CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Metabolism of human milk oligosaccharides by infant-associated bifidobacteria
Authors	James, Kieran
Publication date	2018
Original Citation	James, K. 2018. Metabolism of human milk oligosaccharides by infant-associated bifidobacteria. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2018, Kieran James. - http://creativecommons.org/licenses/by-nc-nd/3.0/
Download date	2026-11-03 08:40:42
Item downloaded from	https://hdl.handle.net/10468/6097



UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Metabolism of human milk oligosaccharides by infant-associated bifidobacteria



Ollscoil na hÉireann, Corcaigh

THE NATIONAL UNIVERSITY OF IRELAND, CORK

A thesis presented to the National University of Ireland for the

Degree of Doctor of Philosophy by

Kieran James, B.Sc

School of Microbiology

University College Cork

Supervisor: Prof. Douwe van Sinderen

Head of School: Prof. Gerald F. Fitzgerald

April 2018

TABLE OF CONTENTS

Table of contents i

Declaration iv

List of figure legends v

List of table legends viii

List of publications xi

Abbreviations xii

General abstract xvii

Chapter I: General Introduction 1

1.1 Summary 2

1.2 Introduction 3

1.3 Human Milk Oligosaccharides (HMOs) 8

1.4 Infant-Associated Bifidobacteria 13

1.5 Metabolism of Human Milk Oligosaccharides by Infant-Associated Bifidobacteria 17

1.6 Conclusion 30

1.7 Thesis outline 31

1.8 Figures 32

1.9 References 34

Chapter II: *Bifidobacterium breve* UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways 62

2.1 Abstract 63

2.2 Introduction 64

2.3 Materials and Methods 67

2.4 Results 75

2.5 Discussion 85

2.6 Acknowledgements 88

2.7 Tables and Figures 89

2.8 Supplemental Material 99

2.9 References 106

Chapter III: *Bifidobacterium breve* UCC2003 employs multiple transcriptional regulators to control metabolism of particular human milk oligosaccharides. 115

3.1 Abstract 116

3.2 Introduction 117

3.3 Materials and Methods 120

3.4 Results 127

3.5 Discussion 135

3.6 Acknowledgements 138

3.7 Tables and Figures 139

3.8 Supplemental Material 147

3.9 References 159

Chapter IV: Screening of infant faecal samples for human milk oligosaccharide-utilising *Bifidobacterium* isolates 168

4.1 Abstract 169

4.2 Introduction 170

4.3 Materials and Methods 173

4.4 Results 176

4.5 Discussion 181

4.6 Acknowledgements 184

4.7 Tables and Figures 186

4.8 Supplemental Material 190

4.9 References 191

Chapter V: Molecular dissection of a gene cluster from *Bifidobacterium kashiwanohense* involved in the metabolism of fucosyllactose, a major human milk oligosaccharide component 200

5.1 Abstract 201

5.2 Introduction 202

5.3 Materials and Methods 205

5.4 Results 215

5.5 Discussion 226

5.6 Acknowledgements 229

5.7 Tables and Figures 230

5.8 Supplemental Material 239

5.9 References 245

Chapter VI: General Discussion and Future Perspectives 256
6.1 References 261

Acknowledgements 265

Declaration

I hereby declare that the research presented in this thesis is my own work and effort, and that it has not been submitted for any other degree, either at University College Cork or elsewhere. Whatever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature and acknowledgement of collaborative research and discussions.

This work was completed under the guidance of Prof. Douwe van Sinderen at the APC Microbiome Institute & School of Microbiology, Biosciences Institute, University College Cork.

Signature:…….

Date:…….

LIST OF FIGURE LEGENDS

Figure 1.1. Schematic diagram of selected HMO structures. (pg. 32)

Figure 1.2. Schematic representation of the differing HMO utilisation strategies in *B. bifidum*, *B. breve*, *B. longum* subsp. *infantis* and *B. kashiwanohense*. (pg. 33)

Figure 2.1. Schematic structures of Type I HMO moiety LNT, and Type II HMO moiety LNnT. (pg. 91)

Figure 2.2. Schematic representation of the gene loci involved in the utilisation of LNT, LNnT and their substituents in *B. breve* UCC2003, as based on transcriptome analysis. (pg. 92)

Figure 2.3. HPAEC chromatogram profiles of LNT and LNnT, when incubated LntA and/or NahA. (pg. 93)

Figure 2.4. Growth of *B. breve* UCC2003 wild type and mutants on LN(n)T and LN(n)T-derived substrates. (pg. 95)

Figure 2.5. Heatmap representing the distribution of homologs of key LN(n)T/LNB utilisation genes from *B. breve* UCC2003, *B. longum* subsp. *longum* JCM1217 and *B. bifidum* PRL2010, across the *Bifidobacterium* genus. (pg. 97)

Figure 2.6. Schematic representation of the proposed model for the metabolism of free LNT, LNnT and LNB by *B. breve* UCC2003. (pg. 98)

Figure S2.1. Growth of 16 *B. breve* strains on LNT, LNnT and lactose. (pg. 102)

Figure S2.2. HPAEC chromatogram profiles of LNT and LNnT, when incubated LacZ2 and/or NahA. (pg. 103)

Figure S2.3. Growth of *B. breve* UCC2003 wild type, as well as select mutants and their complemented strains, on LN(n)T and LN(n)T-derived substrates. (pg. 105)

Figure 3.1. Schematic representation of HMO metabolism-associated loci in *B. breve* UCC2003. (pg. 140)

Figure 3.2. EMSA images showing LntR, NahR, and NagR1 interactions with a range of DNA fragments from the regions in the proximity of their predicted target promoters. (pg. 142)

Figure 3.3. WebLogos generated for the operator motif consensus sequences of LntR, NahR and NagR1. (pg. 143)

Figure 3.4. EMSA images showing LntR, NahR, and NagR1 interactions with promoter-containing DNA fragments, with the addition of gradients of their respective inducers. (pg. 144)

Figure 3.5. Schematic representation of the proposed model for transcriptomic regulation of LNT, LNnT and LNB metabolism by *B. breve* UCC2003. (pg. 146)

Figure S3.1. Primer extension images and schematic representations of the *lntP1, lntS, nahS, nahA, nagB3, nagK, lnpB* and *gltA* promoter regions. (pg. 152-155)

Figure S3.2. EMSA images showing LntR, NahR, and NagR1 interactions with promoter-containing DNA fragments, with the addition of a range of potential carbohydrate inducers. (pg. 156)

Figure S3.3. EMSA images showing LntR, NahR, and NagR1 interactions with promoter-containing DNA fragments, with the addition of gradients of Gal-1-P and Gal-6-P. (pg. 158)

Figure 4.1. Percentages of each isolated *Bifidobacterium* species obtained in total during the screening of infant faecal samples, as well as for each selective carbohydrate source used. (pg. 188)

Figure 4.2. Heatmap representing GH profiles of *B. breve* UCC2003, *B. longum* subsp. *longum* NCC2705, *B. longum* subsp. *infantis* ATCC 15697, *B. bifidum* PRL2010, *B. kashiwanohense* JCM 15439 and *B. pseudocatenulatum* DSM 20438. (pg. 189)

Figure 5.1. Growth of *B. kashiwanohense* APCKJ1 on a range of carbohydrate sources. (pg. 232)

Figure 5.2. HPAEC-PAD chromatogram profiles of *B. kashiwanohense* APCKJ1 fermentations of 2'FL and 3-FL. (pg. 233)

Figure 5.3. Schematic representation of the gene loci involved in the utilisation of 2-FL or 3-FL in *B. kashiwanohense* APCKJ1, and L-fucose in *B. breve* UCC2003. (pg. 234)

Figure 5.4. HPAEC-PAD chromatogram profiles of 2'-FL and 3-FL, when incubated with FucA1 or FucA2, at various timepoints. (pg. 236)

Figure 5.5. Growth of *B. kashiwanohense* APCKJ1, *B. breve* UCC2003 wild type and recombinant *B. breve* UCC2003 strains expressing APCKJ1 genes, on lactose, 2'-FL and 3-FL. (pg. 237)

Figure 5.6. Schematic representation of the pathway for the utilisation of fucosyllactose in *B. kashiwanohense* APCKJ1. (pg. 238)

Figure S5.1. Multiple sequence alignment of *B. kashiwanohense* APCKJ1 gene *galR* with other predicted LacI-encoding genes. (pg. 243)

Figure S5.2. HPAEC-PAD chromatogram profiles of 2'-FL and 3-FL, when incubated with FucA1 and FucA2 together, at various timepoints. (pg. 244)

LIST OF TABLE LEGENDS

Table 2.1. *B. breve* UCC2003 genes upregulated in transcription during growth in mMRS medium supplemented with 1 % LNT, LNnT, LNB, lactosamine-HCl, or lactose as the sole carbohydrate. (pg. 90)

Table S2.1. Oligonucleotide primers used in Chapter II. (pg. 99)

Table S2.2. Bacterial plasmids and strains used in Chapter II. (pg. 100)

Table S2.3. Locus tags, names and product functions of *B. breve* UCC2003 and *B. bifidum* PRL2010 genes selected for bioinformatic analysis of LN(n)T/LNB utilisation-associated homologs across the *Bifidobacterium* genus. (pg. 101)

Table 3.1. *B. breve* UCC2003 regulator mutant genes upregulated in transcription during growth in mMRS medium supplemented with 1 % ribose as the sole carbohydrate, as compared to the wild type (control). (pg. 139)

Table S3.1. Oligonucleotide primers used in Chapter III. (pg. 147-148)

Table S3.2. Bacterial plasmids and strains used in Chapter III. (pg. 149)

Table S3.3. Locations and sizes of fragments used in electromobility shift assays, in relation to the promoter regions' respective transcription start sites. (pg. 150)

Table S3.4. Predicted operator sequence used for the *in silico* generation of operator sequence motif consensuses. (pg. 151)

Table 4.1. Total number of isolates of *Bifidobacterium* belonging to each species obtained during the screening of infant faeces. (pg.186)

Table 4.2. Number of isolates of *Bifidobacterium* belonging to each species obtained from each individual faecal sample. (pg. 187)

Table S4.1. Oligonucleotide primers used in Chapter IV. (pg. 190)

Table 5.1. *B. kashiwanohense* APCKJ1 genes that are transcriptionally upregulated during growth in mMRS medium supplemented with 2-FL or 3-FL as the sole carbohydrate, as compared to growth in mMRS supplemented with sorbitol. (pg. 230)

Table 5.2. *B. breve* UCC2003-*fucA1-fucSP1P2* genes that were transcriptionally upregulated during growth in mMRS medium supplemented with 2-FL or lactose as the sole carbohydrate, as compared to growth in mMRS supplemented with ribose. (pg. 231)

Table S5.1. Oligonucleotide primers used in Chapter V. (pg. 239)

Table S5.2. Bacterial plasmids and strains used in Chapter V. (pg. 240)

Table S5.3. *B. longum* subsp. *infantis* DSM20088 genes involved in the catabolism of fucosyllactose, and their homologs in *B. kashiwanohense* APCKJ1 and *B. breve* UCC2003, based on a blastP search of the APCKJ1 and UCC2003 genomes. (pg. 241)

Table S5.4. Scores for APCKJ1 genome sequencing reads and quality. (pg. 242)

This thesis is dedicated to my father.

LIST OF PUBLICATIONS

Chapter II: **James, K., O'Connell Motherway, M., Bottacini, F. & van Sinderen, D. (2016).** *Bifidobacterium breve* UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways. Sci Rep, **6**: p.38560.

Chapter III: **James, K., O'Connell Motherway, M., Penno, C., O'Brien, R.L., & van Sinderen, D.** *Bifidobacterium breve* UCC2003 employs multiple transcriptional regulators to control metabolism of particular human milk oligosaccharides. AEM, Vol. **84**, No. **9**.

Chapter V: **James, K., Bottacini, F., Vigoureux, M., Egan, M., O'Connell Motherway, M. & van Sinderen, D.** Molecular dissection of a gene cluster from *Bifidobacterium kashiwanohense* involved in the metabolism of fucosyllactose, a major human milk oligosaccharide component. Manuscript in preparation, April, 2018.

ABBREVIATIONS

ABC-type transporters: ATP-Binding Cassette Transporters

ADP: Adenosine Diphosphate

APC: Alimentary Pharmabiotic Centre

APS: Ammonium Persulfate

ATCC: American Type Culture Collection

ATP: Adenosine Triphosphate

BLAST: Basic Local Alignment Search Tool

BLASTn: Basic Local Alignment Search Tool for nucleotide sequences

BLASTp: Basic Local Alignment Search Tool for protein sequences

CAZy: Carbohydrate-Active enZYmes

CCR: Carbon Catabolite Repression

cDNA: complementary DNA

CFU: Colony Forming Units

CGH: Comparative Genome Hybridisation

Cm: Chloramphenicol

COG: Cluster of Orthologous Groups of proteins

DF-LN(n)T: Difucosyl-Lacto-N-(neo)tetraose

DNA: Deoxyribonucleic Acid

DS-LN(n)T: Disialyl-Lacto-N-(neo)tetraose

DSM: German Collection of Microorganisms and Cell Cultures

EDTA: Ethylenediaminetetraacetic Acid

Em: Erythromycin

EMSA: Electrophoretic Mobility Shift Assay

EtBr: Ethidium Bromide

FOS: Fructo-oligosaccharides

F6P: Fructose-6-Phosphate

F6PPK: Fructose-6-Phosphate Phosphoketolase

Gal: Galactose

GalNAc: N-acetylgalactosamine

Gal-1-P: Galactose-1-Phosphate

Gal-6-P: Galactose-6-Phosphate

GEO: Gene Expression Omnibus

GH: Glycosyl Hydrolase

GIT: Gastrointestinal Tract

Glc: Glucose

GlcNAc: N-acetylglucosamine

Glc-6-P: Glucose-6-Phosphate

GM17: Glucose-M17 medium

GNB: Galacto-N-biose

GO: Gene Ontology

GOS: Galacto-oligosaccharides

HMO: Human Milk Oligosaccharides

HPAEC-PAD: High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection

IPTG: Isopropyl-β-D-thiogalactopyranoside]

IRD700: Infrared Dye 700 nanometers (emission wavelength)

JCM: Japanese Collection of Microorganisms

Kan: Kanamycin

KDa: kilo Daltons

LAB: Lactic Acid Bacteria

LacNAc: N-acetyllactosamine

Lactosamine HCl: Lactosamine Hydrochloride

LB: Luria Bertani medium

LMG: Belgian Coordinated Collection of Microorganisms

LNB: Lacto-N-biose

LnbP: Lacto-N-biose phosphorylase

LNDFH: Lacto-N-difucohexaose

LNFP: Lacto-N-fucopentose

LNH: Lacto-N-hexaose

LNT: Lacto-N-tetraose

LNnT: Lacto-N-neotetraose

LN(n)T: Lacto-N-(neo)tetraose (i.e. LNT/LNnT)

MCS: Multiple Cloning Site

MEME: multiple expectation maximisation for motif elicitation

MOPS: Morpholinepropanesulfonic acid

MRS: de Man, Rogosa and Sharpe medium

mMRS: modified de Man, Rogosa and Sharpe medium

mRNA: messenger RNA

Mup: Mupirocin

MW: Molecular Weight

NagC: N-acetylglucosamine repressor

NCBI: National Centre for Biotechnology Information

NCFB: National Collection of Food Bacteria

NCIMB: National Collection of Industrial and Marine Bacteria

Neu5Ac: N-acetylneuramic (sialic) acid

NCTC: National Collection of Type Cultures

NIZO: Nizo Food Research

Nys: Nystatin

OD: Optical Density

ORF: Open Reading Frame

PCR: Polymerase Chain Reaction

PEP-PTS: Phosphoenolpyruvate-dependent Phosphotransferase System

Pfam: Protein families (database)

Poly[d(I-C)]: Poly- deoxyinosinic-deoxycytidylic acid

PRL: Culture collection of probiogenomics, University of Parma

RBS: Ribosome Binding Site

RCA: Reinforced Clostridial Agar

RCM: Reinforced Clostridial Medium

RNA: Ribonucleic Acid

ROK: Repressor Open Reading frame Kinase

RT-PCR: Reverse-transcription-PCR

SCFA: Short Chain Fatty Acids

SDS-PAGE: Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis

SignalP: Signal Peptide (prediction software)

Strep: Streptomycin

Spec: Spectinomycin

Tet: Tetracycline

TF: Transcription Factor

TMHMM: Tied Mixture Hidden Markov Model (transmembrane helix prediction software)

TOS: Trans-galacto-oligosaccharides

Tris-HCl: Tris(hydroxymethyl)aminomethane- Hydrochloride

UCC: University College Cork

wt/vol: weight/volume

X-gal: 5-bromo-4-chloro-3-indolyl-D-galactopyranoside

XylR: Xylose Repressor

2'-FL: 2'-fucosyllactose

3-FL: 3-fucosyllactose

3'-SL: 3'-sialyllactose

6'-SL: 6'-sialyllactose

GENERAL ABSTRACT

Bifidobacteria are Gram-positive, anaerobic bacteria belonging to the Actinobacteria phylum, and are commensals of the mammalian, avian and occasionally insect gastrointestinal tracts. In humans, bifidobacteria are typically highly abundant in the intestinal microbiota of healthy breastfed infants, in particular a small number of infant-associated species, including *Bifidobacterium breve*, *Bifidobacterium longum* subsp. *infantis*, *Bifidobacterium bifidum* and *Bifidobacterium kashiwanohense*. A key adaptation, facilitating the establishment and dominance of these species in the breastfed infant gut microbiota, is the ability to consume and metabolise specific glycans only found in breastmilk, and indigestible for the infant, known as Human Milk Oligosaccharides (HMOs). Fascinatingly, the adaptation to utilise HMO as a substrate by bifidobacteria is almost exclusively reserved for infant-associated species, and even more fascinatingly, these species have developed strikingly varying strategies for the consumption of different HMO components. Strategies for the metabolism of various individual HMO glycans in different species of bifidobacteria shall be discussed in this thesis, with particular focus on *B. breve* UCC2003 and the novel isolate *B. kashiwanohense* APCKJ1.

Chapter II of this thesis is comprised of work elucidating the mechanisms and components of LNT, LNnT and LNB utilisation pathways in *Bifidobacterium breve* UCC2003. Using a combination of experimental approaches, the enzymatic machinery involved in the metabolism of LNT, LNnT and LNB is identified and characterised. Furthermore, the distribution across the genus, of homologs for the key genes involved in the utilisation of these substrates, is analysed.

Chapter III focuses on identifying the regulatory network responsible for the transcriptional control of the genes involved in LN(n)T and LNB metabolism, as described in Chapter II. Three transcriptional regulators and corresponding operator and associated (inducible) promoter sequences are characterised, the latter governing transcription of the genetic elements involved in LN(n)T/LNB metabolism. Furthermore, identification of the transcriptional

effectors reveals the presence of a series of positive-feedback loops, inducing expression in the presence of breakdown products of key HMO-derived metabolites.

In Chapter IV, *Bifidobacterium* isolates are obtained by screening a number of faecal samples from breastfed infants, using HMO components fucosyllactose and sialyllactose as selective carbohydrates, and compared with isolates obtained using lactose or GOS. A range of bifidobacterial species were obtained, varying between the selective carbohydrate used, and supporting the notion of selective HMO consumption. Analysis of the glycosyl hydrolase profiles of representative strains of the species obtained reveals interesting correlations with their preferential carbohydrate-based selection.

In Chapter V, a novel *B. kashiwanohense* isolate, APCKJ1, which was isolated in the work of Chapter IV, is demonstrated to consume fucosyllactose, and the mechanisms of its utilisation of both fucosyllactose and L-fucose is examined. Using a combination of approaches, the main cellular machinery involved in the uptake and degradation of fucosyllactose is characterised, and heterologous expression of these genes in a *B. breve* UCC2003 host reveals not only the mechanisms of the utilisation of fucosyllactose, but a potentially functional pathway for the catabolism of fucose, in both strains.

The work presented in this thesis represents novel information on the metabolism of HMO glycans in bifidobacteria, particularly in the species *B. breve* and *B. kashiwanohense*, as well as key insights into the strategies of HMO utilisation by infant-associated bifidobacteria in general, as an adaptation to the GIT of breastfed infants.

Chapter I

General Introduction

1.1 Summary

Consumption of breastmilk is considered to be an important factor that determines the development and composition of the mammalian infant gut microbiota. A specific subset of glycans secreted in human mothers' milk, known as human milk oligosaccharides (HMOs), are believed to be particularly instrumental in the compositional development of the neonatal gut microbiota in humans. Among other functions, HMOs act as growth factors for specific (beneficial, see below) bacteria, and are thus considered to be prebiotics, thereby supporting the establishment of a (healthy) infant-type microbiota. Of the bacterial components of the neonatal (human) gut microbiota, members the genus *Bifidobacterium* are dominant, in particular a small number of species, which are believed to bestow a range of health benefits to the infant host. It is thus hardly surprising that these infant-associated bifidobacterial species possess the ability to directly or indirectly consume HMOs as their sole carbohydrate source. A number of different strategies for HMO utilisation have been observed in these species, targeting various HMO components of this highly heterogeneous glycan mixture. In this review, we discuss the structures and biological functions of HMOs, and the role of bifidobacteria in the gut microbiota of the breastfed infant, as well as describing the various approaches employed by infant-associated *Bifidobacterium* species for the utilisation of HMOs in order to establish stable populations in this highly specific niche.

1.2 Introduction

1.2.1 General features of the infant gut microbiota. Among the many ecological niches within and on the human body, the most densely populated with microbial species is the intestine. Composed of trillions of microbes, this complex community is known as the gut microbiota [1, 2]. This anaerobic ecosystem is rich in nutrients, and populated by microbial members of the three domains of life, *Eukarya*, *Bacteria* and *Archaea*, as well as viruses, all of which may be present as autochthonous or transient inhabitants of this niche [3]. The composition of this community is affected by a range of environmental conditions within the host, which means that the microbiota will not only vary between hosts, but also within a single host during its life span [4]. These microorganisms are believed to influence each other through countless, yet mostly unexplored interactions, while they also affect and are affected by the host via a range of interactions, both symbiotic and parasitic [5]. Such microbe-microbe and microbe-host interactions, and thus the composition of this community, are increasingly being implicated in host health, acting as a critical factor in the development or prevention of numerous health conditions [6, 7]. Food breakdown and nutrient liberation, direct pathogen inhibition or exclusion, immune system priming and modulation, and promotion of host cell differentiation are just some of the functions attributed to (elements of) the gut microbiota. The concept of the overall ecosystem of the human body as a so-called holobiont means that the various associated communities are dynamic, and change with and in response to the host, in an attempt to achieve the most harmonious symbioses possible [2, 8]; however, it is also the changeable nature of these communities that allow disease states to develop [6].

The composition and interactions of this community are known to change over the lifetime of a (mammalian) host [9-11], with a very clear development from an infant-type to an adult-type microbiota as the infant progresses in age and changes dietary habits [12-15]. It is thought that members of the early gut microbiota significantly impact on the health of the infant host [16]. Until recently, the limitations of culture-based techniques meant that only a small proportion of all members of the human gut microbiota had been cultured, let

alone studied in pure culture [17]. This led to the rapid development and use of culture-independent approaches for the determination of microbiota composition [2, 18]. Recent developments in culturomics, however, are aimed at closing the gap between cultivation-dependent and cultivation-independent knowledge of the microbiota by enabling *in vitro* cultivation and characterisation of an ever increasing number of microbes [19, 20].

The infant gut microbiota is typically viewed as that of subjects aged <1 year, is considered to have a low diversity, yet being highly dynamic and unstable as compared to the adult-type microbiota (often described as present in individuals aged >1 year, although this can vary depending on the individual) [13, 21]. Despite this, it has been observed that across individuals, the dominant genera of the infant gut microbiota are *Bifidobacterium*, *Veillonella*, *Streptococcus*, *Citrobacter*, *Escherichia*, *Bacteroides* and *Clostridium* [13, 22], with in particular bifidobacteria being dominant in breastfed infants [21, 23-26].

1.2.2 Factors affecting the development of the infant gut microbiota. A range of factors, both environmental and host-specific, play a role in the development and composition of the gut microbiota in the neonatal gut, which is essentially a blank canvas for the establishment of a complex microbial community [14]. Accordingly, these factors have a major bearing on the overall health status of the infant host, and their development as they progress in age [21].

Recent studies have suggested that establishment of the infant gut microbiota begins as early as during pregnancy [27-29] through microbial exposure *in utero*, although these claims are still controversial. The three biggest determinants of neonatal microbiota composition as we currently know them are delivery mode, gestational age and feeding mode. In full term infants, delivery mode is the first major determinant of neonatal gut microbiota composition [30]. Contact with the maternal faecal and vaginal microbiotas during vaginal delivery results in the colonisation of the neonatal gut with vaginal commensals such as *Lactobacillus* and *Prevotella* [31-33], whereas

caesarean-section (C-section)-delivered neonates are more frequently colonised with microbes associated with the maternal epidermal microbiota and nosocomial microbes [12, 14, 32, 34], such as Proteobacteria, Firmicutes (including *Staphylococcus* and *Clostridium*) and *Propionibacterium* [33, 35-37]. As well as a reduced diversity, the microbiotas of C-section-delivered infants (compared to vaginally-delivered infants) have also been shown to display reduced relative abundance of *Bifidobacterium* and *Bacteroides*, and increased relative abundance of some groups of *Clostridium* [13, 31, 32, 35]. The differences between the compositions of naturally-delivered and C-section-delivered infants do, however, tend to vanish over time, particularly after 12 months [38, 39].

Gestational age also plays a major role in the composition of the neonatal microbiota. Preterm infants are classed as neonates born before 37 weeks gestation [40, 41], and because of developmental immaturity, are often subject to extended stays in hospital following birth, which may include artificial feeding, antibiotic administration and other drug treatment regimes, all of which are known to affect the composition of the gut microbiota. Greater abundances of bacteria of the *Enterobacteriaceae* family and species of the *Enterococcus*, *Lactobacillus* and *Staphylococcus* genera have been observed in the microbiota of pre-term than full-term neonates, as well as reduced abundances of *Bacteroides* and *Bifidobacterium* species [13, 42-49] initially following birth, with the absence of bifidobacteria noted as far as 7 days post-birth in pre-term neonates [50].

Last of the three major factors to affect the microbiota of the neonatal gut is the manner of feeding. Numerous studies have identified significant differences in the gut microbiota composition of breastfed and formula-fed infants [13, 51, 52], which is believed to be caused by the presence of unique compounds present in breastmilk, including specific carbohydrates known as human milk oligosaccharides (HMOs; see below). Breastfed infants have been demonstrated to possess a gut microbiota with increased abundances of bifidobacteria and lactobacilli [24, 53-56], whereas that of formula-fed infants displays a greater diversity, including species of *Escherichia*, *Clostridium*, *Bacteroides*, *Prevotella* and *Enterococcus* [38, 49, 52-54, 56,

57]. It has furthermore been observed that formula-fed infants adopt an adult-type microbiota sooner in life as compared to their breastfed counterparts [34]. In breastfed infants, the weaning stage from exclusive breastfeeding to solid foods encourages the development of the gut microbiota from the dynamic infant-type composition to the more stable, yet much more complex adult-type microbiota [22, 58, 59], which functions more specifically in the metabolism of nutrients derived from an omnivorous diet, including plant-derived carbohydrates [25].

Use of the once-popular term ‘dysbiosis’ has recently become unfashionable for some [60], as the concept of a universal template for a healthy, ‘normal’ microbiota is undefinable due to natural inter- and intra-personal variations in microbiota composition. However, it has been well-documented that certain changes in the gut microbiota composition predispose to the development of disease states in the host, whether pathogenic or immunogenic [6]. In the infant gut microbiota, factors such as pre-term birth and its associated (disturbed) microbiota have been implicated in the development of infectious diseases, including necrotising enterocolitis (NEC) [61] and neonatal sepsis [62, 63]. Immune disorders such as asthma [64], allergy [65] and type 1 diabetes [66], as well as increased risk of obesity in later life [67] have all been correlated with a variety of factors, such as C-section birthing, that are potentially disruptive to the development of the healthy infant microbiota.

1.2.3 Pre- and pro-biotics, and their role in infant microbiota development. As our knowledge of the presence and roles of various member of the infant gut microbiota increases, so too increases the potential for their use in and/or modulation by intervention treatments. The realisation of this potential has come to fruition through the development of the concepts of probiotics and prebiotics.

Probiotics are defined as “live microorganisms that, when administered in adequate amounts, confer a health benefit on the host” [68]. The use of probiotics as clinical interventions in infants is progressively becoming a

more tangible possibility, for example in the treatment of conditions such as neonatal sepsis [69] or NEC [70]. Such interventions rely on direct interactions between the probiotic strains and the pathogens themselves, or the modulation of the immune system for the prevention of disease.

The definition of the term 'prebiotic' has undergone a number of iterations since its original description [71-74], but the most recent consensus on the term defines it as "a substrate that is selectively utilized by host microorganisms conferring a health benefit" [75]. To be considered a prebiotic, an ingredient must: (I) resist gastric activity, hydrolysis by mammalian enzymes and gastrointestinal absorption; (II) be fermented by (particular elements of) the intestinal microbiota; and (III) stimulate growth and/or activity of intestinal bacteria associated with health and well-being [71]. By the rules of its strictest definition, only two commercially-used food ingredients currently meet the criteria to be classed as prebiotics, namely (trans-)galactooligosaccharides (TOS/GOS) and fructooligosaccharides (FOS), including inulin [76]. Inclusion of a mixture TOS and inulin in infant formula has been shown to increase the abundance of faecal bifidobacteria in both preterm and term infants [77, 78]. Furthermore, the potential of synbiotics, which are defined as "mixtures of probiotics and prebiotics that beneficially affect the host by improving the survival and implantation of live microbial dietary supplements in the gastrointestinal tract, by selectively stimulating the growth and/or by activating the metabolism of one or a limited number of health-promoting bacteria, thus improving host welfare" [79], is showing promise for the treatment of infant conditions such as NEC [80] and neonatal sepsis [81].

While the idea of pre- and pro-biotics as clinical or commercially-available interventions may be considered a relatively recent invention, both of these phenomena have naturally evolved in mammals, including humans. The process of vaginal birthing encourages the vertical transfer of microbes from mother to infant, and recent work has also shown the potential transfer of indigenous breastmilk bacteria from mother to infant [82, 83]. Breastmilk functions as a prebiotic through the presence of HMOs, which encourage the establishment of specific microbes by selective utilisation of these sugars [74,

84, 85] (see below). Breastmilk itself can therefore be viewed as the archetypal synbiotic, conveying both vital microbes and prebiotic substrates to the neonatal host.

1.3 Human Milk Oligosaccharides (HMOs)

1.3.1 General features. HMOs constitute a heterogeneous mix of structurally diverse unconjugated glycans that are highly abundant in and rather unique to human milk [74, 86-88]. Over 200 different structures have been identified [89], with oligosaccharide concentration and compositions varying between individual women and over the course of lactation [87, 90]. In fact, a recent study found that HMO concentration and composition profiles vary substantially across populations from different geographical regions [91]. While colostrum, the thick secretion produced by the mammary glands in the days leading up to and following birth, has been found to contain as much as 20-25 g HMOs per litre [92, 93], the concentration of HMOs in human breastmilk is typically 5-20 g/L [87, 92-97]. This, however, still exceeds the total amount of protein in breastmilk, and in fact is 100 to 1000 times greater than the total oligosaccharide concentration of bovine milk [88].

1.3.2 Structure and composition. HMO glycans are structures composed of 5 monosaccharides: glucose (Glc), galactose (Gal), N-acetylglucosamine (GlcNAc), fucose (Fuc) and Sialic/N-acetylneuramic acid (Sia/Neu5Ac) [74, 87]. Lactose (Galβ1-4Glc) is the base glycan forming the reducing end of all HMO structures, and can be elongated through a β1-3 or β1-6 linkage with lacto-N-biose, (LNB; Galβ1-3GlcNAc) or its isomer N-acetyllactosamine (LacNAc; Galβ1-4GlcNAc) [87]. This results in the generation of the tetrasaccharides lacto-N-tetraose (LNT; Galβ1-3GlcNAcβ1-3Galβ1-4Glc) or lacto-N-neotetraose (LNnT; Galβ1-4GlcNAcβ1-3Galβ1-4Glc), respectively, the latter of which can be further elongated by the addition of further LNB or LacNAc moieties [74, 88]. HMO structures with LNB linked to the terminal lactose (at the non-reducing end) are classed as Type 1 structures, while those with LacNAc at this position are classed as Type 2 HMOs. Type 1 HMOs

have been found to predominate over Type 2 HMOs in human breastmilk [98]. Interestingly, this is in contrast to other lactating mammals (which produce similar milk oligosaccharides), where Type 2 structures dominate over (or are present in the absence of) Type 1 structures [99-101]. β1-6 linkages between disaccharide units allows chain branching of chains [74]. Lactose or larger chains can be fucosylated via an α1-2, α1-3 or α1-4 linkage, or sialylated via an α2-3 or α2-6 linkage, with many HMO structures occurring in several isomeric forms [74, 87]. The mechanisms of HMO synthesis are as yet poorly understood [74]. A schematic of a selection of HMO structures and their components is shown in Figure 1.1.

On average, fucosylated structures are thought to comprise 50-80 % of all HMOs in mature human milk, with sialylated HMOs comprising 10-20 % [89, 90, 92-97, 102, 103]. These values, however, are a composite of data obtained from a range of individual studies in various geographical locations, and don't reflect the extent of inter-individual variations in fucosylation and sialylation. The level of fucosylation of HMOs corresponds to the mother's Lewis blood group status [95, 104-107]. Women expressing the enzyme fucosyltransferase 2 (FUT2), which catalyses the addition of Fuc via an α1-2 linkage to Lewis blood group epitopes and HMOs, are known as Secretors, and produce high levels of α1-2-fucosylated HMOs in their breastmilk [104]. Likewise, women expressing the enzyme fucosyltransferase 3 (FUT3) are known as Lewis-positive, and produce high levels of α1-3/4-fucosylated HMOs [105, 108]. The milk of non-Secretors and/or Lewis-negative mothers contains little to no α1-2-fucosylated and/or α1-3/4-fucosylated HMOs, respectively [104, 105], and accordingly, HMO composition of Secretor +/- Lewis +/- women varies greatly [95, 106, 107]. Typically, α1-2-fucosylated HMOs is considerably more prevalent than α1-3/4-fucosylated HMOs [91]. Subtle variation in sialylation does also occur, though it has been suggested that this is not correlated to genotype, yet is due to inter-individual variation in expression patterns of the metabolic machinery required for sialylation [88]. The various concentrations of individual HMO structures thus does vary between individuals, although recent work has demonstrated patterns of different HMO profiles more common to specific geographical regions and

ethnicities [91]. For example, roughly 70 % of Caucasian women are Secretors [88].

Once ingested, HMOs resist degradation in the upper gastrointestinal tract by gastric acid and pancreatic enzymes *en route* to the distal small intestine and colon [109, 110]. As they do not serve any direct nutritional function to humans (since they are not metabolised by the infant), it is in the gut that HMOs perform their beneficial activities for the neonate, as outlined in the two sections below. The vast majority of HMOs are either metabolised by infant gut microbes, or excreted intact in the faeces and urine [111-114].

1.3.3 HMOs as the archetypal prebiotic. The best-studied beneficial HMO function is its role as a prebiotic. HMOs meet all criteria necessary for their definition as a prebiotic [71, 76], encouraging the growth of particular, beneficial microbes in the infant gut. Only a specific set of bacteria are capable of successfully metabolising HMOs, and using them as a substrate for growth, and HMOs are therefore believed to select a specific infant-type microbiota through the provision of a significant metabolic competitive advantage [84, 87, 115, 116]. HMOs, when provided as the sole carbohydrate source, have been demonstrated to encourage the growth of a number of species of *Bacteroides* [117] and *Bifidobacterium* [118-120]. It is this 'bifidogenic' effect, in particular, that is thought to serve as the main driver for the dominance of bifidobacteria in the breastfed neonatal gut microbiota [24]. The mechanisms and implications of this bifidogenic effect will be discussed later. Interestingly, a recent study demonstrated the powerful microbiota-determining effect of HMOs, where it was found that different HMO compositions were in fact correlated with distinct microbiota populations in breastfed infants [121].

1.3.4 Other beneficial properties of HMOs. One of the other important functions of HMOs in shaping the neonatal gut microbiota and benefiting the health of the infant host is that they can prevent or reduce the chance of infection caused by pathogens. As well as competitively excluding

pathogenic microbes through the encouragement of non-pathogenic commensals, HMOs may also prevent infection through direct interaction with pathogens [122]. HMOs are known to prevent adhesion of pathogenic bacteria, protozoans and viruses to the mucosal surfaces, by acting as decoy receptors for pathogen binding, thus preventing colonisation [123, 124]. This anti-adhesive mechanism has been demonstrated to prevent mucosal adhesion by bacterial pathogens such as *Campylobacter jejuni* [125, 126], protozoan pathogens such as *Entamoeba histolytica* [127] and possibly viral pathogens such as human immunodeficiency virus (HIV) [128]. In fact, another study showed the ability of some strains of the common mastitis-causing bacterium *Streptococcus* to bind the HMO 2'-fucosyllactose (2-FL) [129], indicating the potential for HMO to act as a natural anti-mastitis agent in lactating mothers. HMOs have also been demonstrated to function as antimicrobials, preventing the proliferation of pathogens, such as Group B *Streptococcus* [130] and *Candida albicans* [131].

Apart from their prebiotic function, another indirect method of prevention of infection possessed by HMOs is their ability to modulate host cellular responses, such as intestinal epithelial cell apoptosis, proliferation and differentiation [132]. Furthermore, HMOs have been shown to induce changes in the host intestinal epithelial glycocalyx [133], which may prevent adhesion of pathogens to this surface. Immune cells have also been shown to be affected by HMOs [134, 135], thus suggesting immunomodulatory effects that may benefit the infant. Finally, the HMO component sialic acid has been implicated in brain development and cognition in infants [136].

1.3.5 Potential for HMOs as nutritional supplements and clinical therapeutics for infants. Given the benefits that HMOs provide to the infant gut microbiota and, directly or indirectly, the infant itself, the use of these glycans for therapeutic and/or commercial seems to have a lot of potential. As prebiotics, HMOs shape the overall infant gut microbiota and its associated interactions, and thus may present a more effective intervention compared to single-strain or multi-strain probiotics. As a therapeutic, HMO

components are currently enjoying interest as a possible treatment for NEC in infants. Breastfed infants are 6-10 time less likely to develop NEC than their formula-fed counterparts [137-139], and more recent work has shown promise for the prevention of NEC by individual HMO components such as 2-FL and disialyllacto-N-tetraose in both rat and murine models [140-143].

The use of HMOs as (part of) commercially available supplements also presents an attractive prospect. The inclusion of non-human oligosaccharides, such as GOS, fructo-oligosaccharides (FOS) and inulin, in supplements and formula milk for infants, for the modulation of infant gut microbiota is a wide-spread practice [144, 145]. However, the evident structure-specific effects of HMOs mean that the benefits conferred by these structurally distinct glycans cannot fully mimic the beneficial effects of HMO on the infant. For this reason, the inclusion of actual HMO structures in supplements or formula milk may at some point become feasible. Recent work has demonstrated the safety of the inclusion of specific HMOs in infant formula [146], and even observed lower inflammatory cytokines in infants that had been receiving formula milk supplemented with 2-FL, as compared to that of formula-fed controls, and similar to that of breastfed infants [147]. The main prohibitive factor in the advancement of this technology, however, is the present inability to artificially (chemically or biotechnologically) synthesise a wide range of HMO glycans in sufficient quantities, although this is becoming a more tangible possibility, with a number of companies moving into this area. Furthermore, we currently do not know if the administration of high quantities of individual HMO glycans in (infant) humans, as opposed to the heterogeneous mix of glycans naturally found in breastmilk, is associated with any health risks [2].

1.4 Infant-Associated Bifidobacteria

1.4.1 General features of bifidobacteria. As mentioned above, the infant gut microbiota is heavily enriched in species of *Bifidobacterium* [13, 22], particularly in healthy, full-term, breastfed infants [21, 23-26]. First isolated from the faeces of a breastfed infant by Tissier in 1899, bifidobacteria are Gram-positive, saccharolytic, typically Y-shaped anaerobes, whose chromosomes have a high G-C content. Belonging to the *Bifidobacteriaceae* family and the Actinobacteria phylum [148-151], the genus *Bifidobacterium* at present includes 59 different taxa [2, 45, 151]. These species fall within 7 phylogenetic clusters; namely (I) the *B. asteroides* group, (II) the *B. pseudolongum* group, (III) the *B. longum* group, (IV) the *B. bifidum* group, (V) the *B. adolescentis* group, (VI) the *B. pullorum* group and (VII) the *B. boum* group [150]. Species of bifidobacteria can also be clustered into seven ecological niches of origin, namely the human GIT, human blood, human oral cavity, non-human mammals, birds, social insects and wastewater [152]. As of December 2017, there are 56 complete bifidobacterial genome sequences in the database of the National Centre for Biotechnology Information (located at the following website: https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi), representing 40 species. Bifidobacterial genomes typically possess a G-C content that varies between 58 and 61 %, ranging in size from 1.73 to 3.16 Mb, the largest being that of *B. scardovii* JCM 12489 [153-155].

A recent study of the distribution of bifidobacterial species in a wide range of hosts across the mammalian branch of the tree of life revealed a high prevalence (>85 %) of 4 *Bifidobacterium* species across the 67 mammalian taxa studied [156]. These four highly ubiquitous species, namely *B. longum*, *B. adolescentis*, *B. bifidum* and *B. pseudolongum*, are also commonly associated with the gut microbiota of humans. Despite the presence of bifidobacteria in both infant and adult microbiota, there is considerable variation between the typical species make-up of either [157, 158]. Following the depletion of oxygen in the infant gut by facultative anaerobes, bifidobacteria become dominant in the associated gut microbiota [159]. Their relative abundance decreases as the infant develops and the microbiota

progresses from an infant-type to an adult-type [158, 160]. The abundance of bifidobacteria in the adult gut microbiota remains stable at a relative abundance of between 2 % and 15 % [158]. A number of studies have observed that the most prevalent species of *Bifidobacterium* in the adult gut microbiota is *B. longum* subsp. *longum* [161, 162]. While a study by Matsuki *et al*. also detected the high prevalence of this species in the microbiota of healthy adults, they observed a greater abundance of the species *B. adolescentis* and *B. catenulatum* [162]. Levels of *B. breve* and *B. bifidum*, which are among the most abundant species in the infant gut microbiota (see below), were observed to markedly decrease as hosts progressed in age, and their microbiota developed into an adult-type [163]. The bifidobacterial component of the gut microbiota generally remains stable throughout adulthood [158, 160], but a general decrease in the relative abundance of *Bifidobacterium* species has been observed in elderly individuals [164-167]. A multitude of purported benefits are correlated to the presence of bifidobacteria in the adult host, ranging from management of inflammatory gut conditions, to improved neurocognition, to potential antitumor effects [168-170], and as such, a number of adult-targeted probiotic supplements containing bifidobacteria are now commercially available.

1.4.2 Bifidobacteria in the infant gut. The infant gut microbiota typically possesses a significantly higher relative abundance of bifidobacteria than that of the adult [24, 161]. As with the entire infant gut microbiota, the colonisation and establishment of bifidobacteria in this niche is dependent on several extrinsic factors. Vertical transmission from the mother is believed to be an important contributor to the bifidobacterial component of the neonatal gut microbiota, with a number of maternal sources implicated in this process, including the vagina, faeces, breast milk, amniotic fluid and placenta [171-173]. Significantly higher (relative abundance) levels of bifidobacteria have been detected in vaginally-born infants as compared to their C-section-born counterparts [31]. As mentioned above, considerably lower levels of *Bifidobacterium* species have been observed in pre-term as compared to full-term infants [13, 42-49, 174].

Perhaps the most significant factor in the establishment and persistence of high levels of bifidobacteria in the infant gut microbiota is breastfeeding. In general, the most abundant species of *Bifidobacterium* in the neonatal gut microbiota are *B. longum*, *B. breve* and *B. bifidum* [24, 159, 161], and to a lesser extent *B. pseudocatenulatum* and *B. adolescentis* [24, 54, 175-177]. *B. kashiwanohense* is also considered an infant-associated species, but appears to be much less frequently detected, with only two strains classified at present [178, 179]. Evidence exists for the correlation of breastfeeding and formula-feeding with the abundance of particular bifidobacterial species in the infant [26]. This is likely due to the 'bifidogenic' effect of breastmilk, in particular HMOs, which only specific infant-associated species of *Bifidobacterium* are capable of metabolising. *B. breve*, *B. bifidum*, *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* have been detected in both breastfed and formula-fed infants [180, 181], Interestingly, one study found that *B. longum* subsp. *infantis* was more abundant in breastfed infants, whereas *B. longum* subsp. *longum* was found to be more abundant in formula-fed neonates [54]. Additionally, *B. adolescentis*, a typical adult-associated species, has been identified in the faecal microbiota of formula-fed infants, but appears to be rarely found in that of their breastfed counterparts [181], an observation consistent with the notion that an adult-type microbiota develops earlier in formula-fed infants as compared to age-matched breastfed infants [34]. Interestingly however, a very recent study, on inheritance of maternal bifidobacteria in infants by vertical transmission, observed the abundance of typically adult-associated *Bifidobacterium* species *B. adolescentis* and *B. catenulatum* in infant faecal samples [177]. This puts into question the currently-accepted notion of typical infant-associated and adult-associated species, though it would be interesting to assess the feeding status of the infants in question, which may have a bearing on these results.

1.4.3 Health benefits of bifidobacteria in the infant host. As the healthy infant gut microbiota is heavily enriched for bifidobacteria, it is hardly surprising that this bacterial group is thought to provide various benefits to the infant host. From a metabolic perspective, the microbiota, and thus

bifidobacteria as a dominant component in the case of infants, may act as a virtual organ, degrading otherwise non-digestible carbohydrates, thereby releasing metabolites such as short-chain fatty acids (SCFAs) [182, 183], which can be absorbed by and benefit the host [184-186]. Bifidobacteria have also been demonstrated to exhibit antimicrobial activity against pathogens such as *Salmonella typhimurium* and *Listeria monocytogenes* [187, 188], and have been shown to inhibit growth of *Enterobacteriaceae* within colicky infant microbiota cultures [189]. Through interactions with host cells, bifidobacteria are also thought to benefit the infant host with anti-inflammatory activity; growth *B. longum* subsp. *infantis* and *B. bifidum* in the presence of HMOs has been shown to promote the expression of anti-inflammatory cytokines in eukaryotic cells [190], while *B. longum* subsp. *infantis* alone has been demonstrated to reduce intestinal permeability [191].

Studies using murine models have demonstrated the potential for the use of bifidobacterial strains as probiotics with reported effects on coeliac disease [192], obesity [193] and infection-induced psychology [194]. Human studies involving treatment with bifidobacteria have also reported benefits in infants [195], with effects ranging from reduced allergic responses [196], the prevention of NEC and infant colic [197-199], to the prevention of diarrhoea [197] and improved enteral feeding in low birth-weight neonates [200, 201], all supporting the potential of bifidobacteria as a bio-therapeutic that can be used for a variety of disorders/diseases/conditions.

1.5 Metabolism of Human Milk Oligosaccharides by Infant-Associated Bifidobacteria

1.5.1 Carbohydrate metabolism in bifidobacteria. Perhaps the most useful tool at the disposal of bifidobacteria in order to successfully colonize and persist in the human gut is their ability to metabolize a wide array of both plant- and host-derived glycans [202]. The carbohydrate-containing portion of the human diet can be divided into two general categories in the context of digestibility in the GIT: (i) simple sugars, absorbed in the host intestine directly or following hydrolysis by host-coded enzymes; and (ii) complex carbohydrates of plant (e.g cellulose, arabinoxylan) or human/animal (e.g. HMO, glycogen, mucins) origin that are resistant to degradation by host-encoded enzymes and absorption [74, 88, 203-205]. Thus, whether of host-, carer-, or diet-derived origin [206], a plethora of carbohydrates reach the small and large intestine and will then be metabolized by elements of the microbiota. It is believed that this abundance of carbohydrate substrates allows bifidobacteria to thrive, owing to their dynamic and highly specialised saccharolytic metabolism.

Bifidobacterial carbohydrate metabolism depends on systems to internalize the carbohydrates, while hydrolytic abilities are also required, being present as intra- and/or extracellular enzymes. A recent study of predicted proteins in different bifidobacterial species found that 13.7 % of identified cluster of orthologous groups (COGs) were associated with carbohydrate metabolism, across the 47 (sub)species assessed, whereas carbohydrate metabolism-associated COGs were only found to comprise 8.0 % of the COGs for the entire gut microbiome [202]. Interestingly, in this same study, the core carbohydrate metabolism-associated COGs were found to represent just 5.5 % of all identified bifidobacterial COGs, suggesting strong selective pressure for the acquisition of accessory genes necessary for the utilisation of specific carbohydrates, allowing a species of *Bifidobacterium* to thrive within a particular niche.

While most gut bacteria rely on the glycolytic pathway for energy generation when metabolising carbohydrates [207], bifidobacteria utilise the fructo-6-phosphate phosphoketolase pathway (also known as the Bifid Shunt) as their

central catabolic pathway [208, 209]. Fructose-6-phosphoketolase (F6PPK) is the key enzyme in this pathway, and its presence can be used as a signature property for members of the *Bifidobacterium* genus [210]. This pathway yields 2.5 ATP molecules for every mole of glucose, making it more efficient in energy production than carbohydrate fermentation pathways of other gut commensals, including lactic acid bacteria [211]. Fermentation of 1 mole of hexose sugar via the Bifid Shunt theoretically yields 1.5 moles of acetate and 1 mole of lactate, while the fermentation of pentose sugars typically yields one mole each of acetate and lactate, although factors such as carbon source, cell growth phase and external pH can affect these ratios [207, 211]. This doesn't, however, account for the energy cost of the internalisation of carbohydrate substrates. The end products of other monosaccharide catabolism pathways in Bifidobacteria, such as the Leloir pathway (glucose-6-phosphate) and the amino-sugar metabolism pathway (fructose-6-phosphate), are subsequently shuttled through the central catabolic pathway of the Bifid shunt [208].

In nearly all bifidobacteria, carbohydrate transport is predominantly mediated using ATP-requiring ATP-binding cassette (ABC) transporters, proton symporters and proton-motive force-driven permeases, with energy cost-neutral phosphoenolpyruvate-phosphotransferase systems (PEP-PTS) being much less common than other transport systems [212-215], as exemplified in *B. longum* subsp. *longum* NCC2705 [216]. This may be due to the metabolic preferences of bifidobacteria for the more abundantly available complex carbohydrates in the intestinal niche, as ABC transporters have been demonstrated to transport a variety of oligosaccharides [217], whereas PEP-PTS systems are normally only associated with the transport of mono- and certain di-saccharides [218, 219]. Interestingly, this notion is supported by the finding that *B. bifidum* PRL2010 atypically (for bifidobacteria) possesses more PEP-PTSs than ABC transporters, however, since *B. bifidum* extracellularly hydrolyses complex HMO structures, members of this species internalize the released mono- and di-saccharide constituents (see for further details and references below).

As the majority of the carbohydrates available to bifidobacteria in the intestine arrive in the form of di-, tri- or more complex oligo-saccharides, a crucial family of proteins are those which degrade these carbohydrates into their constituent monosaccharides, known as glycosyl hydrolases (GHs). Found in all three kingdoms (Archaebacteria, Eubacteria and Eukaryota), GHs hydrolyse the glycosidic bond between monosaccharides [220, 221], of which there are 148 families, listed in the CAZy database (available at http://www.cazy.org/Glycoside-Hydrolases.html) [221, 222]. A recent study of the carbohydrate-active enzymes encoded by the *Bifidobacterium* pangenome revealed the presence of proteins belonging to 57 GH families encoded among 47 species [202]. *B. scardovii*, *B. biavatii*, *B. saeculare* and *B. dentium* were each found to possess a GH index (the number of GHs predicted in each genome normalized by genome size) that is higher than the bifidobacterial average. It has previously been suggested that the large GH arsenal encoded by *B. dentium* Bd1 is due to the availability of a wider array of complex carbohydrate sources in its niche, the oral cavity [214]. While predicted GHs required for the hydrolysis of both plant-derived and human/host-derived glycans were detected across the pangenome, those associated with the degradation of human-derived glycans were found to be particularly prevalent in *B. longum* subsp. *infantis* and *B. bifidum* [202]. These include GH33 exo-sialidases, GH29 and GH95 fucosidases, GH20 hexosaminidases and lacto-N-biosidases, GH20 lacto-N-biosidases, GH38 and GH125 α-mannosidases, and GH101 and GH129 α-N-acetylgalactosaminidases. This is not surprising, as both of these species are commonly infant-associated, and it would follow that they preferentially target the human-derived carbohydrates present in breastmilk for metabolism (see below).

1.5.2 Transcriptional regulation of carbohydrate metabolism in bifidobacteria. As with other bacteria, bifidobacteria appear to employ the regulatory process known as carbon catabolite repression (CCR) to, when presented with multiple carbon sources, preferentially utilise the substrate yielding the greatest amount of energy, while inhibiting the expression or

activity of proteins involved in the uptake and catabolism of other substrates [223, 224]. This typically occurs either via the inhibition of expression of genes involved in the uptake or utilisation of secondary substrates [225], or via the inhibition of proteins involved in the uptake of secondary substrates [223, 226].

The phenomenon of CCR is of vital importance to (bifido)bacterial metabolism, particularly those found in niches such as the intestine, where multiple carbohydrate sources are likely to be available, and utilisation of the most energy-efficient substrates is key to remain competitive, and establish and maintain a stable population. So far in bifidobacteria, CCR appears to predominantly occur by the inhibition of expression of genes involved in secondary substrate uptake and catabolism. Down-regulation of expression of genes involved in fructose metabolism was observed in *B. breve* UCC2003 grown in the presence of ribose and/or glucose [227]. Likewise, *B. longum* subsp. *longum* NCC2705 was shown to preferentially consume lactose over glucose, with the downregulation in expression of a glucose-specific permease [228], and *B. longum* subsp. *infantis* ATCC15697 was demonstrated to downregulate transcription of two fucosidase-encoding genes (and the concurrent upregulation of another) when grown on a complex mixture of HMOs as opposed to lactose [229].

These transcriptional regulatory processes are presumed to be mediated by the activity of transcription factors (TFs), which can act as transcriptional activators and/or repressors. Bacterial TFs are typically DNA binding proteins which bind to an operator sequence in the region around a regulated promoter, which impedes (or promotes) the binding of RNA polymerase, and thus the initiation of transcription. This is best characterised for the LacI regulator protein, which represses transcription of the *lac* (lactose uptake and utilisation) operon in *E. coli* [230, 231]. Binding of an effector molecule to the repressor protein induces a conformational change in the repressor, which prevents its binding to the operator sequence, and thus allowing transcription to proceed. In the case of the *lac* operon from *E. coli*, the inducer molecule is allolactose, a product of the β-galactosidase activity on lactose [231]. This ensures the expression of the *lac* operon only when the breakdown product of

lactose, and thus lactose itself is present [232]. Multiple LacI-type regulators are found in *E. coli*, functioning in the same manner, but with distinct specific operator sequences, target promoters and effectors [233-235].

LacI-type TFs are highly common in bifidobacteria, and in fact are more prevalent in *Bifidobacterium* species than any other genus in the Actinobacteria phylum [236]. This abundance of encoded LacI-type repressors in bifidobacterial genomes is believed to enable rapid adaptation of these species to changes in available carbohydrate sources [215]. However, a number of TFs type for other regulator types, including ROK-type, RpiR-type, DeoR/SorC-type, BglG-type, TetR-type, AraC-type and GntR-type regulators, have been identified in bifidobacteria (although their involvement in the control of carbohydrate-related metabolic pathways has not been established yet for all) [237].

In bifidobacteria, transcriptional repression is best studied in *B. breve* UCC2003. To date, seven predicted LacI-type regulators have been characterised in UCC2003: $LacI_{fos}$, regulating transcription of the *fos* operon [238]; GalR, regulating transcription of the galactan utilisation operon [239]; CldR, regulating cellodextrin utilisation [240]; RbsR, regulating transcription of the ribose utilisation cluster [227]; MelR1 and MelR2, regulating transcription of the melezitose utilisation cluster [241]; and AtsR1, regulating utilisation of N-acetylglucosamine-6-sulfate (GlcNAc-6-S) [242]. A review of bifidobacterial LacI regulators found that 69 % of the LacI-type TFs studied regulate one or two operons, with 90 % of the investigated LacI-type TFs representing local regulators, controlling a single metabolic pathway, usually associated with the metabolism of one carbohydrate (where it involves more carbohydrates, these are then related in terms of monosaccharide identity and glycosodic linkage) [236]. It was also found that 20 % of the operons regulated by LacI-type TFs contain multiple operator sites, and 75 % of the LacI-type TF binding motifs are located between 30 and 140 bp upstream of the closest regulated gene. These findings agree with the experimentally characterised LacI-type repressor systems from *B. breve* UCC2003 (see above). The DNA binding motifs for the LacI operator sequences were found to share the central CG pair in the centre of the binding

sites for other LacI-type TFs [236, 243], with the only exception being the predicted binding site of $LacI_{fos}$ [238].

However, LacI-type regulators are not the only repressors found in *B. breve* UCC2003. NanR, a GntR-type repressor, was identified as regulating utilisation of sialic acid [244], while AtsR2, a ROK-type repressor, was also implicated in the regulation of GlcNAc-6-S utilisation [242]. Furthermore, a LacI-type repressor, ROK-type repressor and an N-acetylglucosamine (NagC)-type repressor have all been implicated in the regulation of HMO utilisation in UCC2003 (see Chapter III of this thesis) [245]. In fact, a recent study analysing transcriptional regulons in 10 *Bifidobacterium* genomes revealed 268 predicted TFs belonging to the LacI, ROK, DeoR, AraC, GntR and TetR regulator families, forming 64 orthologous groups of regulators [237]. Transcription factors for LacI-type and ROK-type regulators were found to be the most common, however. Interestingly, this study predicted a LacI-type regulator, AraQ, as a global regulator for genes involved in central carbohydrate metabolism in bifidobacteria.

Despite the abundance of repressors studied in *B. breve* UCC2003, as of yet only one carbohydrate utilisation-associated transcriptional regulator acting as an activator has been characterised in bifidobacteria: a ROK (repressor open reading frame kinase)-type regulator RafR, which was found to activate transcription of the raffinose-related sugar utilisation cluster [241].

1.5.3 HMO metabolism by infant-associated bifidobacteria as an adaptation to the infant gut. It is hardly surprising that the wide-ranging carbohydrate-utilising capabilities of bifidobcateria extend to the utilisation of milk-derived glycans, as they are among the most abundant species of the infant gut, particularly in breastfed neonates [24]. As already mentioned, the repertoire of carbohydrates metabolised by bifidobacteria includes glycans of both plant and human origin [202], the latter including HMOs. Thus, it would follow that the bifidobacterial species most commonly found in the gut microbiota of breastfed infants possess metabolic capabilities for the degradation and utilisation of HMO structures. A number of studies have

assessed the ability of a range of *Bifidobacterium* species to consume various individual and combined HMO glycans as a growth substrate, and found that the only species demonstrating significant growth on these sugars were those commonly associated with the infant gut microbiota, namely *B. bifidum*, *B. breve*, *B. longum* subsp. *infantis* and, on occasion, *B. longum* subsp. *longum* [118-120, 246-248]. More studies have also demonstrated the ability of *B. kashiwanohense* isolates to utilise fucosyllactose as a growth substrate [249, 250], and recently species occasionally found in the infant gut microbiota, such as *B. pseudocatenulatum* and *B. longum* subsp. *suis*, have been shown to utilise certain HMO substrates to a limited degree [177, 249]. These studies present strong evidence for the utilisation of HMOs by infant-associated bifidobacteria as a highly specialised adaptation to the environment of the breastfed infant gut [86, 251]. A recent study in fact correlated the presence of specific sets of GHs and transporters encoded in the genomes of infant-associated bifidobacteria with each strain's ability to consume various HMO substrates [177], further supporting the notion of HMO utilisation as a highly specific adaptation by infant-associated species of *Bifidobacterium*.

In particular, the utilisation of the HMO component lacto-N-biose (LNB) is viewed as a strong indicator of overall HMO consumption by bifidobacteria [252]. The pathway for the utilisation of galacto-N-biose (GNB) and LNB has only been identified in infant-associated bifidobacteria, including *B. breve*, *B. bifidum*, *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* [252-256]. This pathway, whose central enzyme is a GH112 GNB/LNB phosphorylase (GLNBP), functions in the degradation of LNB into glucose-1-phosphate and N-acetylglucosamine, which then enter the energy-generating Bifid Shunt and amino-sugar metabolising pathways, respectively [257]. Despite the relatively low abundance of free LNB in HMO [91], LNB can be found as a component of many larger HMO structures, especially due to the aforementioned predominance of Type 1 over Type 2 HMOs [98]. Thus, the presence of the GNB/LNB pathway in bifidobacterial species can be viewed as a proxy for HMO utilisation in some form, and perhaps an indicator of an infant-associated species, as has been previously suggested in the 'LNB hypothesis' [255]. This, however, may be something of an oversimplification,

as different species of infant-associated bifidobacteria utilise considerably varying strategies for HMO utilisation, targeting distinct HMO glycans and their components, as will be discussed below.

1.5.4 Strategies for HMO utilisation by infant-associated bifidobacteria vary. The selective consumption of various HMO components by different species of infant-associated bifidobacteria explains the abundance and prevalence of a number of different species in the same niche [258]. While all of the species share the same central metabolic pathway, differences in their catabolic capabilities of higher substrate structures creates a form of metabolic partitioning, enabling them to target different HMO components. Excluding the work presented in this thesis, the species best characterised for their strategies of HMO utilisation are *B. bifidum* and *B. longum* subsp. *infantis*, as well as some limited knowledge of HMO metabolism in *B. longum* subsp. *longum* [257, 258]. Within the context of this thesis, the metabolism of HMO components LNT, LNnT and LNB in *B. breve* UCC2003 is described in detail (Chapter II, Chapter III) [259], as well as the elucidation of HMO metabolism pathways in *B. kashiwanohense*. In the following section, the current knowledge of HMO metabolism in these infant-associated species is discussed. A simplified schematic representing the (currently) known metabolic strategies for HMO utilisation by *B. bifidum*, *B. breve*, *B. longum* subsp. *infantis* and *B. kashiwanohense* is shown in Figure 1.2.

1.5.5 HMO metabolism by *Bifidobacterium bifidum*. *B. bifidum* typically targets complex HMO structures, carrying out the extracellular hydrolysis of these HMOs, including LN(n)T and fucosylated and sialylated structures, prior to the import and metabolism of (most of) the resultant mono- and di-saccharides [258]. This extracellular hydrolysis is enabled by the expression of an array of secreted GHs, including α-sialidases, α-fucosidases, β-galactosidases, lacto-N-biosidases and N-acetylhexosaminidases [257, 260, 261]. *B. bifidum* extracellularly releases fucose and sialic acid residues from complex fucosylated and/or sialylated HMOs through the activity of secreted

fucosidases and sialidases, respectively. *B. bifidum* strains have been shown to express 2 extracellular α-fucosidases; one a GH95 enzyme targeting alpha-1,2-linked HMOs (e.g. 2-FL) and one a GH29 enzyme targeting □-1,3/4-fucosylated HMOs (e.g. 3-FL), respectively, liberating fucose [262-264]. Likewise, a GH33 α-sialidase expressed by *B. bifidum* has been shown to liberate sialic acid residues from sialylated HMOs (e.g. 3/6-sialyllactose) [265, 266]. Interestingly, *B. bifidum* was shown not to utilise fucose or sialic acid residues [118, 119, 267], which can instead be consumed by other members of the microbiota. A recent study has indeed demonstrated the ability of *B. breve* UCC2003 to cross-feed on the sialic acid released by the degradation of sialyllactose by *B. bifidum* PRL2010 [268].

Whether free, or released by the hydrolysis of fucosylated/sialylated HMOs, the tetrasaccharides lacto-N-tetraose (LNT) and lacto-N-neotetraose (LNnT) are also extracellularly hydrolysed by *B. bifidum*. LNT is hydrolysed at its central β-1,3-link by an extracellular GH20 lacto-N-biosidase into lactose and LNB, the latter being transported into the cell and then degraded by two distinct LNB phosphorylases (LNBP), releasing galactose 1-phosphate and GlcNAc, which enter the Leloir and amino-sugar metabolising pathways, respectively [269] (and then routed through the Bifid Shunt). An extracellular GH42 β-galactosidase cleaves LNnT at its Galβ-1,4 residue, liberating galactose and lacto-N-triose [270]. The lacto-N-triose is then further hydrolysed by an extracellular GH20 N-acetylhexosaminidase, releasing GlcNAc and lactose. Even longer HMO structures are similarly hydrolysed by β-galactosidases in a sequential fashion, releasing constituent mono- and di-saccharides which *B. bifidum* can successfully consume [271]. Whether released by the hydrolysis of fucosyllactose, sialyllactose or more complex structures such as LN(n)T, liberated lactose is either degraded by extracellular β-galactosidases (of GH42 or GH2 family) and the resulting glucose and galactose internalised, or the lactose is internalised intact and hydrolysed by intracellular β-galactosidases (of the GH2 family) [271]. The resulting galactose and glucose then enter the Leloir and Bifid Shunt Pathways, respectively.

1.5.6 HMO metabolism by *B. longum* subsp. *infantis*. *B. longum* subsp. *infantis* internalises particular, intact small-mass HMOs, and degrades these then into monosaccharides for further metabolic processing [120, 258]. Studies on bifidobacterial consumption of various HMOs have shown the ability of *B. longum* subsp. *infantis* to consume small-chain HMOs structures, typically DP≤8 (i.e. composed of ≤8 monosaccharide units) with a molecular mass <1400 Da [120], such as sialyllactose, fucosyllactose and LN(n)T (as well as their fucosylated/sialylated versions) for growth [118-120, 177]. These structures are internalised by *B. longum* subsp. *infantis* [272], and then degraded into their constituent monosaccharides by various GHs [257]. Fucosylated HMOs are hydrolysed at their fucosyl linkages by a number of both GH29 and GH95 α-fucosidases, liberating L-fucose [229]. Sialic acid residues are liberated from sialylated HMOs by the hydrolytic activity of two GH33 α-sialidases [248]. Liberated sialic acid and L-fucose are both further metabolised by *B. longum* subsp. *infantis* for energy generation [250, 273].

Free LN(n)T is internalised by the cell, or may be released by the intracellular hydrolysis of other complex, yet internalized HMOs, as described above. LN(n)T is then degraded sequentially from its non-reducing end in a similar fate to LNnT in *B. bifidum* metabolism [257]. A GH42 β-galactosidase hydrolyses the Galβ1-3GlcNAc residue at the non-reducing end of LNT, releasing galactose and lacto-N-triose [274]. Likewise, both the aforementioned GH42 β-galactosidase and a GH2 β-galactosidase hydrolyse the Galβ1-4GlcNAc residue at the non-reducing end of LNnT, releasing the same products [274]. Galactose enters the Leloir pathway, while lacto-N-triose is hydrolysed by a GH20 N-acetylhexosaminidase into N-acetylglucosamine (which enters the amino-sugar metabolic pathway) and lactose [275]. Lactose is then hydrolysed by β-galactosidases into glucose and galactose, destined for the Bifid Shunt and Leloir pathways, respectively [274]. Free LNB, if present in the environment, is taken up by *B. longum* subsp. *infantis*, and is phosphorolysed by a GH112 LNBP enzyme, releasing galactose and N-acetylglucosamine [252]. These versatile HMO utilisation capabilities of *B. longum* subsp. *infantis* are enabled by the presence of a large 43-kb gene cluster within its genome [272, 273]. This ‘HMO-utilisation

island' encodes the two fucosidases, one of the sialidases, the LN(n)T-specific β-galactosidases and N-acetylhexosaminidase mentioned above. While no intact equivalent to this locus has been identified in other *Bifidobacterium* species [276], segments encoding the genes for fucosyllactose and/or fucose utilisation have been identified in *B. breve*, *B. longum* subsp. *suis*, *B. longum* subsp. *longum* and *B. kashiwanohense* [249, 250, 277].

1.5.7 HMO utilisation by *Bifidobacterium breve*. *B. breve* internalises intact neutral HMO moieties such as LNT, LNnT and LNB, and sequentially hydrolyses these oligosaccharides for the subsequent metabolism of their monosaccharide constituents [259] (Chapter II of this thesis). Previously, it was thought that *B. breve* simply behaved as a scavenger, 'mopping up' HMO-derived monosaccharides released by the extracellular hydrolysis of larger HMO structures by other members of the infant gut microbiota [86]. This data was supported by the finding that many strains of *B. breve* studied do not grow on fucosylated or small-mass sialylated HMOs, including fucosyllactose and sialyllactose [118-120], and that *B. breve* is capable of cross-feeding on sialic acid and fucose released from the extracellular hydrolysis of sialylated and fucosylated structures by *B. bifidum* [268, 278]. In a recent study, however, we demonstrated not only that *B. breve* UCC2003 is capable of internalising and metabolising small-mass neutral HMOs such as free LNT, LNnT and LNB for growth, but that the pathways for the utilisation of these three glycans in *B. breve* are in fact identical to those in *B. longum* subsp. *infantis* [259]. The mechanistics of this pathway will be described in detail in Chapter II [259], and the transcriptional regulation of the associated genetic loci will be described in Chapter III (manuscript submitted for review). As for its scavenging abilities, *B. breve* does indeed possess pathways for the uptake and utilisation of both free L-fucose [250] and free sialic acid [268].

Interestingly, recent work has also identified the ability of a small number of *B. breve* strains to utilise fucosylated HMO and longer-chain sialylated

HMOs for growth [247, 279]. In the strains tested, this growth seemed linked with the activities of a GH29 α-fucosidase and a GH33 α-sialidase, respectively. As homologs of both of these genes are in fact encoded in the genomes of most *B. breve* strains [177, 202, 247, 279], the inability of many *B. breve* strains to utilise fucosylated and/or sialylated HMOs may be due to the absence of other essential components for their metabolism, such as transporter proteins. Nonetheless, the adaptations of such *B. breve* strains for the consumption these glycans demonstrates the metabolic versatility of this species as a *bona fide* HMO-utiliser.

1.5.8 HMO metabolism in other *Bifidobacterium* species. While *B. longum* subsp. *infantis*, *B. bifidum* and *B. breve* are the most commonly found species of bifidobacteria found in the gut of the breastfed infant, there are, as mentioned above, other species which are associated with this niche, such as *B. longum* subsp. *longum* and *B. kashiwanohense*. The hypothesis of the bifidogenic effect of HMO strongly suggests the ability of such species to also utilise at least some HMOs for growth. The known HMO-metabolising capabilities of these species are outlined below.

Strains of *B. longum* subsp. *longum* generally grow poorly on HMO substrates [118-120, 247]. However, *B. longum* subsp. *longum* strain JCM1254 was shown to utilise LNB in the same fashion as *B. bifidum*, employing the activity of an extracellular GH136 lacto-N-biosidase [269, 280]. Additionally, *B. longum* subsp. *longum* strain SC596 was recently shown to utilise fucosylated HMOs, such as fucosyllactose, due to a gene cluster [277], which encodes the components necessary for the import of fucosylated structures, two α-fucosidases (GH29 and GH95) and the enzymes necessary for L-fucose metabolism. Both the observed utilisation of fucosylated HMOs in this strain and the cluster of genes responsible for this ability are highly similar to that employed by strains of *B. longum* subsp. *infantis*. These two distinct and apparently atypical adaptations in two different strains of *B. longum* subsp. *longum* highlight the versatility of

Bifidobacterium species in their carbohydrate metabolising capabilities to colonise and persist in a given niche.

Only recently have the capabilities of HMO metabolism in *B. kashiwanohense* begun to be explored and understood. Two studies in the last number of years have identified the ability of *B. kashiwanohense* to grow using fucosyllactose as a substrate, yet also observed the inability of the strains tested to utilise the liberated L-fucose [249, 250]. The authors suggested that only the lactose was internally degraded and consumed, and that the L-fucose was not utilised due to the absence of genes encoding key components of the L-fucose utilisation pathway. However, in Chapter IV (manuscript in preparation), we describe the ability of a novel *B. kashiwanohense* strain, isolated from the faeces of a breastfed infant, to utilise fucosyllactose for growth, but do not observe the accumulation of L-fucose in the media following growth. We characterise the internalisation of fucosyllactose by *B.* kashiwanohense, as well as its hydrolysis by two α-fucosidases (GH29 and GH95), and furthermore identify a putative pathway for the utilisation of L-fucose (Chapter IV), similar to that previously observed in *B. longum* subsp. *infantis* and *B. breve* [249, 250].

Little investigation has taken place of HMO metabolism by *B. pseudocatenulatum*. A recent study observed the ability of one strain to consume and grow on LNT [177]. Another study has indicated that the GH content of this species is rather high as compared with that of other *Bifidobacterium* species [202], which may indicate the ability to degrade a range of carbohydrates, possibly including certain HMOs. Further research on this species is therefore required.

1.6 Conclusion

The range of carbohydrate-metabolising capabilities of bifidobacteria is one of the key factors supporting the ability of these species to thrive in the various niches in which they are found. Whether this entails the metabolism of intact carbohydrate sources in the oral cavity by *B. dentium*, the degradation and consumption of indigestible plant-derived dietary polysaccharides and oligosaccharides in the adult gut by adult-type species, or the utilisation of HMOs in the gut of the breastfed neonate by infant-associated species, the repertoire of GHs and other carbohydrate utilisation pathways expressed by bifidobacteria is very specifically tailored to the niche they inhabit. In no group is this better exemplified than in the infant-associated species of *Bifidobacterium*. Not only do strains of *B. breve*, *B. longum* subsp. *infantis*, *B. bifidum* and *B. kashiwanohense* possess the capability to utilise (particular) HMOs, but they each employ specific strategies for the utilisation of different HMO components. Not only does this metabolic partitioning underpin the adaptability of these species to their niche, perhaps allowing them to co-exist in the same host, but it also displays their ability to take part in metabolic resource-sharing, as observed with *B. bifidum* and species such as *B. breve*. The overall metabolic adaptations of these species to the availability of HMO in the gut of breastfed infants, as well as the highly 'bifidogenic' effect of these glycans, provides strong evidence for the co-evolution of bifidobacteria and humans, and only serves to highlight their importance to the infant gut microbiota, and overall infant health and development.

1.7 Thesis outline

Chapter II of this thesis is comprised of work elucidating the mechanisms and components of LNT, LNnT and LNB utilisation pathways in *Bifidobacterium breve* UCC2003. Chapter III focuses on identifying the network responsible for the transcriptional regulation of LN(n)T and LNB metabolism, as described in Chapter II. In Chapter IV, *Bifidobacterium* isolates are obtained by screening of the faeces of breastfed infants, using HMO components fucosyllactose and sialyllactose as selective carbohydrates, and compared with isolates obtained using lactose or GOS. In Chapter V, a novel *B. kashiwanohense* isolate obtained during Chapter IV is demonstrated to consume fucosyllactose, and the mechanisms of its utilisation of both fucosyllactose and L-fucose is examined.

1.8 Figures

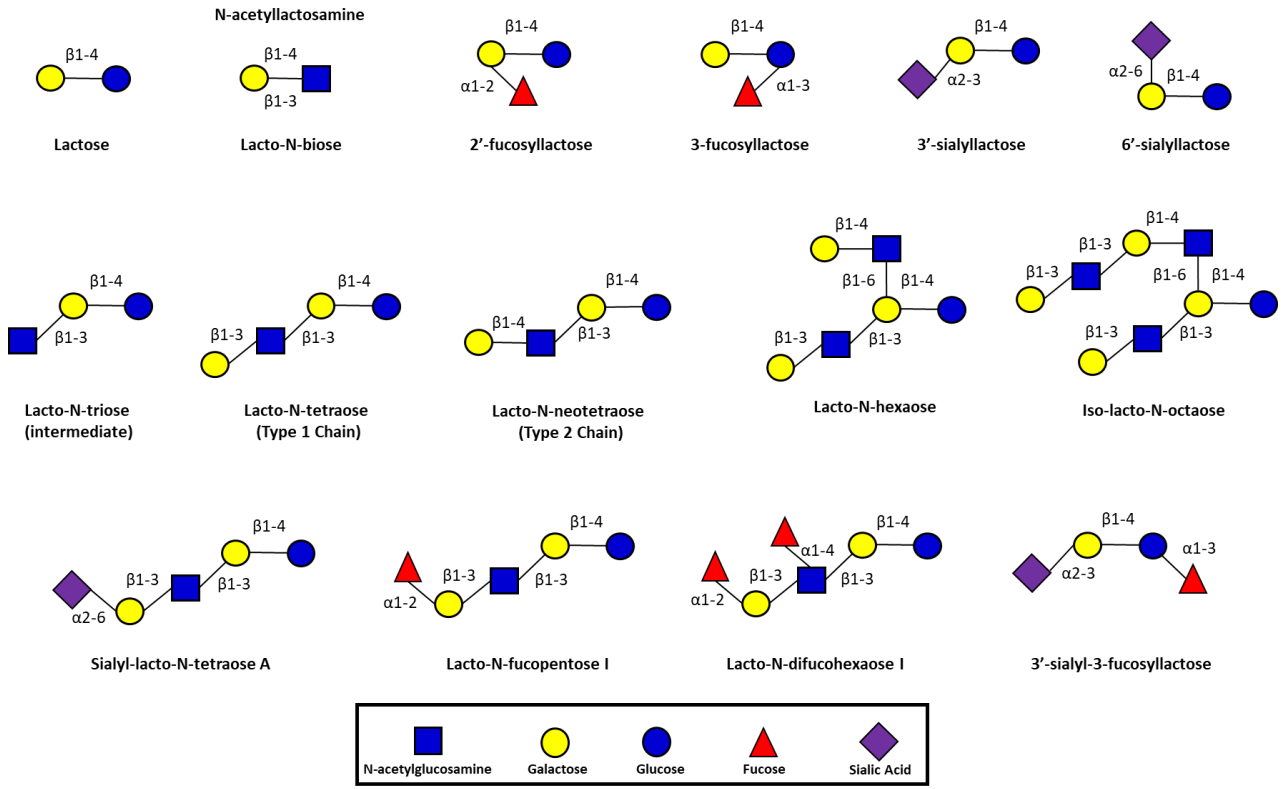


Figure 1.1. Schematic diagram of selected HMO structures. The addition of fucose (in a α1-2 or α1-3 linkage) or sialic acid (in a α2-3 or α2-6) linkage to lactose results in fucosyllactose or sialyllactose, respectively. Alternatively, lactose can be elongated with LNB (Type 1 structures) or N-acetyllactosamine (Type 2 structures). Further additions in a β1-3 linkage extends the chain (*para* HMO) or a β1-6 linkage introduces chain branching (*iso* HMO). Further fucose and/or sialic acid subunits may also be added to larger structures. Monosaccharide key is shown at the bottom of the figure.

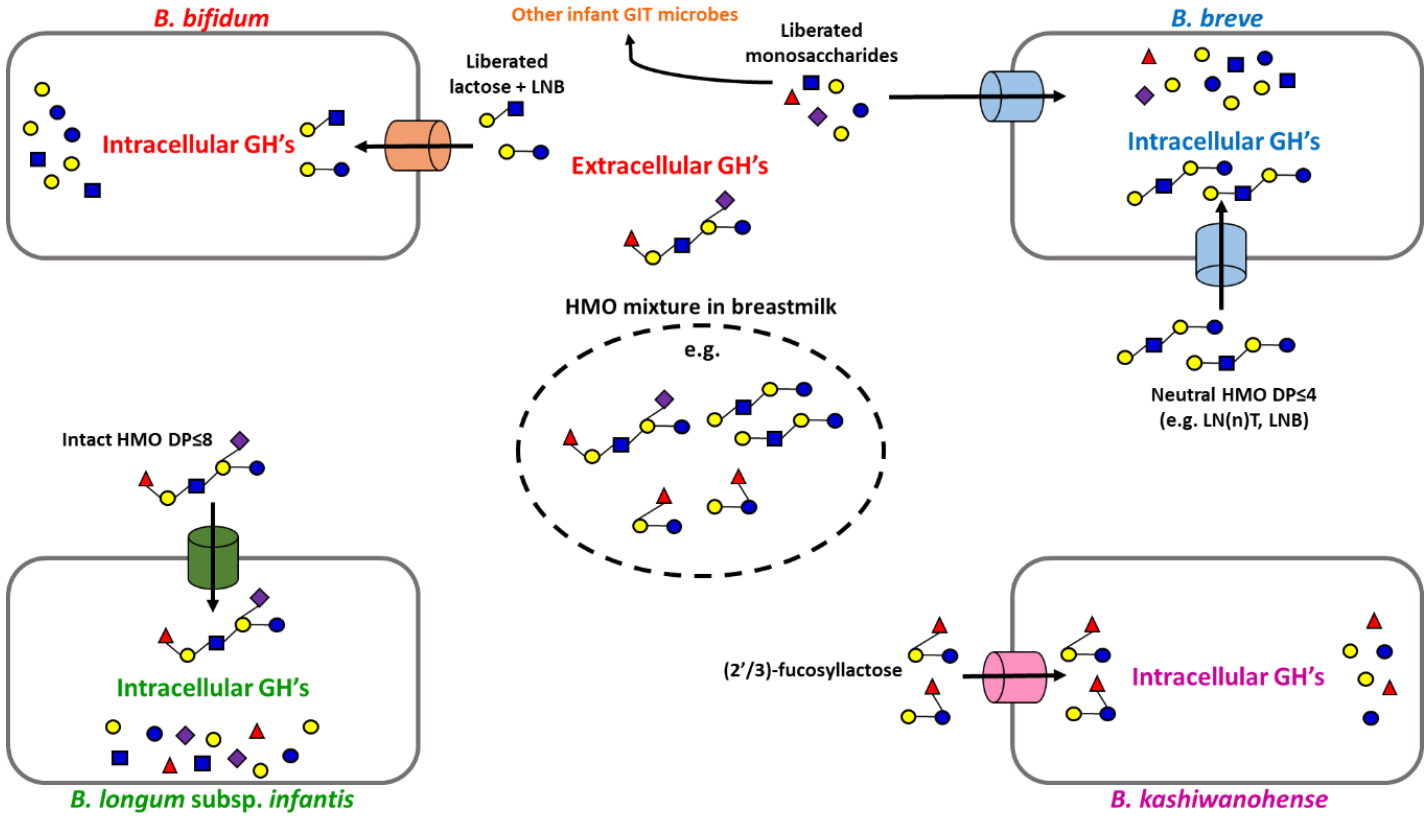


Figure 1.2. Schematic representation of the differing HMO utilisation strategies in *B. bifidum*, *B. breve*, *B. longum* subsp. *infantis* and *B. kashiwanohense*, as described in the text in section 1.5. *B. bifidum* releases extracellular glycosyl hydrolases, and internalises small-mass neutral liberated components, such as lactose and LNB. *B. breve* consumes and internally degrades intact neutral such as LN(n)T, as well as monosaccharides liberated by extracellular degradation of HMO. *B. longum* subsp. *infantis* internalises intact HMO structures (both acidic and neutral), and internally degrades them using an arsenal of glycosyl hydrolases. *B. kashiwanohense* internalises fucosyllactose, and degrades these structures using intracellular glycosyl hydrolases.

1.9 References

1. **Lozupone, C.A., J.I. Stombaugh, J.I. Gordon, J.K. Jansson, and R. Knight,** *Diversity, stability and resilience of the human gut microbiota.* Nature, **2012**. **489**(7415): p. 220-230.
2. **Milani, C., S. Duranti, F. Bottacini, E. Casey, F. Turroni, J. Mahony, C. Belzer, S. Delgado Palacio, S. Arboleya Montes, L. Mancabelli, G.A. Lugli, J.M. Rodriguez, L. Bode, W. de Vos, M. Gueimonde, A. Margolles, D. van Sinderen, and M. Ventura,** *The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota.* Microbiology and Molecular Biology Reviews, **2017**. **81**(4): 66 e00036-17.
3. **Ventura, M., F. Turroni, C. Canchaya, E.E. Vaughan, P.W. O'Toole, and D. van Sinderen,** *Microbial diversity in the human intestine and novel insights from metagenomics.* Front Biosci (Landmark Ed), **2009**. **14**: p. 3214-21.
4. **Ursell, L.K., J.C. Clemente, J.R. Rideout, D. Gevers, J.G. Caporaso, and R. Knight,** *The interpersonal and intrapersonal diversity of human-associated microbiota in key body sites.* The Journal of allergy and clinical immunology, **2012**. **129**(5): p. 1204-1208.
5. **Ventura, M., S. O'Flaherty, M.J. Claesson, F. Turroni, T.R. Klaenhammer, D. van Sinderen, and P.W. O'Toole,** *Genome-scale analyses of health-promoting bacteria: probiogenomics.* Nat Rev Microbiol, **2009**. **7**(1): p. 61-71.
6. **Relman, D.A.,** *The human microbiome: ecosystem resilience and health.* Nutrition reviews, **2012**. **70**(Suppl 1): p. S2-S9.
7. **Clemente, J.C., L.K. Ursell, L.W. Parfrey, and R. Knight,** *The impact of the gut microbiota on human health: an integrative view.* Cell, **2012**. **148**(6): p. 1258-70.
8. **Bordenstein, S.R. and K.R. Theis,** *Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes.* PLoS Biol, **2015**. **13**(8): p. e1002226.
9. **Yatsunenko, T., F.E. Rey, and M.J. Manary,** *Human gut microbiome viewed across age and geography.* Nature, **2012**. **486**: p. 222–227.
10. **O'Toole, P.W. and M.J. Claesson,** *Gut microbiota: Changes throughout the lifespan from infancy to elderly.* International Dairy Journal, **2010**. **20**(4): p. 281-291.
11. **Yassour, M., T. Vatanen, H. Siljander, A.M. Hamalainen, T. Harkonen, S.J. Ryhanen, E.A. Franzosa, H. Vlamakis, C. Huttenhower, D. Gevers, E.S. Lander,**

M. Knip, and R.J. Xavier, *Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability.* Sci Transl Med, **2016**. **8**(343): p. 343ra81.

12. **Bokulich, N.A., J. Chung, T. Battaglia, N. Henderson, M. Jay, H. Li, D.L. A, F. Wu, G.I. Perez-Perez, Y. Chen, W. Schweizer, X. Zheng, M. Contreras, M.G. Dominguez-Bello, and M.J. Blaser,** *Antibiotics, birth mode, and diet shape microbiome maturation during early life.* Sci Transl Med, **2016**. **8**(343): p. 343ra82.
13. **Hill, C.J., D.B. Lynch, K. Murphy, M. Ulaszewska, I.B. Jeffery, C.A. O'Shea, C. Watkins, E. Dempsey, F. Mattivi, K. Tuohy, R.P. Ross, C.A. Ryan, P.W. O' Toole, and C. Stanton,** *Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort.* Microbiome, **2017**. **5**(1): p. 4.
14. **Rodriguez, J.M., K. Murphy, C. Stanton, R.P. Ross, O.I. Kober, N. Juge, E. Avershina, K. Rudi, A. Narbad, M.C. Jenmalm, J.R. Marchesi, and M.C. Collado,** *The composition of the gut microbiota throughout life, with an emphasis on early life.* Microb Ecol Health Dis, **2015**. **26**: p. 26050.
15. **Avershina, E., K. Lundgard, M. Sekelja, C. Dotterud, O. Storro, T. Oien, R. Johnsen, and K. Rudi,** *Transition from infant- to adult-like gut microbiota.* Environ Microbiol, **2016**. **18**(7): p. 2226-36.
16. **Lim, E.S., D. Wang, and L.R. Holtz,** *The Bacterial Microbiome and Virome Milestones of Infant Development.* Trends Microbiol, **2016**. **24**(10): p. 801-10.
17. **Biteen, J.S., P.C. Blainey, Z.G. Cardon, M. Chun, G.M. Church, P.C. Dorrestein, S.E. Fraser, J.A. Gilbert, J.K. Jansson, R. Knight, J.F. Miller, A. Ozcan, K.A. Prather, S.R. Quake, E.G. Ruby, P.A. Silver, S. Taha, G. van den Engh, P.S. Weiss, G.C. Wong, A.T. Wright, and T.D. Young,** *Tools for the Microbiome: Nano and Beyond.* ACS Nano, **2016**. **10**(1): p. 6-37.
18. **Weinstock, G.M.,** *Genomic approaches to studying the human microbiota.* Nature, **2012**. **489**(7415): p. 250-256.
19. **Browne, H.P., S.C. Forster, B.O. Anonye, N. Kumar, B.A. Neville, M.D. Stares, D. Goulding, and T.D. Lawley,** *Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation.* Nature, **2016**. **533**: p. 543-546.
20. **Lagier, J.-C., S. Khelaifia, M.T. Alou, S. Ndongo, N. Dione, P. Hugon, A. Caputo, F. Cadoret, S.I. Traore, E.H. Seck, G. Dubourg, G. Durand, G. Mourembou, E. Guilhot, A. Togo, S. Bellali, D. Bachar, N. Cassir, F. Bittar, J. Delerce, M. Mailhe, D. Ricaboni, M. Bilen, N.P.M. Dangui Nieko, N.M. Dia Badiane, C. Valles, D.**

Mouelhi, K. Diop, M. Million, D. Musso, J. Abrahão, E.I. Azhar, F. Bibi, M. Yasir, A. Diallo, C. Sokhna, F. Djossou, V. Vitton, C. Robert, J.M. Rolain, B. La Scola, P.-E. Fournier, A. Levasseur, and D. Raoult, *Culture of previously uncultured members of the human gut microbiota by culturomics.* Nature Microbiology, **2016**. **1**: p. 16203.

21. **Arrieta, M.-C., L.T. Stiemsma, N. Amenyogbe, E.M. Brown, and B. Finlay,** *The Intestinal Microbiome in Early Life: Health and Disease.* Frontiers in Immunology, **2014**. **5**: p. 427.
22. **Valles, Y., A. Artacho, A. Pascual-Garcia, M.L. Ferrus, M.J. Gosalbes, J.J. Abellan, and M.P. Francino,** *Microbial succession in the gut: directional trends of taxonomic and functional change in a birth cohort of Spanish infants.* PLoS Genet, **2014**. **10**(6): p. e1004406.
23. **Murphy, K., D. Curley, T.F. O'Callaghan, C.-A. O'Shea, E.M. Dempsey, P.W. O'Toole, R.P. Ross, C.A. Ryan, and C. Stanton,** *The Composition of Human Milk and Infant Faecal Microbiota Over the First Three Months of Life: A Pilot Study.* Scientific Reports, **2017**. **7**: p. 40597.
24. **Turroni, F., C. Peano, D.A. Pass, E. Foroni, M. Severgnini, M.J. Claesson, C. Kerr, J. Hourihane, D. Murray, F. Fuligni, M. Gueimonde, A. Margolles, G. De Bellis, P.W. O'Toole, D. van Sinderen, J.R. Marchesi, and M. Ventura,** *Diversity of Bifidobacteria within the Infant Gut Microbiota.* PLoS ONE, **2012**. **7**(5): p. e36957.
25. **Bergstrom, A., T.H. Skov, M.I. Bahl, H.M. Roager, L.B. Christensen, K.T. Ejlerskov, C. Molgaard, K.F. Michaelsen, and T.R. Licht,** *Establishment of intestinal microbiota during early life: a longitudinal, explorative study of a large cohort of Danish infants.* Appl Environ Microbiol, **2014**. **80**(9): p. 2889-900.
26. **Roger, L.C., A. Costabile, D.T. Holland, L. Hoyles, and A.L. McCartney,** *Examination of faecal Bifidobacterium populations in breast- and formula-fed infants during the first 18 months of life.* Microbiology, **2010**. **156**(Pt 11): p. 3329-41.
27. **Jimenez, E., L. Fernandez, M.L. Marin, R. Martin, J.M. Odriozola, C. Nueno-Palop, A. Narbad, M. Olivares, J. Xaus, and J.M. Rodriguez,** *Isolation of commensal bacteria from umbilical cord blood of healthy neonates born by cesarean section.* Curr Microbiol, **2005**. **51**(4): p. 270-4.
28. **DiGiulio, D.B., R. Romero, H.P. Amogan, J.P. Kusanovic, E.M. Bik, F. Gotsch, C.J. Kim, O. Erez, S. Edwin, and D.A. Relman,** *Microbial prevalence, diversity and*

abundance in amniotic fluid during preterm labor: a molecular and culture-based investigation. PLoS One, **2008**. **3**(8): p. e3056.

29. **Aagaard, K., J. Ma, K.M. Antony, R. Ganu, J. Petrosino, and J. Versalovic,** *The placenta harbors a unique microbiome.* Sci Transl Med, **2014**. **6**(237): p. 237ra65.
30. **Munyaka, P.M., E. Khafipour, and J.E. Ghia,** *External influence of early childhood establishment of gut microbiota and subsequent health implications.* Front Pediatr, **2014**. **2**: p. 109.
31. **Dominguez-Bello, C.E.K., M. Contreras, M. Magris, G. Hidalgo, N. Fierer, and R. Knight,** *Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns.* Proc Natl Acad Sci U S A, **2010**. **107**(26)**:** p. 11971–11975.
32. **Biasucci, G., M. Rubini, S. Riboni, L. Morelli, E. Bessi, and C. Retetangos,** *Mode of delivery affects the bacterial community in the newborn gut.* Early Hum Dev, **2010**. **86 Suppl 1**: p. 13-5.
33. **Madan, J.C., S.F. Farzan, P.L. Hibberd, and M.R. Karagas,** *Normal neonatal microbiome variation in relation to environmental factors, infection and allergy.* Current opinion in pediatrics, **2012**. **24**(6): p. 753-759.
34. **Bäckhed, F., J. Roswall, and Y. Peng,** *Dynamics and stabilization of the human gut microbiome during the first year of life.* Cell Host Microbe, **2015**. **17**(5): p. 690-703.
35. **Del Chierico, F., P. Vernocchi, A. Petrucca, P. Paci, S. Fuentes, G. Praticò, G. Capuani, A. Masotti, S. Reddel, A. Russo, C. Vallone, G. Salvatori, E. Buffone, F. Signore, G. Rigon, A. Dotta, A. Miccheli, W.M. de Vos, B. Dallapiccola, and L. Putignani,** *Phylogenetic and Metabolic Tracking of Gut Microbiota during Perinatal Development.* PLOS ONE, **2015**. **10**(9): p. e0137347.
36. **Di Mauro, A., J. Neu, G. Riezzo, F. Raimondi, D. Martinelli, R. Francavilla, and F. Indrio,** *Gastrointestinal function development and microbiota.* Italian Journal of Pediatrics, **2013**. **39**(15): p. 15-15.
37. **Scholtens, P.A., R. Oozeer, R. Martin, K.B. Amor, and J. Knol,** *The early settlers: intestinal microbiology in early life.* Annu Rev Food Sci Technol, **2012**. **3**: p. 425-47.
38. **Martin, R., H. Makino, A. Cetinyurek Yavuz, K. Ben-Amor, M. Roelofs, E. Ishikawa, H. Kubota, S. Swinkels, T. Sakai, K. Oishi, A. Kushiro, and J. Knol,** *Early-Life Events, Including Mode of Delivery and Type of Feeding, Siblings and Gender, Shape the Developing Gut Microbiota.* PLoS One, **2016**. **11**(6): p. e0158498.

39. **Jakobsson, H.E., T.R. Abrahamsson, M.C. Jenmalm, K. Harris, C. Quince, C. Jernberg, B. Bjorksten, L. Engstrand, and A.F. Andersson,** *Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by caesarean section.* Gut, **2014**. **63**(4): p. 559-66.

40. **Ruiz, L., L. Moles, M. Gueimonde, and J.M. Rodriguez,** *Perinatal Microbiomes' Influence on Preterm Birth and Preterms' Health: Influencing Factors and Modulation Strategies.* J Pediatr Gastroenterol Nutr, **2016**. **63**(6): p. e193-e203.

41. **Moutquin, J.M.,** *Classification and heterogeneity of preterm birth.* Bjog, **2003**. **110 Suppl 20**: p. 30-3.

42. **Rouge, C., O. Goldenberg, L. Ferraris, B. Berger, F. Rochat, A. Legrand, U.B. Gobel, M. Vodovar, M. Voyer, J.C. Roze, D. Darmaun, H. Piloquet, M.J. Butel, and M.F. de La Cochetiere,** *Investigation of the intestinal microbiota in preterm infants using different methods.* Anaerobe, **2010**. **16**(4): p. 362-70.

43. **Jacquot, A., D. Neveu, F. Aujoulat, G. Mercier, H. Marchandin, E. Jumas-Bilak, and J.C. Picaud,** *Dynamics and clinical evolution of bacterial gut microflora in extremely premature patients.* J Pediatr, **2011**. **158**(3): p. 390-6.

44. **Arboleya, S., A. Binetti, N. Salazar, N. Fernandez, G. Solis, A. Hernandez-Barranco, A. Margolles, C.G. de Los Reyes-Gavilan, and M. Gueimonde,** *Establishment and development of intestinal microbiota in preterm neonates.* FEMS Microbiol Ecol, **2012**. **79**(3): p. 763-72.

45. **Arboleya, S., B. Sanchez, C. Milani, S. Duranti, G. Solis, N. Fernandez, C.G. de los Reyes-Gavilan, M. Ventura, A. Margolles, and M. Gueimonde,** *Intestinal microbiota development in preterm neonates and effect of perinatal antibiotics.* J Pediatr, **2015**. **166**(3): p. 538-44.

46. **Cong, X., W. Xu, S. Janton, W.A. Henderson, A. Matson, J.M. McGrath, K. Maas, and J. Graf,** *Gut Microbiome Developmental Patterns in Early Life of Preterm Infants: Impacts of Feeding and Gender.* PLoS One, **2016**. **11**(4): p. e0152751.

47. **Berrington, J.E., C.J. Stewart, S.P. Cummings, and N.D. Embleton,** *The neonatal bowel microbiome in health and infection.* Curr Opin Infect Dis, **2014**. **27**(3): p. 236-43.

48. **Berrington, J.E., C.J. Stewart, N.D. Embleton, and S.P. Cummings,** *Gut microbiota in preterm infants: assessment and relevance to health and disease.* Arch Dis Child Fetal Neonatal Ed, **2013**. **98**(4): p. F286-90.

49. **Gritz, E.C. and V. Bhandari,** *The Human Neonatal Gut Microbiome: A Brief Review.* Frontiers in Pediatrics, **2015**. **3**(17): p. 17.

50. **Butel, M.J., A. Suau, F. Campeotto, F. Magne, J. Aires, L. Ferraris, N. Kalach, B. Leroux, and C. Dupont,** *Conditions of bifidobacterial colonization in preterm infants: a prospective analysis.* J Pediatr Gastroenterol Nutr, **2007**. **44**(5): p. 577-82.

51. **O'Sullivan, A., M. Farver, and J.T. Smilowitz,** *The Influence of Early Infant-Feeding Practices on the Intestinal Microbiome and Body Composition in Infants.* Nutrition and Metabolic Insights, **2015**. **8**(Suppl 1): p. 1-9.

52. **Gomez-Llorente, C., J. Plaza-Diaz, M. Aguilera, S. Munoz-Quezada, M. Bermudez-Brito, P. Peso-Echarri, R. Martinez-Silla, M.I. Vasallo-Morillas, L. Campana-Martin, I. Vives-Pinera, M.J. Ballesta-Martinez, and A. Gil,** *Three main factors define changes in fecal microbiota associated with feeding modality in infants.* J Pediatr Gastroenterol Nutr, **2013**. **57**(4): p. 461-6.

53. **Jain, N. and W.A. Walker,** *Diet and host-microbial crosstalk in postnatal intestinal immune homeostasis.* Nat Rev Gastroenterol Hepatol, **2015**. **12**(1): p. 14-25.

54. **Guaraldi, F. and G. Salvatori,** *Effect of Breast and Formula Feeding on Gut Microbiota Shaping in Newborns.* Frontiers in Cellular and Infection Microbiology, **2012**. **2**(94): p. 94.

55. **Jost, T., C. Lacroix, C.P. Braegger, and C. Chassard,** *New Insights in Gut Microbiota Establishment in Healthy Breast Fed Neonates.* PLoS ONE, **2012**. **7**(8): p. e44595.

56. **Bezirtzoglou, E., A. Tsiotsias, and G.W. Welling,** *Microbiota profile in feces of breast- and formula-fed newborns by using fluorescence in situ hybridization (FISH).* Anaerobe, **2011**. **17**(6): p. 478-82.

57. **Harmsen, H.J., A.C. Wildeboer-Veloo, G.C. Raangs, A.A. Wagendorp, N. Klijn, J.G. Bindels, and G.W. Welling,** *Analysis of intestinal flora development in breast-fed and formula-fed infants by using molecular identification and detection methods.* J Pediatr Gastroenterol Nutr, **2000**. **30**(1): p. 61-7.

58. **Fallani, M., S. Amarri, A. Uusijarvi, R. Adam, S. Khanna, M. Aguilera, A. Gil, J.M. Vieites, E. Norin, D. Young, J.A. Scott, J. Dore, and C.A. Edwards,** *Determinants of the human infant intestinal microbiota after the introduction of first complementary foods in infant samples from five European centres.* Microbiology, **2011**. **157**(Pt 5): p. 1385-92.

59. **Koenig, J.E., A. Spor, N. Scalfone, A.D. Fricker, J. Stombaugh, R. Knight, L.T. Angenent, and R.E. Ley,** *Succession of microbial consortia in the developing infant gut microbiome.* Proc Natl Acad Sci U S A, **2011**. **108 Suppl 1**: p. 4578-85.

60. **Hooks, K.B. and M.A. O'Malley,** *Dysbiosis and Its Discontents.* mBio, **2017**. **8**(5).

61. **Torrazza, R.M. and J. Neu,** *The altered gut microbiome and necrotizing enterocolitis.* Clin Perinatol, **2013**. **40**(1): p. 93-108.

62. **Madan, J.C., R.C. Salari, D. Saxena, L. Davidson, G.A. O'Toole, J.H. Moore, M.L. Sogin, J.A. Foster, W.H. Edwards, P. Palumbo, and P.L. Hibberd,** *Gut microbial colonisation in premature neonates predicts neonatal sepsis.* Arch Dis Child Fetal Neonatal Ed, **2012**. **97**(6): p. F456-62.

63. **Mai, V., R.M. Torrazza, M. Ukhanova, X. Wang, Y. Sun, N. Li, J. Shuster, R. Sharma, M.L. Hudak, and J. Neu,** *Distortions in development of intestinal microbiota associated with late onset sepsis in preterm infants.* PLoS One, **2013**. **8**(1): p. e52876.

64. **Thavagnanam, S., J. Fleming, A. Bromley, M.D. Shields, and C.R. Cardwell,** *A meta-analysis of the association between Caesarean section and childhood asthma.* Clin Exp Allergy, **2008**. **38**(4): p. 629-33.

65. **Bager, P., J. Wohlfahrt, and T. Westergaard,** *Caesarean delivery and risk of atopy and allergic disease: meta-analyses.* Clin Exp Allergy, **2008**. **38**(4): p. 634-42.

66. **Cardwell, C.R., L.C. Stene, G. Joner, O. Cinek, J. Svensson, M.J. Goldacre, R.C. Parslow, P. Pozzilli, G. Brigis, D. Stoyanov, B. Urbonaite, S. Sipetic, E. Schober, C. Ionescu-Tirgoviste, G. Devoti, C.E. de Beaufort, K. Buschard, and C.C. Patterson,** *Caesarean section is associated with an increased risk of childhood-onset type 1 diabetes mellitus: a meta-analysis of observational studies.* Diabetologia, **2008**. **51**(5): p. 726-35.

67. **Pei, Z., J. Heinrich, E. Fuertes, C. Flexeder, B. Hoffmann, I. Lehmann, B. Schaaf, A. von Berg, and S. Koletzko,** *Cesarean delivery and risk of childhood obesity.* J Pediatr, **2014**. **164**(5): p. 1068-1073.e2.

68. **Hill, C., F. Guarner, G. Reid, G.R. Gibson, D.J. Merenstein, B. Pot, L. Morelli, R.B. Canani, H.J. Flint, S. Salminen, P.C. Calder, and M.E. Sanders,** *Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic.* Nat Rev Gastroenterol Hepatol, **2014**. **11**(8): p. 506-14.

69. **Rao, S.C., G.K. Athalye-Jape, G.C. Deshpande, K.N. Simmer, and S.K. Patole,** *Probiotic Supplementation and Late-Onset Sepsis in Preterm Infants: A Meta-analysis.* Pediatrics, **2016**. **137**(3): p. e20153684.

70. **Dermyshi, E., Y. Wang, C. Yan, W. Hong, G. Qiu, X. Gong, and T. Zhang,** *The "Golden Age" of Probiotics: A Systematic Review and Meta-Analysis of Randomized and Observational Studies in Preterm Infants.* Neonatology, **2017**. **112**(1): p. 9-23.

71. **Gibson, G.R., H.M. Probert, J.V. Loo, R.A. Rastall, and M.B. Roberfroid,** *Dietary modulation of the human colonic microbiota: updating the concept of prebiotics.* Nutr Res Rev, **2004**. **17**(2): p. 259-75.

72. **Gibson, G.R. and M.B. Roberfroid,** *Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics.* J Nutr, **1995**. **125**(6): p. 1401-12.

73. **Roberfroid, M.,** *Prebiotics: the concept revisited.* J Nutr, **2007**. **137**(3 Suppl 2): p. 830s-7s.

74. **Bode, L.,** *Human milk oligosaccharides: Every baby needs a sugar mama.* Glycobiology, **2012**. **22**(9): p. 1147-1162.

75. **Gibson, G.R., R. Hutkins, M.E. Sanders, S.L. Prescott, R.A. Reimer, S.J. Salminen, K. Scott, C. Stanton, K.S. Swanson, P.D. Cani, K. Verbeke, and G. Reid,** *Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics.* Nat Rev Gastroenterol Hepatol, **2017**. **14**(8): p. 491-502.

76. **Roberfroid, M.,** *Prebiotics: The Concept Revisited.* The Journal of Nutrition, **2007**. **137**(3): p. 830S-837S.

77. **Moro, G., I. Minoli, M. Mosca, S. Fanaro, J. Jelinek, B. Stahl, and G. Boehm,** *Dosage-related bifidogenic effects of galacto- and fructooligosaccharides in formula-fed term infants.* J Pediatr Gastroenterol Nutr, **2002**. **34**(3): p. 291-5.

78. **Boehm, G., M. Lidestri, P. Casetta, J. Jelinek, F. Negretti, B. Stahl, and A. Marini,** *Supplementation of a bovine milk formula with an oligosaccharide mixture increases counts of faecal bifidobacteria in preterm infants.* Archives of Disease in Childhood Fetal and Neonatal Edition, **2002**. **86**(3): p. F178-F181.

79. **Kolida, S. and G.R. Gibson,** *Synbiotics in health and disease.* Annu Rev Food Sci Technol, **2011**. **2**: p. 373-93.

80. **Johnson-Henry, K.C., T.R. Abrahamsson, R.Y. Wu, and P.M. Sherman,** *Probiotics, Prebiotics, and Synbiotics for the Prevention of Necrotizing Enterocolitis.* Adv Nutr, **2016**. **7**(5): p. 928-37.

81. **Panigrahi, P., S. Parida, N.C. Nanda, R. Satpathy, L. Pradhan, D.S. Chandel, L. Baccaglini, A. Mohapatra, S.S. Mohapatra, P.R. Misra, R. Chaudhry, H.H. Chen, J.A. Johnson, J.G. Morris, N. Paneth, and I.H. Gewolb,** *A randomized synbiotic trial to prevent sepsis among infants in rural India.* Nature, **2017**. **548**(7668): p. 407-412.

82. **Pannaraj, P.S., F. Li, C. Cerini, J.M. Bender, S. Yang, A. Rollie, H. Adisetiyo, S. Zabih, P.J. Lincez, K. Bittinger, A. Bailey, F.D. Bushman, J.W. Sleasman, and G.M. Aldrovandi,** *Association Between Breast Milk Bacterial Communities and Establishment and Development of the Infant Gut Microbiome.* JAMA Pediatr, **2017**. **171**(7): p. 647-654.

83. **Martin, V., A. Maldonado-Barragan, L. Moles, M. Rodriguez-Banos, R.D. Campo, L. Fernandez, J.M. Rodriguez, and E. Jimenez,** *Sharing of bacterial strains between breast milk and infant feces.* J Hum Lact, **2012**. **28**(1): p. 36-44.

84. **Brüssow, H.,** *Human microbiota: 'The philosophers have only interpreted the world in various ways. The point, however, is to change it'.* Microbial Biotechnology, **2015**. **8**(1): p. 11-12.

85. **Musilova, S., V. Rada, E. Vlkova, and V. Bunesova,** *Beneficial effects of human milk oligosaccharides on gut microbiota.* Benef Microbes, **2014**. **5**(3): p. 273-83.

86. **Sela, D.A. and D.A. Mills,** *Nursing our microbiota: molecular linkages between bifidobacteria and milk oligosaccharides.* Trends Microbiol, **2010**. **18**(7): p. 298-307.

87. **Kunz, C., S. Rudloff, W. Baier, N. Klein, and S. Strobel,** *OLIGOSACCHARIDES IN HUMAN MILK: Structural, Functional, and Metabolic Aspects.* Annual Review of Nutrition, **2000**. **20**(1): p. 699-722.

88. **Bode, L.,** *The functional biology of human milk oligosaccharides.* Early Hum Dev, **2015**. **91**(11): p. 619-22.

89. **Ninonuevo, M.R., Y. Park, H. Yin, J. Zhang, R.E. Ward, B.H. Clowers, J.B. German, S.L. Freeman, K. Killeen, R. Grimm, and C.B. Lebrilla,** *A strategy for annotating the human milk glycome.* J Agric Food Chem, **2006**. **54**(20): p. 7471-80.

90. **Kobata, A.,** *Structures and application of oligosaccharides in human milk.* Proc Jpn Acad Ser B Phys Biol Sci, **2010**. **86**(7): p. 731-47.

91. **McGuire, M.K., C.L. Meehan, M.A. McGuire, and J.E. Williams,** *What's normal? Oligosaccharide concentrations and profiles in milk produced by healthy women vary geographically.* Am J Clin Nutr., **2017**. **105**(5): p. 1086-1100.

92. **Coppa, G.V., P. Pierani, L. Zampini, I. Carloni, A. Carlucci, and O. Gabrielli,** *Oligosaccharides in human milk during different phases of lactation.* Acta Paediatr Suppl, **1999**. **88**(430): p. 89-94.

93. **Gabrielli, O., L. Zampini, T. Galeazzi, L. Padella, L. Santoro, C. Peila, F. Giuliani, E. Bertino, C. Fabris, and G.V. Coppa,** *Preterm milk oligosaccharides during the first month of lactation.* Pediatrics, **2011**. **128**(6): p. e1520-31.

94. **Newburg, D.S., Z. Shen, and C.D. Warren,** *Quantitative analysis of human milk oligosaccharides by capillary electrophoresis.* Adv Exp Med Biol, **2000**. **478**: p. 381-2.

95. **Chaturvedi, P., C.D. Warren, M. Altaye, A.L. Morrow, G. Ruiz-Palacios, L.K. Pickering, and D.S. Newburg,** *Fucosylated human milk oligosaccharides vary between individuals and over the course of lactation.* Glycobiology, **2001**. **11**(5): p. 365-72.

96. **Davidson, B., J.K. Meinzen-Derr, C.L. Wagner, D.S. Newburg, and A.L. Morrow,** *Fucosylated oligosaccharides in human milk in relation to gestational age and stage of lactation.* Adv Exp Med Biol, **2004**. **554**: p. 427-30.

97. **Bao, Y., L. Zhu, and D.S. Newburg,** *Simultaneous quantification of sialyloligosaccharides from human milk by capillary electrophoresis.* Anal Biochem, **2007**. **370**(2): p. 206-14.

98. **Urashima, T., S. Asakuma, F. Leo, K. Fukuda, M. Messer, and O.T. Oftedal,** *The Predominance of Type I Oligosaccharides Is a Feature Specific to Human Breast Milk.* Advances in Nutrition, **2012**. **3**(3): p. 473S-482S.

99. **Urashima, T., T. Saito, T. Nakamura, and M. Messer,** *Oligosaccharides of milk and colostrum in non-human mammals.* Glycoconjugate Journal, **2001**. **18**(5): p. 357-371.

100. **Urashima, T., M. Kobayashi, S. Asakuma, Y. Uemura, I. Arai, K. Fukuda, T. Saito, T. Mogoe, H. Ishikawa, and Y. Fukui,** *Chemical characterization of the oligosaccharides in Bryde's whale (Balaenoptera edeni) and Sei whale (Balaenoptera borealis lesson) milk.* Comp Biochem Physiol B Biochem Mol Biol, **2007**. **146**(2): p. 153-9.

101. **Urashima, T., G. Odaka, S. Asakuma, Y. Uemura, K. Goto, A. Senda, T. Saito, K. Fukuda, M. Messer, and O.T. Oftedal,** *Chemical characterization of oligosaccharides in chimpanzee, bonobo, gorilla, orangutan, and siamang milk or colostrum.* Glycobiology, **2009**. **19**(5): p. 499-508.

102. **Wu, S., R. Grimm, J.B. German, and C.B. Lebrilla,** *Annotation and structural analysis of sialylated human milk oligosaccharides.* J Proteome Res, **2011**. **10**(2): p. 856-68.

103. **Wu, S., N. Tao, J.B. German, R. Grimm, and C.B. Lebrilla,** *Development of an annotated library of neutral human milk oligosaccharides.* J Proteome Res, **2010**. **9**(8): p. 4138-51.

104. **Kumazaki, T. and A. Yoshida,** *Biochemical evidence that secretor gene, Se, is a structural gene encoding a specific fucosyltransferase.* Proc Natl Acad Sci U S A, **1984**. **81**(13): p. 4193-7.

105. **Johnson, P.H. and W.M. Watkins,** *Purification of the Lewis blood-group gene associated alpha-3/4-fucosyltransferase from human milk: an enzyme transferring fucose primarily to type 1 and lactose-based oligosaccharide chains.* Glycoconj J, **1992**. **9**(5): p. 241-9.

106. **Stahl, B., S. Thurl, J. Henker, M. Siegel, B. Finke, and G. Sawatzki,** *Detection of four human milk groups with respect to Lewis-blood-group-dependent oligosaccharides by serologic and chromatographic analysis.* Adv Exp Med Biol, **2001**. **501**: p. 299-306.

107. **Thurl, S., M. Munzert, J. Henker, G. Boehm, B. Muller-Werner, J. Jelinek, and B. Stahl,** *Variation of human milk oligosaccharides in relation to milk groups and lactational periods.* Br J Nutr, **2010**. **104**(9): p. 1261-71.

108. **Xu, Z., L. Vo, and B.A. Macher,** *Structure-function analysis of human alpha1,3-fucosyltransferase. Amino acids involved in acceptor substrate specificity.* J Biol Chem, **1996**. **271**(15): p. 8818-23.

109. **Engfer, M.B., B. Stahl, B. Finke, G. Sawatzki, and H. Daniel,** *Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract.* Am J Clin Nutr, **2000**. **71**(6): p. 1589-96.

110. **Gnoth, M.J., C. Kunz, E. Kinne-Saffran, and S. Rudloff,** *Human milk oligosaccharides are minimally digested in vitro.* J Nutr, **2000**. **130**(12): p. 3014-20.

111. **Albrecht, S., H.A. Schols, E.G. van den Heuvel, A.G. Voragen, and H. Gruppen,** *CE-LIF-MS n profiling of oligosaccharides in human milk and feces of breast-fed babies.* Electrophoresis, **2010**. **31**(7): p. 1264-73.

112. **Albrecht, S., H.A. Schols, E.G. van den Heuvel, A.G. Voragen, and H. Gruppen,** *Occurrence of oligosaccharides in feces of breast-fed babies in their first six months of life and the corresponding breast milk.* Carbohydr Res, **2011**. **346**(16): p. 2540-50.

113. **Albrecht, S., H.A. Schols, D. van Zoeren, R.A. van Lingen, L.J. Groot Jebbink, E.G. van den Heuvel, A.G. Voragen, and H. Gruppen,** *Oligosaccharides in feces of breast- and formula-fed babies.* Carbohydr Res, **2011**. **346**(14): p. 2173-81.

114. **Rudloff, S., G. Pohlentz, L. Diekmann, H. Egge, and C. Kunz,** *Urinary excretion of lactose and oligosaccharides in preterm infants fed human milk or infant formula.* Acta Paediatr, **1996**. **85**(5): p. 598-603.

115. **Charbonneau, Mark R., D. O'Donnell, Laura V. Blanton, Sarah M. Totten, Jasmine C.C. Davis, Michael J. Barratt, J. Cheng, J. Guruge, M. Talcott, James R. Bain, Michael J. Muehlbauer, O. Ilkayeva, C. Wu, T. Struckmeyer, D. Barile, C. Mangani, J. Jorgensen, Y.-m. Fan, K. Maleta, Kathryn G. Dewey, P. Ashorn, Christopher B. Newgard, C. Lebrilla, David A. Mills, and Jeffrey I. Gordon,** *Sialylated Milk Oligosaccharides Promote Microbiota-Dependent Growth in Models of Infant Undernutrition.* Cell. **164**(5): p. 859-871.

116. **Marcobal, A., M. Barboza, J.W. Froehlich, D.E. Block, J.B. German, C.B. Lebrilla, and D.A. Mills,** *Consumption of Human Milk Oligosaccharides by Gut-related Microbes.* Journal of agricultural and food chemistry, **2010**. **58**(9): p. 5334-5340.

117. **Marcobal, A., M. Barboza, and E.D. Sonnenburg,** *Bacteroides in the infant gut consume milk oligosaccharides via mucus-utilization pathways.* Cell Host Microbe, **2011**. **10**(5): p. 507-14.

118. **Ward, R.E., M. Ninonuevo, D.A. Mills, C.B. Lebrilla, and J.B. German,** *In vitro fermentability of human milk oligosaccharides by several strains of bifidobacteria.* Mol Nutr Food Res, **2007**. **51**(11): p. 1398-405.

119. **Asakuma, S., E. Hatakeyama, T. Urashima, E. Yoshida, T. Katayama, K. Yamamoto, H. Kumagai, H. Ashida, J. Hirose, and M. Kitaoka,** *Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria.* J Biol Chem, **2011**. **286**(40): p. 34583-92.

120. **LoCascio, R.G., M.R. Ninonuevo, S.L. Freeman, D.A. Sela, R. Grimm, C.B. Lebrilla, D.A. Mills, and J.B. German,** *Glycoprofiling of Bifidobacterial Consumption of Human Milk Oligosaccharides Demonstrates Strain Specific, Preferential Consumption of Small Chain Glycans Secreted in Early Human Lactation.* Journal of Agricultural and Food Chemistry, **2007**. **55**(22): p. 8914-8919.

121. **Wang, M., M. Li, S. Wu, C.B. Lebrilla, R.S. Chapkin, I. Ivanov, and S.M. Donovan,** *Fecal Microbiota Composition of Breast-fed Infants is Correlated with*

Human Milk Oligosaccharides Consumed. Journal of pediatric gastroenterology and nutrition, **2015**. **60**(6): p. 825-833.

122. **Newburg, D.S., G.M. Ruiz-Palacios, and A.L. Morrow,** *Human milk glycans protect infants against enteric pathogens.* Annu Rev Nutr, **2005**. **25**: p. 37-58.
123. **Simon, P.M., P.L. Goode, A. Mobasseri, and D. Zopf,** *Inhibition of Helicobacter pylori binding to gastrointestinal epithelial cells by sialic acid-containing oligosaccharides.* Infect Immun, **1997**. **65**(2): p. 750-7.
124. **Gustafsson, A., A. Hultberg, R. Sjostrom, I. Kacskovics, M.E. Breimer, T. Boren, L. Hammarstrom, and J. Holgersson,** *Carbohydrate-dependent inhibition of Helicobacter pylori colonization using porcine milk.* Glycobiology, **2006**. **16**(1): p. 1-10.
125. **Ruiz-Palacios, G.M., L.E. Cervantes, P. Ramos, B. Chavez-Munguia, and D.S. Newburg,** *Campylobacter jejuni binds intestinal H(O) antigen (Fuc alpha 1, 2Gal beta 1, 4GlcNAc), and fucosyloligosaccharides of human milk inhibit its binding and infection.* J Biol Chem, **2003**. **278**(16): p. 14112-20.
126. **Morrow, A.L., G.M. Ruiz-Palacios, M. Altaye, X. Jiang, M.L. Guerrero, J.K. Meinzen-Derr, T. Farkas, P. Chaturvedi, L.K. Pickering, and D.S. Newburg,** *Human milk oligosaccharides are associated with protection against diarrhea in breast-fed infants.* J Pediatr, **2004**. **145**(3): p. 297-303.
127. **Jantscher-Krenn, E., T. Lauwaet, L.A. Bliss, S.L. Reed, F.D. Gillin, and L. Bode,** *Human milk oligosaccharides reduce Entamoeba histolytica attachment and cytotoxicity in vitro.* Br J Nutr, **2012**. **108**(10): p. 1839-46.
128. **Hong, P., M.R. Ninonuevo, B. Lee, C. Lebrilla, and L. Bode,** *Human milk oligosaccharides reduce HIV-1-gp120 binding to dendritic cell-specific ICAM3-grabbing non-integrin (DC-SIGN).* Br J Nutr, **2009**. **101**(4): p. 482-6.
129. **Lane, J.A., R.K. Mehra, S.D. Carrington, and R.M. Hickey,** *Development of biosensor-based assays to identify anti-infective oligosaccharides.* Anal Biochem, **2011**. **410**(2): p. 200-5.
130. **Lin, A.E., C.A. Autran, A. Szyszka, T. Escajadillo, M. Huang, K. Godula, A.R. Prudden, G.J. Boons, A.L. Lewis, K.S. Doran, V. Nizet, and L. Bode,** *Human milk oligosaccharides inhibit growth of group B Streptococcus.* J Biol Chem, **2017**. **292**(27): p. 11243-11249.

131. **Gonia, S., M. Tuepker, T. Heisel, C. Autran, L. Bode, and C.A. Gale,** *Human Milk Oligosaccharides Inhibit Candida albicans Invasion of Human Premature Intestinal Epithelial Cells.* J Nutr, **2015**. **145**(9): p. 1992-8.

132. **Kuntz, S., S. Rudloff, and C. Kunz,** *Oligosaccharides from human milk influence growth-related characteristics of intestinally transformed and non-transformed intestinal cells.* Br J Nutr, **2008**. **99**(3): p. 462-71.

133. **Angeloni, S., J.L. Ridet, N. Kusy, H. Gao, F. Crevoisier, S. Guinchard, S. Kochhar, H. Sigrist, and N. Sprenger,** *Glycoprofiling with micro-arrays of glycoconjugates and lectins.* Glycobiology, **2005**. **15**(1): p. 31-41.

134. **Eiwegger, T., B. Stahl, P. Haidl, J. Schmitt, G. Boehm, E. Dehlink, R. Urbanek, and Z. Szepfalusi,** *Prebiotic oligosaccharides: in vitro evidence for gastrointestinal epithelial transfer and immunomodulatory properties.* Pediatr Allergy Immunol, **2010**. **21**(8): p. 1179-88.

135. **Eiwegger, T., B. Stahl, J. Schmitt, G. Boehm, M. Gerstmayr, J. Pichler, E. Dehlink, C. Loibichler, R. Urbanek, and Z. Szepfalusi,** *Human milk--derived oligosaccharides and plant-derived oligosaccharides stimulate cytokine production of cord blood T-cells in vitro.* Pediatr Res, **2004**. **56**(4): p. 536-40.

136. **Wang, B.,** *Sialic acid is an essential nutrient for brain development and cognition.* Annu Rev Nutr, **2009**. **29**: p. 177-222.

137. **Lucas, A. and T.J. Cole,** *Breast milk and neonatal necrotising enterocolitis.* Lancet, **1990**. **336**(8730): p. 1519-23.

138. **Schanler, R.J., C. Lau, N.M. Hurst, and E.O. Smith,** *Randomized trial of donor human milk versus preterm formula as substitutes for mothers' own milk in the feeding of extremely premature infants.* Pediatrics, **2005**. **116**(2): p. 400-6.

139. **Sisk, P.M., C.A. Lovelady, R.G. Dillard, K.J. Gruber, and T.M. O'Shea,** *Early human milk feeding is associated with a lower risk of necrotizing enterocolitis in very low birth weight infants.* J Perinatol, **2007**. **27**(7): p. 428-33.

140. **Autran, C.A., M.H. Schoterman, E. Jantscher-Krenn, J.P. Kamerling, and L. Bode,** *Sialylated galacto-oligosaccharides and 2'-fucosyllactose reduce necrotising enterocolitis in neonatal rats.* Br J Nutr, **2016**. **116**(2): p. 294-9.

141. **Jantscher-Krenn, E., M. Zherebtsov, C. Nissan, K. Goth, Y.S. Guner, N. Naidu, B. Choudhury, A.V. Grishin, H.R. Ford, and L. Bode,** *The human milk oligosaccharide disialyllacto-N-tetraose prevents necrotising enterocolitis in neonatal rats.* Gut, **2012**. **61**(10): p. 1417-25.

142. **Autran, C.A., B.P. Kellman, J.H. Kim, E. Asztalos, A.B. Blood, E.C. Spence, A.L. Patel, J. Hou, N.E. Lewis, and L. Bode,** *Human milk oligosaccharide composition predicts risk of necrotising enterocolitis in preterm infants.* Gut, **2017**(0): p. e1-7.

143. **Good, M., C.P. Sodhi, Y. Yamaguchi, H. Jia, P. Lu, W.B. Fulton, L.Y. Martin, T. Prindle, D.F. Nino, Q. Zhou, C. Ma, J.A. Ozolek, R.H. Buck, K.C. Goehring, and D.J. Hackam,** *The human milk oligosaccharide 2'-fucosyllactose attenuates the severity of experimental necrotising enterocolitis by enhancing mesenteric perfusion in the neonatal intestine.* Br J Nutr, **2016**. **116**(7): p. 1175-1187.

144. **Macfarlane, G.T., H. Steed, and S. Macfarlane,** *Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics.* J Appl Microbiol, **2008**. **104**(2): p. 305-44.

145. **Rijnierse, A., P.V. Jeurink, B.C. van Esch, J. Garssen, and L.M. Knippels,** *Food-derived oligosaccharides exhibit pharmaceutical properties.* Eur J Pharmacol, **2011**. **668 Suppl 1**: p. S117-23.

146. **Puccio, G., P. Alliet, C. Cajozzo, E. Janssens, G. Corsello, N. Sprenger, S. Wernimont, D. Egli, L. Gosoniu, and P. Steenhout,** *Effects of Infant Formula With Human Milk Oligosaccharides on Growth and Morbidity: A Randomized Multicenter Trial.* Journal of Pediatric Gastroenterology and Nutrition, **2017**. **64**(4): p. 624-631.

147. **Goehring, K.C., B.J. Marriage, J.S. Oliver, J.A. Wilder, E.G. Barrett, and R.H. Buck,** *Similar to Those Who Are Breastfed, Infants Fed a Formula Containing 2'-Fucosyllactose Have Lower Inflammatory Cytokines in a Randomized Controlled Trial.* The Journal of Nutrition, **2016**. **146**(12): p. 2559-2566.

148. **Lee, J.H. and D.J. O'Sullivan,** *Genomic insights into bifidobacteria.* Microbiol Mol Biol Rev, **2010**. **74**(3): p. 378-416.

149. **Ventura, M., C. Canchaya, A. Tauch, G. Chandra, G.F. Fitzgerald, K.F. Chater, and D. van Sinderen,** *Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum.* Microbiol Mol Biol Rev, **2007**. **71**(3): p. 495-548.

150. **Lugli, G.A., C. Milani, F. Turroni, S. Duranti, C. Ferrario, A. Viappiani, L. Mancabelli, M. Mangifesta, B. Taminiau, V. Delcenserie, D. van Sinderen, and M. Ventura,** *Investigation of the evolutionary development of the genus Bifidobacterium by comparative genomics.* Appl Environ Microbiol, **2014**. **80**(20): p. 6383-94.

151. **Lugli, G.A., C. Milani, F. Turroni, S. Duranti, L. Mancabelli, M. Mangifesta, C. Ferrario, M. Modesto, P. Mattarelli, K. Jiri, D. van Sinderen, and M. Ventura,**

Comparative genomic and phylogenomic analyses of the Bifidobacteriaceae family. BMC Genomics, **2017**. **18**(1): p. 568.

152. **Turroni, F., F. Serafini, M. Mangifesta, S. Arioli, D. Mora, D. van Sinderen, and M. Ventura,** *Expression of sortase-dependent pili of Bifidobacterium bifidum PRL2010 in response to environmental gut conditions.* FEMS Microbiol Lett, **2014**. **357**(1): p. 23-33.
153. **Ventura, M., F. Turroni, G.A. Lugli, and D. van Sinderen,** *Bifidobacteria and humans: our special friends, from ecological to genomics perspectives.* J Sci Food Agric, **2014**. **94**(2): p. 163-8.
154. **Milani, C., G.A. Lugli, S. Duranti, F. Turroni, F. Bottacini, M. Mangifesta, B. Sanchez, A. Viappiani, L. Mancabelli, B. Taminiau, V. Delcenserie, R. Barrangou, A. Margolles, D. van Sinderen, and M. Ventura,** *Genomic encyclopedia of type strains of the genus Bifidobacterium.* Appl Environ Microbiol, **2014**. **80**(20): p. 6290-302.
155. **Toh, H., K. Oshima, A. Nakano, N. Yamashita, E. Iioka, R. Kurokawa, H. Morita, and M. Hattori,** *Complete Genome Sequence of Bifidobacterium scardovii Strain JCM 12489(T), Isolated from Human Blood.* Genome Announcements, **2015**. **3**(2): p. e00285-15.
156. **Milani, C., M. Mangifesta, L. Mancabelli, G.A. Lugli, K. James, S. Duranti, F. Turroni, C. Ferrario, M.C. Ossiprandi, D. van Sinderen, and M. Ventura,** *Unveiling bifidobacterial biogeography across the mammalian branch of the tree of life.* Isme j, **2017**. **11**(12): p. 2834-2847.
157. **Turroni, F., E. Foroni, P. Pizzetti, V. Giubellini, A. Ribbera, P. Merusi, P. Cagnasso, B. Bizzarri, G.L. de'Angelis, F. Shanahan, D. van Sinderen, and M. Ventura,** *Exploring the Diversity of the Bifidobacterial Population in the Human Intestinal Tract.* Applied and Environmental Microbiology, **2009**. **75**(6): p. 1534-1545.
158. **Odamaki, T., K. Kato, H. Sugahara, N. Hashikura, S. Takahashi, J.Z. Xiao, F. Abe, and R. Osawa,** *Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study.* BMC Microbiol, **2016**. **16**(90): p. 90.
159. **Favier, C.F., E.E. Vaughan, W.M. De Vos, and A.D. Akkermans,** *Molecular monitoring of succession of bacterial communities in human neonates.* Appl Environ Microbiol, **2002**. **68**(1): p. 219-26.
160. **Arboleya, S., C. Watkins, C. Stanton, and R.P. Ross,** *Gut Bifidobacteria Populations in Human Health and Aging.* Frontiers in Microbiology, **2016**. **7**(1204).

161. **Gueimonde, M., L. Debor, S. Tölkkö, E. Jokisalo, and S. Salminen,** *Quantitative assessment of faecal bifidobacterial populations by real-time PCR using lanthanide probes.* Journal of Applied Microbiology, **2007**. **102**(4): p. 1116-1122.

162. **Matsuki, T., K. Watanabe, J. Fujimoto, Y. Kado, T. Takada, K. Matsumoto, and R. Tanaka,** *Quantitative PCR with 16S rRNA-gene-targeted species-specific primers for analysis of human intestinal bifidobacteria.* Appl Environ Microbiol, **2004**. **70**(1): p. 167-73.

163. **Chaplin, A.V., A.G. Brzhozovskii, T.V. Parfenova, L.I. Kafarskaia, N.N. Volodin, A.N. Shkoporov, E.N. Ilina, and B.A. Efimov,** *[Species Diversity of Bifidobacteria in the Intestinal Microbiota Studied Using MALDI-TOF Mass-Spectrometry]*. Vestn Ross Akad Med Nauk, **2015**(4): p. 435-40.

164. **Gueimonde, M., A. Ouwehand, Pitk, auml, auml, Kaisu, T. Strandberg, H. Finne-Soveri, and S. Salminen,** *Fecal Bifidobacterium Levels in Elderly Nursing Home Patients: Are Levels as Expected?* Bioscience and Microflora, **2010**. **29**(2): p. 111-113.

165. **Biagi, E., M. Candela, S. Fairweather-Tait, C. Franceschi, and P. Brigidi,** *Aging of the human metaorganism: the microbial counterpart.* Age (Dordr), **2012**. **34**(1): p. 247-67.

166. **Claesson, M.J., I.B. Jeffery, and S. Conde,** *Gut microbiota composition correlates with diet and health in the elderly.* Nature, **2012**. **488(7410)**: p. 178–184.

167. **Voreades, N., A. Kozil, and T.L. Weir,** *Diet and the development of the human intestinal microbiome.* Front Microbiol, **2014**. **5**(494): p. 494.

168. **Sivan, A., L. Corrales, N. Hubert, J.B. Williams, K. Aquino-Michaels, Z.M. Earley, F.W. Benyamin, Y. Man Lei, B. Jabri, M.-L. Alegre, E.B. Chang, and T.F. Gajewski,** *Commensal Bifidobacterium promotes antitumor immunity and facilitates anti–PD-L1 efficacy.* Science, **2015**. **350**(6264): p. 1084-1089.

169. **Allen, A.P., G. Clarke, J.F. Cryan, E.M.M. Quigley, and T.G. Dinan,** *Bifidobacterium infantis 35624 and other probiotics in the management of irritable bowel syndrome. Strain specificity, symptoms, and mechanisms.* Curr Med Res Opin, **2017**. **33**(7): p. 1349-1351.

170. **Whorwell, P.J., L. Altringer, J. Morel, Y. Bond, D. Charbonneau, L. O'Mahony, B. Kiely, F. Shanahan, and E.M. Quigley,** *Efficacy of an encapsulated probiotic Bifidobacterium infantis 35624 in women with irritable bowel syndrome.* Am J Gastroenterol, **2006**. **101**(7): p. 1581-90.

171. **Makino, H., A. Kushiro, E. Ishikawa, H. Kubota, A. Gawad, T. Sakai, K. Oishi, R. Martin, K. Ben-Amor, J. Knol, and R. Tanaka,** *Mother-to-infant transmission of intestinal bifidobacterial strains has an impact on the early development of vaginally delivered infant's microbiota.* PLoS One, **2013**. **8**(11): p. e78331.

172. **Collado, M.C., S. Rautava, J. Aakko, E. Isolauri, and S. Salminen,** *Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid.* Scientific Reports, **2016**. **6**: p. 23129.

173. **Milani, C., L. Mancabelli, G.A. Lugli, S. Duranti, F. Turroni, C. Ferrario, M. Mangifesta, A. Viappiani, P. Ferretti, V. Gorfer, A. Tett, N. Segata, D. van Sinderen, and M. Ventura,** *Exploring Vertical Transmission of Bifidobacteria from Mother to Child.* Appl Environ Microbiol, **2015**. **81**(20): p. 7078-87.

174. **Barrett, E., C. Kerr, K. Murphy, O. O'Sullivan, C.A. Ryan, E.M. Dempsey, B.P. Murphy, P.W. O'Toole, P.D. Cotter, G.F. Fitzgerald, R.P. Ross, and C. Stanton,** *The individual-specific and diverse nature of the preterm infant microbiota.* Arch Dis Child Fetal Neonatal Ed, **2013**. **98**(4): p. F334-40.

175. **Benno, Y., K. Sawada, and T. Mitsuoka,** *The intestinal microflora of infants: composition of fecal flora in breast-fed and bottle-fed infants.* Microbiol Immunol, **1984**. **28**(9): p. 975-86.

176. **Solís, G., C.G. Los Reyes-Gavilan, N. Fernández, A. Margolles, and M. Gueimonde,** *Establishment and development of lactic acid bacteria and bifidobacteria microbiota in breast-milk and the infant gut.* Anaerobe, **2010**. **16**(3): p. 307-10.

177. **Duranti, S., G.A. Lugli, L. Mancabelli, F. Armanini, F. Turroni, K. James, P. Ferretti, V. Gorfer, C. Ferrario, C. Milani, M. Mangifesta, R. Anzalone, M. Zolfo, A. Viappiani, E. Pasolli, I. Bariletti, R. Canto, R. Clementi, M. Cologna, T. Crifo, G. Cusumano, S. Fedi, S. Gottardi, C. Innamorati, C. Mase, D. Postai, D. Savoi, M. Soffiati, S. Tateo, A. Pedrotti, N. Segata, D. van Sinderen, and M. Ventura,** *Maternal inheritance of bifidobacterial communities and bifidophages in infants through vertical transmission.* Microbiome, **2017**. **5**(1): p. 66.

178. **Vazquez-Gutierrez, P., C. Lacroix, C. Chassard, J. Klumpp, C. Jans, and M.J.A. Stevens,** *Complete and Assembled Genome Sequence of Bifidobacterium kashiwanohense PV20-2, Isolated from the Feces of an Anemic Kenyan Infant.* Genome Announcements, **2015**. **3**(1): p. e01467-14.

179. **Morita, H., A. Nakano, H. Onoda, H. Toh, K. Oshima, H. Takami, M. Murakami, S. Fukuda, T. Takizawa, T. Kuwahara, H. Ohno, S. Tanabe, and M. Hattori,**

Bifidobacterium kashiwanohense sp. nov., isolated from healthy infant faeces. Int J Syst Evol Microbiol, **2011**. **61**(Pt 11): p. 2610-5.

180. **Mevissen-Verhage, E.A., J.H. Marcelis, M.N. de Vos, W.C. Harmsen-van Amerongen, and J. Verhoef,** *Bifidobacterium, Bacteroides, and Clostridium spp. in fecal samples from breast-fed and bottle-fed infants with and without iron supplement.* J Clin Microbiol, **1987**. **25**(2): p. 285-9.
181. **Klaassens, E.S., R.J. Boesten, M. Haarman, J. Knol, F.H. Schuren, E.E. Vaughan, and W.M. de Vos,** *Mixed-Species Genomic Microarray Analysis of Fecal Samples Reveals Differential Transcriptional Responses of Bifidobacteria in Breast- and Formula-Fed Infants.* Applied and Environmental Microbiology, **2009**. **75**(9): p. 2668-2676.
182. **Wall, R., R.P. Ross, C.A. Ryan, S. Hussey, B. Murphy, G.F. Fitzgerald, and C. Stanton,** *Role of Gut Microbiota in Early Infant Development.* Clinical Medicine. Pediatrics, **2009**. **3**: p. 45-54.
183. **Roberfroid, M., G.R. Gibson, L. Hoyles, A.L. McCartney, R. Rastall, I. Rowland, D. Wolvers, B. Watzl, H. Szajewska, B. Stahl, F. Guarner, F. Respondek, K. Whelan, V. Coxam, M.-J. Davicco, L. Léotoing, Y. Wittrant, N.M. Delzenne, P.D. Cani, A.M. Neyrinck, and A. Meheust,** *Prebiotic effects: metabolic and health benefits.* British Journal of Nutrition, **2010**. **104**(S2): p. S1-S63.
184. **Roy, C.C., C.L. Kien, L. Bouthillier, and E. Levy,** *Short-chain fatty acids: ready for prime time?* Nutr Clin Pract, **2006**. **21**(4): p. 351-66.
185. **Hinnebusch, B.F., S. Meng, J.T. Wu, S.Y. Archer, and R.A. Hodin,** *The effects of short-chain fatty acids on human colon cancer cell phenotype are associated with histone hyperacetylation.* J Nutr, **2002**. **132**(5): p. 1012-7.
186. **Galvez, J., M.E. Rodriguez-Cabezas, and A. Zarzuelo,** *Effects of dietary fiber on inflammatory bowel disease.* Mol Nutr Food Res, **2005**. **49**(6): p. 601-8.
187. **Lievin, V., I. Peiffer, S. Hudault, F. Rochat, D. Brassart, J. Neeser, and A. Servin,** *Bifidobacterium strains from resident infant human gastrointestinal microflora exert antimicrobial activity.* Gut, **2000**. **47**(5): p. 646-652.
188. **Toure, R., E. Kheadr, C. Lacroix, O. Moroni, and I. Fliss,** *Production of antibacterial substances by bifidobacterial isolates from infant stool active against Listeria monocytogenes.* J Appl Microbiol, **2003**. **95**(5): p. 1058-69.
189. **Simone, M., C. Gozzoli, A. Quartieri, G. Mazzola, D. Di Gioia, A. Amaretti, S. Raimondi, and M. Rossi,** *The Probiotic Bifidobacterium breve B632 Inhibited the*

Growth of Enterobacteriaceae within Colicky Infant Microbiota Cultures. BioMed Research International, **2014**. **2014**(301053**)**: p. 7.

190. **Chichlowski, M., G. De Lartigue, J.B. German, H.E. Raybould, and D.A. Mills,** *Bifidobacteria isolated from infants and cultured on human milk oligosaccharides affect intestinal epithelial function.* J Pediatr Gastroenterol Nutr, **2012**. **55**(3): p. 321-7.
191. **Underwood, M.A., J.B. German, C.B. Lebrilla, and D.A. Mills,** *Bifidobacterium longum subspecies infantis: champion colonizer of the infant gut.* Pediatr Res, **2015**. **77**(1-2): p. 229-35.
192. **De Angelis, M., C.G. Rizzello, A. Fasano, M.G. Clemente, C. De Simone, M. Silano, M. De Vincenzi, I. Losito, and M. Gobbetti,** *VSL#3 probiotic preparation has the capacity to hydrolyze gliadin polypeptides responsible for Celiac Sprue.* Biochim Biophys Acta, **2006**. **1762**(1): p. 80-93.
193. **An, H.M., S.Y. Park, D.K. Lee, J.R. Kim, M.K. Cha, S.W. Lee, H.T. Lim, K.J. Kim, and N.J. Ha,** *Antiobesity and lipid-lowering effects of Bifidobacterium spp. in high fat diet-induced obese rats.* Lipids in Health and Disease, **2011**. **10**(116): p. 116.
194. **Bercik, P., A.J. Park, D. Sinclair, A. Khoshdel, J. Lu, X. Huang, Y. Deng, P.A. Blennerhassett, M. Fahnestock, D. Moine, B. Berger, J.D. Huizinga, W. Kunze, P.G. McLean, G.E. Bergonzelli, S.M. Collins, and E.F. Verdu,** *The anxiolytic effect of Bifidobacterium longum NCC3001 involves vagal pathways for gut–brain communication.* Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society, **2011**. **23**(12): p. 1132-1139.
195. **Di Gioia, D., I. Aloisio, G. Mazzola, and B. Biavati,** *Bifidobacteria: their impact on gut microbiota composition and their applications as probiotics in infants.* Appl Microbiol Biotechnol, **2014**. **98**(2): p. 563-77.
196. **Kuitunen, M., K. Kukkonen, K. Juntunen-Backman, R. Korpela, T. Poussa, T. Tuure, T. Haahtela, and E. Savilahti,** *Probiotics prevent IgE-associated allergy until age 5 years in cesarean-delivered children but not in the total cohort.* J Allergy Clin Immunol, **2009**. **123**(2): p. 335-41.
197. **Aloisio, I., C. Santini, B. Biavati, G. Dinelli, A. Cencic, W. Chingwaru, L. Mogna, and D. Di Gioia,** *Characterization of Bifidobacterium spp. strains for the treatment of enteric disorders in newborns.* Appl Microbiol Biotechnol, **2012**. **96**(6): p. 1561-76.
198. **Lin, H.C., C.H. Hsu, H.L. Chen, M.Y. Chung, J.F. Hsu, R.I. Lien, L.Y. Tsao, C.H. Chen, and B.H. Su,** *Oral probiotics prevent necrotizing enterocolitis in very low birth*

weight preterm infants: a multicenter, randomized, controlled trial. Pediatrics, **2008**. **122**(4): p. 693-700.

199. **Repa, A., M. Thanhaeuser, D. Endress, M. Weber, A. Kreissl, C. Binder, A. Berger, and N. Haiden,** *Probiotics (Lactobacillus acidophilus and Bifidobacterium infantis) prevent NEC in VLBW infants fed breast milk but not formula [corrected]*. Pediatr Res, **2015**. **77**(2): p. 381-8.
200. **Totsu, S., C. Yamasaki, M. Terahara, A. Uchiyama, and S. Kusuda,** *Bifidobacterium and enteral feeding in preterm infants: cluster-randomized trial.* Pediatr Int, **2014**. **56**(5): p. 714-9.
201. **Yamasaki, C., S. Totsu, A. Uchiyama, H. Nakanishi, K. Masumoto, Y. Washio, K. Shuri, S. Ishida, K. Imai, and S. Kusuda,** *Effect of Bifidobacterium administration on very-low-birthweight infants.* Pediatr Int, **2012**. **54**(5): p. 651-6.
202. **Milani, C., G.A. Lugli, S. Duranti, F. Turroni, L. Mancabelli, C. Ferrario, M. Mangifesta, A. Hevia, A. Viappiani, M. Scholz, S. Arioli, B. Sanchez, J. Lane, D.V. Ward, R. Hickey, D. Mora, N. Segata, A. Margolles, D. van Sinderen, and M. Ventura,** *Bifidobacteria exhibit social behavior through carbohydrate resource sharing in the gut.* Scientific Reports, **2015**. **5**(15782): p. 15782.
203. **Bond, J.H., B.E. Currier, H. Buchwald, and M.D. Levitt,** *Colonic conservation of malabsorbed carbohydrate.* Gastroenterology, **1980**. **78**(3): p. 444-7.
204. **Englyst, H.N., S.M. Kingman, and J.H. Cummings,** *Classification and measurement of nutritionally important starch fractions.* Eur J Clin Nutr, **1992**. **46 Suppl 2**: p. S33-50.
205. **Lattimer, J.M. and M.D. Haub,** *Effects of Dietary Fiber and Its Components on Metabolic Health.* Nutrients, **2010**. **2**(12): p. 1266-1289.
206. **Salyers, A.A., J.R. Vercellotti, S.E. West, and T.D. Wilkins,** *Fermentation of mucin and plant polysaccharides by strains of Bacteroides from the human colon.* Applied and Environmental Microbiology, **1977**. **33**(2): p. 319-322.
207. **Macfarlane, S. and G.T. Macfarlane,** *Regulation of short-chain fatty acid production.* Proc Nutr Soc, **2003**. **62**(1): p. 67-72.
208. **de Vries, W. and A.H. Stouthamer,** *Pathway of glucose fermentation in relation to the taxonomy of bifidobacteria.* J Bacteriol, **1967**. **93**(2): p. 574-6.
209. **Scardovi V, T.L.,** *The fructose 6 phosphate shunt as peculiar pattern of hexose degradation in the genus Bifidobacterium.* Ann Microbiol **1965**(15): p. 19–29.

210. **Rada, V.,** *Detection of Bifidobacteriumspecies by enzymatic methods and antimicrobial susceptibility testing.* Biotechnology Techniques, **1997**. **11**(12): p. 909-912.

211. **Palframan, R.J., G.R. Gibson, and R.A. Rastall,** *Carbohydrate preferences of Bifidobacterium species isolated from the human gut.* Curr Issues Intest Microbiol, **2003**. **4**(2): p. 71-5.

212. **Postma, P.W., J.W. Lengeler, and G.R. Jacobson,** *Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria.* Microbiol Rev, **1993**. **57**(3): p. 543-94.

213. **Pokusaeva, K., G.F. Fitzgerald, and D. van Sinderen,** *Carbohydrate metabolism in Bifidobacteria.* Genes & Nutrition, **2011**. **6**(3): p. 285-306.

214. **Ventura, M., F. Turroni, A. Zomer, E. Foroni, V. Giubellini, F. Bottacini, C. Canchaya, M.J. Claesson, F. He, M. Mantzourani, L. Mulas, A. Ferrarini, B. Gao, M. Delledonne, B. Henrissat, P. Coutinho, M. Oggioni, R.S. Gupta, Z. Zhang, D. Beighton, G.F. Fitzgerald, P.W. O'Toole, and D. van Sinderen,** *The Bifidobacterium dentium Bd1 Genome Sequence Reflects Its Genetic Adaptation to the Human Oral Cavity.* PLOS Genetics, **2009**. **5**(12): p. e1000785.

215. **Schell, M.A., M. Karmirantzou, B. Snel, D. Vilanova, B. Berger, G. Pessi, M.-C. Zwahlen, F. Desiere, P. Bork, M. Delley, R.D. Pridmore, and F. Arigoni,** *The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract.* Proceedings of the National Academy of Sciences, **2002**. **99**(22): p. 14422-14427.

216. **Parche, S., J. Amon, I. Jankovic, E. Rezzonico, M. Beleut, H. Barutcu, I. Schendel, M.P. Eddy, A. Burkovski, F. Arigoni, and F. Titgemeyer,** *Sugar transport systems of Bifidobacterium longum NCC2705.* J Mol Microbiol Biotechnol, **2007**. **12**(1-2): p. 9-19.

217. **Schneider, E.,** *ABC transporters catalyzing carbohydrate uptake.* Res Microbiol, **2001**. **152**(3-4): p. 303-10.

218. **Deutscher, J., C. Francke, and P.W. Postma,** *How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria.* Microbiol Mol Biol Rev, **2006**. **70**(4): p. 939-1031.

219. **Lengeler, J.W., R.J. Mayer, and K. Schmid,** *Phosphoenolpyruvate-dependent phosphotransferase system enzyme III and plasmid-encoded sucrose transport in Escherichia coli K-12.* J Bacteriol, **1982**. **151**(1): p. 468-71.

220. **Cantarel, B.L., P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, and B. Henrissat,** *The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics.* Nucleic Acids Research, **2009**. **37**(Database issue): p. D233-D238.

221. **Henrissat, B.,** *A classification of glycosyl hydrolases based on amino acid sequence similarities.* Biochemical Journal, **1991**. **280**(Pt 2): p. 309-316.

222. **Lombard, V., H. Golaconda Ramulu, E. Drula, P.M. Coutinho, and B. Henrissat,** *The carbohydrate-active enzymes database (CAZy) in 2013.* Nucleic Acids Research, **2014**. **42**(Database issue): p. D490-D495.

223. **Gorke, B. and J. Stulke,** *Carbon catabolite repression in bacteria: many ways to make the most out of nutrients.* Nat Rev Microbiol, **2008**. **6**(8): p. 613-24.

224. **Stulke, J. and W. Hillen,** *Carbon catabolite repression in bacteria.* Curr Opin Microbiol, **1999**. **2**(2): p. 195-201.

225. **Cohn, M. and K. Horibata,** *Inhibition by glucose of the induced synthesis of the β-galactoside-enzyme system of Escherichia coli. Analysis of maintenance.* Journal of Bacteriology, **1959**. **78**(5): p. 601-612.

226. **Dills, S.S., A. Apperson, M.R. Schmidt, and M.H. Saier,** *Carbohydrate transport in bacteria.* Microbiological Reviews, **1980**. **44**(3): p. 385-418.

227. **Pokusaeva, K., A.R. Neves, A. Zomer, M. O'Connell-Motherway, J. MacSharry, P. Curley, G.F. Fitzgerald, and D. van Sinderen,** *Ribose utilization by the human commensal Bifidobacterium breve UCC2003.* Microb Biotechnol, **2010**. **3**(3): p. 311-23.

228. **Parche, S., M. Beleut, E. Rezzonico, D. Jacobs, F. Arigoni, F. Titgemeyer, and I. Jankovic,** *Lactose-over-Glucose Preference in Bifidobacterium longum NCC2705: glcP, Encoding a Glucose Transporter, Is Subject to Lactose Repression.* Journal of Bacteriology, **2006**. **188**(4): p. 1260-1265.

229. **Sela, D.A., D. Garrido, L. Lerno, S. Wu, K. Tan, H.-J. Eom, A. Joachimiak, C.B. Lebrilla, and D.A. Mills,** *Bifidobacterium longum subsp. infantis ATCC 15697 α-Fucosidases Are Active on Fucosylated Human Milk Oligosaccharides.* Applied and Environmental Microbiology, **2012**. **78**(3): p. 795-803.

230. **Gilbert, W. and B. Müller-Hill,** *ISOLATION OF THE LAC REPRESSOR.* Proceedings of the National Academy of Sciences of the United States of America, **1966**. **56**(6): p. 1891-1898.

231. **Jacob, F. and J. Monod,** *Genetic regulatory mechanisms in the synthesis of proteins.* Journal of Molecular Biology, **1961**. **3**(3): p. 318-356.

232. **Matthews, K.S. and J.C. Nichols,** *Lactose repressor protein: functional properties and structure.* Prog Nucleic Acid Res Mol Biol, **1998**. **58**: p. 127-64.

233. **Swint-Kruse, L. and K.S. Matthews,** *Allostery in the LacI/GalR Family: Variations on a Theme.* Current opinion in microbiology, **2009**. **12**(2): p. 129-137.

234. **Weickert, M.J. and S. Adhya,** *The galactose regulon of Escherichia coli.* Mol Microbiol, **1993**. **10**(2): p. 245-51.

235. **Weickert, M.J. and S. Adhya,** *A family of bacterial regulators homologous to Gal and Lac repressors.* J Biol Chem, **1992**. **267**(22): p. 15869-74.

236. **Ravcheev, D.A., M.S. Khoroshkin, O.N. Laikova, O.V. Tsoy, N.V. Sernova, S.A. Petrova, A.B. Rakhmaninova, P.S. Novichkov, M.S. Gelfand, and D.A. Rodionov,** *Comparative genomics and evolution of regulons of the LacI-family transcription factors.* Frontiers in Microbiology, **2014**. **5**(294).

237. **Khoroshkin, M.S., S.A. Leyn, D. Van Sinderen, and D.A. Rodionov,** *Transcriptional Regulation of Carbohydrate Utilization Pathways in the Bifidobacterium Genus.* Frontiers in Microbiology, **2016**. **7**: p. 120.

238. **Ryan, S.M., G.F. Fitzgerald, and D. van Sinderen,** *Transcriptional Regulation and Characterization of a Novel β-Fructofuranosidase-Encoding Gene from Bifidobacterium breve UCC2003.* Applied and Environmental Microbiology, **2005**. **71**(7): p. 3475-3482.

239. **O'Connell Motherway, M., G.F. Fitzgerald, and D. van Sinderen,** *Metabolism of a plant derived galactose-containing polysaccharide by Bifidobacterium breve UCC2003.* Microbial Biotechnology, **2011**. **4**(3): p. 403-416.

240. **Pokusaeva, K., M. O'Connell-Motherway, A. Zomer, J. Macsharry, G.F. Fitzgerald, and D. van Sinderen,** *Cellodextrin utilization by bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2011**. **77**(5): p. 1681-90.

241. **O'Connell, K.J., M.O. Motherway, A. Liedtke, G.F. Fitzgerald, R. Paul Ross, C. Stanton, A. Zomer, and D. van Sinderen,** *Transcription of two adjacent carbohydrate utilization gene clusters in Bifidobacterium breve UCC2003 is controlled by LacI- and repressor open reading frame kinase (ROK)-type regulators.* Appl Environ Microbiol, **2014**. **80**(12): p. 3604-14.

242. **Egan, M., H. Jiang, M. O'Connell Motherway, S. Oscarson, and D. van Sinderen,** *Glycosulfatase-Encoding Gene Cluster in Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2016**. **82**(22): p. 6611-6623.

243. **Gilbert, W. and A. Maxam,** *The Nucleotide Sequence of the lac Operator.* Proceedings of the National Academy of Sciences of the United States of America, **1973**. **70**(12 Pt 1-2): p. 3581-3584.

244. **Egan, M., M. O'Connell Motherway, and D. van Sinderen,** *A GntR-type transcriptional repressor controls sialic acid utilization in Bifidobacterium breve UCC2003.* FEMS Microbiol Lett, **2015**. **362**(4).

245. **James, K., M.O.C. Motherway, C. Penno, R.L. O'Brien, and D. van Sinderen,** *Bifidobacterium breve UCC2003 employs multiple transcriptional regulators to control metabolism of particular human milk oligosaccharides.* Applied and Environmental Microbiology, **2018**. **84**(9).

246. **Locascio, R.G., M.R. Ninonuevo, S.R. Kronewitter, S.L. Freeman, J.B. German, C.B. Lebrilla, and D.A. Mills,** *A versatile and scalable strategy for glycoprofiling bifidobacterial consumption of human milk oligosaccharides.* Microb Biotechnol, **2009**. **2**(3): p. 333-42.

247. **Ruiz-Moyano, S., S.M. Totten, D.A. Garrido, J.T. Smilowitz, J.B. German, C.B. Lebrilla, and D.A. Mills,** *Variation in Consumption of Human Milk Oligosaccharides by Infant Gut-Associated Strains of Bifidobacterium breve.* Applied and Environmental Microbiology, **2013**. **79**(19): p. 6040-6049.

248. **Sela, D.A., Y. Li, L. Lerno, S. Wu, A.M. Marcobal, J.B. German, X. Chen, C.B. Lebrilla, and D.A. Mills,** *An Infant-associated Bacterial Commensal Utilizes Breast Milk Sialyloligosaccharides.* The Journal of Biological Chemistry, **2011**. **286**(14): p. 11909-11918.

249. **Bunesova, V., C. Lacroix, and C. Schwab,** *Fucosyllactose and L-fucose utilization of infant Bifidobacterium longum and Bifidobacterium kashiwanohense.* BMC Microbiology, **2016**. **16**(1): p. 248.

250. **Schwab, C., H.-J. Ruscheweyh, V. Bunesova, V.T. Pham, N. Beerenwinkel, and C. Lacroix,** *Trophic Interactions of Infant Bifidobacteria and Eubacterium hallii during L-Fucose and Fucosyllactose Degradation.* Frontiers in Microbiology, **2017**. **8**: p. 95.

251. **Katayama, T.,** *Host-derived glycans serve as selected nutrients for the gut microbe: human milk oligosaccharides and bifidobacteria.* Bioscience, Biotechnology, and Biochemistry, **2016**. **80**(4): p. 621-632.

252. **Xiao, J.Z., S. Takahashi, M. Nishimoto, T. Odamaki, T. Yaeshima, K. Iwatsuki, and M. Kitaoka,** *Distribution of in vitro fermentation ability of lacto-N-biose I, a*

major building block of human milk oligosaccharides, in bifidobacterial strains. Appl Environ Microbiol, **2010**. **76**(1): p. 54-9.

253. **Nishimoto, M. and M. Kitaoka,** *Identification of N-Acetylhexosamine 1-Kinase in the Complete Lacto-N-Biose I/Galacto-N-Biose Metabolic Pathway in Bifidobacterium longum.* Applied and Environmental Microbiology, **2007**. **73**(20): p. 6444-6449.
254. **Fushinobu, S.,** *Unique sugar metabolic pathways of bifidobacteria.* Biosci Biotechnol Biochem, **2010**. **74**(12): p. 2374-84.
255. **Kitaoka, M., J. Tian, and M. Nishimoto,** *Novel putative galactose operon involving lacto-N-biose phosphorylase in Bifidobacterium longum.* Appl Environ Microbiol, **2005**. **71**(6): p. 3158-62.
256. **Kitaoka, M. and K. Hayashi,** *Carbohydrate-Processing Phosphorolytic Enzymes.* Trends in Glycoscience and Glycotechnology, **2002**. **14**(75): p. 35-50.
257. **Kitaoka, M.,** *Bifidobacterial enzymes involved in the metabolism of human milk oligosaccharides.* Adv Nutr, **2012**. **3**(3): p. 422s-9s.
258. **Sela, D.A.,** *Bifidobacterial utilization of human milk oligosaccharides.* Int J Food Microbiol, **2011**. **149**(1): p. 58-64.
259. **James, K., M.O. Motherway, F. Bottacini, and D. van Sinderen,** *Bifidobacterium breve UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways.* Sci Rep, **2016**. **6**: p. 38560.
260. **Turroni, F., F. Bottacini, E. Foroni, I. Mulder, J.-H. Kim, A. Zomer, B. Sánchez, A. Bidossi, A. Ferrarini, V. Giubellini, M. Delledonne, B. Henrissat, P. Coutinho, M. Oggioni, G.F. Fitzgerald, D. Mills, A. Margolles, D. Kelly, D. van Sinderen, and M. Ventura,** *Genome analysis of Bifidobacterium bifidum PRL2010 reveals metabolic pathways for host-derived glycan foraging.* Proceedings of the National Academy of Sciences of the United States of America, **2010**. **107**(45): p. 19514-19519.
261. **Zhurina, D., A. Zomer, M. Gleinser, V.F. Brancaccio, M. Auchter, M.S. Waidmann, C. Westermann, D. van Sinderen, and C.U. Riedel,** *Complete Genome Sequence of Bifidobacterium bifidum S17.* Journal of Bacteriology, **2011**. **193**(1): p. 301-302.
262. **Katayama, T., A. Sakuma, T. Kimura, Y. Makimura, J. Hiratake, K. Sakata, T. Yamanoi, H. Kumagai, and K. Yamamoto,** *Molecular cloning and characterization of Bifidobacterium bifidum 1,2-alpha-L-fucosidase (AfcA), a novel inverting glycosidase (glycoside hydrolase family 95).* J Bacteriol, **2004**. **186**(15): p. 4885-93.

263. **Ashida, H., A. Miyake, M. Kiyohara, J. Wada, E. Yoshida, H. Kumagai, T. Katayama, and K. Yamamoto,** *Two distinct alpha-L-fucosidases from Bifidobacterium bifidum are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates.* Glycobiology, **2009**. **19**(9): p. 1010-7.

264. **Nagae, M., A. Tsuchiya, T. Katayama, K. Yamamoto, S. Wakatsuki, and R. Kato,** *Structural basis of the catalytic reaction mechanism of novel 1,2-alpha-L-fucosidase from Bifidobacterium bifidum.* J Biol Chem, **2007**. **282**(25): p. 18497-509.

265. **Kiyohara, M., K. Tanigawa, T. Chaiwangsri, T. Katayama, H. Ashida, and K. Yamamoto,** *An exo-alpha-sialidase from bifidobacteria involved in the degradation of sialyloligosaccharides in human milk and intestinal glycoconjugates.* Glycobiology, **2011**. **21**(4): p. 437-47.

266. **Katayama, T., K. Fujita, and K. Yamamoto,** *Novel bifidobacterial glycosidases acting on sugar chains of mucin glycoproteins.* J Biosci Bioeng, **2005**. **99**(5): p. 457-65.

267. **Marcobal, A. and J.L. Sonnenburg,** *Human milk oligosaccharide consumption by intestinal microbiota.* Clin Microbiol Infect, **2012**. **18 Suppl 4**: p. 12-5.

268. **Egan, M., M. O'Connell Motherway, M. Ventura, and D. van Sinderen,** *Metabolism of sialic acid by Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2014**. **80**(14): p. 4414-26.

269. **Wada, J., T. Ando, M. Kiyohara, H. Ashida, M. Kitaoka, M. Yamaguchi, H. Kumagai, T. Katayama, and K. Yamamoto,** *Bifidobacterium bifidum lacto-N-biosidase, a critical enzyme for the degradation of human milk oligosaccharides with a type 1 structure.* Appl Environ Microbiol, **2008**. **74**(13): p. 3996-4004.

270. **Miwa, M., T. Horimoto, M. Kiyohara, T. Katayama, M. Kitaoka, H. Ashida, and K. Yamamoto,** *Cooperation of beta-galactosidase and beta-N-acetylhexosaminidase from bifidobacteria in assimilation of human milk oligosaccharides with type 2 structure.* Glycobiology, **2010**. **20**(11): p. 1402-9.

271. **Moller, P.L., F. Jorgensen, O.C. Hansen, S.M. Madsen, and P. Stougaard,** *Intra- and extracellular beta-galactosidases from Bifidobacterium bifidum and B. infantis: molecular cloning, heterologous expression, and comparative characterization.* Appl Environ Microbiol, **2001**. **67**(5): p. 2276-83.

272. **Garrido, D., J.H. Kim, J.B. German, H.E. Raybould, and D.A. Mills,** *Oligosaccharide Binding Proteins from Bifidobacterium longum subsp. infantis Reveal a Preference for Host Glycans.* PLoS ONE, **2011**. **6**(3): p. e17315.

273. **Sela, D.A., J. Chapman, A. Adeuya, J.H. Kim, F. Chen, T.R. Whitehead, A. Lapidus, D.S. Rokhsar, C.B. Lebrilla, J.B. German, N.P. Price, P.M. Richardson, and D.A. Mills,** *The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome.* Proc Natl Acad Sci U S A, **2008**. **105**(48): p. 18964-9.

274. **Yoshida, E., H. Sakurama, M. Kiyohara, M. Nakajima, M. Kitaoka, H. Ashida, J. Hirose, T. Katayama, K. Yamamoto, and H. Kumagai,** *Bifidobacterium longum subsp. infantis uses two different beta-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides.* Glycobiology, **2012**. **22**(3): p. 361-8.

275. **Garrido, D., S. Ruiz-Moyano, and D.A. Mills,** *Release and utilization of N-acetyl-D-glucosamine from human milk oligosaccharides by Bifidobacterium longum subsp. infantis.* Anaerobe, **2012**. **18**(4): p. 430-5.

276. **Zivkovic, A.M., J.B. German, C.B. Lebrilla, and D.A. Mills,** *Human milk glycobiome and its impact on the infant gastrointestinal microbiota.* Proc Natl Acad Sci U S A, **2011**. **108 Suppl 1**: p. 4653-8.

277. **Garrido, D., S. Ruiz-Moyano, N. Kirmiz, J.C. Davis, S.M. Totten, D.G. Lemay, J.A. Ugalde, J.B. German, C.B. Lebrilla, and D.A. Mills,** *A novel gene cluster allows preferential utilization of fucosylated milk oligosaccharides in Bifidobacterium longum subsp. longum SC596.* Sci Rep, **2016**. **6**: p. 35045.

278. **Egan, M., M.O. Motherway, M. Kilcoyne, M. Kane, L. Joshi, M. Ventura, and D. van Sinderen,** *Cross-feeding by Bifidobacterium breve UCC2003 during co-cultivation with Bifidobacterium bifidum PRL2010 in a mucin-based medium.* BMC Microbiol, **2014**. **14**(282): p. 282.

279. **Matsuki, T., K. Yahagi, H. Mori, H. Matsumoto, T. Hara, S. Tajima, E. Ogawa, H. Kodama, K. Yamamoto, T. Yamada, S. Matsumoto, and K. Kurokawa,** *A key genetic factor for fucosyllactose utilization affects infant gut microbiota development.* Nat Commun, **2016**. **7**: p. 11939.

280. **Sakurama, H., M. Kiyohara, J. Wada, Y. Honda, M. Yamaguchi, S. Fukiya, A. Yokota, H. Ashida, H. Kumagai, M. Kitaoka, K. Yamamoto, and T. Katayama,** *Lacto-N-biosidase Encoded by a Novel Gene of Bifidobacterium longum Subspecies longum Shows Unique Substrate Specificity and Requires a Designated Chaperone for Its Active Expression.* The Journal of Biological Chemistry, **2013**. **288**(35): p. 25194-25206.

Chapter II

Bifidobacterium breve UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways.

Dr. Francesca Bottacini carried out the bioinformatics analysis.

This chapter was published in:

James, K., O'Connell Motherway, M., Bottacini, F. & van Sinderen, D. (2016). *Bifidobacterium breve* UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways. Sci Rep, **6**: p.38560.

2.1 Abstract

In this study, we demonstrate that the *B. breve* strain UCC2003 possesses specific metabolic pathways for the utilisation of lacto-N-tetraose (LNT) and lacto-N-neotetraose (LNnT), which represent the backbone moieties of Type I and Type II human milk oligosaccharides (HMOs), respectively. Using a combination of experimental approaches, the enzymatic machinery involved in the metabolism of LNT and LNnT was identified and characterised. Homologs of the key genetic loci involved in the utilisation of these HMO substrates were identified in *B. breve*, *B. bifidum*, *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* using bioinformatic analyses, and were shown to be variably present among other members of the *Bifidobacterium* genus, with a distinct pattern of conservation among human-associated bifidobacterial species.

2.2 Introduction

Consumption of maternal breast milk, or the lack thereof, influences the gut microbiota composition of the neonate [1-3]. Incorrect development or disruption of this microbial community contributes to disorders such as Necrotising Enterocolitis, infantile diarrhoea and Group B streptococcal neonatal infection [4-8]. Strikingly, the faecal microbiota of healthy breastfed infants is enriched for certain species of the *Bifidobacterium* genus [9], which are high-G+C Gram-positive anaerobes and members of the *Actinobacteria* phylum. Naturally found as symbionts of the mammalian, avian or insect digestive tract, bifidobacteria enjoy substantial scientific attention due to their purported beneficial properties [2, 10-16].

While lactose (Galβ1-4Glc) comprises the main carbohydrate component of human breast milk and colostrum (~90 %), human milk oligosaccharides (HMOs) constitute the next most significant carbohydrate fraction, ahead of glycolipids [2, 17], and are typically found at a concentration of ≥4 g/L (and as high as 15 g/L) [2, 18-20]. HMOs represent a heterogeneous glycan mix, of which > 200 distinct structures have been identified [18]. The majority of these HMO structures are classified into 2 types. The abundant Type I HMOs contain lacto-N-tetraose (LNT; Fig. 2.1) (Galβ1-3GlcNAcβ1-3Galβ1-4Glc), which is composed of a lactose coupled to lacto-N-biose (LNB) (Galβ1-3GlcNAc). Type II HMOs contain the LNT isomer lacto-N-neotetraose (LNnT; Fig. 2.1) (Galβ1-4GlcNAcβ1-3Galβ1-4Glc), which is composed of lactose linked to N-acetyllactosamine (LacNAc) (Galβ1-4GlcNAc), an isomer of LNB. Larger Type I and II HMOs may contain further LNB or LacNAc subunits, and can be fucosylated or sialylated [2, 18, 21].

Despite the abundance of HMOs in breast milk, these glycans cannot be metabolised by the infant, and it is currently believed that they facilitate the establishment of an infant-specific gut microbiota, with bifidobacteria being particularly abundant [16, 22]. Common among the latter are *Bifidobacterium bifidum*, *Bifidobacterium longum* subsp. *infantis* and subsp. *longum*, *Bifidobacterium breve*, *Bifidobacterium pseudocatenulatum* and *Bifidobacterium kashiwanohense*[9, 23-27]. Unsurprisingly, it has been

shown that certain bifidobacterial species can metabolize (particular) HMOs [18, 28-34]. Previous studies have elucidated some of the metabolic pathways for HMO utilisation by *B. bifidum* and *B. longum* subsp. *infantis*, with particular focus on LNT and LNnT [28]. *B. longum* subsp. *infantis* internalises particular, intact small-mass HMOs, including (precursors of) LN(n)T [18, 28, 35], which are in turn hydrolysed into lacto-N-triose and galactose, by two HMO type-specific β-galactosidases (i.e. one enzyme acting on LNT, the other on LNnT) [29]. Lacto-N-triose is further hydrolysed by an N-acetylhexosaminidase into N-acetylglucosamine (GlcNAc) and lactose, the latter of which is then hydrolysed by a β-galactosidase [30] into galactose and glucose, to enter the Leloir and fructose-6-phosphate (F6P) phosphoketolase pathways and amino-sugar metabolising pathway (for GlcNAc) [28, 31], all of which feed into the overall *Bifidobacteriaceae*-specific metabolic pathway known as the Bifid Shunt.

B. bifidum possesses two distinct, and apparently unique pathways for the metabolism of LN(n)T. Large type I and II HMOs are degraded by extracellular fucosidases, sialidases and glycosyl hydrolases to release LNT and LNnT [28]. LNT is hydrolysed at its central β-1,3-link by an extracellular glycosyl hydrolase into lactose and LNB, the latter being transported into the cell and then degraded by two distinct LNB phosphorylases (LNBP), releasing galactose 1-phosphate and GlcNAc [32, 33]. The released lactose is either hydrolysed by an extracellular β-galactosidase into galactose and glucose (which are both internalized by the cell), or transported into the cell, where it is similarly hydrolysed by intracellular β-galactosidases [28, 32, 36]. These monosaccharides are then further metabolised by the same pathways as those described for *B. longum* subsp. *infantis*. This type I HMO metabolism has also been observed in some species of *B. longum* subsp. *longum* [32, 37]. In addition, *B. bifidum* possesses a separate pathway to degrade and utilise LNnT. An extracellular β-galactosidase cleaves LNnT at its Galβ-1,4 residue, liberating galactose and lacto-N-triose [34]. The lacto-N-triose is then further hydrolysed by an extracellular N-acetylhexosaminidase, releasing GlcNAc and lactose, with the latter further hydrolysed by the aforementioned extracellular β-galactosidases into glucose and galactose (which are

transported into the cell) [36], or internalised and then degraded as described. Once within the cell, these monosaccharides are metabolised as mentioned above [34].

It should be noted that a specific pathway exists for LNB metabolism, known as the GNB/LNB pathway. In *B. bifidum*, as mentioned above, LNB is phosphorolysed into monosaccharides by either of two different LNBP enzymes, while in *B. infantis*, LNB is phosphorolysed by a single LNBP enzyme, whose gene shares homology with both *B. bifidum* LNBP genes [28, 31]. This GNB/LNB pathway appears to be present in bifidobacterial species commonly found in infant faeces [28]. The apparent absence of this GNB/LNB pathway and, specifically, the LNBP-encoding gene in adult-associated bifidobacteria (such as *B. adolescentis*) is manifested through their inability to utilise LNB or other HMOs as a carbon source for growth, and may therefore explain, at least in part, their absence or low abundance in the microbiota of breast-fed infants[18].

Little information exists regarding HMO utilisation by *B. breve*, although it has been suggested that *B. breve* acts as a 'scavenger' through cross-feeding on HMO-derived monosaccharides that are released due to the extracellular hydrolytic activities produced by other infant gut microbiota members [18]. However, more recent studies have suggested that *B. breve* is able to utilise particular HMOs, such as fucosyllactose, LNT and sialyl-LNT, or derived structures such as LNB and sialic acid [17, 38-40].

In this study, we show that *B. breve* possesses the metabolic machinery for the degradation and utilisation of LNT and LNnT. Furthermore, we assess the presence and distribution of key gene loci involved in LNT and LNnT utilisation across members of the *Bifidobacterium* genus.

2.3 Materials and Methods

Bacterial strains, plasmids, and culture conditions. Bacterial strains and plasmids used in this study are listed in Supplemental Table S2.2. *B. breve* UCC2003 was routinely cultured in either de Man Rogosa and Sharpe medium (MRS medium; Difco, BD, Le Pont de Claix, France) supplemented with 0.05 % cysteine-HCl or reinforced clostridial medium (RCM; Oxoid Ltd., Basingstoke, England). Carbohydrate utilization by bifidobacterial strains was examined in modified de Man Rogosa and Sharpe (mMRS) medium prepared from first principles[41], and excluding a carbohydrate source. Prior to inoculation, the mMRS medium was supplemented with cysteine-HCl (0.05 %, wt/vol) and a particular carbohydrate source (1 %, wt/vol). It has previously been shown that mMRS does not support growth of *B. breve* UCC2003 in the absence of an added carbohydrate [42]. Carbohydrates used were lactose (Sigma Aldrich, Steinheim, Germany), LNB (Elicityl Oligotech, Crolles, France), lactosamine-hydrochloride (lactosamine-HCl) (Glycom, Lyngby, Denmark), LNT (Glycom, Lyngby, Denmark; Elicityl Oligotech, Crolles, France) and LNnT (Glycom, Lyngby, Denmark). A 1 % wt/vol concentration of carbohydrate was considered sufficient to analyse the growth capabilities of a strain on a particular carbon source. The addition of these carbohydrates did not significantly alter the pH of the medium, and therefore subsequent pH adjustment was not required.

B. breve cultures were incubated under anaerobic conditions in a modular atmosphere-controlled system (Davidson and Hardy, Belfast, Ireland) at 37°C. *Lactococcus lactis* strains were cultivated in M17 broth (Oxoid Ltd., Basingstoke, England) containing 0.5 % glucose[43] at 30°C. *Escherichia coli* strains were cultured in Luria-Bertani (LB) broth[44] at 37°C with agitation. Where appropriate, growth media contained tetracycline (Tet; 10μg ml^{-1}), chloramphenicol (Cm; 5 μg ml^{-1} for *L. lactis* and *E. coli*, 2.5 μg ml^{-1} for *B. breve*), erythromycin (Em; 100 μg ml^{-1}) or kanamycin (Kan; 50 μg ml^{-1}). Recombinant *E. coli* cells containing (derivatives of) pORI19 were selected on LB agar containing Em and Kan, and supplemented with X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (40 μg ml^{-1}) and 1 mM IPTG (isopropyl-β-D-galactopyranoside). In order to determine bacterial growth

profiles and final optical densities, 5 ml of freshly prepared mMRS medium, including a particular carbohydrate (see above), was inoculated with 50 µl (1 %) of a stationary phase culture of *B. breve* UCC2003. Uninoculated mMRS medium was used as a negative control. Cultures were incubated anaerobically at 37°C for 24h, and the optical density at 600 nm (OD600) was determined manually, or using a PowerWave microplate spectrophotometer (BioTek Instruments, Inc., USA) in conjunction with Gen5 microplate software for Windows, at the end of this period, as described previously[42, 45].

***Bifidobacterium breve* Growth Assays.** Growth profiles of sixteen distinct *Bifidobacterium breve* strains from the UCC collection (listed in Supplemental Table S2.2) on LNT or LNnT, as the sole carbohydrate source, using lactose as a positive control, were determined in mMRS, using the microplate spectrophotometer, as described above. LNB and lactosamine-HCl were not included in these assays, as sufficient (and affordable) quantities of these carbohydrate substrates could not be obtained.

Growth profiles of insertion mutant, Tn5 transposon mutant, and complementation strains of *B. breve* UCC2003 generated in this and other studies, were determined, manually, as described above, adopting LNT, LNnT or LNB as carbohydrate source and in each case using lactose as a positive control.

Nucleotide sequence analysis. Sequence data were obtained from the Artemis-mediated[46] genome annotations of *B. breve* UCC2003[47]. Database searches were performed using non-redundant sequences accessible at the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) using the basic local alignment search tool (BLAST)[48, 49]. Sequences were verified and analysed using the SeqMan and SeqBuilder programs of the DNAStar software package (version 10.1.2; DNAStar, Madison, WI, USA). Gene product (protein) localisation and signal

peptide predictions were made using the TMHMM, v. 2.0 and SignalP, v. 4.1[50] servers, respectively, available at http://www.cbs.dtu.dk/.

DNA Manipulations. Chromosomal DNA was isolated from *B. breve* UCC2003 as previously described[51]. Plasmid DNA was isolated from *E. coli*, *L. lactis* and *B. breve* using the Roche High Pure Plasmid Isolation kit (Roche Diagnostics, Basel, Switzerland). An initial lysis step was performed using 30 mg ml^{-1} of lysozyme for 30 minutes at 37°C prior to plasmid isolation from *L. lactis* or *B. breve*. DNA manipulations were essentially performed as described previously[44]. All restriction enzymes and T4 DNA ligase were used according to the supplier's instructions (Roche Diagnostics, Basel, Switzerland). Synthetic single stranded oligonucleotide primers used in this study (Supplemental Table S2.1) were synthesized by Eurofins (Ebersberg, Germany). Standard PCRs were performed using Taq PCRmaster mix (Qiagen) or Extensor Hi-Fidelity PCR Master Mix (Thermo Scientific, Waltham, United States) in a Life Technologies Proflex PCR System (Thermo Scientific, Waltham, United States). PCR products were visualized by ethidium bromide (EtBr) staining following agarose gel electrophoresis (1 % agarose). *B. breve* colony PCR reactions were performed as described previously [52]. PCR fragments were purified using the Roche high Pure PCR purification kit (Roche Diagnostics, Basel, Switzerland). Plasmid DNA was isolated using the Roche High Pure Plasmid Isolation kit (Roche Diagnostics, Basel, Switzerland). Plasmid DNA was introduced into *E. coli* by electroporation as described previously[44]. *B. breve* UCC2003[53] and *L. lactis*[54] were transformed by electroporation according to published protocols. The correct orientation and integrity of all plasmid constructs (see also below) were verified by DNA sequencing, performed at Eurofins (Ebersberg, Germany).

Analysis of global gene expression using *B. breve* DNA microarrays. Global gene expression was determined during log-phase growth of *B. breve* UCC2003 in mMRS supplemented with either LNT, LNnT, LNB,

lactosamine-HCl or lactose. The obtained transcriptome was compared to that determined for log-phase *B. breve* UCC2003 cells when grown in mMRS supplemented with ribose. DNA microarrays containing oligonucleotide primers representing each of the 1864 identified open reading frames on the genome of *B. breve* UCC2003 were designed and obtained from Agilent Technologies (Palo Alto, Ca., USA). Methods for cell disruption, RNA isolation, RNA quality control, complementary DNA synthesis and labelling were performed as described previously [55]. Labelled cDNA was hybridized using the Agilent Gene Expression hybridization kit (part number 5188-5242) as described in the Agilent Two-Colour Microarray-Based Gene Expression Analysis v4.0 manual (publication number G4140-90050). Following hybridization, microarrays were washed in accordance with Agilent's standard procedures and scanned using an Agilent DNA microarray scanner (model G2565A). Generated scans were converted to data files with Agilent's Feature Extraction software (Version 9.5). DNA-microarray data were processed as previously described [56-58]. Differential expression tests were performed with the Cyber-T implementation of a variant of the t-test[59].

Construction of *B. breve* UCC2003 insertion mutants. An internal fragment of Bbr_0530 (designated here as *lntS*) (465 bp, representing codon numbers 61 through to 216 of the 420 codons of this gene), Bbr_1554 (designated here as *nahS*) (488 bp, representing codon numbers 92 through to 255 of the 442 codons of this gene), and Bbr_1556 (designated here as *nahA*) (443 bp, representing codon numbers 70 through to 218 of the 660 codons of this gene) were amplified by PCR using *B. breve* UCC2003 chromosomal DNA as a template and primer pairs IM530F and IM530R, IM1554F and IM1554R, or IM1556F and IM1556R, respectively (Supplemental Table S2.1). The insertion mutants were constructed using a previously described approach [52]. Site-specific recombination of potential tet-resistant mutant isolates was confirmed by colony PCR using primer combinations tetWFw and tetWRv to verify *tetW* gene integration, and primers 530confirm1 or 530confirm2, 1554Confirm1 or 1554Confirm2, and 1556confirm1 or 1556confirm2 (positioned upstream of the selected internal fragments of

Bbr_0530, Bbr_1554 and Bbr_1556, respectively) in combination with primer tetWFw to confirm integration at the correct chromosomal location (Supplemental Table S2.1).

Complementation of *B. breve* insertion mutants. DNA fragments encompassing Bbr_0529 (designated here as *lntA*), Bbr_1554 (*nahS*) and Bbr_1556 (*nahA*) were generated by PCR amplification from *B. breve* UCC2003 chromosomal DNA using Q5 High-Fidelity Polymerase (New England BioLabs, Herefordshire, United Kingdom) and primer pairs: 529pNZ44F and 529pNZ44R, 1554PCB1.2F and 1554PBC1.2R, and 1556PBC1.2F and 1556PBC1.2R, respectively (Supplemental Table S2.1).

The resulting *lntA*-encompassing fragment was digested with PstI and XbaI, and ligated to the similarly digested pNZ44 [60]. The ligation mixture was introduced into *L. lactis* NZ9000 by electrotransformation and transformants were then selected based on chloramphenicol resistance. The plasmid content of a number of Cm-resistant transformants was screened by restriction analysis. The integrity of the cloned insert of one of the recombinant plasmids, designated pNZ44-lntA, was confirmed by sequencing. The *lntA*-coding sequence, together with the constitutive p44 lactococcal promoter, specified by pNZ44, was amplified by PCR from pNZ44-lntA using Q5 High-Fidelity DNA polymerase and primer combination P44 Forward and 529pNZ44R (Supplemental Table S2.1). The resulting DNA fragment was digested with EcoRV and XbaI, and ligated to the similarly digested pBC1.2[61], generating pBC1.2-lntA.

PCR-generated DNA fragments encompassing *nahS* and *nahA*, including the p44 promoter upstream of each, were digested with BamHI and XbaI, and ligated to the similarly digested pBC1.2 to generate pBC1.2-nahS or pBC1.2-nahA, respectively. The ligation mixtures were introduced into *E. coli* XL1-blue by electrotransformation and transformants selected based on tetracycline and chloramphenicol resistance. Transformants were checked for plasmid content using colony PCR, restriction analysis of plasmid DNA, and verified by sequencing. Plasmids pBC1.2-lntA, pBC1.2-nahS or pBC1.2-

nahA were introduced into the insertion mutant *B. breve* UCC2003-lntA, *B. breve* UCC2003-nahS and *B. breve* UCC2003-nahA [62], respectively, by electrotransformation and transformants were selected based on tetracycline and chloramphenicol resistance.

Construction of overexpression vectors, protein overproduction and purification. For the construction of the plasmid pNZ-nahA, a DNA fragment encompassing the predicted N-acetylhexosaminidase-encoding gene *nahA* was generated by PCR amplification from chromosomal DNA of *B. breve* UCC2003 using Q5 High-Fidelity DNA polymerase and the primer combination 1556F and 1556R (Supplemental Table S2.1). An in-frame N-terminal His10-encoding sequence was incorporated into the forward primer 1556F to facilitate downstream protein purification. The generated amplicons were digested with PvuII and XbaI, and ligated into the ScaI and XbaI-digested, nisin-inducible translational fusion plasmid pNZ8150[63]. The ligation mixtures were introduced into *L. lactis* NZ9000 by electrotransformation and transformants were then selected based on chloramphenicol resistance. The plasmid content of a number of Cm-resistant transformants was screened by restriction analysis and the integrity of positively identified clones was verified by sequencing.

Nisin-inducible gene expression and protein overproduction was performed as described previously [64-66]. In brief, 400 ml of M17 broth supplemented with 0.5 % (wt/vol) glucose was inoculated with a 2 % inoculum of a particular *L. lactis* strain, followed by incubation at 30°C until an OD600 of 0.5 was reached, at which point protein expression was induced by addition of cell-free supernatant of a nisin-producing strain[67], followed by continued incubation for a further 2 hours. Cells were harvested by centrifugation and protein purification achieved as described previously[64]. Protein concentrations were determined using the Bradford method[68].

Assay of individual and combined β-Galactosidase activities. The individual or sequential hydrolytic activities specified by *lntA* (corresponding

to Bbr_0529), *nahA* (corresponding to Bbr_1556), *lacZ2* (corresponding to Bbr_0010) and *lacZ6* (corrsponding to Bbr_1552) were determined essentially as described previously [66], using LNT, LNnT or lactose as a substrate. Briefly, a 50-µl volume of each purified protein (protein concentration of 0.5 mg/ml) was added to 20 mM morpholinepropanesulfonic acid (MOPS) (pH 7.0) buffer and 1 mg ml^{-1} (wt/vol) of one of the above-mentioned sugars in a final volume of 1 ml, followed by incubation for 24 hours at 37°C. When sequential activities were assessed, a sample was heated to 85°C for 15 minutes following 12 hour incubation with the first enzyme and a given substrate, before the addition of the addition of a second enzyme, which was then followed by a further 12-hour incubation at 37°C. All samples were subject to a final enzyme denaturation step at 85°C for 15 minutes, before storage at -20°C.

HPAEC-PAD analysis. For HPAEC-PAD analysis, a Dionex (Sunnyvale, CA) ICS-3000 system was used. Carbohydrate fractions from the above-mentioned hydrolysis assays (25 µl aliquots) were separated on a CarboPac PA1 analytical-exchange column (dimensions, 250 mm by 4 mm) with a CarboPac PA1 guard column (dimensions, 50 mm by 4 mm) and a pulsed electrochemical detector (ED40) in PAD mode (Dionex). Elution was performed at a constant flow-rate of 1.0 ml/min at 30°C using the following eluents for the analysis: eluent A, 200 mM NaOH; eluent B, 100 mM NaOH plus 550 mM Na acetate; eluent C, Milli-Q water. The following linear gradient of sodium acetate was used with 100 mM NaOH: from 0 to 50 min, 0 mM; from 50 to 51 min, 16 mM; from 51 to 56 min, 100 mM; from 56 to 61 min, 0 mM. Chromatographic profiles of standard carbohydrates were used for comparison of the results of their breakdown by LntA, LacZ2, LacZ6 and NahA proteins. Chromeleon software (version 6.70; Dionex Corporation) was used for the integration and evaluation of the chromatograms obtained. A 1 mg/ml stock solution of each of the carbohydrates, as well as their putative breakdown products (where available) used as reference standards was prepared by dissolving the particular sugar in Milli-Q water.

Bioinformatic Analysis. Based on the analysis of the microarray results and functional characterisation of gene loci from *B. breve* UCC2003, as well as previously published data on HMO utilisation by *B. longum* subsp. *infantis* and *B. bifidum* and *B. longum* subsp. *longum* [18, 28-37], four genes were identified as crucial for the utilisation of Type I central moieties LNT and LNB, and Type II HMO moiety LNnT. On-line available genomic data sets of bifidobacteria were first retrieved from the NCBI website (http://www.ncbi.nlm.nih.gov) and aligned using an all-vs-all BLASTP approach[48], using 70 % of iterative similarity across all available *Bifidobacterium* species over 50 % of protein length and a 0.0001 e-value as a significance cut-off. The resulting alignment was subsequently clustered in MCL families of orthologous genes using the mclblastline algorithm[69]. The resulting output was used to first build a presence/absence binary matrix, and then the genes of interest were selected and represented in a heatmap employing a code colour grading that represents the degree of sequence similarity, with species ordered by origin of isolation. Bbr_1556 (*nahA*) and Bbr_1587 (*lnbP*) were selected from *B. breve* UCC2003, BLLJ_1505 (*lnbX*) from *B. longum* subsp. *longum* JCM1217 and BBPR_1438 (*lnbB*) was selected from *B. bifidum* PRL2010.

Microarray data accession number. The microarray data obtained in this study have been deposited in NCBI's Gene Expression Omnibus database and are accessible through GEO Series accession number GSE84710.

2.4 Results

Growth of *B. breve* strains on LNT and LNnT. In order to determine if *B. breve* strains are capable of LNT and/or LNnT metabolism, growth in modified MRS medium (mMRS) supplemented with either 1 % (wt/vol) LNT, LNnT or lactose (as a positive control) was assessed for sixteen *B. breve* strains by measuring the OD_{600nm} following 24 hours of anaerobic growth at 37°C. All tested *B. breve* strains were generally observed to grow well (final OD_{600nm}> 0.8) on both LNT and LNnT, with some variability between strains on one or both HMO substrates (Supplemental Fig. S2.1).

Transcriptome analysis of *B. breve* UCC2003 grown on LNT and LNB. In order to identify genes that are involved in the metabolism of the Type I HMO central moiety LNT and its constituent component LNB, global gene expression was determined by microarray analysis during growth of *B. breve* UCC2003 in mMRS supplemented with LNT or LNB, and compared to the transcriptome of the strain when grown in mMRS supplemented with ribose. Ribose was selected as a suitable transcriptomic reference, as the metabolic pathway and gene expression profile for growth of UCC2003 on ribose is known and has been employed previously as a reference[39, 70]. Genes that were shown to be significantly upregulated in transcription above the designated cut-off (fold-change >2.5, P<0.001) are shown in Table 2.1. The genes upregulated in expression included those corresponding to the loci Bbr_0526-530, Bbr_1551-1553, Bbr_1554-1560 and Bbr_1585-1590. The possible involvement of these genes in LNT/LNB metabolism will be further discussed below.

Genetic organisation of the genes involved in metabolism of LNT. Based on the microarray results and functional prediction of these LNT (and LNB)-upregulated genes, we implicate the gene clusters Bbr_0526-0530, Bbr_1554-1560, Bbr_1585-1590 and possibly Bbr_1551-1553 (outlined in Fig. 2.2) in LNT and LNB metabolism in *B. breve* UCC2003.

Bbr_0527 and Bbr_0528 (designated here as *lntP1* and *lntP2*, respectively) are both predicted to encode permease components of an ABC transporter system. Also located in this cluster (Fig. 2.2) is Bbr_0529 (designated *lntA*), which encodes a predicted β-galactosidase of the GH42 glycosyl hydrolase family. Bbr_0530 (designated here as *lntS*) encodes a putative solute-binding protein of an ABC transporter system. Located immediately upstream of this cluster, Bbr_0526 (designated *lntR*) encodes a putative LacI-type transcriptional regulator. We have previously implicated the Bbr_0526-530 gene cluster, in the metabolism of galacto-oligosaccharides [62], where the genes were designated *gosR* (*lntR*), *gosD* (*lntP1*), *gosE* (*lntP2*), *gosG* (*lntA*) and *gosC* (*lntS*). Here, we chose to re-designate this cluster as the *lnt* cluster, as its primary function appears to be in LNT and LNnT metabolism (see below).

Bbr_1551 (designated here as *lacS*) encodes a galactoside symporter, and is predicted to function in the transport of lactose and galacto-oligosaccharides into the cell. The *lacS* gene is located in a cluster that also contains genes Bbr_1552 (designated *lacZ6*), a β-galactosidase (GH2) previously shown to be involved in galacto-oligosaccharide metabolism [62], and Bbr_1553 (designated *lacI*), a lacI-type regulator.

Bbr_1555 (designated *nahR*) is predicted to encode a Nag-type transcriptional regulator. Bbr_1556 (designated *nahA*) encodes a putative β-N-acetylhexosaminidase (GH20). Upstream of *nahR* is a gene encoding a putative solute binding protein (Bbr_1554 and designated here as *nahS*), while located downstream of *nahA* are Bbr_1558 (*nahP1*), Bbr_1559 (*nahP2*) and Bbr_1560 (*nahT*), which are predicted to specify two permeases and an ATP-binding protein; respectively (Fig. 2.2).

Bbr_1587 (designated here as *lnbP*) encodes a clear homolog (89.84 % similarity to BBPR_1055 of *B.bifidum* PRL2010, and 97.62 % to Blon_2174 of *B. longum* subsp. *infantis* ATCC 15697) of the previously characterised LNBP, which belongs to the 1,3-β-Galactosyl-N-acetylhexosamine phosphorylase family (GH112)[31, 71, 72]. The presumed function of this protein in UCC2003 is the cleavage and concomitant phosphorylation of

LNB, and its passage into the GNB/LNB pathway [71-75]. The *lnbP* gene is located in the cluster Bbr_1585-1590, which also includes genes encoding a UDP-glucose 4-epimerase (Bbr_1585; *galE*), a phosphotransferase family protein (Bbr_1586; *nahK*), two permease proteins (Bbr_1588 and Bbr_1589; *galP1* and *galP2*, respectively), and a solute-binding protein (Bbr_1590; *galS*) (Fig. 2.2), which are all predicted to function in the metabolism of GNB/LNB.

Heterologous expression, purification and biochemical characterisation of LntA and NahA, and enzymatic activity on LNT. In order to investigate the predicted enzymatic activities encoded by *lntA* (Bbr_0529) and *nahA* (Bbr_1556) on core Type I HMO structure LNT, the corresponding LntA and NahA proteins were purified as His-tagged versions ($LntA_{His}$ and $NahA_{His}$; see Materials and Methods). Biochemical and substrate specificity characterisations were performed by incubating $LntA_{His}$ and $NahA_{His}$ on their own or in combination with LNT, and analysing the reaction products by HPAEC-PAD. Purified $LntA_{His}$ was shown to remove the galactose moiety at the non-reducing end of the substrate LNT (Fig 2.3A), demonstrating hydrolytic activity towards Galβ-1,3GlcNAc in Type I HMO structures, and indicating a key role in the hydrolysis and utilisation of LNT.

When $NahA_{His}$ was incubated alone with LNT, no degradation of the tetrasaccharide structure was observed (Fig 2.3A). However, when $LntA_{His}$, and $NahA_{His}$ were together incubated with LNT, complete breakdown of LNT to the monosaccharide constituents was observed. When LNT was incubated first with $LntA_{His}$, followed by an enzymatic heat denaturation step, and then incubated with $NahA_{His}$, different reaction product profiles were observed. Samples taken following the initial denaturation prior to the addition of $NahA_{His}$ showed the presence of lacto-N-triose and galactose. Samples were then taken following subsequent incubation with $NahA_{His}$ indicated the presence of galactose, GlcNAc and lactose.

These results show that LntA hydrolyses LNT, releasing galactose and lacto-N-triose. Lacto-N-triose is then hydrolysed by NahA, liberating Lactose and

GlcNAc. The lactose is then further broken down by LntA (and probably other β-Galactosidases *in vivo*), releasing galactose and glucose.

Phenotypic analysis of *B. breve* strains harbouring mutations of genes implicated in LNT metabolism. In order to investigate if disruption of individual genes of the *lnt* cluster affect the ability of UCC2003 to utilise LNT, a number of insertion mutants, which either had been generated previously, or which were constructed here, were assessed. An insertion mutant was constructed in *lntS*, resulting in strain *B. breve* UCC2003-lntS. Insertion mutants in *lntP1* and *lntA*, generating strains *B. breve* UCC2003-lntP1 and *B. breve* UCC2003-lntA, respectively, had been generated in a previous study (then designated *B. breve* UCC2003-gosD and *B. breve* UCC2003-gosG, respectively)[62]. These strains were analysed for their ability to grow in mMRS supplemented with LNT or LNB, with lactose controls, as compared to *B. breve* UCC2003. A complete lack of growth was observed for *B. breve* UCC2003-lntA in media containing LNT, in contrast to normal growth by the wild type in the same media (Fig 2.4A). Growth of this mutant strain was not impaired in media containing LNB (not shown). As expected, reintroduction of the *lntA* gene on plasmid pBC1.2 under the control of the constitutive p44 promoter [76] (see Materials and Methods) in the UCC2003-lntA mutant restored the mutant's inability to grow on LNT (Supplemental Fig. S2.3). Thus, transcriptome data, substrate hydrolysis profiles and mutant growth results demonstrate that this β-galactosidase is specifically required for the hydrolysis of the Type I HMO central moiety LNT at its Galβ1-3GlcNAc residue, liberating galactose and lacto-N-triose for further metabolic processing. Insertion mutants *B. breve* UCC2003-*lntP1* and *B. breve* UCC2003-*lntS* were shown to reach the same final optical density as wild type strain UCC2003 during growth in mMRS supplemented with LNT (Fig 2.4A), indicating that either these predicted transport components do not play a role in LNT metabolism, or that there are compensatory transport systems for this substrate.

In order to investigate if disruption of *lnbP* affects the ability of UCC2003 to utilise LNT and/or LNB, a Tn5 transposon insertion mutant of *lnbP* (designated *B. breve* UCC2003-lnbP) was adopted from a previous study[74] and compared to wild type *B. breve* UCC2003 for its ability to grow in mMRS broth supplemented with LNT or LNB, or lactose as control. As expected, and in contrast to the wild type control, *B. breve* UCC2003-*lnbP* displayed a near total inability to grow on LNB (Fig 2.4B). This mutant reached final OD_{600nm} levels on LNT and lactose that are comparable to the wild type strain (Fig 2.4B), confirming the crucial role of *lnbP* in LNB metabolism, while it also shows that *lntA* plays no direct *in vivo* role in LNB metabolism.

In order to investigate if disruption of *lacS*, *nahS* or *nahA* affects the ability of UCC2003 to utilise LNT, insertion mutants in these genes were assessed. The insertional mutant *B. breve* UCC2003-lacS[62] did not show any significant difference in the final OD reached following growth on LNT as compared with the wild type (Fig 2.4A). The insertional mutant in *nahS* (generated in this study), designated *B. breve* UCC2003-nahS, also did not exhibit a difference in final OD following growth in media containing LNT (as compared to the wild type, Fig 2.4A). This suggests that either *nahS* is not involved in LNT transport, or that while *nahS* and the other transport system components of the *nah* locus may be involved in the transport of LNT into the cell, their function is compensated by the activity of one or more other transport systems. The insertion mutant in *nahA* (generated in this study), designated *B. breve* UCC2003-nahA, was shown to exhibit a complete lack of growth in LNT-containing media (Fig 2.4A). Reintroduction of the *nahA* gene *in trans* on plasmid pBC1.2, under the transcriptional control of its own promoter [76] (see Materials and Methods), in the UCC2003-nahA mutant restored the ability to grow on LNT (Supplemental Fig. S2.3). These findings demonstrate that *nahA* is crucial for LNT metabolism, being responsible for the hydrolysis of lacto-N-triose, thereby liberating lactose and GlcNAc.

Transcriptome analysis of *B. breve* UCC2003 grown on LNnT, lactosamine and lactose. In order to investigate which genes are involved in

the metabolism of Type II central moiety LNnT and its constituent component LacNAc, global gene expression was determined by microarray analysis during growth of *B. breve* UCC2003 in mMRS supplemented with each respective sugar, as well as lactose (which also possesses a Galβ-1,4 residue) as compared with gene expression during growth in mMRS supplemented with ribose (NB. We used a hydrochloride salt of lactosamine instead of LacNAc, as the latter was not commercially available in an affordable quantity). Genes that were shown to be significantly upregulated in transcription above the designated cut-off (fold-change >2.5, $P<0.001$) are shown in Table 2.1. The genes upregulated in expression included those located in the loci Bbr_0526-530, Bbr_1551-1553, Bbr_1554-1560 and Bbr_1585-1590. Possible involvement of these genes in LNnT/LacNAc metabolism are assessed below.

Genetic organisation of the genes involved in metabolism of LNnT. Based on the results of the microarray analyses performed and functional annotation of LNnT/LacNAc-upregulated genes, we propose that the products of the gene clusters Bbr_0526-0530, Bbr_1551-1553 and Bbr_1554-1560 (schematically outlined in Fig. 2.2) are involved in the metabolism of LNnT and LacNAc (present as central moieties in Type II HMO) in *B. breve* UCC2003.

Heterologous expression, purification and biochemical characterisation of LntA, LacZ2, LacZ6, NahA, and enzymatic activity on Type II HMO structure LNnT. In order to investigate the predicted individual and combined enzymatic activities of the protein products *lacZ2* (Bbr_0010), *lntA* (Bbr_0529), *lacZ6* (Bbr_1552) and *nahA* (Bbr_1556) on core Type II HMO structure LNnT, the corresponding His-tagged protein products were overproduced and purified. Biochemical and substrate specificity characterisations were performed by incubating individual enzymes or combinations thereof with a particular substrate, and analysing the reaction products by HPAEC against a number of substrate standards and reaction

controls. Purified $LntA_{His}$ was capable of removing the galactose moiety at the non-reducing end of LNnT (Fig 2.3B), as well as lactose (data not shown), demonstrating a triple specificity for Galβ-1,3GlcNAc, Galβ-1,4GlcNAc and Galβ-1,4Glc glycosidic linkages, and thus both Type I and Type II HMO central moieties and lactose. Purified $LacZ2_{His}$ and $LacZ6_{His}$ were, under the conditions applied, also capable of hydrolysing LNnT and lactose (data not shown).

When both $LntA_{His}$ and $NahA_{His}$ were incubated with LNnT, complete hydrolysis of LNnT to its constituent monosaccharides was observed. When LNnT was incubated first with $LntA_{His}$, followed by an enzymatic heat denaturation step, and then incubated with $NahA_{His}$, different reaction product profiles were observed. Samples taken following the initial denaturation prior to the addition of $NahA_{His}$ showed the presence of lacto-N-triose and galactose. Samples taken following the addition of $NahA_{His}$ and subsequent incubation showed the presence of galactose, GlcNAc and lactose (Fig 2.3B).

Similar results from separate and combined reactions were obtained using $LacZ2_{His}$ or $LacZ6_{His}$, together with $NahA_{His}$ on the substrate LNnT (Supplemental Fig. S2.2).

These results agree with the model for Type II HMO metabolism proposed here; where LntA, LacZ2 and/or LacZ6 (and perhaps other β-galactosidases) hydrolyse LNnT, releasing galactose and lacto-N-triose, unlike LNT, which neither purified LacZ2 nor LacZ6 displayed the ability to hydrolyse (data not shown). Lacto-N-triose is then hydrolysed by NahA, liberating lactose and GlcNAc. Lactose is then further broken down by LntA (and other β-Galactosidases, including LacZ2 and LacZ6, *in vivo*), releasing galactose and glucose. Therefore, these two β-Galactosidases may carry out hydrolysis of Type II HMO central moieties (in conjunction with LntA).

Phenotypic analysis of *B. breve* strains harbouring mutations of genes implicated in LNnT metabolism. In order to investigate if disruption of individual genes of the *lnt* cluster affects the ability of UCC2003 to utilise LNnT, a number of insertion mutants were assessed. *B. breve* UCC2003-

lntP1, *B. breve* UCC2003-lntS and *B. breve* UCC2003-lntA were analysed for their ability to grow in mMRS supplemented with LNnT with lactose controls, as compared to wild type *B. breve* UCC2003. *B. breve* UCC2003-lntA reached the same final optical density following growth in media containing LNnT compared to wild type UCC2003 (Fig 2.4A). Transcriptome data and carbohydrate hydrolysis assays (see above) demonstrated that *lntA* is involved in LNnT metabolism, but its function can also be carried out by other glycosyl hydrolases, as mentioned. The insertion mutants *B. breve* UCC2003-lntP1 and *B. breve* UCC2003-lntS did not show any significant impairment in growth on either LNnT or lactose, as compared with the wild type (Fig 2.4A), indicating that either these predicted transport components do not play a role in LN(n)T metabolism or that there are additional transport systems that allow internalisation LNnT. It is most likely that *lntP1*, *lntP2* and *lntS* do indeed also function in the transport of extracellular LNnT into the cytoplasm, but that their function can be supplemented or indeed supplanted by other cellular transport systems. The *lnbP* insertion mutant, *B. breve* UCC2003-*lnbP*, was shown to reach final OD_{600nm} levels on either lactose or LNnT that were comparable to those of the wild type strain (Fig 2.4B), suggesting *lnbP* plays no direct role in LNnT utilisation.

In order to investigate if disruption of each of the five individual genes *lacZ2*, *lacZ6*, *lacS*, *nahS* and *nahA* affects the ability of UCC2003 to utilise LNnT, mutants of these genes were assessed. The *nahS* and *nahA* insertion mutants mentioned above were used for this purpose, in addition to insertion mutants in the genes Bbr_1551 (*lacS*), Bbr_1552 (*lacZ6*), constructed in a previous study[62], as well as a Tn5 transposon mutant in Bbr_0010 (*lacZ2*), from a separate study[77]. These mutant strains were analysed for their ability to grow in mMRS supplemented with LNnT with a lactose (and ribose, in the case of *B. breve* UCC2003-lacZ2 and *B. breve* UCC2003-lacS) control, as compared to *B. breve* UCC2003. The insertion mutants *B. breve* UCC2003-lacZ2 and *B. breve* UCC2003-lacZ6 did not show any significant impairment in growth on either LNnT, as compared to the wild type, as neither did the insertional mutant *B. breve* UCC2003-lacS (Fig 2.4A). While these growth analyses indicate that *lacZ2* and *lacZ6* are not essential for LNnT metabolism,

the hydrolysis assay results and LNnT-dependent transcriptional induction of *lacZ6* (described above) suggest that these two glycosyl hydrolases play a role in LNnT hydrolysis, presumably in concert with *lntA*. The insertion mutant *B. breve* UCC2003-nahS was shown to exhibit a near total lack of growth in media supplemented with LNnT, as compared to the wild type grown in the same media (Fig 2.4A). As expected, reintroduction of the *nahS* gene on plasmid pBC1.2, under the regulation of its own promoter [76] (see Materials and Methods), in the UCC2003-nahS mutant reverted the mutant's (near complete) inability to grow on LNnT (Supplemental Fig. S2.3). This indicates that *nahS* encodes the solute binding protein predominantly required for the uptake of LNnT, and that the *nah* locus-encoded transport system is of critical importance for the transport of the Type II central tetrasaccharide into the cell. In contrast to the wild type, *B. breve* UCC2003-nahA failed to grow in media containing LNnT (Fig 2.4A). As expected, reintroduction of the *nahA* gene on plasmid pBC1.2, under the regulation of its own promoter [76] (see Materials and Methods), in the UCC2003-nahA mutant reverted the mutant's inability to grow on LNnT (Supplemental Fig. S2.3). This result demonstrates the essential role of the *nahA* product in LNnT metabolism, by hydrolysing lacto-N-triose at its GalNacβ1-3Gal linkage, liberating lactose and GlcNAc.

Growth of the insertion mutants was not impaired on lactose, where all strains reached final OD_{600nm} levels comparable to that reached by UCC2003, except for *B. breve* UCC2003-lacZ2 and *B. breve* UCC2003-lacS, as the interrupted genes in these mutants are known to be crucial for lactose metabolism[62, 77] (Fig 2.4A). *B. breve* UCC2003-lacZ2 and *B. breve* UCC2003-lacS did reach final OD_{600nm} values similar to that of UCC2003 when grown on ribose (data not shown).

Distribution of HMO central moiety utilisation-associated genes across the *Bifidobacterium* genus. Two signature genes, encoding glycosyl hydrolases essential to the catabolic pathways of HMO central moieties LNT, LNnT and LNB, in *B. breve,* were identified based on the above results for *B.*

breve UCC2003 (Supplemental Fig 2.3). The *nahA* gene was identified as being crucial for the degradation of both LNT and LNnT (through the hydrolysis of lacto-N-triose), and *lnbP* was identified as essential for the utilisation of LNB. Additionally, one signature gene, *lnbB*, was identified in *B. bifidum* as encoding the key glycosyl hydrolase required for the metabolism of LNT, based on previous literature[32] (Supplemental Fig 2.3). Another distinct glycosyl hydrolase required for the metabolism of LNT in this same way, *lnbX*, was identified in a strain of *B. longum* subsp. *longum* by Sakurama et al.[37], and thus was also selected (Supplemental Fig 2.3). No genes were selected from *B. longum* subsp. *infantis*, as this species and *B. breve* appear to share the same functional homologs and thus appear to utilize the same pathways for the metabolism of LN(n)T and LNB. The deduced amino acid sequences of these four genes were employed as the reference sequences in a multiple alignment of all available *Bifidobacterium* genomes retrieved from the NCBI database, as described, and represented in a heatmap, based on a cut-off of 70 % iterative similarity over 50 % protein length, and an e-value of <0.0001 (Fig 2.5). The representation obtained, ordered by origin of isolation, reveals the distribution of these key genes, and thus the metabolic pathways, required for LN(n)T/LNB utilisation across the *Bifidobacterium* genus. The *B. bifidum* gene *lnbB* and *B. longum* subsp. *longum* gene *lnbX* (whose products are responsible for the hydrolysis of LNT into LNB and lactose) appear to be individually unique to *B. bifidum* and *B. longum* subsp. *infantis*, respectively, with no clear homologs in any other species of *Bifidobacterium* including each other. The two *B. breve* signature genes used in the search yielded multiple significant hits for homologs, but to differing degrees. The analysis identified a relatively small number, i.e. four homologous *nahA* genes across the genus: one in the infant-associated species *B. longum* subsp. *infantis*, two in marmoset-associated and one tamarin-associated species- *B. callittrichos*, *B. ruteri*, and *B. saguini*, respectively. On the other hand, *lnbP* yielded 16 significant matches to homologous genes in other *Bifidobacterium* species, isolated from both human and non-human-related sources. The significant differences in the conservation of these genes across the genus are indicative of the importance of specific glycan moiety-utilising pathways in bifidobacteria.

2.5 Discussion

The role of HMOs as a selective substrate, for specific bacterial species in the neonatal gut, is now widely proposed as one of the key factors in the development of a healthy microbiota in early life. The high proportion of bifidobacteria, specifically the species *B. breve*, *B. longum* subsp. *infantis* and *B. bifidum*, in the microbiota of breastfed infants indicates their ability to utilise these carbohydrates as growth factors. Our findings allow us to propose a model for the utilisation of HMOs LN(n)T by *B. breve* (Fig 2.6). In this model, LN(n)T is internalised by the cell and subsequently degraded by intracellular pathways into monosaccharides for energy production. While *B. breve* is able to metabolize smaller HMO components, such as fucose, fucosyllactose and sialic acid, which are released through extracellular hydrolysis of larger molecules [16, 17, 22, 39, 40], our findings clearly show that *B. breve* can also utilise larger HMO structures. This expands our view of this gut commensal from being merely a scavenger, to an active and direct HMO utilizer.

Our multi-pronged approaches reveal the activities of individual components and thus the overall pathways that facilitate LNT and LNnT utilisation by *B. breve* UCC2003. LntA exhibits a triple specificity for the Galβ-1,3GlcNAc and Galβ-1,4GlcNAc linkages of LNT and LNnT, as well as the Galβ-1,4Glc moiety of lactose, as previously suggested [62]. However, the ability of UCC2003-lntA to grow on LNnT, but not on LNT, indicates that while LNT can only be intracellularly hydrolysed by LntA, the hydrolysis of LNnT is not exclusively attributable to this β-galactosidase, and can be degraded by other cellular glycosyl hydrolases such as LacZ6 and LacZ2, releasing lacto-N-triose for hydrolysis by NahA, and galactose. It should be noted that both *lacZ2* and *lacZ6* have previously been shown to be involved in the metabolism of lactose and galacto-oligosaccharides [62]. Interestingly, these results mirror those previously shown by Yoshida et al.[29], who, in *B. longum* subsp. *infantis* ATCC15697, demonstrated the preferential activities of one GH42 family glycosyl hydrolase (Bga42A) in hydrolysing Type I HMOs and one GH2 family glycosyl hydrolase (Bga2A) in hydrolysing Type II HMOs and lactose.

The transcriptomic results clearly implicate *nahS, lntS, lntP1* and *lntP2* in the utilisation on both LNT and LNnT, being involved in the internalisation of these sugars into the cell. The inability of the UCC2003-nahS mutant to grow on LNnT suggests the role of the *nah* locus-encoded transporter as the sole system responsible for LNnT internalisation by UCC2003. In contrast, since UCC2003-nahS displays growth on LNT comparable to that of the wild type strain, the *nah* transport system may not be involved in LNT transport, or may be, but with its function aided by one or more additional transport systems. The ability of the UCC2003-lntP1 and UCC2003-lntS to grow in mMRS supplemented with LNT demonstrates that the *lnt* transport system is also not exclusively, or potentially, at all, responsible for LNT internalisation. We therefore suggest that either these two transport systems may have at least partially overlapping substrate specificities, or that another yet undetermined transport system is partially, or wholly responsible for the internalisation of LNT.

While LNT and LNnT enter the *B. breve* cell as distinct isomers, their degradation products are identical and are shuttled through the same metabolic routes for energy production, i.e. two galactose molecules and one glucose to the Leloir and F6P phosphoketolase pathways and one GlcNAc to the amino-sugar metabolising pathway (thus both directly and indirectly feeding into the Bifid Shunt) (Fig 2.6).

Although found within the molecular structure of LNT, free LNB is not released during the degradation of the Type I HMO central moiety by *B. breve* UCC2003. However, despite the relative low abundance of free LNB in human breast milk and thus the breastfed infant gut, *B. breve* UCC2003 possesses a distinct pathway for LNB utilisation. It has previously been suggested that LNB metabolism can be seen as a proxy for HMO utilisation by Bifidobacteria [17, 28], which explains the presence of this pathway in *B. breve*, which likely 'sweeps up' the free LNB released by the extracellular hydrolysis of larger HMO structures by other microbiota members, such as *B. bifidum*, and then utilise it via the GNB/LNB pathway.

The general model of LN(n)T and LNB utilisation in *B. breve* is mirrored by that found in *B. longum* subsp. *infantis*. The functional equivalent of *lacZ2* and *lacZ6* in *B. longum* subsp. *infantis* ATCC15697 is *bga2a* [29]. The ortholog of *lntA* is *bga42a*[29], and the counterpart of *nahA* is *nagZ* [30]. A copy of *lnbP* is also found in *B. longum* subsp. *infantis* ATCC15697 [31]. In contrast, HMO central moiety utilisation in *B. bifidum* diverges somewhat from the *B. breve*/*B. longum* subsp. *infantis* model. While *B. bifidum* possesses functionally equivalent orthologs of *lacZ2*/*lacZ6* (*bbgIII*)[78], and a *nahA* (*bbhI* and *bbhII*)[34], all of these appear to be extracellular proteins, as opposed to the (predicted) intracellular localisation of their *B. breve* counterparts. However, the biggest defining factor separating the *B. breve*/*B. longum* subsp. *infantis* model and the *B. bifidum* model appears to be the hydrolysis pathway of Type I central moiety LNT. In *B. bifidum* PRL2010 LnbB hydrolyses LNT extracellularly at its GlcNAcβ1-3Gal linkage, releasing LNB, which is internalised, and phosphorolysed by *lnbP1* and *lnbP2*; and lactose, which is hydrolysed by β-galactosidase activities.

The absence of *B. bifidum lnbB* (and *B. longum* subsp. *longum* JCM1217 *lnbX*) homologs in other *Bifidobacterium* species highlights its unique function in HMO metabolism. This agrees with previous knowledge, as already mentioned, of *B. bifidum* (and one strain of *B. longum* subsp. *longum*) utilizing a significantly different pathway for LNT utilisation, and of HMO metabolism as compared to *B. breve* and *B. longum* subsp. *infantis*. A similarity search for the two *B. breve* LN(n)T/LNB signature genes, *nahA* and *lnbP*, among bifidobacteria shows the presence of homologs in various species of *Bifidobacterium*, but the extent of their conservation differs considerably. The sixteen species of *Bifidobacterium* that were shown to possess an *lnbP* homolog, have been isolated from a range of environments, including faecal samples of (both infant and adult) humans, primates and other mammals. The conservation of this gene across various bifidobacterial species points toward the importance of this gene in the GNB/LNB pathway, which is common to many *Bifidobacterium* species, and has previously been shown to function in roles such as mucin metabolism [28, 74, 79, 80]. Interestingly, clear homologs of the *nahA* gene are only found in four other

bifidobacterial species, the human isolate *B. longum* subsp. *infantis*, and three other primate-associated species *B. callitrichos* and *B. reuteri*, originally isolated from marmoset faeces, and *B. sanguini*, originally isolated from tamarin faeces. As LNT and similar oligosaccharide structures can be found in the glycome of primate milk[81], the *nahA* homologs are expected to play a similar role in *Bifidobacterium* species associated with these hosts. This suggests a common adaptation of the LNT/LNnT-utilisation pathway among bifidobacteria associated with the primate gut, using their respective milk oligosaccharides as substrates, thereby explaining co-evolution with and colonisation of this host.

2.6 Acknowledgements

The authors would like to sincerely thank Glycom A/S (Lyngby, Denmark) for the provision of purified HMO samples used in this study under their donation program. This study was funded in part by the Irish Research Council, under the Postgraduate Research Project Award; Project ID GOIPG/2013/651. In addition, the authors are supported by Science Foundation Ireland (SFI) (Grant No. SFI/12/RC/2273) and Mary O'Connell Motherway is a recipient of a HRB postdoctoral fellowship (Grant No. PDTM/20011/9).

2.7 Tables and Figures

Table 2.1. *B. breve* UCC2003 genes upregulated in transcription during growth in mMRS medium supplemented with 1 % LNT, LNnT, LNB, lactosamine-HCl, or lactose as the sole carbohydrate.

Gene ID	Gene name	Function	Fold upregulation[b] during growth on:				
			LNT	LNnT	LNB	Lactosamine-HCl	Lactose
Bbr_0417	galC	Solute-binding protein of ABC transporter system for sugars	—	—	—	10.93	—
Bbr_0418	galD	Permease protein of ABC transporter system for sugars	3.23	5.18	—	6.98	—
Bbr_0419	galE	Permease protein of ABC transporter system for sugars	3.10	7.88	3.18	9.80	3.07
Bbr_0420	galG	GH42 lacZ4 Beta-galactosidase	—	—	—	3.85	—
Bbr_0421	galR	Transcriptional regulator, LacI family	—	—	—	—	—
Bbr_0422	galA	GH53 galA Endogalactanase	—	—	—	4.07	—
Bbr_0490	Bbr_0490	Transcriptional regulator, DeoR family	—	3.27	—	2.60	—
Bbr_0491	galT	Galactose-1-phosphate uridylyltransferase	—	—	5.08	—	—
Bbr_0492	galK	Galactokinase	—	—	3.57	—	—
Bbr_0526	**lntR**	**Transcriptional regulator, LacI family**	**—**	**3.95**	**—**	**10.16**	**—**
Bbr_0527	**lntP1**	**Permease protein of ABC transporter system for sugars**	**6.81**	**10.05**	**—**	**8.57**	**3.48**
Bbr_0528	**lntP2**	**Permease protein of ABC transporter system for sugars**	**3.23**	**10.64**	**3.03**	**11.15**	**3.35**
Bbr_0529	**lntA**	**GH42 Beta-galactosidase**	**6.61**	**12.58**	**—**	**9.53**	**3.64**
Bbr_0530	**lntS**	**Solute-binding protein of ABC transporter system for sugars**	**4.15**	**5.20**	**2.97**	**17.17**	**3.31**
Bbr_0845	glgP2	glgP2 Glycogen phosphorylase	—	—	—	2.73	—
Bbr_0846	nagA1	nagA1 N-acetylglucosamine-6-phosphate deacetylase	—	—	—	4.76	—
Bbr_0847	nagB2	nagB2 Glucosamine-6-phosphate isomerase	—	—	—	5.82	—
Bbr_0848	Bbr_0848	Sugar kinase, ROK family	3.30	5.12	—	10.69	—
Bbr_0849	Bbr_0849	NagC/XylR-type transciptional regulator	—	2.90	—	14.40	—
Bbr_0850	Bbr_0850	Aldose 1-epimerase family protein	—	—	—	7.39	—
Bbr_0851	Bbr_0851	Glucose/fructose transport protein	2.91	3.75	—	16.43	—
Bbr_0852	atsA2	Sulfatase family protein	—	—	—	4.18	—
Bbr_0853	atsB2	atsB Arylsulfatase regulator (Fe-S oxidoreductase)	—	—	—	2.70	—
Bbr_0854	Bbr_0854	Conserved hypothetical membrane spanning protein with DUF81 domain	—	—	—	4.03	—
Bbr_0855	Bbr_0855	Hypothetical protein	—	—	—	7.31	—
Bbr_0856	Bbr_0856	Conserved hypothetical membrane spanning protein	—	—	—	3.82	—
Bbr_1247	nagA2	CE9 nagA2 N-acetylglucosamine-6-phosphate deacetylase	2.67	5.06	3.92	—	—
Bbr_1248	nagB3	nagB3 Glucosamine-6-phosphate isomerase	4.11	5.95	8.51	—	—
Bbr_1249	Bbr_1249	Transcriptional regulator, ROK family	—	—	—	3.84	—
Bbr_1250	Bbr_1250	Sugar kinase, ROK family	2.44	7.70	—	8.75	—
Bbr_1251	Bbr_1251	N-acetylglucosamine repressor	—	—	—	7.71	—
Bbr_1252	pfkB	Fructokinase	—	—	—	—	—
Bbr_1550	Bbr_1550	Hypothetical protein	—	2.90	—	28.72	2.81
Bbr_1551	**lacS**	**Galactoside symporter**	**13.10**	**73.82**	**—**	**43.73**	**31.62**
Bbr_1552	**LacZ6**	**GH2 Beta-galactosidase**	**44.01**	**105.71**	**—**	**39.65**	**11.01**
Bbr_1553	**lacI**	**Transcriptional regulator, LacI family**	**—**	**4.56**	**—**	**15.60**	**—**
Bbr_1554	**nahS**	**Solute-binding protein of ABC transporter system (lactose)**	**5.71**	**15.05**	**9.91**	**13.10**	**—**
Bbr_1555	**nahR**	**NagC/XylR-type transciptional regulator**	**4.86**	**12.53**	**—**	**21.74**	**—**
Bbr_1556	**nahA**	**GH20 nagZ Beta-N-acetylhexosaminidase**	**2.65**	**3.74**	**—**	**2.90**	**—**
Bbr_1558	**nahP**	**Permease protein of ABC transporter system**	**—**	**—**	**4.97**	**—**	**—**
Bbr_1559	**nahT1**	**ATP-binding protein of ABC transporter system**	**—**	**—**	**—**	**2.66**	**—**
Bbr_1560	**nahT2**	**ATP-binding protein of ABC transporter system**	**—**	**—**	**3.42**	**2.51**	**—**
Bbr_1585	**galE**	**UDP-glucose 4-epimerase**	**—**	**3.11**	**3.33**	**4.12**	**—**
Bbr_1586	**nahK**	**Phosphotransferase family protein**	**4.55**	**10.69**	**3.19**	**—**	**—**
Bbr_1587	**lnbP**	**GH112 lacto-N-biose phorylase**	**—**	**3.86**	**6.52**	**—**	**—**
Bbr_1588	**galP1**	**Permease protein of ABC transporter system for sugars**	**3.38**	**4.20**	**6.38**	**—**	**—**
Bbr_1589	**galP2**	**Permease protein of ABC transporter system for sugars**	**2.78**	**5.47**	**4.05**	**—**	**—**
Bbr_1590	**galS**	**Solute-binding protein of ABC transporter system for sugars**	**4.41**	**4.45**	**16.49**	**—**	**—**

The level of transcription is shown as a fold-value of increase in transcription on each carbohydrate, as compared to a ribose control, with a cut-off of a minimum 2.5-fold increase in transcription. Genes within the 4 loci focused on in this study are shown in bold script.

a Based on comparative transcriptome analysis using B. breve UCC2003 grown on 1% LNT, LNnT, or LNB compared to growth on ribose. Microarray data were obtained using B. breve UCC2003 grown on 1% LNT, LNnT, or LNB and were compared with array data obtained when B. breve UCC2003 was grown on ribose as a control.

b The cutoff point is 2.5-fold, with a P value of _0.001. —, value below the cutoff.

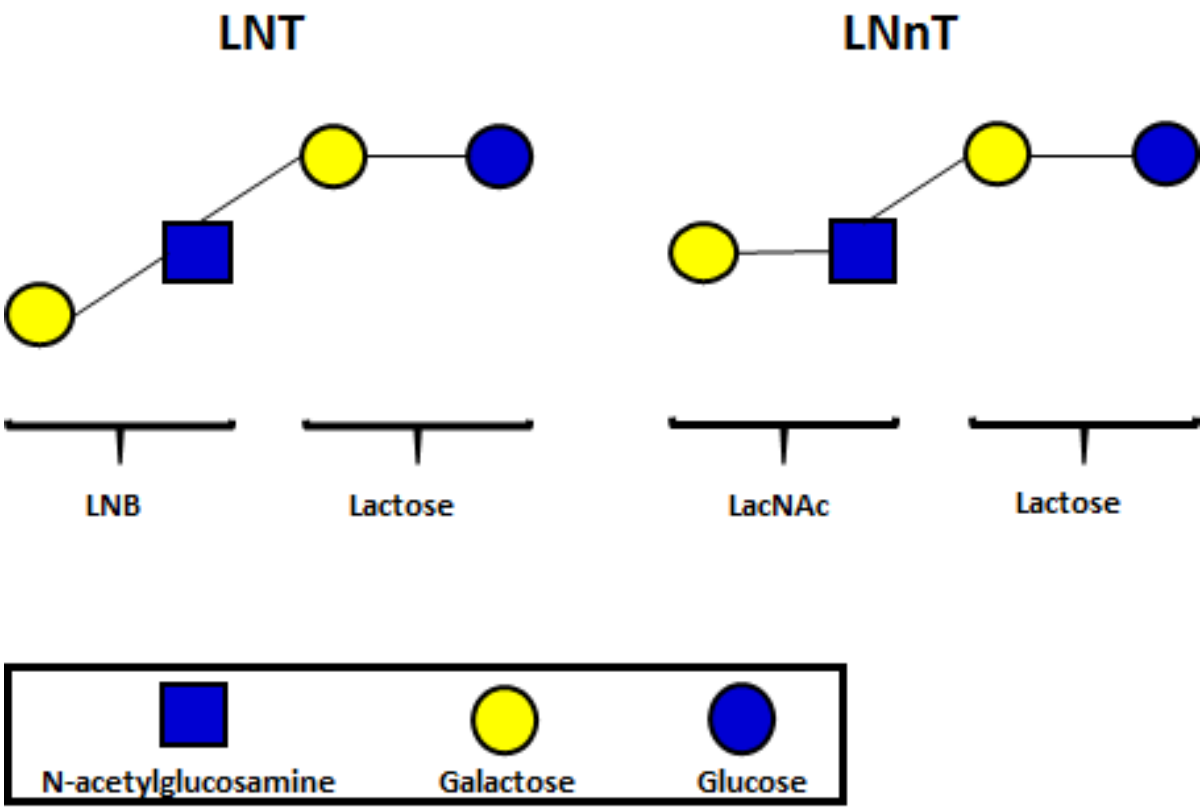


Figure 2.1. Schematic structures of Type I HMO moiety LNT, and Type II HMO moiety LNnT.

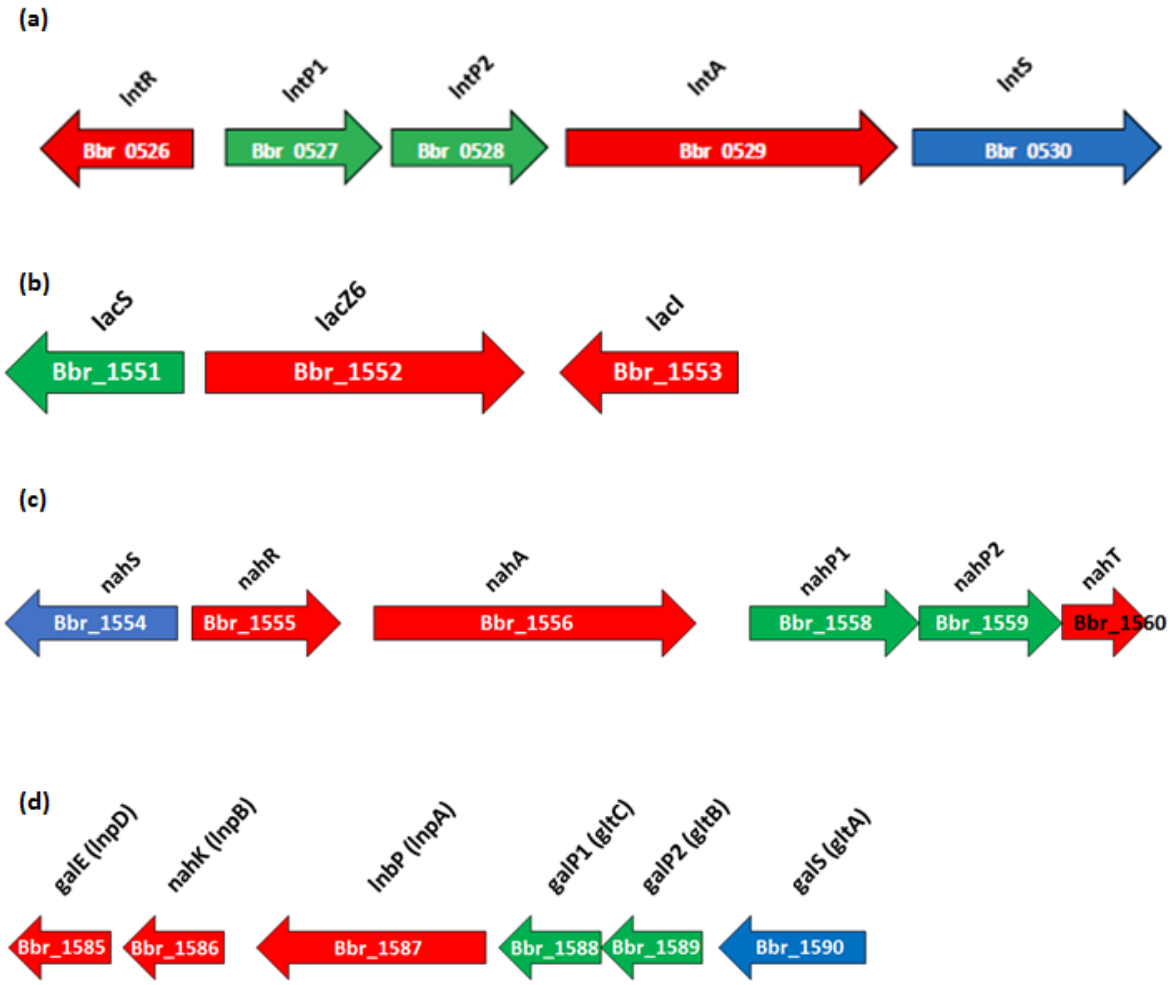


Figure 2.2. Schematic representation of the gene loci involved in the utilisation of LNT, LNnT and their substituents in *B. breve* UCC2003, as based on transcriptome analysis. The length of the arrows is proportional to the size of the open reading frame and the gene locus name, which is indicative of its putative function, is given at the top. Genes shown in red are predicted to encode proteins with an intracellular localisation, genes shown in green are predicted to encode proteins with a transmembrane localisation, and genes shown in blue are predicted to encode proteins with an extracellular localisation and a signal peptide sequence. The predicted function of each gene product is given in Table 2.1.

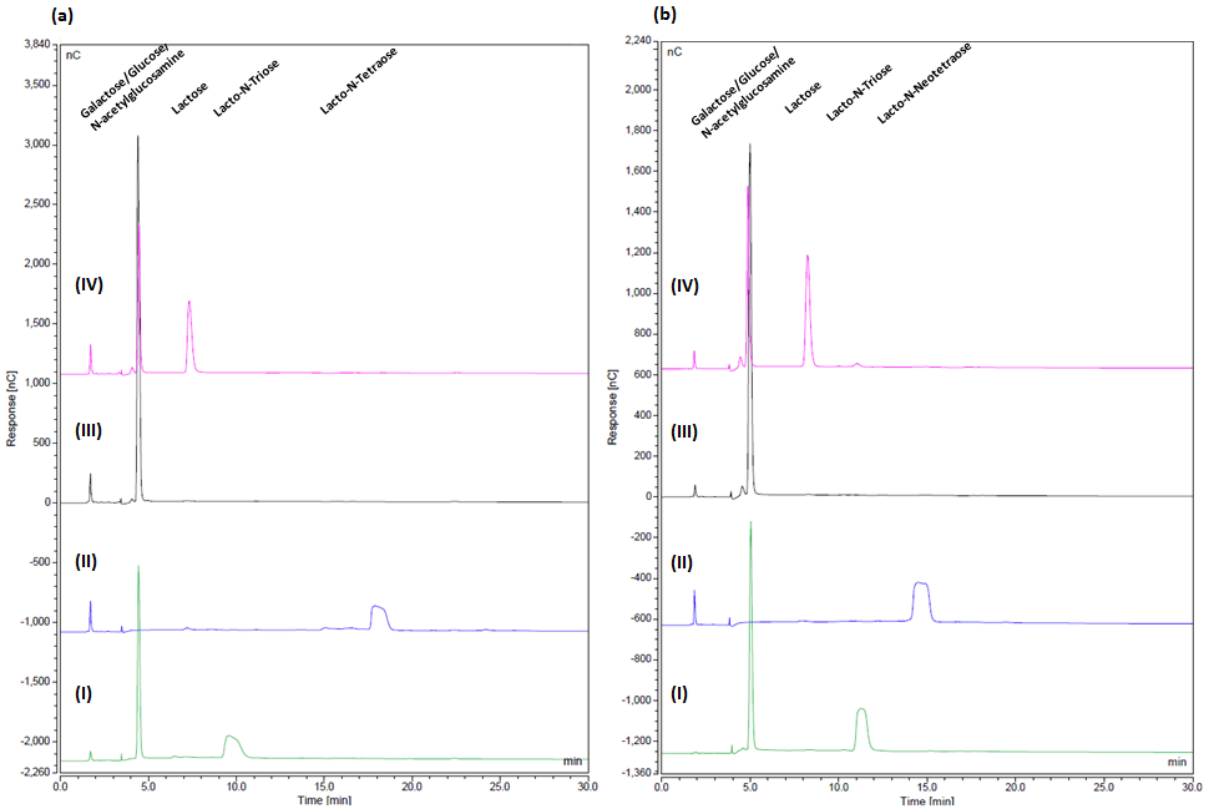


Figure 2.3. HPAEC chromatogram profiles of **(a)** LNT and **(b)** LNnT, when incubated in MOPS buffer (pH7) with: (I) LntA alone, (II) NahA alone, (III) LntA and NahA together, and (IV) LntA, followed by a denaturation step and the subsequent addition of NahA.

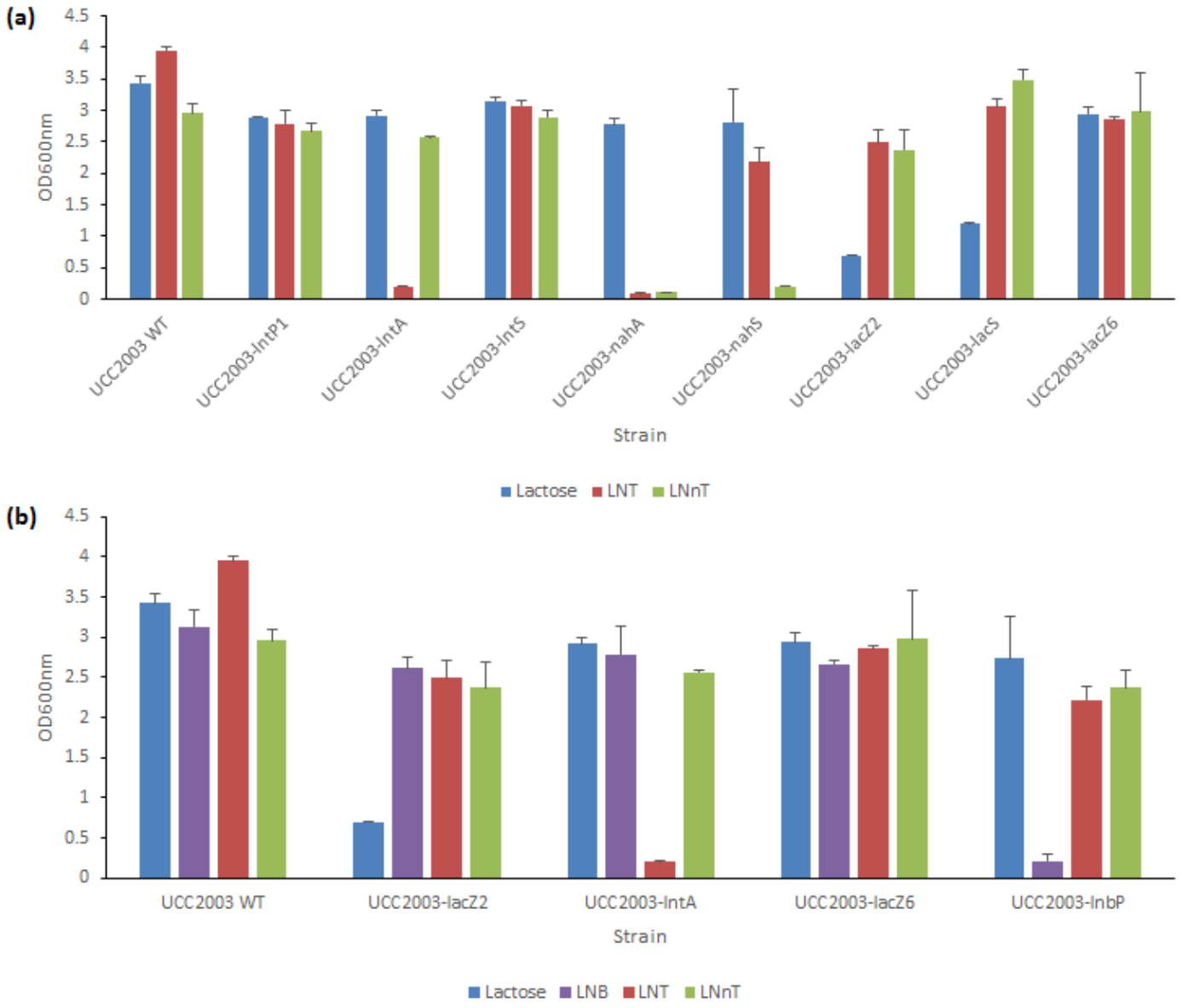
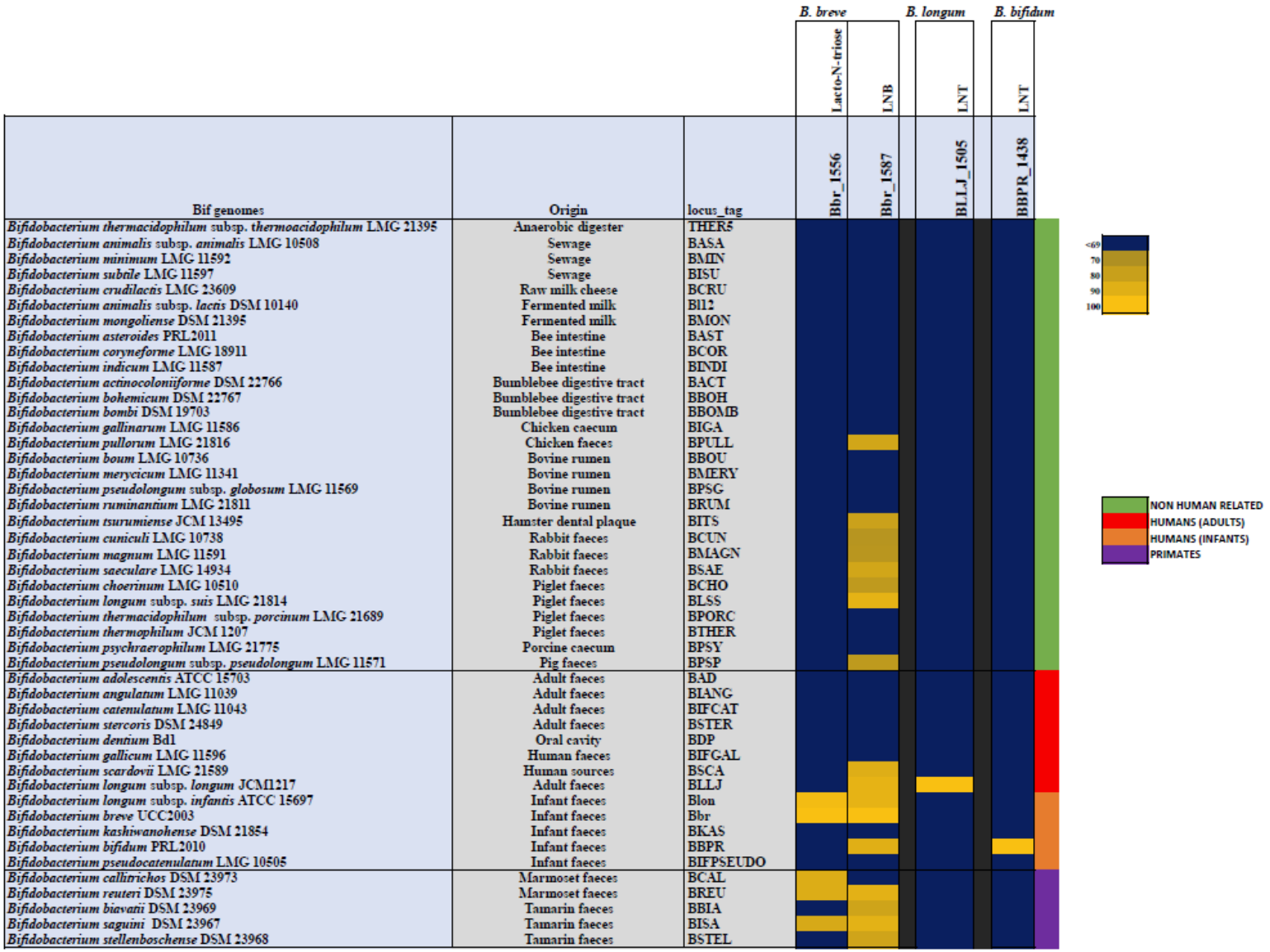

(a)
OD600nm
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
UCC2003 WT
UCC2003-lntP1
UCC2003-lntA
UCC2003-lntS
UCC2003-nahA
UCC2003-nahS
UCC2003-lacZ2
UCC2003-lacS
UCC2003-lacZ6
Strain
Lactose
LNT
LNnT
(b)
OD600nm
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
UCC2003 WT
UCC2003-lacZ2
UCC2003-lntA
UCC2003-lacZ6
UCC2003-lnbP
Strain
Lactose
LNB
LNT
LNnT

Figure 2.4. (a) Final OD_{600nm} values after 24 hours of growth of wild type *B. breve* UCC2003 and mutants *B. breve* UCC2003-lntP1, *B. breve* UCC2003-lntA, *B. breve* UCC2003-lntS, *B. breve* UCC2003-nahA, *B. breve* UCC2003-nahS, *B. breve* UCC2003-lacZ2, *B. breve* UCC2003-lacS and *B. breve* UCC2003-lacZ6 in modified MRS containing 1 % (wt/vol) lactose, 1 % (wt/vol) LNT or 1 % (wt/vol) LNnT as the sole carbon source. **(b)** Final OD_{600nm} values after 24 hours of growth of wild type *B. breve* UCC2003, and mutants *B. breve* UCC2003-lacZ2, *B. breve* UCC2003-lntA, *B. breve* UCC2003-lacZ6, and *B. breve* UCC2003-lnbP in modified MRS containing 1 % (wt/vol) lactose, 1% LNB, 1 % (wt/vol) LNT, 1 % (wt/vol) LNnT as the sole carbon source.

The results are the mean values obtained manually from two separate experiments (due to the limited availability of certain carbohydrates). Error bars represent the standard deviation.



Bif genomes	Origin	locus_tag
Bifidobacterium thermacidophilum subsp. *thermoacidophilum* LMG 21395	Anaerobic digester	THER5
Bifidobacterium animalis subsp. *animalis* LMG 10508	Sewage	BASA
Bifidobacterium minimum LMG 11592	Sewage	BMIN
Bifidobacterium subtile LMG 11597	Sewage	BISU
Bifidobacterium crudilactis LMG 23609	Raw milk cheese	BCRU
Bifidobacterium animalis subsp. *lactis* DSM 10140	Fermented milk	Bl12
Bifidobacterium mongoliense DSM 21395	Fermented milk	BMON
Bifidobacterium asteroides PRL2011	Bee intestine	BAST
Bifidobacterium coryneforme LMG 18911	Bee intestine	BCOR
Bifidobacterium indicum LMG 11587	Bee intestine	BINDI
Bifidobacterium actinocoloniiforme DSM 22766	Bumblebee digestive tract	BACT
Bifidobacterium bohemicum DSM 22767	Bumblebee digestive tract	BBOH
Bifidobacterium bombi DSM 19703	Bumblebee digestive tract	BBOMB
Bifidobacterium gallinarum LMG 11586	Chicken caecum	BIGA
Bifidobacterium pullorum LMG 21816	Chicken faeces	BPULL
Bifidobacterium boum LMG 10736	Bovine rumen	BBOU
Bifidobacterium merycicum LMG 11341	Bovine rumen	BMERY
Bifidobacterium pseudolongum subsp. *globosum* LMG 11569	Bovine rumen	BPSG
Bifidobacterium ruminantium LMG 21811	Bovine rumen	BRUM
Bifidobacterium tsurumiense JCM 13495	Hamster dental plaque	BITS
Bifidobacterium cuniculi LMG 10738	Rabbit faeces	BCUN
Bifidobacterium magnum LMG 11591	Rabbit faeces	BMAGN
Bifidobacterium saeculare LMG 14934	Rabbit faeces	BSAE
Bifidobacterium choerinum LMG 10510	Piglet faeces	BCHO
Bifidobacterium longum subsp. *suis* LMG 21814	Piglet faeces	BLSS
Bifidobacterium thermacidophilum subsp. *porcinum* LMG 21689	Piglet faeces	BPORC
Bifidobacterium thermophilum JCM 1207	Piglet faeces	BTHER
Bifidobacterium psychraerophilum LMG 21775	Porcine caecum	BPSY
Bifidobacterium pseudolongum subsp. *pseudolongum* LMG 11571	Pig faeces	BPSP
Bifidobacterium adolescentis ATCC 15703	Adult faeces	BAD
Bifidobacterium angulatum LMG 11039	Adult faeces	BIANG
Bifidobacterium catenulatum LMG 11043	Adult faeces	BIFCAT
Bifidobacterium stercoris DSM 24849	Adult faeces	BSTER
Bifidobacterium dentium Bd1	Oral cavity	BDP
Bifidobacterium gallicum LMG 11596	Human faeces	BIFGAL
Bifidobacterium scardovii LMG 21589	Human sources	BSCA
Bifidobacterium longum subsp. *longum* JCM1217	Adult faeces	BLLJ
Bifidobacterium longum subsp. *infantis* ATCC 15697	Infant faeces	Blon
Bifidobacterium breve UCC2003	Infant faeces	Bbr
Bifidobacterium kashiwanohense DSM 21854	Infant faeces	BKAS
Bifidobacterium bifidum PRL2010	Infant faeces	BBPR
Bifidobacterium pseudocatenulatum LMG 10505	Infant faeces	BIFPSEUDO
Bifidobacterium callitrichos DSM 23973	Marmoset faeces	BCAL
Bifidobacterium reuteri DSM 23975	Marmoset faeces	BREU
Bifidobacterium biavatii DSM 23969	Tamarin faeces	BBIA
Bifidobacterium saguini DSM 23967	Tamarin faeces	BISA
Bifidobacterium stellenboschense DSM 23968	Tamarin faeces	BSTEL

Figure 2.5. Heatmap representing the distribution of homologs of two genes from *B. breve* UCC2003, one gene from *B. longum* subsp. *longum* JCM1217 and one gene from *B. bifidum* PRL2010 across the *Bifidobacterium* genus. Gene products from the representative strain genomes of all online-available *Bifidobacterium* species with a significant homology of 70 % iterative similarity over 50 % of protein length are represented in the matrix, which employs a code colour grading that represents the degree of sequence similarity, with species ordered by origin of isolation. Bbr_1556 (*nahA*) and Bbr_1587 (*lnbP*) were selected from *B. breve* UCC2003, BLLJ_1505 (*lnbX*) was selected from *B. longum* subsp. *longum* JCM1217, and BBPR_1438 (*lnbB*) was selected from *B. bifidum* PRL2010.

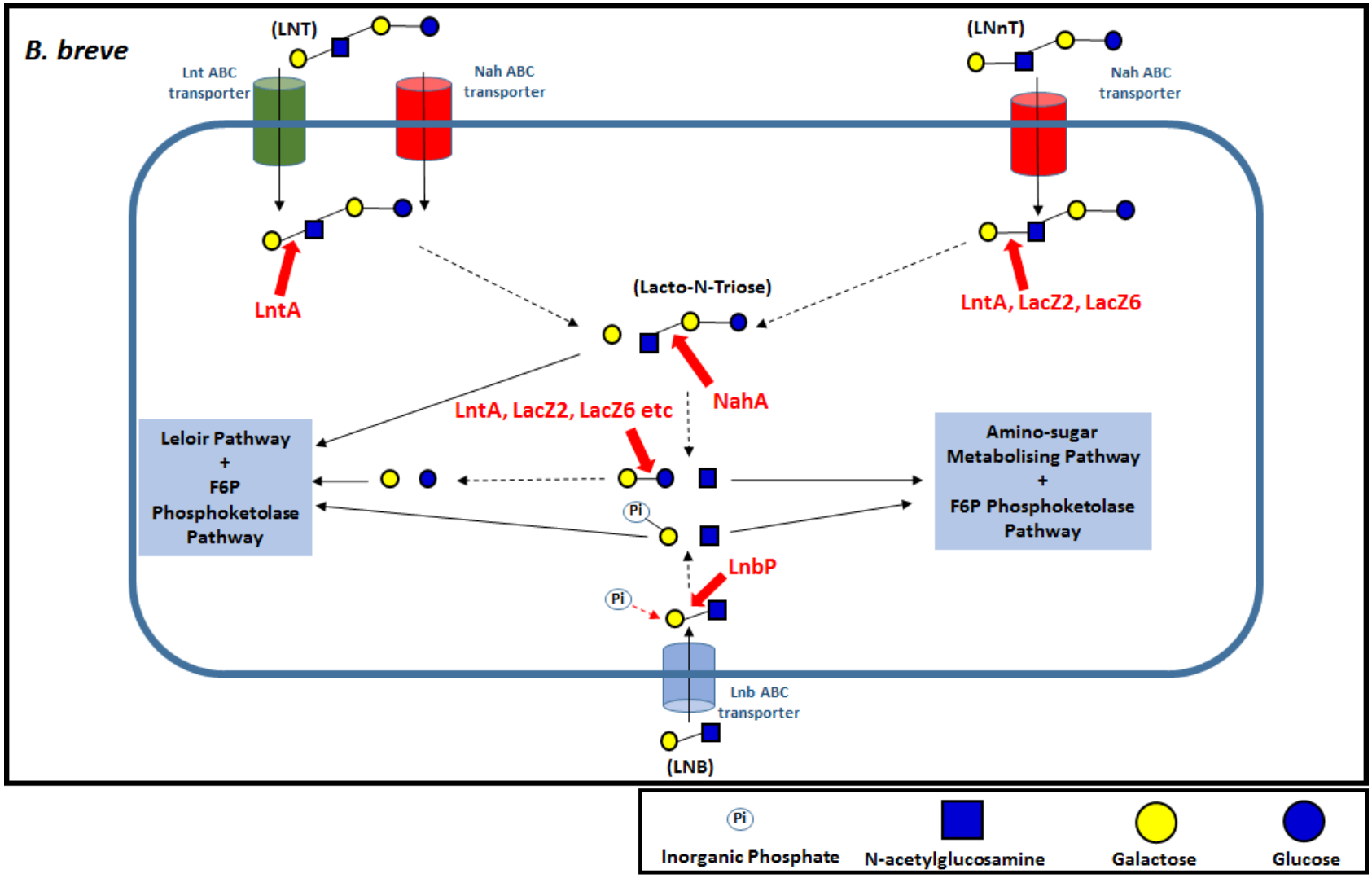


Figure 2.6. Schematic representation of the proposed model for the metabolism of free LNT, LNnT and LNB by *B. breve* UCC2003.

2.8 Supplemental Material

Supplemental Table S2.1. Oligonucleotide primers used in this work.

Purpose	Primer	Sequence (5'-3')
Cloning of Bbr_1556 in pNZ8150	1556F	tgcatc**cagctg**atg*catcaccatcaccatcaccatcaccatcac*aacggtcaaacatctctgcaa
	1556R	tgcgca**tctaga**ggattggaagcgcagatggtc
Cloning of internal 465bp fragment of Bbr_0530 in pORI19	IM530F	ctggtc**aagctt**gaaggagaccggcatcaagg
	IM530R	ctggtc**tctaga**gtctggaacggcttggcgc
Cloning of internal 488bp fragment of Bbr_1554 in pORI19	IM1554F	ctggtc**aagctt**gaaggagaccggcatcaagg
	IM1554R	ctggtc**tctaga**gtctggaacggcttggcgc
Cloning of internal 443bp fragment of Bbr_1556 in pORI19	IM1556F	ctggtc**aagctt**cgaatacctcgccatccc
	IM1556R	ctggtc**tctaga**ccagtgacttcatccagtcgtc
Amplification of *tet*W	tetWFw	tcagctgtcgacatgctcatgtacggtaaggaagca
	tetWRv	gcgacggtcgaccataacttctgattgttgccg
Confirmation of site specific homologous recombination	526confirm1	gcgctagctgttacaatggtc
	526confirm2	gccatttccaacccctctc
	530confirm1	gaggaaggagaacagaacgatg
	530confirm2	gcgaaaggaatggcttcatag
	1554confrim1	ctgcgacgccaatgccgacaag
	1554confrim2	ccgctactgcaggtcgccatcatcaagcac
	1556confrim1	gcagctcatcgacgacatccgc
	1556confrim2	gcgcgcgcagtcggtgac
Cloning of 2086bp fragment containing Bbr_0529 into pNZ44	529pNZ44F	ctggtc**ctgcag**gactcccatttgcccacacgtag
	529pNZ44R	cgagc**tctaga**gatttctcctcgacactg
Cloning of p44-529 in pBC1.2	P44 Forward	ctggtcc**ggatcc**gagataatgccgactgtac
Cloning of 1637bp fragment containing Bbr_1554 into PBC1.2	1554PBC1.2F	ctggtccggatccgccggaacaacagtctgc
	1554PBC1.2R	cgagctctagaattcaccgtttggattattgg
Cloning of 2360bp fragment containing Bbr_1556 into PBC1.2	1556PBC1.2F	ctggtccggatccccacgcaggtggacaagt
	1556PBC1.2R	cgagctctagagaaatctcacaatcgcagcc

Restriction sites incorporated into oligonucleotide primer sequences are indicated in bold, and His-tag sequences incorporated into nucleotide primer sequences are indicated in italics.

Supplemental Table S2.2. Bacterial plasmids and strains used in this work.

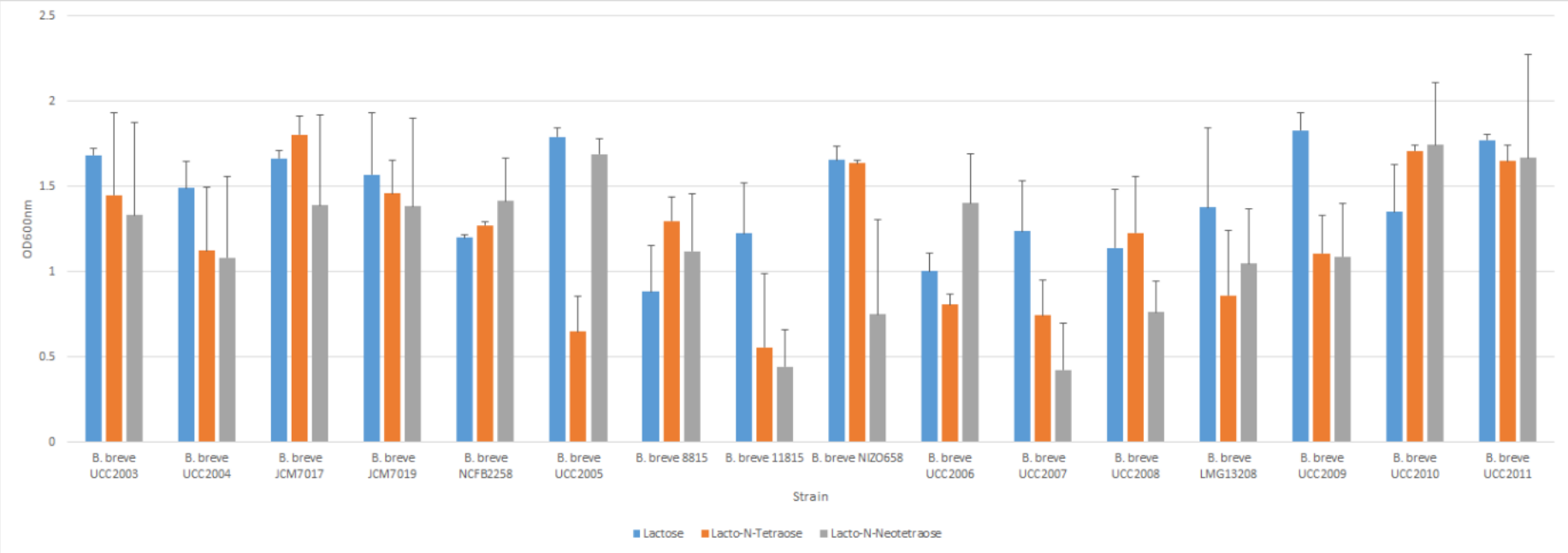
Strain or plasmid	Relevant Features (antibiotic resistances are given in brackets)	Reference or Source
Strains		
Escherichia coli strains		
E. coli EC101	Cloning host, repA$^+$ (km^r)	[54]
E. coli EC101-pNZ-M.BbrII + M.BbrIII	EC101 harbouring pNZ8048 derivative containing *bbrIIM* and *bbrIIIM*	[52]
E. coli XL1-Blue	*supE44 hsdR17 recA1 gyrA96 thi relA1 lac* F' [*proAB+ lacI*q *lacZΔM15* Tn*10*(Tet^r)]	Stratagene
E.coli XL1-blue-pBC1.2-529	XL1-blue containing pBC1.2-529 (Cm^r)	This study
Lactococcus lactis strains		
L. lactis NZ9000	MG1363, pepN::nisRK, nisin-inducible overexpression host	[67]
L. lactis NZ9700	Nisin-producing strain (Cm^r)	[67]
L. lactis NZ9000-pNZ-lacZ2	NZ9000 conatining pNZ-010 (Cm^r)	[62]
L. lactis NZ9000-pNZ-lntA	NZ9000 conatining pNZ-529 (Cm^r)	[62]
L. lactis NZ9000-pNZ-lacZ6	NZ9000 conatining pNZ-1552 (Cm^r)	[62]
L. lactis NZ9000-pNZ-nahA	NZ9000 conatining pNZ-1556 (Cm^r)	This study
L. lactis NZ9000-pNZ44-lntA	NZ9000 containing pNZ44-529 (Cm^r)	This study
Bifidobacterium sp. Strains		
B. breve UCC2003	Isolate from nursling stool	[53]
B. breve UCC2003-lacZ2	Tn5-transposon mutant in Bbr_0010 of UCC2003 (Tet^r)	[77]
B. breve UCC2003-lntP1	pORI19-tet-bbr_0527 insertion mutant of UCC2003 (Tet^r)	[62]
B. breve UCC2003-lntA	pORI19-tet-bbr_0529 insertion mutant of UCC2003 (Tet^r)	[62]
B. breve UCC2003-lacS	pORI19-tet-bbr_1551 insertion mutant of UCC2003 (Tet^r)	[62]
B. breve UCC2003-lacZ6	pORI19-tet-bbr_1552 insertion mutant of UCC2003 (Tet^r)	[62]
B. breve UCC2003-lnbP	pORI19-tet-bbr_1587 insertion mutant of UCC2003 (Tet^r)	[74]
B. breve UCC2003-lntS	pORI19-tet-bbr_0530 insertion mutant of UCC2003 (Tet^r)	This study
B. breve UCC2003-nahS	pORI19-tet-bbr_1554 insertion mutant of UCC2003 (Tet^r)	This study
B. breve UCC2003-nahA	pORI19-tet-bbr_1556 insertion mutant of UCC2003 (Tet^r)	This study
B. breve UCC2003-lntA-PBC1.2-lntA	pORI19-tet-529 insertion mutant of UCC2003 harbouring complementation construct pBC1.2-529 (Tet^r)(Cm^r)	This study
B. breve UCC2003-nahS-PBC1.2-nahS	pORI19-tet-1554 insertion mutant of UCC2003 harbouring complementation construct pBC1.2-1554 (Tet^r)(Cm^r)	This study
B. breve UCC2003-nahA-PBC1.2-nahA	pORI19-tet-1556 insertion mutant of UCC2003 harbouring complementation construct pBC1.2-1556 (Tet^r)(Cm^r)	This study
B. breve UCC2004	Isolate from infant faeces	UCC
B. breve JCM7017	Isolate from human faeces	JCM
B. breve JCM7019	Isolate from infant faeces	JCM
B. breve NCFB2258	Isolate from infant intestine	NCFB
B. breve UCC2005	Isolate from nursling stool	UCC
B. breve 8815	Isolate from human faeces	NCIMB
B. breve 11815	Isolate from infant intestine	NCIMB
B. breve NIZO658	Isolate from human faeces	NIZO
B. breve UCC2006	Mother's milk	UCC
B. breve UCC2007	Mother's milk	UCC
B. breve UCC2008	Isolate from infant/adult faeces	UCC
B. breve LMG13208	Isolate from infant faeces	LMG
B. breve UCC2009	Isolate from infant/adult faeces	UCC
B. breve UCC2010	Mother's milk	UCC
B. breve UCC2011	Mother's milk	UCC
Plasmids		
pAM5	pBC1-puC19-(Tet^r)	[61]
pORI19	Em^r, repA$^-$, ori$^+$, cloning vector (Em^r)	[54]
pORI19-tet-lntS	Internal 465 bp fragment of *bbr_0530* and tetW cloned in pORI19 (Em^r)	This study
pORI19-tet-nahS	Internal 443 bp fragment of *bbr_1554* and tetW cloned in pORI19 (Em^r)	This study
pORI19-tet-nahA	Internal 443 bp fragment of *bbr_1556* and tetW cloned in pORI19 (Em^r)	This study
pNZ8150	Cm^r, nisin inducible translational fusion vector (Cm^r)	[63]
pNZ-nahA	Cm^r, pNZ8150 derivative containing translational fusion of Bbr_1556 encoding DNA fragment to nisin inducible promoter (Cm^r)	This study
pNZ44	pNZ8048 containing constitutive p44 promoter from Lactococcal chromosome (Cm^r)	[60]
pNZ44-lntA	pNZ44 harbouring Bbr_0529 downstream of p44 promoter (Cm^r)	This study
pBC1.2	pBC1-pSC101-(Cm^r)	[61]
pBC1.2-lntA	pBC1-pSC101-Cmr harbouring Bbr_0529 downstream of p44 promoter (Cm^r)	This study
pBC1.2-nahS	pBC1-pSC101-Cmr harbouring Bbr_1554 (Cm^r)	This study
pBC1.2-nahA	pBC1-pSC101-Cmr harbouring Bbr_1556 (Cm^r)	This study

Cm^r, Em^r,Km^r and Tet^r, resistance to chloramphenicol, erythromycin, kanamycin and tetracycline, respectively.

Collection of the University of Goteborg**;** CIP**,** Collection de l'Institut Pasteur; DSM**,** German Collection of Microorganisms and Cell Cultures; JCM, Japan Collection of Microorganisms; LMG, Belgian Co-ordinated Collection of Microorganisms; NCDO, National Collection of Dairy Organisms; NCFB, National Collection of Food Bacteria; NCIMB, National Collection of Industrial and Marine Bacteria; NCTC, National Collection of Type Cultures; UCC, University College Cork Culture Collection.

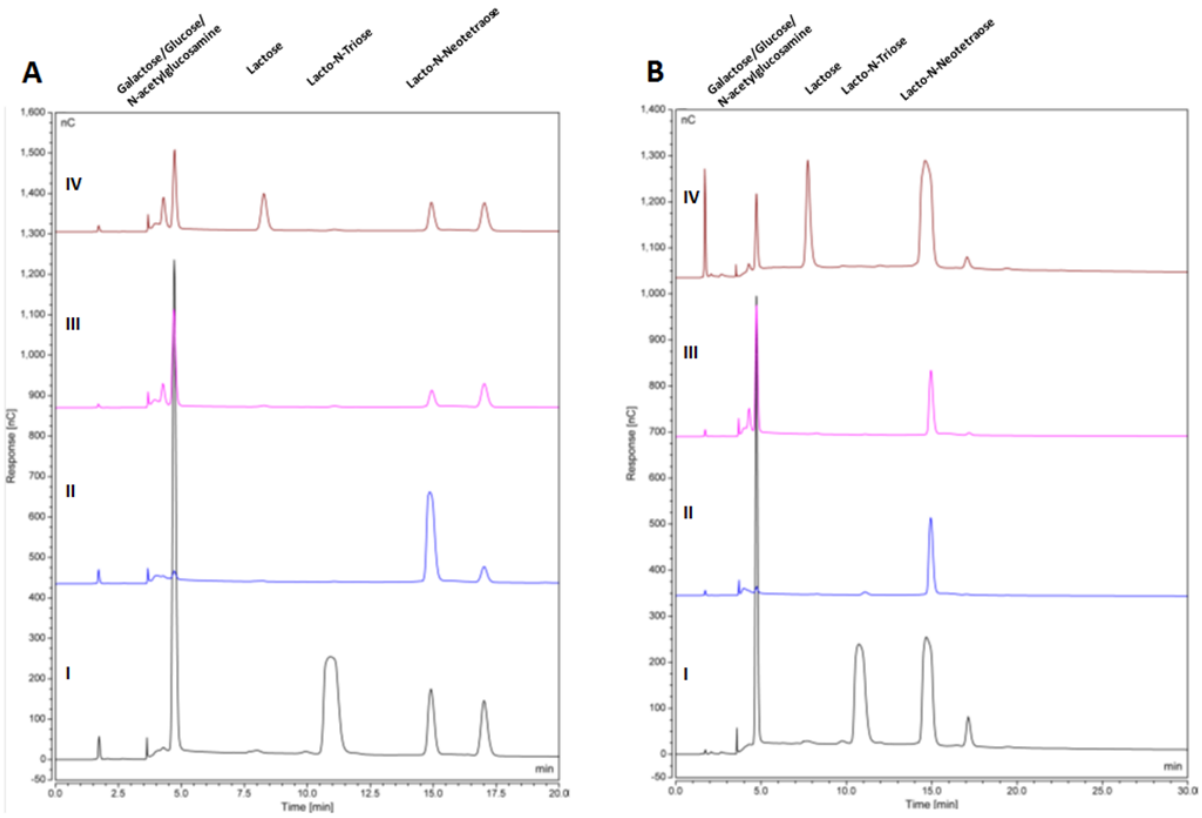
Supplemental Table S2.3. Locus tags, names and product functions of *B. breve* UCC2003 and *B. bifidum* PRL2010 genes selected for bioinformatic analysis of LN(n)T/LNB utilisation-associated homologs across the *Bifidobacterium* genus.

Gene ID	Gene Name	Location of Activity	Function in HMO Metabolism
***Bifidobacterium breve* UCC2003**			
Bbr_1556	nahA	Intracellular	nagZ Beta-N-acetylhexosaminidase; hydrolyses Lacto-N-Triose at GlcNAcβ1-3Gal residue, releasing GlcNAc and galactose.
Bbr_1587	lnbP	Intracellular	Lacto-N-biose phosphorylase; hydrolyses LNB, releasing GlcNAc and phosphorylated galactose.
***Bifidobacterium bifidum* PRL2010**			
BBPR_1438	lnbB	Extracellular	Lacto-N-biosidase; hydrolyses LNT at its GlcNAcβ1-3Gal residue, releasing LNB and lactose.
***Bifidobacterium longum* subsp. *longum* JCM1217**			
BLLJ_1505	lnbX	Extracellular	Lacto-N-biosidase; hydrolyses LNT at its GlcNAcβ1-3Gal residue, releasing LNB and lactose.

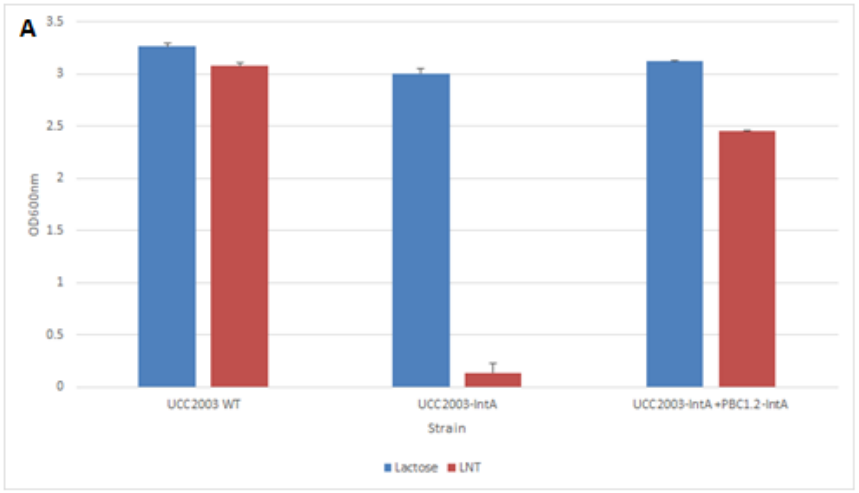


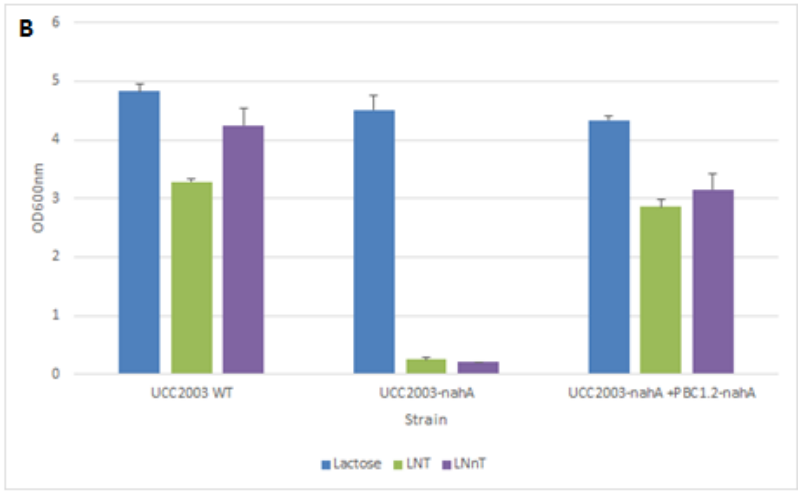
Supplemental Figure S2.1. Final OD_{600nm} values (measured using PowerWave microplate spectrophotometer) after 24 hours of growth of 16 *B. breve* strains in modified MRS containing 1 % (wt/vol) LNT, 1 % (wt/vol) LNnT or 1 % (wt/vol) lactose as the sole carbon source.

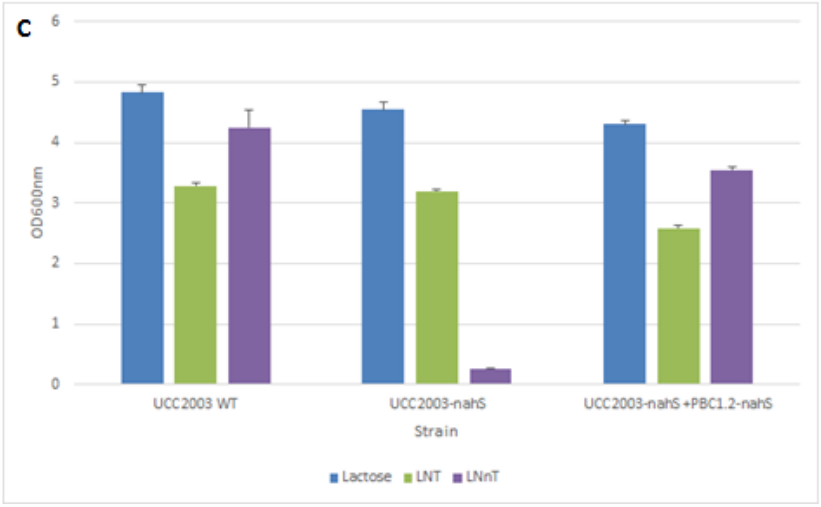
The results are the mean values obtained manually from two separate experiments (due to the limited availability of certain carbohydrates). Error bars represent the standard deviation.



Supplemental Figure S2.2. (A) HPAEC chromatogram profiles of LNnT, when incubated in MOPS buffer (pH7) with: (I) LacZ2 alone, (II) NahA alone, (III) LacZ2 and NahA together, and (IV) LacZ2, followed by a denaturation step and the subsequent addition of NahA. **(B)** HPAEC chromatogram profiles of LNnT, when incubated in MOPS buffer (pH7) with: (I) LacZ6 alone, (II) NahA alone, (III) LacZ6 and NahA together, and (IV) LacZ6, followed by a denaturation step and the subsequent addition of NahA.


A
3.5
3
2.5
2
1.5
1
0.5
0
OD600nm
UCC2003 WT
UCC2003-lntA
UCC2003-lntA +PBC1.2-lntA
Strain
Lactose
LNT


B
6
5
4
3
2
1
0
OD600nm
UCC2003 WT
UCC2003-nahA
UCC2003-nahA +PBC1.2-nahA
Strain
Lactose
LNT
LNnT


C
6
5
4
3
2
1
0
OD600nm
UCC2003 WT
UCC2003-nahS
UCC2003-nahS +PBC1.2-nahS
Strain
Lactose
LNT
LNnT

Supplemental Figure S2.3. (A) Final OD600 values (measured manually) after 24 hours of growth of wild type B. breve UCC2003, the insertion mutant *B. breve* UCC2003-lntA, and the complementation strain *B. breve* UCC2003-lntA+PBC1.2-lntA in modified MRS containing 1 % (wt/vol) lactose or 1 % (wt/vol) LNT as the sole carbon source. The results are the mean values obtained from two separate experiments*. Error bars represent the standard deviation. **(B)** Final OD600 values (measured manually) after 24 hours of growth of wild type B. breve UCC2003, the insertion mutant *B. breve* UCC2003-nahA, and the complementation strain *B. breve* UCC2003-nahA+PBC1.2-nahA in modified MRS containing 1 % (wt/vol) lactose or 1 % (wt/vol) LNT as the sole carbon source. The results are the mean values obtained from two separate experiments*. Error bars represent the standard deviation. **(C)** Final OD600 values (measured manually) after 24 hours of growth of wild type B. breve UCC2003, the insertion mutant *B. breve* UCC2003-nahS, and the complementation strain *B. breve* UCC2003-nahS+PBC1.2-nahS in modified MRS containing 1 % (wt/vol) lactose or 1 % (wt/vol) LNT as the sole carbon source. The results are the mean values obtained from two separate experiments*. Error bars represent the standard deviation.

*Experiments were carried out in duplicate due to a limited amount of carbohydrate material.

2.9 References:

1. **Engfer, M.B., B. Stahl, B. Finke, G. Sawatzki, and H. Daniel,** *Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract.* The American Journal of Clinical Nutrition, **2000**. **71**(6): p. 1589-1596.
2. **Kunz, C., S. Rudloff, W. Baier, N. Klein, and S. Strobel,** *OLIGOSACCHARIDES IN HUMAN MILK: Structural, Functional, and Metabolic Aspects.* Annual Review of Nutrition, **2000**. **20**(1): p. 699-722.
3. **Brüssow, H.,** *Human microbiota: 'The philosophers have only interpreted the world in various ways. The point, however, is to change it'.* Microbial Biotechnology, **2015**. **8**(1): p. 11-12.
4. **Morowitz, M.J., V. Poroyko, M. Caplan, J. Alverdy, and D.C. Liu,** *Redefining the Role of Intestinal Microbes in the Pathogenesis of Necrotizing Enterocolitis.* Pediatrics, **2010**. **125**(4): p. 777-785.
5. **Sim, K., A.G. Shaw, P. Randell, M.J. Cox, Z.E. McClure, M.-S. Li, M. Haddad, P.R. Langford, W.O.C.M. Cookson, M.F. Moffatt, and J.S. Kroll,** *Dysbiosis Anticipating Necrotizing Enterocolitis in Very Premature Infants.* Clinical Infectious Diseases, **2015**. **60**(3): p. 389-397.
6. **Di Gioia, D., I. Aloisio, G. Mazzola, and B. Biavati,** *Bifidobacteria: their impact on gut microbiota composition and their applications as probiotics in infants.* Appl Microbiol Biotechnol, **2014**. **98**(2): p. 563-77.
7. **Amisano, G., S. Fornasero, G. Migliaretti, S. Caramello, V. Tarasco, and F. Savino,** *Diarrheagenic Escherichia coli in acute gastroenteritis in infants in North-West Italy.* New Microbiol, **2011**. **34**(1): p. 45-51.
8. **Matamoros, S., C. Gras-Leguen, F. Le Vacon, G. Potel, and M.F. de La Cochetiere,** *Development of intestinal microbiota in infants and its impact on health.* Trends Microbiol, **2013**. **21**(4): p. 167-73.
9. **Turroni, F., C. Peano, D.A. Pass, E. Foroni, M. Severgnini, M.J. Claesson, C. Kerr, J. Hourihane, D. Murray, F. Fuligni, M. Gueimonde, A. Margolles, G. De Bellis, P.W. O'Toole, D. van Sinderen, J.R. Marchesi, and M. Ventura,** *Diversity of Bifidobacteria within the Infant Gut Microbiota.* PLoS ONE, **2012**. **7**(5): p. e36957.
10. **Simone, M., C. Gozzoli, A. Quartieri, G. Mazzola, D. Di Gioia, A. Amaretti, S. Raimondi, and M. Rossi,** *The Probiotic Bifidobacterium breve B632 Inhibited the*

Growth of Enterobacteriaceae within Colicky Infant Microbiota Cultures. BioMed Research International, **2014**. **2014**(301053): p. 301053.

11. **Di Gioia, D., I. Aloisio, G. Mazzola, and B. Biavati,** *Bifidobacteria: their impact on gut microbiota composition and their applications as probiotics in infants.* Applied Microbiology and Biotechnology, **2014**. **98**(2): p. 563-577.
12. **Kau, A.L., P.P. Ahern, N.W. Griffin, A.L. Goodman, and J.I. Gordon,** *Human nutrition, the gut microbiome, and immune system: envisioning the future.* Nature, **2011**. **474**(7351): p. 327-336.
13. **Maslowski, K.M. and C.R. Mackay,** *Diet, gut microbiota and immune responses.* Nat Immunol, **2011**. **12**(1): p. 5-9.
14. **Sivan, A., L. Corrales, N. Hubert, J.B. Williams, K. Aquino-Michaels, Z.M. Earley, F.W. Benyamin, Y. Man Lei, B. Jabri, M.-L. Alegre, E.B. Chang, and T.F. Gajewski,** *Commensal Bifidobacterium promotes antitumor immunity and facilitates anti–PD-L1 efficacy.* Science, **2015**. **350**(6264): p. 1084-1089.
15. **Subramanian, S., L.V. Blanton, Steven A. Frese, M. Charbonneau, David A. Mills, and Jeffrey I. Gordon,** *Cultivating Healthy Growth and Nutrition through the Gut Microbiota.* Cell. **161**(1): p. 36-48.
16. **Charbonneau, Mark R., D. O'Donnell, Laura V. Blanton, Sarah M. Totten, Jasmine C.C. Davis, Michael J. Barratt, J. Cheng, J. Guruge, M. Talcott, James R. Bain, Michael J. Muehlbauer, O. Ilkayeva, C. Wu, T. Struckmeyer, D. Barile, C. Mangani, J. Jorgensen, Y.-m. Fan, K. Maleta, Kathryn G. Dewey, P. Ashorn, Christopher B. Newgard, C. Lebrilla, David A. Mills, and Jeffrey I. Gordon,** *Sialylated Milk Oligosaccharides Promote Microbiota-Dependent Growth in Models of Infant Undernutrition.* Cell. **164**(5): p. 859-871.
17. **Asakuma, S., E. Hatakeyama, T. Urashima, E. Yoshida, T. Katayama, K. Yamamoto, H. Kumagai, H. Ashida, J. Hirose, and M. Kitaoka,** *Physiology of Consumption of Human Milk Oligosaccharides by Infant Gut-associated Bifidobacteria.* The Journal of Biological Chemistry, **2011**. **286**(40): p. 34583-34592.
18. **Sela, D.A. and D.A. Mills,** *Nursing our microbiota: molecular linkages between bifidobacteria and milk oligosaccharides.* Trends Microbiol, **2010**. **18**(7): p. 298-307.
19. **Asakuma, S., T. Urashima, M. Akahori, H. Obayashi, T. Nakamura, K. Kimura, Y. Watanabe, I. Arai, and Y. Sanai,** *Variation of major neutral oligosaccharides levels in human colostrum.* Eur J Clin Nutr, **2007**. **62**(4): p. 488-494.

20. **Jantscher-Krenn, E. and L. Bode,** *Human milk oligosaccharides and their potential benefits for the breast-fed neonate.* Minerva Pediatr, **2012**. **64**(1): p. 83-99.
21. **Sela, D.A.,** *Bifidobacterial utilization of human milk oligosaccharides.* International Journal of Food Microbiology, **2011**. **149**(1): p. 58-64.
22. **LoCascio, R.G., M.R. Ninonuevo, S.L. Freeman, D.A. Sela, R. Grimm, C.B. Lebrilla, D.A. Mills, and J.B. German,** *Glycoprofiling of Bifidobacterial Consumption of Human Milk Oligosaccharides Demonstrates Strain Specific, Preferential Consumption of Small Chain Glycans Secreted in Early Human Lactation.* Journal of Agricultural and Food Chemistry, **2007**. **55**(22): p. 8914-8919.
23. **Benno, Y., K. Sawada, and T. Mitsuoka,** *The intestinal microflora of infants: composition of fecal flora in breast-fed and bottle-fed infants.* Microbiol Immunol, **1984**. **28**(9): p. 975-86.
24. **Guaraldi, F. and G. Salvatori,** *Effect of Breast and Formula Feeding on Gut Microbiota Shaping in Newborns.* Frontiers in Cellular and Infection Microbiology, **2012**. **2**(94): p. 94.
25. **Solís, G., C.G. de los Reyes-Gavilan, N. Fernández, A. Margolles, and M. Gueimonde,** *Establishment and development of lactic acid bacteria and bifidobacteria microbiota in breast-milk and the infant gut.* Anaerobe, **2010**. **16**(3): p. 307-310.
26. **Morita, H., A. Nakano, H. Onoda, H. Toh, K. Oshima, H. Takami, M. Murakami, S. Fukuda, T. Takizawa, T. Kuwahara, H. Ohno, S. Tanabe, and M. Hattori,** *Bifidobacterium kashiwanohense sp. nov., isolated from healthy infant faeces.* Int J Syst Evol Microbiol, **2011**. **61**(Pt 11): p. 2610-5.
27. **Vazquez-Gutierrez, P., C. Lacroix, T. Jaeggi, C. Zeder, M.B. Zimmerman, and C. Chassard,** *Bifidobacteria strains isolated from stools of iron deficient infants can efficiently sequester iron.* BMC Microbiol, **2015**. **15**(1): p. 3.
28. **Kitaoka, M.,** *Bifidobacterial enzymes involved in the metabolism of human milk oligosaccharides.* Adv Nutr, **2012**. **3**(3): p. 422s-9s.
29. **Yoshida, E., H. Sakurama, M. Kiyohara, M. Nakajima, M. Kitaoka, H. Ashida, J. Hirose, T. Katayama, K. Yamamoto, and H. Kumagai,** *Bifidobacterium longum subsp. infantis uses two different beta-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides.* Glycobiology, **2012**. **22**(3): p. 361-8.
30. **Garrido, D., S. Ruiz-Moyano, and D.A. Mills,** *Release and utilization of N-acetyl-D-glucosamine from human milk oligosaccharides by Bifidobacterium longum subsp. infantis.* Anaerobe, **2012**. **18**(4): p. 430-5.

31. **Xiao, J.Z., S. Takahashi, M. Nishimoto, T. Odamaki, T. Yaeshima, K. Iwatsuki, and M. Kitaoka,** *Distribution of in vitro fermentation ability of lacto-N-biose I, a major building block of human milk oligosaccharides, in bifidobacterial strains.* Appl Environ Microbiol, **2010**. **76**(1): p. 54-9.

32. **Wada, J., T. Ando, M. Kiyohara, H. Ashida, M. Kitaoka, M. Yamaguchi, H. Kumagai, T. Katayama, and K. Yamamoto,** *Bifidobacterium bifidum lacto-N-biosidase, a critical enzyme for the degradation of human milk oligosaccharides with a type 1 structure.* Appl Environ Microbiol, **2008**. **74**(13): p. 3996-4004.

33. **Wada, J., R. Suzuki, S. Fushinobu, M. Kitaoka, T. Wakagi, H. Shoun, H. Ashida, H. Kumagai, T. Katayama, and K. Yamamoto,** *Purification, crystallization and preliminary X-ray analysis of the galacto-N-biose-/lacto-N-biose I-binding protein (GL-BP) of the ABC transporter from Bifidobacterium longum JCM1217.* Acta Crystallogr Sect F Struct Biol Cryst Commun, **2007**. **63**(Pt 9): p. 751-3.

34. **Miwa, M., T. Horimoto, M. Kiyohara, T. Katayama, M. Kitaoka, H. Ashida, and K. Yamamoto,** *Cooperation of beta-galactosidase and beta-N-acetylhexosaminidase from bifidobacteria in assimilation of human milk oligosaccharides with type 2 structure.* Glycobiology, **2010**. **20**(11): p. 1402-9.

35. **Sela, D.A., J. Chapman, A. Adeuya, J.H. Kim, F. Chen, T.R. Whitehead, A. Lapidus, D.S. Rokhsar, C.B. Lebrilla, J.B. German, N.P. Price, P.M. Richardson, and D.A. Mills,** *The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome.* Proceedings of the National Academy of Sciences of the United States of America, **2008**. **105**(48): p. 18964-18969.

36. **Møller, P.L., F. Jørgensen, O.C. Hansen, S.M. Madsen, and P. Stougaard,** *Intra- and Extracellular β-Galactosidases from Bifidobacterium bifidum and B. infantis: Molecular Cloning, Heterologous Expression, and Comparative Characterization.* Applied and Environmental Microbiology, **2001**. **67**(5): p. 2276-2283.

37. **Sakurama, H., M. Kiyohara, J. Wada, Y. Honda, M. Yamaguchi, S. Fukiya, A. Yokota, H. Ashida, H. Kumagai, M. Kitaoka, K. Yamamoto, and T. Katayama,** *Lacto-N-biosidase Encoded by a Novel Gene of Bifidobacterium longum Subspecies longum Shows Unique Substrate Specificity and Requires a Designated Chaperone for Its Active Expression.* The Journal of Biological Chemistry, **2013**. **288**(35): p. 25194-25206.

38. **Ruiz-Moyano, S., S.M. Totten, D. Garrido, J.T. Smilowitz, J.B. German, C.B. Lebrilla, and D.A. Mills,** *Variation in consumption of human milk oligosaccharides by infant-gut associated strains of Bifidobacterium breve.* Applied and Environmental Microbiology, **2013**. **79**(19): p. 6040–6049.

39. **Egan, M., M. O'Connell Motherway, M. Ventura, and D. van Sinderen,** *Metabolism of sialic acid by Bifidobacterium breve UCC2003.* Applied and Environmental Microbiology, **2014**. **80**(14): p. 4414-26.

40. **Matsuki, T., K. Yahagi, H. Mori, H. Matsumoto, T. Hara, S. Tajima, E. Ogawa, H. Kodama, K. Yamamoto, T. Yamada, S. Matsumoto, and K. Kurokawa,** *A key genetic factor for fucosyllactose utilization affects infant gut microbiota development.* Nat Commun, **2016**. **7**(11939).

41. **De Man JC, R.M., Sharpe ME,** *A medium for the cultivation of 688 lactobacilli.* J.Appl. Bacteriol., **1960**. **23**: p. 130-135.

42. **Watson, D., M. O'Connell Motherway, M.H. Schoterman, R.J. van Neerven, A. Nauta, and D. van Sinderen,** *Selective carbohydrate utilization by lactobacilli and bifidobacteria.* J Appl Microbiol, **2013**. **114**(4): p. 1132-46.

43. **Terzaghi, B.E. and W.E. Sandine,** *Improved medium for lactic streptococci and their bacteriophages.* Appl Microbiol, **1975**. **29**(6): p. 807-13.

44. **Sambrook J, F.E., Maniatis T.**, *Molecular cloning: a laboratory manual, 2nd ed.* **1989**: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

45. **McLaughlin, H.P., M.O. Motherway, B. Lakshminarayanan, C. Stanton, R. Paul Ross, J. Brulc, R. Menon, P.W. O'Toole, and D. van Sinderen,** *Carbohydrate catabolic diversity of bifidobacteria and lactobacilli of human origin.* Int J Food Microbiol, **2015**. **203**: p. 109-21.

46. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M.A. Rajandream, and B. Barrell,** *Artemis: sequence visualization and annotation.* Bioinformatics, **2000**. **16**(10): p. 944-5.

47. **O'Connell Motherway, M., A. Zomer, S.C. Leahy, J. Reunanen, F. Bottacini, M.J. Claesson, F. O'Brien, K. Flynn, P.G. Casey, J.A. Moreno Munoz, B. Kearney, A.M. Houston, C. O'Mahony, D.G. Higgins, F. Shanahan, A. Palva, W.M. de Vos, G.F. Fitzgerald, M. Ventura, P.W. O'Toole, and D. van Sinderen,** *Functional genome analysis of Bifidobacterium breve UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor.* Proceedings of the National Academy of Sciences, **2011**. **108**(27): p. 11217-11222.

48. **Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman,** *Basic local alignment search tool.* J Mol Biol, **1990**. **215**(3): p. 403-10.

49. **Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman,** *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res, **1997**. **25**(17): p. 3389-402.

50. **Petersen, T.N., S. Brunak, G. von Heijne, and H. Nielsen,** *SignalP 4.0: discriminating signal peptides from transmembrane regions.* Nat Meth, **2011**. **8**(10): p. 785-786.

51. **O'Riordan, K. and G.F. Fitzgerald,** *Molecular characterisation of a 5.75-kb cryptic plasmid from Bifidobacterium breve NCFB 2258 and determination of mode of replication.* FEMS Microbiol Lett, **1999**. **174**(2): p. 285-94.

52. **O'Connell Motherway, M., J. O'Driscoll, G.F. Fitzgerald, and D. Van Sinderen,** *Overcoming the restriction barrier to plasmid transformation and targeted mutagenesis in Bifidobacterium breve UCC2003.* Microb Biotechnol, **2009**. **2**(3): p. 321-32.

53. **Maze, A., M. O'Connell-Motherway, G.F. Fitzgerald, J. Deutscher, and D. van Sinderen,** *Identification and characterization of a fructose phosphotransferase system in Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2007**. **73**(2): p. 545-53.

54. **Law, J., G. Buist, A. Haandrikman, J. Kok, G. Venema, and K. Leenhouts,** *A system to generate chromosomal mutations in Lactococcus lactis which allows fast analysis of targeted genes.* J Bacteriol, **1995**. **177**(24): p. 7011-8.

55. **Zomer, A., M. Fernandez, B. Kearney, G.F. Fitzgerald, M. Ventura, and D. van Sinderen,** *An interactive regulatory network controls stress response in Bifidobacterium breve UCC2003.* J Bacteriol, **2009**. **191**(22): p. 7039-49.

56. **Garcia de la Nava, J., D.F. Santaella, J. Cuenca Alba, J. Maria Carazo, O. Trelles, and A. Pascual-Montano,** *Engene: the processing and exploratory analysis of gene expression data.* Bioinformatics, **2003**. **19**(5): p. 657-8.

57. **van Hijum, S., A. de Jong, R. Baerends, H. Karsens, N. Kramer, R. Larsen, C. den Hengst, C. Albers, J. Kok, and O. Kuipers,** *A generally applicable validation scheme for the assessment of factors involved in reproducibility and quality of DNA-microarray data.* BMC Genomics, **2005**. **6**(1): p. 1-10.

58. **van Hijum, S.A., J. Garcia de la Nava, O. Trelles, J. Kok, and O.P. Kuipers,** *MicroPreP: a cDNA microarray data pre-processing framework.* Appl Bioinformatics, **2003**. **2**(4): p. 241-4.

59. **Long, A.D., H.J. Mangalam, B.Y. Chan, L. Tolleri, G.W. Hatfield, and P. Baldi,** *Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in Escherichia coli K12.* J Biol Chem, **2001**. **276**(23): p. 19937-44.

60. **McGrath, S., G.F. Fitzgerald, and D. van Sinderen,** *Improvement and optimization of two engineered phage resistance mechanisms in Lactococcus lactis.* Appl Environ Microbiol, **2001**. **67**(2): p. 608-16.

61. **Álvarez-Martín, P., M. O'Connell-Motherway, D. van Sinderen, and B. Mayo,** *Functional analysis of the pBC1 replicon from Bifidobacterium catenulatum L48.* Applied Microbiology and Biotechnology, **2007**. **76**(6): p. 1395-1402.

62. **O'Connell Motherway, M., M. Kinsella, G.F. Fitzgerald, and D. van Sinderen,** *Transcriptional and functional characterization of genetic elements involved in galacto-oligosaccharide utilization by Bifidobacterium breve UCC2003.* Microb Biotechnol, **2013**. **6**(1): p. 67-79.

63. **Mierau, I. and M. Kleerebezem,** *10 years of the nisin-controlled gene expression system (NICE) in Lactococcus lactis.* Appl Microbiol Biotechnol, **2005**. **68**(6): p. 705-17.

64. **O'Connell Motherway, M., G.F. Fitzgerald, and D. van Sinderen,** *Metabolism of a plant derived galactose-containing polysaccharide by Bifidobacterium breve UCC2003.* Microbial Biotechnology, **2011**. **4**(3): p. 403-416.

65. **Pokusaeva, K., M. O'Connell-Motherway, A. Zomer, J. Macsharry, G.F. Fitzgerald, and D. van Sinderen,** *Cellodextrin utilization by bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2011**. **77**(5): p. 1681-90.

66. **O'Connell, K.J., M. O'Connell Motherway, J. O'Callaghan, G.F. Fitzgerald, R.P. Ross, M. Ventura, C. Stanton, and D. van Sinderen,** *Metabolism of four alpha-glycosidic linkage-containing oligosaccharides by Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2013**. **79**(20): p. 6280-92.

67. **de Ruyter, P.G., O.P. Kuipers, and W.M. de Vos,** *Controlled gene expression systems for Lactococcus lactis with the food-grade inducer nisin.* Applied and Environmental Microbiology, **1996**. **62**(10): p. 3662-3667.

68. **Bradford, M.M.,** *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Analytical Biochemistry, **1976**. **72**(1–2): p. 248-254.

69. **Enright, A.J., S. Van Dongen, and C.A. Ouzounis,** *An efficient algorithm for large-scale detection of protein families.* Nucleic Acids Research, **2002**. **30**(7): p. 1575-1584.

70. **Pokusaeva, K., A.R. Neves, A. Zomer, M. O'Connell-Motherway, J. MacSharry, P. Curley, G.F. Fitzgerald, and D. van Sinderen,** *Ribose utilization by the human commensal Bifidobacterium breve UCC2003.* Microb Biotechnol, **2010**. **3**(3): p. 311-23.

71. **Nishimoto, M. and M. Kitaoka,** *Identification of N-Acetylhexosamine 1-Kinase in the Complete Lacto-N-Biose I/Galacto-N-Biose Metabolic Pathway in Bifidobacterium longum.* Applied and Environmental Microbiology, **2007**. **73**(20): p. 6444-6449.

72. **Kitaoka, M., J. Tian, and M. Nishimoto,** *Novel Putative Galactose Operon Involving Lacto-N-Biose Phosphorylase in Bifidobacterium longum.* Applied and Environmental Microbiology, **2005**. **71**(6): p. 3158-3162.

73. **Turroni, F., E. Foroni, F. Serafini, A. Viappiani, B. Montanini, F. Bottacini, A. Ferrarini, P.L. Bacchini, C. Rota, M. Delledonne, S. Ottonello, D. van Sinderen, and M. Ventura,** *Ability of Bifidobacterium breve To Grow on Different Types of Milk: Exploring the Metabolism of Milk through Genome Analysis.* Applied and Environmental Microbiology, **2011**. **77**(20): p. 7408-7417.

74. **Egan, M., M. O'Connell Motherway, M. Kilcoyne, M. Kane, L. Joshi, M. Ventura, and D. van Sinderen,** *Cross-feeding by Bifidobacterium breve UCC2003 during co-cultivation with Bifidobacterium bifidum PRL2010 in a mucin-based medium.* BMC Microbiology, **2014**. **14**(282): p. 282.

75. **Derensy-Dron, D., F. Krzewinski, C. Brassart, and S. Bouquelet,** *Beta-1,3-galactosyl-N-acetylhexosamine phosphorylase from Bifidobacterium bifidum DSM 20082: characterization, partial purification and relation to mucin degradation.* Biotechnol Appl Biochem, **1999**. **29 (Pt 1)**: p. 3-10.

76. **Alvarez-Martin, P., M. O'Connell Motherway, F. Turroni, E. Foroni, M. Ventura, and D. van Sinderen,** *A two-component regulatory system controls autoregulated serpin expression in Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2012**. **78**(19): p. 7032-41.

77. **Ruiz, L., M.O. Motherway, N. Lanigan, and D. van Sinderen,** *Transposon mutagenesis in Bifidobacterium breve: construction and characterization of a Tn5 transposon mutant library for Bifidobacterium breve UCC2003.* PLoS One, **2013**. **8**(5): p. e64699.

78. **Goulas, T.K., A.K. Goulas, G. Tzortzis, and G.R. Gibson,** *Molecular cloning and comparative analysis of four beta-galactosidase genes from Bifidobacterium bifidum NCIMB41171.* Appl Microbiol Biotechnol, **2007**. **76**(6): p. 1365-72.

79. **Suzuki, R., T. Katayama, M. Kitaoka, H. Kumagai, T. Wakagi, H. Shoun, H. Ashida, K. Yamamoto, and S. Fushinobu,** *Crystallographic and mutational analyses of substrate recognition of endo-alpha-N-acetylgalactosaminidase from Bifidobacterium longum.* J Biochem, **2009**. **146**(3): p. 389-98.

80. **Fujita, K., F. Oura, N. Nagamine, T. Katayama, J. Hiratake, K. Sakata, H. Kumagai, and K. Yamamoto,** *Identification and molecular cloning of a novel glycoside hydrolase family of core 1 type O-glycan-specific endo-alpha-N-acetylgalactosaminidase from Bifidobacterium longum.* J Biol Chem, **2005**. **280**(45): p. 37415-22.

81. **Tao, N., S. Wu, J. Kim, H.J. An, K. Hinde, M.L. Power, P. Gagneux, J.B. German, and C.B. Lebrilla,** *Evolutionary Glycomics: Characterization of Milk Oligosaccharides in Primates.* Journal of proteome research, **2011**. **10**(4): p. 1548-1557.

Chapter III

Bifidobacterium breve UCC2003 employs multiple transcriptional regulators to control metabolism of particular human milk oligosaccharides.

EMSAs were carried out with the assistance of Rebecca Louise O'Brien.

Primer extension analysis was carried out with the assistance of Dr. Christophe Penno, for sample loading and imaging.

This chapter was published in:

James, K., O'Connell Motherway, M., Penno, C., O'Brien, R.L. & van Sinderen, D. (2018). *Bifidobacterium breve* UCC2003 employs multiple transcriptional regulators to control metabolism of particular human milk oligosaccharides. AEM Vol. **84**, No. **9**.

3.1 Abstract

Bifidobacterial carbohydrate metabolism has been studied in considerable detail for a variety of both plant and human-derived glycans, particularly involving the bifidobacterial prototype *Bifidobacterium breve* UCC2003. We recently elucidated the metabolic pathways by which the human milk oligosaccharide (HMO) constituents lacto-N-tetraose (LNT), lacto-N-neotetraose (LNnT) and lacto-N-biose (LNB) are utilized by *B. breve* UCC2003. However, to date no work has been carried out on the regulatory mechanisms that control expression of the genetic loci involved in these HMO metabolic pathways. In the current study, we describe the characterization of three transcriptional regulators and corresponding operator and associated (inducible) promoter sequences, the latter governing transcription of the genetic elements involved in LN(n)T/LNB metabolism. The activity of these regulators is dependent on the release of specific monosaccharides, which are believed to act as allosteric effectors, and which are derived from the corresponding HMOs targeted by the particular locus.

3.2 Introduction

Bifidobacteria represent high-G+C, Gram-positive, anaerobic members of the phylum Actinobacteria, and are common commensals of the mammalian, avian and insect gut. In humans, they are particularly abundant and prevalent among the gut microbiota of healthy, vaginally-delivered, breastfed infants [1], and are thought to confer a multitude of benefits to the neonatal host [2-4]. For this reason, as well as because of their purported health-promoting activities in adults, bifidobacteria are used as functional ingredients in a variety of foods and therapeutic products. The use of prebiotics is also becoming commonplace for the improvement of both adult and infant (gut) health. A prebiotic has been defined as 'a non-digestible food ingredient that beneficially affects the host by selectively stimulating the growth and/or activity of one or a limited number of bacteria in the colon, and thus improves host health' [5].

The archetypal prebiotic would appear to be human breastmilk, and in particular its bifidogenic constituents known as Human Milk Oligosaccharides (HMOs). HMOs represent specific glycans present in human breastmilk that are thought to shape, at least partly, the compositional structure of the neonatal gut microbiota [6, 7]. HMOs represent, after lactose, the second-largest carbohydrate component of breastmilk [6, 8], and constitute a heterogeneous mix of at least 200 distinct glycan structures [9]. The majority of complex HMO structures can be classified into one of two types, depending on their backbone composition. The more abundant Type I HMOs contain the core tetrasaccharide lacto-N-tetraose (LNT) within their structure (Galβ1-3GlcNAcβ1-3Galβ1-4Glc). Type II HMOs contain lacto-N-neotetraose (LNnT), a stereoisomer of LNT, within their backbone (Galβ1-4GlcNAcβ1-3Galβ1-4Glc). Lacto-N-biose (LNB; Galβ1-3GlcNAc) is a subunit of LNT and other Type I HMO structures, and can be released by the degradation of these sugars [10].

The effects of the specific human breastmilk components on the prevalence, abundance and activity of members of the infant gut microbiota are currently enjoying an increasing amount of scientific and commercial attention, due to the beneficial roles they are believed to play in infant health and development

[11, 12]. Understanding the pathways by which specific HMOs are metabolized by particular microbial species that inhabit the infant gut is important, although our knowledge regarding these processes is still in its infancy, particularly with regards to the manner in which they affect microbiota development.

It is not surprising that the dominant *Bifidobacterium* species found among the neonatal gut microbiota can utilise various HMO components as their sole carbohydrate source [9]. These species chiefly include strains of *Bifidobacterium bifidum*, *Bifidobacterium longum* subsp. *infantis* and *Bifidobacterium breve*. HMO utilization by *B. bifidum* and *B. longum* subsp. *infantis* is relatively well characterised. *B. bifidum* extracellularly hydrolyses complex HMO structures, including LNT and LNnT, employing secreted glycosyl hydrolases, followed by the internalisation and intracellular degradation/metabolism of (most of) the resulting mono- and di-saccharides, such as LNB [10, 13-19]. *B. longum* subsp. *infantis* internalises intact LNT, LNnT and LNB, and uses a series of sequential hydrolytic/phosphorolytic reactions acting from the non-reducing end of the carbohydrate structures to degrade them into their monosaccharide components for further metabolic processing [9, 10, 20-23]. However, *B. infantis* has also been demonstrated to take up and utilise fucosyl- and sialyl-lactose [24-26].

The metabolic pathways of LNT, LNnT and LNB have recently been elucidated in the prototype strain *B. breve* UCC2003 [27]. In the latter study, converging pathways of LNT and LNnT catabolism were identified, where monosaccharide moieties are sequentially released from the non-reducing end of either sugar by hydrolytic reactions. The genetic units responsible for the uptake and breakdown of these structures are the *lnt* locus (corresponding to locus tags Bbr_0526-0530) and the *nah* locus (locus tags Bbr_1554-1560) (Table 3.1 and Fig. 3.1). The *lnt* locus encodes proteins that are responsible for the internalisation of LNT and intracellular hydrolysis of both LNT and LNnT, releasing a galactose (Gal) moiety from their non-reducing end, and at the same time liberating the trisaccharide lacto-N-triose (GlcNAcβ1-3Galβ1-4Glc). The *nah* locus specifies an LNT/LNnT uptake system, while it furthermore encodes a glycosyl hydrolase which liberates GlcNAc from the

non-reducing end of lacto-N-triose, leaving lactose, which itself is further degraded by lactose-specific glycosyl hydrolases. Additionally, the gene products of the *lnp/glt* locus (corresponding to locus tags Bbr_1585-1590; Table 3.1 and Fig. 3.1) are responsible for the internalisation and subsequent phosphorolysis of free LNB, releasing its constituent monosaccharides Gal-1-phosphate and GlcNAc [27, 28]. We also identified the transcriptional upregulation of genes in the *nag* locus (locus tags Bbr_1247-1252; Table 3.1 and Fig. 3.1) during growth on LNT, LNnT and LNB, indicating their role in the utilisation of these sugars, specifically in the multi-step metabolism of GlcNAc. The *nag* locus has previously been implicated in the metabolism of sialic acid and mucin-derived N-glycans, both of which contain GlcNAc as well [29, 30]. While the degradation routes of these key HMO structures have thus been identified, the regulatory mechanisms that control expression of these pathways have remained unexplored, both for *B. breve* and HMO-utilising *Bifidobacterium* species as a whole.

In the current study, we identified and characterized the genes encoding transcriptional regulators responsible for control of gene expression in four key HMO-associated loci in *B. breve* UCC2003 during growth on LNT, LNnT or LNB.

3.3 Materials and Methods

Bacterial strains, plasmids, and culture conditions. Bacterial strains and plasmids used in this study are listed in Supplemental Table S3.2. *B. breve* UCC2003 was routinely cultured in either de Man Rogosa and Sharpe medium (MRS medium; Difco, BD, Le Pont de Claix, France) supplemented with 0.05 % cysteine-HCl or reinforced clostridial medium (RCM; Oxoid Ltd., Basingstoke, England). Growth of bifidobacterial strains for transcriptional and primer extension analyses was carried out in modified de Man Rogosa and Sharpe (mMRS) medium, which was prepared from first principles (using individual components) [31], and which does not contain a fixed carbohydrate source. Prior to inoculation, the mMRS medium was supplemented with cysteine-HCl (0.05 %, wt/vol) and a particular carbohydrate source (1 %, wt/vol). It has previously been shown that mMRS does not support growth of *B. breve* UCC2003 in the absence of an added carbohydrate [32]. Carbohydrates used were ribose (Sigma Aldrich, Steinheim, Germany), LNB (Elicityl Oligotech, Crolles, France) and LNnT (Glycom, Lyngby, Denmark). A 1 % wt/vol concentration of carbohydrate was considered sufficient to encourage adequate growth for RNA harvesting. The addition of these carbohydrates did not significantly alter the pH of the medium, and therefore subsequent pH adjustment was not required.

B. breve cultures were incubated under anaerobic conditions in a modular atmosphere-controlled system (Davidson and Hardy, Belfast, Ireland) at 37°C. *Lactococcus lactis* strains were cultivated in M17 broth (Oxoid Ltd., Basingstoke, England) containing 0.5 % glucose [33] at 30°C. *Escherichia coli* strains were cultured in Luria-Bertani (LB) broth [34] at 37°C with agitation. Where appropriate, growth media contained tetracycline (Tet; 10 μg ml^{-1}), chloramphenicol (Cm; 5 μg ml^{-1} for *L. lactis* and *E. coli*, 2.5 μg ml^{-1} for *B. breve*), erythromycin (Em; 100 μg ml^{-1}) or kanamycin (Kan; 50 μg ml^{-1}). Recombinant *E. coli* EC101 cells containing (derivatives of) pORI19 were selected on LB agar containing Em and Kan, and supplemented with X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (40 μg ml^{-1}) and 1 mM IPTG (isopropyl-β-D-galactopyranoside).

Nucleotide sequence analysis. Sequence information was obtained from the Artemis-mediated [35] genome annotations of *B. breve* UCC2003 [36]. Database searches were performed using non-redundant sequences accessible at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) using the basic local alignment search tool (BLAST) [37, 38]. Sequences were verified and analysed using the SeqMan and SeqBuilder programs of the DNAStar software package (version 10.1.2; DNAStar, Madison, WI, USA).

DNA Manipulations. Chromosomal DNA was isolated from *B. breve* UCC2003 as previously described [39]. Plasmid DNA was isolated from *Escherichia coli*, *Lactococcus lactis* and *B. breve* using the Roche High Pure Plasmid Isolation kit (Roche Diagnostics, Basel, Switzerland). An initial lysis step was performed using 30 mg ml^{-1} of lysozyme for 30 minutes at 37°C prior to plasmid isolation from *L. lactis* or *B. breve*. Procedures for DNA manipulations were essentially performed as described previously [34]. All restriction enzymes and T4 DNA ligase were used according to the supplier's instructions (Roche Diagnostics, Basel, Switzerland). Synthetic single stranded oligonucleotide primers used in this study (Supplemental Table S3.1) were synthesized by Eurofins (Ebersberg, Germany). Standard PCRs were performed using Taq PCRmaster mix (Qiagen) or Extensor Hi-Fidelity PCR Master Mix (Thermo Scientific, Waltham, United States) in a Biometra T3000 thermocycler Biometra, Göttingen, Germany) or a Life Technologies Proflex PCR System (Thermo Scientific, Waltham, United States). PCR products were visualized by ethidium bromide (EtBr) staining following agarose gel electrophoresis (1 % agarose). *B. breve* colony PCR reactions were performed as described previously [40]. PCR fragments were purified using the Roche high Pure PCR purification kit (Roche Diagnostics, Basel, Switzerland). Plasmid DNA was isolated using the Roche High Pure Plasmid Isolation kit (Roche Diagnostics, Basel, Switzerland). Plasmid DNA was introduced into *E. coli* by electroporation as described previously [34]. *B. breve* UCC2003 [41] and *L. lactis* [42] were transformed by electroporation according to published protocols. Correct orientation of DNA inserts and

integrity of all plasmid constructs (see also below) were verified by DNA sequencing, performed at Eurofins (Ebersberg, Germany).

Construction of *B. breve* UCC2003 insertion mutants. Internal fragments of Bbr_0526 (designated here as *lntR*) (367 base pairs [bp] representing codon numbers 40 through to 162 of the 320 codons of this gene), Bbr_1249 (designated here as *nagR1*) (502 bp representing codon numbers 64 through to 231 of the 375 codons of this gene), Bbr_1251 (designated here as *nagR2*) (507 bp representing codon numbers 62 through to 230 of the 405 codons of this gene), and Bbr_1555 (designated here as *nahR*) (448 bp representing codon numbers 74 through to 223 of the 380 codons of this gene) were amplified by PCR using *B. breve* UCC2003 chromosomal DNA as a template and primer pairs 526LacIInsFHindIII and 526LacIInsRXbaI, 1249LacIInsFHindIII and 1249LacIInsRXbaI, 1251LacIInsFHindIII and 1251LacIInsRXbaI, or 1555LacIInsFHindIII and 1555LacIInsRXbaI (Supplemental Table S3.1), respectively. The insertion mutants were constructed using a previously described approach [40], generating mutant strains *B. breve* UCC2003-lntR, *B. breve* UCC2003-nagR1, *B. breve* UCC2003-nagR2 and *B. breve* UCC2003-nahR, which carried disrupted *lntR*, *nagR1*, *nagR2* and *nahR* genes, respectively (Supplemental Table S3.2). Site-specific recombination of potential tet-resistant mutant isolates was confirmed by colony PCR using primer combinations TetWF and TetWR to verify *tetW* gene integration, and primers Bbr_526ConfirmP1 or Bbr_526ConfirmP2, Bbr_1249ConfirmP1 or Bbr_1249ConfirmP2, Bbr_1251ConfirmP1 or Bbr_1251ConfirmP2, and Bbr_1555ConfirmP1 or Bbr_1555ConfirmP2 (positioned upstream of the selected internal fragments of Bbr_0526, Bbr_1249, Bbr_1251 and Bbr_1555, respectively) in combination with primer TetWF to confirm integration at the correct chromosomal location (Supplemental Table S3.1).

Analysis of global gene expression using *B. breve* DNA microarrays. Global gene expression was determined during log-phase growth of the

insertional mutant strains *B. breve* UCC2003-*lntR*, *B. breve* UCC2003-*nagR1*, *B. breve* UCC2003-*nagR2* and *B. breve* UCC2003-*nahR* in mMRS supplemented with ribose. The generated transcriptome data sets were compared to the transcriptome information obtained for log-phase wild-type *B. breve* UCC2003 cells when grown in mMRS supplemented with ribose. Ribose was selected as a suitable transcriptomic reference, as the metabolic pathway and gene expression profile for growth of UCC2003 on ribose is known and has been employed previously as a reference [43, 44]. DNA microarrays containing oligonucleotide primers representing each of the 1864 identified open reading frames on the genome of *B. breve* UCC2003 were designed and obtained from Agilent Technologies (Palo Alto, Ca., USA). Methods for cell disruption, RNA isolation, RNA quality control, complementary DNA synthesis and labelling were performed as, described previously [45]. Two independent biological replicates were used for each array using a Cy3/Cy5 dye-swap, as described previously [45]. Labelled cDNA was hybridized using the Agilent Gene Expression hybridization kit (part number 5188-5242) as described in the Agilent Two-Colour Microarray-Based Gene Expression Analysis v4.0 manual (publication number G4140-90050). Following hybridization, microarrays were washed in accordance with Agilent's standard procedures and scanned using an Agilent DNA microarray scanner (model G2565A). Generated scans were converted to data files with Agilent's Feature Extraction software (Version 9.5). DNA-microarray data were processed as previously described [46-48]. Differential expression tests were performed with the Cyber-T implementation of a variant of the t-test [49].

Construction of overexpression vectors, protein overproduction and purification. For the construction of plasmids pNZ-lntR, pNZ-nagR1 and pNZ-nahR, DNA fragments encompassing *lntR*, *nagR1* and *nahR* were generated by PCR amplification from chromosomal DNA of *B. breve* UCC2003 using Q5 High-Fidelity DNA polymerase and primer combinations 526PurFSmaI and 526PurRXbaI, 1249PurFPvuII and 1249PurRXbaI, or 1555PurFEcoRV and 1555PurXbaI, respectively (Supplemental Table S3.1).

An in-frame N-terminal His10-encoding sequence was incorporated into the forward primers 526PurFSmaI, 1249PurFPvuII and 1555PurFEcoRV to facilitate downstream protein purification. The generated amplicons were digested with SmaI and XbaI, PvuII and XbaI, or EcoRV and XbaI, respectively, and ligated into the ScaI and XbaI digested, nisin-inducible translational fusion plasmid pNZ8150 [50]. The ligation mixtures were introduced into *L. lactis* NZ9000 by electrotransformation and transformants were then selected based on chloramphenicol (Cm) resistance. The plasmid content of a number of Cm-resistant transformants was screened by restriction analysis and the integrity of positively identified clones was verified by sequencing.

Nisin-inducible gene expression and protein overproduction was performed as described previously [51-53]. In brief, 50 ml of M17 broth supplemented with 0.5 % (wt/vol) glucose was inoculated with a 2 % inoculum of a particular *L. lactis* strain, followed by incubation at 30°C until an OD600 of 0.5 was reached, at which point protein expression was induced by addition of cell-free supernatant of a nisin-producing strain [54], followed by continued incubation for a further 2 hours. Cells were harvested by centrifugation, and crude cell extract was obtained as described previously [55]. Although protein purification of LntR-His, NahR-His and NagR1-His was achieved using His tag affinity chromatography, the purification procedure appeared to render the proteins inactive in subsequent electrophoretic mobility shift assays (EMSAs). For this reason, crude cell extracts, prepared in a 10 mM Tris-HCl lysis buffer (pH 7.0), were adopted for the EMSAs (see below).

Electrophoretic mobility shift assay (EMSA). DNA fragments representing different portions of the promoter regions upstream of *lntP1* (locus tag Bbr_0527) and *lntS* (locus tag Bbr_0530), *nagB3* (locus tag Bbr_1248) and *nagK* (locus tag Bbr_1250), *lnpB* (locus tag Bbr_1586) and *gltA* (locus tag Bbr_1590) and *nahS* (Bbr_1554) and *nahA* (Bbr_1556) were prepared by PCR using IRD700-labelled primers pairs (Integrated DNA Technologies, Coralville, Indiana, United States) (Supplemental Table S3.1). EMSAs were

performed essentially as described previously [51, 56]. In all cases, binding reactions were carried out in a final volume of 20 µl in the presence of poly[d(I-C)] in binding buffer (20 mM Tris-HCl, 5 mM MgCl2, 0.5 mM DTT, 1 mM EDTA, 100 mM KCl, 10 % glycerol). Varying amounts of crude protein extract, ranging from 140 ng to 180 ng, of the LntR-, NahR-, or NagR1-(over)producing *L. lactis* NZ9000 strain constructed, and a fixed amount of DNA probe (0.1 pmol) was mixed on ice and subsequently incubated for 15 min at 37°C. In order to assess if the binding activity of LntR, NahR, or NagR1 is modulated by a carbohydrate ligand, various carbohydrates including galactose, galactose-1-phosphate, galactose-6-phosphate (all Sigma Aldrich, Steinheim, Germany), LNT (Glycom, Lyngby, Denmark), LNnT (Glycom, Lyngby, Denmark), LNB (Elicityl Oligotech, Crolles, France), glucose, N-acetylglucosamine, N-acetylglucosamine-6-phosphate or lactose (all Sigma Aldrich, Steinheim, Germany) ranging in concentration from 50 to 0.0625 mM, were included to the binding reaction buffer. Samples were loaded onto a 6 % non-denaturing PAA gel prepared in TAE buffer (40 mM Tris acetate [pH 8.0], 2 mM EDTA) and run in a 0.5-to-2.0 x gradient of TAE at 100 V for 90 min in an Atto Mini PAGE system (Atto Bioscience and Biotechnology, Tokyo, Japan). Signals were detected using Odyssey Infrared Imaging System (Li-Cor Biosciences UK Ltd, Cambridge, UK) and captured using the supplied software Odyssey V3.0.

Primer extension analysis. Total RNA was isolated from *B. breve* UCC2003, grown in mMRS supplemented with 1 % LNnT or 1 % LNB, to early exponential phase, using a previously described Macaloid method [57]. RNA samples were treated with RNase-free DNase (Ambion). Primer extension was performed by annealing 1 pmol of IRD700 synthetic 18-mer oligonucleotides to 15 µg of RNA as described previously [58]. Sequence ladders of the presumed promoter regions immediately upstream of *lntP1*, *lntS*, *nagB3*, *nagK*, *lnpB*, *gltA*, *nahS* or *nahA*, amplified from both UCC2003 genomic DNA which were run alongside the primer extension products, were produced using the same primer as the primer extension reaction and employing the Thermo Sequenase Primer Cycle Sequencing Kit; Amersham. Separation was achieved on a 6.5 % Li-Cor Matrix KB Plus acrylamide gel.

Signal detection and image capture was performed by means of a Li-Cor sequencing instrument (Li-Cor Biosciences).

Operator Consensus Sequence Prediction, using MEME and WEBLOGO online software tools. Co-regulated promoter regions were assessed for the presence of operator sequences by the use of the MEME (Multiple Em for Motif Elicitation) online tool (http://meme-suite.org/tools/meme) [59], which were the visualized by the WebLogo online tool (http://weblogo.berkeley.edu/logo.cgi) [60, 61]. Sequences used for consensus sequence prediction are given in Supplemental Table S3.4.

Microarray data accession number. The microarray data obtained in this study have been deposited in NCBI's Gene Expression Omnibus database and are accessible through GEO Series accession number GSE105108.

3.4 Results

Identification of putative transcriptional regulator-encoding genes in the vicinity of HMO-utilisation loci. In a previous study, we had observed that genes within four chromosomal loci exhibit transcriptional induction during growth of *B. breve* UCC2003 on LNT, LNnT or LNB as the sole carbohydrate source [27]. This indicates that these genes are subject to transcriptional regulation, which was presumed to be either directly or indirectly controlled by the presence of these HMO substrates. The four loci concerned are: the *lnt* locus (Bbr_0526-530), the *nah* locus (Bbr_1554-1560), the *nag* locus (Bbr_1247-1250) and the *lnp/glt* locus (Bbr_1585-1590) (Fig. 3.1; see Table 3.1 for a description of [predicted] functions). Detailed scrutiny of these four loci and neighbouring regions showed that the *lnt* and *nah* loci are flanked by or contain a predicted regulator-encoding gene, respectively: *lntR* (Bbr_0526), encoding a LacI-type repressor, and *nahR* (Bbr_1555), encoding a NagC/XylR-type repressor (Fig. 3.1A and 3.1B). The *nag* locus is associated with two genes, *nagR1* (Bbr_1249) and *nagR2* (Bbr_1251), both predicted to encode ROK/NagC family-type repressors, while no regulator-encoding gene was observed in close vicinity of the *lnp/glt locus* (Fig. 3.1C and 3.1D). NagC/XylR-type and ROK/NagC-type repressors are both members of the large family of ROK-type transcriptional regulators [62]. The four identified putative regulator-encoding genes were thus selected as candidates for mutagenesis in order to ascertain their role, if any, in the transcriptional regulation of the *lnt*, *nah*, *nag* and *lnp*/*glt* loci.

Generation and transcriptomic analysis of insertional mutants in putative HMO-associated regulator-encoding genes. Individual insertional mutants were constructed in *lntR*, *nahR*, *nagR1* and *nagR2*, resulting in *B. breve* strains UCC2003-lntR, UCC2003-nahR, UCC2003-nagR1 and UCC2003-nagR2, respectively (see Materials and Methods). In order to identify promoters/genes that are subject to transcriptional control of these predicted regulators, global gene transcription data was obtained from microarray-based analyses performed on the *B. breve* UCC2003-lntR, UCC2003-nahR, UCC2003-nagR1 and UCC2003-nagR2 insertion mutants,

grown in mMRS supplemented with ribose, as compared to the transcriptome of the UCC2003 wild type strain grown under the same conditions.

Transcriptome analysis of the *lntR* mutant revealed the upregulation of the adjacent *lntP1*, *lntP2*, *lntA* and *lntS* genes of the *lnt* locus (Table 3.1 and Fig. 3.1), when this mutant was grown on ribose (as compared to wild type UCC2003), all of which were also previously found to be upregulated in expression during growth of wild type UCC2003 on LNT or LNnT [27]. This corroborates the notion that LntR is a LacI-type repressor and that this protein negatively regulates the LNT/LNnT-dependent transcription of genes within the *lnt* cluster. Conversely, the array data obtained for the *nahR* mutant only revealed transcriptional upregulation (compared to the UCC2003 control) of the *nahS* gene (Table 3.1), when grown on ribose. This is consistent with previously observed expression patterns in UCC2003, with the exception of *nahA*, which may have been expected to exhibit transcriptional upregulation in the *nahR* mutant, as its expression was increased during growth on LNT and LNnT in *B. breve* UCC2003 [27]. These results suggest that NahR, a NagC/XylR-type repressor, is responsible for the transcriptional regulation of at least one gene of the *nah* cluster. For the *nagR1* mutant, upregulation of *nagA2*, *nagB3* and *nagK* (but not *nagR2* or *nagK2*), as well as all of the genes of the *lnp/glt* locus (Table 3.1) was observed when grown on ribose (as compared to the UCC2003 control). These results suggest that NagR1, a ROK/NagC family-type repressor, is responsible for the transcriptional regulation of (part of) the *nag* and *lnp-glt* clusters. This is consistent with transcriptomic data previously obtained for wild type UCC2003 during growth on LN(n)T and LNB, which demonstrated the transcriptional upregulation of genes in both of these loci [27]. When the transcriptome of UCC2003-nagR2 was compared to that of UCC2003 when grown on ribose, the *nagR2* mutant exhibited increased transcription of genes in the *mal* locus (locus tags Bbr_0118-0123), which is known to be involved in maltooligosaccharide metabolism [63, 64], and Bbr_1719-1721 (predicted to function in fatty acid metabolism) [65] (data not shown), none of which are predicted to function in HMO metabolism, nor were shown to be upregulated in our previous wild type arrays on LNT, LNnT or LNB [27]. These results

thus show that NagR2 is not involved in the transcriptional control of the loci responsible for LNT, LNnT or LNB metabolism, and no further investigation of this regulator was carried out. The *lntR*, *nahR* and *nagR1* genes, however, were selected for further study, as described below, in order to further elucidate their regulatory activity and specificity.

Promoter mapping through identification of the transcription start sites. Based on the transcriptome findings, we presumed that LntR, NahR and NagR1 act as transcriptional regulators of (certain genes of the) *lnt*, *nah* and *nag/lnp/glt* loci, respectively. Gene expression patterns observed for the regulator gene mutants, and examination of the genetic layout and transcriptome profiles of these loci allowed us to assign putative promoter-containing regions within each locus. In order to verify these predicted promoter regions, the associated transcription start sites (TSS) were experimentally determined by primer extension analyses.

The *lnt* locus was deduced to contain at least two promoters: one just upstream of *lntP1* (Fig. 3.2A) and one in front of *lntS* (Fig. 3.2B). The *lntP1* and *lntS* genes on the *B. breve* UCC2003 genome encode a permease and a solute-binding protein of an ABC-transporter system, respectively, and exhibit an increase in transcription upon growth on LNT, LNnT, LNB, lactosamine or lactose [27]. The transcription start sites (TSS) of the presumed *lntP1* and *lntS* promoters were determined by primer extension analysis using RNA extracted from *B. breve* UCC2003 grown in mMRS supplemented with 1 % LNnT. An extension product was identified 41 nucleotides 5' of the predicted translational start site of the *lntP1* gene (Supplemental Fig. S3.1A), while the TSS for the *lntS* gene was identified 154 nucleotides 5' of the predicted translational start site (Supplemental Fig. S3.1B). In both cases, the TSS was preceded by -10 and -35 hexamers that resemble (bifidobacterial) consensus vegetative promoter recognition sequences [66, 67].

The *nah* locus was deduced to contain at least two promoters: one just upstream of *nahS* (Fig. 3.2C) and one in front of *nahA* (Fig. 3.2D). The *nahS* and *nahA* genes on the *B. breve* UCC2003 genome encode a solute-binding

protein of an ABC-transporter system, and a GH20 *N*-acetylhexosaminidase, respectively. While an increase in transcription was only observed for *nahS* in the *nahR* mutant-based array, both this gene and *nahA* were found to be subject to transcriptional induction when wild-type UCC2003 is grown on LNT, LNnT or lactosamine [27]. The TSSs of the presumed *nahS* and *nahA* promoters were determined by primer extension analysis using RNA extracted from *B. breve* UCC2003 grown in mMRS supplemented with 1 % LNnT. An extension product was identified 59 nucleotides 5' of the predicted translational start site for the *nahS* gene (Supplemental Fig. S3.1C), while the TSS upstream of *nahA* was identified 74 nucleotides 5' of the predicted *nahA* translational start site (Supplemental Fig. S3.1D). The *nahS* upstream region contained a -10 and a -35 hexamer just upstream of the TSS resembling bifidobacterial promoter sequences [66, 67], while in the case of the *nahA* promoter region the TSS is preceded by a sequence that resembles a canonical -10 promoter sequence, although no associated -35 hexamer could be identified.

The *nag* and *lnp/glt* loci were each deduced to contain at least two promoters, just upstream of the genes: *nagB3* (Fig. 3.2E) and *nagK* (Fig. 3.2F), and *lnpB* (Fig. 3.2G) and *gltA* (Fig. 3.2H), respectively, based on the associated genetic lay-out coupled to transcription patterns of the *nagR1* mutant or when UCC2003 was grown on LNB [27]. The transcription start sites (TSS) of the presumed *nagB3*, *nagK*, *lnpB* and *gltA* promoters were determined by primer extension analysis using RNA extracted from *B. breve* UCC2003 grown in mMRS supplemented with 1 % LNB. An extension product was identified 155 nucleotides 5' of the predicted translational start site for the *nagB3* gene (Supplemental Fig. S3.1E), while the transcriptional start site of *nagK* was identified 35 nucleotides 5' of the predicted translational start site (Supplemental Fig. S3.1F). An extension product was identified 43 nucleotides 5' of the predicted translational start site for the *lnpB* gene (Supplemental Fig. S3.1G), while the transcription start site for the *gltA* gene was identified 44 nucleotides 5' of the predicted translational start site (Supplemental Fig. S3.1H). All four regions contained -10 and -35 hexamers

just upstream of the TSS that resembled bifidobacterial vegetative promoter recognition sequences.

Identification of regulator-operator interactions using electromobility shift assays and *in silico* analysis. In order to establish if the LntR, NahR and NagR1 proteins directly and specifically interact with operator sequences within the identified promoter regions of the *lnt*, *nah*, and *nag*/*lnp*/*glt* gene clusters, respectively, electrophoretic mobility shift assays (EMSAs) were performed. For the purpose of performing EMSAs, the *lntR*, *nahR* and *nagR1* genes were first individually cloned into the nisin-inducible vector pNZ8150 with an N-terminal His tag-encoding sequence to facilitate protein expression and purification in *L. lactis* NZ9000 (see Materials and Methods). As had been noted previously for other regulators from bifidobacteria [52, 55, 68], LntR, NahR and NagR1 could be obtained as purified proteins, but had lost their DNA binding activity during some stage of the purification process. Thus, instead of purified protein, crude cell extracts of (nisin-induced) *L. lactis* NZ9000 pNZ-*lntR*$_{His,}$ *L. lactis* NZ9000 pNZ-*nahR*$_{His}$ and *L. lactis* NZ9000 pNZ-*nagR1*$_{His}$ were used to carry out the EMSAs. Crude cell extract obtained from nisin-induced *L. lactis* NZ9000 pNZ8150 (empty vector) incubated with the respective DNA fragments was used as a negative control. The DNA fragments used were various short amplicons representing different segments of the putative promoter regions (Fig. 3.2, Supplemental Table S3.3).

LntR-containing crude extract was shown to specifically bind to the IRD700-labelled DNA fragments lntP1a and lntP1b, but not with lntP1c (Fig. 3.2A, Supplemental Table S3.3). A double mobility-shift was observed for fragment lntP1a, indicative of two distinct LntR binding sites being present on this fragment, while a single mobility shift was visible for fragment lntP1b. Similarly, LntR was able to bind to IRD700-labelled DNA fragments lntSb and lntSa, in the latter case being visible as a double mobility-shift (suggesting the presence of two distinct LntR-binding sites), while no binding was observed with lntSc (Fig. 3.2B, Supplemental Table S3.3). Inspection and comparison of the four fragments in which binding was observed

revealed the presence of at least one complete conserved sequence, representing an inverted repeat, in all four fragments, while two such conserved sequences were observed in fragments lntP1a and lntSa (being consistent with the observed double mobility-shift). Comparative analysis of these inverted repeats identified a 14-nucleotide consensus sequence (Fig. 3.3A), containing a conserved 'CG' at its centre, which is a well-documented conserved feature of operator sequences bound by LacI-type regulators [69, 70]. This consensus sequence furthermore contains a conserved 5' 'TG' and 3' 'CA' at its flanking ends, a feature previously documented in operator sequences identified for other LacI-type regulators encoded by *B. breve* UCC2003 [43, 51, 55]. In both promoter regions, one such presumed operator sequence was found closely downstream of or partially overlapping the predicted -10 element of the promoter region, while the second was found closely upstream of the predicted -35 element of the promoter region (Supplemental Fig. S3.1A, S3.1B). The positions of these identified operators are consistent with LntR acting as a repressor for the identified *lnt* promoters [71, 72].

The results obtained with the *L. lactis* NZ9000 pNZ-*nahR*$_{His}$ crude extract demonstrated specific binding to the IRD700-labelled DNA fragments nahSa and nahSb, but not with nahSc (Fig. 3.2C, Supplemental Table S3.3). Furthermore, binding was observed for IRD700-labelled DNA fragment nahAa, but not to fragment nahAb or nahAc (Fig. 3.2D, Supplemental Table S3.3). Sequence inspection and comparison of the NahR-bound DNA fragments revealed the presence of an inverted repeat sequence, which was common to these fragments, yet not present in fragments to which NahR did not bind. These inverted repeat elements therefore represent putative operator sequences required for the NahR protein. Further analysis of these inverted repeats identified a 10-nucleotide consensus sequence (Fig. 3.3B). Conserved 5' C and 3' G nucleotides at the extreme flanks of this consensus sequence have previously been observed for operator sequences of certain NagC/XylR-type regulators [73]. The presumed operator upstream of *nahS* overlaps with the downstream end of the predicted -10 promoter element (Supplemental Fig. S3.1C), while the *nahA*-associated operator was found to be roughly 110

bp upstream of the predicted -10 element (Supplemental Fig. S3.1D). The position of the identified *nahS* operator, and the consensus obtained between this and the putative operator identified for *nahA*, confirm the function of NahR as a repressor for *nahS*. While binding of NahR may occur at the *nahA* operator, its binding does not appear to directly interfere with the *nahA* promoter. This agrees with the lack of upregulation in *nahA* expression observed for the *nahR* mutant, though the transcriptional role of NahR in this case, if any, is not clear. The mapped locations of the operator, -10 and -35 sequences are shown in Supplemental Figure S1.

The results obtained with the crude extract obtained from nisin-induced *L. lactis* NZ900 pNZ-*nagR1*$_{\text{His}}$ revealed specific binding to the IRD700-labelled DNA fragments nagB3a, nagB3b and nagB3c (with a weak apparent double-shift observed for nagB3a), but not to fragment nagB3d (Fig. 3.2E, Supplemental Table S3.3). Specific binding was identified for IRD700-labelled DNA fragment nagKa, while no binding was detected when fragment nagKb was used (Fig. 3.2F, Supplemental Table S3.3). Binding of the NagR protein was also demonstrated for the IRD700-labelled DNA fragments lnpBa and lnpBc, but not with lnpBb (Fig. 3.2G, Supplemental Table S3.3). Finally, NagR1 was shown to bind IRD700-labelled DNA fragments gltAa, gltAb and gltAc (Fig. 3.2H, Supplemental Table S3.3). Inspection and comparison of the *nagB3*, *nagK*, *lnpB*, and gltA-associated fragments in which binding was observed revealed the presence of a common sequence, representing an inverted repeat (with two repeats present in the fragment nagB3a, consistent with the observed double-shift), while being absent within fragments for which no binding was observed. These sequence motifs are presumed to act as operator sequences for the NagR1 protein. *In silico* analysis of these inverted repeat sequences revealed a 23-nucleotide consensus (Fig. 3.3C). Interestingly, while this obtained consensus motif bears little resemblance to many previously proposed binding motifs for ROK/NagC family-type repressors from other bacteria [74], a substantial degree of similarity can be observed to motifs identified previously for other ROK/NagC-type regulators encoded by *B. breve* UCC2003 [30, 55]. The putative *nagB3*, *nagK*, *lnpB* and *gltA* operators were all found to be

overlapping with or encompassing the predicted -10 or -35 elements (Supplemental Fig. S3.1E, S3.1F, S3.1G, S3.1H). The positions of these identified operators corroborate the notion that NagR1 acts as a transcriptional repressor of its target genes (i.e. *nagB3*, *nagK*, *lnpB* and *gltA*).

Identification of transcriptional effectors. In order to identify effectors that control the binding activity of LntR, NahR and NagR1, we performed EMSAs with fragments containing the binding motifs for each regulator, in the presence of a range of carbohydrates, including: lactose, LNB, LNT, LNnT, galactose, galactose-6-phosphate (Gal-6-P), galactose-1-phosphate (Gal-1-P), GlcNAc, N-acetylglucosamine-6-phosphate (GlcNAc-6P), GalNAc or glucose (at a standard concentration of 20 mM) (Supplemental Fig. S3.2). These carbohydrates were chosen as they include both the complete structures and various components (or breakdown products) of LNT, LNnT or LNB. Carbohydrates which did not elicit any effect on fragment binding by the regulator (at a concentration of 20 mM) were assumed not to represent transcriptional effectors for that particular regulator. If an inhibition in binding was observed at 20 mM, the EMSA was repeated at a range of descending concentrations (or in some cases higher concentrations were used for a related molecule [e.g. Gal, Gal-1-P and Gal-6-P]). For LntR, galactose was found to reduce binding of this regulator to its DNA targets at a concentration of 10 mM or less (Fig. 3.4A). Gal-6-P and Gal-1-P were also found to reduce target DNA binding of LntR, but at considerably higher, and perhaps biologically irrelevant concentrations of ≥20 mM (Supplemental Figure S3.3). For NahR, only GlcNAc was found to reduce interaction between NahR and its DNA target at a minimum concentration of 0.0625 mM (Fig. 3.4B), while in the case of NagR1, GlcNAc-6-P was found to prevent NagR1-binding activity at a minimum concentration of 1 mM (Fig. 3.4C, 3.4D).

3.5 Discussion

The dominance of (certain) bifidobacteria within the breastfed neonatal gut microbiota [1] is substantially aided by the ability of these infant-associated species to utilise indigestible HMO residues as a carbon source [9]. Our previous work demonstrated that consumption and utilisation of LNT, LNnT or LNB by *B. breve* UCC2003 is facilitated by interrelated catabolic pathways [27]. While pathways for HMO utilisation in other *Bifidobacterium* species have been identified and elucidated [13, 21, 22], very little work has been carried out with regard to their regulation. Our results reveal molecular details of the transcriptional regulation of *B. breve* UCC2003 loci responsible for LN(n)T/LNB metabolism, and provide insights into how metabolism of these HMOs is controlled in *B. breve* UCC2003.

In the current study, we identified four transcriptional regulators, three of which were shown to be involved in regulating LN(n)T/LNB metabolism in UCC2003 (Fig 3.5). Microarray analysis of insertional mutants in *lntR*, *nahR*, *nagR1* and *nagR2* identified genes under the regulation of each encoded regulator. LntR and NahR were shown to represent 'local' regulators, i.e. controlling transcription of genes adjacent to *lntR* and *nahR*, respectively. In contrast, NagR1 regulates transcription of not only the 'local' *nag* locus, but also of the genetically unlinked *lnp/glt* locus. We also investigated the transcriptome effect of a mutation in Bbr_1251 (*nagR2*), however, the affected genes are not believed to be involved in HMO metabolism, but apparently in malto-oligosaccharide and fatty acid metabolism. While LacI-type, NagC/XylR-type and ROK/NagC-type regulators have all previously been identified and characterised in *B. breve* UCC2003 [30, 51, 52, 55], functional analysis of regulators in other bifidobacteria is comparatively undocumented. However, a recent study identified homologous transcription factors for those of LntR, NagR1and NagR2 in a range of different *Bifidobacterium* species [75].

Details of promoter and operator sequences specific to the LntR, NahR and NagR1 regulators were elucidated using a combination of electromobility shift and primer extension analyses. These operator results, for the most part, agree with those predicted by Khoroshkin et al. [75]. The operator sequences

predicted in their study concur with our experimentally determined data, both in approximate location and number, for both LntR and NagR1, with the exception of one additional predicted operator for LntR, and two for NagR1. An additional NagR1 operator sequence was predicted upstream of *gltA*, however, this did not appear to be functional, based on the lack of a double mobility-shift in the EMSA's for that region. This operator may indeed be a non-functional relic resulting from a duplication event. Khoroshkin et al. [75] also predicted an operator sequence upstream of the gene Bbr_1884 for NagR1 binding, though we did not examine this. However, based on the predicted functions of this gene in the Bifid Shunt, it may also be tied into the overall regulation of LNB and LacNAc metabolism carried out by NagR1. An additional LntR operator was predicted upstream of *lntR* itself, which may function in *lntR* transcriptional auto-regulation. The observed lack of upregulation in *nahA* transcription for the *nahR* mutant appears discordant with the transcriptional increase of this gene that was previously observed in wild type UCC2003 during growth on LN(n)T [27], as well as the presence of the functional *nahA* operator sequence for NahR binding identified in this study. However, this may be explained if transcriptional induction of *nahA* is mediated by both LntR and NahR. This possibility is corroborated by the presence of an inverted repeat sequence resembling an LntR operator and this intriguing possibility merits further experimental investigation.

Perhaps most interestingly of all was the identification of the effectors for each transcriptional regulator. Binding of LntR to its targets is impeded by Gal, NahR-mediated operator binding is prevented by the presence of GlcNAc, while NagR1-operator interaction is prevented by the presence of GlcNAc-6-P. In each case, the genes under transcriptional control by their respective regulator encode the metabolic machinery responsible for the release (and/or generation) of the effector monosaccharide from the substrate at that metabolic step. For example, Gal is released from the non-reducing end of LN(n)T through the hydrolytic activity of LntA, which is encoded by the *lnt* locus [27]. Transcriptional repression of this locus is thus believed to be relieved by the presence of the released monosaccharide, which is presumed to interact with the allosteric effector site typical of LacI-type

repressors [55, 71, 76, 77]. A similar scenario applies to GlcNAc release which acts as the effector for the NahR regulator that controls transcription of *nahS*, and to GlcNAc-6-P which governs the activity of the NagR1, the presumed transcriptional regulator of the *lnp/glt* and *nag* loci. The possible dual regulation of *nahA* transcription, as mentioned above, would mean that both the presence of the *lnt* locus activity product (and LntR effector) galactose, and the *nah* locus activity product (and NahR effector) GlcNAc, are required for the induction of *nahA* expression. This provides an extra level of transcriptional and thus metabolic control, ensuring the expression of *nahA* strictly during LN(n)T metabolism, despite GlcNAc release during metabolism of other sugars, such as LNB, sialic acid and sulphated GlcNAc [27, 29, 30]. Interestingly, in the case of GlcNAc-6-P and NagR1, the *lnp/glt* locus is required for the degradation of LNB, while the activity of the *nag* locus results in the generation of GlcNAc-6-P from liberated GlcNAc, during both HMO and sialic acid metabolism [29]. This may not be surprising, as sialic acid residues are commonly found in HMO [6], and more importantly, GlcNAc is a breakdown product of LNB (as well as LNT and LNnT).

Interestingly, previous work has shown that transcriptional induction takes place of the *lnt* locus during growth of UCC2003 on galacto-oligosaccharides (GOS) [78]. This would appear to disagree with the high degree of specificity of transcriptional induction by effectors of these HMO-associated loci. However, it is worth noting that GOS consist mainly of galactose [79, 80], and that the intracellular release of galactose during GOS metabolism by UCC2003 would be sufficient to cause transcriptional induction of the *lnt* locus.

Thus, the presence and initial degradation of such a structure (i.e. LNT, LNnT or LNB) indirectly induces further expression of the locus required for its degradation, until the sugar is no longer available, at which point the absence of inducers will cause a return to transcriptional repression. Initial internalisation and degradation is likely facilitated by a low level of ‘leaky’ gene expression of the locus. In the case of LNT and LNnT degradation, this regulation is a two-step process, at the level of LN(n)T degradation first (by the *lnt* locus), and then at the level of (LN(n)T breakdown product) lacto-N-

triose degradation (by the *nah* locus). Regulation of LNB metabolism is managed in a single step, at the level of LNB phosphorolysis and GlcNAc phosphorylation (by the *lnp/glt* and *nag* loci). We see that all three regulators in this transcriptional control network belong to distinct families of regulator proteins, despite functioning in similar roles as saccharide-controlled repressors. In conclusion, our results reveal a tightly controlled system for transcriptional regulation of genes encoding the metabolic machinery required for (certain) HMO metabolism in *B. breve* UCC2003. Such tight regulation is necessary for infant-associated bifidobacteria such as *B. breve*, where switching metabolic processing to and from milk-derived sugars such as HMO and lactose, and plant-derived carbohydrate sources [81] is a regular occurrence during the weaning period. Moreover, this suggests the evolution of specific catabolic responses to the presence of and for the utilisation of specific HMO moieties by *B. breve*, and poses the question as to whether such regulatory systems have similarly evolved in other infant-associated *Bifidobacterium* species.

3.6 Acknowledgements

The authors would like to sincerely thank Glycom A/S (Lyngby, Denmark) for the provision of purified HMO samples used in this study under their donation program. We thank Dr. Takane Katayama for indicating an error within the manuscript prior to its official publication.This study was funded in part by the Irish Research Council, under the Postgraduate Research Project Award; Project ID GOIPG/2013/651. In addition, the authors are supported by Science Foundation Ireland (SFI) (Grant No. SFI/12/RC/2273); M.O.C.M. is a recipient of a HRB postdoctoral fellowship (Grant No. PDTM/20011/9).

3.7 Tables and Figures

Table 3.1. *B. breve* UCC2003 regulator mutant genes upregulated in transcription during growth in mMRS medium supplemented with 1 % ribose as the sole carbohydrate, as compared to the wild type (control).

Gene ID	Gene name	Function	Fold upregulation[a,b] during growth on:		
			UCC2003-lntR	UCC2003-nahR	UCC2003-nagR1
Bbr_0526	lntR	Transcriptional regulator, LacI family	N/A	-	-
Bbr_0527	lntP1	Permease protein of ABC transporter system for sugars	3.84	-	-
Bbr_0528	lntP2	Permease protein of ABC transporter system for sugars	3.77	-	-
Bbr_0529	lntA	GH42 Beta-galactosidase	2.78	-	-
Bbr_0530	lntS	Solute-binding protein of ABC transporter system for sugars	5.66	-	-
Bbr_1247	nagA2	CE9 nagA2 N-acetylglucosamine-6-phosphate deacetylase	-	-	6.70
Bbr_1248	nagB3	nagB3 Glucosamine-6-phosphate isomerase	-	-	9.11
Bbr_1249	nagR1	Transcriptional regulator, ROK family	-	-	N/A
Bbr_1250	nagK	Sugar kinase, ROK family	-	-	2.29
Bbr_1251	nagR2	Transcriptional regulator, ROK family	-	-	-
Bbr_1252	nagK2	Sugar kinase, pfkB family	-	-	-
Bbr_1554	nahS	Solute-binding protein of ABC transporter system (lactose)	-	17.44	-
Bbr_1555	nahR	NagC/XylR-type transciptional regulator	-	N/A	-
Bbr_1556	nahA	GH20 nagZ Beta-N-acetylhexosaminidase	-	-	-
Bbr_1558	nahP	Permease protein of ABC transporter system	-	-	-
Bbr_1559	nahT1	ATP-binding protein of ABC transporter system	-	-	-
Bbr_1560	nahT2	ATP-binding protein of ABC transporter system	-	-	-
Bbr_1585	lnpD	UDP-glucose 4-epimerase	-	-	3.06
Bbr_1586	lnpB	Phosphotransferase family protein	-	-	3.36
Bbr_1587	lnpA	GH112 lacto-N-biose phorylase	-	-	2.90
Bbr_1588	gltC	Permease protein of ABC transporter system for sugars	-	-	2.91
Bbr_1589	gltB	Permease protein of ABC transporter system for sugars	-	-	3.02
Bbr_1590	gltA	Solute-binding protein of ABC transporter system for sugars	-	-	5.07

The level of transcription is shown as a fold-value of increase in transcription on each carbohydrate, as compared to a ribose control, with a cut-off of a minimum 2.0-fold increase in transcription.

a Based on comparative transcriptome analysis using B. breve UCC2003-lntR, B. breve UCC2003-nahR and B. breve UCC2003-nagR1 grown on 1% ribose, as compared to wild-type B. breve UCC2003 grown under the same conditions as a control.

b The cutoff point is 2.0-fold, with a P value of _0.001. —, value below the cutoff.

N/A indicates that the fold-value for increase in transcription for this gene is not included, as this is the gene in which the mutation was made, and thus does not accurately represent its natural transcription under these conditions.

The level of transcription is not given for the regulator-encoding genes containing the mutations in their respective arrays, as their transcription has been interrupted, and thus cannot be considered as reliable.

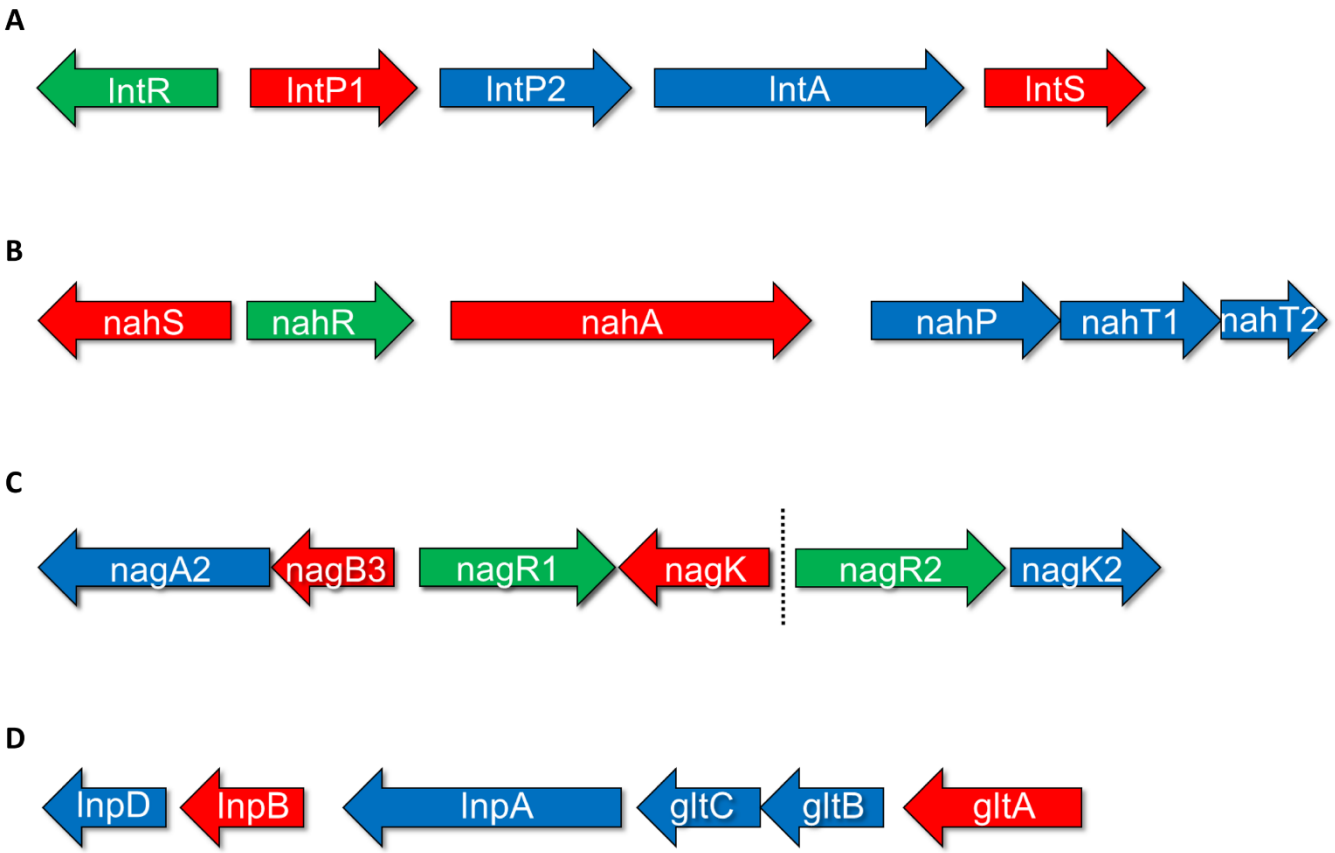


Figure 3.1. Schematic representation of HMO metabolism-associated loci in *B. breve* UCC2003, as identified previously [22]. (a) The genes of the *lnt* locus. (b) The genes of the *nah* locus. (c) The genes of the *nag* locus, and adjacent genes *nagR2* and *nagK2*. (d) The genes of the *lnp*/*glt* locus. The length of the arrows is proportional to the size of the open reading. Genes shown in red possess a predicted promoter in their upstream intergenic region. Genes shown in green are predicted to encode a regulator protein. Genes shown in blue were identified as not possessing a predicted promoter in their upstream intergenic region.

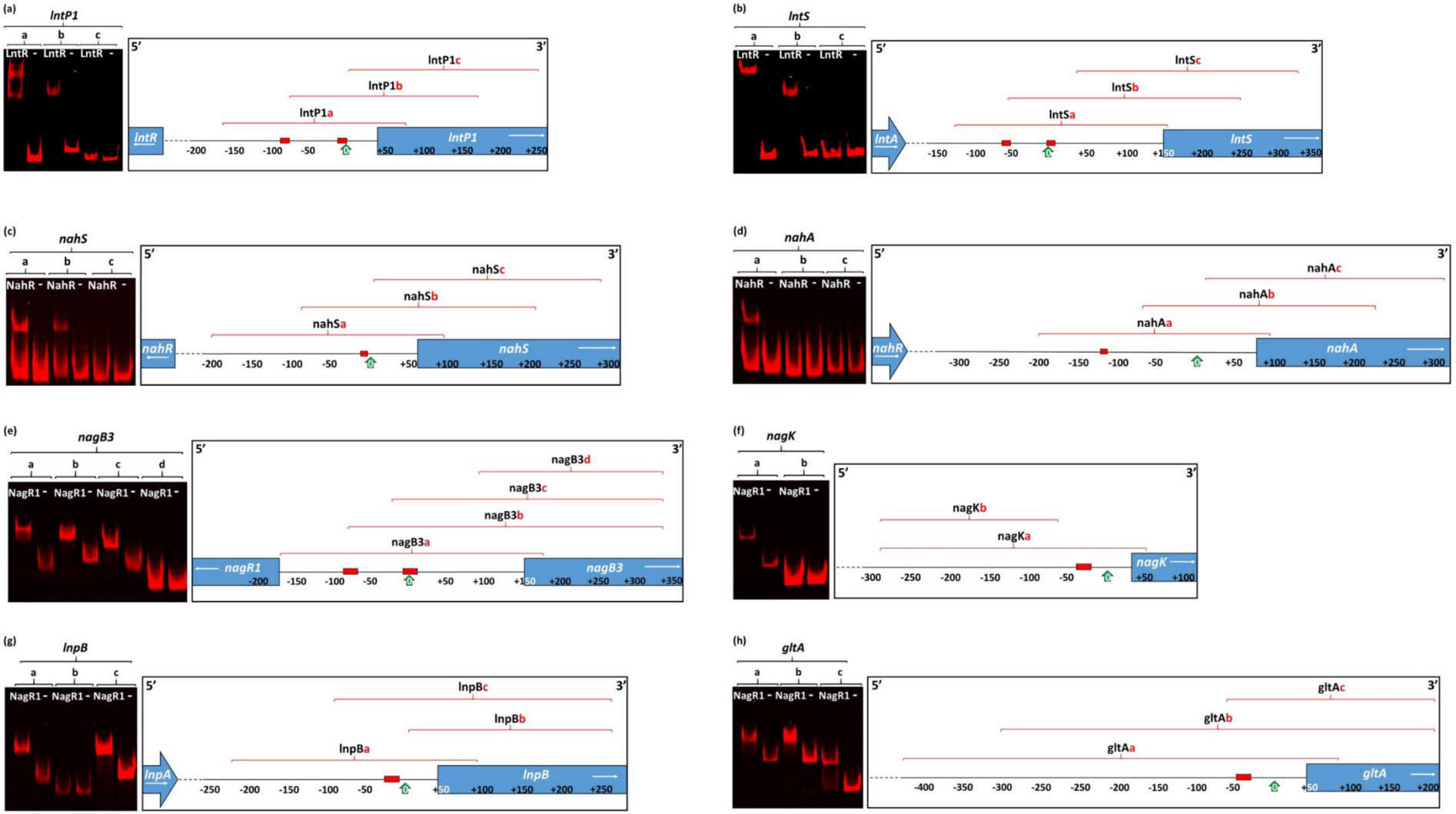

(a)
IntP1
a
b
c
LntR -
LntR -
LntR -
5′
3′
IntP1c
IntP1b
IntP1a
IntR
IntP1
-200 -150 -100 -50 +50 +100 +150 +200 +250
(b)
IntS
a
b
c
LntR -
LntR -
LntR -
5′
3′
IntSc
IntSb
IntSa
IntA
IntS
-150 -100 -50 +50 +100 +150 +200 +250 +300 +350
(c)
nahS
a
b
c
NahR -
NahR -
NahR -
5′
3′
nahSc
nahSb
nahSa
nahR
nahS
-200 -150 -100 -50 +50 +100 +150 +200 +250 +300
(d)
nahA
a
b
c
NahR -
NahR -
NahR -
5′
3′
nahAc
nahAb
nahAa
nahR
nahA
-300 -250 -200 -150 -100 -50 +50 +100 +150 +200 +250 +300
(e)
nagB3
a
b
c
d
NagR1 -
NagR1 -
NagR1 -
NagR1 -
5′
3′
nagB3d
nagB3c
nagB3b
nagB3a
nagR1
nagB3
-200 -150 -100 -50 +50 +100 +150 +200 +250 +300 +350
(f)
nagK
a
b
NagR1 -
NagR1 -
5′
3′
nagKb
nagKa
nagK
-300 -250 -200 -150 -100 -50 +50 +100
(g)
lnpB
a
b
c
NagR1 -
NagR1 -
NagR1 -
5′
3′
lnpBc
lnpBb
lnpBa
lnpA
lnpB
-250 -200 -150 -100 -50 +50 +100 +150 +200 +250
(h)
gltA
a
b
c
NagR1 -
NagR1 -
NagR1 -
5′
3′
gltAc
gltAb
gltAa
gltA
-400 -350 -300 -250 -250 -200 -150 -100 -50 +50 +100 +150 +200

Figure 3.2. EMSA images showing LntR (a and b), NahR (c and d), and NagR1 (e to h) interactions with a range of DNA fragments from the regions in the proximity of their predicted target promoters, in order to identify their approximate locations. The locations and sizes of fragments used, in relation to the promoter regions' respective transcription start sites, are given in Table S1 in the supplemental material. The panels at the right schematically represent the locations of the DNA fragments used in relation to the locations of the putative operator sequences (red boxes), transcription start sites (green arrows), and genes (arrows in blue boxes). In each panel, "-" indicates a negative control, where an equivalent amount of the crude cell extract from NZ9000 harboring empty plasmid pNZ8150 was added instead of the crude extract from the regulator-expressing NZ9000 strain.

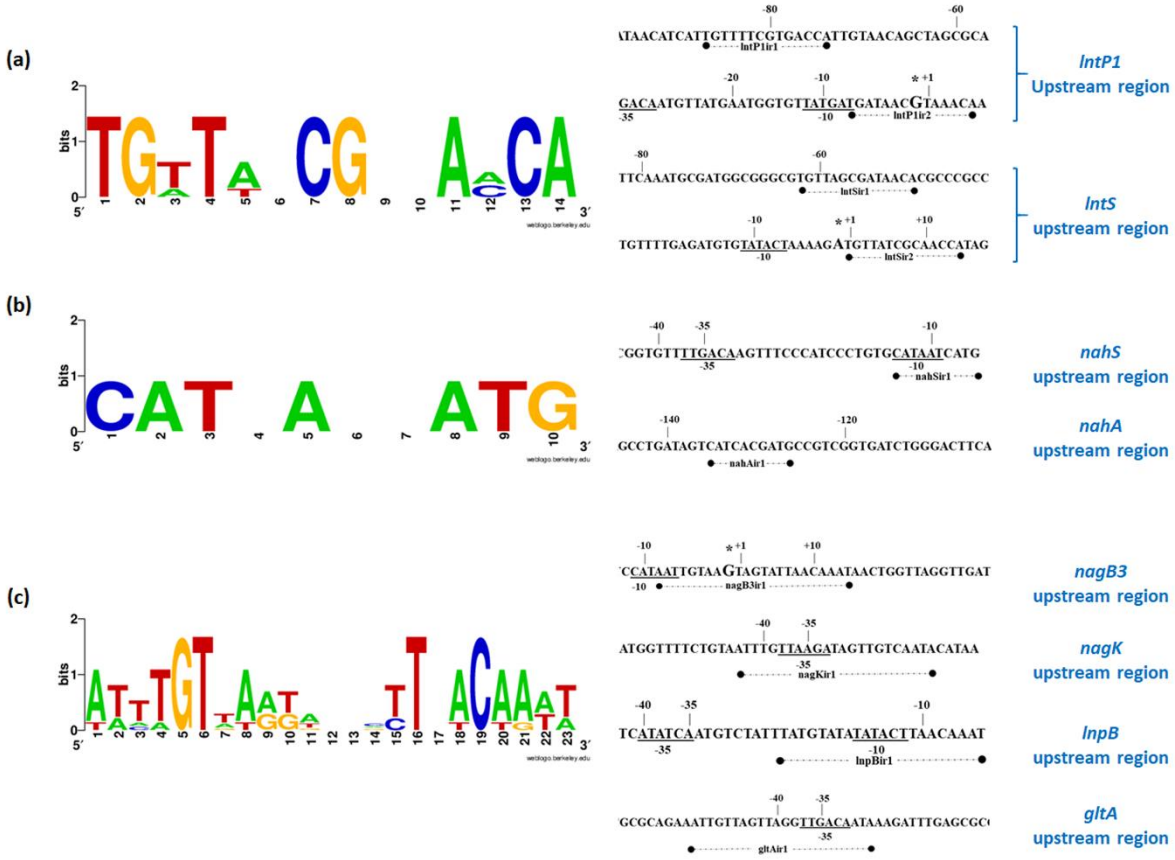


Figure 3.3. WebLogo representation of the operator motif consensus sequences for (a) the LacI-type regulator LntR, (b) the NahC/XylR-typr regulator NahR and (c) the ROK/NagC-type regulator NagR1, predicted using *in silico* analysis. Predicted operator sequences identified in the intergenic regions containing the co-regulated promoters for each regulator using the MEME online tool. Motif consensuses were generated by inputting these predicted operator sequences to the WebLogo online tool. The locations and sequences of each operator are shown alongside their respective consensus sequence.

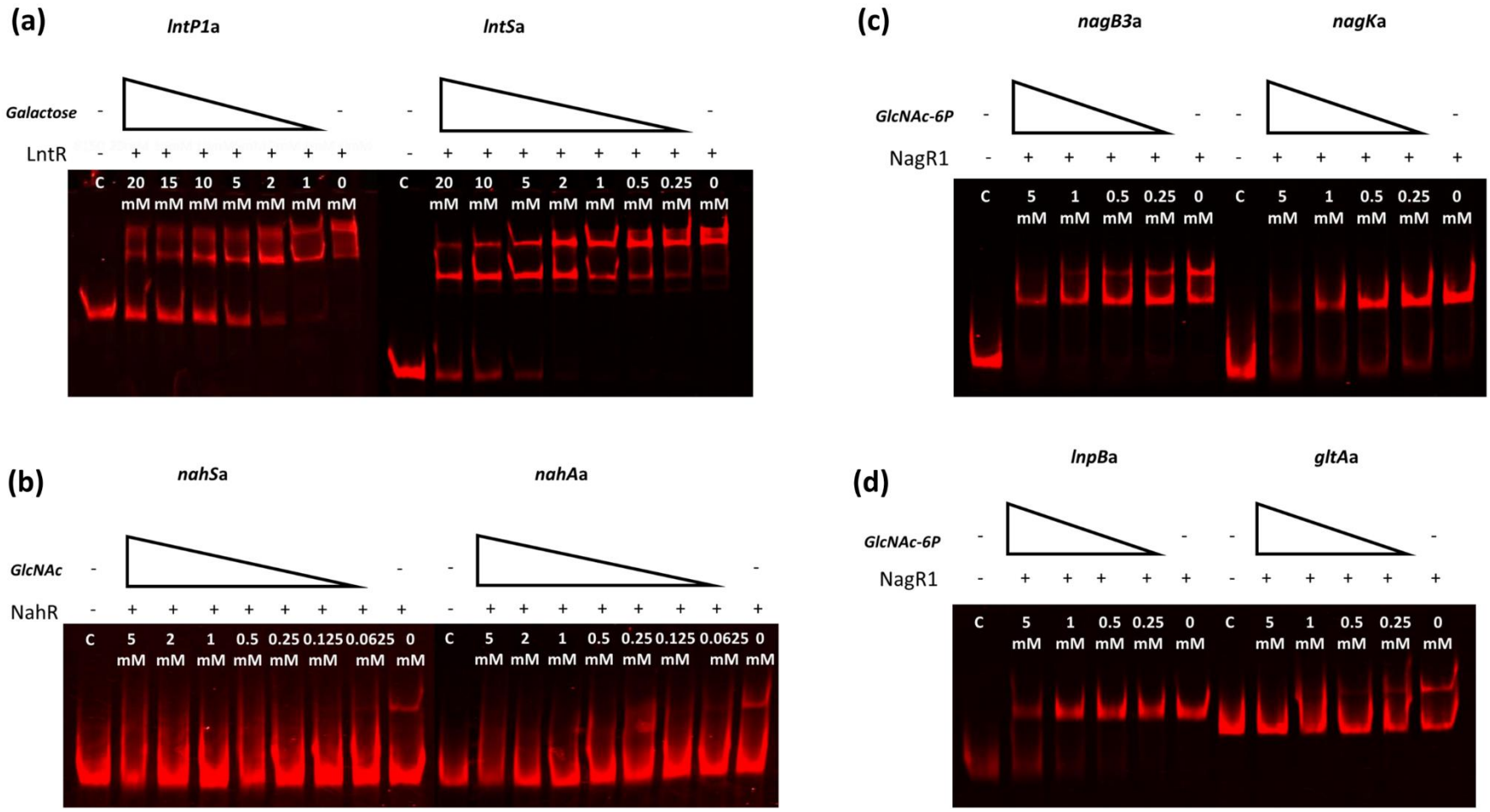


Figure 3.4. EMSA images showing (a) LntR, (b) NahR, and (c) and (d) NagR1 interactions with promoter-containing DNA fragments, with the addition of a gradient of their respective inducers, ranging from 0mM-20mM. In each panel, 'C' indicates a negative control, where an equivalent amount crude cell extract from NZ9000 harbouring empty pNZ8150 was added instead of crude extract from the regulator-expressing NZ9000 strain.

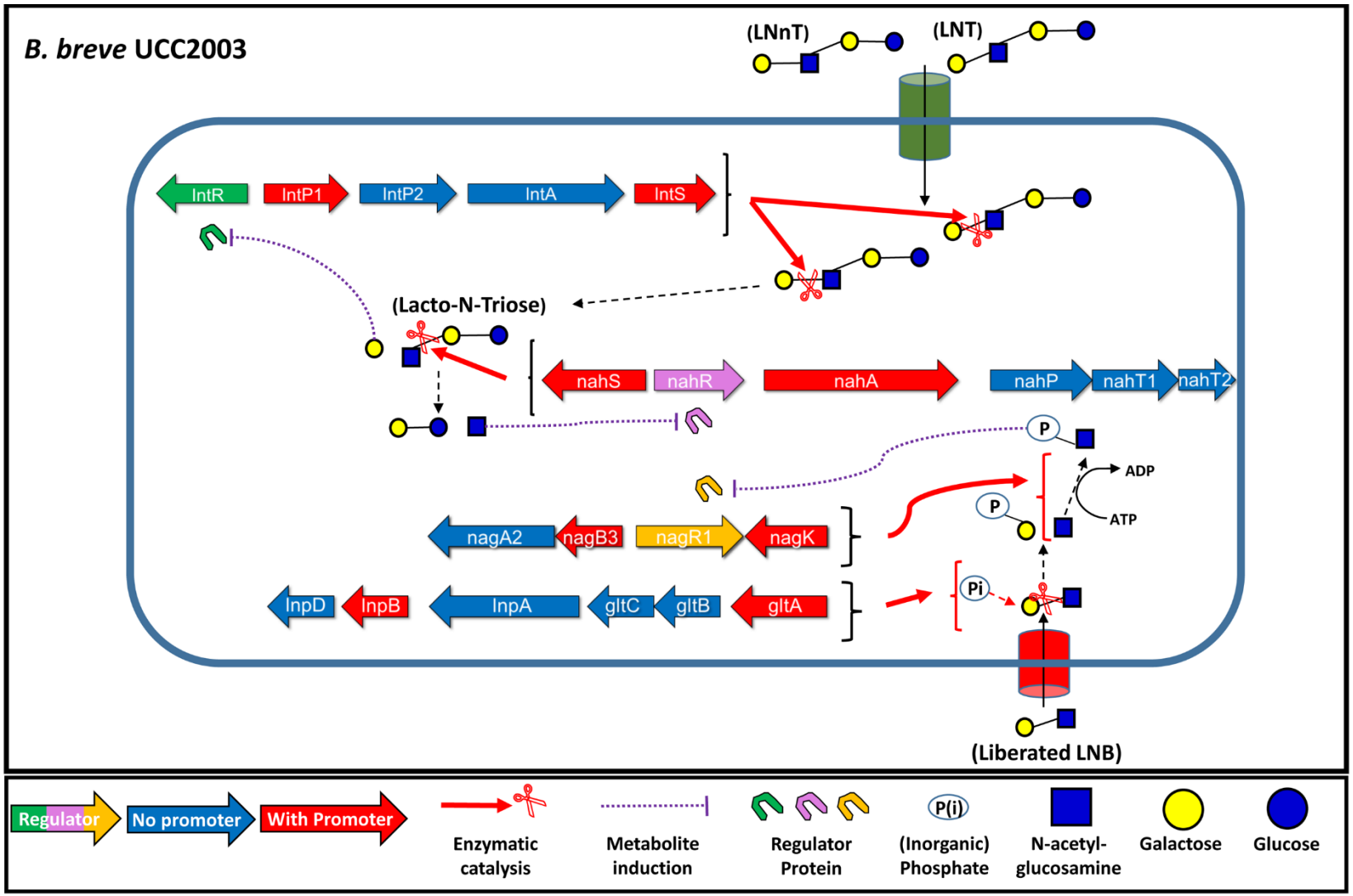

B. breve UCC2003
(LNnT)
(LNT)
IntR
IntP1
IntP2
IntA
IntS
(Lacto-N-Triose)
nahS
nahR
nahA
nahP
nahT1
nahT2
P
ADP
ATP
nagA2
nagB3
nagR1
nagK
lnpD
lnpB
lnpA
gltC
gltB
gltA
Pi
(Liberated LNB)
Regulator
No promoter
With Promoter
Enzymatic catalysis
Metabolite induction
Regulator Protein
P(i)
(Inorganic) Phosphate
N-acetyl-glucosamine
Galactose
Glucose

Figure 3.5. Schematic representation of the proposed model for transcriptomic regulation of LNT, LNnT and LNB metabolism by *B. breve* UCC2003. LNT and LNnT are internalised and intracellularly degraded through sequential hydrolysis and release of monosaccharides from their non-reducing ends. These released monosaccharides (or their metabolites) act as effectors relieving transcriptional repression of the loci encoding the cellular components responsible for liberation of these glycans. As such, liberated galactose relieves transcriptional repression of the *lnt* locus, and N-acetylglucosamine relieves transcriptional repression of the *nah* locus. Similarly, intracellular degradation of LNB (derived from the extracellular hydrolysis of complex HMO structures by other infant GIT microbes) releases N-acetylglucosamine and galactose-1-phosphate. GlcNAc is further converted into GlcNAc-6-P, which relieves transcriptional repression of the *lnp*/*glt* and *nag* loci.

3.8 Supplemental Material

Supplemental Table S3.1. Oligonucleotide primers used in this work.

Purpose	Primer	Sequence (5'-3')
Cloning of Bbr_0526 in pNZ8150	526F	tgcatc**cccggg**atg*catcaccatcaccatcaccatcaccatcac*gcgagaccaacacaggtttcc
	526R	tgcgca**tctaga**cgtttcccgtataccattaatcag
Cloning of Bbr_1555 in pNZ8150	1555F	tgcatc**gatatc**atg*catcaccatcaccatcaccatcaccatcac*tacgctaaatccaatccc
	1555R	tgcgca**tctaga**cggcggcacggtgatctg
Cloning of Bbr_1249 in pNZ8150	1249F	tgcatc**cagctg**atg*catcaccatcaccatcaccatcaccatcac*tcgtatcccggtcttgcc
	1249R	tgcatc**cagctg**atgtcgtatcccggtcttgcc
Cloning of internal 465bp fragment of Bbr_0526 in pORI19	IM526F	ctggtc**aagctt**cgttgaagccgcgatgga
	IM526R	ctggtc**tctaga**gtcaacggtggggcagtg
Cloning of internal 488bp fragment of Bbr_1555 in pORI19	IM1555F	ctggtc**aagctt**gctggccatcgatacggac
	IM1555R	ctggtc**tctaga**ctcgtcgttcagcagcac
Cloning of internal 443bp fragment of Bbr_1249 in pORI19	IM1249F	ctggtc**aagctt**cgaagaaggcctattgcg
	IM1249R	ctggtc**tctaga**cagcagaatcgccgaacc
Cloning of internal 488bp fragment of Bbr_1251 in pORI19	IM1251F	ctggtc**aagctt**gaagagaccggcgacctgg
	IM1251R	ctggtc**tctaga**gccattgtcgatgacgcc
Amplification of *tet*W	tetWFw	tcagctgtcgacatgctcatgtacggtaaggaagca
	tetWRv	gcgacggtcgaccataacttctgattgttgccg
Confirmation of site specific homologous recombination	526confirm1	gcgctagctgttacaatggtc
	526confirm2	gccatttccaacccctctc
	1555confirm1	tacgctaaatccaatccc
	1555confirm2	gacgcaagggccaacaaccgc
	1249confrim1	catacagccgccacggcac
	1249confrim2	tcgtatcccggtcttgcc
	1251confrim1	gcagacgatactgcacgcg
	1251confrim2	gtcaagcatctctaccac
Amplification of Bbr_0527 promoter fragments with IRD700-labelled oligonucleotides	527IRDfa	ctcgcccctcgcttgtctctc
	527IRDra	gcataggcacggcagcgac
	527IRDfb	attgttttcgtgaccattg
	527IRDrb	gaataatgaacacgaacacg
	527IRDfc	caattttggtcaaccttcg
	527IRDrc	cgcgcgtagttctcgac
Amplification of Bbr_0530 promoter fragments with IRD700-labelled oligonucleotides	530IRDfa	gccgaacggtgtgctggtgg
	530IRDra	cttcatcgttctgttctccttc
	530IRDfb	cgataacacgcccgccatc
	530IRDrb	gctggacttgccgctatc
	530IRDfc	cttcatagagccacttc
	530IRDrc	ctcgaagtccttggcaac
Amplification of Bbr_1554 promoter fragments with IRD700-labelled oligonucleotides	1554IRDfa	gtcgctgggattggatttagcg
	1554IRDra	gtggctatgactgcgcgc
	1554IRDfb	cggctttcaggataacaccca
	1554IRDrb	ggatttggcggcgcgatc
	1554IRDfc	ccaaacaaaatagttgctacggc
	1554IRDrc	gttgagtgcggtgtaggtctcc
Amplification of Bbr_1556 promoter fragments with IRD700-labelled oligonucleotides	1556IRDfa	ctggacggctgctcaaagc
	1556IRDra	gcagagatgtttgaccgttcat
	1556IRDfb	ggtgacgacgccactctgc
	1556IRDrb	gtggtttgcggttgccct
	1556IRDfc	gccatctcaggaccgaacg
	1556IRDrc	ggtcgtcaaggtgatgaatcc
Amplification of Bbr_1248 promoter fragments with IRD700-labelled oligonucleotides	1248IRDfa	cctcctgcctgaacgatg
	1248IRDra	gacaatgatgatttccggc
	1248IRDfb	gttagggaacttcactaatacattcc
	1248IRDrb	cctgggagatgtcgatcgactc
	1248IRDfc	gtccgtacgtccataattgtaagtag
	1248IRDrc	cctgggagatgtcgatcgactc
	1248IRDfd	gatgggcggctttgggcag
	1248IRDrd	cctgggagatgtcgatcgactc

Restriction sites incorporated into oligonucleotide primer sequences are indicated in bold, and His-tag sequences incorporated into nucleotide primer sequences are indicated in italics.

Supplemental Table S3.1. Oligonucleotide primers used in this work.

Purpose	Primer	Sequence (5'-3')
Amplification of Bbr_1250 promoter fragments with IRD700-labelled oligonucleotides	1250IRDfa	gatgccgttgtggtagagatg
	1250IRDra	ggtgttatcagtcattgcctatcc
	1250IRDfb	gcgtgtcgcgtatgaggc
	1250IRDrb	ggtgttatcagtcattgcctatcc
Amplification of Bbr_1586 promoter fragments with IRD700-labelled oligonucleotides	1586IRDfa	cggttcgtcgaaaatccaag
	1586IRDra	cagtgcgaagtgtgaggcg
	1586IRDfb	gccgcttattgcggctttatag
	1586IRDrb	ctttgagggcagaagtaactagttc
	1586IRDfc	gtatgcgcgttcgtccac
	1586IRDrc	ctttgagggcagaagtaactagttc
Amplification of Bbr_1590 promoter fragments with IRD700-labelled oligonucleotides	1590IRDfa	ggcccgctggcagattag
	1590IRDra	ggcaagagcagccacgatg
	1590IRDfb	gacagatgtctgagcggtc
	1590IRDrb	gaatcgggcagacggtgc
	1590IRDfc	cgcgcagaaattgttagttagg
	1590IRDrc	gaatcgggcagacggtgc
Amplification of region containing Bbr_0527 promoter region for sequencing ladders	527promF	gcattgctgtcattcgccacac
	527promR	gaataatgaacacgaacacg
Amplification of region containing Bbr_0530 promoter region for sequencing ladders	530promF	gcgtgcggatgaaactgg
	530promR	gtctggaacggcttggcgc
Amplification of region containing Bbr_1554 promoter region for sequencing ladders	1554promF	cgtttcctcgaccccagttc
	1554promR	gaatgtgtccttgagcttggc
Amplification of region containing Bbr_1556 promoter region for sequencing ladders	1556promF	ctggacggctgctcaaagc
	1556promR	ggtcgtcaaggtgatgaatcc
Amplification of region containing Bbr_1248 promoter region for sequencing ladders	1248promF	ggaggctttggcggtacgg
	1248promR	cctgggagatgtcgatcgactc
Amplification of region containing Bbr_1250 promoter region for sequencing ladders	1250promF	gatgccgttgtggtagagatg
	1250promR	ggtgccacccacatcaacac
Amplification of region containing Bbr_1586 promoter region for sequencing ladders	1586promF	gcgagaccttcgaccttcagcc
	1586promR	cggcacgagattgtaagacac
Amplification of region containing Bbr_1590 promoter region for sequencing ladders	1590promF	ggcccgctggcagattag
	1590promR	gaatcgggcagacggtgc
527 promoter for primer extension analysis	527PE	gcataggcacggcagcgac
530 promoter for primer extension analysis	530PE	cttcatcgttctgttctccttc
1554 promoter for primer extension analysis	1554PE	gttcatgttggtcttctttcc
1556 promoter for primer extension analysis	1556PE	gcagagatgtttgaccgttcat
1248 promoter for primer extension analysis	1248PE	ctgcccaaagccgcccatc
1250 promoter for primer extension analysis	1250PE	ggtgttatcagtcattgcctatcc
1586 promoter for primer extension analysis	1586PE	gcgatgtcaaatagtgtttcc
1590 promoter for primer extension analysis	1590PE	ggcaagagcagccacgatg

Restriction sites incorporated into oligonucleotide primer sequences are indicated in bold, and His-tag sequences incorporated into nucleotide primer sequences are indicated in italics.

Supplemental Table S3.2. Bacterial plasmids and strains used in this work.

Strain or plasmid	Relevant Features (antibiotic resistances are given in brackets)	Reference or Source
Strains		
Escherichia coli strains		
E. coli EC101	Cloning host, repA$^+$ (km^r)	[42]
E. coli EC101-pNZ-M.BbrII + M.BbrIII	EC101 harbouring pNZ8048 derivative containing *bbrIIM* and *bbrIIIM*	[40]
Lactococcus lactis strains		
L. lactis NZ9000	MG1363, pepN::nisRK, nisin-inducible overexpression host	[54]
L. lactis NZ9700	Nisin-producing strain	[54]
L. lactis NZ9000-pNZ-lntR	(Cm^r), NZ9000 conatining pNZ-lntR	This study
L. lactis NZ9000-pNZ-nahR	(Cm^r), NZ9000 conatining pNZ-nahR	This study
L. lactis NZ9000-pNZ-nagR1	(Cm^r), NZ9000 conatining pNZ-nagR1	This study
Bifidobacterium sp. Strains		
B. breve UCC2003	Isolate from nursling stool	[41]
B. breve UCC2003-lntR	(Tet^r), pORI19-tet-bbr_0526 insertion mutant of UCC2003	This study
B. breve UCC2003-nahR	(Tet^r), pORI19-tet-bbr_1555 insertion mutant of UCC2003	This study
B. breve UCC2003-nagR1	(Tet^r), pORI19-tet-bbr_1249 insertion mutant of UCC2003	This study
B. breve UCC2003-nagR2	(Tet^r), pORI19-tet-bbr_1251 insertion mutant of UCC2003	This study
Plasmids		
pAM5	pBC1-puC19-(Tet^r)	[82]
pORI19	(Tet^r), (Em^r), repA$^-$, ori$^+$, cloning vector	[42]
pORI19-tet-lntR	(Tet^r), (Em^r), Internal 367 bp fragment of *bbr_0526* and tetW cloned in pORI19	This study
pORI19-tet-nahR	(Tet^r), (Em^r), Internal 448 bp fragment of *bbr_1554* and tetW cloned in pORI19	This study
pORI19-tet-nagR1	(Tet^r), (Em^r), Internal 502 bp fragment of *bbr_1249* and tetW cloned in pORI19	This study
pORI19-tet-nagR2	(Tet^r), (Em^r), Internal 507 bp fragment of *bbr_1251* and tetW cloned in pORI19	This study
pNZ8150	(Cm^r, nisin inducible translational fusion vector	[50]
pNZ-lntR	(Cm^r), pNZ8150 derivative containing translational fusion of Bbr_0526 encoding DNA fragment to nisin inducible promoter	This study
pNZ-nahR	(Cm^r), pNZ8150 derivative containing translational fusion of Bbr_1555 encoding DNA fragment to nisin inducible promoter	This study
pNZ-nagR1	(Cm^r), pNZ8150 derivative containing translational fusion of Bbr_1249 encoding DNA fragment to nisin inducible promoter	This study

Cm^r, Em^r,Km^r and Tet^r, resistance to chloramphenicol, erythromycin, kanamycin and tetracycline, respectively.

UCC, University College Cork Culture Collection.

Supplemental Table S3.3. Locations and sizes of fragments used in electromobility shift assays, in relation to the promoter regions' respective transcription start sites.

Fragment Name	Fragment Start (Position)	Fragment End (Position)	Fragment Size (bp)	Number of identified inverted repeats	Shift Observed (+/-)
lntP1a	-174	+74	249	2	++
lntP1b	-89	+173	263	2*	+
lntP1c	+5	+258	254	0	-
lntSa	-132	+159	292	2	++
LntSb	-56	+252	309	2*	+
lntSc	+30	+337	308	0	-
nahSa	-197	+90	288	1	+
nahSb	-91	+211	303	1	+
nahSFc	+4	+289	286	0	-
nahAa	-206	+95	302	1	+
nahAb	-70	+226	297	0	-
nahAc	+11	+310	300	0	-
nagB3a	-176	+175	352	2	++
nagB3b	-84	+338	423	2*	+
nagB3c	-22	+338	361	1	+
nagB3d	+94	+338	245	0	-
nagKa	-283	+49	333	1	+
nagKb	-283	-59	225	0	-
lnpBa	-227	+93	321	1	+
lnpBb	+4	+262	259	0	-
lnpBc	-91	+262	354	1	+
gltAa	-419	+83	503	1	+
gltAb	-304	+205	510	1	+
gltAc	-59	+205	265	1	+

Binding of the amplified DNA fragment by the regulator protein, and thus formation of a DNA-protein complex is observed as a shift in band position in the EMSA gel.

A positive result for a band shift is represented by a '+'. A double-shift is represented by a '++'. No observed shift is represented by a '-'.

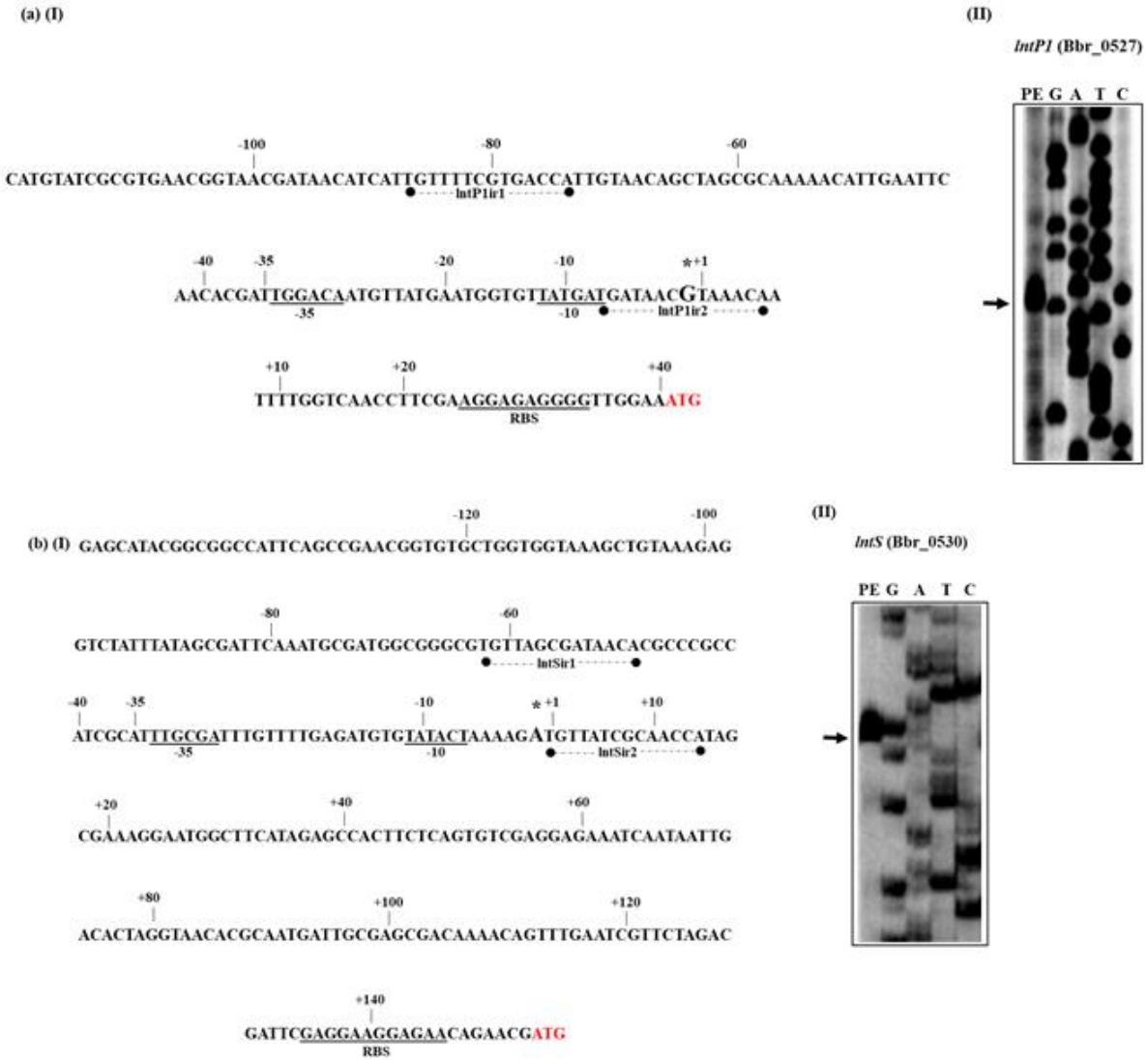
Fragment positions are given in relation to the deduced transcription start site (0).

'*' indicates that one of the identified inverted repeat sequences is partially interrupted by or in close proximity to the terminus of the fragment.

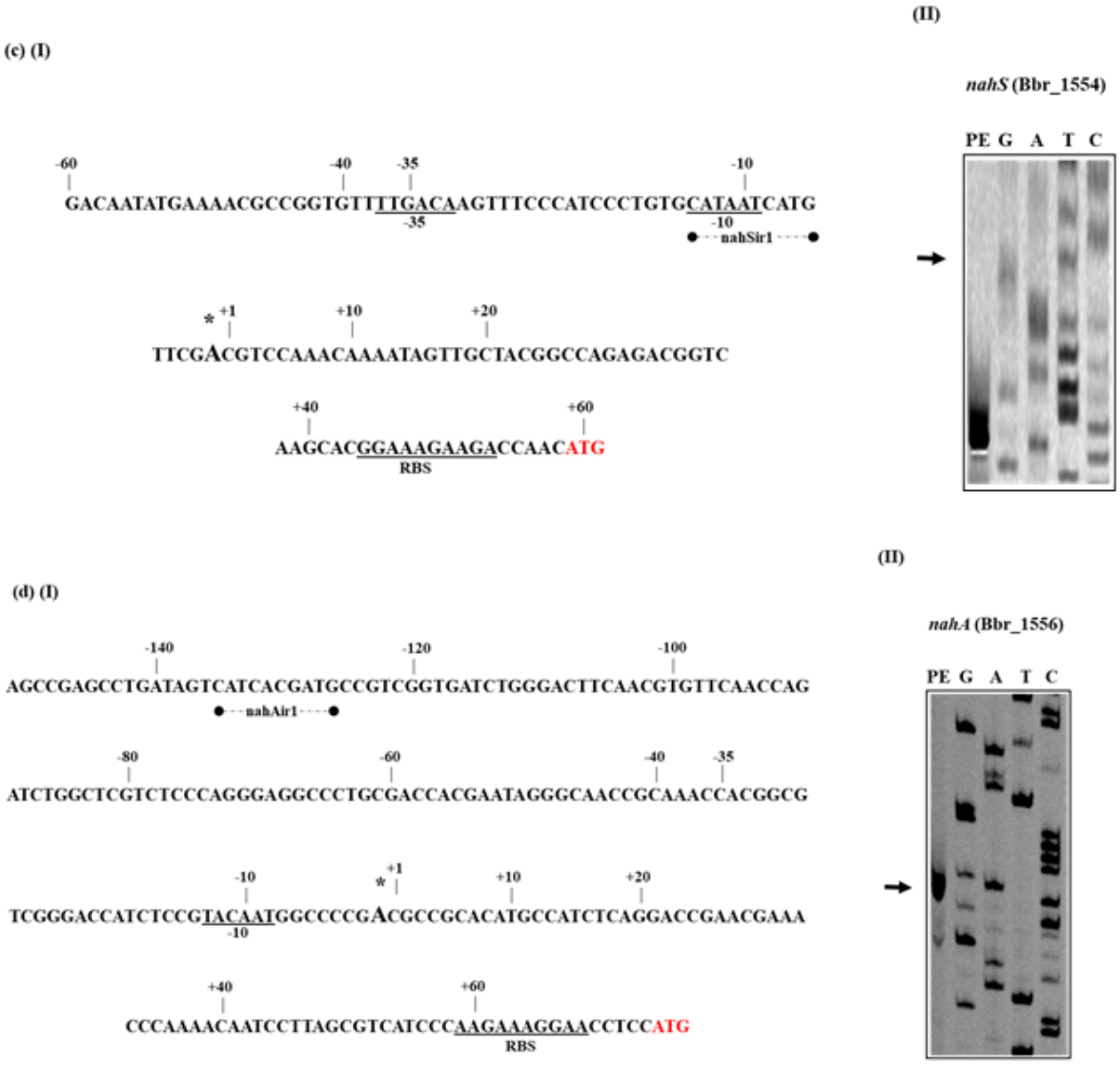
Supplemental Table S3.4. Predicted operator sequence used for the *in silico* generation of operator sequence motif consensuses.

Regulator	Operator Name	Sequence
LntR	lntP1ir1	tgttttcgtgacca
	lntP1ir2	tgataacgtaaaca
	lntS1ir1	tgttagcgataaca
	lntS1ir2	tgttatcgcaacca
NahR	nahSir1	cataatcatg
	nahAir1	catcacgatg
NagR1	nagB3ir1	aattgtaagtagtattaacaaat
	nagB3ir2	atctgttagggaacttcactaat
	nagKir1	atttgttaagatagttgtcaata
	lnpBir1	ttatgtatatatacttaacaaat
	gltAir1	aattgttagttaggttgacaata
	gltAir2[A]	ctttgtaaggagtattaataaaa

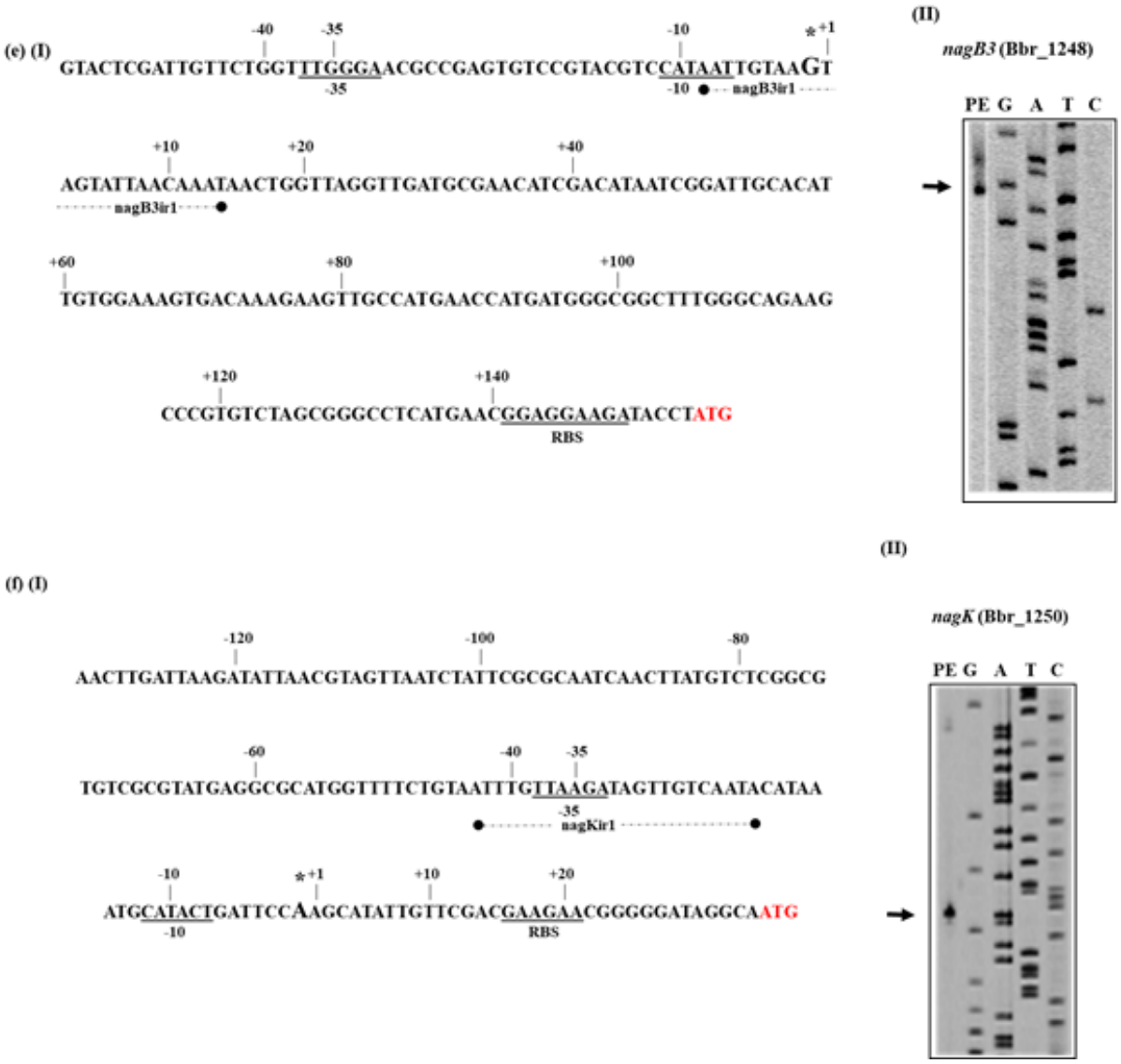
[A] Not identified by EMSA analysis.



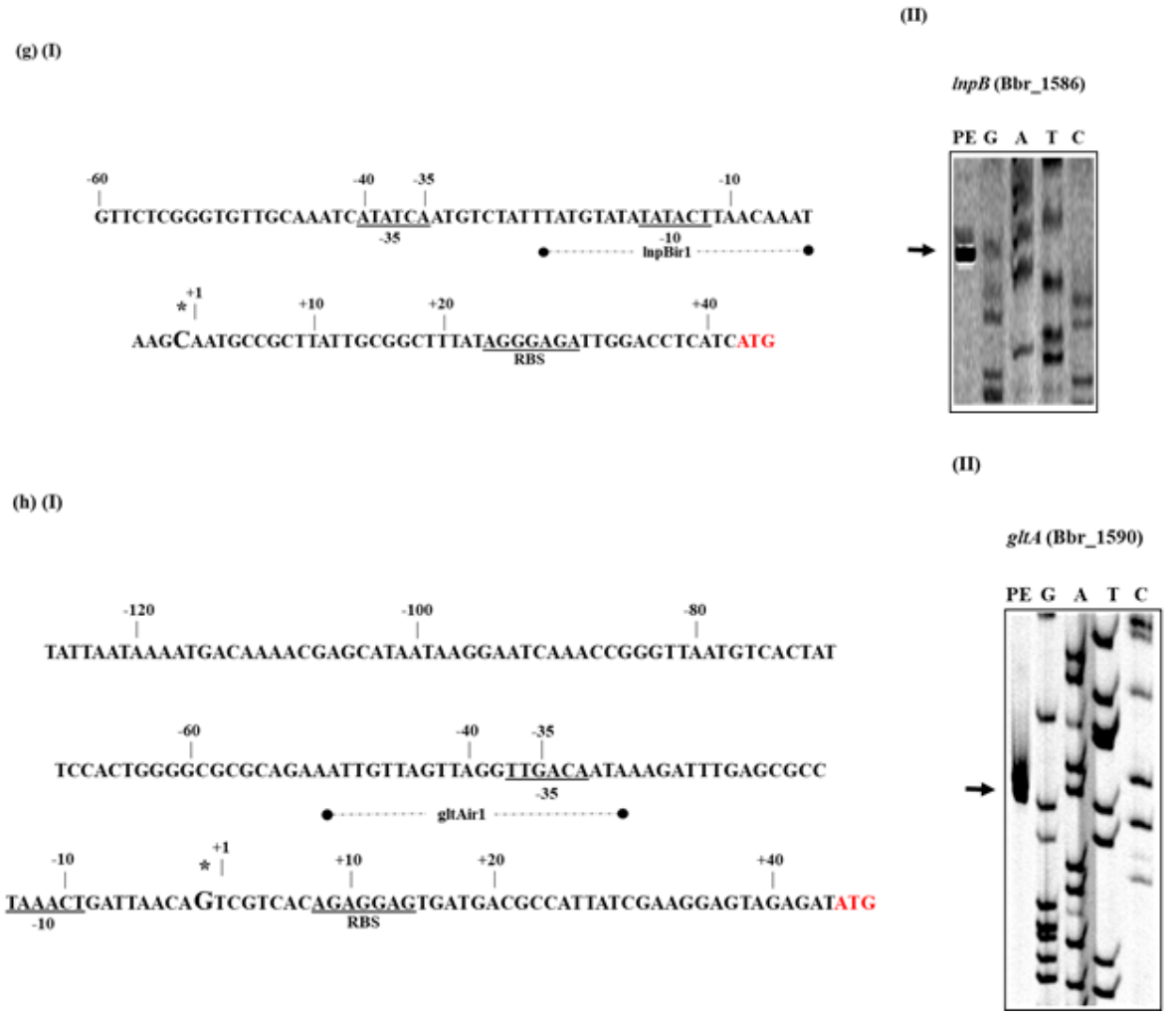
Supplemental Figure S3.1. Schematic representations (I) of the *lntP1* (a) and *lntS* (b) promoter regions for the binding of LntR. Underlining indicates the 10 and 35 hexamers, and ribosome binding sites deduced from the primer extension results; experimentally determined transcriptional start sites (TSS) are indicated by asterisks. The dashed lines under sequences with names in boldface indicate the inverted-repeat lntP1ir1, lntP1ir2, lntSir1 and lntSir2 binding sequences displayed above the lollipops. The arrows in the right panels (II) indicate the primer extension products.



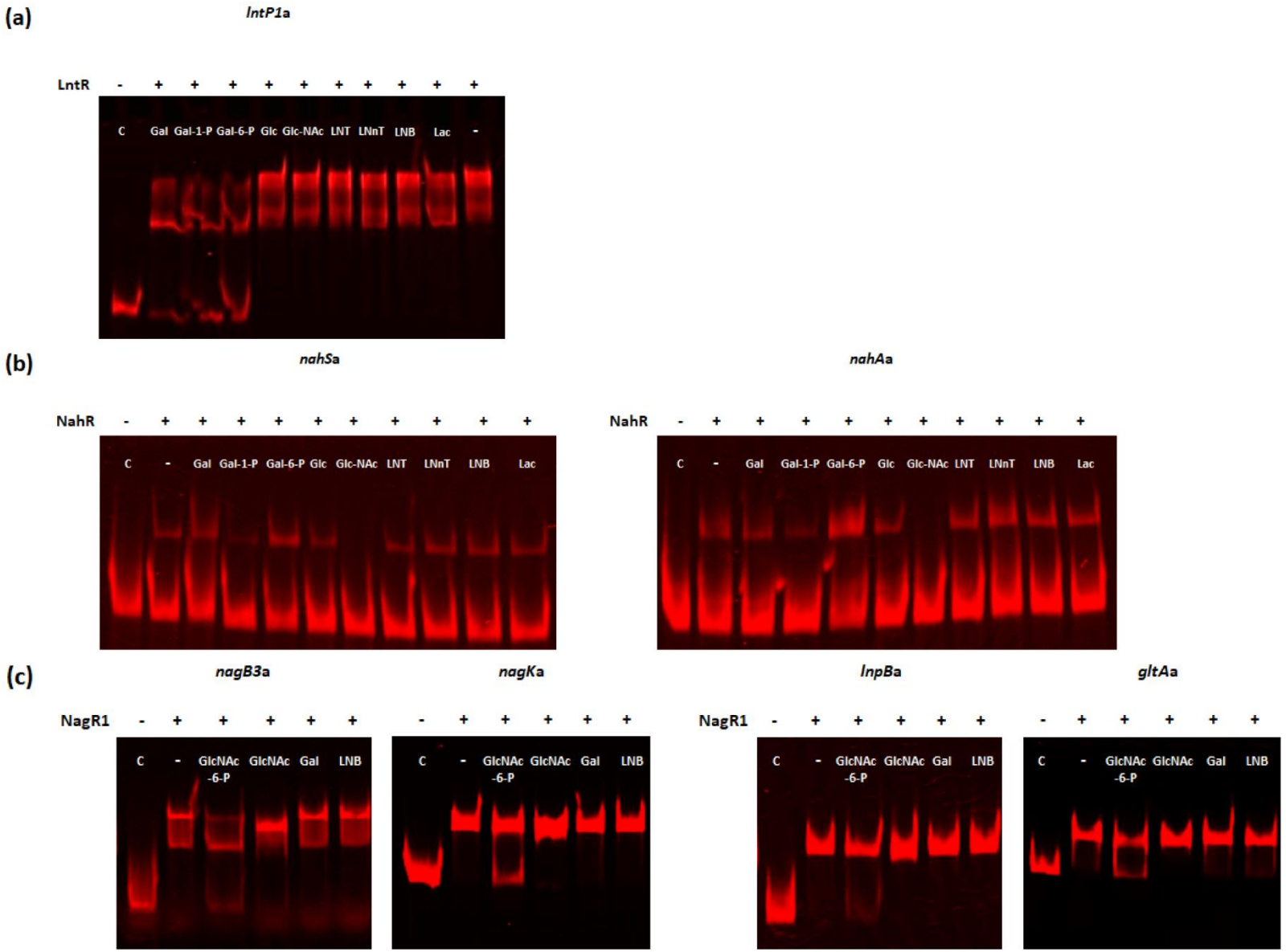
Supplemental Figure S3.1. Schematic representations (I) of the *nahS* (c) and *nahA* (d) promoter regions for the binding of NahR. Underlining indicates the 10 and 35 hexamers, and ribosome binding sites deduced from the primer extension results; experimentally determined transcriptional start sites (TSS) are indicated by asterisks. The dashed lines under sequences with names in boldface indicate the inverted-repeat nahSir1 and nahAir1 binding sequences displayed above the lollipops. The arrows in the right panels (II) indicate the primer extension products.



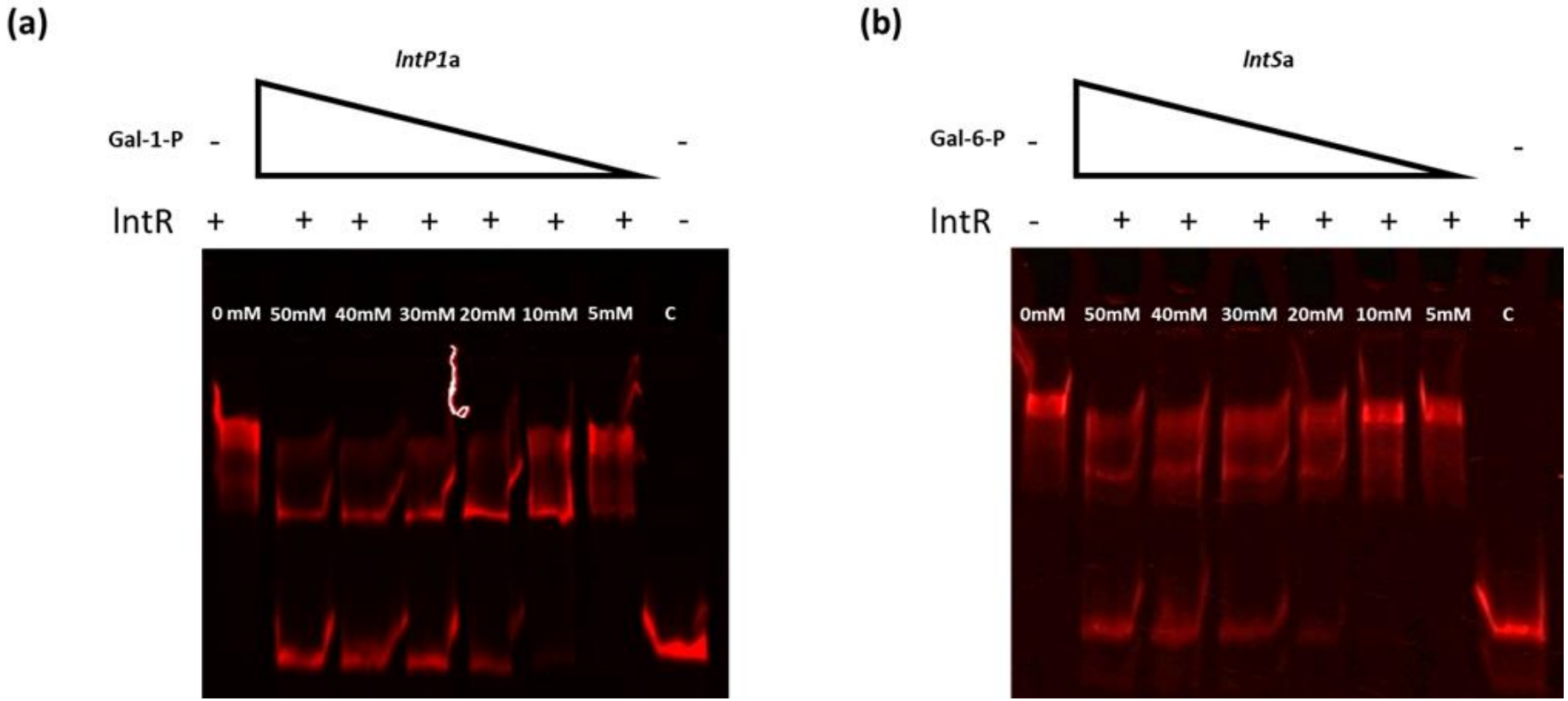
Supplemental Figure S3.1. Schematic representations (I) of the *nagB3* (e) and *nagK* (f) promoter regions for the binding of NagR1. Underlining indicates the 10 and 35 hexamers, and ribosome binding sites deduced from the primer extension results; experimentally determined transcriptional start sites (TSS) are indicated by asterisks. The dashed lines under sequences with names in boldface indicate the inverted-repeat nagB3ir1 and nagKir1 binding sequences displayed above the lollipops. The arrows in the right panels (II) indicate the primer extension products.



Supplemental Figure S3.1. Schematic representation (I) of the *lnpB* (g) and *gltA* (h) promoter regions for the binding of NagR1. Underlining indicates the 10 and 35 hexamers, and ribosome binding sites deduced from the primer extension results; experimentally determined transcriptional start sites (TSS) are indicated by asterisks. The dashed lines under sequences with names in boldface indicate the inverted-repeat lnpBir1 and gltAir1 binding sequences displayed above the lollipops. The arrows in the right panels (II) indicate the primer extension products.


(a)
lntP1a
LntR - + + + + + + + + + +
C Gal Gal-1-P Gal-6-P Glc Glc-NAc LNT LNnT LNB Lac -
(b)
nahSa
nahAa
NahR - + + + + + + + + + +
C - Gal Gal-1-P Gal-6-P Glc Glc-NAc LNT LNnT LNB Lac
NahR - + + + + + + + + + +
C - Gal Gal-1-P Gal-6-P Glc Glc-NAc LNT LNnT LNB Lac
(c)
nagB3a
nagKa
lnpBa
gltAa
NagR1 - + + + + +
C - GlcNAc -6-P GlcNAc Gal LNB
- + + + + +
C - GlcNAc -6-P GlcNAc Gal LNB
NagR1 - + + + + +
C - GlcNAc -6-P GlcNAc Gal LNB
- + + + + +
C - GlcNAc -6-P GlcNAc Gal LNB

Supplemental Figure S3.2. EMSA images showing (a) LntR, (b) NahR, and (c) NagR1 interactions with promoter-containing DNA fragments, with the addition of a range of potential carbohydrate inducers, at a concentration of 20mM. In each panel, 'C' indicates a negative control, where an equivalent amount crude cell extract from NZ9000 harbouring empty pNZ8150 was added instead of crude extract from the regulator-expressing NZ9000 strain.



Supplemental Figure S3.3. EMSA images showing LntR interactions with promoter-containing DNA fragments, with the addition of a gradient (a) Gal-1-P and (b) Gal-6-P, ranging from 0mM-50mM. In each panel, 'C' indicates a negative control, where an equivalent amount crude cell extract from NZ9000 harbouring empty pNZ8150 was added instead of crude extract from the regulator-expressing NZ9000 strai

3.9 References:

1. **Turroni, F., C. Peano, D.A. Pass, E. Foroni, M. Severgnini, M.J. Claesson, C. Kerr, J. Hourihane, D. Murray, F. Fuligni, M. Gueimonde, A. Margolles, G. De Bellis, P.W. O'Toole, D. van Sinderen, J.R. Marchesi, and M. Ventura,** *Diversity of Bifidobacteria within the Infant Gut Microbiota.* PLoS ONE, **2012**. **7**(5): p. e36957.
2. **Sivan, A., L. Corrales, N. Hubert, J.B. Williams, K. Aquino-Michaels, Z.M. Earley, F.W. Benyamin, Y. Man Lei, B. Jabri, M.-L. Alegre, E.B. Chang, and T.F. Gajewski,** *Commensal Bifidobacterium promotes antitumor immunity and facilitates anti–PD-L1 efficacy.* Science, **2015**. **350**(6264): p. 1084-1089.
3. **Maslowski, K.M. and C.R. Mackay,** *Diet, gut microbiota and immune responses.* Nat Immunol, **2011**. **12**(1): p. 5-9.
4. **Milani, C., S. Duranti, F. Bottacini, E. Casey, F. Turroni, J. Mahony, C. Belzer, S. Delgado Palacio, S. Arboleya Montes, L. Mancabelli, G.A. Lugli, J.M. Rodriguez, L. Bode, W. de Vos, M. Gueimonde, A. Margolles, D. van Sinderen, and M. Ventura,** *The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota.* Microbiology and Molecular Biology Reviews, **2017**. **81**(4).
5. **Gibson, G.R. and M.B. Roberfroid,** *Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics.* J Nutr, **1995**. **125**(6): p. 1401-12.
6. **Kunz, C., S. Rudloff, W. Baier, N. Klein, and S. Strobel,** *OLIGOSACCHARIDES IN HUMAN MILK: Structural, Functional, and Metabolic Aspects.* Annual Review of Nutrition, **2000**. **20**(1): p. 699-722.
7. **Brüssow, H.,** *Human microbiota: 'The philosophers have only interpreted the world in various ways. The point, however, is to change it'.* Microbial Biotechnology, **2015**. **8**(1): p. 11-12.
8. **Asakuma, S., E. Hatakeyama, T. Urashima, E. Yoshida, T. Katayama, K. Yamamoto, H. Kumagai, H. Ashida, J. Hirose, and M. Kitaoka,** *Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria.* J Biol Chem, **2011**. **286**(40): p. 34583-92.
9. **Sela, D.A. and D.A. Mills,** *Nursing our microbiota: molecular linkages between bifidobacteria and milk oligosaccharides.* Trends Microbiol, **2010**. **18**(7): p. 298-307.

10. **Kitaoka, M.,** *Bifidobacterial enzymes involved in the metabolism of human milk oligosaccharides.* Adv Nutr, **2012**. **3**(3): p. 422s-9s.
11. **Good, M., C.P. Sodhi, Y. Yamaguchi, H. Jia, P. Lu, W.B. Fulton, L.Y. Martin, T. Prindle, D.F. Nino, Q. Zhou, C. Ma, J.A. Ozolek, R.H. Buck, K.C. Goehring, and D.J. Hackam,** *The human milk oligosaccharide 2'-fucosyllactose attenuates the severity of experimental necrotising enterocolitis by enhancing mesenteric perfusion in the neonatal intestine.* Br J Nutr, **2016**. **116**(7): p. 1175-1187.
12. **Bode, L.,** *Human milk oligosaccharides: every baby needs a sugar mama.* Glycobiology, **2012**. **22**(9): p. 1147-62.
13. **Wada, J., T. Ando, M. Kiyohara, H. Ashida, M. Kitaoka, M. Yamaguchi, H. Kumagai, T. Katayama, and K. Yamamoto,** *Bifidobacterium bifidum lacto-N-biosidase, a critical enzyme for the degradation of human milk oligosaccharides with a type 1 structure.* Appl Environ Microbiol, **2008**. **74**(13): p. 3996-4004.
14. **Moller, P.L., F. Jorgensen, O.C. Hansen, S.M. Madsen, and P. Stougaard,** *Intra- and extracellular beta-galactosidases from Bifidobacterium bifidum and B. infantis: molecular cloning, heterologous expression, and comparative characterization.* Appl Environ Microbiol, **2001**. **67**(5): p. 2276-83.
15. **Miwa, M., T. Horimoto, M. Kiyohara, T. Katayama, M. Kitaoka, H. Ashida, and K. Yamamoto,** *Cooperation of beta-galactosidase and beta-N-acetylhexosaminidase from bifidobacteria in assimilation of human milk oligosaccharides with type 2 structure.* Glycobiology, **2010**. **20**(11): p. 1402-9.
16. **Ashida, H., A. Miyake, M. Kiyohara, J. Wada, E. Yoshida, H. Kumagai, T. Katayama, and K. Yamamoto,** *Two distinct alpha-L-fucosidases from Bifidobacterium bifidum are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates.* Glycobiology, **2009**. **19**(9): p. 1010-7.
17. **Kiyohara, M., K. Tanigawa, T. Chaiwangsri, T. Katayama, H. Ashida, and K. Yamamoto,** *An exo-alpha-sialidase from bifidobacteria involved in the degradation of sialyloligosaccharides in human milk and intestinal glycoconjugates.* Glycobiology, **2011**. **21**(4): p. 437-47.
18. **Wada, J., R. Suzuki, S. Fushinobu, M. Kitaoka, T. Wakagi, H. Shoun, H. Ashida, H. Kumagai, T. Katayama, and K. Yamamoto,** *Purification, crystallization and preliminary X-ray analysis of the galacto-N-biose-/lacto-N-biose I-binding protein (GL-BP) of the ABC transporter from Bifidobacterium longum JCM1217.* Acta Crystallogr Sect F Struct Biol Cryst Commun, **2007**. **63**(Pt 9): p. 751-3.

19. **Sakurama, H., M. Kiyohara, J. Wada, Y. Honda, M. Yamaguchi, S. Fukiya, A. Yokota, H. Ashida, H. Kumagai, M. Kitaoka, K. Yamamoto, and T. Katayama,** *Lacto-N-biosidase Encoded by a Novel Gene of Bifidobacterium longum Subspecies longum Shows Unique Substrate Specificity and Requires a Designated Chaperone for Its Active Expression.* The Journal of Biological Chemistry, **2013**. **288**(35): p. 25194-25206.

20. **Sela, D.A., J. Chapman, A. Adeuya, J.H. Kim, F. Chen, T.R. Whitehead, A. Lapidus, D.S. Rokhsar, C.B. Lebrilla, J.B. German, N.P. Price, P.M. Richardson, and D.A. Mills,** *The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome.* Proc Natl Acad Sci U S A, **2008**. **105**(48): p. 18964-9.

21. **Yoshida, E., H. Sakurama, M. Kiyohara, M. Nakajima, M. Kitaoka, H. Ashida, J. Hirose, T. Katayama, K. Yamamoto, and H. Kumagai,** *Bifidobacterium longum subsp. infantis uses two different beta-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides.* Glycobiology, **2012**. **22**(3): p. 361-8.

22. **Garrido, D., S. Ruiz-Moyano, and D.A. Mills,** *Release and utilization of N-acetyl-D-glucosamine from human milk oligosaccharides by Bifidobacterium longum subsp. infantis.* Anaerobe, **2012**. **18**(4): p. 430-5.

23. **Xiao, J.Z., S. Takahashi, M. Nishimoto, T. Odamaki, T. Yaeshima, K. Iwatsuki, and M. Kitaoka,** *Distribution of in vitro fermentation ability of lacto-N-biose I, a major building block of human milk oligosaccharides, in bifidobacterial strains.* Appl Environ Microbiol, **2010**. **76**(1): p. 54-9.

24. **Sela, D.A., Y. Li, L. Lerno, S. Wu, A.M. Marcobal, J.B. German, X. Chen, C.B. Lebrilla, and D.A. Mills,** *An Infant-associated Bacterial Commensal Utilizes Breast Milk Sialyloligosaccharides.* The Journal of Biological Chemistry, **2011**. **286**(14): p. 11909-11918.

25. **Sela, D.A., D. Garrido, L. Lerno, S. Wu, K. Tan, H.-J. Eom, A. Joachimiak, C.B. Lebrilla, and D.A. Mills,** *Bifidobacterium longum subsp. infantis ATCC 15697 α-Fucosidases Are Active on Fucosylated Human Milk Oligosaccharides.* Applied and Environmental Microbiology, **2012**. **78**(3): p. 795-803.

26. **Bunesova, V., C. Lacroix, and C. Schwab,** *Fucosyllactose and L-fucose utilization of infant Bifidobacterium longum and Bifidobacterium kashiwanohense.* BMC Microbiology, **2016**. **16**: p. 248.

27. **James, K., M.O. Motherway, F. Bottacini, and D. van Sinderen,** *Bifidobacterium breve UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways.* Sci Rep, **2016**. **6**: p. 38560.

28. **Egan, M., M.O. Motherway, M. Kilcoyne, M. Kane, L. Joshi, M. Ventura, and D. van Sinderen,** *Cross-feeding by Bifidobacterium breve UCC2003 during co-cultivation with Bifidobacterium bifidum PRL2010 in a mucin-based medium.* BMC Microbiol, **2014**. **14**: p. 282.

29. **Egan, M., M. O'Connell Motherway, M. Ventura, and D. van Sinderen,** *Metabolism of sialic acid by Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2014**. **80**(14): p. 4414-26.

30. **Egan, M., H. Jiang, M. O'Connell Motherway, S. Oscarson, and D. van Sinderen,** *Glycosulfatase-Encoding Gene Cluster in Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2016**. **82**(22): p. 6611-6623.

31. **De Man JC, R.M., Sharpe ME,** *A medium for the cultivation of 688 lactobacilli.* J.Appl. Bacteriol., **1960**. **23**: p. 130-135.

32. **Watson, D., M. O'Connell Motherway, M.H. Schoterman, R.J. van Neerven, A. Nauta, and D. van Sinderen,** *Selective carbohydrate utilization by lactobacilli and bifidobacteria.* J Appl Microbiol, **2013**. **114**(4): p. 1132-46.

33. **Terzaghi, B.E. and W.E. Sandine,** *Improved medium for lactic streptococci and their bacteriophages.* Appl Microbiol, **1975**. **29**(6): p. 807-13.

34. **Sambrook J, F.E., Maniatis T.**, *Molecular cloning: a laboratory manual, 2nd ed.* **1989**: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

35. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M.A. Rajandream, and B. Barrell,** *Artemis: sequence visualization and annotation.* Bioinformatics, **2000**. **16**(10): p. 944-5.

36. **O'Connell Motherway, M., A. Zomer, S.C. Leahy, J. Reunanen, F. Bottacini, M.J. Claesson, F. O'Brien, K. Flynn, P.G. Casey, J.A. Moreno Munoz, B. Kearney, A.M. Houston, C. O'Mahony, D.G. Higgins, F. Shanahan, A. Palva, W.M. de Vos, G.F. Fitzgerald, M. Ventura, P.W. O'Toole, and D. van Sinderen,** *Functional genome analysis of Bifidobacterium breve UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor.* Proceedings of the National Academy of Sciences, **2011**. **108**(27): p. 11217-11222.

37. **Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman,** *Basic local alignment search tool.* J Mol Biol, **1990**. **215**(3): p. 403-10.

38. **Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman,** *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res, **1997**. **25**(17): p. 3389-402.

39. **O'Riordan, K. and G.F. Fitzgerald,** *Molecular characterisation of a 5.75-kb cryptic plasmid from Bifidobacterium breve NCFB 2258 and determination of mode of replication.* FEMS Microbiol Lett, **1999**. **174**(2): p. 285-94.

40. **O'Connell Motherway, M., J. O'Driscoll, G.F. Fitzgerald, and D. Van Sinderen,** *Overcoming the restriction barrier to plasmid transformation and targeted mutagenesis in Bifidobacterium breve UCC2003.* Microb Biotechnol, **2009**. **2**(3): p. 321-32.

41. **Maze, A., M. O'Connell-Motherway, G.F. Fitzgerald, J. Deutscher, and D. van Sinderen,** *Identification and characterization of a fructose phosphotransferase system in Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2007**. **73**(2): p. 545-53.

42. **Law, J., G. Buist, A. Haandrikman, J. Kok, G. Venema, and K. Leenhouts,** *A system to generate chromosomal mutations in Lactococcus lactis which allows fast analysis of targeted genes.* J Bacteriol, **1995**. **177**(24): p. 7011-8.

43. **Pokusaeva, K., A.R. Neves, A. Zomer, M. O'Connell-Motherway, J. MacSharry, P. Curley, G.F. Fitzgerald, and D. van Sinderen,** *Ribose utilization by the human commensal Bifidobacterium breve UCC2003.* Microb Biotechnol, **2010**. **3**(3): p. 311-23.

44. **Egan, M., M. O'Connell Motherway, M. Ventura, and D. van Sinderen,** *Metabolism of sialic acid by Bifidobacterium breve UCC2003.* Applied and Environmental Microbiology, **2014**.

45. **Zomer, A., M. Fernandez, B. Kearney, G.F. Fitzgerald, M. Ventura, and D. van Sinderen,** *An interactive regulatory network controls stress response in Bifidobacterium breve UCC2003.* J Bacteriol, **2009**. **191**(22): p. 7039-49.

46. **Garcia de la Nava, J., D.F. Santaella, J. Cuenca Alba, J. Maria Carazo, O. Trelles, and A. Pascual-Montano,** *Engene: the processing and exploratory analysis of gene expression data.* Bioinformatics, **2003**. **19**(5): p. 657-8.

47. **van Hijum, S., A. de Jong, R. Baerends, H. Karsens, N. Kramer, R. Larsen, C. den Hengst, C. Albers, J. Kok, and O. Kuipers,** *A generally applicable validation scheme*

for the assessment of factors involved in reproducibility and quality of DNA-microarray data. BMC Genomics, **2005**. **6**(1): p. 1-10.

48. **van Hijum, S.A., J. Garcia de la Nava, O. Trelles, J. Kok, and O.P. Kuipers,** *MicroPreP: a cDNA microarray data pre-processing framework.* Appl Bioinformatics, **2003**. **2**(4): p. 241-4.
49. **Long, A.D., H.J. Mangalam, B.Y. Chan, L. Tolleri, G.W. Hatfield, and P. Baldi,** *Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in Escherichia coli K12.* J Biol Chem, **2001**. **276**(23): p. 19937-44.
50. **Mierau, I. and M. Kleerebezem,** *10 years of the nisin-controlled gene expression system (NICE) in Lactococcus lactis.* Appl Microbiol Biotechnol, **2005**. **68**(6): p. 705-17.
51. **O'Connell Motherway, M., G.F. Fitzgerald, and D. van Sinderen,** *Metabolism of a plant derived galactose-containing polysaccharide by Bifidobacterium breve UCC2003.* Microbial Biotechnology, **2011**. **4**(3): p. 403-416.
52. **Pokusaeva, K., M. O'Connell-Motherway, A. Zomer, J. Macsharry, G.F. Fitzgerald, and D. van Sinderen,** *Cellodextrin utilization by bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2011**. **77**(5): p. 1681-90.
53. **O'Connell, K.J., M. O'Connell Motherway, J. O'Callaghan, G.F. Fitzgerald, R.P. Ross, M. Ventura, C. Stanton, and D. van Sinderen,** *Metabolism of four alpha-glycosidic linkage-containing oligosaccharides by Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2013**. **79**(20): p. 6280-92.
54. **de Ruyter, P.G., O.P. Kuipers, and W.M. de Vos,** *Controlled gene expression systems for Lactococcus lactis with the food-grade inducer nisin.* Applied and Environmental Microbiology, **1996**. **62**(10): p. 3662-3667.
55. **O'Connell, K.J., M.O. Motherway, A. Liedtke, G.F. Fitzgerald, R. Paul Ross, C. Stanton, A. Zomer, and D. van Sinderen,** *Transcription of two adjacent carbohydrate utilization gene clusters in Bifidobacterium breve UCC2003 is controlled by LacI- and repressor open reading frame kinase (ROK)-type regulators.* Appl Environ Microbiol, **2014**. **80**(12): p. 3604-14.
56. **Hamoen, L.W., A.F. Van Werkhoven, J.J. Bijlsma, D. Dubnau, and G. Venema,** *The competence transcription factor of Bacillus subtilis recognizes short A/T-rich sequences arranged in a unique, flexible pattern along the DNA helix.* Genes Dev, **1998**. **12**(10): p. 1539-50.

57. **Kuipers, O.P., M.M. Beerthuyzen, R.J. Siezen, and W.M. De Vos,** *Characterization of the nisin gene cluster nisABTCIPR of Lactococcus lactis. Requirement of expression of the nisA and nisI genes for development of immunity.* Eur J Biochem, **1993**. **216**(1): p. 281-91.

58. **Ventura, M., R. Zink, G.F. Fitzgerald, and D. van Sinderen,** *Gene structure and transcriptional organization of the dnaK operon of Bifidobacterium breve UCC 2003 and application of the operon in bifidobacterial tracing.* Appl Environ Microbiol, **2005**. **71**(1): p. 487-500.

59. **Bailey, T.L. and C. Elkan,** *Fitting a mixture model by expectation maximization to discover motifs in biopolymers.* Proc Int Conf Intell Syst Mol Biol, **1994**. **2**: p. 28-36.

60. **Crooks, G.E., G. Hon, J.M. Chandonia, and S.E. Brenner,** *WebLogo: a sequence logo generator.* Genome Res, **2004**. **14**(6): p. 1188-90.

61. **Schneider, T.D. and R.M. Stephens,** *Sequence logos: a new way to display consensus sequences.* Nucleic Acids Research, **1990**. **18**(20): p. 6097-6100.

62. **Titgemeyer, F., J. Reizer, A. Reizer, and M.H. Saier, Jr.,** *Evolutionary relationships between sugar kinases and transcriptional repressors in bacteria.* Microbiology, **1994**. **140 (Pt 9)**: p. 2349-54.

63. **O'Connell Motherway, M., G.F. Fitzgerald, S. Neirynck, S. Ryan, L. Steidler, and D. van Sinderen,** *Characterization of ApuB, an extracellular type II amylopullulanase from Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2008**. **74**(20): p. 6271-9.

64. **Ruiz, L., M.O.C. Motherway, N. Lanigan, and D. van Sinderen,** *Transposon Mutagenesis in Bifidobacterium breve: Construction and Characterization of a Tn5 Transposon Mutant Library for Bifidobacterium breve UCC2003.* PLoS ONE, **2013**. **8**(5): p. e64699.

65. **Ruiz, L., A. Zomer, M. O'Connell-Motherway, D. van Sinderen, and A. Margolles,** *Discovering novel bile protection systems in Bifidobacterium breve UCC2003 through functional genomics.* Appl Environ Microbiol, **2012**. **78**(4): p. 1123-31.

66. **Hawley, D.K. and W.R. McClure,** *Compilation and analysis of Escherichia coli promoter DNA sequences.* Nucleic Acids Research, **1983**. **11**(8): p. 2237-2255.

67. **Bottacini, F., A. Zomer, C. Milani, C. Ferrario, G.A. Lugli, M. Egan, M. Ventura, and D. van Sinderen,** *Global transcriptional landscape and promoter mapping of the gut commensal Bifidobacterium breve UCC2003.* BMC Genomics, **2017**. **18**(1): p. 991.

68. **Egan, M., M. O'Connell Motherway, and D. van Sinderen,** *A GntR-type transcriptional repressor controls sialic acid utilization in Bifidobacterium breve UCC2003.* FEMS Microbiol Lett, **2015**. **362**(4): p.1-9.

69. **Ravcheev, D.A., M.S. Khoroshkin, O.N. Laikova, O.V. Tsoy, N.V. Sernova, S.A. Petrova, A.B. Rakhmaninova, P.S. Novichkov, M.S. Gelfand, and D.A. Rodionov,** *Comparative genomics and evolution of regulons of the LacI-family transcription factors.* Frontiers in Microbiology, **2014**. **5**(294): p.294.

70. **Gilbert, W. and A. Maxam,** *The Nucleotide Sequence of the lac Operator.* Proceedings of the National Academy of Sciences of the United States of America, **1973**. **70**(12 Pt 1-2): p. 3581-3584.

71. **Matthews, K.S. and J.C. Nichols,** *Lactose repressor protein: functional properties and structure.* Prog Nucleic Acid Res Mol Biol, **1998**. **58**: p. 127-64.

72. **Lewis, M.,** *The lac repressor.* Comptes Rendus Biologies, **2005**. **328**(6): p. 521-548.

73. **Plumbridge, J.,** *DNA binding sites for the Mlc and NagC proteins: regulation of nagE, encoding the N-acetylglucosamine-specific transporter in Escherichia coli.* Nucleic Acids Research, **2001**. **29**(2): p. 506-514.

74. **Brechemier-Baey, D., L. Dominguez-Ramirez, J. Oberto, and J. Plumbridge,** *Operator recognition by the ROK transcription factor family members, NagC and Mlc.* Nucleic Acids Res, **2015**. **43**(1): p. 361-72.

75. **Khoroshkin, M.S., S.A. Leyn, D. Van Sinderen, and D.A. Rodionov,** *Transcriptional Regulation of Carbohydrate Utilization Pathways in the Bifidobacterium Genus.* Frontiers in Microbiology, **2016**. **7**: p. 120.

76. **Weickert, M.J. and S. Adhya,** *A family of bacterial regulators homologous to Gal and Lac repressors.* J Biol Chem, **1992**. **267**(22): p. 15869-74.

77. **Weickert, M.J. and S. Adhya,** *The galactose regulon of Escherichia coli.* Mol Microbiol, **1993**. **10**(2): p. 245-51.

78. **O'Connell Motherway, M., M. Kinsella, G.F. Fitzgerald, and D. van Sinderen,** *Transcriptional and functional characterization of genetic elements involved in galacto-oligosaccharide utilization by Bifidobacterium breve UCC2003.* Microb Biotechnol, **2013**. **6**(1): p. 67-79.

79. **Barboza, M., D.A. Sela, C. Pirim, R.G. LoCascio, S.L. Freeman, J.B. German, D.A. Mills, and C.B. Lebrilla,** *Glycoprofiling Bifidobacterial Consumption of Galacto-Oligosaccharides by Mass Spectrometry Reveals Strain-Specific, Preferential*

Consumption of Glycans. Applied and Environmental Microbiology, **2009**. **75**(23): p. 7319-7325.

80. **Coulier, L., J. Timmermans, R. Bas, R. Van Den Dool, I. Haaksman, B. Klarenbeek, T. Slaghek, and W. Van Dongen,** *In-depth characterization of prebiotic galacto-oligosaccharides by a combination of analytical techniques.* J Agric Food Chem, **2009**. **57**(18): p. 8488-95.
81. **Pokusaeva, K., G.F. Fitzgerald, and D. van Sinderen,** *Carbohydrate metabolism in Bifidobacteria.* Genes & Nutrition, **2011**. **6**(3): p. 285-306.
82. **Álvarez-Martín, P., M. O'Connell-Motherway, D. van Sinderen, and B. Mayo,** *Functional analysis of the pBC1 replicon from Bifidobacterium catenulatum L48.* Applied Microbiology and Biotechnology, **2007**. **76**(6): p. 1395-1402.

Chapter IV

Screening of infant faecal samples for human milk oligosaccharide-utilising *Bifidobacterium* isolates.

Faecal Samples were provided by the InfantMet Project.

Valentina Ambrogi and Mariane Vigoureux assisted with the faecal sample screening, in material preparation, isolate subculuturing and ITS PCR analysis.

Dr. Francesca Bottacini carried out the GH profile analysis.

B. breve isolates obtained in the work of this chapter were used in the publication:

Bottacini, F., Morrissey, R., Roberts, R.J., James, K., van Breen, J., Egan, M., Lambert, J., van Limpt, K., Knol, J., O'Connell Motherway, M., & van Sinderen, D. (2018). Comparative genome and methylome analysis reveals restriction/modification system diversity in the gut commensal *Bifidobacterium breve*. NAR Volume **46**, Issue **8**.

4.1 Abstract

Utilisation of Human Milk Oligosaccharides (HMOs) represents an important adaptation that facilitates the establishment and dominance of bifidobacteria in the intestinal microbiota of the breastfed neonate. This adaptation comes in the form of a range of distinct metabolic strategies employed by different infant-associated *Bifidobacterium* species. The particular arsenal of glycosyl hydrolases (GHs) encoded by the genomes of these species is believed to be a reflection of their HMO metabolizing abilities. In the current study, faecal samples of exclusively breastfed infants were screened for bifidobacteria based on their ability to grow on various glycans, including lactose, GOS and the HMO sugars 3'-sialyllactose, 2'-fucosyllactose and 3-fucosyllactose, as selective carbohydrate sources. A range of bifidobacterial species were obtained, varying between the selective carbohydrate used, and supporting the notion of selective HMO consumption. We analysed GH profiles of representative strains of each of the identified species, correlating these profiles with their preferential carbohydrate-based selection, and thereby highlighting the importance of carbohydrate utilisation strategies for bifidobacterial colonisation in the infant gut.

4.2 Introduction

Bifidobacteria are among the first colonisers of the infant gut, and have been shown to dominate the intestinal microbiota of, in particular, breastfed neonates [1, 2]. Being Gram-positive, anaerobic members of the Actinobacteria phylum, *Bifidobacterium* species are commensals of the mammalian, avian and insect gut, and currently garner a great deal of scientific and commercial interest because of their purported benefits to infant host health and development [3-8].

Breastfeeding is known to impact on the microbiota composition of the infant gut [9-12]. Perhaps the most influential component of human breastmilk in terms of its impact on infant microbiota composition is a group of related glycans, commonly referred to as Human Milk Oligosaccharides (HMOs). They represent a heterogeneous mix of at least 200 distinct glycan structures [13], constituting, after lactose, the second-largest carbohydrate component of breastmilk [10, 14]. All HMO structures contain (at the reducing end) lactose as the base, which is elongated through the addition of a lacto-N-biose (LNB) moiety or its stereoisomer N-acetyllactosamine (LacNAc), creating the tetrasaccharides lacto-N-tetraose (LNT) or lacto-N-neotetraose (LNnT), respectively. These, in turn, may be elongated by the addition of further LNB or LacNAc residues at the reducing end. Any of these structures, including lactose itself (though lactose is not considered an HMO), may be fucosylated or sialylated [10, 13, 15].

Selective consumption of HMOs provides a major advantage to the establishment of specific microbes in the breastfed neonatal gut microbiota, and few do this as successfully as bifidobacteria [5, 13, 16, 17]. This is reflected in the enrichment of specific infant-associated species in the breastfed neonatal gut microbiota [1], in particular *Bifidobacterium bifidum, Bifidobacterium longum* subsp. *infantis*, *Bifidobacterium longum* subsp. *longum, Bifidobacterium breve, Bifidobacterium pseudocatenulatum* and (though less frequently) *Bifidobacterium kashiwanohense* [1, 18-22]. *B. bifidum* and certain strains of *B. longum* subsp. *longum* have been shown to carry out the extracellular hydrolysis of complex HMOs, including LN(n)T

and fucosylated and sialylated structures, prior to the import and metabolism of (some of) the resultant mono- and di-saccharides, [23-30], while *B. kashiwanohense* is thought to internalise and intracellularly hydrolyse small fucosylated HMO structures, such as fucosyllactose [2]. *B. breve* and *B. longum* subsp. *infantis* are known to internalise intact neutral HMO moieties such as LNT, LNnT and LNB, and sequentially hydrolyse these oligosaccharides for the subsequent metabolism of their monosaccharide constituents [13, 23, 31-35]. However, *B. infantis* has also been demonstrated to take up and utilise small-mass fucosylated and sialylated HMO structures. These HMO moieties, such as fucosyl-/sialyl-lactose and fucosylated or sialylated LN(n)T, are internalised [36], and then degraded into their constituent monosaccharides by an array of glycosyl hydrolases, including fucosidases and sialidases [23, 31, 37-39]. At present, little to nothing is known about HMO metabolism in *B. pseudocatenulatum*, except for its ability to consume LNT [40]. Recent work has correlated differences in neonatal microbiota composition with variations in the specific HMO components consumed by the infant and thus microbiota itself [41]. Thus, it should follow that such differences extend to the *Bifidobacterium* component of the neonatal microbiota. Previous work has highlighted considerable differences in metabolic capabilities and strategies for carbohydrate utilisation by different bifidobacterial species [42], and as such, the propagation or isolation of specific species of *Bifidobacterium* should be dependent on the available carbohydrate source.

In this chapter, we describe the generation of a strain bank of 502 *Bifidobacterium* isolates, belonging to the species *B. breve*, *B. bifidum*, *B. longum* (subsp. *infantis* and *longum*), *B. pseudocatenulatum* and *B. kashiwanohense* from the faeces of exclusively breastfed infants in the first 4-8 weeks of life, using various sugars as the selective carbohydrate. The employed selective carbohydrates include the HMO components 3'-sialyllactose (3'-SL), 2'-fucosyllactose (2'-FL) and 3-fucosyllactose (3-FL), as well as the sugars lactose and galacto-oligosaccharides (GOS). Lactose [42] and GOS [23, 43, 44] are substrates efficiently used by a wide range of *Bifidobacterium* species, both of infant and adult origin, and thus provide an

excellent comparison for the species isolated using these substrates to those obtained using HMO, especially as lactose comprises the dominant carbohydrate component of breastmilk [10], while GOS is a prebiotic frequently included in infant formula milk [45, 46]. The species identity of the isolates obtained was determined, revealing interesting differences between bifidobacterial species isolated using HMO and those obtained using less selective substrates. Furthermore, we compare the distribution of glycoside hydrolases (GHs) involved in carbohydrate metabolism in representative strains (previously sequenced other publications) of all species obtained in the screening, with particular focus on HMO utilisation, revealing insights into key metabolic adaptations common to infant-associated *Bifidobacterium* species.

4.3 Materials and Methods

Faecal sample handling and strain isolation. Faecal samples were obtained from the APC Microbiome Institute human clinical sample collection (stored at -80°C), originally collected for the InfantMet Cohort [47]. Mothers were approached for consent between February 2012 and May 2014 at the Cork University Maternity Hospital, with ethical approval provided by the Cork University Hospital Research Ethics Committee (ethical approval reference: ECM (w) 07/02/2012). Faecal samples were selected based on infant age at the time of collection (age range: 4-8 weeks) and the requirement of exclusive breastfeeding up the point of sample collection. This information was obtained from the InfantMet subject metadata [47]. *Bifidobacterium* culture isolates were obtained using a method based on that employed in the InfantMet study [47], though with the following significant modifications. All sample preparation and platings were carried out under anaerobic conditions in a modular atmosphere-controlled system (Davidson and Hardy, Belfast, Ireland) at 37 °C. One gram of frozen faecal sample was thawed anaerobically at 37°C, and resuspended in 10 ml PBS (Sigma Aldrich, Ireland) supplemented with 0.05 % L-cysteine hydrochloride (Sigma Aldrich, Ireland). Selection of bifidobacteria was performed by spread-plating 10 aliquots of 1 ml of the faecal resuspension on: modified de Man Rogosa and Sharpe (mMRS) agar prepared from first principles [48], supplemented with 1 % (wt/vol) 2'-fucosyllactose, 3-fucosyllactose, 3'-sialyllactose (all Glycom A/S, Lyngby, Denmark) or lactose (Sigma Aldrich, Ireland), as well as 0.05 % L-cytseine HCl, 100 μg/ml mupirocin (Oxoid, Fannin, Ireland) and 50 units nystatin suspension (Sigma Aldrich, Ireland). As a control, faecal resuspensions were also plated on TOS Propionate Agar (Merck KGaA, Darmstadt, Germany), supplemented with 0.05 % L-cysteine HCl, 100 μg/ml mupirocin and 50 units nystatin suspension. Agar plates were incubated anaerobically at 37 °C for 72 h. Emerging colonies from selected plates were resuspended in 1 ml of PBS supplemented with 0.05 % L-cysteine HCl, after which the resuspensions from around 10 plates of a single sample were pooled and homogenised, and serial-diluted in the aforementioned PBS, and dilutions between $1x10^{-4}$ and $1x10^{-8}$ were spread-plated on the same type of

HMO/lactose-supplemented modified de Man Rogosa and Sharpe (mMRS) agar or TOS propionate agar from which the colonies were obtained, as described above, and incubated anaerobically at 37 °C for 72 h. A representative number of single, isolated colonies were selected from the serial dilution plates, and again re-streaked onto the same type of agar plates as isolated from, and incubated anaerobically at 37 °C for 48 h. This sub-cultivation process was repeated twice, after which single isolated colonies were inoculated into modified de Man Rogosa and Sharpe (mMRS) medium prepared from first principles, supplemented with 1 % lactose and 0.05 % L-cysteine HCl. Pure cultures of were stocked and stored at -80°C.

Confirmation of *Bifidobacterium* isolates, and species identification. Pure cultures were confirmed as bifidobacteria using the fructose-6-phosphate phosphoketolase assay [49], as well as employing PCR reactions based on the ITS (Intergenic Spacer region between the 16S and 23S rRNA-encoding genes) and *ruvAB* sequences, using primer pairs Bifspp and 23Sbif [50], and ruvABf and ruvABr, respectively (Supplemental Table S4.1). The amplicons from the ITS- and *ruvAB*-based PCR reactions were sequenced (MWG Eurofins, Ireland) using the same primers as those used for the PCR, and the resulting sequences were assembled (Seqman, Seqbuilder, DNAStar Suite) and subject to a nucleotide BLAST (NCBI) in order to determine their species identity.

Bacterial strains and culture conditions. *Bifidobacterium* cultures were routinely cultured in modified de Man Rogosa and Sharpe (mMRS) medium prepared from first principles [48], supplemented with cysteine-HCl (0.05 %, wt/vol) and 1 % lactose. *Bifidobacterium* cultures were incubated under anaerobic conditions in a modular atmosphere-controlled system (Davidson and Hardy, Belfast, Ireland) at 37°C.

Comparison of glycoside hydrolase content in infant-associated bifidobacteria. Publicly available genome sequences from bifidobacterial strains were selected to represent each species isolated during the faecal screening. The glycosyl hydrolase (GH) content of these representative strains were obtained from the information stored within the publicly available CAZy database (http://www.cazy.org/), and compared. The strains used for this purpose were: *B. breve* UCC2003 [51], *B. longum* subsp. *longum* NCC2705 [52], *B. longum* subsp. *infantis* ATCC 15697 [31], *B. bifidum* PRL2010 [53], *B. kashiwanohense* JCM 15439 [54] and *B. pseudocatenulatum* DSM 20438 [55].

4.4 Results

Generation of a strain bank of infant-derived bifidobacteria. Using selective growth media we isolated bifidobacteria from seventeen faecal samples, which had been derived from seventeen exclusively breastfed (at the time of sampling, according to proforma data provided by the mothers), healthy and full-term 4 to 8 week-old infants. This subject selection ensured the presence of a microbial gut population that was highly dependent on breast milk for growth, whilst also allowing for the development of a 'typical' infant gut microbiota.

Samples were treated as described in the InfantMet study [47], with a number of modifications, as described in the Materials and Methods section. The inclusion of mupirocin in all media used ensured the selection of bifidobacteria [56, 57], while the inclusion of nystatin ensured fungal inhibition [58]. The use of anaerobic conditions during sample preparation, plating, resuspension and secondary plating ensured minimal oxidative stress to the bifidobacteria present, thus minimizing loss of potential isolates. Resuspension and re-plating of obtained bacteria following the initial plating and incubation was conducted to eliminate false positives due to growth on residual lactose (or other carbohydrates) present in the faeces prior to selection on a particular carbohydrate source. The mMRS used was supplemented with lactose, or various HMO components as the sole carbohydrate source, in order to select for *Bifidobacterium* isolates capable of such a particular HMO. The HMO components included: 3'-sialyllactose (3'-SL), 2'-fucosyllactose (2'-FL) and 3-fucosyllactose (3-FL). TOS propionate agar, containing galactooligosaccharides (GOS) as the carbohydrate source, was also used as a control medium to generally select for bifidobacteria.

A bank of 502 *Bifidobacterium* isolates was generated, of which 252 were isolated using lactose, 157 isolated using GOS and 93 isolated using HMO (Fig. 4.1). This bank of isolates, following species identification, was shown to consist of 177 *B. breve*, 108 *B. bifidum*, 7 *B. kashiwanohense*, 177 *B. longum* (subsp. *longum*/*infantis*) and 33 *B. pseudocatenulatum* isolates (Fig. 4.1, Fig. 4.2). Of these, 252 were isolated from mMRS agar with 1 % lactose,

93 were isolated from mMRS agar with 1 % HMO (3'-SL, 2'-FL or 3-FL) and 157 were isolated from TOS (GOS) propionate agar. While comparing the numbers of each species within the total bank of isolates obtained does not provide us with a great deal of information (as equal number of isolates were not obtained using each substrate), examining the relative proportions of each species obtained for a given substrate separately reveals a very interesting correlation between the carbohydrate(s) present and the species isolated. When using HMO (3'-SL, 2'-FL or 3-FL) as the selective carbohydrate, over 75 % of the 93 isolates obtained were identified as *B. bifidum* (Fig. 4.1). The remaining isolates were made up of equal proportions (~7.5 %) of *B. breve*, *B. kashiwanohense* and *B. pseudocatenulatum* strains, and a mere 1 % were shown to belong to *B. longum* spp. (Fig. 4.1). This starkly contrasts with the profile of the isolates obtained using lactose as a selective carbohydrate, where no *B. kashiwanohense* isolates were obtained, while just 2 % of the isolates were identified as *B. bifidum* (Fig. 4.1). Conversely, *B. breve* and *B. longum* spp. dominated these isolates, making up 44 % and 51 %, respectively, of the 252 isolates characterized (Fig. 4.1). The isolates obtained using TOS (GOS) propionate agar provided something of a more balanced species profile. While no *B. kashiwanohense* strains were isolated using this medium either, *B. breve*, *B. longum* spp., *B. bifidum* and *B. pseudocatenulatum* composed 38 %, 31 %, 20 % and 11 % of the 157 isolates obtained, respectively (Fig. 4.1).

These results show that, as expected, the utilized carbohydrate source and medium base influenced the frequency at which certain bifidobacterial species were isolated from these faecal samples. There appears to be a strong bias for the selection of *B. bifidum* and *B. kashiwanohense* when faecal samples were screened using the HMO's 3'-SL, 2'-FL or 3-FL as the selective substrate. This is echoed in the much lower occurrence of *B. bifidum* and the absence of *B. kashiwanohense* among isolates obtained using lactose or GOS as the selective carbohydrate. In contrast, *B. breve* and *B. longum* spp. appear to dominate among the isolates obtained using lactose as a carbohydrate. *B. pseudocatenulaum* was obtained using all 3 substrates, composing 3-11 % of the isolates. Of the 3 substrates used, GOS appears to have generated the most

balanced proportion of *Bifidobacterium* species isolates. However, in the case of the GOS-derived isolates, it must be noted that the composition of the TOS propionate medium base differs from that of the mMRS medium base. As such, other components than just the carbohydrate source may have affected the species obtained using this medium in comparison to mMRS supplemented with lactose or HMO.

Distribution of glycoside hydrolases in infant-associated bifidobacteria. In order to find a correlation between infant-associated *Bifidobacterium* species obtained by carbohydrate-based selection and the array of glycosyl hydrolases (GHs) expressed by these species, an analysis of the distribution of GH families across the species we isolated in the infant faecal screening was carried out. One representative strain of each species was selected for this purpose and compared for their GH-encoding content; namely *B. breve* UCC2003 [51], *B. longum* subsp. *longum* NCC2705 [52], *B. longum* subsp. *infantis* ATCC 15697 [31], *B. bifidum* PRL2010 [53], *B. kashiwanohense* JCM 15439 [54] and *B. pseudocatenulatum* DSM 20438 [55].

Analysis of the GH distribution across the genomes of these 6 species reveals a significant variation in the array of GH families (Fig. 4.2). A core set of highly-conserved GH families was observed across all 6 species, including: GH2 and GH42 β-galactosidases, GH3 β-glucosidases, GH13 α-amylases and GH 23 lysozymes. The conservation of GH2 and GH42 β-galactosidases across these 6 infant-derived species is unsurprising, as both these classes of GH have been shown to be involved in lactose hydrolysis [23, 25, 26, 31, 32, 35], which comprises the largest carbohydrate component of breastmilk. GH2 β-galactosidases have also been previously implicated in the utilisation of GOS [59], which may explain the relatively balanced proportion of species obtained using this substrate. Furthermore, both GH42 and GH2 β-galactosidases have also been demonstrated to degrade the neutral HMO tetrasaccharide LNnT in *Bifidobacterium* species, while GH42 β-galactosidases have also been demonstrated to degrade LNT [26, 32, 35]. Here, *B. breve* was found to be particularly rich in GH2 β-galactosidases.

GH3 β-glucosidases typically act on plant-derived oligosaccharides and polysaccharides, as has previously been observed in both *B. longum* subsp. *infantis* [60] and *B. longum* subsp. *longum* [61]. GH13 α-amylases represent the most dominant GH family of the *Bifidobacterium* glycobiome [62], and also act on a wide range of plant-derived carbohydrates [42]. This is interesting, as such plant-derived glycans represent the dominant carbohydrates in the adult mammalian diet [63]. GH23 lysozymes are typically associated with the degradation of exopolysaccharides such as chitin [64].

Examining the differences between the GH content of these species, however, provides some interesting insights into their specific carbohydrate utilisation capabilities. The presence of GH33 sialidases appears strictly conserved in *B. breve*, *B. longum* subsp. *infantis* and *B. bifidum*. This is remarkable, as *B. bifidum* comprised the vast majority of isolates obtained using 3'-SL as the selective carbohydrate. *B. bifidum* is known to encode sialidases capable of sialyllactose degradation [28]. A small number of *B. breve* isolates were also obtained using 3'-SL as the selective substrate, and while, however, *B. breve* UCC2003 is not capable of degrading sialyllactose, it has been previously shown to utilise sialic acid released in the extracellular hydrolysis of sialyllactose by *B. bifidum* [65, 66]. The sialidase expressed by *B. breve* may in fact function in the hydrolysis of other sialylated HMO structures, such as sialyl-LN(n)T. While GH95 α-L-fucosidases appear to be distributed across most of the 6 species, GH29 α-L-fucosidase-encoding genes are only present in the genomes of *B. longum* subsp. *infantis*, *B. bifidum* and *B. kashiwanohense*. Of the 6 species examined here, these three are the only ones that have been demonstrated to typically degrade fucosyllactose [2, 23, 27, 29, 37, 39]. This agrees with the make-up of species isolated here using 2'-FL or 3-FL as the selective carbohydrate. *B. bifidum* was by a substantial margin the dominant species isolated using fucosyllactose, and this substrate was also the only substrate from which *B. kashiwanohense* was isolated. This suggests the conservation of a GH29 rather than a GH95 α-L-fucosidase as key to the utilisation of fucosyllactose, the latter perhaps primarily

functioning in the degradation of larger fucosylated HMO structures, such as fucosyl-LN(n)T.

Shifting the focus to differences in the distribution of GH families not associated with carbohydrates used in the screening also reveals some interesting patterns among the 6 species. The species *B. longum* subsp. *longum* and *B. pseudocatenulatum*, which are frequently found in the microbiota of both adults and infants, appear rich in GH family proteins associated with the metabolism of plant-derived carbohydrates, often absent from the strictly infant-associated species *B. longum* subsp. *infantis* and *B. bifidum*, including GH31 α-glucosidases, GH120 β-xylosidases and GH127 β-L-arabinofuranosidases. *B. breve* and *B. kashiwanohense* appear to possess a mix of both HMO-specific and plant-derived glycan-specific GH family proteins.

4.5 Discussion

Bifidobacteria employ a range of strategies to allow them to utilise an array of carbohydrates. This wide variation in carbohydrate metabolic capabilities has been observed in both adult and infant-associated *Bifidobacterium* species [13, 16, 42, 67], with particular specialisations in the latter group enabling the selective development of these species in the environment of the breastfed infant gut. While lactose comprises the predominant carbohydrate component of breastmilk, and is metabolised by a wide range of bifidobacteria [42], the adaptation to consume HMO sugars is believed to play a major role in the establishment of *B. breve*, *B. bifidum*, *B. longum* subsp. *infantis*, *B. kashiwanohense*, *B. longum* subsp. *longum* and *B. pseudocatenulatum* as part of the (breastfed) neonatal gut microbiota [1, 18-22]. In particular, the consumption of sialylated and fucosylated HMO glycans is a strategy efficiently employed by infant-associated *Bifidobacterium* species [5, 13, 16, 17]. Here, we observe significant variation in *Bifidobacterium* species isolated using the HMOs fucosyllactose or sialyllactose, as compared lactose or GOS.

Analysis of the strains isolated using each selective carbohydrate source clearly shows the influence of carbohydrate source on the species of *Bifidobacterium* obtained from infant faecal samples. The strong bias for selection of *B. bifidum* in the case of the selective sugars 2'-FL, 3-FL or 3'-SL supports the notion of *B. bifidum* being an efficient utiliser of fucosylated and sialylated HMOs through the activity of its extracellular fucosidases and sialidases [13, 23, 67]. While by no means the dominant species obtained using this substrate, *B. kashiwanohense* was only obtained using 2'-FL as the selective carbohydrate. This is not surprising, however, as not only has previous work demonstrated the ability of *B. kashiwanohense* to consume fucosyllactose [2], but it has also highlighted that this species is rarely isolated (only previously isolated twice, in both cases from infant faeces) [22, 54]. The inclusion of lactose or GOS in the media appears to encourage growth, and thus isolation of *B. longum* spp. and *B. breve*. This makes sense in the context of their metabolic capabilities, as while all strains obtained in this screening process are capable of utilising lactose [42], *B. breve* and *B. longum*

subsp. *infantis* can also target and efficiently utilise other galactose-rich small-molecular weight glycans such as LNT and LNnT [13, 23, 31-35]. While the primer sets used did not enable the subspeciation of *B. longum* spp. isolates into the *longum* or *infantis* subspecies, it would be plausible that many of these isolates belong to subsp. *longum*. *B. longum* subsp. *infantis* is more commonly associated with the microbiota of breastfed infants, whereas subsp. *longum* is more common in the microbiota of adults and formula-fed infants [1, 19, 50, 68]. However, the former are capable of utilising free fucosyl- and sialyl-lactose [2, 69], and a greater proportion of *B. longum* isolates would be expected using these HMO substrates if they indeed were of the subspecies *infantis*. The prevalence of *B. pseudocatenulatum* across the isolates obtained from all used substrates suggests its ability to grow equally well on HMO or alternative carbohydrate sources. This would agree with previous work which has identified it as a commensal member of both the infant and adult gut microbiota [1, 50, 68].

Of course, the main caveats to the results obtained from this screening are that not all samples were screened using the exact same set of carbohydrates, and some samples may intrinsically be richer in certain *Bifidobacterium* species, which would create a natural bias for the selection of particular species on a sample-by-sample basis. Nonetheless, the compiled results show a trend toward the selection of specific species based on the carbohydrate source used, namely HMO or non-HMO.

Examining the results of the screening in the context of the GH content of a representative strain for each species further strengthens the argument for the partitioning of specific carbohydrate metabolism into different infant-associated species, particularly with regard to HMO components. While they do not represent the GH content of all strains of each species, these strains represent the general features of each as an indication of their carbohydrate utilisation capabilities. The observed conservation of fucosidase- and sialidase-encoding genes, particularly in *B. bifidum*, agrees with its ability to utilise fucosylated and sialylated HMOs [16, 27, 28]. Likewise, the observed presence of fucosidases in *B. kashiwanohense* may correlate with the ability of the species to utilise fucosyllactose [2]. However, as only two genomes of

this species are currently publicly available, this is yet to be considered as a common feature of *B. kashiwanohense* [54, 70]. *B. longum* subsp. *infantis* also expresses sialidases and fucosidases, and has been demonstrated to utilise both fucosyllactose and sialyllactose [2, 37, 38]. However, it may be the case that *B. bifidum* and to a lesser extent, *B. kashiwanohense*, may simply be more efficient utilisers of these particular HMO moieties, accounting for the former's dominance and latter's presence in the isolates obtained using fucosyl- or sialyl-lactose as the selective carbohydrate. While *B. longum* subsp. *infantis* is capable of metabolising free fucosyllactose and sialyllactose [2, 69], this species is known to internalise and intracellularly degrade larger, complex fucosylated and/or sialylated HMO (such as fucosyl-/sialyl-LN(n)T) [14, 16, 17], and may preferentially consume these over fucosyl-/sialyl-lactose. This may account for the lower abundance of *B. longum* isolates obtained, if they are belonging to subsp. *infantis*, compared to *B. bifidum* or *B. kashiwanohense*, in the isolates obtained using fucosyllactose or sialyllactose. However, as mentioned, it is highly likely that these *B. longum* isolates belong to the subsp. *longum*, which does typically not consume HMO substrates [14, 16, 17].

Analysing the overall GH content of the six analysed genomes reveals an interesting pattern in terms of the origin of the carbohydrates targeted. While all six species possess GH family proteins which target lactose (which is expected as it represents the dominant carbohydrate in breastmilk ahead of HMO), an enrichment of GH family proteins targeting HMO-derived sugars is observed in *B. bifidum* and *B. longum* subsp. *infantis*, and an enrichment of GH family proteins targeting plant-derived sugars is observed in *B. longum* subsp. *longum* and *B. pseudocatenulatum*, with *B. breve* and *B. kashiwanohense* encoding members of both GH families. This may provide an advantage during exclusive breastfeeding in early infancy for specialist HMO-consuming species such as *B. bifidum* and *B. longum* subsp. *infantis*. Subsequent weaning would be expected to present an opportunity for species capable of both HMO and plant-derived sugar metabolism such as *B. breve* and *B. kashiwanohense*, while subsequent to the transition to exclusively solid foods may favour more adult-associated species, such as *B. longum* subsp.

longum and *B. pseudocatenulatum*, which specialise in the metabolism of a range of plant-derived sugars. Such differences in the species of *Bifidobacterium* present in the gut microbiota over the course of infancy have previously been observed [1, 50, 71, 72].

However, this analysis of GH distribution in these strains of course has its limitations. We only included a small number of species from the *Bifidobacterium* genus in this analysis; just the six species we obtained in our isolations, all of which are associated with the infant gut microbiota (as well as the adult for *B. longum* subsp. *longum* and *B. pseudocatenulatum*). Furthermore, we examined the GH content of just a single representative strain for each species. In order to obtain a more comprehensive picture of the distribution of GH's across the genus, such an analysis would need to be repeated using a much wider range of *Bifidobacterium* species, from a range of different environments, and as many strains of each as possible.

On the whole, our work has allowed the development of a methodology for the isolation of HMO-utilising bifidobacteria from infant faeces, and this method can easily be adjusted for the selection of bifidobacteria under different selective conditions and/or from different sources. Our findings furthermore support the notion that HMO-utilisation abilities of bifidobacteria, particularly with regards to fucosyllactose and sialyllactose, impact on their establishment in the breastfed infant gut microbiota. Recent work has indeed demonstrated that variations in HMO composition influence the overall gut microbiota structure of infants [41]. Individual HMO components, or combinations thereof, may be used in the future as supplements or therapeutics to formula-fed infants, to aid in the development of a normal gut microbiota by encouraging the establishment of key infant-associated *Bifidobacterium* species.

4.6 Acknowledgements

The authors would like to sincerely thank Glycom A/S (Lyngby, Denmark) for the provision of purified HMO samples used in this study under their

donation program. Additionally, the authors would like to sincerely thank all those involved in the InfantMet Project for the provision of infant faecal sampled for the screening of bifidobacteria. This study was funded in part by the Irish Research Council, under the Postgraduate Research Project Award; Project ID GOIPG/2013/651. In addition, the authors are supported by Science Foundation Ireland (SFI) (Grant No. SFI/12/RC/2273).

4.7 Tables and Figures

Table 4.1. Total number of isolates of *Bifidobacterium* belonging to each species obtained during the screening of infant faeces. The number of isolates belonging to each species, and total number of isolates obtained are given for each carbohydrate substrate used, as well as the overall number of isolates of each species.

Carbohydrate	*B. breve*	*B. bifidum*	*B. kashiwanohense*	*B. longum spp.*	*B. pseudocatenulatum*	Total Isolates
Lactose	110	5	0	128	9	252
TOS	60	32	0	48	17	157
HMO*	7	71	7	1	7	93
Total	177	108	7	177	33	502

*HMO substrates used were 3-SL, 2-FL or 3-FL

Table 4.2. Number of isolates of *Bifidobacterium* belonging to each species obtained from each individual faecal sample. The number of isolates belonging to each species, and total number of isolates obtained are given for each carbohydrate substrate used, as well as the overall number of isolates of each species.

Sample	Carbohydrate	*B. breve*	*B. bifidum*	*B. kashiwanohense*	*B. longum spp.*	*B. pseudocatenulatum*	Total Bifidobacteria
091W4	Lactose	2	0	0	6	8	16
	TOS	0	0	0	11	12	23
	Total	**2**	**0**	**0**	**17**	**20**	**39**
095W4	Lactose	0	0	0	23	0	23
	TOS	0	0	0	4	0	4
	Total	**0**	**0**	**0**	**27**	**0**	**27**
180W8	Lactose	1	0	0	23	0	24
	TOS	0	0	0	4	0	4
	Total	**1**	**0**	**0**	**27**	**0**	**28**
082W4	Lactose	29	0	0	0	0	29
	TOS	3	0	0	0	0	3
	Total	**32**	**0**	**0**	**0**	**0**	**32**
200W4	Lactose	22	0	0	0	0	22
	TOS	8	0	0	0	0	8
	Total	**30**	**0**	**0**	**0**	**0**	**30**
144W4	Lactose	27	3	0	0	1	31
	TOS	6	0	0	0	0	6
	Total	**33**	**3**	**0**	**0**	**1**	**37**
215W4	Lactose	5	0	0	0	0	5
	TOS	8	0	0	0	1	9
	Total	**13**	**0**	**0**	**0**	**1**	**14**
060W4	Lactose	0	0	0	67	0	67
	Total	**0**	**0**	**0**	**67**	**0**	**67**
017W4	3-SL	5	32	0	0	1	38
	TOS	27	9	0	0	0	36
	Total	**32**	**41**	**0**	**0**	**1**	**74**
131W4	2-FL	0	0	7	0	0	7
	Total	**0**	**0**	**7**	**0**	**0**	**7**
139W4	Lactose	19	2	0	4	0	25
	Total	**19**	**2**	**0**	**4**	**0**	**25**
165W4	Lactose	5	0	0	5	0	10
	Total	**5**	**0**	**0**	**5**	**0**	**10**
003W4	2-FL	2	0	0	0	3	5
	TOS	8	0	0	0	0	8
	Total	**10**	**0**	**0**	**0**	**3**	**13**
005W4	3-FL	0	7	0	0	1	8
	3-SL	0	2	0	0	0	2
	TOS	0	12	0	2	0	14
	Total	**0**	**21**	**0**	**2**	**1**	**24**
007W4	2-FL	0	19	0	1	0	20
	TOS	0	7	0	9	4	20
	Total	**0**	**26**	**0**	**10**	**4**	**40**
061W4	2-FL	0	0	0	0	2	2
	TOS	0	0	0	18	0	18
	Total	**0**	**0**	**0**	**18**	**2**	**20**
070W4	2-FL	0	11	0	0	0	11
	TOS	0	4	0	0	0	4
	Total	**0**	**15**	**0**	**0**	**0**	**15**

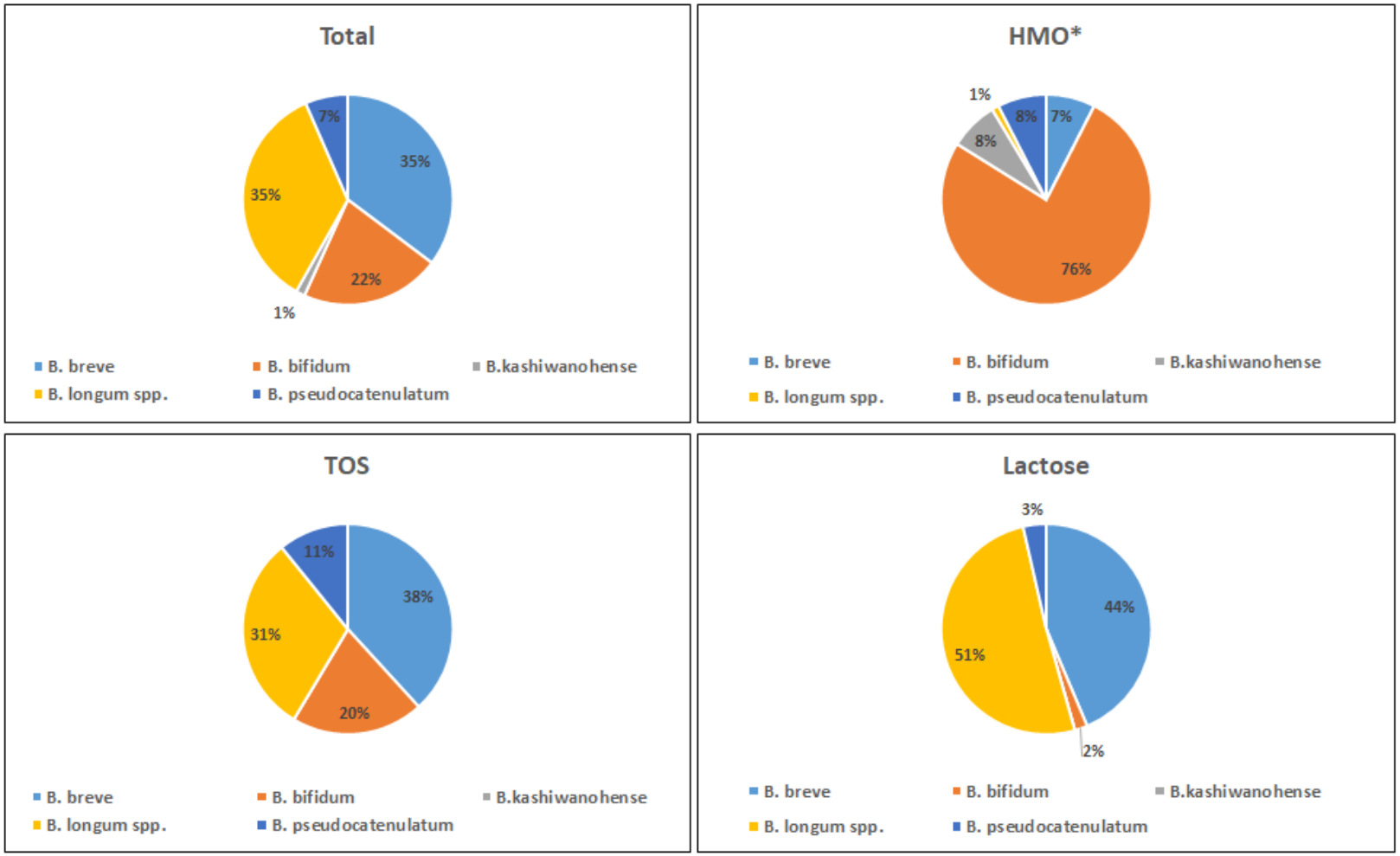


Figure 4.1. Propotions of each isolated *Bifidobacterium* species obtained in total during the screening of infant faecal samples, as well as for each selective carbohydrate source used.

*HMO includes samples screened using 2-FL, 3-FL or 3-SL as the selective carbohydrate source.

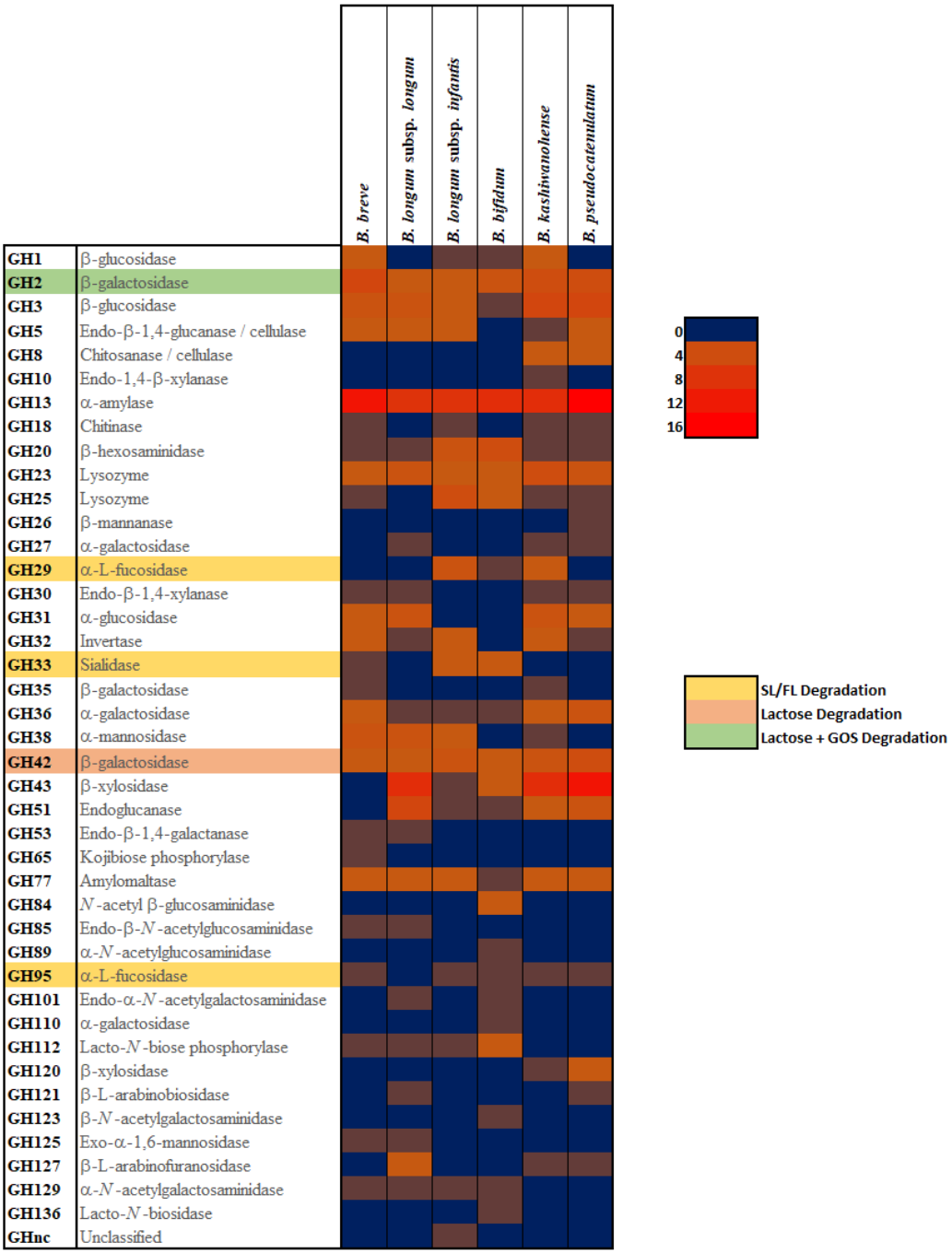


Figure 4.2. Heatmap representing the distribution of encoded glycosyl hydrolases, belonging to each GH family, across the genomes of *B. breve* UCC2003, *B. longum* subsp. *longum* NCC2705, *B. longum* subsp. *infantis* ATCC 15697, *B. bifidum* PRL2010, *B. kashiwanohense* JCM 15439 and *B. pseudocatenulatum* DSM 20438. A code colour grading is employed, which represents the number of glycosyl hydrolases of the same family encoded in each genome. Glycosyl hydrolase families which are known to function in the degradation of the HMO's 2-FL, 3-FL or 3-SL, lactose and/or GOS are indicated by shading.

4.8 Supplemental Material

Supplemental Table S4.1. Oligonucleotide primers used in this work.

Purpose	Primer	Sequence (5'-3')
Amplification of the ITS region for *Bifidobacterium* isolate species identification	Bifspp 23Sbif	ggtgtgaaagtccatcgct gtctgccaaggcatccacca
Amplification of the ruvAB region for *Bifidobacterium* isolate species identification	ruvabf ruvabr	ygcncagaaratcatyct agrtcrtgsacdatvgccc

4.9 References

1. **Turroni, F., C. Peano, D.A. Pass, E. Foroni, M. Severgnini, M.J. Claesson, C. Kerr, J. Hourihane, D. Murray, F. Fuligni, M. Gueimonde, A. Margolles, G. De Bellis, P.W. O'Toole, D. van Sinderen, J.R. Marchesi, and M. Ventura,** *Diversity of Bifidobacteria within the Infant Gut Microbiota.* PLoS ONE, **2012**. **7**(5): p. e36957.
2. **Bunesova, V., C. Lacroix, and C. Schwab,** *Fucosyllactose and L-fucose utilization of infant Bifidobacterium longum and Bifidobacterium kashiwanohense.* BMC Microbiology, **2016**. **16**(1): p. 248.
3. **Simone, M., C. Gozzoli, A. Quartieri, G. Mazzola, D. Di Gioia, A. Amaretti, S. Raimondi, and M. Rossi,** *The Probiotic Bifidobacterium breve B632 Inhibited the Growth of Enterobacteriaceae within Colicky Infant Microbiota Cultures.* BioMed Research International, **2014**. **2014**(301053**)**: p. 1-7.
4. **Di Gioia, D., I. Aloisio, G. Mazzola, and B. Biavati,** *Bifidobacteria: their impact on gut microbiota composition and their applications as probiotics in infants.* Appl Microbiol Biotechnol, **2014**. **98**(2): p. 563-77.
5. **Charbonneau, Mark R., D. O'Donnell, Laura V. Blanton, Sarah M. Totten, Jasmine C.C. Davis, Michael J. Barratt, J. Cheng, J. Guruge, M. Talcott, James R. Bain, Michael J. Muehlbauer, O. Ilkayeva, C. Wu, T. Struckmeyer, D. Barile, C. Mangani, J. Jorgensen, Y.-m. Fan, K. Maleta, Kathryn G. Dewey, P. Ashorn, Christopher B. Newgard, C. Lebrilla, David A. Mills, and Jeffrey I. Gordon,** *Sialylated Milk Oligosaccharides Promote Microbiota-Dependent Growth in Models of Infant Undernutrition.* Cell. **164**(5): p. 859-871.
6. **Milani, C., S. Duranti, F. Bottacini, E. Casey, F. Turroni, J. Mahony, C. Belzer, S. Delgado Palacio, S. Arboleya Montes, L. Mancabelli, G.A. Lugli, J.M. Rodriguez, L. Bode, W. de Vos, M. Gueimonde, A. Margolles, D. van Sinderen, and M. Ventura,** *The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota.* Microbiology and Molecular Biology Reviews, **2017**. **81**(4): p. e00036-17.
7. **Good, M., C.P. Sodhi, Y. Yamaguchi, H. Jia, P. Lu, W.B. Fulton, L.Y. Martin, T. Prindle, D.F. Nino, Q. Zhou, C. Ma, J.A. Ozolek, R.H. Buck, K.C. Goehring, and D.J. Hackam,** *The human milk oligosaccharide 2'-fucosyllactose attenuates the severity of experimental necrotising enterocolitis by enhancing mesenteric perfusion in the neonatal intestine.* Br J Nutr, **2016**. **116**(7): p. 1175-1187.

8. **Bode, L.,** *Human milk oligosaccharides: every baby needs a sugar mama.* Glycobiology, **2012**. **22**(9): p. 1147-62.
9. **Engfer, M.B., B. Stahl, B. Finke, G. Sawatzki, and H. Daniel,** *Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract.* The American Journal of Clinical Nutrition, **2000**. **71**(6): p. 1589-1596.
10. **Kunz, C., S. Rudloff, W. Baier, N. Klein, and S. Strobel,** *OLIGOSACCHARIDES IN HUMAN MILK: Structural, Functional, and Metabolic Aspects.* Annual Review of Nutrition, **2000**. **20**(1): p. 699-722.
11. **Brüssow, H.,** *Human microbiota: 'The philosophers have only interpreted the world in various ways. The point, however, is to change it'.* Microbial Biotechnology, **2015**. **8**(1): p. 11-12.
12. **Bezirtzoglou, E., A. Tsiotsias, and G.W. Welling,** *Microbiota profile in feces of breast- and formula-fed newborns by using fluorescence in situ hybridization (FISH).* Anaerobe, **2011**. **17**(6): p. 478-82.
13. **Sela, D.A. and D.A. Mills,** *Nursing our microbiota: molecular linkages between bifidobacteria and milk oligosaccharides.* Trends Microbiol, **2010**. **18**(7): p. 298-307.
14. **Asakuma, S., E. Hatakeyama, T. Urashima, E. Yoshida, T. Katayama, K. Yamamoto, H. Kumagai, H. Ashida, J. Hirose, and M. Kitaoka,** *Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria.* J Biol Chem, **2011**. **286**(40): p. 34583-92.
15. **Sela, D.A.,** *Bifidobacterial utilization of human milk oligosaccharides.* International Journal of Food Microbiology, **2011**. **149**(1): p. 58-64.
16. **Ward, R.E., M. Ninonuevo, D.A. Mills, C.B. Lebrilla, and J.B. German,** *In vitro fermentability of human milk oligosaccharides by several strains of bifidobacteria.* Mol Nutr Food Res, **2007**. **51**(11): p. 1398-405.
17. **LoCascio, R.G., M.R. Ninonuevo, S.L. Freeman, D.A. Sela, R. Grimm, C.B. Lebrilla, D.A. Mills, and J.B. German,** *Glycoprofiling of Bifidobacterial Consumption of Human Milk Oligosaccharides Demonstrates Strain Specific, Preferential Consumption of Small Chain Glycans Secreted in Early Human Lactation.* Journal of Agricultural and Food Chemistry, **2007**. **55**(22): p. 8914-8919.
18. **Benno, Y., K. Sawada, and T. Mitsuoka,** *The intestinal microflora of infants: composition of fecal flora in breast-fed and bottle-fed infants.* Microbiol Immunol, **1984**. **28**(9): p. 975-86.

19. **Guaraldi, F. and G. Salvatori,** *Effect of Breast and Formula Feeding on Gut Microbiota Shaping in Newborns.* Frontiers in Cellular and Infection Microbiology, **2012**. **2**(94): p. 94.
20. **Solís, G., C.G. Los Reyes-Gavilan, N. Fernández, A. Margolles, and M. Gueimonde,** *Establishment and development of lactic acid bacteria and bifidobacteria microbiota in breast-milk and the infant gut.* Anaerobe, **2010**. **16**(3): p. 307-10.
21. **Morita, H., A. Nakano, H. Onoda, H. Toh, K. Oshima, H. Takami, M. Murakami, S. Fukuda, T. Takizawa, T. Kuwahara, H. Ohno, S. Tanabe, and M. Hattori,** *Bifidobacterium kashiwanohense sp. nov., isolated from healthy infant faeces.* Int J Syst Evol Microbiol, **2011**. **61**(Pt 11): p. 2610-5.
22. **Vazquez-Gutierrez, P., C. Lacroix, T. Jaeggi, C. Zeder, M.B. Zimmerman, and C. Chassard,** *Bifidobacteria strains isolated from stools of iron deficient infants can efficiently sequester iron.* BMC Microbiol, **2015**. **15**(1): p. 3.
23. **Kitaoka, M.,** *Bifidobacterial enzymes involved in the metabolism of human milk oligosaccharides.* Adv Nutr, **2012**. **3**(3): p. 422s-9s.
24. **Wada, J., T. Ando, M. Kiyohara, H. Ashida, M. Kitaoka, M. Yamaguchi, H. Kumagai, T. Katayama, and K. Yamamoto,** *Bifidobacterium bifidum lacto-N-biosidase, a critical enzyme for the degradation of human milk oligosaccharides with a type 1 structure.* Appl Environ Microbiol, **2008**. **74**(13): p. 3996-4004.
25. **Moller, P.L., F. Jorgensen, O.C. Hansen, S.M. Madsen, and P. Stougaard,** *Intra- and extracellular beta-galactosidases from Bifidobacterium bifidum and B. infantis: molecular cloning, heterologous expression, and comparative characterization.* Appl Environ Microbiol, **2001**. **67**(5): p. 2276-83.
26. **Miwa, M., T. Horimoto, M. Kiyohara, T. Katayama, M. Kitaoka, H. Ashida, and K. Yamamoto,** *Cooperation of beta-galactosidase and beta-N-acetylhexosaminidase from bifidobacteria in assimilation of human milk oligosaccharides with type 2 structure.* Glycobiology, **2010**. **20**(11): p. 1402-9.
27. **Ashida, H., A. Miyake, M. Kiyohara, J. Wada, E. Yoshida, H. Kumagai, T. Katayama, and K. Yamamoto,** *Two distinct alpha-L-fucosidases from Bifidobacterium bifidum are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates.* Glycobiology, **2009**. **19**(9): p. 1010-7.
28. **Kiyohara, M., K. Tanigawa, T. Chaiwangsri, T. Katayama, H. Ashida, and K. Yamamoto,** *An exo-alpha-sialidase from bifidobacteria involved in the degradation of*

sialyloligosaccharides in human milk and intestinal glycoconjugates. Glycobiology, **2011**. **21**(4): p. 437-47.

29. **Garrido, D., S. Ruiz-Moyano, N. Kirmiz, J.C. Davis, S.M. Totten, D.G. Lemay, J.A. Ugalde, J.B. German, C.B. Lebrilla, and D.A. Mills,** *A novel gene cluster allows preferential utilization of fucosylated milk oligosaccharides in Bifidobacterium longum subsp. longum SC596.* Sci Rep, **2016**. **6**: p. 35045.
30. **Sakurama, H., M. Kiyohara, J. Wada, Y. Honda, M. Yamaguchi, S. Fukiya, A. Yokota, H. Ashida, H. Kumagai, M. Kitaoka, K. Yamamoto, and T. Katayama,** *Lacto-N-biosidase Encoded by a Novel Gene of Bifidobacterium longum Subspecies longum Shows Unique Substrate Specificity and Requires a Designated Chaperone for Its Active Expression.* The Journal of Biological Chemistry, **2013**. **288**(35): p. 25194-25206.
31. **Sela, D.A., J. Chapman, A. Adeuya, J.H. Kim, F. Chen, T.R. Whitehead, A. Lapidus, D.S. Rokhsar, C.B. Lebrilla, J.B. German, N.P. Price, P.M. Richardson, and D.A. Mills,** *The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome.* Proc Natl Acad Sci U S A, **2008**. **105**(48): p. 18964-9.
32. **Yoshida, E., H. Sakurama, M. Kiyohara, M. Nakajima, M. Kitaoka, H. Ashida, J. Hirose, T. Katayama, K. Yamamoto, and H. Kumagai,** *Bifidobacterium longum subsp. infantis uses two different beta-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides.* Glycobiology, **2012**. **22**(3): p. 361-8.
33. **Garrido, D., S. Ruiz-Moyano, and D.A. Mills,** *Release and utilization of N-acetyl-D-glucosamine from human milk oligosaccharides by Bifidobacterium longum subsp. infantis.* Anaerobe, **2012**. **18**(4): p. 430-5.
34. **Xiao, J.Z., S. Takahashi, M. Nishimoto, T. Odamaki, T. Yaeshima, K. Iwatsuki, and M. Kitaoka,** *Distribution of in vitro fermentation ability of lacto-N-biose I, a major building block of human milk oligosaccharides, in bifidobacterial strains.* Appl Environ Microbiol, **2010**. **76**(1): p. 54-9.
35. **James, K., M.O. Motherway, F. Bottacini, and D. van Sinderen,** *Bifidobacterium breve UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways.* Sci Rep, **2016**. **6**: p. 38560.

36. **Garrido, D., J.H. Kim, J.B. German, H.E. Raybould, and D.A. Mills,** *Oligosaccharide Binding Proteins from Bifidobacterium longum subsp. infantis Reveal a Preference for Host Glycans.* PLoS ONE, **2011**. **6**(3): p. e17315.

37. **Sela, D.A., D. Garrido, L. Lerno, S. Wu, K. Tan, H.-J. Eom, A. Joachimiak, C.B. Lebrilla, and D.A. Mills,** *Bifidobacterium longum subsp. infantis ATCC 15697 α-Fucosidases Are Active on Fucosylated Human Milk Oligosaccharides.* Applied and Environmental Microbiology, **2012**. **78**(3): p. 795-803.

38. **Sela, D.A., Y. Li, L. Lerno, S. Wu, A.M. Marcobal, J.B. German, X. Chen, C.B. Lebrilla, and D.A. Mills,** *An Infant-associated Bacterial Commensal Utilizes Breast Milk Sialyloligosaccharides.* The Journal of Biological Chemistry, **2011**. **286**(14): p. 11909-11918.

39. **Schwab, C., H.-J. Ruscheweyh, V. Bunesova, V.T. Pham, N. Beerenwinkel, and C. Lacroix,** *Trophic Interactions of Infant Bifidobacteria and Eubacterium hallii during L-Fucose and Fucosyllactose Degradation.* Frontiers in Microbiology, **2017**. **8**: p. 95.

40. **Duranti, S., G.A. Lugli, L. Mancabelli, F. Armanini, F. Turroni, K. James, P. Ferretti, V. Gorfer, C. Ferrario, C. Milani, M. Mangifesta, R. Anzalone, M. Zolfo, A. Viappiani, E. Pasolli, I. Bariletti, R. Canto, R. Clementi, M. Cologna, T. Crifo, G. Cusumano, S. Fedi, S. Gottardi, C. Innamorati, C. Mase, D. Postai, D. Savoi, M. Soffiati, S. Tateo, A. Pedrotti, N. Segata, D. van Sinderen, and M. Ventura,** *Maternal inheritance of bifidobacterial communities and bifidophages in infants through vertical transmission.* Microbiome, **2017**. **5**(1): p. 66.

41. **Wang, M., M. Li, S. Wu, C.B. Lebrilla, R.S. Chapkin, I. Ivanov, and S.M. Donovan,** *Fecal Microbiota Composition of Breast-fed Infants is Correlated with Human Milk Oligosaccharides Consumed.* Journal of pediatric gastroenterology and nutrition, **2015**. **60**(6): p. 825-833.

42. **Pokusaeva, K., G.F. Fitzgerald, and D. van Sinderen,** *Carbohydrate metabolism in Bifidobacteria.* Genes & Nutrition, **2011**. **6**(3): p. 285-306.

43. **Roberfroid, M., G.R. Gibson, L. Hoyles, A.L. McCartney, R. Rastall, I. Rowland, D. Wolvers, B. Watzl, H. Szajewska, B. Stahl, F. Guarner, F. Respondek, K. Whelan, V. Coxam, M.-J. Davicco, L. Léotoing, Y. Wittrant, N.M. Delzenne, P.D. Cani, A.M. Neyrinck, and A. Meheust,** *Prebiotic effects: metabolic and health benefits.* British Journal of Nutrition, **2010**. **104**(S2): p. S1-S63.

44. **Garrido, D., S. Ruiz-Moyano, R. Jimenez-Espinoza, H.-J. Eom, D.E. Block, and D.A. Mills,** *Utilization of galactooligosaccharides by Bifidobacterium longum subsp. infantis isolates.* Food microbiology, **2013**. **33**(2): p. 262-270.

45. **Macfarlane, G.T., H. Steed, and S. Macfarlane,** *Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics.* J Appl Microbiol, **2008**. **104**(2): p. 305-44.

46. **Rijnierse, A., P.V. Jeurink, B.C. van Esch, J. Garssen, and L.M. Knippels,** *Food-derived oligosaccharides exhibit pharmaceutical properties.* Eur J Pharmacol, **2011**. **668 Suppl 1**: p. S117-23.

47. **Hill, C.J., D.B. Lynch, K. Murphy, M. Ulaszewska, I.B. Jeffery, C.A. O'Shea, C. Watkins, E. Dempsey, F. Mattivi, K. Tuohy, R.P. Ross, C.A. Ryan, P.W. O' Toole, and C. Stanton,** *Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort.* Microbiome, **2017**. **5**(1): p. 4.

48. **Watson, D., M. O'Connell Motherway, M.H. Schoterman, R.J. van Neerven, A. Nauta, and D. van Sinderen,** *Selective carbohydrate utilization by lactobacilli and bifidobacteria.* J Appl Microbiol, **2013**. **114**(4): p. 1132-46.

49. **Scardovi, V.**, *Genus Bifidobacterium,* in *Bergey's manual of systematic bacteriology,* P. H. A. Sneath and M.E.S. N. S. Mair, and J. G. Holt, Editors. **1986**, Williams and Wilkins: Baltimore, Md. p. 1418–1434.

50. **Turroni, F., E. Foroni, P. Pizzetti, V. Giubellini, A. Ribbera, P. Merusi, P. Cagnasso, B. Bizzarri, G.L. de'Angelis, F. Shanahan, D. van Sinderen, and M. Ventura,** *Exploring the Diversity of the Bifidobacterial Population in the Human Intestinal Tract.* Applied and Environmental Microbiology, **2009**. **75**(6): p. 1534-1545.

51. **O'Connell Motherway, M., A. Zomer, S.C. Leahy, J. Reunanen, F. Bottacini, M.J. Claesson, F. O'Brien, K. Flynn, P.G. Casey, J.A. Moreno Munoz, B. Kearney, A.M. Houston, C. O'Mahony, D.G. Higgins, F. Shanahan, A. Palva, W.M. de Vos, G.F. Fitzgerald, M. Ventura, P.W. O'Toole, and D. van Sinderen,** *Functional genome analysis of Bifidobacterium breve UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor.* Proceedings of the National Academy of Sciences, **2011**. **108**(27): p. 11217-11222.

52. **Schell, M.A., M. Karmirantzou, B. Snel, D. Vilanova, B. Berger, G. Pessi, M.-C. Zwahlen, F. Desiere, P. Bork, M. Delley, R.D. Pridmore, and F. Arigoni,** *The genome sequence of Bifidobacterium longum reflects its adaptation to the human*

gastrointestinal tract. Proceedings of the National Academy of Sciences, **2002**. **99**(22): p. 14422-14427.

53. **Turroni, F., F. Bottacini, E. Foroni, I. Mulder, J.-H. Kim, A. Zomer, B. Sánchez, A. Bidossi, A. Ferrarini, V. Giubellini, M. Delledonne, B. Henrissat, P. Coutinho, M. Oggioni, G.F. Fitzgerald, D. Mills, A. Margolles, D. Kelly, D. van Sinderen, and M. Ventura,** *Genome analysis of Bifidobacterium bifidum PRL2010 reveals metabolic pathways for host-derived glycan foraging.* Proceedings of the National Academy of Sciences of the United States of America, **2010**. **107**(45): p. 19514-19519.
54. **Morita, H., H. Toh, A. Nakano, K. Oshima, M. Takagi, W. Suda, S. Tanabe, and M. Hattori,** *Complete Genome Sequence of Bifidobacterium kashiwanohense JCM 15439(T), Isolated from Feces from a Healthy Japanese Infant.* Genome Announcements, **2015**. **3**(2): p. e00255-15.
55. **Morita, H., H. Toh, K. Oshima, A. Nakano, K. Arakawa, Y. Takayama, R. Kurokawa, K. Takanashi, K. Honda, and M. Hattori,** *Complete genome sequence of Bifidobacterium pseudocatenulatum JCM 1200T isolated from infant feces.* Journal of Biotechnology, **2015**. **210**: p. 68-69.
56. **Rada, V.,** *Detection of Bifidobacteriumspecies by enzymatic methods and antimicrobial susceptibility testing.* Biotechnology Techniques, **1997**. **11**(12): p. 909-912.
57. **Simpson, P.J., G.F. Fitzgerald, C. Stanton, and R.P. Ross,** *The evaluation of a mupirocin-based selective medium for the enumeration of bifidobacteria from probiotic animal feed.* J Microbiol Methods, **2004**. **57**(1): p. 9-16.
58. **Ghannoum, M.A. and L.B. Rice,** *Antifungal Agents: Mode of Action, Mechanisms of Resistance, and Correlation of These Mechanisms with Bacterial Resistance.* Clinical Microbiology Reviews, **1999**. **12**(4): p. 501-517.
59. **O'Connell Motherway, M., M. Kinsella, G.F. Fitzgerald, and D. van Sinderen,** *Transcriptional and functional characterization of genetic elements involved in galacto-oligosaccharide utilization by Bifidobacterium breve UCC2003.* Microb Biotechnol, **2013**. **6**(1): p. 67-79.
60. **Matsumoto, T., S. Shimada, Y. Hata, T. Tanaka, and A. Kondo,** *Multi-functional glycoside hydrolase: Blon_0625 from Bifidobacterium longum subsp. infantis ATCC 15697.* Enzyme and Microbial Technology, **2015**. **68**(Supplement C): p. 10-14.
61. **Lee, J.H., Y.J. Hyun, and D.H. Kim,** *Cloning and characterization of α-L-arabinofuranosidase and bifunctional α-L-arabinopyranosidase/β-D-*

galactopyranosidase from Bifidobacterium longum H-1. Journal of Applied Microbiology, **2011**. **111**(5): p. 1097-1107.

62. **Milani, C., F. Turroni, S. Duranti, G.A. Lugli, L. Mancabelli, C. Ferrario, D. van Sinderen, and M. Ventura,** *Genomics of the Genus Bifidobacterium Reveals Species-Specific Adaptation to the Glycan-Rich Gut Environment.* Applied and Environmental Microbiology, **2016**. **82**(4): p. 980-991.
63. **Kaoutari, A.E., F. Armougom, J.I. Gordon, D. Raoult, and B. Henrissat,** *The abundance and variety of carbohydrate-active enzymes in the human gut microbiota.* Nature Reviews Microbiology, **2013**. **11**(7): p. 497.
64. **Ueda, M., K. Ohata, T. Konishi, A. Sutrisno, H. Okada, M. Nakazawa, and K. Miyatake,** *A novel goose-type lysozyme gene with chitinolytic activity from the moderately thermophilic bacterium Ralstonia sp. A-471: cloning, sequencing, and expression.* Appl Microbiol Biotechnol, **2009**. **81**(6): p. 1077-85.
65. **Egan, M., M. O'Connell Motherway, M. Ventura, and D. van Sinderen,** *Metabolism of sialic acid by Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2014**. **80**(14): p. 4414-26.
66. **Egan, M., M.O. Motherway, M. Kilcoyne, M. Kane, L. Joshi, M. Ventura, and D. van Sinderen,** *Cross-feeding by Bifidobacterium breve UCC2003 during co-cultivation with Bifidobacterium bifidum PRL2010 in a mucin-based medium.* BMC Microbiol, **2014**. **14**(282): p. 282.
67. **Sela, D.A.,** *Bifidobacterial utilization of human milk oligosaccharides.* Int J Food Microbiol, **2011**. **149**(1): p. 58-64.
68. **Avershina, E., O. Storrø, T. Øien, R. Johnsen, R. Wilson, T. Egeland, and K. Rudi,** *Bifidobacterial succession and correlation networks in a large unselected cohort of mothers and their children.* Appl Environ Microbiol, **2013**. **79**.
69. **Thongaram, T., J.L. Hoeflinger, J. Chow, and M.J. Miller,** *Human milk oligosaccharide consumption by probiotic and human-associated bifidobacteria and lactobacilli.* J Dairy Sci, **2017**. **100**(10): p. 7825-7833.
70. **Vazquez-Gutierrez, P., C. Lacroix, C. Chassard, J. Klumpp, C. Jans, and M.J.A. Stevens,** *Complete and Assembled Genome Sequence of Bifidobacterium kashiwanohense PV20-2, Isolated from the Feces of an Anemic Kenyan Infant.* Genome Announcements, **2015**. **3**(1): p. e01467-14.

71. **Avershina, E., K. Lundgard, M. Sekelja, C. Dotterud, O. Storro, T. Oien, R. Johnsen, and K. Rudi,** *Transition from infant- to adult-like gut microbiota.* Environ Microbiol, **2016**. **18**(7): p. 2226-36.

72. **Avershina, E., O. Storro, T. Oien, R. Johnsen, R. Wilson, T. Egeland, and K. Rudi,** *Bifidobacterial succession and correlation networks in a large unselected cohort of mothers and their children.* Appl Environ Microbiol, **2013**. **79**(2): p. 497-507.

Chapter V

Molecular analysis of a gene cluster from *Bifidobacterium kashiwanohense* involved in the metabolism of fucosyllactose, a major human milk oligosaccharide component.

Mariane Vigoureux aided in the isolation of the strain APCKJ1.

Dr. Francesca Bottacini assisted with the genome assembly, annotation and methylome analysis.

Dr. Muireann Egan assisted in the design of the APCKJ1 microarray slides.

This chapter is in preparation as a manuscript as:

James, K., Bottacini, F., Vigoureux, M., Egan, M., O'Connell Motherway, M. & van Sinderen, D. Molecular dissection of a gene cluster from *Bifidobacterium kashiwanohense* involved in the metabolism of fucosyllactose, a major human milk oligosaccharide component. Manuscript in preparation, April, 2018.

5.1 Abstract

A number of bifidobacterial species are found at a particularly high prevalence and abundance in faecal samples of healthy breastfed infants, a phenomenon that is believed to be, at least partially, due to the ability of bifidobacteria to metabolize Human Milk Oligosaccharides (HMOs) as their sole carbohydrate substrate. *Bifidobacterium kashiwanohense* is a rather unexplored species that has occasionally been isolated from infant faeces. We isolated a novel strain of *B. kashiwanohense*, named APCKJ1, from the faeces of a breastfed infant based on its ability to utilise the HMO component fucosyllactose as its sole carbohydrate source. In the current study, we identified and annotated the genome sequence of this strain, which facilitated the analysis of fucosyllactose metabolism in *B. kashiwanohense* APCKJ1. Employing transcriptomic and growth analyses, combined with *in vitro* hydrolysis assays, and heterologous expression the pathway for fucosyllactose metabolism in this species/strain was elucidated. The generated findings furthermore allowed a wider perspective of bifidobacterial fucosyllactose and L-fucose metabolism.

5.2 Introduction

Breast-feeding is known to have an impact on the development and composition of the intestinal microbiota of the neonate [1-3], with significant differences reported between the gut microbiota of breast-fed and formula-fed infants [4]. It is currently believed that Human Milk Oligosaccharides (HMOs) are among the most influential components of human breastmilk and colostrum in terms of their impact on microbiota composition by acting as a selective growth substrate [2, 3, 5], with recent work correlating differences in the neonatal microbiota with variations in HMO composition of corresponding mother's milk [6]. HMOs are, after lactose, the most abundant carbohydrate component of breastmilk [2, 7], and are composed of a heterogeneous mix of at least 200 distinct glycan structures [8]. HMOs consist of a lactose molecule linked (through a variety of different glycosidic connections) to sialic acid, L-fucose, lacto-N-biose (LNB) or N-acetyllactosamine (LacNAc), in the latter two cases creating the tetrasaccharides lacto-N-tetraose (LNT) or lacto-N-neotetraose (LNnT), respectively. LN(n)T may in turn be elongated by additional LNB or LacNAc residues at the reducing end, and/or may be fucosylated or sialylated [2, 8, 9]. HMOs can be viewed as a model prebiotic, since it complies with the definition of the latter as 'a nondigestible food ingredient that beneficially affects the host by selectively stimulating the growth and/or activity of one or a limited number of bacteria in the colon, and thus improves host health' [10].

Bifidobacteria are Gram-positive, anaerobic members of the phylum Actinobacteria, and are common commensals of the mammalian, avian and insect gut. The faecal microbiota of healthy breastfed infants is enriched for bifidobacteria [11], which are believed to provide a range of health benefits to their (neonatal) host [5, 12-14]. The dominance of bifidobacteria in this niche is attributed to the presence of a small number of key species, which are capable of selectively utilising (particular) HMOs as their sole carbohydrate source [5, 8, 15, 16]. These infant-associated species include *Bifidobacterium bifidum*, *Bifidobacterium longum* subsp. *infantis* and *Bifidobacterium breve*. *B. bifidum* carries out extracellular hydrolysis of complex HMOs, including LN(n)T, and fucosylated and sialylated structures,

prior to the import and metabolism of (many of) the resultant mono- and di-saccharides [17-23]. *B. breve* and *B. longum* subsp. *infantis* internalise (certain) intact HMO moieties, such as LNT, LNnT and LNB, and sequentially degrade them for subsequent metabolism of the generated monosaccharides [8, 17, 24-28]. Furthermore, *B. infantis* internalizes and internally metabolizes fucosyl- and sialyl-HMO, including sialyl-/fucosyl-lactose (which is in contrast to *B. bifidum*, which degrades these HMOs extracellularly) [29-31].

Bifidobacterium kashiwanohense is a species that has previously been isolated from infant faeces [32, 33], though it is not considered a common member of the neonatal gut microbiota. Very little is known about this species and its metabolic capabilities. The ability of two *B. kashiwanohense* strains to grow on the HMO components 2'-fucosyllactose (2'-FL) and 3-fucosyllactose (3-FL) was recently demonstrated, although these strains did not appear to metabolize (the fucosyllactose component) L-fucose [31]. While the presence of fucosylated HMOs in breastmilk has been shown to vary greatly, depending on genetic and geographical factors [34], fucosylated HMOs are generally thought to represent ~50-80 % of all HMOs found in breastmilk of healthy women [35]. Furthermore, free 2'-FL and 3-FL constitute between 12-45 % and between 0.5-3 % of the total HMO content, respectively [34]. 2'-FL (Fucα1-2[Galβ1-4]Glc) consists of L-fucose linked by an α1-2 bond to the galactose residue in lactose, while its isomer 3-FL (Fucα1-3Galβ1-4Glc) consists of L-fucose linked by an α1-3 bond to the glucose moiety in lactose [36]. The adaptation to grow on 2'-FL and/or 3-FL is therefore considered to be an important factor in infant gut microbiota development [37], representing an important trait of bifidobacterial species such as *B. kashiwanohense* since it supports their ability to establish themselves as part of the infant microbiota. Metabolite formation profiles for *B. kashiwanohense* following growth on 2'-FL or 3-FL suggest that the inability of *B. kashiwanohense* to grow on L-fucose is due to the lack of two enzymes, L-fucose mutarotase and fucose permease, involved in the metabolism of this sugar [31]. The latter study, and another recent study [38], however, demonstrated L-fucose metabolism during growth on

fucosyllactose by *B. breve*, *B. longum* subsp. *infantis* and *B. longum* subsp. *suis*, suggesting the presence of a common pathway for L-fucose metabolism among these species, and apparently not being functional in *B. kashiwanohense*.

In the current study, we isolated and sequenced the genome of a novel *B. kashiwanohense* strain, APCKJ1, from faeces of an exclusively breastfed infant, using 2'-FL as a selective carbohydrate. Transcriptomic and comparative genome analysis revealed two adjacent genetic loci involved in fucosyllactose metabolism in APCKJ1, one of which encodes 2 distinct α-fucosidases. Heterologous expression of and *in vitro* hydrolysis assays using these fucosidases revealed their ability to hydrolyse 2'-FL and 3-FL with differing specificities. We furthermore introduced the *B. kashiwanohense* APCKJ1 genes encoding one of these two α-fucosidases, as well as associated genes predicted to encode a fucosyllactose transport system, into *B. breve* UCC2003. The resulting recombinant strain was shown to utilise 2'-FL, and was transcriptomically analysed during growth on these substrates, identifying key genes involved in L-fucose metabolism in *B. breve* UCC2003. This work has not only elucidated the mechanism of fucosyllactose metabolism in *B. kashiwanohense*, but also suggests the presence of an L-fucose metabolism pathway in APCKJ1, revealing general insights into bifidobacterial L-fucose metabolism.

5.3 Materials and Methods

Strain isolation from infant faeces, and species identification. Pure cultures of *Bifidobacterium* strains, capable of utilising 2'-FL or 3-FL as their sole carbohydrate source, were isolated from the faeces of exclusively breastfed neonates, as described in Chapter IV. Individual isolates were confirmed as bifidobacteria based on the fructose-6-phosphate phosphoketolase activity test [39], and their associated species identity was determined by sequencing of their ITS region [40] (see Chapter IV).

Genome sequencing, annotation and methylome analysis of novel isolate *Bifidobacterium kashiwanohense* APCKJ1. Genome sequencing of the *B. kashiwanohense* isolate was performed by GATC Biotech Ltd. (Germany) using Pacific Biosciences SMRT RSII technology. Raw sequencing reads were *de novo* assembled using the Hierarchical Genome Assembly Process (HGAP) protocol RS_Assembly.2 implemented in the SMRT Smart Analysis portal v.2.3 with default parameters (https://github.com/PacificBiosciences/SMRT-Analysis).

Open Reading Frame (ORF) prediction and automatic annotation was performed using Prodigal v2.0 (http://prodigal.ornl.gov) for gene predictions, BLASTP v2.2.26 (cut-off e-value of 0.0001) [41] for sequence alignments against a combined bifidobacterial genome-based database, and MySQL relational database to assign annotations. Predicted functional assignments were manually revised and edited using similarity searches against the non-redundant protein database curated by the National Centre for Biotechnology Information (ftp://ftp.ncbi.nih.gov/BLAST/db/) and PFAM database (http://pfam.sanger.ac.uk), which allowed a more detailed, *in silico* characterization of hypothetical proteins. GenBank editing and manual inspection was performed using Artemis v18 (http://www.sanger.ac.uk/resources/soft-ware/artemis/). Transfer RNA genes were identified employing tRNAscan-SE v1.4 and ribosomal RNA genes were detected based on the software package Rnammer v1.2 [42] supported by BLASTN v2.2.26.

Methylome analysis was performed using a combination of SMRT sequencing and comparative genome analysis. Base modification and methylated motif detection was performed employing the SMRT Analysis portal following *de novo* genome assembly using the "RS_Modification_and_Motif_Analysis.1" protocol. Methylation motifs with a score of 40 (corresponding to a p-value of 0.0001) or higher were considered specific and were taken for further analysis (https://github.com/PacificBiosciences/Bioinformatics-Training/wiki/Methylome-Analysis-Technical-Note).

The deduced protein products of all identified ORFs were searched for similarity to known R/M systems (including orphan methylases) using BLASTP [41] alignments against the REBASE database (http://rebase.neb.com/rebase/rebase.html) [43]. Significant BLASTP hits were selected using a cut-off E-value of <0.0001 and exhibiting over 30 % similarity across at least 80 % of the sequence length. Further manual refinement of these predictions included discarding of false positive BLASTP hits and refinement using PFAM database (http://pfam.sanger.ac.uk).

Nucleotide sequence analysis. Sequence data were obtained from the Artemis-mediated [44] genome annotations of *B. kashiwanohense* APCKJ1, *B. kashiwanohense* JCM15439 [45], *B. kashiwanohense* PV20-2 [46] and *B. breve* UCC2003 [47]. Database searches were performed using non-redundant sequences accessible at the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) using the basic local alignment search tool (BLAST) [41, 48]. Sequences were verified and analysed using the SeqMan and SeqBuilder programs of the DNAStar software package (version 10.1.2; DNAStar, Madison, WI, USA).

Multiple sequence alignment. Multiple sequence alignments of amino acid sequences were created with the online resource Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/), using the default settings [49, 50]. The generated alignment was visualised with Genedoc v 2.7.0 [51], using

physiochemical display mode (amino acids coloured based on shared physical and chemical properties).

Bacterial strains, plasmids, culture conditions and *Bifidobacterium* growth assays. Bacterial strains and plasmids used in this study are listed in Supplemental Table S5.2. *B. breve* and *B. kashiwanohense* cultures were incubated under anaerobic conditions in a modular atmosphere-controlled system (Davidson and Hardy, Belfast, Ireland) at 37°C. *B. breve* UCC2003 was routinely cultured in either de Man Rogosa and Sharpe medium (MRS medium; Difco, BD, Le Pont de Claix, France) supplemented with 0.05 % cysteine-HCl and 1 % lactose, or reinforced clostridial medium (RCM; Oxoid Ltd., Basingstoke, England). *B. kashiwanohense* APCKJ1 was routinely cultured in modified de Man Rogosa and Sharpe (mMRS) supplemented with 0.05 % cysteine-HCl and 1 % lactose. *Lactococcus lactis* strains were cultivated in M17 broth (Oxoid Ltd., Basingstoke, England) containing 0.5 % glucose [52] at 30°C. *Escherichia coli* strains were cultured in Luria-Bertani (LB) broth [53] at 37°C with agitation. Where appropriate, growth media contained chloramphenicol (Cm; 5 μg ml^{-1} for *L. lactis* and *E. coli*, 2.5 μg ml^{-1} for *B. breve*) or streptomycin (Strep; 400 μg ml^{-1}).

Carbohydrate utilization by bifidobacterial strains was examined in modified de Man Rogosa and Sharpe (mMRS) medium prepared from first principles [54], and excluding a carbohydrate source. Prior to inoculation, the mMRS medium was supplemented with cysteine-HCl (0.05 %, wt/vol) and a particular carbohydrate source (1 %, wt/vol). It has previously been shown that mMRS does not support growth of *B. breve* UCC2003 in the absence of an added carbohydrate [54]. Carbohydrates used were lactose, glucose, D-ribose, sorbitol, cellobiose, raffinose, melibiose (obtained from Sigma Aldrich, Steinheim, Germany), and LNT, LNnT, lactosamine-hydrochloride (lactosamine HCl), N-acetylneuramic acid (sialic acid), L-fucose, 3-sialyllactose (3-SL), 6-sialyllactose (6-SL), 2'-FL and 3-FL (obtained from Glycom, Lyngby, Denmark). A 1 % wt/vol concentration of carbohydrate was considered sufficient to analyse the growth capabilities of a strain on a

particular carbon source. Addition of these carbohydrates did not significantly alter the pH of the medium, and therefore subsequent pH adjustment was not required. In order to determine bacterial growth profiles and final optical densities, 5 ml of freshly prepared mMRS medium, including a particular carbohydrate (see above), was inoculated with 50 µl (1 %) of a stationary phase culture of *B. breve* UCC2003 or *B. kashiwanohense* APCKJ1. Uninoculated mMRS medium was used as a negative control. Cultures were incubated anaerobically at 37°C for 24h, and the optical density at 600 nm (OD_{600nm}) was determined manually. Growth profiles of *B. kashiwanohense* APCKJ1, wild type *B. breve* UCC2003 and recombinant *B. breve* UCC2003 were determined manually in this fashion.

HPLC analysis. HPLC analysis was used to identify the fermentation product profiles following the growth of *B. kashiwanohense* APCKJ1 in media containing 2'-FL, 3-FL or lactose as a sole carbohydrate source. All strains were cultivated on mMRS medium supplemented with 1 % 2'-FL, 3-FL or lactose. The collected samples were then prepared for HPLC analysis by centrifugation at 5000 rpm for 5 min, the resulting supernatants were filter sterilized (0.45 µm filter, Costar Spin-X Column) and stored at −20°C prior to analysis. An Agilent 1200 HPLC system (Agilent Technologies, Santa Clara, CA) with a refractive index detector was used to the metabolites generated as a result of each carbohydrate fermentation. Metabolite peaks were identified based on 2'-FL, 3-FL, L-fucose (Glycom, Lyngby, Denmark), lactose, lactic acid, acetic acid, formic acid, sodium pyruvate and 1,2-propanediol (Sigma Aldrich, Steinheim, Germany) retention times of known standards at a concentration of 10 mM. Non-fermented mMRS medium containing carbohydrates (as well as without any carbohydrate source added) served as controls. A REXEX 8 µ 8 % H organic acid column (300 mm × 7.8 mm, Phenomenex, Torrance, CA, USA) was utilized and maintained at 65°C. Elution was performed for 25 min using a 0.01 M H_2SO_4 solution at a constant flow rate of 0.6 mL/min.

Analysis of global gene expression using *B. breve* and *B. kashiwanohense* DNA microarrays. Global gene transcription was determined during log-phase growth of *B. kashiwanohense* APCKJ1 in mMRS supplemented with 2'-FL. The obtained transcriptome was compared to that determined for log-phase *B. kashiwanohense* APCKJ1 cells when grown in mMRS supplemented with sorbitol. The transcriptome for growth of APCKJ1 on 2'-FL, as compared to a lactose control, was also obtained by the same method. DNA microarrays containing oligonucleotide primers representing each of the 2110 identified open reading frames on the genome of *B. kashiwanohense* APCKJ1 were designed and obtained from Agilent Technologies (Palo Alto, Ca., USA).

Global gene expression was determined during log-phase growth of recombinant strain *B. breve* UCC2003-*fucA1fucSP1P2* in mMRS supplemented with either 2'-FL or lactose. The obtained transcriptome was compared to that determined for log-phase *B. breve* UCC2003-*fucA1fucSP1P2* cells when grown in mMRS supplemented with ribose.

DNA microarrays containing oligonucleotide primers representing each of the 1864 identified open reading frames on the genome of *B. breve* UCC2003, as well as separate DNA microarrays containing oligonucleotide primers representing each of the 2110 identified open reading frames on the genome of *B. kashiwanohense* APCKJ1, were designed and obtained from Agilent Technologies (Palo Alto, Ca., USA). Methods for cell disruption, RNA isolation, RNA quality control, complementary DNA synthesis and labelling were performed as described previously [55]. Labelled cDNA was hybridized using the Agilent Gene Expression hybridization kit (part number 5188-5242) as described in the Agilent Two-Colour Microarray-Based Gene Expression Analysis v4.0 manual (publication number G4140-90050). Following hybridization, microarrays were washed in accordance with Agilent's standard procedures and scanned using an Agilent DNA microarray scanner (model G2565A). Generated scans were converted to data files with Agilent's Feature Extraction software (Version 9.5). DNA-microarray data were processed as previously described [56-58]. Differential expression tests were

performed with the Cyber-T implementation of a variant of the student t-test [59].

DNA Manipulations. Chromosomal DNA was isolated from *B. breve* UCC2003 and *B. kashiwanohense* APCKJ1 as previously described [60]. Plasmid DNA was isolated from *E. coli*, *L. lactis* and *B. breve* using the Roche High Pure Plasmid Isolation kit (Roche Diagnostics, Basel, Switzerland). An initial lysis step was performed using 30 mg ml^{-1} of lysozyme for 30 minutes at 37°C prior to plasmid isolation from *L. lactis*, *B. breve* or *B. kashiwanohense*. DNA manipulations were essentially performed as described previously [53]. All restriction enzymes and T4 DNA ligase were used according to the supplier's instructions (Roche Diagnostics, Basel, Switzerland). Synthetic single stranded oligonucleotide primers used in this study (Supplemental Table S5.1) were synthesized by Eurofins (Ebersberg, Germany). Standard PCRs were performed using Extensor Hi-Fidelity PCR Master Mix (Thermo Scientific, Waltham, United States) or Q5 High-Fidelity 2X Mastermix (New England BioLabs, Herefordshire, United Kingdom) in a Life Technologies Pro-FLex PCR System (Thermo Scientific, Waltham, United States). PCR products were visualized by ethidium bromide (EtBr) staining following agarose gel electrophoresis (1 % agarose). *B. breve* colony PCR reactions were performed as described previously [61]. PCR fragments were purified using the Roche high Pure PCR purification kit (Roche Diagnostics, Basel, Switzerland). Plasmid DNA was isolated using the Roche High Pure Plasmid Isolation kit (Roche Diagnostics, Basel, Switzerland). Plasmid DNA was introduced into *E. coli* by electroporation as described previously [53]. *B. breve* UCC2003 [62] and *L. lactis* [63] were transformed by electroporation according to published protocols. The correct fragment orientation and integrity of all plasmid constructs (see also below) were verified by DNA sequencing, performed at Eurofins (Ebersberg, Germany).

Construction of overexpression vectors, protein overproduction and purification. For the construction of plasmids pNZ-FucA1 and pNZ-FucA2, DNA fragments encompassing the predicted fucosidase-encoding genes

fucA1 (corresponding to BKKJ1_2069) and *fucA2* (corresponding to BKKJ1_2070) were generated by PCR amplification from chromosomal DNA of *B. kashiwanohense* APCKJ1 using Q5 High-Fidelity DNA polymerase and the primer combinations 2069F and 2069R, or 2070F and 2070R, respectively (Supplemental Table S5.1). An in-frame N-terminal His10-encoding sequence was incorporated into the forward primers 2069F and 2070F to facilitate downstream protein purification. The generated amplicons were digested with SmaI and XbaI, and ligated into the ScaI and XbaI-digested, nisin-inducible translational fusion plasmid pNZ8150 [64]. The ligation mixtures were introduced into *L. lactis* NZ9000 by electrotransformation and transformants were then selected based on chloramphenicol resistance. The plasmid content of a number of Cm-resistant transformants was screened by restriction analysis and the integrity of positively identified clones was verified by sequencing.

Nisin-inducible gene expression and protein overproduction was performed as described previously [65-67]. In brief, 400 ml of M17 broth supplemented with 0.5 % (wt/vol) glucose was inoculated with a 2 % inoculum of a particular *L. lactis* strain, followed by incubation at 30°C until an OD_{600nm} of 0.5 was reached, at which point protein expression was induced by addition of cell-free supernatant of a nisin-producing strain [68], followed by continued incubation for a further 2 hours. Cells were harvested by centrifugation and protein purification achieved as described previously [65]. Protein concentrations were determined using the Bradford method [69].

Assay of individual and combined fucosidase activities. The individual or sequential hydrolytic activities specified by FucA1 and FucA2 were determined essentially as described previously [67], using either 2'-FL or 3-FL as a substrate. Briefly, a 50 µl volume of each purified protein (protein concentration of 0.5 mg/ml) was added to 20 mM morpholinepropanesulfonic acid (MOPS) (pH 7.0) buffer and 1 mg ml^{-1} (wt/vol) of one of the above-mentioned sugars in a final volume of 1 ml, followed by incubation for 24 hours at 37°C. In order to determine enzyme affinity, each enzyme was

incubated individually or in combination, with 2'-FL or 3-FL, as described above, with 200 µl samples taken at the following time points: 20 minutes, 1 hour, 4 hours and 24 hours. All samples were subject to a final enzyme denaturation step at 85°C for 15 minutes, before storage at -20°C.

HPAEC-PAD analysis. For HPAEC-PAD analysis, a Dionex (Sunnyvale, CA) ICS-3000 system was used. Carbohydrate fractions from the above-mentioned hydrolysis assays (25 µl aliquots) were separated on a CarboPac PA1 analytical-exchange column (dimensions, 250 mm by 4 mm) with a CarboPac PA1 guard column (dimensions, 50 mm by 4 mm) and a pulsed electrochemical detector (ED40) in PAD mode (Dionex). Elution was performed at a constant flow-rate of 1.0 ml/min at 30°C using the following eluents for the analysis: eluent A, 200 mM NaOH; eluent B, 100 mM NaOH plus 550 mM Na acetate; eluent C, Milli-Q water. The following linear gradient of sodium acetate was used with 100 mM NaOH: from 0 to 50 min, 0 mM; from 50 to 51 min, 16 mM; from 51 to 56 min, 100 mM; from 56 to 61 min, 0 mM. Chromatographic profiles of standard carbohydrates were used for comparison of the results of their breakdown by FucA1 or FucA2 proteins. Chromeleon software (version 6.70; Dionex Corporation) was used for the integration and evaluation of the chromatograms obtained. A 1 mg/ml stock solution of each of the carbohydrates, as well as their putative breakdown products (where available) used as reference standards was prepared by dissolving the particular sugar in Milli-Q water. Chromatographic profiles of mMRS containing standard carbohydrates were used for comparison of the results of their fermentation by *B. kashiwanohense* APCKJ1. Chromeleon software (version 6.70; Dionex Corporation) was used for the integration and evaluation of the chromatograms obtained. mMRS containing 1 % of 2'-FL or 3-FL, as well as their putative breakdown products (where available) used as reference standards.

Generation of a recombinant *B. breve* strain. DNA fragments encompassing *fucA1* (corresponding to locus tag BKKJ1_2069); and *fucS*,

fucP1 and *fucP2* (corresponding to locus tags BKKJ1_2076-2078, including their presumed promoter region located upstream of BKKJ1_2078; this fragment was designated *fucSP1P2*) were generated by PCR amplification from *B. kashiwanohense* APCKJ1 using Q5 High-Fidelity Polymerase (New England BioLabs, Herefordshire, United Kingdom) and primer pairs 2069pNZ44F and 2069pNZ44R, and 2076-78pCB1.2F and 2076-78pBC1.2R, respectively. Similarly, the *strR* gene (encoding streptomycin resistance and corresponding to CNCMI4321_0987 from *B. breve* CNCMI4321, unpublished data) was amplified using *B. breve* CNCMI4321 chromosomal DNA and primer pair Strep^{R}F and Strep^{R}R. Primer sequences are listed in Supplemental Table S5.1.

The resulting *fucA1*-encompassing fragment was digested with KpnI and XbaI, and ligated to the similarly digested pNZ44 [70]. The ligation mixture was introduced into *L. lactis* NZ9000 by electrotransformation and transformants were then selected based on chloramphenicol resistance. The plasmid content of a number of Cm-resistant transformants was screened by restriction analysis. The integrity of the cloned insert of one of the recombinant plasmids, designated pNZ44-*fucA1*, was confirmed by sequencing. The *fucA1*-coding sequence and pNZ44 plasmid backbone, including the constitutive p44 lactococcal promoter, specified by pNZ44, but excluding the chloramphenicol resistance-encoding cassette, was amplified by PCR from pNZ44-*fucA1* using Q5 High-Fidelity DNA polymerase and primer combination pNZ44noCmF and pNZ44noCmR (Supplemental Table S5.1). The resulting DNA fragment was digested with SpeI and SmaI, and ligated with the similarly digested *strR*-encompassing fragment, to generate pNZ44-*fucA1-strR*. Using the same method, a DNA fragment containing the empty pNZ44 sequence, excluding the chloramphenicol resistance-encoding cassette, was amplified from unaltered pNZ44, digested and ligated with the similarly digested *strR*-encompassing fragment, to generate pNZ44-*strR*. The ligation mixtures were introduced into *E. coli* EC101 by electrotransformation and transformants were selected based on streptomycin resistance. Transformants were checked for plasmid content using colony PCR, restriction analysis of plasmid DNA, and verified by sequencing.

Plasmids pNZ44-*fucA1-strR* and pNZ44-*strR* were introduced into *E. coli* EC101 harbouring pNZ-M.BbrII-M.BbrIII by electroporation, as previously described [61], and transformants were selected based on Cm and Strep resistance. Methylation of the plasmid complement of transformants by the M.BbrIII (isoschizomer of PstI) was confirmed by their observed resistance to PstI restriction [61]. Methylated pNZ44-*fucA1-strR* was introduced by electrotransformation into *B. breve* UCC2003 and transformants were selected based on streptomycin resistance, generating *B. breve* strain UCC2003-*fucA1*.

The DNA fragment encompassing *fucSP1P2*, including its presumed promoter region, was digested with XmaI and XbaI, and ligated to the similarly digested pBC1.2 to generate pBC1.2-*fucSP1P2*. The ligation mixture was introduced into *E. coli* EC101 by electrotransformation and transformants selected based on chloramphenicol resistance. Transformants were checked for plasmid content using colony PCR, restriction analysis of plasmid DNA, and verified by sequencing. pBC1.2-*fucSP1P2* was introduced into *B. breve* UCC2003-*fucA1* by electrotransformation and transformants were selected based on streptomycin and chloramphenicol resistance, generating the strain *B. breve* UCC2003-*fucA1-fucSP1P2*.

In addition, the 'empty' cloning vector pBC1.2 was introduced into strain *B. breve* UCC2003-*fucA1* (to generate *B. breve* UCC2003-*fucA1*-pBC1.2), while 'empty' cloning vector pNZ44-*strR* was introduced into strain *B. breve* UCC2003-*fucSP1P2* (to generate *B. breve* UCC2003-*fucSP1P2*-pNZ44-*strR*), which were then used as controls.

Comparative analysis of fucosyllactose and fucose utilisation genes. Sequences of proteins, proposed to be involved in the utilisation of fucosyllactose or fucose, encoded by the genomes of *B. kashiwanohense* APCKJ1, *B longum* subsp. *infantis* DSM20088 and *B. breve* UCC2003 were compared, and their homologs were identified in the protein sequences encoded by each other's genomes using BLASTP [41] at default settings. Cut-off values of a minimum Bit Score of 200 bits, a minimum identity of 50 %

coverage and minimum e-value of 0.0001 were used for the identification of homologous proteins. In the cases of non-identical predicted functions, automatic protein annotations were checked for alternative annotations using the Pfam (http://pfam.xfam.org) [71] and KEGG (http://www.genome.jp/kegg/) [72] databases.

Nucleotide sequence and microarray data accession numbers. The genome sequence of *B. kashiwanohense* APCKJ1 was deposited at the GenBank under the accession no. CP026729. The microarray data obtained in this study have been deposited in NCBI's Gene Expression Omnibus database and are accessible through GEO Series accession number GSE107439.

5.4 Results

Isolation of a novel *B. kashiwanohense* isolate. A single stool sample each from six exclusively breast-feeding neonates, aged 4-8 weeks, was selected for the isolation of fucosyllactose-utilising *Bifidobacterium* isolates. Pure cultures of fucosyllactose-utilising *Bifidobacterium* species were isolated from these faecal samples, based on the presence of mupirocin (which selects for bifidobacteria [73, 74]) and their ability to grow on 2'-FL or 3-FL as their sole carbohydrate source.

Of the 53 bifidobacterial isolates obtained in our screening of six infant faecal samples using 2'-FL or 3-FL, only seven isolates, all from the same faecal sample using 2'-FL as the selective carbohydrate (and the only *Bifidobacterium* isolates obtained from this sample), were found to belong to the species *B. kashiwanohense* (see Chapter IV for further details on other isolates). As this is an infrequently encountered species (from infant faecal samples), yet clearly exhibiting an ability to grow on 2'-FL as its sole carbohydrate source, one of these isolates, designated APCKJ1, was chosen for further characterisation.

Carbohydrate Utilisation by *B. kashiwanohense* APCKJ1. In order to further investigate the carbohydrate utilization capabilities of *B. kashiwanohense* APCKJ1, growth assays were carried out for this strain inoculated in modified mMRS medium supplemented with 1 % (wt/vol) of one of 16 different sugars, including nine HMO components, the results of which are shown in Figure 5.1. Growth was assessed by measuring the OD_{600nm} following 24 hours of anaerobic growth at 37 °C. APCKJ1 was shown to reach its highest level of cell density following growth on lactose (OD_{600nm}>3.0), while also reaching high optical densities following growth on the monosaccharides glucose, ribose, sorbitol, as well as the disaccharide melibiose and the trisaccharide raffinose. Good growth (final OD_{600nm}>1.5) was observed for two of the nine HMO components tested: the isomers 2'-FL and 3-FL. This confirmed the observation that APCKJ1 is capable of metabolising 2'-FL as its sole carbohydrate source. L-fucose, however, did not support growth of APCKJ1 to any substantial degree, suggesting the strain is incapable of consuming this component of fucosyllactose when present extracellularly in its free form. It should be mentioned that some sugars did not support growth to any substantial level (final OD_{600nm}<0.5), yet provided optical densities higher than the negative control; this may be due to the presence of minor impurities, such as lactose, in some of the sugar preparations (probably due to the synthesis process which uses lactose as a feeding material for the synthesis of the carbohydrate), thus causing this residual growth.

The fermentation supernatants of APCKJ1 when grown on 2'-FL or 3-FL (as well as lactose) were analysed by HPLC, in order to determine the fates of these glycans, by the presence or absence of specific metabolic end-products. Samples were centrifuged and filtered to remove cellular debris, and then analysed for their composition by HPLC. This analysis appears to reveal full degradation of both 2'-FL and 3-FL, and substantial catabolism of their immediate breakdown products lactose and L-fucose (Fig. 5.2). Using uninoculated mMRS containing lactose, 2'-FL, 3-FL or no carbohydrate source, and the fermentation supernatant of APCKJ1 in mMRS with lactose as controls, we can examine of the metabolic end products of APCKJ1 when

grown in mMRS containing 2'-FL or 3-FL. This analysis reveals no accumulation of lactose, and a low level of L-fucose release in the growth medium by APCKJ1. In addition, we observe the accumulation of a number of metabolic products. As with the lactose fermentate, chromatographic peaks matching those of lactate and acetate standards are observed for both the 2'-FL and 3-FL-associated fermentates. Lactate and acetate are known metabolic end products of both lactose and fucosyllactose in bifidobacteria [31, 38, 75]. However, in the chromatograms obtained from spent medium of APCKJ1 grown on 2'-FL and 3-FL, we also observe the presence of several additional peaks that indicate the production of metabolic compounds not observed for lactose fermentation. One of the largest among these peaks is one, with a retention time of approximately 18.1 minutes, which matches with that of the standard of 1,2-propanediol, which has been put forward as an end product of L-fucose metabolism [31, 38]. These findings suggest that *B. kashiwanohense* APCKJ1 is not only capable of internalising and degrading 2'-FL/3-FL and utilising its lactose component, but that it also utilises (at least some of the) L-fucose. However, mass spectrometry or nuclear magnetic resonance spectroscopy of these accumulating products, detected by HPLC analysis, would be required to definitively identify them.

General genome features of *B. kashiwanohense* APCKJ1, and transcriptome analysis during growth on 2'-FL and 3-FL. The genome of the *B. kashiwanohense* isolate APCKJ1 was sequenced, revealing a 2,445,409-bp circular molecule, containing 53 tRNA genes and 5 rRNA operons. The G+C content of the genome is 56.19 %, with a total of 2,110 predicted protein encoding sequences (CDS). Scores for the sequencing reads and quality are given in Supplemental Table S5.4. Methylome analysis revealed the presence of m6A-type methylation at the third position of the methylation motif 'GGAGT', most commonly associated with a Type IIS restriction-modification (R/M) system [76-78]. A type II R-M system methyltransferase was predicted to be encoded on the APCKJ1 genome at the gene BKKJ1_0217. A rebase blast (http://rebase.neb.com/cgi-bin/pacbiolist;

http://tools.neb.com/genomes) [43] of its amino acid sequence revealed up to 98 % identity to other type II methyltransferases.

In order to investigate which genes are involved in the metabolism of 2'-FL and 3-FL, microarrays were designed based on the predicted gene content of *B. kashiwanohense* APCKJ1, and global gene expression was determined by microarray analysis during growth of APCKJ1 in mMRS supplemented with each respective sugar, and compared with the transcriptome obtained during growth in mMRS supplemented with sorbitol. Genes that were shown to be significantly up- or down-regulated in transcription above the designated cut-off (fold-change >5.0, $P < 0.001$) are shown in Table 5.1. During growth on 2'-FL or 3-FL (and compared to the transcriptome of the strain when grown on sorbitol), the transcriptionally upregulated genes included *galA*, a predicted β-galactosidase located in the gene cluster corresponding to locus tags BKKJ1_2062-2068 (based on their predicted functions designated here as the *gal* locus), and all the genes (with the exception of a predicted regulator, *fucR*) of an adjacent gene cluster corresponding to locus tags BKKJ1_2069-2079 (and designated here as the *fuc* locus) (Fig. 5.3A). This locus was found to contain two adjacent predicted α-fucosidase-coding genes, BKKJ1_2069 and BKKJ1_2070 (designated *fucA1* and *fucA2*, respectively). The hydrolytic activities of these 2 genes' products were thus characterised, as described in the next section. Based on their annotations, the *gal* and *fuc* loci are predicted to function in lactose and fucosyllactose/L-fucose metabolism, respectively. The genes significantly downregulated in transcription belong to three distinct gene clusters (corresponding to locus tags BKKJ1_0067-0069, BKKJ1_0206-0208 and BKKJ1_0336-0341), which, as expected, appear to be involved in sorbitol metabolism based on BLAST analysis (Table 5.1).

Global gene expression was also determined for APCKJ1 during growth in mMRS supplemented with 2'-FL, as compared with gene expression during growth in mMRS supplemented with lactose, in order to identify genes potentially involved in the metabolism of the L-fucose released by the degradation of fucosyllactose. A transcription profile, similar to that of APCKJ1 grown on 2'-FL compared to the sorbitol control, was observed, with the exception of the absence of upregulation of *galA*, but instead the

upregulation of the adjacent galactoside symporter-coding gene *galS*. Interestingly, significant transcriptional upregulation was noted of gene BKKJ1_0429 (designated here as *prpO*), predicted to encode 1,2-propanediol oxidoreductase. 1,2-propanediol has previously been identified as a product of L-fucose metabolism in certain bifidobacteria [31, 38], and the presence of this gene may be consistent with the results of the fermentation product analysis which suggests the utilisation of L-fucose derived from 2'-FL/3-FL by APCKJ1. The *gal* locus was observed to possess an apparently truncated LacI-type repressor-coding gene, *galR* (Fig. 5.3A). A multiple sequence alignment of the amino acid sequences of *galR* and other predicted LacI regulator-coding genes indeed revealed a large proportion of the sequence missing in *galR*, as compared to other genes encoding LacI-type regulators (Supplemental Fig. S5.1). If *galR* is indeed expressed as a non-functional product otherwise required for transcriptional repression of the genes of the *gal* locus, it would be expected that the genes responsible for lactose metabolism are constitutively expressed. Interestingly, no such LacI-type regulator-coding gene, truncated or intact, can be identified in the homologous region of *B. kashiwanohense* DSM21854 [31].

Heterologous expression, purification and biochemical characterisation of FucA1 and FucA2, and enzymatic activity on 2'-FL and 3-FL. In order to investigate the predicted fucosidase activities encoded by *fucA1* (BKKJ1_2069) and *fucA2* (BKKJ1_2070), the corresponding FucA1 and FucA2 proteins were purified as His-tagged versions (and designated here as FucA1His and FucA2His, respectively; see Materials and Methods). Assessment of substrate specificity was performed by incubating FucA1His and FucA2His on their own or in combination with 2'-FL or 3-FL, and analysing the reaction products by HPAEC-PAD. Purified FucA1His was shown to fully liberate lactose and L-fucose from both 2'-FL and 3-FL (Fig. 5.4), demonstrating hydrolytic activity towards both the Fucα1-2Gal linkage in 2'-FL and the Fucα1-3Glc connection in 3-FL, thus indicating a key role in the hydrolysis and utilisation of both sugars. Purified FucA2His was shown to fully cleave 3-FL, but not 2'-FL, into lactose and L-fucose (Fig. 5.4),

demonstrating hydrolytic activity solely towards the Fucα1-3Glc residue. As would be expected, given these results, the combined activities of both FucA1His and FucA2His, when incubated together with either sugar, degraded both 2'-FL and 3-FL into lactose and L-fucose (Supplemental Fig. S5.2).

By analysing the reaction products of these incubations at different time points (of 20 minutes, 1 hour, 4 hours and 24 hours), we gained further insights into the reaction rate and specificity of these enzymes, particularly FucA1. The reaction products of both FucA2His alone (Fig. 5.4), or FucA1His and FucA2His in combination (Supplemental Fig. S5.2) with either 2'-FL or 3-FL were identical across all time points, indicating a high hydrolytic specificity and rate of FucA1His for 2'-FL, and similarly of the FucA2His enzyme for 3-FL. However, while the reaction profile of FucA1His alone with 2'-FL was identical across all time points, as expected, this was not similarly observed for the reaction products of FucA2His incubated alone with 3-FL. While FucA1His was shown to fully degrade 3-FL into lactose and L-Fucose by the 4-hour time point, small amounts of un-degraded 3-FL could still be observed at the 1-hour and 20-minute time points, indicating a lower catalytic rate for FucA1 with 3-FL compared to that with 2'-FL as a substrate (Fig. 5.4). All together, these enzymatic assays point towards a specificity of FucA2 to 3-FL alone, and dual activity of FucA1 towards 2'-FL as well as 3-FL, albeit that this enzyme exhibits, at least under the conditions tested, a higher catalytic rate for 2'-FL as compared to 3-FL.

Phenotypic analysis of *B. breve* UCC2003 strains harbouring *B. kashiwanohense* APCKJ1 genes implicated in 2'-FL/3-FL metabolism. It has previously been shown that the infant-associated species *B. breve* is capable of utilising L-fucose as a growth substrate to a limited degree [16, 79]. While a small number of strains of *B. breve* have been demonstrated to consume 2'-FL [31, 80], the majority of studied strains, including UCC2003, are not capable of using 2'-FL or 3-FL as a substrate to any appreciable extent [15, 16, 31]. Thus, *B. breve* UCC2003 is capable of successfully metabolising

the two individual components that make up fucosyllactose (i.e. lactose and L-fucose), yet not the 2'-FL or 3-FL compounds themselves, whereas *B. kashiwanohense* APCKJ1 is capable of utilising 2'-FL, 3-FL and lactose, but not L-fucose. We therefore elected to heterologously express the *B. kashiwanohense* APCKJ1 *fuc* gene cluster, which we predicted to be required for (uptake and hydrolysis of) fucosyllactose, in *B. breve* UCC2003 in an attempt to confer the ability for 2'-FL and 3-FL utilisation to this latter strain. The *fucA1* gene was selected because of its ability to hydrolyse both 2'-FL and 3-FL (see above), and the genes *fucS*, *fucP1* and *fucP2* (corresponding to locus tags BKKJ1_2076-2078) were chosen, as these adjacent genes are predicted to encode the (2'-FL and 3-FL) transporter components of the *fuc* locus (Table 1). Accordingly, *fucA1* was cloned into pNZ44-*strR* using an *L. lactis* NZ900 cloning host, and this construct was passaged through an *E. coli* DNA-methylating strain, as previously described [61], to facilitate introduction into *B. breve* UCC2003. This produced strain *B. breve* UCC2003-*fucA1*, where the *fucA1* gene is expressed under the control of the constitutive p44 promoter [81]. The DNA fragment containing the *fucS*, *fucP1* and *fucP2* genes (collectively designated *fucSP1P2* and including the presumed promoter region) was cloned into pBC1.2 using an *E. coli* EC101 cloning host. This construct was introduced into *B. breve* UCC2003-*fucA1*, thereby generating strain *B. breve* UCC2003-*fucA1-fucSP1P2*.

This recombinant strain was assessed for its ability to grow in mMRS supplemented with either 2'-FL or 3-FL, or with a lactose control, as compared to the following strains: wild-type UCC2003, *B. breve* UCC2003-*fucA1*-pBC1.2, UCC2003-*fucSP1P2*-pNZ44-*strepR*, and wild-type *B. kashiwanohense* APCKJ1 (Fig. 5.5). The recombinant strains *B. breve* UCC2003-*fucA1*-pBC1.2 and UCC2003-*fucSP1P2*-pNZ44-*strepR* were generated and used as negative controls, as each only express either of the two components predicted as required for 2'-FL/3-FL utilisation *fucA1* and *fucSP1P2*, which are both expressed by *B. breve* UCC2003-*fucA1-fucSP1P2*. While wild-type UCC2003, as well as the recombinant UCC2003 strains expressing either *fucA1* or *fucSP1P2* alone, displayed a complete lack of growth on 2'-FL and 3-FL (final OD_{600nm}<0.5), the recombinant UCC2003

strain expressing both *fucA1* and *fucSP1P2* demonstrated good growth on either 2'-FL or 3-FL (final OD_{600nm}>2.0), which is comparable with that obtained by APCKJ1 grown on either of these substrates (Fig. 5.5). As expected all of the tested cultures grew to a high cell density on lactose (OD_{600nm}>3.0).

Thus, the transcriptome data, substrate hydrolysis profiles and growth results of various recombinant strains corroborate the notion that in *B. kashiwanohense* APCKJ1 the fucosidase encoded by *fucA1* is specifically required for 2'-FL hydrolysis, and capable of hydrolysing 3-FL, while the solute binding protein and two permeases encoded by *fucS*, *fucP1* and *fucP2*, respectively, are responsible for the uptake and internalisation of 2'-FL and 3-FL. These results also demonstrate the ability of *B. breve* UCC2003 to successfully internalise and utilise 2'-FL and 3-FL for growth, with the heterologous and concomitant expression of this transport system and fucosidase from *B. kashiwanohense* APCKJ1.

Transcriptome analysis of recombinant *B. breve* UCC2003 grown on 2'-FL. In order to better understand the metabolism of fucosyllactose by the recombinant *B. breve* strain generated in this study, global gene expression was determined by microarray analysis during growth of UCC2003-*fucA1-fucSP1P2* in mMRS supplemented with 2'-FL, as compared with gene expression during growth in mMRS supplemented with ribose. As lactose is a major component of 2'-FL, microarray analysis was also carried out to compare gene expression during growth of UCC2003-*fucA1-fucSP1P2* in mMRS supplemented with lactose, as compared to gene expression during growth in mMRS supplemented with ribose. Genes that were shown to be significantly upregulated in transcription above the designated cut-off (fold-change >2.0, $P < 0.001$) are displayed in Table 5.2.

Among the genes that showed increased transcription when UCC2003-*fucA1-fucSP1P2* was grown on lactose or 2'-FL (as compared to when this strain was grown on ribose) were genes from loci known to function in lactose metabolism, namely the *lnt* (Bbr_0526-0530) [28] and *lac* loci (Bbr_1551-

1553) [82]. Among the genes that were uniquely upregulated in transcription during growth on 2'-FL were those of the gene cluster Bbr_1739-1748 (designated here as the *fcu* locus) (Fig. 5.3B). These results indicate that while expression of the non-native *B. kashiwanohense* APCKJ1 genes *fucA1*, *fucS*, *fucP1* and *fucP2* in *B. breve* UCC2003 allows the latter strain to take up 2'-FL and hydrolyse it into lactose and L-fucose, expression of native UCC2003 genes of the *lnt*, *lac* and *fcu* loci enable subsequent utilisation of these carbohydrates.

Elucidation of a fucosyllactose metabolic pathway in *B. kashiwanohense* and *B. breve*. The results of the APCKJ1 transcriptome analyses indicate involvement of the *fuc* locus and *prpO* in fucosyllactose metabolism, and the combined results of the *in vitro* hydrolytic assays and recombinant UCC2003 growth analysis implicate the APCKJ1 fucosidase-coding genes *fucA1* and *fucA2*, as well as the transporter-encoding genes *fucS*, *fucP1* and *fucP2* in the internalisation and hydrolysis of 2'-FL and 3-FL. However, the presence of further genes of the *fuc* locus, as well as *prpO*, upregulated in transcription during growth on 2'-FL or 3-FL suggests the involvement of these genes in the metabolism of the liberated L-fucose, given that the genes for lactose metabolism appear to be constitutively expressed. Likewise, the transcriptional upregulation of genes in the UCC2003 *fcu* locus during growth of the recombinant strain on 2'-FL implicate these genes in L-fucose metabolism in UCC2003. This suggests that both *B. kashiwanohense* APCKJ1 and *B. breve* UCC2003 possess functional pathways for the utilisation of (intracellular) L-fucose.

A pathway for the anaerobic catabolism of L-fucose, into lactate and pyruvate, was recently proposed in *B. longum* subsp. *infantis* DSM20088 [31] based on a fucose-dependent metabolic route as previously identified in *Xanthomonas campestris* [83]. The six enzymes of this pathway were identified as: (i) an L-fucose mutarotase, (ii) an L-fucose dehydrogenase, (iii) an L-fuconolactone hydrolase, (iv) an L-fuconate dehydratase, (v) an L-2-keto-3-deoxy-fuconate-4-dehydrogenase, and (vi) an L-2,4-diketo-3-deoxy-

fuconate hydrolase. Additionally, the presence of an alternative L-fucose catabolism pathway was proposed by the former study [31], diverging at the L-2-keto-3-deoxy-fuconate metabolite step, and instead yielding pyruvate and 1,2-propanediol as end products for further metabolism, through the activity of an L-2-keto-3-deoxy-fuconate aldolase and an L-1,2-propanediol oxidoreductase. A similar pathway, generating end products of pyruvate and 1,2-propanediol, yet for D-fucose utilisation, has previously been described for a number of pseudomonads [84-87] and *E. coli* [88, 89].

Thus, the genes of the proposed L-fucose utilisation pathway in *B. longum* subsp. *infantis* DSM20088 were used in a search for homologs in the genomes of both *B. kashiwanohense* APCKJ1 and *B. breve* UCC2003. The eight genes encoding the above-mentioned enzymatic activities, as well as two fucosidase-encoding genes, from *B. longum* subsp. *infantis* DSM20088 were used in the BLASTP search of the APCKJ1 and UCC2003 genomes. Significant homologs identified in the APCKJ1 and UCC2003 genomes, whose predicted annotations did not match the annotations of their DSM20088 equivalents based on their function in the fucose/fucosyllactose utilisation pathways [31], were checked for alternative annotations closer to their predicted function in the pathways, in the Pfam (http://pfam.xfam.org) [71] and KEGG (http://www.genome.jp/kegg/) [72] databases, based on the families of their coded proteins and similar enzymatic functions. This similarity search for the ten aforementioned DSM2008 genes in the APCKJ1 genome revealed the presence of significant homologs for all 10 genes (Supplemental Table S5.3). The homologs in APCKJ1 included seven of the genes of the *fuc* locus, and the gene *prpO*, all of which were transcriptionally upregulated in the APCKJ1 arrays during growth on 2'-FL. The remaining 2 homologs, encoding a predicted L-fucose dehydrogenase and an L-2,4-diketo-3-deoxy-fuconate hydrolase, were located elsewhere in the genome, and were not observed as transcriptionally upregulated in the APCKJ1 arrays, possibly indicating their constitutive expression. These results therefore suggest that all genes proposed as necessary for L-fucose metabolism in *B. kashiwanohense* APCKJ1, as observed in *B. breve*, *B. longum* subsp. *infantis*, and *B. longum* subsp. *suis* [31, 38], yielding L-lactate and pyruvate (Fig. 5.6).

Furthermore, these results also indicate the presence of the proposed secondary pathway for L-fucose metabolism in APCKJ1, yielding 1,2-propanediol and pyruvate [31, 38] (Fig. 5.6). The potential presence of both these pathways in APCKJ1 is supported by the observed accumulation of lactate, acetate (a pyruvate breakdown product) and 1,2-propanediol in the 2'-FL and 3-FL fermentates analysed by HPLC, as mentioned above.

Homologs for eight of the ten assessed DSM20088 genes were identified in the UCC2003 genome (Supplemental Table S5.3), four of which were located in the *fcu* locus and upregulated in expression in the arrays of the recombinant UCC2003 grown on 2'-FL; namely *fcuH* (Bbr_1741), *fcuF* (Bbr_1744), *fcuD* (Bbr_1743) and *fcuG* (Bbr_1740). The lack of increased expression of the remaining four homologous genes in the recombinant UCC2003 array; namely *fclA* (Bbr_1288), *fcuD* (Bbr_0398), *fcsH* (Bbr_1783) and *fucO* (Bbr_1505); suggests that they are constitutively transcribed. The two DSM20088 genes lacking homologs in UCC2003 were the GH29 α-fucosidase-coding gene and the L-fucose mutarotase-coding gene. Despite the absence of an L-fucose mutarotase-encoding gene, *B. breve* UCC2003 has previously been shown to consume L-fucose released from HMO structures [79], which corroborates the observed recombinant array results, and the notion that *B. breve* UCC2003 possesses a functional L-fucose utilisation pathway (Fig. 5.6). The apparently absent L-fucose mutarotase, therefore, is likely not prohibitive in the metabolism of free L-fucose. While homologs for the three genes encoding components of the APCKJ1 transport system, *fucP1*, *fucP2* and *fucS*, and a GH95 fucosidase-coding gene *fucA1* were also identified in UCC2003 (not shown), it is apparent that they do not act on 2'-FL or 3-FL, as demonstrated in the recombinant UCC2003 growth assays. It is possible that they may target other fucosylated HMO structures, such as fucosyl-LN(n)T, as *B. breve* is known to efficiently consume free LN(n)T [28]. Likewise, the conspicuous absence of a homolog for the UCC2003 L-fucose permease, *fcuP*, in APCKJ1 (not shown) may explain the inability of APCKJ1 to utilise free L-fucose.

5.5 Discussion

B. kashiwanohense presents an example of an apparently infant-specific *Bifidobacterium* species, although only two publications have reported its isolation from infant faeces [32, 33], and it is thus not considered to represent a common member of the neonatal gut microbiota. The bifidobacterial component of the infant gut microbiota is typically dominated by the species *B. breve*, *B. longum* subsp. *longum* and *B. bifidum*, with *B. pseudocatenulatum* common, and *B. longum* subsp. *infantis* and *B. longum* subsp. *suis* occasionally also featuring [11, 90-92]. However, adaptations to life in the infant gut by *B. kashiwanohense* suggest that this is its main niche. The ability of a *B. kashiwanohense* isolate to efficiently sequester iron has previously been described [32], while it has also been reported that the same isolate is able to grow on the HMOs 2'-FL and 3-FL as its sole carbohydrate source. HMO utilisation by certain members of the *Bifidobacterium* genus appears to be the prerogative of infant-associated species [8, 16, 17, 27, 31], and as fucosylated sugars typically represent a substantial proportion of all HMO [2, 34, 35], the ability to utilise fucosyllactose in *B. kashiwanohense* represents a specific adaptation to the neonate intestinal niche. However, the precise mechanism of fucosyllactose metabolism in *B. kashiwanohense* is as yet poorly understood.

Among the isolates obtained during the screening of infant faecal samples, using 2'-FL as the selective carbohydrate, was the novel isolate of *B. kashiwanohense* APCKJ1. This strain was subsequently demonstrated as capable of consuming both the HMO components 2'-FL and 3-FL as its sole carbohydrate source. Understanding the mechanisms of fucosyllactose metabolism in *B. kashiwanohense* APCKJ1 became the focus of this investigation.

The transcriptome results for APCKJ1 strongly indicate the presence of a cluster dedicated to the metabolism of fucosyllactose. Composed of the adjacent *gal* and *fuc* loci, this roughly 23 kb region represents a potential 'HMO island' within the APCKJ1 genome, encoding the metabolic machinery necessary for the utilisation of fucosyllated HMO. A similar such

region has previously been described in *B. longum* subsp. *infantis* [24, 93, 94] and *B. longum* subsp. *longum* [95], and reinforces the notion of specific adaptations to HMO utilisation by infant-associated bifidobacteria. In particular, the increased expression of the *fuc* locus during growth on 2'-FL and 3-FL implicates the 11 associated genes in the metabolism of fucosyllactose. Interestingly, the array results also indicate the constitutive expression of the genes involved in lactose metabolism in APCKJ1, the chief candidates of which are those of the adjacent *gal* locus, which appears due to the truncated, non-functional LacI-type repressor, GalR, encoded within the *gal* locus. This gene is in fact entirely absent in *B. kashiwanohense* DSM21854 and *B. longum* subsp. *infantis* DSM20088 [31], suggesting the constitutive expression of the equivalent loci in these two strains.

The combined results of the *in vitro* hydrolysis assays with the heterologously expressed FucA1 and FucA2 fucosidases, and of the growth results obtained with the recombinant *B. breve* UCC2003 strain confirm the predicted functions of some of the genes of the *fuc* locus. Our findings are consistent with the notion that the three adjacent genes *fucS*, *fucP1* and *fucP2* encode the transport components necessary for the uptake of 2'-FL and 3-FL, while *fucA1* and *fucA2* encode the fucosidases responsible for the hydrolysis of 2'-FL and 3-FL, respectively, releasing lactose and L-fucose. While FucA1 is capable of hydrolysing both 2'-FL and 3-FL, it possesses higher catalytic efficiency with the former, FucA2 can only hydrolyse the latter. Such redundancy in the degradation of isomeric HMO structures by multiple glycosyl hydrolases is not unusual in bifidobacteria [17, 25, 28], but the observation that the two identified APCKJ1 fucosidases are able to hydrolyse 3-FL is interesting, as 3-FL has been reported to be much less abundant in breastmilk than its isomer 2'-FL [96]. However, the true target substrate for FucA2 may in fact be another fucosylated HMO structure, such as fucosyl-LN(n)T, and thus its ability to hydrolyse 3-FL simply being due to the similarity of the linkage in these two glycans. A similar result of good growth would be expected for 3-FL, yet with no growth on 2'-FL, if the recombinant growth assay was replicated with recombinant UCC2003 expressing *fucSP1P2*, and *fucA2* instead of *fucA1*.

While these results reveal the mechanisms of fucosyllactose uptake and intracellular hydrolysis in *B. kashiwanohense* APCKJ1, they do not reveal the fate of L-fucose in the cell. The combined results of the HPLC analysis of the APCKJ1 2'-FL/3-FL fermentations and APCKJ1 microarrays elude to the utilisation of the L-fucose released by the hydrolysis of 2'-FL and 3-FL, and the comparative analysis with *B. longum* subsp. *infantis* DSM20088 revealed all the genetic components necessary in *B. kashiwanohense* APCKJ1 for the utilisation of L-fucose by two diverging pathways. This therefore represents the first identification of a functional fucose utilisation pathway in *B. kashiwanohense*, a species which has previously been suggested to simply release the fucose component of fucosyllactose extracellularly [31, 38]. This notion is supported by the (predicted) identification of identical pathways for L-fucose utilisation in *B. breve* UCC2003 here, as this species has recently been characterised for its utilisation of L-fucose [38]. The one major difference between *B. breve* UCC2003 and *B. kashiwanohense* APCKJ1 in their metabolism of fucose is the inability of APCKJ1 to utilise free L-fucose (only utilising L-fucose released by cytoplasmic 2'-FL/3-FL hydrolysis) and the inability of UCC2003 to utilise 2'-FL or 3-FL (only utilising free L-fucose). These distinctive metabolic approaches perhaps represent key adaptations to the infant gut, allowing for the efficient targeting of particular HMO components. While the ability to utilise fucosyllactose is undoubtedly a powerful approach for the establishment of a population in the breastfed neonatal gut, it is a strategy employed by multiple species of infant-associated Bifidobacteria, including *B. bifidum*, *B. longum* subsp. *infantis*, *B. longum* subsp. *longum* and some species of *B. breve*, all of which are capable of utilising other HMO components, such as sialyllactose, LN(n)T, LNB, sialic acid and L-fucose [8, 17-30, 95, 97]. Thus, the narrow specialisation in HMO metabolism by *B. kashiwanohense* may in fact also be its greatest hindrance in becoming dominant in the breastfed infant gut microbiota.

Nonetheless, the results obtained in this study demonstrate the adaptation of fucosyllactose utilisation as a key factor for the establishment and survival of *B. kashiwanohense* in the gut of the breastfed neonate, particularly the ability of the strain APCKJ1 to utilise both lactose and L-fucose components.

Furthermore, this sheds light on L-fucose metabolism in *B. breve*, revealing a common pathway among infant-associated species of bifidobacteria capable of metabolising L-fucose. On the whole, this serves to reinforce the effectiveness of HMO components such as fucosyllactose as prebiotics for the development of a specific infant gut microbiota, and the ability of infant-associated microbiota members, including bifidobacteria, to colonise this environment.

5.6 Acknowledgements

The authors would like to sincerely thank Glycom A/S (Lyngby, Denmark) for the provision of purified HMO samples used in this study under their donation program. Additionally, the authors would like to sincerely thank all those involved in the InfantMet Project for the provision of infant faecal sampled for the screening of bifidobacteria. This study was funded in part by the Irish Research Council, under the Postgraduate Research Project Award; Project ID GOIPG/2013/651. In addition, the authors are supported by Science Foundation Ireland (SFI) (Grant No. SFI/12/RC/2273).

5.7 Tables and Figures

Table 5.1. *B. kashiwanohense* APCKJ1 genes that are transcriptionally upregulated during growth in mMRS medium supplemented with 2-FL or 3-FL as the sole carbohydrate, as compared to growth in mMRS supplemented with sorbitol.

			Fold up-/down-regulation[a,b] during growth on:		
Gene ID	**Gene name**	**Function**	**2-FL vs Sorbitol**	**2-FL vs Lactose**	**3-FL vs Sorbitol**
BKKJ1_0067	BKKJ1_0067	ABC transporter substrate-binding protein	-8.12	-	-6.88
BKKJ1_0068	BKKJ1_0068	ABC transporter permease	-5.91	-	-6.61
BKKJ1_0069	BKKJ1_0069	ABC transporter permease	-5.71	-	-6.01
BKKJ1_0206	BKKJ1_0206	6-phosphogluconate dehydrogenase	-13.31	-	-11.41
BKKJ1_0207	BKKJ1_0207	hypothetical protein	-17.67	-	-15.14
BKKJ1_0208	BKKJ1_0208	putative gluconokinase	-17.21	-	-14.71
BKKJ1_0336	BKKJ1_0336	hypothetical protein	-16.99	-	-26.87
BKKJ1_0338	BKKJ1_0338	hypothetical protein	-27.69	-	-70.61
BKKJ1_0339	BKKJ1_0339	xylitol (sorbitol) dehydrogenase	-34.61	-	-76.79
BKKJ1_0340	BKKJ1_0340	transcriptional regulator	-36.34	-	-75.91
BKKJ1_0341	BKKJ1_0341	aldehyde-alcohol dehydrogenase 2	-22.03	-	-19.96
BKKJ1_0429	*prpO*	L-1 2-propanediol oxidoreductase	-	5.90	-
BKKJ1_2062	*galE*	Acetyltransferase	-	-	-
BKKJ1_2063	*galT1*	ABC transporter ATP-binding protein	-	-	-
BKKJ1_2064	*galT2*	ABC transporter permease	-	-	-
BKKJ1_2065	*galT3*	ABC transporter permease	-	-	-
BKKJ1_2066	*galS*	galactoside symporter	-	5.35	-
BKKJ1_2067	*galA*	beta-galactosidase	6.47	-	7.92
BKKJ1_2068	*galR*	Truncated LacI-type transcriptional regulator	-	-	-
BKKJ1_2069	*fucA1*	GH95 alpha-1 3/4-fucosidase	7.45	6.88	8.96
BKKJ1_2070	*fucA2*	GH29 alpha-1 3/4-fucosidase	11.04	10.43	13.07
BKKJ1_2071	*fucM*	L-fucose mutarotase	12.80	12.17	15.17
BKKJ1_2072	*fucG*	L-2-keto-3-deoxy-fuconate aldolase	13.14	10.86	15.94
BKKJ1_2073	*fucH*	L-fuconolactone hydrolase	14.15	8.21	16.93
BKKJ1_2074	*fucD*	L-keto-3-deoxy-fuconate-4-dehydrogenase	15.95	15.37	21.90
BKKJ1_2075	*fucF*	L-fuconate dehydratase	17.29	13.90	20.72
BKKJ1_2076	*fucS*	ABC transporter solute binding protein	19.68	30.20	26.53
BKKJ1_2077	*fucP1*	ABC transporter permease	20.79	22.82	25.66
BKKJ1_2078	*fucP2*	ABC transporter permease	21.89	30.41	29.15
BKKJ1_2079	*fucR*	LacI family transcriptional regulator	-	-	-

The level of transcription is shown as a fold-value of increase in transcription on each carbohydrate, as compared to a sorbitol or lactose control, with a cut-off of a minimum 5.0-fold increase in transcription.

a Based on comparative transcriptome analysis using *B. kashiwanohense* APCKJ1 grown on 1 % 2-FL or 3-FL compared to growth on sorbitol or lactose. Microarray data were obtained using *B. kashiwanohense* APCKJ1 grown on 1 % 2-FL or 3-FL and were compared with array data obtained when *B. kashiwanohense* APCKJ1 was grown on sorbitol or lactose as a control.

b The cut-off point 5.0-fold, with a P value of <0.001. a — sign menas that the obtained fold value was below the cut-off.

Table 5.2. *B. breve* UCC2003-*fucA1-fucSP1P2* genes that were transcriptionally upregulated during growth in mMRS medium supplemented with 2-FL or lactose as the sole carbohydrate, as compared to growth in mMRS supplemented with ribose.

			Fold upregulation[a,b] during growth on:	
Gene ID	**Gene name**	**Function**	**2-FL**	**Lactose**
Bbr_0526	*lntR*	LacI Transcriptional Regulator	-	-
Bbr_0527	*lntP1*	ABC transporter permease	2.61	3.81
Bbr_0528	*lntP2*	ABC transporter permease	3.35	2.12
Bbr_0529	*lntA*	GH42 Beta-galactosidase	2.48	2.21
Bbr_0530	*lntS*	ABC transporter solute binding protein	-	-
Bbr_1551	*lacS*	Galactoside symporter	18.72	22.27
Bbr_1552	*lacZ6*	GH2 Beta-galactosidase	12.98	9.05
Bbr_1553	*lacI*	LacI Transcriptional Regulator	-	-
Bbr_1739	*fcuT1*	Transporter protein	-	-
Bbr_1740	*fcuG*	L-2-keto-3-deoxy-fuconate aldolase	4.16	-
Bbr_1741	*fcuH*	L-fuconolactone hydrolase	4.69	-
Bbr_1742	*fcuP*	L-fucose permease	4.49	-
Bbr_1743	*fcuD*	L-keto-3-deoxy-fuconate-4-dehydrogenase	3.70	-
Bbr_1744	*fcuF*	L-fuconate dehydratase	7.12	-
Bbr_1745	*fcuR*	LacI Transcriptional Regulator	-	-
Bbr_1746	*fcuT2*	Transporter protein	-	-
Bbr_1747	*fcuT3*	ABC transporter ATP-binding protein	-	2.30
Bbr_1748	*fcuE*	Acetyltransferase	-	-

The level of transcription is shown as a fold-value of increase in transcription on each carbohydrate, as compared to a ribose control, with a cut-off of a minimum 2.0-fold increase in transcription.

a Based on comparative transcriptome analysis using *B. breve* UCC2003-*fuA1-fucSP1P2*grown on 1% 2-FL or lactose compared to growth on ribose. Microarray data were obtained using *B. breve* UCC2003-*fuA1-fucSP1P2* grown on 1% 2-FL or lactose and were compared with array data obtained when *B. breve* UCC2003-*fuA1-fucSP1P2* was grown on ribose as a control.

b The cut-off point 2.0-fold, with a P value of <0.001. The — sign menas that the obtained fold value was below the cut-off.

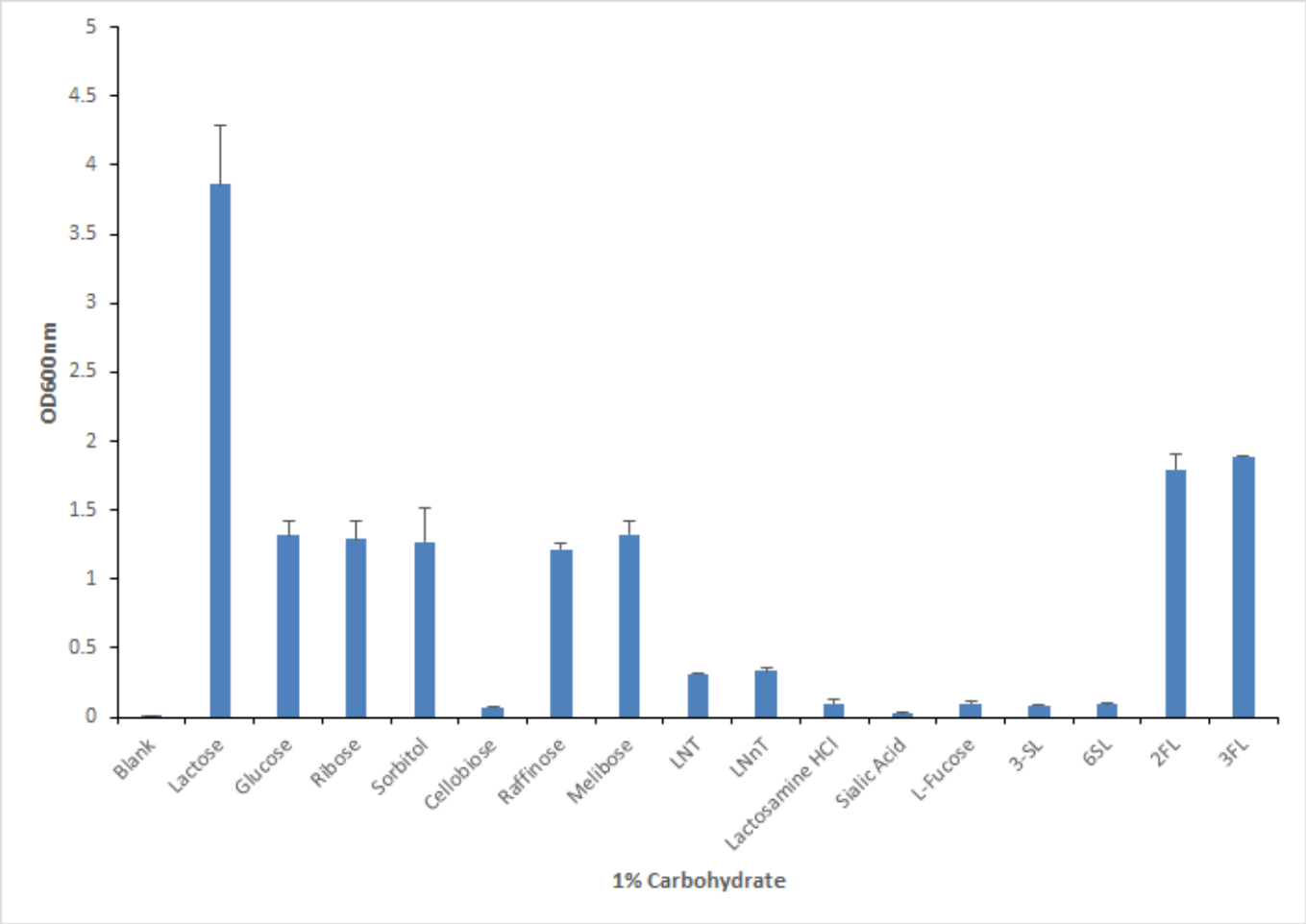


Figure 5.1. Final OD600nm values after 24 hours of growth of wild type *B. kashiwanohense* APCKJ1 in modified MRS containing a range of carbohydrates at 1% (wt/vol) as the sole carbon source. The results are the mean values obtained manually from two separate experiments (due to the limited availability of certain carbohydrates). Error bars represent the standard deviation.

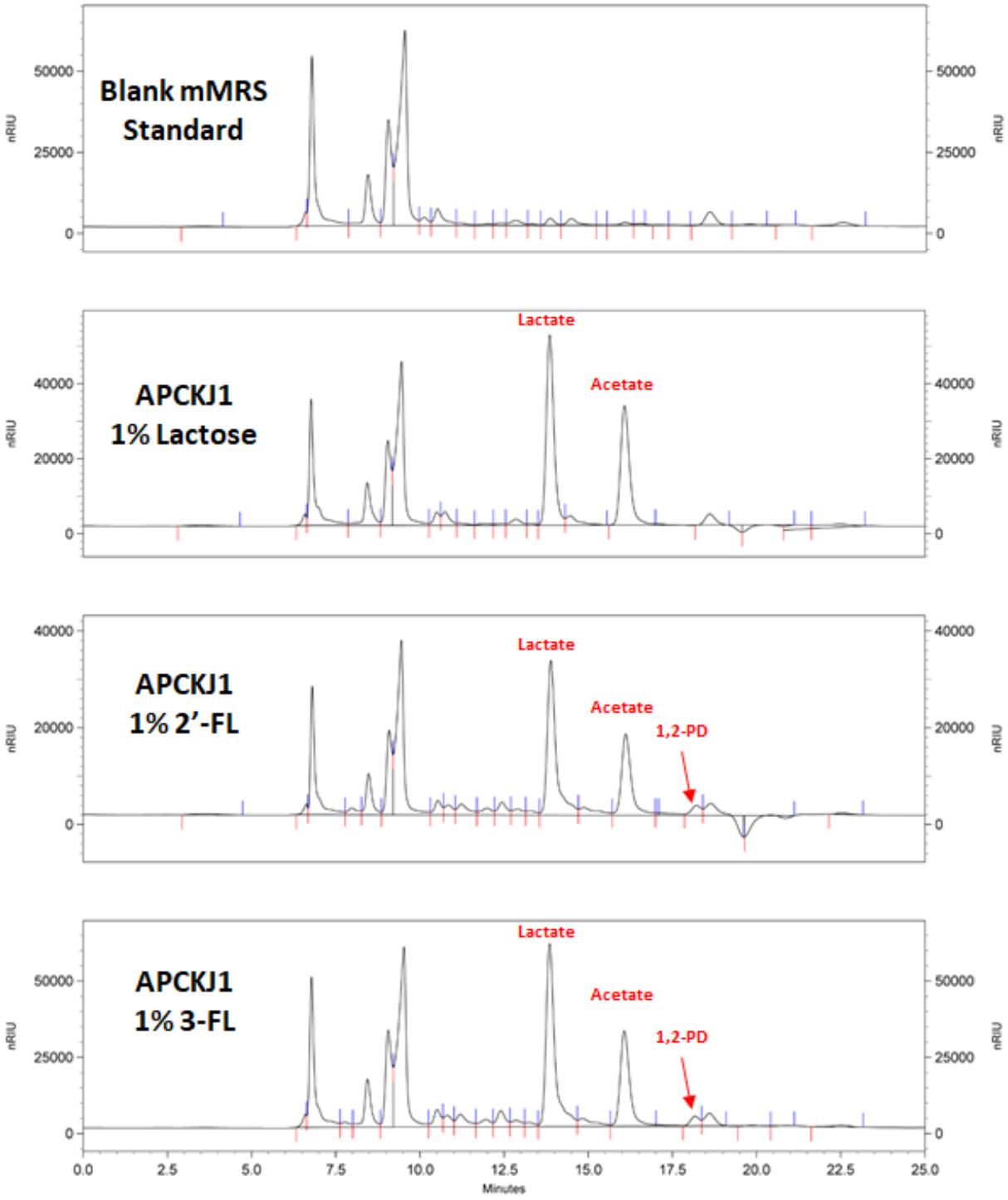


Figure 5.2. HPLC chromatogram profiles of fermentations of mMRS containing 1% lactose, 1% 2'-FL or 1% 3-FL inoculated with *B. kashiwanohense* APCKJ1, following 24 hours growth anaerobically at 37°C, as well as a blank mMRS control.

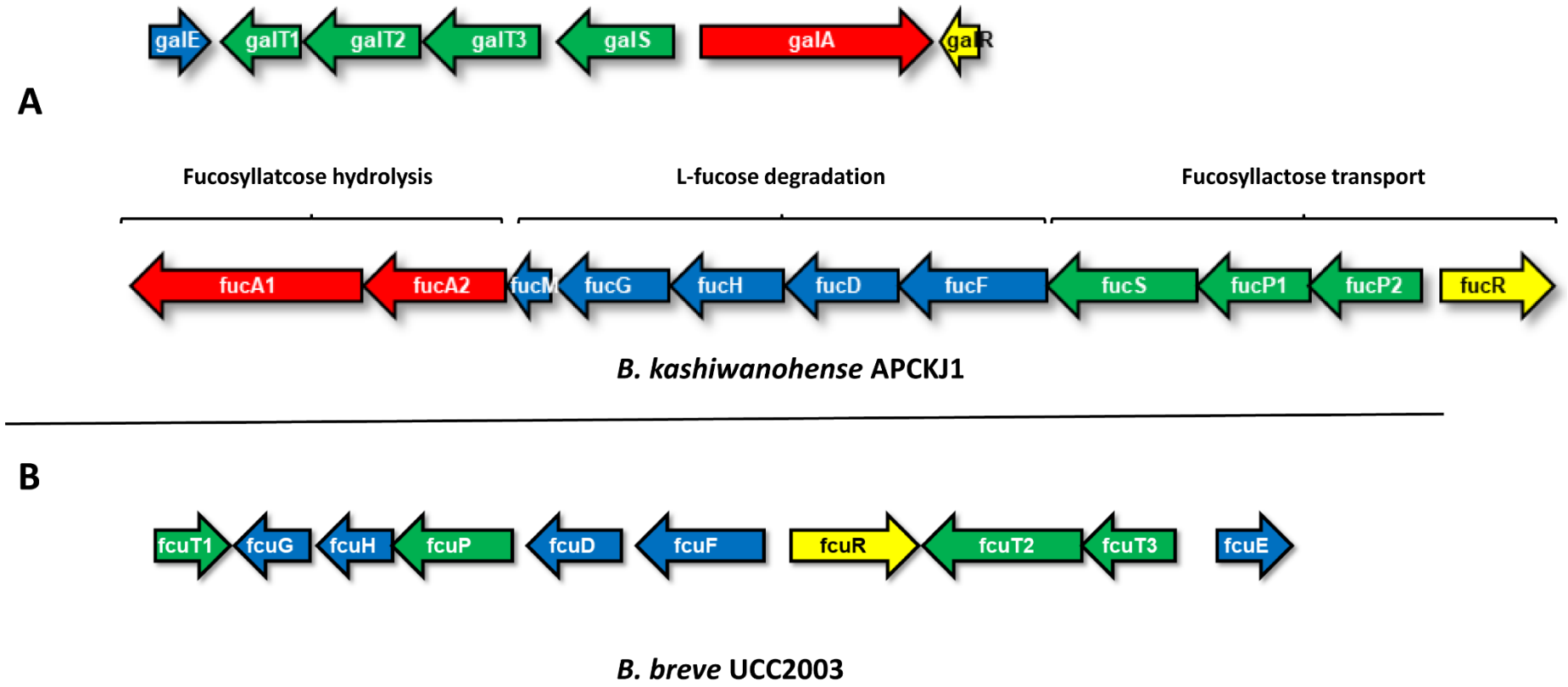
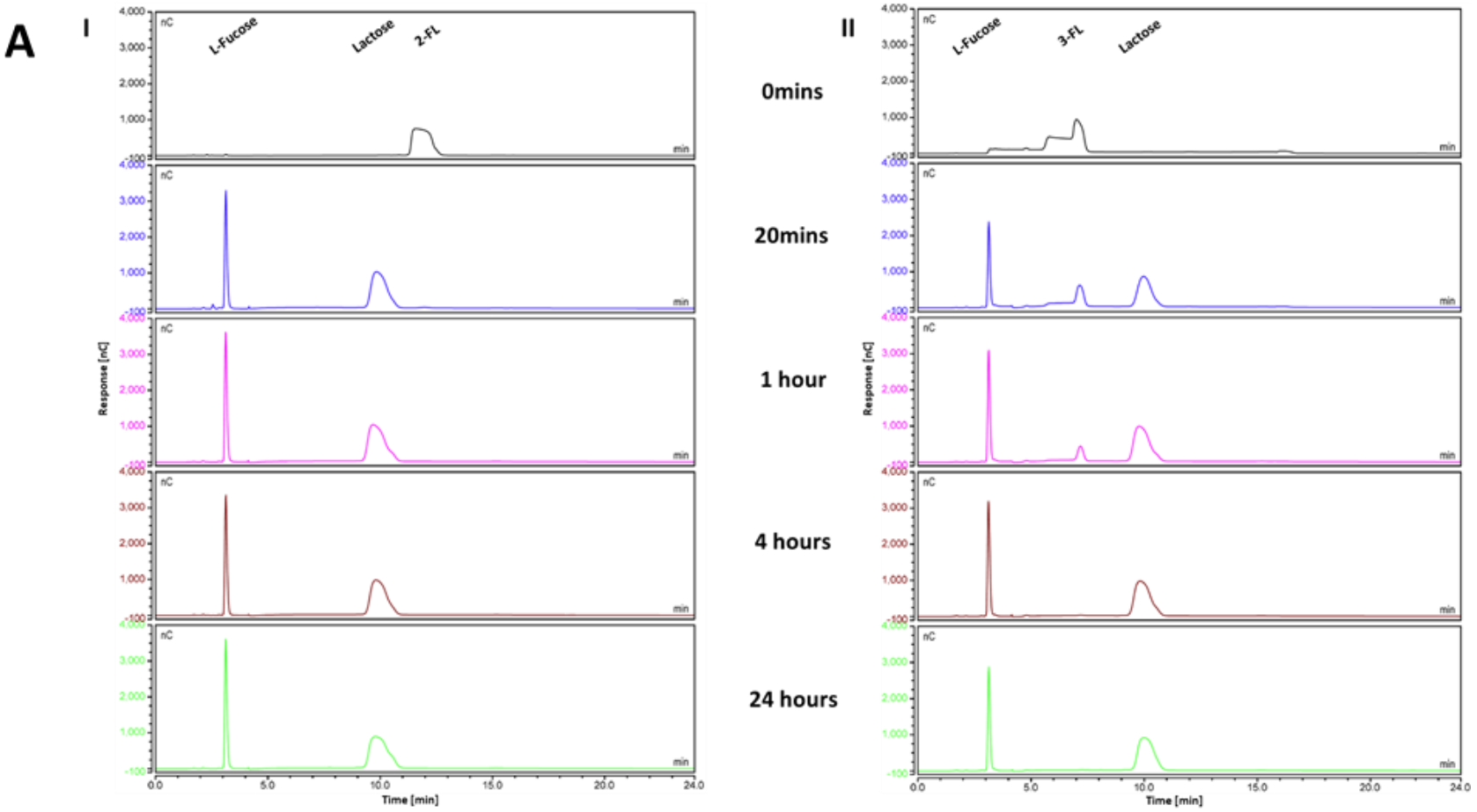


Figure 5.3. Schematic representation of the gene loci involved in the utilisation of: **(A)** 2-FL or 3-FL in *B. kashiwanohense* APCKJ1, and **(B)** L-fucose in *B. breve* UCC2003; as based on transcriptome analysis. The length of the arrows is proportional to the size of the open reading frame and the gene locus name, which is indicative of its putative function, is given inside the arrows. Genes shown in red are predicted to encode proteins with a hydrolytic function, genes shown in yellow are predicted to encode proteins with a regulatory function, genes shown in green are predicted to encode proteins with a transport function and genes shown in blue are predicted to encode proteins with another metabolic function.


A
I
II
L-Fucose
Lactose
2-FL
3-FL
0mins
20mins
1 hour
4 hours
24 hours
nC
min
Response [nC]
Time [min]

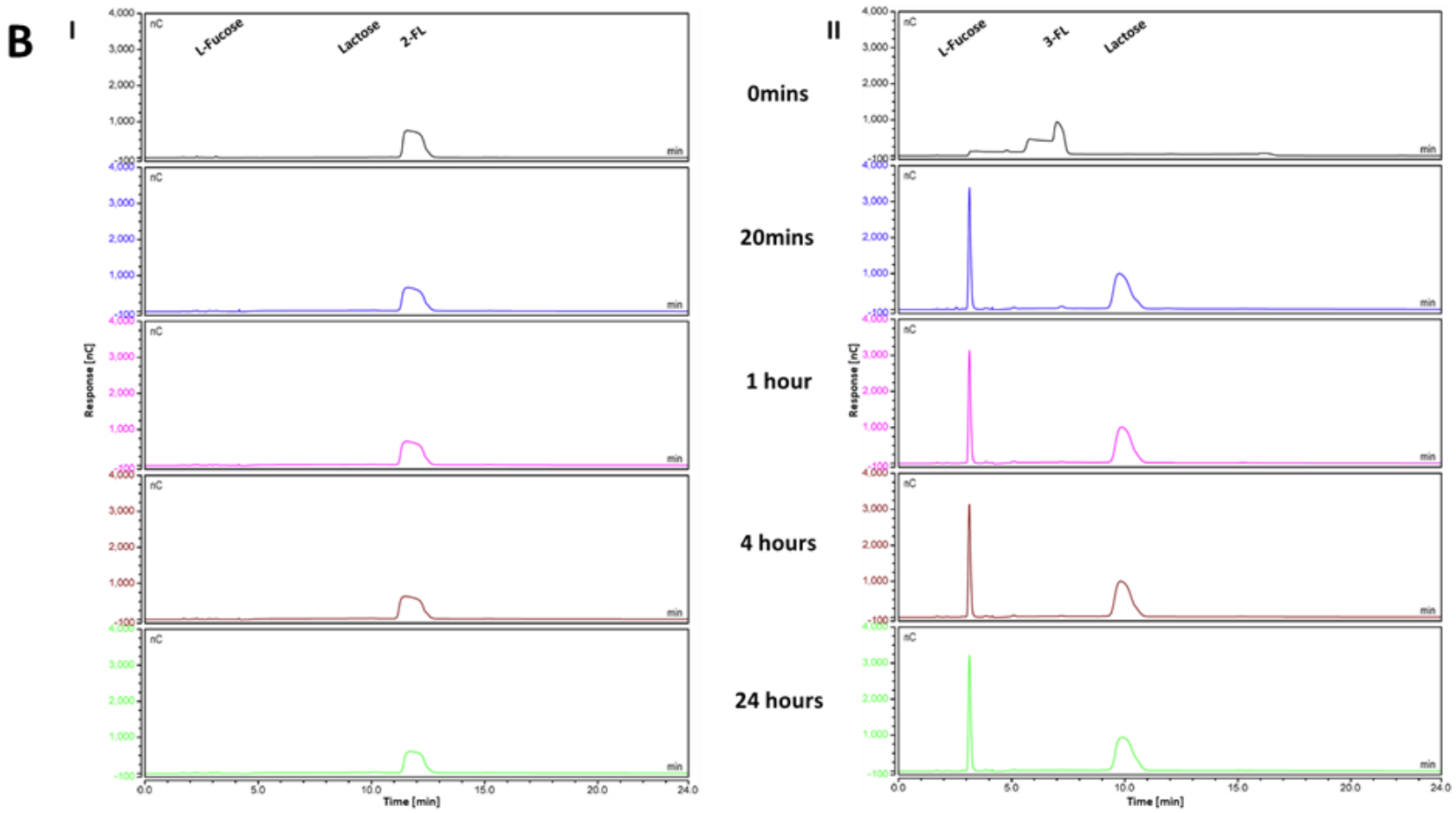


Figure 5.4. A. HPAEC-PAD chromatogram profiles of **(I)** 2-FL and **(II)** 3-FL, when incubated in MOPS buffer (pH7) with FucA1 at the time-points 0 minutes, 20 minutes, 1 hour, 4 hours and 24 hours. **B.** HPAEC-PAD chromatogram profiles of **(I)** 2-FL and **(II)** 3-FL, when incubated in MOPS buffer (pH7) with FucA2 at the time-points 0 minutes, 20 minutes, 1 hour, 4 hours and 24 hours.

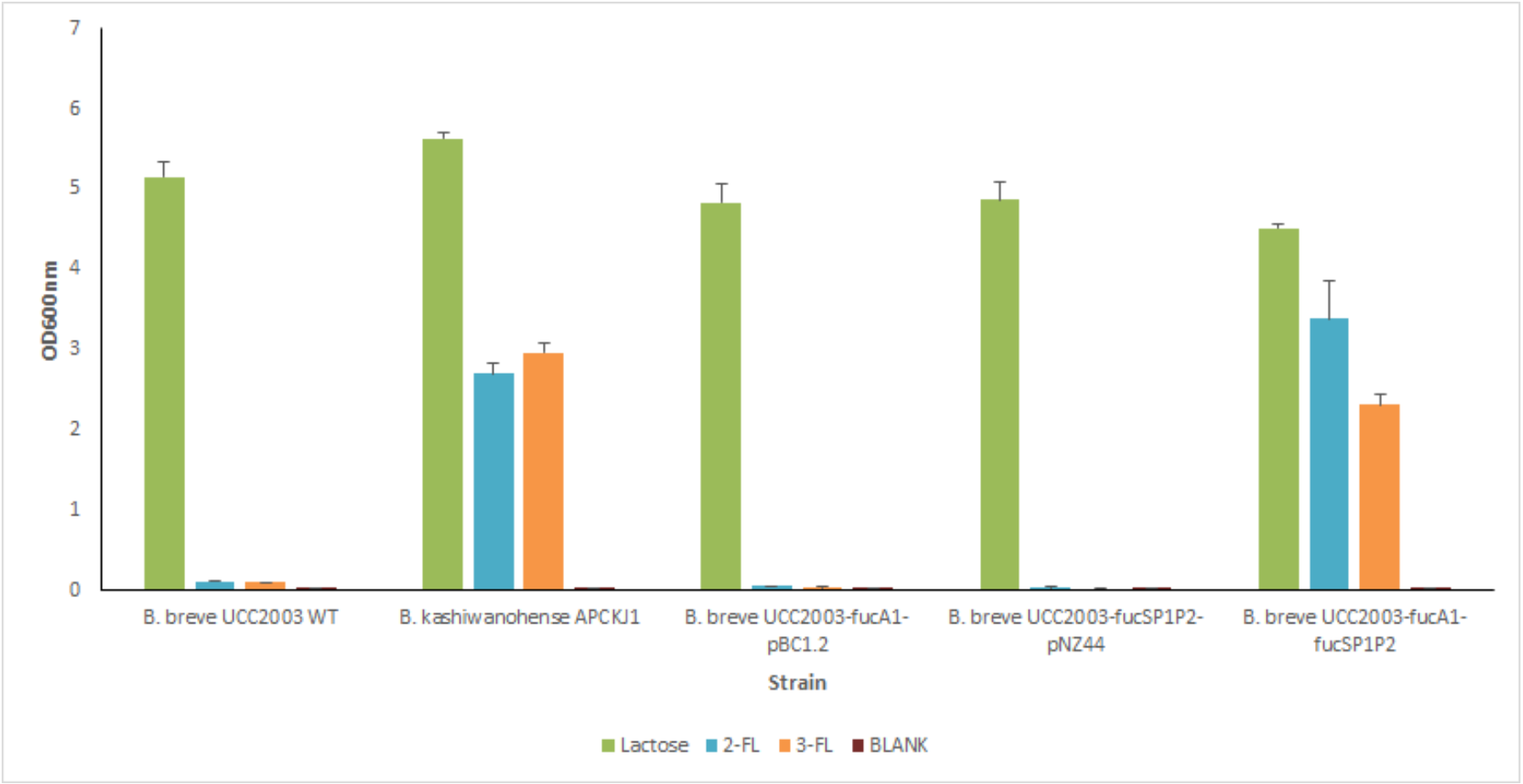


Figure 5.5. Final OD600nm values after 24 hours of growth of wild type *B. breve* UCC2003, wild type *B. kashiwanohense* APCKJ1 and recombinant *B. breve* UCC2003 strains in modified MRS containing 1% (wt/vol) lactose, 2-FL or 3-FL as the sole carbon source. The results are the mean values obtained manually from three separate experiments. Error bars represent the standard deviation.

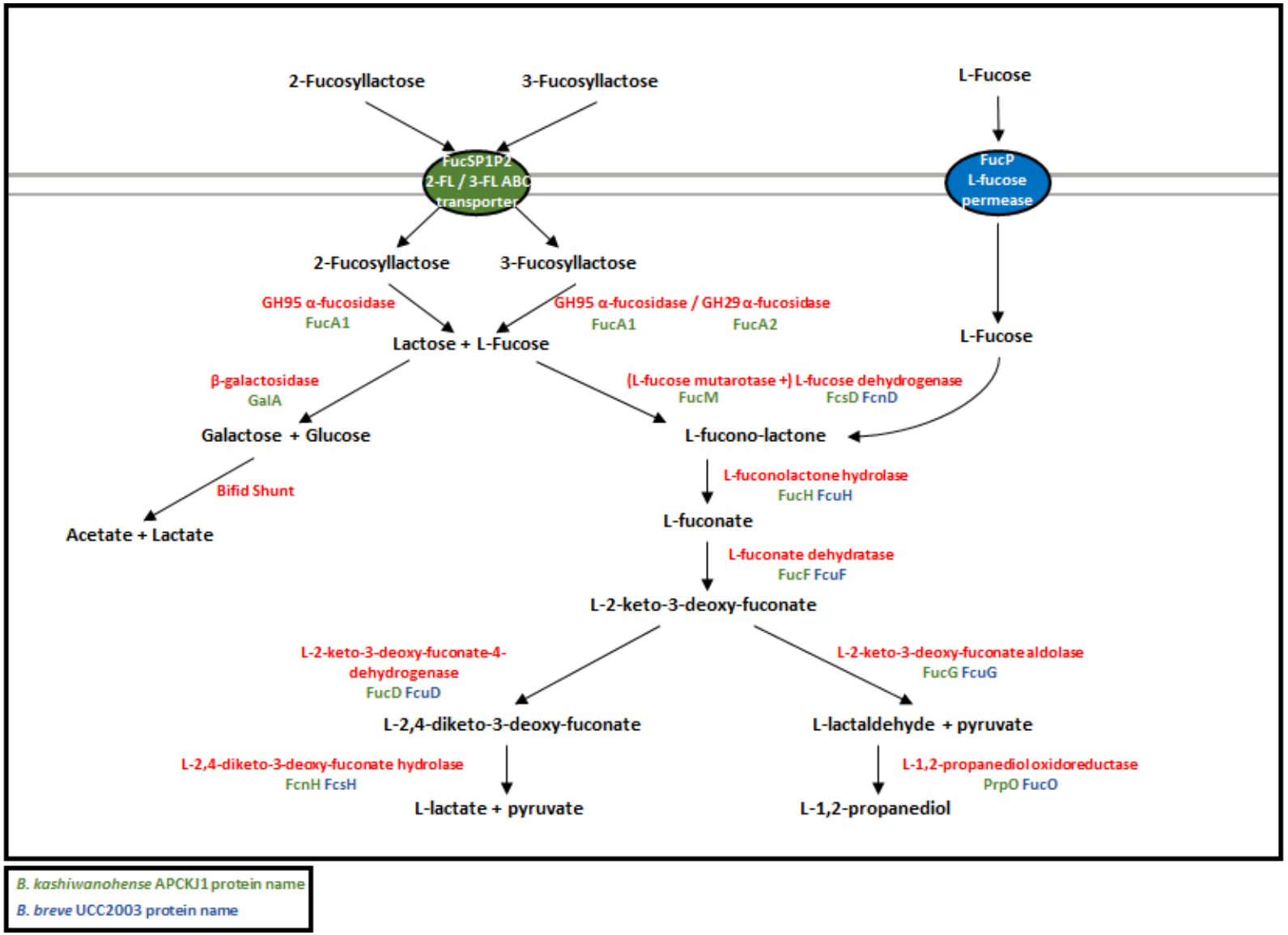


Figure 5.6. Schematic representation of the pathway for the utilisation of fucosyllactose in *B. kashiwanohense* APCKJ1.

5.8 Supplemental Material

Supplemental Table S5.1. Oligonucleotide primers used in this work.

Purpose	Primer	Sequence (5'-3')
Amplification of the ITS region for *Bifidobacterium* isolate species identification	Bifspp	ggtgtgaaagtccatcgct
	23Sbif	gtctgccaaggcatccacca
Cloning of BKKJ1_2069 in pNZ8150	2069F	tgcatc**cccggg**atgcatcaccatcaccatcaccatcaccatcacaaactcacattcgatggaatc
	2069R	tgcgca**tctaga**cgtaacggatataacgcaatacg
Cloning of BKKJ1_2070 in pNZ8150	2070F	tgcatc**cccggg**atgcatcaccatcaccatcaccatcaccatcacagcaatccaacaaatgatggt
	2070R	tgcgca**tctaga**agtttcatggtgacgtatcgcc
Cloning of 2379bp fragment containing BKKJ1_2069 into pNZ44	2069pNZ44F	ctggtc**ggtacc**ggcgatacgtcaccatgaaact
	2069pNZ44R	tgcgca**tctaga**ataacgcaatacgttaacgccg
Amplification of 2069pNZ44 without Cm^R cassette	pNZ44noCmF	ctggtc**actagt**ttgatatgcctcctaaatttttat
	pNZ44noCmR	cgagc**cccggg**tgagataatgccgactgtactttt
Amplification and cloning of $Strep^R$ cassette CNCMI4321_0987 in 2069pNZ44	$Strep^R$F	ctggtc**actagt**atactgcatttcccggtg
	$Strep^R$R	cgagc**cccggg**atcactgttcccgcctct
Cloning of BKKJ1_2076-2078 in pBC1.2	2076-78pBC1.2F	ctggtc**cccggg**gcccgctgttctctggatg
	2076-78pBC1.2R	tgcgca**tctaga**cgatgcgcttccttctttg

Restriction sites incorporated into oligonucleotide primer sequences are indicated in bold, and His-tag sequences incorporated into nucleotide primer sequences are indicated in italics.

Supplemental Table S5.2. Bacterial plasmids and strains used in this work.

Strain or plasmid	Relevant Features (antibiotic resistances are given in brackets)	Reference or Source
Strains		
Escherichia coli strains		
E. coli EC101	Cloning host, repA$^+$ (Km^r)	[63]
E. coli EC101-pNZ-M.BbrII + M.BbrIII	EC101 harbouring pNZ8048 derivative containing *bbrIIM* and *bbrIIIM* (Cm^r)	[61]
E .coli EC101-pBC1.2-fucT	XL1-blue containing pBC1.2-fucT (Cm^r)	This study
E. coli EC101-pNZ-M.BbrII + M.BbrIII +pNZ44-fucA1	EC101 harbouring pNZ8048 derivative containing *bbrIIM* ,*bbrIIIM* and pNZ44-fucA1-(Strep^r)	This study
Lactococcus lactis strains		
L. lactis NZ9000	MG1363, pepN::nisRK, nisin-inducible overexpression host	[68]
L. lactis NZ9700	Nisin-producing strain (Cm^r)	[68]
L. lactis NZ9000-pNZ-fucA1	NZ9000 containing pNZ-fucA1(Cm^r)	This study
L. lactis NZ9000-pNZ-fucA2	NZ9000 containing pNZ-fucA2 (Cm^r)	This study
L. lactis NZ9000-pNZ44-fucA1	NZ9000 containing pNZ44-fucA1 (Cm^r)	This study
Bifidobacterium sp. Strains		
B. kashiwanohense APCKJ1	Isolate from nursling stool	This study
B. breve UCC2003	Isolate from nursling stool	[62]
B. breve CNCMI4321	Isolate from nursling stool	Danone
B. breve UCC2003-fucA1-fucSP1P2	UCC2003 harbouring pNZ44-fucA1-Strep^R and pBC1.2-fucSP1P2 (Cm^r) (Strep^r)	This study
B. breve UCC2003-fucA1-pBC1.2	UCC2003 harbouring pNZ44-fucA1-Strep^R and pBC1.2 (Cm^r) (Strep^r)	This study
B. breve UCC2003-fucSP1P2-pNZ44-StrepR	UCC2003 harbouring pNZ44 -Strep^R and pBC1.2-fucSP1P2 (Cm^r) (Strep^r)	This study
Plasmids		
pBC1.2	pBC1-pSC101-(Cm^r)	[98]
pBC1.2-fucT	(Cm^r), pBC1-pSC101-Cmr harbouring fucT and its indigenous promoter	This study
pNZ8150	(Cm^r), nisin inducible translational fusion vector	[64]
pNZ-fucA1	(Cm^r), pNZ8150 derivative containing translational fusion of BKKJ_2069 encoding DNA fragment to nisin inducible promoter	This study
pNZ-fucA2	(Cm^r), pNZ8150 derivative containing translational fusion of BKKJ_2070 encoding DNA fragment to nisin inducible promoter	This study
pNZ44	(Cm^r), pNZ8048 containing constitutive p44 promoter from Lactococcal chromosome	[70]
pNZ44-fucA1	(Cm^r), pNZ44 harbouring BKKJ_2069 downstream of p44 promoter	This study
pNZ44-StrepR	**(S**trep^r), pNZ44 harbouring CNCMI4321_0987	This study
pNZ44-fucA1-StrepR	**(S**trep^r), pNZ44 harbouring BKKJ_2059 downstream of p44 promoter, and CNCMI4321_0987	This study

Cm^r, Km^r and Strep^r, resistance to chloramphenicol, kanamycin and streptomycin, respectively.

UCC, University College Cork Culture Collection.

Supplemental Table S5.3. *B. longum* subsp. *infantis* DSM20088 genes involved in the catabolism of fucosyllactose, and their homologs in *B. kashiwanohense* APCKJ1 and *B. breve* UCC2003, based on a blastP search of the APCKJ1 and UCC2003 genomes.

***B. longum* subsp. *infantis* DSM20088**		***B. kashiwanohense* APCKJ1**		***B. breve* UCC2003**	
Gene	**Predicted Function**	**Gene**	**BLASTP Result**	**Gene**	**BLASTP Result**
Blon_2335	GH95 α-fucosidase	BKKJ1_2069 ***fucA1***	77%, 1285, 0.0	Bbr_1288 ***fcuA***	77%, 1285, 0.0
Blon_2336	GH29 α-fucosidase	BKKJ1_2070 ***fucA2***	86%, 879, 0.0	-	-
Blon_2337	L-fucose mutarotase	BKKJ1_2071 ***fucM***	87%, 262, 2e-62	-	-
Blon_0406	L-fucose dehydrogenase	BKKJ1_0238* ***fcsD***	80%, 567, 0.0	Bbr_0398 ***fcnD***	94%, 664, 0.0
Blon_2306	L-fuconolactone hydrolase	BKKJ1_2073 ***fucH***	96%, 513, 0.0	Bbr_1741 ***fcuH***	50%, 248, 7e-87
Blon_2340	L-fuconate dehydratase	BKKJ1_2075 ***fucF***	98%, 867, 0.0	Bbr_1744 ***fcuF***	78%, 703, 0.0
Blon_2339	L-2-keto-3-deoxy-fuconate-4-dehydrogenase	BKKJ1_2074 ***fucD***	94%, 487, 0.0	Bbr_1743 ***fcuD***	76%, 390, 6e-143
Blon_2375	L-2,4-diketo-3-deoxy-fuconate hydrolase	BKKJ1_1881* ***fcnH***	87%, 495, 0.0	Bbr_1783 ***fcsH***	96%, 548, 0.0
Blon_2338	L-2-keto-3-deoxy-fuconate aldolase	BKKJ1_2072 ***fucG***	89%, 556, 0.0	Bbr_1740 ***fcuG***	70%, 429, 3e-157
Blon_0540	L-1,2-propanediol oxidoreductase	BKKJ1_0429 ***prpO***	95%, 750, 0.0	Bbr_1505 ***fucO***	98%, 765, 0.0

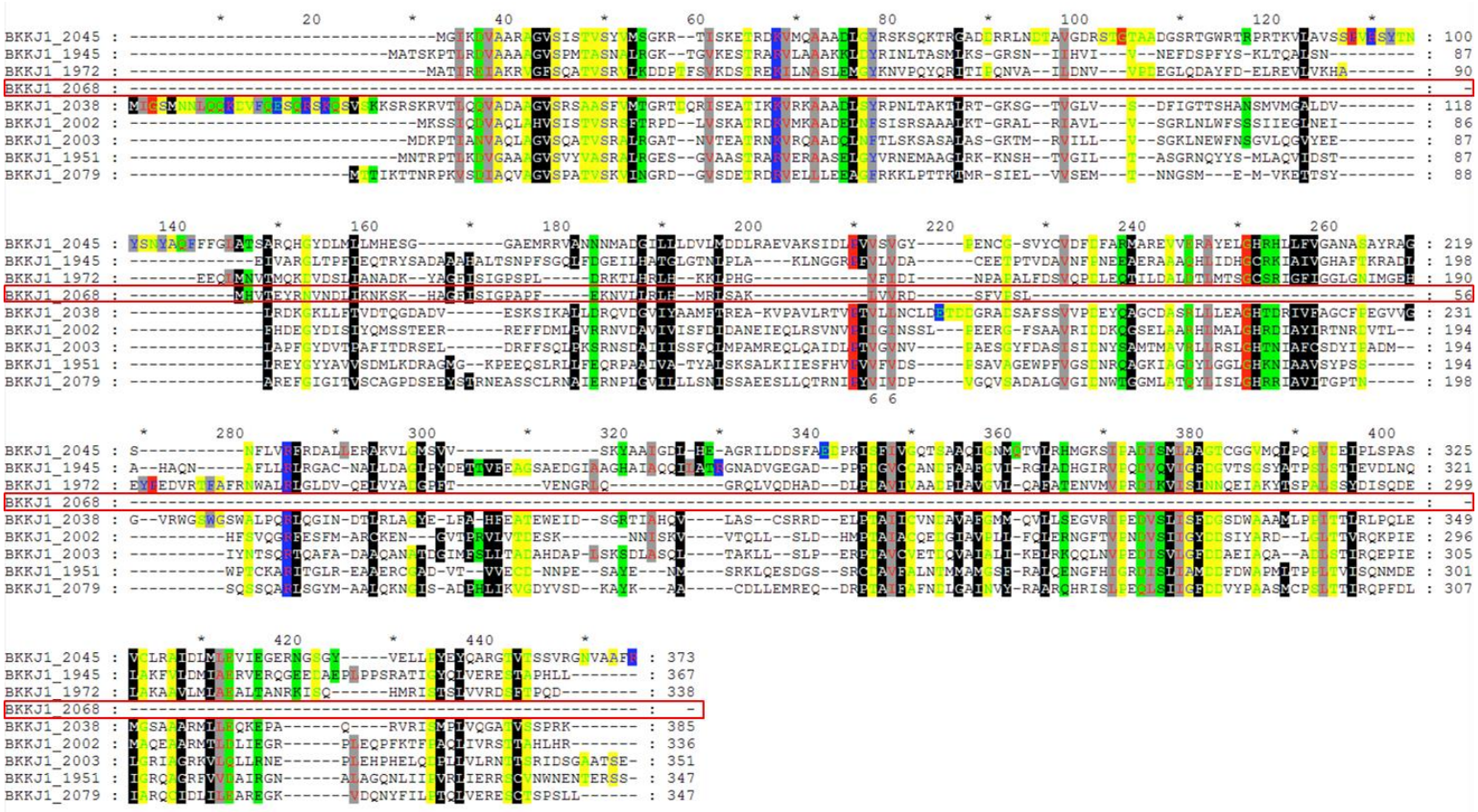
Values in the BLASTP column represent match identity, Bit Score and e-value.

Cut-off values of a minimum Bit Score of 200 bits, a minimum identity of 50% coverage and minimum e-value of 0.0001 were employed.

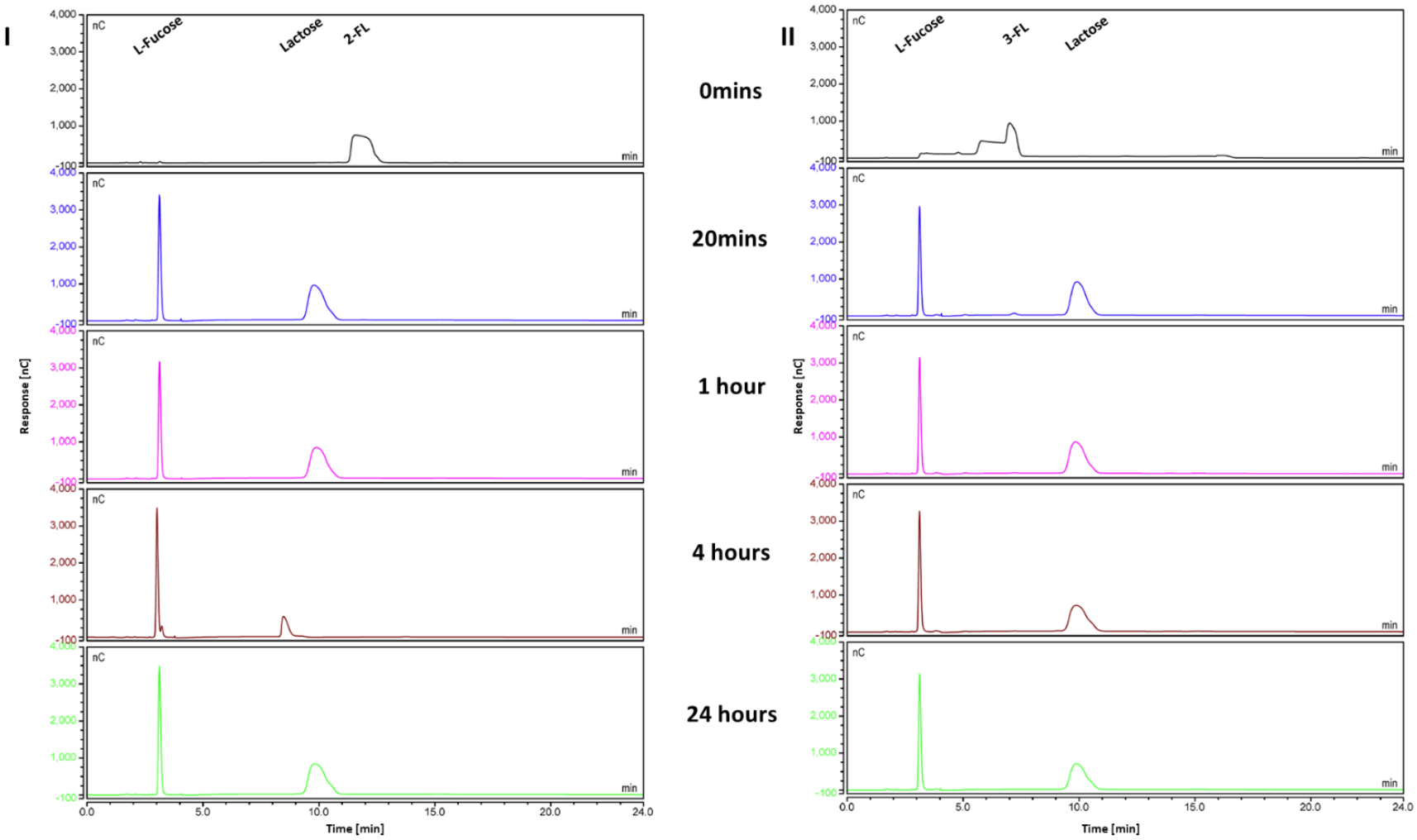
*Denotes genes not upregulated in transcription during growth on 2-FL or 3-FL.

Supplemental Table S5.4. Scores for APCKJ1 genome sequencing reads and quality.

Sequencing Project	Number of Reads	Coverage	Mapped N50 (Long Reads)	Mapped N50 (Short Reads)	Polymerase Read Quality
B. kashiwanohense APCKJ1 Genome	61974	245.02x	17573bp	7971bp	0.853



Supplemental Figure S5.1. Multiple sequence alignment of the amino acid sequences of the *B. kashiwanohense* APCKJ1 predicted truncated lacI-encoding gene *galR* (BKKJ1_2068; highlighted in red) and other predicted LacI-encoding genes in the APCKJ1 genome. A multiple sequence aligmnent was generated using the online resource Clustal Omega, with the default settings, and this alignment was visualised with Genedoc v 2.7.0, using physiochemical display mode (amino acids coloured based on shared physical and chemical properties). The alignment clearly shows a large proportion of the *galR* sequence missing when compared witht the sequences of other LacI regulator-encoding genes.



Supplemental Figure S5.2. HPAEC-PAD chromatogram profiles of **(I)** 2-FL and **(II)** 3-FL, when incubated in MOPS buffer (pH7) with FucA1 and FucA2 together at the time-points 0 minutes, 20minutes, 1 hour, 4 hours and 24 hours.

5.9 References

1. **Engfer, M.B., B. Stahl, B. Finke, G. Sawatzki, and H. Daniel,** *Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract.* The American Journal of Clinical Nutrition, **2000**. **71**(6): p. 1589-1596.
2. **Kunz, C., S. Rudloff, W. Baier, N. Klein, and S. Strobel,** *OLIGOSACCHARIDES IN HUMAN MILK: Structural, Functional, and Metabolic Aspects.* Annual Review of Nutrition, **2000**. **20**(1): p. 699-722.
3. **Brüssow, H.,** *Human microbiota: 'The philosophers have only interpreted the world in various ways. The point, however, is to change it'.* Microbial Biotechnology, **2015**. **8**(1): p. 11-12.
4. **Bezirtzoglou, E., A. Tsiotsias, and G.W. Welling,** *Microbiota profile in feces of breast- and formula-fed newborns by using fluorescence in situ hybridization (FISH).* Anaerobe, **2011**. **17**(6): p. 478-82.
5. **Charbonneau, Mark R., D. O'Donnell, Laura V. Blanton, Sarah M. Totten, Jasmine C.C. Davis, Michael J. Barratt, J. Cheng, J. Guruge, M. Talcott, James R. Bain, Michael J. Muehlbauer, O. Ilkayeva, C. Wu, T. Struckmeyer, D. Barile, C. Mangani, J. Jorgensen, Y.-m. Fan, K. Maleta, Kathryn G. Dewey, P. Ashorn, Christopher B. Newgard, C. Lebrilla, David A. Mills, and Jeffrey I. Gordon,** *Sialylated Milk Oligosaccharides Promote Microbiota-Dependent Growth in Models of Infant Undernutrition.* Cell. **164**(5): p. 859-871.
6. **Wang, M., M. Li, S. Wu, C.B. Lebrilla, R.S. Chapkin, I. Ivanov, and S.M. Donovan,** *Fecal Microbiota Composition of Breast-fed Infants is Correlated with Human Milk Oligosaccharides Consumed.* Journal of pediatric gastroenterology and nutrition, **2015**. **60**(6): p. 825-833.
7. **Asakuma, S., E. Hatakeyama, T. Urashima, E. Yoshida, T. Katayama, K. Yamamoto, H. Kumagai, H. Ashida, J. Hirose, and M. Kitaoka,** *Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria.* J Biol Chem, **2011**. **286**(40): p. 34583-92.
8. **Sela, D.A. and D.A. Mills,** *Nursing our microbiota: molecular linkages between bifidobacteria and milk oligosaccharides.* Trends Microbiol, **2010**. **18**(7): p. 298-307.
9. **Sela, D.A.,** *Bifidobacterial utilization of human milk oligosaccharides.* International Journal of Food Microbiology, **2011**. **149**(1): p. 58-64.

10. **Gibson, G.R. and M.B. Roberfroid,** *Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics.* J Nutr, **1995**. **125**(6): p. 1401-12.
11. **Turroni, F., C. Peano, D.A. Pass, E. Foroni, M. Severgnini, M.J. Claesson, C. Kerr, J. Hourihane, D. Murray, F. Fuligni, M. Gueimonde, A. Margolles, G. De Bellis, P.W. O'Toole, D. van Sinderen, J.R. Marchesi, and M. Ventura,** *Diversity of Bifidobacteria within the Infant Gut Microbiota.* PLoS ONE, **2012**. **7**(5): p. e36957.
12. **Simone, M., C. Gozzoli, A. Quartieri, G. Mazzola, D. Di Gioia, A. Amaretti, S. Raimondi, and M. Rossi,** *The Probiotic Bifidobacterium breve B632 Inhibited the Growth of Enterobacteriaceae within Colicky Infant Microbiota Cultures.* BioMed Research International, **2014**. **2014**(301053): p. 7.
13. **Di Gioia, D., I. Aloisio, G. Mazzola, and B. Biavati,** *Bifidobacteria: their impact on gut microbiota composition and their applications as probiotics in infants.* Appl Microbiol Biotechnol, **2014**. **98**(2): p. 563-77.
14. **Milani, C., S. Duranti, F. Bottacini, E. Casey, F. Turroni, J. Mahony, C. Belzer, S. Delgado Palacio, S. Arboleya Montes, L. Mancabelli, G.A. Lugli, J.M. Rodriguez, L. Bode, W. de Vos, M. Gueimonde, A. Margolles, D. van Sinderen, and M. Ventura,** *The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota.* Microbiology and Molecular Biology Reviews, **2017**. **81**(4): p. e00036-17.
15. **LoCascio, R.G., M.R. Ninonuevo, S.L. Freeman, D.A. Sela, R. Grimm, C.B. Lebrilla, D.A. Mills, and J.B. German,** *Glycoprofiling of Bifidobacterial Consumption of Human Milk Oligosaccharides Demonstrates Strain Specific, Preferential Consumption of Small Chain Glycans Secreted in Early Human Lactation.* Journal of Agricultural and Food Chemistry, **2007**. **55**(22): p. 8914-8919.
16. **Ward, R.E., M. Ninonuevo, D.A. Mills, C.B. Lebrilla, and J.B. German,** *In vitro fermentability of human milk oligosaccharides by several strains of bifidobacteria.* Mol Nutr Food Res, **2007**. **51**(11): p. 1398-405.
17. **Kitaoka, M.,** *Bifidobacterial enzymes involved in the metabolism of human milk oligosaccharides.* Adv Nutr, **2012**. **3**(3): p. 422s-9s.
18. **Wada, J., T. Ando, M. Kiyohara, H. Ashida, M. Kitaoka, M. Yamaguchi, H. Kumagai, T. Katayama, and K. Yamamoto,** *Bifidobacterium bifidum lacto-N-biosidase, a critical enzyme for the degradation of human milk oligosaccharides with a type 1 structure.* Appl Environ Microbiol, **2008**. **74**(13): p. 3996-4004.

19. **Moller, P.L., F. Jorgensen, O.C. Hansen, S.M. Madsen, and P. Stougaard,** *Intra- and extracellular beta-galactosidases from Bifidobacterium bifidum and B. infantis: molecular cloning, heterologous expression, and comparative characterization.* Appl Environ Microbiol, **2001**. **67**(5): p. 2276-83.
20. **Sakurama, H., M. Kiyohara, J. Wada, Y. Honda, M. Yamaguchi, S. Fukiya, A. Yokota, H. Ashida, H. Kumagai, M. Kitaoka, K. Yamamoto, and T. Katayama,** *Lacto-N-biosidase Encoded by a Novel Gene of Bifidobacterium longum Subspecies longum Shows Unique Substrate Specificity and Requires a Designated Chaperone for Its Active Expression.* The Journal of Biological Chemistry, **2013**. **288**(35): p. 25194-25206.
21. **Miwa, M., T. Horimoto, M. Kiyohara, T. Katayama, M. Kitaoka, H. Ashida, and K. Yamamoto,** *Cooperation of beta-galactosidase and beta-N-acetylhexosaminidase from bifidobacteria in assimilation of human milk oligosaccharides with type 2 structure.* Glycobiology, **2010**. **20**(11): p. 1402-9.
22. **Ashida, H., A. Miyake, M. Kiyohara, J. Wada, E. Yoshida, H. Kumagai, T. Katayama, and K. Yamamoto,** *Two distinct alpha-L-fucosidases from Bifidobacterium bifidum are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates.* Glycobiology, **2009**. **19**(9): p. 1010-7.
23. **Kiyohara, M., K. Tanigawa, T. Chaiwangsri, T. Katayama, H. Ashida, and K. Yamamoto,** *An exo-alpha-sialidase from bifidobacteria involved in the degradation of sialyloligosaccharides in human milk and intestinal glycoconjugates.* Glycobiology, **2011**. **21**(4): p. 437-47.
24. **Sela, D.A., J. Chapman, A. Adeuya, J.H. Kim, F. Chen, T.R. Whitehead, A. Lapidus, D.S. Rokhsar, C.B. Lebrilla, J.B. German, N.P. Price, P.M. Richardson, and D.A. Mills,** *The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome.* Proc Natl Acad Sci U S A, **2008**. **105**(48): p. 18964-9.
25. **Yoshida, E., H. Sakurama, M. Kiyohara, M. Nakajima, M. Kitaoka, H. Ashida, J. Hirose, T. Katayama, K. Yamamoto, and H. Kumagai,** *Bifidobacterium longum subsp. infantis uses two different beta-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides.* Glycobiology, **2012**. **22**(3): p. 361-8.
26. **Garrido, D., S. Ruiz-Moyano, and D.A. Mills,** *Release and utilization of N-acetyl-D-glucosamine from human milk oligosaccharides by Bifidobacterium longum subsp. infantis.* Anaerobe, **2012**. **18**(4): p. 430-5.

27. **Xiao, J.Z., S. Takahashi, M. Nishimoto, T. Odamaki, T. Yaeshima, K. Iwatsuki, and M. Kitaoka,** *Distribution of in vitro fermentation ability of lacto-N-biose I, a major building block of human milk oligosaccharides, in bifidobacterial strains.* Appl Environ Microbiol, **2010**. **76**(1): p. 54-9.
28. **James, K., M.O. Motherway, F. Bottacini, and D. van Sinderen,** *Bifidobacterium breve UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways.* Sci Rep, **2016**. **6**: p. 38560.
29. **Sela, D.A., Y. Li, L. Lerno, S. Wu, A.M. Marcobal, J.B. German, X. Chen, C.B. Lebrilla, and D.A. Mills,** *An Infant-associated Bacterial Commensal Utilizes Breast Milk Sialyloligosaccharides.* The Journal of Biological Chemistry, **2011**. **286**(14): p. 11909-11918.
30. **Sela, D.A., D. Garrido, L. Lerno, S. Wu, K. Tan, H.-J. Eom, A. Joachimiak, C.B. Lebrilla, and D.A. Mills,** *Bifidobacterium longum subsp. infantis ATCC 15697 α-Fucosidases Are Active on Fucosylated Human Milk Oligosaccharides.* Applied and Environmental Microbiology, **2012**. **78**(3): p. 795-803.
31. **Bunesova, V., C. Lacroix, and C. Schwab,** *Fucosyllactose and L-fucose utilization of infant Bifidobacterium longum and Bifidobacterium kashiwanohense.* BMC Microbiology, **2016**. **16**(1): p. 248.
32. **Vazquez-Gutierrez, P., C. Lacroix, T. Jaeggi, C. Zeder, M.B. Zimmerman, and C. Chassard,** *Bifidobacteria strains isolated from stools of iron deficient infants can efficiently sequester iron.* BMC Microbiol, **2015**. **15**(1): p. 3.
33. **Morita, H., A. Nakano, H. Onoda, H. Toh, K. Oshima, H. Takami, M. Murakami, S. Fukuda, T. Takizawa, T. Kuwahara, H. Ohno, S. Tanabe, and M. Hattori,** *Bifidobacterium kashiwanohense sp. nov., isolated from healthy infant faeces.* Int J Syst Evol Microbiol, **2011**. **61**(Pt 11): p. 2610-5.
34. **McGuire, M.K., C.L. Meehan, M.A. McGuire, J.E. Williams, J. Foster, D.W. Sellen, E.W. Kamau-Mbuthia, E.W. Kamundia, S. Mbugua, S.E. Moore, A.M. Prentice, L.J. Kvist, G.E. Otoo, S.L. Brooker, W.J. Price, B. Shafii, C. Placek, K.A. Lackey, B. Robertson, S. Manzano, L. Ruíz, J.M. Rodríguez, R.G. Pareja, and L. Bode,** *What's normal? Oligosaccharide concentrations and profiles in milk produced by healthy women vary geographically.* The American Journal of Clinical Nutrition, **2017**. **105**(5): p. 1086-1100.

35. **Bode, L.,** *Human milk oligosaccharides: every baby needs a sugar mama.* Glycobiology, **2012**. **22**(9): p. 1147-62.
36. **Bode, L.,** *Human milk oligosaccharides: Every baby needs a sugar mama.* Glycobiology, **2012**. **22**(9): p. 1147-1162.
37. **Matsuki, T., K. Yahagi, H. Mori, H. Matsumoto, T. Hara, S. Tajima, E. Ogawa, H. Kodama, K. Yamamoto, T. Yamada, S. Matsumoto, and K. Kurokawa,** *A key genetic factor for fucosyllactose utilization affects infant gut microbiota development.* Nat Commun, **2016**. **7**: p. 11939.
38. **Schwab, C., H.-J. Ruscheweyh, V. Bunesova, V.T. Pham, N. Beerenwinkel, and C. Lacroix,** *Trophic Interactions of Infant Bifidobacteria and Eubacterium hallii during L-Fucose and Fucosyllactose Degradation.* Frontiers in Microbiology, **2017**. **8**: p. 95.
39. **Scardovi, V.**, *Genus Bifidobacterium*, in *Bergey's manual of systematic bacteriology*, P. H. A. Sneath and M.E.S. N. S. Mair, and J. G. Holt, Editors. **1986**, Williams and Wilkins: Baltimore, Md. p. 1418–1434.
40. **Turroni, F., E. Foroni, P. Pizzetti, V. Giubellini, A. Ribbera, P. Merusi, P. Cagnasso, B. Bizzarri, G.L. de'Angelis, F. Shanahan, D. van Sinderen, and M. Ventura,** *Exploring the Diversity of the Bifidobacterial Population in the Human Intestinal Tract.* Applied and Environmental Microbiology, **2009**. **75**(6): p. 1534-1545.
41. **Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman,** *Basic local alignment search tool.* J Mol Biol, **1990**. **215**(3): p. 403-10.
42. **Lagesen, K., P. Hallin, E.A. Rodland, H.H. Staerfeldt, T. Rognes, and D.W. Ussery,** *RNAmmer: consistent and rapid annotation of ribosomal RNA genes.* Nucleic Acids Res, **2007**. **35**(9): p. 3100-8.
43. **Roberts, R.J., T. Vincze, J. Posfai, and D. Macelis,** *REBASE--a database for DNA restriction and modification: enzymes, genes and genomes.* Nucleic Acids Res, **2015**. **43**(Database issue): p. D298-9.
44. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M.A. Rajandream, and B. Barrell,** *Artemis: sequence visualization and annotation.* Bioinformatics, **2000**. **16**(10): p. 944-5.
45. **Morita, H., H. Toh, A. Nakano, K. Oshima, M. Takagi, W. Suda, S. Tanabe, and M. Hattori,** *Complete Genome Sequence of Bifidobacterium kashiwanohense JCM 15439(T), Isolated from Feces from a Healthy Japanese Infant.* Genome Announcements, **2015**. **3**(2): p. e00255-15.

46. **Vazquez-Gutierrez, P., C. Lacroix, C. Chassard, J. Klumpp, C. Jans, and M.J.A. Stevens,** *Complete and Assembled Genome Sequence of Bifidobacterium kashiwanohense PV20-2, Isolated from the Feces of an Anemic Kenyan Infant.* Genome Announcements, **2015**. **3**(1): p. e01467-14.

47. **O'Connell Motherway, M., A. Zomer, S.C. Leahy, J. Reunanen, F. Bottacini, M.J. Claesson, F. O'Brien, K. Flynn, P.G. Casey, J.A. Moreno Munoz, B. Kearney, A.M. Houston, C. O'Mahony, D.G. Higgins, F. Shanahan, A. Palva, W.M. de Vos, G.F. Fitzgerald, M. Ventura, P.W. O'Toole, and D. van Sinderen,** *Functional genome analysis of Bifidobacterium breve UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor.* Proceedings of the National Academy of Sciences, **2011**. **108**(27): p. 11217-11222.

48. **Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman,** *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res, **1997**. **25**(17): p. 3389-402.

49. **Sievers, F., A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J.D. Thompson, and D.G. Higgins,** *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.* Mol Syst Biol, **2011**. **7**(1): p. 539.

50. **Goujon, M., H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern, and R. Lopez,** *A new bioinformatics analysis tools framework at EMBL-EBI.* Nucleic Acids Res, **2010**. **38**(Web Server issue): p. W695-9.

51. **Nicholas, K.B. and H.B.J. Nicholas,** *GeneDoc: a tool for editing and annotating multiple sequence alignments.* Distributed by the author. http://www.psc.edu/biomed/genedoc, **1997**.

52. **Terzaghi, B.E. and W.E. Sandine,** *Improved medium for lactic streptococci and their bacteriophages.* Appl Microbiol, **1975**. **29**(6): p. 807-13.

53. **Sambrook J, F.E., Maniatis T.**, *Molecular cloning: a laboratory manual, 2nd ed.* **1989**: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

54. **Watson, D., M. O'Connell Motherway, M.H. Schoterman, R.J. van Neerven, A. Nauta, and D. van Sinderen,** *Selective carbohydrate utilization by lactobacilli and bifidobacteria.* J Appl Microbiol, **2013**. **114**(4): p. 1132-46.

55. **Zomer, A., M. Fernandez, B. Kearney, G.F. Fitzgerald, M. Ventura, and D. van Sinderen,** *An interactive regulatory network controls stress response in Bifidobacterium breve UCC2003.* J Bacteriol, **2009**. **191**(22): p. 7039-49.

56. **Garcia de la Nava, J., D.F. Santaella, J. Cuenca Alba, J. Maria Carazo, O. Trelles, and A. Pascual-Montano,** *Engene: the processing and exploratory analysis of gene expression data.* Bioinformatics, **2003**. **19**(5): p. 657-8.

57. **van Hijum, S., A. de Jong, R. Baerends, H. Karsens, N. Kramer, R. Larsen, C. den Hengst, C. Albers, J. Kok, and O. Kuipers,** *A generally applicable validation scheme for the assessment of factors involved in reproducibility and quality of DNA-microarray data.* BMC Genomics, **2005**. **6**(1): p. 1-10.

58. **van Hijum, S.A., J. Garcia de la Nava, O. Trelles, J. Kok, and O.P. Kuipers,** *MicroPreP: a cDNA microarray data pre-processing framework.* Appl Bioinformatics, **2003**. **2**(4): p. 241-4.

59. **Long, A.D., H.J. Mangalam, B.Y. Chan, L. Tolleri, G.W. Hatfield, and P. Baldi,** *Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in Escherichia coli K12.* J Biol Chem, **2001**. **276**(23): p. 19937-44.

60. **O'Riordan, K. and G.F. Fitzgerald,** *Molecular characterisation of a 5.75-kb cryptic plasmid from Bifidobacterium breve NCFB 2258 and determination of mode of replication.* FEMS Microbiol Lett, **1999**. **174**(2): p. 285-94.

61. **O'Connell Motherway, M., J. O'Driscoll, G.F. Fitzgerald, and D. Van Sinderen,** *Overcoming the restriction barrier to plasmid transformation and targeted mutagenesis in Bifidobacterium breve UCC2003.* Microb Biotechnol, **2009**. **2**(3): p. 321-32.

62. **Maze, A., M. O'Connell-Motherway, G.F. Fitzgerald, J. Deutscher, and D. van Sinderen,** *Identification and characterization of a fructose phosphotransferase system in Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2007**. **73**(2): p. 545-53.

63. **Law, J., G. Buist, A. Haandrikman, J. Kok, G. Venema, and K. Leenhouts,** *A system to generate chromosomal mutations in Lactococcus lactis which allows fast analysis of targeted genes.* J Bacteriol, **1995**. **177**(24): p. 7011-8.

64. **Mierau, I. and M. Kleerebezem,** *10 years of the nisin-controlled gene expression system (NICE) in Lactococcus lactis.* Appl Microbiol Biotechnol, **2005**. **68**(6): p. 705-17.

65. **O'Connell Motherway, M., G.F. Fitzgerald, and D. van Sinderen,** *Metabolism of a plant derived galactose-containing polysaccharide by Bifidobacterium breve UCC2003.* Microbial Biotechnology, **2011**. **4**(3): p. 403-416.

66. **Pokusaeva, K., M. O'Connell-Motherway, A. Zomer, J. Macsharry, G.F. Fitzgerald, and D. van Sinderen,** *Cellodextrin utilization by bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2011**. **77**(5): p. 1681-90.

67. **O'Connell, K.J., M. O'Connell Motherway, J. O'Callaghan, G.F. Fitzgerald, R.P. Ross, M. Ventura, C. Stanton, and D. van Sinderen,** *Metabolism of four alpha-glycosidic linkage-containing oligosaccharides by Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2013**. **79**(20): p. 6280-92.

68. **de Ruyter, P.G., O.P. Kuipers, and W.M. de Vos,** *Controlled gene expression systems for Lactococcus lactis with the food-grade inducer nisin.* Applied and Environmental Microbiology, **1996**. **62**(10): p. 3662-3667.

69. **Bradford, M.M.,** *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Analytical Biochemistry, **1976**. **72**(1–2): p. 248-254.

70. **McGrath, S., G.F. Fitzgerald, and D. van Sinderen,** *Improvement and optimization of two engineered phage resistance mechanisms in Lactococcus lactis.* Appl Environ Microbiol, **2001**. **67**(2): p. 608-16.

71. **Finn, R.D., P. Coggill, R.Y. Eberhardt, S.R. Eddy, J. Mistry, A.L. Mitchell, S.C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, G.A. Salazar, J. Tate, and A. Bateman,** *The Pfam protein families database: towards a more sustainable future.* Nucleic Acids Research, **2016**. **44**(D1): p. D279-D285.

72. **Kanehisa, M., Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe,** *KEGG as a reference resource for gene and protein annotation.* Nucleic Acids Research, **2016**. **44**(D1): p. D457-D462.

73. **Rada, V.,** *Detection of Bifidobacteriumspecies by enzymatic methods and antimicrobial susceptibility testing.* Biotechnology Techniques, **1997**. **11**(12): p. 909-912.

74. **Simpson, P.J., G.F. Fitzgerald, C. Stanton, and R.P. Ross,** *The evaluation of a mupirocin-based selective medium for the enumeration of bifidobacteria from probiotic animal feed.* J Microbiol Methods, **2004**. **57**(1): p. 9-16.

75. **Pokusaeva, K., G.F. Fitzgerald, and D. van Sinderen,** *Carbohydrate metabolism in Bifidobacteria.* Genes & Nutrition, **2011**. **6**(3): p. 285-306.

76. **Friedrich, T., M. Fatemi, H. Gowhar, O. Leismann, and A. Jeltsch,** *Specificity of DNA binding and methylation by the M.FokI DNA methyltransferase.* Biochimica et

Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, **2000**. **1480**(1): p. 145-159.

77. **Leismann, O., M. Roth, T. Friedrich, W. Wende, and A. Jeltsch,** *The Flavobacterium okeanokoites adenine-N6-specific DNA-methyltransferase M.FokI is a tandem enzyme of two independent domains with very different kinetic properties.* European Journal of Biochemistry, **1998**. **251**(3): p. 899-906.
78. **Pingoud, A., G.G. Wilson, and W. Wende,** *Type II restriction endonucleases—a historical perspective and more.* Nucleic Acids Research, **2014**. **42**(12): p. 7489-7527.
79. **Egan, M., M.O. Motherway, M. Kilcoyne, M. Kane, L. Joshi, M. Ventura, and D. van Sinderen,** *Cross-feeding by Bifidobacterium breve UCC2003 during co-cultivation with Bifidobacterium bifidum PRL2010 in a mucin-based medium.* BMC Microbiol, **2014**. **14**(282): p. 282.
80. **Ruiz-Moyano, S., S.M. Totten, D.A. Garrido, J.T. Smilowitz, J.B. German, C.B. Lebrilla, and D.A. Mills,** *Variation in Consumption of Human Milk Oligosaccharides by Infant Gut-Associated Strains of Bifidobacterium breve.* Applied and Environmental Microbiology, **2013**. **79**(19): p. 6040-6049.
81. **Alvarez-Martin, P., M. O'Connell Motherway, F. Turroni, E. Foroni, M. Ventura, and D. van Sinderen,** *A two-component regulatory system controls autoregulated serpin expression in Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2012**. **78**(19): p. 7032-41.
82. **O'Connell Motherway, M., M. Kinsella, G.F. Fitzgerald, and D. van Sinderen,** *Transcriptional and functional characterization of genetic elements involved in galacto-oligosaccharide utilization by Bifidobacterium breve UCC2003.* Microb Biotechnol, **2013**. **6**(1): p. 67-79.
83. **Yew, W.S., A.A. Fedorov, E.V. Fedorov, J.F. Rakus, R.W. Pierce, S.C. Almo, and J.A. Gerlt,** *Evolution of enzymatic activities in the enolase superfamily: L-fuconate dehydratase from Xanthomonas campestris.* Biochemistry, **2006**. **45**(49): p. 14582-97.
84. **Dahms, A.S. and R.L. Anderson,** *d-Fucose Metabolism in a Pseudomonad : I. Oxidation of d-fucose to d-fucono-δ-lactone by a d-aldohexose dehydrogenase.* Journal of Biological Chemistry, **1972**. **247**(7): p. 2222-2227.
85. **Dahms, A.S. and R.L. Anderson,** *d-Fucose Metabolism in a Pseudomonad : II. Oxidation of d-fucose to d-fucono-γ-lactone by an l-arabino-aldose dehydrogenase and hydrolysis of the lactone by a lactonase.* Journal of Biological Chemistry, **1972**. **247**(7): p. 2228-2232.

86. **Dahms, A.S. and R.L. Anderson,** *d-Fucose Metabolism in a Pseudomonad : III. conversion of d-fuconate to 2-keto-3-deoxy-d-fuconate by a dehydratase.* Journal of Biological Chemistry, **1972**. **247**(7): p. 2233-2237.

87. **Dahms, A.S. and R.L. Anderson,** *d-Fucose Metabolism in a Pseudomonad : IV. Cleavage of 2-keto-3-deoxy-d-fuconate to pyruvate and d-lactaldehyde by 2-keto-3-deoxy-l-arabonate aldolase.* Journal of Biological Chemistry, **1972**. **247**(7): p. 2238-2241.

88. **Cocks, G.T., J. Aguilar, and E.C.C. Lin,** *Evolution of l-1,2-Propanediol Catabolism in Escherichia coli by Recruitment of Enzymes for l-Fucose and l-Lactate Metabolism.* Journal of Bacteriology, **1974**. **118**(1): p. 83-88.

89. **Hacking, A.J., J. Aguilar, and E.C. Lin,** *Evolution of propanediol utilization in Escherichia coli: mutant with improved substrate-scavenging power.* Journal of Bacteriology, **1978**. **136**(2): p. 522-530.

90. **Benno, Y., K. Sawada, and T. Mitsuoka,** *The intestinal microflora of infants: composition of fecal flora in breast-fed and bottle-fed infants.* Microbiol Immunol, **1984**. **28**(9): p. 975-86.

91. **Guaraldi, F. and G. Salvatori,** *Effect of Breast and Formula Feeding on Gut Microbiota Shaping in Newborns.* Frontiers in Cellular and Infection Microbiology, **2012**. **2**: p. 94.

92. **Solís, G., C.G. Los Reyes-Gavilan, N. Fernández, A. Margolles, and M. Gueimonde,** *Establishment and development of lactic acid bacteria and bifidobacteria microbiota in breast-milk and the infant gut.* Anaerobe, **2010**. **16**(3): p. 307-10.

93. **LoCascio, R.G., P. Desai, D.A. Sela, B. Weimer, and D.A. Mills,** *Broad conservation of milk utilization genes in Bifidobacterium longum subsp. infantis as revealed by comparative genomic hybridization.* Appl Environ Microbiol, **2010**. **76**(22): p. 7373-81.

94. **Zivkovic, A.M., J.B. German, C.B. Lebrilla, and D.A. Mills,** *Human milk glycobiome and its impact on the infant gastrointestinal microbiota.* Proc Natl Acad Sci U S A, **2011**. **108 Suppl 1**: p. 4653-8.

95. **Garrido, D., S. Ruiz-Moyano, N. Kirmiz, J.C. Davis, S.M. Totten, D.G. Lemay, J.A. Ugalde, J.B. German, C.B. Lebrilla, and D.A. Mills,** *A novel gene cluster allows preferential utilization of fucosylated milk oligosaccharides in Bifidobacterium longum subsp. longum SC596.* Sci Rep, **2016**. **6**: p. 35045.

96. **McGuire, M.K., C.L. Meehan, M.A. McGuire, and J.E. Williams,** *What's normal? Oligosaccharide concentrations and profiles in milk produced by healthy women vary geographically.* **2017**. **105**(5): p. 1086-1100.

97. **Wada, J., R. Suzuki, S. Fushinobu, M. Kitaoka, T. Wakagi, H. Shoun, H. Ashida, H. Kumagai, T. Katayama, and K. Yamamoto,** *Purification, crystallization and preliminary X-ray analysis of the galacto-N-biose-/lacto-N-biose I-binding protein (GL-BP) of the ABC transporter from Bifidobacterium longum JCM1217.* Acta Crystallogr Sect F Struct Biol Cryst Commun, **2007**. **63**(Pt 9): p. 751-3.

98. **Álvarez-Martín, P., M. O'Connell-Motherway, D. van Sinderen, and B. Mayo,** *Functional analysis of the pBC1 replicon from Bifidobacterium catenulatum L48.* Applied Microbiology and Biotechnology, **2007**. **76**(6): p. 1395-1402.

Chapter VI

General Discussion and Future Perspectives

The bifidogenic effect of HMOs is now well accepted to be of significant importance to the establishment and persistence of bifidobacteria as part of the (breastfed) infant gut microbiota. Several microbiota studies have demonstrated the ability of (the consumption of) breastmilk to significantly increase relative abundance levels of bifidobacteria in the infant gut [1-4]. These observations are in line with the fact that various infant-associated species of *Bifidobacterium* are known to metabolise individual HMO components [5-8]. A significant understanding, including the precise metabolic pathways and associated enzymatic activities, of the mechanisms of HMO utilisation in *B. bifidum*, *B. longum* subsp. *infantis* has been acquired [6, 9-11]. In addition, as part of the work described in this thesis, knowledge on the mechanisms of HMO utilisation by *B. breve* [12] and *B. kashiwanohense* has vastly improved.

As our understanding of these HMO utilisation preferences and pathways in infant-associated bifidobacteria increases, it has provided a much more in depth appreciation of the bifidogenic effect of HMOs. It furthermore, as has been suggested previously, highlights the existence of remarkable co-evolution events between humans and certain bifidobacterial species as part of a breastfeed-driven symbiosis [13, 14]. Since lactation is a general feature of mammalian biology, it is likely that this co-evolution also exists in other mammals. The presence of various indigestible milk oligosaccharides has been observed in milk and colostrum samples of a range of other mammals [15], and bifidobacteria have similarly been isolated from the faeces of a wide array of mammals [16]. Analysis of the distribution of particular LN(n)T/LNB utilisation genes across the *Bifidobacterium* genus, as discussed in Chapter II [12], provides corroborating evidence for the existence of this co-evolutionary phenomenon in other mammals, particularly among primates. The observation that homologs of the *nahA* gene, encoding N-acetylhexosaminidase responsible for the hydrolysis of the LN(n)T component lacto-N-triose into lactose and GlcNAc, are present in *Bifidobacterium* species isolated from marmosets and tamarins, indeed suggests similar co-evolutionary symbioses in other lactating mammals. This demonstrates the importance of maternal milk as a tool for the development

of a specific gut microbiota in the mammalian neonate, with bifidobacteria as a dominant component, and in turn, underscores the importance of bifidobacteria for a healthy development of these neonates.

The health benefits that bifidobacteria have been claimed to endow their human (and, by extension, mammalian) infant host are wide-ranging and extensive. As discussed in the introduction, such benefits range from the generation of essential nutrients from dietary sources, protection from pathogens, to immunomodulatory effects, and these so-called probiotic properties are the basis of wide-spread commercial exploitation of Bifidobacteria for infant health [17]. Likewise, the numerous benefits of HMOs to the infant have also been well-documented [18, 19], in particular their role as a bifidogenic prebiotic. It is expected that the use of HMOs as beneficial supplements and in formaula for infants will become more commercially viable, thereby creating opportunities for the generation of new prebiotics to modulate the infant gut microbiota, and thus to improve infant health. The use of prebiotics such as GOS and inulin in infant formula milk as bifidogenic agents is well-documented [20]. However, the use of HMOs is believed to provide an opportunity for the selection of a more specific, breastfed-type microbiota that is not observed in infants that are formula-fed [21, 22]. Furthermore, it may also create opportunities for the use of synbiotics, where (particular) HMOs are combined with infant-associated bifidobacterial cultures, to help modulate the infant microbiota. Several studies have documented how the infant gut microbiota composition is influenced by factors such as C-section delivery, pre-term birth or antibiotic usage [23-25]. The use of synbiotic formulations may be a powerful approach to restore or create a *Bifidobacterium*-rich (specifically of infant-associated species) gut microbiota characteristic of healthy, breastfed neonates [1, 25].

The two major caveats to these potential applications are the possible risks of using individual HMOs (as opposed to the highly complex mix of HMOs in breast milk, the composition of which changes during lactation), and the current inability to commercially produce sufficient quantities of HMOs. With regards to the former point, HMO research is a relatively new field, particularly studying its effects when used as a therapeutic agent in humans.

As such, the full range of short-term and long-term effects in individuals are unknown at this time. HMOs are naturally present in breastmilk as a complex, heterogeneous mixture [26], and therefore the inclusion of high quantities of one or a few HMOs in the infant diet may unnaturally skew the microbiota, and select for an atypical, or potentially even harmful composition, through the portential disproportionate proliferation of opportunistic pathogens or accumulation of deleterious metabolites. Indeed, it has been observed that differing HMO compositions in milk have a direct effect on the gut microbiota composition in infants [2], demonstrating the power of the prebiotic effect of (particular) HMOs. Regarding the scarce availability of synthetically/enzymatically-produced HMOs, the major limiting factor is the inability, at present, to generate large-scale batch quantities of high-purity HMOs. However, advances in techniques for such commercially viable HMO syntheses, particularly by biotechnological methods [27, 28], show promise for addressing this limitation in the future.

Before such advances are realised into commercial and/or therapeutic applications, it is important to understand the mechanisms by which these bifidogenic effects occur. As stated, the pathways of HMO utilisation in the dominant infant-associated *Bifidobacterium* species; *B. bifidum*, *B. longum* subsp. *infantis* and *B. breve*, as well as the less-prevalent species *B. kashiwanohense*, have now been well-characterised. This research has demonstrated the striking specificity of each species for a very particular set of HMO components, whether that is the rather broad range HMO structures metabolized by *B. bifidum* through its extracellular enzyme approach [6, 9]; the specific small-mass HMOs internalised and intracellularly digested by *B. longum* subsp. *infantis* [11], the neutral, simple HMO glycans metabolised by *B. breve* (Chapter II, [12]), or strictly small-mass fucosylated HMOs consumed by *B. kashiwanohense* (Chapter V, [8]). What this reveals is a form of resource partitioning among infant-associated *Bifidobacterium* species for specific HMO components, which may, in fact, allow the co-existence of multiple species within the intestinal niche of a single infant host. Such resource partitioning in microbial communities is by no means a new concept, and has been observed in other environments [29, 30]. The case observed here

with the selective consumption of HMOs by bifidobacteria in the breastfed infant gut, however, is an interesting case of this concept applied to species within a single genus. This phenomenon, with respect to the community of HMO-utilising bifidobacteria within the breastfed infant gut, is even further increased in the depth of its complexity, with the occurrence of HMO cross-feeding, as has observed between *B. bifidum* and *B. breve* [31].

Further study of HMO metabolism in less common, infant-derived *Bifidobacterium* species, such as *B. longum* subsp. *longum* and *B. pseudocatenulatum*, as well as in other infant gut-associated bacteria outside of the *Bifidobacterium* genus, is necessary in order to gain an even more complete picture of the relationship between HMOs and bifidobacteria. Additionally, while the work in Chapter III of this thesis goes some way to describe the mechanisms of transcriptional regulation of HMO metabolism in *B. breve* UCC2003, there is still little knowledge of this vital process in other infant-associated bifidobacteria, and such research is necessary to fully understand the control mechanisms that regulate these metabolic processes at a more in depth level.

Altogether, the presented work of this thesis is important for our understanding of HMO utilisation in infant-associated bifidobacteria, in particular in the species *B. breve* and *B. kashiwanohense*. However, it only represents a tiny segment of the overall picture in the context of the relationship between HMOs and *Bifidobacterium* species, and moreover the overall microbiota, in the infant gut. As our knowledge of the patterns and mechanisms of utilisation by these species expands, we not only reveal more of the fascinating natural bifidogenic, prebiotic nature of HMOs, but we also step closer to the use of HMOs alone, or in combination with bifidobacteria, to create food products and therapeutics which can effect significant beneficial changes in the infant gut microbiota.

6.1 References

1. **Turroni, F., C. Peano, D.A. Pass, E. Foroni, M. Severgnini, M.J. Claesson, C. Kerr, J. Hourihane, D. Murray, F. Fuligni, M. Gueimonde, A. Margolles, G. De Bellis, P.W. O'Toole, D. van Sinderen, J.R. Marchesi, and M. Ventura,** *Diversity of Bifidobacteria within the Infant Gut Microbiota.* PLoS ONE, **2012**. **7**(5): p. e36957.
2. **Wang, M., M. Li, S. Wu, C.B. Lebrilla, R.S. Chapkin, I. Ivanov, and S.M. Donovan,** *Fecal Microbiota Composition of Breast-fed Infants is Correlated with Human Milk Oligosaccharides Consumed.* Journal of pediatric gastroenterology and nutrition, **2015**. **60**(6): p. 825-833.
3. **Bezirtzoglou, E., A. Tsiotsias, and G.W. Welling,** *Microbiota profile in feces of breast- and formula-fed newborns by using fluorescence in situ hybridization (FISH).* Anaerobe, **2011**. **17**(6): p. 478-82.
4. **Charbonneau, Mark R., D. O'Donnell, Laura V. Blanton, Sarah M. Totten, Jasmine C.C. Davis, Michael J. Barratt, J. Cheng, J. Guruge, M. Talcott, James R. Bain, Michael J. Muehlbauer, O. Ilkayeva, C. Wu, T. Struckmeyer, D. Barile, C. Mangani, J. Jorgensen, Y.-m. Fan, K. Maleta, Kathryn G. Dewey, P. Ashorn, Christopher B. Newgard, C. Lebrilla, David A. Mills, and Jeffrey I. Gordon,** *Sialylated Milk Oligosaccharides Promote Microbiota-Dependent Growth in Models of Infant Undernutrition.* Cell. **164**(5): p. 859-871.
5. **Ward, R.E., M. Ninonuevo, D.A. Mills, C.B. Lebrilla, and J.B. German,** *In vitro fermentability of human milk oligosaccharides by several strains of bifidobacteria.* Mol Nutr Food Res, **2007**. **51**(11): p. 1398-405.
6. **Asakuma, S., E. Hatakeyama, T. Urashima, E. Yoshida, T. Katayama, K. Yamamoto, H. Kumagai, H. Ashida, J. Hirose, and M. Kitaoka,** *Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria.* J Biol Chem, **2011**. **286**(40): p. 34583-92.
7. **LoCascio, R.G., M.R. Ninonuevo, S.L. Freeman, D.A. Sela, R. Grimm, C.B. Lebrilla, D.A. Mills, and J.B. German,** *Glycoprofiling of Bifidobacterial Consumption of Human Milk Oligosaccharides Demonstrates Strain Specific, Preferential Consumption of Small*

Chain Glycans Secreted in Early Human Lactation. Journal of Agricultural and Food Chemistry, **2007**. **55**(22): p. 8914-8919.

8. **Bunesova, V., C. Lacroix, and C. Schwab,** *Fucosyllactose and L-fucose utilization of infant Bifidobacterium longum and Bifidobacterium kashiwanohense.* BMC Microbiology, **2016**. **16**(1): p. 248.
9. **Sela, D.A.,** *Bifidobacterial utilization of human milk oligosaccharides.* Int J Food Microbiol, **2011**. **149**(1): p. 58-64.
10. **Kitaoka, M.,** *Bifidobacterial enzymes involved in the metabolism of human milk oligosaccharides.* Adv Nutr, **2012**. **3**(3): p. 422s-9s.
11. **Underwood, M.A., J.B. German, C.B. Lebrilla, and D.A. Mills,** *Bifidobacterium longum subspecies infantis: champion colonizer of the infant gut.* Pediatr Res, **2015**. **77**(1-2): p. 229-35.
12. **James, K., M.O. Motherway, F. Bottacini, and D. van Sinderen,** *Bifidobacterium breve UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways.* Sci Rep, **2016**. **6**: p. 38560.
13. **Katayama, T.,** *Host-derived glycans serve as selected nutrients for the gut microbe: human milk oligosaccharides and bifidobacteria.* Bioscience, Biotechnology, and Biochemistry, **2016**. **80**(4): p. 621-632.
14. **Sela, D.A. and D.A. Mills,** *The marriage of nutrigenomics with the microbiome: the case of infant-associated bifidobacteria and milk.* Am J Clin Nutr, **2014**. **99**(3): p. 697s-703s.
15. **Urashima, T., T. Saito, T. Nakamura, and M. Messer,** *Oligosaccharides of milk and colostrum in non-human mammals.* Glycoconjugate Journal, **2001**. **18**(5): p. 357-371.
16. **Milani, C., M. Mangifesta, L. Mancabelli, G.A. Lugli, K. James, S. Duranti, F. Turroni, C. Ferrario, M.C. Ossiprandi, D. van Sinderen, and M. Ventura,** *Unveiling bifidobacterial biogeography across the mammalian branch of the tree of life.* Isme j, **2017**. **11**(12): p. 2834-2847.
17. **Di Gioia, D., I. Aloisio, G. Mazzola, and B. Biavati,** *Bifidobacteria: their impact on gut microbiota composition and their applications as probiotics in infants.* Appl Microbiol Biotechnol, **2014**. **98**(2): p. 563-77.
18. **Bode, L.,** *Human milk oligosaccharides: Every baby needs a sugar mama.* Glycobiology, **2012**. **22**(9): p. 1147-1162.

19. **Bode, L.,** *The functional biology of human milk oligosaccharides.* Early Hum Dev, **2015**. **91**(11): p. 619-22.
20. **Roberfroid, M.,** *Prebiotics: The Concept Revisited.* The Journal of Nutrition, **2007**. **137**(3): p. 830S-837S.
21. **Guaraldi, F. and G. Salvatori,** *Effect of Breast and Formula Feeding on Gut Microbiota Shaping in Newborns.* Frontiers in Cellular and Infection Microbiology, **2012**. **2**: p. 94.
22. **O'Sullivan, A., M. Farver, and J.T. Smilowitz,** *The Influence of Early Infant-Feeding Practices on the Intestinal Microbiome and Body Composition in Infants.* Nutrition and Metabolic Insights, **2015**. **8**(Suppl 1): p. 1-9.
23. **Arboleya, S., B. Sanchez, C. Milani, S. Duranti, G. Solis, N. Fernandez, C.G. de los Reyes-Gavilan, M. Ventura, A. Margolles, and M. Gueimonde,** *Intestinal microbiota development in preterm neonates and effect of perinatal antibiotics.* J Pediatr, **2015**. **166**(3): p. 538-44.
24. **Bokulich, N.A., J. Chung, T. Battaglia, N. Henderson, M. Jay, H. Li, D.L. A, F. Wu, G.I. Perez-Perez, Y. Chen, W. Schweizer, X. Zheng, M. Contreras, M.G. Dominguez-Bello, and M.J. Blaser,** *Antibiotics, birth mode, and diet shape microbiome maturation during early life.* Sci Transl Med, **2016**. **8**(343): p. 343ra82.
25. **Martin, R., H. Makino, A. Cetinyurek Yavuz, K. Ben-Amor, M. Roelofs, E. Ishikawa, H. Kubota, S. Swinkels, T. Sakai, K. Oishi, A. Kushiro, and J. Knol,** *Early-Life Events, Including Mode of Delivery and Type of Feeding, Siblings and Gender, Shape the Developing Gut Microbiota.* PLoS One, **2016**. **11**(6): p. e0158498.
26. **McGuire, M.K., C.L. Meehan, M.A. McGuire, and J.E. Williams,** *What's normal? Oligosaccharide concentrations and profiles in milk produced by healthy women vary geographically.* **2017**. **105**(5): p. 1086-1100.
27. **Han, N.S., T.J. Kim, Y.C. Park, J. Kim, and J.H. Seo,** *Biotechnological production of human milk oligosaccharides.* Biotechnol Adv, **2012**. **30**(6): p. 1268-78.
28. **Prudden, A.R., L. Liu, C.J. Capicciotti, M.A. Wolfert, S. Wang, Z. Gao, L. Meng, K.W. Moremen, and G.J. Boons,** *Synthesis of asymmetrical multiantennary human milk oligosaccharides.* Proc Natl Acad Sci U S A, **2017**. **114**(27): p. 6954-6959.

29. **Bryson, S., Z. Li, F. Chavez, P.K. Weber, J. Pett-Ridge, R.L. Hettich, C. Pan, X. Mayali, and R.S. Mueller,** *Phylogenetically conserved resource partitioning in the coastal microbial loop.* The Isme Journal, **2017**. **11**: p. 2781.
30. **Kramer, S., D. Dibbern, J. Moll, M. Huenninghaus, R. Koller, D. Krueger, S. Marhan, T. Urich, T. Wubet, M. Bonkowski, F. Buscot, T. Lueders, and E. Kandeler,** *Resource Partitioning between Bacteria, Fungi, and Protists in the Detritusphere of an Agricultural Soil.* Frontiers in Microbiology, **2016**. **7**: p. 1524
31. **Egan, M., M. O'Connell Motherway, M. Ventura, and D. van Sinderen,** *Metabolism of sialic acid by Bifidobacterium breve UCC2003.* Appl Environ Microbiol, **2014**. **80**(14): p. 4414-26.

ACKNOWLEDGEMENTS

It is my pleasure to acknowledge the many people who have contributed to the completion of this thesis.

I would, first and foremost, like to express my sincere gratitude to my supervisor, Professor Douwe van Sinderen. Thank you for your wise guidance, encouragement, patience and enthusiasm over the course of these four years. You have been an excellent mentor, and always ensured that I got the best out of myself in every aspect of the PhD.

Thank you to Dr. Francesca Bottacini for the enormous time and help you provided me, with just about anything and everything relating to bioinformatics. Thank you also to Dr. Mary O'Connell Motherway for all your help and advice throughout the PhD with many different experiments, and to Dr. Christophe Penno for the help with a great number of primer extensions.

I thank all the staff of the School of Microbiology, in particular Paddy O'Reilly for providing any obscure strain or chemical I could possibly think of; and also Dan Walsh, Colette Crowley, Colette O'Gorman, Jonathon McLintock, Mary Hough, Carmel Shortiss and Aine Murphy.

I wish to acknowledge the Irish Research Council for the financial contribution that made this PhD possible, through the School of Microbiology and the APC Microbiome Institute.

I would also like to thank Glycom A/S for the generous donation of (many, many grams) of high-purity HMO samples, which made the experimental work of this thesis possible.

I thank all those working on the InfantMet project as well, in particular Prof. Paul O'Toole and Dr. Cian Hill, for the provision of infant faecal samples for the screening work in this thesis.

To Dr. Paul Simpson, I would like to sincerely thank you for starting my interest in microbiology, and for the enthusiasm for research you passed on to me.

To my friends and colleagues in Lab 4.25 and the Fifth Floor over the course of the last 4 years, thank you for being there through all the trials and tribulations of research, and for all the fun we've shared working together. Thanks also to my friends and colleagues in other labs, including everyone in Lab 340 and 405. To my long-time bench cohabiter and friend Ruth, thank you for the hours of laughs and sage advice when I needed it most. To my friends also finishing their PhDs at this time; Alli, Emer and Noreen; thank you for sharing the highs and lows of the PhD with me, and still having a lot of fun in the process; I really couldn't have done it without you. To my favourite summer student (and now PhD student and good friend) Valentina, thank you for always finding a way to make me laugh, even when you don't mean it. To my other favourite Italian friend, Francesca, thanks for all the advice, fun and food, and thanks to Sandra for the reminder that the 'sesh' can solve all woes. Thanks to Ana for the daily conversations entirely in made-up accents, and to Muireann, for always giving a valued opinion of the freshest kicks. Thanks also to Maria, Rocio, Justin, Valerio, Rebecca and Christophe.

To my good friends outside of work, thank you for your constant encouragement and reminder that there is a world outside the PhD; in particular Dan, Caoilfhionn, Josh, Loughlin and Ciaran. To my girlfriend, Katie, thanks for your endless support, and patiently putting up with my stressed self in the last few months of the PhD.

Finally, to my mother Claire, brother Daniel and sister Sinead, I know I wouldn't have reached this point without your love and support throughout the entirety of my education. You have always been there for me, and I cannot thank you enough. To my father Robin, I know seeing me finish this PhD would have made you proud; I dedicate this thesis to you.