi, Y. (2019). Prebiotics: definition, types, sources, mechanisms, and clinical applications. *Foods*, *8*(3), 92. https://doi.org/10.3390/foods8030092

Davis, J. C. C., Totten, S. M., Huang, J. O., Nagshbandi, S., Kirmiz, N., Garrido, D. A., Lewis, Z. T., Wu, L. D., Smilowitz, J. T., German, J. B., Mills, D. A., & Lebrilla, C. B. (2016). Identification of oligosaccharides in feces of breast-fed infants and their correlation with the gut microbial community. *Molecular & Cellular Proteomics*,

15(9), 2987–3002. https://doi.org/10.1074/mcp.M116.060665

De Bruyn, F., Beauprez, J., Maertens, J., Soetaert, W., & De Mey, M. (2013). Unraveling the Leloir pathway of Bifidobacterium bifidum: significance of the rridylyltransferases. *Applied and Environmental Microbiology*, *79*(22), 7028–7035. https://doi.org/10.1128/AEM.02460-13

De Weirdt, R., & Van de Wiele, T. (2015). Micromanagement in the gut: microenvironmental factors govern colon mucosal biofilm structure and functionality. *NPJ Biofilms and Microbiomes*, *1*(1), 15026. https://doi.org/10.1038/npjbiofilms.2015.26

Dedon, L. R., Özcan, E., Rani, A., & Sela, D. A. (2020). *Bifidobacterium infantis* metabolizes 2′Fucosyllactose-derived and free fucose through a common catabolic pathway resulting in 1,2-Propanediol secretion. *Frontiers in Nutrition*, *7*. https://doi.org/10.3389/fnut.2020.583397

Dekaboruah, E., Suryavanshi, M. V., Chettri, D., & Verma, A. K. (2020). Human microbiome: an academic update on human body site specific surveillance and its possible role. *Archives of Microbiology*, *202*(8), 2147–2167. https://doi.org/10.1007/s00203-020-01931-x

Demouveaux, B., Gouyer, V., Gottrand, F., Narita, T., & Desseyn, J. L. (2018). Gel-forming mucin interactome drives mucus viscoelasticity. *Advances in Colloid and Interface Science*, *252*(1), 69–82. https://doi.org/10.1016/j.cis.2017.12.005

Depommier, C., Everard, A., Druart, C., Plovier, H., Van Hul, M., Vieira-Silva, S., Falony, G., Raes, J., Maiter, D., Delzenne, N. M., de Barsy, M., Loumaye, A., Hermans, M. P., Thissen, J. P., de Vos, W. M., & Cani, P. D. (2019). Supplementation with *Akkermansia muciniphila* in overweight and obese human volunteers: a proof-of-concept exploratory study. *Nature Medicine*, *25*(7), 1096–1103. https://doi.org/10.1038/s41591-019-0495-2

Desai, M. S., Seekatz, A. M., Koropatkin, N. M., Kamada, N., Hickey, C. A., Wolter, M., Pudlo, N. A., Kitamoto, S., Terrapon, N., Muller, A., Young, V. B., Henrissat, B., Wilmes, P., Stappenbeck, T. S., Núñez, G., & Martens, E. C. (2016). A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. *Cell*, *167*(5), 1339--1353.e21. https://doi.org/10.1016/j.cell.2016.10.043

Dhanisha, S. S., Guruvayoorappan, C., & Abeesh, P. (2018). Mucins : structural diversity , biosynthesis , its role in pathogenesis and as possible therapeutic targets. *Critical Reviews in Oncology / Hematology*, *122*(June 2017), 98–122. https://doi.org/10.1016/j.critrevonc.2017.12.006

Díaz, R., Torres-Miranda, A., Orellana, G., & Garrido, D. (2021). Comparative Genomic Analysis of Novel Bifidobacterium longum subsp. longum Strains Reveals Functional Divergence in the Human Gut Microbiota. *Microorganisms*, *9*(9), 1906. https://doi.org/10.3390/microorganisms9091906

DiLeo, M. V., Strahan, G. D., den Bakker, M., & Hoekenga, O. A. (2011). Weighted

correlation network analysis (WGCNA) applied to the tomato fruit metabolome. *PLoS ONE*, *6*(10), e26683. https://doi.org/10.1371/journal.pone.0026683

Ding, M., Yang, B., Ross, R. P., Stanton, C., Zhao, J., Zhang, H., & Chen, W. (2021). Crosstalk between sIgA-coated bacteria in infant gut and early-life health. *Trends in Microbiology*, *29*(8), 725–735. https://doi.org/10.1016/j.tim.2021.01.012

Donovan, S. M. (2019). Human milk proteins: composition and physiological significance. In S. M. Donovan, J. B. German, B. Lönnerdal, & A. Lucas (Eds.), *Human Milk: Composition, Clinical Benefits and Future Opportunities* (1st ed., pp. 93–101). 2019 Nestlé Nutrition Institute, Switzerland/S. Karger AG, Basel. https://doi.org/10.1159/000490298

Downs, D. (2003). Genomics and bacterial metabolism. *Current Issues in Molecular Biology*, *5*(1), 17–25. https://doi.org/10.21775/cimb.005.017

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*(11), 3247. https://doi.org/10.3390/nu12113247

Duran-Pinedo, A. E., Paster, B., Teles, R., & Frias-Lopez, J. (2011). Correlation network analysis applied to complex biofilm communities. *PLoS ONE*, *6*(12), e28438. https://doi.org/10.1371/journal.pone.0028438

Dvorak, P., Hlavac, V., Mohelnikova-Duchonova, B., Liska, V., Pesta, M., & Soucek, P. (2017). Downregulation of ABC transporters in non-neoplastic tissues confers better prognosis for pancreatic and colorectal cancer patients. *Journal of Cancer*, *8*(11), 1959–1971. https://doi.org/10.7150/jca.19364

Earley, H., Lennon, G., Balfe, Á., Coffey, J. C., Winter, D. C., & O'Connell, P. R. (2019). The abundance of *Akkermansia muciniphila* and its relationship with sulphated colonic mucins in health and ulcerative colitis. *Scientific Reports*, *9*(1), 15683. https://doi.org/10.1038/s41598-019-51878-3

Egan, M., O'Connell-Motherway, M., & van Sinderen, D. (2015). A GntR-type transcriptional repressor controls sialic acid utilization in *Bifidobacterium breve* UCC2003. *FEMS Microbiology Letters*, *362*(4), 1–9. https://doi.org/10.1093/femsle/fnu056

Egan, M., O'Connell Motherway, M., Kilcoyne, M., Kane, M., Joshi, L., Ventura, M., & van Sinderen, D. (2014). Cross-feeding by *Bifidobacterium breve* UCC2003 during co-cultivation with *Bifidobacterium bifidum* PRL2010 in a mucin-based medium. *BMC Microbiology*, *14*(1), 282. https://doi.org/10.1186/s12866-014-0282-7

Everard, A., Lazarevic, V., Gaïa, N., Johansson, M., Ståhlman, M., Backhed, F., Delzenne, N. M., Schrenzel, J., François, P., & Cani, P. D. (2014). Microbiome of prebiotic-treated mice reveals novel targets involved in host response during obesity. *ISME Journal*, *8*(10), 2116–2130. https://doi.org/10.1038/ismej.2014.45

Fiat, A. M., & Jollès, P. (1989). Caseins of various origins and biologically active casein

peptides and oligosaccharides: structural and physiological aspects. *Molecular and Cellular Biochemistry*, *87*(1), 5-30.

Fletcher, C. M., Coyne, M. J., Villa, O. F., Chatzidaki-Livanis, M., & Comstock, L. E. (2009). A general *O*-glycosylation system important to the physiology of a major human intestinal symbiont. *Cell*, *137*(2), 321–331. https://doi.org/10.1016/j.cell.2009.02.041

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*(4), 289–306. https://doi.org/10.4161/gmic.19897

Flores-Bautista, E., Hernandez-Guerrero, R., Huerta-Saquero, A., Tenorio-Salgado, S., Rivera-Gomez, N., Romero, A., Ibarra, J. A., & Perez-Rueda, E. (2020). Deciphering the functional diversity of DNA-binding transcription factors in Bacteria and Archaea organisms. *PLOS ONE*, *15*(8), e0237135. https://doi.org/10.1371/journal.pone.0237135

Fujita, K., Fusako, O., Nagamine, N., Katayama, T., Hiratake, J., Sakata, K., Kumagai, H., & Yamamoto, K. (2005). Identification and molecular cloning of a novel glycoside hydrolase Family of Core 1 type *O*-glycan-specific Endo-α-*N*-acetylgalactosaminidase from *Bifidobacterium longum*. *Journal of Biological Chemistry*, *280*(45), 37415–37422.

Galán-Vásquez, E., & Perez-Rueda, E. (2019). Identification of modules with similar gene regulation and metabolic functions based on co-expression data. *Frontiers in Molecular Biosciences*, *6*(139). https://doi.org/10.3389/fmolb.2019.00139

Garrido, D., Barile, D., & Mills, D. A. (2012). A molecular basis for bifidobacterial enrichment in the infant gastrointestinal tract. *Advances in Nutrition*, *3*(3), 415S-421S. https://doi.org/10.3945/an.111.001586

Garrido, D., Nwosu, C., Ruiz-Moyano, S., Aldredge, D., German, J. B., Lebrilla, C. B., & Mills, D. A. (2012a). Endo-*β*-*N*-acetylglucosaminidases from infant gut-associated bifidobacteria release complex *N*-glycans from human milk glycoproteins. *Molecular & Cellular Proteomics*, *11*(9), 775–785. https://doi.org/10.1074/mcp.M112.018119

Garrido, D., Nwosu, C., Ruiz-Moyano, S., Aldredge, D., German, J. B., Lebrilla, C. B., & Mills, D. A. (2012b). Endo-β-*N*-acetylglucosaminidases from infant gut-associated Bifidobacteria release complex *N*-glycans from human milk glycoproteins. *Molecular & Cellular Proteomics*, *11*(9), 775–785. https://doi.org/10.1074/mcp.M112.018119

Garrido, D., Nwosu, C., Ruiz-Moyano, S., Aldredge, D., German, J. B., Lebrilla, C. B., & Mills, D. A. (2012c). Endo-*β*-*N*-acetylglucosaminidases from Infant Gut-associated Bifidobacteria Release Complex *N*-glycans from Human Milk Glycoproteins. *Molecular & Cellular Proteomics*, *11*(9), 775–785. https://doi.org/10.1074/mcp.M112.018119

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*(1), 35045. https://doi.org/10.1038/srep35045

Garrido, D., Ruiz-Moyano, S., Lemay, D. G., Sela, D. A., German, J. B., & Mills, D. A. (2015a). Comparative transcriptomics reveals key differences in the response to milk oligosaccharides of infant gut-associated bifidobacteria. *Scientific Reports*, *5*(1), 13517. https://doi.org/10.1038/srep13517

Garrido, D., Ruiz-Moyano, S., Lemay, D. G., Sela, D. A., German, J. B., & Mills, D. A. (2015b). Comparative transcriptomics reveals key differences in the response to milk oligosaccharides of infant gut-associated bifidobacteria. *Scientific Reports*, *5*. https://doi.org/10.1038/srep13517

Garrido, D., Ruiz-Moyano, S., Lemay, D. G., Sela, D. A., German, J. B., & Mills, D. A. (2015c). Comparative transcriptomics reveals key differences in the response to milk oligosaccharides of infant gut-associated Bifidobacteria. *Scientific Reports*, *5*(1), 13517. https://doi.org/10.1038/srep13517

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*(4), 430–435. https://doi.org/10.1016/j.anaerobe.2012.04.012

Gibson, G., Probert, H. M., Loo, J. Van, Rastall, R. A., & Roberfroid, M. B. (2017). Dietary modulation of the human colonic microbiota: updating the concept of prebiotics. *Nutrition Research Reviews*, *14*(2), 259–275. https://doi.org/10.1079/NRR200479

Gibson, G. R., Hutkins, R., Sanders, M. E., Prescott, S. L., Reimer, R. A., Salminen, S. J., Scott, K., Stanton, C., Swanson, K. S., Cani, P. D., Verbeke, K., & Reid, G. (2017). Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. *Nature Reviews Gastroenterology & Hepatology*, *14*(8), 491–502. https://doi.org/10.1038/nrgastro.2017.75

Gibson, G. R., McCartney, A. L., & Rastall, R. A. (2005). Prebiotics and resistance to gastrointestinal infections. *British Journal of Nutrition*, *93*(S1), S31–S34. https://doi.org/10.1079/BJN20041343

Goehring, K. C., Kennedy, A. D., Prieto, P. A., & Buck, R. H. (2014). Direct evidence for the presence of human milk oligosaccharides in the circulation of breastfed infants. *PLoS ONE*, *9*(7), e101692. https://doi.org/10.1371/journal.pone.0101692

Goettig, P. (2016). Effects of glycosylation on the enzymatic activity and mechanisms of proteases. In *International Journal of Molecular Sciences* (Vol. 17, Issue 12). MDPI AG. https://doi.org/10.3390/ijms17121969

González-Morelo, K. J., Galán-Vásquez, E., Melis, F., Pérez-Rueda, E., & Garrido, D. (2023). Structure of co-expression networks of *Bifidobacterium* species in response to human milk oligosaccharides. *Frontiers in Molecular Biosciences*, *10*. https://doi.org/10.3389/fmolb.2023.1040721

González-Morelo, K. J., Vega-Sagardía, M., & Garrido, D. (2020a). Molecular insights

into *O*-linked glycan utilization by gut microbes. *Frontiers in Microbiology*, *11*. https://doi.org/10.3389/fmicb.2020.591568

González-Morelo, K. J., Vega-Sagardía, M., & Garrido, D. (2020b). Molecular Insights Into *O*-Linked Glycan Utilization by Gut Microbes. *Frontiers in Microbiology*, *11*, 2834. https://doi.org/10.3389/fmicb.2020.591568

Gralka, M., Szabo, R., Stocker, R., & Cordero, O. X. (2020). Trophic interactions and the drivers of microbial community assembly. *Current Biology*, *30*(19), R1176–R1188. https://doi.org/10.1016/j.cub.2020.08.007

Graziani, F., Pujol, A., Nicoletti, C., Dou, S., Maresca, M., Giardina, T., Fons, M., & Perrier, J. (2016). *Ruminococcus gnavus* E1 modulates mucin expression and intestinal glycosylation. *Journal of Applied Microbiology*, *120*(5), 1403–1417. https://doi.org/10.1111/jam.13095

Groux-Degroote, S., Cavdarli, S., Uchimura, K., Allain, F., & Delannoy, P. (2020). Glycosylation changes in inflammatory diseases. *Advances in Protein Chemistry and Structural Biology*, *119*, 111–156. https://doi.org/10.1016/bs.apcsb.2019.08.008

Guan, B., Zhang, Z., Chai, Y., Amantai, X., Chen, X., Cao, X., & Yue, X. (2022). N-glycosylation of milk proteins: A review spanning 2010–2022. *Trends in Food Science & Technology*, *128*, 1–21. https://doi.org/10.1016/j.tifs.2022.07.017

Guan, Q., & Zhang, J. (2017). Recent advances: the imbalance of cytokines in the pathogenesis of inflammatory bowel disease. *Mediators of Inflammation*, *2017*, 1–8. https://doi.org/10.1155/2017/4810258

Gutiérrez, N., & Garrido, D. (2019). Species deletions from microbiome consortia reveal key metabolic interactions between gut microbes. *MSystems*, *4*(4). https://doi.org/10.1128/mSystems.00185-19

Hall, A. B., Yassour, M., Sauk, J., Garner, A., Jiang, X., Arthur, T., Lagoudas, G. K., Vatanen, T., Fornelos, N., Wilson, R., Bertha, M., Cohen, M., Garber, J., Khalili, H., Gevers, D., Ananthakrishnan, A. N., Kugathasan, S., Lander, E. S., Blainey, P., … Huttenhower, C. (2017). A novel *Ruminococcus gnavus* clade enriched in inflammatory bowel disease patients. *Genome Medicine*, *9*(1), 1–12. https://doi.org/10.1186/s13073-017-0490-5

Hasan, N., & Yang, H. (2019). Factors affecting the composition of the gut microbiota, and its modulation. *PeerJ*, *7*, e7502. https://doi.org/10.7717/peerj.7502

Henrick, B. M., Chew, S., Casaburi, G., Brown, H. K., Frese, S. A., Zhou, Y., Underwood, M. A., & Smilowitz, J. T. (2019). Colonization by B. *infantis* EVC001 modulates enteric inflammation in exclusively breastfed infants. *Pediatric Research*, *86*(6), 749–757. https://doi.org/10.1038/s41390-019-0533-2

Hidalgo, M. E., Pires, M. S., & Risso, P. H. (2010). A study on bovine kappa-casein aggregation after the enzymatic action of chymosin. *Colloids and Surfaces B: Biointerfaces*, *76*(2), 556–563. https://doi.org/10.1016/j.colsurfb.2009.12.018

Hills, R., Pontefract, B., Mishcon, H., Black, C., Sutton, S., & Theberge, C. (2019). Gut microbiome: profound implications for diet and disease. *Nutrients*, *11*(7), 1613.

https://doi.org/10.3390/nu11071613

Hirata, T., & Kizuka, Y. (2021). N-*Glycosylation* (pp. 3–24). https://doi.org/10.1007/978-3-030-70115-4_1

Hirmas, B., Gasaly, N., Orellana, G., Vega-Sagardía, M., Saa, P., Gotteland, M., & Garrido, D. (2022). Metabolic modeling and bidirectional culturing of two gut microbes reveal cross-feeding interactions and protective effects on intestinal cells. *MSystems*, *7*(5). https://doi.org/10.1128/msystems.00646-22

Horvath, S. (2011). *Weighted Network Analysis|Applications in genomics and systems biology* (1st ed.). Springer New York. https://doi.org/10.1007/978-1-4419-8819-5

Hou, K., Wu, Z.-X., Chen, X.-Y., Wang, J.-Q., Zhang, D., Xiao, C., Zhu, D., Koya, J. B., Wei, L., Li, J., & Chen, Z.-S. (2022). Microbiota in health and diseases. *Signal Transduction and Targeted Therapy*, *7*(1), 135. https://doi.org/10.1038/s41392-022-00974-4

Idota, T., Kawakami, H., & Nakajima, I. (1994). Growth-promoting Effects of N -Acetylneuraminic Acid-containing Substances on Bifidobacteria. *Bioscience, Biotechnology, and Biochemistry*, *58*(9), 1720–1722. https://doi.org/10.1271/bbb.58.1720

Itano, N. (2019). Implications of altered *O*-glycosylation in tumour immune evasion. *The Journal of Biochemistry*, *165*(5), 387–390. https://doi.org/10.1093/jb/mvz003

Itzkowitz, S. H., Yuan, M., Montgomery, C. K., Kjeldsen, T., Takahashi, H. K., Bigbee, W. L., & Kim, Y. S. (1989). Expression of Tn, Sialosyl-Tn, and T Antigens in human colon cancer. *Cancer Research*, *49*(1), 197–204.

Jha, R. K., Udupa, S., Rai, A. K., Rani, P., Singh, P. R., Govind, S., & Nagaraja, V. (2020). Conditional down-regulation of GreA impacts expression of rRNA and transcription factors, affecting *Mycobacterium smegmatis* survival. *Scientific Reports*, *10*(1), 5802. https://doi.org/10.1038/s41598-020-62703-7

Jia, R., Zhao, H., & Jia, M. (2020). Identification of co-expression modules and potential biomarkers of breast cancer by WGCNA. *Gene*, *750*, 144757. https://doi.org/10.1016/j.gene.2020.144757

Jin, C., Kenny, D. T., Skoog, E. C., Padra, M., Adamczyk, B., Vitizeva, V., Thorell, A., Venkatakrishnan, V., Lindén, S. K., & Karlsson, N. G. (2017). Structural diversity of human gastric mucin glycans. *Molecular & Cellular Proteomics*, *16*(5), 743–758. https://doi.org/10.1074/mcp.M117.067983

Johansson, M. E. V., Jakobsson, H. E., Holmén-Larsson, J., Schütte, A., Ermund, A., Rodríguez-Piñeiro, A. M., Arike, L., Wising, C., Svensson, F., Bäckhed, F., & Hansson, G. C. (2015). Normalization of host intestinal mucus layers requires long-term microbial colonization. *Cell Host and Microbe*. https://doi.org/10.1016/j.chom.2015.10.007

Junker, B. H., & Schreiber, F. (2008). *Analysis of Biological Networks* (B. H. Junker & F. Schreiber (eds.)). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470253489

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*(1), 27–30. https://doi.org/10.1093/nar/28.1.27

Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., & Tanabe, M. (2019). New approach for understanding genome variations in KEGG. *Nucleic Acids Research*, *47*(D1), D590–D595. https://doi.org/10.1093/nar/gky962

Karav, S., Le Parc, A., Leite Nobrega de Moura Bell, J. M., Frese, S. A., Kirmiz, N., Block, D. E., Barile, D., & Mills, D. A. (2016). Oligosaccharides released from milk glycoproteins are selective growth substrates for infant-associated bifidobacteria. *Applied and Environmental Microbiology*, *82*(12), 3622–3630. https://doi.org/10.1128/AEM.00547-16

Katoh, T., Maeshibu, T., Kikkawa, K., Gotoh, A., Tomabechi, Y., Nakamura, M., Liao, W.-H., Yamaguchi, M., Ashida, H., Yamamoto, K., & Katayama, T. (2017). Identification and characterization of a sulfoglycosidase from *Bifidobacterium bifidum* implicated in mucin glycan utilization. *Bioscience, Biotechnology, and Biochemistry*, *81*(10), 2018–2027. https://doi.org/10.1080/09168451.2017.1361810

Katoh, T., Ojima, M. N., Sakanaka, M., Ashida, H., Gotoh, A., & Katayama, T. (2020). Enzymatic adaptation of *Bifidobacterium bifidum* to Host glycans, viewed from glycoside hydrolyases and carbohydrate-binding modules. *Microorganisms*, *8*(4), 481. https://doi.org/10.3390/microorganisms8040481

Keel, K., Harte, F. M., Berbejillo, J., & López-Pedemonte, T. (2022). Functionality of glycomacropeptide glycated with lactose and maltodextrin. *Journal of Dairy Science*, *105*(11), 8664–8676. https://doi.org/10.3168/jds.2022-21959

Kesimer, M., & Sheehan, J. K. (2012). Mass spectrometric analysis of mucin core proteins. *Methods in Molecular Biology*, *842*(1), 67–79. https://doi.org/10.1007/978-1-61779-513-8_4

Kethireddipalli, P., & Hill, A. R. (2015). Rennet coagulation and cheesemaking properties of thermally processed milk: overview and recent developments. *Journal of Agricultural and Food Chemistry*, *63*(43), 9389–9403. https://doi.org/10.1021/jf504167v

Khoroshkin, M. S., Leyn, S. A., Van Sinderen, D., & Rodionov, D. A. (2016). Transcriptional regulation of carbohydrate utilization pathways in the *Bifidobacterium* genus. *Frontiers in Microbiology*, *7*(1). https://doi.org/10.3389/fmicb.2016.00120

Kim, J.-H., An, H. J., Garrido, D., German, J. B., Lebrilla, C. B., & Mills, D. A. (2013). Proteomic analysis of *Bifidobacterium longum* subsp. *infantis* reveals the metabolic insight on consumption of prebiotics and host glycans. *PLoS ONE*, *8*(2), e57535.

Kirmiz, N., Robinson, R. C., Shah, I. M., Barile, D., & Mills, D. A. (2018). Milk glycans and their interaction with the infant-gut microbiota. *Annual Review of Food Science and Technology*, *9*(1), 429–450. https://doi.org/10.1146/annurev-food-030216-030207

Kitaoka, M. (2012a). Bifidobacterial enzymes involved in the metabolism of human milk oligosaccharides. *Advances in Nutrition*, *3*(3), 422S-429S. https://doi.org/10.3945/an.111.001420

Kitaoka, M. (2012b). Bifidobacterial Enzymes Involved in the Metabolism of Human Milk Oligosaccharides. *Advances in Nutrition*, *3*(3), 422S-429S. https://doi.org/10.3945/an.111.001420

Kiyohara, M., Nakatomi, T., Kurihara, S., Fushinobu, S., Suzuki, H., Tanaka, T., Shoda, S., Kitaoka, M., Katayama, T., Yamamoto, K., & Ashida, H. (2012). α-*N*-acetylgalactosaminidase from infant-associated Bifidobacteria belonging to novel glycoside hydrolase Family 129 is implicated in alternative mucin degradation pathway. *Journal of Biological Chemistry*, *287*(1), 693–700.

Koutsioulis, D., Landry, D., & Guthrie, E. P. (2008a). Novel endo-α-*N*-acetylgalactosaminidases with broader substrate specificity. *Glycobiology*, *18*(10), 799–805. https://doi.org/10.1093/glycob/cwn069

Koutsioulis, D., Landry, D., & Guthrie, E. P. (2008b). Novel endo-α-*N*-acetylgalactosaminidases with broader substrate specificity. *Glycobiology*, *18*(10), 799–805. https://doi.org/10.1093/glycob/cwn069

Kukurba, K. R., & Montgomery, S. B. (2015). RNA sequencing and analysis. *Cold Spring Harbor Protocols*, *2015*(11), pdb.top084970. https://doi.org/10.1101/pdb.top084970

La Rosa, S. L., Ostrowski, M. P., Vera-Ponce de León, A., McKee, L. S., Larsbrink, J., Eijsink, V. G., Lowe, E. C., Martens, E. C., & Pope, P. B. (2022). Glycan processing in gut microbiomes. *Current Opinion in Microbiology*, *67*, 102143. https://doi.org/10.1016/j.mib.2022.102143

Lamacchia, C., Musaico, D., Henderson, M. E., Bergillos-meca, T., Roul, M., Landriscina, L., Decina, I., Corona, G., & Costabile, A. (2018). Temperature-treated gluten proteins in Gluten-Friendly ™ bread increase mucus production and gut-barrier function in human intestinal goblet cells. *Journal of Functional Foods*, *48*(1), 507–514. https://doi.org/10.1016/j.jff.2018.07.047

Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1), 559. https://doi.org/10.1186/1471-2105-9-559

Lawson, M. A. E., O'Neill, I. J., Kujawska, M., Gowrinadh Javvadi, S., Wijeyesekera, A., Flegg, Z., Chalklen, L., & Hall, L. J. (2020). Breast milk-derived human milk oligosaccharides promote *Bifidobacterium* interactions within a single ecosystem. *ISME Journal*, *14*(2), 635–648. https://doi.org/10.1038/s41396-019-0553-2

Lecuit, M., & Eloit, M. (2017). The Viruses of the Gut Microbiota. In *The Microbiota in Gastrointestinal Pathophysiology* (pp. 179–183). Elsevier. https://doi.org/10.1016/B978-0-12-804024-9.00021-5

Lee, S. M., Donaldson, G. P., Mikulski, Z., Boyajian, S., Ley, K., & Mazmanian, S. K. (2013). Bacterial colonization factors control specificity and stability of the gut microbiota. *Nature*, *501*(7467), 426–429. https://doi.org/10.1038/nature12447

Lewis, M. (2005). The lac repressor. *Comptes Rendus Biologies*, *328*(6), 521–548. https://doi.org/10.1016/j.crvi.2005.04.004

Li, E. W. Y., & Mine, Y. (2004). Technical note: Comparison of chromatographic profile

of glycomacropeptide from cheese whey isolated using different methods. *Journal of Dairy Science*, *87*(1), 174–177. https://doi.org/10.3168/jds.S0022-0302(04)73155-0

Li, Z., & Chai, W. (2019). Mucin *O*-glycan microarrays. *Current Opinion in Structural Biology*, *56*, 187–197. https://doi.org/10.1016/j.sbi.2019.03.032

Liao, Y., Wang, Y., Cheng, M., Huang, C., & Fan, X. (2020). Weighted gene coexpression network analysis of features that control cancer stem cells reveals prognostic biomarkers in lung adenocarcinoma. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.00311

Lillehoj, E. P., Kato, K., Lu, W., & Kim, K. C. (2013). Cellular and molecular biology of airway mucins. In *International review of cell and molecular biology* (Vol. 303, pp. 139–202). Elsevier. https://doi.org/10.1016/B978-0-12-407697-6.00004-0

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. *Journal of Infectious Diseases*, *209*(3), 389–398. https://doi.org/10.1093/infdis/jit464

Liu, B., & Newburg, D. S. (2013). Human milk glycoproteins protect infants against human pathogens. *Breastfeeding Medicine*, *8*(4), 354–362. https://doi.org/10.1089/bfm.2013.0016

Liu, Y., Yu, X., Zhao, J., Zhang, H., Zhai, Q., & Chen, W. (2020). The role of MUC2 mucin in intestinal homeostasis and the impact of dietary components on MUC2 expression. *International Journal of Biological Macromolecules*, *164*, 884–891. https://doi.org/10.1016/j.ijbiomac.2020.07.191

Liu, Z., Li, L., Wang, Q., Sadiq, F. A., Lee, Y., Zhao, J., Zhang, H., Chen, W., Li, H., & Lu, W. (2021). Transcriptome analysis reveals the genes involved in *Bifidobacterium Longum* FGSZY16M3 biofilm formation. *Microorganisms*, *9*(2), 385. https://doi.org/10.3390/microorganisms9020385

Lopez-Siles, M., Enrich-Capó, N., Aldeguer, X., Sabat-Mir, M., Duncan, S. H., Garcia-Gil, L. J., & Martinez-Medina, M. (2018). Alterations in the abundance and co-occurrence of *Akkermansia muciniphila* and *Faecalibacterium prausnitzii* in the Colonic mucosa of inflammatory bowel disease subjects. *Frontiers in Cellular and Infection Microbiology*, *8*. https://doi.org/10.3389/fcimb.2018.00281

Luis, A. S., Briggs, J., Zhang, X., Farnell, B., Ndeh, D., Labourel, A., Baslé, A., Cartmell, A., Terrapon, N., Stott, K., Lowe, E. C., McLean, R., Shearer, K., Schückel, J., Venditto, I., Ralet, M.-C., Henrissat, B., Martens, E. C., Mosimann, S. C., … Gilbert, H. J. (2018). Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic *Bacteroides*. *Nature Microbiology*, *3*(2), 210–219. https://doi.org/10.1038/s41564-017-0079-1

Luis, A. S., & Hansson, G. C. (2023). Intestinal mucus and their glycans: A habitat for thriving microbiota. *Cell Host & Microbe*, *31*(7), 1087–1100. https://doi.org/10.1016/j.chom.2023.05.026

MacMillan, J. L., Vicaretti, S. D., Noyovitz, B., Xing, X., Low, K. E., Inglis, G. D.,

Zaytsoff, S. J. M., Boraston, A. B., Smith, S. P., Uwiera, R. R. E., Selinger, L. B., Zandberg, W. F., & Abbott, D. W. (2019). Structural analysis of broiler chicken small intestinal mucin *O*-glycan modification by *Clostridium* <i>perfringens<i>. *Poultry Science*, *98*(10), 5074–5088. https://doi.org/10.3382/ps/pez297

Magnelli, P. E., Bielik, A. M., & Guthrie, E. P. (2011). Identification and characterization of protein glycosylation using specific endo- and exoglycosidases. *Journal of Visualized Experiments*, *58*(1), 8–12. https://doi.org/10.3791/3749

Manso, M. A., & López-Fandiño, R. (2004). κ-casein macropeptides from cheese whey: physicochemical, biological, nutritional, and technological features for possible uses. *Food Reviews International*, *20*(4), 329–355. https://doi.org/10.1081/FRI-200033456

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*(5), 507–514. https://doi.org/10.1016/j.chom.2011.10.007

Markowiak, P., & Śliżewska, K. (2017). Effects of probiotics, prebiotics, and symbiotics on human health. *Nutrients*, *9*(9), 1021. https://doi.org/10.3390/nu9091021

Martens, E. C., Chiang, H. C., & Gordon, J. I. (2008). Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host & Microbe*, *4*(5), 447–457. https://doi.org/10.1016/j.chom.2008.09.007

Martin, A. J. M., Serebrinsky-Duek, K., Riquelme, E., Saa, P. A., & Garrido, D. (2023). Microbial interactions and the homeostasis of the gut microbiome: the role of *Bifidobacterium*. *Microbiome Research Reports*, *2*(3), 17. https://doi.org/10.20517/mrr.2023.10

Martin, C., Lin, P.-R., & Blackburn, G. (2016). Review of infant feeding: key features of breast Milk and infant formula. *Nutrients*, *8*(5), 279. https://doi.org/10.3390/nu8050279

Martinez, F., Balciunas, E. M., Converti, A., Cotter, P. D., & de Souza Oliveira, R. P. (2013). Bacteriocin production by *Bifidobacterium* spp. A review. *Biotechnology Advances*, *31*(4), 482–488. https://doi.org/10.1016/j.biotechadv.2013.01.010

Marx, C., Bridge, R., Wolf, A. K., Rich, W., Kim, J. H., & Bode, L. (2014). Human Milk Oligosaccharide composition differs between donor milk and mother's own milk in the NICU. *Journal of Human Lactation*, *30*(1), 54–61. https://doi.org/10.1177/0890334413513923

Masi, A. C., & Stewart, C. J. (2022). Untangling human milk oligosaccharides and infant gut microbiome. *IScience*, *25*(1), 103542. https://doi.org/10.1016/j.isci.2021.103542

Medina, D. A., Pinto, F., Ovalle, A., Thomson, P., & Garrido, D. (2017). Prebiotics Mediate Microbial Interactions in a Consortium of the Infant Gut Microbiome. *International Journal of Molecular Sciences*, *18*(2095), 1–16. https://doi.org/10.3390/ijms18102095

Medina, D., Pinto, F., Ovalle, A., Thomson, P., & Garrido, D. (2017). Prebiotics mediate

microbial interactions in a consortium of the infant gut microbiome. *International Journal of Molecular Sciences*, *18*(10), 2095. https://doi.org/10.3390/ijms18102095

Mendis, P. M., & Jackson, G. P. (2022). Structural characterization of human milk oligosaccharides using ultrahigh performance liquid chromatography–helium charge transfer dissociation mass spectrometry. *Glycobiology*, *32*(6), 483–495. https://doi.org/10.1093/glycob/cwac010

Mendoza, S. N., Cañón, P. M., Contreras, Á., Ribbeck, M., & Agosín, E. (2017). Genome-scale reconstruction of the metabolic network in *Oenococcus oeni* to assess wine malolactic fermentation. *Frontiers in Microbiology*, *8*(534). https://doi.org/10.3389/fmicb.2017.00534

Meng, X., Wu, H., Wang, W., Lan, T., Yang, W., Yu, D., & Fang, X. (2020). A purified aspartic protease from *Akkermansia muciniphila* plays an important role in degrading Muc2. *International Journal of Molecular Sciences*, *21*(1), 72. https://doi.org/10.3390/ijms21010072

Mészáros, Z., Nekvasilová, P., Bojarová, P., Křen, V., & Slámová, K. (2021). Advanced glycosidases as ingenious biosynthetic instruments. *Biotechnology Advances*, *49*, 107733. https://doi.org/10.1016/j.biotechadv.2021.107733

Milani, C., Turroni, F., Duranti, S., Lugli, G. A., Mancabelli, L., Ferrario, C., van Sinderen, D., & Ventura, M. (2016). Genomics of the genus *Bifidobacterium* reveals species-specific adaptation to the glycan-rich gut environment. *Applied and Environmental Microbiology*, *82*(4), 980–991. https://doi.org/10.1128/AEM.03500-15

Miyazaki, T., Ikegaya, M., & Alonso-Gil, S. (2022). Structural and mechanistic insights into the substrate specificity and hydrolysis of GH31 *α*-*N*-acetylgalactosaminidase. *Biochimie*, *195*, 90–99. https://doi.org/10.1016/j.biochi.2021.11.007

Morio, A., Xu, J., Masuda, A., Kinoshita, Y., Hino, M., Morokuma, D., Goda, H. M., Okino, N., Ito, M., Mon, H., Fujita, R., Kusakabe, T., & Lee, J. M. (2019). Expression, purification, and characterization of highly active endo-α-*N*-acetylgalactosaminidases expressed by silkworm-baculovirus expression system. *Journal of Asia-Pacific Entomology*, *22*(2), 404–408. https://doi.org/10.1016/j.aspen.2019.01.009

Morozumi, M., Wada, Y., Tsuda, M., Tabata, F., Ehara, T., Nakamura, H., & Miyaji, K. (2023). Cross-feeding among bifidobacteria on glycomacropeptide. *Journal of Functional Foods*, *103*, 105463. https://doi.org/10.1016/j.jff.2023.105463

Morrow, A. L., Ruiz-Palacios, G. M., Jiang, X., & Newburg, D. S. (2005). Human-milk glycans that inhibit pathogen binding protect breast-feeding infants against infectious diarrhea. *The Journal of Nutrition*, *135*(5), 1304–1307. https://doi.org/10.1093/jn/135.5.1304

Nagpal, R., Kumar, M., Yadav, A. K., Hemalatha, R., Yadav, H., Marotta, F., & Yamashiro, Y. (2016). Gut microbiota in health and disease: an overview focused on metabolic inflammation. *Beneficial Microbes*, *7*(2), 181–194. https://doi.org/10.3920/bm2015.0062

Neelima, Sharma, R., Rajput, Y. S., & Mann, B. (2013). Chemical and functional properties of glycomacropeptide (GMP) and its role in the detection of cheese whey adulteration in milk: a review. *Dairy Science & Technology*, *93*(1), 21–43. https://doi.org/10.1007/s13594-012-0095-0

Nishiyama, K., Yamamoto, Y., Sugiyama, M., Takaki, T., Urashima, T., Fukiya, S., Yokota, A., Okada, N., & Mukai, T. (2017). *Bifidobacterium bifidum* extracellular sialidase enhances adhesion to the mucosal surface and supports carbohydrate assimilation. *MBio*, *8*(5), 1–15. https://doi.org/10.1128/mBio.00928-17

O' Riordan, N., Callaghan, J. O., Buttò, L. F., Kilcoyne, M., Joshi, L., & Hickey, R. M. (2018). Bovine glycomacropeptide promotes the growth of *Bifidobacterium longum* ssp. *infantis* and modulates its gene expression. *Journal of Dairy Science*, *101*(8), 6730–6741. https://doi.org/10.3168/jds.2018-14499

O'Callaghan, A., & van Sinderen, D. (2016). Bifidobacteria and their role as members of the human gut microbiota. *Frontiers in Microbiology*, *7*(1), 925. https://doi.org/10.3389/fmicb.2016.00925

O'Riordan, N., O'Callaghan, J., Buttò, L. F., Kilcoyne, M., Joshi, L., & Hickey, R. M. (2018). Bovine glycomacropeptide promotes the growth of *Bifidobacterium longum* ssp. *infantis* and modulates its gene expression. *Journal of Dairy Science*, *101*(8), 6730–6741. https://doi.org/10.3168/jds.2018-14499

Ojima, M. N., Jiang, L., Arzamasov, A. A., Yoshida, K., Odamaki, T., Xiao, J., Nakajima, A., Kitaoka, M., Hirose, J., Urashima, T., Katoh, T., Gotoh, A., van Sinderen, D., Rodionov, D. A., Osterman, A. L., Sakanaka, M., & Katayama, T. (2022). Priority effects shape the structure of infant-type *Bifidobacterium* communities on human milk oligosaccharides. *The ISME Journal*. https://doi.org/10.1038/s41396-022-01270-3

Okburan, G., & Kızıler, S. (2023). Human milk oligosaccharides as prebiotics. *Pediatrics & Neonatology*, *64*(3), 231–238. https://doi.org/10.1016/j.pedneo.2022.09.017

Oliveira, D., Wilbey, R. A., Grandison, A. S., & Roseiro, L. B. (2015). Milk oligosaccharides: a review. *International Journal of Dairy Technology*, *68*(3), 305–321. https://doi.org/10.1111/1471-0307.12209

Ouwerkerk, J. P., van der Ark, K. C. H., Davids, M., Claassens, N. J., Finestra, T. R., de Vos, W. M., & Belzer, C. (2016). Adaptation of *Akkermansia muciniphila* to the oxic-anoxic interface of the mucus layer. *Applied and Environmental Microbiology*, *82*(23), 6983–6993. https://doi.org/10.1128/AEM.01641-16

Ozsolak, F., & Milos, P. M. (2011). RNA sequencing: advances, challenges and opportunities. *Nature Reviews Genetics*, *12*(2), 87–98. https://doi.org/10.1038/nrg2934

Padra, M., Adamczyk, B., Benktander, J., Flahou, B., Skoog, C., Padra, J. T., Smet, A., Jin, C., Samuelsson, T., Haesebrouck, F., Karlsson, N. G., Teneberg, S., & Lindén, S. K. (2018). *Helicobacter suis* binding to carbohydrates on human and porcine gastric mucins and glycolipids occurs via two modes. *Virulence*, *9*(1), 898–918.

https://doi.org/10.1080/21505594.2018.1460979

Parc, A. Le, Karav, S., Maria, J., Nobrega, L., Bell, D. M., Frese, S. A., Mills, D. A., Block, D. E., & Barile, D. (2015). A novel endo-β-*N*-acetylglucosaminidase releases specific *N*-glycans depending on different reaction conditions. *Biotechnology Progress*, *31*(5), 1323–1330. https://doi.org/10.1002/btpr.2133

Parhi, P., Song, K. P., & Choo, W. S. (2021). Effect of inulin and fructooligosaccharide supplementation on the growth and survival of *Lactobacillus casei* in model sugar systems. *Journal of Food Processing and Preservation*, *45*(3). https://doi.org/10.1111/jfpp.15228

Passos, M. do C. F., & Moraes-Filho, J. P. (2017). Intestinal microbiota in digestive diseases. *Arquivos de Gastroenterologia*, *54*(3), 255–262. https://doi.org/10.1590/s0004-2803.201700000-31

Peredo-Lovillo, A., Romero-Luna, H. E., & Jiménez-Fernández, M. (2020). Health promoting microbial metabolites produced by gut microbiota after prebiotics metabolism. *Food Research International*, *136*, 109473. https://doi.org/10.1016/j.foodres.2020.109473

Peterson, C. T., Perez Santiago, J., Iablokov, S. N., Chopra, D., Rodionov, D. A., & Peterson, S. N. (2022). Short-Chain fatty acids modulate healthy gut microbiota composition and functional potential. *Current Microbiology*, *79*(5), 128. https://doi.org/10.1007/s00284-022-02825-5

Peterson, R., Cheah, W. Y., Grinyer, J., & Packer, N. (2013). Glycoconjugates in human milk: Protecting infants from disease. *Glycobiology*, *23*(12), 1425–1438. https://doi.org/10.1093/glycob/cwt072

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, *29*(9), 45e – 45. https://doi.org/10.1093/nar/29.9.e45

Pinho, S. S., & Reis, C. A. (2015). Glycosylation in cancer: mechanisms and clinical implications. *Nature Publishing Group*, *16*(August). https://doi.org/10.1038/nrc3982

Pluvinage, B., Massel, P. M., Burak, K., & Boraston, A. B. (2019). Structural and functional analysis of four Family 84 glycoside hydrolases from the opportunistic pathogen *Clostridium perfringens*. *Glycobiology*, *30*(1), 58–68. https://doi.org/10.1093/glycob/cwz069

Png, C. W., Lindén, S. K., Gilshenan, K. S., Zoetendal, E. G., McSweeney, C. S., Sly, L. I., McGuckin, M. A., & Florin, T. H. J. (2010). Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. *American Journal of Gastroenterology*, *105*(11), 2420–2428. https://doi.org/10.1038/ajg.2010.281

Podolsky, D. K. (1985). Oligosaccharide structures of human colonic mucin. *The Journal of Biological Chemistry*, *260*(14), 8262–8271.

Praharaj, A. B., Dehury, B., Mahapatra, N., Kar, S. K., & Behera, S. K. (2018). Molecular dynamics insights into the structure, function, and substrate binding mechanism of

mucin desulfating sulfatase of gut microbe *Bacteroides fragilis*. *Journal of Cellular Biochemistry*, *119*(4), 3618–3631. https://doi.org/10.1002/jcb.26569

Pruss, K. M., Marcobal, A., Southwick, A. M., Dahan, D., Smits, S. A., Ferreyra, J. A., Higginbottom, S. K., Sonnenburg, E. D., Kashyap, P. C., Choudhury, B., Bode, L., & Sonnenburg, J. L. (2021). Mucin-derived *O*-glycans supplemented to diet mitigate diverse microbiota perturbations. *The ISME Journal*, *15*(2), 577–591. https://doi.org/10.1038/s41396-020-00798-6

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. *Journal of Pediatric Gastroenterology and Nutrition*, *64*(4), 624–631. https://doi.org/10.1097/MPG.0000000000001520

Pudlo, N. A., Urs, K., Kumar, S. S., German, J. B., Mills, D. A., & Martens, E. C. (2015). Symbiotic human gut bacteria with variable metabolic priorities for host mucosal glycans. *MBio*, *6*(6), e01282-15. https://doi.org/10.1128/mBio.01282-15

Purcell, R. V., Pearson, J., Aitchison, A., Dixon, L., Frizelle, F. A., & Keenan, J. I. (2017). Colonization with enterotoxigenic *Bacteroides fragilis* is associated with early-stage colorectal neoplasia. *PLoS ONE*, *12*(2), e0171602. https://doi.org/10.1371/journal.pone.0171602

Quintana-Hayashi, M., Padra, M., Padra, J., Benktander, J., & Lindén, S. (2018). Mucus-pathogen interactions in the gastrointestinal tract of farmed animals. *Microorganisms*, *6*(2), 55. https://doi.org/10.3390/microorganisms6020055

Raev, S. A., Amimo, J. O., Saif, L. J., & Vlasova, A. N. (2023). Intestinal mucin-type *O* -glycans: the major players in the host-bacteria-rotavirus interactions. *Gut Microbes*, *15*(1). https://doi.org/10.1080/19490976.2023.2197833

Ramos, J. L., Martínez-Bueno, M., Molina-Henares, A. J., Terán, W., Watanabe, K., Zhang, X., Gallegos, M. T., Brennan, R., & Tobes, R. (2005). The TetR family of transcriptional repressors. *Microbiology and Molecular Biology Reviews*, *69*(2), 326–356. https://doi.org/10.1128/MMBR.69.2.326-356.2005

Ravcheev, D. A., Khoroshkin, M. S., Laikova, O. N., Tsoy, O. V., Sernova, N. V., Petrova, S. A., Rakhmaninova, A. B., Novichkov, P. S., Gelfand, M. S., & Rodionov, D. A. (2014). Comparative genomics and evolution of regulons of the LacI-family transcription factors. *Frontiers in Microbiology*, *5*. https://doi.org/10.3389/fmicb.2014.00294

Ravcheev, D. A., & Thiele, I. (2017). Comparative genomic analysis of the human gut microbiome reveals a broad distribution of metabolic pathways for the degradation of host-synthetized mucin glycans and utilization of mucin-derived monosaccharides. *Frontiers in Genetics*, *8*(1), 1–22. https://doi.org/10.3389/fgene.2017.00111

Rawi, M. H., Zaman, S. A., Pa'ee, K. F., Leong, S. S., & Sarbini, S. R. (2020). Prebiotics metabolism by gut-isolated probiotics. *Journal of Food Science and Technology*, *57*(8), 2786–2799. https://doi.org/10.1007/s13197-020-04244-5

Rezaei, Z., Ranjbaran, J., Safarpour, H., Nomiri, S., Salmani, F., Chamani, E., Larki, P., Brunetti, O., Silvestris, N., & Tavakoli, T. (2022). Identification of early diagnostic biomarkers via WGCNA in gastric cancer. *Biomedicine & Pharmacotherapy*, *145*, 112477. https://doi.org/10.1016/j.biopha.2021.112477

Ringot-Destrez, B., Kalach, N., Mihalache, A., Gosset, P., Michalski, J. C., Léonard, R., & Robbe-Masselot, C. (2017). How do they stick together? Bacterial adhesins implicated in the binding of bacteria to the human gastrointestinal mucins. In *Biochemical Society Transactions*. https://doi.org/10.1042/BST20160167

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*(1), 979. https://doi.org/10.3389/fmicb.2016.00979

Robbe, C., Capon, C., Maes, E., Rousset, M., Zweibaum, A., Zanetta, J.-P., & Michalski, J.-C. (2003). Evidence of regio-specific glycosylation in human intestinal mucins. *Journal of Biological Chemistry*, *278*(47), 46337–46348. https://doi.org/10.1074/jbc.M302529200

Roberts, G., Tarelli, E., Homer, K. A., Philpott-Howard, J., & Beighton, D. (2000). Production of an endo-β-*N*-acetylglucosaminidase activity mediates growth of *Enterococcus faecalis* on a high-mannose-type glycoprotein. *Journal of Bacteriology*, *182*(4), 882–890. https://doi.org/10.1128/JB.182.4.882-890.2000

Rojas, E., & Torres, G. (2013). Isolation and recovery of glycomacropeptide from milk whey by means of thermal treatment. *Food Science and Technology*, *33*(1), 14–20. https://doi.org/10.1590/s0101-20612013005000027

Ross, S. A., Lane, J. A., Kilcoyne, M., Joshi, L., & Hickey, R. M. (2015). The milk fat globule membrane. In V. K. Gupta, M. G. Tuohy, A. O'Donovan, & M. Lohani (Eds.), *Biotechnology of Bioactive Compounds* (pp. 631–668). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118733103.ch26

Rossez, Y., Maes, E., Lefebvre Darroman, T., Gosset, P., Ecobichon, C., Joncquel Chevalier Curt, M., Boneca, I. G., Michalski, J. C., & Robbe-Masselot, C. (2012). Almost all human gastric mucin *O*-glycans harbor blood group A, B or H antigens and are potential binding sites for *Helicobacter pylori*. *Glycobiology*, *22*(9), 1193–1206. https://doi.org/10.1093/glycob/cws072

Ruan, W., Engevik, M. A., Spinler, J. K., & Versalovic, J. (2020). Healthy human gastrointestinal microbiome: composition and function after a decade of exploration. *Digestive Diseases and Sciences*, *65*(3), 695–705. https://doi.org/10.1007/s10620-020-06118-4

Saa, P., Urrutia, A., Silva-Andrade, C., Martín, A. J., & Garrido, D. (2022). Modeling approaches for probing cross-feeding interactions in the human gut microbiome. *Computational and Structural Biotechnology Journal*, *20*, 79–89. https://doi.org/10.1016/j.csbj.2021.12.006

Saito, T., & Itoh, T. (1992). Variations and distributions of *O*-glycosidically linked sugar

chains in bovine κ-casein. *Journal of Dairy Science*, *75*(7), 1768–1774. https://doi.org/10.3168/jds.S0022-0302(92)77936-3

Sakanaka, M., Gotoh, A., Yoshida, K., Odamaki, T., Koguchi, H., Xiao, J., Kitaoka, M., & Katayama, T. (2020). 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*(1), 71. https://doi.org/10.3390/nu12010071

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. zhong, Kitaoka, M., Fukiya, S., Yokota, A., Leggio, L. Lo, Hachem, M. A., & Katayama, T. (2019). Evolutionary adaptation in fucosyllactose uptake systems supports bifidobacteria-infant symbiosis. *Science Advances*, *5*(8), 1–16. https://doi.org/10.1126/sciadv.aaw7696

Sanders, D. J., Inniss, S., Sebepos-Rogers, G., Rahman, F. Z., & Smith, A. M. (2021). The role of the microbiome in gastrointestinal inflammation. *Bioscience Reports*, *41*(6). https://doi.org/10.1042/BSR20203850

Sawicki, C., Livingston, K., Obin, M., Roberts, S., Chung, M., & McKeown, N. (2017). Dietary fiber and the human gut microbiota: Application of evidence mapping methodology. *Nutrients*, *9*(2), 125. https://doi.org/10.3390/nu9020125

Schmid, A. K. (2018). Conserved principles of transcriptional networks controlling metabolic flexibility in archaea. *Emerging Topics in Life Sciences*, *2*(4), 659–669. https://doi.org/10.1042/ETLS20180036

Schroeder, B. O., Birchenough, G. M. H., Ståhlman, M., Arike, L., Johansson, M. E. V., Hansson, G. C., & Bäckhed, F. (2018). Bifidobacteria or fiber protects against diet-induced microbiota-mediated colonic mucus deterioration. *Cell Host and Microbe*. https://doi.org/10.1016/j.chom.2017.11.004

Schulz, B. L., Sloane, A. J., Robinson, L. J., Prasad, S. S., Lindner, R. A., Robinson, M., Bye, P. T., Nielson, D. W., Harry, J. L., Packer, N. H., & Karlsson, N. G. (2007). Glycosylation of sputum mucins is altered in cystic fibrosis patients. *Glycobiology*, *17*(7), 698–712. https://doi.org/10.1093/glycob/cwm036

Scott, K. P., Gratz, S. W., Sheridan, P. O., Flint, H. J., & Duncan, S. (2013). The influence of diet on the gut microbiota. *Pharmacological Research*, *69*(1), 52–60. https://doi.org/10.1016/j.phrs.2012.10.020

Sekirov, I., Russell, S. L., Antunes, L. C. M., & Finlay, B. B. (2010). Gut microbiota in health and disease. *Physiological Reviews*, *90*(3), 859–904. https://doi.org/10.1152/physrev.00045.2009

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*(48), 18964–18969.

https://doi.org/10.1073/pnas.0809584105

Sela, D., Li, Y., Lerno, L., Wu, S., Marcobal, A. A., German, J. B., Chen, X., Lebrilla, C., & Mills, D. A. (2011). An infant-associated bacterial commensal utilizes breast milk sialyloligosaccharides. *Journal of Biological Chemistry*, *286*(14), 11909–11918. https://doi.org/10.1074/jbc.M110.193359

Seregin, S. S., Golovchenko, N., Schaf, B., Chen, J., Pudlo, N. A., Mitchell, J., Baxter, N. T., Zhao, L., Schloss, P. D., Martens, E. C., Eaton, K. A., & Chen, G. Y. (2017). NLRP6 protects Il10−/− mice from colitis by limiting colonization of *Akkermansia muciniphila*. *Cell Reports*. https://doi.org/10.1016/j.celrep.2017.03.080

Shaw, A. G., Cornwell, E., Sim, K., Thrower, H., Scott, H., Brown, J. C. S., Dixon, R. A., & Kroll, J. S. (2020). Dynamics of toxigenic *Clostridium perfringens* colonisation in a cohort of prematurely born neonatal infants. *BMC Pediatrics*, *20*(1), 75. https://doi.org/10.1186/s12887-020-1976-7

Sheridan, P., Martin, J. C., Lawley, T. D., Browne, H. P., Harris, H. M. B., Bernalier-Donadille, A., Duncan, S. H., O'Toole, P. W., P. Scott, K., & J. Flint, H. (2016). Polysaccharide utilization loci and nutritional specialization in a dominant group of butyrate-producing human colonic Firmicutes. *Microbial Genomics*, *2*(2). https://doi.org/10.1099/mgen.0.000043

Shimada, Y., Watanabe, Y., Wakinaka, T., Funeno, Y., Kubota, M., Chaiwangsri, T., Kurihara, S., Yamamoto, K., Katayama, T., & Ashida, H. (2015). α-*N*-acetylglucosaminidase from *Bifidobacterium bifidum* specifically hydrolyzes α-linked *N*-acetylglucosamine at nonreducing terminus of *O*-glycan on gastric mucin. *Applied Microbiology and Biotechnology*, *99*(9), 3941–3948. https://doi.org/10.1007/s00253-014-6201-x

Shin, J., Noh, J. R., Chang, D. H., Kim, Y. H., Kim, M. H., Lee, E. S., Cho, S., Ku, B. J., Rhee, M. S., Kim, B. C., Lee, C. H., & Cho, B. K. (2019). Elucidation of *Akkermansia muciniphila* probiotic traits driven by mucin depletion. *Frontiers in Microbiology*, *10*. https://doi.org/10.3389/fmicb.2019.01137

Shokryazdan, P., Faseleh Jahromi, M., Navidshad, B., & Liang, J. B. (2017). Effects of prebiotics on immune system and cytokine expression. *Medical Microbiology and Immunology*, *206*(1), 1–9. https://doi.org/10.1007/s00430-016-0481-y

Sicard, J.-F., Le Bihan, G., Vogeleer, P., Jacques, M., & Harel, J. (2017). Interactions of intestinal bacteria with components of the intestinal mucus. *Frontiers in Cellular and Infection Microbiology*, *7*(1), 1–12. https://doi.org/10.3389/fcimb.2017.00387

Stanley, P., Taniguchi, N., & Aebi, M. (2017). *N*-glycans. In A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, & P. H. Seeberger (Eds.), *Essentials of Glycobiology* (3rd. Editi). Cold Spring Harbor Laboratory Press.

Stuart, J. M., Segal, E., Koller, D., & Kim, S. K. (2003). A gene-coexpression network for global discovery of conserved genetic modules. *Science*, *302*(5643), 249–255. https://doi.org/10.1126/science.1087447

Sun, Q., Zhao, H., Zhang, C., Hu, T., Wu, J., Lin, X., Luo, D., Wang, C., Meng, L., Xi, L., Li, K., Hu, J., Ma, D., & Zhu, T. (2017). Gene co-expression network reveals shared modules predictive of stage and grade in serous ovarian cancers. *Oncotarget*, *8*(26), 42983–42996. https://doi.org/10.18632/oncotarget.17785

Sunds, A. V., Poulsen, N. A., & Larsen, L. B. (2019). Short communication: application of proteomics for characterization of caseinomacropeptide isoforms before and after desialidation. *Journal of Dairy Science*, *102*(10), 8696–8703. https://doi.org/10.3168/jds.2019-16617

Tailford, L. E., Crost, E. H., Kavanaugh, D., & Juge, N. (2015). Mucin glycan foraging in the human gut microbiome. *Frontiers in Genetics*, *6*. https://doi.org/10.3389/fgene.2015.00081

Takihara, H., & Okuda, S. (2023). Glycan-related genes in human gut microbiota exhibit differential distribution and diversity in carbohydrate degradation and glycan synthesis. *Frontiers in Molecular Biosciences*, *10*. https://doi.org/10.3389/fmolb.2023.1137303

Thomä-Worringer, C., Sørensen, J., & López-Fandiño, R. (2006). Health effects and technological features of caseinomacropeptide. *International Dairy Journal*, *16*(11), 1324–1333. https://doi.org/10.1016/j.idairyj.2006.06.012

Thompson, P., Medina, D. A., & Garrido, D. (2018). Human milk oligosaccharides and infant gut bifidobacteria: Molecular strategies for their utilization. *Food Microbiology*, *75*, 37–46. https://doi.org/10.1016/j.fm.2017.09.001

Thomson, P., Medina, D. A., & Garrido, D. (2018a). Human milk oligosaccharides and infant gut bifidobacteria: molecular strategies for their utilization. *Food Microbiology*, *75*, 37–46. https://doi.org/10.1016/j.fm.2017.09.001

Thomson, P., Medina, D. A., & Garrido, D. (2018b). Human milk oligosaccharides and infant gut bifidobacteria: Molecular strategies for their utilization. *Food Microbiology*, *75*, 37–46. https://doi.org/10.1016/j.fm.2017.09.001

Thomson, P., Santibañez, R., Aguirre, C., Galgani, J. E., & Garrido, D. (2019). Short-term impact of sucralose consumption on the metabolic response and gut microbiome of healthy adults. *British Journal of Nutrition*, *122*(8), 856–862. https://doi.org/10.1017/S0007114519001570

Thurl, S., Munzert, M., Boehm, G., Matthews, C., & Stahl, B. (2017). Systematic review of the concentrations of oligosaccharides in human milk. *Nutrition Reviews*, *75*(11), 920–933. https://doi.org/10.1093/nutrit/nux044

Tsevelkhoroloo, M., Shim, S. H., Lee, C.-R., Hong, S.-K., & Hong, Y.-S. (2021). LacI-family transcriptional regulator DagR acts as a repressor of the agarolytic pathway genes in *Streptomyces coelicolor* A3(2). *Frontiers in Microbiology*, *12: 658657*. https://doi.org/10.3389/fmicb.2021.658657

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*(9), 2574–2590. https://doi.org/10.1038/s41396-021-00937-7

Turroni, F., Bottacini, F., Foroni, E., Mulder, I., Jae-Han, K., Zomer, A., Sánchez, B., Bidossi, A., Ferrarini, A., Giubellini, V., Delledonne, M., Henrissat, B., Coutinho, P., Oggioni, M., Fitzgerald, G. F., Mills, D., Margolles, A., Denise, K., Van Douwe, S., & Ventura, M. (2010). Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-derived glycan foraging. *Proceedings of the National Academy of Sciences*, *107*(45), 19514–19519.

Turroni, F., Milani, C., Duranti, S., Mahony, J., van Sinderen, D., & Ventura, M. (2018). Glycan utilization and cross-feeding activities by Bifidobacteria. *Trends in Microbiology*, *26*(4), 339–350. https://doi.org/10.1016/j.tim.2017.10.001

Turroni, F., Milani, C., Ventura, M., & van Sinderen, D. (2022). The human gut microbiota during the initial stages of life: insights from bifidobacteria. *Current Opinion in Biotechnology*, *73*, 81–87. https://doi.org/10.1016/j.copbio.2021.07.012

van Dam, S., Võsa, U., van der Graaf, A., Franke, L., & de Magalhães, J. P. (2017). Gene co-expression analysis for functional classification and gene–disease predictions. *Briefings in Bioinformatics*, *19*(4), 575–592. https://doi.org/10.1093/bib/bbw139

van der Ark, K. C. H., Aalvink, S., Suarez-Diez, M., Schaap, P. J., de Vos, W. M., & Belzer, C. (2018). Model-driven design of a minimal medium for *Akkermansia muciniphila* confirms mucus adaptation. *Microbial Biotechnology*, *11*(3), 476–485. https://doi.org/10.1111/1751-7915.13033

van Passel, M. W. J., Kant, R., Zoetendal, E. G., Plugge, C. M., Derrien, M., Malfatti, S. A., Chain, P. S. G., Woyke, T., Palva, A., de Vos, W. M., & Smidt, H. (2011). The genome of *Akkermansia muciniphila*, a dedicated intestinal mucin degrader, and its use in exploring intestinal metagenomes. *PLoS ONE*, *6*(3), e16876. https://doi.org/10.1371/journal.pone.0016876

van Tassell, M. L., & Miller, M. J. (2011). *Lactobacillus* adhesion to mucus. In *Nutrients*. https://doi.org/10.3390/nu3050613

Vandeputte, D., Kathagen, G., D'hoe, K., Vieira-Silva, S., Valles-Colomer, M., Sabino, J., Wang, J., Tito, R. Y., De Commer, L., Darzi, Y., Vermeire, S., Falony, G., & Raes, J. (2017). Quantitative microbiome profiling links gut community variation to microbial load. *Nature*, *551*(7681), 507–511. https://doi.org/10.1038/nature24460

Varki, A. (2017). Biological roles of glycans. *Glycobiology*, *27*(1), 3–49. https://doi.org/10.1093/glycob/cww086

Varki, A., & Lowe, J. B. (2009). Biological roles of glycans. In A. Vark, R. Cummings, & J. Esko (Eds.), *Essentials of Glycobiology.* (2nd Ed.). Cold Spring Harbor Laboratory Press. https://www.ncbi.nlm.nih.gov/books/NBK1897/

Vega-Sagardía, M., Cabezón, E. C., Delgado, J., Ruiz-Moyano, S., & Garrido, D. (2023). Screening microbial interactions during inulin utilization reveals strong competition and proteomic changes in Lacticaseibacillus paracasei M38. *Probiotics and Antimicrobial Proteins*. https://doi.org/10.1007/s12602-023-10083-5

Vega-Sagardía, M., Delgado, J., Ruiz-Moyano, S., & Garrido, D. (2023). Proteomic

analyses of Bacteroides ovatus and Bifidobacterium longum in xylan bidirectional culture shows sugar cross-feeding interactions. *Food Research International*, *170*, 113025. https://doi.org/10.1016/j.foodres.2023.113025

Vernocchi, P., Del Chierico, F., & Putignani, L. (2020). Gut microbiota metabolism and interaction with food components. *International Journal of Molecular Sciences*, *21*(10), 3688. https://doi.org/10.3390/ijms21103688

Vliegenthart, J. F. (2017). Review The complexity of glycoprotein-derived glycans. *Proceedings of the Japan Academy*, *93*(2), 64–86.

Walsh, C., Lane, J. A., van Sinderen, D., & Hickey, R. M. (2020). Human milk oligosaccharides: shaping the infant gut microbiota and supporting health. *Journal of Functional Foods*, *72*, 104074. https://doi.org/10.1016/j.jff.2020.104074

Wandall, H. H., Nielsen, M. A. I., King-Smith, S., de Haan, N., & Bagdonaite, I. (2021). Global functions of *O*-glycosylation: promises and challenges in *O*-glycobiology. *The FEBS Journal*, *288*(24), 7183–7212. https://doi.org/10.1111/febs.16148

Wang, M., Zhang, X. Y., Guo, R. R., Cai, Z. P., Hu, X. C., Chen, H., Wei, S., Voglmeir, J., & Liu, L. (2018). Cloning, purification and biochemical characterization of two β-*N*-acetylhexosaminidases from the mucin-degrading gut bacterium *Akkermansia muciniphila*. *Carbohydrate Research*, *457*, 1–7. https://doi.org/10.1016/j.carres.2017.12.007

Wang, Y. (2020). Sialidases from *Clostridium perfringens* and their inhibitors. *Frontiers in Cellular and Infection Microbiology*, *9*. https://doi.org/10.3389/fcimb.2019.00462

Wardman, J. F., Bains, R. K., Rahfeld, P., & Withers, S. G. (2022). Carbohydrate-active enzymes (CAZymes) in the gut microbiome. *Nature Reviews Microbiology*, *20*(9), 542–556. https://doi.org/10.1038/s41579-022-00712-1

Wexler, A. G., & Goodman, A. L. (2017). An insider's perspective: Bacteroides as a window into the microbiome. *Nature Microbiology*, *2*(5), 17026. https://doi.org/10.1038/nmicrobiol.2017.26

Wilson, B., & Whelan, K. (2017). Prebiotic inulin-type fructans and galacto-oligosaccharides: definition, specificity, function, and application in gastrointestinal disorders. *Journal of Gastroenterology and Hepatology*, *32*, 64–68. https://doi.org/10.1111/jgh.13700

Wrzosek, L., Miquel, S., Noordine, M. L., Bouet, S., Chevalier-Curt, M. J., Robert, V., Philippe, C., Bridonneau, C., Cherbuy, C., Robbe-Masselot, C., Langella, P., & Thomas, M. (2013). *Bacteroides thetaiotaomicron* and *Faecalibacterium prausnitzii* influence the production of mucus glycans and the development of goblet cells in the colonic epithelium of a gnotobiotic model rodent. *BMC Biology*, *11*(1), 1–13. https://doi.org/10.1186/1741-7007-11-61

Yamada, T., Hino, S., Iijima, H., Genda, T., Aoki, R., Nagata, R., Han, K.-H., Hirota, M., Kinashi, Y., Oguchi, H., Suda, W., Furusawa, Y., Fujimura, Y., Kunisawa, J., Hattori, M., Fukushima, M., Morita, T., & Hase, K. (2019). Mucin *O*-glycans facilitate symbiosynthesis to maintain gut immune homeostasis. *EBioMedicine*, *48*, 513–525.

https://doi.org/10.1016/j.ebiom.2019.09.008

Yamamoto, T., Ugai, H., Nakayama-Imaohji, H., Tada, A., Elahi, M., Houchi, H., & Kuwahara, T. (2018). Characterization of a recombinant *Bacteroides fragilis* sialidase expressed in *Escherichia coli*. *Anaerobe*. https://doi.org/10.1016/j.anaerobe.2018.02.003

Yassour, M., Jason, E., Hogstrom, L. J., Arthur, T. D., Tripathi, S., Siljander, H., Selvenius, J., Oikarinen, S., Hyöty, H., Virtanen, S. M., Ilonen, J., Ferretti, P., Pasolli, E., Tett, A., Asnicar, F., Segata, N., Vlamakis, H., Lander, E. S., Huttenhower, C., … Xavier, R. J. (2018). Strain-level analysis of mother-to-child bacterial transmission during the first few months of life. *Cell Host and Microbe*, *24*(1), 146-154.e4. https://doi.org/10.1016/j.chom.2018.06.007

Yokoi, T., Nishiyama, K., Kushida, Y., Uribayashi, K., Kunihara, T., Fujimoto, R., Yamamoto, Y., Ito, M., Miki, T., Haneda, T., Mukai, T., & Okada, N. (2022). *O*-acetylesterase activity of *Bifidobacterium bifidum* sialidase facilities the liberation of sialic acid and encourages the proliferation of sialic acid scavenging *Bifidobacterium breve*. *Environmental Microbiology Reports*, *14*(4), 637–645. https://doi.org/10.1111/1758-2229.13083

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. *Scientific Reports*, *10*(1), 15919. https://doi.org/10.1038/s41598-020-72792-z

Zhai, L., Huang, C., Ning, Z., Zhang, Y., Zhuang, M., Yang, W., Wang, X., Wang, J., Zhang, L., Xiao, H., Zhao, L., Asthana, P., Lam, Y. Y., Chow, C. F. W., Huang, J., Yuan, S., Chan, K. M., Yuan, C.-S., Lau, J. Y.-N., … Bian, Z. (2023). *Ruminococcus gnavus* plays a pathogenic role in diarrhea-predominant irritable bowel syndrome by increasing serotonin biosynthesis. *Cell Host & Microbe*, *31*(1), 33-44.e5. https://doi.org/10.1016/j.chom.2022.11.006

Zhang, B., & Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology*, *4*(1). https://doi.org/10.2202/1544-6115.1128

Zhang, B., Li, L.-Q., Liu, F., & Wu, J.-Y. (2022). Human milk oligosaccharides and infant gut microbiota: Molecular structures, utilization strategies and immune function. *Carbohydrate Polymers*, *276*, 118738. https://doi.org/10.1016/j.carbpol.2021.118738

Zhang, L., Li, Y., Li, R., Yang, X., Zheng, Z., Fu, J., Yu, H., & Chen, X. (2023). Glycoprotein *in vitro N*-Glycan processing using enzymes expressed in *E. coli*. *Molecules*, *28*(6), 2753. https://doi.org/10.3390/molecules28062753

Zhang, S., Li, T., Xie, J., Zhang, D., Pi, C., Zhou, L., & Yang, W. (2021). Gold standard for nutrition: a review of human milk oligosaccharide and its effects on infant gut microbiota. *Microbial Cell Factories*, *20*(1), 108. https://doi.org/10.1186/s12934-021-01599-y

Zhang, Y., Wang, L., Ocansey, D. K. W., Wang, B., Wang, L., & Xu, Z. (2021). Mucin-type O-Glycans: Barrier, microbiota, and immune anchors in inflammatory bowel disease. *Journal of Inflammation Research*, *14*, 5939–5953. https://doi.org/10.2147/JIR.S327609

Zheng, X., Su, Y., Chen, Y., Huang, H., & Shen, Q. (2018). Global transcriptional responses of denitrifying bacteria to functionalized single-walled carbon nanotubes revealed by weighted gene-coexpression network analysis. *Science of The Total Environment*, *613–614*, 1240–1249. https://doi.org/10.1016/j.scitotenv.2017.09.193

Zhu, B., Wang, X., & Li, L. (2010). Human gut microbiome: the second genome of human body. *Protein & Cell*, *1*(8), 718–725. https://doi.org/10.1007/s13238-010-0093-z

Zúñiga, M., Monedero, V., & Yebra, M. J. (2018a). Utilization of host-derived glycans by intestinal *Lactobacillus* and *Bifidobacterium* species. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.01917

Zúñiga, M., Monedero, V., & Yebra, M. J. (2018b). Utilization of Host-Derived Glycans by Intestinal Lactobacillus and Bifidobacterium Species. *Frontiers in Microbiology*, *9*(1), 1–23. https://doi.org/10.3389/fmicb.2018.01917

6. SUPPLEMENTARY MATERIAL

6.1 SUPPLEMENTARY FIGURES

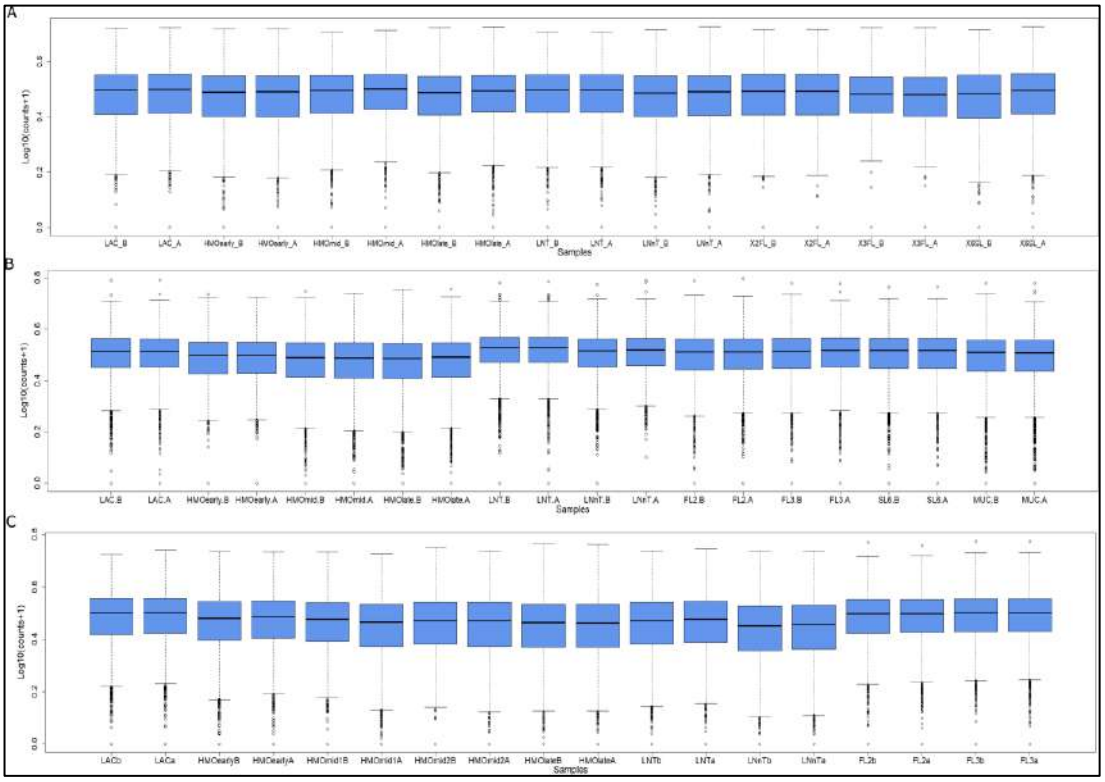

Figure S1- Normalized data of samples. A) *Bifidobacterium longum* subsp. *infantis* ATCC 15697, B) *Bifidobacterium bifidum* SC555, C) *Bifidobacterium longum* subsp *longum* SC596.

Normalized samples were taken from NCBI Geo Datasets. For WGCNA analysis, $\log_{10}$ normalize read counts were used for this purpose.

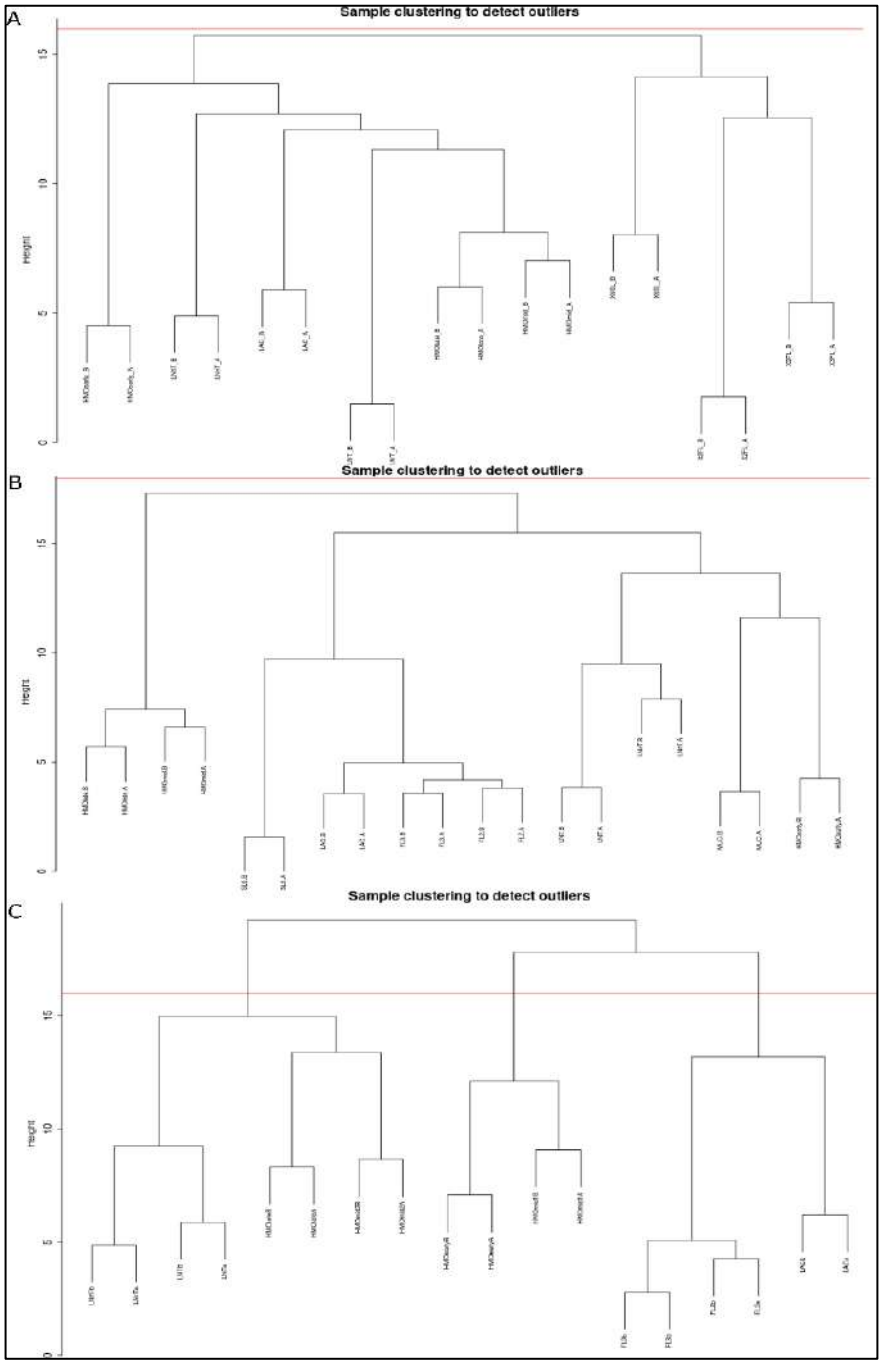

Figure S2- Sample clustering to detect outliers. A) *B. longum* subsp. *infanti*s ATCC 15697, B) *B. bifidum* SC 555, c) *B. longum* subsp. *longum* SC596.

RNA-seq datasets (GSE58773, GSE87697, GSE59053) were evaluated by sample clustering according to the Euclidean distance between different samples observed for each bacterium (Figure 1). No outliers were detected in the clusters; therefore, 56 samples were used to construct a hierarchical clustering tree (dendogram).

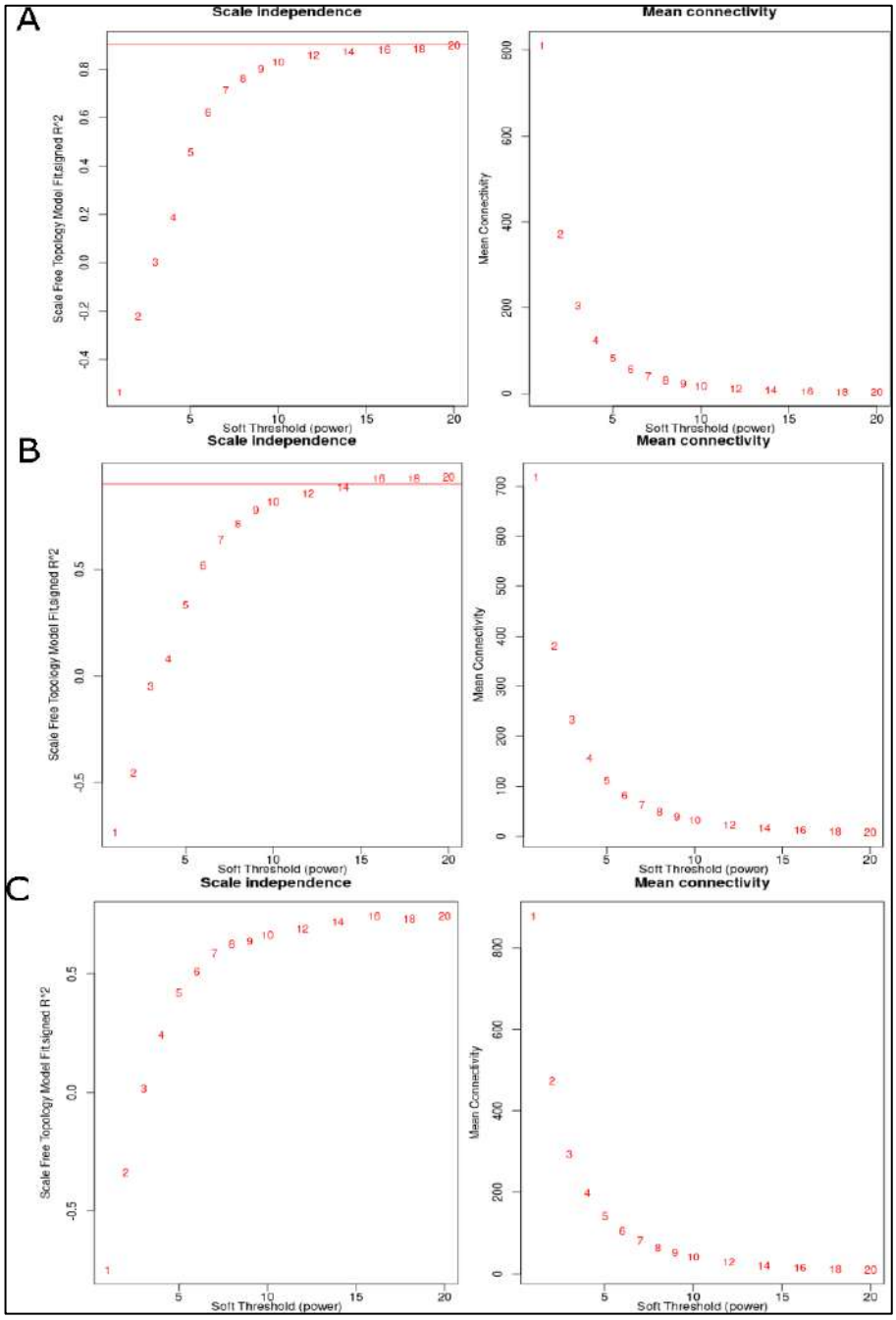

Figure S3- Network topology for different soft-thresholding powers on each Bifidobacterium member. A) *Bifidobacterium longum* subsp. *infantis* ATCC 15697, B) *Bifidobacterium bifidum* SC 555, C) *Bifidobacterium longum* subsp. *longum* SC596.

Numbers in the plots indicate the corresponding soft thresholding powers. The approximate scale-free topology can be attained at the soft-thresholding power of 12, 12, 16, respectively.

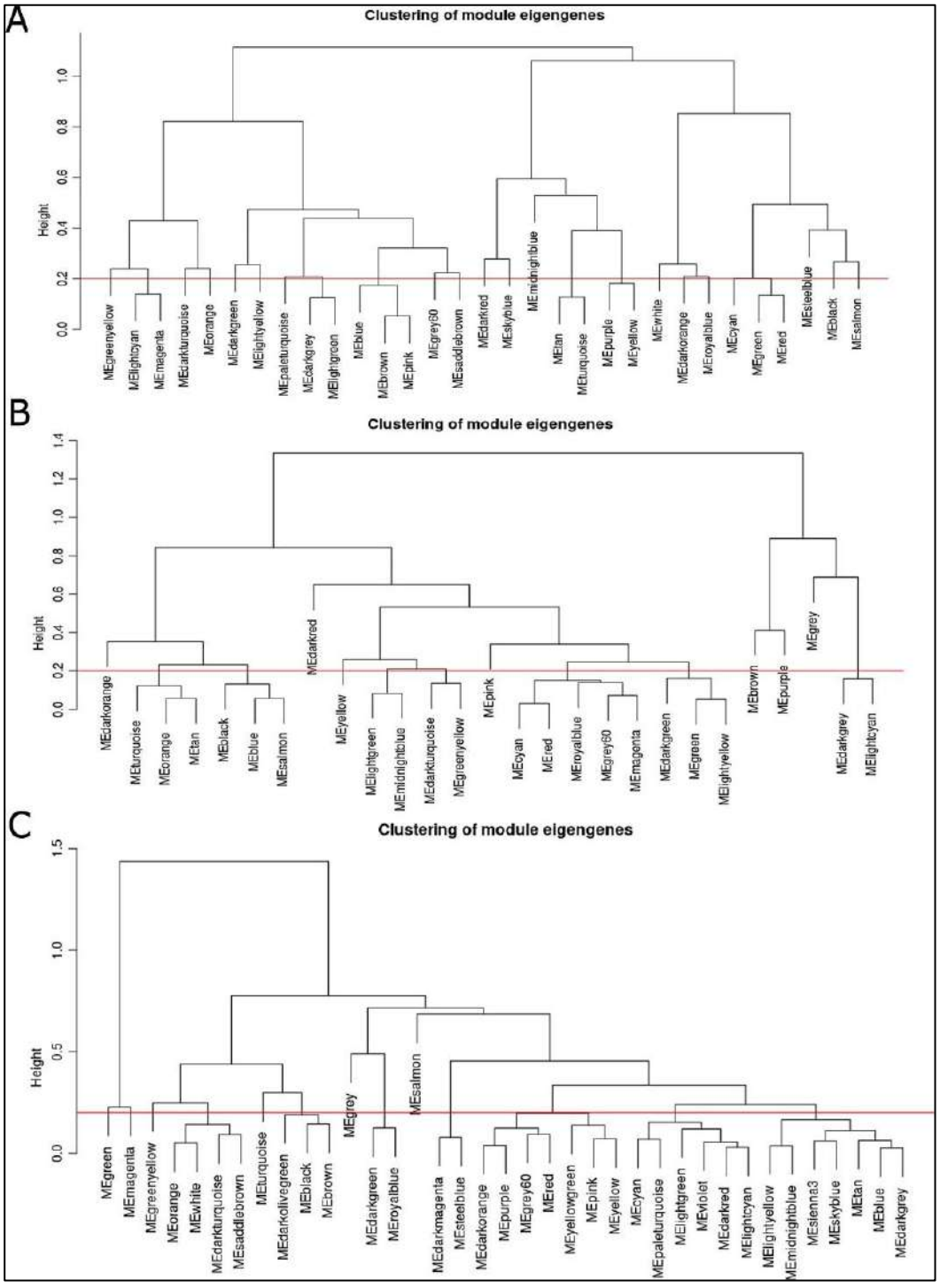

Figure S4- Dendrogram of consensus module eigengenes obtained by WGCNA on the consensus correlation.

The red line is the merging threshold, and groups of eigengenes below the threshold represent modules whose expressions profiles should be merged due to their similarity.

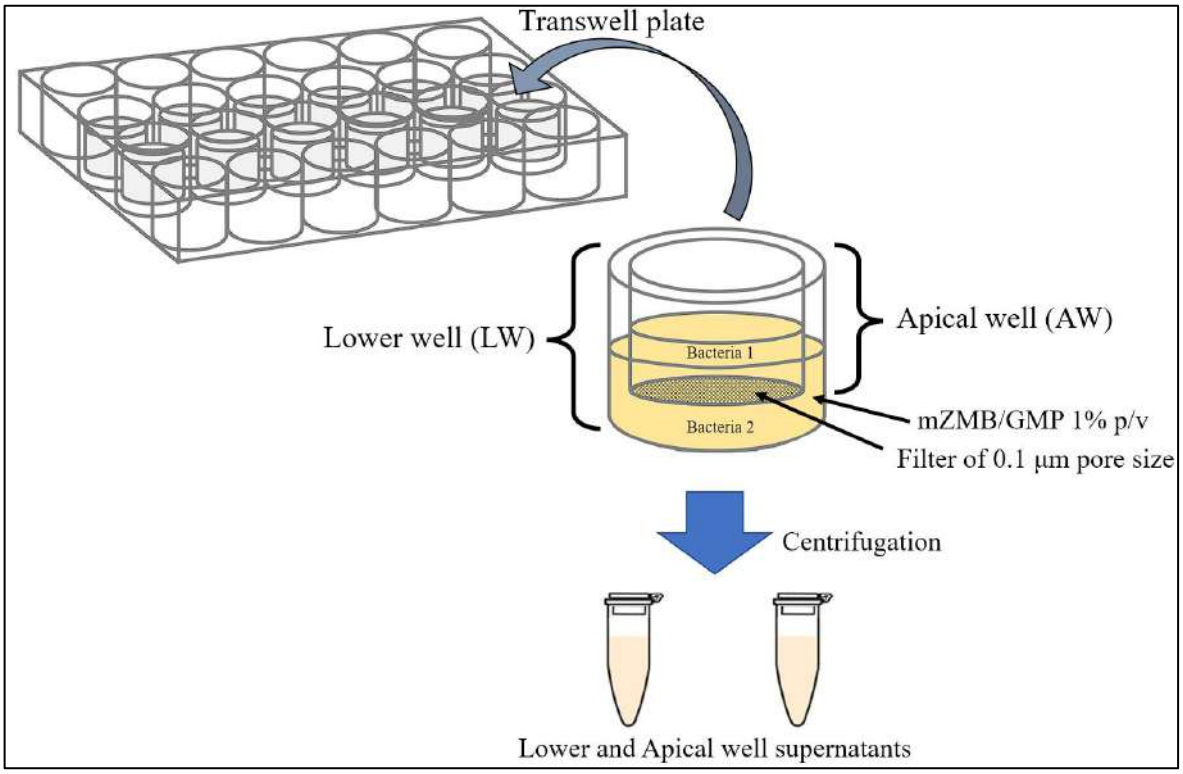

Figure S5- The schematization of bidirectional culture.

Bidirectional assays were made using Transwell plates locating GMP-degrading bacteria in the lower well and non-degrading bacteria in the apical well. The cell density of the cultures used for apical and lower wells was equalized to 1.0 OD_{620nm} before the experiment. Bacterial combinations for inoculation were made per duplicate.

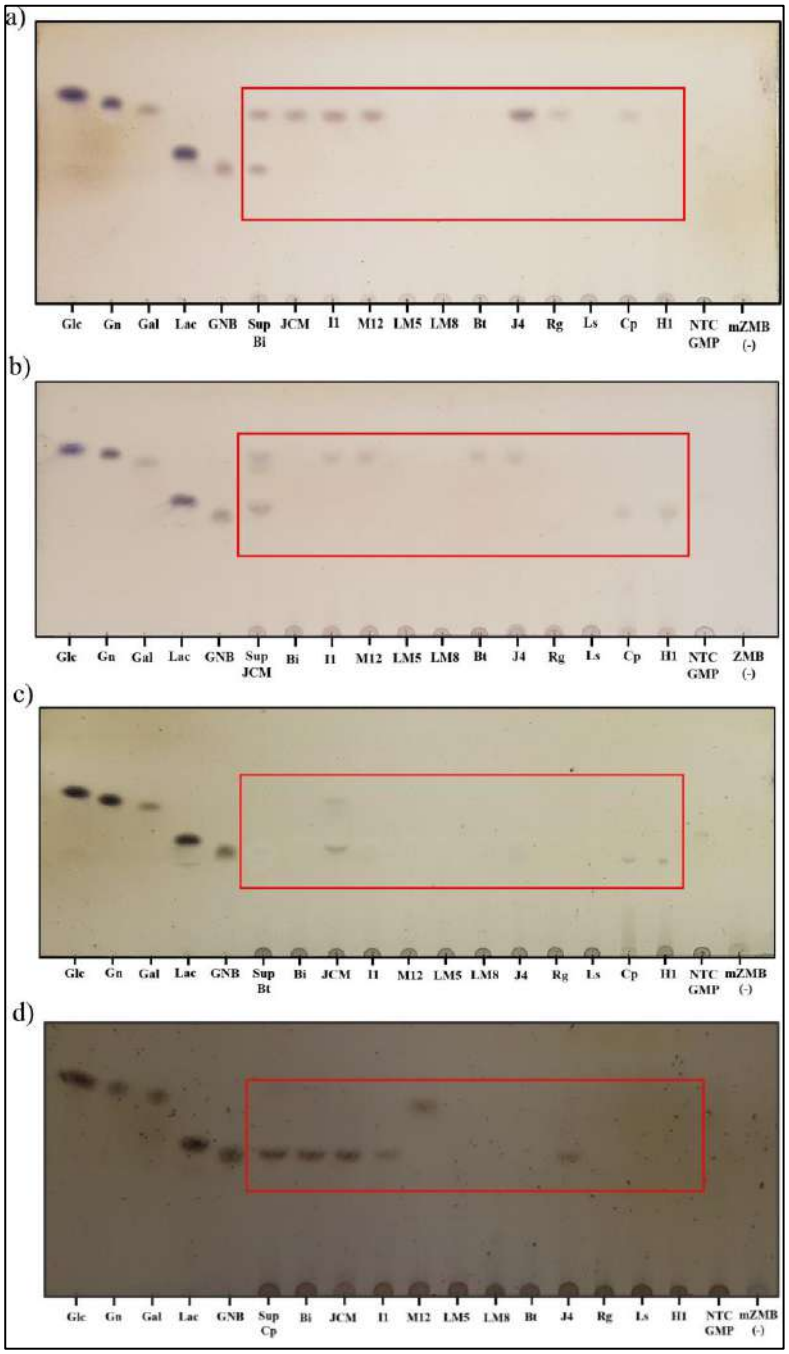

Figure S6- Thin Layer Chromatography of Unidirectional assays.

Supernatants were recovered after centrifugation. Standard sugar such as glucose (Glc), *N*-acetylgalactosamine (Gn), galactose (Gal), lactose (Lac), and galacto-*N*-biose (GNB) were used at 1mg/mL. *Binf*: *B. infantis* ATCC 15697; JCM: *B. bifidum* JCM 1254; I1: *B. breve* I1; M12: *B. longum* subsp. *longum* M12; LM5: *S. thermophilus* LM5; LM8: *S. thermophilus* LM8; Bt: *B. thetaiotaomicron* VPI 5482; Rg: *R. gnavus* CC55_001C; Ls: *L. symbiosum*; Cp: C. *perfringens* ATCC 13124; H1: *E. faecalis*. Sup: supernatant.

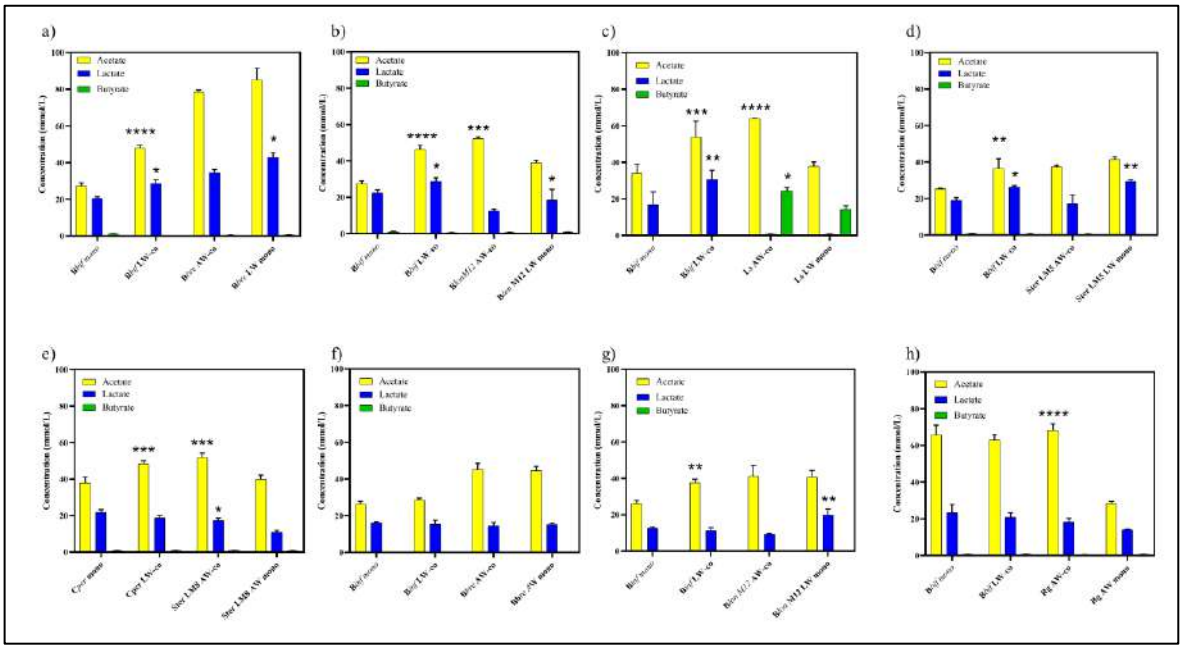

Figure S7- Determination of SCFA in bidirectional culture using GMP.

Supernatants were recovered by centrifugation and analyzed after 48 h of co-culture growth to quantify lactate, acetate, butyrate, succinate, succinate, and propionate concentrations.

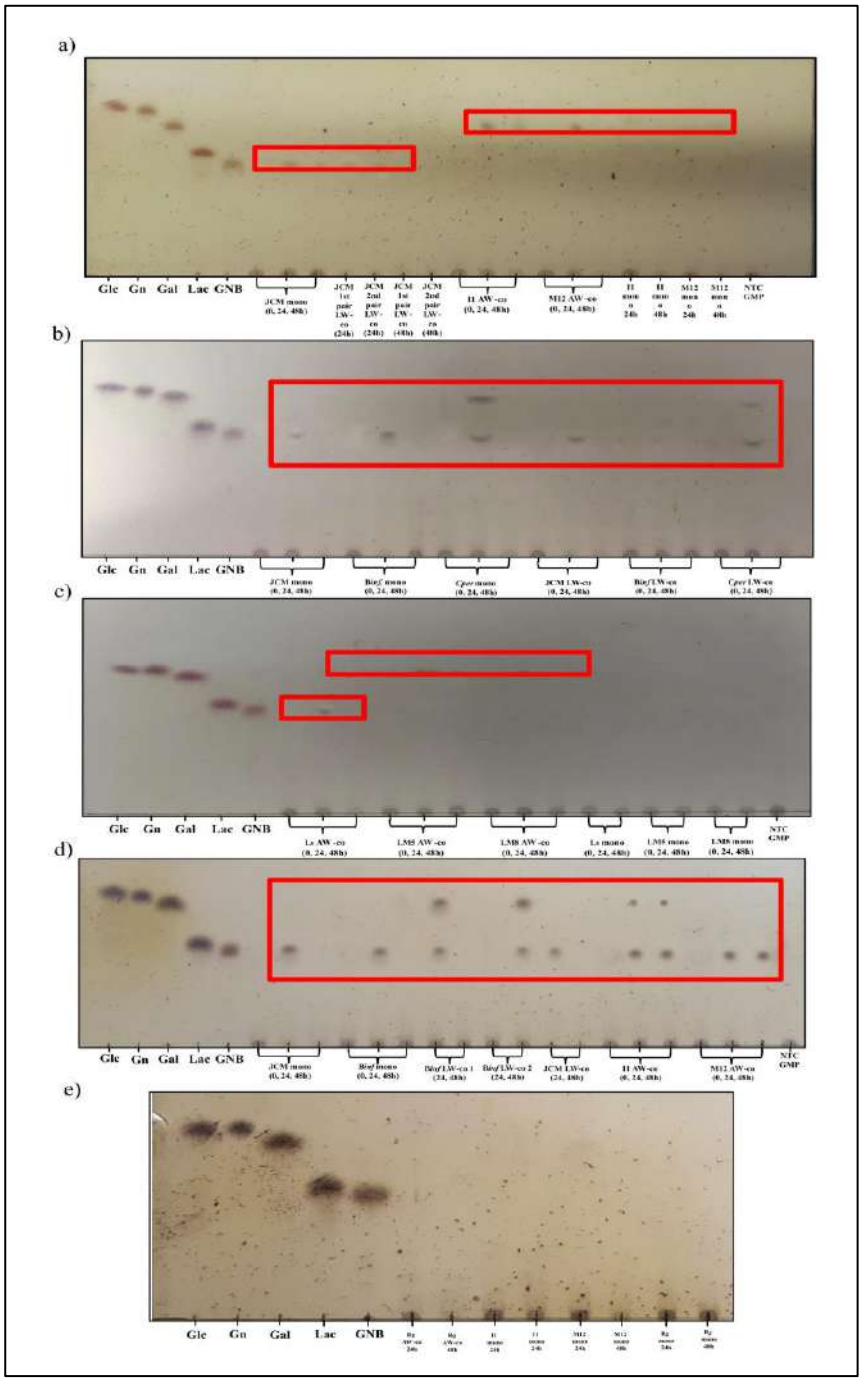

Figure S8- Thin Layer Chromatography of bidirectional assays.

Supernatants of individual and co-culture growths were recovered after centrifugation at 4000 x *g* for 2 min. Standard sugar such as glucose (Glc), *N*-acetylgalactosamine (Gn), galactose (Gal), lactose (Lac), and galacto-*N*-biose (GNB) were used at 1mg/mL. *Binf*: *B. infantis* ATCC 15697; JCM: *B. bifidum* JCM 1254; I1: *B. breve* I1; M12: *B. longum* subsp. *longum* M12; LM5: *S. thermophilus* LM5; LM8: *S. thermophilus* LM8; Rg: *R. gnavus* CC55_001C; Ls: *L. symbiosum* WAL-14673; Cp: C. *perfringens* ATCC 13124.

6.2 SUPPLEMENTARY TABLES

Table S1. RNA-seq samples from different *Bifidobacterium* strains growth in various substrates.

***Bifidobacterium longum* subsp. *infantis* ATCC 15697**	***Bifidobacterium bifidum* SC555**	**B. *longum* subsp. *longum* SC596**
GSM1419252, lactose treatment, replicate A	GSM1425041, lactose treatment, replicate A	GSM2338941, lactose treatment, replicate A
GSM1419253, lactose treatment, replicate B	GSM142504, lactose treatment, replicate B	GSM2338942, lactose treatment, replicate B
GSM1419254, pooled HMO, early time point, replicate A	GSM142504, pooled HMO, early time point, replicate A	GSM2338943, pooled HMO, early time point, replicate A
GSM1419255, pooled HMO, early time point, replicate B	GSM1425044, pooled HMO, early time point, replicate B	GSM2338944, pooled HMO, early time point, replicate B
GSM1419256, pooled HMO, mid-time point, replicate A	GSM1425045, pooled HMO, mid-time point, replicate A	GSM2338945, pooled HMO, mid1 time point, replicate A
GSM1419257, pooled HMO, mid-time point, replicate B	GSM1425046, pooled HMO, mid-time point, replicate B	GSM2338946, pooled HMO, mid1 time point, replicate B
GSM1419258, pooled HMO, late time point, replicate A	GSM1425047, pooled HMO, late time point, replicate A	GSM2338947, pooled HMO, mid2 time point, replicate A
GSM1419259, pooled HMO, late time point, replicate B	GSM1425048, pooled HMO, late time point, replicate B	GSM2338948, pooled HMO, mid2 time point, replicate B
GSM1419260, lacto-*N*-tetraose treatment, replicate A	GSM1425049, lacto-*N*-tetraose treatment, replicate A	GSM2338949, pooled HMO, late time point, replicate A
GSM1419261, lacto-*N*-tetraose treatment, replicate B	GSM1425050, lacto-*N*-tetraose treatment, replicate B	GSM2338950, pooled HMO, late time point, replicate B
GSM1419262, lacto-*N*-neotetraose treatment, replicate A	GSM1425051, lacto-*N*-neotetraose treatment, replicate A	GSM2338951, lacto-N-tetraose treatment, replicate A
GSM1419263, lacto-*N*-neotetraose treatment, replicate B	GSM1425052, lacto-*N*-neotetraose treatment, replicate B	GSM2338952, lacto-*N*-tetraose treatment, replicate B
GSM1419264, 2'fucosyllactose treatment, replicate A	GSM1425053, 2'fucosyllactose treatment, replicate A	GSM2338953, lacto-*N*-neotetraose treatment, replicate A
GSM1419265, 2'fucosyllactose treatment, replicate B	GSM1425054, 2'fucosyllactose treatment, replicate B	GSM2338954, lacto-*N*-neotetraose treatment, replicate B
GSM1419266, 3-fucosyllactose treatment, replicate A	GSM1425055, 3-fucosyllactose treatment, replicate A	GSM2338955, 2'fucosyllactose treatment, replicate A
GSM1419267, 3-fucosyllactose treatment, replicate B	GSM1425056, 3-fucosyllactose treatment, replicate B	GSM2338956, 2'fucosyllactose treatment, replicate B
GSM1419268, 6'sialyllactose treatment, replicate A	GSM1425057, 6'sialyllactose treatment, replicate A	GSM2338957, 3-fucosyllactose treatment, replicate A
GSM1419269, 6'sialyllactose treatment, replicate B	GSM1425058, 6'sialyllactose treatment, replicate B	GSM2338958, 3-fucosyllactose treatment, replicate B
	GSM1425059, mucin treatment, replicate A	
	GSM1425060, mucin treatment, replicate B	

Table S2. Genes in module Greenyellow from *B. infantis**.

Table S3. Genes in module Blue from *B. infantis**.

Table S4. Genes in module Darkturquoise for *B. bifidum* SC555*.

Table S5. Genes in module Purple for *B. bifidum* SC555*.

Table S6. Genes in Green module for *B. longum* SC596*.

Table S7. Genes in Darkgreen module for *B. longum* SC596*.

Table S8. Hub genes for each module in *B. longum* subsp. *infantis* ATCC 15697.

Supplementary Material available in

https://www.frontiersin.org/articles/10.3389/fmolb.2023.1040721/full

Table S9. Hub genes for each module in *B. bifidum* SC 555.

Module	**Locus Taq**	**Degree**	**Coding protein**
Cyan	BBIF_01407	56	16S rRNA (adenine1518-N6/adenine1519-N6)-dimethyltransferase)
DarkTurquoise	BBIF_01738	20	ABC-2 Type transport system, ATP-binding protein
Black	BBIF_01747	76	UDP-*N*-Acetylmuramyl tripeptide synthase
Turquoise	BBIF_00702	61	16S rRNA (cytidine1402-2'-O)-methyltransferase
Brown	BBIF_01233	15	TetR/AcrR family transcriptional regulator
Pink	BBIF_00964	28	Polyphosphate: nucleotide phosphotransferase
Purple	BBIF_00840	21	Large subunit ribosomal L18
	BBIF_00527	21	PTS system
Darkgrey	BBIF_00921	12	Nucleoside-diphosphate-sugar epimerase
Darkred	BBIF_00795	8	Putative FmdB Regulatory Protein
Darkorange	BBIF_016016	7	Glycosyltransferase involved in cell wall biosynthesis
	BBIF_00538	7	Hypothetical protein
	BBIF_00474	7	FtsX transport permease

Table S10. Hub genes for each module in *B. longum* SC596

Module	**Locus Taq**	**Degree**	**Coding protein**
Cyan	BLNG_00490	19	MraZ protein
Darkorange	BLNG_02236	18	Amino acid DMT transport membrane protein YddG family
Darkolivegreen	BLNG_01776	17	ADP-pyrophosphatase
Orange	BLNG_01683	14	Helix turn Helix protein
Turquoise	BLNG_00139	36	Manganese transport protein
Darkgreen	BLNG_02244	11	ATP-dependent Zn protease
Green	BLNG_00718	14	Nitroreductase family protein
Magenta	BLNG_01958 BLNG_01309	21 21	Regulator of protease activity Nicotinate phosphoribosyltransferase
Greenyellow	BLNG_02106	29	Hypothetical protein
Salmon	BLNG_01438	40	LSU ribosomal protein L5P
Darkmagenta	BLNG_00479	15	Cell division protein FtsQ

Table S11. Description of different bacterial strains of the human gut microbiota used for Bidirectional culture.

Pairs	**GMP-degrader bacteria**	**Non-GMP-degrader bacteria**
a)	*B. bifidum* JCM 1254	*B. breve* I1
b)	*B. bifidum* JCM 1254	*B. longum* M12
c)	*B. bifidum* JCM 1254	*L. symbiosum* Wal-14673
d)	*B. infantis* ATCC 15697	*S. thermophilus* LM5
e)	*C. perfringens* ATCC 13124	*S. thermophilus* LM8
f)	*B. infantis* ATCC 15697	*B. longum* M12
g)	*B. infantis* ATCC 15697	*B. breve* I1
h)	*B. bifidum* JCM 1254	*R. gnavus* CC55_001C

Table S12. General information of the primers designed for fold change measurement by qPCR.

Bacteria	**Locus Taq**	**Direction**	**Direction**	**Tm (°C)**	**Sequence**	**Gene Size (bp)**
B. infantis ATCC 15697	Blon_2348	Exo-alpha-sialidase	F	59.96	CAGGCCAAGGTTCTGCTGTA	1185
			R	60.04	GAGCCAGTCGAGCGAGATAC	
	Blon_1831	Glycoside hydrolase, family 25	F	59.90	TATGGCCGATGGTGTACGTG	1233
			R	59.96	TGTAGGTCGATGGTCTGGGT	
	Blon_1479	Alpha-*N*-Acetylgalactosaminidase	F	59.96	AACTCCATGGAATGCCGGTT	1896
			R	59.69	GCGCAAAGGTGACCAATGAA	
	Blon_0646	Sialidase	F	59.98	GTCTTGTATGCGACGAACGC	2283
			R	59.89	GAGGTAGATGTTGCCCGAGG	
	Blon_2177	Multiple sugar transport system substrate-binding protein (GNB transporter)	F	61.90	TACAAGGTCGACGGCAACTC	1317
			R	63.20	AAGGAACTCCATGGCTTCGG	
B. bifidum JCM 1254	Ga012833 0_104513	Endo-α-N-Acetylgalactosaminidase	F	63.1	AAATGTACCCGGAAGCCAGG	1866
			R	60.4	GTTGCGTCATACTCGTTGCC	
	Ga012833 0_10636	Exo-α-sialidase	F	60.9	GTCGCGTGATGCTCAATTCC	2508
			R	62.1	ACAGCAGACCATAACCACCG	
	Ga012833 0_100248	ABC Transport protein	F	61.9	TCGACAACTGGCGTAAAGCG	2226
			R	59.9	AGCACGTCATCGAAGCTGAT	