DIRECT VISUALIZATION OF YEAST MANNAN METABOLISM IN BOVINE-ADAPTED <u>BACTEROIDES THETAIOTAOMICRON</u> STRAINS AT THE SINGLE CELL LEVEL USING FLUORESCEINAMINE – YEAST MANNAN CONJUGATES

LEEANN KLASSEN
Bachelor of Science, University of Lethbridge, 2017

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

MASTER OF SCIENCE

in

BIOLOGICAL SCIENCES

Department of Biological Sciences
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

© Leeann Klassen, 2019

DIRECT VISUALIZATION OF YEAST MANNAN METABOLISM IN BOVINE-ADAPTED
BACTEROIDES THETAIOTAOMICRON STRAINS AT THE SINGLE CELL LEVEL USING
FLUORESCEINAMINE – YEAST MANNAN CONJUGATES

LEEANN KLASSEN

Date of Defence: May 10, 2019

Dr. D. Wade Abbott	Research Scientist	Ph.D.
Dr. Dmytro P. Yevtushenko	Associate Professor	Ph.D.
Thesis Co-Supervisors		
Dr. Trushar R. Patel	Assistant Professor	Ph.D.
Thesis Examination Committee Member		
Dr. Trevor Alexander	Research Scientist	Ph.D.
Thesis Examination Committee Member		
Dr. Bastien Castagner	Assistant Professor	Ph.D.
External Examiner		
McGill Univeristy		
Montreal, Quebec		
Dr. Igor Kovalchuk	Professor	Ph.D.
Chair, Thesis Examination Committee		

Dedication

This is dedicated to my family, all who worked hard to provide me with the means to

explore my curiosity, encouraged my personal development, and taught me to extol reason.

Abstract

The rumen of beef cattle houses a diverse community of microorganisms that impact feed digestion, nutrient accessibility, host health, and waste production. Fibrolytic enzymes, probiotics, and prebiotics are promising candidates for next-generation feed additives, with the goal of inducing beneficial changes to the rumen microbiome. Current methods to investigate the mechanistic interactions between bacteria and feed glycans are "indirect" and lack the required sensitivity to validate probiotic-prebiotic application. To address these limitations, I have extended fluorescent glycan conjugates (FGCs) to visualize polysaccharide uptake in bacterial isolates and complex ecosystems at the single-cell level. Yeast α-mannan (YM) was conjugated to 6-aminofluorescein and fed to pure cultures of *Bacteroides thetaiotaomicron* VPI-5482, a well-studied intestinal symbiont that metabolizes YM, and closely related bovine-adapted bacterial isolates. Uptake of FGCs, coupled to complementary genomic and transcriptomic analysis, provided direct evidence of individual genotypes endowed with YM metabolic potential in pure and complex culture.

Acknowledgements

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AA	Auxiliary Activity
ANI	Average Nucleotide Identity
ANIb	Average Nucleotide Identity based on BLAST+
B. theta	*Bacteroides thetaiotaomicron*
B. theta^{Bov}	Bovine-adapted *Bacteroides thetaiotaomicron* strain isolates
BHI	Brain Heart Infusion
BLAST	Basic Local Alignment Search Tool
CAZymes	Carbohydrate Active enZymes
CE	Carbohydrate Esterase
CNBr	Cyanogen Bromide
CP	Crude Protein
DDGS	Dried Distiller's Grains with Solubles
DGS	Distiller's Grains with Solubles
DHP	4-hydroxy-4(H)-pyridone
DOC	Dissolved Organic Carbon
ECF	Extracytoplasmic Function
ETEC	Enterotoxigenic *Escherichia coli*
FA	Formaldehyde
FCR	Feed Conversion Ratio
FGC	Fluorescent Glycan Conjugate
FISH	Fluorescence *in situ* Hybridization
FLA	Fluoresceinamine, 6-aminofluorescein
FLA-YM	Fluoresceinamine – Yeast Mannan conjugate
FOS	Fructooligosaccharide
GH	Glycoside Hydrolase
GHG	Greenhouse Gas
GIT	Gastrointestinal Tract
GlcNAc	N-acetylglucosamine
GT	Glycosyl Transferase
HG	High Grower
HMNG	High Mannose *N*-Glycan
HTCS	Hybrid Two Component System
L. leucocephala	*Leucaena leucocephala*
Man	Mannose
MAN-PUL	Yeast Mannan Polysaccharide Utilization Locus
MFS	Major Facility Superfamily
MG	Medium Grower
MM	Minimal Medium
MOS	Mannanoligosaccharide
NaOH	Sodium Hydroxide
NDF	Neutral Detergent Fiber
PL	Polysaccharide Lyase
PULs	Polysaccharide Utilization Loci
RFI	Residual Feed Intake
RGII	Rhamnogalacturonan-II

RNA-seq	RNA sequencing
S. cerevisiae	*Saccharomyces cerevisiae*
SACCHARIS	Sequence Analysis and Clustering of CarboHydrate Active enzymes for Rapid Informed prediction of Specificity
SCFA	Short Chain Fatty Acid
SGBP	Surface Glycan Binding Protein
SR-SIM	Super Resolution – Structured Illumination Microscopy
Sus	Starch utilization system
TBDT	Ton-B Dependent Transporter
TLC	Thin Layer Chromatography
TPM	Transcript per kilobase per million
WDGS	Wet Distiller's Grains with Solubles
YM	Yeast Mannan
YM-MM	Yeast Mannan in Minimal Medium

Chapter 1

Literature Review

1.1 Feed efficiency

Ruminants produce meat and milk from vegetal biomass, providing significant sources of protein, vitamins, and minerals for humans. As such, ruminant livestock are a vital part of food security for the increasing global population of consumers [1]. Canada ranked sixth in beef exporting nations in 2017 (excluding live slaughter) [2]; exporting over 640,000 cattle to the United States alone [3]. Beef production and exportation face increasing challenges, such as limiting the use of antibiotics [4, 5], minimizing land and water usage [6], and reducing greenhouse gas (GHG) emissions [7]. To address these issues, organizations, such as the Beef Cattle Research Council (BCRC), are investing in research that aims to provide novel solutions [8, 9]. "Feed Grains" and "Feed Efficiency" are top priorities to the beef industry, advocating for the development of innovative technologies to increase feed conversion efficiency and protein output, while limiting the use of finite resources and the unnecessary production of GHGs and other waste byproducts.

Feed efficiency is the amount of valued product generated from a portion of feed intake, most commonly measured as feed conversion ratio (FCR) or residual feed intake (RFI) [10]. FCR is the ratio of consumed feed to live-weight increase, while RFI measures the difference between actual feed consumed and theoretical feed intake. More efficient cattle gain more weight from consuming less feed than expected, which translates into reduced feed costs. Feed is the major production cost in the cattle industry, with an estimated 50-70% of funds spent on feedstuff. Much of the energy stored in the fiber- and carbon-rich feed is not digested or is wasted as methane, a powerful GHG that contributes to climate change. Indeed, increasing feed efficiency and fiber digestibility has been shown to have positive economic and environmental impacts through improved animal performance with reduced feed intake and methane emissions [9, 11]. Boaitey *et al.* estimated that with every 1% reduction in feed intake (kg fed day-1), net returns and GHG

emissions will increase by 2.3% and decrease by 33.46 tonnes, respectively [11]. However, improvements to feed efficiency are constrained due to a limited understanding of feed utilization in the ruminant digestive system and of efficient methods to manipulate this ecosystem for increased live-weight gain.

1.2 The rumen

The gastrointestinal tract (GIT) of mammals contributes to many physiological processes [12], such as nutrient absorption, waste excretion, and pathogen regulation. Ruminants, such as *Bovinae* (cattle, bison) and *Ovinae* (sheep), have evolved efficient digestive systems composed of four compartments that precede the small and large intestines: the rumen, reticulum, omasum, and abomasum. The rumen is the primary fermentation chamber; microbial depolymerization of complex feed fibers begins in this compartment. The esophagus facilitates bidirectional movement of feed and is critical for regurgitation and rumination when the consumed biomaterial requires further mastication; a process commonly referred to as "chewing the cud". The reticulum follows the rumen and is sometimes considered an anterior sac of the rumen, as movement of ingesta is less restricted between the two compartments. The omasum has specialized folds that increase surface area to promote water and fatty acid absorption. Finally, the abomasum most resembles the stomach of monogastric animals due to its production of hydrochloric acid and digestive enzymes.

Ruminants primarily consume a diet of plant glycans. Thus, to optimize access to carbon stored in plant fibers, ruminants rely on the rumen microbiota, as it is critical for the digestion and metabolism of feed by the animal. Rumen microorganisms, most notably bacteria, anaerobic fungi, and fibrolytic protists, depolymerize polysaccharides into oligosaccharides and monosaccharides which are transported into microbial cells for further saccharification or direct entry into glycolysis and fermentation pathways. Short chain fatty acids (SCFAs) are byproducts of microbial fermentation and are often used by host epithelial cells or transported to the liver for gluconeogenesis. It is estimated that up to 70% of bovine caloric demand is derived from SCFAs

2

[13]. As such, the rumen community is crucial for feed digestion and energy production in ruminant animals.

1.2.1 The rumen microbiome

The foregut chambers, primarily the rumen and reticulum, accommodate a diverse community of microorganisms associated with nutrient accessibility [14], host health [15], and waste production [11]. Ruminants are robust herbivores due to the evolution of this key symbiotic relationship. The rumen provides the microflora an anaerobic, carbon rich environment and, in turn, microbial metabolites provide the host with absorbable energy and nutrients [13]. Foregut fermentation is not unusual in other mammals, however individual microbiomes exhibit significant variation and diversity due to evolutionary divergence [16]. The ruminant digestive system is more elaborate than monogastric systems due to the physiology of the rumen, process of rumination, and expanded role of the microbiome in fiber digestion and feed efficiency.

The bovine diet is a cornucopia of carbohydrates that sustain a dense community of bacteria, fungi, protozoa, and archaea [12, 17, 18]. Henderson *et al.* reported the microbiome compositions of 32 ruminant species across the globe and found that four bacterial phyla dominated the majority of samples: Bacteroidetes, Firmicutes, Fibrobacter and Proteobacteria [19]. The most abundant bacterial taxa in the samples were *Prevotella, Butyrivibrio, Ruminococcus, Lachnospiraceae, Bacteroidales,* and *Clostridiales*, accounting for 67.1% of total bacterial sequencing data. The composition of the rumen microbiome is influenced by many intrinsic, environmental, or therapeutic factors, such as diet, host genetics, and antibiotics [20, 21].

Traditionally, cataloguing the species present in the rumen microbiota was performed using culture-dependent methods developed by Robert Hungate, who is often referred to as the father of rumen microbiology [22]. Despite the large extent of studies performed, bacteria cultivated using Hungate's anaerobic isolation techniques are estimated to represent less than 20% of bacterial communities in the rumen [23]. Recent advances in culturing methods and sequencing technologies

have resulted in geometric increases in the amount of sequencing data, providing previously unobtainable information on the diversity of microorganisms that live in these communities.

In addition to revealing the identity of an unprecedented number of species from the rumen microbiota, next-generation sequencing techniques, such as shotgun whole genome metagenomics and metatranscriptomics analyses, have also resulted in the identification of extensive datasets of uncharacterized genes. Bioinformatic pipelines and phylogenetics have become key toward deconvoluting the accurate prediction of gene functions [24]. For example, using metagenomics, Hess *et al.* determined the abundance of bovine rumen bacteria that encode expansive collections of carbohydrate-active enzymes (CAZymes), identifying over 27,000 putative genes [25]. More recent investigations have expanded this number to over 69,000 using genome assemblies [26]. These studies underpin the vast catalytic potential of the rumen microbiome, particularly in the context of improving feed digestibility.

1.3 Feed additives

Despite the robust catalytic capabilities of the rumen microbiota, conversion of dietary polysaccharides into milk and meat is inefficient, resulting in significant loss of potential energy from feed [14]. In this respect, manipulating the rumen microbiome to refine feed efficiency can address recent challenges affiliated with sustainable beef production [15]. Probiotics [20], prebiotics [27], and fibrolytic enzymes [28] are promising feed interventions for improving animal productivity, as each have the potential to increase the metabolizable energy from feed and promote growth.

1.3.1 Probiotics

Probiotics are living microorganisms that, when ingested in adequate quantities, improve GIT health [20]. This can be achieved through multiple avenues [29, 30], such as: supporting a healthy microbial community, displacing pathogens [31], improving digestion and weight gain [32, 33], managing pH [34], producing SCFAs [35] and promoting mucosal immunity [36]. Efficacy of

probiotics depends on many parameters, such as dose, length of administration, and life stage of the animal [30]. Consistent treatment of probiotics promotes rapid colonization of beneficial bacteria that can increase fiber digestibility, exercise host immunity, and prevent pathogen settlement early in the life of the animal, with decreased effects later in life [30, 37]. Although significant advances have been made in the field of rumen microbiology, understanding the function of rumen microbiota remains elusive.

Results of rumen manipulation experiments are often ineffective, with only a few examples of compelling successes [38]. The 'mimosine story' is a common example of an investigation into the function of rumen bacteria that resulted in the development of a novel probiotic [38]. *Leucaena leucocephala* is a nutrient rich perennial legume plant and common cattle feed in tropical countries [39]. In the 1960s, cattle farmers in Australia were interested in adopting the leguminous plant as a feedstock for cattle; however, *L. leucocephala* contains the compound mimosine, which is converted to the toxin 4-hydroxy-4(H)-pyridone (DHP) in the rumen. DHP had a negative impact on Australian cattle, including weight loss, salivation, loss of appetite, and decreased animal performance [40]. Interestingly, goat and cattle populations in Maui, Hawaii displayed resistance to the toxicity of this compound. The protection came from the presence of *Synergistes jonesii* in the rumen microbiota of resistant populations, which facilitates the degradation of DHP [41]. Introducing DHP-degrading bacteria into the microbiomes of non-resistant ruminants neutralized the toxic side effects of *L. leucocephala* [40, 42]. This is an excellent example of how the rumen ecosystem can be manipulated with live microbial adjuncts to improve animal performance and welfare.

1.3.2 Prebiotics

Prebiotics are dietary compounds that are indigestible to host enzymes and shift the GIT microbial community to a composition that is more beneficial for the host [43]. Generally, prebiotics are selectively used to stimulate a bloom of beneficial bacteria that provide host benefits,

such as changing pH through SCFA production [34], priming the community for members more adept at utilizing feed [30], improving animal performance [27], and colonization resistance [44]. The theory behind prebiotics is supported by the nutrient-niche hypothesis described by Freter *et al.*: host diet, as it relates to nutrient accessibility in the GIT, drives microbial species to adopt an ecological role (*i.e.* niche) that provides them with a competitive advantage [45]. Furthermore, the link between host dietary carbohydrates and microbiota composition and activity is well-established [39, 46, 47]. In this regard, prebiotics, typically carbohydrates, support a healthy GIT through the modulation growth and/or displacement of select microbial species [30].

Fructooligosaccharides (FOS) and α-mannanoligosaccharides (MOS) are the most widely used and studied bovine prebiotics, and are often used as supplements in calf milk replacer to increase carcass weight and alter SCFA production [27, 48]. MOSs\ also mimic the structure of host epithelial high-mannose *N*-glycans (HMNGs) and are used to disrupt the attachment and effacement of pathogens to host epithelial cells [27, 49]. However, MOS has been shown to have other effects on the rumen of cattle as well. For example, Diaz *et al.* supplemented grain-based diets of Holstein steers (n=3) with a live yeast (*Saccharomyces cerevisiae*) probiotic or MOS, which resulted in an increase of ruminal pH and decreased levels of lipopolysaccharide and serum amyloid A in the blood, all of which lead to decreased inflammation [34]. High grain diets tend to decrease rumen pH and promote inflammation, thus live yeast and MOS could be used to treat these side effects.

Prebiotics impact the digestive systems of mono- and polygastric animals differently [30]. This is likely attributed to microbial and host-specific processes, host physiology, and diet [50]. Notably, combining probiotic and prebiotic supplements, a process referred to as synbiotics [43, 51], may improve their efficacy by introducing a unique nutrient that is specific to the metabolic capabilities of the probiotic. Selective prebiotics provide a competitive advantage to the probiotic strain and increase their abundance in the GIT ecosystem. Further investigation into the role of

probiotics, prebiotics, and synbiotics in the rumen ecosystem is required before they will be fully adopted as solutions for contemporary livestock production.

1.3.3 Fibrolytic enzymes

Mammals only genetically encode enzymes that aid in the depolymerization of a few dietary carbohydrates, such as starch and sucrose [52]. Therefore, ruminants rely on enzymes possessed by rumen microorganisms to digest high molecular weight, complex carbohydrates. Diverse, competitive microbial ecosystems have favored the evolution of unique enzymes, pathways, and foraging strategies for harnessing energy from complex carbohydrates [53, 54]. In this regard, application of exogenous enzymes to feed glycans can facilitate depolymerization and the release of absorbable energy from feedstocks [55]. Holtshausen *et al.* found a 10.7% increase in milk production efficiency of Holstein dairy cattle fed an alfalfa diet supplemented with a high concentration (1.0 mL kg_{-1}) of the developmental fibrolytic enzyme, Econase [56]. This effect was concentration-dependent, as the low enzyme treatment (0.5 mL kg_{-1}) did not result in a significant change in efficiency. Application of fibrolytic enzyme to the diets of finishing steers that contained 30% wheat DDGS, was shown by He *et al.* to improve starch digestion, protein metabolism, feed efficiency, and cattle health [57]. In this regard, improvements to feed digestibility and efficiency may result from supplementing enzyme activities that are limited within the rumen.

1.4 Carbohydrates

Carbohydrates are the most prevalent class of biomolecules and are fundamental to the structural integrity and physiology of all living organisms. Polysaccharides are vital structural components of plant fibers common in livestock feed and are the major contributor to plant biomass. Common plant polysaccharides include cellulose, xylan, pectin, and starch. Some of most structurally complex carbohydrates are rhamnogalacturonan II (RGII) [58] and yeast α-mannan [59], which have diverse composition, stereochemical linkages, and / or branching. To digest fungal

cell wall and plant fibers, which are especially complex glycans, a wide diversity of enzyme activities is required.

1.4.1 Yeast mannan

The cell wall of the common baker's yeast, *Saccharomyces cerevisiae,* is a matrix of β-glucan, chitin, and mannan; which also have propensity to cross-link with each other [60]. β-glucans are β-1,3 and β-1,6-linked glucose polymers of the yeast cell wall that extend up to ~1,500 residues. Chitin is a homopolymer of β-1,4-linked *N*-acetylglucosamine (GlcNAc) and makes up less than 2% of the yeast cell wall dry mass [60, 61]. *N*-linked and *O*-linked mannoproteins comprise approximately 40% of the yeast cell's dry mass. Yeast α-mannan (YM) is a complex polysaccharide that forms *N*-linkages to cell wall proteins and is a homopolymer of α-mannose residues [61] (Figure 1.1). YM has an extensive α-1,6 linked backbone (>150 units), adorned with α-1,2-mannose side chains capped by α-1,3 linked mannose [59]. Phosphodiester bonds are occasionally integrated between a secondary side chain branching from the C6 of the α-1,2-mannose residues. The side chains of YM are densely packed, which restricts enzyme accessibility [59]. Thus, complete YM saccharification requires the combined effort of functionally distinct enzymes (Figure 1.2).
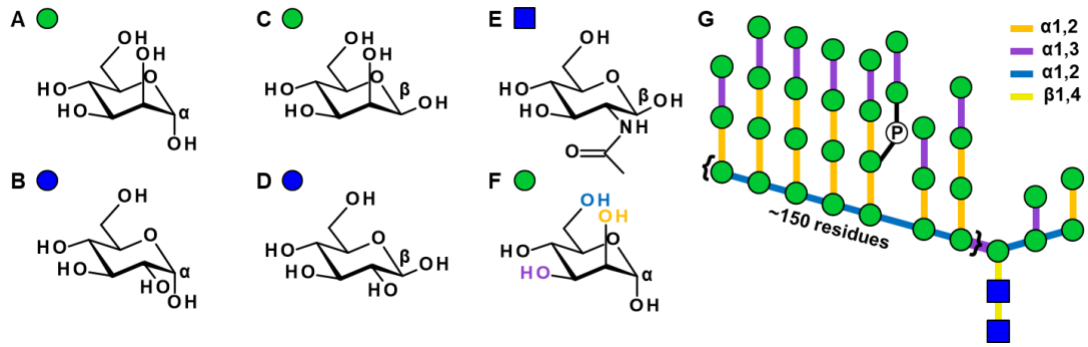
Figure 1.1 Glycosidic linkages between mannose residues of yeast α-mannan (YM).
(**A**) α-mannose. (**B**) α-glucose. (**C**) β-mannose. (**D**) β-glucose. (**E**) *N*-acetylglucosamine
(GlcNAc). (**F**) α-mannose with O2, O3, and O4 hydroxyl groups involved in YM linkages are
highlighted using the colour scheme shown in (G). (**G**) Structural schematic of YM. Coloured
bars indicate discriminatory linkages, Green circles = mannose, Blue circles = glucose, Blue
squares = GlcNAc.
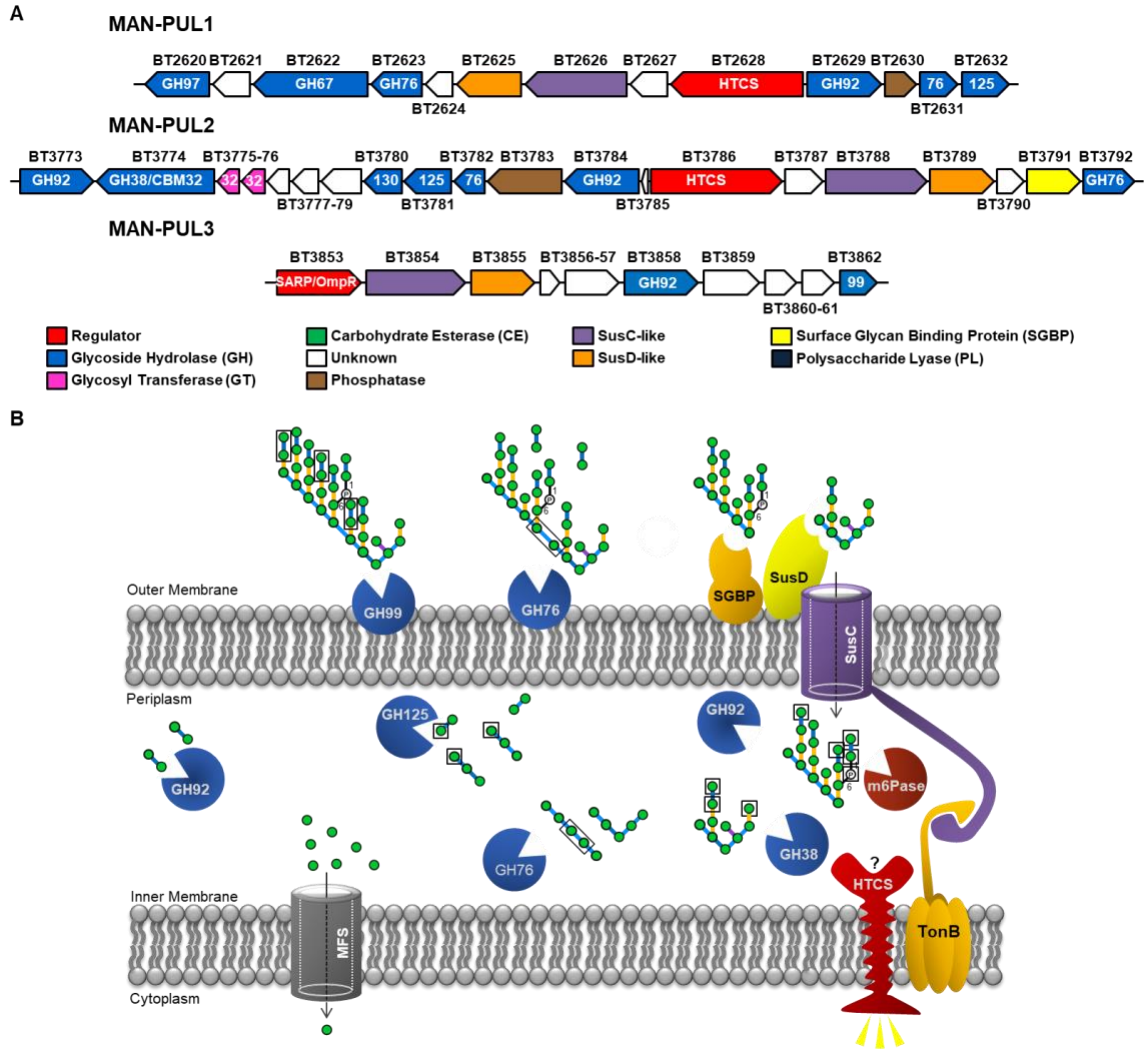
Figure 1.2 Model *Bacteroides thetaiotaomicron* utilization of mannan from *Saccharomyces cerevisiae.*

(**A**) Three MAN-PULs encoded in the genome of B. theta VPI-5482 [59]. Gene direction is indicated by arrows, and sizes representative relative length. Colour scheme indicates different functions for genes, which are labeled in the figure legend. (**B**) Model of YM and YM-product flow through the YM-degradation system. Depolymerization of YM initiates on the surface of the cell and products are transported through the SusC-like transporter into the periplasm. Complete depolymerization of YM-products by α-mannanases, α-mannosidases, and glycosidic phosphatases occurs in the periplasm. Mannose is transported into the cytoplasm by an MFS transporter. A signature YM oligosaccharide binds to the HTCS regulator and increases transcription of MAN-PULs.

1.4.2 Yeast mannan derived commercial products

Interest in YM and its polymers began in the 1980's due to its potential therapeutic properties against enterobacteria [62-64]. Enteric pathogens, such as enterotoxigenic *Escherichia coli* (ETEC)*,* produce lectins that bind to mannose-containing glycans that decorate the surface of mammalian gastrointestinal epithelial cells [65]. Polymers of YM mimic host mannans and bind to the lectins attached to the bacterial cell wall; this prevents the pathogen from colonizing and effacing the host epithelial tissue.

YM is a component of commercial prebiotics, such as Alltech's Bio-Mos® [64], as well as dried distillers grains with soulbles (DDGS) [66]. Bio-Mos® has been sold commercially since the early 1990s and is produced by the extraction of MOS from the cell surface of *S. cerevisiae*. Similarly, DDGS contains *S. cerevisiae* natively found on maize (corn), among other yeasts and enzymes, and is a co-product of corn fermentation and ethanol production. DDGS has been used as an alternative feed for livestock and is thought to be a nutritious feed alternative containing neutral detergent fiber (NDF) and crude protein (CP) [28].

1.5 Carbohydrate-active enzymes

CAZymes are enzymes that are categorized into five classes that catalyze the saccharification, biosynthesis, or modification of glycans and their derivatives: glycosyl transferases (GTs), polysaccharide lyases (PLs), carbohydrate esterases (CEs), auxiliary activities (AAs), and glycoside hydrolases (GHs) [67]. Diversity in the monosaccharide composition, stereochemical linkage, and branching of complex carbohydrates has promoted the evolution of highly specific CAZymes. GTs catalyze the formation of glycosidic linkages. PLs act by β-elimination to cleave polysaccharides containing uronic-acid groups. CEs catalyze the hydrolysis of carbohydrate esters. AA is a class of enzymes with redox activity that work in tandem with other

CAZymes. The largest family of CAZymes is the GHs, which depolymerize carbohydrate substrates by hydrolysing glycosidic linkages.

Hydrolysis of a glycosidic linkage is catalyzed by two amino acid residues in GH enzymes; a general acid (*e.g.* glutamic acid) acts as the proton donor and a nucleophile/base (*e.g.* aspartic acid) [67]. There are two primary catalytic mechanisms of GHs: single displacement and double displacement, which either invert or retain the anomeric configuration at C1, respectively. According to the CAZy database [67, 68], GHs are divided into over 160 sequence-related families [68], further categorized into subfamilies that often display similar catalytic specificities [67]. Members of the same family have conserved tertiary structure, mechanism, and catalytic residues; however, this does not necessarily translate into functional redundancy. In polyspecific families, members can be further partitioned into subfamilies that can differ in mode of action (e.g. endo-verses exo-hydrolysis) or substrate specificity [69-72]. This property can make functional characterization of CAZymes within the same family challenging. For example GH92s, an enzyme family that has been shown to be heavily involved in the digestion of α-mannans, have been shown to be active on α-1,2, α-1,3, α-1,4, α-1,6 linked and α-1-mannosyl-phosphate substrates; whereas members of GH76 and GH125 families are currently classified as exclusive endo-α-mannanases and exo-α-1,6-mannosidases, respectively [59, 73]. Functional redundancy within a family, however, may simply reflect a limited sample size of characterized activities. *In silico* tools, such as SACCHARIS (Sequence Analysis and Clustering of CarboHydrate Active enzymes for Rapid Informed prediction of Specificity) [24], will facilitate rapid and accurate prediction of CAZyme function, especially in polyspecific families. Some microbes encode polysaccharide utilization loci (PULs), which are co-regulated gene clusters that can include GHs, regulators, carbohydrate-binding proteins, and transporters that work together to depolymerize a specific carbohydrate [53].

1.6 Polysaccharide utilization loci

The starch utilization system (Sus) of *Bacteroides thetaiotaomicron* (*B. theta*) was the first PUL to be characterized [53, 74, 75]. PULs are distinguished by the presence of a TonB-dependent transport (TBDT) and surface glycan binding protein (SGBP) complex [76], commonly referred to as the *susC/D*-like pair, which is responsible for binding to carbohydrates on the outer cell membrane and importing them into the periplasm [53]. SusR-like, hybrid two-component system (HTCS), and ECF-sigma/ant-sigma are the three systems that can regulate a PUL [53]. GHs depolymerize a polysaccharide into signature products that are recognized by the periplasmic domain of the regulator and cause a signal cascade that results in PUL upregulation. Complete depolymerization of oligosaccharides produces monosaccharides that are imported into the cytoplasm of the cell through a major facility superfamily (MFS) transporter. PULs encode all the necessary CAZymes to degrade chemically distinct polysaccharides and these systems represent relatively unexplored repositories of enzymes and enzyme activities that was previously limited by our inability to culture the vast majority of gut microbes. However, with new –omics techniques and next-generation sequencing, we can now use bioinformatics to map PULs and the CAZy database to annotate predicted functions of CAZymes [53, 67, 68, 76]. While PULs have been found in many commensal bacterial species residing in the gut, members of Bacteroidetes encode a plethora of PULs in their genomes [53], with over 30,000 predicted PULs for the specialized degradation of complex carbohydrates identified in the phylum [67, 68, 76].

1.6.1 A model of YM depolymerization

Typically, one polysaccharide activates one PUL [53]. YM is an exception to this rule as it was shown to activate multiple PULs within the same bacterium (Figure 1.2A) [59]. In addition, YM-PULs are promiscuous as they target α-mannans derived from the cell walls of different yeast species. This may reflect the complexity of mannans found across yeast species as *S. cerevisiae, Schizosaccharomyces pombe,* and *Candida albicans* display structural heterogeneity, and therefore,

diverse enzymes are necessary to recognize and cleave these substrates. Importantly, the subtle changes to PUL architecture in strain variants may result in differential foraging strategies and provide individual strains with fitness advantages within complex communities.

Cuskin *et al.,* characterized three PULs encoded in *B. theta* VPI-5482 that are responsible for depolymerizing YM: MAN-PUL1, MAN-PUL2, and MAN-PUL3 [59] (Figure 1.2A). Together, the PULs encode 15 enzymes, SGBP, SusC-like transporter, and HTCS regulator/sensors that coordinate the complete saccharification of *S. cerevisiae* YM (Figure 1.2). The extracellular enzyme GH99, an endo-α-1,2-mannosidase, is bound to the outer membrane of *B. theta* VPI-5482 and hydrolyses the YM branches into man-α-1,3-man disaccharides, decluttering the backbone. Upon removal of a few branches, an endo-α-1,6-mannanase (GH76) can cleave the exposed backbone to generate products that can be imported through the SusC-like transporter. Once in the periplasm, stochastic hydrolysis of YM occurs. Specifically, mannose-6-phosphatases cleave the phosphodiester bond, and endo-α-1,2/3-mannosidases (GH92), exo-α-1,2-mannosidases (GH38), endo-α-1,6-mannanases (GH76), and exo-α-1,6-mannosidases (GH125) hydrolyse the remaining linkages. Finally, released mannose residues are imported into the cytoplasm through an MFS transporter.

MAN-PULs may impose an energetic debt that hinders cell fitness in YM deficient environments. As reported in Cuskin *et al.*, mutant strains lacking the three MAN-PULs or MAN-PUL2 outcompete the wild type strain in gnotobiotic mice fed a glycan-free diet [59]. Alternatively, when mice are fed a diet rich in YM, the wild type *B. theta* strain outcompetes mutant strains. These results, alongside the expansive collection of YM utilizing machinery, support the hypothesis that YM metabolism provides a competitive advantage to *B. theta* VPI-5482 [77]. Although well-studied for human intestinal symbionts, investigation of YM utilization and its derivatives, such as Bio-Mos® and DDGS, in ruminant microbes is lacking. Understanding the relationship between rumen microbiota and feed ingredients is essential to develop and validate novel probiotic and prebiotic treatments.

1.7 Fluorescent glycan conjugates

Traditional techniques to study interactions between the microbiota and dietary carbohydrates rely on indirect methods, such as growth kinetics, genomics, and enzymology [59, 78]. Generally, these techniques are slow, expensive, and indirect. In this regard, the development of fluorescent polysaccharides (*i.e.* fluorescent glycan conjugates; FGCs) may provide a novel molecular tool to visualize polysaccharide utilization by rumen microbiota, which will facilitate selection of species with particular metabolic capabilities that can be further investigated for novel CAZymes.

Production of fluorescent carbohydrates is lagging behind fluorescently labelled proteins and nucleic acids [79]. This is likely attributed to the structural complexity and variability of complex carbohydrates, making site-specific labelling difficult. There is also concern regarding label interference of carbohydrate-protein interactions. Regardless of these challenges, many fluorescence techniques have been developed to study carbohydrates in biological systems, including boronic acid-based fluorescent detection, fluorogenic labelling agents, and fluorescently-labelled glycolipids [79]. Many of these techniques label the reducing end of oligosaccharides, with a focus on mono- and oligosaccharides. An alternative and highly versatile technique was established by Glabe *et al.* in 1983, in which the fluorescent probe fluoresceinamine (FLA) was conjugated to a variety of polysaccharides [80]. The hydroxyl group of a carbohydrate becomes activated when mixed with cyanogen bromide. Subsequently, FLA is coupled to the activated product via its amino group. These fluorescein derivatives are stable and exhibit activity equal to underivatized polysaccharides when used in inhibition assays.

FGCs were initially used to study the interaction between polysaccharides and marine bacterial enzymes. In the 1990's, fluorescent polysaccharides were injected into samples of marine sediments, providing a means to measure the hydrolysis rates of extracellular bacterial enzymes found naturally in these samples [81, 82]. Using this approach, Arnosti found differential rates of hydrolysis of two labelled polysaccharides, FLA-laminarin and FLA-pullulan, across a range of

sediment depths [81], which showcased the utility of FGCs as tools to investigate complex interactions between bacteria and carbohydrates. The two FGCs used in the study were found to be easily and rapidly detected, unaffected by background noise of dissolved organic carbon (DOC), reliable over time and within a broad range of concentrations, and effective at environmentally relevant concentrations. Arnosti further revealed the versatility of fluorescein derivatives by refining labelling procedures and demonstrating their efficacy with a more diverse range of substrates, including sulfated polysaccharides [83]. More recently, metabolic strategies and potentials of marine bacteria were investigated and will be described below.

1.7.1 Investigating microbial communities with FGCs

It was assumed that FGCs were too bulky to be imported into the intracellular space of bacterial cells, thus hydrolysis by extracellular and/or cell-surface enzymes was the sole activity being measured [81]. As the depolymerization of polysaccharides took place outside the cell, byproducts could be shared with near-by microbes; a nutrition acquisition strategy known as the 'distributive' mechanism. Two decades later, with use of FGCs, this hypothesis was challenged; Reintjes *et al.,* incubated ocean water collected from five oceanic provinces with one of three fluorescently-labelled polysaccharides and, using fluorescence *in situ* hybridization (FISH), were able to visualize uptake of three FGCs by marine bacteria using super-resolution structured illumination microscopy (SR-SIM) [84]. The polysaccharides used in the study were laminarin, xylan, and chondroitin sulphate, structurally diverse glycans commonly utilized by marine bacteria. The high-resolution imaging of SR-SIM provided images displaying the uptake of FGCs into the periplasm or paryphoplasm of marine bacteria identified using FISH that belong to three distinct phyla: Bacteroidetes*,* Planctomycetes*,* and Proteobacteria. Using this technique, Reintjes *et al.* demonstrated that for these bacteria, the glycans used in the study are hydrolysed first by extracellular enzymes before being transported into the cell, preventing carbon loss to the

surrounding microbes. Previous to this data, an alternative polysaccharide uptake strategy had only been suggested via extrapolations from metagenomic data [85].

Reintjes *et al.,* hypothesized that the uptake of high-molecular-weight polysaccharides is homologous to the 'selfish' mechanism described by Cuskin *et al.* in 2015 [59]. This foraging strategy is bacteria and glycan specific and is dependent on the architecture of the pathway and bacterial cell wall morphology. 'Selfish' mechanisms result in limited depolymerization of a polysaccharide on the cell surface and rapid importation of products into the periplasm for further enzymatic cleavage. By changing the site of hydrolysis to the periplasm, the cell prevents the diffusion of lost products. Evolution of this strategy, as described in the Sus-like paradigm of gut *Bacteroides* [53, 77], is suggested to increase fitness by preventing loss of high energy products, opposing the 'distributive' strategy of nutrient uptake. The impact of this finding is significant, as it called for a re-evaluation of previous models of carbohydrate dynamics in marine ecosystems and refocusing how bacterial carbohydrate metabolism is measured.

1.7.2 FGCs in microbiome research

FGCs have proven beneficial in studies surveying enzyme activity and tracing complex carbohydrate dynamics in environmental samples. The high-resolution detectability, versatility, lack of interference of substrate function, and structural stability of fluorescein derivatives make them promising candidates as tools to investigate microbe-glycan interactions in complex communities, such as within the rumen. Use of these first-in-class molecular probes could provide direct evidence of bacterial genotypes endowed with specific glycan metabolic pathways, and provide a technique to correlate metabolic phenotypes with genotypes. In this manner, direct visualization of FGC uptake at the single-cell level could help determine the distribution of functionally divergent pathways within the genomes of rumen bacteria.

Probiotic, prebiotic, and exogenous enzyme applications in the livestock industry, although promising strategies, have been hindered by slow characterization of sequencing data and lack of

reliable molecular tools to inform selection of relevant microbes. This thesis considers FGCs as a first-in-class molecular tool to provide rapid detection, visualization, and quantification of bacterial-glycan interactions that can be used to inform microbiome research useful for ruminant digestive health and the beef industry.

1.8 Hypothesis

As a fungal cell wall polysaccharide, YM is present in many feedstuffs provided to livestock and is a nutrient available in diverse ecosystems, including the rumen microbiome. The ability of rumen bacteria to grow with YM as a carbohydrate source suggests the bacteria encode the requisite machinery (CAZymes, transporters, and regulators) to utilize and metabolize this complex glycan substrate. These interactions between bacteria and nutrient source can be taken advantage of to develop new tools for directly monitoring polysaccharide dynamics in complex samples and visualizing bacterial interactions with the substrate. This project aims to test the hypothesis:

Bovine-adapted *B. theta* strains have PUL systems tailored for the metabolism of YM in the rumen, and the accumulation of YM-FGCs in the periplasm can be visualized directly using fluorescence.

1.9 Objectives

1. Use growth profiling, whole genome sequencing, and RNA-seq to characterize utilization of YM by *B. theta* strains isolated from bovine rumen and fecal samples (Figure 1.3B)

2. Produce fluoresceinamine – YM conjugates (FLA-YM) for use as molecular-based probes.

3. Visualize the uptake of FLA-YM by bovine-adapted *B. theta* strains using super resolution fluorescence microscopy (Figure 1.3C).

A   **B**   **C**

Enrich on YM   Growth Profiles

Bioinformatics

Extract Rumen Sample   Cultivate Isolates   Fluorescence Microscopy

Rumen
Reticulum
Omasum
Abomasum
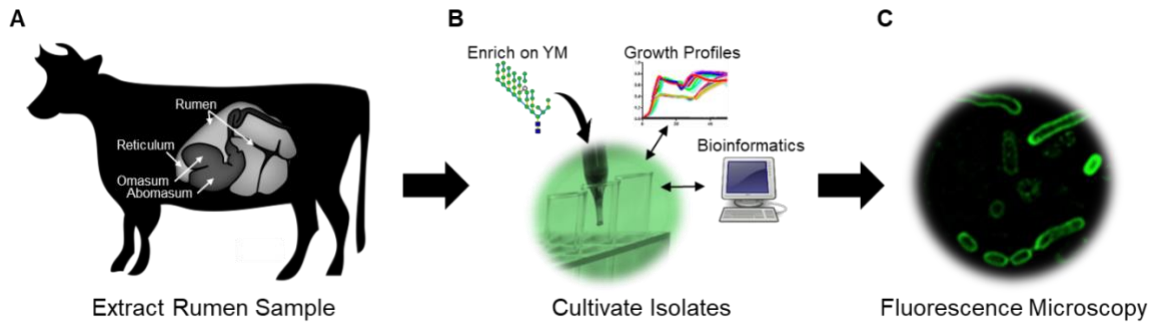
Figure 1.3 Graphical representation of thesis objectives.
(A) Collection of bacterial isolates from the rumen of cattle. **(B)** Selective isolation and characterization of rumen bacteria growth profiles, including genomic and transcriptomic analysis. **(C)** Visualize YM utilization capabilities using FLA-YM.

Chapter 2

Mechanistic insights into differential growth phenotypes of bovine-adapted *Bacteroides thetaiotaomicron* strains using comparative genomics, transcriptomics, and fluorescent glycan conjugates

2.1 Introduction

Ruminants have evolved foregut digestive systems specialized for the metabolism of plant cell wall material. The rumen houses a dense and diverse microbial community consisting of bacteria, ciliated protozoa, anaerobic fungi, and bacteriophages [86]. A core rumen bacterial microbiome has been described [87], which is predominantly comprised of *Prevotella*, *Butyrivibrio*, *Ruminococcus*, *Lachnospiraceae*, *Ruminococcaceae*, *Bacteroidales*, and *Clostridiales* [19, 88]. Previously, the bovine rumen was estimated to contain over 27,000 carbohydrate-active enzyme (CAZyme) genes [25, 89], a number that has been recently increased to over 69,000 genes identified from assembled genomes [26]. Despite this diversity of sequence and catalytic function, the microbial conversion of plant cell walls is suboptimal with only 10-35% of consumed carbohydrates being converted into meat and milk [14]. In this regard, improving the efficiency of microbial digestion with direct-fed microbials, such as *Bacteroides* spp., and selective feeds may help address the emerging challenges associated with sustainable livestock production [20]. Despite current efforts, our elementary understanding of bacteria-carbohydrate interactions in the rumen has limited the successes of supplementing livestock feed with live microbial adjuncts (*i.e.* probiotics) and selective carbohydrates (*i.e.* prebiotics) to improve digestion of feedstuffs.

The genomes of *Bacteroides* spp. encode dedicated carbohydrate metabolic systems called polysaccharide utilization loci (PULs) [90]. The starch utilization system of *Bacteroides thetaiotaomicron* VPI-5482 (*B. theta* VPI-5482) was the first PUL to be described [91]. Recently, a series of PULs found in *Bacteroides* spp. that colonize diverse ecosystems have been identified that metabolize glycans with diverse chemistries [18]. PULs are distinguished by the presence of a

TonB-dependent transporter coupled to a surface glycan binding protein (SGBP), referred to as the SusC/D-like complex, and other associated proteins that modify or bind the target glycan [53, 59]. Each of these gene products function together in an orchestrated cascade to release and assimilate monosaccharides from the target substrate. PULs can operate through a 'distributive' mechanism, which releases products (*i.e.* 'public goods' [92]) to the community; or a 'selfish' mechanism [59], which limits product loss by concentrating substrate depolymerisation within the periplasm [53]. Recently, PUL-prediction [76] and whole-PUL characterization [58, 59, 93, 94] have become high-priority approaches for the discovery of new CAZyme families and catalytic activities. Genomic and biochemical characterization of distinct PULs has proven indispensable for establishing the molecular basis of unique metabolic abilities, insights that differentiate the saccharolytic potential of *Bacteroides* at the species [94, 95] and strain level [78, 96, 97].

Glycoside hydrolases (GHs) are a class of CAZymes that cleave glycosidic bonds by acid-base catalysis [98]. GHs have been reported that hydrolyse carbohydrates with diverse compositions, degrees of polymerization, linkage chemistries, branching, and rheological properties [99]. Currently, GHs are divided into over 160 sequence related families and are taxonomic clusters that display conserved folds, mechanisms, and catalytic residues. Computational tools designed to improve predictive accuracy in CAZyme families will help to streamline the characterization of enzymes that are essential for dismantling dietary glycans in the rumen [24, 76, 100].

Yeast α-mannan (YM) is a mannose-rich glycoprotein found on the surface of *Saccharomyces cerevisiae* [59] and is synthesized by a collection of glycosyl transferases that act in sequence [101]. YM consists of a Man_8-$GlcNAc_2$ core that is conserved with mammalian high-mannose *N*-glycans (HMNGs). Extending from this core, there is an α-1,6-mannan backbone decorated by short α-1,2 side-chains terminated with α-1,3 mannosyl residues. These side-chains can be further branched through 1,6-mannophosphate esters [101]. The glycan composition and linkage chemistry of mannnoproteins vary between species [77, 102]; for example,

21

Schizosaccharomyces pombe has truncated side-chains and is enriched in α-1,2-galactose. YM and its derivative oligosaccharides (*i.e.* α-mannanoligosaccharides, MOS) have been shown to induce a variety of beneficial physiological effects in livestock [103-105]. YM is a component of several commercial prebiotics, including Bio-Mos® [106], as well as alternative feedstocks, such as distillers grains with solubles (DGS) [107]. DGS is a by-product of bioethanol production from grains (*e.g.* corn) for fuel and beverages. Wet or dried DGS (WDGS or DDGS, respectively) residues can be used directly as alternative livestock feeds [28, 108] or further fractionated to increase nutritional content or recover coproducts [109].

Digestion of YM requires the combined action of several enzymes active on distinct linkages, including α-mannanases, α-mannosidases, and glycosidic phosphatases [77]. These enzymes are commonly classified in GH76, GH92, and GH125 families [59, 73, 110]. Previously, the human intestinal bacterium *B. theta* VPI-5482 was shown to possess three PULs (MAN-PUL1,2,3) that fostered growth on YM and additional α-mannans derived from other yeast species through a selfish mechanism [59]. Comparative growth profiles for a library of Bacteroidetes demonstrated that YM metabolism is dispersed throughout the phylum, but is was most proficient in a collection of *B. theta* strains [59]. Examination of *Bacteroides xylanisolvens* NLAE-zl P352 strains isolated from pigs reared on a diet infused with DGS revealed that the MAN-PUL1 cassette was encoded on a transposable element, suggesting that aspects of YM metabolism can be exchanged between strains [59]. This is consistent with reports that specialized metabolic abilities can be transferred to intestinal *Bacteroides* spp. from species that occupy ecologically distinct habitats [78, 96, 111]. The acquisition of CAZymes or pathways that endow microorganisms with suitable metabolic machinery for unlocking underexploited nutrient streams facilitates their persistence in highly competitive ecosystems.

Surveying microbiome communities at the genus and species levels does not provide a comprehensive assessment of metabolic proficiency. In this regard, coupling biochemical system-

based approaches to microbial genetics can provide more accurate predictions of metabolic abilities. Cross-disciplinary tools, such as fluorescent glycan conjugates (FGCs) [83, 84], which provide 'direct' evidence of glycan-metabolism relationships within individual cells, can illuminate competitive nutrition dynamics at the single-cell level. Here, I have conducted a multi-layered study of the evolution and function of YM metabolism in bovine-adapted *B. theta* strains (*B. theta*Bov) . These strains adopt one of two growth phenotypes, referred to as "High Growers" (HG) and "Medium Growers" (MG). Using genomics, transcriptomics, and CAZyme fingerprinting analyses, I have identified two major MAN-PUL architectures; however, these do not correspond with the HG and MG phenotype. Therefore, I have adapted the use of FGCs to develop a method for quantifying glycan uptake to reveal new insights in transport efficiencies and growth phenotypes.

2.2 Results

2.2.1 Isolation and growth profiling of YM-utilizing *B. theta*Bov strains

To isolate pure bacterial cultures that can utilize YM, rumen and fecal communities were enriched with YM, Bio-Mos®, or DDGS and streaked on selective YM (YM-MM) plates. Single colonies were observed on the YM-MM plates within 24 h and new colonies formed up to 96 h (Figure 2.1A). In total, 50 bacterial strains were collected from the enriched rumen and fecal cultures. Pure cultures of each mannan-degrading (MD) isolate was subjected to multiple rounds of streak plating, assigned a reference number (*e.g.* isolate #8 = MD8), and characterized by 16S analysis. 16S rRNA gene sequences were run through the NCBI BLASTN [112] database to determine their closest relatives. Nine isolates were further studied due to their propensity to grow on YM (OD$_{600}$ > 0.3): MD8 (rumen enriched with Bio-Mos®), MD11 (rumen enriched with YM), MD13 (rumen enriched with YM), MD17 (rumen enriched with YM), MD28 (feces enriched with YM), MD33 (feces enriched with Bio-Mos®), MD35 (feces enriched with Bio-Mos®), MD40 (feces enriched with Bio-Mos®), and MD51 (feces enriched with DDGS). All nine isolates shared the highest homology to *B. theta*; however, there was variation at the strain level (Table 2.1). MD

23

isolates are classified and denoted as bovine-adapted strains of *B. theta*, or *B. theta*Bov. These data are unsurprising, as *Bacteroides* spp. are among the predominant bacterial colonizers of the gut in mammals, and *B. theta* strains (*e.g.* KPPR-3 and NLAE-21-C579) have been previously isolated from bovine rumen and fecal samples [25, 113]; however, characterization of carbohydrate metabolism was minimal.

Table 2.1 Alignment of 16S rRNA gene sequences from mannan-degrading *B. theta* strains isolated from bovine rumen and fecal samples blasted against the NCBI database.

	Top Blast Hit(s)	Query cover (%)	Identity (%)
MD11_{MG}	*Bacteroides thetaiotaomicron* VPI-5482 and strain 7330	100	99.7
MD28_{MG}	*Bacteroides thetaiotaomicron* strain 7330	100	99.7
MD33_{MG}	*Bacteroides thetaiotaomicron* VPI-5482	100	99.8
MD35_{MG}	*Bacteroides thetaiotaomicron* VPI-5482 and strain 7330	100	99.7
MD8_{HG}	*Bacteroides thetaiotaomicron* VPI-5482 and strain 7330	100	99.7
MD13_{HG}	*Bacteroides thetaiotaomicron* VPI-5482 and strain 7330	100	99.7
MD17_{HG}	*Bacteroides thetaiotaomicron* strain 7330	100	99.7
MD40_{HG}	*Bacteroides thetaiotaomicron* strain DMF	100	100
MD51_{HG}	*Bacteroides thetaiotaomicron* strain 7330	100	100

Related bacterial strain(s) with the highest identity are shown.
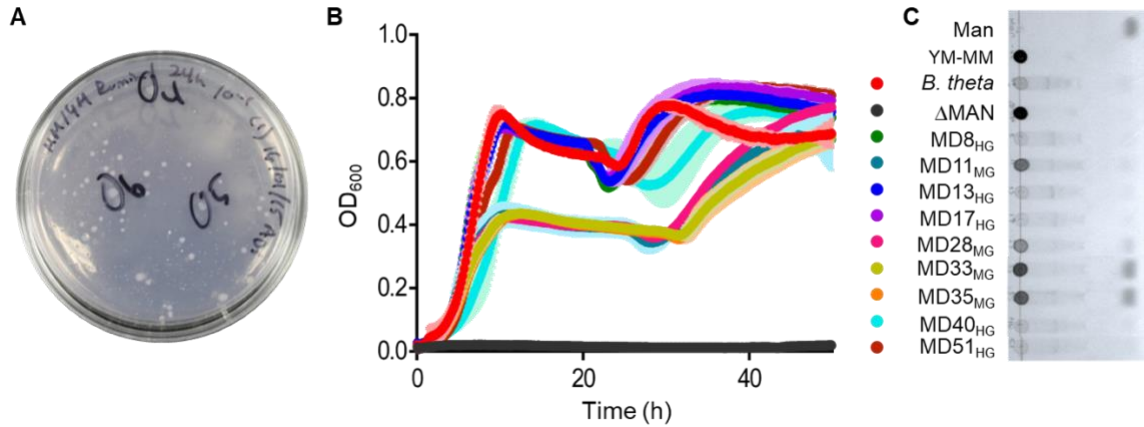
Figure 2.1 YM utilization by *B. theta*Bov isolates.
(A) Colonies of rumen bacteria grown on YM-MM plates. Plate shown represents a 1:10₆ dilution of rumen sample. Figure courtesy of Adam D. Smith. **(B)** Growth profiles of *B. theta* VPI-5482, *B. theta* ΔMAN-PUL1/2/3, and nine rumen isolates grown on 0.5% YM. **(C)** TLC analysis of post-growth supernatants (50 h) of MD isolates grown on 0.5% YM. *B. theta* = strain VPI-5482, ΔMAN = *B. theta* ΔMAN-PUL1/2/3.

The proficiency of bovine isolates to utilize YM was investigated in pure cultures grown in liquid media with YM as the sole carbon source. 41 of the 50 isolates did not display reproducible growth in liquid YM culture and/or were classified as "Low Growers", having a peak $OD_{600} < 0.2$ after 50 h of growth (data not shown). Nine rumen isolates were observed having a peak $OD_{600} \geq 0.4$ (Figure 2.1B). Based upon their growth phenotypes, these nine isolates, along with *B. theta* VPI-5482, were divided into two categories: "Medium Growers" (MG) and "High Growers" (HG). MD11$_{MG}$, MD28$_{MG}$, MD33$_{MG}$, and MD35$_{MG}$ classified as MG (*i.e.* MD##$_{MG}$); whereas, MD8$_{HG}$, MD13$_{HG}$, MD17$_{HG}$, MD40$_{HG}$, MD51$_{HG}$, and *B. theta* VPI-5482 were classified as HG (*i.e.* MD##$_{HG}$). Both phenotypes displayed biphasic growth, which was consistent with previous observations [59]. A decline in OD_{600} followed the first growth peak, lasting ~20 h and ~13 h for MG and HG, respectively, followed by a second growth phase. The initial MG peak reached an OD_{600} ~0.4 at 9 h, with isolates ranging from 0.75-0.8 at ~50 h. The HG strains reached a peak OD_{600} of ~0.7 during the first phase of growth and reached OD_{600} of ~0.8 at 30 h.

Post-growth supernatants were analyzed by thin layer chromatography (TLC) to assess the depolymerisation profile of YM in the media (Figure 2.1C). Pure mannose was used as a control. The YM-MM negative control demonstrates that there are no free mannooligosaccharides or mannose in the substrate, and YM can be visualized as a dense signal at the origin. The *B. theta* MAN-PUL1/2/3 deletion mutant (ΔMAN-PUL1/2/3) did not grow and demonstrated that the knock-out mutant cannot release products into the media. *B. theta* VPI-5482, MD8$_{HG}$, MD13$_{HG}$, MD17$_{HG}$, MD40$_{HG}$, and MD51$_{HG}$ generated similar product profiles, with a noticeable loss of YM staining and faint detection of oligosaccharides and mannose. The MGs displayed a different product profile; each of the strains had residual YM in the post-growth media and the accumulation of mannose was clearly visible for MD28$_{MG}$, MD33$_{MG}$, and MD35$_{HG}$, which suggest there may be differences in how YM is sequentially metabolised by these strains. To determine if these phenotypes could be explained by their genetic content, the genomes of all nine rumen isolates were sequenced.

2.2.2 PUL delineation and comparative CAZomics

Whole genome sequencing and *de novo* assembly was performed to identify genes implicated in YM utilization systems. SPAdes assembly output is shown in Table 2.2. 16S rRNA gene sequencing is limited in its ability to delineate bacterial species and strains [114]. To further resolve species identity, average nucleotide identity (ANI) based on BLAST+ was calculated using the SPAdes assembled contigs and *B. theta* reference genomes in JSpeciesWS (Table 2.3) [115, 116]. The results were similar to the 16S rRNA gene sequence data, as each bovine isolate shared the highest (>97%) identity with multiple strains of *B. theta* (*i.e.* 7330 and VPI-5482). These results confirm that the bovine isolates are indeed strains of *B. theta*.

Reconstruction of the three YM-specific PULs and alignment with *B. theta* VPI-5482 determined a high-level of synteny between all strains (Figure 2.2A). MAN-PUL 2 and 3 are absolutely conserved; MAN-PUL1 is only present in *B. theta* VPI-5482, MD33MG, and MD35MG. The HMNG-PUL, which is not active on YM [59], is present in all nine isolate genomes. Based upon genomic sequence, therefore, PUL architecture and the presence or absence of MAN-PULs do not explain the differences in growth phenotypes.

Table 2.2 SPAdes *de novo* genome assembly output of nine *B. theta*Bov isolates sequenced by Illumina MiSeq PE150bp.

		MD8$_{HG}$	MD11$_{MG}$	MD13$_{HG}$	MD17$_{HG}$	MD28$_{MG}$	MD33$_{MG}$	MD35$_{MG}$	MD40$_{HG}$	MD51$_{HG}$
SPAdes	Number of contigs	920	62	61	59	63	72	70	62	62
	Largest contig size (bp)	868,373	589,645	781,280	928,731	838,230	759,138	723,774	519,222	868,373
	N50	157,167	236,443	229,189	208,464	236,443	150,716	145,695	189,229	229,189

Number of contigs does not include contigs ≤ 1000 bp.

Table 2.3 ANIb output of _B. theta_Bov assembled contigs blasted to characterized _B. theta_ strains from the JSpeciesWS genome reference database.

	MD11$_{MG}$	MD28$_{MG}$	MD33$_{MG}$	MD35$_{MG}$	MD8$_{HG}$	MD13$_{HG}$	MD17$_{HG}$	MD40$_{HG}$	MD51$_{HG}$
B. theta strain	VPI-5482	VPI-5482	KLE1254	KLE1254	7330	7330	7330	7330	7330
ANIb (%)	98.03	98.05	99.16	99.15	98.32	98.33	98.33	98.34	98.33
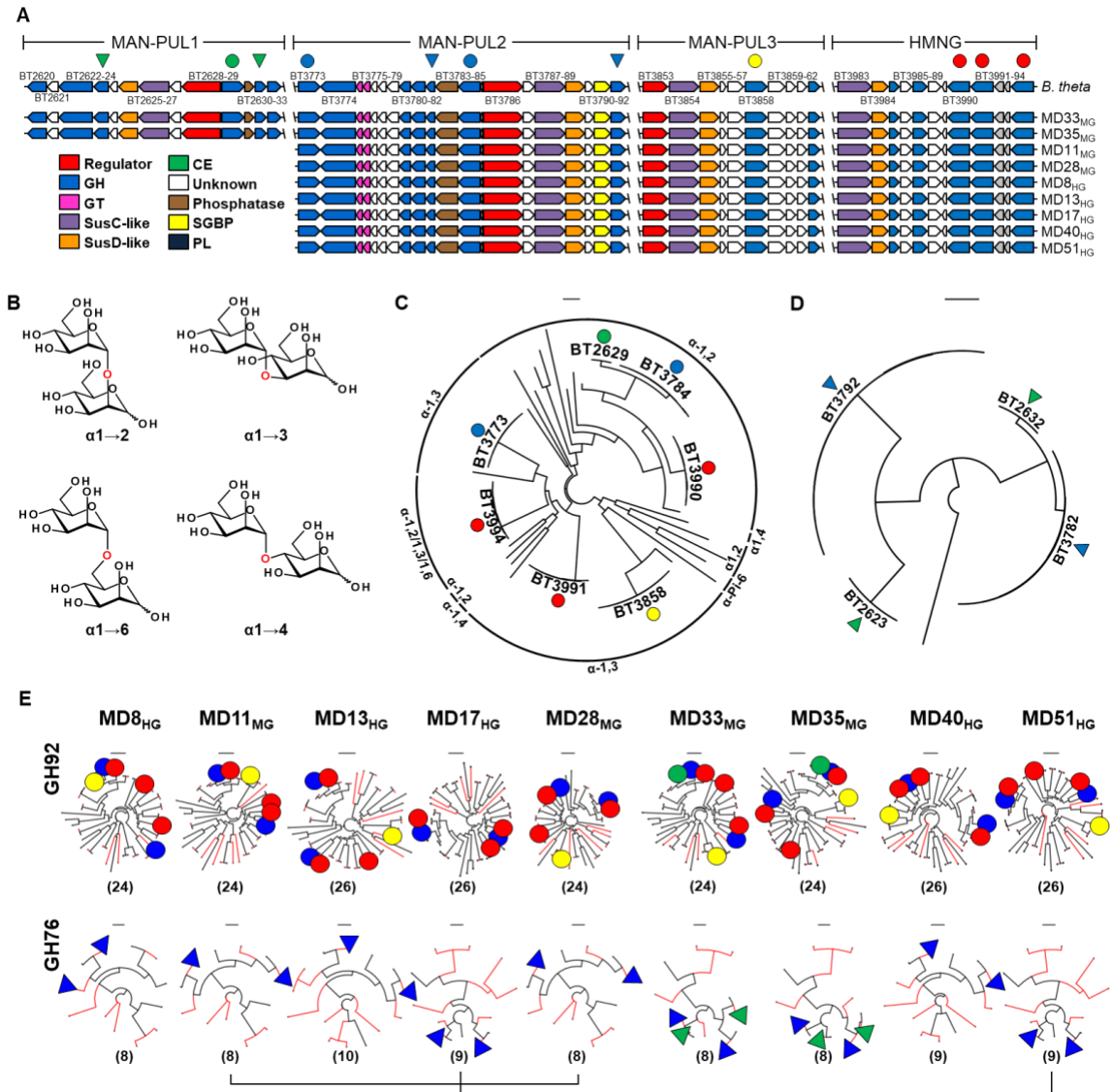
Figure 2.2 GH76 and GH92 profiles across *B. theta*Bov isolates.
(**A**) Synteny of MAN-PUL1,2,3 and HMNG-PUL of *B. theta* and nine *B. theta*Bov strains. Length of genes are to scale and arrows indicate directionality. GH76s are indicated with coloured triangles, whereas GH92s are indicated with coloured circles. MAN-PUL1 = Green, MAN-PUL2 = Blue, MAN-PUL3 = Yellow, and the HMNG-PUL = Red. (**B**) Mannose linkages within YM targeted by GH76 or GH92 enzymes. The red oxygen represents the cleaved glycosidic bond. Phylogenetic trees of all GH92 (coloured circles) (**C**) and GH76 (coloured triangles). (**D**) Using SACCHARIS [24], GH76 and GH92 sequences from bovine isolate MAN-PULs are embedded into trees of characterized sequences extracted from CAZy [67, 68]. Activities assigned to characterized enzymes within each clade for GH92 are listed; GH76 has only been described to possess endo-α-1,6-mannanase activity. (**E**) GH92 (top) and GH76 (bottom) fingerprints for each bovine isolate. Black lines represent characterized GH enzymes published in the CAZy database; red lines represent each family member identified in the strain. Bracketed numbers indicate the total number of enzymes within each strain. Identical trees are connected with lines. Trees courtesy of Jeffrey Tingley.

To determine if there were differences in the amino acid sequence and evolution of homologous enzymes, GH76 and GH92, two GH families known to be involved in YM depolymerisation (Figure 2.2B; [59]), were analyzed by SACCHARIS (Figure 2.2C & D) [24]. For both GH76 and GH92, every sequence within MAN-PUL1, 2, and 3 displayed the highest level of amino acid sequence conservation with its syntenic homolog. Thus, each PUL is under selective pressure to function as an intact system. To investigate if this is relationship is conserved for other α-mannan degrading PULs, I used a genome-wide approach, previously referred to as CAZome fingerprinting [24], which predicts the entire range of catalytic activities encoded within a family or genome. As shown in Figure 2.2E, each bovine isolate encodes between 24 and 26 GH92s and 8 or 9 GH76s within its genome. When total enzymes from each family were embedded into phylogenetic trees comprised of all characterized enzymes within each GH92 and GH76 family, there were differences between the strains (Figure 2.2E). Only MD11MG and MD28MG, and MD17HG and MD51HG displayed identical phylogenetic trees for GH76; whereas, every GH92 tree was unique. Notably, these patterns are the result of sequence divergence, and total enzyme numbers (Figure 2.2E). Similar differences were observed for other α-mannan active enzyme families (GH38, GH99, and GH125; data not shown), suggesting that despite the high-level of conservation within the MAN-PUL proteins, there may be metabolic specialization in α-mannan consumption that exists between these strains.

2.2.3 RNA-seq: assembly, quantitation, and comparative analysis

RNA-seq was performed on *B. theta* VPI-5482, MD33MG, and MD40HG cells cultured on both mannose and YM to determine if differential patterns in MAN-PUL expression could provide insights into the MG and HG growth phenotypes. Log2 expression ratios were calculated using DESeq2 [117] to normalize YM transcript expression to mannose. MAN-PUL2 and 3 displayed increased expression in all three bacteria and MAN-PUL1 similarly increased in *B. theta* VPI-5482 and MD33MG (Figure 2.3A); this is consistent with what was previously reported for *B. theta* VPI-

5482 [95]. The MAN-PUL *gh*, *susC*-like, and *susD*-like genes displayed the highest $\log_2$ expression values across the *B. theta* VPI-5482, MD33MG, and MD40HG transcriptomes (Figure 2.3A & B), which is consistent with previous PUL expression data [118]. Within the MAN-PULs, highest levels of expression ($\log_2$ fold-change) were observed in the MAN-PUL3 *susC*-like (5.4) and *susD*-like (5.5) genes in *B. theta* VPI-5482; BT_2629 (GH92; 7.7) and BT_3792 (GH76; 7.7) homologs in MD33MG; and the BT_3784 (GH92; 7.7) homolog in MD40HG (Figure 2.3B). Statistical analyses of the MAN-PUL1,2,3 *susC*-like TPM values showed a significant difference ($p < 0.05$) between the YM expression of *B. theta* VPI-5482 and the two *B. theta*Bov strains, with no statistical difference ($p > 0.05$) of expression of the same genes in the mannose treatment. The MD33MG transcriptome was noticeably more dynamic than the other two strains, having higher numbers of genes with increased or decreased expression (Figure 2.3A). As a HG, *B. theta* VPI-5482 was predicted to have a similar expression profile as MD40HG. Somewhat surprisingly, *B. theta* VPI-5482 MAN-PUL expression was at least a 2 $\log_2$ fold-change lower relative to MD33MG and MD40HG (Figure 2.3B).

In addition to activation of the MAN-PULs, each transcriptome had regions that were repressed (Figure 2.3A). *B. theta* VPI-5482 displayed down-regulation ($\log_2$ fold-change < -3.5) of many genes, including PUL22, which is predicted to target dietary fructans (BT1758-1763) [94, 119]. Similarly, MD33MG and MD40HG repressed a homologous cluster of fructan-degrading genes when grown on YM ($\log_2$ fold-change < -3.5). This suggests that there may be a YM-catabolite repression effect between these two polysaccharides. Interestingly, this effect was not conserved when the cultures were grown on mannose. In MAN-PUL2, BT_3775-BT_3776, which encodes a biosynthetic gene cluster, displayed the lowest expression levels of all the genes. This is consistent with what was reported previously (Figure 2.3B) [95].

PUL regulators are typically expressed at basal levels and regulated differently than enzymes and transporters [95, 120]. This was observed for BT_3853 in all three bacterial strains, which encodes the Sarp/OmpR regulator gene from MAN-PUL3 ($-0.2 < \log_2$ fold-change > 0.3; p

> 0.05). In contrast, BT_2628, the HTCS from MAN-PUL1; and BT_3786, the HTCS from MAN-PUL2 were both moderately upregulated ($\log_2$ fold-change > 2.6; Figure 2.3B). However, induction of regulator gene expression is not uncommon; previously, the PUL43 HTCS was also reported to be activated when induced by oat-spelt xylan relative to xylose ($\log_2$ fold-change = 3.30) [118].
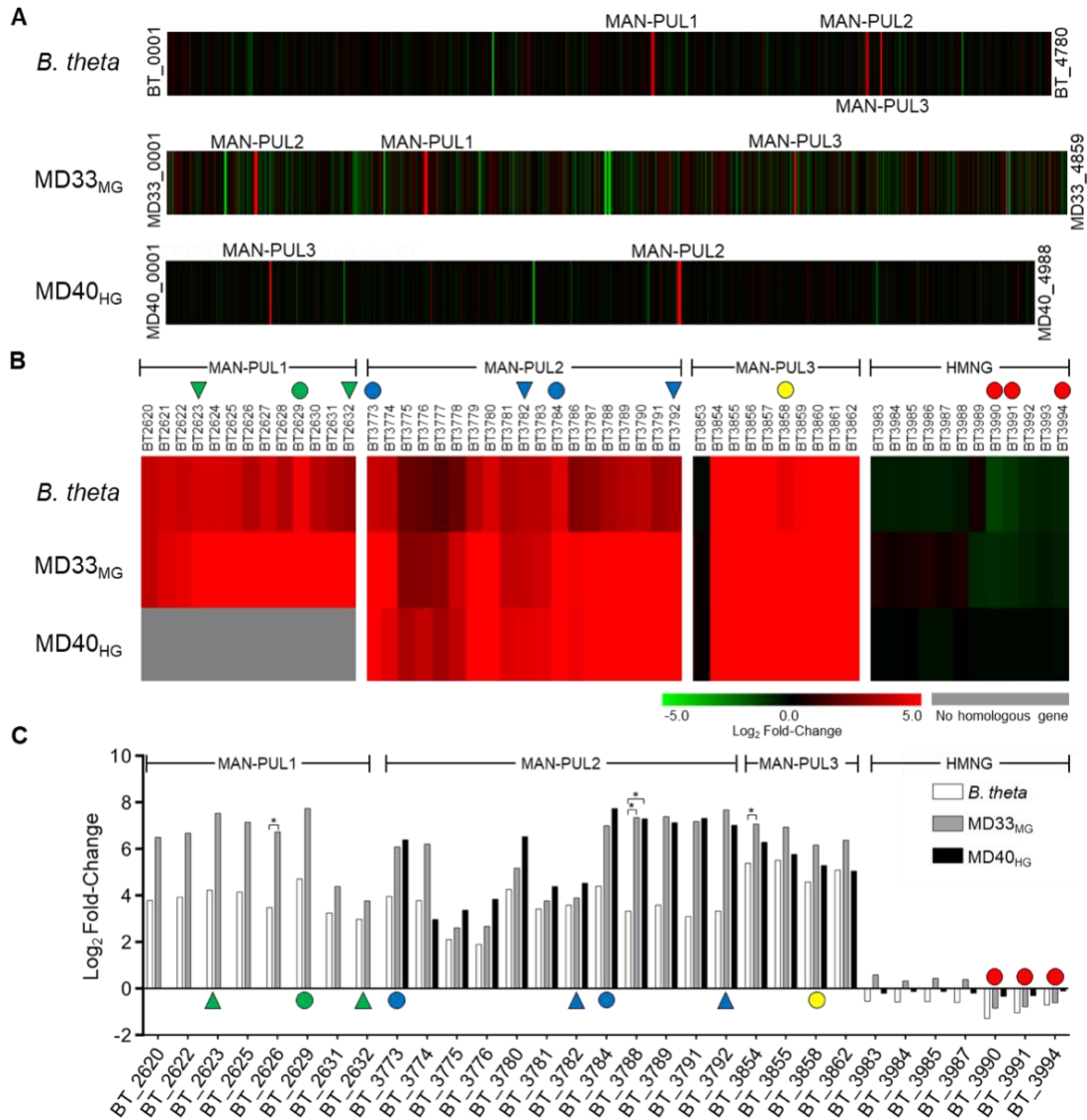
Figure 2.3 RNA-Seq analysis of B. theta VPI-5482, MD33MG, and MD40HG cultured in YM.

(A) Log$_2$ expression ratios of each transcript of *B. theta* VPI-5482, MD33$_{MG}$, and MD40$_{HG}$. Length of transcriptomes is indicative of genome size (*B. theta* VPI-5482 = 6.26 Mb, MD33$_{MG}$ = 6.28 Mb, MD40$_{HG}$ = 6.16 Mb). **(B)** Expression levels of MAN-PUL1,2,3 and HMNG-PUL genes. Red indicates an increase in transcript expression compared to the control, while green is a decreased expression; grey represents gene transcripts not present in the corresponding genome. **(C)** Bar graph represents log$_2$ fold-change of CAZymes and SusC/D/E-like proteins in MAN-PUL1,2,3 and HMNG-PUL encoded in the genomes of *B. theta* VPI-5482 (white), MD33$_{MG}$ (grey), and MD40$_{HG}$ (black). Log$_2$ fold-change compares gene expression of YM cultures normalized to mannose, N=3, p≤0.01; except BT3853 and HMNG-PUL gene expression is not significantly (p > 0.05) different between the two treatments. *Significantly different (p < 0.05) TPM expression of the *susC*-like genes between *B. theta* VPI-5482 and *B. theta*$_{Bov}$ strains, courtesy of Timothy Schwinghamer; statistical comparison of other MAN-PUL genes between strains were not calculated.

2.2.4 Relatedness of transporters in MG and HG strains of *B. theta*Bov

The similar syntenic structures and gene sequences with YM-specific PULs, and the similar patterns in MAN-PUL expression between strains show that the phenotypic growth differences do not arise from MAN-PUL architecture or expression. To further investigate the cause of the two MG and HG phenotypes, differences in transport efficiency were assessed. The sequences from the MAN-PUL2 SusC-like (BT3788) and SusD-like (BT3789) homologs from each strain were aligned (Figure 2.4A). The SusCD-like trees displayed a different clade pattern than the 16S rRNA tree (Figure 2.4A-C). In the 16S rRNA gene tree, there is no apparent relationship between sequence relatedness and growth phenotypes. The closest relative to *B. theta* VPI-5482 (a HG) is MD33MG, and the sequences from the HG and MG strains are distributed stochastically throughout the tree. In contrast, for the SusC-like and SusD-like amino acid alignments, there is a clear partitioning that relates to growth phenotype (Figure 2.4B & C). HGs share 100% sequence identity (Table 2.4 shows SusC-like alignments; SusD-like data not shown) and form a distinct clade in the phylogenetic trees produced using amino acid sequences; MGs similarly cluster with each other, with the exception of one amino acid change in the SusD-like sequence. This relationship suggests that changes in SusC-like or SusD-like primary structure, a complex that is responsible for transport of YM-products into the periplasm, may contribute to the differential growth phenotypes. To visualize the amino acid differences, a theoretical model of the MD33MG SusCD-like complex was predicted based on a previously published *B. theta* structure [121] (Figure 2.4D & E). Colour differentiation was used to highlight the amino acid changes between the HG and MG proteins, and displayed multiple sites of amino acid differences distributed throughout the proteins. Amino acid changes could affect structural and functional variability of the SusCD-like complex between the MG and HGs, resulting in different uptake efficiencies of the surface YM products. Therefore, I decided to investigate if of FGC accumulation is different between the *B. theta* VPI-5482, MD33MG, and MD40HG strains.
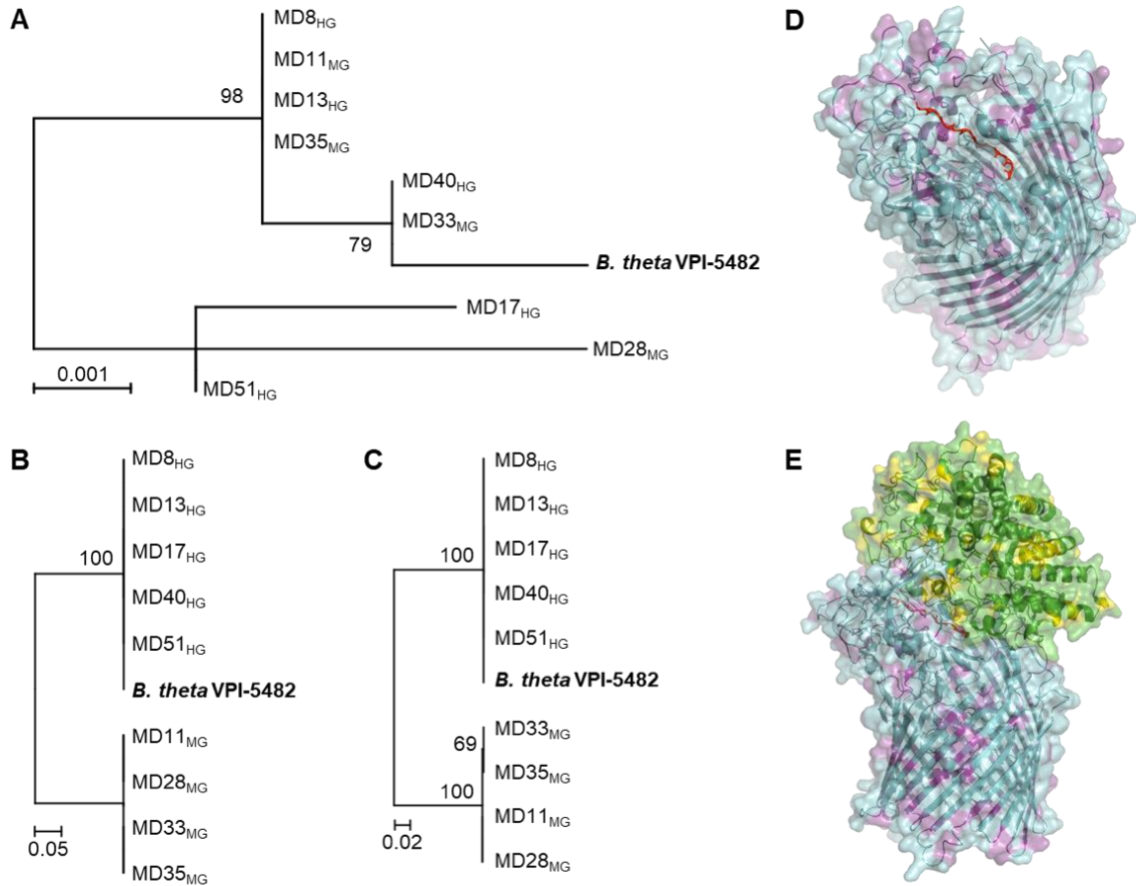
Figure 2.4 Sequence comparisons between *B. theta* and bovine isolates.
(A) 16S rRNA gene sequence comparison of *B. theta* VPI-5482 and *B. theta*Bov strains. Scale represents number of nucleotide changes across horizontal axis. Amino acid sequence relatedness of the **(B)** SusC-like and **(C)** SusD-like proteins encoded in MAN-PUL2. Bootstrap values indicated at branch points. Scale represents number of amino acid substitutions. Trees courtesy of Jeffrey Tingley. **(D)** Homology model of MD33MG SusC-like protein and **(E)** SusC/D-like pair (modelled to PDB: 5FQ6 [121]) displaying cartoon and surface of SusC-like transporter in cyan, with residues that are divergent between MG and HGs highlighted in magenta; MG SusD-like protein shown in green with amino acid differences highlighted in yellow. Peptide ligand from the reference SusCD complex displayed as red sticks. Models courtesy of Dr. Darryl R. Jones.

Table 2.4 Percent identity matrix of the MAN-PUL2 SusC-like amino acid sequences from isolated *B. theta*Bov strains generated by MUSCLE [76].

	MD11MG	MD28MG	MD33MG	MD35MG	MD8HG	MD13HG	MD17HG	MD40HG	MD51HG	*B. theta*
MD11MG	100	100	100	100	77	77	77	77	77	77
MD28MG	100	100	100	100	77	77	77	77	77	77
MD33MG	100	100	100	100	77	77	77	77	77	77
MD35MG	100	100	100	100	77	77	77	77	77	77
MD8HG	77	77	77	77	100	100	100	100	100	100
MD13HG	77	77	77	77	100	100	100	100	100	100
MD17HG	77	77	77	77	100	100	100	100	100	100
MD40HG	77	77	77	77	100	100	100	100	100	100
MD51HG	77	77	77	77	100	100	100	100	100	100
***B. theta*[1]**	77	77	77	77	100	100	100	100	100	100

[1]*B. theta* strain VPI-5482.

2.2.5 Uptake of fluorescent glycan conjugates by *B. theta*Bov strains

FGCs have been used to study interactions with marine bacteria [84] and the human intestinal bacterium *B. theta* VPI-5482 [122]. To determine if FGCs can be used to study the interactions between YM and bovine-adapted strains of *B. theta*. MD33MG and MD40HG were incubated with 6-aminofluorescein-YM conjugates (FLA-YM) or unlabelled YM (YM-MM) for 72 h. Following incubations with the probes, cells were stained with DAPI and Nile Red to visualize cellular DNA and membranes. Subsequently, samples were imaged by super-resolution structured illumination microscopy (SR-SIM). Cells cultured in YM-MM or cells sampled at 0 h in FLA-YM did not show an increase in cellular fluorescence (Figure 2.5). FLA-YM signal accumulated within MD33MG and MD40HG cells by 24 h and maintained fluorescence after 72 h incubation (Figure 2.5). FLA-YM signal was at the periphery of the cells and co-localized with the membrane. Heterogeneity of FLA-YM uptake within populations was observed, suggesting there is heterogeneity in cellular responses between cells. These results are consistent with previous results of FLA-YM accumulation in *B. theta* VPI-5482 [122]. In this regard, FGCs can be used to study differential metabolic responses in clonal populations.
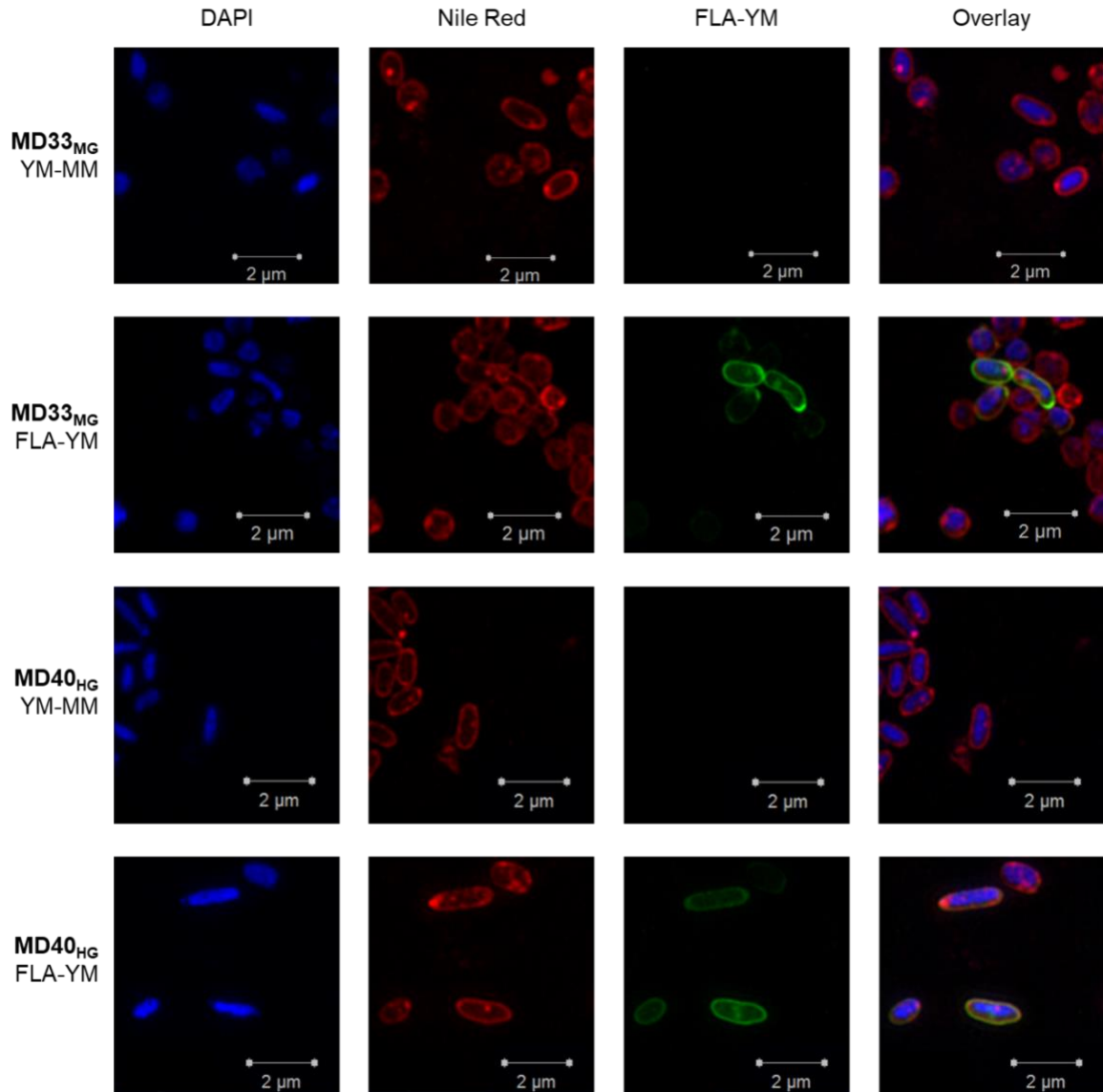
Figure 2.5 Single-cell visualization of FLA-YM uptake in MD33$_{MG}$ and MD40$_{HG}$.
Cells counter-stained to localize DNA (DAPI) and membranes (Nile Red) within individual cells. FLA-YM accumulation was observed heterogeneously in *B. theta*$_{Bov}$ strains incubated with FLA-YM. Time point taken 72 h post-incubation with unlabelled YM (YM-MM) or FLA-YM. Samples visualized by SR-SIM. Images courtesy of Dr. Greta Reintjes.

2.2.6 Rate of FLA-YM uptake by *B. theta*Bov isolates

To determine if uptake can be differentiated between the two different growth phenotypes, samples of *B. theta* VPI-5482, MD33MG and MD40HG were collected at 5 min, 10 min, 15 min, 20 min, 30 min, 1 h, 2 h, 4 h, and 24 h. Cellular accumulation and uptake of FLA-YM was determined using flow cytometry (10,000 events measured per sample). *B. theta* VPI-5482, MD33MG, and MD40HG cells start to uptake FLA-YM within 5 min of incubation (Figure 2.6A); however, the three strains display differences in FLA-YM accumulation (Figure 2.6B) and rate of uptake (Figure 2.6C). The fluorescence intensity measures the accumulation of FLA-YM in each individual cell in a sample (Figure 2.6B). Fluorescence of the three strains increases from 0 to 120 min, peaks at 120 min, and declines from 120 min to 1440 min (Figure 2.6B). MD33MG had the lowest fluorescence intensity values at each time point, with a peak value of 5011. MD40HG displayed the highest fluorescence intensity, peaking at 13,581. *B. theta* VPI-5482 showed a fluorescence intensity between the two bovine strains for each time point, with a peak value of 9,786. Large standard deviations in each sample were observed due to heterogeneity of fluorescence intensity between individual cells (Figure 2.6B). The percent of cells displaying uptake of FLA-YM showed a similar pattern to single-cell accumulation, with more MD40HG cells showing uptake more rapidly (22% at 5 min), *B. theta* VPI-5482 cells having an intermediate rate of uptake (9% at 5 min), and MD33MG cells showing a slower rate (2% at 5 min) (Figure 2.6C). A difference in FLA-YM uptake was observed between the three strains, with HG displaying a higher fluorescence intensity and percentage of fluorescent cells than the MG at each time point.
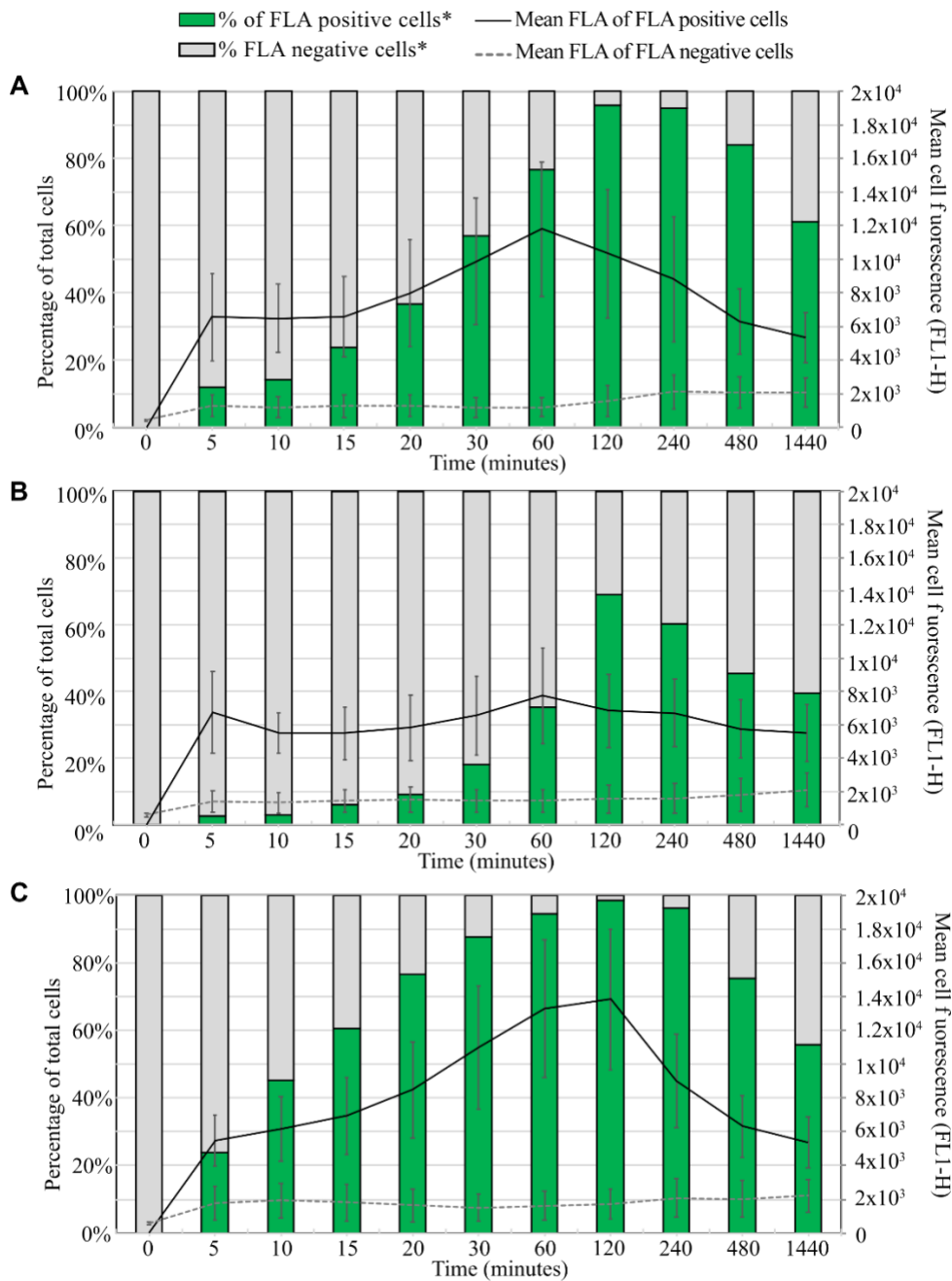
Figure 2.6 Quantification of FLA-YM uptake over time using flow cytometry.
Bar graph represents percentage of total **(A)** *B. theta* VPI-5482, **(B)** MD33MG, and **(C)** MD40HG
cells incubating with FLA-YM for 0 to 1440 min. Green = FLA-positive cells, grey = FLA-
negative cells. Line graph shows the mean cell fluorescence of FLA-positive (solid line) and
FLA-negative (broken line) cells. Standard deviation bars shown. N = 10,000. Flow cytometry
data courtesy of Dr. Greta Reintjes.

2.3 Discussion

The rumen houses a dense and diverse microbial community [86] consisting of a bacterial microbiome [87] that specializes in the saccharification of plant and fungal polysaccharides. Despite sequence diversity and catalytic potential of over 69,000 predicted CAZymes encoded in the rumen microbiome [26], bacterial utilization of dietary glycans is suboptimal [14]. In this regard, improving the efficiency of microbial digestion through direct-fed microbials (*i.e.* probiotics) and selective feed additives (*i.e.* prebiotics) may help address the emerging challenges associated with sustainable livestock production [20].

Effects of probiotic and prebiotic supplements on live weight gain and feed efficiency in ruminants have been met with inconclusive results. Research trials tend to overlook mechanisms of action through which probiotics and prebiotics act, often focusing instead on health and productivity measures, such as ruminal pH, live weight gain, and volatile fatty acid production [56, 108, 123, 124]. Very little empirical data is available for interpreting the beneficial outcomes of these dietary interventions for livestock. Advanced methods, such as culturomics, genomics, transcriptomics, and bioinformatics will be useful for developing evidence-based probiotics and prebiotics for beef production.

To discover bacteria within the rumen with a propensity for digesting YM, rumen samples were enriched with YM and incubated anaerobically. Bacterial isolates were selected for and isolated from these communities by dilution plating on nutrient restrictive, YM-containing media (Figure 2.1A). Nine of the 50 defined mannan-degrading isolates showed a higher degree of YM metabolic potential based on growth analysis ($OD_{600} > 0.4$). Interestingly, two definitive growth phenotypes were observed, referred to here as MG (Medium Grower) and HG (High Grower) (Fig. 2.1B). Analysis of the spent supernatants of cultures grown on YM (Fig.2.1C) revealed that HGs, including *B. theta* VPI-5482, depleted more of the polymeric substrate, and released very little oligosaccharides into the media. This profile is consistent with selfish metabolism that was previously reported for *B. theta* VPI-5482 [59]. Sharing of liberated carbohydrate products, or

"public goods"[125], within a community has been referred to as distributive metabolism [53, 77]. Previously, *Bifidobacterium adolescentis*, a bacterium that cannot metabolize xylan, proliferated when cultured with *Bacteroides ovatus*, a xylanolytic bacterium. It was determined that *B. adolescentis* grew on simple xylan products released into the media [97]. A growing body of evidence suggests that this feeding strategy is dependent on the complexity of the glycan substrate [97]. In Bacteroidetes, structurally complex glycans appear to be transported into their periplasm where they are saccharified by a collection of enzymes tailored for the recognition and cleavage of substrate specific linkages and chemical modifications [58, 59, 84]. In contrast, the MG strains consumed less YM substrate and released mannose into the media (Figure 2.1C), suggesting they are less efficient at converting YM into energy and they may foster the growth of other bacteria on mannose derived from YM in a community.

To determine if these growth phenotypes resulted from the acquisition of enzymes with different functions or PULs with different architectures, the genomes of four MG (MD11$_{MG}$, MD28$_{MG}$, MD33$_{MG}$, and MD35$_{MG}$) and five HG (MD8$_{HG}$, MD13$_{HG}$, MD17$_{HG}$, MD40$_{HG}$, and MD51$_{HG}$) strains were sequenced and assembled. Alignment of the MAN-PULs from each genome revealed a high-level of synteny between the nine *B. theta*$_{Bov}$ strains and *B. theta* VPI-5482 (Figure 2.2A). MAN-PUL2, MAN-PUL3 and HMNG-PUL were completely conserved in gene content and pathway structure between all the isolates (Figure 2.2A). Somewhat surprisingly, MAN-PUL1 was only observed in *B. theta* VPI-5482, MD33$_{MG}$, and MD35$_{MG}$. It was anticipated that the acquisition of an additional PUL (*i.e.* MAN-PUL1) may provide a fitness advantage associated with the HG phenotype. However, the observation that MAN-PUL1 is found in *B. theta* VPI-5482 (a HG) and in only two of the four MG strains, suggests that the enzymes (BT2620, BT2622, BT2623, BT2629, BT2630, BT2631 and BT2632) and transporter (BT2626) encoded with this pathway do not catalyze rate-limiting reactions or transport processes during YM metabolism. Furthermore, MAN-PUL2 was previously determined to be the cardinal PUL for YM metabolism [59]. *B. theta* VPI-5482 MAN-PUL2 deletion mutants had a complete loss of growth phenotype;

in contrast, deletion of both MAN-PUL1 and 3 produced only a delayed growth profile [59]. YM metabolism is widespread in human gut Bacteroidetes, which was attributed to the consumption of domesticated yeast [77]. Growth proficiency of these bacteria was also reported to vary between species and strains, which may have resulted from multiple gene transfer events between gut *Bacteroides* spp.. The data presented here, however, suggests that differential growth phenotypes and modes of feeding (*i.e.* selfish vs. distributive), cannot be accurately predicted solely from the analysis of PUL structure.

The sequential roles of enzymes within a pathway and the targeting of enzymes to different compartments of the cell can have significant consequences for how substrates are modified. In this regard, "keystone" [126] or "vanguard" [93] enzymes have been described as foundational enzymes for providing access to a complex glycan substrate. Keystone enzymes tend to be surface exposed, endo-acting enzymes that cleave large, complex glycans into products that can be transported into the cell. For YM digestion, BT2623 and BT3792 have been identified as surface exposed endo-α-mannanases (GH76s) required for the generation of YM-fragments that can be transported. Without these activities, YM would be refractory to transport, and the cell would be unable to use YM subunits as an energy source. Sequence divergence of the cell surface enzymes could change their activity and the depolymerization profile of polysaccharides outside the cell. Thus, methods to improve the accuracy of functional prediction of CAZymes are extremely useful for helping to decode large sequence datasets. In this regard, clustering family members into sub-families has been shown to provide insight into sequence-function relationships [69-72]. Similarly, Jones *et al.* has developed an automated pipeline that embeds unknown user sequences into phylogenetic trees built from biochemically characterized family sequences [24]. This pipeline, called "SACCHARIS", can aid in the functional prediction of uncharacterized sequences from genomic and transcriptomic datasets. In this study, SACCHARIS was used to phylogenetically compare rumen *B. theta*Bov and previously characterized α-mannan-active GH sequences.

To determine if sequence heterogeneity within MAN-PUL GHs could be responsible for influencing the sequence or rate of YM modification, GH92 and GH76 sequences were extracted from all of the MG and HG genomes. GH92 is a polyspecific family known to be active on diverse α-mannosyl linkages [73]; GH76 is a family of endo-α-mannanases (Figure 2.2B) [24, 59]. When aligned with the characterized CAZy sequences from each family, the MAN-PUL sequences from the rumen isolates display strict conservation with the *B. theta* VPI-5482 GH92s (Figure 2.2C) and GH76s (Figure 2.2D), including the keystone GH76s BT2623 and BT3792. This pattern is not conserved across the entire genome, however, as each strain displayed a unique pair of GH92/GH76 CAZyme fingerprints (Figure 2.2E). Unique profiles were also observed for other α-mannosyl cleaving GH families (i.e. GH38, GH99, and GH125; data not shown), suggesting that although there is strict conservation of YM metabolism encoded within MAN-PUL1,2,3, there is strain-level specialization in the consumption of other fungal or host α-mannan substrates. In this regard, CAZyme fingerprinting is a valuable approach for identifying enzyme additions and deletions that exist in related strains; information that can streamline biochemical characterization.

Following the observation that GH sequences encoded within MAN-PULs were conserved, I reasoned that the MG and HG growth phenotype may result from differential regulation of the MAN-PULs. Increases in the transcript levels of SusC/D-like or keystone enzymes (*e.g.* GH76), may improve the conversion efficiency of YM into mannose. Differences in transcript levels could result from substrate recognition, promoter structure, HTCS affinity, or catabolite effects [119, 127-129]. MAN-PULs are described as promiscuous PULs because they can be activated by structurally diverse α-mannans [59]. Phenotypic growth differences could be explained by nutrient specialization within *B. theta*Bov strains. Decreased MAN-PUL transcript expression in MG isolates would be indicative of PULs that are less sensitive to YM from *S. cerevisiae*; whole transcriptomes of HG and MG strains were sequenced and analyzed to test this hypothesis.

As expected, the MAN-PUL1,2,3 in *B. theta* VPI-5482 and MD33MG, and MAN-PUL2,3 in MD40HG; were activated when grown on YM (Figure 2.3A & B) with similar expression patterns

for GHs and transporters (Figure 2.3C). *B. theta* VPI-5482 is model bacterium, well adapted for use in laboratory environments. In this regard, along with the a 2-fold decrease in transcript expression of MAN-PUL2 proteins relative to MD33MG and MD40HG (Figure 2.3C), suggests that *B. theta*Bov strains, which were selectively isolated solely for their ability to degrade YM, may be better adapted for metabolism of α-mannans. However, the significantly ($p < 0.05$) lower transcript expression of *B. theta* VPI-5482 MAN-PUL genes, specifically the *susC*-like genes relative to MD33MG, does not correlate with the differential growth phenotypes. A closer look at the individual transcripts revealed that every MAN-PUL gene is activated by YM, except for the MAN-PUL3 HTCS, BT_3853 (Figure 2.3B). This is a common observation, as regulator genes tend to be independently regulated at basal levels [95, 120]. Moderate HTCS transcript activation was observed for MAN-PUL1 and 2, however, a result that is consistent with what was described for BXY_29350 [118] and BT_3172 [127]. Although the impacts of HTCS induction on PUL function are not known, it may be important for regulating or augmenting the expression of keystone enzymes or transport machinery.

Previously, it was shown that the amino acid homology of a SusD-like protein involved in fructan utilization is relatively low between two strains of *B. theta*, despite few taxonomic differences seen in their 16S rRNA taxonomy [130]. This suggested that PUL genes can evolve independently, which was further supported in the sequence comparisons of the *B. theta*Bov isolates. The clear divergence of the MAN-PUL2 SusC-like and SusD-like amino acid sequences between HG and MG isolates did not reflect 16S phylogeny (Figure 2.4A-C). In addition, this divergence was not seen in the MAN-PUL3 SusC-like and SusD-like protiens, which showed 99% identity shared between all nine isolates. Thus, differences in the MAN-PUL2 transporter may play a role in the differential growth phenotypes of the bovine isolates. Alignment of the primary structure of the HG and MG proteins determined that there are regions of highly conserved sequence punctuated by regions of heterogeneity (data not shown). A theoretical model of the MAN-PUL2 SusC-like and SusD-like proteins displays the sections of amino acid divergence between the HG and MG

sequences (Figure 2.4D & E). Modifications to protein structure could result in functional variability between the HG and MG SusC/D-like complexes. Differences in transport proteins may prevent sufficient import of GH76 end-products, which would impact the rate of depolymerization into small oligosaccharides and mannose. This is supported by the pronounced mannose signal observed in the MG relative to HG supernatants (Figure 2.1C). *B. theta* VPI-5482 displays a similar growth profile (Fig. 2.1B) and feeding strategy (Figure 2.1C) to the HG group of bovine-adapted *B. theta*Bov isolates, while the MG phenotype adopts a distributive, less efficient feeding strategy (Figure 2.1C). RNA-seq analysis displayed similar patterns of MAN-PUL expression between *B. theta* VPI-5482, MD33MG, and MD40HG, with few outliers, suggesting differential expression of MAN-PUL proteins is not the cause of the growth phenoypes. Comparative sequence analysis highlighted amino acid differences between the MAN-PUL2 SusC- and SusD-like proteins in the MG and HGs that may contribute to less efficient import of depolymerized YM, which was further investigated using FGCs.

Previously, FGCs, paired with SR-SIM, have been used to study interactions between glycans and marine bacteria from the phyla Bacteroidetes and Planctomycetes [84], and more recently, between YM and RGII FGCs and *B. theta* VPI-5482 [122]. Specifically, FLA-YM and FLA-RGII uptake into the periplasmic space of *B. theta* VPI-5482 was visualized. Therefore, to determine if FGCs could be used in a similar manner to study glycan uptake in rumen bacteria, MD33MG and MD40HG were incubated with FLA-YM. SR-SIM analysis determined that both strains contained FLA-YM signal (Figure 2.5), similar to what was observed by Hehemann *et al.* [122]. Importantly, the FLA co-localized with the Nile Red membrane stain, suggesting that the probe accumulates in the periplasmic space.

The observation that interactions of FLA-YM with HG and MG strains can be visualized (Figure 2.5) suggested that FGCs could be used to study rates of uptake. Therefore, I decided to conduct a time course analysis on MD33MG and MD40HG FLA-YM uptake. Following activation on YM, MD40HG displayed a more rapid uptake and accumulation of FLA-YM, with the highest

48

fluorescence intensity (2,683) and percentage of stained cells (22%) relative to the other strains after 5 min; whereas, MD33$_{MG}$, required 30 min to reach a similar level (2,203) of intensity and fluorescent cell count (16%) (Figure 2.5B & C). These results suggest that transport of YM products is rate-limiting and supports the hypothesis that uptake efficiencies may be responsible for differential growth phenotypes. Potentially, transport inefficiencies could be attributed to SusC/D-like effects, as their sequences are differentially conserved between the MAN-PUL2 transport proteins encoded in the *B. theta*$_{Bov}$ MG and HG strains.

Based upon the fructan PUL and the MAN-PULs, an emerging pattern suggests that the selection pressures governing carbohydrate metabolism operate independently of taxonomy. The acquisition of PULs, or combinations of PULs (Fig. 2.2A), within *Bacteroides* spp. has been hypothesized to occur by horizontal gene transfer and is linked to spatial and cultural dietary habits [78, 111]. Functional diversification within Bacteroidetes, therefore, is consistent with the "Nutrient Niche Hypothesis" proposed by Freter et al. [45], which suggests that metabolic abilities are determined by the creation and filling of ecological nutrient niches. In this regard, the rumen represents a complex and dynamic system to study the evolution of carbohydrate metabolism shaped by responses to dietary regimes.

2.4 Conclusion

This study presents a combination of omic-based and fluorescent glycan methods to study differential YM growth phenotypes displayed by two groups of rumen *B. theta*$_{Bov}$ isolates. The HG phenotype is associated with higher growth densities, selfish metabolism, and faster YM uptake kinetics; whereas, the MG phenotype displays an early growth plateau, a release of mannose into the media, and slower YM uptake kinetics. Somewhat surprisingly, these physiological properties and feeding strategies cannot be explained by PUL architecture, CAZyme sequence, or transcript expression levels. Analysis of the proteins involved in YM transport revealed there is sequence heterogeneity in both the SusC-like and SusD-like proteins of MAN-PUL2; differences which may

affect uptake efficiencies. This possibility was supported by the use of FGCs to visualize YM interactions and quantify its transport kinetics. Importantly, this is the first time FGCs have been used to visualize uptake of glycans by rumen bacteria at the single-cell level and to study uptake kinetics of glycans by *B. theta* strains. In this regard, this study represents a methodological platform to study interactions between chemically distinct glycans with bacteria from diverse ecosystems, and to develop evidence-based probiotics and prebiotics for applications in livestock.

2.5 Materials and Methods

2.5.1 Isolation of bovine-adapted mannan-degraders

Bovine rumen and fecal samples were collected and enriched by Dr. Trevor Alexander's group. *In vitro* batch culture experiments using ruminal and fecal inoculant were enriched with the following substrates: Bio-Mos® (1% W/V), corn DDGS (1% W/V; DDGS), or YM (1% W/V). Three biological replicates for each sample were incubated anaerobically at 39°C on a rotary shaker for 24 h. Bacteria were isolated from the enriched batch cultures by streaking onto nutrient restricted media supplemented with 0.5% YM to select for YM-degraders. Serial dilutions of the batch cultures (10_0-10_{-6}) were similarly streaked on YM plates. Plates were incubated anaerobically (atmosphere: 85% N_2, 10% CO_2, 5% H_2) for up to four days. After 24-96 h of incubation, single colonies were selected and inoculated in nutrient rich Brain Heart Infusion (BHI) medium: 3.7% Bacto™ Brain Heart Infusion (BD; 237500), 8.3 mM L-cysteine, 10mL/L hemin (0.77 mM hemin in 0.01M NaOH), and 0.2% $NaHCO_3$. Overnight cultures were centrifuged and resuspended in 0.8 mL glycerol (50%) and stored at -80 °C. In total, 50 YM-degrading bacterial isolates were characterized for their propensity to metabolize YM. Nine of these isolates were selected for detailed analysis in this study.

2.5.2 Growth profiling of bovine isolates

Bovine isolates, wild type *B. theta* VPI-5482, and a mutant *B. theta* strain lacking MAN-PULs 1, 2, and 3 (ΔMAN-PUL1/2/3; donated by Dr. Elizabeth Lowe, Newcastle University) [59]

were cultured overnight in tryptone-yeast-glucose (TYG) medium : 1% Bacto™ Tryptone (BD; 211705), 0.5% Yeast Extract Bacteriological (VWR; J850), 4.1 mM L-cysteine, 0.2% glucose, 0.1 M KPO_4 pH 7.2, 2.2 µM vitamin K_3, 40 µL/mL TYG Salts (2 mM $MgSO_4 \cdot 7H_2O$, 119 mM $NaHCO_3$, and 34.2 mM NaCl), 28.8 µM $CaCl_2$, 1.4 µM $FeSO_4$, 4.4 µM resazurin, 1 µL/mL (v/v) histidine/hematin (1.9 mM hematin, and 200 mM L-histidine, 1000X stock solution). All growth incubations were performed in anaerobic conditions (atmosphere: 85% N_2, 10% CO_2, 5% H_2) at 37°C in a vinyl anaerobic chamber (Coy Lab Products). The overnight cultures (OD_{600} 1.0-1.4) were diluted to an of OD_{600} 0.05 in 2X *Bacteroides* minimal medium (MM), pH 7.2: 200 mL/L (v/v) 10X Bacteroides salts solution pH 7.2 (999 mM KH_2PO_4, 30 mM NaCl, 17 mM $(NH_4)_2SO_4$), 20 mL/L (v/v) Balch's Vitamins pH 7.0 (36.5 µM ρ-aminobenzoic acid, 4.5 µM folic acid, 8.2 µM biotin, 40.6 µM nicotinic acid, 10.5 µM calcium pantothenate, 13.3 µM riboflavin, 14.8 µM thiamine HCl, 48.6 µM vitamin B_6, 73.8 nM vitamin B_{12}, and 24.2 mM thioctic acid), 20 mL/L (v/v) Amino Acid Solution (5 mg/mL amino acids: alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine), 20 mL/L (v/v) Purine/Pyrimidine Solution pH 7.0 (1 mg/mL adenine, guanine, thymine, cytosine, and uracil), 20 mL/L (v/v) Trace Mineral Solution pH 7.0 (1.7 M ethylenediaminetetraacetic acid, 12.2 M $MgSO_4 \cdot 7H_2O$, 3 $MnSO_4 \cdot H_2O$, 17.1 M NaCl, 359.7 mM $FeSO_4 \cdot 7H_2O$, 901.1 mM $CaCl_2$, 347.7 mM $ZnSO_4 \cdot 7H_2O$, 40.1 mM $CuSO_4 \cdot 5H_2O$, 161.7 mM H_3BO_3, 41.3 mM $Na_2MoO_4 \cdot 2H_2O$, and 84.1 mM $NiCl_2 \cdot 6H_2O$), 4.4 µM vitamin K_3, 2.9 µM $FeSO_4 \cdot 7H_2O$, 14.4 µM $CaCl_2$, 2 mM $MgCl_2 \cdot 6H_2O$, 7.4 pM vitamin B_{12}, 16.5 mM L-cysteine, and 2 µL/mL (v/v) histidine/hematin. Wells of 300 µLflat-bottomed 96-well microtiter plates (Falcon) were filled with 100 µL of sterilized 1% (w/v) YM (Sigma, St. Loius, USA; M7504) or mannose with 100 µL inoculant (n=4). Negative control wells consisted of 100 µL 2X MM combined with 100 µL 1% (w/v) YM or mannose and were used to normalize curves against buffer effects. 100 µL bacterial suspension was used as inoculant to get starting OD_{600nm} ~0.025. Plates were sealed with polyurethane Breathe-Easy gas-permeable

membranes (Sigma; Z390059). Absorbance (600nm) of each well was measured with a Biotek Eon microplate reader and recorded on Biotek Gen5 software every 10 minutes for 50 hours. Mean ($\pm$ standard deviation) of each condition (n = 4) was visualized using GraphPad Prism 6.

Post-growth cultures were harvested and centrifuged. Supernatants were taken and 6 µL was ran on a silica sheet in 2:1:1 (butanol:d_2H2O:acetic acid) running buffer. The plate was dried at ambient temperature. Orcinol (diluted to 1% in a solution of 70:3 ethanol:sulfuric acid) was used as a stain. Once the plate was dry, it was activated in an oven at 120˚C and imaged.

2.5.3 Genome sequencing, assembly, and annotation of *B. theta*Bov strains

16S rRNA gene sequencing was done on 50 bovine bacterial isolates to determine taxonomic classification using universal primers: 5'-AGR GTT TGA TCM TGG CTC AG-3' and 5'-GGT TAC CTT GTT ACG ACT T-3'. Of these isolates, nine were chosen for whole genome sequencing using Illumina MiSeq PE150bp based on growth profiles ($OD_{600} > 0.3$) and 16S rRNA sequencing results. Genomes were assembled using SPAdes *de novo* assembly [131]. The K-mer value in SPAdes was chosen from (21, 33, 55, 77 - defaults for 150bp reads). Quality reporting of the assemblies was done using Quast [132]. SPAdes assembly N50s, largest contigs, and number of contigs shown in Table 2.2. Isolate contigs were blasted against the reference genome *B. theta* VPI-5482 for MAN-PUL1,2,3, and the HMNG-PUL using NCBI BLAST (2.7.1) [133]. SPAdes contig assemblies were aligned with the JSpeciesWS reference *B. theta* genomes to calculate average nucleotide identity based on BLAST+ (ANIb) [115, 116] (Table 2.3).

2.5.4 PUL delineation and comparative CAZomics

Isolate contigs were run through EMBOSS GetORF [134] in order to determine open reading frames; this data was run through the dbCAN [100] HMMscan in order to pick out CAZyme sequences. CAZyme sequences were put through SACCHARIS [24] (Sequence Analysis and Clustering of CarboHydrate Active enzymes for Rapid Informed prediction of Specificity), a novel in-house designed pipeline that produces phylogenetic trees of gene sequences based on percent

identity. User sequences of CAZymes and associated modules, such as CBMs, can be paired with known CAZy sequences from the CAZY database [135], trimmed to their catalytic domain with dbCAN, aligned with MUSCLE [136], and fitted for the best phylogenetic tree using Prottest3 [137] to find the appropriate amino acid replacement model, and either RAxML [138] or FastTree [139] to generate the final tree. GHs from families 38, 76, 92, 99, and 125 identified in the genomes of the MD isolates were analyzed by SACCHARIS. Phylogenetic trees were developed using FastTree, and newick file outputs were viewed using FigTree [140] (Figure 2.2).

2.5.5 RNA-seq: assembly, quantitation, and comparative analysis

B. theta VPI-5482, MD33MG, and MD40HG were each inoculated into three tubes containing 5 mL of TYG media. Overnight cultures of *B. theta* VPI-5482, MD33MG, and MD40HG (OD$_{600}$ 1.0-1.4) were diluted with 2X MM to an OD$_{600}$ 0.05. 100 µL of 1% YM or mannose was added to six wells of diluted culture aliquoted from the same overnight stock, to a final volume of 200 µL. This step was repeated three times for a total of three replicates (each divided among six wells) for each bacterial strain and treatment. Cells were harvested during exponential phase (OD$_{600}$ 0.4-0.8). Six wells were pooled and added to an equal volume of RNAprotect (Qiagen). The mixture was vortexed for 5 sec and incubated for 5 min at room temperature. After incubation, the protected cells were centrifuged (10 min; 2,800 x *g*). The supernatant was decanted and pellets were stored at -80°C until further processing.

RNA was extracted and purified using a GeneJET RNA Purification kit (Thermo Scientific) within 1 week of storage. RNA was sent to Génome Québec for Illumina HiSeq 4000 PE100bp sequencing. Using Geneious v11.1.2 [141], each set of reads was mapped to their previously assembled genomic sequence, or in the case of *B. theta* VPI-5482, to the genomic sequence from the NCBI database (NC_004663). Expression levels were calculated as transcript expression (transcript per kilobase per million; TPM) for each growth treatment. Ambiguously mapped reads were counted as partial matches. The Geneious DESeq2 [117] plugin was used to

compare the expression levels between the two treatments, producing log2 expression ratios and p-values. A generalized linear mixed model was used to calculate statistical differences of TPM values between the MAN-PUL1,2,3 *susC*-like genes of each bacterial strain.

2.5.6 Sequence comparison and modelling of SusCD-like proteins

MUSCLE was used to align MAN-PUL2 and 3 SusC-like and SusD-like amino acid sequences of each isolate and calculate percent identity (Table 2.4). 16S rRNA gene and MAN-PUL2 SusC-like and SusD-like amino acid phylogenetic trees were generated using the Maximum Likelihood method and Tamura-Nei model [142]. Evolutionary analyses were performed by MEGA X [143]. Trees with the highest log likelihood are shown in Figure 2.4A, B, and C. MD33MG SusC-like and SusD-like protein structures were predicted based on PDB 5FQ6 (BT2261-2264 SusCD complex) [121] using the PHYRE2 server [144] and modelled with PyMOL (v2.0) as cartoon and surface features. Residues identified in the MUSCLE alignment to be different between HG and MG SusC-like and SusD-like sequences were differentially coloured (Figure 2.4D & E).

2.5.7 Generation of FLA-YM conjugates

A previously defined protocol [83, 84] was used to generate fluorescently labelled YM (FLA-YM), with slight variations. To chemically activate the polysaccharide, 350 µL 0.81 M Cyanogen bromide (CNBr; 97%; Sigma C91492) was added to 2 mL 2% YM. For ~5 min, pH was monitored and maintained above 9.5 with additions of 0.25 M NaOH. Activated YM was separated from CNBr using Sephadex® G-50 gel filtration medium in a column coupled to a Bio-Rad BioLogic LP Multistatic peristaltic pump (flow rate: 1 mL min-1). The mobile phase consisted of 0.2 M sodium tetraborate decahydrate pH 8.0 (≥99.5%, Sigma S9640). Activated YM was eluted into a vial containing 2.0 mg Fluoresceinamine Isomer II (FLA; ~95%; Sigma O7985) wrapped in aluminum foil and incubated for ~24 h at ambient temperature. To remove excess FLA and purify labelled YM, the reaction mixture was loaded onto Sartorius Vivaspin 15R columns (5,000 MWCO; VS15RH11) and centrifuged (210 x *g*). Columns were repeatedly topped up with distilled H2O and

centrifuged until a clear filtrate was observed. The purified FLA-YM was lyophilized, covered in aluminum foil and stably stored (~4°C) until further use.

2.5.8 Uptake of FLA-YM by strains of *B. theta*Bov

Wild type *B. theta* VPI-5482, *B. theta* ΔMAN-PUL1/2/3, and rumen isolates MD33MG, and MD40HG were inoculated in TYG and grown as described above. Cells were harvested in exponential phase (OD_{600} ~1.0) and centrifuged (4,700 x *g*) for 5 min, the supernatant was removed, and pellets resuspended in 2 mL 2X MM for the first two washes. After the third centrifugation, pellets were resuspended in 2 mL MM with 0.5% YM (*B. theta* VPI-5482, MD33MG, and MD40HG) or 0.5% glucose + YM (*B. theta* ΔMAN-PUL1/2/3) as the sole carbon source (not conjugated to FLA) to prime bacterial metabolism of YM. After ~18 h incubation, cultures were centrifuged and washed three times, with the final resuspension in 2 mL 2X MM. 300 µL of the resuspended pellet was aliquoted into 0.2% unlabeled YM or FLA-YM. 20 µL of the 2X MM resuspension was used as the 0 h time point, as the cells were not exposed to FLA-YM. 40 µL aliquots of each condition were taken at two additional time points: 24 h and 72 h. The cells were centrifuged (10 min; 2,300 x *g*) and the pellet was fixed in 1% formaldehyde (FA; Sigma; F8775) in 1X phosphate-buffered saline pH 7.4 (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4), at 4°C for 18-24 h. The fixed cells were centrifuged (10 min; 2,300 x *g*) and washed in 1X PBS. The samples were centrifuged and stored at 4°C in the dark until visualized by SR-SIM (Figure 2.5).

2.5.9 Rate of FLA-YM uptake by *B. theta*Bov isolates

B. theta VPI-5482, MD33MG, and MD40HG were grown in TYG and prepared as described above. After 24 h growth, cultures were divided into 2 mL 0.5% YM or mannose to create activated and non-activated populations, respectively. Cells were harvested in exponential phase (OD_{600} 0.6-1.0), centrifuged (10 min; 2,300 x *g*), and resuspended in 2 mL 2X MM. 300 µL of the resuspension was added to 1 mL 2X MM. 20 µL of each treatment sample was aliquoted into 1 mL 1% FA and used as the T0 timepoint. FLA-YM or YM was then added to each culture and additional time

points of 40 µL were taken at 5 min, 10min, 15 min, 20 min, 30 min, 1 h, 2 h, 4 h, 8 h, and 24 h

and visualized by SR-SIM (Figure 2.6).

2.5.10 Visualization of FLA-YM uptake

Super-resolution microscopy were performed as described in [122]. Briefly, 30 µL of fixed cells were fixed to poly-D-lysine-coated #1.5 coverslips (0.17mm thick) at 40°C. The coverslip was rinsed with d_2H_2O to remove residual salts. Samples were counter-stained for 10 min and 25 min with 1 ng µL$_{-1}$ WS 4',6-diamidino-2-phenylindole (DAPI) and 2 ng µL$_{-1}$ WS Nile red, respectively. Cells were washed with 18 MΩcm$_{-1}$ after each stain and mounted with 4:1 Citifluor/VectaShield mounting solution. Cells were then visualized using a Zeiss ELYRA PS.l (Carl Zeiss). Lasers with wavelengths 561, 488, and 406 nm and optical filters BP 573-613, BP 502-538, and BP 420-480 + LP 750 were used. Using Plan-Apochromat x63/1.4 oil objective, Z-stack images were taken. Images were processed using the ZEN2011 (Carl Zeiss, Germany) software.

Chapter 3

Future Directions

This project has used a combination of genomics, transcriptomics, bioinformatics, and FGCs to investigate the interaction between YM and *B. theta* strains isolated from rumen and fecal samples of cattle, referred to here as *B. theta*Bov. This work builds upon previous studies that used FLA-YM and FLA-RGII to visualize uptake in *B. theta* VPI-5482 [122], and a variety of FGCs to visualize glycan uptake in pure cultures of marine bacteria [84]. Beyond these applications, I foresee that FGCs could be used in future studies, including to visualize single-cell glycan metabolism in complex communities, generate metabolically enriched metagenomes, and tract passage of FGCs through the gastrointestinal tract of mice using non-invasive whole body imaging. Each of these applications will be discussed in more detail below.

3.1 Visualization of individual cell glycan metabolism in complex communities

In this and previous studies [84, 122], FGCs were used in pure culture. Adding FGCs to complex microbial communities, however, has the potential to discriminate between cells that interact with and metabolize the labelled glycan. Subsequently, taxa within the community can be identified using fluorescence *in situ* hybridization (FISH) and enumerated using flow cytometry to determine the proportions of taxonomic groups in different samples. In this regard, I have conducted a preliminary study to explore the utility of FLA-YM to label microbial cells within rumen communities. Samples were extracted from cannulated sheep that were either fed a control diet (did not contain α-mannan), or one of two diets that were enriched with either DDGS or Bio-Mos®. Preliminary SR-SIM images revealed a diversity of cell morphologies and that FLA-YM signal is enriched in samples that were supplemented with DDGS and Bio-Mos® (Figure 3.1). Therefore, it appears that this approach may provide an effective way to directly visualize interactions between glycans and intestinal bacteria, which could be useful for identifying new prebiotics.
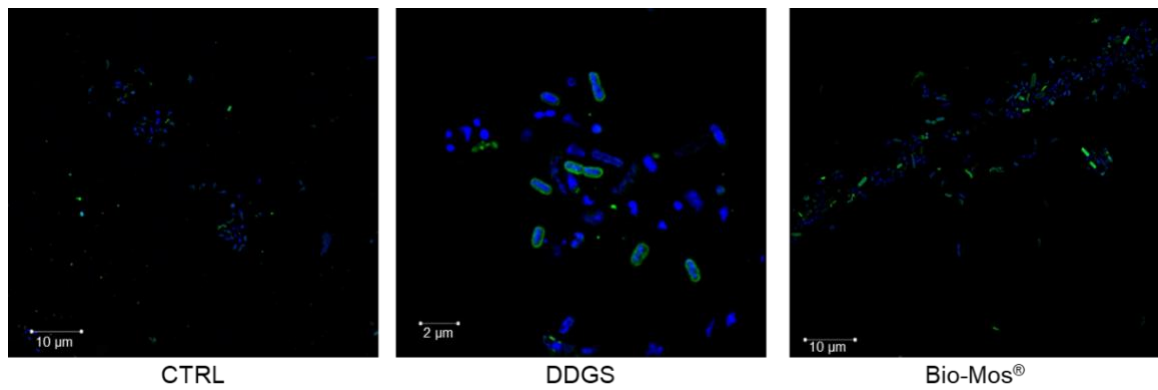
Figure 3.1 Rumen communities fed different diets incubated with FLA-YM.
Sheep reared on control diets (CTRL), or diets enriched with DDGS or Bio-Mos®. Samples were incubated with FLA-YM (green) for 24 h anaerobically at 37˚C and visualized by SR-SIM. Samples counter-stained with DAPI (blue) to localize cellular DNA. Images courtesy of Dr. Greta Reintjes.

3.2 Fluorescence assisted cell-sorting to generate metabolically enriched metagenomes

FGC-labeled microorganisms within complex communities (Figure 3.1) can also be sorted and characterized. I hypothesize that these sorted microbiomes would be enriched with genes encoding transport and enzyme machinery tailored for the metabolism of FGC-associated glycans. Fluorescent assisted cell sorting and sequencing (FACSeq) of intestinal microbiomes, therefore, provides a culture-independent method to enrich natural microbial communities for CAZyme sequences of interest and provide a novel method for CAZyme bioprospecting.

3.3 Non-invasive whole-body imaging

FGCs have the potential to be used as *in vivo* molecular probes to monitor the passage of prebiotic glycans through the gastrointestinal tract of mice. This information may be useful for measuring transit time and the spatial distribution of glycan metabolism. To investigate the potential of this technology, a preliminary study has been conducted. Conventional c57bl/6 mice with an agnotobiotic gut microbiota were fed a diet ± YM for two weeks and then gavaged with a sample that contained ± 3% FLA-YM. Non-invasive whole-body imaging was performed at 3 h (Figure 3.2A). Fluorescent signal was detected in mice reared on diets ± YM (Figure 3.2A). There did appear to be a retention of signal in the small intestine in the mice that were fed YM in their diets (Figure 3.2B). In the future, follow-up experiments that include 3 h, 5 h, 9 h, and 24 h time points will be done to determine if activated communities delay the passage of FGCs through the GI tract over longer periods of time.

To investigate if changes in FGC uptake by intestinal bacteria was influenced by diet, cecum samples were collected, fixed, and visualized using epifluorescence. There was a higher density of FGC signal in the activated communities (Figure 3.2C), which was similar to what was observed in pure culture (Figure 2.5) and sheep rumen communites (Figure 3.1). Therefore, it appears that FGCs represent versatile tools to investigate glycan-bacterial interactions and can be used at levels ranging from the entire digestive system down to single-cells. In the future, coupling

FGC applications to -omics technologies, such as FACSeq or meta-transcriptomics and meta-proteomics will provide next-generation approaches to characterize the function of microbiomes.
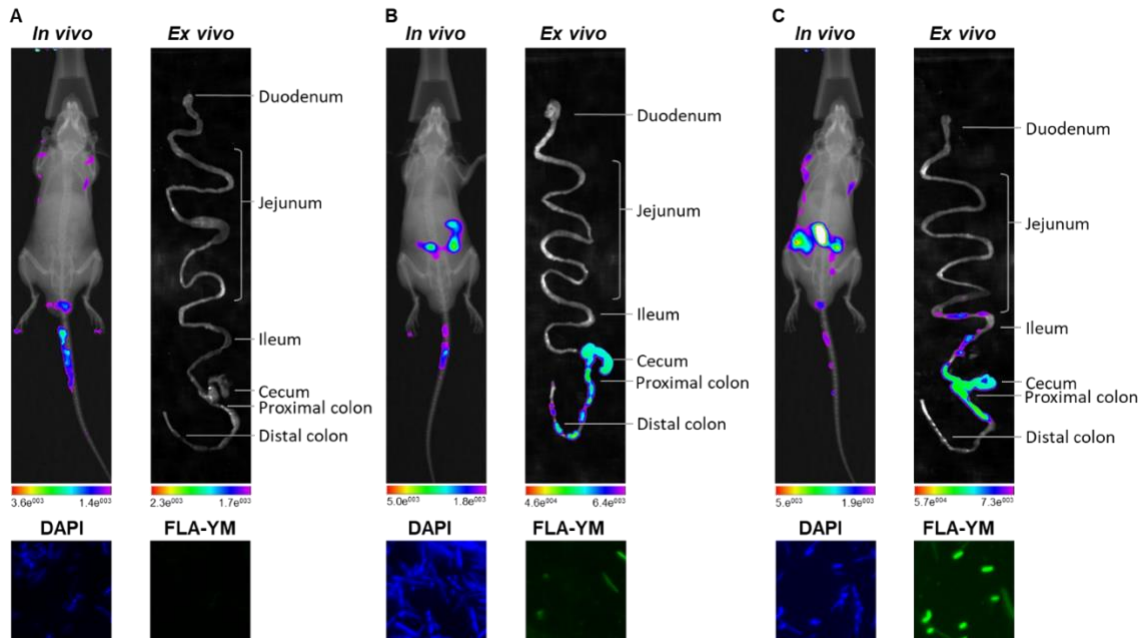
Figure 3.2 Visualization of FLA-YM transport through murine gastrointestinal tracts.
(A) Mouse fed control diet and gavaged with H_2O. **(B)** Mouse fed control diet and gavaged with FLA-YM. **(C)** Mouse fed YM diet and gavaged with FLA-YM. **(Top Left)** Whole body imaging of mice overlaying fluorescence and x-ray scans (ventral side of mice are shown). **(Top Right)** *Ex vivo* fluorescence scan of dissected gastrointestinal tracts (major anatomy labelled) showing FLA-YM signal. Fluorescence intensity scale shown beneath each image; grey = no fluorescence detected, red = lowest fluorescence signal, magenta = highest fluorescence signal. Images courtesy of Dr. Bjoern Petri. Bacteria extracted from mice ceca were visualized by epifluorescence microscopy; cells were counterstained with DAPI **(Bottom Left)** to identify individual cells accumulating FLA-YM **(Bottom Right)** signal.

References

1.   United Nations, D.o.E.a.S.A., Population Division, *World Population Prospects: The 2017 Revision, Key Findings, and Advance Tables*. 2017.
2.   Agriculture, U.S.D.o., *Livestock and Poultry: World Markets and Trade*. 2018.
3.   Agriculture and Agri-Food Canada, A.I.D., Market Information Section. *Red meat and livestock export reports*. 2019 [cited 2019 March 4]; Available from: http://www.agr.gc.ca/eng/industry-markets-and-trade/canadian-agri-food-sector-intelligence/red-meat-and-livestock/red-meat-and-livestock-market-information/exports/?id=1415860000065.
4.   Beef Cattle Research Council, A.B.P. *The way you purchase antibiotics is changing*. 2018 [cited 2019 February 28]; Available from: http://www.beefresearch.ca/blog/prescription-required-for-livestock-cattle-antibiotics-december-2018-repost/.
5.   Canada, P.H.A.o., *Tackling antimicrobial resistance and antimicrobial use: A pan-Candian framework for action*. 2017.
6.   Beef Cattle Research Council, A.B.P., *Producing beef using less water*, T.M.a.K. Beauchemin, Editor. 2017.
7.   Hristov, A.N., Oh, J., Lee, C., Meinen, R., Montes, F., Ott, T., Firkins, J., Rotz, A., Dell, C., Adesogan, A., Yang, W., and J. Tricarico, Kebreab, E., Waghorn, G., Dijkstra, J. & Oosting, S., *Mitigation of greenhouse gas emissions in livestock production: A review of technical options for non-CO2 emissions*, B.H. Pierre J. Gerber, and Harinder P.S. Makkar, Editor. 2013: FAO, Rome, Italy.
8.   (CRSB), C.R.f.S.B., *National beef sustainability assessment and strategy summary report.* Calgary, AB: CRSB, 2016.
9.   Council, B.C.R., *Priority Area Review: Feed Grains & Feed Efficiency.* 2017.
10.   Maddock, T.D.a.L., G. C., *The Economic Impact of Feed Efficiency in Beef Cattle.* Animal Sciences Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida, Florida, 2012.
11.   Boaitey, A., Goddard, E.; Mohapatra, S. and Crowley, J., *Feed efficiency estimates in cattle: The economic and environmental impacts of reranking.* Sustainable Agriculture Research, 2017. **6**(2): p. 35-47.
12.   Steele, M.A., et al., *Development and physiology of the rumen and the lower gut: Targets for improving gut health.* J Dairy Sci, 2016. **99**(6): p. 4955-4966.
13.   Bergman, E.N., *Energy contributions of volatile fatty acids from the gastrointestinal tract in various species.* Physiol Rev, 1990. **70**(2): p. 567-90.
14.   Varga, G.A. and E.S. Kolver, *Microbial and animal limitations to fiber digestion and utilization.* J Nutr, 1997. **127**(5 Suppl): p. 819s-823s.
15.   Malmuthuge, N. and L.L. Guan, *Understanding host-microbial interactions in rumen: searching the best opportunity for microbiota manipulation.* J Anim Sci Biotechnol, 2017. **8**: p. 8.
16.   Nishida, A.H. and H. Ochman, *Rates of gut microbiome divergence in mammals.* Mol Ecol, 2018. **27**(8): p. 1884-1897.
17.   Baumler, A.J. and V. Sperandio, *Interactions between the microbiota and pathogenic bacteria in the gut.* Nature, 2016. **535**(7610): p. 85-93.
18.   Wexler, H.M., *Bacteroides: the good, the bad, and the nitty-gritty.* Clin Microbiol Rev, 2007. **20**(4): p. 593-621.
19.   Henderson, G., et al., *Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range.* Sci Rep, 2015. **5**: p. 14567.

20.   Uyeno, Y., S. Shigemori, and T. Shimosato, *Effect of Probiotics/Prebiotics on Cattle Health and Productivity.* Microbes Environ, 2015. **30**(2): p. 126-32.

21.   Mizrahi, I. and E. Jami, *Review: The compositional variation of the rumen microbiome and its effect on host performance and methane emission.* Animal, 2018. **12**(s2): p. s220-s232.

22.   Hungate, R.E., *The rumen and its microbes*. 1966, New York, NY: Academic Press.

23.   Zeineldin, M., et al., *Synergetic action between the rumen microbiota and bovine health.* Microb Pathog, 2018. **124**: p. 106-115.

24.   Jones, D.R., et al., *SACCHARIS: an automated pipeline to streamline discovery of carbohydrate active enzyme activities within polyspecific families and de novo sequence datasets.* Biotechnol Biofuels, 2018. **11**: p. 27.

25.   Hess, M., et al., *Metagenomic discovery of biomass-degrading genes and genomes from cow rumen.* Science, 2011. **331**(6016): p. 463-7.

26.   Stewart, R.D., et al., *Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen.* Nat Commun, 2018. **9**(1): p. 870.

27.   Ghosh, S. and R.K. Mehla, *Influence of dietary supplementation of prebiotics (mannanoligosaccharide) on the performance of crossbred calves.* Trop Anim Health Prod, 2012. **44**(3): p. 617-22.

28.   He, Z.X., et al., *Effect of wheat dried distillers grains with solubles and fibrolytic enzymes on ruminal fermentation, digestibility, growth performance, and feeding behavior of beef cattle.* J Anim Sci, 2015. **93**(3): p. 1218-28.

29.   Thomas, F., et al., *Environmental and gut bacteroidetes: the food connection.* Front Microbiol, 2011. **2**: p. 93.

30.   Gaggia, F., P. Mattarelli, and B. Biavati, *Probiotics and prebiotics in animal feeding for safe food production.* Int J Food Microbiol, 2010. **141 Suppl 1**: p. S15-28.

31.   Barba-Vidal, E., et al., *Evaluation of the Probiotic Strain Bifidobacterium longum subsp. Infantis CECT 7210 Capacities to Improve Health Status and Fight Digestive Pathogens in a Piglet Model.* Front Microbiol, 2017. **8**: p. 533.

32.   Timmerman, H.M., et al., *Health and growth of veal calves fed milk replacers with or without probiotics.* J Dairy Sci, 2005. **88**(6): p. 2154-65.

33.   Angelakis, E., *Weight gain by gut microbiota manipulation in productive animals.* Microb Pathog, 2017. **106**: p. 162-170.

34.   Garcia Diaz, T., et al., *Use of live yeast and mannan-oligosaccharides in grain-based diets for cattle: Ruminal parameters, nutrient digestibility, and inflammatory response.* PLoS One, 2018. **13**(11): p. e0207127.

35.   Fischbach, M.A. and J.L. Sonnenburg, *Eating for two: how metabolism establishes interspecies interactions in the gut.* Cell Host Microbe, 2011. **10**(4): p. 336-47.

36.   Salminen, S., E. Isolauri, and E. Salminen, *Clinical uses of probiotics for stabilizing the gut mucosal barrier: successful strains and future challenges.* Antonie Van Leeuwenhoek, 1996. **70**(2-4): p. 347-58.

37.   Kabir, A.M., et al., *Prevention of Helicobacter pylori infection by lactobacilli in a gnotobiotic murine model.* Gut, 1997. **41**(1): p. 49-55.

38.   John Wallace, R., *Gut microbiology - broad genetic diversity, yet specific metabolic niches.* Animal, 2008. **2**(5): p. 661-8.

39.   Huws, S.A., et al., *Addressing Global Ruminant Agricultural Challenges Through Understanding the Rumen Microbiome: Past, Present, and Future.* Front Microbiol, 2018. **9**: p. 2161.

40.   Jones, R.J. and R.G. Megarrity, *Successful transfer of DHP-degrading bacteria from Hawaiian goats to Australian ruminants to overcome the toxicity of Leucaena.* Aust Vet J, 1986. **63**(8): p. 259-62.

41.   Allison, M.J.M., W.R.; McSweeny, C.S.; and Stahl, D.A., *Syntergistis jonesii, gen. nov., sp. nov.: A rumen bacterium that degrades toxic pyridinediols.* System. Appl. Microbiol., 1992. **15**: p. 522-529.

42.   Hammond, A.C., *Leucaena toxicosis and its control in ruminants.* J Anim Sci, 1995. **73**(5): p. 1487-92.

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

44.   Spring, P., Wenk, C., Dawson, K. A., and Newman, K. E., *The Effects of Dietary Mannanoligosaccharides on Cecal Parameters and the Concentrations of Enteric Bacteria in the Ceca of Salmonella-Challenged Broiler Chickens.* Poultry Science, 2000. **79**: p. 205-211.

45.   Freter, R., et al., *Mechanisms that control bacterial populations in continuous-flow culture models of mouse large intestinal flora.* Infect Immun, 1983. **39**(2): p. 676-85.

46.   Faith, J.J., et al., *Predicting a human gut microbiota's response to diet in gnotobiotic mice.* Science, 2011. **333**(6038): p. 101-4.

47.   Martinez, I., et al., *Resistant starches types 2 and 4 have differential effects on the composition of the fecal microbiota in human subjects.* PLoS One, 2010. **5**(11): p. e15046.

48.   Grand, E., et al., *Effects of short-chain fructooligosaccharides on growth performance of preruminant veal calves.* J Dairy Sci, 2013. **96**(2): p. 1094-101.

49.   Semjen, G. and P. Galfi, *Factors influencing the adherence of strains of Streptococcus bovis and Escherichia coli isolated from ruminal epithelium.* Vet Res Commun, 1990. **14**(3): p. 181-91.

50.   Ley, R.E., et al., *Evolution of mammals and their gut microbes.* Science, 2008. **320**(5883): p. 1647-51.

51.   Markowiak, P. and K. Slizewska, *The role of probiotics, prebiotics and synbiotics in animal nutrition.* Gut Pathog, 2018. **10**: p. 21.

52.   Perry, G.H., et al., *Diet and the evolution of human amylase gene copy number variation.* Nat Genet, 2007. **39**(10): p. 1256-60.

53.   Grondin, J.M., et al., *Polysaccharide Utilization Loci: Fueling Microbial Communities.* J Bacteriol, 2017. **199**(15).

54.   Patel, D.D., et al., *Microbial and Carbohydrate Active Enzyme profile of buffalo rumen metagenome and their alteration in response to variation in the diet.* Gene, 2014. **545**(1): p. 88-94.

55.   Beauchemin, K.A., Colombatto, D.; Morgavi, D. P. and Yang, W. Z., *Use of exogenous fibrolytic enzymes to improve feed utilzation by ruminants.* Animal Science, 2003. **81**(E. Suppl. 2): p. E37-E47.

56.   Holtshausen, L., et al., *Improved milk production efficiency in early lactation dairy cattle with dietary addition of a developmental fibrolytic enzyme additive.* J Dairy Sci, 2011. **94**(2): p. 899-907.

57.   He, Z.X., et al., *Using a fibrolytic enzyme in barley-based diets containing wheat dried distillers grains with solubles: ruminal fermentation, digestibility, and growth performance of feedlot steers.* J Anim Sci, 2014. **92**(9): p. 3978-87.

58.   Ndeh, D., et al., *Complex pectin metabolism by gut bacteria reveals novel catalytic functions.* Nature, 2017. **544**(7648): p. 65-70.

59.   Cuskin, F., et al., *Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism.* Nature, 2015. **517**(7533): p. 165-169.

60.   Aguilar-Uscanga, B. and J.M. Francois, *A study of the yeast cell wall composition and structure in response to growth conditions and mode of cultivation.* Lett Appl Microbiol, 2003. **37**(3): p. 268-74.

61.   Orlean, P., *Architecture and biosynthesis of the Saccharomyces cerevisiae cell wall.* Genetics, 2012. **192**: p. 775-818.

62.   Dimitroglou, A., et al., *Dietary mannan oligosaccharide supplementation modulates intestinal microbial ecology and improves gut morphology of rainbow trout, Oncorhynchus mykiss (Walbaum).* J Anim Sci, 2009. **87**(10): p. 3226-34.

63.   Firon, N., et al., *Aromatic alpha-glycosides of mannose are powerful inhibitors of the adherence of type 1 fimbriated Escherichia coli to yeast and intestinal epithelial cells.* Infect Immun, 1987. **55**(2): p. 472-6.

64.   Spring, P., Wenk, C., Connolly, A., and Kiers, A. , *A review of 733 published trials on Bio-Mos®, a mannan oligosaccharide, and Actigen®, a second generation mannose rich fraction, on farm and companion animals.* Applied Animal Nutrition, 2015. **3**(e7): p. 1-11.

65.   Kolenda, R., M. Burdukiewicz, and P. Schierack, *A systematic review and meta-analysis of the epidemiology of pathogenic Escherichia coli of calves and the role of calves as reservoirs for human pathogenic E. coli.* Front Cell Infect Microbiol, 2015. **5**: p. 23.

66.   Scholey, D.V., E.J. Burton, and P.E. Williams, *The bio refinery; producing feed and fuel from grain.* Food Chem, 2016. **197**(Pt A): p. 937-42.

67.   Lombard, V., et al., *The carbohydrate-active enzymes database (CAZy) in 2013.* Nucleic Acids Res, 2014. **42**(Database issue): p. D490-5.

68.   www.cazy.org, *Carbohydrate-Active enZYmes.* Accessed February 12, 2019.

69.   Aspeborg, H., et al., *Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5).* BMC Evol Biol, 2012. **12**: p. 186.

70.   Stam, M.R., et al., *Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins.* Protein Eng Des Sel, 2006. **19**(12): p. 555-62.

71.   St John, F.J., J.M. Gonzalez, and E. Pozharski, *Consolidation of glycosyl hydrolase family 30: a dual domain 4/7 hydrolase family consisting of two structurally distinct groups.* FEBS Lett, 2010. **584**(21): p. 4435-41.

72.   Mewis, K., et al., *Dividing the Large Glycoside Hydrolase Family 43 into Subfamilies: a Motivation for Detailed Enzyme Characterization.* Appl Environ Microbiol, 2016. **82**(6): p. 1686-92.

73.   Zhu, Y., et al., *Mechanistic insights into a Ca2+-dependent family of alpha-mannosidases in a human gut symbiont.* Nat Chem Biol, 2010. **6**(2): p. 125-32.

74.   Tancula, E., et al., *Location and characterization of genes involved in binding of starch to the surface of Bacteroides thetaiotaomicron.* J Bacteriol, 1992. **174**(17): p. 5609-16.

75.   D'Elia, J.N. and A.A. Salyers, *Effect of regulatory protein levels on utilization of starch by Bacteroides thetaiotaomicron.* J Bacteriol, 1996. **178**(24): p. 7180-6.

76.   Terrapon, N., et al., *PULDB: the expanded database of Polysaccharide Utilization Loci.* Nucleic Acids Res, 2018. **46**(D1): p. D677-d683.

77.   Abbott, D.W., et al., *Coevolution of yeast mannan digestion: Convergence of the civilized human diet, distal gut microbiome, and host immunity.* Gut Microbes, 2015. **6**(5): p. 334-9.

78.   Pluvinage, B., et al., *Molecular basis of an agarose metabolic pathway acquired by a human intestinal symbiont.* Nat Commun, 2018. **9**(1): p. 1043.

79.   Yan, H., R.S. Yalagala, and F. Yan, *Fluorescently labelled glycans and their applications.* Glycoconj J, 2015. **32**(8): p. 559-74.

80.   Glabe, C.G., P.K. Harty, and S.D. Rosen, *Preparation and properties of fluorescent polysaccharides.* Anal Biochem, 1983. **130**(2): p. 287-94.

81.   Arnosti, C., *Measurement of depth- and site-related differences in polysaccharide hydrolysis rates in marine sediments.* Geochimica et Cosmochimica Acta, 1995. **59**(20): p. 4247-4257.

82.	Arnosti, C., *A new method for measuring polysaccharide hydrolysis rates in marine environments.* Organic Geochemistry, 1996. **25**(1/2): p. 105-115.

83.	Arnosti, C., *Fluorescent derivatization of polysaccharides and carbohydrate-containing biopolymers for measurement of enzyme activities in complex media.* J Chromatogr B Analyt Technol Biomed Life Sci, 2003. **793**(1): p. 181-91.

84.	Reintjes, G., et al., *An alternative polysaccharide uptake mechanism of marine bacteria.* Isme j, 2017. **11**(7): p. 1640-1650.

85.	Teeling, H., et al., *Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom.* Science, 2012. **336**(6081): p. 608-11.

86.	White, B.A., et al., *Biomass utilization by gut microbiomes.* Annu Rev Microbiol, 2014. **68**: p. 279-96.

87.	Weimer, P.J., *Redundancy, resilience, and host specificity of the ruminal microbiota: implications for engineering improved ruminal fermentations.* Front Microbiol, 2015. **6**: p. 296.

88.	Henderson, G., et al., *Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range.* Sci Rep, 2015. **5**: p. 14567.

89.	Hess, M., Sczybra, A., Egan, R., Kim, T-W., Chokhawala, H., Schroth, G., Luo, S., Clark, D. S., Chen, F., Zhang, T., Mackie, R. I.,Pennacchio, L. A., Tringe, S. G., Visel, A., Woyke, T., Wange, Z., and Rubin, E. M., *Metagenomic discovery of biomass-degrading genes and genomes from cow rumen.* Science, 2011. **331**(6016): p. 463-467.

90.	Foley, M.H., D.W. Cockburn, and N.M. Koropatkin, *The Sus operon: a model system for starch uptake by the human gut Bacteroidetes.* Cell Mol Life Sci, 2016. **73**(14): p. 2603-17.

91.	Anderson, K.L. and A.A. Salyers, *Biochemical evidence that starch breakdown by Bacteroides thetaiotaomicron involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes.* J Bacteriol, 1989. **171**(6): p. 3192-8.

92.	Rakoff-Nahoum, S., M.J. Coyne, and L.E. Comstock, *An ecological network of polysaccharide utilization among human intestinal symbionts.* Curr Biol, 2014. **24**(1): p. 40-9.

93.	Larsbrink, J., et al., *A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes.* Nature, 2014. **506**(7489): p. 498-502.

94.	Sonnenburg, E.D., et al., *Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations.* Cell, 2010. **141**(7): p. 1241-52.

95.	Martens, E.C., et al., *Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts.* PLoS Biol, 2011. **9**(12): p. e1001221.

96.	Hehemann, J.H., et al., *Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes.* Proc Natl Acad Sci U S A, 2012. **109**(48): p. 19786-91.

97.	Rogowski, A., et al., *Glycan complexity dictates microbial resource allocation in the large intestine.* Nat Commun, 2015. **6**: p. 7481.

98.	Davies, G. and B. Henrissat, *Structures and mechanisms of glycosyl hydrolases.* Structure, 1995. **3**(9): p. 853-9.

99.	Martens, E.C., et al., *The devil lies in the details: how variations in polysaccharide fine-structure impact the physiology and evolution of gut microbes.* J Mol Biol, 2014. **426**(23): p. 3851-65.

100.	Yin, Y., et al., *dbCAN: a web resource for automated carbohydrate-active enzyme annotation.* Nucleic Acids Research, 2012. **40**(W1): p. W445-W451.

101.	Orlean, P., *Architecture and biosynthesis of the Saccharomyces cerevisiae cell wall.* Genetics, 2012. **192**(3): p. 775-818.

102. Abbott, D.W., et al., *Coevolution of yeast mannan digestion: Convergence of the civilized human diet, distal gut microbiome, and host immunity.* Gut Microbes, 2015. **6**(5): p. 334–339.

103. Miguel, J.C., Rodriguez-Zas, S.L., Pettigrew, J.E. , *Efficacy of a mannan oligosaccharide (Bio-Mos®) for improving nursery pig performance.* Journal of Swine Health and Production, 2004. **12**: p. 296-307.

104. Wohlt, J.E., T.T. Corcione, and P.K. Zajac, *Effect of yeast on feed intake and performance of cows fed diets based on corn silage during early lactation.* Journal of Dairy Science, 1998. **81**(5): p. 1345–1352.

105. Finck, D.N., et al., *Yeast supplementation alters the performance and health status of receiving cattle.* The Professional Animal Scientist, 2014. **30**(3): p. 333-341.

106. Spring, P., Wenk, C., Connolly, A., and Kiers, A., *A review of 733 published trials on Bio-Mos®, a mannan oligosaccharide, and Actigen®, a second generation mannose rich fraction, on farm and companion animals.* Journal of Applied Animal Nutrition, 2015. **3**(7).

107. Scholey, D.V., Burton, E. J., and Williams, P. E. V., *The bio refinery; producing feed and fuel from grain.* Food Chemistry, 2016. **197**: p. 937-942.

108. He, Z.X., Walker, N. D., McAllister, T. A., and Yang, W. Z., *Effect of wheat dried distillers grains with solubles and fibrolytic enzymes on ruminal fermentation, digestibility, growth performans, and feeding behavior of beef cattle.* Journal of Animal Science, 2015. **93**: p. 1218-1228.

109. Berger, L. and V. Singh, *Changes and evolution of corn coproducts for beef cattle.* J Anim Sci, 2010. **88**(13 Suppl): p. E143-50.

110. Gregg, K.J., et al., *Analysis of a new family of widely distributed metal-independent alpha-mannosidases provides unique insight into the processing of N-linked glycans.* J Biol Chem, 2011. **286**(17): p. 15586-96.

111. Hehemann, J.H., et al., *Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota.* Nature, 2010. **464**(7290): p. 908-12.

112. Information, N.C.f.B. *Basic Local Alignment Searach Tool* 2019-03-26]; Available from: https://blast-ncbi-nlm-nih-gov.ezproxy.uleth.ca/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome.

113. Seshadri, R., et al., *Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection.* Nat Biotechnol, 2018. **36**(4): p. 359-367.

114. Rossello-Mora, R. and R. Amann, *Past and future species definitions for Bacteria and Archaea.* Syst Appl Microbiol, 2015. **38**(4): p. 209-16.

115. Richter, M., et al., *JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison.* Bioinformatics, 2016. **32**(6): p. 929-31.

116. Software, J. [cited 2019 April 3]; Available from: http://jspecies.ribohost.com/jspeciesws/#analyse.

117. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.

118. Despres, J., et al., *Xylan degradation by the human gut Bacteroides xylanisolvens XB1A(T) involves two distinct gene clusters that are linked at the transcriptional level.* BMC Genomics, 2016. **17**: p. 326.

119. Martens, E.C., H.C. Chiang, and J.I. Gordon, *Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont.* Cell Host Microbe, 2008. **4**(5): p. 447-57.

120. Despres, J., et al., *Unraveling the pectinolytic function of Bacteroides xylanisolvens using a RNA-seq approach and mutagenesis.* BMC Genomics, 2016. **17**: p. 147.

121. Glenwright, A.J., et al., *Structural basis for nutrient acquisition by dominant members of the human gut microbiota.* Nature, 2017. **541**(7637): p. 407-411.

122. Hehemann, J.H.R., G.; Klassen, L.; Smith, A.D.; Ndeh, D.; Arnosti, C.; Amann, R.; Abbott, D.W., *Single cell fluorescence imaging of glycan uptake by intestinal bacteria.* ISME, 2019. **In Print**.

123. He, Z.E., Ferlisi, B., Eckert, E., Brown, H.E., Aguilar, A., and Steele, M.A., *Supplementing a yeast probiotic to pre-weaning Holstein calves: Feed intake, growth and fecal biomarkers of gut health.* Animal Feed Science and Technology, 2017. **226**: p. 81-87.

124. He, Z.X., et al., *Effect of wheat dried distillers grains and enzyme supplementation on growth rates, feed conversion ratio and beef fatty acid profile in feedlot steers.* Animal, 2015. **9**(10): p. 1740-6.

125. Rakoff-Nahoum, S., M.J. Coyne, and L.E. Comstock, *An ecological network of polysaccharide utilization among human intestinal symbionts.* Current Biology, 2014. **24**(1): p. 40-49.

126. Cartmell, A., et al., *A surface endogalactanase in Bacteroides thetaiotaomicron confers keystone status for arabinogalactan degradation.* Nat Microbiol, 2018. **3**(11): p. 1314-1326.

127. Sonnenburg, E.D., et al., *A hybrid two-component system protein of a prominent human gut symbiont couples glycan sensing in vivo to carbohydrate metabolism.* Proc Natl Acad Sci U S A, 2006. **103**(23): p. 8834-9.

128. Lynch, J.B. and J.L. Sonnenburg, *Prioritization of a plant polysaccharide over a mucus carbohydrate is enforced by a Bacteroides hybrid two-component system.* Mol Microbiol, 2012. **85**(3): p. 478-91.

129. Rogers, T.E., et al., *Dynamic responses of Bacteroides thetaiotaomicron during growth on glycan mixtures.* Mol Microbiol, 2013. **88**(5): p. 876-90.

130. Joglekar, P., et al., *Genetic Variation of the SusC/SusD Homologs from a Polysaccharide Utilization Locus Underlies Divergent Fructan Specificities and Functional Adaptation in Bacteroides thetaiotaomicron Strains.* mSphere, 2018. **3**(3).

131. Bankevich, A., et al., *SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing.* J Comput Biol, 2012. **19**(5): p. 455-77.

132. Gurevich, A., et al., *QUAST: quality assessment tool for genome assemblies.* Bioinformatics, 2013. **29**(8): p. 1072-5.

133. Altschul, S.F., et al., *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.

134. Rice, P., I. Longden, and A. Bleasby, *EMBOSS: the European Molecular Biology Open Software Suite.* Trends Genet, 2000. **16**(6): p. 276-7.

135. Cantarel, B.L., et al., *The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics.* Nucleic Acids Res, 2009. **37**(Database issue): p. D233-8.

136. Edgar, R.C., *MUSCLE: multiple sequence alignment with high accuracy and high throughput.* Nucleic Acids Res, 2004. **32**(5): p. 1792-7.

137. Darriba, D., et al., *ProtTest 3: fast selection of best-fit models of protein evolution.* Bioinformatics, 2011. **27**(8): p. 1164-5.

138. Stamatakis, A., *RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.* Bioinformatics, 2014. **30**(9): p. 1312-3.

139. Price, M.N., P.S. Dehal, and A.P. Arkin, *FastTree 2--approximately maximum-likelihood trees for large alignments.* PLoS One, 2010. **5**(3): p. e9490.

140. Rambaut, A., *FigTree v1. 3.1.* http://tree. bio. ed. ac. uk/software/figtree/, 2009.

141. Kearse, M., et al., *Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.* Bioinformatics, 2012. **28**(12): p. 1647-9.

142.   Tamura, K. and M. Nei, *Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees.* Mol Biol Evol, 1993. **10**(3): p. 512-26.

143.   Kumar, S., et al., *MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms.* Mol Biol Evol, 2018. **35**(6): p. 1547-1549.

144.   Kelley, L.A., et al., *The Phyre2 web portal for protein modeling, prediction and analysis.* Nat Protoc, 2015. **10**(6): p. 845-58.