Mapping glycan binding profiles of the gut microbes using novel liquid glycan array

by

Pravina Yadav

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Animal Science

Department of Agricultural, Food and Nutritional Science
University of Alberta

Abstract

The role of glycan binding of gut microbes has been shown in colonization and interactions between microbiota and host. While glycan binding of selected pathogens has been explored in a targeted way, there has been limited research on glycan binding patterns of gut commensals. This has partially been limited by tools available to assess glycan binding. In this thesis a novel Liquid glycan array (LiGA), which comprises a library of glycosylated M13 phage particles with silent DNA barcodes in the phage genome, was used to test glycan binding profiles of gut bacteria. The potential role of glycan binding in host specificity of *Limosilactobacillus reuteri* was explored by comparing 16 *L. reuteri* strains consisting of 4 isolates from each of murine, porcine, poultry and human lineages. Many of the enriched glycans were shared amongst *L. reuteri* isolates, however, there was no evidence that isolates from the same host shared greater glycan binding similarity, thus glycan binding profile did not appear to be a key determinant of host specificity. However, some strains showed notable unique glycan enrichment. In particular a poultry isolate *L. reuteri* JCM1081 showed the drastically higher enrichment of Fucα1-2Galβ-Sp (Fold change (FC) = 88, FDR < 0.0001) and Galβ1-4Glcβ-P4 (FC = 46, FDR < 0.0001). The unique binding profile of this isolate conforms with observations of enhanced adhesive capacity to gut epithelial cells.

Next, to understand the profiles of glycan binding of other bacterial species, we tested glycan binding of taxonomically diverse bacteria from three different phyla Firmicutes/Bacillota, Bacteroidetes/Bacteroidota, Proteobacteria/Pseudomonadota consisting of 9 different species *L. reuteri, Escherichia coli, Bacteroides dorei, Bacteroides thetaiotamicron, Bacteroides fragilis,*

Bacteroides vulgatus, Limosilactobacillus mucosae, Citrobacter fruendii, and *Clostridium ramosum*. Results indicate that the taxonomic closeness leads to similar glycan binding of gut bacteria with few exceptions.

We established that LiGA is an effective new tool for characterizing glycan binding of bacteria. Exploring glycan binding profiles of more commensal bacteria using this novel approach will provide new insights into intestinal microbial ecology and provide strategies to manipulate the microbial community through the provision of glycans.

Preface

This thesis is an original work by Pravina Yadav. This work was funded by Alberta Innovates grant held by Dr. Ratmir Derda.

Author contribution: conception and design of experiments: Benjamin Willing, Ratmir Derda, and Pravina Yadav. Performance of the experiments: Pravina Yadav. Analysis of data: Pravina Yadav, Mirat Sojitra. Contribution of reagents/materials/analysis tools: Pravina Yadav, Benjamin Willing and Ratmir Derda.

Acknowledgements

Pursuing this degree during pandemic came with its own unique experiences and challenges. There are so many people I want to thank for making this a wonderful journey.

First, I want to express my deepest gratitude to my supervisor, Dr. Benjamin Willing for giving me this research opportunity. I am grateful for his great insights, guidance, and support throughout this journey.

I would like to thank Dr. Ratmir Derda for serving on my supervisory committee, giving me invaluable advice and support throughout the project. I would like thank Mirat Sojitra for helping me with data analysis.

I would also like to thank Dr. Jens Walter for provision of *L. reuteri* isolates and Dr. Eytan Wine for the provision of human gut isolates.

I would really like to thank my lab mates without whom it would have been impossible. I want to thank Dr. Andrew Forgie, Dr. Tingting Ju, Rajibur Rahman, Camila Marcolla, Dr. Benjamin Bourrie, Edward (Yi) Fan, Tausha Prisnee, Maanasa Mudoor Sooresh, Natalie Diether, and Hannah Lantz for their support throughout the project.

I also want to thank all my friends especially Bhumika, Yuvraj, Arathi, and Kaushik for supporting me always. Last but not the least, I want to thank my parents, my brother and my family for their encouragement and faith in my abilities.

Table of contents

List of tables

List of figures

List of abbreviations

LiGA: Liquid glycan array

CFU: Colony forming unit

PFU: Plaque forming assay

PCR: Polymerase chain reaqction

DE: differential enrichment

ETEC: Enterotoxigenic *E. coli*

AIEC: Adherent invasive *E. coli*

GBP: Glycan binding protein

SGBP: Surface glycan binding protein

FDR: False discovery rate

FC: Fold change

S.D.: Standard deviation

MUB: mucus binding protein

Ef-tu: Elongation factor-tu

MapA: Mucus adhesion promoting protein A

1 Chapter 1: Literature review

1.1 Introduction

Glycans in the gastrointestinal (GI) environment facilitate the interaction of bacteria with the host, and microbes can utilize these glycans for adherence and colonization (Ofek et al., 2005). The GI tract presents abundant glycans in the forms of the epithelial glycocalyx, dietary glycans, and mucus layer. The epithelial glycocalyx is the further most layer from the lumen and is made of glycoproteins similar to the mucus layer glycoproteins.

Besides glycans, bacterial surface proteins also play a crucial role in interactions with the host, some of these proteins are characterized as glycan-binding proteins (GBPs). GBPs are mostly identified as lectins and adhesins.

A complete spectrum of glycans involved in host-bacterial interaction is known for a few bacteria and *Helicobacter pylori* is one of the best described. The role of adhesins of *H. pylori* has been studied in the context of its pathology (Ilver et al., 1998a). Recently, glycan array analysis of *Campylobacter jejuni* determined the full spectrum of the glycans involved in host bacterial interaction (Christopher J. Day et al., 2013; Christopher James Day et al., 2012). In the context of commensals, the glycan-binding preferences can be used in deciphering the mechanism of colonization, evaluating the probiotic potential of beneficial microbes, and understanding of carbohydrate-binding modules in glycan recognition. GBPs in *Bifidobacteria* (Garrido et al., 2011; Servin & Coconnier, 2003), *Lactobacilli* (Takao Mukai et al., 2004; Petrova et al., 2016)*, and*

Ruminococcus gnavus (Owen et al., 2017) have been studied, but GBPs for most of the gut commensal remain unknown.

1.1.1 Glycans /sugars/ oligosaccharides in the gut

Colonic mucus layer is primarily made of mucin-2-glycoprotein (MUC2), that contains large amount of hydroxy amino acids, serine (Ser), and threonine (Thr), that act as attachment sites for hydroxy linked carbohydrate chains also known as *O*-linked glycans (Schroeder, 2019) (Johansson et al., 2011). These glycans mainly consist of N-acetyl galactosamine (GalNAc), galactose (Gal), N-acetylglucosamine (GlcNAc), fucose (Fuc) and N-acetylneuraminic acid (Neu5Ac) residues (Holmén Larsson et al., 2013). Additionally, dietary glycans which come in forms of long polysaccharide chains (eg. Cellulose, pectins, resistant starch), oligosaccharides, and disaccharides (e.g. glucose, lactose) play important role in shaping gut microbiota (Coker et al., 2021; Schroeder, 2019).

1.1.2 Bacterial glycan-binding proteins (GBPs)

In gut microbes, glycan-binding proteins (GBPs) exist as lectins, adhesins, and carbohydrate-binding modules (CBMs) of carbohydrate-active enzymes (CAZymes). Bacterial lectins occur commonly in the form of elongated, multi-sub-unit protein appendages, known as fimbriae (Varki et al., 2009). Some bacteria have multiple adhesins with different carbohydrate-binding specificity, which help in defining a microbe's ecological niche (Esko & Sharon, 2009).

However, CAZymes are generally studied regarding dietary carbohydrate utilization as they break the glycosidic bonds joining sugar residues. By definition, CAZymes contain a catalytic domain that can break glycosidic bonds, and the most common super-family of CAZymes is the glycoside hydrolases (GHs) family (Cantarel et al., 2009). GHs often contain numerous accessory modules, most common of which are carbohydrate-binding modules (CBMs) that non-catalytically mediate enzyme-carbohydrate interaction (Boraston et al., 2004). CBMs help in concentrating enzymes on carbohydrate substrates and thus enhance the catalytic activity but this mechanism is largely based on the non-surface attached enzymes (A. K. Singh et al., 2014). Besides, there have been recent highlights on the surface attached CAZymes with CBMs functioning as adhesion factors. CBMs in surface anchored CAZymes mediate interactions between the bacteria and host. *Ruminococcus gnavus* and *Streptococcus pneumoniae* (pneumococcus) express surface anchored sialidase (RgNanH) and b-galactosidase (BgaA) respectively (Owen et al., 2017; A. K. Singh et al., 2014). These surface proteins are hypothesized to be novel adhesins and further studies will be warranted to understand the role of CBMs as adhesins. Few bacteria and their GBPs are listed in Table 1.1.

1.1.3 Diversity of glycan structures expressed on the host GI tract

Glycan mediated host-bacterial interaction calls for the understanding of the diversity of glycan expression across the GI tract. Glycosylation in the GI tract is mainly present in the form of the mucus glycosylation and the epithelial glycosylation (glycocalyx) (Figure 1.1). Literature suggests that sources of selective pressure to create glycan diversity include the need to evade pathogenic

microbes and the need to sustain relationships with symbionts (Hooper & Gordon, 2001). For example, a bacterial species could initiate pathogenic interactions with their hosts by first binding to glycan structures, and as an evasive tactic the host could eliminate expression of the structures, either by inactivation of glycosyltransferases that mediate its production or by activation of a glycosyltransferase that utilizes the structure as an acceptor (Hooper et al., 1999). Gut symbiont, *Bacteroides thetaiotamicron* presents a model system for gut commensals regulating the expression of fucosylated glycans on host epithelium using the fucose sensing system FucR (Hooper et al., 1999). *Bacteroides thetaiotamicron* has also been shown to contribute to mucus fucosylation (Kandori et al., 1996).

The mucus layer mainly consists of two layers, a loosely adherent layer and a layer firmly attached to the mucosa (Figure 1.1). The mucus layer is made of mucin, glycoproteins which are heavily glycosylated and consist of approximately 50-80% carbohydrates (Karlsson *et al.*, 1997). The protective properties of mucus are significantly impacted by the high carbohydrate content of mucins, as mucus glycans present the binding sites for bacteria and function as the barrier to limit their invasion to the epithelium (Moran et al., 2011). Mucus layer glycans are also modulated with the changes in microbiota similar to the glycocalyx remodelling. Intestinal Muc2 mucin glycosylation is affected by microbiota as shown by comparing Muc2 glycosylation in germ free (GF) and conventionally raised (ConvR) mice (Arike et al., 2017). Also, *Salmonella* has been seen to degrade host glycocalyx and induce the expression of host genes for glycan remodelling (Arabyan et al., 2016). Spatial and temporal changes in glycosylation are speculated to be affected

by gut microbiota, however, further studies are needed to shed light on glycan mediated host-bacterial interactions and their impact on host glycan diversity.

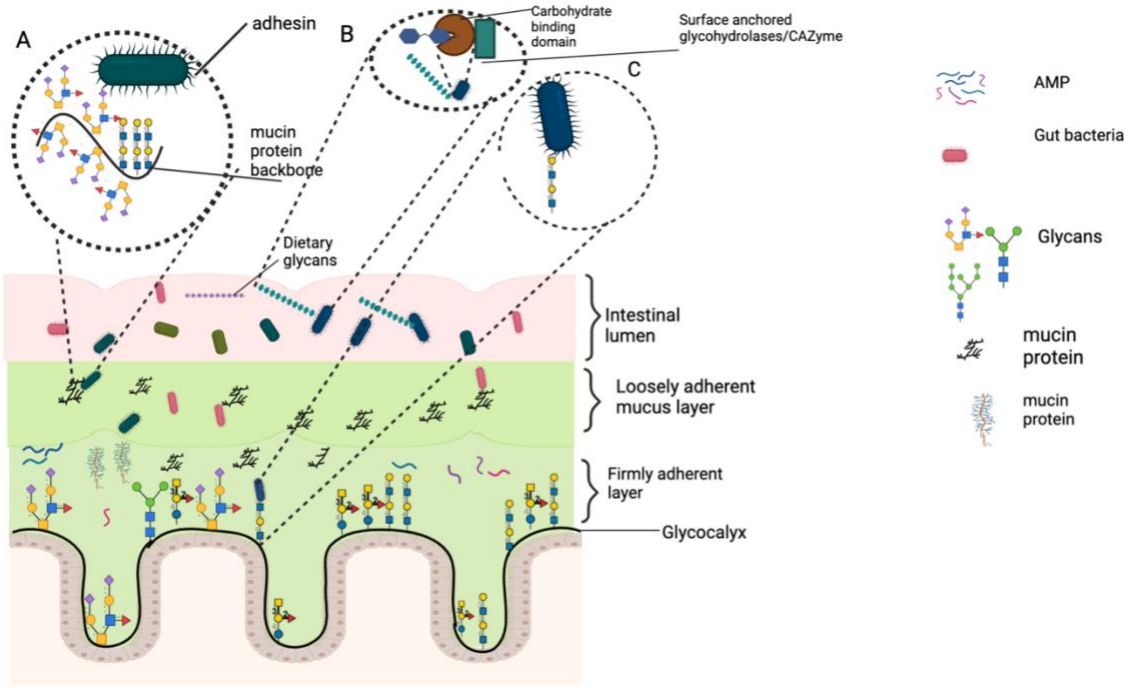

Figure 1.1 Bacterial interactions with glycans in the intestinal environment. AMP: antimicrobial peptides. CAZyme: Carbohydrate active enzyme. A) Bacterial adhesin binding to mucus glycans. B) Recognition of dietary glycans by carbohydrate binding module of the CAzyme/ glycosylhyrolases C) Pathogen binding to epithelial glycocalyx after breaching the mucosal barrier. This figure was created in Biorender platform.

1.1.4 Glycan binding proteins in host-bacterial interactions

Glycan binding proteins play a crucial role in governing host-microbial interactions (Hooper & Gordon, 2001). GBPs of bacteria that bind the glycoconjugates of gut epithelium are scarcely studied, but there has been advancement in the understanding of proteins that play roles in binding to host and host glycoconjugates. These proteins have been studied considering their role in colonization of commensals, infections by pathogens, and protective roles of probiotics and prebiotics.

1.1.4.1 Microbiota colonization and glycan-binding proteins

Bacterial adhesion to the host is one mechanism of colonization (Ofek & Doyle, 1994). Adhesion to mucus and epithelial surfaces is mediated by the bacterial surface proteins, lectins, adhesins, and attachment pili (Taylor et al., 1987). The mechanism of bacterial adhesion to host tissue has been well studied, mostly in pathogens. Pathogenic bacteria like *E. coli* possess numerous lectins with diverse sugar specificities allowing them to bind mucin as well as other glycoproteins and extracellular matrix components of epithelial cells (Mouricout et al., 1995). Bacterial surface proteins, pili, and cell wall components such as LPS (lipopolysaccharides) may act as adhesins (Moran et al., 2011; Rogemond & Guinet, 1986). Adhesion mechanisms are thought to be controlled by hydrophobic interactions, cation-bridging whereby divalent cations counteract the repulsion of negatively charged surfaces of bacteria, and host receptor-ligand binding (Hooper & Gordon, 2001; Liévin-Le Moal & Servin, 2006). Among these mechanisms, the binding of bacterial lectins to the corresponding glycosylated receptors associated with host cells is one of the extensively studied mechanisms (Moran et al., 2011).

1.1.4.1.1 Colonization of *Lactobacilli* and *Bifidobacteria* (gram-positive bacteria/commensals)

*Bifidobacteria a*nd *Lactobacilli* adhere to particular regions of the epithelium and mucus layer with the surface proteins. Adhesion mechanisms of these have been in focus to evaluate their probiotic potential and efficiency as the adherence to mucosal surfaces is an essential step for probiotics to colonize and persist in GIT (Servin & Coconnier, 2003).

Lactobacilli that colonize the small intestine and stomach present a model system for studying adhesion by commensals, with exopolysaccharides, pili, and cell-wall anchored proteins (Sengupta et al., 2013). Cell surface anchored mucus binding proteins (MUBs) of Lactobacilli play a crucial role in adherence (MacKenzie et al., 2010). Although the carbohydrate-binding ability of lactobacilli is important to elucidate mechanisms of adhesion to the epithelial cells of digestive tracts, little is understood about glycan-binding proteins in *Limosictobacillus species.* Three *Limnosilactobacillis reuteri* strains have been shown to bind to different sugar residues of glycoconjugates using Haemagglutination (HA), HA inhibition assays and TLC overlay assay (T. Mukai et al., 1998). This binding was further shown to be mediated by mucus-binding proteins (MUBs) of *L. reuteri* (Roos & Jonsson, 2002). Further, the role of MUBs in strain-specific adherence and recognition of mucus elements has been explored (MacKenzie et al., 2010). Strain specific diversity in adhesion and biofilm formation factors likely confer the host specificity of *L. reuteri* (Frese et al., 2013). The nature of interactions between MUBs and host have been suggested to be the lectin-carbohydrate interactions (Ledder et al., 2008). Although a direct interaction of MUB with specific glycans remains to be demonstrated, but competitive adhesion

assays showed that such an interaction can be significantly reduced by the addition of specific glycans (Bumbaca et al., 2006; Pretzer et al., 2005; Roos & Jonsson, 2002), suggesting a lectin-type mediated interaction. Further studies are required to investigate the direct interactions of MUBs with specific glycans.

Like *lactobacilli, bifidobacteria* also show binding to glycolipids. *Bifidobacterium bifidum*'s cell-surface proteinaceous components bind to the carbohydrate moieties of intestinal glycolipid and its binding ability to intestinal glycolipids play a crucial role for colonization of the intestinal mucosal surface (Takao Mukai et al., 2004). Recently, the novel major pilin subunit protein FimM of *Bifidobacterium longum* BBMN68 has been tested for its role in bacterial adhesion to intestinal epithelial cells (Xiong et al., 2020), although its carbohydrate binding ability has not been tested. In *Bifidobacterium infantis*, solute binding proteins (SBPs) have been characterized for their role in binding to intestinal glycoconjugates (Garrido et al., 2011).

1.1.4.1.2 Colonization of *H. pylori*

H. pylori is gram-negative bacterium found in more than half of the human population (Hooper & Gordon, 2001). Although it is prevalent in humans, it produces significant pathology in only a small subset of people. Because of its ability induce pathology in a minority of the colonized individuals, speculation has been that *H. pylori* exists as commensal and only emerges as pathogen in response to host, microbial and environmental factors (Hooper & Gordon, 2001). *H. pylori* presents the model for how glycans produced in gastric epithelial cells may affect the fate of

colonization (Hooper & Gordon, 2001). The role of adhesins in colonization and infection strategy of *H. pylori* has been extensively illustrated using in-vivo studies, which indicate that position of *H. pyroli* in the commensal-pathogen continuum may be determined in part by the repertoire of the glycans expressed in the gastric ecosystem of its host and by the microbe's ability to express the corresponding adhesins (Hooper and Gordon, 2001). Adhesins help this organism to attach to the gastric epithelium that leads to the processes for evading host immunity and invading mucus layer (Ilver et al., 1998b; Magalhães & Reis, 2010). Among the known lectin-like adhesins of *H. pyroli*, blood group-binding adhesin (BabA), sialic acid binding adhesin (SabA) and LabA (also known as lacdiNAc- binding adhesin) are well studied. BabA binds to lewis b (Le^b) H-1 antigens, Sab A binds to sialyl-Le^x (Magalhães & Reis, 2010), LabA binds to lacdiNAc structures (GalNAcβ1-4GlcNAc) conjugated to MUC5AC mucins in the gastric mucosal layer (Rossez et al., 2014). Le^b mediated attachment of *H. pyroli* led to the hypothesis that Le^b mediated binding to epithelium decides whether colonization will result in a pathogenic relationship. Further studies established that Le^b mediated binding significantly alter the outcome of the infection (Hooper & Gordon, 2001). Bacteria bind directly to the epithelium only in Le^b positive transgenic animals. The immune response elicited by this binding leads to the more severe gastritis (Guruge et al., 1998). These three adhesins are most important for colonization and others are still to be studied (Figure 1.2).

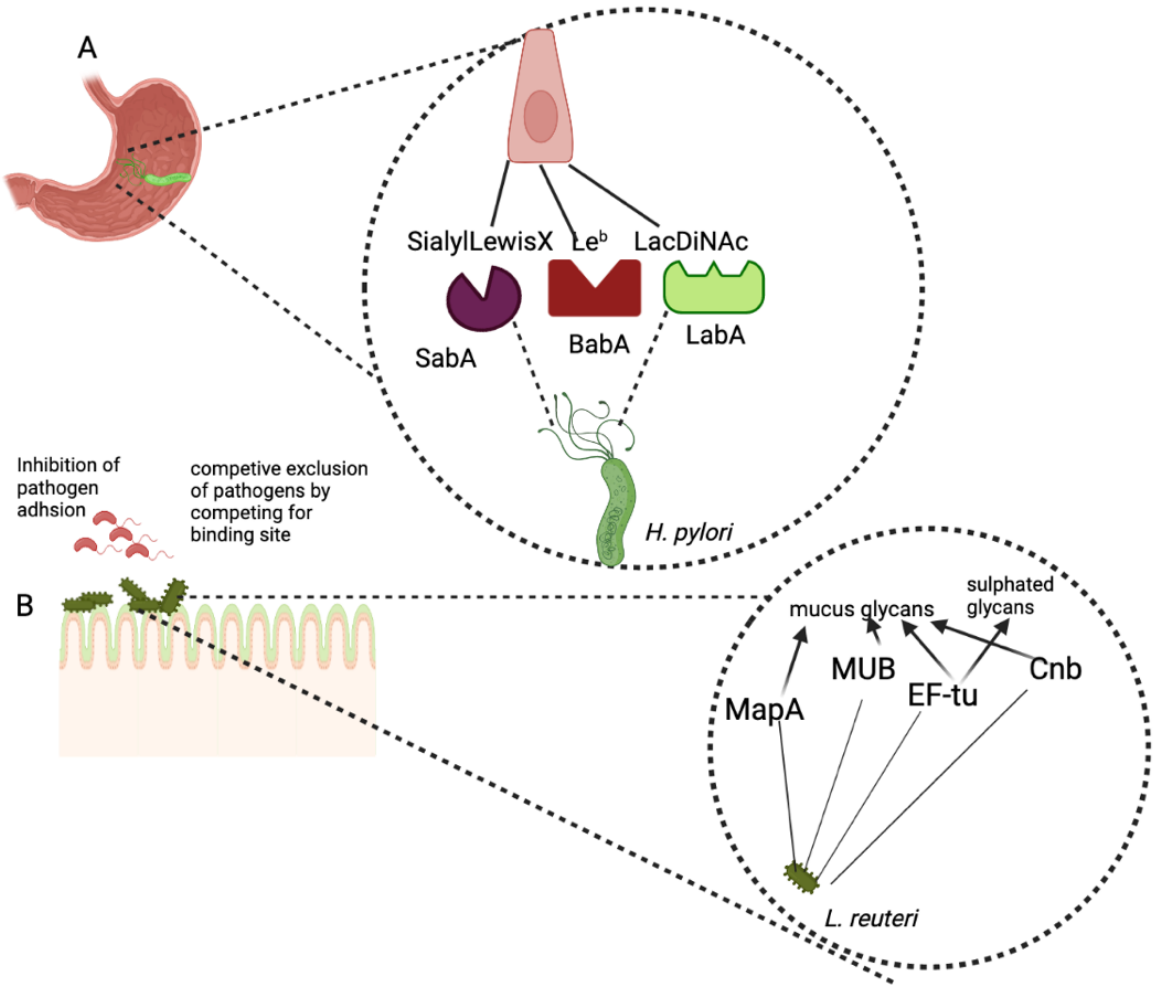

Figure 1.2 Examples of glycan mediated interactions of pathogenic and commensal gut bacteria with the host. A) *Helicobacter pyroli* binds to the gastric epithelium glycans with the given sugar residues for colonization. B) *Limosilactobacillus reuteri* bind to mucus glycans and inhibits pathogen-host interaction. Several surface mucus binding proteins of *L. reuteri* are shown. MapA: mucus adhesion promoting protein A, MUB: mucus binding protein, EF-tu: Elongation factor Tu, Cnb: collagen binding protein. This figure was created in Biorender platform.

1.1.4.1.3 Colonization of C. jejuni

Campylobacter jejuni is a human pathogen, and its ability to recognize a broad range of host cell surface glycosylation has been shown to be crucial for adherence and infectivity (Christopher J. Day et al., 2009). *C. jejuni* can be either infectious or commensal in different hosts depending on the host and environmental factors (Christopher J. Day et al., 2009). *In vivo* studies have shown fucosylated glycans of human breast milk proteins and free fucosylated oligosaccharides can protect against the incidence *of C. jejuni* infections (Morrow et al., 2005). It was suggested that feeding fucosylated glycans confer protection by binding to *C. jejuni* preventing binding to host glycans. Glycan array analysis was used to demonstrate that *C. jejuni* binds to fucose structures (Christopher J. Day et al., 2013). Initial interactions for *C. jejuni* 11168 were speculated to be mediated by the sialylated and mannosylated structures such as those found on human glycoprotein MUC1, a releasable decoy abundant in human intestinal mucosa (Juge, 2012; McAuley et al., 2007). Persistent *C. jejuni* infection in crypts of the intestinal epithelium seems to depend on fucose and galactose, the structures more readily abundant on the gel forming mucin such as MUC2 (Christopher J. Day et al., 2009; Christopher James Day et al., 2012; Juge, 2012). Glycan array results also showed that *C. jejuni* recognizes mannose and sialic acid more often after environmental stress and binds to galactose and fucose in host-like conditions (Christopher J. Day et al., 2013). Day et al (2013) illustrated the glycan binding profiles of twelve different strains of *C. jejuni* to determine the glycan binding differences between the invasive and non-invasive strains. They showed that there is a similarity in general types of structures of glycan recognized. *C. jejuni* recognizes a broad range of both α and ß linked galactose, which might explain the broad

host range. Glycan array results suggest that *C. jejuni* has fucose and galactose binding lectins. Further confirmatory studies will help in the understanding and characterization of these lectin-like components on the surface of *C. jejuni*.

1.1.4.2 Fimbriae in E. coli infection

Fimbriae are proteinaceous adhesion components expressed on the bacterial envelope, evolutionarily adapted by *Escherichia coli* strains for the colonization of epithelial cells. Lectin domains of fimbrial adhesins of enterotoxigenic *E. coli* (ETEC) have been identified using glycan arrays of the consortium for functional glycomics (CFG) (Lonardi et al., 2013). Fimbriae, F17 and F18 have been seen on the surface of enterotoxigenic and septicemic *E.coli* strains (El Mazouari et al., 1994; Hahn et al., 2000; Kapitany et al., 1979); these comprise of few thousand copies of major pilin (F17A, and FedA respectively), several minor pilin proteins and a single two domain adhesin at the tip (F17G (Lintermans et al., 1988), and FedF (Smeds et al., 2001) respectively). F17G selectively recognize glycans with a terminal GlcNAc, abundant in intestinal mucins (Lonardi et al., 2013).

F18-fimbriated *E. coli* are associated with postweaning diarrhea in pigs (Kaper et al., 2004). The adhesion of F18-fimbriated *E. coli* to the susceptible pigs is mediated by the minor fimbrial subunit FedF (Coddens et al., 2009). FedF recognizes epithelial glycosphingolipids having blood group determinants ABH determinants on type 1 core, and blood group A type 4 heptaglycosylceramide (Coddens et al., 2009). Additionally, epithelial receptor (F18R) was identified to mediate the

binding of F18-fimbriated bacteria with glycosphingolipids. Investigation of the glycosphingolipids recognized by the F18-fimbriated bacteria revealed that these bacteria can also interact with blood group A type 4 tetraglycosylceramide (Coddens et al., 2009). Complete spectrum of the glycans involved in adhesion of these fimbriated *E. coli* would enhance the understanding of their infectivity.

1.1.5 Sugars in protective roles against pathogens

Many intestinal pathogens, including *Salmonella* and enteropathogenic *E. coli,* express adhesins that recognize the carbohydrates moieties expressed on epithelial cells (Quintero-Villegas et al., 2013). A strategy for inhibiting bacterial adherence employs oligosaccharides that mimic epithelial binding sites (Shoaf-Sweeney & Hutkins, 2008).

Galactooligosaccharides (GOS) and other prebiotics have been shown to inhibit pathogen adherence to epithelial cells *in vitro* (Kittana et al., 2018). Prebiotics are non-digestible, fermentable oligosaccharides capable of modulating the gut microbiota (Gibson et al., 2017), they may also act as molecular decoys that competitively inhibit pathogen adherence to epithelial cells *in vitro.* GOS supplementation significantly reduced the adherence of the murine pathogen *Citrobacter rodentium* to the surface of cultured epithelial cells in a dose dependent manner (Kittana et al., 2018). However, the mouse studies revealed that treatment with GOS neither reduced the adherence of *C. rodentium* to the distal colon nor decreased its dissemination to systemic organs. The protection against *C. rodentium*–induced colonic tissue damage is provided

by dietary GOS supplementation in an anti-adherence-independent fashion (Kittana et al., 2018). These mechanisms suggest that the well-established and highly reported anti adherence observed for GOS *in vitro* do not explain the protective role of GOS observed *in vivo* (Kittana et al., 2018).

Additionally, antiadhesive effect of human milk oligosaccharides (HMOs) has been studied for a number of bacteria (Andersson et al., 1986; Simon et al., 1997). *In vitro* studies showed the effect of human milk oligosaccharides in inhibiting the adhesion of enteropathogenic *E. coli, Vibrio cholerae,* and *Salmonella fryis* to Caco-2 cells. The capacity of fucosyl-oligosaccharides to inhibit the adhesion of *E. coli* O119 is confirmed alongside 3-FL as it also resulted in a reduction in adhesion (Coppa et al., 2006). Constituting single monosaccharides of HMOs (glucose, galactose, sialic acid , N-acetylglucosamine, fucose), did not significantly inhibit the adhesion to Caco-2 cells of *E. coli, Vibrio cholerae,* and *Salmonella fryis* (Coppa et al., 2006), whereas, intact oligosaccharides were effective in inhibiting the adhesion of *V. cholera* and *E. coli* O119, but not of *S. fyris* (Coppa et al., 2006). Collectively, the basis of the anti-adherence mechanism has not been discussed in detail, one of the possible mechanisms is competitive adhesion by oligosaccharides for sugar binding proteins on bacterial cells that have yet to be validated with biochemical assays. Oligosaccharides that can act as decoys for bacteria are summarized in Table 1.2.

Table 1.1 Glycans and glycan binding bacteria.

Bacteria	Glycan/glycan component	Glycan structures	Receptor (on bacteria, such as protein	References
Campylobacter jejuni	1. Galactose 2.Mannose 3. Sialic acid 4. Chitin structures	1. Gal 2.Man 3.Neu5Ac 4. (-GlcNAc-)$_n$		(Day et al., 2013)
Helicobacter pylori	1.Lewis b (Le^b)H-1 antigens 2. sialyl-Le^x 3. lacdiNAc	1. α-Fuc(1-2)-β-Gal-(1-3)-(α-Fuc-1-4)-GlcNAc 2. Neu5Acα2-3Galβ1-4[Fucα1-3]GlcNAcβ 3.(GalNAcβ1-4GlcNAc)	Blood group-binding adhesin (BabA) Sialic acid binding adhesin (Sab A) lacdiNAc- binding adhesin (LabA)	(Ilver et al., 1998b), (Magalhães & Reis, 2010), (Rossez et al., 2014).
Lactobaccilus reuteri JCM1081	1.Fucose 2. Galβ1-3GalNAcβ 3. Asialo-GM1 4.Sulfatide 5. Galactosylceramide 6.Lactosylceramide	1.Fuc 2.Galβ1-3GalNAcβ 3.Galα1-3GalNAcβ-4Galβ1-4Glcβ1-1'Cer 4. HSO3-3Galβ1-1'Cer 5.Galβ1-1'Cer 6.Galβ1-4Glcβ1-1'Cer	MUB (mucus binding protein)	(T. Mukai et al., 1998)

F-18 fimbriated *E. coli*	1. H5 type 1 2. B6 type 1 3. A6 type 1 4. A7 type 1 5. B7 type 1 6. A8 type 1 7.A9 type 1	1. Fucα2Galβ4GlcNAcβ3Galβ4Glcβ1Cer 2. Galα3(Fucα2)Galβ3GlcNAcβ3Galβ4Glcβ1Cer 3.GalNAcα3(Fucα2)Galβ3GlcNAcβ3Galβ4Glcβ1Cer 4.GalNAcα3(Fucα2)Galβ3(Fucα4)GlcNAcβ3Galβ4Glcβ1Cer 5.Galα3(Fucα2)Galβ3(Fucα4)GlcNAclβ3GallβGlclβ1Cer 6.GalNAcα3(Fucα2)Galβ3GlcNAcβGalβ3GlcNAcβ3Galβ4Glcβ1Cer 7.GalNAcα3(Fucα2)Galβ3GalNAc3(Fucα2)Galβ3GlcNAcβ3Galβ4Glcβ1Cer		(Coddens et al., 2009)
Bifidobacterium bifidum EB102	1.Galactosylceramide 2.Lactosylceramide 3. Asialo-GM1 4. Asialo-GM2 5. Sulfatide	1. Galβ1-1'Cer 2.Galβ1-4Glcβ1-1'Cer 3.Galβ1-3GalNAc1-4Galβ1-4Glcβ1-1'Cer 4. GalNAcß1-4Galß1-4Glcß1-1Cer 5.HSO3-3Galß1-1'Cer		(Takao Mukai et al., 2004)
R. gnavus ATCC 29149	Sialic acid	Neu5Ac	RgCBM40	(Owen et al., 2017)

B. animalis subsp. lactis Bl-04	1. Xylotetraose 2. arabinoxylotriose	1. Xylβ1-4Xylβ1-4Xylβ1-4Xyl	BlAXBP	(Ejby et al. 2013)
B. longum subsp. *Infantis*	1.lacto-N-biose (LNB) 2. polylactosamines	1.Galβ1–3-GlcNAc 2.(-Galβ1-4GlcNAcβ1-3-)$_n$	F1SBPs	(Sela et al., 2008), (Garrido et al., 2011)
B. bifidum	1.lacto-N-biose (LNB), 2.Galacto-N-biose (GNB) 3.Lacto-N-tetraose (LNT)	1. Galβ1-3GlcNAc 2.Galβ1-3GalNAc 3. Galβ1-3GlcNAcβ1-3Galβ1-4Glc	Homolog of F1SBP	(Suzuki et al., 2008)
Streptococcus pneumoniae (pneumococcus)	1. lactose 2. N-acetylglucosamine	1. Galβ1-4Glc 2. GlcNAc	beta-galactosidase BgaA	(A. K. Singh et al., 2014)

Table 1.2 Protection by oligosaccharides against pathogens.

Decoy carbohydrates	Bacteria	Mechanism
Lacto-N-tetraose (LNT), lacto-N-neotetraose (LNnT):	*S. pneumoniae* (Andersson et al., 1986)	Anti-adherence based
Fucosyl-tetra-pentasaccharides:	*E. coli* (Cravioto et al., 1979)	Anti-adherence based
Fucosyloligosaccharides:	*V. cholerae, E. coli* 0119(Coppa et al., 2006)	Anti-adherence based
Fucosyl-oligosaccharides	*L. monocytogenes*(Coppa et al., 2006)	Anti-adherence based
Galacto-oligosaccharides	*C. rodentium* (Kittana et al., 2018).	Anti-adherence mechanism independent

1.1.6 Glycan recognition and utilization by gut bacteria

Many members of the gut microbiota have developed glycan utilization strategies that are compatible with their specific cell structures and nutrient niches (Briggs et al., 2021). Glycan utilization systems within the dominant phyla of Bacteroidetes, Firmicutes, and Actinobacteria have some common functional components:1) non-catalytic-glycan binding proteins at the cell surface that initiate substrate attachment and assist with transport; 2) ensembles of linkage specific CAZymes, including polysaccharide lyases and glycoside hydrolases; 3) carbohydrate transporters; and 4) carbohydrate sensors/transcriptional regulators (Briggs et al., 2021). Although above strategies are used by gut microbes for glycan utilization, non-catalytic glycan binding proteins are mainly discussed in the following sections. Here, we summarize the surface glycan-binding proteins and their role in dietary and host derived glycan-utilization. Glycan binding proteins are scarcely studied for most gut commensals yet there are some species of gut symbionts that have been adequately characterized for their surface glycan binding proteins and glycan utilization system. Moreover, we provide examples of surface binding proteins in glycan recognition and glycan utilization from the dominant phyla with *Ruminococcus gnavus, Bacteroides,* and *Bifidobacteria*.

1.1.6.1 Glycan recognition and utilization in Ruminococcus gnavus

The adaptation of intestinal bacteria to the mucosal environment is well exemplified by *Ruminococcus gnavus* and dependent on mucosal glycan recognition and utilization characteristics (Owen et al., 2017). The niche created by the mucus layer glycans provides bacteria with

attachment sites as well as a source of nutrients. Two key glycan residues are found in the mucus layer, sialic acid and fucose (Lewis & Lewis, 2012; Robbe et al., 2003). The regional distribution of sialic acid throughout the length of the gut is dynamic, with the ratio of sialic acid to fucose increasing aborally (Robbe et al., 2003). Notably, the ratio of sialic acid to fucose shows a reversed pattern in the mouse (Holmén Larsson et al., 2013). Sialic acids such as N-acetyl neuraminic acid (Neu5Ac) and fucose residues in the terminating positions on mucin glycans are the most common target for commensal and pathogenic bacteria alike (Juge et al., 2016; Lewis & Lewis, 2012). Bacteria access these sialic acids with sialidases and utilize the sialic acid for catabolism, biofilm formation, or incorporation into surface glycan structure, which impacts host immune modulation (Juge et al., 2016; Lewis & Lewis, 2012; Ouwerkerk et al., 2013; Tailford et al., 2015). The sialidases also contain the carbohydrate-binding modules (CBM), such as sialic acid-specific CBM40 (Moustafa et al., 2004; Tailford, et al., 2015). The discovery of intramolecular trans-sialidase in *Ruminococcus gnavus* led to the understanding of the novel mechanisms of mucosal adaptation. The IT-sialidase from *R. gnavus* (Rg) ATCC 29149 (RgNanH) comprises a catalytic glycoside hydrolase domain, RgGH33, and a carbohydrate-binding module, RgCBM40 (Tailford et al., 2015).

R*g*CBM40 increases binding affinity to mucin with increasing sialic acid content indicating that binding of *R. gnavus* ATCC 29149 to intestinal mucus is sialic acid-mediated (Owen et al., 2017). It has been suggested that CBMs not only mediate the attachment of CAZymes to glycans present on host tissues but allow the adherence of the bacterium to the mucus layer (A. K. Singh et al., 2014) . The potential avidity effect of CBM40-mediated binding of sialylated mucins *in vivo* may

favor a mechanism by which CBM40 helps targeting the bacteria towards sialic acid-rich regions of the GI tract, therefore promoting bacterial colonization within the outer mucus layer (Owen et al., 2017).

There is a diverse repertoire of the CAZymes with CBM in the gut microbiota (Kaoutari et al., 2013), although the specificity of binding for most CBMs is poorly characterized. Bioinformatic analyses of bacterial genomes showed that the canonical RgCBM40 domain is widespread among Firmicutes, yet is distinct from CBM of CAZyme amongst Bacteroidetes (Kaoutari et al., 2013). It is speculated that CBMs are important in shaping the spatial distribution of symbiotic bacteria among physical niches in the gut.

1.1.6.2 Sugar utilization and recognition in Bacteroides

The sugar utilization system, the mechanism of colonization, and the mucosal adaptation of *Bacteroides* are the most extensively studied. Glycan utilization potential of these bacteria and their colonization ability in the gut are regulated by the surface glycans and surface proteins to some extent. Early studies revealed that many strains of the *Bacteroides* species possess very broad glycan degradation/utilization potential, with some strains able to target dozens of different complex glycans (Salyers et al., 1977; Tomlin et al., 1988).

In the quest to understand the glycan utilization and role of glycan-bacteria interactions in these mechanisms, there have been ample efforts to characterize the structure and glycan-binding preferences of surface proteins of bacteria. Advances in omics technologies and human microbiome project have shed light on glycan utilization systems and the associated saccharolytic

genes in *Bacteroides* species. These bacteria have acquired the ability to digest complex glycans because they are an abundant nutrient source in the distal colon (Sonnenburg et al., 2005). In *Bacteroides* species, polysaccharide utilization is driven by the binding of the putative outer membrane receptor complex, and then it is translocated into the periplasm, where it is degraded by enzymes (Reeves et al., 1997), rather than by extracellular polysaccharide degrading enzymes.

Bacteroides have a highly organized system for glycan utilization where all genes regulating the degradation of the dietary or host-derived glycans are assembled in a single polysaccharide utilization locus (PUL) (Martens et al., 2009). The most well-studied PUL-encoded glycan uptake systems are the starch utilization system (Sus) and Sus-like system, which play an essential role in dietary glycan and host-glycan utilization (Martens et al., 2009) and it encodes for eight proteins, SusRABCDEFG. Despite being well-characterized structural and functional characteristics of sus-like proteins, strategies of recognition of glycans by these proteins are not adequately understood. A recent review (R. P. Singh, 2019) discussed glycan recognition by different *Bacteroides* species from the view of Surface Glycan Binding Proteins (SGBPs) and it listed their glycan-binding preferences. SGBPs of starch are particularly well characterized in *B. thetaiotamicron* and are grouped as SusEFG. SusEF have lesser role in glycan binding as compared to SusD (Reeves et al., 1997), while SusG is a starch-degrading enzyme (Shipman et al., 1999). Super-resolution imaging of SusG has shown that SusG interacts with other membrane proteins (i.e. SusDEF) in presence of glucose, maltose, and amylopectin (Karunatilaka et al., 2014) indicating that SusG is also important to bind to these sugars. While the glycan binding of *B. thetaiotamicron* is well understood and is the model organism for Bacteroides-glycan interactions, details of glycan-

binding by SGBPs in other *Bacteroides* species is limited due to the complexity of structural and functional characterization of these glycan binding proteins.

In addition to binding to dietary polysaccharides, mucin-glycan foraging of *Bacteroides* species have also been linked to sus-like proteins. In *Bacteroides thetaiotamicron* SusD-like protein BT1043 is an outer membrane protein associated with O-glycan utilization of host mucin (Martens et al., 2008). Structural and functional characterization of a SusD like protein NanU, a SusD family protein from *B. fragilis* has shown high binding affinity to sialic acid (Phansopa et al., 2014). However, more studies will be required to characterize the full spectrum of the glycan-binding by these Sus-like proteins. A recent study used fluorescently labeled, microscopic glass beads containing different bound glycans to examine carbohydrate-based adhesion of *Bacteroides* strains (Patnode et al., 2021). New high throughput methods like these along with various high throughput glycan screening platforms will enhance understanding of glycan recognition by glycan utilization machinery of bacteria.

1.1.6.3 Glycan recognition and utilization in Bifidobacteria

Bifidobacteria constitute approximately 60-90% of total gut microbiota in early life (Odamaki et al., 2016). Unlike *Bacteroides, Bifidobacteria* preferentially utilize short oligosaccharides (Sela et al., 2008). Similar to *Bacteroides, Bifidobacteria* possess membrane-associated glycohydrolase (GHs) for digesting long oligo/polysaccharides, and solute-binding proteins (SBPs) for recognizing those digested short oligosaccharides before importing into the cytoplasm (R. P.

Singh, 2019). A conserved SBP (BlAXBP) of *B. animalis subsp. lactis* Bl-04 is composed of two domains which show specificity in binding to xylotetraose and arabinoxylotriose (Ejby et al., 2013). Another SBP (BlG16BP) of Bl-04 shows generality in binding to glycans with an α-(1,6)- galactoside/glucoside link (Ejby et al., 2016). It has been suggested that such sugar-binding genericity of SBPs (B1AXBP and BlG16BP) may be an adaptive strategy to persist in a nutrient-poor environment (R. P. Singh, 2019). Additionally, other SBPs, *Bl*MnBP1, and *Bl*MnBP2 are characterized to have a different affinity towards mannan-oligosaccharides. The diversification of the SBPs highlights the ability of these proteins to modulate the oligosaccharide uptake preferences at the time of adaptation to specific ecological niches (Ejby et al., 2019).

Genomic analysis of a milk oligosaccharide utilizing bacterial strain, *B. longum* subsp. *Infantis* ATCC15697, suggests the potential of encoding notable numbers of family 1 SBP (solute binding protein) (F1SBP) of ABC transports, 7 of these are encoded by 43-kbp gene cluster for utilizing diverse human milk oligosaccharides (Sela et al., 2008). Glycan array analysis of 11 F1SBPs showed that these proteins have binding affinities towards oligosaccharides belonging to lacto-N-biose (LNB, Galβ1–3-GlcNAc) and branched or unbranched polylactosamines (Garrido et al., 2011). Notably, some of the F1SBPs can bind to multiple glycans, such that Blon_2177, Blon_2347, Blon_2344 and Blon_0883 can recognize multiple glycans including mucus glycans (Garrido et al., 2011).

A homolog of F1SBP in mucus-glycan utilizing strains of *B. bifidum* can also recognize lacto-N-biose (LNB), galacto-N-biose (GNB) and lacto-N-tetraose (LNT) (Suzuki et al., 2008).

1.1.7 Glycan array technologies for bacteria

Different microarray technologies have been used to explore the bacterial surface glycans. Antibody and lectin microarrays have played an important role in illustrating bacterial glycosignatures, enabling differentiation among strains and to evaluate the difference in the glycan structures due to changes in environmental conditions (Campanero-Rhodes et al., 2020). Several lectin microarrays have been utilized for the analysis of glycosylation of bacteria such as lectin microarray (Gao et al., 2010; Liu et al., 2016). Several bacterial glycan microarrays have been designed for serodiagnosis of infections and bacterial glycan arrays for the identification of novel vaccine candidates and bacterial glycan arrays for the study of bacterial glycan binding proteins. The list of glycans printed in the arrays includes polysaccharides, lipopoly/lipooligosaccharides, (lipo) teichoic acids, and peptidoglycans, as well as sequence-defined oligosaccharide fragments (Campanero-Rhodes et al., 2020). For observing host- bacteria glycan- glycan interactions, purified bacterial glycans have been used on the microarray printed host glycans (Christopher J. Day et al., 2015). However, the potential of the carbohydrate microarrays for studying host - bacteria interactions is inadequately explored due to limited library of probes that are used for building the array.

1.1.7.1 DNA encoded Next generation glycan array

Despite some limitations, traditional microarray can theoretically be used to directly assay glycan-binding preferences of intact bacterial cells. However, it captures limited amount of adhesion because the two dimensional surface of slides can provide partial interaction of large intact cells

(Yan et al., 2019). To enhance the robustness and high throughput ability of glycan microarray, the next generation glycan microarray (NGGM) has been developed recently, which is based on the artificial DNA-coding of glycans (Yan et al., 2019). NGGM is a mixture of glycans and glycoconjugates which are coded with unique oligonucleotide code. Its solution phase binding assay makes it amenable to screen glycan binding of intact cells.

1.1.7.2 Phage-display glycan arrays

With advances in phage-display technologies, engineered phages displaying monosaccharides and oligosaccharides, referred as "glycophages", have been designed as alternative to previously mentioned glycan libraries (Çelik et al., 2015). In these glycophage libraries, there are few limitations, such as authors were not able to quantify the density of glycans on different glycophages, and it can only be used for monovalent display of the glycans.

1.1.7.2.1 Liquid glycan array (Genetically encoded multivalent phage-display enabled liquid glycan array)

The genetically encoded multivalent glycan array on M13 phage was developed and recently published by the Derda group (Sojitra et al., 2021). Unlike the glycophage display system that employs the biosynthesis of glycans, LiGA decouples DNA encoding and glycan display from biosynthesis. Chemical manufacturing of LiGA allows the repurposing of many chemical approaches previously employed in construction of the traditional PGAs. LiGA is built on

filamentous M13 phage particles with silent DNA barcodes inside the phage genome. The authors chemically conjugated glycans to a subset of ~2700 copies of major coat protein pVIII to produce a multivalent display of ~150-1500 copies of glycans. LiGA has been used to test glycan binding of pure lectins, glycan binding proteins and mammalian cells but it has not been previously used for testing glycan binding of bacteria cells.

1.1.8 Research Objectives

Objective_1: To understand the glycan binding preferences of different *L. reuteri* isolates using Liquid Glycan Array (LiGA).

Hypothesis: We hypothesize that glycan binding profiles differ between *L. reuteri* from different hosts.

Objective_2: To understand the glycan binding preferences of bacteria from different taxonomic groups using Liquid Glycan Array (LiGA).

Hypothesis:

We hypothesize that taxonomically closer gut bacteria have similar glycan binding profiles.

2 Chapter 2: Glycan binding Profiles of the *Limosilactobacillus reuteri*

2.1 Introduction

Vertebrates have evolved in association with the microbes in their digestive tracts. Gut microbes not only enhance the metabolic ability of the host through the nutrient provision but also exclude pathogens and help in the host immune system (Dethlefsen et al., 2007). It is evident that vertebrates benefit from the symbiotic relationship with the gut microbes and this symbiotic relationship plays a role in the evolution of vertebrates. The taxonomic profile of the vertebrate microbiota is largely host-specific (Dethlefsen et al., 2007; Ley et al., 2008) and in some cases, congruent with the evolution of host species. This seeming relatedness between the microbial community composition and host phylogeny has been postulated as an evidence for co-evolution (Fraune et al., 2010; Ley et al., 2008). Much of this information is derived from the 16s rRNA gene analysis, this gene has evolved too slowly, especially given the more recent diversification of contemporary vertebrate species compared to their bacterial symbionts.

However, definitive evidence for strong associations of specific lineages with vertebrate hosts over evolutionary time scales has not been provided by 16s rRNA data (Frese et al., 2011). Co-diversification of specific bacterial lineages hosts has been determined by analysis of fast evolving gene phylogenies (Falush et al., 2003; Moeller et al., 2016). However, even in the few well-known

cases, the mechanisms underlying the evolution of these microbes and the outcomes of this evolution remain unclear.

Trends of community similarity provide evidence for co-diversification of gut communities with their hosts, suggesting that there are host specific evolutionary interactions between hosts and their microbiota (Ley et al., 2008). Additionally, some gut microbes are specialists, such as *Helicobacter pylori*, which is highly host specific and has been used to track human migrations over long time spans (Linz et al., 2007). On other hand, many microbes in the mammalian gut are shared across host species (Ley et al., 2008), which shows that some members of the gut are generalists inhabiting multiple ecosystems. This evolutionary strategy is illustrated by the commensal *Escherichia coli*, which have broad host range and alternate between niches within the environment and their vertebrate hosts (Tenaillon et al., 2010; Touchon et al., 2009). But there are not a lot of vertebrate gut symbionts for which host specificity has been illustrated. Little is known about the mechanisms by which gut microbes can evolve stable associations with their hosts that would allow for reciprocal evolutionary interactions between bacterial lineages and host genotypes (Frese et al., 2011). The gram-positive bacterium *Limosilactobacillus. reuteri* has been used as a model organism to study the evolutionary strategy of vertebrate gut symbionts because this organism colonizes the GIT of vertebrates as diverse as humans, pigs, mice, rats, and chickens. In rodents, pigs and chickens, it is one of the dominant species in the GI tract and forms the biofilm like associations with the stratified squamous epithelial lining of the proximal regions of the digestive tract (Brooks et al., 2003; Leser et al., 2002; Salzman et al., 2002; Tannock, 1992; Walter, 2008). It is observed that strains of *L. reuteri* from global sources comprised distinct phylogenetic

clusters that can be detected with the Multilocus Sequence Analysis (MLSA) and Amplified fragment length polymorphism (AFLP), and these clades show significant association with host origin (Jensen et al., 2014). Besides these genotypic patterns, an adaptive evolutionary process is also demonstrated by phenotypic characteristics of *L. reuteri* strains in terms of the ecological performance in the gut and adhesion to epithelial cells (Oh et al., 2010; Schreiber et al., 2009; Suegara, 1975). Genomic approaches in combination with experiments in animal models offer mechanistic insight into the evolution and ecology of microbial symbionts of vertebrates. Previous study showed that only rodent strains of *L. reuteri* colonize the gut of reconstituted Lactobacillus free mice in high number, while isolates from humans, swine, and chicken form either lower populations or fail to colonize (Frese et al., 2011). A microarray analysis of 57 *L. reuteri* strains revealed specific gene combination in host adapted lineages of *L. reuteri.* Among genomic features associated with the host origin, large surface proteins of rodent lineage showed specificity towards this lineage and these eleven large surface proteins were rare in isolates from pigs and poultry and absent from lineage human isolates of MLSA lineage II. Several of these surface proteins are predicted to be involved in epithelial adhesion and biofilm formation. Proteins Lr_69656, Lr_70131, Lr_70134, Lr_70135, Lr_70581, and Lr_71380 contained putative mucin-binding MucBP domains and other domains involved in the extracellular matrix binding (Frese et al., 2011).

Although adhesion of these surface proteins to epithelial cells is known to be mediated by glycans, direct binding of *L. reuteri* with specific glycans remains to be demonstrated. To determine whether glycan binding profile was similar between *L. reuteri* from the same host, glycan binding

of 16 *L. reuteri* isolates consisting of 4 isolates from each of murine, porcine, poultry and human lineages was tested. Glycan binding was assessed using a novel Liquid Glycan Array (LiGA) which comprises a library of glycosylated M13 phage particles with silent DNA barcodes in the phage genome (Sojitra et al., 2021). Validation of glycan binding of LiGA was carried out using mannose binding lectin, ConA and wild type and fimH-mutant *E. coli*. We hypothesized that the glycan binding profiles of *L. reuteri* are host specific.

2.2 Materials and methods

2.2.1 Bacterial culture

To determine whether glycan binding profile was similar between *L. reuteri* from the same host, glycan binding of 16 *L. reuteri* isolates consisting of 4 isolates from each of murine, porcine, poultry and human lineages was tested. Bacterial strains used in this study are listed in Table 2.1. *L. reuteri* strains were grown in de Man, Rogosa, and Sharpe (MRS) medium (Difco) supplemented with 5% maltose and 10% fructose under anaerobic conditions (5% CO2, 5% H2, and 90% N2) in anaerobic chamber. All E. coli strains were grown in Luria-Bertani (LB) broth with agitation. Each strain was grown and the optical density (measured using spectrophotometer [4001/4, Thermo Fisher Scientific]) and CFU/ml were noted at 12 hrs, 18 hrs, and 24 hrs.

2.2.2 Sanger sequencing

Each bacterial strain was tested by the colony PCR using the sanger sequencing using 16S rRNA primers using 8F and 926R primer (Coolen et al., 2005). Each 50 µl PCR reaction mixture consisted of 2 µl of 10 µM forward primer 8F, 2 µl of 10 µM reverse primer 926R, 2 µl of 10 mM deoxynucleotide triphosphate mix (Invitrogen, Carlsbad, CA), 5 µl of 10x Taq polymerase buffer (Invitrogen), 2 µl of 50 mM $MgCl_2$ (Invitrogen), 0.5 µl of 1 U/µl Taq polymerase (Invitrogen), and a small amount of colony. The thermal cycling program included an initial 10 min denaturation step at 94°C, 40 cycles of 94°C for 30 s, 56°C for 30 s, and 72°C for 1 min, and a final 7 min extension at 72°C. PCR products were visualized by 1% agarose gel electrophoresis followed by the SYBR Safe DNA gel staining (Invitrogen). Sequencing results were verified using nucleotide blast.

Table 2.1 List of *L. reuteri* strains used in this study.

Strain	Host Origin	Source
L. reuteri DSM20016. Human	Human	JGI 2671180761
L. reuteri mm2-3. Human	Human	JGI 2502171171
L. reuteri Sd2112.human	Human	JGI 650716048
L. reuteri M27415.human	Human	JGI 2687453659
L. reuteri I5007.pig	Pig	JGI 2554235423
L. reuteri Lp167. Pig	Pig	JGI 2599185361
L. reuteri ATCC 53608. Pig	Pig	EMBL LN906634
L. reuteri.---pig	pig	Willing's lab
L. reuteri Jcm 1081.poultry	Poultry	JGI 2684623011
L. reuteri 1366.poultry	Poultry	JGI 2684623010

L. reuteri AP3.Poultry	Poultry	GCA_014145445
L. reuteri CSF8.poultry	Poultry	JGI 2684623009
L. reuteri Ml3.mouse	Rodent	JGI 2506381016
L. reuteri Tmw1.656. rat	Rodent	JGI 2534682350
L. reuteri Lpuph1.rat	Rodent	JGI 2506381017
L. reuteri 100-23. rat	Rodent	JGI 2500069000

2.2.3 Liquid Glycan array

Liquid glycan array LiGA-ED was prepared by Mirat Sojitra from Dr. Ratmir Derda's lab . It contains 96 (counting different glycosylation density as unique structure) the different glycan structures. Glycan structures present in the LiGA-ED are listed in Table 2.2.

Table 2.2 List of glycans in Liquid glycan array LiGA-ED.

Glycan-no.	Modification	Common name	Glycan-density (glycans/phage)
1.	Galb1-4Glcβ-P4	Lac-peg4	1080
2.	Manα-S6	aMan	840
3.	Manα1-6[Manα1-3]Manα-S6	(Man)3	<8
4.	Manα1-6[Manα1-3]Manα-S6	(Man)3	1300
5.	Manα1-6[Manα1-3]Manα-S6	(Man)3	1730
6.	Fucα1-2Galβ1-4Glcβ-Sp	2'FL	950
7.	Fucα1-2Galβ1-3GlcNAcβ-Sp	H type 1	700
8.	Fucα1-2(Galβ1-4GlcNAcβ1-3)2β-Sp	H2	430
9.	Fucα1-2(Galβ1-4GlcNAcβ1-3)3β-Sp	H3	190
10.	Fucα1-2Galβ1-4GlcNAcβ-Sp	H-type 2	540
11.	GalNAcα1-3[Fucα1-2]Galβ1-3GlcNAcβ-Sp	A tetra type 1	700
12.	GalNAcα1-3[Fucα1-2]Galβ1-4GlcNAcβ-Sp	A tetra type 2	920
13.	GalNAcα1-3[Fucα1-2]Galβ1-4Glcβ-Sp	A tetra L	590
14.	Galα1-3[Fucα1-2]Galβ1-4Glcβ-Sp	B tetra L	920

15.	GlcNAcβ1-3Galβ1-4GlcNAcβ-Sp	GNLN	810
16.	GalNAcβ1-3Galα1-4Galβ1-4Glcβ-Sp	Globoside-P	1030
17.	GalNAcβ1-3Galα1-4Galβ1-4Glcβ-Sp	Globoside-P	730
18.	GlcNAcβ-Sp	GN	1050
19.	GalNAcβ1-3Galα1-4Galβ1-4GlcNAcβ-Sp	P1 tetra	970
20.	GalNAcβ1-4GlcNAcβ-Sp	LacDiNAc	50
21.	GlcNAcβ1-3Galβ1-3GlcNAcβ-Sp	3'GN type1	860
22.	Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ-Sp	LNnT	240
23.	Galα1-3[Fucα1-2]Galβ1-3GlcNAcβ-Sp	B tetra type 1	620
24.	Galα1-3[Fucα1-2]Galβ1-4[Fucα1-3]GlcNAcβ-Sp	2'F-B type 2	220
25.	Galα1-3[Fucα1-2]Galβ1-4[Fucα1-3]GlcNAcβ-Sp	2'F-B type 2	760
26.	Galα1-3[Fucα1-2]Galβ1-4GlcNAcβ-Sp	B tetra type 2	970
27.	Galβ1-4[Fucα1-3]GlcNAcβ-Sp	Lex	810
28.	(Galβ1-4[Fucα1-3]GlcNAcβ1-3)2β-Sp	Di-Lex	410
29.	Galβ1-3GlcNAcβ1-3Galβ1-4[Fucα1-3]GlcNAcβ-Sp	Lec-LeX	570
30.	Galβ1-3[Fucα1-4]GlcNAcβ1-3Galβ1-4[Fuca1-3]GlcNAcβ-Sp	LeALex	350
31.	GlcNAcβ1-3Galβ1-4Glcβ-Sp	LNT-2	430
32.	Fucα1-2Galβ1-4[Fucα1-3]GlcNAcβ1-3Galβ1-4[Fucα1-3]GlcNAcβ-Sp	Ley-Lex	350
33.	Fucα1-2Galβ1-4[Fucα1-3]GlcNAcβ1-3(Galβ1-4[Fucα1-3]GlcNAcβ1-3)2β-Sp	Ley-Di-Lex	510
34.	Galα1-3Galβ1-4[Fucα1-3]GlcNAcβ-Sp	Gala3Lex	620
35.	KDNα2-3Galβ1-4GlcNAcβ-Sp	3'KDNLN	510
36.	Galα1-3Galβ1-3GlcNAcβ-Sp	Gala3-type1	350
37.	Galα1-4Galβ1-4Glcβ-Sp	Pk	860
38.	Galα1-4Galβ1-4GlcNAcβ-Sp	P1 tri	620
39.	Galα1-3Galβ1-4GlcNAcβ-Sp	B2 tri	350
40.	Galα1-3Galβ1-4Glcβ-Sp	Galili-tri	1000
41.	Fucα1-2Galβ-Sp	Di-N3	970
42.	GalNAcα1-3[Fucα1-2]Galβ-Sp	Tri-AN3	1080
43.	Neu5Acα2-3Galβ1-3GlcNAcβ-Sp	3'SLec	430
44.	Neu5Acα2-3GalNAcβ1-4GlcNAcβ-Sp	3'SLDN	460
45.	Neu5Acα2-3[GalNAcβ1-4]Galβ1-4Glcβ-Sp	GM2	590
46.	Neu5Acα2-3[Galβ1-3GalNAcβ1-4]Galβ1-4Glcβ-Sp	GM1	460

47.	Galβ1-3GlcNAcβ-Sp	Lec	680
48.	Galβ1-4GlcNAcβ-Sp	LacNAc, LN	970
49.	Galβ1-3[Fucα1-4]GlcNAcβ-Sp	LeA	950
50.	(Galβ1-4[Fucα1-3]GlcNAcβ1-3)3β-Sp	Tri-Lex	430
51.	(Galβ1-4GlcNAcβ1-3)2β-Sp	Di-LN	650
52.	(Galβ1-4GlcNAcβ1-3)3β-Sp	Tri-LN	380
53.	Neu5Acα2-8Neu5Acα2-8Neu5Acα2-8Neu5Acα2-3[GalNAcβ1-4]Galβ1-4Glcβ-Sp	GQ2	140
54.	Neu5Acα2-8Neu5Acα2-8Neu5Acα2-3[GalNAcβ1-4]Galβ1-4Glcβ-Sp	GT2	160
55.	Neu5Acα2-8Neu5Acα2-8Neu5Acα2-3Galβ1-4Glcβ-Sp	GT3	510
56.	Neu5Acα2-8Neu5Acα2-3[GalNAcb1-4]Galβ1-4Glcβ-Sp	GD2	460
57.	Neu5Acα2-8Neu5Acα2-8Neu5Acα2-8Neu5Acα2-3Galβ1-4Glcβ-Sp	TetraSLac	80
58.	Neu5Acα2-3[GalNAcβ1-4]Galβ1-4GlcNAcβ-Sp	CT|Sda	350
59.	Neu5Acα2-3[Neu5Acα2-3Galβ1-3GalNAcβ1-4]Galβ1-4Glcβ-Sp	GD1a	110
60.	Neu5Acα2-3Galβ1-3[Fucα1-4]GlcNAcβ1-3Galβ1-4[Fucα1-3]GlcNAcβ-Sp	3'SLeA-Lex	570
61.	Neu5Acα2-3Galβ1-3GlcNAcβ1-3Galβ1-4GlcNAcβ-Sp	3'SLecLN	860
62.	Neu5Acα2-3Galβ1-4GlcNAcβ1-3Galβ1-3GlcNAcβ-Sp	3'SLN-Lec	410
63.	Neu5Gcα2-3Galβ1-3GlcNAcβ-Sp	3'SLec (Gc)	350
64.	Neu5Acα2-3(Galβ1-4GlcNAcβ1-3)2β-Sp	3'S-Di-LN	1350
65.	Neu5Acα2-3(Galβ1-4GlcNAcβ1-3)3β-Sp	3'STri-LN	160
66.	Neu5Acα2-3(Galβ1-4[Fucα1-3]GlcNAcβ1-3)3β-Sp	3'S-Tri-LeX	160
67.	Neu5Acα2-3Galβ1-4Glcβ-Sp	GM3	540
68.	Neu5Gcα2-6Galβ1-4GlcNAcβ-Sp	6'SLN (Gc)	430
69.	Neu5Acα2-6GalNAcβ1-4GlcNAcβ-Sp	6'SLDN	620
70.	Neu5Acα2-6Galb1-4Glcβ-Sp	6'SL	620
71.	Neu5Acα2-6Galβ1-4GlcNAcβ1-3Galβ1-3GlcNAcβ-Sp	6SLN-Lec	300
72.	Neu5Acα2-6(Galβ1-4GlcNAcβ1-3)2β-Sp	6'S-Di-LN	300
73.	Neu5Acα2-8Neu5Acα2-3Galβ1-4Glcβ-Sp	GD3	320
74.	Neu5Gcα2-3Galβ1-4GlcNAcβ-Sp	3'SLN (Gc)	570

75.	Neu5Gcα2-3Galβ1-4Glcβ-Sp	3'SL (Gc)	490
76.	KDNα2-3Galβ1-3GlcNAcβ-Sp	3'-KDNLec	760
77.	Galβ1-4[Fuca1-3]GlcNAcβ1-3Galb1-3[Fuca1-4]GlcNAcβ-Sp	Lex-LeA	410
78.	Neu5Acα2-3Galβ1-3GlcNAcβ1-3Galβ1-3GlcNAcβ-Sp	3'S-Di-Lec	270
79.	Neu5Acα2-3(Galβ1-3[Fucα1-4]GlcNAcβ1-3)2β-Sp	3'S-Di-LeA	160
80.	Galα1-4Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ-Sp	P1 penta	620
81.	Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4GlcNAcβ-Sp	P1 penta	620

2.2.4 Plaque assay of input library and bound library

For determining the input library phage count, the plaque assay was conducted using *E. coli* K-12 strain which is susceptible to the M13 bacteriophage. For the plaque assay, *E. coli* K-12 was grown in the Luria-Bertani (LB) broth with agitation at 37 °C. Approximately 200 µl of the *E. coli* K-12 broth culture was used for each plaque assay. For input library (5×10^8 PFU/ml), 1:10, 1:100, and 1:1000 dilutions were used for plaque assay. Apart from the input phage for the glycan binding assay, the bound LiGA phage was also quantitated by the plaque assay titration. We took 2 µl of resuspended phage in HEPES buffer (50 mM HEPES, 150 mM NaCl, 2 mM $CaCl_2$, 10 mM $MgCl_2$) after the last washing step in the binding. 2 µl from each binding assay sample was directly mixed with the 198 µl of the *E. coli* K-12 and 3 ml of the top agar (0.8% LB agar, 50 °C) was added and after swirling the mixture, it was poured and spread on LB agar plates with the X-gal, IPTG, 40 µg/ml each. Another 2 µl of phage resuspended in HEPES buffer was used for serial dilution in 198 µl of the HEPES buffer and 2 µl of this diluted mixture was used in plaque assay to determine the titer of bound phage.

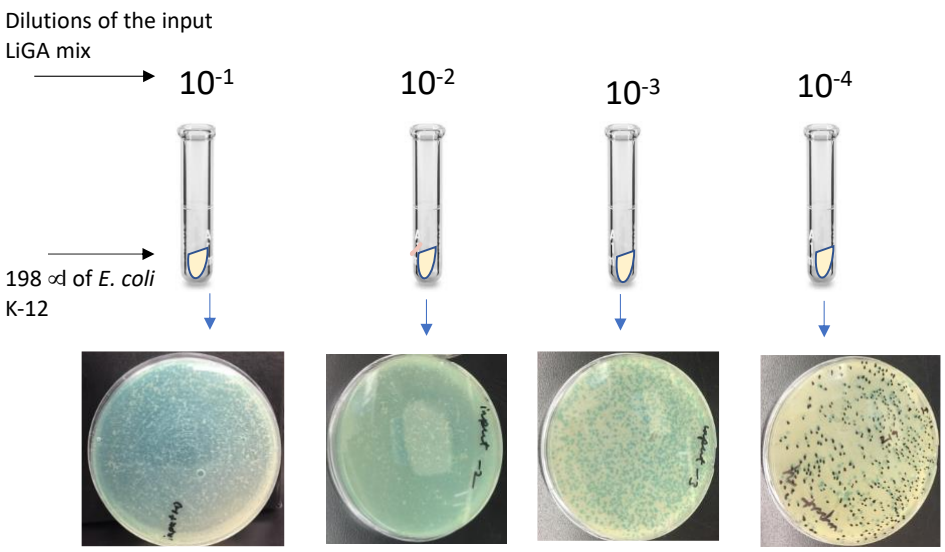

Figure 2.1 Schematic to enumerate input Library phage titer using Plaque assay.

2.2.5 Glycan binding assay for concanavalin A (ConA) coated 96 well-plate

To validate the glycan binding and integrity of LiGA, mannose binding lectin ConA was tested. Lyophilized ConA lectin (Sigma-Aldrich, C2010) was dissolved in HEPES buffer (50 mM HEPES, 150 mM NaCl, 2 mM $CaCl_2$) and store in 4 °C. Lectin solution (70 µl of 20 µg/ml) was deposited in 4 wells of the 96 well plate (Corning 9017 96- Well microplate, Cole-Parmer) and plate was sealed with 96 well plate transparent seal. Lectin coated plate is incubated at 4 °C overnight. Next day, at room temperature, the seal was removed and the protein solution from the wells was discarded. Further, wells were washed by pipetting 150 µl washing solution (50 mM HEPES, 150 mM NaCl, 2 mM $CaCl_2$, 0.1% tween-20). Then, the wells were blocked by 75 µl of

blocking solution (50 mM HEPES, 150 mM NaCl, 2 mM $CaCl_2$, 1% BSA) along with the 4 additional wells as control. After incubating for 1 h at room temperature, the blocking solution was discarded and well were washed three times with washing solution. LiGA library (50 µl of 5×10^8 PFU/ml) was added to each well and plate was incubated at room temperature for 1 h. Further, the unbound LiGA library was discarded, and wells were washed two times using washing solution, and the last wash was done by 150 µl of binding solution. For eluting the bound LiGA, 50 µl of HCl is added to each well and incubated for 9 min at room temperature, and eluted LiGA with HCl solution was transferred immediately to the Eppendorf tubes containing 5x HF buffer. The eluted phage solution was stored in 4 °C until first step PCR is performed, first step PCR was done on the same day of assay (Sojitra et al., 2021).

2.2.6 Glycan binding assay for bacterial Cells

For each bacterial strain, 8 biological replicates were grown by inoculating single colony in 5 ml of broth media. After growing the bacteria in the anaerobic chamber, the cells were harvested by centrifuging (2-5 min 10K RPM, 4 °C) the 1 ml of broth culture in microcentrifuge tube. The supernatant was washed out and the pellet was washed by HEPES buffer (50 mM HEPES, 150 mM NaCl, 2 mM $CaCl_2$, 10 mM $MgCl_2$) to remove media stuck to the cells, further pellet was resuspended in the 1 ml of HEPES buffer and volume containing 10^8 CFU was taken for assay. The aliquot of cells was spun down **(**2 min 10K RPM = 9391g, 4 °C**)** and after removing the supernatant completely, the cells were kept on ice for assay. Phage mix was prepared from the LiGA library stock containing 5×10^{10} PFU/ml by diluting in the HEPES buffer (50 mM HEPES,

150 mM NaCl, 2 mM $CaCl_2$, 10 mM $MgCl_2$) to 5×10^8 PFU/ml so that 100 µl contains 5×10^7 PFU/ml. For the binding assay, 100 µl of 5×10^8 PFU/ml LiGA mix was added to the tube containing the bacterial cells and cells were completely resuspended by inverting the tube. The cell and Liga mix was kept on ice for 1 h for binding to occur and the tube was inverted every 15 min to mix the solution. After incubation, the mixture was centrifuged (2 min 10K RPM, 4 °C) to remove the unbound LiGA as supernatant. Cells were then washed three times with 1 ml of cold (4 °C) HEPES buffer (50 mM HEPES, 150 mM NaCl, 2 mM $CaCl_2$, 10 mM $MgCl_2$). After the last washing step, cells were resuspended in 100 µl of HEPES buffer and 4 µl was kept for the plaque assay and the remaining mixture was centrifuged to remove the supernatant. The cell pellet was used for phage DNA extraction using the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific). Extracted DNA was eluted in the 25 µl of elution buffer. The eluted DNA was directly be used for the PCR.

2.2.7 Illumina sequencing

2.2.7.1 PCR amplification of the template of bound LiGA to cells

The extracted DNA was used for first step PCR, 12 µl of the eluted DNA was used for the PCR. DNA template of eluted phage after panning procedure was amplified in total volume of 30 µl with 5x Phusion® buffer (ThermoFisher Scientific, cat. no. F-518), 10 mM of each dNTPs, 10 µM NF10 forward primer, 10 µM '96' reverse primer, DMSO (Sigma D2650-100ML) and one unit of Phusion® High-Fidelity DNA Polymerase (NEB, #M0530S). In amplification of naïve libraries, volume of template (phage solution) was 4 µl. Cycling was performed using the following

thermocycler settings: a) 95 °C 3 min, b) 95 °C 10 s, c) 58 °C 30 s, d) 72 °C 20 s, e) repeat b-d for 35 cycles), f) 72 °C 5 min, f) 4 °C hold. The PCR products were verified by 2% (w/v) agarose gel in Tris-Acetate-EDTA buffer at 107 volts for ~30 min using a low molecular weight DNA ladder as standard (FroggaBio, # DM001-R500).

2.2.7.2 Indexing PCR

Illumina indexing barcodes were added using the second PCR of the template DNA from first step PCR. It was amplified in 50 µl with 5x Phusion® buffer (ThermoFisher Scientific, cat. no. F-518), 10 mM dNTPs, 10 µM SDB (silent DNA barcode) forward primer, 10 µM SDB reverse primer, and one unit of Phusion® High-Fidelity DNA Polymerase (NEB, #M0530S).

Indexed PCR products were pooled allowing 20 ng of each product in the mixture. The mixture was purified by eGel, quantified by Qubit (Thermo Fisher) and sequenced using the Illumina NextSeq paired-end 500/550 High Output Kit v2.5 (2x75 Cycles) (Sojitra et al., 2021). Data was automatically uploaded to BaseSpace™ Sequence Hub. Processing of the data is described below.

2.2.7.3. Processing of Illumina data

This protocol is adapted from (Sojitra et al., 2021). The Gzip compressed FASTQ files were downloaded from BaseSpace™ Sequence Hub. The files were converted into tables of DNA sequences and their counts per experiment. Briefly, FASTQ files were parsed into separate files based on unique multiplexing barcodes within the reads. Reads that did not contain an identifiable

multiplex barcode were discarded. Several quality control steps were performed based on i) reads that contained a Phred quality score = 0 in any position were also discarded (ii) mapping the forward and reverse primer regions was done allowing no more than one base substitution each, (iii) alignment of the forward and reverse read-end overlap was performed allowing no mismatches between forward and reverse read in the overlap region. Reads that did not match criteria (ii) and (iii) were discarded. The two ends of read-pairs that pass the filtering criteria were joined and trimmed to the DNA sequences located between the priming regions; the reads were organized in a tab-delimited text file containing the unique DNA sequences and their copy numbers. Technical replicates were combined in the same file. Using an SDB-lookup table, DNA sequences were mapped to SDB and were translated to glycans using a LiGA-specific lookup table ("LiGA dictionary"). The file with DNA reads, raw counts, and mapped glycans, were uploaded to http://ligacloud.ca/ server. Each experiment has a unique alphanumeric name (e.g., 20210813-87EDcsfGT-OB).

2.2.8 Data analysis

Data analysis was performed in R–Bioconductor and was adapted from (Sojitra et al., 2021). Comparisons and testing differences for significance in the LiGA data were performed as described in publication (Matochko et al., 2014) using DE implemented in edgeR (Robinson & Smyth, 2008). For DE analysis, three factors were considered: (1) modeling of the observed counts using a negative binomial model, (2) BH (Benjamini & Hochberg, 1995) adjustment to control the FDR at $\alpha = 0.05$ and (3) normalization of data across multiple samples using TMM normalization

(Robinson & Oshlack, 2010). To assess the significance of glycan binding in a specific experiment, the DE of the levels of the DNA barcode associated with that glycan in 'test' sets of the DNA read was compared to that of the levels of the same read in 'control' sets. For example, in bacterial cells, the 'test' dataset was association of the LiGA with the cells, whereas the control was naïve library. For binding to lectins, the 'control' dataset was association of the LiGA with blank carriers (blank wells in plates). Before DE analysis, 'test' and 'control' datasets were retrieved from the server at http://ligacloud. ca/ as tables of glycans, DNA and raw sequencing counts. DNA reads that could not be mapped to any entries in the LiGA dictionary were deleted.

2.3 Results

2.3.1 Glycan binding of controls using Liquid glycan array (LiGA)

To establish that the co-incubation conditions worked, we compared LiGA binding in a fimH mutant as compared to wildtype control with the expectation of differential mannose binding. Differential enrichment (DE) analysis (Figure. 2.2, b) identified the mannose binding of *E.coli* BW25113 (positive control) when differential enrichment (DE) analysis was done taking *E.coli* BW25113 cells as 'test' and naïve library as 'control'. Mannose binding was consistent when differential enrichment analysis was done taking *E.coli* BW25113 cells as 'test' and *E.coli* BW25113-fimH mutant as 'control' (Figure 2.2, d). We considered fold change ≥ 2, and $Q<0.05$ as significant binding for the particular glycan in LiGA library. A secondary control that can be used to check for assay variability was mannose binding lectin, ConA. It showed consistent

binding with all mannose residues (Manα1- 6(Manα1-3) Manα1- (1300 and 1700 per phage) and α-Man present in the LiGA except the Man3-(<8 per phage) (Figure 2.2, a).

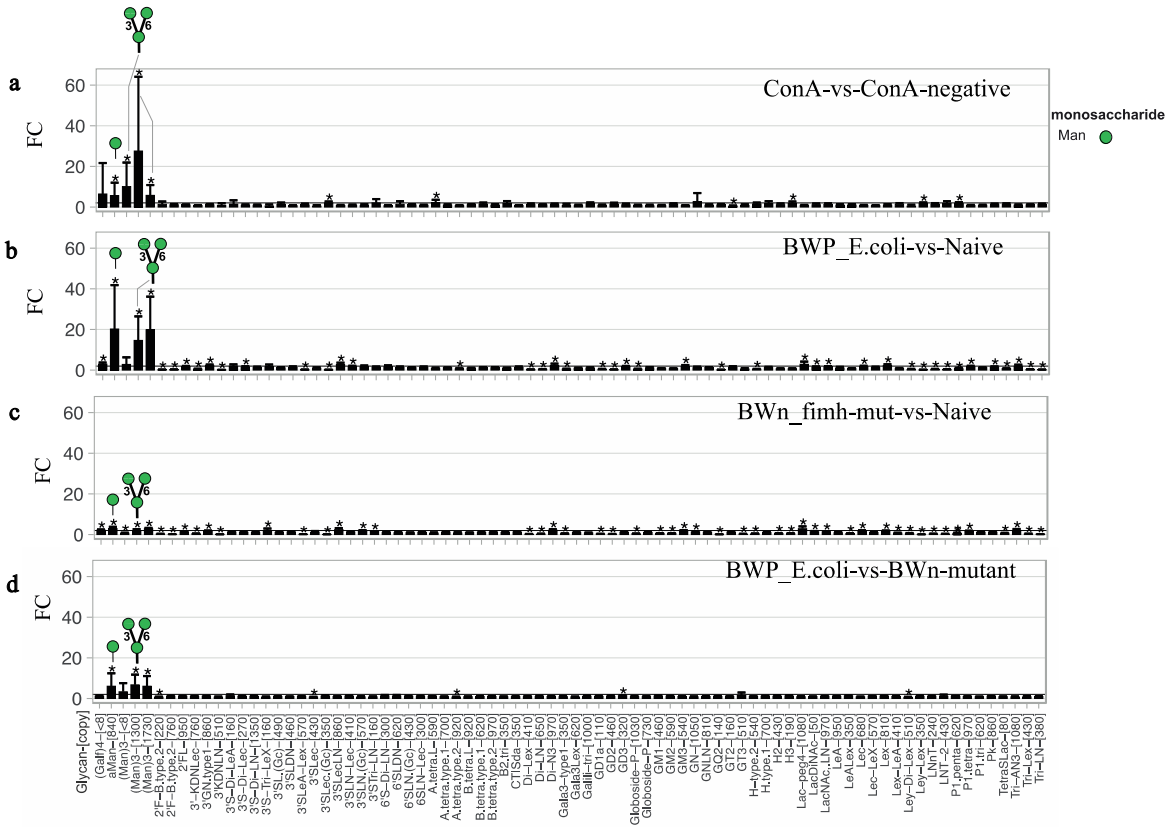

Figure 2.2 Differential enrichment analysis of controls. a) Differential enrichment analysis of the ConA (n = 4 test samples, n = 4 control samples), b) *E.coli* BW25113-fimH mutant (n = 6 test samples (*E.coli* BW25113-fimH mutant cells, n = 4 control samples (naïve library), c) *E.coli* BW25113, *E.coli* (n = 12 test samples (*E.coli* BW25113 cells), n = 4 control samples (naïve library)), and d) *E.coli* BW25113 (test, n = 6), *E.coli* BW25113-fimH mutant (control, n = 6). All respective n values are independent biological replicates, and FC + s.d. are reported with asterisks indicating FDR < 0.05. FC was calculated by edgeR DE analysis using the negative binomial model, TMM normalization and BH correction for FDR.

2.3.2 Comparison of glycan binding of the *L. reuteri* strains from different hosts

As shown in the Figure 2.3, the strains of *L. reuteri* coming from the same host lineage did not cluster based on their glycan binding profiles. Despite unique glycan binding profile of each strain, there are noticeable similarities in the glycan binding of these isolates (Figure 2.3, Figure 2.4, Figure 2.5, Figure 2.6, and Figure 2.7). Each isolate bound to at least a mannose structure among three different mannose motifs in the glycan array. In addition, each strain bound to terminal galactose containing glycans. Although 8/16 isolates bound to Di-N3 (Fucα1-2Galβ-Sp), the fold change value for differential enrichment was highest for *L. reuteri* JCM1081. Similarly, 10/16 isolates showed significant binding to Lac-peg4 (Galβ1-4Glcβ-P4). Only *L. reuteri* JCM1081 showed binding to H type 1 (Fucα1-2Galβ1-3GlcNAcβ-Sp) and Di-LN ((Galβ1-4GlcNAcβ1-3)2β-sp). On other hand, only *L. reuteri* limo bound to Tri-AN3 (GalNAcα1-3[Fucα1-2]Galβ-Sp) and only *L. reuteri* AP3 bound to 3' S-Di-Lec (Neu5Acα2-3Galβ1-3GlcNAcβ1-3Galβ1-3GlcNAcβ-Sp) (Figure 2.3). All strains showed binding to the galactosamine and glucosamine containing glycans and also all binds to the sialic acid containing glycans (Figure 2.3).

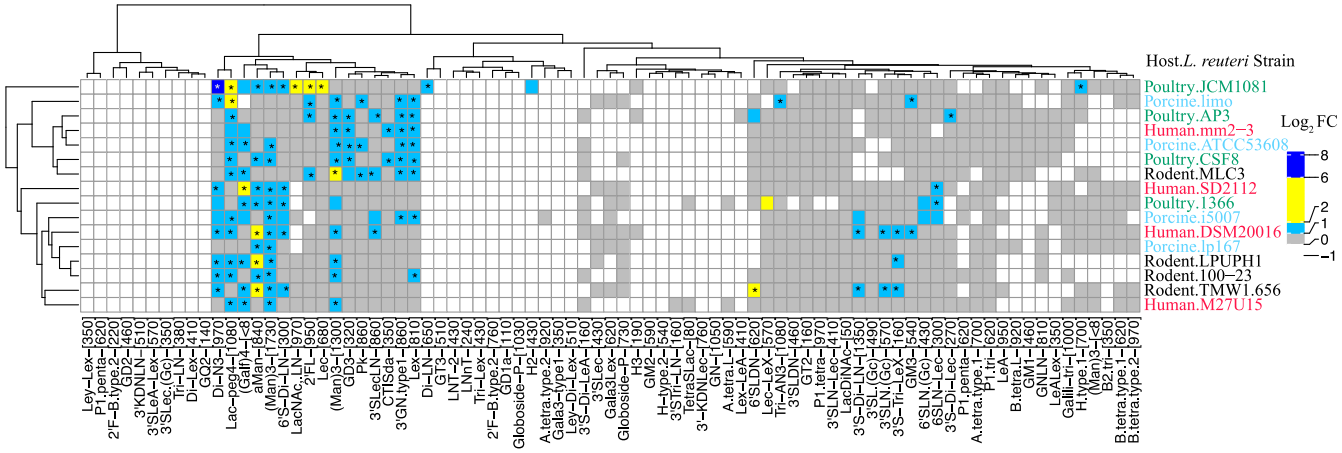

Figure 2.3 Heatmap of glycan binding of 16 *L. reuteri* strains using the hierarchical clustering of k-means of the Euclidean distance. * Represents FC≥ 2, FDR ≤ 0.05, n = 7. Log_2FC was calculated by edgeR DE analysis using the negative binomial model, TMM normalization and BH correction for FDR. Isolates from different hosts are marked in distinct colors. Heatmap was drawn using pheatmap package in R.

2.3.3 Unique binding ability of *L. reuteri* JCM1081 among other *L. reuteri* strains

We tested four different poultry *L. reuteri* strains, *L. reuteri* JCM1081, *L. reuteri* CSF8, *L. reuteri* 1366 and *L. reuteri* AP3. Differential enrichment (DE) analysis (Figure 2.4) identified 7 glycans associating with *L. reuteri* JCM1081 cells. *L. reuteri* JCM1081 showed binding to glycans Di-N3 (Fucα1-2Galβ-Sp) (FC = 88) and Lac-peg4 (Galβ1-4Glcβ-P4) (FC = 46) (Figure 2.3 and Figure 2.4) with greater binding affinity than the other isolates which lead it to be outlier in the hierarchical clustering of all the isolates (Figure 2.3).

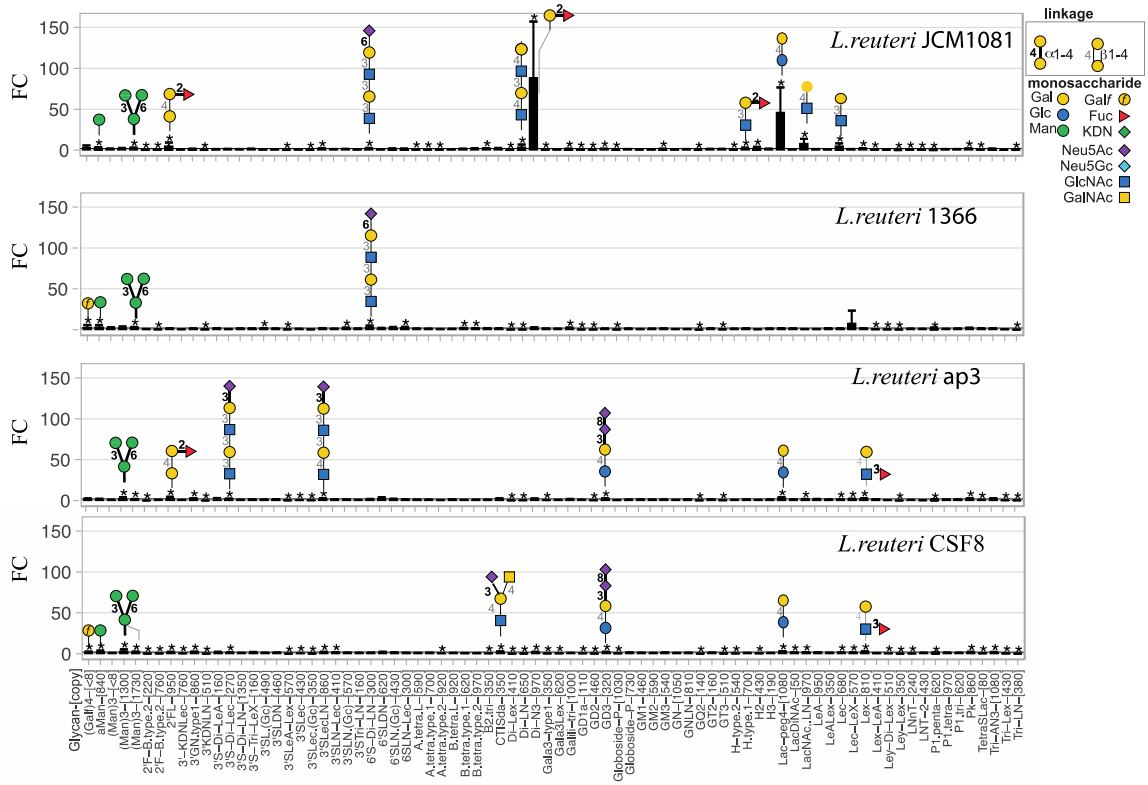

Figure 2.4 Glycan binding of the poultry strains. FC was calculated by Bioconductor edgeR DE analysis using the negative binomial model, TMM normalization and BH correction for FDR. Error bars represent s.d. propagated from the variance of the TMM-normalized sequencing data. Glycan notations and color codes of α and β linkages are shown in the legend. * represents FDR ≤ 0.05, n = 7.

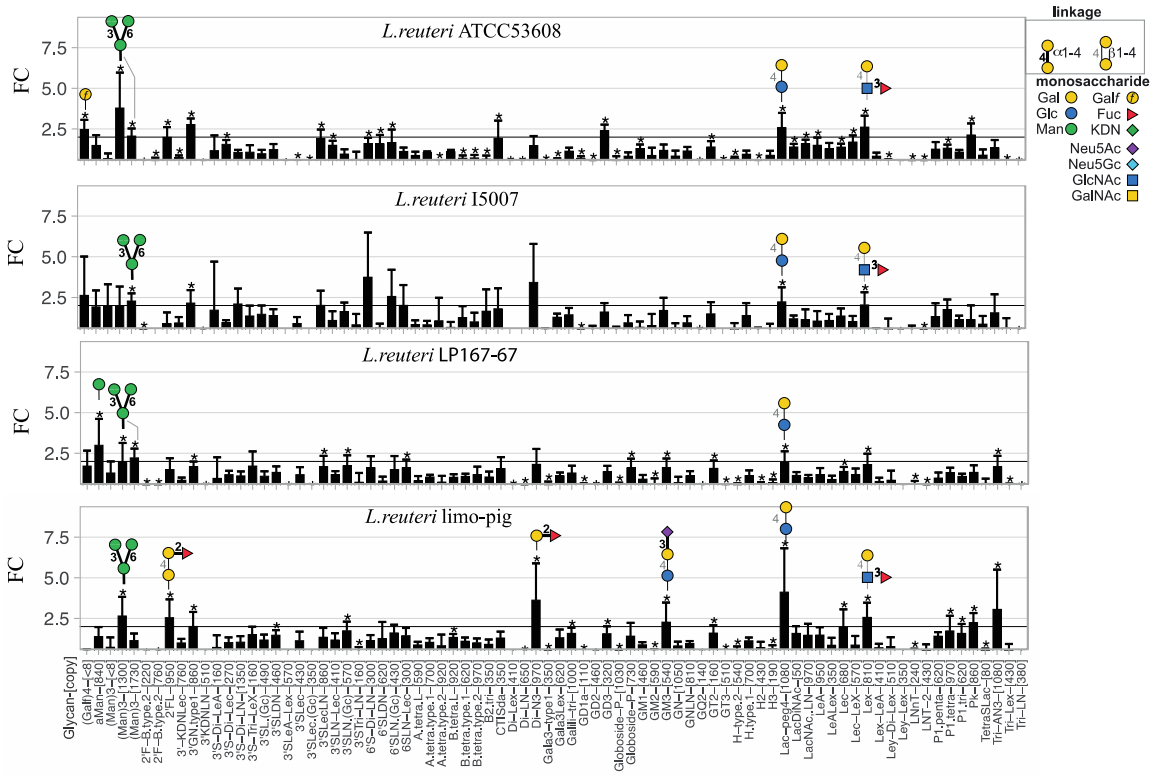

Figure 2.5 Glycan binding of the porcine strains. FC was calculated by Bioconductor edgeR DE analysis using the negative binomial model, TMM normalization and BH correction for FDR. Error bars represent s.d. propagated from the variance of the TMM-normalized sequencing data. Glycan notations and color codes of α and β linkages are shown in the legend. * represents FDR ≤ 0.05, n = 7.

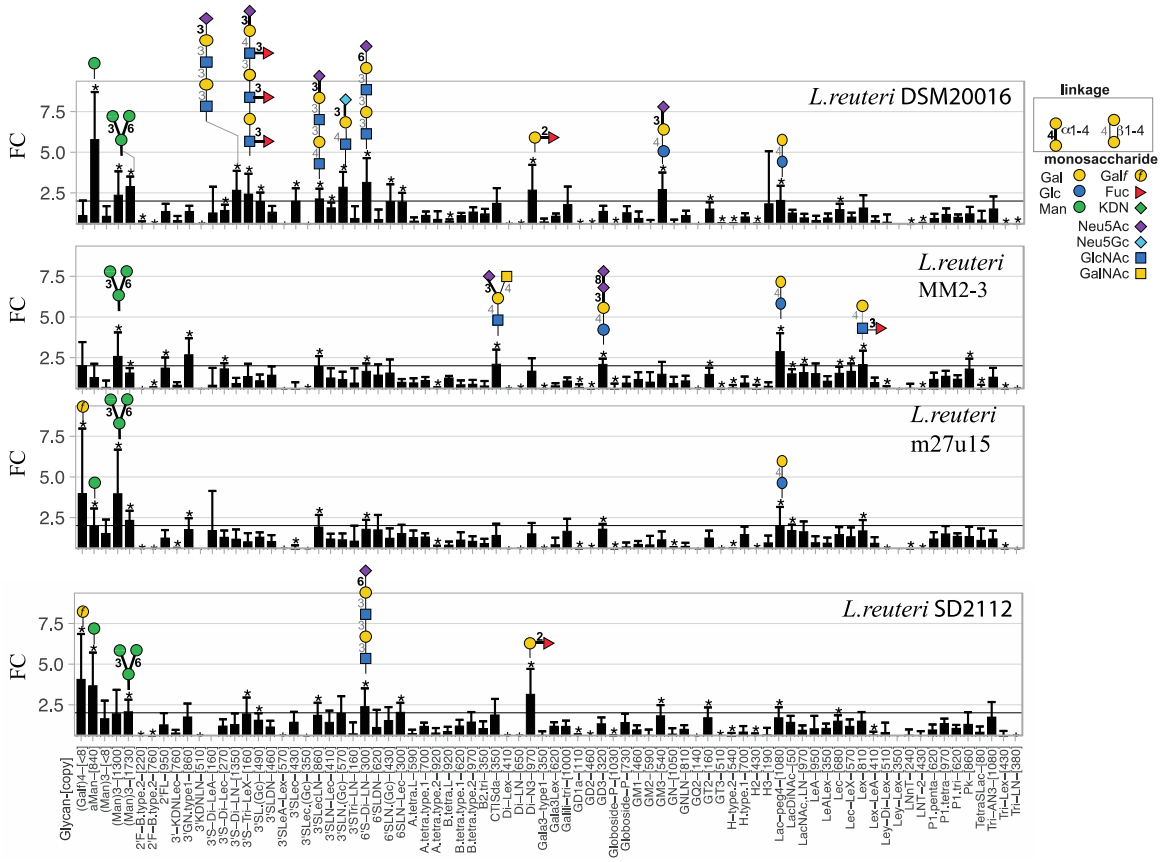

Figure 2.6 Glycan binding of the human strains. FC was calculated by Bioconductor edgeR DE analysis using the negative binomial model, TMM normalization and BH correction for FDR. Error bars represent s.d. propagated from the variance of the TMM-normalized sequencing data. Glycan notations and color codes of α and β linkages are shown in the legend. * represents FDR ≤ 0.05, n = 7.

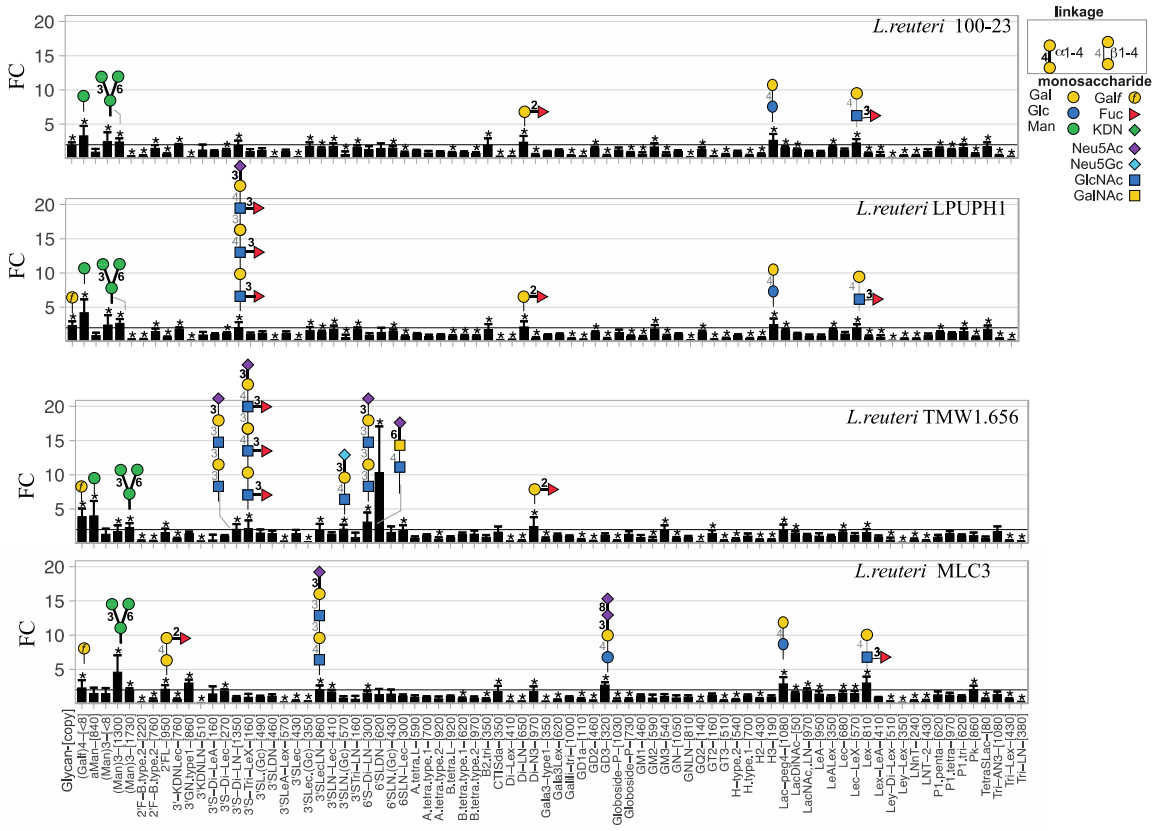

Figure 2.7 Glycan binding of the rodent strains. FC was calculated by Bioconductor edgeR DE analysis using the negative binomial model, TMM normalization and BH correction for FDR. Error bars represent s.d. propagated from the variance of the TMM-normalized sequencing data. Glycan notations and color codes of α and β linkages are shown in the legend. * represents FDR ≤ 0.05, n = 7.

2.4 Discussion

Glycan binding of *L. reuteri* has been previously shown using the TLC overlay assay, inhibition assay by free carbohydrates and evaluating the different surface proteins of these bacteria. However, this is the first study to screen the glycan binding of these bacteria using LiGA. Genetic analysis of the *L. reuteri* from different hosts has shown the clustering of the isolates from similar host and experimental evaluation of the host specificity has been done *in vivo*. Host specificity of many pathogens are based on the glycan binding; therefore, we evaluated if the glycan binding of commensal *L. reuteri* is host specific. These results indicate that glycan binding profiles of *L. reuteri* vary in a strain specific but not host specific manner. Results showed similarity in binding of some glycans such as galactose, mannose, and sialic acid in all 16 isolates.

Mannose is an integral part of the gut epithelial glycosylation and microbes have adapted to bind to mannose to get access to the gut epithelium. Pathogens such as the ETEC express type 1 fimbriae that are involved in the mannose specific adhesion to epithelial cells (Reid & Burton, 2002). Apart from *E.coli*, other pathogens such as *Salmonella enterica* serovar Enteritidis, *Vibrio cholerae,* and *Pseudomonas aeruginosa* (Bhattacharjee & Srivastava, 1979; Imberty et al., 2004) also target the host surface mannose containing glycoconjugates to infect. Other *Limosilactobacillus* strains have shown binding to mannose but very few strains of *L. reuteri* have shown binding to mannose. In previous study, *L. reuteri* 1063 (*L. reuteri* ATCC53608) strain had shown binding to mannose (Roos & Jonsson, 2002). This characteristic potentially aids in the ability of *L. reuteri* to protect their host from these pathogens by competitive exclusion at mannose receptor of the gut epithelial surface. Pig isolate *L. reuteri* lp167-67 bound to mannose structures only and it can be assumed

that only mannose binding adhesin is present on the surface of this bacterium while other isolates bound to several other glycans containing fucose, sialic acid and galactose. These data indicate that strains used in this study possess mannose specific adhesin (Msa) protein as an adhesin that helps in competitive exclusion of pathogens in the intestine. In case of the *campylobacter jejuni,* mannose contributes its initial recognition of host tissue for adherence after environmental challenge and may not be needed for continued colonization (Christopher J. Day et al., 2013). The *Limosilactobacillus* adhesion to epithelial layer is suggested to be mannose mediated and it has been seen to inhibit the interaction of *Salmonella* serovars and ETEC from the epithelial lining, because ETEC and *Salmonella* serovars are also seen to bind the high mannose glycans of the epithelial layer (Garcia-Gonzalez et al., 2018; Yue et al., 2012).

It is suggested that glycan binding preferences are the result of the adaptation to gut environment glycans. Commonly, the mucin O-glycan chains are composed of an α-linked N-acetylgalactosamine residue linked to serine or threonine, and extended by carbohydrate residues such as galactose, N-acetylglucosamine, fucose, or sialic acid, but not mannose, glucose, or xylose moieties (Brockhausen et al., 2009; Van Tassell & Miller, 2011). Mucin from the pig stomach is primarily composed of core 2 (Galβ1-3(GlcNAc β1-6)GalNAc-) and core 1 (Galβ1-3GalNAc) mucins (Karlsson et al., 1997), suggesting the presence of a higher proportion of NAG (N-acetyl-glucosamine/ N-acetyl-galactosamine). Monosaccharide analyses of purified pig mucin (Type III, Sigma) by gas chromatography-mass spectrometry (GC-MS) have shown that it contains N-glycans with 9.1% fucose, 5.4% mannose, 34% galactose, 28.9% GlcNAc, and 22.4% GalNAc (Gunning et al., 2013). Intestinal mucin contains much less mannose, however, mannose is

abundant in the gut epithelial layer and plays a critical role in immune activities in the gut (Park et al., 2017). Gut epithelial enterocyte glycosylation has been evaluated in different diets and Caco-2 cells have been used to study changes in the mannosylation levels in response to diet (Park et al., 2017). Mannose is also abundantly present in dietary components and it has been seen in bacterial surface glycosylation (Górska et al., 2016; Poroyko et al., 2011). It has been shown that intestinal brush borders have less galactose as compared to the mannose (Gebhard & Gebert, 1999). In our study, each strain binding to terminal galactose containing glycans and mannose indicates that these bacteria interact with the mucin glycans and the epithelial glycans.

Previous studies emphasize on the mucus binding proteins (MUBs) which show binding to sialic acid containing glycolipid such as aialo-GM1 (T. Mukai et al., 1998; Takao Mukai et al., 1992, 2002) . Apart from MUBs, there can be other proteinous components such as Cazymes to mediate the glycan binding. The cazyme annotation of these *L. reuteri* strains shows that their genomes contain several glycosylhydrolases (GHs) and some of those GHs have been characterized for the carbohydrate binding modules (CBMs). There are no direct studies evaluating the glycan binding specificities of these GHs/ CAzymes in *L. reuteri,* however, the surface anchored sialidase of the *Ruminococcus gnavus* has been characterized for the sialic acid binding in-vitro (Owen et al., 2017).

These strains' binding to Le^x containing structures and to lipid glycoprotein glycans such as GM3 (Neu5Acα2-3Galβ1-4Glcβ-Sp) and GD3 (Neu5Acα2-8Neu5Acα2-3Galβ1-4Glcβ-Sp) has been previously shown (Takao Mukai et al., 2004). 7/16 (3 rodent, 2 human, 1 poultry and 1 porcine) strains bound to the galactofuranose (Galf), which has not been shown to bind *L. reuteri*

previously. Galf has been studied in context of the pathogenicity by microorganisms (Górska et al., 2016) but in context of gut intestinal tract, it has been minimally explored. Only two strains (rodent *L. reuteri* TMW1.656 and human *L. reuteri* DSM20016) bind to N-glycolylneuraminic acid containing 3'SLN(Gc) (Neu5Gcα2-3Galβ1-4GlcNAcβ-Sp). The glycocalyx of human cells differs from that of many other mammals by the lack of the N-glycolylneuraminic acid (Neu5Gc) and increased abundance of its precursor N-acetylneuraminic acid (Neu5Ac) (Altman & Gagneux, 2019). Uptake and incorporation of dietary Neu5Gc into salivary mucin O-glycans has been observed in humans (Tangvoranuntakul et al., 2003), and the adsorption of Neu5Gc from the diet may be the source of the O-glycans containing N-glycolylneuraminic acid. Previous reports have shown that alterations in chicken diet affect mucin glycosylation in goblet cells and also *C. jejuni* colonization in broiler chicks (Fernandez et al., 2000). Our results suggest that most of the strains do not bind to (Neu5Gc) as it is rarely present in many animal species.

Most (10/16) strains bound to the fucose terminating glycans and all strains bound to fucose containing glycans which suggests the presence of mucus binding protein (MUB), which binds to fucose. The binding MUB of *L. reuteri* 1063 (*L. reuteri* ATCC53608) to mucus was significantly decreased with addition of fucose (MacKenzie et al., 2010) which suggest that MUB has specificity to bind fucose.

Strong glycan binding of poultry *L. reuteri* JCM1081 conforms with observations of enhanced adhesive capacity to gut epithelial cells (Wang et al., 2008). Its binding to galactose containing glycans conforms with previous studies where it had been shown to bind gangliotetraosylceramide (asialo-Neu5Acα2-3[Galβ1-3GalNAcβ1-4]Galβ1-4Glcβ-cer) and sulphated galactoceramide

(Gal-cer) (Takao Mukai et al., 2002). *L. reuteri* JCM1081 inhibits *Helicobacter pylori* binding to glycolipids receptor, including sulfatide, as demonstrated by thin-layer chromatography (Takao Mukai et al., 2002). Later, a surface protein of *L. reuteri* JCM1081 was found which mediated the glycan binding to the sulphatide glycans (Wang et al., 2008). Further, it was identified as EF-tu (47-kDa) which is a cytoplasmic protein that interacts with the various partners during the elongation cycle of protein synthesis (Nishiyama et al., 2013). It had also been shown in the other *Limolactobacilli* (e.g. *L. johnsonii* NCC533, *L. plantarum* strains) to bind the intestinal cells and mucins (Granato et al., 2004; Ramiah et al., 2007). EF-tu in *L. reuteri* JCM1081 bound to sulphated glycolipids, but not to the non-sulphated glycolipids (Nishiyama et al., 2013). In our glycan array analysis, we found that these bacteria bound to non-sulphated glycans, which suggests that there are other glycan binding factors on its surface along with EF-tu.

In a recent study, two pig isolates *L. reuteri* ZJ617 and *L. reuteri* ZJ615 had been shown to differ in their adhesive ability to intestinal mucus in *in vitro* (Deng et al., 2020). Differential abundance of cell wall-associated glyceraldehyde-3-phosphate dehydrogenases (cw-GAPDH) was correlated with adhesive ability of *L. reuteri* ZJ617 (high-adhesive) and *L. reuteri* ZJ615 (low-adhesive). Interaction between the cw-GAPDH of lactobacilli and host mucin might affect bacterial adhesion to the intestine. cw-GAPDH had been shown to GalNAcα1-O-Ser, GalNAcαSer, and Galβ3GalNAc using the molecular docking (Deng et al., 2020). EF-Tu also had been shown to bind sulfated carbohydrates but not sialic acids (Nishiyama et al., 2013). Conversely, our results showed that *L. reuteri* JCM1081 binds to sialic acid containing glycan (6'S-Di-LN: Neu5Acα2-6(Galβ1-4GlcNAcβ1-3)2β-Sp (FC = 3 fold)). Since these characterized surface proteins partially

depict glycan binding of this bacterium, there must be more surface factors which contribute to its glycan binding and still to be categorized.

Overall, despite some similarities in glycan binding of all the strains, there is strain specific glycan binding of *L. reuteri*. Similar strain level glycan binding specificity has been shown in the *Bacteroides* strains (Patnode et al., 2021). Mucus-binding proteins (MUBs) have been revealed as one of the effector molecules involved in mechanisms of the adherence of lactobacilli to the host. MUBs in *L. reuteri* show the strain specific diversity in adhesion to host (MacKenzie et al., 2010). Our study is limited with respect to the glycans available on the LiGA library, to further investigate glycan binding of *L. reuteri* strains, inclusion of more glycans of other hosts (e.g., poultry, pig) to the LiGA is warranted as LiGA mostly has glycans from humans and mice.

2.5 Conclusion

The results indicate that glycan binding profiles of *L. reuteri* vary in a strain specific but not host specific manner. Since these are screening results of the glycan array, further validation will be required through isolation of the surface proteins and functional characterization of these proteins.

3 Chapter 3: Glycan binding profiles of gut microbes from different taxonomic groups

3.1 Introduction

The human gut is home to trillions of microbes that exhibit a phenomenal strain level diversity (Human microbiome project consortium, 2012). There has been significant progress in understanding the genomic basis of this diversity (Karcher et al., 2020; Yaffe & Relman, 2020). However, study of functional basis of this diversity is still in nascent stage (Yang et al., 2020). This diversity might be related to capacity of organisms to physically interact with and utilize various components of the diets consumed by their hosts. The ability of gut microbes to forge competitive or cooperative relationships with one another is influenced by their physical proximity (Patnode et al., 2021).

Glycan mediated adhesion is important for bacteria to increase their access to nutrients in diverse ecosystems such as adhesion of marine microbes to chitin (Sun et al., 2015) and adhesion of cellulolytic species to cellulose (Miron et al., 2001). Moreover, pathogenic microbes adhere to host cells to enable access to nutrients (Poole et al., 2018); in many cases, adhesins bind to specific host glycan structures (Kalas et al., 2018; Le Guennec et al., 2020). Besides pathogenic microbes, commensals such as *Lactobacilli* and *Bacteroides* interact with host glycans using mucus binding proteins (MUBs) and surface glycan binding proteins (SGBPs) respectively.

In our previous experiment (Chapter 2), we observed the strain level variation in the glycan binding of the 16 strains of *L. reuteri*. To further understand the profiles of glycan binding of other bacterial species, we tested glycan binding of taxonomically diverse bacteria from three different phyla Firmicutes/Bacillota, Bacteroidetes/Bacteroidota, Proteobacteria/Pseudomonadota consisting of 9 different species *L. reuteri, E. coli, B. dorei, B. thetaiotamicron, B. fragilis, B. vulgatus, L. mucosae, Citrobacter fruendii, Clostridium ramosum*. We hypothesized that taxonomic closeness leads to similar glycan binding of gut bacteria.

3.2 Materials and methods

3.2.1 Bacterial culture

Bacterial strains used in this study are listed in Table 3.1. *L. reuteri* strains were grown in de Man, Rogosa, and Sharpe (MRS) medium (Difco) supplemented with 5% maltose and 10% fructose under anaerobic conditions (5% CO_2, 5% H_2, and 90% N_2) in anaerobic chamber (model and make). All other bacteria were grown in the media indicated in Table 3.2 in the anaerobic chamber. *E. coli* strains were grown aerobically in Luria-Bertani (LB) broth with agitation.

3.2.2 Colony forming units (CFU) and OD enumeration for the bacterial strains

Each strain was grown and the optical density (Table 3.1) (measured using spectrophotometer [4001/4, Thermo Fisher Scientific]) and CFU/ml were noted at 12 h, 18 h, 24 h and 48 h.

Table 3.1 CFU and OD enumeration.

Bacteria	CFU/ml of broth	OD	Incubation time (growth time)
L. reuteri	1.6 × 10^9	1.11	18 h
E. coli	1.8 × 10^9	1.1	18 h
B. vulgatus	3 × 10^8	1.0	48 h
B. fragilis	6 × 10^8	1.0	48 h
B. dorei	5× 10^8	1.2	48 h
Mice E. coli	1.8 × 10^9	1.1	18 h
E. coli-rat	1.8 × 10^9	1.1	18 h
Clotridium ramosum	1.6× 10^8	1.07	18 h
Citrobacter freundii	5× 10^8	1.09	18 h
AIEC	1.8 × 10^9	1.1	18 h
EHEC	1.8 × 10^9	1.1	18 h
Bacteroides thetaiotamicron	1.8 × 10^9	1.08	24 h
Lactobacillus mucosae	1.8 × 10^9	1.09	18 h

Table 3.2 List of bacterial strains with their growth media and host origin.

Bacteria	Growth media	Host
B. vulgatus	BHI + 0.5 L-cysteine + 5% yeast extract	Human
B. fragilis	BHI + 0.5 L-cysteine + 5% yeast extract	Human
B. dorei	BHI + 0.5 L-cysteine + 5% yeast extract	Human
Mice E. coli	L.B.	Mice
E. coli-rat	L.B.	Rat
Clotridium ramosum	BHI + 0.5 L-cysteine/ FAA	Human
Citrobacter freundii	BHI + 0.5 L-cysteine/ FAA	Human
AIEC	L.B.	Human
EHEC	L.B.	Human
Bacteroides thetaiotamicron	BHI + 0.5 L-cysteine + yeast extract	Pig
Bacteroides thetaiotamicron	BHI+0.5 L-cysteine + yeast extract	Human
Lactobacillus mucosae	MRS + 5% maltose + 10% fructose	Pig

3.2.3 Glycan binding assay for ConA coated 96 well plate

ConA Protocol is described in Chapter 2. Briefly, to validate the glycan binding and integrity of LiGA, mannose binding lectin ConA was tested. ConA solution was deposited in the wells of the 96 well plate and incubated overnight at 4 °C. Later, glycan binding assay was performed at room temperature.

3.2.4 Glycan binding assay for bacterial Cells

Protocol is described in Chapter 2. Briefly, bacterial cells were grown in media mentioned in

Table 3.2. Cells were washed and glycan binding assay was performed at 4 °C. DNA associated with LiGA bound to bacteria cells was extracted using plasmid mini prep kit.

3.2.5 Illumina sequencing

Protocol is described in Chapter 2. Briefly, the extracted DNA from the glycan binding assay was used for first step of PCR. Indexing PCR was performed using the PCR product as PCR template. Indexed PCR products were pooled, and mixture was purified by eGel, quantified by Qubit and sequenced using Illumina NextSeq paired-end 500/550 High Output Kit v2.5 (22x75 Cycles). Data analysis was performed as following.

3.2.6 Processing of Illumina data

Processing of Illumina data was performed as mentioned in Chapter 2.

3.2.7 Data analysis

Protocol is described in Chapter 2. In brief, comparisons and testing differences for significance in LiGA data were done using Differential enrichment (DE) in edgeR (Matochko et al., 2014; Robinson & Smyth, 2008). The significance of glycan binding in a specific experiment was assessed by comparing DE of the levels of the DNA barcode associated with that glycan in 'test'

sets of the DNA read and that levels of same read in 'control' sets. In bacterial cells, the 'test' dataset was LiGA associated with the cells and 'control' was naïve library. Before DE analysis, 'test' and 'control' datasets were retrieved from the server at http://ligacloud. ca/ as tables of glycans, DNA and raw sequencing counts.

3.3 Results

3.3.1 Comparison of glycan binding profiles of diverse gut bacterial isolates

In our previous experiment (Chapter 2), we observed the strain level variations in the glycan binding of the 16 strains of *L. reuteri*. To further investigate pattern/trend in the glycan binding of other bacterial species, we tested glycan binding of taxonomically diverse bacteria from three different phyla Firmicutes/Bacillota, Bacteroidota/Bacteroidetes, Proteobacteria/Pseudomonadota (consisting of 9 different species *L. reuteri, E. coli, Bacteroides dorei*, *B. thetaiotamicron*, *B. fragilis*, *B. vulgatus, L. mucosae, Citrobacter fruendii,* and *Clostridium ramosum*) (Figure 3.2) and from multiples host species (Figure 3.1). The strains were grown until late-log phase, the bacterial cells were washed in HEPES buffer then incubated with the LiGA library-ED and glycan binding assay was performed. This LiGA library contains 81 host glycan structures which represents the majority of host glycan structures present in gut epithelium and mucus layer. As shown in Figure 3.1, hierarchical clustering (using complete-linkage method) according to the glycan binding profiles of all bacteria showed 3 primary clades in the dendrogram. Clade 1 consists of *C. freundii* and *Clostridium ramosum*, clade 2 is divided in 3 sub-clades; first sub-clade has *L. mucosae*, *B.*

fragilis, and *L. reuteri* limo, second sub-clade contains both *B. thetaiotamicron* from pig and human and *B. vulgatus* and third sub-clade has all *E. coli* strains. Third clade contained 15 *L. reuteri* strains and *B. dorei* strain. Clade 1 bacteria *C. freundii* and *C. ramosum* bound to Man3[<8] (Manα1-6[Manα1-3]Manα-S6), Lac-diNAc (GalNAcβ1-4GlcNAc-Sp), and 6SLN (gc) (Neu5Gcα2-6Galβ1-4GlcNAcβ-Sp). These glycans are not bound by other bacteria except the porcine *B. thetaiotamicron* binding to LacdiNAc and *L. mucosae* binding to 6 SLN (Gc). Both *C. freundii* and *C. ramosum* bound to (Galf)4 (FC = 6) and GD3 (FC = 6) (Neu5Acα2-8Neu5Acα2-3Galβ1-4Glcβ-Sp) more strongly than other bacteria. These different glycan binding features of two strains led the clustering of these two in a separate clade. Third sub-clade of the clade 2 has all the *E. coli* strains due to their similar binding to Di-N3 (Fucα1-2Galβ-Sp) (blood group O), mannose (Manα1-6[Manα1-3]Manα-S6, αMan), Lac-peg4 (Galβ1-4Glcβ-P4), Tri-AN3 (GalNAcα1-3[Fucα1-2]Galβ-Sp) (tri-saccharide blood group A), and Le^x (Galβ1-4[Fucα1-3]GlcNAcβ-Sp). However, there are some differences in glycan binding of *E. coli* commensals and *E. coli* pathogenic strains as B tetra type 2 (Galα1-3[Fuca1-2]Galβ1-4GlcNAcβ-Sp) and B tetra type 1 (Galα1-3[Fucα1-2]Galβ1-3GlcNAcβ-Sp), are bound by AIEC and ETEC but not by the commensal strains. Also, GM3 (Neu5Acα2-3Galβ1-4Glcβ-Sp) was bound by commensals but not by the pathogenic strains. B2-tri (Galα1-3Galβ1-4GlcNAcβ-Sp) was only bound by ETEC. Subclade 2 and 3 of clade 2 has four Bacteroidetes and two Firmicutes due to similar glycan binding.

The heatmap (Figure 3.1) illustrates the similarity in glycan binding of bacteria from similar taxonomic groups shown by cluster of pathogenic and commensal *E. coli* strains (third sub-clade

of the clade 2) and 15 strains of *L. reuteri* (Clade 3). However, there are a few exceptions, such as *Citrobacter freundii and C. ramosum* cluster together despite being taxonomically distant.

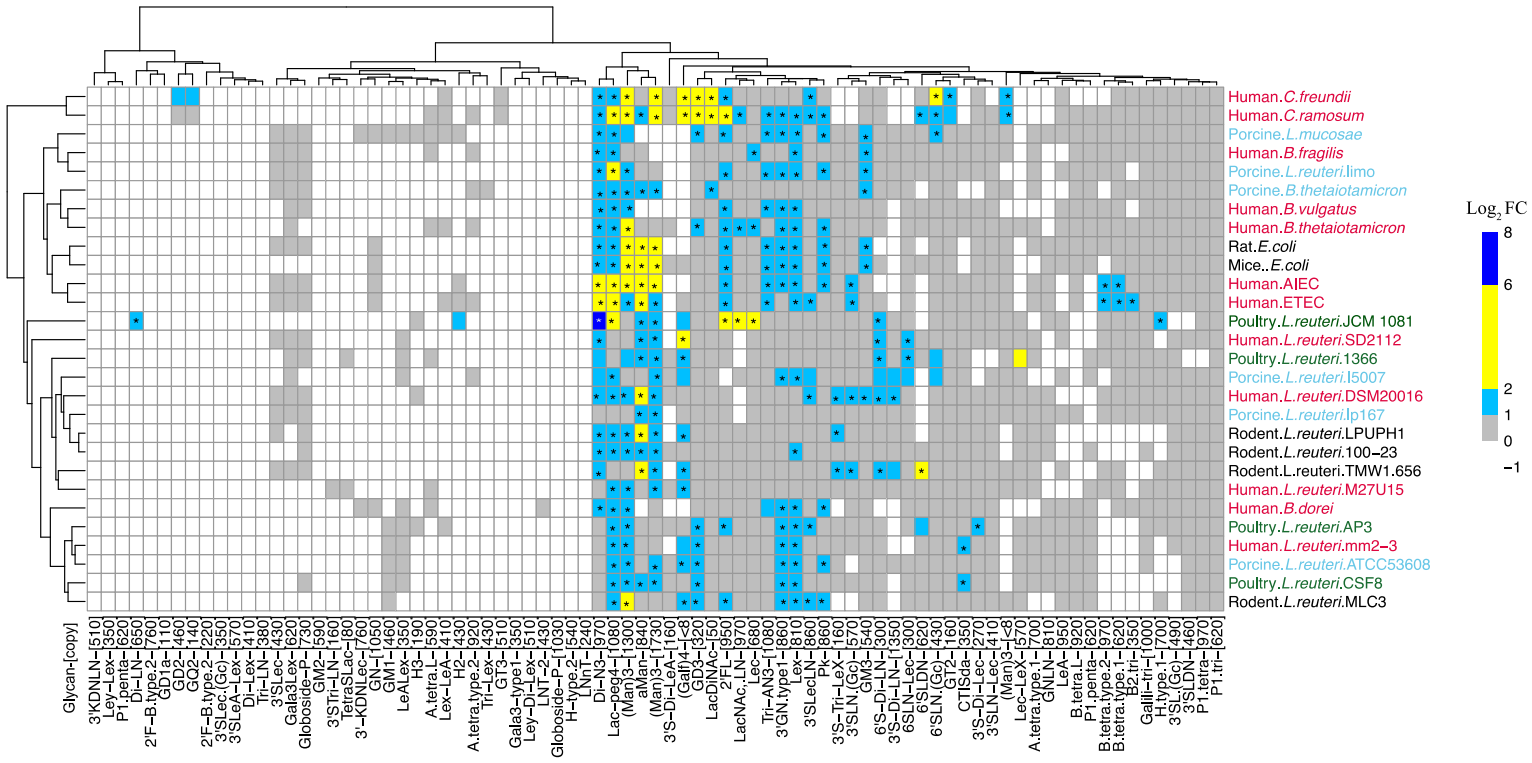

Figure 3.1 Summary of glycan binding of different isolates from three phyla, 5 genera and 9 different species. Bacteria are from phyla, Proteobacteria (Pseudomonadota), Bacteroidetes (Bacteroidota), and Firmicutes (Bacillota). Heatmap is drawn using the hierarchical clustering of k-means of the euclidean distance with complete linkage clustering method. (* represents FC≥ 2, FDR ≤ 0.05, n=7). Log_2FC was calculated by edgeR DE analysis using the negative binomial model, TMM normalization and BH correction for FDR. Heatmap was drawn using pheatmap package in R. Isolates from different hosts are marked in distinct colors.

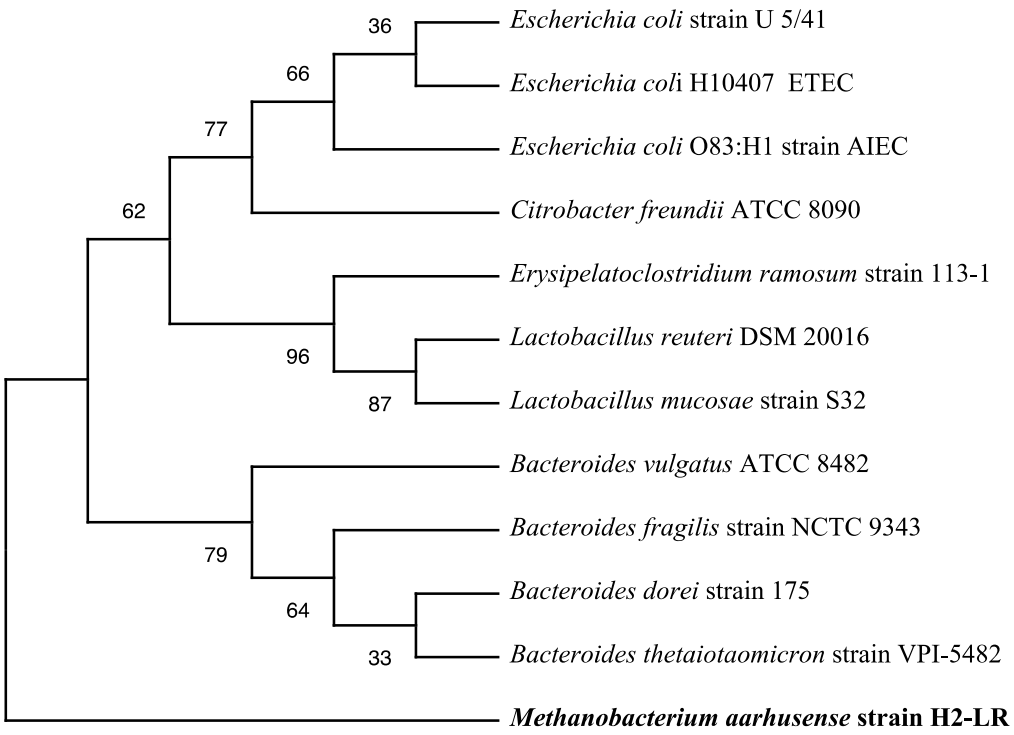

Figure 3.2 Maximum likelihood phylogenetic tree based on 16S rRNA gene sequences of bacterial species used in this study. Sequences were obtained from NCBI. Archean M*ethanobacterium aarhusense* strain H2-LR was used as outgroup. For each node bootstrap values (1,000 replicates) are shown. Tree was generated in MEGA software version 11.

3.3.2 Glycan binding of *Escherichia coli* isolates

Adhesive invasive *E. coli* (AIEC) and Enterotoxigenic *E. coli* (ETEC) both bound to total 13 glycan structures in the array (Figure 3.3, c, d), rat *E. coli* and mice *E. coli* bound to 11 glycan structures in the array (Figure 3.3, a, b). All strains bound to 2'FL, Di-N3, Tri-AN3, Pk, Lex, Lac-peg, mannose (Figure 3.3). All *E. coli* strains bound to mannose structures with good affinity (fold change ranging from 4 to 24).

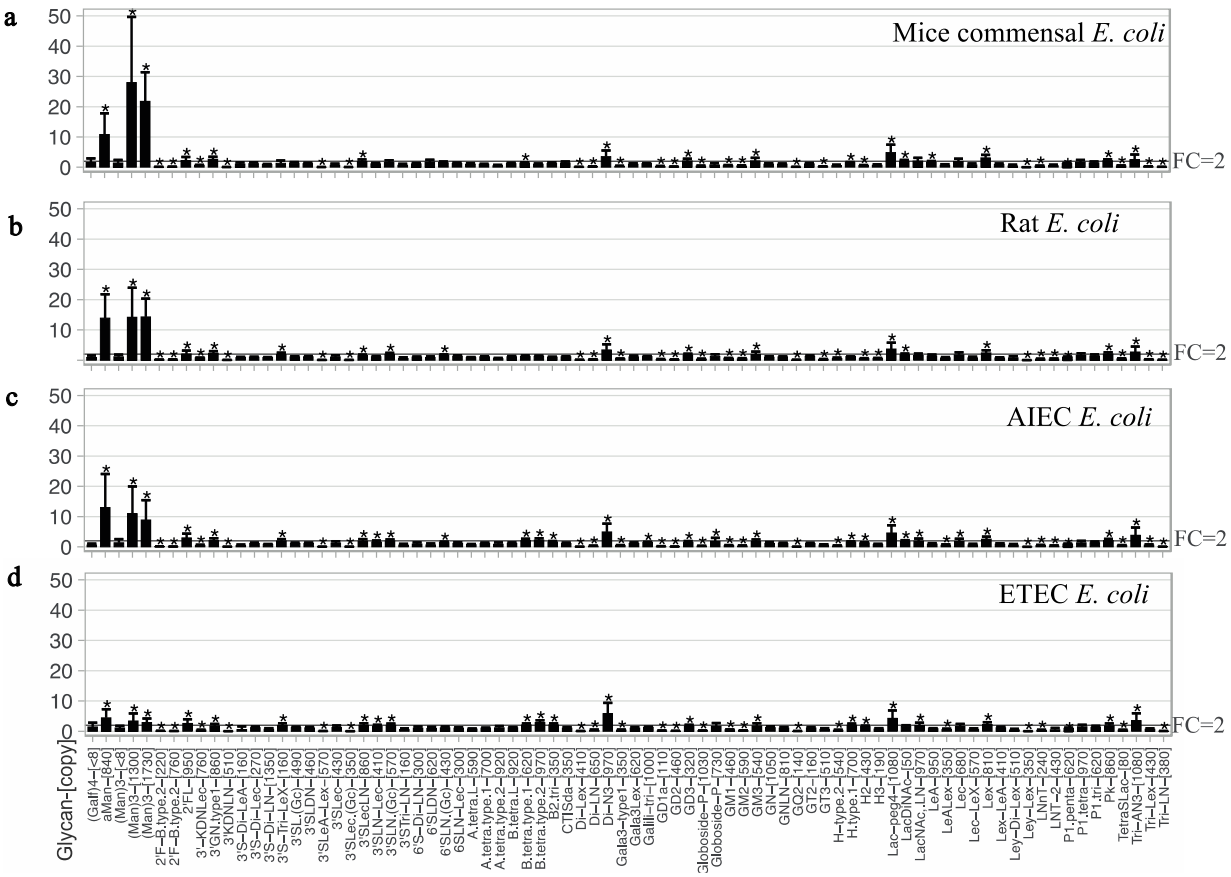

Figure 3.3 Glycan binding of the *E. coli* strains. FC was calculated by Bioconductor edgeR DE analysis using the negative binomial model, TMM normalization and BH correction for FDR. Error bars represent s.d. propagated from the variance of the TMM-normalized sequencing data. * represents FDR ≤ 0.05, n = 7. a) Mice commensal *E. coli,* b) Rat *E. coli* c) Adherent invasive *E. coli* (AIEC) and d) Enterotoxigenic *E. coli* (ETEC).

3.3.3 Glycan binding of *Bacteroides* strains

Differential enrichment analysis of 5 *Bacteroides* isolates is shown in Figure 3.4. Porcine *B. thetaiotamicron* bound to 7 glycan motifs including mannose, Di-N3 and LacdiNAc (Figure 3.4, b), while human *B. thetaiotamicron* strain 10 glycan structures including mannose and Di-N3 (Figure 3.4, a). All *Bacteroides* strains bound to galactose, mannose and fucose containing glycans, while only *B. thetaiotamicron* and *B. fragilis* bound to sialic acid containing glycans GD3 (Neu5Aca2-8Neu5Aca2-3Galb1-4Glcb-Sp) and GM3 (Neu5Aca2-3Galb1-4Glcb-Sp). *B. vulgatus* and *B. dorei* did not show binding to any sialic acid containing glycan.

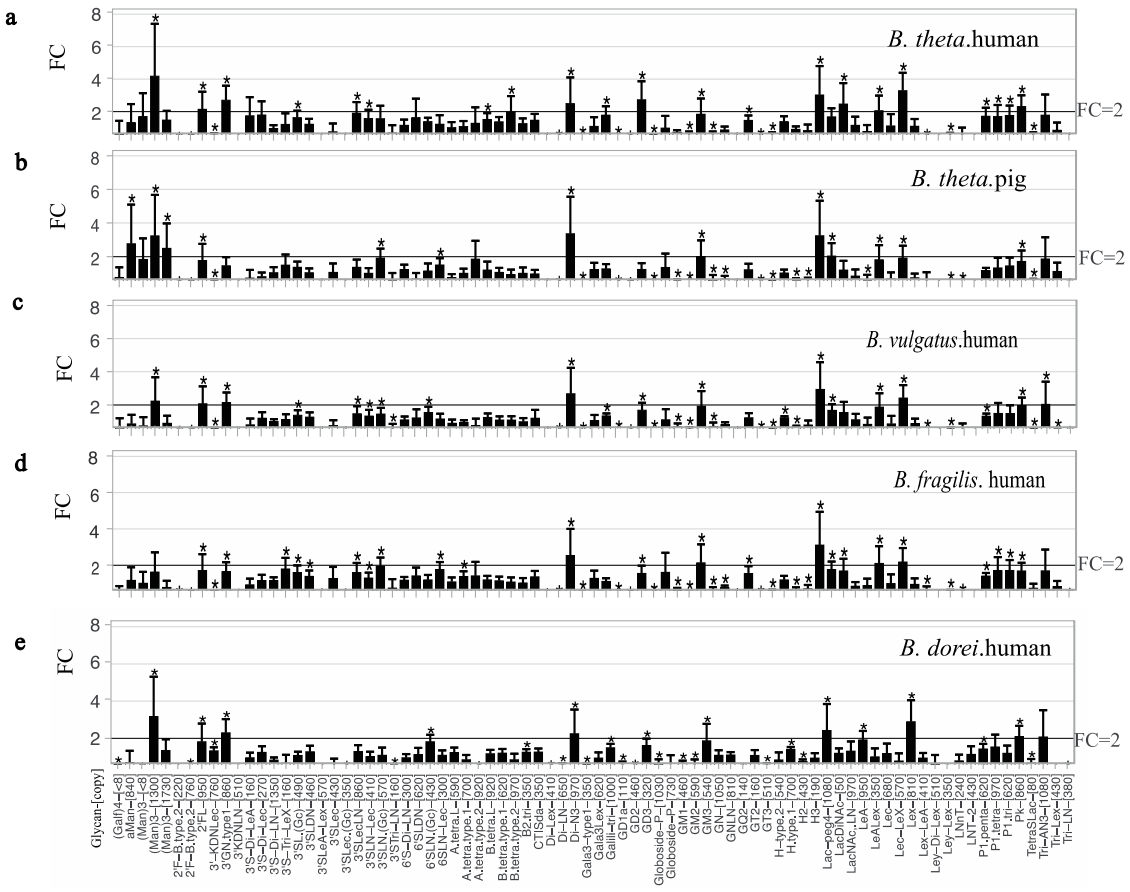

Figure 3.4 Glycan binding of the *Bacteroides* strains. FC was calculated by Bioconductor edgeR DE analysis using the negative binomial model, TMM normalization and BH correction for FDR. Error bars represent s.d. propagated from the variance of the TMM-normalized sequencing data. * represents FDR ≤ 0.05, n = 7. a) *Bacteroides theiotamicron* human strain, b) *Bacteroides theiotamicron* pig strain c) *Bacteroides vulgatus* human strain, d) *Bacteroides fragilis* and e) *Bacteroides dorei* human strain.

3.3.4 Glycan binding of *C. freundii* and *C. ramosum*

Differential enrichment analysis of the *C. freundii* and *C. ramosum* is shown in Figure 3.5. *C. freundii* bound to 12 glycans (Figure 3.5, a) and *C. ramosum* bound to 18 glycans (Figure 3.5, b). *C. freundii* and *C. ramosum* have shown strong binding to LacdiNAc (FC = 20 and FC = 15 respectively).

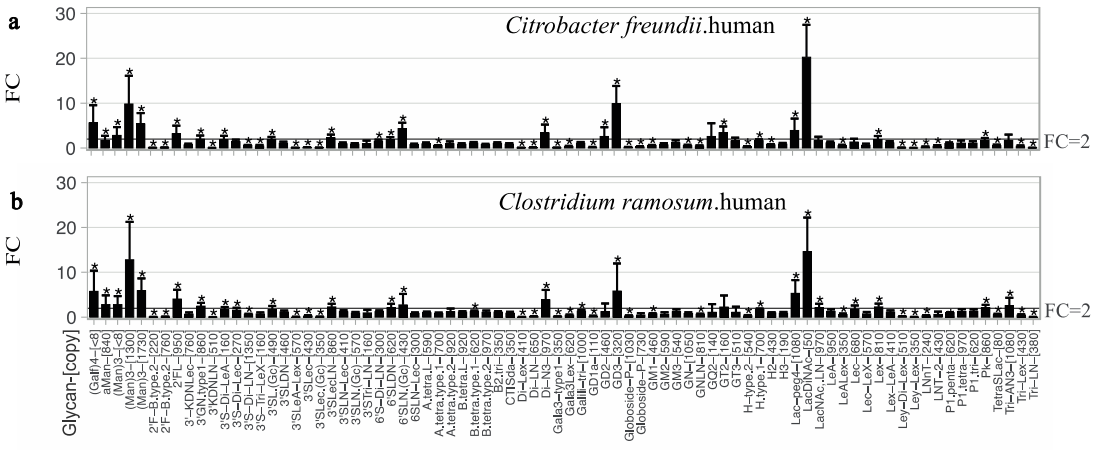

Figure 3.5 Glycan binding profiles of *Citrobacter freudii* and *Clostridium ramosum* (*Erysipelatoclostridium ramosum*). FC was calculated by Bioconductor edgeR DE analysis using the negative binomial model, TMM normalization and BH correction for FDR. Error bars represent s.d. propagated from the variance of the TMM-normalized sequencing data. * represents FDR ≤ 0.05, n = 7. a) *Citrobacter freundii* human strain, b) *Clostridium ramosum* human strain.

3.4 Discussion

The results suggest that there is certain level of similarity in glycan binding of bacteria from similar taxonomic groups highlighted by glycan binding of pathogenic and commensal *E. coli* strains from human and rodents. However, the results also indicate that there are strain level differences in glycan binding of bacteria underlined by *L. reuteri* strains, and *Bacteroides* strains.

The results showed that all four gut *E. coli* strains bind to mannose residues, which supports the presence of fimH protein on the surface of *E.coli* (Bouckaert et al., 2006). In our glycan array analysis, *E. coli* showed binding to a series of glycans including mannose, galactose, fucose and sialic acid containing glycans. Recently, a study showed interaction of different fimbrial proteins of *E. coli* to glycans with terminal mannose, galactose, fucose and sialic acid glycans (Day et al., 2021). Docking analysis showed *E. coli* binding to Lewis A (Galβ1-3[Fucα1-4]GlcNAcβ-Sp) (Mottram et al., 2018); however, our glycan array data did not show glycan binding to Lewis A. Alternatively, binding to Lewis X (Galβ1-4[Fucα1-3]GlcNAcβ-Sp) was observed. It might be result of strain level differences in glycan binding. Binding of AIEC and ETEC to the B tetra type 1 (Galα1-3[Fucα1-2]Galβ1-3GlcNAcβ-Sp) (Blood group B antigen tetraose type 1) and B tetra type 2 (Galα1-3[Fucα1-2]Galβ1-4GlcNAcβ-Sp) has been previously shown in the FedF adhesin of ETEC (Coddens et al., 2009). In our study, ganglioside GM3 bound commensal *E. coli* strains and in a previous study, *E. coli* K99 fimbriae have shown binding to GM3 in intestinal tissues (Esko & Sharon, 2009).

All *Bacteroides* strains bound to galactose, mannose and fucose containing glycans while only *B. thetaiotamicron* and *B. fragilis* bound to sialic acid containing glycans GD3 (Neu5Acα2-8Neu5Acα2-3Galβ1-4Glcβ-Sp) and GM3 (Neu5Acα2-3Galβ1-4Glcβ-Sp). In *B. thetaiotamicron,* SusD-like protein BT1043 (outer membrane protein) has been associated with O-glycan utilization of host mucin (Martens et al., 2008). A SusD like protein NanU, a SusD family protein from *B. fragilis* has also shown high binding affinity to sialic acid (Phansopa et al., 2014). *B. vulgatus* and *B. dorei* did not bind to any sialic acid containing glycan. Possible explanation is that sialic acid is mainly present in the host epithelium and mucus layer glycans and rarely in the dietary glycans, thus these dietary glycan recognizing/metabolizing bacteria have evolved to interact mildly with host glycans containing sialic acid (Coker et al., 2021). Mechanism of interaction of these bacteria with host has not been studied, therefore, further study will be required to show that *B. dorei* and *B. vulgatus* do not bind to sialic acid. *Bacteroides* binding to the host mannose and galactose containing (Lac-peg4, Di-N3 etc.) structures in glycan array is speculated to represent the *Bacteroides* binding to dietary mannan $(Man)_n$ and galactan $(Gal)_n$ (Patnode et al., 2021). It is speculated that the glycan-binding that we observed may also indicate that binding of members of *Bacteroides* is mediated by cell surface lectins; for example lectins incorporated into pilli that extend beyond polysaccharide capsule (Berne et al., 2018; Xu et al., 2016), and surface anchored CaZymes/GHs (with carbohydrate binding modules).

C. freundii and *C. ramosum* have shown strong binding to LacdiNAc (GalNAcβ1-4GlcNAcβ-Sp). LacdiNAc is especially expressed in MUC5AC gastric mucins (Rossez et al., 2014) and LacdiNAc has been associated with the tissue specificity of the *H. pylori* in gastric tissues. Tissue tropism of

H. pylori has been suggested to be the result of strong acidic conditions of the stomach (Wyatt et al., 1990). BabA and SabA adhesins of *H. pylori* have been characterized to bind Le^b and sialylated Le^a or Le^x respectively. However, functional binding of these adhesins cannot explain tissue tropism as these glycan motifs are not restricted to gastric mucosa and are present on other parts of GI tract. Conversely, LabA (LacdiNAc binding adhesin) of *H. pylori* explains its restricted specificity to the gastric mucous surface in GI tract as LacdiNAc co-localizes with MUC5AC of gastric mucosa (Rossez et al., 2014). LacdiNAc has also been associated with diverse micro-environments (other than GI tract) such as hepatic granuloma tissues (Van de Vijver et al., 2006) and other diseased tissues (Haga et al., 2019; Van de Vijver et al., 2006). *C. freundii* and *C. ramosum* are known for systemic infection capability. *Erysipelatoclostridium ramosum* (also known as *C. ramosum*) is associated with disease in humans (Yutin & Galperin, 2013) and *C. freundii* is bacteremic/speticemic, many organs and tissues are affected besides the gut (Gelberg et al., 2017). It can be speculated that systemic infection ability of these bacteria is contributed by their LacdiNAc binding adhesins. Further validation of adhesin is warranted as with exception of binding of *H. pylori* to LacdiNAc, direct binding of this glycan motif with other microbes has not been shown. In our glycan array analysis of porcine *B. thetaiotamicron*, its binding to LacdiNAc is unanticipated as this bacterium is neither seen in gastric mucosa nor colonizes other tissues except GI tract; therefore, further investigation of its LacdiNAc binding is warranted.

3.5 Conclusion

This study showed novel glycan binding profiles of gut microbes and glycan binding profiles of more gut microbes can be further explored. Study of glycan binding profiles of these microbes has many potential applications. It can enable the selection of strains with glycan-mediated adhesion phenotypes. Glycan-binding relationships can be exploited to intentionally foster competition between microbes when both competitors adhere to the same carbohydrate, or two different diets or host derived glycans.

4 Chapter 4: General discussion

4.1 Summary and conclusions

This study presents a novel way in which glycan binding of bacteria can be studied in biochemical assay integrated with the next-generation sequencing. Previous studies have presented several methods to study the glycan binding of pathogens including the extraction of surface protein and characterization of the proteins using *in silico* and *in vitro* methods (Christopher J. Day et al., 2013, 2021; T. Mukai et al., 1998). Our results showed that there is an overlap in glycan binding of pathogenic microbes and commensals; this overlap in glycan binding of pathogens and commensals conform with the mechanism of competitive exclusion by probiotics. Glycan binding of tested *L. reuteri* strains has similarity with the previously studied glycan binding of pathogens like *H. pylori* (Ilver et al., 1998a; Jin et al., 2018). This may support the antiadhesive effects of *L. reuteri* JCM 1081 against *Helicobacter pylori* as it has also been shown to bind to fucosylated and galactosylated glycans. ETEC and AIEC bound to mannose structures and blood group glycans in the array. Previously, the fimbrial proteins of several *E. coli* have shown binding to the mannose and blood group glycans (Coddens et al., 2009; Christopher J. Day et al., 2021; Smeds et al., 2001). Moreover, pathogens *C. freundii* and *C. ramosum* showed a greater number of glycans (12 and 18 respectively) enriched as compared to the other bacteria. Genome screening of several vaginal bacterial species showed that those associated with the infection and inflammation possess a larger repertoire of CBPs (carbohydrate binding proteins) (Bonnardel et al., 2021). To establish the

correlation between the repertoire of glycan binding and pathogenicity, testing of more bacteria will be required.

4.1.1 Understanding the role of glycans in host specificity of *L. reuteri* isolates

L. reuteri has been used to study the evolutionary strategy of vertebrate gut symbionts as it colonizes the GIT of vertebrates as diverse as humans, pigs, mice, rats, and chickens (Walter, 2008). The evolutionary patterns detected indicate a long-term association of *L. reuteri* lineages with particular vertebrate species and host-driven diversification (Duar et al., 2017).

Among genomic features associated with the host origin, large surface proteins of rodent lineage showed specificity towards this lineage and these large surface proteins were rare in isolates from pigs and poultry and absent from lineage of human isolates. Several of these surface proteins are predicted to be involved in epithelial adhesion (Frese et al., 2011). Although these surface proteins are associated with host adaptation and adhesion of these surface proteins to epithelial cells is known to be mediated by glycans, role of glycans in host adaptation of *L. reuteri* remains to be demonstrated. We hypothesized that the glycan binding profiles of *L. reuteri* are host specific. We showed that there is no consistency between glycan binding profiles of *L. reuteri* from the same host and the glycan binding profiles are strain specific. In another study, surface glycan binding proteins such as mucus-binding proteins (MUBs) also showed the strain specific diversity in adhesion to the host (MacKenzie et al., 2010).These glycan-binding relationships can be used to characterize the GBPs of the *L. reuteri*.

4.1.2 Taxonomy and glycan binding profiles of gut microbes

In our previous experiment, we observed the strain level variation in the glycan binding of *L. reuteri* strains. To further understand the pattern of glycan binding of other bacterial species, we tested the glycan binding of taxonomically diverse bacteria from three different phyla Firmicutes/Bacillota, Bacteroidetes/Bacteroidota, Proteobacteria/Pseudomonadota consisting of 9 different species. We hypothesized that the taxonomic closeness leads to similar glycan binding of gut bacteria. Results supported the hypothesis showing taxonomically similar bacteria have similar glycan binding profiles highlighted by glycan binding of pathogenic and commensal *E. coli* strains from human and rodents. However, our results have emphasized that there are strain level differences in glycan binding of bacteria highlighted by *L. reuteri* strains, and *Bacteroides* strains. It suggests that the gut community which contains strains with diverse glycan binding ability would have greater capacity to gain access to host glycans. It also suggests that syntrophic relationships can be sustained between the bacterial strains that have evolved to cohabitate on surface of host-derived structures (Patnode et al., 2021). Study of glycan binding profiles of these microbes has many potential applications. It can enable the selection of strains with specific glycan-mediated adhesion phenotypes. Glycan-binding relationships can be exploited to intentionally foster competition between microbes when both competitors adhere to the same carbohydrate, or two different diets or host derived glycans.

4.2 Limitations

One of the limitations is the involvement of DNA extraction from the bound LiGA; there is some loss of the phage DNA during the DNA extraction using the geneJET plasmid mini prep kit, that leads to loss of some bound glycans which results in lower fold enrichment than actual enrichment of some glycans.

Another limitation is the limited number of glycan structures present in LiGA-ED (LiGA with modification based on ED.xlsx has 81 glycan structures). A LiGA with greater number of glycans including dietary glycans and bacterial surface glycans will be helpful in mapping more comprehensive glycan binding profiles of bacteria. Bacterial surface glycan binding proteins (GBPs) and glycans interact with hundreds of diverse glycan structures (Christopher J. Day et al., 2015; Christopher James Day et al., 2012). Previously, to map the glycan binding of the bacterial lectin, a mammalian glycan library with 609 glycan structure has been used (Petrova et al., 2016).

It has been shown that the dynamic nature of glycan binding profiles is one of the limitations of glycan binding assay (Berne et al., 2018). Initial attachment of bacteria is governed by the extracellular *N*-glycosylation patterns of the host, which can be shaped by dietary and environmental stimuli (Park et al., 2017). It has been shown that the dietary and environmental conditions keep changing in the gut and it can affect the selection of different bacteria for colonization (Gamage et al., 2020). Our glycan binding assay was performed at 4 °C and 7.4 pH as these conditions allow the completion of glycan binding assay without causing truncation of glycans by bacterial glycosylhydrolases (GHs). However, it would be worthwhile to test binding

at physiological temperature. Moreover, it is studied that under acidic conditions, surface GBP EF-Tu from *L. reuteri* JCM1081 exhibited binding activity for acidic oligosaccharides purified from sulfated glycolipids or mucin (Kinoshita et al., 2008; Nishiyama et al., 2013). EF-Tu showed little or no affinity for negatively charged sialic acid, but specifically bound to the mucin sulfate group or sulfated blood-group antigen (Nishiyama et al., 2013). Thus, testing the glycan binding at range of pH conditions will help in understanding effect of pH on glycan binding profiles of different bacteria.

Another limitation is the handling of the cells through washing steps. In glycan binding assay, bacterial cells require very efficient handling of cells without vortexing as it can lead to some disfiguring of bacterial cells leading to abrupt loss of glycan binding.

4.3 Future directions

Our study contributed important information regarding the use of LIGA technology for screening glycan binding of bacteria.

Among all the strains tested *L. reuteri* JCM1081 bound to the blood group glycans with highest affinity. Further study of the surface GBPs expression will shed light on factors leading to the stronger binding of this strain as compared to the other *L. reuteri* strains. Expression of potential surface GBPs such as MapA (mucus adhesion promoting protein), EF-tu (elongation factor-tu), MUBs (mucus binding proteins) and cw-GAPDH (cell wall associated glyceraldehyde-3-phosphate dehydrogenase) (Deng et al., 2020; Kinoshita et al., 2008; Ramiah et al., 2007) should

be carried out to understand if the expression of GBP in *L. reuteri* JCM1081 is greater than other *L. reuteri* strains.

Although our experiments presented the glycan binding profiles of diverse gut microbes, further association of GBPs to glycans will be required with number of methods including prediction of the lectome of these bacteria (Bonnardel et al., 2021).

Another important future direction will be to decipher the bacteria-glycan interactions in the temporal and spatial dynamics of GI tract mucus. Our study mapped glycan binding profiles of gut bacteria *in vitro* using LiGA, further *in vivo*—injection of LiGA in the gut and isolation of subtype of bacteria and associate LiGA—will identify the glycan-specific interactions and homing preferences in complex bacterial communities in the gut.

Bibliography

Altman, M. O., & Gagneux, P. (2019). Absence of Neu5Gc and Presence of Anti-Neu5Gc Antibodies in Humans—An Evolutionary Perspective. *Frontiers in Immunology*, *10*(789). https://doi.org/10.3389/fimmu.2019.00789

Andersson, B., Porras, O., Hanson, L. A., Lagergard, T., & Svanborg-Eden, C. (1986). Inhibition of Attachment of Streptococcus pneumoniae and Haemophilus influenzae by Human Milk and Receptor Oligosaccharides. *Journal of Infectious Diseases*, *153*(2), 232–237. https://doi.org/10.1093/infdis/153.2.232

Arabyan, N., Park, D., Foutouhi, S., Weis, A. M., Huang, B. C., Williams, C. C., Desai, P., Shah, J., Jeannotte, R., Kong, N., Lebrilla, C. B., & Weimer, B. C. (2016). Salmonella Degrades the Host Glycocalyx Leading to Altered Infection and Glycan Remodeling. *Scientific Reports*, *6*(29525). https://doi.org/10.1038/srep29525

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

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berne, C., Ellison, C. K., Ducret, A., & Brun, Y. V. (2018). Bacterial adhesion at the single-cell

level. *Nature Reviews Microbiology*, *16*(10), 616–627. https://doi.org/10.1038/s41579-018-0057-5

Bhattacharjee, J. W., & Srivastava, B. S. (1979). Adherence of wild-type and mutant strains of Vibrio cholerae to normal and immune intestinal tissue. *Bulletin of the World Health Organization*, *57*(1), 123–128. http://www.ncbi.nlm.nih.gov/pubmed/311709

Bonnardel, F., Haslam, S. M., Dell, A., Feizi, T., Liu, Y., Tajadura-Ortega, V., Akune, Y., Sykes, L., Bennett, P. R., MacIntyre, D. A., Lisacek, F., & Imberty, A. (2021). Proteome-wide prediction of bacterial carbohydrate-binding proteins as a tool for understanding commensal and pathogen colonisation of the vaginal microbiome. *Npj Biofilms and Microbiomes*, *7*(1), 49. https://doi.org/10.1038/s41522-021-00220-9

Boraston, A. B., Bolam, D. N., Gilbert, H. J., & Davies, G. J. (2004). Carbohydrate-binding modules: Fine-tuning polysaccharide recognition. *Biochemical Journal*, *382*(3), 769–781. https://doi.org/10.1042/BJ20040892

Bouckaert, J., Mackenzie, J., De Paz, J. L., Chipwaza, B., Choudhury, D., Zavialov, A., Mannerstedt, K., Anderson, J., Piérard, D., Wyns, L., Seeberger, P. H., Oscarson, S., De Greve, H., & Knight, S. D. (2006). The affinity of the FimH fimbrial adhesin is receptor-driven and quasi-independent of Escherichia coli pathotypes. *Molecular Microbiology*, *61*(6), 1556–1568. https://doi.org/10.1111/j.1365-2958.2006.05352.x

Briggs, J. A., Grondin, J. M., & Brumer, H. (2021). Communal living: glycan utilization by the human gut microbiota. *Environmental Microbiology*, *23*(1), 15–35.

https://doi.org/10.1111/1462-2920.15317

Brockhausen, I., Dowler, T., & Paulsen, H. (2009). Site directed processing: role of amino acid sequences and glycosylation of acceptor glycopeptides in the assembly of extended mucin type O-glycan core 2. *Biochimica et Biophysica Acta*, *1790*(10), 1244–1257. https://doi.org/10.1016/j.bbagen.2009.05.020

Brooks, S. P. J., McAllister, M., Sandoz, M., & Kalmokoff, M. L. (2003). Culture-independent phylogenetic analysis of the faecal flora of the rat. *Canadian Journal of Microbiology*, *49*(10), 589–601. https://doi.org/10.1139/w03-075

Bumbaca, D., Littlejohn, J. E., Nayakanti, H., Lucas, A. H., Rigden, D. J., Galperin, M. Y., & Jedrzejas, M. J. (2006). Genome-based identification and characterization of a putative mucin-binding protein from the surface of Streptococcus pneumoniae. *Proteins: Structure, Function, and Bioinformatics*, *66*(3), 547–558. https://doi.org/10.1002/prot.21205

Campanero-Rhodes, M. A., Palma, A. S., Menéndez, M., & Solís, D. (2020). Microarray Strategies for Exploring Bacterial Surface Glycans and Their Interactions With Glycan-Binding Proteins. *Frontiers in Microbiology*, *10*(2909). https://doi.org/10.3389/fmicb.2019.02909

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research*, *37*(Database), D233–D238. https://doi.org/10.1093/nar/gkn663

Çelik, E., Ollis, A. A., Lasanajak, Y., Fisher, A. C., Gür, G., Smith, D. F., & Delisa, M. P. (2015).

Glycoarrays with engineered phages displaying structurally diverse oligosaccharides enable high-throughput detection of glycan-protein interactions. *Biotechnology Journal*, *10*(1), 199–209. https://doi.org/10.1002/biot.201400354

Coddens, A., Diswall, M., Ångström, J., Breimer, M. E., Goddeeris, B., Cox, E., & Teneberg, S. (2009). Recognition of blood group ABH type 1 determinants by the FedF adhesin of F18-fimbriated Escherichia coli. *Journal of Biological Chemistry*, *284*(15), 9713–9726. https://doi.org/10.1074/jbc.M807866200

Coker, J. K., Moyne, O., Rodionov, D. A., & Zengler, K. (2021). Carbohydrates great and small, from dietary fiber to sialic acids: How glycans influence the gut microbiome and affect human health. *Gut Microbes*, *13*(1), 1–18. https://doi.org/10.1080/19490976.2020.1869502

Coolen, M. J. L., Post, E., Davis, C. C., & Forney, L. J. (2005). Characterization of microbial communities found in the human vagina by analysis of terminal restriction fragment length polymorphisms of 16S rRNA genes. *Applied and Environmental Microbiology*, *71*(12), 8729–8737. https://doi.org/10.1128/AEM.71.12.8729-8737.2005

Coppa, G. V., Zampini, L., Galeazzi, T., Facinelli, B., Ferrante, L., Capretti, R., & Orazio, G. (2006). Human Milk Oligosaccharides Inhibit the Adhesion to Caco-2 Cells of Diarrheal Pathogens: Escherichia coli, Vibrio cholerae, and Salmonella fyris. *Pediatric Research*, *59*(3), 377–382. https://doi.org/10.1203/01.pdr.0000200805.45593.17

Cravioto, A., Gross, R. J., Scotland, S. M., & Rowe, B. (1979). An adhesive factor found in strains of Escherichia coli belonging to the traditional infantile enteropathogenic serotypes. *Current*

Microbiology, *3*(2), 95–99. https://doi.org/10.1007/BF02602439

Day, Christopher J., Lo, A. W., Hartley-Tassell, L. E., Argente, M. P., Poole, J., King, N. P., Tiralongo, J., Jennings, M. P., & Schembri, M. A. (2021). Discovery of Bacterial Fimbria–Glycan Interactions Using Whole-Cell Recombinant Escherichia coli Expression. *MBio*, *12*(1), 1–10. https://doi.org/10.1128/mBio.03664-20

Day, Christopher J., Tiralongo, J., Hartnell, R. D., Logue, C. A., Wilson, J. C., von Itzstein, M., & Korolik, V. (2009). Differential carbohydrate recognition by Campylobacter jejuni strain 11168: Influences of temperature and growth conditions. *PLoS ONE*, *4*(3). https://doi.org/10.1371/journal.pone.0004927

Day, Christopher J., Tram, G., Hartley-Tassell, L. E., Tiralongo, J., & Korolik, V. (2013). Assessment of glycan interactions of clinical and avian isolates of Campylobacter jejuni. *BMC Microbiology*, *13*(1). https://doi.org/10.1186/1471-2180-13-228

Day, Christopher J., Tran, E. N., Semchenko, E. A., Tram, G., Hartley-Tassell, L. E., Ng, P. S. K., King, R. M., Ulanovsky, R., McAtamney, S., Apicella, M. A., Tiralongo, J., Morona, R., Korolik, V., & Jennings, M. P. (2015). Glycan:glycan interactions: High affinity biomolecular interactions that can mediate binding of pathogenic bacteria to host cells. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(52), E7266-75. https://doi.org/10.1073/pnas.1421082112

Day, Christopher James, Semchenko, E. A., & Korolik, V. (2012). Glycoconjugates Play a Key Role in Campylobacter jejuni Infection: Interactions between Host and Pathogen. *Frontiers*

in Cellular and Infection Microbiology, *2*(9). https://doi.org/10.3389/fcimb.2012.00009

Deng, Z., Dai, T., Zhang, W., Zhu, J., Luo, X. M., Fu, D., Liu, J., & Wang, H. (2020). Glyceraldehyde-3-phosphate dehydrogenase increases the adhesion of Lactobacillus reuteri to host mucin to enhance probiotic effects. *International Journal of Molecular Sciences*, *21*(24), 1–16. https://doi.org/10.3390/ijms21249756

Dethlefsen, L., McFall-Ngai, M., & Relman, D. A. (2007). An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature*, *449*(7164), 811–818. https://doi.org/10.1038/nature06245

Duar, R. M., Frese, S. A., Lin, X. B., Fernando, S. C., Burkey, T. E., Tasseva, G., Peterson, D. A., Blom, J., Wenzel, C. Q., Szymanski, C. M., & Walter, J. (2017). Experimental Evaluation of Host Adaptation of Lactobacillus reuteri to Different Vertebrate Species. *Applied and Environmental Microbiology*, *83*(12). https://doi.org/10.1128/AEM.00132-17

Ejby, M., Guskov, A., Pichler, M. J., Zanten, G. C., Schoof, E., Saburi, W., Slotboom, D. J., & Abou Hachem, M. (2019). Two binding proteins of the ABC transporter that confers growth of Bifidobacterium animalis subsp. lactis ATCC27673 on β-mannan possess distinct manno-oligosaccharide-binding profiles. *Molecular Microbiology*, *112*(1), 114–130. https://doi.org/10.1111/mmi.14257

Ejby, Morten, Fredslund, F., Andersen, J. M., Vujičić Žagar, A., Henriksen, J. R., Andersen, T. L., Svensson, B., Slotboom, D. J., & Abou Hachem, M. (2016). An ATP Binding Cassette Transporter Mediates the Uptake of α-(1,6)-Linked Dietary Oligosaccharides in

Bifidobacterium and Correlates with Competitive Growth on These Substrates. *The Journal of Biological Chemistry*, *291*(38), 20220–20231. https://doi.org/10.1074/jbc.M116.746529

Ejby, Morten, Fredslund, F., Vujicic-Zagar, A., Svensson, B., Slotboom, D. J., & Abou Hachem, M. (2013). Structural basis for arabinoxylo-oligosaccharide capture by the probiotic Bifidobacterium animalis subsp. lactisBl-04. *Molecular Microbiology*, *90*(5), 1100–1112. https://doi.org/10.1111/mmi.12419

El Mazouari, K., Oswald, E., Hernalsteens, J. P., Lintermans, P., & De Greve, H. (1994). F17-like fimbriae from an invasive Escherichia coli strain producing cytotoxic necrotizing factor type 2 toxin. *Infection and Immunity*, *62*(6), 2633–2638. https://doi.org/10.1128/iai.62.6.2633-2638.1994

Esko, J. D., & Sharon, N. (2009). Microbial Lectins: Hemagglutinins, Adhesins, and Toxins. In *Essentials of Glycobiology* (2nd ed.). http://www.ncbi.nlm.nih.gov/pubmed/20301238

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics*, *164*(4), 1567–1587. https://doi.org/10.1093/genetics/164.4.1567

Fernandez, F., Hinton, M., & Van Gils, B. (2000). Evaluation of the effect of mannan-oligosaccharides on the competitive exclusion of Salmonella Enteritidis colonization in broiler chicks. *Avian Pathology*, *29*(6), 575–581. https://doi.org/10.1080/03079450020016823

Fraune, S., Augustin, R., Anton-Erxleben, F., Wittlieb, J., Gelhaus, C., Klimovich, V. B., Samoilovich, M. P., & Bosch, T. C. G. (2010). In an early branching metazoan, bacterial colonization of the embryo is controlled by maternal antimicrobial peptides. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(42), 18067–18072. https://doi.org/10.1073/pnas.1008573107

Frese, S. A., Benson, A. K., Tannock, G. W., Loach, D. M., Kim, J., Zhang, M., Oh, P. L., Heng, N. C. K., Patil, P. B., Juge, N., MacKenzie, D. A., Pearson, B. M., Lapidus, A., Dalin, E., Tice, H., Goltsman, E., Land, M., Hauser, L., Ivanova, N., … Walter, J. (2011). The evolution of host specialization in the vertebrate gut symbiont Lactobacillus reuteri. *PLoS Genetics*, *7*(2). https://doi.org/10.1371/journal.pgen.1001314

Frese, S. A., MacKenzie, D. A., Peterson, D. A., Schmaltz, R., Fangman, T., Zhou, Y., Zhang, C., Benson, A. K., Cody, L. A., Mulholland, F., Juge, N., & Walter, J. (2013). Molecular Characterization of Host-Specific Biofilm Formation in a Vertebrate Gut Symbiont. *PLoS Genetics*, *9*(12). https://doi.org/10.1371/journal.pgen.1004057

Gamage, H. K. A. H., Chong, R. W. W., Bucio-Noble, D., Kautto, L., Hardikar, A. A., Ball, M. S., Molloy, M. P., Packer, N. H., & Paulsen, I. T. (2020). Changes in dietary fiber intake in mice reveal associations between colonic mucin O-glycosylation and specific gut bacteria. *Gut Microbes*, *12*(1), 1802209. https://doi.org/10.1080/19490976.2020.1802209

Gao, J., Liu, D., & Wang, Z. (2010). Screening Lectin-Binding Specificity of Bacterium by Lectin Microarray with Gold Nanoparticle Probes. *Analytical Chemistry*, *82*(22), 9240–9247.

https://doi.org/10.1021/ac1022309

Garcia-Gonzalez, N., Prete, R., Battista, N., & Corsetti, A. (2018). Adhesion Properties of Food-Associated Lactobacillus plantarum Strains on Human Intestinal Epithelial Cells and Modulation of IL-8 Release. *Frontiers in Microbiology*, *9*(2392), 1–11. https://doi.org/10.3389/fmicb.2018.02392

Garrido, D., Kim, J. H., German, J. B., Raybould, H. E., & Mills, D. A. (2011). Oligosaccharide binding proteins from bifidobacterium longum subsp. infantis reveal a preference for host glycans. *PLoS ONE*, *6*(3). https://doi.org/10.1371/journal.pone.0017315

Gebhard, A., & Gebert, A. (1999). Brush cells of the mouse intestine possess a specialized glycocalyx as revealed by quantitative lectin histochemistry: Further evidence for a sensory function. *Journal of Histochemistry and Cytochemistry*, *47*(6), 799–808. https://doi.org/10.1177/002215549904700609

Gibson, G. R., Hutkins, R. W., & Sanders, M. E. (2017). The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on fermented foods. *Nature Reviews Gastroenterology and Hepatology*, *18*(3), 196–208. https://doi.org/10.1038/s41575-020-00390-5

Górska, S., Sandstrőm, C., Wojas-Turek, J., Rossowska, J., Pajtasz-Piasecka, E., Brzozowska, E., & Gamian, A. (2016). Structural and immunomodulatory differences among lactobacilli exopolysaccharides isolated from intestines of mice with experimentally induced inflammatory bowel disease. *Scientific Reports*, *6*(37613). https://doi.org/10.1038/srep37613

Granato, D., Bergonzelli, G. E., Pridmore, R. D., Marvin, L., Rouvet, M., & Corthésy-Theulaz, I. E. (2004). Cell Surface-Associated Elongation Factor Tu Mediates the Attachment of Lactobacillus johnsonii NCC533 (La1) to Human Intestinal Cells and Mucins. *Infection and Immunity*, *72*(4), 2160–2169. https://doi.org/10.1128/IAI.72.4.2160-2169.2004

Gunning, A. P., Kirby, A. R., Fuell, C., Pin, C., Tailford, L. E., & Juge, N. (2013). Mining the "glycocode" - Exploring the spatial distribution of glycans in gastrointestinal mucin using force spectroscopy. *FASEB Journal*, *27*(6), 2342–2354. https://doi.org/10.1096/fj.12-221416

Guruge, J. L., Falk, P. G., Lorenz, R. G., Dans, M., Wirth, H. P., Blaser, M. J., Berg, D. E., & Gordon, J. I. (1998). Epithelial attachment alters the outcome of Helicobacter pylori infection. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(7), 3925–3930. https://doi.org/10.1073/pnas.95.7.3925

Haga, Y., Uemura, M., Baba, S., Inamura, K., Takeuchi, K., Nonomura, N., & Ueda, K. (2019). Identification of Multisialylated LacdiNAc Structures as Highly Prostate Cancer Specific Glycan Signatures on PSA. *Analytical Chemistry*, *91*(3), 2247–2254. https://doi.org/10.1021/acs.analchem.8b04829

Hahn, E., Wild, P., Schraner, E. M., Bertschinger, H. U., Häner, M., Müller, S. A., & Aebi, U. (2000). Structural analysis of F18 fimbriae expressed by porcine toxigenic Escherichia coli. *Journal of Structural Biology*, *132*(3), 241–250. https://doi.org/10.1006/jsbi.2000.4323

Holmén Larsson, J. M., Thomsson, K. A., Rodríguez-Piñeiro, A. M., Karlsson, H., & Hansson, G. C. (2013). Studies of mucus in mouse stomach, small intestine, and colon. III. Gastrointestinal

Muc5ac and Muc2 mucin O -glycan patterns reveal a regiospecific distribution. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *305*(5), G357–G363. https://doi.org/10.1152/ajpgi.00048.2013

Hooper, L. V., & Gordon, J. I. (2001). Glycans as legislators of host-microbial interactions: spanning the spectrum from symbiosis to pathogenicity. *Glycobiology*, *11*(2), 1R-10R. https://doi.org/10.1093/glycob/11.2.1R

Hooper, L. V., Xu, J., Falk, P. G., Midtvedt, T., & Gordon, J. I. (1999). A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(17), 9833–9838. https://doi.org/10.1073/pnas.96.17.9833

Ilver, D., Arnqvist, A., Ögren, J., Frick, I. M., Kersulyte, D., Incecik, E. T., Berg, D. E., Covacci, A., Engstrand, L., & Borén, T. (1998a). Helicobacter pylori adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science*, *279*(5349), 373–377. https://doi.org/10.1126/science.279.5349.373

Ilver, D., Arnqvist, A., Ögren, J., Frick, I. M., Kersulyte, D., Incecik, E. T., Berg, D. E., Covacci, A., Engstrand, L., & Borén, T. (1998b). Helicobacter pylori adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science*, *279*(5349), 373–377. https://doi.org/10.1126/science.279.5349.373

Imberty, A., Wimmerová, M., Mitchell, E. P., & Gilboa-Garber, N. (2004). Structures of the lectins from Pseudomonas aeruginosa: Insights into the molecular basis for host glycan recognition.

Microbes and Infection, *6*(2), 221–228. https://doi.org/10.1016/j.micinf.2003.10.016

Jensen, H., Roos, S., Jonsson, H., Rud, I., Grimmer, S., van Pijkeren, J.-P., Britton, R. A., & Axelsson, L. (2014). Role of Lactobacillus reuteri cell and mucus-binding protein A (CmbA) in adhesion to intestinal epithelial cells and mucus in vitro. *Microbiology*, *160*(4), 671–681. https://doi.org/10.1099/mic.0.073551-0

Jin, C., Barone, A., Borén, T., & Teneberg, S. (2018). Helicobacter pylori–binding nonacid glycosphingolipids in the human stomach. *Journal of Biological Chemistry*, *293*(44), 17248–17266. https://doi.org/10.1074/jbc.RA118.004854

Johansson, M. E. V., Larsson, J. M. H., & Hansson, G. C. (2011). The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host–microbial interactions. *Proceedings of the National Academy of Sciences*, *108*(supplement_1), 4659–4665. https://doi.org/10.1073/pnas.1006451107

Juge, N. (2012). Microbial adhesins to gastrointestinal mucus. *Trends in Microbiology*, *20*(1), 30–39. https://doi.org/10.1016/j.tim.2011.10.001

Juge, N., Tailford, L., & Owen, C. D. (2016). Sialidases from gut bacteria: a mini-review. *Biochemical Society Transactions*, *44*(1), 166–175. https://doi.org/10.1042/BST20150226

Kalas, V., Hibbing, M. E., Maddirala, A. R., Chugani, R., Pinkner, J. S., Mydock-McGrane, L. K., Conover, M. S., Janetka, J. W., & Hultgren, S. J. (2018). Structure-based discovery of glycomimetic FmlH ligands as inhibitors of bacterial adhesion during urinary tract infection.

Proceedings of the National Academy of Sciences, *115*(12), E2819–E2828. https://doi.org/10.1073/pnas.1720140115

Kandori, H., Hirayama, K., Takeda, M., & Doi, K. (1996). Histochemical, lectin-histochemical and morphometrical characteristics of intestinal goblet cells of germfree and conventional mice. *Experimental Animals*, *45*(2), 155–160. https://doi.org/10.1538/expanim.45.155

Kaoutari, A. El, Armougom, F., Gordon, J. I., Raoult, D., & Henrissat, B. (2013). The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nature Reviews Microbiology*, *11*(7), 497–504. https://doi.org/10.1038/nrmicro3050

Kaper, J. B., Nataro, J. P., & Mobley, H. L. T. (2004). Pathogenic Escherichia coli. *Nature Reviews Microbiology*, *2*(2), 123–140. https://doi.org/10.1038/nrmicro818

Kapitany, R. A., Forsyth, G. W., Scoot, A., Mckenzie, S. F., & Worthington, R. W. (1979). Isolation and partial characterization of two different heat-stable enterotoxins produced by bovine and porcine strains of enterotoxigenic Escherichia coli. *Infection and Immunity*, *26*(1), 173–177. https://doi.org/10.1128/iai.26.1.173-177.1979

Karcher, N., Pasolli, E., Asnicar, F., Huang, K. D., Tett, A., Manara, S., Armanini, F., Bain, D., Duncan, S. H., Louis, P., Zolfo, M., Manghi, P., Valles-Colomer, M., Raffaetà, R., Rota-Stabelli, O., Collado, M. C., Zeller, G., Falush, D., Maixner, F., … Segata, N. (2020). Analysis of 1321 Eubacterium rectale genomes from metagenomes uncovers complex phylogeographic population structure and subspecies functional adaptations. *Genome Biology*, *21*(1), 138. https://doi.org/10.1186/s13059-020-02042-y

Karlsson, N. G., Nordman, H., Karlsson, H., Carlstedt, I., & Hansson, G. C. (1997). Glycosylation differences between pig gastric mucin populations: A comparative study of the neutral oligosaccharides using mass spectrometry. *Biochemical Journal*, *326*(3), 911–917. https://doi.org/10.1042/bj3260911

Karunatilaka, K. S., Cameron, E. A., Martens, E. C., Koropatkin, N. M., & Biteen, J. S. (2014). Superresolution Imaging Captures Carbohydrate Utilization Dynamics in Human Gut Symbionts. *MBio*, *5*(6), 1–10. https://doi.org/10.1128/mBio.02172-14

Kinoshita, H., Wakahara, N., Watanabe, M., Kawasaki, T., Matsuo, H., Kawai, Y., Kitazawa, H., Ohnuma, S., Miura, K., Horii, A., & Saito, T. (2008). Cell surface glyceraldehyde-3-phosphate dehydrogenase (GAPDH) of Lactobacillus plantarum LA 318 recognizes human A and B blood group antigens. *Research in Microbiology*, *159*(9–10), 685–691. https://doi.org/10.1016/j.resmic.2008.07.005

Kittana, H., Quintero-Villegas, M. I., Bindels, L. B., Gomes-Neto, J. C., Schmaltz, R. J., Segura Munoz, R. R., Cody, L. A., Moxley, R. A., Hostetter, J., Hutkins, R. W., & Ramer-Tait, A. E. (2018). Galactooligosaccharide supplementation provides protection against citrobacter rodentium-induced colitis without limiting pathogen burden. *Microbiology (United Kingdom)*, *164*(2), 154–162. https://doi.org/10.1099/mic.0.000593

Le Guennec, L., Virion, Z., Bouzinba-Ségard, H., Robbe-Masselot, C., Léonard, R., Nassif, X., Bourdoulous, S., & Coureuil, M. (2020). Receptor recognition by meningococcal type IV pili relies on a specific complex N-glycan. *Proceedings of the National Academy of Sciences of*

the United States of America, *117*(5), 2606–2612. https://doi.org/10.1073/pnas.1919567117

Ledder, R. G., Timperley, A. S., Friswell, M. K., Macfarlane, S., & McBain, A. J. (2008). Coaggregation between and among human intestinal and oral bacteria. *FEMS Microbiology Ecology*, *66*(3), 630–636. https://doi.org/10.1111/j.1574-6941.2008.00525.x

Leser, T. D., Amenuvor, J. Z., Jensen, T. K., Lindecrona, R. H., Boye, M., & Moøller, K. (2002). Culture-independent analysis of gut bacteria: The pig gastrointestinal tract microbiota revisited. *Applied and Environmental Microbiology*, *68*(2), 673–690. https://doi.org/10.1128/AEM.68.2.673-690.2002

Lewis, A. L., & Lewis, W. G. (2012). Host sialoglycans and bacterial sialidases: A mucosal perspective. *Cellular Microbiology*, *14*(8), 1174–1182. https://doi.org/10.1111/j.1462-5822.2012.01807.x

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., Schlegel, M. L., Tucker, T. A., Schrenzel, M. D., Knight, R., & Gordon, J. I. (2008). Evolution of Mammals and Their Gut Microbes. *Science*, *320*(5883), 1647–1651. https://doi.org/10.1126/science.1155725

Liévin-Le Moal, V., & Servin, A. L. (2006). The front line of enteric host defense against unwelcome intrusion of harmful microorganisms: Mucins, antimicrobial peptides, and Microbiota. *Clinical Microbiology Reviews*, *19*(2), 315–337. https://doi.org/10.1128/CMR.19.2.315-337.2006

Lintermans, P., Pohl, P., Deboeck, F., Bertels, A., Schlicker, C., Vandekerckhove, J., Van Damme, J., Van Montagu, M., & De Greve, H. (1988). Isolation and nucleotide sequence of the F17-A gene encoding the structural protein of the F17 fimbriae in bovine enterotoxigenic Escherichia coli. *Infection and Immunity*, *56*(6), 1475–1484. https://doi.org/10.1128/iai.56.6.1475-1484.1988

Linz, B., Balloux, F., Moodley, Y., Manica, A., Liu, H., Roumagnac, P., Falush, D., Stamer, C., Prugnolle, F., Van Der Merwe, S. W., Yamaoka, Y., Graham, D. Y., Perez-Trallero, E., Wadstrom, T., Suerbaum, S., & Achtman, M. (2007). An African origin for the intimate association between humans and Helicobacter pylori. *Nature*, *445*(7130), 915–918. https://doi.org/10.1038/nature05562

Liu, Y., Li, N., Zhao, X., Yue, B., He, S., Gao, Z., Zhou, S., & Zhang, M. (2016). A C-type lectin that inhibits bacterial infection and facilitates viral invasion in black rockfish, Sebastes schlegelii. *Fish & Shellfish Immunology*, *57*, 309–317. https://doi.org/10.1016/j.fsi.2016.08.053

Lonardi, E., Moonens, K., Buts, L., de Boer, A., Olsson, J., Weiss, M., Fabre, E., Guérardel, Y., Deelder, A., Oscarson, S., Wuhrer, M., & Bouckaert, J. (2013). Structural Sampling of Glycan Interaction Profiles Reveals Mucosal Receptors for Fimbrial Adhesins of Enterotoxigenic Escherichia coli. *Biology*, *2*(3), 894–917. https://doi.org/10.3390/biology2030894

MacKenzie, D. A., Jeffers, F., Parker, M. L., Vibert-Vallet, A., Bongaerts, R. J., Roos, S., Walter, J., & Juge, N. (2010). Strain-specific diversity of mucus-binding proteins in the adhesion and

aggregation properties of Lactobacillus reuteri. *Microbiology*, *156*(11), 3368–3378. https://doi.org/10.1099/mic.0.043265-0

Magalhães, A., & Reis, C. A. (2010). Helicobacter pylori adhesion to gastric epithelial cells is mediated by glycan receptors. *Brazilian Journal of Medical and Biological Research*, *43*(7), 611–618. https://doi.org/10.1590/S0100-879X2010007500049

Martens, E. C., Chiang, H. C., & Gordon, J. I. (2008). Mucosal Glycan Foraging Enhances Fitness and Transmission of a Saccharolytic Human Gut Bacterial Symbiont. *Cell Host and Microbe*, *4*(5), 447–457. https://doi.org/10.1016/j.chom.2008.09.007

Martens, E. C., Koropatkin, N. M., Smith, T. J., & Gordon, J. I. (2009). Complex glycan catabolism by the human gut microbiota: The bacteroidetes sus-like paradigm. *Journal of Biological Chemistry*, *284*(37), 24673–24677. https://doi.org/10.1074/jbc.R109.022848

Matochko, W. L., Cory Li, S., Tang, S. K. Y., & Derda, R. (2014). Prospective identification of parasitic sequences in phage display screens. *Nucleic Acids Research*, *42*(3), 1784–1798. https://doi.org/10.1093/nar/gkt1104

McAuley, J. L., Linden, S. K., Chin, W. P., King, R. M., Pennington, H. L., Gendler, S. J., Florin, T. H., Hill, G. R., Korolik, V., & McGuckin, M. A. (2007). MUC1 cell surface mucin is a critical element of the mucosal barrier to infection. *Journal of Clinical Investigation*, *117*(8), 2313–2324. https://doi.org/10.1172/JCI26705

Miron, J., Ben-Ghedalia, D., & Morrison, M. (2001). Invited review: Adhesion mechanisms of

rumen cellulolytic bacteria. *Journal of Dairy Science*, *84*(6), 1294–1309. https://doi.org/10.3168/jds.S0022-0302(01)70159-2

Moeller, A. H., Caro-Quintero, A., Mjungu, D., Georgiev, A. V., Lonsdorf, E. V., Muller, M. N., Pusey, A. E., Peeters, M., Hahn, B. H., & Ochman, H. (2016). Cospeciation of gut microbiota with hominids. *Science*, *353*(6297), 380–382. https://doi.org/10.1126/science.aaf3951

Moran, A. P., Gupta, A., & Joshi, L. (2011). Sweet-talk: Role of host glycosylation in bacterial pathogenesis of the gastrointestinal tract. *Gut*, *60*(10), 1412–1425. https://doi.org/10.1136/gut.2010.212704

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

Mottram, L., Liu, J., Chavan, S., Tobias, J., Svennerholm, A. M., & Holgersson, J. (2018). Glyco-engineered cell line and computational docking studies reveals enterotoxigenic Escherichia coli CFA/I fimbriae bind to Lewis a glycans. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-29258-0

Mouricout, M., Milhavet, M., Durié, C., & Grange, P. (1995). Characterization of glycoprotein glycan receptors for Escherichia coli F17 fimbrial lectin. *Microbial Pathogenesis*, *18*(4), 297–306. https://doi.org/10.1016/s0882-4010(05)80006-3

Moustafa, I., Connaris, H., Taylor, M., Zaitsev, V., Wilson, J. C., Kiefel, M. J., Von Itzstein, M.,

& Taylor, G. (2004). Sialic acid recognition by Vibrio cholerae neuraminidase. *Journal of Biological Chemistry*, *279*(39), 40819–40826. https://doi.org/10.1074/jbc.M404965200

Mukai, T., Kaneko, S., & Ohori, H. (1998). Haemagglutination and glycolipid-binding activities of Lactobacillus reuteri. *Letters in Applied Microbiology*, *27*(3), 130–134. https://doi.org/10.1046/j.1472-765X.1998.00418.x

Mukai, Takao, Arihara, K., & Itoh, H. (1992). Lectin-like activity of Lactobacillus acidophilus strain JCM 1026. *FEMS Microbiology Letters*, *98*(1–3), 71–74. https://doi.org/10.1111/j.1574-6968.1992.tb05492.x

Mukai, Takao, Asasaka, T., Sato, E., Mori, K., Matsumoto, M., & Ohori, H. (2002). Inhibition of binding of Helicobacter pylori to the glycolipid receptors by probiotic Lactobacillus reuteri. *FEMS Immunology and Medical Microbiology*, *32*(2), 105–110. https://doi.org/10.1111/j.1574-695X.2002.tb00541.x

Mukai, Takao, Kaneko, S., Matsumoto, M., & Ohori, H. (2004). Binding of Bifidobacterium bifidum and Lactobacillus reuteri to the carbohydrate moieties of intestinal glycolipids recognized by peanut agglutinin. *International Journal of Food Microbiology*, *90*(3), 357–362. https://doi.org/10.1016/S0168-1605(03)00317-9

Nishiyama, K., Ochiai, A., Tsubokawa, D., Ishihara, K., Yamamoto, Y., & Mukai, T. (2013). Identification and characterization of sulfated carbohydrate-binding protein from Lactobacillus reuteri. *PLoS ONE*, *8*(12), 1–10. https://doi.org/10.1371/journal.pone.0083703

Odamaki, T., Kato, K., Sugahara, H., Hashikura, N., Takahashi, S., Xiao, J. Z., Abe, F., & Osawa, R. (2016). Age-related changes in gut microbiota composition from newborn to centenarian: A cross-sectional study. *BMC Microbiology*, *16*(1), 1–12. https://doi.org/10.1186/s12866-016-0708-5

Ofek, I., & Doyle, R. J. (1994). Principles of Bacterial Adhesion. In *Bacterial Adhesion to Cells and Tissues* (pp. 1–15). Springer US. https://doi.org/10.1007/978-1-4684-6435-1_1

Ofek, I., Hasty, D. L., Abraham, S. N., & Sharon, N. (2005). Role of Bacterial Lectins in Urinary Tract Infections. *Genes and Proteins Underlying Microbial Urinary Tract Virulence*, 183–192. https://doi.org/10.1007/0-306-46840-9_25

Oh, P. L., Benson, A. K., Peterson, D. A., Patil, P. B., Moriyama, E. N., Roos, S., & Walter, J. (2010). Diversification of the gut symbiont lactobacillus reuteri as a result of host-driven evolution. *ISME Journal*, *4*(3), 377–387. https://doi.org/10.1038/ismej.2009.123

Ouwerkerk, J. P., De Vos, W. M., & Belzer, C. (2013). Glycobiome: Bacteria and mucus at the epithelial interface. *Best Practice and Research: Clinical Gastroenterology*, *27*(1), 25–38. https://doi.org/10.1016/j.bpg.2013.03.001

Owen, C. D., Tailford, L. E., Monaco, S., Šuligoj, T., Vaux, L., Lallement, R., Khedri, Z., Yu, H., Lecointe, K., Walshaw, J., Tribolo, S., Horrex, M., Bell, A., Chen, X., Taylor, G. L., Varki, A., Angulo, J., & Juge, N. (2017). Unravelling the specificity and mechanism of sialic acid recognition by the gut symbiont Ruminococcus gnavus. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-02109-8

Park, D., Xu, G., Barboza, M., Shah, I. M., Wong, M., Raybould, H., Mills, D. A., & Lebrilla, C. B. (2017). Enterocyte glycosylation is responsive to changes in extracellular conditions: implications for membrane functions. *Glycobiology*, *27*(9), 847–860. https://doi.org/10.1093/glycob/cwx041

Patnode, M. L., Guruge, J. L., Castillo, J. J., Couture, G. A., Lombard, V., Terrapon, N., Henrissat, B., Lebrilla, C. B., & Gordon, J. I. (2021). Strain-level functional variation in the human gut microbiota based on bacterial binding to artificial food particles. *Cell Host and Microbe*, *29*(4), 664-673.e5. https://doi.org/10.1016/j.chom.2021.01.007

Petrova, M. I., Imholz, N. C. E., Verhoeven, T. L. A., Balzarini, J., Van Damme, E. J. M., Schols, D., Vanderleyden, J., & Lebeer, S. (2016). Lectin-Like Molecules of Lactobacillus rhamnosus GG Inhibit Pathogenic Escherichia coli and Salmonella Biofilm Formation. *PLoS ONE*, *11*(8). https://doi.org/10.1371/journal.pone.0161337

Phansopa, C., Roy, S., Rafferty, J. B., Douglas, C. W. I., Pandhal, J., Wright, P. C., Kelly, D. J., & Stafford, G. P. (2014). Structural and functional characterization of NanU, a novel high-affinity sialic acid-inducible binding protein of oral and gut-dwelling Bacteroidetes species. *Biochemical Journal*, *458*(3), 499–511. https://doi.org/10.1042/BJ20131415

Poole, J., Day, C. J., Von Itzstein, M., Paton, J. C., & Jennings, M. P. (2018). Glycointeractions in bacterial pathogenesis. *Nature Reviews Microbiology*, *16*(7), 440–452. https://doi.org/10.1038/s41579-018-0007-2

Poroyko, V., Morowitz, M., Bell, T., Ulanov, A., Wang, M., Donovan, S., Bao, N., Gu, S., Hong,

L., Alverdy, J. C., Bergelson, J., & Liu, D. C. (2011). Diet creates metabolic niches in the "immature gut" that shape microbial communities. *Nutricion Hospitalaria*, *26*(6), 1283–1295. https://doi.org/10.1590/S0212-16112011000600015

Pretzer, G., Snel, J., Molenaar, D., Wiersma, A., Bron, P. A., Lambert, J., De Vos, W. M., Van Der Meer, R., Smits, M. A., & Kleerebezem, M. (2005). Biodiversity-based identification and functional characterization of the mannose-specific adhesin of Lactobacillus plantarum. *Journal of Bacteriology*, *187*(17), 6128–6136. https://doi.org/10.1128/JB.187.17.6128-6136.2005

Quintero-Villegas, M. I., Aam, B. B., Rupnow, J., Sørlie, M., Eijsink, V. G. H., & Hutkins, R. W. (2013). Adherence inhibition of enteropathogenic escherichia coli by chitooligosaccharides with specific degrees of acetylation and polymerization. *Journal of Agricultural and Food Chemistry*, *61*(11), 2748–2754. https://doi.org/10.1021/jf400103g

Ramiah, K., van Reenen, C. A., & Dicks, L. M. T. (2007). Expression of the mucus adhesion genes Mub and MapA, adhesion-like factor EF-Tu and bacteriocin gene plaA of Lactobacillus plantarum 423, monitored with real-time PCR. *International Journal of Food Microbiology*, *116*(3), 405–409. https://doi.org/10.1016/j.ijfoodmicro.2007.02.011

Reeves, A. R., Wang, G. R., & Salyers, A. A. (1997). Characterization of four outer membrane proteins that play a role in utilization of starch by Bacteroides thetaiotaomicron. *Journal of Bacteriology*, *179*(3), 643–649. https://doi.org/10.1128/jb.179.3.643-649.1997

Reid, G., & Burton, J. (2002). Use of Lactobacillus to prevent infection by pathogenic bacteria.

Microbes and Infection, *4*(3), 319–324. https://doi.org/10.1016/S1286-4579(02)01544-7

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

Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, *11*(3), R25. https://doi.org/10.1186/gb-2010-11-3-r25

Robinson, M. D., & Smyth, G. K. (2008). Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics*, *9*(2), 321–332. https://doi.org/10.1093/biostatistics/kxm030

Rogemond, V., & Guinet, R. M. F. (1986). Lectinlike adhesins in the Bacteroides fragilis group. *Infection and Immunity*, *53*(1), 99–102. https://doi.org/10.1128/iai.53.1.99-102.1986

Roos, S., & Jonsson, H. (2002). A high-molecular-mass cell-surface protein from Lactobacillus reuteri 1063 adheres to mucus components. *Microbiology*, *148*(2), 433–442. https://doi.org/10.1099/00221287-148-2-433

Rossez, Y., Gosset, P., Boneca, I. G., Magalhães, A., Ecobichon, C., Reis, C. A., Cieniewski-Bernard, C., Curt, M. J. C., Léonard, R., Maes, E., Sperandio, B., Slomianny, C., Sansonetti, P. J., Michalski, J. C., & Robbe-Masselot, C. (2014). The lacdiNAc-specific adhesin LabA

mediates adhesion of Helicobacter pylori to human gastric mucosa. *Journal of Infectious Diseases*, *210*(8), 1286–1295. https://doi.org/10.1093/infdis/jiu239

Salyers, A. A., West, S. E. H., Vercellotti, J. R., & Wilkins, T. D. (1977). Fermentation of mucins and plant polysaccharides by anaerobic bacteria from the human colon. *Applied and Environmental Microbiology*, *34*(5), 529–533. https://doi.org/10.1128/aem.34.5.529-533.1977

Salzman, N. H., de Jong, H., Paterson, Y., Harmsen, H. J. M., Welling, G. W., & Bos, N. A. (2002). Analysis of 16S libraries of mouse gastrointestinal microflora reveals a large new group of mouse intestinal bacteria. *Microbiology*, *148*(11), 3651–3660. https://doi.org/10.1099/00221287-148-11-3651

Schreiber, O., Petersson, J., Phillipson, M., Perry, M., Roos, S., & Holm, L. (2009). Lactobacillus reuteri prevents colitis by reducing P-selectin-associated leukocyte- and platelet-endothelial cell interactions. *American Journal of Physiology - Gastrointestinal and Liver Physiology*, *296*(3), 534–542. https://doi.org/10.1152/ajpgi.90470.2008

Schroeder, B. O. (2019). Fight them or feed them: How the intestinal mucus layer manages the gut microbiota. *Gastroenterology Report*, *7*(1), 3–12. https://doi.org/10.1093/gastro/goy052

Sela, D. A., Chapman, J., Adeuya, A., Kim, J. H., Chen, F., Whitehead, T. R., Lapidus, A., Rokhsar, D. S., Lebrilla, C. B., German, J. B., Price, N. P., Richardson, P. M., & Mills, D. A. (2008). The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome. *Proceedings of the National Academy of*

Sciences of the United States of America, *105*(48), 18964–18969. https://doi.org/10.1073/pnas.0809584105

Sengupta, R., Altermann, E., Anderson, R. C., McNabb, W. C., Moughan, P. J., & Roy, N. C. (2013). The role of cell surface architecture of lactobacilli in host-microbe interactions in the gastrointestinal tract. *Mediators of Inflammation*, *2013*. https://doi.org/10.1155/2013/237921

Servin, A. L., & Coconnier, M. H. (2003). Adhesion of probiotic strains to the intestinal mucosa and interaction with pathogens. *Bailliere's Best Practice and Research in Clinical Gastroenterology*, *17*(5), 741–754. https://doi.org/10.1016/S1521-6918(03)00052-0

Shipman, J. A., Cho, K. H., Siegel, H. A., & Salyers, A. A. (1999). Physiological characterization of SusG, an outer membrane protein essential for starch utilization by Bacteroides thetaiotaomicron. *Journal of Bacteriology*, *181*(23), 7206–7211. https://doi.org/10.1128/jb.181.23.7206-7211.1999

Shoaf-Sweeney, K. D., & Hutkins, R. W. (2008). Chapter 2 Adherence, Anti-Adherence, and Oligosaccharides. Preventing Pathogens from Sticking to the Host. *Advances in Food and Nutrition Research*, *55*(08), 101–161. https://doi.org/10.1016/S1043-4526(08)00402-6

Simon, P. M., Goode, P. L., Mobasseri, A., & Zopf, D. (1997). Inhibition of Helicobacter pylori binding to gastrointestinal epithelial cells by sialic acid-containing oligosaccharides. *Infection and Immunity*, *65*(2), 750–757. https://doi.org/10.1128/iai.65.2.750-757.1997

Singh, A. K., Pluvinage, B., Higgins, M. A., Dalia, A. B., Woodiga, S. A., Flynn, M., Lloyd, A.

R., Weiser, J. N., Stubbs, K. A., Boraston, A. B., & King, S. J. (2014). Unravelling the multiple functions of the architecturally intricate Streptococcus pneumoniae β-galactosidase, BgaA. *PLoS Pathogens*, *10*(9), e1004364. https://doi.org/10.1371/journal.ppat.1004364

Singh, R. P. (2019). Glycan utilisation system in Bacteroides and Bifidobacteria and their roles in gut stability and health. *Applied Microbiology and Biotechnology*, *103*(18), 7287–7315. https://doi.org/10.1007/s00253-019-10012-z

Smeds, A., Hemmann, K., Jakava-Viljanen, M., Pelkonen, S., Imberechts, H., & Palva, A. (2001). Characterization of the adhesin of Escherichia coli F18 fimbriae. *Infection and Immunity*, *69*(12), 7941–7945. https://doi.org/10.1128/IAI.69.12.7941-7945.2001

Sojitra, M., Sarkar, S., Maghera, J., Rodrigues, E., Carpenter, E. J., Seth, S., Ferrer Vinals, D., Bennett, N. J., Reddy, R., Khalil, A., Xue, X., Bell, M. R., Zheng, R. B., Zhang, P., Nycholat, C., Bailey, J. J., Ling, C. C., Lowary, T. L., Paulson, J. C., … Derda, R. (2021). Genetically encoded multivalent liquid glycan array displayed on M13 bacteriophage. *Nature Chemical Biology*, *17*(7), 806–816. https://doi.org/10.1038/s41589-021-00788-5

Sonnenburg, J. L., Xu, J., Leip, D. D., Chen, C. H., Westover, B. P., Weatherford, J., Buhler, J. D., & Gordon, J. I. (2005). Glycan foraging in vivo by an intestine-adapted bacterial symbiont. *Science*, *307*(5717), 1955–1959. https://doi.org/10.1126/science.1109051

Suegara, N. et al. (1975). Specific adhesion of lactobacilli to keratinized epithelial cells of the rat stomach in vitro. *American Journal of Clinical Nutrition*, *30*(11), 1777–1780. https://doi.org/10.1093/ajcn/30.11.1777

Sun, S., Tay, Q. X. M., Kjelleberg, S., Rice, S. A., & McDougald, D. (2015). Quorum sensing-regulated chitin metabolism provides grazing resistance to Vibrio cholerae biofilms. *ISME Journal*, *9*(8), 1812–1820. https://doi.org/10.1038/ismej.2014.265

Suzuki, R., Wada, J., Katayama, T., Fushinobu, S., Wakagi, T., Shoun, H., Sugimoto, H., Tanaka, A., Kumagai, H., Ashida, H., Kitaoka, M., & Yamamoto, K. (2008). Structural and thermodynamic analyses of solute-binding protein from Bifidobacterium longum specific for core 1 disaccharide and lacto-N-biose. *Journal of Biological Chemistry*, *283*(19), 13165–13173. https://doi.org/10.1074/jbc.M709777200

Tailford, L. E., Crost, E. H., Kavanaugh, D., & Juge, N. (2015). Mucin glycan foraging in the human gut microbiome. *Frontiers in Genetics*, *6*(81). https://doi.org/10.3389/fgene.2015.00081

Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., & Muchmore, E. (2003). Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(21), 12045–12050. https://doi.org/10.1073/pnas.2131556100

Tannock, G. W. (1992). The Lactic Microflora of Pigs, Mice and Rats. *The Lactic Acid Bacteria Volume 1*, *1*, 21–48. https://doi.org/10.1007/978-1-4615-3522-5_2

Taylor, M. E., Leaning, M. S., & Summerfield, J. A. (1987). Uptake and processing of glycoproteins by rat hepatic mannose receptor. *American Journal of Physiology-Endocrinology and Metabolism*, *252*(5), E690–E698.

https://doi.org/10.1152/ajpendo.1987.252.5.E690

Tenaillon, O., Skurnik, D., Picard, B., & Denamur, E. (2010). The population genetics of commensal Escherichia coli. *Nature Reviews Microbiology*, *8*(3), 207–217. https://doi.org/10.1038/nrmicro2298

Tomlin, J., Read, N. W., Edwards, C. A., & Duerden, B. I. (1988). The digestion of guar gum by individual strains of colonic bacteria. *Microbial Ecology in Health and Disease*, *1*(3), 163–167. https://doi.org/10.3109/08910608809141532

Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiapello, H., Clermont, O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., El Karoui, M., Frapy, E., … Denamur, E. (2009). Organised genome dynamics in the Escherichia coli species results in highly diverse adaptive paths. *PLoS Genetics*, *5*(1). https://doi.org/10.1371/journal.pgen.1000344

Van de Vijver, K. K., Deelder, A. M., Jacobs, W., Van Marck, E. A., & Hokke, C. H. (2006). LacdiNAc- and LacNAc-containing glycans induce granulomas in an in vivo model for schistosome egg-induced hepatic granuloma formation. *Glycobiology*, *16*(3), 237–243. https://doi.org/10.1093/glycob/cwj058

Van Tassell, M. L., & Miller, M. J. (2011). Lactobacillus adhesion to mucus. *Nutrients*, *3*(5), 613–636. https://doi.org/10.3390/nu3050613

Varki, A., Etzler, M. E., Cummings, R. D., & Esko, J. D. (2009). Discovery and Classification of

Glycan-Binding Proteins. In *Essentials of Glycobiology*. http://www.ncbi.nlm.nih.gov/pubmed/20301249

Walter, J. (2008). Ecological role of lactobacilli in the gastrointestinal tract: Implications for fundamental and biomedical research. *Applied and Environmental Microbiology*, *74*(16), 4985–4996. https://doi.org/10.1128/AEM.00753-08

Wang, B., Wei, H., Yuan, J., Li, Q., Li, Y., Li, N., & Li, J. (2008). Identification of a surface protein from Lactobacillus reuteri JCM1081 that adheres to porcine gastric mucin and human enterocyte-like HT-29 cells. *Current Microbiology*, *57*(1), 33–38. https://doi.org/10.1007/s00284-008-9148-2

Wyatt, J. I., Rathbone, B. J., Sobala, G. M., Shallcross, T., Heatley, R. V., Axon, A. T. R., & Dixon, M. F. (1990). Gastric epithelium in the duodenum: Its association with Helicobacter pylori and inflammation. *Journal of Clinical Pathology*, *43*(12), 981–986. https://doi.org/10.1136/jcp.43.12.981

Xiong, Y., Zhai, Z., Lei, Y., Xiao, B., & Hao, Y. (2020). A Novel Major Pilin Subunit Protein FimM Is Involved in Adhesion of Bifidobacterium longum BBMN68 to Intestinal Epithelial Cells. *Frontiers in Microbiology*, *11*(590435). https://doi.org/10.3389/fmicb.2020.590435

Xu, Q., Shoji, M., Shibata, S., Naito, M., Sato, K., Elsliger, M. A., Grant, J. C., Axelrod, H. L., Chiu, H. J., Farr, C. L., Jaroszewski, L., Knuth, M. W., Deacon, A. M., Godzik, A., Lesley, S. A., Curtis, M. A., Nakayama, K., & Wilson, I. A. (2016). A Distinct Type of Pilus from the Human Microbiome. *Cell*, *165*(3), 690–703. https://doi.org/10.1016/j.cell.2016.03.016

Yaffe, E., & Relman, D. A. (2020). Tracking microbial evolution in the human gut using Hi-C reveals extensive horizontal gene transfer, persistence and adaptation. *Nature Microbiology*, *5*(2), 343–353. https://doi.org/10.1038/s41564-019-0625-0

Yan, M., Zhu, Y., Liu, X., Lasanajak, Y., Xiong, J., Lu, J., Lin, X., Ashline, D., Reinhold, V., Smith, D. F., & Song, X. (2019). Next-Generation Glycan Microarray Enabled by DNA-Coded Glycan Library and Next-Generation Sequencing Technology. *Analytical Chemistry*, *91*(14), 9221–9228. https://doi.org/10.1021/acs.analchem.9b01988

Yang, C., Mogno, I., Contijoch, E. J., Borgerding, J. N., Aggarwala, V., Li, Z., Siu, S., Grasset, E. K., Helmus, D. S., Dubinsky, M. C., Mehandru, S., Cerutti, A., & Faith, J. J. (2020). Fecal IgA Levels Are Determined by Strain-Level Differences in Bacteroides ovatus and Are Modifiable by Gut Microbiota Manipulation. *Cell Host & Microbe*, *27*(3), 467-475.e6. https://doi.org/10.1016/j.chom.2020.01.016

Yue, M., Rankin, S. C., Blanchet, R. T., Nulton, J. D., Edwards, R. A., & Schifferli, D. M. (2012). Diversification of the Salmonella Fimbriae: A model of macro- and microevolution. *PLoS ONE*, *7*(6). https://doi.org/10.1371/journal.pone.0038596

Yutin, N., & Galperin, M. Y. (2013). A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environmental Microbiology*, *15*(10), 2631–2641. https://doi.org/10.1111/1462-2920.12173