A snapshot into the uptake and utilization of potential oligosaccharide prebiotics by probiotic lactobacilli and bifidobacteria as accessed by transcriptomics, functional genomics, and recombinant protein characterization

Andersen, Joakim Mark

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

A snapshot into the uptake and utilization of potential oligosaccharide prebiotics by probiotic lactobacilli and bifidobacteria as accessed by transcriptomics, functional genomics, and recombinant protein characterization

Ph.D. thesis (2012)

Joakim Mark Andersen

Enzyme and Protein Chemistry (EPC),

Department of Systems Biology,

Technical University of Denmark (DTU)

Supervisors:

R&D Director Dr. Rodolphe Barrangou, DuPont Nutrition and Health, WI, USA

Assoc. Prof. Maher Abou Hachem, EPC, Department of Systems Biology, DTU, DK

Prof. Todd Klaenhammer, Department of Food Science, North Carolina State University, NC, USA

Prof. Birte Svensson, EPC, Department of Systems Biology, DTU, DK

Preface

The present thesis summarizes the results of my Ph.D. project carried out in the Enzyme and Protein Chemistry group (EPC), Department of Systems Biology, Technical University of Denmark from February 2009 to May 2012 under supervision of Professor Birte Svensson and Associate Professor Maher Abou Hachem. The Ph.D. project was performed in collaboration with Dupont Nutrition and Health (former Danisco A/S) and Department of Food Science, North Carolina State University (NCSU) under the supervision of Dr Rodolphe Barrangou and Professor Todd Klaenhammer, respectively. The Ph.D. stipend was funded by DTU Systems Biology and DuPont (former Danisco A/S) and the project was supported by the Danish Strategic Research Council for the project "*Gene discovery and molecular interactions in prebiotics/probiotics systems. Focus on carbohydrate prebiotics*" (project no. 2101-07-0105) and the North Carolina Dairy Foundation for the work conducted at NCSU. Collaboration with Professor Hanne Frøkiær (Copenhagen University) was initiated to explore the immune-modulation of *L. acidophilus* NCFM although not completed the work outline is summarized.

The work of this Ph.D. project has resulted in the following publication and manuscripts:

Joakim Mark Andersen, Rodolphe Barrangou, Maher Abou Hachem, Sampo Lahtinen, Yong Jun Goh, Birte Svensson, Todd R. Klaenhammer (2011). Transcriptional and functional analysis of galactooligosaccharide uptake by *lacS* in *Lactobacillus acidophilus*. Proc Natl Acad Sci USA. 108: 17785−17790

Joakim Mark Andersen, Rodolphe Barrangou, Maher Abou Hachem, Sampo Lahtinen, Yong Jun Goh, Birte Svensson, Todd R. Klaenhammer. Transcriptional analysis of prebiotic uptake and catabolism by *Lactobacillus acidophilus* NCFM. Submitted to PLoS ONE.

Joakim Mark Andersen, Rodolphe Barrangou, Maher Abou Hachem, Sampo Lahtinen, Yong Jun Goh, Birte Svensson, Todd R. Klaenhammer. Mapping the uptake and catabolic pathways of prebiotic utilization in *Bifidobacterium animalis* subsp. *lactis* Bl-04 by differential transcriptomics. In preparation for BMC genomics.

Joakim Mark Andersen, Morten Ejby, Jonas Rosager Henriksen, Thomas Lars Andresen, Maher Abou Hachem, Birte Svensson. Dual substrate specificity of a prebiotic transporter from *Bifidobacterium animalis* subsp. *lactis* Bl-04. In preparation.

During the current Ph.D. project the following posters have been presented and are summarized in Appendix 6.6:

Andersen, J.M., Majumder, A., Fredslund, F., Ejby, M., van Zanten, G.C., Barrangou, R., Goh, Y.J., Lahtinen, S.J., Lo Leggio, L., Coutinho, P.M., Jacobsen, S., Abou Hachem, M., Klaenhammer, T.R., Svensson, B.: *Prebiotic galacto-oligosaccharide utilization by* Lactobacillus acidophilus *NCFM. Establishment of a methodological platform for protein discovery.* 7th Danish Conference on Biotechnology and Molecular Biology, Vejle (Denmark), May 2012.

Andersen, J.M., Barrangou, R., Abou Hachem, M., Svensson, B., Goh, Y., Klaenhammer, T.R.: *Gene induction patterns of prebiotic metabolic loci within* Lactobacillus acidophilus *NCFM*. Symposium for Biotechnological Research 2011, Kgs. Lyngby (Denmark), November 2011. (1st Poster prize)

Andersen, J.M., Barrangou, R., Abou Hachem, M., Svensson, B., Goh, Y., Klaenhammer, T. R.: *Gene induction patterns of prebiotic metabolic loci within* Lactobacillus acidophilus *NCFM*. 9th Carbohydrate Bioengineering Meeting, Lisbon (Portugal), May 2011.

Andersen, J.M., and Barrangou, R., Abou Hachem, M., Svensson, B., Goh, Y. and Klaenhammer, T.R.: *Transcriptional analysis of prebiotic utilization by* Lactobacillus acidophilus *NCFM*. American Society for Microbiology 110th General Meeting: San Diego (USA), May 2010.

Additionally, I have been co-authoring the following publications involving characterization of protein-carbohydrate interactions and bacterial utilization of candidate prebiotic:

Vigsnaes L.K., Nakai H, Hemmingsen L,**. Andersen J.M.**, Lahtinen S.J., Rasmussen L.E., Abou Hachem M., Petersen B.O., Duus J.Ø., Meyer A.S., Licht T.R., and Svensson B. In vitro *growth of individual human gut bacteria on potential prebiotic oligosaccharides produced by chemoenzymatic synthesis*. Submitted to J. Agric. Food Chem. (2012)

Abou Hachem M., Fredslund, F., **Andersen, J. M**., Larsen, R.J, Majumder, A., Ejby, M., Van Zanten, G., Lahtinen, S. J., Barrangou, B., Klaenhammer, T., Jacobsen, S., Coutinho, P.M., Lo Leggio, L., Svensson, B. (2011) *Raffinose family oligosaccharide utilisation by probiotic bacteria: insight into substrate recognition, molecular architecture and diversity of GH36 α-galactosidases*. Biocatalysis and Biotransformation (doi: 10.3109/10242422.2012.674717) (**Proceeding**)

Nielsen, M.M., Bozonnet, S.,Seo, E.S., Mótyán, J., **Andersen, J.M.**, Dilokpimol, A., Abou Hachem, M., Naested, H., and Svensson, B. (2009) "*Two Secondary Carbohydrate Binding Sites on the Surface of Barley α-Amylase 1 Have Distinct Functions and Display Synergy in Hydrolysis of Starch*". Biochemistry, 48: 7686–7697.

Nielsen, M.M., Seo, E.S., Dilokpimol A., **Andersen, J.M.**, Abou Hachem, M., Naested, H., Willemoës, M., Bozonnet, S., Kandra, L., Gyémánt, G., Haser, R., Aghajari, N., and Svensson, B. (2008) *"Roles of multiple surface sites, long substrate binding clefts, and carbohydrate binding modules in the action of amylolytic enzymes on polysaccharide substrates"*. Biocatal. Biotransform. 59–67.

Seo, E.-S., Nielsen, M.M., **Andersen, J.M.**, Vester-Christensen, M.B., Jensen, J.M., Christiansen, C., Dilokpimol, A., Abou Hachem, M., Hägglund, P., Maeda, K., Finnie, C., Blennow, A. & Svensson, B. (2008) *"α-Amylases. Interaction with polysaccharide substrates, proteinaceous inhibitors and regulatory proteins"*. Workshop on carbohydrate-active enzymes. Center for New Bio-Materials in Agriculture, Seoul National University, September 2008 (Ed. K.-H. Park, Woodhead Publishing Limited), pp. 20–36. (**Proceeding**)

Acknowledgement

First and foremost, I would like to thank my supervisors R&D Director Rodolphe Barrangou, Associate Professor Maher Abou Hachem, Professor Todd Klaenhammer and Professor Birte Svensson. They have been supervising through exceptional guidance, inspirational discussions and at times with the right amount of patience. A very special thanks to Birte Svensson for her commitment, before and during my Ph.D. project, allowing me to pursue an academic career in her research group.

I would like to thank my fellow colleagues on the 'FøSu project': Especially Ph.D. students Gabriella van Zanten, Anne Knudsen and Morten Ejby for kind collaborations, discussions and laughs. Former Post docs Folmer Fredslund and Avishek Majumder for inspirational academic work and good times in the office. Associate Professor Susanne Jacobsen (EPC, DTU), Senior Scientist Sampo Lahtinen (DuPont, Finland), Post Doc Louise K. Vigsnæs and Professor Tine R. Licht (National Food Institute, DTU), Assistant Professor Jonas Rosager Henriksen and Senior Researcher Thomas Lars Andresen (Department of Micro- and Nanotechnology, DTU) and Technician Anni Mehlsen and Professor Hanne Frøkiær (Department of Basic Sciences and Environment, Copenhagen University) are all thanked for their collaborations and discussions during my work.

Everybody at the TRK lab, Department of Food Science, North Carolina State University, are sincerely thanked for making my one year visit an unforgettable experience both science-wise and personally. A particular thank to Yong Jun Goh for her skilled guidance during preparation and analysis of the DNA microarray samples and constructions of gene deletion mutants.

All my dear colleagues, present and former, at the Enzyme and Protein Chemistry (Department of Systems Biology, DTU) are thanked for all the great moments during the last three years and making our laboratory a pleasant place to work, even in late evenings.

A warm thank to my friends and family, who have supported me through three busy years of ups and downs – mostly ups.

Summary

Microorganisms that when administered in sufficient amounts exert a beneficial effect to the host are defined as probiotics. The positive clinical effects of probiotics, mainly belonging to the *Bifidobacterium* and *Lactobacillus* genera in treatments of irritated bowel disorders, gut infections and lifestyle diseases are currently well documented. Selective utilization, of primarily non-digestible carbohydrates, termed prebiotics, by probiotics has been identified as an attribute of probiotic action, however the molecular mechanisms of prebiotics utilization and in particular the specificities of carbohydrate transporters and glycoside hydrolases that confer this remain largely unknown, limiting a robust understanding for the basis of selective utilization of known prebiotics and the discovery and documentation of novel prebiotics.

The aim of this Ph.D. thesis was to identify the genes involved with uptake and catabolism of potential prebiotics by the probiotics *Lactobacillus acidophilus* NCFM and *Bifidobacterium animalis* subsp. *lactis* Bl-04 as model organisms, using DNA whole genome microarrays and by *in silico* pathway re-construction to identify key genes for further functional analysis by gene deletions and recombinant protein characterization.

Transcriptional analysis was used to measure the global gene expression, in both bacteria, grown on glucose and various prebiotics and potential prebiotics covering diverse types of glycoside linkages and compositions: β-galacto-oligosaccharides, cellobiose, gentiobiose, isomaltose, panose, raffinose, stachyose and selected strain-specific potential prebiotics – *L. acidophilus* NCFM: barley β-glucan hydrolysate, lactitol, isomaltulose and polydextrose, while for *B. lactis* Bl-04: maltotriose, melibiose, xylobiose and xylo-oligosaccharides were used.

The differential transcriptional analysis of *L. acidophilus* NCFM revealed upregulation of genes encoding phosphoenolpyruvate-dependent sugar phosphotransferase systems mainly associated with disaccharide uptake, galactoside pentose hexuronide permease and ATP-binding cassette transporters were upregulated by dominantly oligosaccharides. Glycoside hydrolases from families 1, 2, 4, 13, 32, 36, 42, and 65 were found associated with the various transporters for carbohydrate catabolism.

The differential transcriptional analysis of *B. lactis* Bl-04 identified carbohydrate transporters of the major facilitator superfamily and galactoside pentose hexuronide permeases for disaccharide uptake and ATP-binding cassette transporters mainly for uptake of oligosaccharides. These transporters were found in gene clusters with glycoside hydrolases from families 1, 2, 13, 36, 42, 43 and 77.

Based on gene landscape analysis and the transcriptional findings, reconstruction of utilization pathways were done *in silico*. Hereafter the role of essential gene products in uptake of β-galacto-oligosaccharides putatively facilitated by a galactoside pentose hexuronide permease and the involvement of an ATP-binding cassette transporter and an α-galactosidase for uptake of raffinose family oligosaccharides and catabolism, respectively, were confirmed by gene deletion mutants in *L. acidophilus* NCFM.

The *B. lactis* Bl-04 homologous protein of the *L. acidophilus* NCFM raffinose specific solute binding protein displayed dual substrate specificity for raffinose family oligosaccharides and isomalto-oligosaccharides. The binding affinities (K_D) to a set of α-1,6 glycosides representing both classes of ligands were in the μM range, notably lower than typical values for oligosaccharide binding to solute binding proteins. The binding was enthalpically dominated and the lower affinity owed to a large unfavorable binding entropy suggestive of a high plasticity of the ligand binding site needed to accommodate different ligands varying in size, and monosaccharide composition, but recognizing a core structure comprising an α-D-(1,6)-linked galactose or its glucose C4 epimer. Biochemical characterization of the recombinant protein validated the broad substrate specificity, however the binding affinity was 100–1000 fold lower for the preferred substrates panose and raffinose, than seen for mono-specific carbohydrate transporters previously described although any biological implication of the weaken affinities is yet to be investigated.

In conclusion, differential transcriptomics revealed the global regulated gene response of *L. acidophilus* NCFM and *B. lactis* Bl-04 to potential prebiotic carbohydrates from which novel specificities for carbohydrate transporters and glycoside hydrolases were identified and validated through functional characterization. The work adds to the understanding of how probiotic bacteria can selective utilize prebiotics and how novel prebiotics can be discovered.

Dansk resumé (summary in Danish)

Mikroorganismer, der når de tilføjes i tilstrækkelig dosis udviser en positiv effekt på modtageren, er defineret som probiotika. Det er dokumenteret, at probiotika, hovedsageligt fra genera *Bifidobacterium* og *Lactobacillus*, kan anvendes i behandlingen af irriteret tyktarm, tarminfektioner og livsstilssygdomme. Selektiv udnyttelse af primært ufordøjelige kulhydrater, kaldet præbiotika, er én virkningsmekanisme benyttet af probiotika, dog er de specialiserede molekylære interaktioner primært specificiteter af kulhydrattransportører og glykosid-hydrolaser stortset ukendt. Den manglende viden begrænser opdagelsen samt anvendelsen af nye præbiotika.

Formålet med denne Ph.D. afhandling var at identificere de gener, som er involveret i optag og katabolisme af potentielle præbiotika i de probiotiske bakterier *Lactobacillus acidophilus* NCFM og *Bifidobacterium animalis* subsp. *Lactis* Bl-04, som model organismer, ved at benytte transkriptomanalyse og *in silico* rekonstruktion af metaboliske reaktionsveje for at kortlægge centrale gener til videre funktionel analyse ved hjælp af gen-deletioner og rekombinant proteinkarakterisering.

Transkriptionsanalyse blev anvendt til at måle det globale gen-udtryk i begge bakterier, dyrket med glukose og en række præbiotika samt potentielle præbiotika dækkende en bred vifte af glykosid-bindinger og glykosid-kompositioner: β-galacto-oligosakkarider, cellobiose, gentiobiose, isomaltose, panose, raffinose, stachyose samt udvalgte stamme-specifikke potentielle præbiotika – *L. acidophilus* NCFM: byg β-glykaner, lactitol, isomaltulose og polydextrose, imens de følgende blev benyttet til *B. lactis* Bl-04: maltotriose, melibiose, xylobiose og xylo-oligosakkarider.

Differential transkriptionsanalyse af *L. acidophilus* NCFM afslørede opregulering af gener kodende for phosphoenolpyrovatafhængige sukker-phospho-transferase systemer primært knyttet til optag af disakkarider. En galaktosid pentose hexuronid permease og ATP-bindende kasette transportører var opreguleret af hovedsageligt oligosakkarider. I tilknytning til de forskellige kulhydrattransportører var glykosid-hydrolaser involveret i kulhydratkatabolismen fra familierne 1, 2, 4, 13, 32, 36, 42 og 65 opreguleret.

Differential transkriptionsanalyse af *B. lactis* Bl-04 identificerede kulhydrattransportører klassificeret som 'Major facilitator superfamily' og galaktosid pentose hexuronid permeaser involveret i disakkaridtransport samt ATP-bindende kasette transportører primært for oligosakkaridoptag. Disse transportører blev fundet i genklynger med glykosid-hydrolaser fra familierne: 1, 2, 13, 36, 42, 43 og 77.

Reaktionsveje for oligosakkaridudnyttelse blev genskabt *in silico* ud fra gen-landskabsanalyse og de transkriptionelle resultater. Betydningen af tre formodede essentielle gener kodende for en galaktosid pentose hexuronid permease, en ATP-bindende kasette transportør samt en α-galaktosidase blev undersøgt for deres rolle i optag af β-galakto-oligosakkarider; for optag og katabolisme af raffinose-lignende oligosakkarider eftervist ved gen-deletioner i *L. acidophilus* NCFM.

Det homologe *B. lactis* Bl-04 protein af det raffinose specifikke *solute binding* protein fra *L. acidophilus* NCFM udviste en dobbelt substratspecificitet for raffinose-lignende oligosakkarider og isomaltooligosakkarider. Bindingsaffiniteterne (K_D) for et sæt af α-1,6-glykosider, der repræsentere begge typer af ligander, var i μM skala, hvilket er mærkbart lavere end typiske værdier for oligosakkarid binding til andre *solute binding* proteiner. Bindingen var entalpisk drevet og den lavere affinitet skyldtes en større ufavorabel entropi. Dette var muligvis resultat af ligang-bindingslommen ændrede form, som krævet for at binde de forskellige ligander varierende i længde og glykosid komposition, selvom α-D-(1,6) galaktose, eller den C-4 epimere glukose, blev genkendt i bindingslommen.

Biokemisk karakterisering af rekombinant protein validerede den brede substratspecificitet, dog var affiniteten 100–1000 gange lavere for de fortrukne substrater panose og raffinose end for tidligere beskrevet mono-specifikke kulhydrattransportører omend det endnu ikke er undersøgt om de lavere affiniteter har biologisk relevans.

Som konklusion har differentiel transkriptionsanalyse vist det globalt regulerede genudtryk af *L. acidophilus* NCFM og *B. lactis* Bl-04 i forhold til potentielle præbiotiske kulhydrater, hvorfra nye specificiteter bekræftet ved funktionel karakterisering er fundet for kulhydrattransportører og glykosid-hydrolaser. Dette studie tilføjer forståelse af hvordan probiotiske bakterier selektivt kan omsætte præbiotika samt hvordan nye præbiotika kan udvælges.

List of abbreviations

ABC	ATP-binding cassette transporters
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
AXOS	Arabinose-decorated β-xylo-oligosaccharides
CAZy	Carbohydrate active enzymes
CBM	Carbohydrate binding modules
FOS	Fructo-oligosaccharides
GH	Glycoside hydrolase
GIT	Gastrointestinal tract
GOS	β-galacto-oligosaccharides
GPH	Glycoside-pentoside-hexuronide
IMO	Isomalto-oligosaccharides
Mb	Mega-basepairs
MFS	Major facilitator superfamily
Msm	Multiple sugar metabolism
N.I.	Not investigated
PCR	DNA Polymerase chain reaction
PTS	Phosphoenolpyruvate phosphotransferase
RFO	Raffinose family oligosaccharides
SOE-PCR	Splicing by overlap extension-PCR
TC	Transporter classification
XOS	β-xylo-oligosaccharides

Table of Content

Preface III

Acknowledgement VII

Summary IX

Dansk resumé (summary in Danish) XI

List of abbreviations XIII

Table of Content XV

1 Introduction 1

1.1 Beneficial modulation of the human gastrointestinal tract 1

1.1.1 Probiotics 1

1.1.2 Prebiotics 3

1.1.3 Health, dietary and commercial aspects of pre- and probiotics 7

1.2 The gastrointestinal tract and beyond: the probiotic perspective 7

1.2.1 The gastrointestinal tract as a microbial habitat 7

1.2.2 Homeostasis and microbial diversity in the gastrointestinal tract 8

1.2.3 Molecular mechanisms of probiotic functions 9

1.3 Genomics and phylogenetics of probiotics 10

1.3.1 *Lactobacillus* 10

1.3.1.1 *Lactobacillus acidophilus* NCFM 11

1.3.2 *Bifidobacterium* 12

1.3.2.1 *Bifidobacterium animalis* subsp. *lactis* Bl-04 12

1.4 Molecular elements of pre- and probiotic interactions 13

1.4.1 The paradigm of non-digestible carbohydrate utilization by the gut microbiome 13

1.4.2 Bacterial carbohydrate transport systems 15

1.4.3 Classification and distribution of carbohydrate active enzymes 18

1.5 Experimental methods for gene and protein identification of probiotic properties 20

1.6 Scientific basis and objectives for the current project 21

2 Materials and methods 23

2.1 Databases, prediction and modeling tools 23

2.2 Construction of phylogenetic tree of carbohydrate solute binding proteins 23

2.3 Experimental design of DNA microarray setup. 24
2.4 Generation of *L. acidophilus* NCFM gene deletion mutants....25
2.5 Crystallization setup of recombinant *Bl*16GBP....27
3 Results and discussion....29
3.1 Bioinformatics assessment and comparison of potential prebiotic utilization by *L. acidophilus* NCFM and *B. animalis* subsp. *lactis* Bl-04....29
3.1.1 Mapping of potential prebiotic utilization systems....29
3.1.2 Genome-mining and assessment of the utilization of potential prebiotics by *L. acidophilus* NCFM....30
3.1.3 Function deduction and interplay of genes identified within *L. acidophilus* NCFM 34
3.1.4 Genome-mining and assessment of the utilization of potential prebiotics by *B. animalis* subsp. *lactis* Bl-04 35
3.1.5 Functional deduction from gene identification within *B. animalis* subsp. *lactis* Bl-04 38
3.1.6 Comparative genomics of potential prebiotic utilization of *L. acidophilus* NCFM and *B. animalis* subsp. *lactis* Bl-04....39
3.1.7 Functional map of carbohydrate-specific solute binding proteins of ABC transporters 41
3.2 Transcriptional analysis of potential prebiotic utilization....45
3.2.1 Selection of potential prebiotics for transcriptomics analysis....45
3.2.2 Summary of potential prebiotics induced differential transcriptomics....46
3.2.3 Comparative pathway analysis of potential prebiotic utilization by *L. acidophilus* NCFM and *B. lactis* Bl-04....48
3.2.4 Correlation of *in silico* predictions and transcriptional observations 50
3.2.5 Experimental design and limitations of DNA microarrays for gene identification.51
3.3 Functional characterization of genes and proteins involved with potential prebiotic uptake and catabolism 52
3.3.1 Functional genomics of *L. acidophilus* NCFM 52
3.3.2 Structure-function relationship of a solute binding protein with dual substrate specificity 55
4 Conclusions and perspectives....57
5 References 59

6 Appendices 81

6.1 Transcriptional analysis of prebiotic uptake and catabolism by *Lactobacillus acidophilus* NCFM.....83

6.2 Mapping the uptake and catabolic pathways of prebiotic utilization in *Bifidobacterium animalis* subsp. *lactis* Bl-04 by differential transcriptomics.....121

6.3 Transcriptional and functional analysis of galactooligosaccharide uptake by *lacS* in *Lactobacillus acidophilus*.....155

6.4 Dual substrate specificity of a prebiotic transporter from *Bifidobacterium animalis* subsp. *lactis* BL-04 163

6.5 Ongoing collaborative work.....197

6.6 Posters contributions.....201

1 Introduction

The use of probiotic microorganisms for improvement of human health (1) has been clinically well-documented within the recent years as reviewed in the following. Probiotics can be supplemented with selectively metabolized prebiotics, mainly carbohydrates, for synergistic effects (2).

Advances within genomics have shed new light into the diversity and functions of the human gastrointestinal tract (3) pushing for more defined consideration in design of probiotic treatments (4). In this context, next-generation probiotic products are estimated to be knowledge-driven, focused on the molecular mechanism of their effects (5) including deeper understanding of selective prebiotic metabolism. Hence it is the purpose of the following sections to introduce pre- and probiotics, their role in the gastrointestinal tract and their mechanisms of actions leading to molecular understanding of the protein facilitating prebiotic uptake and catabolism. This sets the stage for presentation of differential transcriptomics for gene identification induced by potential prebiotics, and the targeted functional genomics and recombinant protein work for characterization of carbohydrate transporters. The presented work adds to the fundamental understanding of probiotic bacteria in particular with respect to their utilization of carbohydrate prebiotics by pathway mapping, comparative gene landscaping of identified genetic loci and biochemical characterization of a novel dual specific carbohydrate transporter.

1.1 Beneficial modulation of the human gastrointestinal tract

1.1.1 Probiotics

Microorganisms positively modulating human health have been long known and were pioneered by the studies of Ilya Mechnikov in the early 20th century (6). The initial work focused on fermented milk containing lactic acid bacteria and their impact on human health and longevity. Subsequent scientific investigations established the knowledge on positive health impact elicited by supplementing diet with beneficial microorganisms and this eventually led to the first definition of probiotics (7) which was followed-up by WHO's definition of probiotics as: "*live*

microorganisms, which when administered in adequate amounts, confer a health benefit on the host." (8). Continuous development in the field produced numerous supporting studies addressing molecular mechanistic and clinical work, which consolidated the above definition of probiotics into a state of consensus lead by the scientific community for continuously critical review (1). Following regulatory demands from governmental agencies, regarding documentation and increased substantiation of probiotic related health claims for commercial products containing probiotics (9, 10). The scientific focus changed to go beyond clinical studies into the areas of genetics and molecular mechanisms of action for probiotics interactions by techniques such as proteomics, transcriptomics, functional genomics and recombinant protein characterization (5). In the following sections probiotics will be introduced, with emphasis on their synergetic effects with carbohydrate prebiotics and role in the gastrointestinal tract leading to the molecular level genomics and identification of genetic loci involved with observed probiotic effects and subsequently including the corresponding gene products and mechanism of action.

Probiotic organisms have been found within the following microbial genera: *Bifidobacterium*, *Enterococcus*, *Escherichia, Lactobaccilus*, *Lactococcus*, *Leuconostoc*, *Pediococcus*, *Ruminococcus*, *Saccharomyces* and *Streptococcus* (11). Yet the main topic for probiotic research, as it is also the case of the present work, has been the bifidobacteria and lactobacilli. Table 1-1 summarizes some of the investigated beneficial roles that probiotic bifidobacteria and lactobacilli exert on the host upon supplementation. It is recognized through interventions for modulation of the microbial activity within the human gastrointestinal tract (GIT) that probiotics have impact on the immune system throughout life and can target illnesses described as major health issues such as dietary related lifestyle diseases and bowel disorders, colonic cancer and infant malnutrition (12–14).

Table 1-1: Selected probiotic clinical trials in humans. The types of probiotic modulation are divided into: short term treatments covering mainly acute infection, age influenced treatments and chronic diseases within the GIT to illustrate the broadness of positive interventions.

Type of probiotic modulation	Intervention focus	References
Short term treatment	Viral infections	(15, 16)
	Urinal tract infections	(17, 18)
	Allergies	(19, 20)
	Acute diarrhea	(21, 22)
Age dependent grouping of subjects	Preterm newborn	(23)
	Infants with diarrhea	(24)
	Elderly and improve immunity	(25)
Irritated bowel syndromes	Ulcerative colitis	(26, 27)
	Crohn's disease	(28, 29)
	Colonic cancer prevention	(30, 31)

The mechanism of probiotic actions include functions as bacterial bulking agent hence hampering colonization potential of opportunistic pathogens and exogenous microorganisms (32), production of secondary metabolites such as short chain fatty acids stimulating the epithelial cell metabolism and turnover (33), acidification of the local GIT environment to suppress viability of undesirable lesser acid tolerant microorganisms (34), and the modulation of the host immune system (35). By the current understanding of probiotics, there is a need to correlate the underlying mechanisms linked to probiotic proliferation and activity in the GIT with the beneficial effects (36). This includes highlighting the selective stimulation by dietary food components bypassing the host digestive breakdown and uptake systems for entry into the colon where the non-digestible fraction, dominantly carbohydrates, is selectively utilized by a sub-population of the microbes residing in the GIT (37).

1.1.2 Prebiotics

The microbial inhabitants of the GIT rely on components of the human diet not digested in the upper GIT of the host. This fraction is mainly oligo and polysaccharides of plant origin (38) which are utilized by the diverse community of microbes residing in the lower GIT. Thus some of these carbohydrates (both oligosaccharides and polysaccharides) were found to be preferentially utilized by probiotic microorganisms with subsequent increase in their cell

numbers and activity in the GIT and hence the term prebiotics was coined to describe this category of compounds (11). Prebiotic lipids and proteins have been mentioned in the literature but are sparsely reported (11) while more recently plant isoflavones (39) and polyphenols (40) have been suggested to exhibit prebiotic effects. The current definition set three criteria for prebiotics (41):

1. Resistance to gastric acidity, hydrolysis by mammalian enzymes and gastrointestinal absorption
2. Ability to be fermented by intestinal microbiota
3. Selective stimulation of the growth and/or activity of intestinal bacteria associated with health and wellbeing

The selective metabolism of prebiotics and selected studies supporting their impact on human health are summarized in Table 1-2. The chemical structures of the prebiotics listed in Table 1-2 and in the following section are summarized in Table 1-4 showing diversity of glycosidic linkages and composition of prebiotics. Numerous *in vivo* human and animal studies have investigated the proposed prebiotic effects. Yet to date, documentation has only been obtained sufficiently for a few carbohydrates to grant a status as prebiotics, namely: β-galacto-oligosaccharides (GOS), lactulose, fructo-oligosaccharides (FOS) and inulin (2). The level of documentation obtained through clinical studies to grant the status of prebiotic is a matter of debate (42). As this debate is yet unsettled, this work will adopt the currently accepted and previously published classification of prebiotics, although selective utilization pathways for novel candidate prebiotic will be proposed later, potentially acting as supportive claims for prebiotics classification.

Dual supplements of pro- and prebiotic, termed synbiotics, have shown great potential for increased efficiency of GIT associated disorders (43–47) substantiating the selective metabolism by probiotics leading to improvement of GIT treatments.

Novel potential prebiotics have been proposed (gentiobiose, panose, polydextrose, raffinose family oligosaccharides (RFO)) although mainly based on *in vitro* methodology hence making more studies needed to fully document them as prebiotics based on *in vivo* studies (42, 58). Table 1-3 summarizes selectively fermented potential prebiotics.

Table 1-2: Selected studies of prebiotic effects *in vivo*.

Prebiotic	Observed positive effect	Reference
FOS	Host immune modulation	(48)
FOS	Prevention of diarrhea	(49)
FOS	Increased fecal bifidobacteria counts	(50)
FOS	Treatment of Crohn's disease	(51)
GOS	Increased bifidobacteria counts	(52)
GOS	Prevention of diarrhea	(53)
GOS	Review of clinical trials	(54)
GOS/FOS	Modulation of vaccine response in mice	(55)
Inulin	Increased bifidobacteria counts	(56)
Inulin	Review of immune-modulation	(48)
Lactulose	Review of prebiotic effects	(57)

Table 1-3: Selected studies proposing novel potential prebiotics.

Carbohydrate	Screening system	Observed effect	Reference
Polydextrose	Fecal fermentations	Improved short chain fatty acid profile production	(59, 60)
Polydextrose	Humans	Increased bifidobacteria and lactobacilli counts	(61)
RFO	Humans	Selective fermentation by bifidobacteria	(62)
RFO	Rats	Improved mineral uptake	(63)
AXOS	Rats	Improved mucin turnover	(64)
XOS	Humans	Increased bifidobacteria and lactobacilli counts	(65)
IMO	Humans	Review of clinical trials	(66)
IMO	Humans	Increased lactobacilli count in rats	(67)
IMO	Humans	Bowel functions in elderly	(68)
Panose	Fecal fermentations	Increased bifidobacteria counts	(58)
Lactitol	Humans	Immuno-modulation	(69)
Xylitol	Fecal fermentations	Improved short chain fatty acid profile production	(59)
Gentio-oligosaccharides	Fecal fermentations	Fermentation by bifidobacteria and lactobacilli	(70, 71)
Pectic-oligosaccharides	Fecal fermentations	Some selective fermentation by bifidobacteria	(72, 73)
α-manno-oligosaccharides	Weaning pigs	Positive immune-modulation and reduced diarrhea	(74, 75)

Dietary prebiotics elicit both a proliferational effect and enhanced activity of probiotics, yet prebiotics are not only linked to selective metabolism. It has been shown *in vitro* how GOS may prevent adhesion of pathogenic *Vibrio cholera* and *Cronobacter sakazakii* to protein receptors on

the surface of epithelial cells (76–78) and inhibit adhesion of *Salmonella enteric* serovar *Typhimurium* via murine enterocytes (79). These effects have been explained by GOS structurally mimicking human surface glycoproteins (80) where prebiotics may also act as decoys for pathogens hence reducing their adhesion to the mucosa barrier being a first step in infection (81). Similar observations for pathogenic inhibition of adhesion has been reported for XOS (82).

Despite the promising nature of prebiotics, recent studies highlight a controversy regarding the use of prebiotics: The capabilities of potential pathogenic *Listeria monocytogenes* to utilize known prebiotics (83), increase the severity of *Salmonella enteric* serovar *Typhimurium* infections when stimulated with prebiotics (84) and the identification of genetic loci enabling FOS utilization, have been identified in *E. coli* (85, 86). These observations reflect a shortage of sufficient documentation of pre- and probiotics (42), motivating discovery of novel prebiotics, combined with further understanding of molecular mechanism of selective prebiotic metabolism and how this applies to the microbiome in the GIT.

Table 1-4: Chemical structures of reported prebiotics and potential prebiotics. The size ranges of each oligosaccharide are defined as footnotes below with the range of raffinose oligosaccharides listed with common names. The size distributions are reported in the respective references to each oligosaccharide listed for documented prebiotics in Table 1-2 and candidate prebiotics in Table 1-3.

Common names	Chemical Structure
Fructo-oligosaccharides (FOS)	[β-D-Fru*f*-(1-2)]$_a$ -(β2,α1)-D-Glc*p*
Inulin	As FOS with a > 20
β-galacto-oligosaccharides (GOS)	[β-D-Gal*p*-(1–4)]$_b$-D-Glc*p*
Lactitol	β-D-Gal*p*-(1–4)-D-Glc-ol
Lactulose	β-D-Gal*p*-(1–4)-D-Fru*f*
Raffinose family oligosaccharides (RFO)	[α-D-Gal*p*-(1–6)]$_c$- D-Glc*p*-(α1,β2)-D-Fru*f*
Melibiose	α-D-Gal*p*-(1–6)- D-Glc*p*
Isomalto-oligosaccharides (IMO)	[α-D-Glc*p*-(1–6)]$_d$-D-Glc*p*
Panose	α-D-Glc*p*(1–6)-α-D-Glc*p*-(1–4)-D-Glc*p*
Gentio-oligosaccharides	[β-D-Glc*p*-(1–6)]$_e$-D-Glc*p*
β-xylo-oligosaccharides (XOS)	[β- D-xyl*f*-(1–4)]$_f$-D-xyl*f*
Arabinose-decorated XOS (AXOS)	α-D-Ara*f*-(1–2) and/or α-D-Ara*f*-(1–3) linked to XOS
Xylitol	β- D-xyl*f*-(1–4)-D-xyl-ol
Polydextrose	Primarily mixed α-glucans, DP=2–30
Pectic-oligosaccharides	[α-L-Rha*p*-(1,2)-α-D-GalA*f*-(1,4)]$_g$
Manno-oligosaccharides	[α- D-Man*p*-(1–4)]-D-Man*p*

a = [1–5]; b = [1–5]; c = [1=raffinose, 2=stachyose, 3=verbascose]; d = [1–5]; e = [1–5]; f = [1–7]; g = [1–4]

1.1.3 Health, dietary and commercial aspects of pre- and probiotics

The above definitions of pro- and prebiotics are well documented. Pre- and probiotics, however are although not confined to humans, indeed markets are emerging in animal feed, driven by regulations to reduce application of antibiotics in livestock production (87) and aquaculture (88), showing widespread use of pre- and probiotics.

The total global sale for probiotic products reached 15.9 billion $ (12.4 billion €) in 2008 and the average annual growth rate has been estimated to around 7% (89) with the main consumer markets being Northern America (90), Japan (91) and Europe (92).

To ensure the efficiency of probiotic products, regarding levels and activity of supplemented functional ingredients and consumers' safety based on manufacturers claims, the application of probiotic supplemented commercial dietary products is being regulated by governmental agencies (*Food and Drug Administration* in Northern America and *European Food Safety Authority* within the European Union) (10). New regulations, as exemplified by the EU regulation No. 1924/2006 (93), have increased the level of documentation needed to propose claims of efficacy for probiotic products emphasizing the needs for further documentation on specific applications of pre- and probiotics (9, 94, 95).

1.2 The gastrointestinal tract and beyond: the probiotic perspective

1.2.1 The gastrointestinal tract as a microbial habitat

The main function of the human gastrointestinal tract (GIT) is to process, digest and absorb nutrients from the diet to supply energy to the various organs of the body. This requires an interplay of α-amylases, acidification, bile salts, proteases and lipases all secreted by the human digestive system through a strict compartmentalization of the GIT, yet these processes are by far insufficient to handle the complexity of the diet and provide nutritional requirements for the human body. Through evolution the lower GIT (colon) has evolved to become a niche habitat (Figure 1-1) for commensal microorganisms (the GIT microbiome in generic terms) dependent on commensalisms and in some cases mutualism with the host or negatively affected under specific conditions by opportunistic pathogenesis (96).

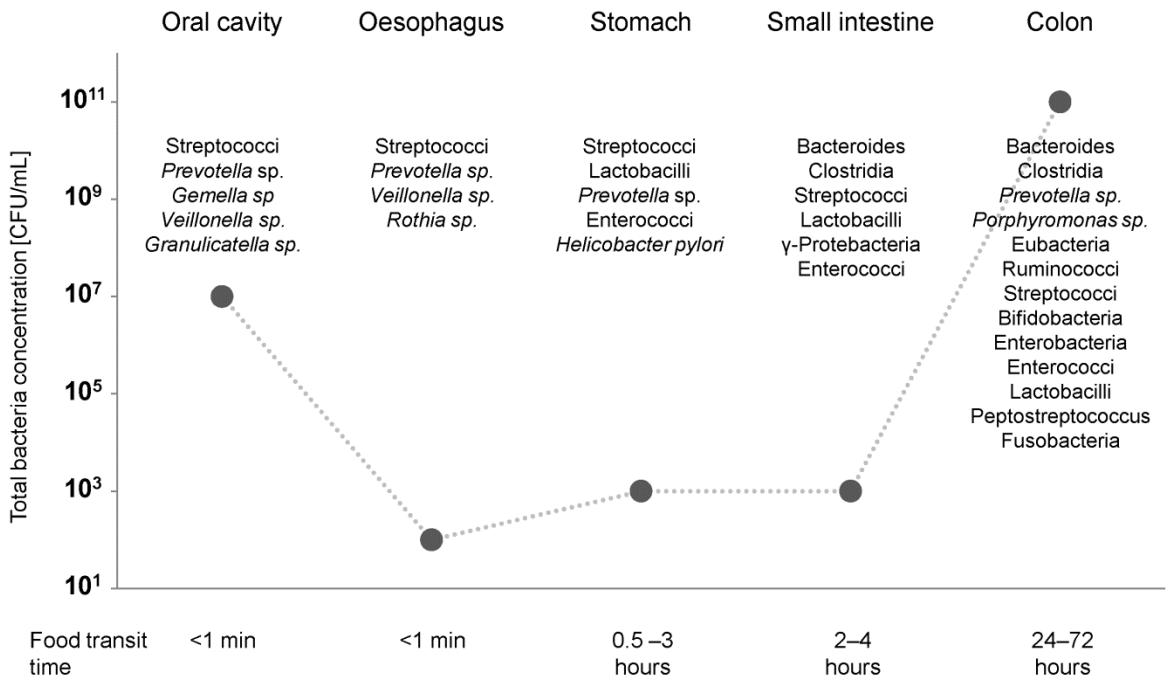


Figure 1-1:Compartments of the GIT, the bacterial genera predominantly found herein and the food transit time showing fermentation in the colon by the duration of food transit enabling microbial colonization (Modified from (97)).

Among the many roles of the GIT microbiota, catabolism of non-digestible components of the human diet is one of the best examples of commensalisms (98, 99). This is illustrated by the estimated 100–150 fold higher number of genes found from sequencing of the GIT microbiome compared to the human genome (100) showing metabolic capabilities beyond the genetic information stored in the human genome and thus how harboring a microbiota can be very advantageous for biosynthesis of vitamins and detoxication of xenobiotics (101), and at times challenging by imbalance of the microbiome leading to inflammatory bowel syndromes (3).

1.2.2 Homeostasis and microbial diversity in the gastrointestinal tract

Given the wide metabolic capabilities of the GIT microbiome and interplay with the host, it is a crucial area of research to classify the microbial span and phylogenetics of the GIT microbiota. This work has been greatly aided with advances within high-throughput sequencing and

bioinformatics in the recent decade (102). A key area in understanding the mutualism and pathogenesis in the GIT is the attempt to classify what can be regarded as a stable microbiome in homeostasis with the host as discussed by Sansonetti and Medzhitov (103) and the measurable changes observed in disease states of the GIT (3). The GIT bacterial composition at the phylum level was recently mapped by ribosomal 16S RNA sequencing to highlight the major phyla and their diversity, dominated by Bacteroidetes and Firmicutes in adults (102). Building from the established reference of the GIT microbiota in homeostasis it has been possible to map the microbiome establishment in neonates, the maturation into adulthood (104) and the decline and changes induced by aging (105). This included incorporating a marked individual variation in microbiome composition resulting from genetic variation and environmental differences of the geographical habitat (106) combined with long-term diet influence (107). The above factors allow microbiome differentiation into enterotypes, which can be applied to link the differences of the microbiome into functional groups and disease states respectively (108) for targeted treatments with probiotics.

Thanks to the gained insight into the diversity and dynamics of the GIT microbiome as outlined above, greater understanding of how disease states arise and negatively modulate the GIT has been obtained (109–112). In turn this can be developed and validated to enable targeted pre- and probiotic treatment (113) taking into account the multitude of mechanism of actions as listed above (Section 1.1.1 and 1.1.2) and highlighted in the following section.

1.2.3 Molecular mechanisms of probiotic functions

Probiotics' mechanism of actions include competitive exclusion of opportunistic pathogens (114, 115), production of secondary metabolites for acidification of the GIT (34) or short chain fatty acids such as propionate which is absorbed through the epithelia and stimulate lower lipogenesis (116). Furthermore proteinacious products such as bacteriocins may inhibit pathogen colonization (117, 118) and bacterial membrane associated proteins (119) or cell-membrane anchored lipids (120) may regulate the host immune response. Notably, prebiotics were indicated to enhance the probiotic functions, beyond stimulating selective growth of probiotics, e.g. FOS mediated increased bacteriocin production (121) and utilization of various prebiotics were shown to increase stress resistance, e.g. oxidative stress, in lactobacilli (122). The following sections

will present the genomics of probiotics, and genetic loci encoding proteins involved with prebiotic utilization, to emphasize the molecular mechanisms of pre- and probiotic interactions.

1.3 Genomics and phylogenetics of probiotics

The advances within bacterial genomics have been essential to link phylogenetics and functional studies leading into a systems biology perspective (3, 123–125). In this light, genomic analysis has proved to be valuable for high through-put screening of bacterial genomes to identify markers of antibiotic resistance and virulence factors within the GIT (126) and comparative assessment of bacterial phage resistance (127) or loci specifically linked to probiotics as reviewed by Ventura et al. (128).

Genome-scale analysis of probiotic strains focusing on bifidobacteria and lactobacilli has been extensively reviewed to highlight genomic features promoting the adaption to the GIT and probiotic mechanism of action (129, 130). In relation to prebiotics, a common feature of probiotic genomes is the large proportion (15–20% of the total numbers of genes) of putatively carbohydrate metabolism genes (131–133). The following will introduce the genera *Lactobacillus* and *Bifidobacterium*, which despite functional similarities in the GIT show phylogenetic distant diversity (128, 134).

1.3.1 Lactobacillus

The *Lactobacillus* genus consists of Gram-positive, low genomic C+G content, acid-tolerant, non-sporulating, aero-tolerant or anaerobic bacteria (135). Lactobacilli have been isolated mainly from natural food fermentations of diary, meat and plants, and from the intestine of animals giving rise to a significant industrial potential for fermented foods products, starter cultures and probiotics (136). Genomic analysis of lactobacilli showed strain differentiation to depend on the original habitat (137) where the GIT associated lactobacilli mainly comprise *L. acidophilus*, *L. gasseri*, *L. johnsonii* and *L. casei* (138). Comparative analysis of strains isolated both from plant material and GIT have highlighted genomic loci associated with niche adaption (129) leading to increased understanding of the protein facilitated mechanisms underlying probiotic actions of lactobacilli from the GIT (139, 140).

The combined approach of genomics (selectively summarized in Table 1-5), functional genomics (141, 142) and high-throughput gene identification (143) has substantiated the potential for identification of genetic loci specific for prebiotic utilization (144).

Table 1-5: Selected probiotic *Lactobacillus* strains. Genome sizes are given in mega-basepairs (Mb)

Strain	Genome size (Mb)	Accession number	Reference
L. acidophilus NCFM	2.0	NC_006814.3	(131)
L. rhamnosus GG	3.0	NC_013198.1	(145)
L. plantarum WFCS1	3.1	NC_004567.1	(146)
L. johnsonii NCC553	2.0	NC_005362.1	(147)
L. gasseri 33323	1.9	NC_008530.1	(148)
L. casei BL23	3.1	NC_010999.1	(149)
L. crispatus ST1	2.0	NC_014106.1	(150)

1.3.1.1 *Lactobacillus acidophilus* NCFM

L. acidophilus NCFM has been reported as a probiotic in clinical studies (44, 151, 152) and in combination with oligosaccharide prebiotics (69). The probiotic character have been analyzed by functional studies to reveal the molecular mechanisms for important probiotic traits, such as bile acid resistance (153, 154), cell adhesion (155, 156) and involvement of lipoteichoic acid in immunomodulation (120). The carbohydrate uptake and catabolism genes comprise 17% of the *L. acidophilus* NCFM genome (131). Broad carbohydrate utilization of *L. acidophilus* NCFM was demonstrated and included transporters for trehalose (141), fructo-oligosaccharides (142), and several mono-, di- and tri-saccharides (143, 157). Yet the potential for *in silico* identification of genes involved in prebiotic utilization is still hampered by the lack of functional studies within strains of lactobacilli harboring multitudes of transporters and families of glycoside hydrolases (137, 158).

The following experimental work (Appendices 6.1 and 6.2) focuses on *L. acidophilus* NCFM as an important representative of the *acidophilus* cluster of GIT associated lactobacilli (159) to elucidate novel potential prebiotic utilization.

1.3.2 *Bifidobacterium*

Bifidobacteria are non-motile, non-sporulating and non-gas producing, anaerobic high genomic C+G Gram-positive, bacteria from the *Actinobacteria* phylum, (160, 161). Bifidobacteria are mainly isolated from ecological niches associated with the human (or animal) GIT indicating their significance in the microbiome (162). Phylogenetic analysis showed clustering of bifidobacteria into the following groups: *B. asteroids, B. adolescentis, B. longum, B. pollorum, B. boum* and *B. pseudolongum* (162, 163) where the latter group harbors the *B. animalis* subsp. *lactis* taxon, utilized commercially for its probiotic characteristics (164).

To date, 53 bifidobacterial genomes are publicly available (May 2012) and selected bifidobacteria strains associated with probiotic effects are listed in Table 1-6. General size of the genomes ranges from 1.9 to 2.9 Mb and display an overall low level of genomic diversity (165) with a core of 1000 common genes estimated from pan-genomics (166).

Table 1-6: Selected genomes of probiotics strains of bifidobacteria.

Strain	Genome size (Mb)	Accession number	Reference
B. animalis subsp. *lactis* Bl-04	1.9	NC_012814	(167)
B. adolescentis ATCC15703	2.1	NC_008618	Only in database
B. longum subsp. *infantis* ATCC15697	2.8	NC_011593	(168)
B. longum subsp. *longum* NCC2705	2.3	NC_004307	(169)
B. bifidum PRL2010	2.2	NC_014638	(170)
B. breve UCC2003	2.4	CP000303.1	(171)

The GIT adaption, linked to defined genetic loci within bifidobacteria has been proposed for complex dietary carbohydrate utilization (169, 172, 173). Alternative adaption, although not yet considered a probiotic characteristic, has been identified for mucin degradation and utilization of *B. bifidum* strains (170). Furthermore colonization of *B. dentium* strains as opportunistic cariogenic pathogens in the oral cavity as been reported through the mechanisms of adhesion and utilization of human saliva-derived compounds (174).

1.3.2.1 *Bifidobacterium animalis* subsp. *lactis* Bl-04

B. animalis subsp. *lactis* strains have documented effects as probiotics (175–179) and are widely used in commercial products (180). Characterization of *B. lactis* strains on the genomics and

molecular level (181–183) has identified elements potentially conferring probiotic characteristics such as oxidative stress tolerance (184, 185), XOS utilization (186) and bile resistance (187).

Although clinical well-documented and with excessive pan-genomic data available showing multiple genes putatively involved with prebiotic utilization, functional work is lacking to substantiate the pre- and probiotic interactions (188–190). The present work focuses on the strain *B. lactis* Bl-04 (167) as a representative of the highly important probiotic *B. animalis* species.

1.4 Molecular elements of pre- and probiotic interactions

1.4.1 The paradigm of non-digestible carbohydrate utilization by the gut microbiome

The impact of the GIT microbiome on the digestion of dietary polysaccharides in the colon, primarily plant cell wall material and starch, has been studied to explain the functionalities of host non-digestible carbohydrates and their metabolic effects on the host (191). Genomic analysis of GIT associated sub-groups of commensal bacteria, exemplified by *Bacteroides thetaiotaomicron* harboring 88 putative polysaccharide utilization loci mainly specific for mucin *O*-glycan utilization (192), showed strains harboring extensive genes encoding secreted hydrolytic enzymes for polysaccharide breakdown (193). Functional studies of selected polysaccharide utilizing GIT bacteria such as *Roseburia inulinivorans* (194) identified gene clusters and pathways for inulin and starch utilization and xylan utilization in *Prevotella bryantii* (195). Other GIT microbes, and most bifidobacteria and lactobacilli (128) do not encode enzymes for polysaccharide utilization to the same extent, but have rather evolved symbiotic relations for cross-feeding of polysaccharide breakdown products (38). The essential interplay of carbohydrate hydrolyzing enzymes and carbohydrate transporter have been reviewed (191, 196), and a schematic overview of the current understanding regarding microbial utilization of host non-digestible dietary polysaccharides is shown in Figure 1-2.

Notably, identification of genetic loci encoding transport systems and carbohydrate hydrolytic enzymes showed gene organization in operons and clusters (197) which are tightly regulated on the Transcriptional level (143, 198–200). This prompts analysis of oligosaccharide transport

systems and genetically associated hydrolytic enzymes to provide a functional rationale for *in silico* predictions of these loci.

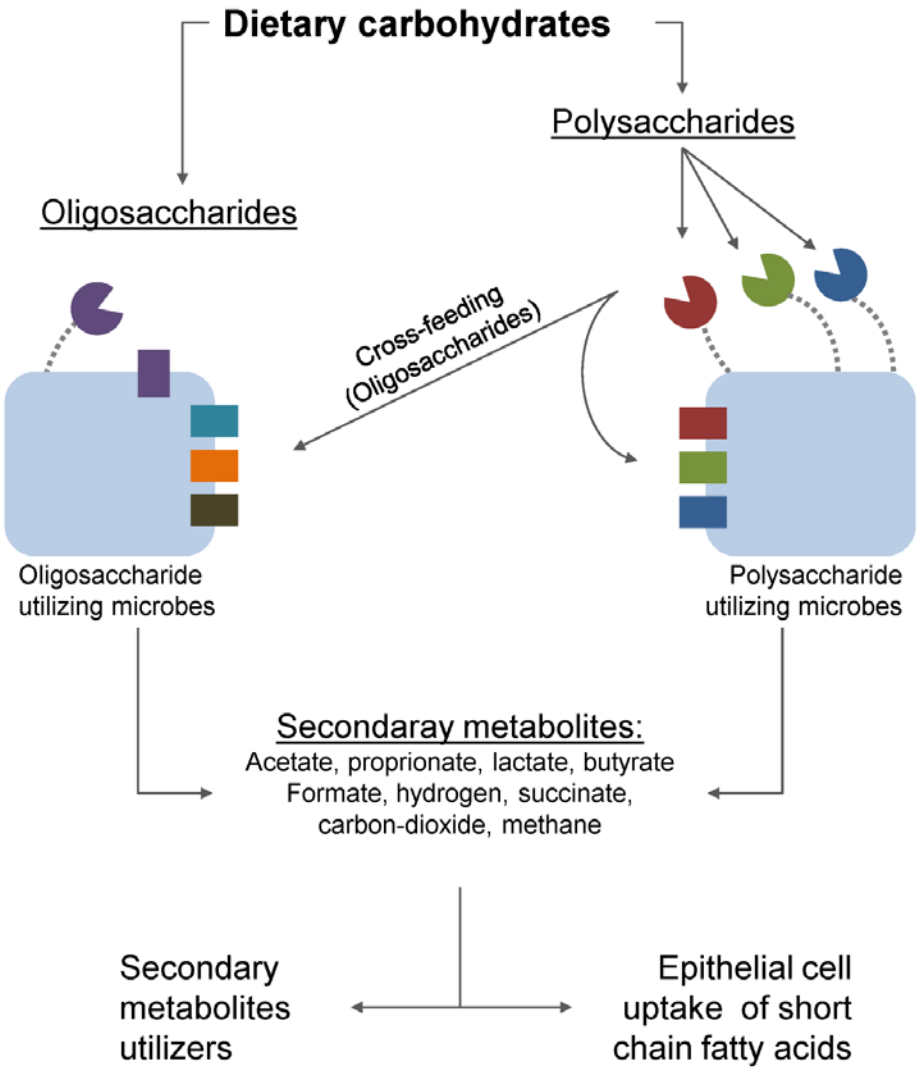


Figure 1-2: Schematic utilization of host non-digestible polysaccharides in the GIT. Extracellular glycoside hydrolases are represented as circular pie shapes and the dotted lines indicate cell attachment elements. Carbohydrate transporters are represented by square blocks protruding of the microbial cells shown in light blue. The colors of glycoside hydrolases and transporters schematically represent different substrate specificities. (Inspired from (193, 201, 202)).

1.4.2 Bacterial carbohydrate transport systems

There are three main classes of carbohydrate transporters identified within probiotic bacteria (132, 203): ATP-binding cassette transporters (ABC); phosphoenolpyruvate phosphotransferase (PTS) permeases; major facilitator superfamily (MFS) permeases, where glycoside-pentoside-hexuronide (GPH) permeases form a sub-group of MFS permeases. Classification and annotation of transporters have been aided by the sequence homology based transporter classification (TC) (204, 205). The following will introduce the above classes of transporters and their protein organization (Figure 1-3).

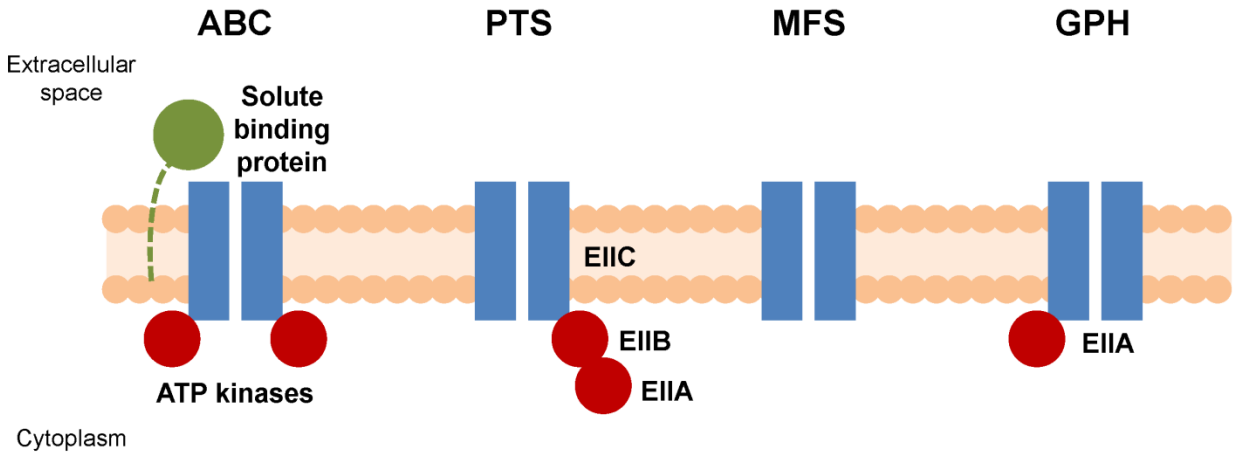


Figure 1-3: Schematic representation of carbohydrate transporters found in probiotics which may transporter di- and oligosaccharides. All permease domains are shown in blue. The substrate capturing solute binding protein of ABC transporters is shown in green where the dotted line represents the cell membrane anchoring domain found in Gram-positive bacteria and absent in Gram-negative bacteria, where the solute binding protein is secreted to the periplasm. ATP kinases, PTS domains EIIA and EIIB all involved with ATP hydrolysis and phosphate coupling are shown in red.

ABC transporters (TC 3.A.1)

ABC transporters are present in organisms from all domains of life and facilitate an ATP energized uptake (or export) of vitamins, carbohydrates, oligo-peptides and amino acids, ions and other organic compounds (206). The broad range of uptake is reflected by diversity in modularity of the domains constituting the transporter (207). Bacterial carbohydrate transporters are typically found as pentamers composed of an extracellular cell membrane attached solute binding protein for Gram-positive bacteria, whereas Gram-negative bacteria secrete the solute

binding protein into the periplasmic space, two membrane-spanning domains forming the permease and two nucleotide binding proteins coupling the hydrolysis of two ATP molecules to energize the transport (207). The molecular mechanism of carbohydrate ABC transporters has been pioneered based on structural work (208, 209) showing how the solute binding protein is the determinant of substrate specificity. Transport occurs via the solute binding protein, which upon substrate binding undergoes a conformational change allowing docking onto the permease sub-unit for release of the substrate into a transmembrane funnel like channel. Substrate translocation is finalized by the ATP coupled conformational change of the permease domains, closing the extracellular facing of the permease while opening the intracellular facing hence releasing the substrate to the cytoplasm.

Comparative analysis of solute binding proteins showed structural differentiation which could be linked to the substrate specificities allowing further differentiation of carbohydrate ABC transporters into monosaccharide and oligosaccharide specific transporters (210). The specificities of oligosaccharide transporters can be analyzed by functional studies involving inactivation of the full ABC transporter by a single solute binding protein gene knock-out (142). The molecular architecture of ABC transporters featuring the solute binding protein as a non-integral part of the transmembrane domain, has allowed recombinant production of solute binding proteins as a screening tool for ABC transporter specificities (211) and for biochemical characterization of carbohydrate affinities, usually in the sub-μM range (212–217). The family of oligosaccharide ABC transporters has been shown to facilitate prebiotic uptake (142, 198) yet annotation of novel ABC transporter is limited by low sequence similarity and lack of functional data.

PTS permeases (TC 4.A.1–4.A.6)

PTS permeases are found in prokaryotes (218) and facilitate an ATP energized uptake of mainly mono- and disaccharides, where the carbohydrate in the process is phosphorylated at the non-reducing O-6 position. The PTS permease is part of the three component PTS system termed: PTS EIIA, EIIB, EIIC and in some cases a EIID domain (219) where the single domains may be encoded in single genes or as a multi domain single protein. The EIIC domain is the

transmembrane domain displaying initial binding of the transported carbohydrate whereas the remaining EIIA and EIIB domains are involved with a cascade reaction for transferring the inorganic phosphate from phosphoenolpyruvate to the carbohydrate. The structure of a cellobiose specific PTS EIIC permease has recently been solved, aiding in understanding substrate binding of PTS permeases (220). *In silico* prediction of PTS permeases is hampered by lack of functional data (158) beyond annotations of novel PTS permeases into six classes of substrate specificities (Glucose/glucoside, fructose-mannitol, lactose/β-glucoside, glucitol, galactitol and mannose-sorbose) (205, 219). These six classes, however, does not represent for the amount of putative PTS permeases identified in some GIT associated strains, hence limiting annotations of novel PTS permeases (131, 137, 158).

MFS (TC 2.A.1) and GPH permeases (TC 2.A.2)

MFS and GPH permeases are secondary active transporters with broad substrate specificity including simple mono- and disaccharides such as melibiose (221), sucrose (222), lactose and galactose (143, 223, 224). The mechanism of MFS substrate binding and transport has been reviewed (225) and structural work has focused on the *Escherichia coli* lactose permease (226–228) indicating a tight binding pocket restricted to transport voluminous substrates beyond disaccharides.

In some Gram-positive bacteria, lactose permeases fused with a C-terminal PTS EIIA domain have been identified, linking regulation of the transporter activity to cellular energy levels (143, 224, 229, 230). Notably, some GIT lactobacilli encode a lactose specific PTS permease (231) whereas other employ a lactose specific GPH permease (143) indicating diversely evolved lactose utilization systems

The identification of carbohydrate transporters has been a key factor for understanding the selective utilization of prebiotic (142, 211). Yet, the push for novel prebiotics and the advances in microbiome genomics, as stated above, continues to necessitate further characterization of novel oligosaccharide transporters.

1.4.3 Classification and distribution of carbohydrate active enzymes

Catabolism of dietary carbohydrates depends on a vast array of microbial enzymes in the GIT and is a key activity of the microbiome (232). These breakdown reactions require enzymes that catalyze the hydrolysis of the glycosidic linkages to release mono- or oligosaccharides for uptake. The known collection of these carbohydrate active enzymes has been categorized in term of sequence similarity into glycoside hydrolase families (GH) via the carbohydrate active enzyme database (CAZy) (233). This classification system allows functional deduction of putative enzyme specificities by amino acid similarity.

Experimental assessment of the glycoside hydrolase repertoire encoded by the microbiome showed a highly dynamic and wide distribution of 73 glycoside hydrolase families (234, 235). In terms of selective prebiotics catabolism, the identification of glycoside hydrolase families predominantly found in probiotics is crucial to assess the catabolic capabilities. Figure 1-4 shows a comparative overview of the glycoside hydrolase families and their abundance found in probiotic bacteria compared to known pathogens (236–239). Notably enrichment of the following glycoside hydrolase families, displaying specificities for prebiotics, was observed: GH2 and GH42 encoding putative β-galactosidases involved with GOS catabolism (54, 240, 241), GH13 encoding putative α-1,6-glucoside specific enzymes for IMO catabolism (Møller et al, J. Bacteriol. 2012, in press), GH32 encoding β-fructosidases for FOS catabolism (142, 242) and GH43 together with GH51 for XOS and arabinoxylan catabolism (202, 243, 244). Further understanding of the prebiotic catabolic potential of glycoside hydrolase families have been provided by biochemical characterizations (140, 202) and *in silico* sub-family separation (245–249) to further improve annotation of genes potentially involved with prebiotic utilization.

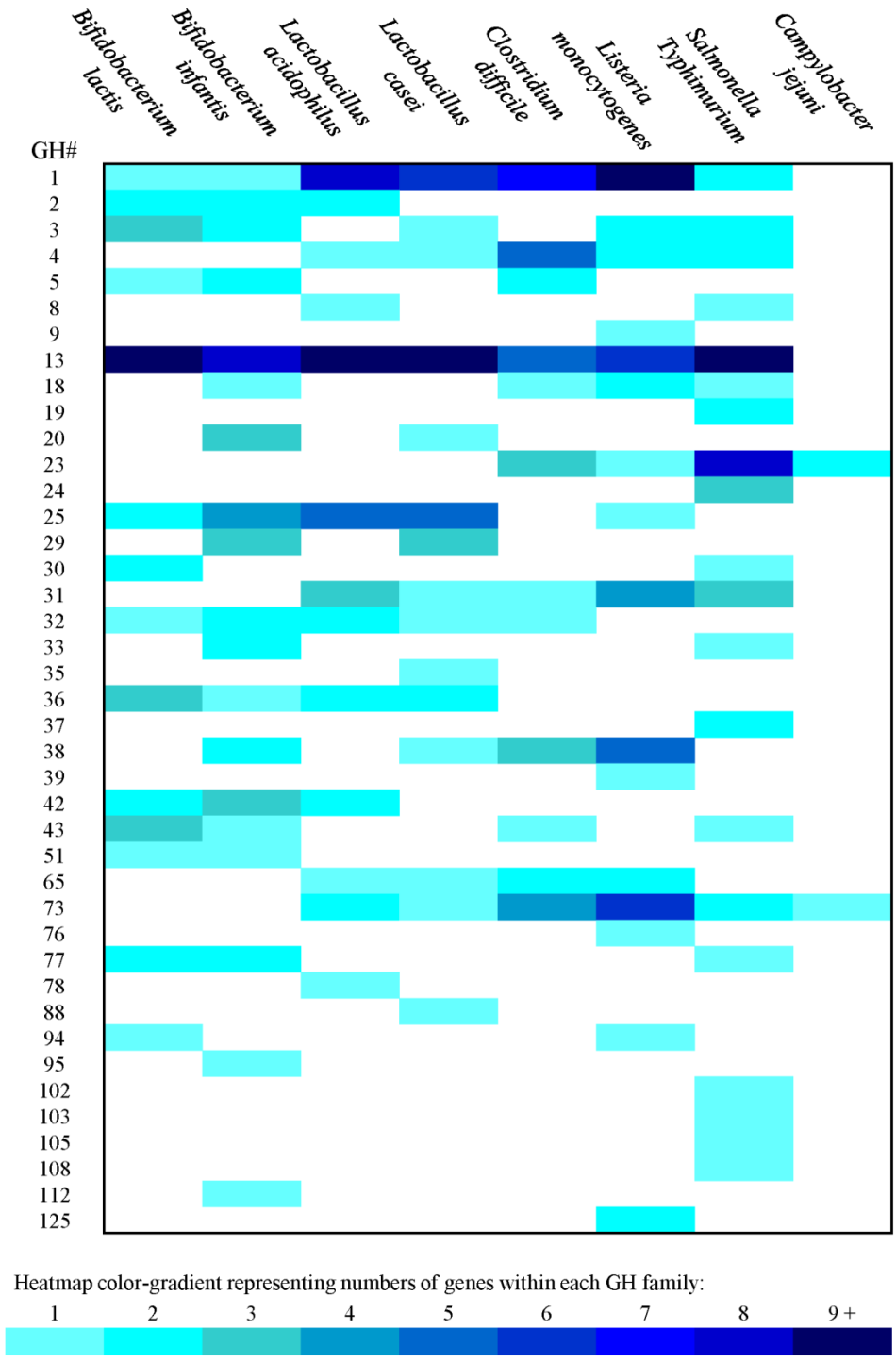


Figure 1-4: Heatmap distribution of glycoside hydrolase families identified through the CAZy database from selected probiotic and pathogenic bacteria. Full strain names: *Bifidobacterium animalis* subsp. *lactis* Bl-04; *Bifidobacterium longum* subsp. *infantis* ATCC 15697; *Lactobacillus acidophilus* NCFM; *Lactobacillus casei* BL23; *Clostridium difficile* CD196; *Listeria monocytogenes* 10403S; *Salmonella enterica* subsp. *enterica serovar Typhimurium* str. UK-1; *Campylobacter jejuni* RM1221

1.5 Experimental methods for gene and protein identification of probiotic properties

Despite the evolutionary phylogenetic distant clustering of probiotic lactobacilli and bifidobacteria (Section 1.3), specific genomic loci linked to utilization of prebiotics was shown to vary on a strain dependent level (250) by adaptive mechanisms of genome reduction (251) and by horizontal gene transfer (234). Hence the initial functional characterization of probiotic strains is to assess their potential for prebiotic utilization by screening of the supported growth by potential prebiotics in mono-culture fermentations (252–255). Assessment of the genetic basis underlying strain phenotypic behavior has become crucial to deconvolute mechanism of probiotic actions, interpret comparative genomics and identify target proteins for molecular understanding of probiotics, as presented in the previous sections.

Differential transcriptomics and proteomics methodologies have enabled high-throughput data generation for global analysis of the genes and protein respectively being upregulated to a defined growth condition or stimulation compared to an untreated control (256). Transcriptional analysis have proved to be suitable for identification of prebiotic induced gene expression as both carbohydrate transporters and hydrolases can be identified (143) whereas a gel-based proteomics approach in general does not enable identification of transmembrane proteins (186). Table 1-7 lists gene and protein identification within probiotics for increased understanding of molecular mechanisms of prebiotic utilization.

However, to put the gene and protein findings into a biological relevance, validation is required either by complementary methods e.g. quantitative-PCR (259) or by more biologically relevant methods such as functional genomics, where phenotypic characterization of targeted gene knock-outs can corroborate the proposed function (153, 260, 261).

Table 1-7: Selected studies employing high-throughput transcriptomics or proteomics for gene and protein identification within probiotic bacteria grown on prebiotic or potential prebiotic carbohydrates.

Method	Strain	Carbohydrates	Reference
Proteomics	*B. animalis* subsp. *lactis* Bl-04	GOS	(Ejby and Majumder, manuscript in preparation)
Proteomics	*L. acidophilus* NCFM	Lactitol	(157)
Proteomics	*L. acidophilus* NCFM	Cellobiose	(Van Zanten and Majumder, manuscript in preparation)
Proteomics	*L. acidophilus* NCFM	Raffinose	(Ejby and Majumder, manuscript in preparation
Proteomics Transcriptomics	*B. animalis* subsp. *lactis* BB-12	XOS	(182, 186)
Transcriptomics	*L. plantarum* WCFS1	FOS	(257)
Transcriptomics	*L. acidophilus* NCFM	Glucose, fructose, galactose, trehalose, lactose, sucrose, raffinose, FOS	(143)
Transcriptomics	*B. longum* NCC2705	Maltose, lactose, raffinose, FOS	(198)
Transcriptomics	*B. longum* LMG 13197	Glucose, GOS, human milk oligosaccharides	(258)

1.6 Scientific basis and objectives for the current project

Only few carbohydrates are classified as prebiotics and several candidate prebiotics are lacking sufficient documentation to gain status as prebiotics, hence it is desirable to expand the knowledge of the interactions of pre- and probiotics. It is the purpose of this thesis to functionally characterize prebiotic utilization by two clinically well-documented and commercially widely used probiotic strains, *Lactobacillus acidophilus* NCFM and *Bifidobacterium animalis* subsp. *lactis* Bl-04 using transcriptional analysis, functional genomics and biochemical characterization of recombinant proteins. A initial step for evaluating the potential prebiotic utilization was *in silico* genome mining of both strains (presented in Section 3.1) showing the types of prebiotics putatively utilizable by *L. acidophilus* NCFM and *B. lactis*

Bl-04. This led to selection of two diverse sets of carbohydrates, covering linkage families and glycoside compositions postulated to be utilized by carbohydrate transporters and glycoside hydrolases with potential for selective utilization of prebiotics. Subsequently, the objective of the thesis was to investigate the differential transcriptomics of the *L. acidophilus* NCFM and *B. lactis* Bl-04 grown on the above carbohydrates. The findings from both bacteria have been presented as separate manuscripts (Appendix 6.1 and 6.2) followed by comparison and discussion of the results within the thesis (Section 3.2) to highlight pathway differences in prebiotic utilization of the two probiotics. The aim of following work was to characterize single key genes identified from the transcriptional work for gaining insight into molecular mechanism of prebiotic uptake. Targeted gene deletion within *L. acidophilus* NCFM confirmed GOS uptake (Appendix 6.3) and RFO utilization (Appendix 6.1). A putative dual specificity IMO/RFO ABC transporter was identified in *B. lactis* Bl-04 and the substrate specificity was characterized (Appendix 6.4). The differences with regard to substrate recognition by transporters are discussed in Section 3.3 and the pending protein structure determination of the IMO/RFO solute binding protein is presented. In summary, this work provides a robust functional basis for selection and design of novel prebiotics and for analysis of selective prebiotics utilization. This data integrates well into the context of advancing the understanding of biomarkers for selective metabolism by probiotics and metagenomics within the GIT.

2 Materials and methods

The material and methods used for the experimental work have been described in the corresponding manuscripts (Appendices 6.1–6.4). Some of the obtained results have not been prepared into manuscripts and are only presented in the thesis hence the subsequent material and method section will expand and supplement the procedures used.

2.1 Databases, prediction and modeling tools

Table 2-1 lists the various bioinformatics online databases, servers and tools used in the present project.

Table 2-1: Collection of databases and bioinformatics tools used throughout the study.

Service	Description	URL (12th of May 2012)	Reference
SignalP	Signal peptide prediction	www.cbs.dtu.dk/services/SignalP/	(262)
BLAST	Sequence homolog detection	blast.ncbi.nlm.nih.gov/	(263)
Genbank	Publicly available nucleotide sequences	www.ncbi.nlm.nih.gov/genbank/	(264)
TCDB	Classification system for membrane transport proteins	www.tcdb.org/	(204)
CAZy	Database of carbohydrate active proteins	www.cazy.org/	(233)
ClustalX	Multiple sequence alignment	www.clustal.org/	(265)
Dendroscope	Visualizing of phylogenetic trees	ab.inf.uni-tuebingen.de/software/dendroscope/	(266)

2.2 Construction of phylogenetic tree of carbohydrate solute binding proteins

The phylogenetics of oligosaccharide solute binding proteins were analyzed (Section 3.1.7) by assembling a collection of homologous sequences. The sequence dataset was compiled from

initially 25 carbohydrate solute binding proteins all identified from previous work by transcriptomics or protein binding studies showing involvement of each protein to a type of oligosaccharides or identified from the current project to be involved with oligosaccharide binding. Sequence homologs for each protein entry were identified by BLAST (263) and restricted to either 100 hits or an e-value of 10^{-10} against the non-redundant database (264) before compiling all hit-sequences in a database collection made publicly available at the National Center for Biotechnology Information (MD, USA):

http://www.ncbi.nlm.nih.gov/sites/myncbi/collections/public/10kLj68I56iVl63rf8w5buCAc

Short link: **http://tinyurl.com/ca5rlrx** (12th of May 2012)

All redundant sequences were removed to result in 1649 unique entries, which were exported from the collection into FASTA format. Additionally, the 25 starting curated sequences were added together with a monosaccharide (fructose) binding protein (all entries listed with gi-number and functional annotation in Table 3-6). All sequences were loaded into ClustalX for multiple sequence alignment (265). The multiple sequences alignment was done using the Blosum series substitution matrix and a gap opening penalty of 2, compared to the standard penalty of 10. The resulting phylogenetic tree file was visualized using Dendroscope (266) and the tree was rooted using the fructose binding protein.

2.3 Experimental design of DNA microarray setup.

The methodology and results of the transcriptional analysis of *L. acidophilus* NCFM and *B. lactis* Bl-04 is presented in Appendices 6.1 and 6.2, respectively. The aspect and the basis for the experimental design were however beyond the scope of the manuscripts and will be assessed in this section to leading to the discussion of powers and limitations of transcriptional analysis (Section 3.2).

The gene expression of *L. acidophilus* NCFM was measured from cultures harvested in the early exponential phase and grown on glucose compared to 11 potential prebiotic oligosaccharides. An experimental dye-swapped loop-design (Figure 2-1) was used to generate two technical replicates from each biological replicate (one replicate culture per carbohydrate) as it was estimated from a previous study (143), using a round robin hybridization design (267), that the

number of samples and replicated would yield sufficient measurements for adequate statistical power in the ANOVA model to allow single gene identification with statistical significance (268).

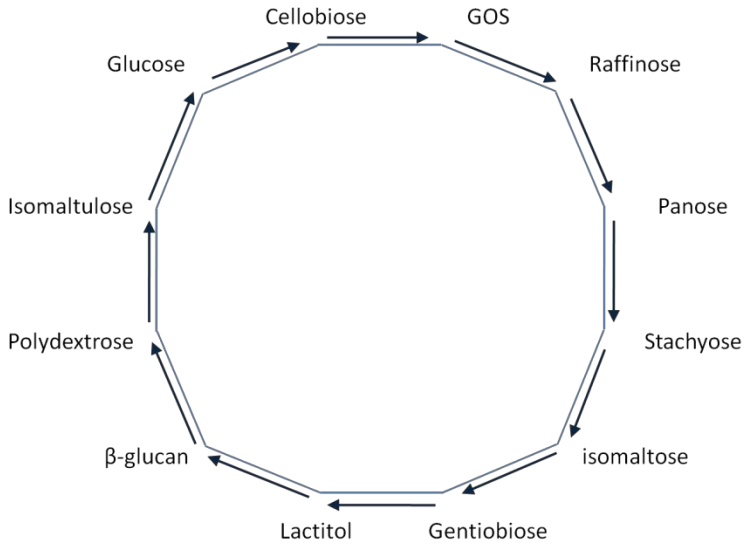


Figure 2-1: Loop-design for pairwise sample hybridization. Arrow heads indicate Cy5 labeled cDNA and the arrow tail indicated Cy3 labeled cDNA exemplified by the hybridization of Cy3 labeled glucose cDNA hybridized with Cy5 labeled cellobiose cDNA to the chip.

The gene expression of *B. lactis* Bl-04 was prepared similarly as for *L. acidophilus* NCFM regarding culture preparation, harvest and RNA isolation as presented in Appendix 6.2. For *B. lactis* Bl-04 however, a single dye RNA-labeling kit was used, resulting in one sample per hybridized chip with no need for dye-swap or loop design, although 24 chips were required compared to the 12 for *L. acidophilus* NCFM.

2.4 Generation of *L. acidophilus* NCFM gene deletion mutants

The construction of *Lactobacillus acidophilus* NCFM *Δupp* isogenic mutants with in-frame DNA excision of the genes LBA1438, LBA1442 and LBA1463 are presented in Appendices 6.1 and 6.3.

Two additional gene deletion mutants, LBA0502 a FOS solute binding protein of an ABC transporter and LBA1866 a putative maltose binding proteins of an ABC transporter were done according to Goh et al. (261). The constructed strains are listed in Table 2-2.

Table 2-2: Strains and plasmids used to construct gene deletion mutants of LBA0502 and LBA1866.

Strain or plasmid	Characteristics	Reference or source
***E. coli* strains**		
NCK1831	EC101: $RepA^+$ JM101; Km^r; *repA* from pWV01 integrated in chromosome; host for pORI-based plasmids	(269)
NCK1911	NCK1831 harboring pTRK935	(261)
NCK2120	NCK1831 harboring pTRK1012	Present work
NCK2128	NCK1831 harboring pTRK1016	Present work
***L. acidophilus* strains**		
NCFM	Human intestinal isolate	(131)
NCK1909	NCFM carrying a 315 bp in-frame deletion in the *upp* gene	(261)
NCK1910	NCK1909 harboring pTRK669, host for pORI-based counter selective integration vector	(261)
NCK2121	NCK1909 carrying a 1212 bp in-frame deletion in the LBA0502 gene	Present work
NCK2129	NCK1909 carrying a 1140 bp in-frame deletion in the LBA1866	Present work
Plasmids		
pTRK669	Ori (pWV01), Cm^r $RepA^+$	(270)
pTRK935	pORI28 derived with an inserted *upp* expression cassette and *lacZ´* from pUC19, serves as counterselective integration vector, Em^r	(261)
pTRK1012	pTRK935 with a mutated copy of LBA0502 cloned into BamHI/EcoRI sites	Present work
pTRK1016	pTRK935 with a mutated copy of LBA1866 cloned into BamHI/EcoRI sites	Present work

The upstream and downstream flanking regions (approximate length of 750 basepair each) of the deletion targets were PCR-amplified either with the 0502A/0502B and 0502C/0502D or 1866A/1866B and 1866C/1866D primer pairs, respectively, and fused by splicing by overlap extension PCR (SOE-PCR). The SOE-PCR products were ligated into pTRK935 linearized with compatible ends (BamHI and EcoRI for all constructs), and transformed into NCK1831. The resulting recombinant plasmids, pTRK1012 and pTRK1016, harbored in NCK2120 and NCK2128, were transformed into NCK1910 harboring pTRK669, for chromosomal integration

and following DNA excision to generate the ΔLBA0502 (NCK2121) or ΔLBA1866 (NCK2129) genotypes respectively. Confirmation of DNA deletion was done by PCR and DNA sequencing using primer pair 0502UP/0502DN and 1866UP/1866DN (Table 2-3).

Table 2-3: Primers used for construction of gene deletion mutants. Restriction sites are highlighted in bold and underlined

LBA0502 upstream flanking region	
0502A	CGC**GGATCC** ACTATGCTACGAAAAGATGGTT
0502B	TGCAACTCCTAATTTCCATT
LBA0502 downstream flanking region	
0502C	AATGGAAATTAGGAGTTGCAGTACAAAAGGTAATGAACGAACA
0502D	CCG**GAATTC**TTCAGCTGCTTCATACAATG
LBA0502 DNA excision control	
0502UP	TTCCAACATTCCTTTTGTTAGC
0502DN	TGGGTCATGATCATTGGTTG
LBA1866 upstream flanking region	
1866A	CGC**GGATCC**ATCAGACTGAAGCGATGACT
1866B	ACCTAAAGCCATTTTCTTCCA
LBA1442 downstream flanking region	
1866C	TGGAAGAAAATGGCTTTAGGTCCAAGTCAATACAAGGCACAA
1866D	CCG**GAATTC**GTTGGCAAGATGGTAAAGAA
LBA1866 DNA excision control	
1866UP	CAAAGACAGCGTGTTGCATT
1866DN	CAGCCCAATACTGGGAAGAA

2.5 Crystallization setup of recombinant *Bl*16GBP

Cloning and production of the protein encoded by Balac_1599 referred to as *Bl*16GBP is presented in Appendix 6.4. Crystallization of the recombinant *Bl*16GBP and data collection was done by Ph.D. student Morten Ejby at the Membrane Enzymology group, Department of Biochemistry, University of Groningen (the Netherlands) in collaboration with Associate Professor Dirk Slotboom and Post Doc. Andreja Vujicic-Zagar.

Recombinant *Bl*16GBP and selenomethionine labeled *Bl*16GBP was produced and purified as described in Appendix 6.4 with the only expectation that selenomethionine labeled protein was grown in selenomethionine containing media as previously described for other proteins (271). Protein stocks in 10 mM 2-(*N*-morpholino)ethanesulfonic acid, pH 6.5 and 150 mM NaCl were concentrated to 15 mg/ml. Crystals of *Bl*16GBP were grown by vapor diffusion in hanging drops. Crystals was only obtained when *Bl*16GBP was in its closed complex conformation with a

ligand (1 mM) either panose or raffinose. Crystal conditions that yielded crystals consisted of a drop set up of 1:1 ratio of protein (*Bl*16GBP 15 mg/ml) and reservoir (0.1 M Tris pH 8.5, 25% PEG 4000 and 0.8 M $MgCl_2$). Crystals grew after 60 h incubation at 5 °C. Due to the PEG in the reservoir solution no further cryoprotectant was applied and the crystals were flash frozen directly in liquid nitrogen. Data was collected to 1.6 Å for *Bl*16GBP in complex with raffinose, 1.9 Å for *Bl*16GBP in complex for with panose and 8 Å for Selenomethionine labeled *Bl*16GBP at the SLS beamline PX III, Villigen, Switzerland.

3 Results and discussion

3.1 Bioinformatics assessment and comparison of potential prebiotic utilization by *L. acidophilus* NCFM and *B. animalis* subsp. *lactis* Bl-04

One of the cornerstones in the definition of prebiotics is the selective metabolism by specific organisms in the GIT (2). Hence in assessment and development of novel prebiotics it is desirable to predict the target probiotic uptake and catabolic systems to link selective metabolism to specific genetic loci. The types of potential prebiotics utilized by probiotic organisms can be predicted based on homology to already known utilization pathways and enable comparison of both differences among probiotic organisms on the genome level, but also by comparison to commensal GIT organisms and pathogens, to reveal taxonomical niche-specific gene clusters for selective targeting with prebiotics.

3.1.1 Mapping of potential prebiotic utilization systems

Ideally by the above approach both predicted and experimentally validated gene clusters encoding prebiotic utilization systems can be mapped in relation to the recently scientifically established GIT metagenome (101, 108) to predict the selective metabolism of the potential prebiotic by the GIT microbiome. However, the reconstruction of putative gene clusters by *in silico* methods primarily based on single gene homology within automatically annotated genomes is hampered by the complexity of predicting the interplay of the full genome or parts thereof. Databases describe most intracellular metabolic pathways of simple metabolic compounds, such as the Kyoto Encyclopedia of Genes and Genomes (272), but the multitude of uptake and hydrolytic pathways for oligosaccharides remain poorly described (158). Hence it is the aim to make a gene landscape analysis of gene clusters encoding putative oligosaccharide transporters and glycoside hydrolases to allow re-construction of the pathways for utilization of potential prebiotic, based on *in silico* predictions of *L. acidophilus* NCFM and *B. lactis* Bl-04 as described in the following.

Initially, all carbohydrate transporters were identified and classified using the transporter classification (TC) database (205), while all glycoside hydrolases were identified and assigned a GH family number using the CAZy database (233). The cellular localization of glycoside

hydrolases were predicted using the signalP tool (262). Gene clusters were generated by gene landscape analysis of carbohydrate transporters and glycoside hydrolases by their neighboring encoded genomic location, and where applicable also with identification of transcriptional regulators. This combined *in silico* gene landscape analysis approach allowed to couple the annotation of both transporters and hydrolases to strengthen the overall prediction of the single clusters, as aided by BLAST homology searching (263) to the Uniprot database of characterized proteins (273).

3.1.2 Genome-mining and assessment of the utilization of potential prebiotics by *L. acidophilus* NCFM

The *in silico* annotation of the uptake and catabolic systems of *L. acidophilus* NCFM was done based on the available genome sequence (131). The constructed gene clusters are shown in Table 3-1 including a putative function in relation to potential prebiotic utilization. Remaining genes encoding glycoside hydrolases and carbohydrate transporters, which could not be assigned into a gene cluster, are listed in Table 3-2 and Table 3-3, respectively.

The majority of the identified glycoside hydrolases and carbohydrate transporters could be functionally connected into gene clusters and appeared to be organized with transcriptional regulators putatively associated with carbohydrate metabolism. This implies how each single gene cluster can be transcribed in response to sensing of available carbohydrates in the GIT. A total of three ABC, one GPH and eight PTS transporter containing gene clusters were identified (Table 3-1). Annotation of both the ABC and GPH systems was supported by earlier transcriptional work (141–143) for robustly assigning their functions. Evaluation of PTS systems in the present analysis could only be deemed reliable if supported by experimental work as in the case of trehalose and sucrose utilization. The remaining PTS permease encoding gene clusters showed specificities for β-glucosides but further experimental work is needed to determine if the gene clusters encode redundancies or specialized functions within β-glucosides utilization.

Table 3-1: *In silico* predicted gene clusters in *L. acidophilus* NCFM with putative involvement in carbohydrate utilization listed with any experimental evidence to support the predictions. All gene clusters are sub-grouped by the type of Transporter Classified numbering (TC) and all glycoside hydrolases are predicted to be localized intracellularly unless otherwise noted. Carbohydrate binding modules (CBM) are given in brackets after the GH family.

ABC transporter encoding gene clusters				
ORF #	**Gene annotation by Altermann et al. (131)**	**GH (CBM)**	**TC**	**Cluster function and level of predictive confidence**
0500	Msm associated regulator			FOS based on transcriptomics and functional genomics (142)
0502	Solute binding protein		3.A.1	
0503	Permease domain		3.A.1	
0504	Permease domain		3.A.1	
0505	β-fructosidase, EC 3.2.1.26	32		
0506	ATP-binding protein		3.A.1	
0507	Sucrose phosphorylase, EC 2.4.1.7	13_18		
1437	Sucrose phosphorylase, EC 2.4.1.7	13_18		Raffinose based on transcriptomics (143)
1438	α-galactosidase, EC 3.2.1.22	36		
1439	ATP-binding protein		3.A.1	
1440	Permease domain		3.A.1	
1441	Permease domain		3.A.1	
1442	Solute binding protein		3.A.1	
1443	Msm associated regulator			
1864	Permease domain		3.A.1	Malto-oligosaccharides based on protein characterization of LBA1870 (274) and some homology to gene clusters from other bacteria (200, 213)
1865	Permease domain		3.A.1	
1866	Solute binding protein		3.A.1	
1867	ATP-binding protein		3.A.1	
1868	Transposase			
1869	β-phosphoglucomutase, EC 5.4.2.6			
1870	Maltose phosphorylase, EC 2.4.1.8	65		
1871	Maltogenic α-amylase, EC 3.2.1.133	13_20 (34)		
1872	Oligo-α-(1,6)-glucosidase, EC 3.2.1.10	13_31		
1873	Acetate kinase			
1874	LacI type regulator			
GPH transporter encoding gene cluster				
1460	Truncated mucin binding protein			Lactose based on transcriptomics (143)
1461	Unknown type regulator			
1462	β-galactosidase, EC 3.2.1.23	42		
1463	GPH permease		2.A.2	
1464	Tranposase			
1465	LacI type regulator			
1467	β-galactosidase, large sub-unit, EC 3.2.1.23	2		
1468	β-galactosidase, small sub-unit, EC 3.2.1.23	2		

PTS transporter encoding gene cluster				
ORF #	**Gene annotation by Altermann et al.** (131)	**GH (CBM)**	**TC**	**Cluster function and level of predictive confidence**
0225	6-phospho-β-glucosidase, EC 3.2.1.21	1		β-glucosides, *in silico* prediction only
0226	Unknown type regulator			
0227	PTS EIIC domain		4.A	
0228	NagC type regulator			
0399	Sucrose operon regulator			Sucrose based on transcripto-mics (143)
0400	Sucrose-6-phosphate hydrolase, EC 3.2.1.B3	32		
0401	PTS EIIABC domain		4.A	
0724	LicT type regulator			Cellobiose based on functional genomics in the related *L. gasseri* (158)
0725	PTS EIIC domain		4.A	
0726	6-phospho-β-glucosidase, EC 3.2.1.21	1		
0874	6-phospho-β-glucosidase, EC 3.2.1.21	1		β-glucosides, *in silico* prediction only
0875	GntR type regulator			
0876	PTS EIIC domain		4.A	
0877	PTS EIIA domain		4.A	
0878	Hypothetical protein			
0879	PTS EIIC domain		4.A	
0880	Hypothetical protein			
0881	6-phospho-β-glucosidase, EC 3.2.1.21	1		
0882	Unknown type regulator			
0883	Hypothetical protein			
0884	PTS EIIC domain		4.A	
0885	6-phospho-β-glucosidase, EC 3.2.1.21	1		
0886	NagC type regulator			
1012	PTS EIIC domain		4.A	Trehalose based on functional genomics (141)
1013	Trehalose regulator			
1014	Trehalose-6-phosphate hydrolase, EC 3.2.1.93	13_29		
1364	β-galactosidase, EC 3.2.1.23	42		β-glucosides, *in silico* prediction only
1365	α-glucosidase, EC 3.2.1.20	31		
1366	6-phospho-β-glucosidase, EC 3.2.1.21	1		
1367	AgC type regulator			
1368	XylR type regulator			
1369	PTS EIIC domain		4.A	
1574	6-phospho-β-glucosidase, EC 3.2.1.21	1		β-glucosides, *in silico* prediction only
1576	PTS EIIAC domain		4.A	
1577	RpiR type regulator			
1705	PTS EIIBC domain		4.A	β-glucosides, *in silico* prediction only
1706	6-phospho-β-glucosidase, EC 3.2.1.21	1		
1707	PTS EIIABC domain		4.A	
1708	β-glucoside type regulator			

Table 3-2: Glycoside hydrolases not co-encoded with a carbohydrate transporter within *L. acidophilus* NCFM. All Glycoside hydrolases are predicted to be localized intracellularly unless highlighted with bold face. Carbohydrate binding modules are given in brackets after the GH family.

ORF #	**Gene annotation** (131)	GH (CBM)
0107	β-glucanase, **extracellular**, EC 3.2.1.4	8
0143	α-glucosidase, EC 3.2.1.20	31 (32)
0176	*N*-Acetylmuramidase, EC 3.5.1.28	73
0264	Glucan-α-(1,6)-glucosidase[1], EC 3.2.1.70	13_31
0527	*N*-Acetylmuramidase, EC 3.5.1.28	73
0680	α-(1,4)-glucan branching enzyme, EC 2.4.1.18	13_9 (48)
0686	Amylopullulanase, 3.2.1.41	13_20
1140	Muramidase fragment, 3.2.1.17	25
1336	6-phospho-β-glucosidase, EC 3.2.1.21	1
1351	Muramidase, EC 3.2.1.17	25
1352	Muramidase fragment EC 3.2.1.17	25
1473	α-L-rhamnosidase, EC 3.2.1.40	78
1689	Maltose-6-phosphate glucosidase	4
1710	Pullulanase[1], **extracellular**, EC 3.2.1.41	13_14 (41, 48)
1812	α-glucosidase, EC 3.2.1.20	31
1918	Muramidase, EC 3.2.1.17	25

[1]Annotation based on biochemical characterization (Abou Hachem, M. personal communication).

Table 3-3: Carbohydrate transporters identified in *L. acidophilus* NCFM and not co-encoded with one or more glycoside hydrolases. The transporters listed by locustag numbers are grouped by their Transporter classification numbering.

ORF#	**Gene annotation** (131)	TC
	MFS	
0045	Unspecified monosaccharide uptake	2.A.2
	PTS	
0146	Monosaccharide regulation, PTS EIIA	4.A
0452	Glucose uptake, PTS EIIAB	4.A
0456	Glucose uptake, PTS EIIC	4.A
0456	Glucose uptake, PTS EIID	4.A
0491	β-glucose uptake, PTS EIIC	4.A
0606	α-glucoside uptake, PTS EIIBC	4.A
0609	α-glucoside uptake, PTS EIIA	4.A
0618	β-glucose uptake, PTS EIIC	4.A
0989	Monosaccharide uptake, PTS EIIC	4.A
1478	Monosaccharide uptake, PTS EIIBC	4.A
1484	Monosaccharide regulation, PTS EIIA	4.A
1777	Fructose uptake, PTS EIIABC	4.A
	The Drug/Metabolite Transporter Superfamily	
1102	Ribose uptake	2.A.7.5
1376	Ribose uptake	2.A.7.5

Notably, 8 of 16 glycoside hydrolases unassigned a gene cluster were predicted to be involved with other function then potential prebiotic hydrolysis such as intracellular glycogen metabolism (LBA0680 and 0686), lysozymes (LBA1351, 1352) and catabolism of modified glycoside or glycopeptides (LBA0176, 0527 1140 and 1918)).

Nine of the 12 carbohydrate transporters, which could not be assigned to a gene cluster were predicted to be involved in monosaccharide uptake and hence would not be involved in uptake of potential prebiotic substrates. The remaining transporters (LBA0491, 0606–0609 and 0618) were all PTS permeases which could possibly be functionally linked to the above glycoside hydrolases (Table 3-2) based on their subtle putative specificities for α- and β-glucosides. The lack of structured gene clusters could be due to a recent gene uptake representing a adaption-mechanism within the GIT (235, 251).

The observation of most glycoside hydrolases and oligosaccharide transporters being found in gene clusters supports the mechanism of *L. acidophilus* NCFM being highly adaptive to exogenous nutritional stimulation on the transcriptional level (143).

3.1.3 Function deduction and interplay of genes identified within *L. acidophilus* NCFM

The *in silico* analysis of the cellular localization of encoded glycoside hydrolases revealed only two cases with signal peptide being present; a putative β-glucanase (LBA0107) and a pullulanase (LBA1710). The sequence analysis identified an encoded membrane attachment domain (bacterial surface layer protein; pfam03217), suggesting that these enzymes act on the outer surface of the bacterial cell. None of these two identified genes could be assigned into a gene cluster for functional association with transporters or intracellular enzymes. However, a putative α-glucan utilization gene cluster was identified (LBA1864–1872) together with a putative α-glucan PTS system (LBA0606, 0609), which indicated a functional link with the extracellular pullulanase (LBA1710), which could release oligomeric α-glucan fragments from partially degraded starch for uptake and intracellular catabolism. Sequence analysis of the putative β-glucanase showed a mutated catalytic nucleophile acid residue (aspartic $acid_{257}$ to asparagine) recognized as motif within a sub-family of GH8 (275). No enzymatic activity has so far been reported for this sub-family of GH8, but analysis of the genomic position of LBA0107 identified

a glycosyl transferase family 2 enzyme (LBA0106) to be encoded adjacent to LBA0107 indicating a potential role in cell wall modifications or extracellular exo-polysaccharide modifications.

In the view of the predicted specialized oligosaccharide transport systems and few extracellular glycoside hydrolases encoded by *L. acidophilus* NCFM it is evident that the organism has adopted a scavenging mechanism for oligosaccharides in the GIT (201). In the light of the predicted potential prebiotic utilization profile, *L. acidophilus* NCFM can clearly be targeted with a number of oligosaccharide prebiotic candidates. Further knowledge and analysis is required to assess, which compounds are the best substrates in the complex niche that *L. acidophilus* NCFM inhabits. Therefore, further work, of systems biology nature applying differential transcriptional and proteomics, is needed to reveal the efficiencies and specificities of the oligosaccharide transport systems. Candidate prebiotics to screen, based on the above analysis, would be of the following types: FOS, GOS, IMO, RFO, breakdown products of resistant starches, and fragmented β-glucans to support selective growth of *L. acidophilus* NCFM.

3.1.4 Genome-mining and assessment of the utilization of potential prebiotics by *B. animalis* subsp. *lactis* Bl-04

The *in silico* annotation of the uptake and catabolic systems of *B. lactis* Bl-04 was based on the recently published genome (167) to identify carbohydrate transporters and glycoside hydrolases for further annotation. All reconstructed gene clusters encoding both carbohydrate transporters and glycoside hydrolases are listed in Table 3-4. Only one putative carbohydrate transporter (Balac_1154) was identified, which could not be associated a gene cluster harboring a glycoside hydrolase encoding gene and thus being potentially implicated in oligosaccharide catabolism. Sequence analysis indicated the putative transporter to be of the MFS type specific for monosaccharides and hence having little relevance for uptake of potential prebiotics. Table 3-5 lists the glycoside hydrolases not associated with a carbohydrate transporter.

Table 3-4: *In silico* predicted gene clusters in *B. lactis* Bl-04 with putative involvement in carbohydrate utilization listed with any experimental evidence to support the predictions. All gene clusters are sub-grouped by the Transporter Classification and all glycoside hydrolases are predicted to be intracellularly. CBMs are given in brackets after the GH family.

ABC transporter encoding gene clusters				
ORF #	**Gene annotation by Barrangou et al. (167)**	**GH (CBM)**	**TC**	**Cluster function and level of predictive confidence**
0483	Solute binding protein		3.A.1	Lactose and GOS based on transcriptomics analysis in the related *B. longum* (198)
0484	β-galactosidase, EC 3.2.1.23	42		
0485	Permease domain		3.A.1	
0486	Permease domain		3.A.1	
0487	LacI type regulator			
0511	xylose isomerase			XOS based on proteomics and transcriptomics analysis *B. animalis* subsp. *lactis* BB-12 (186)
0512	β-xylosidase, EC 3.2.1.37	43		
0513	LacI type regulator			
0514	Solute binding protein		3.A.1	
0515	Permease domain		3.A.1	
0516	Permease domain		3.A.1	
0517	β-xylosidase, EC 3.2.1.37	43		
0518	Hypothetical protein			
0519	Carbohydrate esterase			
0520	β-xylosidase, EC 3.2.1.37	43		
0521	Xylulose kinase			
0522	NagC type regulator			
1562	Pullulanase, EC 3.2.1.41	13_14 (48)		Malto-oligosaccharides based on homology to gene clusters from other bacteria (200, 213)
1563	Permease domain		3.A.1	
1564	Permease domain		3.A.1	
1565	Solute binding protein		3.A.1	
1566	α-(1,4)-glucosidase, EC 3.2.1.20	13_30		
1567	α-(1,4)-glucanotransferase, EC 2.4.1.25	77		
1568	Hypothetical protein			
1569	Permease domain		3.A.1	
1570	Permease domain		3.A.1	
1571	LacI type regulater			
1572	Solute binding protein		3.A.1	
1573	α-glucosidase, EC 3.2.1.20	13_30		
1593	Oligo-α-(1,6)-glucosidase, EC 3.2.1.10	13_31		Raffinose and isomaltose based on transcriptional analysis in *Streptococcus mutans* (276)
1594	Short open reading frame			
1595	Short open reading frame			
1596	α-galactosidase, EC 3.2.1.22	36		
1597	Permease domain		3.A.1	
1598	Permease domain		3.A.1	
1599	Solute binding protein		3.A.1	
1600	NagC type regulator			
1601	α-galactosidase, EC 3.2.1.22	36		

GPH transporter encoding gene clusters				
ORF #	**Gene annotation by Barrangou et al (167)**	**GH (CBM)**	**TC**	**Cluster function and level of predictive confidence**
0475 0476 0477	GPH permease β-galactosidase, EC 3.2.1.23 LacI type regulator	2	2.A.2	Lactose based on transcriptional analysis in *B. longum* (198)
1588 1589 1590	GPH permease β-L-arabinofuranosidase, EC 3.2.1.X[1] LacI type regulator	127	2.A.2	arabinoxylan fragments based on *in silico* predictions
MFS transporter encoding gene clusters				
0052 0053 0054 0055	β-(1,6)-glucanase, EC 3.2.1.75 β-galactosidase, EC 3.2.1.23 MFS permease TetR type regulator	30 42	2.A.1	β-glycosides based on *in silico* predictions
0137 0138 0139	LacI type regulator Sucrose phosphorylase, EC 2.4.1.7 MFS permease	13_18	2.A.1	Sucrose based on *in silico* predictions
1239 1240 1241	LacI type regulator MFS permease Sucrose hydrolase, EC 3.2.1.26	32	2.A.1	Sucrose and FOS based on a Balac_1241 homolog from *Bifidobacterium lactis* (277)

[1] Enzymatic activity is not yet classified.

A total of four ABC, and five MFS (divided into GPH and MFS types) transporter encoding gene clusters were assigned for *B. lactis* Bl-04, with a general tendency for ABC containing clusters to encode multiple glycoside hydrolases suggesting transport of oligosaccharides with such a complexity that additional glycoside hydrolases are required for hydrolysis or that the ABC transporters display multiple specificities and are able to facilitate uptake of a range of oligosaccharides, as discussed later in Section 3.2.3. Biochemical information on MFS permeases, and the GPH sub-group of MFS transporters, is limited for substrate specificity of uptake. Therefore the in-depth analysis of these gene clusters depends mainly on the associated glycoside hydrolases, which in the above table mainly indicate disaccharide hydrolysis and hence disaccharide uptake by the MFS types of transporters.

Table 3-5: Glycoside hydrolases not co-encoded with a carbohydrate transporter within *B. lactis* Bl-04. All Glycoside hydrolases are predicted to be localized intracellularly unless highlighted in bold. Carbohydrate binding modules are given in brackets after the GH family numbering.

ORF #	Gene annotation (167)	GH(CBM)
0049	β-glucosidase, EC 3.2.1.21	3
0065	α-L-arabinofuranosidase, EC 3.2.1.55	51
0151	β-glucosidase, EC 3.2.1.21	1
0268	β-galactosidase, EC 3.2.1.23	2
0373	α-(1,4)-glucanotransferase, EC 2.4.1.25	77
0376	Isoamylase, EC 3.2.1.68	13_11(48)
0924	Truncated pullulanase, EC 3.2.1.41	13_?
0952	α-amylase, EC 3.2.1.1	13_5
0977	Isoamylase, EC 3.2.1.68	13_11(48)
0995	α-(1,4)-glucan branching enzymes, EC 2.4.1.8	13_9(48)
1025	β-*N*-acetyl-hexosaminidase, EC 3.2.1.52	3
1418	Endo-β-(1,6)-glucanase, EC 3.2.1.164	30
1421	Cellobiose phosphorylase, EC 2.4.1.20	94
1450	endo-β-mannosidase, **Extracellular**, EC 3.2.1.25	5(10)
1458	No known activity within sub-family	13_3
1516	Muramidase, EC 3.2.1.17	25
1517	Muramidase, EC 3.2.1.17	25
1537	α-galactosidase, EC 3.2.1.22	36
1551	β-*N*-acetyl-hexosaminidase, EC 3.2.1.52	3

B. lactis Bl-04 encoded 19 glycoside hydrolases not predicted to a gene cluster, four of which were predicted to be specific for glycopeptides (Balac_1025, 1516, 1517 and 1551) albeit not constituting the full pathways for host glycopeptides utilization as identified in *B. bifidum* (170). The 15 remaining enzymes could not be assigned to a gene cluster and any functionality for prebiotic catabolism cannot be hypothesized beyond their putative EC numbering and CAZy classification.

3.1.5 Functional deduction from gene identification within *B. animalis* subsp. *lactis* Bl-04

First step in assessing the utilization of potential prebiotics was to identify any putatively extracellular enzymes within *B. lactis* Bl-04. Remarkably, only a single enzyme was predicted to be secreted, namely the putative endo-β-mannosidase (Balac_1450), which by the gene annotation is expected to release short β-manno-oligosaccharides by hydrolysis of polymeric substrates such as β-mannans. A β-manno-oligosaccharide transporter could not be identified *in silico* to support the uptake of released β-manno-oligosaccharide degradation products and with

respect to the intracellular glycoside hydrolases not found in gene clusters with oligosaccharide transporters (Table 3-5), only one entry were found in a glycoside hydrolase family (Balac_0268, GH2) harboring putative β-mannosidase activity. Thus, the utilization of β-manno-oligosaccharides is either lacking beyond the extracellular endo-β-mannosidase or is facilitated by a to-date unpredictable transporter and intracellular exo-β-mannosidase. As for *L. acidophilus* NCFM, a scavenging profile of prebiotic utilization in the GIT is suggested for *B. lactis* Bl-04, indicating how the encoded transporters are the initial substrate interacting components being an important substrate determining factor for the potential prebiotic utilization.

All identified oligosaccharide transporters could be allocated into gene clusters, which could be grouped based on the associated transporter to be either ABC or GPH/MFS type permease with a tendency of ABC containing gene clusters to encode additional glycoside hydrolases (the locus Balac_0483–0487 excluded) compared to generally one glycoside hydrolase in MFS permease specific gene clusters (the locus Balac_0052–0055 excluded). This observation suggests how ABC transporters within *B. lactis* Bl-04 may transport complex oligosaccharides requiring a multitude of intracellular glycoside hydrolases whereas the MFS type of permeases may facilitate disaccharide uptake.

From the present genomic analysis it is suggested how *B. lactis* Bl-04 holds potential for utilizing oligosaccharide prebiotics such as: β-manno-oligosaccharides, GOS, IMO, RFO and XOS.

3.1.6 Comparative genomics of potential prebiotic utilization of *L. acidophilus* NCFM and *B. animalis* subsp. *lactis* Bl-04

Genomics analysis of the carbohydrate utilization systems of *L. acidophilus* NCFM and *B. lactis* Bl-04 was used to map and evaluate the carbohydrate uptake and catabolic potential of each organism to enable strain comparison.

The established mechanism of action for dietary carbohydrate utilization by the GIT microbiome depends on the genus analyzed (202) where the general consensus for utilization of a given complex carbohydrate is initial extracellular hydrolysis within the GIT releasing shorter oligomeric, or monosaccharide, substrates for uptake and followed by intracellular hydrolysis to

monosaccharides (38, 194, 242). Notably, *in silico* secretome analysis of *L. acidophilus* NCFM and *B. lactis* Bl-04 identified two and one putative secreted glycoside hydrolases, respectively. This particularly low number of membrane attached hydrolytic enzymes suggests an overall alternative scavenging mechanism of prebiotic utilization for both strains, where the transport systems are the initial substrate recognizing component of the bacteria (201, 278). This has implications for the evaluation of the strains. First, from a scientific point of the view, the characterization of carbohydrate transporters becomes essential to understand the selective prebiotic utilization and secondly it can guide the design and selection of oligosaccharide prebiotics to stimulate selective growth of *L. acidophilus* NCFM and *B. lactis* Bl-04 and the related probiotic strains they both represent.

The relative high number of carbohydrate transporters encoded in both strains supports the paradigm of bacterial interplay in the GIT where primary polysaccharide-degrading microorganisms cross-feed oligosaccharides to secondary users for utilization (191). Here the transporters would facilitate wide substrate specificity for readily uptake of available oligosaccharides in a densely populated, competitive environment. Thus understanding of the prebiotic/probiotic interactions lies to a great extent in understanding the carbohydrate transport systems.

Analysis of the encoded putative transporters revealed 17 putative oligosaccharide transporters in *L. acidophilus* NCFM as compared to nine in *B. lactis* Bl-04. This difference in distribution of transporters and the gene cluster associated glycoside hydrolases further highlights the overall routes of carbohydrate utilization. *L. acidophilus* NCFM seemingly processes a 'one transporter, one substrate' mechanism of action as exemplified by the PTS containing gene clusters mainly associated with disaccharide uptake (219) and GH1 enzymes with exo-acting activity (279), linking these pathways to uptake of disaccharides released from polysaccharide breakdown. In comparison *B. lactis* Bl-04 encodes five ABC transporter gene clusters with GH profiles suggesting broad substrate specificity as illustrated by the XOS and IMO/RFO specific gene clusters as shown in Appendix 6.2.

Another striking difference in *L. acidophilus* NCFM compared to *B. lactis* Bl-04 is the gene organization of ABC encoding cluster, where *L. acidophilus* NCFM encodes a transporter specific ATP-binding protein within each of the three identified gene cluster. A similar

organization is lacking for *B. lactis* Bl-04, where only a single oligosaccharide ABC transporter associated ATP binding proteins was identified in the genome (Balac_1610), as also previously reported for *B. longum* NCC2705 (198). The ability of an ATP binding protein with specificity towards multiple ABC transporters was shown previously (280). This suggests divergent regulatory mechanisms of the two organisms' responses to nutritional changes.

3.1.7 Functional map of carbohydrate-specific solute binding proteins of ABC transporters

ABC transporter associated uptake of carbohydrates is widespread in nature with the solute binding protein being the main substrate specific component (210). The extracellular nature of this class of proteins compared to transmembrane carbohydrate transporter types such as MFS or PTS permeases, allows sequence identification of the specific substrate binding domain for functional assignment and biochemical characterization. Furthermore, the previous *in silico* observation of oligosaccharide transport by ABC transporters made the solute binding proteins interesting to study within the scope of prebiotic transport. *In silico* approaches have so far shown how monosaccharide solute binding proteins can be distinguished from oligosaccharide solute binding proteins by structural fold and sequence length and how oligosaccharide solute binding proteins all adopt the same overall structural fold, yet still the overall protein sequence similarity is modest (25–35%) making functional predictions difficult (210).

The functional phylogenetic relationship of prebiotic and related oligosaccharides solute binding proteins from ABC transporters was prepared as described in Section 2.1. This revealed evolutionary grouping of functionalities driven by their niche habitat and taxonomical drift as illustrated in Figure 3-1. The graphical analysis first show how all identified clusters are mono-specific as only one functionally determined specificity was found (except cluster 5C) and how related functions sub-cluster as for GOS (5A) and lacto-*N*-biose (5B), a lactose based potential prebiotic disaccharide isolated from human milk (215). Also, differentiation is observed for maltose-like binding proteins indicated (cluster 4), where specificities are reported for cyclodextrins (281), malto-oligosaccharides (213) and various disaccharides (212) or cellodextrins (250, 282), all together representing a diverse landscape largely driven by protein specificity rather than taxonomy which is however reflected within the single sub-groups.

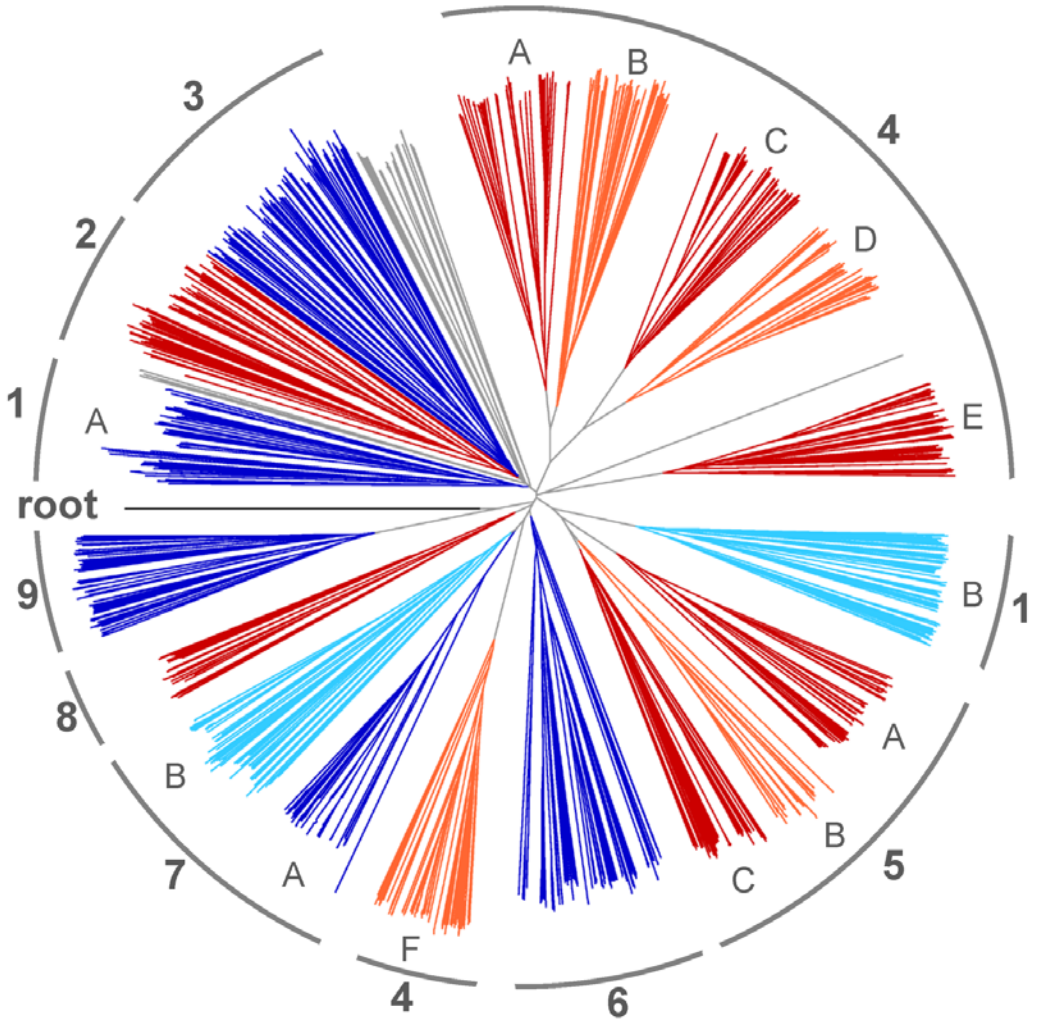


Figure 3-1: Phylogenetic tree of oligosaccharide binding proteins showing the distribution of characterized oligosaccharide binding proteins. The tree has been rooted with a fructose specific solute binding protein as an out-group and manually divided into clusters based on protein functionality (shown by numbers) and sub-clusters (shown by letters and color codes) listed clock-wise from the root. Table 3-6 lists details to each cluster and sub-cluster. Protein entries, which could not be assigned a cluster are shown in grey.

Table 3-6: Identified clusters of oligosaccharide binding proteins from Figure 3-1. Clusters are shown by numbers and if possible sub-clusters are listed with letters and color coding. The experimentally identified oligosaccharide binding proteins used to generate the tree are listed in the corresponding cluster.

Cluster	Sub-cluster	Substrate specificity	Identified Organism	Reference
1	**A**	β-(1,4)-gluco-oligosaccharides	*Clostridium thermocellum* ACTT 27405	(282)
	B	β-(1,4)-gluco-oligosaccharides	*B. breve* UCC2003	(250)
2	**-**	FOS	*L. acidophilus* NCFM	(142)
3	**-**	Arabino-oligosaccharides[1]	*Geobacillus stearothermophilus*	(283)
4	**A**	Maltose	*L. casei* BL23 *L. acidophilus* NCFM	(199) Appendix 6.2
	B	Putative maltose	*B. animalis* subsp. *lactis* Bl-04	Appendix 6.2
	C	α-(1,4)-malto-oligosaccharides	*Listeria monocytogenes* *Streptococcus pneumoniae*	(200) (213)
	D	β-Cyclodextrin, maltose	*Bacillus subtilis*	(281)
	E	Trehalose, maltose, palatinose	*Thermus thermophilus* HB27	(212)
	F	Maltose Maltotriose	*B. longum* NCC2705 *B. animalis* subsp. *lactis* Bl-04	(198) Appendix 6.2
5	**A**	Lactose β-galacto-oligosaccharides	*B. longum* NCC2705 *B. animalis* subsp. *lactis* Bl-04	(198) Appendix 6.2
	B	Lacto-*N*-biose[2]	*B. bifidum* *B. longum*	(211) (215)
	C	Unknown	None specified	-
6	**-**	β-(1,4)-xylo-oligosaccharides	*B. animalis* subsp. *lactis* Bl-04 *Streptomyces thermoviolaceus* OPC-520	Appendix 6.2 (284)
7	**A**	Raffinose Raffinose and isomaltose	*L. acidophilus* NCFM *Streptococcus mutans*	Appendix 6.1 (276)
	B	Raffinose RFO and IMO	*B. longum* NCC2705 *B. animalis* subsp. *lactis* Bl-04	(198) Appendix 6.4
8	**-**	Gal-α-(1,3)-Fuc-α(1,2)-Gal	*Streptococcus pneumonia* SP3-BS71	(214)
9	**-**	Laminaribose[3]	*Clostridium thermocellum* ACTT 27405	(282)
Root	**-**	Fructose	*B. longum* NCC2705	(285)

[1] α-(1,5)-arabino-oligosaccharides DP 2–8. [2] (Gal*p*-β-(1–3)-GlcNAc). [3] (β-D-Glc*p*-(1–3)-D-Glc*p*).

Notably, analysis of species distribution within the maltose binding protein containing sub-clusters (4A–4F) show the sub-clusters 4A, 4B and 4F to be dominated by probiotic lactobacilli and bifidobacteria as compared to sub-cluster 4E (extremophile bacteria), 4C (human pathogenic bacteria) and 3D (soil associated bacteria). The present level of relatively little biochemical characterization beyond that of maltose binding proteins (286) does not allow detailed discrimination of the functional differentiation although analysis of the ABC transporter

associated gene clusters indicates functional distinction between transporters, and the gene clusters where they are encoded from thermophiles and mesophile bacteria (212).

The shortcomings of the sequence analysis of the solute binding proteins can be amended by mapping co-encoded glycoside hydrolases within related gene clusters of solute binding proteins for additional specificity information to assess the functionality of ABC transporters (Appendix 6.2). The observation of niche specificity of the identified clusters, and sub-clusters, relating both to functionality and selected species supports the paradigm of selective metabolism by probiotic organisms, as also observed within sub-cluster 7B, being almost exclusively populated by bifidobacteria, for uptake of the proposed prebiotic lacto-*N*-biose fraction of human milk oligosaccharide (211, 287). Gene landscape analysis of co-encoded glycoside hydrolases to solute binding proteins from sub-cluster 5C, with no experimentally characterized solute binding protein representative or defined taxonomical group, identified three combinations of GH families. The first sub-cluster encoded a GH2, the second sub-cluster a GH42 with a GH31 and lastly the third sub-cluster encoded a GH42 with a GH53 as exemplified by the protein entries YP_001222851, YP_004242545 and YP_003493824, respectively, in comparison to the sub-clusters 5A, encoding a GH42, and 5B, encoding a GH112, respectively. This observation theoretically links sub-cluster 5C to β-galactoside utilization, but also highlights a weakness in homology-based deduction of function, where the potential future work lies in annotation of gene clusters rather than single genes and furthermore experimentally characterize novel solute binding proteins within microbial niche areas to understand part of the underlying mechanisms for selective metabolism of probiotics in comparison to GIT commensal and pathogenic bacteria.

In summary, genome mining within *L. acidophilus* NCFM and *B. lactis* Bl-04 identified putative genes involved in uptake and hydrolysis of oligosaccharides several of which were proposed to possess prebiotic activity. Gene landscape analysis of the identified genes enabled functional assessment of the specificity regarding prebiotic utilization by *L. acidophilus* NCFM and *B. lactis* Bl-04 showing that both likely adopt a scavenging role in the GIT for carbohydrate utilization, and are dependent on other organisms to process polysaccharide into oligosaccharides. Evaluation of identified gene clusters proposed a higher number and more substrate specific transporters for *L. acidophilus* NCFM compared to the fewer transporters for *B. lactis* Bl-04, which however probably have broader specificities. Interestingly, both *L.*

acidophilus NCFM and *B. lactis* Bl-04 encode putative oligosaccharide transporters, where no functional homolog could be identified hence hampering the deduction of functions *in silico* to a speculative level and requiring experimental characterization.

3.2 Transcriptional analysis of potential prebiotic utilization

3.2.1 Selection of potential prebiotics for transcriptomics analysis

From the *in silico* assessment of potential prebiotic utilization by *B. lactis* Bl-04 and *L. acidophilus* NCFM, it was possible to hypothesize metabolic pathways for uptake and hydrolysis of various oligosaccharides. In parallel with the genome mining, the abilities of selected carbohydrate prebiotic candidates to support the growth of *L. acidophilus* NCFM and *B. lactis* Bl-04 were screened in mono bacterial cultures (van Zanten et al, manuscript submitted to PLoS ONE). These growth data showed a comparative overview of potential prebiotics utilizable by the tested probiotics with respect to the total bacterial growth. The growth data of potential prebiotics selected for transcriptional analysis within the current PhD project are summarized in Table 3-7. All potential prebiotics specified were selected based on the presence of predicted pathways and the criteria below:

Human indigestibility. The ability of prebiotics to bypass human digestion and reach the lower GIT is a fundamental part of the definition of prebiotics (41) and hence a key property for potential prebiotics to fulfill. Maltotriose does not fulfill this criterion but the functional glucoside composition is relevant as discussed for the functional glycoside compositions below.

Growth parameters. Oligosaccharides yielding a higher growth than a glucose reference would indicate efficient metabolism and hence a potential efficient synbiotic combination.

Functional glycoside composition. With the character of the present study to primarily map proteins involved with potential prebiotic uptake and catabolism, the selected carbohydrates were all covering related groups of glycoside structures and linkage types, to correlate the transcriptional findings to specific glycoside compositions (included in Table 3-7).

Table 3-7: Growth yield of *L. acidophilus* NCFM and *B. lactis* Bl-04 on carbohydrates used for preparation of cultures for transcriptional analysis (van Zanten et al, manuscript submitted to PLoS ONE). The carbohydrate linkage type defines the functional groups of oligosaccharides. Growth yield given in brackets denote combinations of bacteria and carbohydrates not used for transcriptional analysis. No growth data were obtained on XOS yet *B. lactis* Bl-04 cultures were prepared for transcriptional analysis.

Carbohydrate[1]	Principal carbohydrate linkage type	Relative growth of *L. acidophilus* NCFM[2]	Relative growth of *B. lactis* Bl-04[2]
Glucose	none	100	100
Melibiose	α-galactoside	(4)	95
Raffinose	α-galactoside	133	118
Stachyose	α-galactoside	81	98
Isomaltose	α-glucoside	129	109
Isomaltulose	α-glucoside	133	(13)
Panose	α-glucoside	87	125
Polydextrose	α-glucoside	30	(14)
Maltotriose	α-glucoside	(4)	140
GOS	β-galactoside	80	97
Lactitol	β-galactoside	47	(3)
Cellobiose	β-glucoside	131	40
Gentiobiose	β-glucoside	125	72
Barley β-glucan oligomers	β-glucoside	10	(5)
Xylobiose	β-xyloside	(4)	82
XOS	β-xyloside	(Not tested)	Not tested

[1] The chemical structures for all listed oligosaccharides are given in Table 1-4. [2] Relative to glucose

3.2.2 Summary of potential prebiotics induced differential transcriptomics

The differential transcriptomics analysis of the utilization of the potential prebiotic by *L. acidophilus* NCFM and *B. lactis* Bl-04 have been prepared as separate research articles (Appendices 6.1 and 6.2 respectively). These manuscripts cover the observations from the global transcriptomics, differential gene upregulation and *in silico* comparative analysis of the identified genes in context of selective prebiotic utilization. It is hence the purpose of the following sections to combine the findings of the two studies, to discuss the comparative assessment of the identified genes in relation to uptake and catabolism of potential prebiotics within two phylogenetically distant probiotic bacteria. A main observation for both bacteria, grown on the

different oligosaccharides and glucose, was how the global transcriptome remained largely unchanged regardless of the source of carbohydrate utilized and only single gene clusters were seemingly differentially upregulated when compared to the carbohydrate utilized. Analysis of variance (ANOVA) of the differential transcriptome revealed upregulation of selected loci involved with oligosaccharide uptake and catabolism constituted in operons and structured gene clusters. The encoded set of carbohydrate transporters and glycoside hydrolases are summarized in Table 3-8.

Table 3-8: Upregulated pathways for uptake and hydrolysis of potential prebiotics in *L. acidophilus* NCFM and *B. lactis* Bl-04. The transporter and hydrolase(s) for each identified pathway are listed horizontally and for those oligosaccharides were two pathways were identified, each pathway is shown on a separate line. Oligosaccharides not included in the transcriptional setup for the given strain are listed as not investigated (N.I.).

Carbohydrate	*L. acidophilus* NCFM		*B. lactis* Bl-04	
	Transporter	GH families	Transporter	GH families
Melibiose	N.I.	N.I.	ABC	36
Raffinose	ABC	13_18, 36	ABC	13_18, 36
Stachyose	ABC	13_18, 36	ABC	13_18, 36
Isomaltose	PTS	4	ABC	13_?
Isomaltulose	PTS	4	N.I.	N.I.
Panose	PTS	4, 65	ABC ABC	13_? 13_30, 77
Maltotriose	N.I.	N.I.	ABC	13_30, 77
Polydextrose	PTS ABC	4, 65 32	N.I.	N.I.
GOS	GPH	2, 42	MFS ABC	2 42
Lactitol	GPH	2, 42	N.I.	N.I.
Gentiobiose	PTS	1	MFS	42
Cellobiose	PTS	1	MFS	1
β-glucan oligomers	PTS PTS	1 1, 1	N.I.	N.I.
Xylobiose	N.I.	N.I.	ABC	43, 43, 43
XOS	N.I.	N.I.	ABC	43, 43, 43

The performed transcriptional studies corroborated the wide capabilities of both bacteria to utilize ranges of potential prebiotics and substantiated the *in silico* predictions of oligosaccharide utilization (Section 3.1.6). Especially the identification of several PTS systems for *L. acidophilus* NCFM and MFS permeases for *B. lactis* Bl-04 represents a significant resource for annotation of protein homologs suffering from the limitations of the current lack of biochemical characterizations. The number of identified ABC transporters in both bacteria further verified

and supported the diversity of oligosaccharide solute binding proteins of ABC transporters as presented earlier (Section 3.1.7). Notably, also differences in glycoside hydrolase facilitated catabolism of potential prebiotics were observed as discussed in the following.

3.2.3 Comparative pathway analysis of potential prebiotic utilization by *L. acidophilus* NCFM and *B. lactis* Bl-04

Based on the functional glycoside composition of the selected potential prebiotics (listed in Table 3-7), it is of interest to compare the routes of utilization to map how representative probiotics bacteria potentially excel in selective utilization of prebiotics. The comparison of *L. acidophilus* NCFM and *B. lactis* Bl-04, by composition of the oligosaccharides utilized, will aid in comparative genomics studies, beyond the scope of the current project, to understand the functional differences between probiotic lactobacilli and bifidobacteria, the commensal microbiota and opportunistic pathogens entering the GIT, as they are becoming available from emerging large-scale sequencing projects (108, 113, 288).

α-galactosides (melibiose, raffinose and stachyose). Commonly for both bacteria, an ABC transporter and a GH36 α-galactosidase were upregulated by raffinose and stachyose, and melibiose for *B. lactis* Bl-04. Notably, the *B. lactis* Bl-04 transporter showed a dual specificity by also being upregulated by the α-glucosides isomaltose and panose, which was not observed for *L. acidophilus* NCFM where a PTS transporter was identified (see the α-glucosides section below). The dual specificity was also suggested for a ABC transporter from *Streptococcus mutans* (276, 289), indicating that the dual specificity is not a feature only found within bifidobacteria. Sequence analysis (Section 3.1.7) did not indicate any differentiation of the two types of ABC mediating raffinose transport, hence lacking the predictive power to differentiate whether a novel raffinose transporter would exhibit a dual specificity for RFO and α-glucosides. This currently requires experimental work to be answered. Noticeably, crystal structures of relevant solute binding proteins in complexes with oligosaccharides may disclose structural determinants for future annotations of sub-specificities at the gene sequence level as later discussed (Section 3.3.2).

α-glucosides (isomaltose, isomaltulose, panose, maltotriose, polydextrose). As listed above, *B. lactis* Bl-04 encoded a dual specificity ABC transporter indicated also to transport α-(1,6)-

linkage containing glucosides (isomaltose and panose), but not the α-(1,4) linked maltotriose, which was possibly transported by an annotated dedicated maltose/malto-oligosaccharide ABC transporter highly specific for GIT associated actinobacteria. Yet a complementing transport mechanism was found by a PTS permease (LBA0606–0609) in *L. acidophilus* NCFM for uptake of isomaltose, isomaltulose, panose and possibly also fractions of polydextrose. The putative specificity and potential for trisaccharide uptake of this PTS permease is novel. As was the hydrolytic interplay with a novel GH4 isomaltose-6-phosphate hydrolase, selectively found in GIT associated lactobacilli (290), revealing a metabolic pathway expanding the knowledge of previous studies of α-glucan disaccharide specific PTS permeases coupled with a GH4 enzyme (LBA1689) activity (291–293). Notably *L. acidophilus* NCFM also encodes a putative maltose ABC transporter yet the lack of tested purely α-(1,4) linked glucosides apparently precluded upregulation of this locus. Maltose specific genes were though found to be upregulated in both strains by two different hydrolytic pathways. In *L. acidophilus* NCFM a GH65 maltose phosphorylase (LBA1872) was upregulated with LBA1689 (which is proposed to hydrolyze the phosphorylated α-1,6 linked-glucose from panose to release maltose) whereas in *B. lactis* Bl-04 maltotriose upregulated an α-glucosidase (Balac_1573) and a α-1,4-glucanotransferase (Balac_1567). This highlights strain differences also on the level of catabolic pathways of *L. acidophilus* NCFM and *B. lactic* Bl-04.

β-glucosides (cellobiose, gentiobiose and barley β-glucan hydrolysate). Clear metabolic differentiation was observed for β-glucoside utilization within *L. acidophilus* NCFM where dedicated PTS encoding loci were revealed for differential recognition of β-(1,4) and β-(1,6) glucosidic linkages. With the lack of PTS permeases in *B. lactis* Bl-04, uptake of gentiobiose was facilitated by a MFS transporter and interestingly indicated to be hydrolyzed by a GH42 putative β-galactoside. This indicates the first observation of gentiobiose uptake by an MFS permease and a novel specificity within GH42 to date only harboring β-galactosidases. The *in silico* phylogenetic mapping of the GH42 family showed the gentiobiose specific GH42 to differ significantly from enzymes of known specificities within the GH42 family, supporting this novel observation (Alexander Viborg Holm, unpublished results). No specific cellobiose transporter was found in *B. lactis* Bl-04 neither from the *in silico* nor the transcriptional analysis, in agreement with the relatively low growth observed on cellobiose.

β-galactosides (GOS and lactitol). Both bacteria displayed common catabolic pathways for intracellular GOS by hydrolysis by GH2 and GH42 β-galactosidases, yet *B. lactis* Bl-04 encoded two GOS transporters in different loci (an ABC transporter associated with a GH42 β-galactosidase and an MFS permease associated with the GH2 β-galactosidase) for GOS uptake compared to the GPH permease identified in *L. acidophilus* NCFM. It is unknown from the present study what fraction of GOS chain-lengths are utilized by B. lactis Bl-04, yet the relative high growth yield (Table 3-7) indicates utilization also of GOS with a higher degree of polymerization. From previous knowledge of transporters, it can hypothesized that the two transporters may act in parallel with uptake of short chain GOS by the MFS and longer chain GOS to be facilitated by the ABC transporter. In comparison the GPH permease encoded by *L. acidophilus* NCFM has evolved a wide substrate specificity to comprise lactose, GOS and the sugar alcohol lactitol with hydrolysis by both a GH2 and a GH42 β-galactoside (Appendix 6.3).

β-xylosides (xylobiose and XOS). The ability of *B. lactis* Bl-04 to utilize β-xylosides, compared to *L. acidophilus* NCFM, which cannot utilize XOS, may reflect the organism's adaption to utilization of dietary plant derived human non-digestible carbohydrates. Biophysical characterization of the solute binding protein of the XOS ABC transporter demonstrated binding of arabinosylated xylo-oligosaccharides (Ejby et al., manuscript in preparation). This supports the prediction of GH43 arabinofuranosidases within the XOS utilization locus of *B. lactis* Bl-04, signifying how the locus may also enable the utilization of arabinosylated XOS. No methylated, acetylated or feruloylated XOS substrates were tested to support the predictions of novel putative esterases found within the locus, although these esterifications are commonly found in plant xylan and removed by various xylan acetyl esterases (294, 295), to which the putative estereases identified in *B. lactis* Bl-04 showed, however, less than 30% amino acid sequence identity towards.

3.2.4 Correlation of *in silico* predictions and transcriptional observations

The transcriptional identification of oligosaccharide transporters and hydrolases validates the *in silico* predictions for ABC transporter containing pathways, but also highlighted novel specificities which could not be deduced from amino acid sequence similarity e.g. in the case of PTS transporters where no homologs had been identified in related organisms previously.

However, the interpretation of the differential transcriptomes for heterogeneous oligosaccharide preparations such as polydextrose and barley β-glucan hydrolysate, becomes complicated as it cannot be deduced by the applied methods, which fractions of the carbohydrate sample have been utilized at the time of culture harvest in this setup with induction of multiple transporters. Here data on disaccharides, e.g. isomaltose or cellobiose, assist in determining the linkage and glycoside specificity of single transporters. Such knowledge for the single transporters can in turn support theoretical utilization of more complex oligosaccharide mixtures.

From the genome mining of *B. lactis* Bl-04 a putative extracellular β-mannosidase (Balac_1450) was proposed. Likewise a putative extracellular pullulanse (LBA1710) was identified in *L. acidophilus* NCFM, yet no oligosaccharides being potential substrates for these enzymes were included in the experimental setup, hence no upregulation of these genes was observed. Yet recombinant proteins of both genes showed enzymatic activity towards β-mannan and pullulan, respectively (Personal communication, Abou Hachem, M.) validating the proposed specificities of these enzymes and their possible involvement in potential prebiotic utilization.

3.2.5 Experimental design and limitations of DNA microarrays for gene identification

A key step in analysis of differential transcriptomics is the ANOVA modeling to identify upregulated genes with statistical significance among the total gene transcriptome data. Hence it is a consideration of how many samples, biological- and technical replicates to include in the experimental design to obtain a useful statistically reliable data analysis. Based on a previous study using the same experimental platform for studying gene expression in *L. acidophilus* NCFM (143), it was estimated that a 12 samples setup with two technical replicates would yield sufficient quality data to allow detection of single gene upregulation with statistical significance. No biological replicates are included as bacterial cultures were assumed to be genetically identical and introducing a biological replicate would show data variance based on undesirable differences in culture preparation and handling rather than the potential variance in mRNA preparation and labeling, and hybridization inefficiencies reflected by technical replicates (296, 297). Nonetheless analysis of genes below the significance threshold (p-value = $10^{-4.74}$) for *L. acidophilus* NCFM showed how the putative extracellular pullulanase (LBA1710), despite being

more than 20 fold upregulated by polydextrose compared to glucose, was below the cut off with a p-value of $10^{-4.73}$ suggesting transcriptomics data need experimental validation to support the observations, or lack of observations, based on the *in silico* and transcriptional analysis.

Transcriptional analysis revealed genes with constitutively high gene expression values for both bacteria throughout all tested carbohydrate conditions. Yet such genes with functions for carbohydrate utilization, as the glucose PTS (LBA0452, 0455–0456) transporter in *L. acidophilus* NCFM or the putative β-(1,4)-glucosidase (Balac_0151) from *B. lactis* Bl-04, could not be identified from ANOVA models and thus these genes require functional validation. This could be done by semi-quantitative polymerase chain reaction of the analyzed genes compared to house-keeping genes (such as genes encoded RNA polymerase or ribosomal RNA) with constantly high expression. Especially validation of transcriptional analysis is essential to support the claims of gene annotations presented in the previous sections, as differences in mRNA levels are not linearly related to quantitative levels of functional protein or assist in deduction of protein functions. The following section will evaluate the functional validation of the transcriptional findings by functional genomics and recombinant protein characterization.

3.3 Functional characterization of genes and proteins involved with potential prebiotic uptake and catabolism

Transcriptional analysis of genes involved with the utilization of potential prebiotic by *L. acidophilus* NCFM and *B. lactis* Bl-04 identified putative gene products facilitating uptake and hydrolysis of oligosaccharides. A natural next step in understanding the molecular mechanisms of potential prebiotic utilization is functional characterization of these key proteins. This serves both as validation of the proposed new gene annotations based on the transcriptional findings and to characterize the broadness of protein specificity, being it of transporters or glycoside hydrolases. This section will present protein function deduction by complementing gene deletions as well as characterization of produced recombinant proteins.

3.3.1 Functional genomics of *L. acidophilus* NCFM

The recently developed counter-selective gene replacement system in *L. acidophilus* NCFM (261) enabled the construction of gene excision mutants. This was applied to generate in-frame

gene deletions of putative key gene's coding regions (Table 3-9) involved with potential prebiotic utilization presented in the previous sections. The selection of genes for deletion was primarily based on genome mining (Section 3.1.3) as the experimental work of DNA microarray sample preparation from *L. acidophilus* NCFM cultures grown on prebiotic candidates for transcriptional work was performed simultaneously with the construction of gene deletions.

In the light of the previous work, ABC transporters were hypothesized to be the route of oligosaccharide uptake as discussed in Section 3.2.2, hence substrates such as panose, barley β-glucan hydrolysate, raffinose, stachyose, polydextrose and GOS could be transported by ABC transporters, where two putative multiple sugar metabolism transporters and one maltose transporters were predicted (Table 3-1). With three ABC transporters identified, it was rationalized that these would display broad substrate specificities and hence the corresponding solute binding proteins were deleted for indirect functional characterization (as listed in Table 3-9). Additionally a GH36 α-galactosidase was selected for gene deletion to further investigate RFO metabolism. Finally by gene landscape analysis of the putative lactose specific GPH permease, it was proposed how the GPH transporter could potentially facilitate uptake of GOS hence the permease was also targeted for gene excision.

The phenotypic characterization of the GPH permease role in uptake of the prebiotic GOS and lactitol is presented in Appendix 6.3. Notably, this finding validate the hypothesized GOS uptake by a GPH permease, being the first identified GOS specific transporter within the Lactobacillus genus and adds functional support to the discussion of how lactose is either transported by GPH permeases or lactose specific PTS permeases (231) in lactobacilli. The phenotypic characterization of the GH36 α-galactosidase and raffinose ABC transporter is presented in Appendix 6.1, where the gene deletions serve to validate the two essential steps in the RFO utilization of L. acidophilus NCFM. The remaining two mutants were not yet phenotypic characterized based on the transcriptional findings and time limitations.

Table 3-9: *L. acidophilus* NCFM gene deletion mutants constructed in this project. The deleted genes are given by their locustag numbers and the gene annotation refers to the *in silico* predictions in Section 3.1.3 or experimental other work, if available. The references correspond to the construction and characterization of the mutant, where the mutants of LBA0502 and LBA1866 were only constructed and not tested.

Locus tag	Annotation	Characterized phenotype	Reference
LBA0502	FOS solute binding protein (142)	Not characterized in this study	This work
LBA1438	GH36 α-galactosidase	No growth observed on melibiose, raffinose and stachyose	Appendix 6.1
LBA1442	Raffinose solute binding protein	No growth observed on melibiose, raffinose and stachyose	Appendix 6.1
LBA1463	Lactose GPH permease	No growth observed on lactose, lactitol and GOS	Appendix 6.3
LBA1866	Putative maltose solute binding protein	Not characterized in this study	This work

The differential transcriptomics, presented in Section 3.2.2, depicted a surprising alternative utilization pathway of potential prebiotics beyond what was earlier rationalized *in silico*. Especially the above functionally characterized RFO ABC transporter in *L. acidophilus* NCFM was found to be RFO specific and not to be upregulated by isomaltose as previously found for putative multiple sugar metabolism (Msm) ABC transporter (276). The identified utilization pathway of isomaltose and panose was based on a PTS permease for uptake rather than an ABC transporter as initially predicted (Table 3-1). In general, the impact of PTS permease facilitated transport was greater than expected and suggested novel specificities such as gentiobiose and panose to PTS permeases, which would be candidates for validation by functional genomics using the gene deletion system applied in the current work.

The functional genomics presented, confirms the gene products involvement in utilization of prebiotic GOS and potential prebiotic RFO by *L. acidophilus* NCFM. Future work with these gene deletion mutants, linking their utilization phenotype to a proliferating role in the GIT, could further assess prebiotic utilization as an important probiotic characteristic. This hypothesis could be tested in a complex fermentation setup (59) or using an *in vivo* murine model to test the survivability of the *L. acidophilus* NCFM mutants unable to utilize otherwise selective utilizable substrates, as it was previously reported how a mannose specific PTS permease was linked to elongated gut persistence of *L. johnsonii* NCC553 (298).

3.3.2 Structure-function relationship of a solute binding protein with dual substrate specificity

The identification of the RFO specific ABC transporter in *L. acidophilus* NCFM is in contrast to the transcriptional findings from *B. lactis* Bl-04, presented in Appendix 6.2, where a putative raffinose and isomaltose dual specific ABC transporter (Balac_1597–1599) was identified.

The phylogenetic analysis of oligosaccharide solute binding proteins of ABC transporter (Figure 3-1) showed a divergence of putative raffinose binding proteins into two clusters based on a taxonomical drift rather than functionality, hence not reflecting the functional diversity observed here by the mono-specificity represented by the solute binding protein of *L. acidophilus* NCFM LBA1442 and dual specificity represented by the solute binding protein *Bl*16GBP (Balac_1599) encoded by *B. lactis* Bl-04. The dual substrate specificity of *Bl*16GBP was functionally validated by recombinant protein production and biophysical characterization as presented in Appendix 6.4 to confirm binding of RFO and IMO types of oligosaccharides. However, sequence comparison to explain the difference in substrate binding between the two raffinose binding proteins was hampered by the overall low amino acid sequence identity of 25.7% between LBA1442 and Balac_1599.

To comprehend the molecular architecture underlying the broad specificity of *Bl*16GBP, protein crystallization of recombinant *Bl*16GBP performed (as described in Section 2.4) to identify residues and structural motifs unique to the RFO/IMO class of oligosaccharide solute binding proteins. Protein crystals were only obtained in the presence of either panose or raffinose, which could be explained by how binding of an oligosaccharide stabilizes the assumed flexible protein conformation of the solute binding protein (210) hence enabling tight and homogenous packing of the crystal lattice. Electron density maps were obtained for protein-carbohydrate complexes with panose and raffinose (Table 3-10), moreover a dataset of selenomethionine labeled Balac_1599 in complex with panose was obtained (Morten Ejby, Post doc. Andreja Vujicic-Zagar and Associate Professor Dirk Slotboom, preliminary results). Despite high quality datasets of Bl16GBP in complex with panose or raffinose which theoretically could reach atomic resolution structures of the complexes, a final protein structure of Bl16GBP could not be modeled due to the lack of experimentally determined phases of the electron densities.

Table 3-10: Obtained electron density datasets of *Bl*16GBP protein complexes.

Parameter	*Bl*16GBP	*Bl*16GBP
Ligand/soak	Panose	Raffinose
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions *a b c* (Å)	55.5 90.9 146.6	55.5 90.9 146.6
$\alpha\ \beta\ \gamma$ (°)	90.0 90.0 90.0	90.0 90.0 90.0
No. of molecules/ AU	2	2
Solvent content (%)	41.5 (V_m=2.1)	41.5 (V_m=2.1)
Resolution range (Å)	47.3–1.9	50.2–1.6
R_{sym} (%)	12.1 (64.6)	11.2 (63.8)
$I/\sigma(I)$	12.0 (3.3)	11.9 (3.2)
Completeness (%)	98.6 (96.7)	99.6 (97.7)
Redundancy	7.4 (7.4)	7.4 (7.4)

To overcome this, crystals were prepared of selenomethionine labeled *Bl*16GBP to aid in solving the phase problem, however currently only a too low resolution (~8 Å) dataset have been obtained. Alternatively, molecular replacement, using homologous previously determined protein structures, was attempted without success due to sequence identity below 30% to the most similar structurally reported homologs. Experimental work to determine the phases by e.g. obtaining higher resolution datasets of selenomethionine labeled *Bl*16GBP crystals continue and could lead to structural determination of the *Bl*16GBP in complex with raffinose and panose and thus reveal how the dual substrate recognition is structurally manifested.

The characterization of recombinant protein and gene deletions, corroborate on the gene identifications and annotation from *in silico* genome mining and the two differential transcriptions analysis to substantiate the proposed pathways of prebiotic utilization.

4 Conclusions and perspectives

The various health-benefits related to probiotic microorganisms is a major industrial and scientific area, expanding with the advances of clinical data and deeper understanding of the gastrointestinal microbiome as a significant parameter for human health. However, there is still a gap of knowledge related to the documentation of how probiotic strains confer their probiotic effects through mechanisms such as selective utilization of prebiotics.

The overall aim of the current Ph.D. project was to identify and characterize genes involved with uptake and catabolism of potential carbohydrate prebiotics by two commercial probiotic strains, *L. acidophilus* NCFM and *B. lactis* Bl-04. This was achieved by a strategy of genome mining and gene-landscape analysis, differential transcriptomics followed by selection of key genes to further study by functional genomics and protein characterization.

In silico assessment of the putative carbohydrate utilization systems of *L. acidophilus* NCFM and *B. lactis* Bl-04 revealed how both bacteria essentially lack extracellular glycoside hydrolases and hence scavenge polysaccharide degradation products form GIT symbionts. Thus the multitude of encoded oligosaccharide transport systems is the main determinant of substrate specificity of the bacteria. This lead to selection of potential oligosaccharide-prebiotics covering groups of α- and β-linked glycoside-types of galactosides, glucosides and xylosides used for preparation of cultures for transcriptional analysis.

The differential transcriptomics of *L. acidophilus* NCFM revealed an extensive diversity of upregulated PTS permeases for α-1,6-glucosides and β-glucosides together with a broad specificity β-galactoside GPH permease and ABC transporters associated with uptake of oligosaccharides. Key genes, encoding the putative GOS permease, a raffinose solute binding protein and a GH36 α-galactoside respectively, were selected for deletion and their mutant phenotypes confirmed their roles in prebiotic utilization. The transcriptional analysis of *B. animalis* subsp. *lactis* Bl-04 showed putative gene products related to molecular probiotic functions being constitutively expressed hence linking to the probiotic nature of the bacterium. The differential gene findings identified specific gene cluster involved with oligosaccharide utilization, where novel specificities were found for MFS permeases and ABC transporter encoding gene clusters, displaying overall broad substrate specificity. In comparison of the two

strains, a seemingly homologous raffinose specific ABC transporter was cloned from *B. lactis* Bl-04 and produced. The biochemical characterization confirmed a dual substrate specificity for RFO and IMO, yet future protein structural work will aid in understanding this dual substrate specificity.

The identified pathways within *L. acidophilus* NCFM and *B. lactis* Bl-04 revealed novel insight into prebiotic utilization leading the way to further substantiation of the mechanism of probiotic actions by documentation of protein molecular interactions with dietary carbohydrates.

The current work with oligosaccharide transporter and glycoside hydrolase identification adds a significant contribution to the understanding of oligosaccharide uptake and how carbohydrate catabolism affects the global gene expression with a potential role beyond energy turn-over. The studied two strains represent highly important groups of probiotics. First, *L. acidophilus* NCFM as a member of the *acidophilus*-complex of lactobacilli and a model probiotic, which with the presented findings has been further characterized to support a crucial lack of knowledge for identification of novel carbohydrate transport systems. Second, *B. lactis* Bl-04 as a substantially less functionally characterized, yet clinically documented probiotic strain, serves with its commercial value as an interesting candidate to study on the molecular level of gene expression to understand the basis transcript and response to potential prebiotic carbohydrates.

The current results have been presented in a defined scope of strain comparison, however with the recent availability of the human microbiome metagenomics, it will be possible to use the identified putative gene clusters as biomarkers to screen larger data set and quantify the presence in microbiome phylogenetic niches, thus hypothesize which subset of the microbiome can be selectively stimulated by supplemented (potential) prebiotic.

The immediate future validation of this work will be to show the selection utilization of the applied potential prebiotics in a complex fermentation setup inoculated with either *L. acidophilus* NCFM or *B. lactis* Bl-04 (or both). Any synergistic combination of candidate prebiotic and probiotic will confirm a selective fermentation, dependent on the dose supplemented, and hence establish a functionally substantiated combination of pre- and probiotic for further *in vivo* validation of the synbiotic formulations.

5 References

1. Reid G, *et al* (2003) New scientific paradigms for probiotics and prebiotics. *J Clin Gastroenterol* 37: 105–118.

2. Roberfroid M, *et al* (2010) Prebiotic effects: Metabolic and health benefits. *Br J Nutr* 104 (Suppl. 2): S1–S63.

3. Greenblum S, Turnbaugh PJ & Borenstein E (2012) Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proc Natl Acad Sci USA* 109: 594–599.

4. Reid G, *et al* (2010) Responders and non-responders to probiotic interventions: How can we improve the odds? *Gut Microbes* 1: 200–204.

5. Kleerebezem M & Vaughan EE (2009) Probiotic and gut lactobacilli and bifidobacteria: Molecular approaches to study diversity and activity. *Annu Rev Microbiol* 63: 269–290.

6. Brown H (1995) Ilya mechnikov and his studies on comparative inflammation. *Proc Soc Exp Biol Med* 209: 99–101.

7. Fuller R (1989) Probiotics in man and animals. *J Appl Bacteriol* 66: 365–378.

8. FAO and WHO (2002) Guidelines for the evaluation of probiotics in food. Joint FAO/WHO Working Group Report (London, Ontario, Canada, April 30 and May 1, 2002).

9. Sanders ME (2008) Probiotics: Definition, sources, selection, and uses. *Clin Infect Dis* 46: S58–S61.

10. Hoffman FA, Heimbach JT, Sanders ME & Hibberd PL (2008) Executive summary: Scientific and regulatory challenges of development of probiotics as foods and drugs. *Clin Infect Dis* 46: S53–57.

11. Gibson GR & Roberfroid MB (1995) Dietary modulation of the human colonic microbiota: Introducing the concept of prebiotics. *J Nutr* 125: 1401–1412.

12. Neu J, Lorca G, Kingma SD & Triplett EW (2010) The intestinal microbiome: Relationship to type 1 diabetes. *Endocrinol Metab Clin North Am* 39: 563–571.

13. Burcelin R, Serino M, Chabo C, Blasco-Baque V & Amar J (2011) Gut microbiota and diabetes: From pathogenesis to therapeutic perspective. *Acta Diabetol* 48: 257–273.

14. Rijkers GT, *et al* (2010) Guidance for substantiating the evidence for beneficial effects of probiotics: Current status and recommendations for future research. *J Nutr* 140: S671–S676.

15. de Vrese M, *et al* (2005) Effect of *Lactobacillus gasseri* PA 16/8, *Bifidobacterium longum* SP 07/3, *B. bifidum* MF 20/5 on common cold episodes: A double blind, randomized, controlled trial. *Clin Nutr* 24: 481–491.

16. Berggren A, Lazou Ahren I, Larsson N & Onning G (2011) Randomised, double-blind and placebo-controlled study using new probiotic lactobacilli for strengthening the body immune defence against viral infections. *Eur J Nutr* 50: 203–210.

17. Abad CL & Safdar N (2009) The role of *Lactobacillus* probiotics in the treatment or prevention of urogenital infections - a systematic review. *J Chemother* 21: 243–252.

18. Barrons R & Tassone D (2008) Use of *Lactobacillus* probiotics for bacterial genitourinary infections in women: A review. *Clin Ther* 30: 453–468.

19. Kalliomaki M, Salminen S, Poussa T, Arvilommi H & Isolauri E (2003) Probiotics and prevention of atopic disease: 4-year follow-up of a randomised placebo-controlled trial. *Lancet* 361: 1869–1871.

20. Tamura M, *et al* (2007) Effects of probiotics on allergic rhinitis induced by japanese cedar pollen: Randomized double-blind, placebo-controlled clinical trial. *Int Arch Allergy Immunol* 143: 75–82.

21. Sazawal S, *et al* (2006) Efficacy of probiotics in prevention of acute diarrhoea: A meta-analysis of masked, randomised, placebo-controlled trials. *Lancet Infect Dis* 6: 374–382.

22. Guandalini S (2011) Probiotics for prevention and treatment of diarrhea. *J Clin Gastroenterol* 45: S149–153.

23. Ganguli K & Walker WA (2011) Probiotics in the prevention of necrotizing enterocolitis. *J Clin Gastroenterol* 45: S133–138.

24. Chouraqui JP, Van Egroo LD & Fichot MC (2004) Acidified milk formula supplemented with *Bifidobacterium lactis*: Impact on infant diarrhea in residential care settings. *J Pediatr Gastroenterol Nutr* 38: 288–292.

25. Mane J, *et al* (2011) A mixture of *Lactobacillus plantarum* CECT 7315 and CECT 7316 enhances systemic immunity in elderly subjects. A dose-response, double-blind, placebo-controlled, randomized pilot trial. *Nutr Hosp* 26: 228–235.

26. Huynh HQ, *et al* (2009) Probiotic preparation VSL#3 induces remission in children with mild to moderate acute ulcerative colitis: A pilot study. *Inflamm Bowel Dis* 15: 760–768.

27. Ng SC, *et al* (2010) Immunosuppressive effects via human intestinal dendritic cells of probiotic bacteria and steroids in the treatment of acute ulcerative colitis. *Inflamm Bowel Dis* 16: 1286–1298.

28. Gupta P, Andrew H, Kirschner BS & Guandalini S (2000) Is *Lactobacillus* GG helpful in children with Crohn's disease? Results of a preliminary, open-label study. *J Pediatr Gastroenterol Nutr* 31: 453–457.

29. Rolfe VE, Fortun PJ, Hawkey CJ & Bath-Hextall F (2006) Probiotics for maintenance of remission in Crohn's disease. *Cochrane Database Syst Rev* 4: CD004826.

30. Kumar M, *et al* (2010) Cancer-preventing attributes of probiotics: An update. *Int J Food Sci Nutr* 61: 473–496.

31. Capurso G, Marignani M & Delle Fave G (2006) Probiotics and the incidence of colorectal cancer: When evidence is not evident. *Dig Liver Dis* 38: S277–S282.

32. Gorbach SL, Barza M, Giuliano M & Jacobus NV (1988) Colonization resistance of the human intestinal microflora: Testing the hypothesis in normal volunteers. *Eur J Clin Microbiol Infect Dis* 7: 98–102.

33. Stoidis CN, Misiakos EP, Patapis P, Fotiadis CI & Spyropoulos BG (2010) Potential benefits of pro- and prebiotics on intestinal mucosal immunity and intestinal barrier in short bowel syndrome. *Nutr Res Rev* 24: 21–30.

34. Fukuda S, *et al* (2011) Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* 469: 543–547.

35. Marco ML, Pavan S & Kleerebezem M (2006) Towards understanding molecular modes of probiotic action. *Curr Opin Biotechnol* 17: 204–210.

36. Marco ML, *et al* (2009) Lifestyle of *Lactobacillus plantarum* in the mouse *caecum. Environ Microbiol* 11: 2747–2757.

37. Bird AR, Conlon MA, Christophersen CT & Topping DL (2010) Resistant starch, large bowel fermentation and a broader perspective of prebiotics and probiotics. *Benef Microbes* 1: 423–431.

38. Martens EC, *et al* (2011) Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS Biol* 9: e1001221.

39. Legette LL, *et al* (2011) Genistein, a phytoestrogen, improves total cholesterol, and synergy, a prebiotic, improves calcium utilization, but there were no synergistic effects. *Menopause* 18: 923–931.

40. Queipo-Ortuno MI, *et al* (2012) Influence of red wine polyphenols and ethanol on the gut microbiota ecology and biochemical biomarkers. *Am J Clin Nutr* 95: 1323–1334.

41. Roberfroid M (2007) Prebiotics: The concept revisited. *J Nutr* 137: 830S–837S.

42. Gibson GR, Probert HM, Loo JV, Rastall RA & Roberfroid MB (2004) Dietary modulation of the human colonic microbiota: Updating the concept of prebiotics. *Nutr Res Rev* 17: 259–275.

43. Vandenplas Y & De Hert S.G. (2011) Randomised clinical trial: The synbiotic food supplement probiotical vs. placebo for acute gastroenteritis in children. *Aliment Pharmacol Ther* 34: 862–867.

44. Björklund M, *et al* (2011) Gut microbiota of healthy elderly NSAID users is selectively modified with the administration of *Lactobacillus acidophilus* NCFM and lactitol. *Age* 19: DOI: 10.1007/s11357-011-9294-5.

45. Steed H, *et al* (2010) Clinical trial: The microbiological and immunological effects of synbiotic consumption - a randomized double-blind placebo-controlled study in active Crohn's disease. *Aliment Pharmacol Ther* 32: 872–883.

46. Hedin C, Whelan K & Lindsay JO (2007) Evidence for the use of probiotics and prebiotics in inflammatory bowel disease: A review of clinical trials. *Proc Nutr Soc* 66: 307–315.

47. Zhang F, Hang X, Fan X, Li G & Yang H (2007) Selection and optimization procedure of synbiotic for cholesterol removal. *Anaerobe* 13: 185–192.

48. Seifert S & Watzl B (2007) Inulin and oligofructose: Review of experimental data on immune modulation. *J Nutr* 137: S2563–S2567.

49. Cummings JH, Christie S & Cole TJ (2001) A study of fructo oligosaccharides in the prevention of travellers' diarrhoea. *Aliment Pharmacol Ther* 15: 1139–1145.

50. Bouhnik Y, *et al* (1999) Short-chain fructo-oligosaccharide administration dose-dependently increases fecal bifidobacteria in healthy humans. *J Nutr* 129: 113–116.

51. Benjamin JL, *et al* (2011) Randomised, double-blind, placebo-controlled trial of fructo-oligosaccharides in active Crohn's disease. *Gut* 60: 923–929.

52. Davis LM, Martinez I, Walter J, Goin C & Hutkins RW (2011) Barcoded pyrosequencing reveals that consumption of galactooligosaccharides results in a highly specific bifidogenic response in humans. *PLoS One* 6: e25200.

53. Drakoularakou A, Tzortzis G, Rastall RA & Gibson GR (2010) A double-blind, placebo-controlled, randomized human study assessing the capacity of a novel galacto-oligosaccharide mixture in reducing travellers' diarrhoea. *Eur J Clin Nutr* 64: 146–152.

54. Macfarlane GT, Steed H & Macfarlane S (2008) Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics. *J Appl Microbiol* 104: 305–344.

55. Vos AP, Knol J, Stahl B, M'rabet L & Garssen J (2010) Specific prebiotic oligosaccharides modulate the early phase of a murine vaccination response. *Int Immunopharmacol* 10: 619–625.

56. Costabile A, *et al* (2010) A double-blind, placebo-controlled, cross-over study to establish the bifidogenic effect of a very-long-chain inulin extracted from globe artichoke (*Cynara scolymus*) in healthy human subjects. *Br J Nutr* 104: 1007–1017.

57. Panesar PS & Kumari S (2011) Lactulose: Production, purification and potential applications. *Biotechnol Adv* 29: 940–948.

58. Makelainen H, Hasselwander O, Rautonen N & Ouwehand AC (2009) Panose, a new prebiotic candidate. *Lett Appl Microbiol* 49: 666–672.

59. Makelainen HS, Makivuokko HA, Salminen SJ, Rautonen NE & Ouwehand AC (2007) The effects of polydextrose and xylitol on microbial community and activity in a 4-stage colon simulator. *J Food Sci* 72: 153–159.

60. Makivuokko H, *et al* (2007) The effect of cocoa and polydextrose on bacterial fermentation in gastrointestinal tract simulations. *Biosci Biotechnol Biochem* 71: 1834–1843.

61. Jie Z, *et al* (2000) Studies on the effects of polydextrose intake on physiologic functions in Chinese people. *Am J Clin Nutr* 72: 1503–1509.

62. Hayakawa K, *et al* (1990) Effects of soybean oligosaccharides on human faecal flora. *Microb Ecol Health Dis* 3: 293–303.

63. Tenorio MD, Espinosa-Martos I, Préstamo G & Rupérez P (2010) Soybean whey enhance mineral balance and caecal fermentation in rats. *Eur J Nutr* 49: 155–163.

64. Van den Abbeele P, *et al* (2011) Arabinoxylans and inulin differentially modulate the mucosal and luminal gut microbiota and mucin-degradation in humanized rats. *Environ Microbiol* 13: 2667–2680.

65. Broekaert WF, *et al* (2011) Prebiotic and other health-related effects of cereal-derived arabinoxylans, arabinoxylan-oligosaccharides, and xylooligosaccharides. *Crit Rev Food Sci Nutr* 51: 178–194.

66. Goffin D, *et al* (2011) Will isomalto-oligosaccharides, a well-established functional food in Asia, break through the European and American market? The status of knowledge on these prebiotics. *Crit Rev Food Sci Nutr* 51: 394–409.

67. Ketabi A, Dieleman LA & Ganzle MG (2011) Influence of isomalto-oligosaccharides on intestinal microbiota in rats. *J Appl Microbiol* 110: 1297–1306.

68. Yen CH, Tseng YH, Kuo YW, Lee MC & Chen HL (2011) Long-term supplementation of isomalto-oligosaccharides improved colonic microflora profile, bowel function, and blood cholesterol levels in constipated elderly people - a placebo-controlled, diet-controlled trial. *Nutrition* 27: 445–450.

69. Ouwehand AC, Tiihonen K, Saarinen M, Putaala H & Rautonen N (2009) Influence of a combination of *Lactobacillus acidophilus* NCFM and lactitol on healthy elderly: Intestinal and immune parameters. *Br J Nutr* 101: 367–375.

70. Rycroft CE, Jones MR, Gibson GR & Rastall RA (2001) Fermentation properties of gentio-oligosaccharides. *Lett Appl Microbiol* 32: 156–161.

71. Sanz ML, Cote GL, Gibson GR & Rastall RA (2006) Selective fermentation of gentiobiose-derived oligosaccharides by human gut bacteria and influence of molecular weight. *FEMS Microbiol Ecol* 56: 383–388.

72. Olano-Martin E, Gibson GR & Rastell RA (2002) Comparison of the *in vitro* bifidogenic properties of pectins and pectic-oligosaccharides. *J Appl Microbiol* 93: 505–511.

73. Manderson K, *et al* (2005) *In vitro* determination of prebiotic properties of oligosaccharides derived from an orange juice manufacturing by-product stream. *Appl Environ Microbiol* 71: 8383–8389.

74. White LA, Newman MC, Cromwell GL & Lindemann MD (2002) Brewers dried yeast as a source of mannan oligosaccharides for weanling pigs. *J Anim Sci* 80: 2619–2628.

75. Zhao PY, Jung JH & Kim IH (2012) Effect of mannan oligosaccharides and fructan on growth performance, nutrient digestibility, blood profile, and diarrhea score in weanling pigs. *J Anim Sci* 90: 833–839.

76. Shoaf K, Mulvey GL, Armstrong GD & Hutkins RW (2006) Prebiotic galactooligosaccharides reduce adherence of enteropathogenic *Escherichia coli* to tissue culture cells. *Infect Immun* 74: 6920–6928.

77. Sinclair HR, de Slegte J, Gibson GR & Rastall RA (2009) Galactooligosaccharides (GOS) inhibit *Vibrio cholerae* toxin binding to its GM1 receptor. *J Agric Food Chem* 57: 3113–3119.

78. Quintero M, *et al* (2011) Adherence inhibition of *Cronobacter sakazakii* to intestinal epithelial cells by prebiotic oligosaccharides. *Curr Microbiol* 62: 1448–1454.

79. Searle LE, *et al* (2010) Purified galactooligosaccharide, derived from a mixture produced by the enzymic activity of *Bifidobacterium bifidum*, reduces *Salmonella enterica* serovar *Typhimurium* adhesion and invasion *in vitro* and *in vivo*. *J Med Microbiol* 59: 1428–1439.

80. Lebeer S, Vanderleyden J & De Keersmaecker S,C.J. (2010) Host interactions of probiotic bacterial surface molecules: Comparison with commensals and pathogens. *Nature Reviews. Microbiology* 8: 171–184.

81. de Kivit S, Kraneveld AD, Garssen J & Willemsen LEM (2011) Glycan recognition at the interface of the intestinal immune system: Target for immune modulation via dietary components. *Eur J Pharmacol* 668: S124–S132.

82. Ebersbach T, Andersen JB, Bergström A, Hutkins RW & Licht TR (2012) Xylo-oligosaccharides inhibit pathogen adhesion to enterocytes *in vitro*. *Res Microbiol* 163: 22–27.

83. Ebersbach T, *et al* (2010) Certain dietary carbohydrates promote *Listeria* infection in a guinea pig model, while others prevent it. *Int J Food Microbiol* 140: 218–224.

84. Petersen A, *et al* (2009) Some putative prebiotics increase the severity of *Salmonella enterica* serovar *Typhimurium* infection in mice. *BMC Microbiol* 9: 245.

85. Schouler C, Taki A, Chouikha I, Moulin-Schouleur M & Gilot P (2009) A genomic island of an extraintestinal pathogenic *Escherichia coli* strain enables the metabolism of fructooligosaccharides, which improves intestinal colonization. *J Bacteriol* 191: 388–393.

86. Porcheron G, Kut E, Canepa S, Maurel M & Schouler C (2011) Regulation of fructooligosaccharide metabolism in an extra-intestinal pathogenic *Escherichia coli* strain. *Mol Microbiol* 81: 717–733.

87. Gaggia F, Mattarelli P & Biavati B (2010) Probiotics and prebiotics in animal feeding for safe food production. *Int J Food Microbiol* 141: S15–S28.

88. Kim SK, Bhatnagar I & Kang KH (2012) Development of marine probiotics: Prospects and approach. *Adv Food Nutr Res* 65: 353–362.

89. Khan SH & Ansari FA (2007) Probiotics - the friendly bacteria with market potential in global market. *Pak J Pharm Sci* 20: 76–82.

90. Vanderhoof JA & Young R (2008) Probiotics in the United States. *Clin Infect Dis* 46: S67–S72.

91. Amagase H (2008) Current marketplace for probiotics: A Japanese perspective. *Clin Infect Dis* 46: S73–S75.

92. Saxelin M (2008) Probiotic formulations and applications, the current probiotics market, and changes in the marketplace: A European perspective. *Clin Infect Dis* 46: S76–79.

93. European Parliament and Council (2006) Regulation (EC) no 1924/2006 of the European parliament and of the council of 20 December 2006 on nutrition and health claims made on foods. Official journal of the European Union OJ L 404, (30.12.2006) 12: 3–18.

94. Rijkers GT, *et al* (2011) Health benefits and health claims of probiotics: Bridging science and marketing. *Br J Nutr* 106: 1291–1296.

95. Sanders ME, *et al* (2011) Health claims substantiation for probiotic and prebiotic products. *Gut Microbes* 2: 127–133.

96. Backhed F, Ley RE, Sonnenburg JL, Peterson DA & Gordon JI (2005) Host-bacterial mutualism in the human intestine. *Science* 307: 1915–1920.

97. Manson JM, Rauch M & Gilmore MS (2008) The commensal microbiology of the gastrointestinal tract. *Adv Exp Med Biol* 635: 15–28.

98. Comstock LE (2009) Importance of glycans to the host-*bacteroides* mutualism in the mammalian intestine. *Cell Host Microbe* 5: 522–526.

99. Gill SR, *et al* (2006) Metagenomic analysis of the human distal gut microbiome. *Science* 312: 1355–1359.

100. International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431: 931–945.

101. Zhu B, Wang X & Li L (2010) Human gut microbiome: The second genome of human body. *Protein Cell* 1: 718–725.

102. Eckburg PB, *et al* (2005) Diversity of the human intestinal microbial flora. *Science* 308: 1635–1638.

103. Sansonetti PJ & Medzhitov R (2009) Learning tolerance while fighting ignorance. *Cell* 138: 416–420.

104. Agans R, *et al* (2011) Distal gut microbiota of adolescent children is different from that of adults. *FEMS Microbiol Ecol* 77: 404–412.

105. Hopkins MJ, Sharp R & Macfarlane GT (2002) Variation in human intestinal microbiota with age. *Dig Liver Dis* 34: S12–S18.

106. De Filippo C, *et al* (2010) Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci USA* 107: 14691–14696.

107. Wu GD, *et al* (2011) Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334: 105–108.

108. Arumugam M, *et al* (2011) Enterotypes of the human gut microbiome. *Nature* 473: 174–180.

109. Khoruts A & Sadowsky MJ (2011) Therapeutic transplantation of the distal gut microbiota. *Mucosal Immunol* 4: 4–7.

110. Dethlefsen L & Relman DA (2011) Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc Natl Acad Sci USA* 108: 4554–4561.

111. Hartman AL, *et al* (2009) Human gut microbiome adopts an alternative state following small bowel transplantation. *Proc Natl Acad Sci USA* 106: 17187–17192.

112. Wu X, *et al* (2010) Molecular characterisation of the faecal microbiota in patients with type II diabetes. *Curr Microbiol* 61: 69–78.

113. Cani PD & Delzenne NM (2011) The gut microbiome as therapeutic target. *Pharmacol Ther* 130: 202–212.

114. Collado MC, Meriluoto J & Salminen S (2007) Role of commercial probiotic strains against human pathogen adhesion to intestinal mucus. *Lett Appl Microbiol* 45: 454–460.

115. Callaway TR, *et al* (2008) Probiotics, prebiotics and competitive exclusion for prophylaxis against bacterial disease. *Anim Health Res Rev* 9: 217–225.

116. Hosseini E, Grootaert C, Verstraete W & Van de Wiele T (2011) Propionate as a health-promoting microbial metabolite in the human gut. *Nutr Rev* 69: 245–258.

117. Lee JH, Li X & O'Sullivan DJ (2011) Transcription analysis of a lantibiotic gene cluster from *Bifidobacterium longum* DJO10A. *Appl Environ Microbiol* 77: 5879–5887.

118. Collado MC, Hernandez M & Sanz Y (2005) Production of bacteriocin-like inhibitory compounds by human fecal *Bifidobacterium* strains. *J Food Prot* 68: 1034–1040.

119. Konstantinov SR, *et al* (2008) S layer protein A of *Lactobacillus acidophilus* NCFM regulates immature dendritic cell and T cell functions. *Proc Natl Acad Sci USA* 105: 19474–19479.

120. Mohamadzadeh M, *et al* (2011) Regulation of induced colonic inflammation by *Lactobacillus acidophilus* deficient in lipoteichoic acid. *Proc Natl Acad Sci USA* 108: 4623–4630.

121. Munoz M, Mosquera A, Almeciga-Diaz CJ, Melendez AP & Sanchez OF (2012) Fructooligosaccharides metabolism and effect on bacteriocin production in *Lactobacillus* strains isolated from ensiled corn and molasses. *Anaerobe*: http://dx.doi.org/10.1016/j.anaerobe.2012.01.007.

122. Pan X, Wu T, Zhang L, Cai L & Song Z (2009) Influence of oligosaccharides on the growth and tolerance capacity of lactobacilli to simulated stress environment. *Lett Appl Microbiol* 48: 362–367.

123. Saghatelian A & Cravatt BF (2005) Assignment of protein function in the postgenomic era. *Nat Chem Biol* 1: 130–142.

124. Baltz RH (2011) Strain improvement in actinomycetes in the postgenomic era. *J Ind Microbiol Biotechnol* 38: 657–666.

125. de Vos WM (2011) Systems solutions by lactic acid bacteria: From paradigms to practice. *Microb Cell Fact* 10: S2.

126. Bennedsen M, Stuer-Lauridsen B, Danielsen M & Johansen E (2011) Screening for antimicrobial resistance genes and virulence factors via genome sequencing. *Appl Environ Microbiol* 77: 2785–2787.

127. Horvath P, *et al* (2009) Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int J Food Microbiol* 131: 62–70.

128. Ventura M, *et al* (2009) Genome-scale analyses of health-promoting bacteria: Probiogenomics. *Nat Rev Microbiol* 7: 61–71.

129. Siezen RJ, *et al* (2010) Phenotypic and genomic diversity of *Lactobacillus plantarum* strains isolated from various environmental niches. *Environ Microbiol* 12: 758–773.

130. Ventura M, Turroni F & van Sinderen D (2012) Probiogenomics as a tool to obtain genetic insights into adaptation of probiotic bacteria to the human gut. *Bioeng Bugs* 3: 73–79.

131. Altermann E, *et al* (2005) Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. *Proc Natl Acad Sci USA* 102: 3906–3912.

132. Lorca GL, *et al* (2007) Transport capabilities of eleven gram-positive bacteria: Comparative genomic analyses. *Biochim Biophys Acta* 1768: 1342–1366.

133. Pokusaeva K, Fitzgerald GF & van Sinderen D (2011) Carbohydrate metabolism in bifidobacteria. *Genes & Nutrition* 6: 285–306.

134. Lukjancenko O, Ussery DW & Wassenaar TM (2012) Comparative genomics of *Bifidobacterium*, *Lactobacillus* and related probiotic genera. *Microb Ecol* 63: 651–673.

135. Claesson MJ, van Sinderen D & O'Toole PW (2007) The genus *Lactobacillus* - a genomic basis for understanding its diversity. *FEMS Microbiol Lett* 269: 22–28.

136. Kleerebezem M, *et al* (2010) The extracellular biology of the lactobacilli. *FEMS Microbiol Rev* 34: 199–230.

137. Makarova K, *et al* (2006) Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci USA* 103: 15611–15616.

138. Canchaya C, Claesson MJ, Fitzgerald GF, van Sinderen D & O'Toole PW (2006) Diversity of the genus *Lactobacillus* revealed by comparative genomics of five species. *Microbiology* 152: 3185–3196.

139. Berger B, *et al* (2007) Similarity and differences in the *Lactobacillus acidophilus* group identified by polyphasic analysis and comparative genomics. *J Bacteriol* 189: 1311–1321.

140. Lebeer S, Vanderleyden J & De Keersmaecker S,C.J. (2008) Genes and molecules of lactobacilli supporting probiotic action. *Microbiology and Molecular Biology Reviews* 72: 728–764.

141. Duong T, Barrangou R, Russell WM & Klaenhammer TR (2006) Characterization of the *tre* locus and analysis of trehalose cryoprotection in *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol* 72: 1218–1225.

142. Barrangou R, Altermann E, Hutkins R, Cano R & Klaenhammer TR (2003) Functional and comparative genomic analyses of an operon involved in fructooligosaccharide utilization by *Lactobacillus acidophilus*. *Proc Natl Acad Sci USA* 100: 8957–8962.

143. Barrangou R, *et al* (2006) Global analysis of carbohydrate utilization by *Lactobacillus acidophilus* using cDNA microarrays. *Proc Natl Acad Sci USA* 103: 3816–3821.

144. Altermann E & Klaenhammer TR (2011) Group-specific comparison of four lactobacilli isolated from human sources using differential blast analysis. *Genes & Nutrition* 6: 319–340.

145. Kankainen M, *et al* (2009) Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human-mucus binding protein. *Proc Natl Acad Sci USA* 106: 17193–17198.

146. Kleerebezem M, *et al* (2003) Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc Natl Acad Sci USA* 100: 1990–1995.

147. Pridmore RD, *et al* (2004) The genome sequence of the probiotic intestinal bacterium *Lactobacillus johnsonii* NCC 533. *Proc Natl Acad Sci USA* 101: 2512–2517.

148. Azcarate-Peril MA, *et al* (2008) Analysis of the genome sequence of *Lactobacillus gasseri* ATCC 33323 reveals the molecular basis of an autochthonous intestinal organism. *Appl Environ Microbiol* 74: 4610–4625.

149. Maze A, *et al* (2010) Complete genome sequence of the probiotic *Lactobacillus casei* strain BL23. *J Bacteriol* 192: 2647–2648.

150. Ojala T, *et al* (2010) Genome sequence of *Lactobacillus crispatus* ST1. *J Bacteriol* 192: 3547–3548.

151. Ringel-Kulka T, *et al* (2011) Probiotic bacteria *lactobacillus acidophilus* NCFM and *Bifidobacterium lactis* Bi-07 versus placebo for the symptoms of bloating in patients with functional bowel disorders: A double-blind study. *J Clin Gastroenterol* 45: 518–525.

152. de Vrese M, Kristen H, Rautenberg P, Laue C & Schrezenmeir J (2011) Probiotic lactobacilli and bifidobacteria in a fermented milk product with added fruit preparation reduce antibiotic associated diarrhea and *Helicobacter pylori* activity. *J Dairy Res* 78: 396–403.

153. Pfeiler EA, Azcarate-Peril MA & Klaenhammer TR (2007) Characterization of a novel bile-inducible operon encoding a two-component regulatory system in *Lactobacillus acidophilus*. *J Bacteriol* 189: 4624–4634.

154. Goh YJ & Klaenhammer TR (2010) Functional roles of aggregation-promoting-like factor in stress tolerance and adherence of *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol* 76: 5005–5012.

155. Buck BL, Altermann E, Svingerud T & Klaenhammer TR (2005) Functional analysis of putative adhesion factors in *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol* 71: 8344–8351.

156. O'Flaherty SJ & Klaenhammer TR (2010) Functional and phenotypic characterization of a protein from *Lactobacillus acidophilus* involved in cell morphology, stress tolerance and adherence to intestinal cells. *Microbiology* 156: 3360–3367.

157. Majumder A, *et al* (2011) Proteome reference map of *Lactobacillus acidophilus* NCFM and quantitative proteomics towards understanding the prebiotic action of lactitol. *Proteomics* 11: 3470–3481.

158. Francl AL, Thongaram T & Miller MJ (2010) The PTS transporters of *Lactobacillus gasseri* ATCC 33323. *BMC Microbiol* 10: 77.

159. Zhang Z, Ye Z, Yu L & Shi P (2011) Phylogenomic reconstruction of lactic acid bacteria: An update. *BMC Evolutionary Biology* 11: 1.

160. Klijn A, Mercenier A & Arigoni F (2005) Lessons from the genomes of bifidobacteria. *FEMS Microbiol Rev* 29: 491–509.

161. Ventura M, van Sinderen D, Fitzgerald GF & Zink R (2004) Insights into the taxonomy, genetics and physiology of bifidobacteria. *Antonie Van Leeuwenhoek* 86: 205–223.

162. Turroni F, van Sinderen D & Ventura M (2011) Genomics and ecological overview of the genus *Bifidobacterium*. *Int J Food Microbiol* 149: 37–44.

163. Turroni F, *et al* (2009) Exploring the diversity of the bifidobacterial population in the human intestinal tract. *Appl Environ Microbiol* 75: 1534-1545.

164. Sanders ME (2006) Summary of probiotic activities of *Bifidobacterium lactis* HN019. *J Clin Gastroenterol* 40: 776–783.

165. Ventura M, *et al* (2007) Genomics of actinobacteria: Tracing the evolutionary history of an ancient phylum. *Microbiol Mol Biol Rev* 71: 495–548.

166. Bottacini F, *et al* (2010) Comparative genomics of the genus *Bifidobacterium*. *Microbiology* 156: 3243–3254.

167. Barrangou R, *et al* (2009) Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and Bl-04. *J Bacteriol* 191: 4144–4151.

168. Sela DA, *et al* (2008) The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. *Proc Natl Acad Sci USA* 105: 18964–18969.

169. Schell M, *et al* (2002) The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc Natl Acad Sci USA* 99: 14422–14427.

170. Turroni F, *et al* (2010) Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-derived glycan foraging. *Proc Natl Acad Sci USA* 107: 19514–19519.

171. O'Connell Motherway M, *et al* (2011) Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. *Proc Natl Acad Sci USA* 108: 11217–11222.

172. Locascio RG, Desai P, Sela DA, Weimer B & Mills DA (2010) Comparative genomic hybridization of *Bifidobacterium longum* strains reveals broad conservation of milk utilization genes in subsp. *infantis*. *Appl Environ Microbiol* 76: 7373–7381.

173. Sela DA, *et al* (2008) The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. *Proc Natl Acad Sci USA* 105: 18964–18969.

174. Ventura M, *et al* (2009) The *Bifidobacterium dentium* Bd1 genome sequence reflects its genetic adaptation to the human oral cavity. *PLoS Genetics* 5: e1000785.

175. Paineau D, *et al* (2008) Effects of seven potential probiotic strains on specific immune responses in healthy adults: A double-blind, randomized, controlled trial. *FEMS Immunol Med Microbiol* 53: 107–113.

176. Ouwehand AC, *et al* (2009) Specific probiotics alleviate allergic rhinitis during the birch pollen season. *World J Gastroenterol* 15: 3261–3268.

177. Ishizuka A, *et al* (2012) Effects of administration of *Bifidobacterium animalis* subsp. *lactis* GCL2505 on defecation frequency and bifidobacterial microbiota composition in humans. *J Biosci Bioeng* 113: 587–591.

178. Million M, *et al* (2011) Obesity-associated gut microbiota is enriched in *Lactobacillus reuteri* and depleted in *Bifidobacterium animalis* and *Methanobrevibacter smithii*. *International Journal of Obesity*: doi: 10.1038/ijo.2011.153.

179. Szajewska H, Guandalini S, Morelli L, Van Goudoever JB & Walker A (2010) Effect of *Bifidobacterium animalis* subsp. *lactis* supplementation in preterm infants: A systematic review of randomized controlled trials. *J Pediatr Gastroenterol Nutr* 51: 203–209.

180. Gueimonde M, *et al* (2004) Viability and diversity of probiotic *Lactobacillus* and *Bifidobacterium* populations included in commercial fermented milks. *Food Res Int* 37: 839–850.

181. Masco L, Ventura M, Zink R, Huys G & Swings J (2004) Polyphasic taxonomic analysis of *Bifidobacterium animalis* and *Bifidobacterium lactis* reveals relatedness at the subspecies level: Reclassification of *Bifidobacterium animalis* as *Bifidobacterium animalis* subsp. *animalis* subsp. nov. and *Bifidobacterium* lactis as *Bifidobacterium animalis* subsp. *lactis* subsp. nov. *Int J Syst Evol Microbiol* 54: 1137–1143.

182. Gilad O, *et al* (2012) Insights into physiological traits of *Bifidobacterium animalis* subsp. *lactis* BB-12 through membrane proteome analysis. *Journal of Proteomics* 75: 1190–1200.

183. Briczinski EP, *et al* (2009) Strain-specific genotyping of *Bifidobacterium animalis* subsp. *lactis* by using single-nucleotide polymorphisms, insertions, and deletions. *Appl Environ Microbiol* 75: 7501-7508.

184. Oberg TS, Ward RE, Steele JL & Broadbent JR (2012) Identification of plasmalogens in the cytoplasmic membrane of *Bifidobacterium animalis* subsp. *lactis*. *Appl Environ Microbiol* 78: 880–884.

185. Candela M, *et al* (2010) DnaK from *Bifidobacterium animalis* subsp. *lactis* is a surface-exposed human plasminogen receptor upregulated in response to bile salts. *Microbiology* 156: 1609–1618.

186. Gilad O, *et al* (2010) Combined transcriptome and proteome analysis of *Bifidobacterium animalis* subsp. *lactis* BB-12 grown on xylo-oligosaccharides and a model of their utilization. *Appl Environ Microbiol* 76: 7285–7291.

187. Ruiz L, Sánchez B, Ruas-Madiedo P, de Los Reyes-Gavilán C,G. & Margolles A (2007) Cell envelope changes in *Bifidobacterium animalis* ssp. *lactis* as a response to bile. *FEMS Microbiol Lett* 274: 316–322.

188. Garrigues C, Johansen E & Pedersen MB (2010) Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BB-12, a widely consumed probiotic strain. *J Bacteriol* 192: 2467–2468.

189. Chervaux C, *et al* (2011) Genome sequence of the probiotic strain *Bifidobacterium animalis* subsp. *lactis* CNCM I-2494. *J Bacteriol* 193: 5560–5561.

190. Bottacini F, *et al* (2011) Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BLC1. *J Bacteriol* 193: 6387–6388.

191. Flint HJ, Bayer EA, Rincon MT, Lamed R & White BA (2008) Polysaccharide utilization by gut bacteria: Potential for new insights from genomic analysis. *Nat Rev Microbiol* 6: 121–131.

192. Martens EC, Chiang HC & Gordon JI (2008) Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host & Microbe* 4: 447–457.

193. Flint HJ & Bayer EA (2008) Plant cell wall breakdown by anaerobic microorganisms from the mammalian digestive tract. *Ann N Y Acad Sci* 1125: 280–288.

194. Scott KP, *et al* (2011) Substrate-driven gene expression in *Roseburia inulinivorans*: Importance of inducible enzymes in the utilization of inulin and starch. *Proc Natl Acad Sci USA* 108: 4672–4679.

195. Dodd D, Moon YH, Swaminathan K, Mackie RI & Cann IK (2010) Transcriptomic analyses of xylan degradation by *Prevotella bryantii* and insights into energy acquisition by xylanolytic bacteroidetes. *J Biol Chem* 285: 30261–30273.

196. Martens EC, Koropatkin NM, Smith TJ & Gordon JI (2009) Complex glycan catabolism by the human gut microbiota: The bacteroidetes *sus*-like paradigm. *The Journal of Biological Chemistry* 284: 24673–24677.

197. Snel B, Bork P & Huynen MA (2002) The identification of functional modules from the genomic association of genes. *Proc Natl Acad Sci USA* 99: 5890–5895.

198. Parche S, *et al* (2007) Sugar transport systems of *Bifidobacterium longum* NCC2705. *J Mol Microbiol Biotechnol* 12: 9–19.

199. Monedero V, Yebra MJ, Poncet S & Deutscher J (2008) Maltose transport in *Lactobacillus casei* and its regulation by inducer exclusion. *Res Microbiol* 159: 94–102.

200. Gopal S, *et al* (2010) Maltose and maltodextrin utilization by *Listeria monocytogenes* depend on an inducible ABC transporter which is repressed by glucose. *PloS ONE* 5: e10349.

201. Flint HJ, Duncan SH, Scott KP & Louis P (2007) Interactions and competition within the microbial community of the human colon: Links between diet and health. *Environ Microbiol* 9: 1101–1111.

202. van den Broek LA, Hinz SW, Beldman G, Vincken JP & Voragen AG (2008) *Bifidobacterium* carbohydrases-their role in breakdown and synthesis of (potential) prebiotics. *Mol Nutr Food Res* 52: 146–163.

203. Lee J & O'Sullivan D,J. (2010) Genomic insights into bifidobacteria. *Microbiology and Molecular Biology Reviews* 74: 378–416.

204. Saier MH, Tran CV & Barabote RD (2006) TCDB: The Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Res* 34: 181–186.

205. Saier MH, Yen MR, Noto K, Tamang DG & Elkan C (2009) The Transporter Classification Database: Recent advances. *Nucleic Acids Res* 37: 274–278.

206. Davidson AL, Dassa E, Orelle C & Chen J (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev* 72: 317–364.

207. Eitinger T, Rodionov Da, Grote M & Schneider E (2011) Canonical and ECF-type ATP-binding cassette importers in prokaryotes: Diversity in modular organization and cellular functions. *FEMS Microbiol Rev* 35: 3–67.

208. Oldham ML, Davidson AL & Chen J (2008) Structural insights into ABC transporter mechanism. *Curr Opin Struct Biol* 18: 726–733.

209. Oldham ML, Khare D, Quiocho Fa, Davidson AL & Chen J (2007) Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* 450: 515–521.

210. Berntsson RP, Smits SHJ, Schmitt L, Slotboom D & Poolman B (2010) A structural classification of substrate-binding proteins. *FEBS Lett* 584: 2606–2617.

211. Garrido D, Kim JH, German JB, Raybould HE & Mills DA (2011) Oligosaccharide binding proteins from *Bifidobacterium longum* subsp. *infantis* reveal a preference for host glycans. *PLoS One* 6: e17315.

212. Silva Z, *et al* (2005) The high-affinity maltose/trehalose ABC transporter in the extremely thermophilic bacterium *Thermus thermophilus* HB27 also recognizes sucrose and palatinose. *J Bacteriol* 187: 1210–1218.

213. Abbott DW, *et al* (2010) The molecular basis of glycogen breakdown and transport in *Streptococcus pneumoniae*. *Mol Microbiol* 77: 183–199.

214. Higgins Ma, Abbott DW, Boulanger MJ & Boraston AB (2009) Blood group antigen recognition by a solute-binding protein from a serotype 3 strain of *Streptococcus pneumoniae*. *J Mol Biol* 388: 299–309.

215. Suzuki R, *et al* (2008) Structural and thermodynamic analyses of solute-binding protein from *Bifidobacterium longum* specific for core 1 disaccharide and lacto-N-biose I. *J Biol Chem* 283: 13165–13173.

216. Licht A, *et al* (2011) Crystal structures of the bacterial solute receptor AcbH displaying an exclusive substrate preference for β-D-galactopyranose. *J Mol Biol* 406: 92–105.

217. Walker IH, Hsieh PC & Riggs PD (2010) Mutations in maltose-binding protein that alter affinity and solubility properties. *Appl Microbiol Biotechnol* 88: 187–197.

218. Paulsen IT, Nguyen L, Sliwinski MK, Rabus R & Saier MH,Jr (2000) Microbial genome analyses: Comparative transport capabilities in eighteen prokaryotes. *J Mol Biol* 301: 75–100.

219. Deutscher J, Francke C & Postma PW (2006) How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev* 70: 939–1031.

220. Cao Y, *et al* (2011) Crystal structure of a phosphorylation-coupled saccharide transporter. *Nature* 473: 50–54.

221. Purhonen P, Lundbäck A, Lemonnier R, Leblanc G & Hebert H (2005) Three-dimensional structure of the sugar symporter melibiose permease from cryo-electron microscopy. *J Struct Biol* 152: 76–83.

222. Juergensen K, *et al* (2003) The companion cell-specific Arabidopsis disaccharide carrier AtSUC2 is expressed in nematode-induced syncytia. *Plant Physiol* 131: 61–69.

223. Guan L & Kaback HR (2006) Lessons from lactose permease. *Annu Rev Biophys Biomol Struct* 35: 67–91.

224. Vaughan EE, David S & de Vos WM (1996) The lactose transporter in *Leuconostoc lactis* is a new member of the LacS subfamily of galactoside-pentose-hexuronide translocators. *Appl Environ Microbiol* 62: 1574–1582.

225. Law CJ, Maloney PC & Wang DN (2008) Ins and outs of major facilitator superfamily antiporters. *Annu Rev Microbiol* 62: 289–305.

226. Kaback HR, Smirnova I, Kasho V, Nie Y & Zhou Y (2011) The alternating access transport mechanism in LacY. *J Membr Biol* 239: 85–93.

227. Pendse PY, Brooks BR & Klauda JB (2010) Probing the periplasmic-open state of lactose permease in response to sugar binding and proton translocation. *J Mol Biol* 404: 506–521.

228. Garcia-Celma J, Smirnova IN, Kaback HR & Fendler K (2009) Electrophysiological characterization of LacY. *Proc Natl Acad Sci USA* 106: 7373–7378.

229. Geertsma ER, Duurkens RH & Poolman B (2005) Functional interactions between the subunits of the lactose transporter from *Streptococcus thermophilus*. *J Mol Biol* 350: 102–111.

230. Veenhoff LM, Heuberger EH & Poolman B (2001) The lactose transport protein is a cooperative dimer with two sugar translocation pathways. *EMBO J* 20: 3056–3062.

231. Francl A, Hoeflinger JL & Miller MJ (2012) Identification of lactose phosphotransferase systems in *Lactobacillus gasseri* ATCC 33323 required for lactose utilization. *Microbiology* 158: 944–952.

232. Gosalbes MJ, *et al* (2011) Metatranscriptomic approach to analyze the functional human gut microbiota. *PloS One* 6: e17447.

233. Cantarel BL, *et al* (2009) The carbohydrate-active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res* 37: 233–238.

234. Lozupone CA, *et al* (2008) The convergence of carbohydrate active gene repertoires in human gut microbes. *Proc Natl Acad Sci USA* 105: 15076–15081.

235. Tasse L, *et al* (2010) Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome Res* 20: 1605–1612.

236. Na X & Kelly C (2011) Probiotics in *Clostridium difficile* infection. *J Clin Gastroenterol* 45: S154–S158.

237. Cossart P (2011) Illuminating the landscape of host-pathogen interactions with the bacterium *Listeria monocytogenes*. *Proc Natl Acad Sci USA* 108: 19484–19491.

238. Srikanth CV, Mercado-Lubo R, Hallstrom K & McCormick BA (2011) *Salmonella* effector proteins and host-cell responses. *Cell Mol Life Sci* 68: 3687–3697.

239. Marteau P & Chaput U (2011) Bacteria as trigger for chronic gastrointestinal disorders. *Dig Dis* 29: 166–171.

240. Yoshida E, *et al* (2012) *Bifidobacterium longum* subsp. *infantis* uses two different β-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides. *Glycobiology* 22: 361–368.

241. Iqbal S, Nguyen T, Nguyen TT, Maischberger T & Haltrich D (2010) β-galactosidase from *Lactobacillus plantarum* WCFS1: Biochemical characterization and formation of prebiotic galacto-oligosaccharides. *Carbohydr Res* 345: 1408–1416.

242. Goh YJ, Lee JH & Hutkins RW (2007) Functional analysis of the fructooligosaccharide utilization operon in *Lactobacillus paracasei* 1195. *Appl Environ Microbiol* 73: 5716–5724.

243. Lagaert S, *et al* (2010) Substrate specificity of three recombinant α-L-arabinofuranosidases from *Bifidobacterium adolescentis* and their divergent action on arabinoxylan and arabinoxylan oligosaccharides. *Biochem Biophys Res Commun* 402: 644–650.

244. van den Broek LA, *et al* (2005) Cloning and characterization of arabinoxylan arabinofuranohydrolase-D3 (AXHd3) from *Bifidobacterium adolescentis* DSM20083. *Appl Microbiol Biotechnol* 67: 641–647.

245. Mertz B, Kuczenski RS, Larsen RT, Hill AD & Reilly PJ (2005) Phylogenetic analysis of family 6 glycoside hydrolases. *Biopolymers* 79: 197–206.

246. Stam MR, Danchin EGJ, Rancurel C, Coutinho PM & Henrissat B (2006) Dividing the large glycoside hydrolase family 13 into subfamilies: Towards improved functional annotations of α-amylase-related proteins. *Protein Engineering, Design & Selection* 19: 555–562.

247. Markovic O & Janecek S (2001) Pectin degrading glycoside hydrolases of family 28: Sequence-structural features, specificities and evolution. *Protein Eng* 14: 615–631.

248. Fredslund F, *et al* (2011) Crystal structure of α-galactosidase from *Lactobacillus acidophilus* NCFM: Insight into tetramer formation and substrate binding. *J Mol Biol* 412: 466–480.

249. Zona R, Chang-Pi-Hin F, O'Donohue MJ & Janecek S (2004) Bioinformatics of the glycoside hydrolase family 57 and identification of catalytic residues in amylopullulanase from *Thermococcus hydrothermalis*. *Eur J Biochem* 271: 2863–2872.

250. Pokusaeva K, Connell-motherway M, Zomer A, Fitzgerald GF & Sinderen DV (2011) Cellodextrin utilization by *Bifidobacterium breve* UCC2003. *Appl Environ Microbiol* 77: 1681–1690.

251. Callanan M, *et al* (2008) Genome sequence of *Lactobacillus helveticus*, an organism distinguished by selective gene loss and insertion sequence element expansion. *J Bacteriol* 190: 727–735.

252. Gopal PK, Sullivan PA & Smart JB (2001) Utilisation of galacto-oligosaccharides as selective substrates for growth by lactic acid bacteria including *Bifidobacterium lactis* DR10 and *Lactobacillus rhamnosus* DR20. *Int Dairy J* 11: 19–25.

253. Ignatova T, *et al* (2009) Effect of oligosaccharides on the growth of *Lactobacillus delbrueckii* subsp. *bulgaricus* strains isolated from dairy products. *J Agric Food Chem* 57: 9496–9502.

254. Marcobal A, *et al* (2010) Consumption of human milk oligosaccharides by gut-related microbes. *J Agric Food Chem* 58: 5334–5340.

255. Schwab C & Ganzle M (2011) Lactic acid bacteria fermentation of human milk oligosaccharide components, human milk oligosaccharides and galactooligosaccharides. *FEMS Microbiol Lett* 315: 141–148.

256. Wasinger V (2006) Holistic biology of microorganisms: Genomics, transcriptomics, and proteomics. *Methods Biochem Anal* 49: 3–14.

257. Saulnier DM, Molenaar D, de Vos WM, Gibson GR & Kolida S (2007) Identification of prebiotic fructooligosaccharide metabolism in *Lactobacillus plantarum* WCFS1 through microarrays. *Appl Environ Microbiol* 73: 1753–1765.

258. Gonzalez R, Klaassens ES, Malinen E, de Vos WM & Vaughan EE (2008) Differential transcriptional response of *Bifidobacterium longum* to human milk, formula milk, and galactooligosaccharide. *Appl Environ Microbiol* 74: 4686–4694.

259. King HC & Sinha AA (2001) Gene expression profile analysis by DNA microarrays: Promise and pitfalls. *JAMA* 286: 2280–2288.

260. Pfeiler EA & Klaenhammer TR (2009) Role of transporter proteins in bile tolerance of *Lactobacillus acidophilus*. *Appl Environ Microbiol* 75: 6013–6016.

261. Goh YJ, *et al* (2009) Development and application of a *upp*-based counterselective gene replacement system for the study of the S-layer protein SlpX of *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol* 75: 3093–3105.

262. Petersen TN, Brunak S, von Heijne G & Nielsen H (2011) SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat Methods* 8: 785–786.

263. Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410.

264. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J & Sayers EW (2011) GenBank. *Nucleic Acids Res* 39: 32–37.

265. Larkin MA, *et al* (2007) Clustal W and clustal X version 2.0. *Bioinformatics* 23: 2947–2948.

266. Huson DH, *et al* (2007) Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 8: 460.

267. Churchill GA (2002) Fundamentals of experimental design for cDNA microarrays. *Nat Genet* 32: 490–495.

268. Knapen D, Vergauwen L, Laukens K & Blust R (2009) Best practices for hybridization design in two-colour microarray analysis. *Trends Biotechnol* 27: 406–414.

269. Law J, *et al* (1995) A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *J Bacteriol* 177: 7011–7018.

270. Russell WM & Klaenhammer TR (2001) Efficient system for directed integration into the *Lactobacillus acidophilus* and *Lactobacillus gasseri* chromosomes via homologous recombination. *Appl Environ Microbiol* 67: 4361–4364.

271. Studts JM & Fox BG (1999) Application of fed-batch fermentation to the preparation of isotopically labeled or selenomethionyl-labeled proteins. *Protein Expr Purif* 16: 109–119.

272. Kanehisa M, *et al* (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 36: 480–484.

273. Magrane M & Consortium U (2011) UniProt knowledgebase: A hub of integrated protein data. *Database (Oxford)* 29: bar009.

274. Nakai H, *et al* (2009) The maltodextrin transport system and metabolism in *Lactobacillus acidophilus* NCFM and production of novel α-glucosides through reverse phosphorolysis by maltose phosphorylase. *FEBS J* 276: 7353–7365.

275. Adachi W, *et al* (2004) Crystal structure of family GH-8 chitosanase with subclass II specificity from *Bacillus* sp. K17. *J Mol Biol* 343: 785–795.

276. Ajdic D & Pham VT (2007) Global transcriptional analysis of *Streptococcus mutans* sugar transporters using microarrays. *J Bacteriol* 189: 5049–5059.

277. Ehrmann Ma, Korakli M & Vogel RF (2003) Identification of the gene for β-fructofuranosidase of *Bifidobacterium lactis* DSM10140(T) and characterization of the enzyme expressed in *Escherichia coli*. *Curr Microbiol* 46: 391–397.

278. Flint HJ, O'Toole P,W. & Walker AW (2010) Special issue: The human intestinal microbiota. *Microbiology* 156: 3203–3204.

279. Marana SR (2006) Molecular basis of substrate specificity in family 1 glycoside hydrolases. *IUBMB Life* 58: 63–73.

280. Marion C, Aten AE, Woodiga Sa & King SJ (2011) Identification of an ATPase, MsmK, which energizes multiple carbohydrate ABC transporters in *Streptococcus pneumoniae*. *Infect Immun* 79: 4193–4200.

281. Schanert S, *et al* (2006) Maltose and maltodextrin utilization by *Bacillus subtilis*. *J Bacteriol* 188: 3911–3922.

282. Nataf Y, *et al* (2009) Cellodextrin and laminaribiose ABC transporters in *Clostridium thermocellum*. *J Bacteriol* 191: 203–209.

283. Shulami S, et al (2011) The L-arabinan utilization system of *Geobacillus stearothermophilus*. *J Bacteriol* 193: 2838-50.

284. Tsujibo H, Kosaka M, Ikenishi S, Sato T & Miyamoto K (2004) Molecular characterization of a high-affinity xylobiose transporter of *Streptomyces thermoviolaceus* OPC-520 and its transcriptional regulation. *J Bacteriol*186: 1029–1037.

285. Wei X, *et al* (2012) Fructose uptake in *Bifidobacterium longum* NCC2705 is mediated by an ATP-binding cassette transporter. *The Journal of Biological Chemistry* 287: 357–367.

286. Bordignon E, Grote M & Schneider E (2010) The maltose ATP-binding cassette transporter in the 21st century - towards a structural dynamic perspective on its mode of action. *Mol Microbiol* 77: 1354–1366.

287. Xiao JZ, *et al* (2010) Distribution of *in vitro* fermentation ability of lacto-*N*-biose I, a major building block of human milk oligosaccharides, in bifidobacterial strains. *Appl Environ Microbiol* 76: 54–59.

288. Human Microbiome Jumpstart Reference Strains Consortium, *et al* (2010) A catalog of reference genomes from the human microbiome. *Science* 328: 994–999.

289. Webb AJ, Homer Ka & Hosie AHF (2008) Two closely related ABC transporters in *Streptococcus mutans* are involved in disaccharide and/or oligosaccharide uptake. *J Bacteriol* 190: 168–178.

290. O'Sullivan O, *et al* (2009) Comparative genomics of lactic acid bacteria reveals a niche-specific gene set. *BMC Microbiol* 9: 50.

291. Pikis A, Hess S, Arnold I, Erni B & Thompson J (2006) Genetic requirements for growth of *Escherichia coli* K12 on methyl-α-D-glucopyranoside and the five α-D-glucosyl-D-fructose isomers of sucrose. *The Journal of Biological Chemistry* 281: 17900–17908.

292. Thompson J, Hess S & Pikis A (2004) Genes *malh* and *pagl* of *Clostridium acetobutylicum* ATCC 824 encode NAD^+- and Mn^{2+}-dependent phospho-α-glucosidase(s). *The Journal of Biological Chemistry* 279: 1553–1561.

293. Pikis A, Immel S, Robrish S & Thompson J (2002) Metabolism of sucrose and its five isomers by *Fusobacterium mortiferum*. *Microbiology* 148: 843–852.

294. Mathew S & Abraham TE (2004) Ferulic acid: An antioxidant found naturally in plant cell walls and feruloyl esterases involved in its release and their applications. *Crit Rev Biotechnol* 24: 59–83.

295. Crepin VF, Faulds CB & Connerton IF (2004) Functional classification of the microbial feruloyl esterases. *Appl Microbiol Biotechnol* 63: 647–652.

296. Zakharkin SO, Kim K, Bartolucci AA, Page GP & Allison DB (2006) Optimal allocation of replicates for measurement evaluation studies. *Genomics Proteomics Bioinformatics* 4: 196–202.

297. Katagiri F & Glazebrook J (2009) Overview of mRNA expression profiling using DNA microarrays. *Curr Protoc Mol Biol* Chapter 22: 4.

298. Denou E, *et al* (2008) Identification of genes associated with the long-gut-persistence phenotype of the probiotic *Lactobacillus johnsonii* strain NCC533 using a combination of genomics and transcriptome analysis. *J Bacteriol* 190: 3161–3168.

6 Appendices

6.1 Transcriptional analysis of prebiotic uptake and catabolism by *Lactobacillus acidophilus* NCFM

Currently in review at PloS ONE

Title:
Transcriptional analysis of prebiotic uptake and catabolism by *Lactobacillus acidophilus* NCFM

Authors:
Joakim Mark Andersen[1,2], Rodolphe Barrangou[3], Maher Abou Hachem[1], Sampo J. Lahtinen[4], Yong Jun Goh[2], Birte Svensson[1], Todd R. Klaenhammer[2]

Author affiliation:
[1] Enzyme and Protein Chemistry, Department of Systems Biology, Søltofts Plads Building 224, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark.
[2] Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State University, Box 7624, Raleigh, NC 27695.
[3] DuPont Nutrition and Health, 3329 Agriculture Drive, Madison, WI 53716.
[4] DuPont Nutrition and Health, Sokeritehtaantie 20, FI-02460 Kantvik, Finland.

Corresponding author:
Todd R. Klaenhammer

Mailing address:
Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State University,
Box 7624, Raleigh, NC 27695. Phone: (919) 515-2972. Fax: (919) 513-0014.
E-mail: klaenhammer@ncsu.edu.

Contributions to the publication:

Designed research: JMA, RB, MAH, SL, BS, TRK
Performed research: JMA, YJG
Contributed new reagents or analytic tools: SL, YJG, TRK
Analyzed data: JMA, RB, YJG, TRK
Wrote the paper: JMA, RB, MAH, SL, YJG, BS, TRK

Abstract

The human gastrointestinal tract can be positively modulated by dietary supplementation of probiotic bacteria in combination with prebiotic carbohydrates. Here differential transcriptomics and functional genomics were used to identify genes in *Lactobacillus acidophilus* NCFM involved in the uptake and catabolism of 11 potential prebiotic compounds consisting of α- and β- linked galactosides and glucosides. These oligosaccharides induced genes encoding phosphoenolpyruvate-dependent sugar phosphotransferase systems (PTS), galactoside pentose hexuronide (GPH) permease, and ATP-binding cassette (ABC) transporters. PTS systems were upregulated primarily by di- and tri-saccharides such as cellobiose, isomaltose, isomaltulose, panose and gentiobiose, while ABC transporters were upregulated by raffinose, Polydextrose, and stachyose. A single GPH transporter was induced by lactitol and galactooligosaccharides (GOS). The various transporters were associated with a number of glycoside hydrolases from families 1, 2, 4, 13, 32, 36, 42, and 65, involved in the catabolism of various α- and β-linked glucosides and galactosides. Further subfamily specialization was also observed for different PTS-associated GH1 6-phospho-β-glucosidases implicated in the catabolism of gentiobiose and cellobiose. These findings highlight the broad oligosaccharide metabolic repertoire of *L. acidophilus* NCFM and establish a platform for selection and screening of both probiotic bacteria and prebiotic compounds that may positively influence the gastrointestinal microbiota.

Introduction

The microbiota of the human gastrointestinal tract (GIT) can dramatically affect the immune system of the host through increased allergy resistance (1) and modulation of diabetes, obesity (2, 3) and autoimmune bowel disorders (4). The compositional balance and activity of the microbiota can be positively influenced by probiotic microorganisms (5), or shifted by prebiotic supplementation (6). One effective strategy to promote positive impacts on both commensal and probiotic microbes is GIT modulation with prebiotic substrates (7-9).

Prebiotics are complex carbohydrates that are not digested or absorbed by the host, but catabolized by various commensal and health-promoting members of the GIT bacteria and selectively promoting their growth (10). Currently, a few carbohydrates are widely accepted as prebiotics, specifically GOS (β-galactooligosaccharides), inulin, FOS (fructo-oligosaccharides) and lactulose (11). *In vivo* studies, however, have shown increases in the populations of probiotic microbes due to stimulation by candidate prebiotic carbohydrate compounds, e.g. panose (12), polydextrose (13) and lactitol (14). Advances in the genomics of lactobacilli and bifidobacteria have enabled modeling of transport and catabolic pathways for prebiotic utilization (15). Only a few such proposed models, however, have been experimentally validated (16-18), which hampers accurate functional assignment of novel specificities especially for carbohydrate transporters that are largely uncharacterized biochemically. Recent studies have shown transfer of genes enabling prebiotic catabolism in certain pathogenic strains (19) and growth on prebiotic substrates in mono-cultures of some GIT commensal and pathogenic bacteria (20). These findings emphasize the need to provide functional scientific support for novel prebiotic candidates and to address the molecular basis for selective prebiotic catabolism by probiotic microbes.

The probiotic microbe *Lactobacillus acidophilus* NCFM has been investigated by in-depth functional studies to reveal the molecular mechanisms for important probiotic traits, such as bile acid resistance (21), involvement of lipoteichoic acid in immunomodulation (22), and positive outcomes reported in human intervention studies using *L. acidophilus* NCFM as a probiotic (23, 24) and when supplemented as a synbiotic (25). The potential of *L. acidophilus* NCFM to metabolize a diverse number of oligosaccharides is reflected by the large number of predicted glycoside hydrolases encoded by its genome (26), and by functional studies outlining routes for utilization of various oligosaccharides saccharides (16, 27, 28). Accurate annotation of genes involved in prebiotic utilization is hampered by the paucity of functional studies,

especially of transporters and families of glycoside hydrolases that exhibit a multitude of substrate specificities. The scope of this study was to transcriptionally identify and functionally characterize genomic loci encoding catabolic pathways in *L. acidophilus* NCFM essential for the transport and utilization of a range of potential prebiotics spanning hexose families of α- and β- linked glucosides and galactosides.

Results

Carbohydrate dependent differentially expressed gene clusters

Gene expression was measured in *L. acidophilus* NCFM harvested in the early exponential phase and stimulated by glucose compared to 11 different oligosaccharides (Table 1), representing different hexoses in varying groups of carbohydrate linkages. These groups contained the α-galactosides consisting of, raffinose and stachyose; the α-glucosides, isomaltose, isomaltulose, panose and polydextrose; the β-galactosides, lactitol and GOS; and the β-glucosides, β-glucan oligomers, cellobiose and gentiobiose. The overall gene expression pattern for growth on each carbohydrate was represented by cluster analysis (published online, Figure S1). The most extensive differential gene expression was observed for specific gene clusters, while the overall gene expression pattern remained essentially unchanged, thus indicating that *L. acidophilus* NCFM adaptation to complex carbohydrate metabolism is regulated at the transcriptional level.

Statistical analysis of the global gene expression data was performed by a mixed model ANOVA to identify differentially expressed genes to each oligosaccharide treatment. A range of 1 – 45 genes were statistically differentially expressed (threshold $p = 10^{-4,74}$ for $\alpha=0.05$ using Bonferroni correction) for all treatments. The results of differential gene expression and statistical significance were illustrated by volcano plots that highlighted upregulated genes predicted to be involved in oligosaccharide transport and catabolism (Figure 1) summarized in Table 2 and with a heat map representation of expression of all the identified genes (Figure S2). None of the genes predicted to be involved in oligosaccharide catabolism were upregulated by growth on glucose, consistent with previous findings that glucose is transported by a constitutively expressed phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS)

(LBA0452, LBA0455–LBA0457) (27). Analysis of gene induction patterns by specific oligosaccharides showed a differential expression profile of carbohydrate active proteins (Table 2) depending on the carbohydrate linkages (α- *vs*. β-glycosidic linkages) and the monosaccharide constituents of glucoside and galactoside.

β-galactoside differentially induced genes

In the presence of GOS and lactitol, several genes (Table 2) were upregulated (7.1 – 64 fold) within the locus LBA1457 – LBA1469 encompassing genes encoding a galactose-pentose-hexuronide (GPH) LacS permease and two β-galactosidases (GH2 and GH42; CAZy glycoside hydrolase family (GH) classification (29)) together with the Leloir pathway genes for galactoside metabolism. These data indicate how these oligosaccharides are transported by the LacS permease and hydrolyzed by the action of two different β-galactosidases into galactose, glucose in the case of GOS and galactose and glucitol for lactitol, which are shunted into the Leloir and glycolytic pathways, respectively, as reported previously (18, 28).

β-glucoside differentially induced genes

Cellobiose induced genes within two loci (LBA0724–LBA0726 and LBA0877–LBA0884; 6.1 – 65.8 fold upregulation), both encoding a PTS permease EIIABC and a putative GH1 6-phospho-β-glucosidase. Growth on gentiobiose as a carbon source upregulated (9.2 fold) the PTS permease EIIC (LBA0227), albeit at a lower level by panose (5.4 fold), indicating either a dual specificity of the PTS permease or a more complex transcriptional co-regulation of the transport system. The oligomers obtained by hydrolysis of mixed linkage β-1,3/β-1,4 β-glucan stimulated upregulation of both the cellobiose-induced PTS permease gene cluster mentioned above, and notably the α-glucoside induced gene cluster LBA0606–LBA0609 and LBA1684 (encoding a PTS EIIA component). The patterns of upregulated gene clusters for β-glucosides indicate differential recognition of the β-1,4 and β-1,6 linkages and the specialization of different PTS permeases and their corresponding GH1 enzymes that recognize phosphorylated β-glucosides at the C6 position.

α-glucoside differentially induced genes

Both isomaltose and isomaltulose upregulated the LBA0606-LBA0609 locus (13.4 – 65.8 fold), putatively encoding a PTS permease (EIIABC). This regulatory RpiR family protein and a hypothetical protein, together with LBA1684 (11.7 fold upregulated), annotated as a PTS IIA regulatory components. LBA1689 (65.9 fold upregulated) annotated as a GH4 maltose-6-phosphate glucosidase. This suggested that the two α-1,6 linked glucosides are phosphorylated concomitant with their transport by the PTS EIIABC (LBA0606 and LBA0609) permease, and that these phosphorylated disaccharides are hydrolyzed by a specific intracellular (predicted by SignalP (30)) GH4 disaccharide 6-phospho-α-glucosidase into glucose-6-phosphate and either glucose from isomaltose or fructose from isomaltulose, which enter glycolysis. Notably, the trisaccharide panose elicited a similar upregulation pattern as isomaltose, including upregulation of LBA0606-LBA0609 and LBA1689, and also up- LBA0227. The locus LBA0224-LBA0228 was annotated to include a cellobiose-specific PTS permease EIIC domain, a regulatory protein, and a GH1 6-phospho-β-glucosidase.

The diverse structural elements present in polydextrose constitute a complex oligosaccharide mixture of mostly different α-linked glucosides. Accordingly, a complex upregulation pattern was observed that involved genes encoding both an ABC (LBA500-0504) and a PTS permease (LBA0606) and several hydrolases (LBA0505, LBA1689 and LBA1870). The highest upregulation involved the above PTS permease (LBA0606-0609) together with LBA1870 encoding a GH65 maltose phosphorylase (31) and LBA0505-0506 identified as a part of a locus determined previously as a FOS metabolism operon (16).

α-galactoside differentially induced genes

The tetrasaccharide stachyose induced the gene locus LBA1438 – LBA1442 (9.3 – 35.9 fold upregulated) encoding an ABC transporter, a GH36 α-galactosidase and a part of the Leloir pathway enzymes (LBA1458, LBA1459 and LBA1469). This suggests that stachyose is transported into the cytoplasm by this ABC transporter and initially hydrolyzed into galactose and raffinose, which is further processed to galactose and sucrose that subsequently can be phosphorolyzed by LBA1437 encoding a sucrose phosphorylase (GH13_18). This gene cluster was previously found to be upregulated by raffinose (27). From the DNA microarray presented in the present study, no upregulated genes were involved with

oligosaccharide metabolism by stimulation of raffinose, suggesting glucose as an impurity in the medium or raffinose preparation.

Functional characterization of genes involved with α-galactoside metabolism

To corroborate the identification of gene clusters from *L. acidophilus* NCFM (25, 27, 32) involved in the metabolism of α-galactosides of the raffinose family oligosaccharides, two single gene deletions were constructed within the stachyose induced locus, i.e. ΔLBA1438 (α-galactosidase) and ΔLBA1442 (solute binding protein of the ABC transporter) using the *upp*-based counterselective gene replacement system (33). It was predicted by genome mining that *L. acidophilus* NCFM encoded single locus responsible for the transport and hydrolysis of α-galactosides. Phenotypic confirmation of the roles of these genes was accomplished by constructing mutations in these genes. Mutations of LBA1438 and LBA1442 were in-frame deletions of 92% and 91% of the coding regions, respectively. The α-galactosidase (LBA1438) deletion mutant lost the ability to grow on raffinose (Figure 2B), melibiose (α-D-Gal*p*-(1–6)-D-Glc*p*) and stachyose (data not shown). The ability of the LBA1442 mutant to grow on galactose (Figure 2A), but not raffinose (Figure 2C) provides evidence for the specificity of the transporter for α-galactoside oligosaccharides. The phenotypes of single gene deletion variants confirm that the genes identified through differential transcriptomics are functionally crucial for growth on these prebiotic compounds.

Structure, divergence and function of induced gene clusters

The transcriptional gene induction patterns and the essential roles of single proteins responsible for carbohydrate uptake and catabolism demonstrated how specific gene clusters conferred the ability to utilize the prebiotics investigated in this study. Identification of gene clusters selectively upregulated in response to prebiotic substrates (Figure 3) showed that multiple genes within these operons are typically expressed as single transcripts. However, genes LBA1684 (PTS EIIA component) and LBA1689 (putative maltose-6-phosphate-hydrolase), were predicted *in silico* to be monocistronically transcribed. All gene clusters induced by prebiotic substrates were analyzed for regulatory elements.

Catabolite repression elements (CRE) were found upstream of all non-PTS permease containing transcripts and LBA1684, encoding a PTS EIIA. . The molecular responses to oligosaccharide stimulation are likely mediated through CRE sites via catabolite control protein A (*ccpA*, LBA0431), phosphocarrier protein HPR (*ptsH*, LBA0639), and HPr kinase/phosphorylase (*ptsK* LBA0676) linking the regulation to the phosphorylation cascade of EI through EIIA to the PTS permeases (27).

Amino acid sequence comparisons to previously characterized bacterial PTS EIIC trans-membrane substrate binding-domains (Figure S3), including β-1,4 or β-1,6 glucoside specific PTS permeases, showed a clear segregation of LBA0227 and LBA0725, the latter clustering with a functionally characterized cellobiose PTS permease from *L. gasseri* ATCC 33323 (34), consistent with the observed upregulation of LBA0725 on cellobiose. Notably, the PTS permease EIIC domains LBA0879 and LBA0884 were also upregulated by cellobiose, albeit at a lower level than LBA0725. These two proteins clustered distantly on the phylogenetic tree, indicating functional divergence and a likely preference for structurally-related substrates such as sophorose (β-D-Glc*p*-(1–2)-D-Glc*p*), a candidate prebiotic supporting growth of *L. acidophilus* NCFM (12). A schematic overview (Figure 4) summarizes the uptake and catabolism pathways of potential prebiotic oligosaccharides in *L. acidophilus* NCFM.

Notably, the most highly induced gene in the present study was LBA0608 (Figure S4) encoded a hypothetical protein within a PTS permease locus. No function could be assigned for the protein, which was predicted to have a four transmembrane helical topology using the Phobius prediction tool (35). The same topology was found for LBA0878 (Figure 3D), another hypothetical protein encoded in locus with a PTS permease, but no significant amino acid sequence similarity was found for the two proteins.

Discussion

Carbohydrates supplemented for enrichment of specific commensal or probiotic microbes of the GIT can exert selective increases in certain beneficial populations , and decrease pathogens and symptoms of some GIT disorders. Recent studies of prebiotic catabolism (17, 34, 36) have shown a wide array of metabolic capabilities that cannot be deduced

based on *in silico* gene annotations or even on experimental work in homologous organisms. The pathways and the molecular elements for transport and catabolism of FOS, lactitol and GOS have been analyzed in *L. acidophilus* NCFM (16, 18, 28). This serves as a methodological basis to identify the molecular and genetic foundation for screening of potential prebiotic compounds *in vitro* and/or *in vivo* and specific enrichment of health-promoting bacteria in complex microbial ecosystems (37–39).

Importance of carbohydrate transporter variety

The general structure of the identified gene clusters indicates that typically, a three component system consisting of a regulator, transporter and glycoside hydrolase(s) can be sufficient for utilization of potential prebiotics, irrespective of the type of transporter identified (ABC, GPH, or PTS permease, Figure 3). Remarkably, PTS permeases had higher selectivity towards disaccharides, whereas ABC and GPH permeases appeared to be also induced by the longer oligosaccharides e.g. stachyose, and GOS. Furthermore, similar upregulation patterns of gene expression by widely different prebiotics was surprising, notably the FOS-ABC transporter that was also induced by the mixed linkage polydextrose. This suggests that transporters either possess more than one specificity or less strigent molecular recognition of substrates, indicating a wide range of carbohydrates can be metabolized by *L. acidophilus* NCFM, and likely similar commensal lactic acid bacteria. This capability is also expanded by transporters that possess a broad specificity for oligosaccharides sharing structural elements e.g. the α-1,2 glycosidic linkages found in both FOS and polydextrose.

Gene deletions confirm GOS and α-galactosides utilization

Functional corroboration of the specificity of prebiotic transport loci has been facilitated by their identification using differential transcriptomics. We previously confirmed that the GPH-type LacS permease is involved in uptake of β-galactosides, GOS and lactitol (28). Two associated β-galactosidases were involved (LBA1462, GH42, and LBA1467-68, GH2, Figure 3H). The differential expression levels (Figure 1 and Table 2), suggested that GH42 is the main hydrolase for GOS degradation in *L. acidophilus* NCFM. Gene deletions validated both uptake and catabolism for the α-galactosides

raffinose, stachyose and melibiose by the locus containing the ABC transport system and GH36 α-galactosidase (LBA1437-LBA1442, Figure 3G).

Distinct PTS systems and GH1 hydrolases mediate utilization of cellobiose and gentiobiose

Transcriptomics data suggested that the two β-glucoside disaccharide regio-isomers, cellobiose and gentiobiose, only differing in the glucosidic linkage are internalized by two different PTS systems and hydrolyzed by two different GH1 putative 6-phospho-β-glucosidases having 49% overall sequence identity (Figure 3A and 3C). To validate these findings, the sequences of the PTS transporters and the GH1 hydrolases were analyzed *in silico*. The phylogenetic tree constructed for the PTS systems showed clear segregation of the cellobiose and the gentiobiose induced PTS systems (Figure S3). Notably, the cellobiose induced PTS system clustered together with the functionally characterized cellobiose PTS transporter from *L. gasseri* 33323 that apparently lacks a homolog of the gentiobiose-induced PTS system. Currently there is no biochemical characterization of PTS systems with gentiobiose specificity. Similarly, the two GH1 6-phospho-β-glucosidases from *L. acidophilus* clustered in two distinct GH1 subgroups, whereas a third subgroup was represented by a biochemically and a structurally characterized cellobiose specific GH1 β-1,4-glucosidase (40) (Figure S5). The structure of the 6-phospho-β-glucosidase from *L. plantarum* (PDB: 3QOM, The Midwest Center for Structural Genomics), containing a phosphate ion bound in the active site, has the three conserved residues involved in the recognition of the phosphate moiety of phosphorylated disaccharide substrates (Figure S6A). This, together with sequence alignments (Figure S6A) suggests that the catalytic residues and the phosphate recognition pocket are conserved in LBA0225 and LBA0726, together with all amino acid residues defining the pivotal substrate binding subsite –1, where the non-reducing end 6-phospho-glucosyl residue is bound, are completely conserved (Figure S6B). This is consistent with both putative enzymes being catalytically competent and with their induction together with different PTS transporters, congruent with their recognition of non-reducing end phosphorylated substrates at the C-6 position. Clear differences, however, were observed in amino acid residues of LBA0225 and LBA0726 corresponding to those flanking the putative subsite +1 in the structure of the *L. plantarum* putative 6-phospho-β-glucosidase (Figure S6C), in accordance with the specificity differences suggested by the transcriptomics data. A combination of the GH1 structure-function relationship (Figure S5 and S6) and phylogenetic analysis of PTS permeases (Figure S3) corroborates the transcriptomics findings

implicating two β-glucoside isomers in the differential upregulation of the two loci (LBA0225-0228 and LBA0724-0726, Figure 3A and 3C). Improved annotations of both PTS permeases (34) and GH1 6-phospho-glycosidases have previously been limited due to difficulties in working with the transmembrane PTS permeases and the lack of phosphorylated substrates for GH1 or GH4 enzymes. In this light, both transcriptomics and site-specific gene deletions will serve as powerful tools for further functional characterization. The present data offer an important view of the metabolic diversity for *L. acidophilus* NCFM that is differentiated by the type of transporters, GH families and sub-specificities within single GH families.

Remarkably, the LBA0606-0609 locus encoding a PTS permease, was induced by isomaltose and panose revealing a novel pathway for the transport and hydrolysis of short isomaltooligosaccharides, emerging as potential prebiotics (41). *L. acidophilus* NCFM additionally encodes a canonical GH13 subfamily 31 (GH13_31) glucan-α-1,6-glucosidase homolog to an enzyme from *Streptococcus mutans* shown to be more active on isomaltooligosaccharides longer than isomaltose (42). However, this locus (LBA0264, GH13_31) was not significantly upregulated in the current study. It is possible that this latter enzyme is induced on longer isomaltooligosaccharides, which may be transported via a different route. Such size dependent differentiation of the utilization pathway has been reported for maltooligosaccharides in other Gram positive bacteria (43). Furthermore, the locus contained a putative protein with no predictable function (LBA0608), which was the highest induced gene of the study. Sequence analysis indicated a transmembrane topology potentially linking this gene product to the function of the PTS transporter encoded in the locus.

Comparative genomics of niche specific genes relating to prebiotic utilization

A previous comparative genomics approach predicted LBA1689 orthologs to be selectively found only in GIT associated lactobacilli (44). This would indicate that the identified novel isomaltose catabolism pathway utilizing the novel isomaltose-6-phosphate hydrolase LBA1689 to be a potential target for α-1,6-glucoside probiotics (e.g. panose and polydextrose) and complementing the conventional route of degradation mediated by the putative α-1,6 glucosidase (LBA0264) encoded in the genome of *L. acidophilus* NCFM.

The important potential for GIT adaption of *L. acidophilus* NCFM by genetic loci encoding specific oligosaccharide utilization is further emphasized from genomic comparisons to the phylogenetically related, but milk adapted *L. helveticus* DPC 4571 (45) where loci, identified in the current study, have been lost through evolution and adaption to milk fermentation for the following oligosaccharides: gentiobiose, FOS, raffinose, isomaltose and panose. These observations underscore how prebiotic stimulation can be considered as a species-specific attribute reflecting evolutionary adaptation to nutritionally rich environments, like the GIT, by either gene gain and functional diversification, or gene-loss associated genome simplification (15).

In conclusion, genes involved in the uptake and catabolism of prebiotic compounds by *L. acidophilus* NCFM were identified using differential transcriptomics. This study revealed the extensive ability of *L. acidophilus* NCFM to utilize a diversity of prebiotic compounds, employing a broad range of carbohydrate uptake systems, including ABC, GPH and PTS transporters, as well as an expansive repertoire of hydrolases that can readily catabolize α- and β-linked glucosides and galactosides.

Materials and methods

L. acidophilus NCFM mRNA sample preparation and DNA microarray platform

Whole genome oligonucleotide microarrays were designed as described by Goh et al. (33) with four replicate spots for each of the 1,823 predicted genes. Hybridization quality was assessed as described previously (28). For preparation of cultures for the DNA microarray transcriptome analysis, a semi-synthetic medium (SSM, (16) used for cultivation of *L. acidophilus* NCFM was filtered through a 0.22 µm filter and oxygen was removed by the Hungate method (46). *L. acidophilus* NCFM cultures were propagated in parallel in SSM media supplemented with 1% (w/v) of various carbohydrates as listed for structure and manufacturer in Table 1. Cultures were transferred for five passages on each carbohydrate before harvested at the early logarithmic phase (OD_{600}= 0.35–0.5) by pelleting at 4°C (3,000 x *g*, 15 min) and flash freezing the pellets for storage at -80°C.

Cells were mechanically disrupted by beadbeating and total RNA isolated using Trizol-chloroform extraction (Invitrogen, Carlsbad, CA). Genomic DNA was removed with Turbo DNAse (Ambion, Austin, TX), followed by RNA purification using a RNeasy Mini Kit (Qiagen Inc., Valencia, CA) (33).

Reverse transcription of total RNA, fluorescent labeling of cDNA and hybridizations were performed using 20 µg of total RNA for each replicate as described by Goh et al. (33). Total RNA from each carbohydrate treatment was labeled with both Cyanine3 and Cyanine5 for two technical dye-swapped replicates to each growth condition, and pairwise hybridized using a loop-design for a total of 12 hybridizations.

Hybridized chips were scanned at 10 µm resolution per pixel using a ScanArray Express microarray scanner (Packard BioScience, Meriden, CT) for 16-bit spot intensity quantification. Fluorescent intensities were quantified and background-subtracted using the QuantArray 3.0 software package (Packard Bioscience). Median values were calculated for all ORFs (Open Reading Frames) using tetraplicate intensities and log_2-transformed before imported into SAS JMP Genomics 4.0 (SAS Institute Inc, Cary, NC) for data analysis. The full

data set was interquantile normalized and modeled using a mixed model ANOVA for analysis of the differential gene expression pattern, and visualization using heat maps and volcano plots.

Bacterial strains and growth conditions

All bacterial strains and plasmids used throughout this study are listed in Table 3. *Lactobacillus* broth cultures were cultivated in MRS (Difco Laboratories Inc., Detroit, MI) or semi-defined medium (SDM) (47), supplemented with 0.5% (w/v) glucose (Sigma-Aldrich, St. Louis, MO)) or 1% (w/v) sucrose (Sigma), galactose (Sigma), melibiose, raffinose (BDH chemicals, Poole, England) and stachyose (Sigma) as carbon sources, in non-shaking batch cultures, aerobically at 37 °C or 42 °C. Chloramphenicol (Cm, 5 μg/ml) or/and erythromycin (Em, 2 μg/ml) were used when necessary for selection. *Escherichia coli* strains were cultivated in Brain Heart Infusion medium (Difco) aerobically at 37 °C with aeration, and Em (150 μg/ml) and/or kanamycin (Km, 40 μl/ml) were/was added for selection. Solid media were prepared by the addition of 1.5% (w/v) agar (Difco).

Construction and phenotypic determination of deletion mutants in the α-galactoside gene cluster

Genomic DNA of *L. acidophilus* NCFM was isolated by the method of Walker and Klaenhammer (48) or by the Mo Bio Ultraclean microbial DNA isolation kit (Mo Bio Laboratories, Carlsbad, CA). Plasmid DNA from *E. coli* was isolated using a QIAprep Spin miniprep kit (Qiagen). Restriction enzymes (Roche Molecular Biochemicals, Indianapolis, IN) were applied according to the instructions supplied by the manufacturer. DNA ligation was done using T4 DNA ligase (New England Biolabs, Beverly, MA) as directed by the manufacturers' recommendations. All PCR primers were synthesized by Integrated DNA Technologies (Coralville, IA). PCR reactions, preparation and transformation of competent *L. acidophilus* NCFM and *E. coli* cells, analysis by agarose gel electrophoresis, and in gel purification were done as described by Goh et al. (33).

The construction of a *Δupp* isogenic mutant with in-frame DNA excision of the LBA1438 and LBA1442 coding region was done according to Goh et al. (33). In short, the upstream and downstream flanking regions (approximate length of 750 bp each) of the deletion targets were PCR-amplified either with the 1438A/1438B and 1438C/1438D or 1442A/1442B and 1442C/1442D primer pairs, respectively, and fused by splicing by overlap extension PCR (SOE-PCR). The SOE-PCR products were ligated into pTRK935 linearized with compatible ends (BamHI and EcoRI for all constructs), and transformed into NCK1831. The resulting recombinant plasmids, pTRK1013 and pTRK1014, harbored in NCK2122 and NCK2124, were transformed into NCK1910 harboring pTRK669, for chromosomal integration and following DNA excision to generate the Δ*melA* or Δ*msmE* genotypes respectively. Confirmation of DNA deletion was done by PCR and DNA sequencing using primer pair 1438UP/1438DN and 1442UP/1442DN (see Table S1).

Carbohydrate utilization of the gene deletion mutants was tested by comparative growth to wild type *L. acidophilus* NCFM and NCK1909 (*upp* mutant and parent strain of the Δ*melA* and Δ$msmE_{II}$ mutants). All strains were grown in SDM supplemented with 1% (w/v) glucose before inoculation (1% (v/v)) of an overnight culture into SDM supplemented with 1 % (w/v) of the following carbohydrates in separate batches: raffinose, stachyose, sucrose and galactose. Growth was monitored by measuring optical density (OD_{600}) using a Fluostar spectrophotometer (BMG Labtech, Cary, NC)) in triplicate wells of a 96-well plate (200 µl per well) covered with an airtight seal.

Microarray Data Submission

All raw data have been deposited in the GEO database under accession GSE35968 (can be provided during review) and complies with the MIAME guidelines.

Acknowledgements:

We thank Evelyn Durmaz for technical assistance with sequence-based confirmation of the deleted regions of the LBA1438- and LBA1442-deficient *Lactobacillus acidophilus* NCFM mutants.

Funding will be added during the online submission:

This research was funded by DuPont Nutrition and Health, North Carolina Dairy Foundation and the FøSu grant from the Danish Strategic Research Council to the project "Gene discovery and molecular interactions in prebiotics/probiotics systems. Focus on carbohydrate prebiotics". J.M.A. is funded by a joint Ph.D. stipend from DuPont, the FøSu grant, and the Technical University of Denmark.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Reference list

1. Tang ML, Lahtinen SJ & Boyle RJ (2010) Probiotics and prebiotics: Clinical effects in allergic disease*. Curr Opin Pediatr* 22: 626−634.

2. Kootte RS*, et al* (2012) The therapeutic potential of manipulating gut microbiota in obesity and type 2 diabetes mellitus*. Diabetes Obes Metab* 14:112−120.

3. Scarpellini E*, et al* (2010) Gut microbiota and obesity*. Intern Emerg Med* 5 Suppl 1: S53−6.

4. Hedin C, Whelan K & Lindsay JO (2007) Evidence for the use of probiotics and prebiotics in inflammatory bowel disease: A review of clinical trials*. Proc Nutr Soc* 66: 307−315.

5. Reid G*, et al* (2003) New scientific paradigms for probiotics and prebiotics*. J Clin Gastroenterol* 37: 105−118.

6. Davis LM, Martinez I, Walter J, Goin C & Hutkins RW (2011) Barcoded pyrosequencing reveals that consumption of galactooligosaccharides results in a highly specific bifidogenic response in humans *. PLoS One*. 6(9): e25200. doi:10.1371/journal.pone.0025200.

7. Davis LM, Martinez I, Walter J & Hutkins R (2010) A dose dependent impact of prebiotic galactooligosaccharides on the intestinal microbiota of healthy adults*. Int J Food Microbiol* 144: 285−292.

8. Kukkonen K*, et al* (2007) Probiotics and prebiotic galacto-oligosaccharides in the prevention of allergic diseases: A randomized, double-blind, placebo-controlled trial*. J Allergy Clin Immunol* 119: 192−198.

9. Roberfroid M*, et al* (2010) Prebiotic effects: Metabolic and health benefits*. Br J Nutr* 104 Suppl 2: S1−S63.

10. Callaway TR*, et al* (2008) Probiotics, prebiotics and competitive exclusion for prophylaxis against bacterial disease*. Anim Health Res Rev* 9: 217−225.

11. Macfarlane GT, Steed H & Macfarlane S (2008) Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics*. J Appl Microbiol* 104: 305−344.

12. Makelainen H, Hasselwander O, Rautonen N & Ouwehand AC (2009) Panose, a new prebiotic candidate *. Lett Appl Microbiol* 49: 666−672.

13. Jie Z*, et al* (2000) Studies on the effects of polydextrose intake on physiologic functions in Chinese people*. Am J Clin Nutr* 72: 1503−1509.

14. Bjørklund M*, et al* (2011) Gut microbiota of healthy elderly NSAID users is selectively modified with the administration of *Lactobacillus acidophilus* NCFM and lactitol*. Age (Dordr)* doi: 10.1007/s11357-011-9294-5

15. Makarova K*, et al* (2006) Comparative genomics of the lactic acid bacteria*. Proc Natl Acad Sci U S A* 103: 15611−15616.

16. Barrangou R, Altermann E, Hutkins R, Cano R & Klaenhammer TR (2003) Functional and comparative genomic analyses of an operon involved in fructooligosaccharide utilization by *Lactobacillus acidophilus. Proc Natl Acad Sci U S A* 100: 8957−8962.

17. Goh YJ, Lee JH & Hutkins RW (2007) Functional analysis of the fructooligosaccharide utilization operon in *Lactobacillus paracasei* 1195 *. Appl Environ Microbiol* 73: 5716−5724.

18. Majumder A*, et al* (2011) Proteome reference map of *Lactobacillus acidophilus* NCFM and quantitative proteomics towards understanding the prebiotic action of lactitol*. Proteomics* 11: 3470−3481.

19. Schouler C, Taki A, Chouikha I, Moulin-Schouleur M & Gilot P (2009) A genomic island of an extraintestinal pathogenic *Escherichia coli* strain enables the metabolism of fructooligosaccharides, which improves intestinal colonization*. J Bacteriol* 191: 388−393.

20. Petersen A*, et al* (2009) Some putative prebiotics increase the severity of *Salmonella enterica* serovar *typhimurium* infection in mice*. BMC Microbiol* 9: 245.

21. Pfeiler EA, Azcarate-Peril MA & Klaenhammer TR (2007) Characterization of a novel bile-inducible operon encoding a two-component regulatory system in *Lactobacillus acidophilus. J Bacteriol* 189: 4624−4634.

22. Mohamadzadeh M*, et al* (2011) Regulation of induced colonic inflammation by *Lactobacillus acidophilus* deficient in lipoteichoic acid*. Proc Natl Acad Sci U S A* 108 Suppl 1: 4623−4630.

23. Ringel-Kulka T*, et al* (2011) Probiotic bacteria *Lactobacillus acidophilus* NCFM and *Bifidobacterium lactis* Bi-07 versus placebo for the symptoms of bloating in patients with functional bowel disorders: A double-blind study*. J Clin Gastroenterol* 45: 518−525.

24. Leyer GJ, Li S, Mubasher ME, Reifer C & Ouwehand AC (2009) Probiotic effects on cold and influenza-like symptom incidence and duration in children*. Pediatrics* 124: e172-179.

25. Ouwehand AC, Tiihonen K, Saarinen M, Putaala H & Rautonen N (2009) Influence of a combination of *Lactobacillus acidophilus* NCFM and lactitol on healthy elderly: Intestinal and immune parameters*. Br J Nutr* 101: 367−375.

26. Altermann E*, et al* (2005) Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM*. Proc Natl Acad Sci U S A* 102: 3906−3912.

27. Barrangou R*, et al* (2006) Global analysis of carbohydrate utilization by *Lactobacillus acidophilus* using cDNA microarrays*. Proc Natl Acad Sci U S A* 103: 3816−3821.

28. Andersen JM*, et al* (2011) Transcriptional and functional analysis of galactooligosaccharide uptake by *lacS* in *Lactobacillus acidophilus. Proc Natl Acad Sci U S A* 108: 17785−17790.

29. Cantarel BL*, et al* (2009) The carbohydrate-active EnZymes database (CAZy): An expert resource for glycogenomics*. Nucleic Acids Res* 37: D233−238.

30. Petersen TN, Brunak S, von Heijne G & Nielsen H (2011) SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat Methods* 8: 785−786.

31. Nakai H*, et al* (2009) The maltodextrin transport system and metabolism in *Lactobacillus acidophilus* NCFM and production of novel α-glucosides through reverse phosphorolysis by maltose phosphorylase. *FEBS J* 276: 7353−7365.

32. Fredslund F*, et al* (2011) Crystal structure of α-galactosidase from *Lactobacillus acidophilus* NCFM: Insight into tetramer formation and substrate binding. *J Mol Biol* 412: 466−480.

33. Goh YJ*, et al* (2009) Development and application of a *upp*-based counterselective gene replacement system for the study of the S-layer protein SlpX of *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol* 75: 3093−3105.

34. Francl AL, Thongaram T & Miller MJ (2010) The PTS transporters of *Lactobacillus gasseri* ATCC 33323. *BMC Microbiol* 10: 77.

35. Kall L, Krogh A & Sonnhammer EL (2007) Advantages of combined transmembrane topology and signal peptide prediction-the phobius web server . *Nucleic Acids Res* 35: W429−432.

36. Saulnier DM, Molenaar D, de Vos WM, Gibson GR & Kolida S (2007) Identification of prebiotic fructooligosaccharide metabolism in *Lactobacillus plantarum* WCFS1 through microarrays . *Appl Environ Microbiol* 73: 1753−1765.

37. Rycroft CE, Jones MR, Gibson GR & Rastall RA (2001) Fermentation properties of gentio-oligosaccharides. *Lett Appl Microbiol* 32: 156−161.

38. Grimoud J*, et al* (2010) In vitro screening of probiotics and synbiotics according to anti-inflammatory and anti-proliferative effects. *Int J Food Microbiol* 144: 42−50.

39. Makelainen HS, Makivuokko HA, Salminen SJ, Rautonen NE & Ouwehand AC (2007) The effects of polydextrose and xylitol on microbial community and activity in a 4-stage colon simulator. *J Food Sci* 72: M153−159.

40. Wiesmann C, Hengstenberg W & Schulz GE (1997) Crystal structures and mechanism of 6-phospho-β-galactosidase from *Lactococcus lactis. J Mol Biol* 269: 851−860.

41. Goffin D*, et al* (2011) Will isomalto-oligosaccharides, a well-established functional food in Asia, break through the European and American market? The status of knowledge on these prebiotics. *Crit Rev Food Sci Nutr* 51: 394−409.

42. Saburi W, Mori H, Saito S, Okuyama M & Kimura A (2006) Structural elements in dextran glucosidase responsible for high specificity to long chain substrate . *Biochim Biophys Acta* 1764: 688−698.

43. Abbott DW*, et al* (2010) The molecular basis of glycogen breakdown and transport in *Streptococcus pneumoniae. Mol Microbiol* 77: 183−199.

44. O'Sullivan O*, et al* (2009) Comparative genomics of lactic acid bacteria reveals a niche-specific gene set. *BMC Microbiol* 9: 50.

45. Callanan M*, et al* (2008) Genome sequence of *Lactobacillus helveticus*, an organism distinguished by selective gene loss and insertion sequence element expansion. *J Bacteriol* 190: 727–735.

46. Daniels L & Zeikus JG (1975) Improved culture flask for obligate anaerobes. *Appl Microbiol* 29: 710–711.

47. Kimmel SA & Roberts RF (1998) Development of a growth medium suitable for exopolysaccharide production by *Lactobacillus delbrueckii* ssp. *bulgaricus* RR. *Int J Food Microbiol* 40: 87–92.

48. Walker DC & Klaenhammer TR (1994) Isolation of a novel IS3 group insertion element and construction of an integration vector for *Lactobacillus* spp. *J Bacteriol* 176: 5330–5340.

49. Kingsford CL, Ayanbule K & Salzberg SL (2007) Rapid, accurate, computational discovery of rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* 8: R22.

50. Human Microbiome Jumpstart Reference Strains Consortium*, et al* (2010) A catalog of reference genomes from the human microbiome. *Science* 328: 994–999.

51. Magrane M & Consortium U (2011) UniProt knowledgebase: A hub of integrated protein data. *Database (Oxford)* 2011: doi: 10.1093/database/bar009.

52. Larkin MA*, et al* (2007) Clustal W and clustal X version 2.0. *Bioinformatics* 23: 2947–2948.

53. Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410.

54. Law J*, et al* (1995) A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *J Bacteriol* 177: 7011–7018.

55. Russell WM & Klaenhammer TR (2001) Efficient system for directed integration into the *Lactobacillus acidophilus* and *Lactobacillus gasseri* chromosomes via homologous recombination. *Appl Environ Microbiol* 67: 4361–4364.

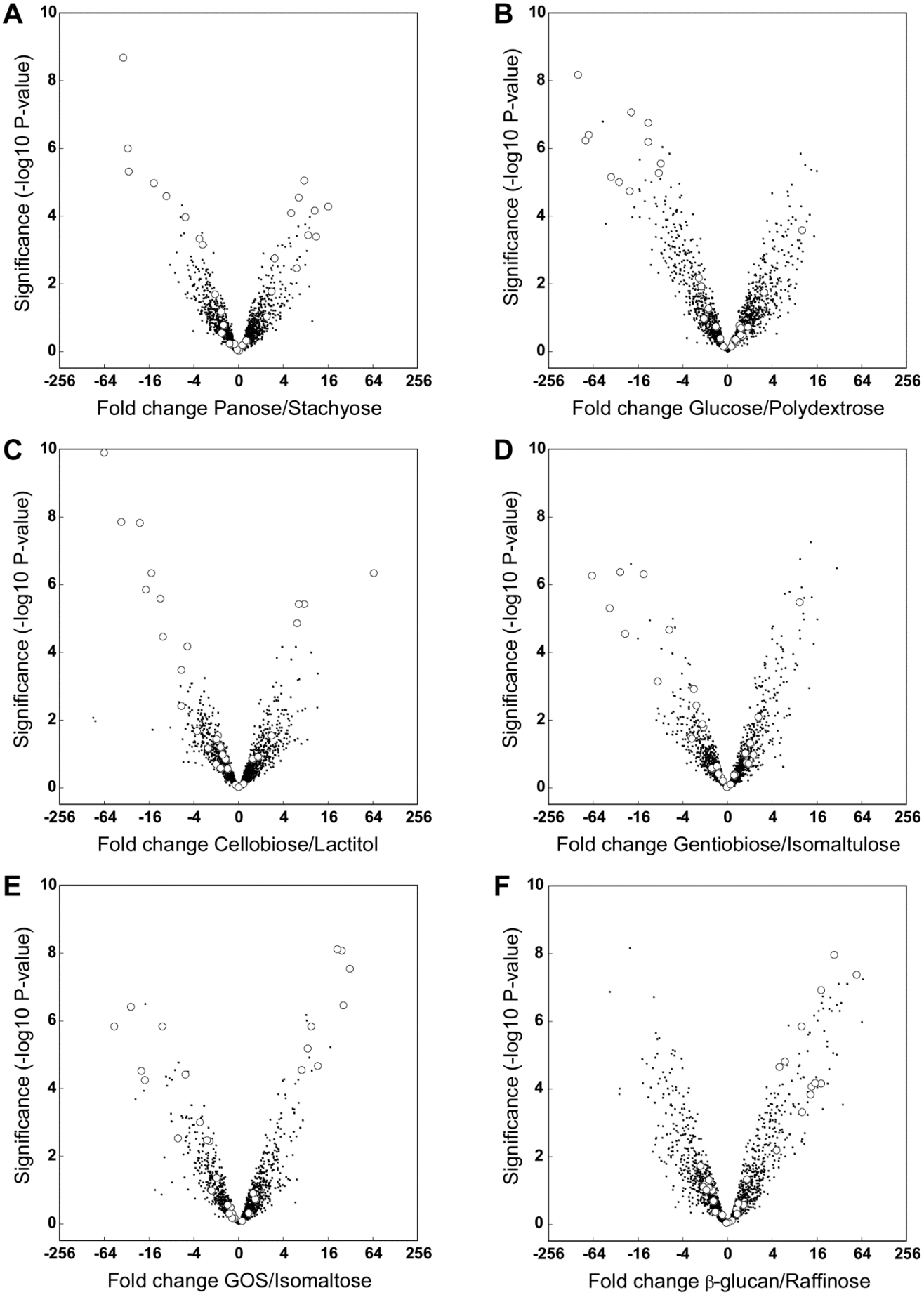

A
Significance (-log10 P-value)
Fold change Panose/Stachyose
B
Significance (-log10 P-value)
Fold change Glucose/Polydextrose
C
Significance (-log10 P-value)
Fold change Cellobiose/Lactitol
D
Significance (-log10 P-value)
Fold change Gentiobiose/Isomaltulose
E
Significance (-log10 P-value)
Fold change GOS/Isomaltose
F
Significance (-log10 P-value)
Fold change β-glucan/Raffinose
10
8
6
4
2
0
-256
-64
-16
-4
0
4
16
64
256

Figure 1: Representative volcano plots of the oligosaccharide-induced differential global transcriptome within *L. acidophilus* NCFM. All genes are shown as black dots (·) and all statistically significant upregulated genes involved with oligosaccharide metabolism (Table 1) are depicted as white circles (○).

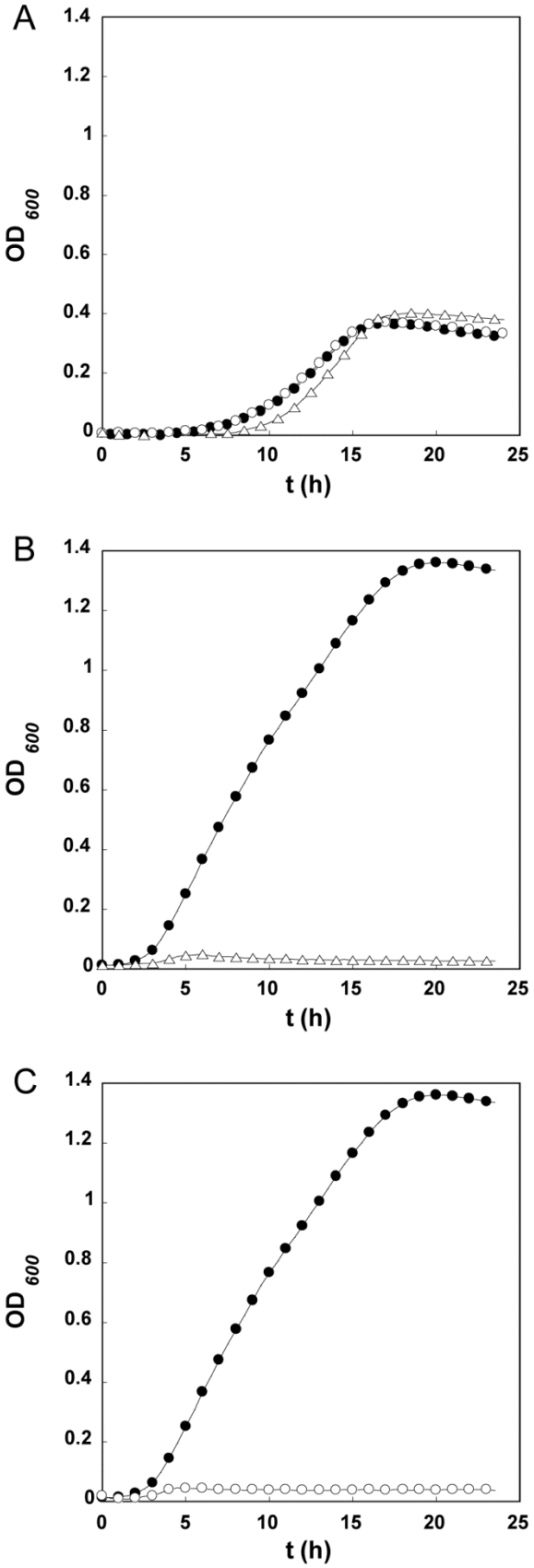


Figure 2: Phenotypic characterization of single gene deletions within *L. acidophilus* NCFM. Growth profiles are shown on galactose (A) and raffinose (B and C) for the mutants within the stachyose-induced gene cluster lacking the GH36 α-

galactosidase ΔLBA1438 (Δ) or the solute binding protein component of the ABC transporter ΔLBA1442 (○) compared to *upp*-wildtype (●).

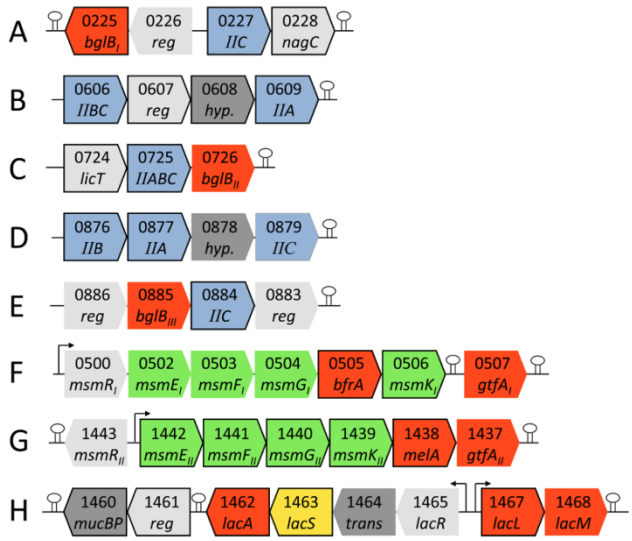


Figure 3: Organization of gene clusters encoding upregulated genes by potential prebiotic oligosaccharide stimulation. All genes are listed with locus tag number and gene name (PTS permeases are shown with domain name; regulators, hypothetical proteins and transposons are abbreviated as reg, hyp. and trans respectively). Gene product functions are colored red for glycoside hydrolases, light grey for transcriptional regulators, blue for PTS permease domains, dark grey for proteins unrelated to carbohydrate metabolism, green for ABC transporter domains and yellow for the GPH permease. All upregulated genes (Table 2) are shown with framed boxes, CRE regulatory sites are represented by arrows and predicted rho-independent transcription terminators (49) by stem loops.

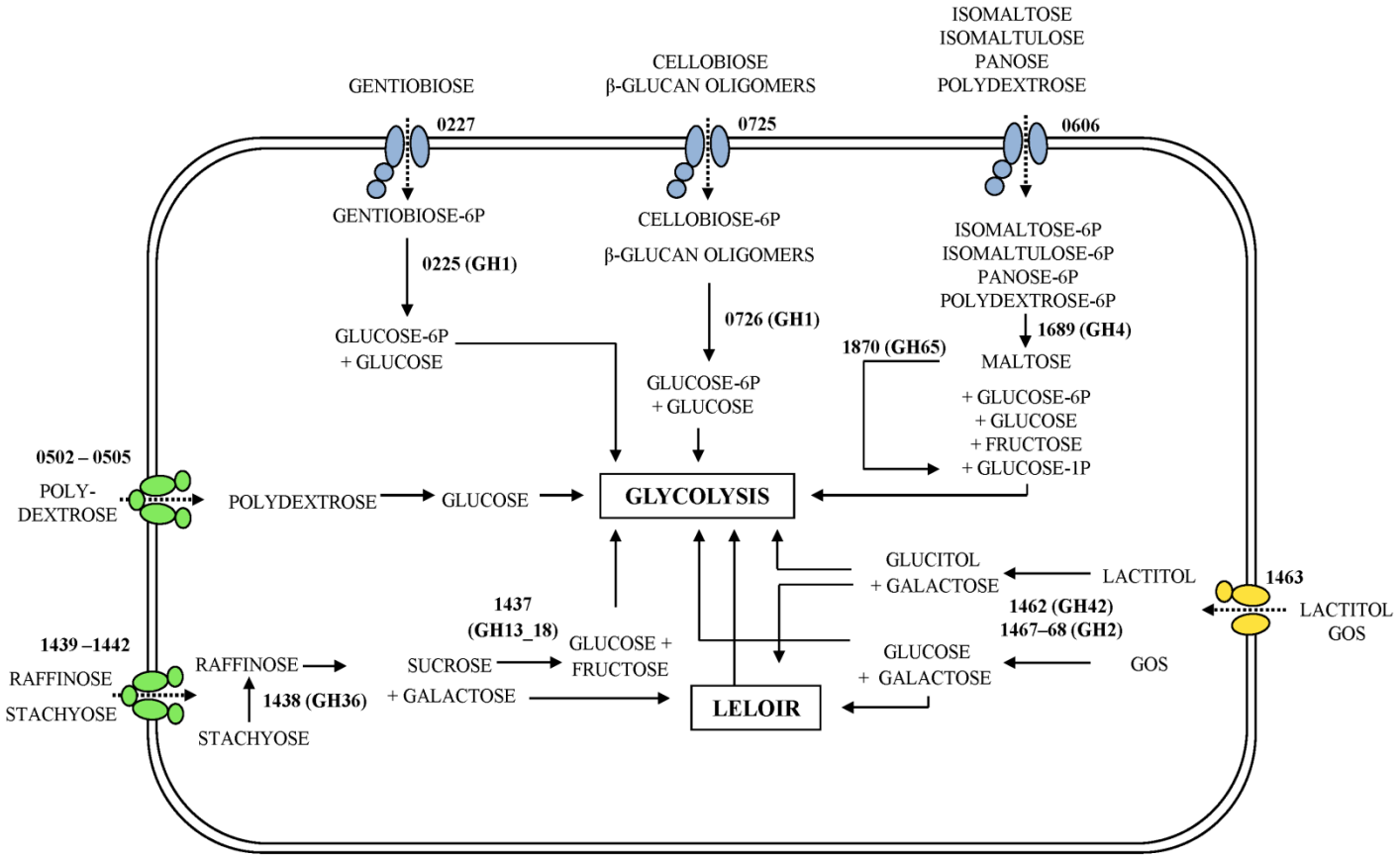


Figure 4: Reconstructed uptake and catabolic pathways in *L. acidophilus* NCFM. Proteins are listed by locus tag LBA numbers, transporters are colored by class (Figure 3) and glycoside hydrolases are listed with GH family number. The Polydextrose fraction transported by the ABC transporter (LBA0502–LBA0505) is uncertain and thus the hydrolytic pathway is marked as unknown. The present data outlines the PTS permease LBA0606 (higher level of induction compared to LBA0502–LBA0505) and associated hydrolytic pathway, as the main route of Polydextrose utilization by *L. acidophilus* NCFM.

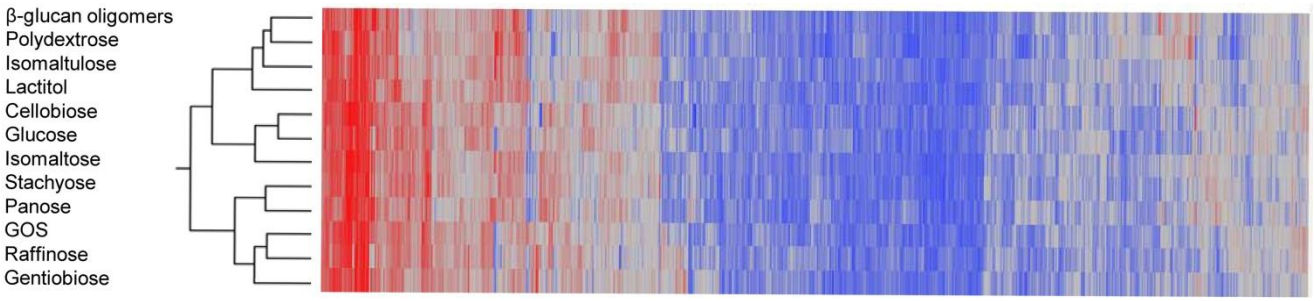


Figure S1: Hierarchical two-way clustering of the global gene expression patterns across 1823 genes for all carbohydrate growth conditions. Up-regulated genes are shown in red while downregulated genes are shown in blue.

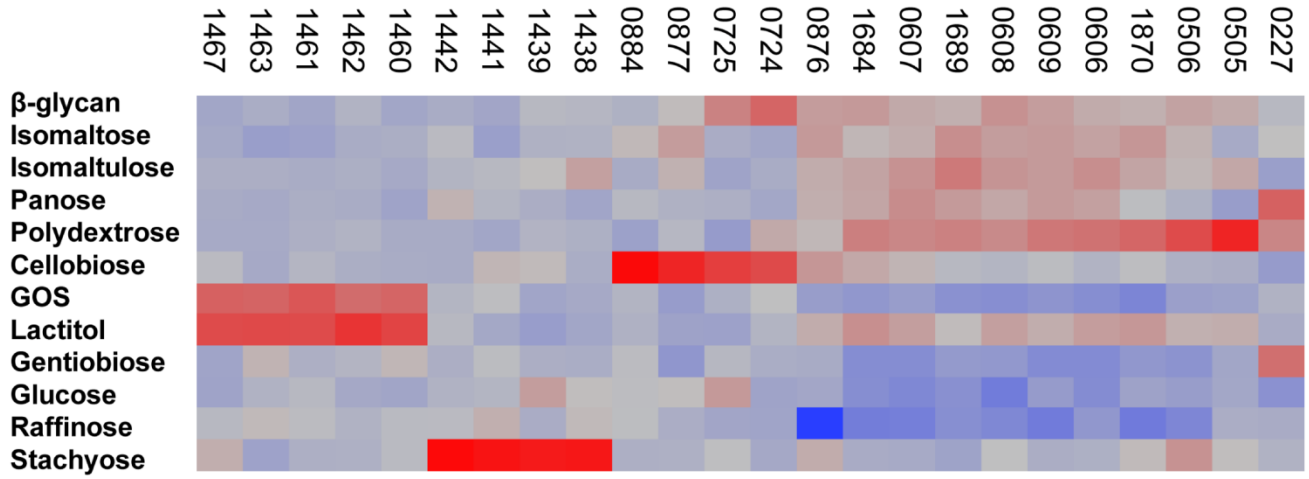


Figure S2: Two-way clustering of identified statistically significant genes involved in carbohydrate utilization listed with locus tag LBA numbers. Up-regulated genes are shown in red while downregulated genes are shown in blue.

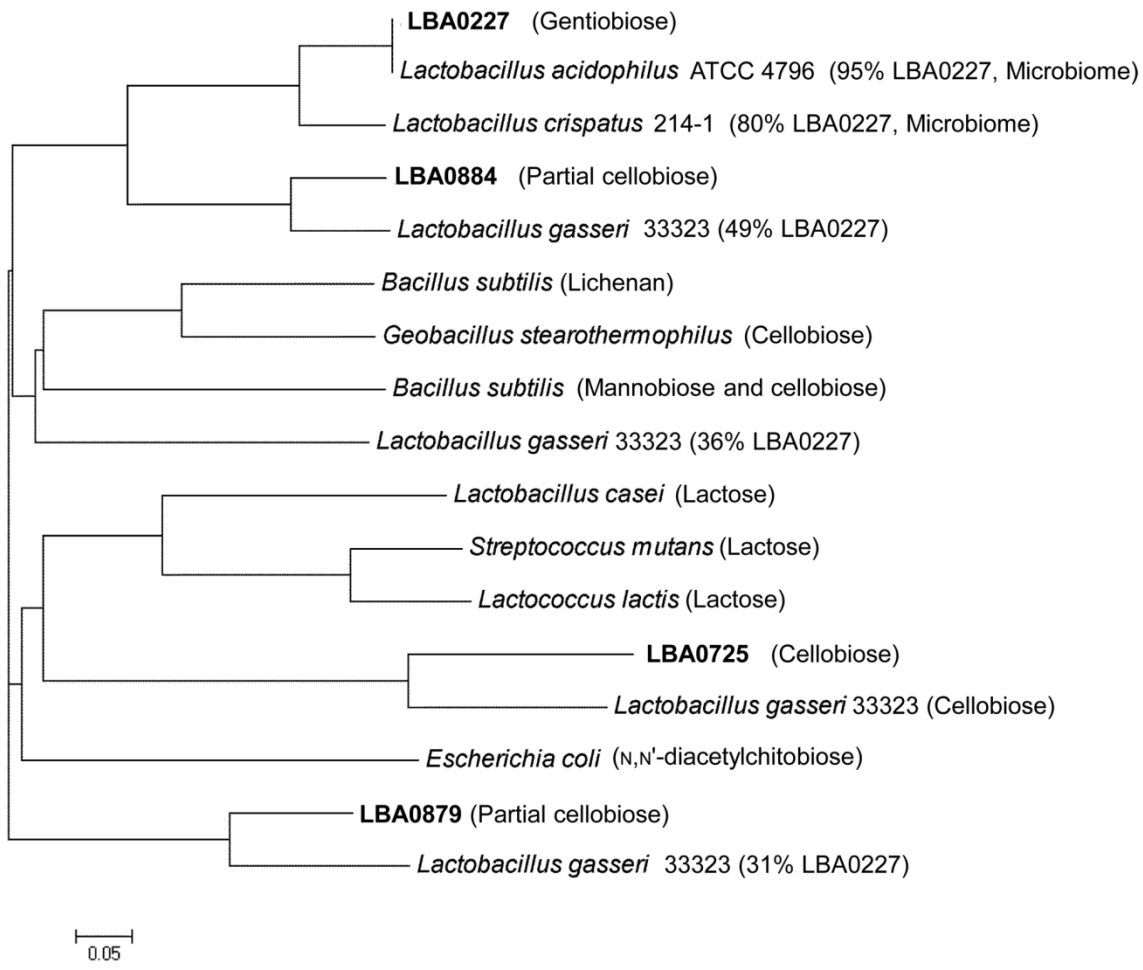


Figure S3: Phylogenetic diversity of β-glucoside specific PTS EIIC domains of *L. acidophilus* NCFM.

Clustering of identified *L. acidophilus* NCFM PTS EIIC domains (highlighted in bold) is visualized by a phylogenetic tree where representative sequences are used to illustrate functional segregation. Reference sequences are from *L. gasseri* 33323 or PTS EIIC homologs (>50 % amino acid identity) from reference genomes of the human microbiome (50). PTS EIIC domain sequences are identified by homology search of the Swiss-prot database (51) and all phylogenetic distances were calculated using ClustalW2 (52). All known substrate specificities are given in parentheses, otherwise amino acid identity to LBA0227 is stated. Uniprot references: *Bacillus subtilis*, lichenan (P46317), *Geobacillus stearothermophilus*, cellobiose (Q45400), *Bacillus subtilis*, mannobiose and cellobiose (O05507), *Lactobacillus casei*, lactose (P24400), *Streptococcus mutants*, lactose (P50976), *Lactococcus lactis*, lactose (P23531) and *Escherichia coli*, *N,N*'-diacetylchitobiose (P17334.2).

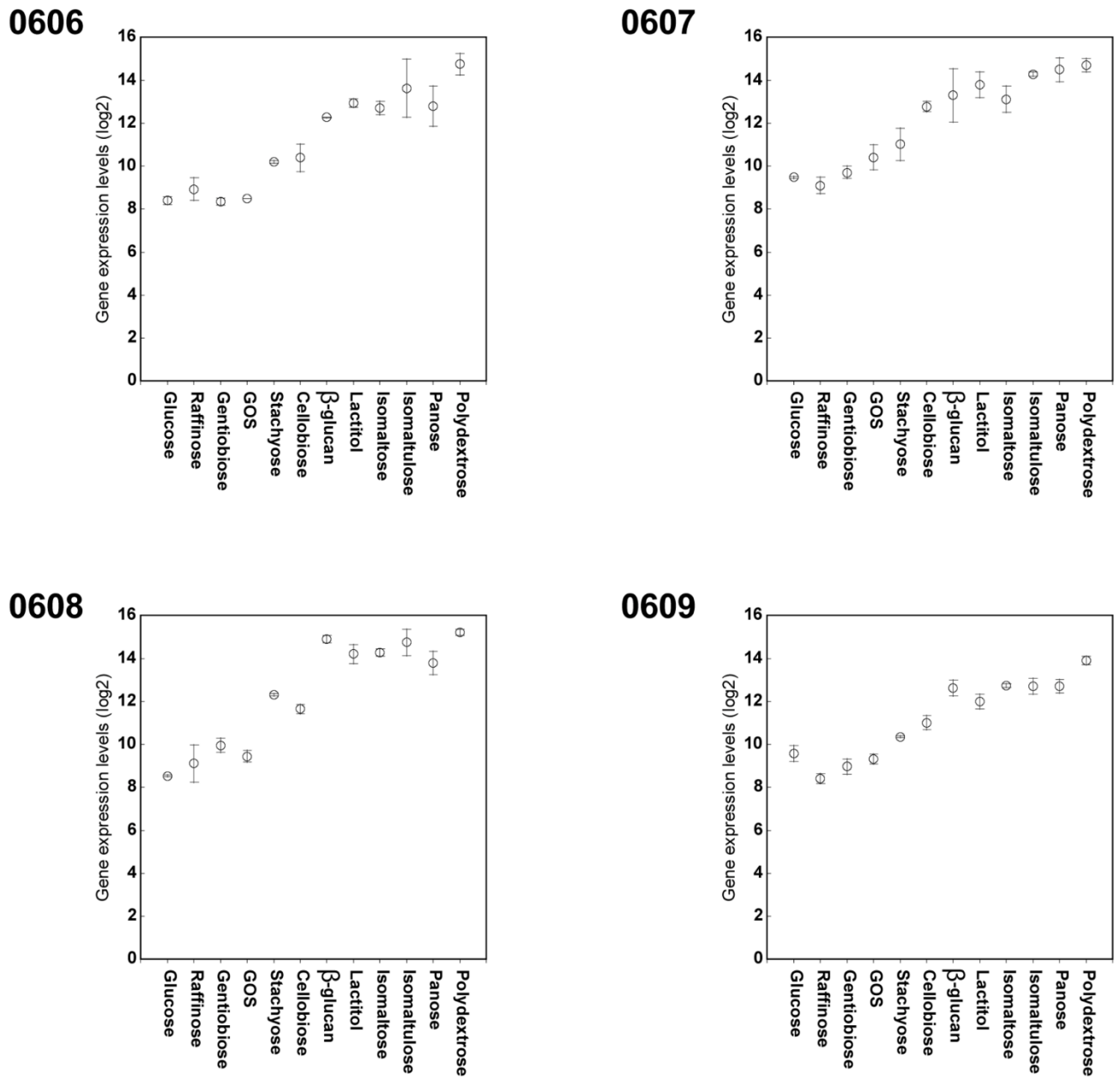


Figure S4: Gene expression levels for highest induced gene (LAB0608) and gene cluster. The locus encoded an α-1,6-glucoside specific PTS EIIBC (LBA0606), a transcriptional regulator (LBA0607), a putative transporter associated protein (LBA0608) and PTS EIIA component (LBA0609) showed consistent high expression of the full locus indicating a functional connection of LBA0608 and PTS permease uptake. Values are given as the mean value (○) of the technical replicates represented by bars.

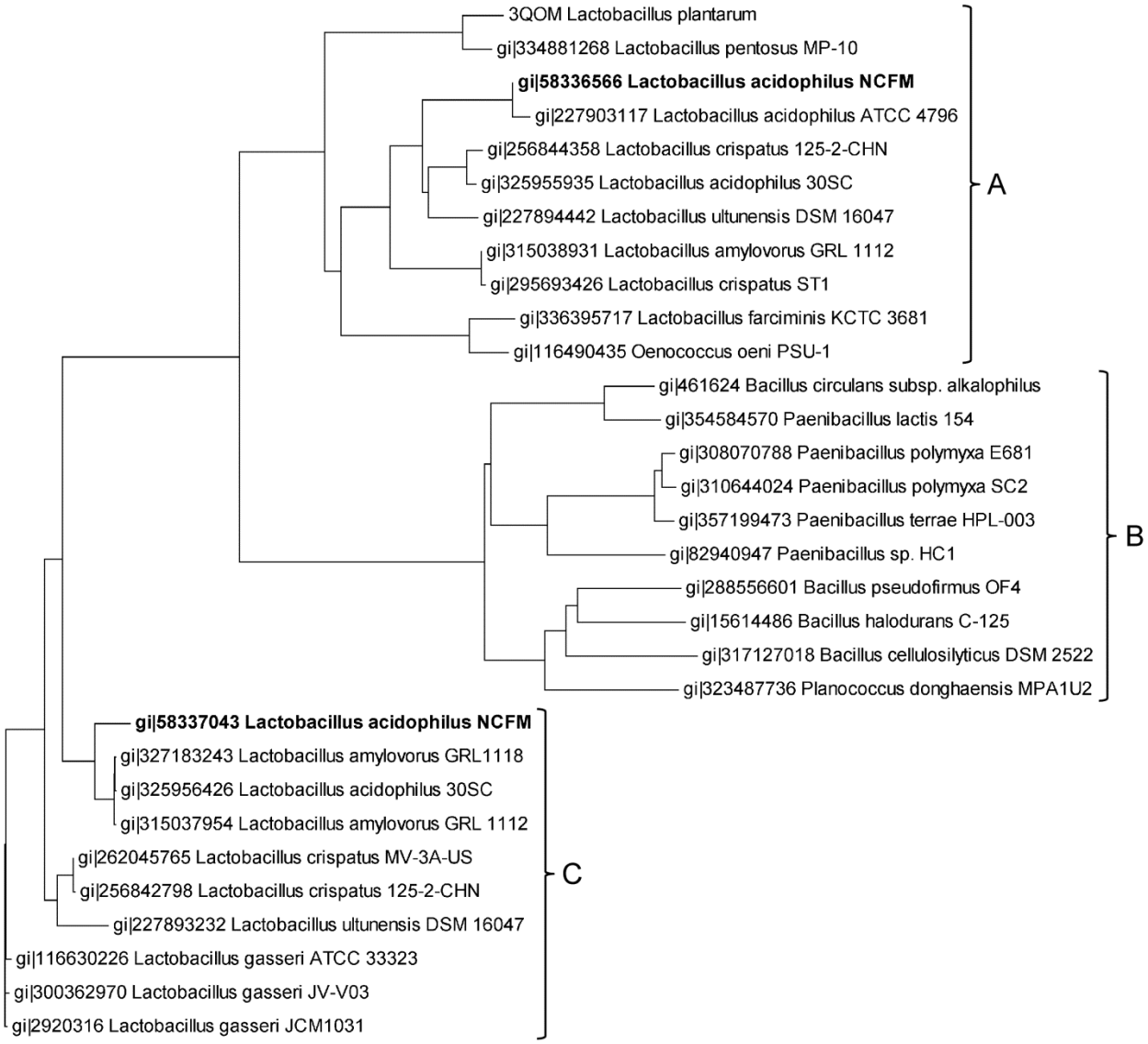


Figure S5: Phylogenetic relationship of the two identified 6-phospho-β-glucosidases (LBA0225 and LBA0726) compared to characterized GH1 enzymes (the *Bacillus circulans* subsp. *alkalophilus* β-1,4-glucosidase (gi: 308070788) and *L. plantarum* 6-phospho-β-glucosidase structure (PDB accession: 3QOM)). The 10 closest homologs, all listed by gi-number, specie and strain name, were identified for LBA0227, LBA726 and the cellobiose specific *Bacillus circulans* subsp. *alkalophilus* β-1,4-glucosidase by BLAST searching against the non-redundant database (53) and all phylogenetic distances were calculated using ClustalW2 (52). Distinct clustering even of related taxa was observed reflecting differential substrate specificity for cluster (A) proposed to be gentiobiose-6-phosphate specific with LBA0227 highlighted in bold, (B) cellobiose specific as represented by *Bacillus circulans* subsp. *alkalophilus* β-1,4-glucosidase and (C) proposed to be cellobiose-6-phosphate specific with LBA0726 highlighted in bold.

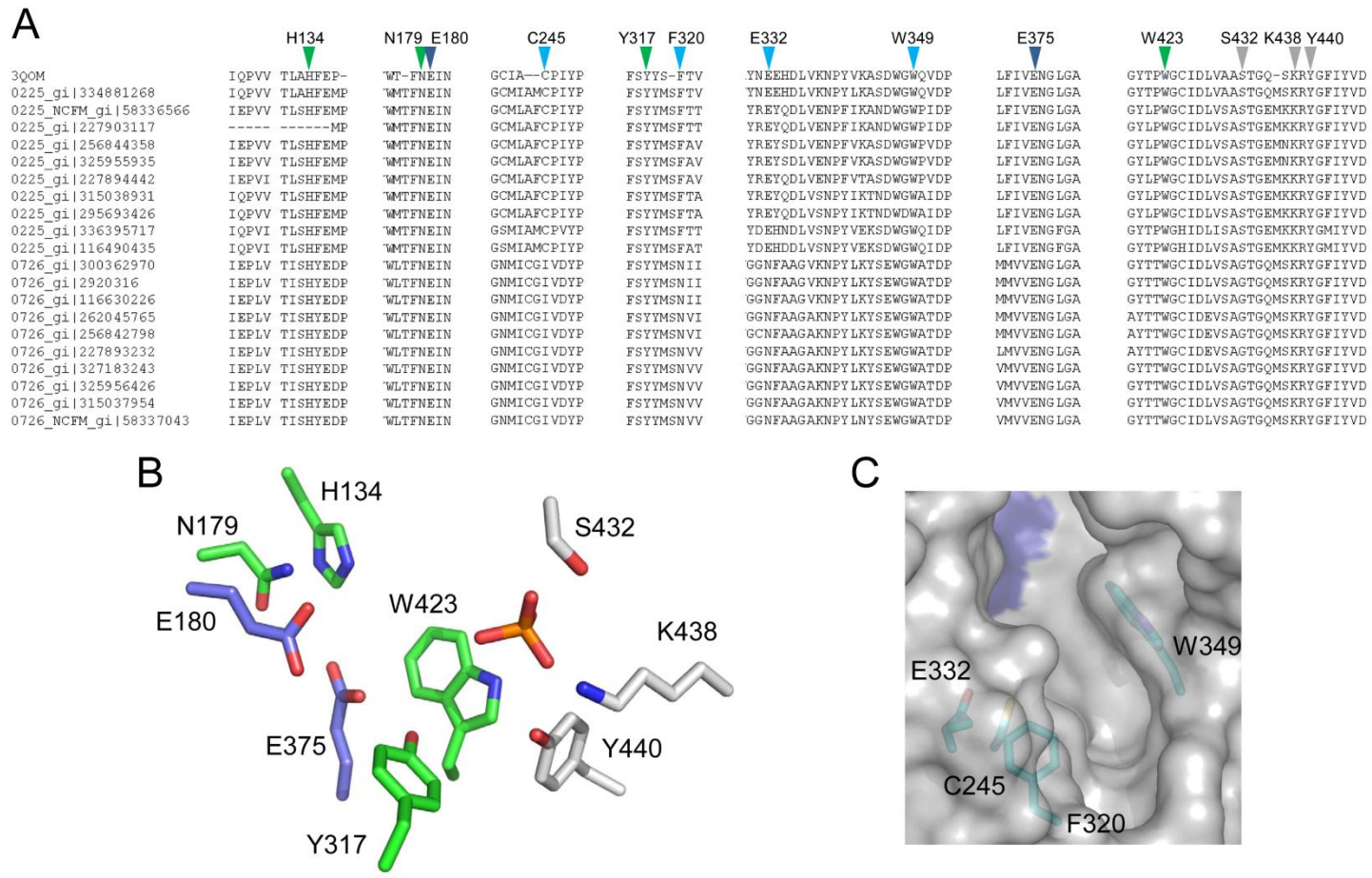


Figure S6: Functionally pivotal residues in 6-phospho-β-glucosidases of GH1. (A) The selected segments of the multi-sequence alignment used to construct the phylogenetic tree (Figure S5 cluster A and C), showing conserved and variable putative substrate interacting residues of LBA0225 and LBA0726. Conserved residues of the –1 subsite are marked with green, the catalytic acid/base (E180) and nucleophile (E375) are marked with purple, the putative +1 subsite is marked with cyan and the residues that recognize the phosphate moiety in the phosphate binding pocket are marked with grey. All numbering corresponds to the *L. plantarum* 6-phospho-β-glucosidase structure (PDB accession: 3QOM) as reference, also used to depict functionally important residues in 6-phospho-β-glucosidases of GH1; (B) highly conserved active site residues are colored as in (A) and shown in sticks. (C) A surface representation of the active site (40% transparency) showing (cyan sticks) the proposed putative subsite +1 specificity determinants distinguishing 6-phospho-β-1,6-glucosides represented by LBA0227, from 6-phospho-β-1,4-glucosides represented by LBA0725. The catalytic residues are surface colored in purple to denote the position of the –1 subsite. Pymol was used for molecular rendering (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC.)

Tables:

Table 1: List of carbohydrates used in this study.

Carbohydrate	Structure[1]	Carbohydrate linkage family	DP [2]	Manufacturer or supplier	Purity (as given by manufacturer or supplier)
Glucose	Glc*p*	-	1	Sigma	> 99 %
GOS	[β-D-Gal*p*-(1–4)]$_n$-D-Glc*p*	β-galactoside	2–6	Dupont	> 94 % DP ≥ 2
Lactitol	β-D-Gal*p*-(1–4)-D-Glc-ol	β-galactoside	2	Dupont	> 99 %
Cellobiose	β-D-Glc*p*-(1–4)-D-Glc*p*	β-glucoside	2	Fluka AG	> 99 %
Gentiobiose[3]	β-D-Glc*p*-(1–6)-D-Glc*p*	β-glucoside	2	Sigma	> 98 %
β-glucan oligomers	[β-D-Glc*p*-(1–4)]$_m$-β-D-Glc*p*-(1–3)-β-D-Glc*p*-[β-D-(1–4)-Glc*p*]$_o$	β-glucoside	DP ≥ 2	Biovelop AB (Sweden)	Essentially free of monosaccharides and cellobiose[4]
Raffinose	α-D-Gal*p*-(1–6)-D-Glc*p*-(α1,β2)-D-Fru*f*	α-galactoside	3	Sigma	> 99 %
Stachyose	[α-D-Gal*p*-(1–6)]$_2$-D-Glc*p*-(α1,β2)-D-Fru*f*	α-galactoside	4	Sigma	> 98 %
Isomaltose	α-D-Glc*p*-(1–6)-D-Glc*p*	α-glucoside	2	Sigma-Aldrich	> 98 %
Isomaltulose	α-D-Glc*p*-(1–6)-D-Fru*f*	α-glucoside	2	Dupont	> 99 %
Panose	α-D-Glc*p*(1–6)-α-D-Glc*p*(1–4)-D-Glc*p*	α-glucoside	3	Sigma	> 98 %
Polydextrose[5]	Primarily mixed α-glucans, reduced ends	α-glucoside	2–30	Dupont	Essentially free of monosaccharides

Footnotes:

[1] n= [1–5], m=[0–2] and o=[0–3], 'n' is based on oligosaccharide product range of transglycosylation for GOS synthesis as previously described (27). 'm' and 'o' are predicted ranges from the theoretical β-glucan repeating polymeric structure and the enzyme used for partial hydrolysis of β-glucan.

[2] Degree of polymerization

[3] Isomaltose free, in-house HPAEC-PAD analysis.

[4] In-house HPAEC-PAD analysis

[5] Polydextrose Litesse® Ultra (Dupont)

Table 2: Statistically significant upregulated genes involved in carbohydrate uptake and catabolism. The genes are listed by ascending locus tag numbers. Only the oligosaccharide that elicited the highest induction level is listed for genes that are upregulated by more than one oligosaccharide.

ORF	Gene cluster identifier[1]	Gene product annotation	Highest inducing oligosaccharide	Inducing linkage type[2]	Volcano plot (Figure 2)	Fold upregulated	$-\log_{10}$ (P-value)
227	A	PTS, EIIC	Gentiobiose	β-glc	2E	9.3	5.48
505	F	β-fructosidase (bfrA), EC 3.2.1.26, GH32	Polydextrose	α-glc	2B	7.9	5.55
506	F	ATP-binding protein (msmK)	Polydextrose	α-glc	2B	11.7	6.75
606	B	PTS permease, EIIBC	Polydextrose	α-glc	2B	81.6	6.23
607	B	Transcriptional regulator, RpiR family	Polydextrose	α-glc	2B	36.9	5.14
608	B	Putative transporter accessory protein	Polydextrose	α-glc	2B	103.3	8.18
609	B	PTS, EIIA	Polydextrose	α-glc	2B	19.9	7.06
724	C	Transcriptional regulator, LicT family	Cellobiose	β-glc	2C	6.1	4.86
725	C	PTS, EIIC	Cellobiose	β-glc	2C	66.0	6.33
876	D	PTS, EIIB	β-glucan oligomers	β-glc	2F	27.1	7.97
877	D	PTS, EIIA	Cellobiose	β-glc	2C	7.6	5.42
884	E	PTS, EIIC	Cellobiose	β-glc	2C	6.4	5.42
1438	G	α-galactosidase (*melA*), EC 3.2.1.22, GH36	Stachyose	α-gal	2E	30.1	5.31
1439	G	ABC, ATP-binding protein ($msmK_{II}$)	Stachyose	α-gal	2E	31.2	5.99
1441	G	ABC, transmembrane permease ($msmF_{II}$)	Stachyose	α-gal	2E	9.3	4.58
1442	G	ABC, substrate-binding protein ($msmE_{II}$)	Stachyose	α-gal	2E	35.9	8.67
1460	H	Putative mucus binding protein (mucBP)	Lactitol	β-gal	2C	11.4	5.59
1461	H	Transcriptional regulator, TetR family	GOS	β-gal	2B	25.5	6.46
1462	H	β-galactosidase (lacA), EC 3.2.1.23, GH42	Lactitol	β-gal	2C	64.0	9.89
1463	H	Lactose permease (lacS)	Lactitol	β-gal	2C	38.2	7.84
1467	H	β-galactosidase large subunit (lacL), EC 3.2.1.23, GH2	GOS	β-gal	2B	24.6	8.06
1684	NA	PTS, EIIA	Polydextrose	α-glc	2B	11.7	6.19
1689	NA	Maltose-6-P glucosidase (malH), EC 3.2.1.122, GH4	Isomaltulose	α-glc	2D	65.9	6.26
1870	NA	Maltose phosphorylase (malP), EC 2.4.1.8, GH65	Polydextrose	α-glc	2B	28.8	5.01

[1] Genes not assigned a gene cluster (Figure 3) are listed as not assigned (NA).
[2] The predominant glycosidic linkage types have been abbreviated as: α-galactosides (α-gal), α-glucosides (α-glc), β-galactosides (β-gal) and β-glucosides (β-glc).

Table 3: Strains and plasmids used in the study

Strain or plasmid	Characteristics	Reference or source
E. coli strains		
NCK1831	EC101: RepA$^+$ JM101; Km^r; *repA* from pWV01 integrated in chromosome; host for pORI-based plasmids	(54)
NCK1911	NCK1831 harboring pTRK935	(33)
NCK2122	NCK1831 harboring pTRK1013	This study
NCK2124	NCK1831 harboring pTRK1014	This study
L. acidophilus strains		
NCFM	Human intestinal isolate	(26)
NCK1909	NCFM carrying a 315 bp in-frame deletion in the *upp* gene	(33)
NCK1910	NCK1909 harboring pTRK669, host for pORI-based counter selective integration vector	(33)
NCK2123	NCK1909 carrying a 2029 bp in-frame deletion in the *melA* gene	This study
NCK2125	NCK1909 carrying a 1141 bp in-frame deletion in the *msmE* gene	This study
Plasmids		
pTRK669	Ori (pWV01], Cm^r RepA$^+$	(55)
pTRK935	pORI28 derived with an inserted *upp* expression cassette and *lacZ´* from pUC19, serves as counterselective integration vector, Em^r	(33)
pTRK1013	pTRK935 with a mutated copy of *melA* cloned into BamHI/EcoRI sites	This study
pTRK1014	pTRK935 with a mutated copy of *msmE$_{II}$* cloned into BamHI/EcoRI sites	This study

Table S1: Primers used for construction of gene deletion mutants. Restriction sites are highlighted in bold and underlined.

LBA1438 upstream flanking region	
1438A	CGC**GGATCC**CGAACCACTATCCAACCTTGA
1438B	CCACCATCTTCAATAGAAAGC
LBA1438 downstream flanking region	
1438C	GCTTTCTATTGAAGATGGTGGACCTTGGCTTTTATGATCCTATTG
14383D	CCG**GAATTC**CCCAAATTTCTGGCTCTACAA
LBA1438 DNA excision control	
1438UP	CACCAAAGTAGGCGATACTGAA
1438DN	ACAGCCCCCTTCAAGTCTTC
LBA1442 upstream flanking region	
1442A	CGC**GGATCC**TTGATGCAAGTAACGCTGAGA
1442B	GTAGCCATCATGACTCCAATTAG
LBA1442 downstream flanking region	
1442C	CTAATTGGAGTCATGATGGCTACGGTAATAAACAACAAATGGTTAATG
1442D	CCG**GAATTC**GGGAGTTCAATCTTCCAGAAA
LBA1442 DNA excision control	
1442UP	AAGGCCAAATGACAATAATGC
1442DN	GCACCTTGAACTAATGGGAAA

6.2 Mapping the uptake and catabolic pathways of prebiotic utilization in *Bifidobacterium animalis* subsp. *lactis* Bl-04 by differential transcriptomics

In preparation for BMC genomics.

Title:

Mapping the uptake and catabolic pathways of prebiotics in *Bifidobacterium animalis* subsp. *lactis* Bl-04 by differential transcriptomics

Authors (in order top to bottom):

Joakim Mark Andersen[1,2] joa@bio.dtu.dk

Rodolphe Barrangou[3] rodolphe.barrangou@danisco.com

Maher Abou Hachem[1] maha@bio.dtu.dk

Sampo J. Lahtinen[4] sampo.lahtinen@danisco.com

Yong Jun Goh[2] yjgoh@ncsu.edu

Birte Svensson[1] bis@bio.dtu.dk

Todd R. Klaenhammer[2] trk@ncsu.edu

Author affiliation:
[1] Enzyme and Protein Chemistry, Department of Systems Biology, Søltofts Plads Building 224, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark.
[2] Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State University, Box 7624, Raleigh, NC 27695.
[3] DuPont Nutrition and Health, 3329 Agriculture Drive, Madison, WI 53716.
[4] DuPont Nutrition and Health, Sokeritehtaantie 20, FI-02460 Kantvik, Finland.

Formatted for:

BMC genomics

Keywords: (3-10)

Bifidobacterium animalis subsp. *lactis*, transcriptomics, ABC transporter, GPH transporter, prebiotics, Glycoside hydrolase

List of abbreviations:

Analysis of variance (ANOVA), ATP binding cassette (ABC), β-galacto-oligosaccharides (GOS),Clusters of orthologous Groups (COG), Glycoside hydrolase (GH), Major Facilitator Superfamily (MFS), Raffinose family oligosaccharides (RFO),Xylo-oligosaccharides (XOS)

Authors' contributions:

Designed research: RB, MAH, SL, BS, TRK
Performed research: JMA, YJG
Contributed new reagents (SL) and analytic tools: YJG, TRK
Analyzed data: JMA, RB, YJG, TRK
Wrote the paper: JMA, RB, MAH, SL, YJG, BS, TRK

Abstract (BMC genomics max 300 words – currently 225)

Background: Probiotic bifidobacteria in combination with carbohydrate prebiotics have documented positive effects on human health regarding gastrointestinal disorders and improved immunity, however the routes of uptake remains unknown for most candidate prebiotics. Differential transcriptomics of *Bifidobacterium animalis* subsp. *lactis* Bl-04, induced by 11 potential prebiotic oligosaccharides was analyzed to identify the genetic loci for uptake and catabolism conferring utilization of the applied α- and β-linked hexoses, and β-xylosides.

Results: The global transcriptome was found to be modulated dependent of the utilized type of glycoside (galactoside, glucoside or xyloside). Carbohydrate transporters of the Major Facilitator superfamily (induced by: gentiobiose and galacto-oligosaccharides) and ATP-binding cassette transporters (upregulated by: cellobiose, β-galacto-oligosaccharides, isomaltose, maltrotriose, melibiose, panose, raffinose, stachyose, xylobiose and β-xylo-oligosaccharides) were differentially upregulated together with glycoside hydrolases from families 1, 2, 13, 36, 42, 43 and 77. Sequence analysis of the identified ABC transporter's solute binding proteins revealed patterns to the broadness and selective prebiotic utilization of bifidobacteria, which currently is a limiting factor to formulate and document novel prebiotics and synbiotics.

Conclusion: This study identifies and emphasizes the extensive capabilities of *Bifidobacterium animalis* subsp. *lactis* Bl-04 to utilize oligosaccharide potential prebiotics. ATP-binding cassette transporters are further emphasized to be involved in prebiotic utilization with dedicated glycoside hydrolases. The identified genetic loci will assist the further substantiation of gene clusters conferring selective utilization of prebiotic by probiotic organisms.

Background

Health-promoting microbes, defined as probiotics (1), have gained increased interest for improvement of human health through clinical studies. Research has shown bifidobacteria to be an important genus for probiotic interventions (2, 3). The areas of beneficial intervention applying bifidobacteria include among others prevention of necrotizing enterocolitis in infants (4), treatment of Crohn's disease (5) and immune functions in elderly (6). Understanding of the mechanism of actions underlying the probiotic character of bifidobacteria on the molecular level is mainly restricted to functional extrapolation from genome sequencing (7). To date, 53 *bifidobacterium* genomes have been deposited publicly and comparative analysis have shown the genetic diversity of bifidobacteria (8), leading to identification of genetic loci for colon adaption and colonization by host mucin degradation in *B. bifidum* (9) and foraging of dietary carbohydrates (10).

Enhancement of probiotic activity within the gastrointestinal tract has been observed by supplementing selectively utilizable carbohydrates (11), defined as prebiotics (12). Prebiotics are dietary non-digestible nutrients, dominantly carbohydrates, resistant to the host digestive system and main commensal microbiome residing in the colon. To date, only a few carbohydrates have been documented as prebiotics, namely: β-galacto-oligosaccharides, lactulose, fructo-oligosaccharides and inulin (13). Novel candidate prebiotics have been proposed primarily based on *in vitro* methodology, hence making more studies needed to fully document them as prebiotics being selectively utilized in the GIT (14, 15). The complex glycoside and linkage composition of prebiotics require broad uptake and hydrolytic pathways which have been proposed *in silico* within bifidobacteria (16, 17). Functional insight into proteins conferring the prebiotic uptake and catabolism rely mainly on the gene annotations failing to depict potentially novel pathways and sequence differences determining the prebiotic selectiveness of bifidobacteria.

Bifidobacterium animalis subsp. *lactis* Bl-04 (*B. lactis* Bl-04) is a member of the *animalis* cluster of bifidobacteria (18) and has documented positive effects as a probiotic in clinical interventions (19, 20), and when

supplemented as a synbiotic combination with certain prebiotics (21). The genome sequence of *B. lactis* Bl-04 extended the insight of putative probiotic abilities, revealing the bacterium to be highly GIT adapted with particular regards to utilization of dietary derived potential prebiotics (22), although *in silico* analysis precludes the broadness of substrate variety. This important knowledge could lead to improve understanding of *B. lactis* Bl-04 for applied use (23), and for design for novel prebiotics selectively stimulating probiotic bifidobacteria.

In the present study, we used differential transcriptomics to identify genetic loci encoding uptake and hydrolytic pathways for potential prebiotics manifested by 11 structurally diverse galactosides, glucosides and xylosides within *B. animalis* subsp. *lactis* Bl-04. This work validates and expands tentative *in silico* predictions of oligosaccharide transporters and specificities of glycoside hydrolases while leading to functional understanding of pre- and probiotic interactions and sets the stage for functional understanding of prebiotic utilization within bifidobacteria.

Results

Oligosaccharide induced global transcriptome profile of *B. lactis* BL-04

Global gene expression profiles were obtained for *B. lactis* Bl-04, exponentially growing on 11 potential prebiotics oligosaccharides (Table 1) and glucose representing α-galactosides (melibiose, raffinose and stachyose), β-galacto-oligosaccharides (GOS), α-glucosides (isomaltose, maltotriose and panose), β-glucosides (cellobiose and gentiobiose) and β-xylosides (xylobiose and xylo-oligosaccharides (XOS)). The gene expression intensities were quantified by whole genome DNA microarrays showing an overall comparable gene expression profile and high technical reproducibility (Figure 1) with only a subset of genes being upregulated differentially to each oligosaccharide although a slight deviation of the GOS and xylobiose samples was observed. The 10% of the highest constitutively expressed genes for all carbohydrates treatments (163 genes) were assigned clusters

of orthologous Groups (COG) categories (24) revealing the main cellular functions of cell growth and energy turn-over (Figure 1). Notably within these genes (listed by *B. lactis* Bl-04 locus tag numbers) were putative functions involved with fibronectin adhesion (Balac_1484-1485), host plasminogen interactions (Balac_1017 and Balac_1557), Phage immunity (25) (Balac_1305), bile-salt hydrolysis and peroxide reduction (Balac_0863 and 0865) and oligosaccharide ABC transporter facilitated uptake by a solute binding protein (Balac_1565) and ATP binding protein associated with oligosaccharide uptake (Balac_1610). All highlighting molecular functions related to probiotics mechanisms in *B. animalis* as previously suggested (26).

Table 1

Figure 1

Notably functional grouping of the global gene expression was observed based on the type of glycoside utilized (galactosides, glucosides or xylosides) from principal component analysis (Figure 2). This depicted a clear differentiation of the global transcriptome based on the type of glycoside utilized indicating how potential prebiotics can affect the global transcriptome and hence physiological functions in *B. lactis* Bl-04. In comparison, single gene clusters, differentially upregulated by specific carbohydrates, were observed in the heat map representation of the global gene expression profile (Figure 1) illustrating the specificity of the transcriptional response of carbohydrate utilization genes as compared to the differential global gene expression by a specific carbohydrates.

Figure 2

Differentially upregulated genes conferring prebiotic utilization

Analysis of the differential upregulation of specific genes mediating prebiotic utilization was done by one-way analysis of variance (ANOVA) and visualized by volcano plots (Figure 3) to identify statistically significant genes (cut off: p-value < $10^{-8,04}$) upregulated to each carbohydrate. An average of 56 genes was more than 2-fold differentially upregulated and above the statistical threshold for each pairwise comparison. Analysis revealed how subsets of genes involved with oligosaccharide utilization were consistently differentially expressed throughout the ANOVA (Table 2, and Figure 4 for real time quantitative-PCR validation of selected genes). This led to reconstruction of six putative gene clusters were based on the differential upregulation of specific genes to specific oligosaccharide treatments. Thus linking gene clusters encoding a transporter and glycoside hydrolase(s) to the uptake and degradation of substrates differing in the degree of polymerization or monosaccharide composition.

Figure 3

Table 2

Figure 4

Moreover, the relative induction of gene clusters involved in carbohydrate uptake and catabolism (Figure 5) evidently supports the identification of the differential specificities of upregulated proteins involved with prebiotics utilization.

Figure 5

Gene cluster analysis and functional assignment

Sequence analysis of the differentially upregulated loci identified in the uptake and hydrolysis of prebiotic oligosaccharides was done to reconstruct six functional gene clusters predicted to be fundamental components

for prebiotic utilization (Figure 6A-F). Notably, transport of prebiotics was facilitated either by Major Facilitator Superfamily (MFS) or ATP binding cassette (ABC) types of transporters co-encoded with one of more glycoside hydrolases of varying glycoside hydrolase (GH) families (27) as a shared structural genomic element essential for prebiotic utilization.

Figure 6

Gene cluster A, differentially up-regulated by gentiobiose, encoded an MFS class transporter, having only 25 % amino acid sequence identity to the closest characterized homolog, being a sucrose permease from *Arabidopsis thaliana* (uniprot: Q9FG00), and an intracellular GH42 putative β-galactosidase, as predicted by SignalP4.0 (28). Interestingly, GH42 enzymes have only been reported active on β-galactoside linkages (29) suggesting a novel specificity of GH42 for the β-glucoside gentiobiose.

A similar gene organization was observed for cluster B, induced by GOS, encoding a GH2 β-galactosidase and an MFS class transporter with homology to a lactose transporter in *B. longum* NCC2705 (30). Likewise cluster C was upregulated by GOS and encoded a heterodimeric ABC transporter permease, a solute binding protein and a GH42 putative β-galactosidase, indicating functional divergence of GH42 enzymes within *B. lactis* Bl-04.

Xylobiose and XOS induced locus D encoding an ABC transporter and one putative β-xylosidase (Balac_0517) and two putative arabinofuranosidases (Balac_0512 and Balac_0520) all three belonging to GH43 and annotated by homology to previously characterized GH43 enzymes (31, 32) suggesting that this gene cluster mediates the transport and hydrolysis of XOS or arabino-xylooligosacchardies into to α-L-arabinofuranose and β-D-xylopyranose for isomerization and phosphorylation into xylulose-5-phosphate for entry into the bifid-shunt (33). Notably also two putative carbohydrate-esterases (Balac_0518 and Balac_0519) were upregulated indicating an additional functionality to process acetylated xylan fragments.

Cluster E showed structural resemblance to a previously identified maltose operons from *B. longum* NCC2705 but differing by the encoded GH13 glycoside hydrolases from those gene clusters identified previously (34-36).

A *multiple sugar metabolism* ABC transporter was identified in cluster F and was induced by both the raffinose family oligosaccharides (RFO) melibiose, raffinose and stachyose representing α-1,6 linked galactosides and isomaltose together with panose representing α-1,6 linked glucosides. As expected, a GH36 α-galactosidase belonging to subfamily I within GH36 conferred the hydrolysis of the transported α-1,6 linked galactosides (37), while a GH13 oligo-α-1,6-glucosidase was responsible for the hydrolysis of α-1,6 linked glucosides. *B. lactis* Bl-04 encoded a total of three GH36 yet the remaining two (Balac_1537 and Balac_1596) were not found to be differentially expressed, indicating alternative functionalities.

In summary, all proposed pathways deduced from the identified gene clusters are shown in figure 7, where potential prebiotic oligosaccharides are internalized and hydrolyzed into products that can readily be further metabolized by the Bifid shunt pathway (33). Notably, a single putative phosphoketalase gene was encoded in *B. lactis* Bl-04, suggesting how this gene product could phosphorylase both fructose-6P and xylulose-5P as the initial step of the bifid shunt, as previously described within *B. lactis* (38).

Figure 7

Discussion

Bifidobacteria have been shown to exert a positive impact on the human gut (39) and may selective utilize oligosaccharides of plant derived prebiotics (40). Despite significant advances in bacterial genomics, understanding of carbohydrate uptake and catabolism mechanisms remain elusive by poor overall annotation of mainly oligosaccharide transporters.

The catabolic adaption potential of *B. lactis* Bl-04 became clear from analysis of the global comparison of prebiotic induced gene expression by principal component analysis (figure 2). The changed global gene expression by the type of glycoside catabolism (galactoside, glucoside or xyloside) was not influenced by the differentially expressed gene clusters involved with prebiotic uptake and catabolism and it is likely that the global expressions, induced by carbohydrate source, involves modulation of the metabolic equilibrium within the bacterium, as it was observed for *B. longum* regarding glycoside induced changes in exopolysaccharide production (41) and pathogenic prevention by acidification when metabolizing fructose rather than glucose (42), showing how the type of glycoside for bifidobacteria can change the overall behavior and potentially probiotic functionality in the GIT. Building on this and the importance for selective utilization of oligosaccharides, we hypothesized the importance of ABC transporters for prebiotic uptake. Interestingly a sole oligosaccharide ABC transporter specific ATP binding protein was found to be constitutively highly expressed suggesting how the single ATP binding protein energizes the multiple identified oligosaccharide ABC transporters, as previously described (43), for readily adaption for utilization to changes in oligosaccharides availability in the GIT. Likewise, various genes encoding proteins linked to proposed probiotic mechanism of actions were found to be highly expressed constantly supporting the probiotic and GIT adapted nature of *B. lactis* Bl-04.

Analysis of the differentially expressed genes involved with prebiotic utilization of *B. lactis* Bl-04 revealed upregulation of explicit gene clusters under transcriptional regulation, as also observed from previously studies of oligosaccharide utilization in probiotic bacteria (30, 44). The uptake of oligosaccharides was facilitated by ABC and MFS types of oligosaccharide transporters, by the lack of PTS permeases in the BL04 genome (22), all associated with glycoside hydrolases.

To differentiate the functionality of ABC and MFS transporters, the putative α-helical topology of the membrane spanning domains of all predicted oligosaccharide transporters in *B. lactis* Bl-04 was mapped (Table 3). Notably it was found how the gentiobiose specific MFS transporter (Balac_0054) was lacking a transmembrane helix,

indicating structural-functional divergence from homologous previously identified MFS permeases (45). Furthermore, one permease protein (Balac_1570) constituting part of the maltotriose upregulated ABC transporter was found to be N-terminally truncated and lacking two helices implicated in heterodimer-formation of the permease domain of the maltose ABC transporter from *Escherichia coli* (46). Comparison to an additional putative *B. lactis* Bl-04 maltose transporter (Balac_1562 – 1564) and experimentally verified maltose ABC transporters (*Lactobacillus casei* (35) and *Streptococcus pneumoniae* (47)) showed how they all harbored the additional two α-helical domain, indicating the maltotriose ABC transporter (Balac_1569, 1570 and 1572) to differ from known maltose ABC transporters by the topology of the permease heterodimer.

In perspective of oligosaccharide utilization, four of the five *in silico* annotated ABC transporters (22), were found to be differentially upregulated while the last putative maltose ABC transporter was found to be constitutively expressed to a comparable level. To elaborate on these findings, and build on the novel specificities and broadness proposed for uptake by ABC transporters facilitated by the solute binding proteins, determining the specificity of ABC transporters (48). The phylogenetics of the ABC transporters solute binding proteins were analyzed and compared them to known protein homologs (Table 5) identified from bifidobacteria and pathogenic GIT associated bacteria (figure 8), hence displaying the functional and taxonomical distribution of oligosaccharide solute binding proteins.

Figure 8

The analysis showed clustering of solute binding proteins identified in the current study with other functionally characterized counter parts. Analysis of each subcluster revealed the taxonomical distribution of functional protein homologs, reflecting how evolutionary adaptation for uptake of lacto-*N*-biose within cluster LacN is close to exclusively found within bifidobacteria while XOS solute binding protein homologs where dominated by soil bacteria and few GIT associated bacteria mainly Actinobacteria, suggesting the XOS utilization of bifidobacteria

to oriented towards dietary plant material as a metabolic niche within the GIT benefitting from xylan utilizing commensal bacteria (49).

From the phylogenetic analysis, a subset of maltose ABC transporters were found to have undergone a convergent evolution, as suggested above for the maltotriose upregulated ABC transporter (Balac_1569, 1570 and 1572). Traditional maltose solute binding protein homologs was found to be widespread by the previously characterized binding proteins (clusters Mal1-4), where a taxanomical subclustering was observed, yet an additional subcluster, Mal5, of maltose binding proteins were identified and represented by the maltotriose upregulated binding protein (Balac_1572). Notably the associated transporter was found to be lacking the additional two α-helical domain of the permease (Table 3) supporting the proposed convergent nature of this type of maltose transport and linking the topological homology to the branches of raffinose and XOS type binding proteins (Figure 8).

Identification of single genes being differentially upregulated by specific oligosaccharides revealed novel protein substrate specificities as compared to the initial gene annotation of the hydrolytic capabilities of *B. lactic* Bl-04 (22). Interestingly, the observation of a GH42 β-galactosidase being induced by the β-1,6-glucoside gentiobiose was intriguing. The GH42 family is only comprised of β-galactosidases so the identification indicates a novel specificity of GH42 which is substantiated by low sequences identity (30%) to any characterized GH42. Additionally the GH42 was co-induced with a MFS type carbohydrate permease with a weak similarity to a plant sucrose permease (50)), suggesting a novel pathway for gentiobiose uptake and catabolism.

No putative glycoside hydrolase was differentially up-regulated on cellobiose however genome mining of *B. lactis* Bl-04 identified a GH1 β-glucosidase (Balac_0151) being constitutively expressed. The β-glucosidase displayed 51% amino acid identity to a GH1 β-glucosidase found to catalyze cellobiose and cellodextrin hydrolysis transported by ABC transporter in *B. brevis* UCC2003 (51) supporting the involvement of the *B. lactis* Bl-04 GH1 for hydrolysis of cellobiose. Furthermore the only transporter differentially up-regulated on cellobiose

was the above ABC transporter also up-regulated by maltotriose (Figure 4) indicating a potential dual specificity of the transporter indicated to have evolved from multiple sugar metabolism-types of oligosaccharide ABC transporters (Figure 8) showing that even without a dedicated cellobiose transporter *B. lactis* Bl04 was evolve a capability to partially transport cellobiose.

The uptake and catabolism of XOS within bifidobacteria was recently proposed (31, 52). Comparative genomic of genes involved with XOS utilization within bifidobacteria (Figure 9) reflected a core gene structure of the XOS ABC transporter with a β-1,4-xylosidase (Balac_0517), while the occurrence of arabino-furanosidases and xylanases of GH8 and GH120 suggested more specie and strain specific adaption.

Figure 9.

Putative oligosaccharide esterases, distantly related to previously identified xylan acetyl esterases (53), were found to be upregulated by XOS and xylobiose in *B. lactis* Bl-04, and conserved among bifidobacteria indicating a specialized mechanism for de-esterification of xylan fragments transported into Bifidobacteria suggesting uptake of both oligomeric, arabinoside substitued and diversely esterified xylosides.

The ABC transporter driven uptake of GOS coupled with co-induction of a GH42 showed homology to a *B. longum* NCC2705 gene cluster upregulated by lactose (30). Interestingly this gene cluster diverges from those identified for human milk oligosaccharide uptake (54) both by the similarity of the associated SBP (Figure 8, LacN *vs* GOS) and the GH encoded in the gene clusters (GH42 versus GH112), indicating how *B. lactis* Bl-04 has evolved a broad oligosaccharide utilization profile for potential prebiotics.

In conclusion, the overall global gene expression was found to be dependent of the type of glycoside utilized (galactosides, glucosides or xylosides) potentially linking the prebiotic catabolism of the bacteria to the overall behavior in the GIT. From the transcriptional analyses we identified the genetic loci within *B. lactis* Bl-04 encoding MFS and ABC transporters with co-occurring glycoside hydrolases for utilization of potential prebiotic

oligosaccharides of α- and β-linkages and varying glycosides composition. This further establishes *B. lactis* Bl-04 as a probiotic bacterium with potential for supplementation with novel prebiotics for increased bifidogenic effects.

Materials and methods

Culture preparation

B. animalis subsp. *lactis* Bl-04 (ATCC SD5219) originally isolated from a human fecal sample (22). Cultures prepared for transcriptional analysis were propagated in 0.22 μm filtered LABSEM media (55) pretreated by the Hungate method for oxygen removal (56). The media was supplemented with 1% (w/v) of the 12 tested carbohydrates (Table 1) and each culture was transferred for five passages, under anaerobic conditions, on each carbohydrate before being harvested in the early logarithmic growth phase (OD_{600}=0.3–0.5) by centrifugation at 4 °C (3.000 g for 15 min) and flash freezing of the cell pellet for storage.

RNA isolation and hybridization setup

Cells were mechanically disrupted by beadbeating and total RNA isolated using Trizol-chloroform extraction (Invitrogen, Carlsbad, CA). Genomic DNA was removed with Turbo DNAse (Ambion, Austin, TX), followed by RNA purification using a RNeasy Mini Kit (Qiagen Inc., Valencia, CA) (57).

Reverse transcription of total RNA, fragmentation and 3' biotin labeling of cDNA was done using 10 μg of total RNA in duplicates for each of the 12 conditions and performed using the Affymetrix GeneChip® system (Affymetrix, Santa Clara, CA). Total RNA was reverse transcribed using random primers and SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, California) and purification of cDNA was done using MinElute PCR Purification kit using a final elution volume of 12 μL (QIAGEN, Inc., Valencia, CA). Following cDNA fragmentation

into 50–100 bp was done using DNase I (GE Healthcare, Waukesha, WI) and biotin-labeling done using GeneChip DNA labeling reagent (Affymetrix) and Terminal deoxynucleotidyl transferase (Promega, Madison, WI).

Labeled cDNA fragments were hybridized at Utah State University using Affymetrix custom-made chips. All extracted data was imported into SAS JMP Genomics (SAS Institute Inc, Cary, NC) before being quantile normalized and modeled using a one-way ANOVA for identification of differentially upregulated genes using a threshold value of α=0.005 and Bonferroni correction.

Real-time quantitative PCR validation of microarray

Real-time quantitative PCR was performed on five selected genes (Table 3) found to be differentially upregulated. The DNAse treated total RNA, identical to the RNA used in microarray sample preparation, was used as template for each of the above 12 growth conditions, measured in triplicates. Experiments were conducted with a QRT-PCR thermal cycler (I-cycler; Bio-Rad, Hercules, CA) in combination with the QuantiTect SYBR Green PCR kit (Qiagen).

Construction of phylogenetic tree of carbohydrate solute binding proteins

The sequence dataset was compiled from oligosaccharide binding proteins all identified from previous work or identified from the current project (Table 5). Sequence homologs for each protein entry was identified by BLAST (58) and restricted to either 100 hits or an e-value of 10^{-3} against the non-redundant database. All redundant sequences were removed and the remaining sequences together with a monosaccharide (fructose) binding proteins were aligned using ClustalX (59) using the Blosum series substitution matrix and a gap opening penalty of 2, compared to the standard penalty of 10. The resulting phylogenetic tree file was visualized using Dendroscope (60). Bootstrap values were calculated by ClustalX using standard conditions (1000 iterations).

Microarray submission

All raw data have been deposited in the GEO database and complies with the MIAME guidelines.

Acknowledgements:

This research was funded by DuPont Nutrition and Health, North Carolina Dairy Foundation and the FøSu grant from the Danish Strategic Research Council to the project "Gene discovery and molecular interactions in prebiotics/probiotics systems. Focus on carbohydrate prebiotics". J.M.A. is funded by a joint Ph.D. stipend from DuPont, the FøSu grant, and the Technical University of Denmark.

Reference list:

1. Reid G*, et al* (2003) New scientific paradigms for probiotics and prebiotics*. J Clin Gastroenterol* 37: 105–118.

2. Turroni F, van Sinderen D & Ventura M (2011) Genomics and ecological overview of the genus *Bifidobacterium. Int J Food Microbiol* 149: 37–44.

3. Szajewska H, Guandalini S, Morelli L, Van Goudoever JB & Walker A (2010) Effect of *Bifidobacterium animalis* subsp. *lactis* supplementation in preterm infants: A systematic review of randomized controlled trials*. J Pediatr Gastroenterol Nutr* 51: 203–209.

4. Ganguli K & Walker WA (2011) Probiotics in the prevention of necrotizing enterocolitis*. J Clin Gastroenterol* 45: S133–S138.

5. Steed H*, et al* (2010) Clinical trial: The microbiological and immunological effects of synbiotic consumption - a randomized double-blind placebo-controlled study in active crohn's disease*. Aliment Pharmacol Ther* 32: 872–883.

6. Ouwehand AC*, et al* (2008) *Bifidobacterium* microbiota and parameters of immune function in elderly subjects*. FEMS Immunol Med Microbiol* 53: 18–25.

7. Ventura M*, et al* (2009) Genome-scale analyses of health-promoting bacteria: Probiogenomics*. Nat Rev Microbiol* 7: 61–71.

8. Bottacini F*, et al* (2010) Comparative genomics of the genus *Bifidobacterium. Microbiology* 156: 3243–3254.

9. Turroni F*, et al* (2010) Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-derived glycan foraging*. Proc Natl Acad Sci U S A* 107: 19514–19519.

10. Schell Ma*, et al* (2002) The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract*. Proc Natl Acad Sci U S A* 99: 14422–14427.

11. Davis LM, Martinez I, Walter J, Goin C & Hutkins RW (2011) Barcoded pyrosequencing reveals that consumption of galactooligosaccharides results in a highly specific bifidogenic response in humans*. PLoS One* 6: e25200.

12. Roberfroid M (2007) Prebiotics: The concept revisited*. J Nutr* 137: S830–S837.

13. Roberfroid M*, et al* (2010) Prebiotic effects: Metabolic and health benefits*. Br J Nutr* 104 Suppl 2: S1–S63.

14. Gibson GR, Probert HM, Loo JV, Rastall RA & Roberfroid MB (2004) Dietary modulation of the human colonic microbiota: Updating the concept of prebiotics*. Nutr Res Rev* 17: 259–275.

15. Makelainen H, Hasselwander O, Rautonen N & Ouwehand AC (2009) Panose, a new prebiotic candidate*. Lett Appl Microbiol* 49: 666–672.

16. van den Broek LA, Hinz SW, Beldman G, Vincken JP & Voragen AG (2008) *Bifidobacterium* carbohydrases-their role in breakdown and synthesis of (potential) prebiotics. *Mol Nutr Food Res* 52: 146–163.

17. Pokusaeva K, Fitzgerald GF & van Sinderen D (2011) Carbohydrate metabolism in bifidobacteria. *Genes & Nutrition* 6: 285–306.

18. Turroni F, van Sinderen D & Ventura M (2011) Genomics and ecological overview of the genus *Bifidobacterium. Int J Food Microbiol* 149: 37–44.

19. Paineau D*, et al* (2008) Effects of seven potential probiotic strains on specific immune responses in healthy adults: A double-blind, randomized, controlled trial. *FEMS Immunol Med Microbiol* 53: 107–113.

20. Ouwehand AC*, et al* (2009) Specific probiotics alleviate allergic rhinitis during the birch pollen season. *World J Gastroenterol* 15: 3261–3268.

21. Bartosch S, Woodmansey EJ, Paterson JC, McMurdo ME & Macfarlane GT (2005) Microbiological effects of consuming a synbiotic containing *Bifidobacterium bifidum*, *Bifidobacterium lactis*, and oligofructose in elderly persons, determined by real-time polymerase chain reaction and counting of viable bacteria. *Clin Infect Dis* 40: 28–37.

22. Barrangou R*, et al* (2009) Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and Bl-04. *J Bacteriol* 191: 4144–4151.

23. Gueimonde M*, et al* (2004) Viability and diversity of probiotic *Lactobacillus* and *Bifidobacterium* populations included in commercial fermented milks. *Food Res Int* 37: 839–850.

24. Tatusov RL*, et al* (2003) The COG database: An updated version includes eukaryotes. *BMC Bioinformatics* 4: 41.

25. Horvath P & Barrangou R (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327: 167–170.

26. Gilad O*, et al* (2012) Insights into physiological traits of *Bifidobacterium animalis* subsp. *lactis* BB-12 through membrane proteome analysis. *Journal of Proteomics* 75: 1190–1200.

27. Cantarel BL*, et al* (2009) The carbohydrate-active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res* 37: 233–238.

28. Petersen TN, Brunak S, von Heijne G & Nielsen H (2011) SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat Methods* 8: 785–786.

29. Shipkowski S & Brenchley JE (2006) Bioinformatic, genetic, and biochemical evidence that some glycoside hydrolase family 42 β-galactosidases are arabinogalactan type I oligomer hydrolases. *Appl Environ Microbiol* 72: 7730–7738.

30. Parche S*, et al* (2007) Sugar transport systems of *Bifidobacterium longum* NCC2705. *J Mol Microbiol Biotechnol* 12: 9–19.

31. Lagaert S*, et al* (2010) Substrate specificity of three recombinant α-L-arabinofuranosidases from *Bifidobacterium adolescentis* and their divergent action on arabinoxylan and arabinoxylan oligosaccharides*. Biochem Biophys Res Commun* 402: 644–650.

32. Lagaert S*, et al* (2007) Recombinant expression and characterization of a reducing-end xylose-releasing exo-oligoxylanase from *Bifidobacterium adolescentis. Appl Environ Microbiol* 73: 5374–5377.

33. Liu D*, et al* (2011) Proteomics analysis of *Bifidobacterium longum* NCC2705 growing on glucose, fructose, mannose, xylose, ribose, and galactose*. Proteomics* 11: 2628–2638.

34. Nakai H*, et al* (2009) The maltodextrin transport system and metabolism in *Lactobacillus acidophilus* NCFM and production of novel α-glucosides through reverse phosphorolysis by maltose phosphorylase*. FEBS J* 276: 7353–7365.

35. Monedero V, Yebra MJ, Poncet S & Deutscher J (2008) Maltose transport in *Lactobacillus casei* and its regulation by inducer exclusion*. Res Microbiol* 159: 94–102.

36. Ajdić D & Pham VT (2007) Global transcriptional analysis of *Streptococcus mutans* sugar transporters using microarrays*. J Bacteriol* 189: 5049–5059.

37. Fredslund F*, et al* (2011) Crystal structure of α-galactosidase from *Lactobacillus acidophilus* NCFM: Insight into tetramer formation and substrate binding*. J Mol Biol* 412: 466–480.

38. Meile L, Rohr LM, Geissmann TA, Herensperger M & Teuber M (2001) Characterization of the D-xylulose 5-phosphate/D-fructose 6-phosphate phosphoketolase gene (*xfp*) from *Bifidobacterium lactis. J Bacteriol* 183: 2929–2936.

39. Cani PD & Delzenne NM (2011) The gut microbiome as therapeutic target*. Pharmacol Ther* 130: 202–212.

40. Macfarlane GT, Steed H & Macfarlane S (2008) Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics*. J Appl Microbiol* 104: 305–344.

41. Audy J, Labrie S, Roy D & Lapointe G (2010) Sugar source modulates exopolysaccharide biosynthesis in *Bifidobacterium longum* subsp. *longum* CRC 002*. Microbiology* 156: 653–664.

42. Fukuda S*, et al* (2011) Bifidobacteria can protect from enteropathogenic infection through production of acetate*. Nature* 469: 543–547.

43. Marion C, Aten AE, Woodiga Sa & King SJ (2011) Identification of an ATPase, MsmK, which energizes multiple carbohydrate ABC transporters in *Streptococcus pneumoniae. Infect Immun* 79: 4193–4200.

44. Barrangou R*, et al* (2006) Global analysis of carbohydrate utilization by *Lactobacillus acidophilus* using cDNA microarrays*. Proc Natl Acad Sci U S A* 103: 3816–3821.

45. Law CJ, Maloney PC & Wang DN (2008) Ins and outs of major facilitator superfamily antiporters*. Annu Rev Microbiol* 62: 289–305.

46. Oldham ML, Khare D, Quiocho Fa, Davidson AL & Chen J (2007) Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* 450: 515–521.

47. Espinosa M & Puyet A (1997) The maltose / maltodextrin regulon of *Streptococcus pneumoniae. Biochemistry (N Y)* 272: 30860–30865.

48. Berntsson RP, Smits SHJ, Schmitt L, Slotboom D & Poolman B (2010) A structural classification of substrate-binding proteins. *FEBS Lett* 584: 2606–2617.

49. Martens EC*, et al* (2011) Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS Biol* 9: e1001221.

50. Juergensen K*, et al* (2003) The companion cell-specific *Arabidopsis* disaccharide carrier AtSUC2 is expressed in nematode-induced syncytia. *Plant Physiol* 131: 61–69.

51. Pokusaeva K, Connell-motherway M, Zomer A, Fitzgerald GF & Sinderen DV (2011) Cellodextrin utilization by *Bifidobacterium breve* UCC2003. 77: 1681–1690.

52. Gilad O*, et al* (2010) Combined transcriptome and proteome analysis of *Bifidobacterium animalis* subsp. *lactis* BB-12 grown on xylo-oligosaccharides and a model of their utilization. *Appl Environ Microbiol* 76: 7285–7291.

53. Kabel Ma*, et al* (2011) Biochemical characterization and relative expression levels of multiple carbohydrate esterases of the xylanolytic rumen bacterium *Prevotella ruminicola* 23 grown on an ester-enriched substrate. *Appl Environ Microbiol* 77: 5671–5681.

54. Fushinobu S (2010) Unique sugar metabolic pathways of bifidobacteria. *Biosci Biotechnol Biochem* 74: 2374–2384.

55. Barrangou R, Altermann E, Hutkins R, Cano R & Klaenhammer TR (2003) Functional and comparative genomic analyses of an operon involved in fructooligosaccharide utilization by *Lactobacillus acidophilus. Proc Natl Acad Sci U S A* 100: 8957–8962.

56. Daniels L & Zeikus JG (1975) Improved culture flask for obligate anaerobes. *Appl Microbiol* 29: 710–711.

57. Goh YJ*, et al* (2009) Development and application of a *upp*-based counterselective gene replacement system for the study of the S-layer protein SlpX of *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol* 75: 3093–3105.

58. Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410.

59. Larkin MA*, et al* (2007) Clustal W and clustal X version 2.0. *Bioinformatics* 23: 2947–2948.

60. Huson DH*, et al* (2007) Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 8: 460.

61. Wei X*, et al* (2012) Fructose uptake in *Bifidobacterium longum* NCC2705 is mediated by an ATP-binding cassette transporter*. The Journal of Biological Chemistry* 287: 357–367.

62. Andersen JM*, et al* (2011) Transcriptional and functional analysis of galactooligosaccharide uptake by *lacS* in *Lactobacillus acidophilus. Proc Natl Acad Sci U S A* 108: 17785–17790.

63. Kall L, Krogh A & Sonnhammer EL (2007) Advantages of combined transmembrane topology and signal peptide prediction - the phobius web server*. Nucleic Acids Res* 35: 429–432.

64. Gopal S*, et al* (2010) Maltose and maltodextrin utilization by *Listeria monocytogenes* depend on an inducible ABC transporter which is repressed by glucose*. PLoS ONE* 5: e10349.

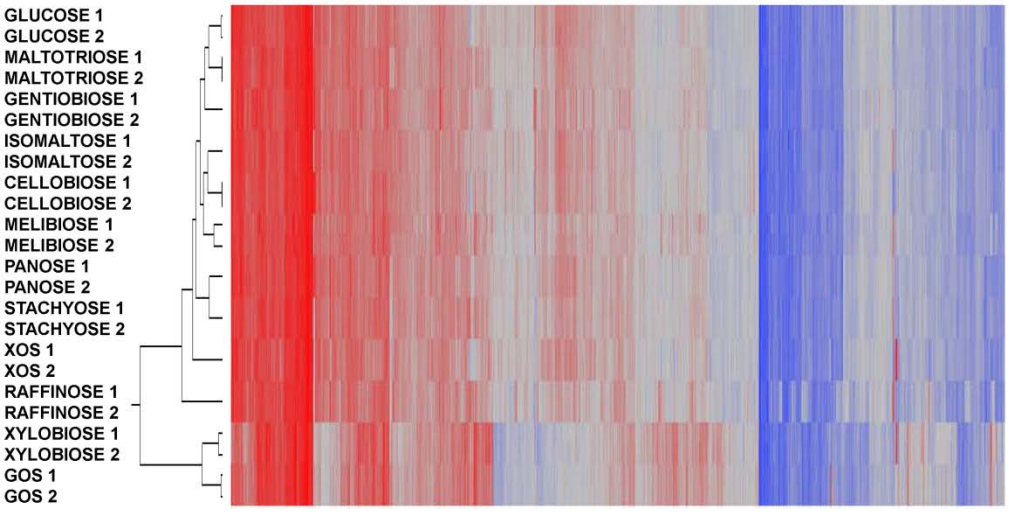
65. Abbott DW*, et al* (2010) The molecular basis of glycogen breakdown and transport in *Streptococcus pneumoniae. Mol Microbiol* 77: 183–199.

66. Schanert S*, et al* (2006) Maltose and maltodextrin utilization by bacillus subtilis*. J Bacteriol* 188: 3911–3922.

67. Garrido D, Kim JH, German JB, Raybould HE & Mills DA (2011) Oligosaccharide binding proteins from *Bifidobacterium longum* subsp. *infantis* reveal a preference for host glycans*. PLoS One* 6: e17315.

68. Suzuki R*, et al* (2008) Structural and thermodynamic analyses of solute-binding protein from *Bifidobacterium longum* specific for core 1 disaccharide and lacto-N-biose I*. J Biol Chem* 283: 13165–13173.

Figures:



Category	Count	%
Information storage and processing		
J	60	36.8
K	7	4.3
L	2	1.2
Cellular processes and signaling		
D	3	1.8
T	2	1.2
M	3	1.8
O	11	6.7
Metabolism		
C	14	8.6
G	12	7.4
E	6	3.7
F	3	1.8
I	1	0.6
P	2	1.2
Poorly characterized		
R	4	2.5
S	10	6.1
Unknown	23	14.1

Figure 1: Two-way clustering of the global gene expression profile and COG distribution of constitutively expressed genes. Gene expression intensities are represented by red coloring: up-regulation, blue coloring: down-regulation. Technical replicates for each carbohydrate are numbered and showed overall high reproducibility. The highest expressed decile of the global transcriptome was assigned COG categories, highlighting the essential metabolism of *B. lactis Bl*-04.

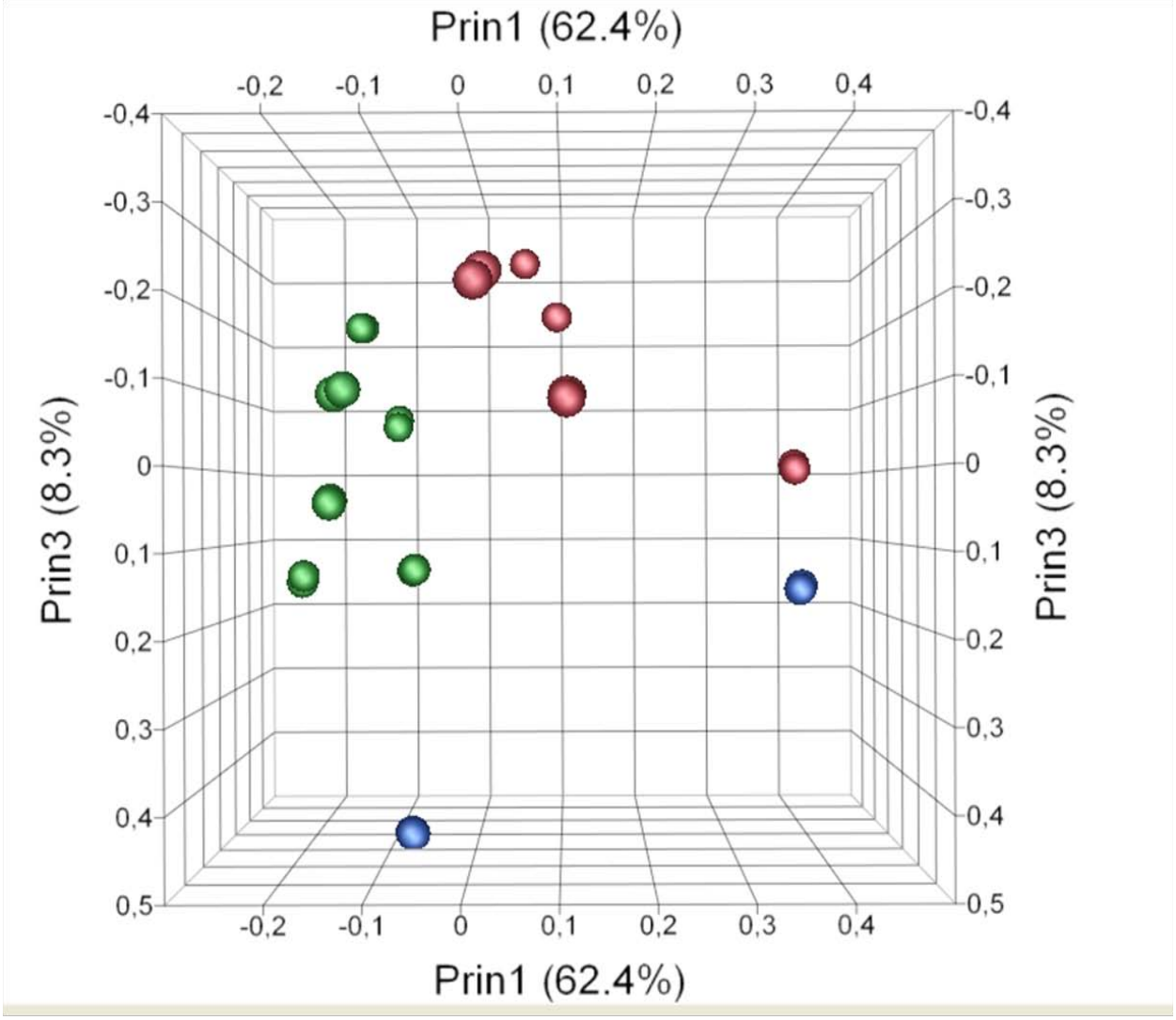


Figure 2: Principal component analysis of the global transcriptome for *B. lactis* BL-04 cultivated with potential prebiotics showing a profound differentiation of the global gene expression profile depending on the type of carbohydrates utilized. Galactosides in red (GOS, melibiose, raffinose, stachyose), glucosides in green (cellobiose, Gentiobiose, glucose, maltotriose, isomaltose, panose) and xylosides in blue (XOS, xylobiose).

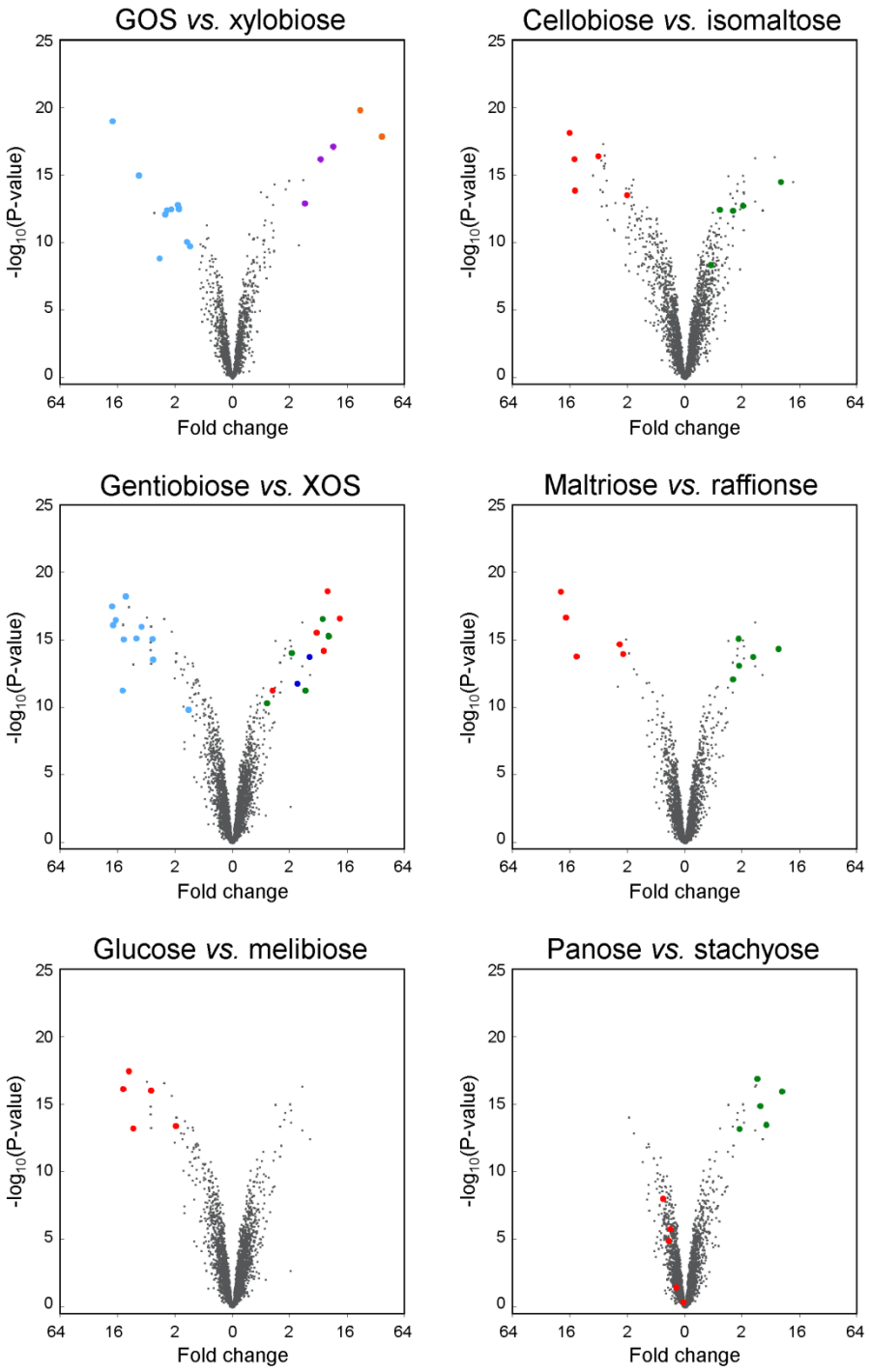


Figure 3: Representative volcano plots of pairwise comparison of oligosaccharide induced differential global transcriptome within *B. animalis* subsp *lactis* BL-04. All genes are shown by solid grey circles, and putative carbohydrate active protein encoding genes being statistically significant up-regulated are highlighted with solid circles and color-coded by gene cluster as by table 2.

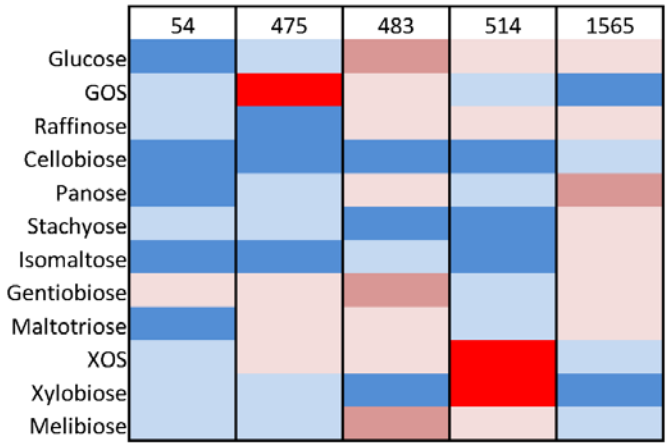


Figure 4: Heatmap representation of rt-qPCR validation of gene expression values. The gene expression value of mRNA quantified for each of the five genes to each of the 12 growth conditions have been color-coded as the relative fold upregulation to the lowest value measured to each gene: Blue (1-2 fold), light blue (2-4), light red (4-8), red (8-16) and strong red (>16).

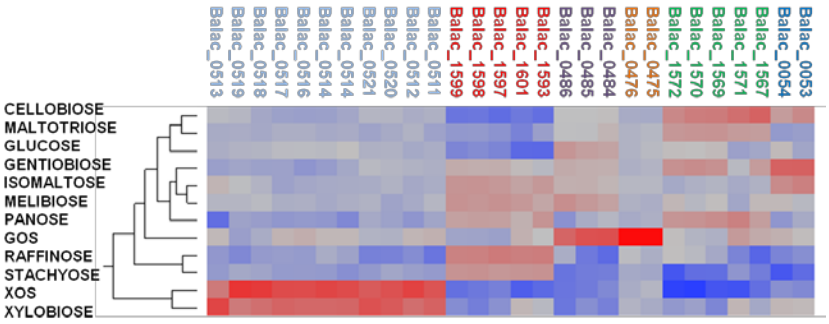


Figure 5: Two-way clustering of the expression profile for genes identified to be differentially upregulated by ANOVA. The coloring of each ORF corresponds to the gene cluster of table 2.

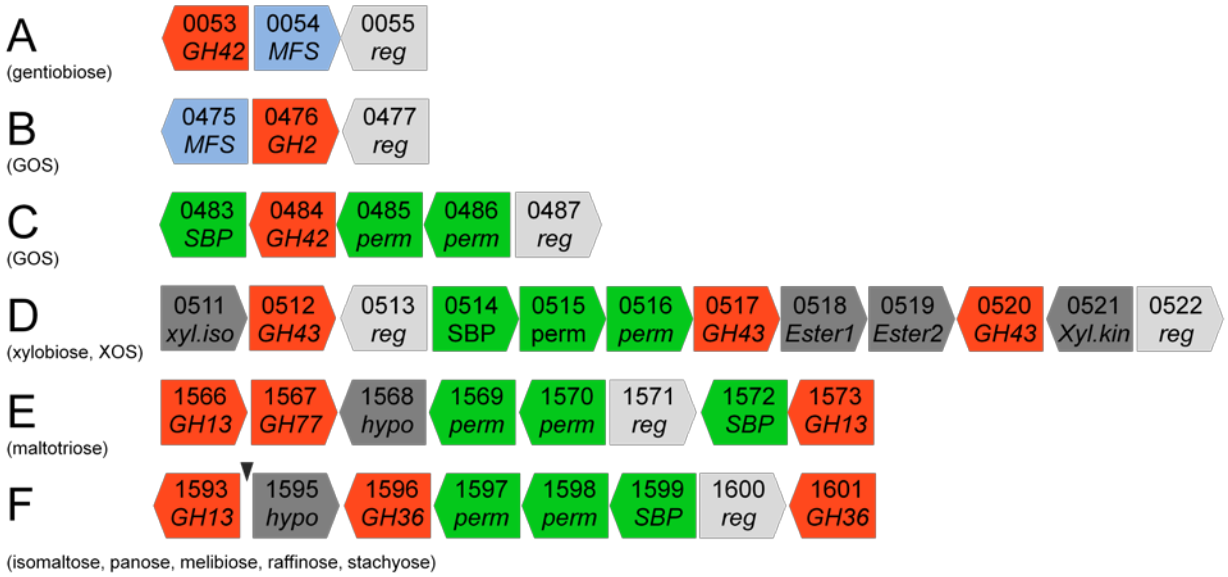


Figure 6: Organization of differentially expressed gene clusters encoding proteins predicted to be involved with prebiotic utilization. Genes are listed with locustag and gene functions are colored as: Glycoside hydrolases in red, ABC transporter solute binding proteins (SBP) and permeases (perm) in green, MFS transporters in blue, transcriptional regulators (reg) in light grey and hypothetical proteins (hypo), carbohydrate esterases (ester1 and ester2), xylose isomerase (xyl.iso) and xylulose kinase (xyl.kin) all in dark gray. Short putative nonfunctional ORF are highlighted by black triangles.

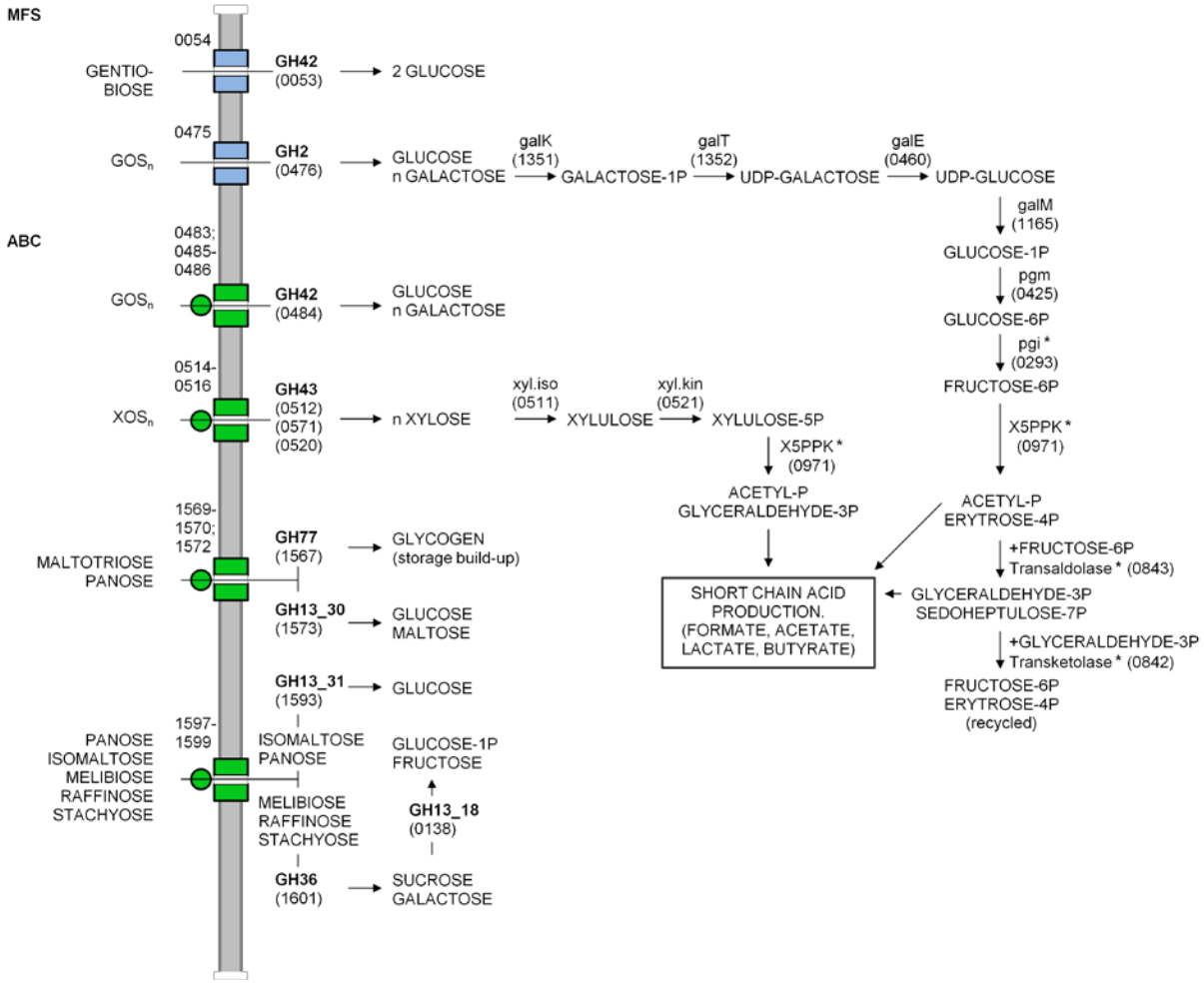


Figure 7: Proposed pathways for oligosaccharide uptake and catabolism in monosaccharides for entry into the bifid shunt. Transporters are colored as in figure 5 and all genes are given by their locustag. The schematic pathways for glucose (entering as glucose-1P), galactose, fructose (entering as fructose-6P) and xylose are shown with the main steps of the bifid shunt. All constitutive highly expressed genes (Figure 1) are denoted with an asterix (*****).

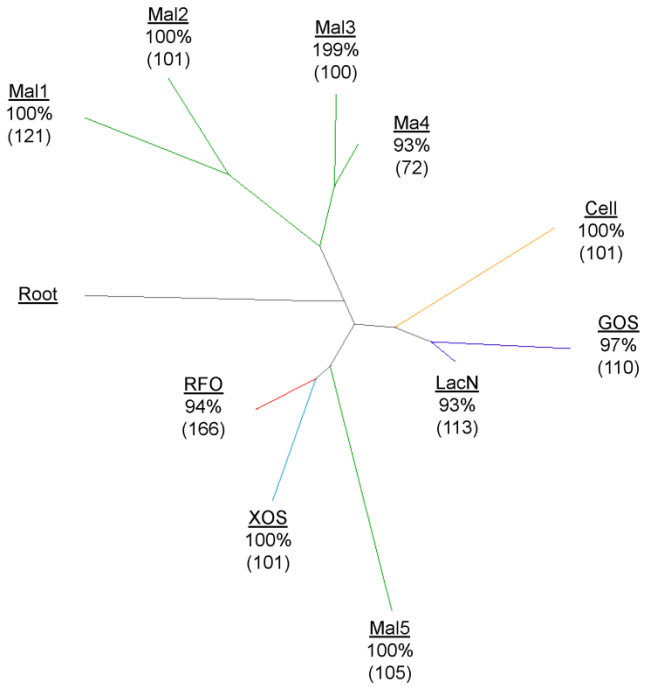


Figure 8: Functional comparison of the identified oligosaccharide solute binding proteins of ABC transporters in *B. lactis* BL-04. The phylogenetic tree was rooted by a characterized fructose binding protein (61) as a functional and structural out group of oligosaccharide binding proteins (48). Sub-clusters were defined by bootstrap values in percentages and the characterized solute binding protein(s) identified within each sub-cluster where the numbers of sequences are listed in brackets. The tree was colored by substrate specificity (maltose binding proteins (green), cellodextrins (orange), GOS and lacto-N-biose (blue), XOS (light blue) and raffinose family oligosaccharides (red)) and sub-clusters were denoted by numbers as given in table 4.

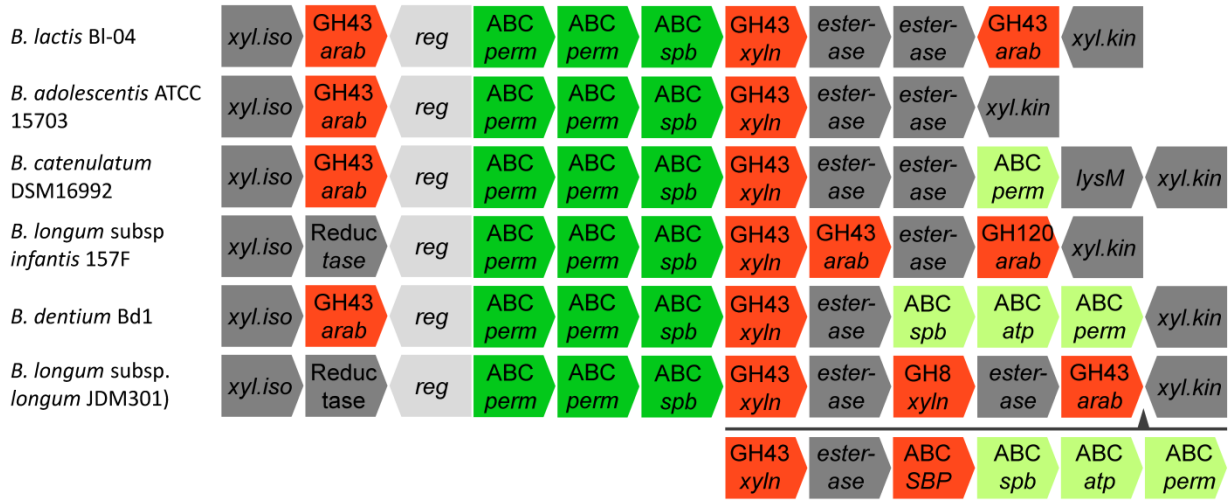


Figure 9: Genomic content and organization of XOS utilization gene clusters identified within bifidobacteria. All strains were ordered top down by highest sequence similarity of the XOS binding protein to the XOS binding protein of *B. lactis* Bl-04 (balac_0514). Gene functions are colored as: Glycoside hydrolases (red), XOS ABC transporters (green), xylose ABC transporters (light green), transcriptional regulators (light grey), and putative XOS esterases, xylose isomerases (xyl.iso), xylulose kinases (xyl.kin) alcohol dehydrogenases (alcohol dehydro.) and a putative secreted amidase (lysM) (all dark gray). Short putative nonfunctional ORF are highlighted by black triangles. All GH43 enzymes were annotated and differentiated by protein similarity to previously characterized xylosidases (xyln) or arabinofuranosidases (arab) together with the GH8 enzyme (31, 32).

Tables:

Table 1: Carbohydrates, used for DNA microarray cultures, listed with glycoside structure and type, supplier and purity

Carbohydrate	Structure[1]	Glycoside type	DP [2]	Manufacturer or supplier	Purity (as given by Manufacturer or supplier)
Glucose	Glc*p*	glucoside	1	Sigma	> 99%
GOS	[β-D-Gal*p*-(1–4)]$_n$-D-Glc*p*	galactoside	2–6	Dupont	> 94% DP $\geq$ 2
Melibiose	α-D-Gal*p*-(1–6)- D-Glc*p*	galactoside	2	Sigma	> 98%
Raffinose	α-D-Gal*p*-(1–6)- D-Glc*p*-(α1,β2)-DFru*f*	galactoside	3	Sigma	> 99%
Stachyose	[α-D-Gal*p*-(1–6)]$_2$- D-Glc*p*-(α1,β2)- D-Fru*f*	galactoside	4	Sigma	> 98%
Isomaltose	α-D-Glc*p*-(1–6)-D-Glc*p*	glucoside	2	Sigma	> 98%
Panose	α-D-Glc*p*(1–6)-α-D-Glc*p*-(1–4)-D-Glc*p*	glucoside	3	Sigma	> 98%
Maltotriose	α-D-Glc*p*-(1–4)-α-D-Glc*p*-(1–4)-D-Glc*p*	glucoside	3	Dupont	> 95%
Cellobiose	β-D-Glc*p*-(1–4)-D-Glc*p*	glucoside	2	Fluka AG	> 99%
Gentiobiose	β-D-Glc*p*-(1–6)-D-Glc*p*	glucoside	2	Sigma	> 98%
Xylobiose	β- D-xyl*f*-(1–4)- D-xyl*f*	xyloside	2	Dupont	> 95%
XOS	[β- D-xyl*f*-(1–4)]$_m$-D-xyl*f*	xyloside	2-7	Shandong Longlive Bio-technology Co., Ltd, (China)	> 90% [3]

[1] n=1-5 as previously described (62), m =1-6 as stated by manufacturer

[2] Degree of polymerization

3 The XOS composition and purity was previously determined (52)

Table 2: Statistically significant upregulated genes involved in carbohydrate uptake and catabolism. The genes are listed by ascending locus tag numbers. Only the oligosaccharide that elicited the highest induction level is listed for genes that are upregulated by more than one oligosaccharide.

ORF	Gene annotation	Inducing CHO type	Volcano plot (Figure 3)	Highest inducing CHO	Gene cluster (figure 5)	Fold upregulated	$-\log_{10}$(P-value)
Balac_0053	β-galactosidase, GH42	Glucoside	C	gentiobiose	A	6.5	13.8
Balac_0054	MFS permease	Glucoside	C	gentiobiose	A	4.8	11.8
Balac_0475	MFS permease	Galactoside	A	GOS	B	21.8	19.9
Balac_0476	β-galactosidase, GH2	Galactoside	A	GOS	B	36.9	17.9
Balac_0484	β-galactosidase, GH42	Galactoside	A	GOS	C	11.4	17.2
Balac_0485	ABC transporter, permease component	Galactoside	A	GOS	C	8.4	16.2
Balac_0486	ABC transporter, permease component	Galactoside	A	GOS	C	5.7	12.9
Balac_0511	Xylose isomerase	xyloside	A,C	XOS	D	13.8	15.1
Balac_0512	α-L-arabinofuranosidase, GH43	xyloside	A,C	XOS	D	6.8	13.6
Balac_0513	Transcriptional regulator (*lacI* type)	xyloside	A,C	Xylobiose	D	3.0	10.1
Balac_0514	ABC transporter, oligosaccharide-binding protein	xyloside	A,C	XOS	D	9.0	16.0
Balac_0515	ABC transporter, permease component	xyloside	A,C	XOS	D	16.8	16.5
Balac_0516	ABC transporter, permease component	xyloside	A,C	XOS	D	18.3	17.5
Balac_0517	β-xylosidase, GH43	xyloside	A,C	XOS	D	17.9	16.1
Balac_0518	Putative carbohydrate esterase	xyloside	A,C	XOS	D	14.2	11.3
Balac_0519	Esterase	xyloside	A,C	XOS	D	6.9	15.1
Balac_0520	α-L-arabinofuranosidase, GH43	xyloside	A,C	XOS	D	10.2	15.2
Balac_0521	Xylulose kinase	xyloside	A,C	Xylobiose	D	18.2	19.0
Balac_1567	4-α-glucanotransferase	glucoside	B,D	Maltriose	E	9.7	14.5
Balac_1569	ABC transporter, permease component	glucoside	B,D	Cellobiose	E	5.2	13.7
Balac_1570	ABC transporter, permease component	glucoside	B,D	Cellobiose	E	3.7	15.1
Balac_1571	Transcriptional regulator (*lacI* type)	glucoside	B,D	Cellobiose	E	3.7	13.1
Balac_1572	ABC transporter, oligosaccharide-binding protein	glucoside	B,D	Cellobiose	E	3.2	12.1
Balac_1593	oligo-1,6-α-glucosidase, GH13	Galactoside, Glucoside	B,D,E	Isomaltose	F	4.5	14.0
Balac_1597	ABC transporter, permease component	Galactoside, Glucoside	B,D,E	Raffinose	F	14.1	13.9
Balac_1598	ABC transporter, permease component	Galactoside, Glucoside	B,D,E	Isomaltose	F	20.1	18.6
Balac_1599	ABC transporter, oligosaccharide-binding protein	Galactoside, Glucoside	B,D,E	Isomaltose	F	17.8	16.7
Balac_1601	α-galactosidase, GH36	Galactoside, Glucoside	B,D,E	Raffinose	F	8.1	16.4

Table 3: Prediction of α-helical topology within oligosaccharide transporters identified in *B. lactis* Bl-04.

ORF	Function	Class	Predicted TMH[1]	Sequence length (aa)
0054	Gentiobiose	MFS	11	384
0139	Sucrose (putative)	MFS	12	537
0475	GOS	GPH homolog	12	505
1240	FOS (putative)	MFS	12	441
1588	arabinofuranosides (putative)	GPH homolog	12	481
0485	GOS	ABC	6	326
0486	GOS	ABC	6	322
0515	XOS	ABC	6	352
0516	XOS	ABC	6	289
1563	Maltose (putative)	ABC	6	322
1564	Maltose (putative)	ABC	8	457
1569	Maltotriose	ABC	6	278
1570	Maltotriose	ABC	6	284
1597	RFO+IMO	ABC	6	301
1598	RFO+IMO	ABC	6	330

[1] Transmembrane α-helices (TMH) predicted using the Phobius tool (63)

Table 4: Identified clusters of oligosaccharide binding proteins from Figure 7. Clusters are shown by numbers and if possible sub-clusters are listed with letters. The experimentally identified oligosaccharide binding proteins used to generate the tree are listed in the corresponding cluster and sub-cluster if possible.

Cluster	Sub-cluster	Substrate specificity	Identified Organism	Reference
Malto-oligosaccharides	1	α-(1,4)-gluco-oligosaccharides	*Listeria monocytogenes* *Streptococcus pneumoniae*	(64) (65)
	2	β-Cyclodextrin and maltose	*Bacillus subtilis*	(66)
	3	Maltose	*L. casei* BL23	(35)
	4	Putative maltose	*B. animalis* subsp *lactis* Bl-04	This study
	5	Maltose Maltotriose	*B. longum* NCC2705 *B. animalis* subsp *lactis* Bl-04	(30) This study
β-glucosides	-	β-(1,4)-gluco-oligosaccharides	*B. breve* UCC2003	(51)
β-galactosides	A	Lactose β-galacto-oligosaccharides	*B. longum* NCC2705 *B. animalis* subsp *lactis* Bl-04	(30) This study
	B	Lacto-*N*-biose	*B. bifidum* *B. longum*	(67) (68)
XOS	-	β-(1,4)-xylo-oligosaccharides	*B. animalis* subsp *lactis* Bl-04	This study
RFO	A	Raffinose and isomaltose	*Streptococcus mutans*	(36)
	B	Raffinose Raffinose and Isomaltose[1]	*B. longum* NCC2705 *B. animalis* subsp *lactis* Bl-04	(30) This study
Root	-	Fructose	*B. longum* NCC2705	(61)

[1] Including melibiose, panose and stachyose

Table 6: Primer pairs used for real time quantitative PCR.

ORF	Primer 5’ – 3’	Product size (bp)
0054	CACACTCGCTCGAGATTC AGGCCAATCATGCATACG	140
0475	GCTGACGATGGGAATGAC GCTCGACGTGTTCTACTC	160
0483	CGTCGGAGTTCTTGATGG CAGGCAGCCTATGACTTC	142
0514	GGCTGACCTTGGATTCTT CTTCTCGCCCATGTAGTTG	145
1565	GAACGCCGTAGATCTTGC ATGTTCGCCAATGACCAG	148

6.3 Transcriptional and functional analysis of galactooligosaccharide uptake by *lacS* in *Lactobacillus acidophilus*

Published: Andersen et al. *Proc. Natl. Acad. Sci*., 108: 17785–17790 (2011)

Transcriptional and functional analysis of galactooligosaccharide uptake by *lacS* in *Lactobacillus acidophilus*

Joakim M. Andersen[a,b], Rodolphe Barrangou[c], Maher Abou Hachem[a], Sampo Lahtinen[d], Yong Jun Goh[b], Birte Svensson[a], and Todd R. Klaenhammer[b,1]

[a]Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark; [c]Danisco USA Inc., Madison, WI 53716; [d]Health and Nutrition, Danisco BioActives, FI-02460, Kantvik, Finland; and [b]Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State University, Raleigh, NC 27695

Contributed by Todd R. Klaenhammer, September 22, 2011 (sent for review May 30, 2011)

Probiotic microbes rely on their ability to survive in the gastrointestinal tract, adhere to mucosal surfaces, and metabolize available energy sources from dietary compounds, including prebiotics. Genome sequencing projects have proposed models for understanding prebiotic catabolism, but mechanisms remain to be elucidated for many prebiotic substrates. Although β-galactooligosaccharides (GOS) are documented prebiotic compounds, little is known about their utilization by lactobacilli. This study aimed to identify genetic loci in *Lactobacillus acidophilus* NCFM responsible for the transport and catabolism of GOS. Whole-genome oligonucleotide microarrays were used to survey the differential global transcriptome during logarithmic growth of *L. acidophilus* NCFM using GOS or glucose as a sole source of carbohydrate. Within the 16.6-kbp *gal-lac* gene cluster, *lacS*, a galactoside-pentose-hexuronide permease-encoding gene, was up-regulated 5.1-fold in the presence of GOS. In addition, two β-galactosidases, LacA and LacLM, and enzymes in the Leloir pathway were also encoded by genes within this locus and up-regulated by GOS stimulation. Generation of a *lacS*-deficient mutant enabled phenotypic confirmation of the functional LacS permease not only for the utilization of lactose and GOS but also lactitol, suggesting a prominent role of LacS in the metabolism of a broad range of prebiotic β-galactosides, known to selectively modulate the beneficial gut microbiota.

lactose permease | catabolite repression element

Increased interest in the ability of the human microbiota of the gastrointestinal tract (GIT) and selected probiotic microbes to impact health has been supported by expanded documentation on resistance to allergies (1), respiratory tract infections (2), and various gastrointestinal conditions such as ulcerative colitis, irritable bowel syndrome, and inflammatory bowel disease (3). Research on probiotic bacteria (4) continues to accumulate further knowledge about the biological mechanisms of action and complex interplay between gut microbes and host health.

The functional attributes of gut microbes and those delivered as probiotics rely on their ability to survive in the GIT, adhere to mucosal surfaces, and metabolize available energy sources from nondigestible dietary compounds (5). Notably, the ability to selectively use a broad range of potentially prebiotic carbohydrates (6), ranging from oligosaccharides to polysaccharides, provides a competitive advantage to the beneficial microbiota during colonization of the GIT and to transient probiotic microbes (7). Prebiotic oligosaccharides are not absorbed by the host and resist degradation by intestinal acids, bile acids, and digestive enzymes, allowing them to travel through the small intestine and colon, where they may be selectively used by beneficial microbes. Commercial β-galactooligosaccharides (GOS) are typically produced by enzymatic transglycosylation using lactose as substrate (8), to yield a mixed-length galactosylated product with a degree of polymerization (DP) ranging from 2 to 6. The oligomeric nature and β-galactoside linkages allow GOS to be used as prebiotic supplements, notably for stimulation of particular lactobacilli and bifidobacteria (9, 10). Specifically, GOS supplements have been shown to exert positive impacts on intestinal *Bifidobacterium* and *Lactobacillus* populations in infants (11), to mitigate irritated bowel syndrome (12), and to reduce the severity and duration of travelers' diarrhea (13). GOS has also been shown to inhibit pathogenic *Vibrio cholerae* and *Cronobacter sakazakii* binding to cell surface receptors of epithelial cells (14, 15) and prevent adhesion of *Salmonella enterica* serovar Typhimurium to murine enterocytes (16).

GOS are acquired naturally through the diet from the degradation of galactan side chains of the rhamnogalacturonan I fraction of pectin (17) and from human milk oligosaccharides (HMOs) that are nondigestible by the host (18, 19). HMOs are hypothesized to promote growth of specific beneficial bacteria in the infant's early GIT colonization (20). Marcobal et al. (21) verified that HMOs can support the growth of *Lactobacillus acidophilus* NCFM, although the genetic complement of *L. acidophilus* NCFM reflects a more specific potential for GOS metabolism compared with other adapted GIT bacteria (22). *L. acidophilus* is a widely used probiotic species, originally isolated by Moro in 1900 from infant feces. The *L. acidophilus* NCFM genome was recently sequenced to reveal that the molecular machinery responsible for carbohydrate uptake and catabolism in NCFM accounts for 17% of the genes present in the genome (23). Broad carbohydrate utilization of *L. acidophilus* NCFM was demonstrated and included transporters for trehalose (24), fructooligosaccharides (25), and several other mono-, di-, and trisaccharides (26).

The current understanding of the molecular and genetic basis for uptake and catabolism of GOS by probiotic lactobacilli is limited to *in silico* predictions based on genome sequencing projects (27). The aim of the present study was to functionally identify the genetic loci responsible for GOS transport and catabolism by *L. acidophilus* NCFM.

Results

GOS-Induced Differential Gene Expression. Global changes in gene expression levels across the transcriptome were used to identify genes differentially expressed in *L. acidophilus* NCFM during

Author contributions: J.M.A., R.B., M.A.H., S.L., B.S., and T.R.K. designed research; J.M.A. and Y.J.G. performed research; S.L., Y.J.G., and T.R.K. contributed new reagents/analytic tools; J.M.A., R.B., Y.J.G., and T.R.K. analyzed data; and J.M.A., R.B., M.A.H., S.L., Y.J.G., B.S., and T.R.K. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: klaenhammer@ncsu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1114152108/-/DCSupplemental.

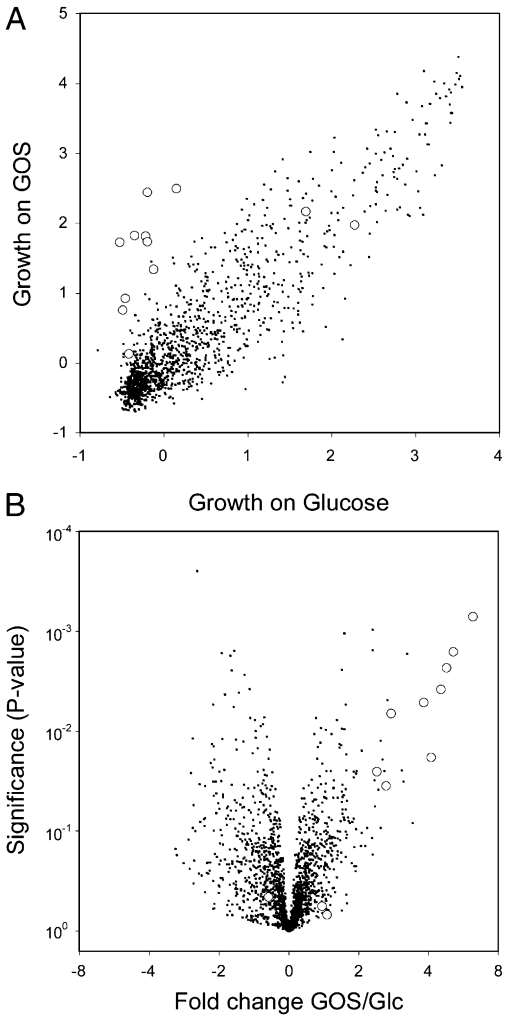


Fig. 1. Differential gene expression profile of GOS vs. glucose utilization by *L. acidophilus* NCFM. Genes involved in lactose metabolism are highlighted by open circles. (*A*) XY scatter plot of the overall normalized logarithmic gene expression profile. (*B*) Comparison of the statistical significance and gene expression differences of GOS (*Right*) vs. glucose (*Left*) depicted as a volcano plot. The *x* axis represents the differential gene induction profile as the ratio of fold difference. The *y* axis indicates statistical significance of expression difference (*P* value from ANOVA).

GOS fermentation in a semisynthetic medium (25). The single differential gene expression profile is depicted as a two-way scatter plot showing an overall linear correlation of GOS and glucose-induced gene expression (Fig. 1*A*). Notably, a subset of genes for lactose metabolism (*lac* genes, shown in white circles) were up-regulated in the presence of GOS, compared with glucose (the full dataset of the *lac* genes are reported in Table S1). Differentially expressed genes of interest were further characterized as statistically relevant ($P < 0.01$ and induction fold >2) in a volcano plot (Fig. 1*B*), confirming GOS induction of the specified *lac* operon (LBA1457–LBA1469). Statistically relevant genes induced by GOS are listed in Table 1 with annotated functions of the up-regulated genes. From Table 1, genes encoded within the 16.6 kbp *lac* operon locus were considered to be potentially involved with GOS metabolism.

The *lac* gene cluster's likely involvement in GOS utilization was consistent with the presence of two β-galactosidase–encoding genes (*lacLM*, LBA1467–1468, and *lacA*, LBA1462) assigned to the glycoside hydrolase family 2 (GH2) and glycoside hydrolase family 42 (GH42), respectively, using the CAZy classification (28).

Table 1. Differentially expressed genes in *L. acidophilus* NCFM identified by DNA microarrays of cells grown in GOS or glucose

Locus tag	Gene annotation	Fold up-regulation	*P* value
GOS-induced genes			
1467	β-galactosidase, large subunit, GH2	6.23	0.0007
1463	Lactose permease	5.10	0.0016
1462	β-galactosidase, GH42	4.79	0.0023
1459	Galactokinase	4.53	0.0038
1469	UDP-galactose-4-epimerase	3.82	0.0051
152	Phosphonate transport system ATP-binding protein	3.23	0.0017
1458	Galactose-1-phosphate uridylyltransferase	2.76	0.0067
1622	S-adenosylmethionine synthetase	2.73	0.0064
965	Hypothetical protein	2.65	0.0049
1952	Putative xanthine-uracil permease	2.29	0.0015
968	30S ribosomal protein	2.29	0.0010
Glucose-induced genes			
1429	Bile efflux transporter	2.50	0.0002
424	Conserved hypothetical protein	2.12	0.0054

Both were predicted to be localized intracellularly using the SignalP tool (29). Enzymatic activity on β-linked galactosides was demonstrated previously for both enzymes when expressed from recombinant constructs in *Escherichia coli* (30, 31). Furthermore, GH2 and GH42 β-galactosidases were proposed by Marcobal et al. (21) to be involved with degradation of HMOs. The identified galactoside-pentose-hexuronide (GPH) permease LacS (LBA1463) showed 83% amino acid sequence identity to the *Lactobacillus helveticus* functionally confirmed lactose permease (32). Two regulatory proteins, LacR (LBA1465), a LacI family regulator, and a noninduced regulator (LBA1461) with an unknown homology association, suggest regulation at the transcriptional level.

No genetic loci involved with carbohydrate metabolism were identified from the list of genes induced by glucose, suggesting that glucose is transported by the constitutively expressed mannose/glucose phospho-enolpyruvate–dependent phosphotransferase system (PEP-PTS) transporter (LBA0452, LBA0454-LBA0456), as suggested previously (26). The transcription analysis indicated that the *lac* operon in *L. acidophilus* NCFM is solely responsible for the metabolism of GOS and potentially other lactose-derived galactosides, because the gene induction profile of GOS is comparable to the lactose-induced *lac* gene expression pattern (26). It also indicates that regulation occurs at the transcriptional level, likely depending upon HPr (*ptsH*, LBA0639), CcpA (*ccpA*, LBA0431), and HPrK/P (*ptsK*, LBA0676), all of which are encoded in the *L. acidophilus* NCFM genome (23) and as previously proposed for carbohydrate utilization in *L. acidophilus* NCFM (26).

Analysis of *lacS* Inactivation. To investigate the potential involvement of the identified GPH permease LacS in GOS uptake, we inactivated the *lacS* gene using a *upp*-based counterselective gene replacement system (33), to create an in-frame deletion of 96% of the *lacS* coding region. The gene deletion had no detectable impact on cell morphology, growth in de Man, Rogosa, and Sharpe medium (MRS) or semi-defined medium (SDM)

using glucose (Fig. 2*A*), sucrose, or galactose as sole carbohydrates, suggesting that the functionality of *lacS* is nonessential for transport of monosaccharides during batch growth. Growth of the *ΔlacS* mutant was significantly impaired on lactose (Fig. 2*B*), confirming the annotation to previously validated *lacS* homologs and the previous findings of lactose induction of *lacS* together with the remaining *lac* genes (26). More significantly, the utilization of GOS (Fig. 2*C*), as well as lactitol, another galactoside prebiotic (34), was also abolished, showing a divergent and broader substrate specificity for GOS, including GOS with a higher degree of polymerization. The identification of a broad specificity transporter combined with the up-regulation of genes encoding two different β-galactosidases based on DNA microarrays illustrates a strong niche adaption by an evolved GPH β-galactosaccharide transporter.

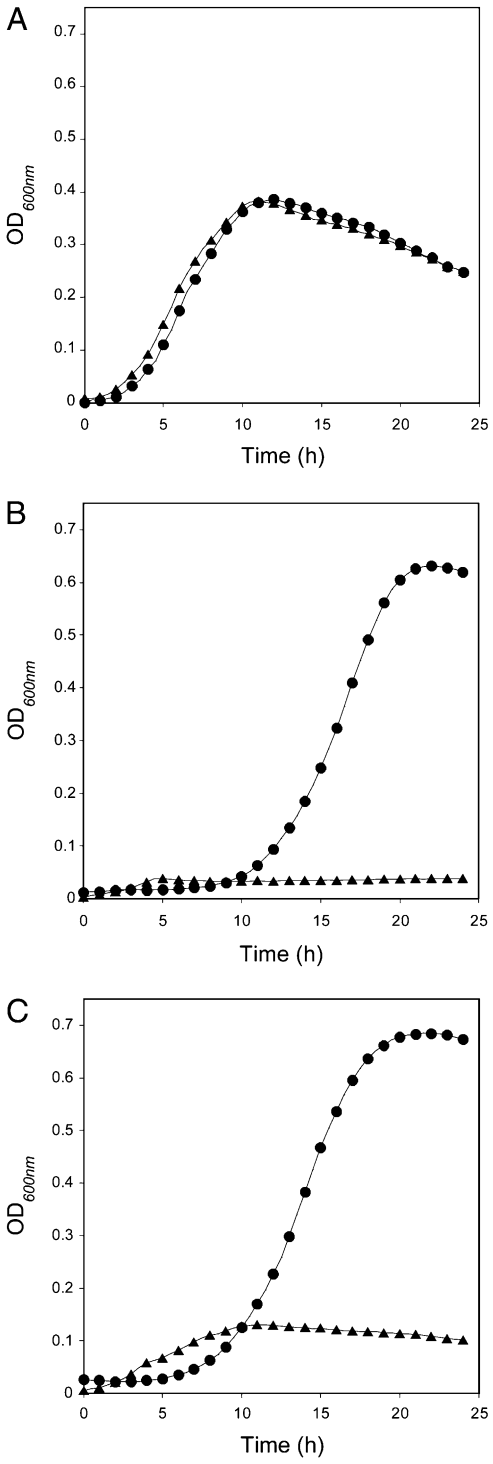


Fig. 2. Phenotype determination of *lacS* deficient mutant (▲) of *L. acidophilus* NCFM compared with wild type (●). (*A*) Growth profile on glucose; (*B*) growth profile on lactose; (*C*) growth profile on GOS.

Complete inhibition of growth by a single gene excision confirmed the hypothesis that the LacS permease was solely responsible for the transport of GOS in *L. acidophilus* NCFM and that no PEP-PTS or ATP-binding cassette (ABC) transporter systems were involved in this process. This finding indicates that the molecular basis for GOS transport and catabolism in other lactobacilli may also rely on GPH transporters and intracellular enzymatic hydrolysis by β-galactosidases from the GH2 and GH42 families before entering the Leloir and glycolysis pathways.

Sequence Analysis of GOS-Induced Gene Cluster. Additional genes surrounding the *lacS* permease and β-galactosidases were annotated in the genome with functions related to lactose and GOS metabolism, indicating a potential polycistronic operon structure for cotranscriptions of 12 genes (Fig. 3). Terminator sites and regulatory catabolite repression element (CRE) sequences were analyzed *in silico*.

The Leloir pathway genes *galM*, *galT*, *galK*, and *galE* were found with putative CRE sites, having palindromic homology to the CRE site of the *L. helveticus* lactose operon (32), yet markedly different from other *lac* CRE sites in *L. acidophilus* NCFM, indicating that these genes can be transcribed independently of the *lac* genes when only galactose is present. The *lacS*, *lacA*, and *lacLM* were all found to be under catabolite repression with two of these CRE sites showing homology to a CRE site found upstream of the *scrB* gene encoding a sucrose hydrolase in *L. acidophilus* NCFM (25). Notably, a CRE site with homology to the *lacR* CRE site was identified upstream of the *mucBP*, indicating cotranscription of *mucBP* simultaneously with the *lac* genes.

Sequence analysis of LacS predicts a two-domain structure with an N-terminal GPH permease and a C-terminal EIIA-like domain, homologous to the enzyme IIA (EIIA) of the PEP-PTS phosphorylation regulation by histidine-containing phosphocarrier protein (HPR) and enzyme I (EI). This indicates rapid regulation of lactose and related galactoside transport by *lacS* on transcriptional level in direct response to a decrease in glucose concentration. The gene locus organization differs from other characterized LacS uptake systems such as in *Lactobacillus bulgaricus* (35), *Leuconostoc lactis* (36), *Streptococcus thermophilus* SMQ-301 (37), and other *Lactobacillus* species (e.g., *Lactobacillus plantarum*, *Lactobacillus johnsonii*, and *Lactobacillus reuteri*) (Fig. S1). The differences in gene arrangement and in the types of encoded glycoside hydrolases reflect a specific adaptation of the varied species of lactic acid bacteria toward a varied β-galactosaccharide metabolism.

Phylogenetic relationships of the above LacS amino acid sequences (Fig. 4*A*) compared with the overall phylogenetic similarity of lactobacilli based on 16S rRNA homologies (38) demonstrates, first, how most *lacS* positive strains are associated with GIT colonization; and second, that the diversity of gene sequences and locus structure follow the evolutionary direction in all but *Lactobacillus delbrueckii* subsp. *bulgaricus* ATCC 11842 (39).

The gene locus organization and LacS sequence homology suggest that the specific locus originated by recent gene transfer from an unrelated precursor, possibly from within a dairy environment. Interestingly, it is observed that *lacS* genes from lactobacilli are present in the loci together with GH42 β-galactosidases for all but *L. delbrueckii* subsp. *bulgaricus* ATCC 11842, which harbors a *lacZ*-GH2 family enzyme. The phylogenetic tree of identified GH42

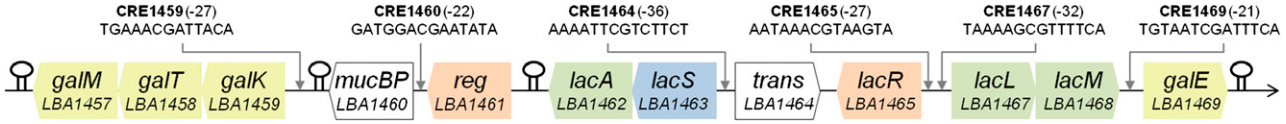


Fig. 3. Gene structure of the GOS-induced genome locus. Predicted ρ-independent transcription terminators (52) are shown as hairpin loops. Regulatory CRE sites are shown above the gene structure, with the upstream base pair distance to the starting codon. Putative functions are indicated by color: carbohydrate permease (blue), transcriptional regulators (red), glycoside hydrolases (green), Leloir pathway (yellow), and genes without a known relationship to carbohydrate metabolism (white boxes). mucBP, mucus-binding domain protein; reg, putative transcriptional regulator; trans, transposase.

β-galactosidases, *lacA*, encoded within *lacS*-containing loci, revealed no marked difference from the tree structure in Fig. 4*A*, indicating the coevolution of LacS with GH42 β-galactosidases (Fig. 4*B*).

Recently available human GIT microbiome sequencing data from the Human Microbiome Project (40) was used to validate the presence of the LacS permease and associated β-galactosidases in the human GIT microbiota by BLAST analysis (41). Both *lacS* and GH42 *lacA* genes were identified with robust statistical significance (threshold e-value $<10^{-15}$) in *L. acidophilus*, *L. helveticus*, *Lactobacillus ultunensis*, and *L. reuteri* strains from a current total of 29 *Lactobacillus* reference genomes.

Discussion

The ability of GOS to selectively promote the growth of selected GIT microbiota further establishes this prebiotic as an attractive nutritional ingredient for foods and dietary supplements. Stimulation of *Bifidobacterium* and *Lactobacillus* species by prebiotic oligosaccharides, including GOS, is well documented by observational studies (13, 42–44). Despite this, only a few studies have confirmed the lactobacilli enrichment by GOS on the strain level (45). In the present study we aimed to identify molecular elements linked to the selective GOS metabolism within *Lactobacillus* to explain in vivo observations of GOS stimulation within the GIT. Whole-genome DNA microarray analysis was performed to differentiate the gene expression pattern of *L. acidophilus* NCFM in the presence of GOS compared with glucose as the sole carbohydrate source. It was found that GOS specifically induced a cluster of genes encoding intracellular proteins involved with galactose and lactose metabolism, notably a LacS permease implicated in GOS transport. The GOS-induced gene cluster was previously identified to be up-regulated by both lactose and bile acids (26, 46), validating a role in metabolism of lactose-derived GOS and suggesting an adaptive combination of GIT-evolved traits for energy metabolism and bile tolerance.

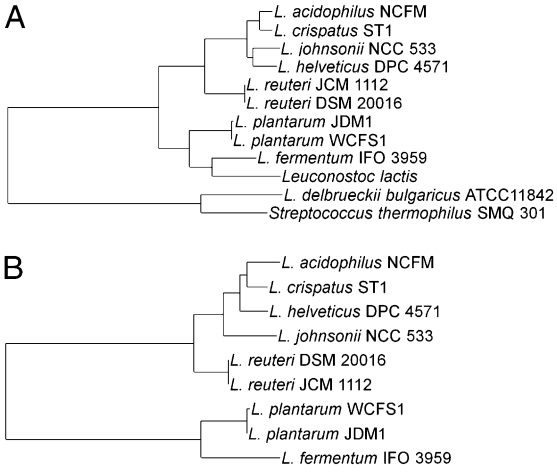


Fig. 4. Unrooted phylogenetic trees of (*A*) LacS and (*B*) LacA.

The environment-adaptive nature of *L. acidophilus* NCFM and the broad specificity of the LacS permease show the potential for delivery of *L. acidophilus* NCFM in dairy-based synbiotic GOS products, whereby a culture prefermented on lactose will rapidly metabolize GOS for increased viability upon exposure in the gut.

This study considered multiple genome sequences of lactose-fermenting lactobacilli to reveal operons encoding either a LacS permease or a PEP-PTS transporter for lactose uptake. Pathway reconstruction positioned these transporters adjacent to β-galactosidases or phospho-β-galactosidases, respectively. However, prediction of potential GOS PEP-PTS transporters was troubled by low sequence similarity to known PEP-PTS transporters families (47). Experimental validation of LacS permease as sole transporter of GOS in *L. acidophilus* NCFM was performed by gene deletion, which eliminated the ability to use lactose, GOS, and lactitol. This serves as the first identified GOS transporter in the *Lactobacillus* genus and is the first evidence that the LacS permease is capable of transporting oligosaccharides such as GOS with a DP of ≥2–6 and modified disaccharides (lactitol).

Bioinformatic identification and analysis of other *lacS* genes and their proximal genetic loci was based on the present study and earlier functional characterization of lactose transport by *lacS* homologs in lactic acid bacteria. Phylogenetic mapping of *lacS* encoding strains revealed that *lacS* is, to date, mainly found in *Lactobacillus* species that are commensals of the human gut. This suggests that transport and metabolism of lactose and complex carbohydrates are important energy sources for intestinal lactobacilli, because GPH permeases compared with ABC and PTS systems do not require ATP for import, allowing an energy-efficient and rapid adaptive transport of GOS. Analysis of the adjacent genes of *lacS* showed three core genes: *lacS*, *lacR*, and β-galactosidase of either GH2 (LacZ or LacLM) or GH42 (LacA) family. Genes without apparent known function for lactose metabolism were also found for some species (e.g., *L. acidophilus* and *L. plantarum*), and interestingly some proximal genes showed putative functional roles for mucin adhesion or rhamno-galactoside metabolism, respectively. This suggests that the base functionality of the *lacS* genetic locus is highly conserved by evolutionary pressure and important for niche survival in the GIT via transport and metabolism of lactose, GOS, and likely fractions of HMOs.

The presence of *lacS* and *lacA* homologs among the intestinal lactobacilli supports the importance of complex galactoside utilization for energy metabolism. The related genetic loci were inclusive within the acidophilus subfamily of lactobacilli (*L. acidophilus*, *L. ultunensis*, and *L. helveticus*), whereas other species (e.g., *Lactobacillus fermentum* and *L. plantarum*) include *lacS*-positive strains and strains that have no homologs of either *lacS* or *lacA*. The retention of *lacS* and *lacA* homologs in *L. helveticus* is consistent with the known genetic lineage and adaptation of these lactobacilli to milk (48). In that process, *L. helveticus* eliminated a number of GIT-related functions (e.g., bile salt hydrolase and mucin binding proteins) but retained the *lac*-related genes while losing most *gal*-related genes except the galactose-1-phosphate uridylyltransferase gene (*galT*).

In conclusion, we identified LacS as the sole transporter for lactose, GOS, and lactitol. A future combination of tran-

Table 2. Bacterial strains and plasmids used in this study

Strain or plasmid	Characteristics	Reference or source
E. coli strains		
NCK1831	EC101: RepA$^+$ JM101; Km^r; *repA* from pWV01 integrated in chromosome; host for pORI-based plasmids	53
NCK1911	NCK1831 harboring pTRK935	33
NCK2126	NCK1831 harboring pTRK1015	Present study
L. acidophilus strains		
NCFM	Human intestinal isolate	23
NCK1909	NCFM carrying a 315 bp in-frame deletion in the *upp* gene	33
NCK1910	NCK1909 harboring pTRK669, host for pORI-based counterselective integration vector	33
NCK2127	NCK1909 carrying a 1831 bp in-frame deletion in the *lacS* gene	Present study
Plasmids		
pTRK669	Ori (pWV01], Cm^r RepA$^+$	54
pTRK935	pORI28 derived with an inserted *upp* expression cassette and *lacZ´* from pUC19, serves as counterselective integration vector, Em^r	33
pTRK1015	pTRK935 with a mutated copy of *lacS* cloned into BamHI/EcoRI sites	Present study

scriptomics, proteomics, and functional genomics analyses will provide a comprehensive platform for study of the molecular interactions between probiotics and prebiotics.

Materials and Methods

Bacterial Strains and Growth Conditions. All bacterial strains and plasmids used throughout this study are listed in Table 2. *Lactobacillus* broth cultures were cultivated in MRS (Difco Laboratories) or SDM (49), supplemented with 0.5% (wt/vol) glucose (Sigma-Aldrich), lactose, lactitol (Danisco), or GOS (94% GOS, DP 2–6; Danisco; Fig. S2) as carbon source, aerobically at 37 °C or 42 °C. Chloramphenicol (5 μg/mL) and erythromycin (2 μg/mL) were used when necessary for selection. *E. coli* strains were cultivated in Brain Heart Infusion medium (Difco) aerobically at 37 °C with aeration, and erythromycin (150 μg/mL) and/or kanamycin (40 μg/mL) was added for selection. Solid media were prepared by the addition of 1.5% (wt/vol) granulated agar (Difco).

***L. acidophilus* NCFM Microarray Platform.** Whole-genome oligonucleotide microarrays were designed as described by Goh et al. (33) with four replicate spots for each of the 1,824 predicted genes. Hybridization quality was assessed by monitoring the Cy3/Cy5 ratio of labeled cDNA, prepared from total RNA, after slide scan to observe a linear correlation between the two fluorophores. For DNA microarray transcriptome study, semisynthetic media (25) used for cultivation of *L. acidophilus* NCFM were filtered through a 0.22-μm filter, and oxygen was removed by the Hungate method (50). *L. acidophilus* NCFM cultures were propagated in parallel in semisynthetic media with 1% (wt/vol) glucose or GOS as carbon source. Cultures were transferred for four passages on each sugar before being harvested at the early logarithmic phase (OD_{600}= 0.35–0.5) by pelleting at 4 °C (3,000 × *g*, 15 min) and flash freezing the pellets for storage at −80 °C.

cDNA Preparation and Microarray Hybridization. Cells were mechanically disrupted by beadbeating, and total RNA was isolated using TRIzol-chloroform extraction (Invitrogen). Genomic DNA was removed with Turbo DNase (Ambion), followed by RNA purification using an RNeasy Mini Kit (Qiagen) (33).

Reverse transcription of total RNA, fluorescent labeling of cDNA, and hybridizations were done using 20 μg of total RNA for each replicate, as described by Goh et al. (33). Total RNA from each of the two carbohydrate treatments was labeled with both Cy3 and Cy5 for two technical replicates to each growth condition.

Microarray Data Acquisition and Analysis. Hybridized chips were scanned at 10-μm resolution per pixel using a ScanArray Express microarray scanner (Packard BioScience) for 16-bit spot intensity quantification. Fluorescent intensities were quantified and background subtracted using the QuantArray 3.0 software package (Packard BioScience). Median values were calculated for all ORF tetraplicate intensities and $\log_2$-transformed before being imported into SAS JMP Genomics 4.0 (SAS Institute) for data analysis. The full dataset was interquantile normalized and modeled using a mixed-model ANOVA for analysis of the differential gene expression pattern and visualization using heat maps and volcano plots.

Construction and Phenotypic Determination of the *lacS* Deletion Mutant. Genomic DNA of *L. acidophilus* NCFM was isolated by the method of Walker and Klaenhammer (51) or by the Mo Bio Ultraclean microbial DNA isolation kit (Mo Bio Laboratories). Plasmid DNA from *E. coli* was isolated using a QIAprep Spin miniprep kit (Qiagen). Restriction enzymes (Roche Molecular Biochemicals) were applied according to the instructions supplied by the manufacturer. DNA ligation was done using T4 DNA ligase (New England Biolabs) as directed by the manufacturer's recommendations. All PCR primers (Table S2) were synthesized by Integrated DNA Technologies. PCR reactions, preparation and transformation of competent *L. acidophilus* NCFM and *E. coli* cells, analysis by agarose gel electrophoresis, and in-gel purification were done as described by Goh et al. (33).

The construction of an Δ*upp* isogenic mutant with in-frame DNA excision of 96% of the *lacS* (LBA1463) coding region was done according to Goh et al. (33). In short, the upstream and downstream flanking regions (approximate length of 750 bp each) of the deletion target were PCR-amplified with the 1463A/1463B and 1463C/1463D primer pairs, respectively, and fused by splicing by overlap extension PCR (SOE-PCR). The SOE-PCR product was cleaved with EcoRI and BamHI before ligation into pTRK935 linearized with compatible ends and transformed into NCK1831. The resulting recombinant plasmid, pTRK1015, harbored in NCK2126, was transformed into NCK1910 harboring pTRK669, for chromosomal integration of pTRK1015 and following DNA excision to generate the Δ*lacS* genotype. Confirmation of DNA deletion was done by PCR and DNA sequencing using primer pair 1463UP and 1463DN (Table S2).

Lactose and GOS utilization of the *lacS* gene deletion mutant was tested by comparative growth to wild-type *L. acidophilus* NCFM and NCK1909 (*upp* mutant and parent strain of the Δ*lacS* mutant). All strains were grown in SDM supplemented with 1% (wt/vol) glucose before inoculation [1% (vol/vol)] of an overnight culture into SDM supplemented with 0.5% (wt/vol) of the following carbohydrates in separate batches: lactose, GOS, lactitol, galactose, sucrose, and glucose. Growth was monitored by optical density using a Fluostar spectrophotometer in triplicate wells of a 96-well plate (200 μL per well) and covered with an airtight seal. All carbohydrates were at least 95% pure.

ACKNOWLEDGMENTS. We thank Evelyn Durmaz for assistance with confirmational sequencing of the deleted region of the *lacS*-deficient *Lactobacillus acidophilus* mutant, and Jesper Wichmann (Danisco) for his contributions to the GOS purification and analysis. This work is supported by Danisco A/S, the North Carolina Dairy Foundation, and the FøSu grant from the Danish Strategic Research Council to the project "Gene discovery and molecular interactions in prebiotics/probiotics systems. Focus on carbohydrate prebiotics." J.M.A. is supported by a joint Ph.D. stipend from Danisco A/S, the FøSu grant, and the Technical University of Denmark.

1. Tang ML, Lahtinen SJ, Boyle RJ (2010) Probiotics and prebiotics: Clinical effects in allergic disease. *Curr Opin Pediatr* 22:626–634.
2. Leyer GJ, Li S, Mubasher ME, Reifer C, Ouwehand AC (2009) Probiotic effects on cold and influenza-like symptom incidence and duration in children. *Pediatrics* 124:e172–e179.
3. Hedin C, Whelan K, Lindsay JO (2007) Evidence for the use of probiotics and prebiotics in inflammatory bowel disease: A review of clinical trials. *Proc Nutr Soc* 66:307–315.
4. Reid G, et al. (2003) New scientific paradigms for probiotics and prebiotics. *J Clin Gastroenterol* 37:105–118.
5. Qin J, et al.; MetaHIT Consortium (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464(7285):59–65.
6. Roberfroid M (2007) Prebiotics: The concept revisited. *J Nutr* 137 (3, Suppl 2): 830S–837S.
7. Callaway TR, et al. (2008) Probiotics, prebiotics and competitive exclusion for prophylaxis against bacterial disease. *Anim Health Res Rev* 9:217–225.
8. Park AR, Oh DK (2010) Galacto-oligosaccharide production using microbial beta-galactosidase: current state and perspectives. *Appl Microbiol Biotechnol* 85:1279–1286.

9. Macfarlane GT, Steed H, Macfarlane S (2008) Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics. *J Appl Microbiol* 104:305–344.
10. Rastall RA (2010) Functional oligosaccharides: Application and manufacture. *Annu Rev Food Sci Technol* 1:305–339.
11. Ben XM, et al. (2004) Supplementation of milk formula with galacto-oligosaccharides improves intestinal micro-flora and fermentation in term infants. *Chin Med J (Engl)* 117:927–931.
12. Silk DB, Davis A, Vulevic J, Tzortzis G, Gibson GR (2009) Clinical trial: The effects of a trans-galactooligosaccharide prebiotic on faecal microbiota and symptoms in irritable bowel syndrome. *Aliment Pharmacol Ther* 29:508–518.
13. Drakoularakou A, Tzortzis G, Rastall RA, Gibson GR (2010) A double-blind, placebo-controlled, randomized human study assessing the capacity of a novel galacto-oligosaccharide mixture in reducing travellers' diarrhoea. *Eur J Clin Nutr* 64(2):146–152.
14. Sinclair HR, de Slegte J, Gibson GR, Rastall RA (2009) Galactooligosaccharides (GOS) inhibit *Vibrio cholerae* toxin binding to its GM1 receptor. *J Agric Food Chem* 57: 3113–3119.
15. Quintero M, et al. (2011) Adherence inhibition of *Cronobacter sakazakii* to intestinal epithelial cells by prebiotic oligosaccharides. *Curr Microbiol* 62:1448–1454.
16. Searle LE, et al. (2010) Purified galactooligosaccharide, derived from a mixture produced by the enzymic activity of *Bifidobacterium bifidum*, reduces *Salmonella enterica* serovar Typhimurium adhesion and invasion in vitro and in vivo. *J Med Microbiol* 59:1428–1439.
17. Jones L, Seymour GB, Knox JP (1997) Localization of pectic galactan in tomato cell walls using a monoclonal antibody specific to (1[->]4)-[beta]-D-galactan. *Plant Physiol* 113:1405–1412.
18. Chaturvedi P, Warren CD, Ruiz-Palacios GM, Pickering LK, Newburg DS (1997) Milk oligosaccharide profiles by reversed-phase HPLC of their perbenzoylated derivatives. *Anal Biochem* 251(1):89–97.
19. Miller JB, McVeagh P (1999) Human milk oligosaccharides: 130 reasons to breast-feed. *Br J Nutr* 82:333–335.
20. Ninonuevo MR, et al. (2006) A strategy for annotating the human milk glycome. *J Agric Food Chem* 54:7471–7480.
21. Marcobal A, et al. (2010) Consumption of human milk oligosaccharides by gut-related microbes. *J Agric Food Chem* 58:5334–5340.
22. Sela DA, et al. (2008) The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. *Proc Natl Acad Sci USA* 105:18964–18969.
23. Altermann E, et al. (2005) Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. *Proc Natl Acad Sci USA* 102:3906–3912.
24. Duong T, Barrangou R, Russell WM, Klaenhammer TR (2006) Characterization of the *tre* locus and analysis of trehalose cryoprotection in *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol* 72:1218–1225.
25. Barrangou R, Altermann E, Hutkins R, Cano R, Klaenhammer TR (2003) Functional and comparative genomic analyses of an operon involved in fructooligosaccharide utilization by *Lactobacillus acidophilus*. *Proc Natl Acad Sci USA* 100:8957–8962.
26. Barrangou R, et al. (2006) Global analysis of carbohydrate utilization by *Lactobacillus acidophilus* using cDNA microarrays. *Proc Natl Acad Sci USA* 103:3816–3821.
27. Makarova K, et al. (2006) Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci USA* 103:15611–15616.
28. Cantarel BL, et al. (2009) The Carbohydrate-Active EnZymes database (CAZy): An expert resource for Glycogenomics. *Nucleic Acids Res* 37(Database issue):D233–D238.
29. Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340:783–795.
30. Nguyen TH, et al. (2007) Characterization and molecular cloning of a heterodimeric beta-galactosidase from the probiotic strain *Lactobacillus acidophilus* R22. *FEMS Microbiol Lett* 269(1):136–144.
31. Pan Q, et al. (2010) Functional identification of a putative beta-galactosidase gene in the special *lac* gene cluster of *Lactobacillus acidophilus*. *Curr Microbiol* 60(1):172–178.
32. Fortina MG, Ricci G, Mora D, Guglielmetti S, Manachini PL (2003) Unusual organization for lactose and galactose gene clusters in *Lactobacillus helveticus*. *Appl Environ Microbiol* 69:3238–3243.
33. Goh YJ, et al. (2009) Development and application of a *upp*-based counterselective gene replacement system for the study of the S-layer protein SlpX of *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol* 75:3093–3105.
34. Ouwehand AC, Tiihonen K, Saarinen M, Putaala H, Rautonen N (2009) Influence of a combination of *Lactobacillus acidophilus* NCFM and lactitol on healthy elderly: Intestinal and immune parameters. *Br J Nutr* 101:367–375.
35. Leong-Morgenthaler P, Zwahlen MC, Hottinger H (1991) Lactose metabolism in *Lactobacillus bulgaricus*: Analysis of the primary structure and expression of the genes involved. *J Bacteriol* 173:1951–1957.
36. Vaughan EE, David S, de Vos WM (1996) The lactose transporter in *Leuconostoc lactis* is a new member of the LacS subfamily of galactoside-pentose-hexuronide translocators. *Appl Environ Microbiol* 62:1574–1582.
37. Vaillancourt K, Moineau S, Frenette M, Lessard C, Vadeboncoeur C (2002) Galactose and lactose genes from the galactose-positive bacterium *Streptococcus salivarius* and the phylogenetically related galactose-negative bacterium *Streptococcus thermophilus*: Organization, sequence, transcription, and activity of the gal gene products. *J Bacteriol* 184:785–793.
38. Ventura M, et al. (2009) Genome-scale analyses of health-promoting bacteria: Probiogenomics. *Nat Rev Microbiol* 7(1):61–71.
39. van de Guchte M, et al. (2006) The complete genome sequence of *Lactobacillus bulgaricus* reveals extensive and ongoing reductive evolution. *Proc Natl Acad Sci USA* 103:9274–9279.
40. Nelson KE, et al.; Human Microbiome Jumpstart Reference Strains Consortium (2010) A catalog of reference genomes from the human microbiome. *Science* 328:994–999.
41. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410.
42. Moro G, et al. (2002) Dosage-related bifidogenic effects of galacto- and fructooligosaccharides in formula-fed term infants. *J Pediatr Gastroenterol Nutr* 34:291–295.
43. Vulevic J, Drakoularakou A, Yaqoob P, Tzortzis G, Gibson GR (2008) Modulation of the fecal microflora profile and immune function by a novel trans-galactooligosaccharide mixture (B-GOS) in healthy elderly volunteers. *Am J Clin Nutr* 88:1438–1446.
44. Davis LM, Martínez I, Walter J, Hutkins R (2010) A dose dependent impact of prebiotic galactooligosaccharides on the intestinal microbiota of healthy adults. *Int J Food Microbiol* 144:285–292.
45. Ben XM, et al. (2008) Low level of galacto-oligosaccharide in infant formula stimulates growth of intestinal Bifidobacteria and Lactobacilli. *World J Gastroenterol* 14: 6564–6568.
46. Pfeiler EA, Azcarate-Peril MA, Klaenhammer TR (2007) Characterization of a novel bile-inducible operon encoding a two-component regulatory system in *Lactobacillus acidophilus*. *J Bacteriol* 189:4624–4634.
47. Saier MH, Jr., Yen MR, Noto K, Tamang DG, Elkan C (2009) The transporter classification database: Recent advances. *Nucleic Acids Res* 37(Database issue):D274–D278.
48. Callanan M, et al. (2008) Genome sequence of *Lactobacillus helveticus*, an organism distinguished by selective gene loss and insertion sequence element expansion. *J Bacteriol* 190:727–735.
49. Kimmel SA, Roberts RF (1998) Development of a growth medium suitable for exopolysaccharide production by *Lactobacillus delbrueckii* ssp. *bulgaricus* RR. *Int J Food Microbiol* 40(1-2):87–92.
50. Daniels L, Zeikus JG (1975) Improved culture flask for obligate anaerobes. *Appl Microbiol* 29:710–711.
51. Walker DC, Klaenhammer TR (1994) Isolation of a novel IS3 group insertion element and construction of an integration vector for *Lactobacillus* spp. *J Bacteriol* 176: 5330–5340.
52. Kingsford CL, Ayanbule K, Salzberg SL (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* 8(2):R22.
53. Law J, et al. (1995) A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *J Bacteriol* 177:7011–7018.
54. Russell WM, Klaenhammer TR (2001) Efficient system for directed integration into the *Lactobacillus acidophilus* and *Lactobacillus gasseri* chromosomes via homologous recombination. *Appl Environ Microbiol* 67:4361–4364.

6.4 Dual substrate specificity of a prebiotic transporter from *Bifidobacterium animalis* subsp. *lactis* BL-04

In preparation.

Dual substrate specificity of a prebiotic transporter from *Bifidobacterium animalis* subsp. *lactis* Bl-04

Joakim Mark Andersen[1], Morten Ejby[1], Jonas Rosager Henriksen[2], Thomas Lars Andresen[2], Maher Abou Hachem[1], Birte Svensson[1]*.

[1]Enzyme and Protein Chemistry, Department of Systems Biology, Søltofts Plads Building 224, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[2]Department of Micro- and Nanotechnology, Produktionstorvet Building 423, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

*To whom correspondence should be addressed: Phone +45 4525 2740; Fax: +45 4588 6307; E-mail: bis@bio.dtu.dk

Keywords

ITC, SPR, biophysical protein characterization, solute binding protein, prebiotics, probiotics, *Bifidobacterium,* ABC transporter.

Abstract (words: 248)

Probiotic microorganisms, such as bifidobacteria, exert beneficial effects on the human host through their presence in the gastrointestinal tract where the proliferation dependent heavily on consumption of complex, mainly oligosaccharides, carbohydrates, termed prebiotics. Functional understanding of the oligosaccharide uptake systems coupled with dedicated glycoside hydrolases are currently lacking functional characterization to understand the molecular mechanisms and broadness of prebiotic utilization.

The aim of the present work was to characterize the substrate recognition of putative dual raffinose oligosaccharide family and isomaltooligosaccharide specific ABC transporter from *Bifidobacterium animalis* subsp. *lactis*, by screening the ligand affinities of the recombinant solute binding protein and deduce a potential mechanism of binding by the thermodynamic finger-prints.

Surface plasmon resonance analysis was used to measure steady-state binding of isomaltooligosaccharide and raffinose family oligosaccharides in the μM-range validating the proposed dual specificity of the transporter, however the affinities found represent a 100 fold decrease in affinity compared to previous reports. Isothermal titration calorimetry suggested this weaker affinity to be driven by increased entropic contributions as a result of an altered ligand binding cleft in dual substrate binding. This was supported by sequence analysis, which allowed prediction of structural changes of binding residues compared to a homolog oligosaccharide binding protein. Noticeably, comparative genomics of the identified locus confined the occurrence of homologous loci to primary bifidobacteria, streptococci and lactobacilli.

In conclusion, the present work reports the detailed substrate affinities and thermodynamics of potential prebiotic uptake by an ABC transporter with novel dual substrate specificity within probiotic bifidobacteria.

Introduction

Probiotic microorganisms (1) have in the recent decade become well-documented by clinical trials for their abilities to promote human health through prevention of bacterial associated diarrhea (2, 3), bowel disorders (4) ranging through treatments of newborn (5) to the elderly (6). The genus *Bifidobacterium* has been shown to harbor several species and strains with probiotics status (7). The probiotic character of bifidobacteria is reflected by their genetically encoded abilities to pass through the human gastrointestinal tract, by displaying high tolerance towards acid and bile, and in the colon to utilize complex nutrients in a competitive habitat (8). A key attribute of the probiotic nature of bifidobacteria is their ability to utilize complex dietary carbohydrate, which is mediated by a battery of carbohydrate uptake and degradation proteins (9–11). Carbohydrates that are selectively metabolized by probiotics have been defined as prebiotics (12) and have shown to increase the probiotic cell numbers and exert positive effects in human intervention studies (13–15). However, the molecular mechanisms of selective metabolism within bifidobacteria are in most cases not well understood beyond *in silico* predictions and a few selected studies highlighting the impact of glycoside hydrolases (16), whereas the function and role of glycoside transporters remain largely unknown.

The majority of bifidobacteria do not possess hydrolytic pathways for polysaccharide utilization (17) as opposed to many other members of the GIT microbiome (18). The bifidobacteria thus have evolved specialized carbohydrate transport systems for uptake of available dietary oligosaccharides and cross-feeding on the polysaccharide breakdown products (19). Recent studies support the importance of oligosaccharide uptake within bifidobacteria and propose primarily ATP binding cassette (ABC) transporters to be involved with uptake of oligosaccharide prebiotics (20–22). The data supporting the functional role so far are only linked to increased bifidobacteria counts *in vivo* (13, 23).

ABC transporters are identified in organism from all domains of life and facilitate an ATP energized uptake (or export) of vitamins, carbohydrates, oligo-peptides, amino acids, ions and various organic

compounds (24). The broad range of uptake is reflected by diversity in modularity of the domains constituting the transporter (25). Gram-positive bacterial carbohydrate transporters are typically found as pentamers composed of an extracellular cell wall attached solute binding protein, two membrane-spanning domains forming the permease and two nucleotide binding proteins coupling the hydrolysis of ATP to energize the transport (25). The substrate specificity of ABC transporters is determined by the solute binding protein (26) and so far characterization of solute binding proteins within bifidobacteria has been limited to a lacto-*N*-biose binding protein (27). Explanation of the specificities for plant derived oligosaccharides prebiotics is currently lacking.

The potential prebiotic utilization of the probiotic *B. animalis* subsp. *lactis* Bl-04 (28) revealed a genomic locus encoding a single ABC transporter adjacent to glycoside hydrolases with putative specificities for the candidate prebiotic groups of raffinose family oligosaccharides (RFO) (29, 30) and isomalto-oligosaccharides (IMO) (31, 32), respectively. These types of potential prebiotics have been shown to increase bifidobacteria counts *in vivo* (29, 31), hence potentially linking the ABC transporter driven utilization of these potential prebiotics to the probiotics nature of *B. lactis* Bl-04.

We therefore hypothesized that the solute binding protein (Balac_1599, in the following referred to as *B. lactis* Bl-04 α-1,6-glycoside binding protein (*Bl*16GBP)) could display broad substrate specificity for RFO and IMO, and the transporter could be a potential link to substantiate the mechanism of selective prebiotic utilization within bifidobacteria. In this light, the aim of the present work was to characterize the substrate binding of the recombinant RFO/IMO solute binding protein and through thermodynamics finger-prints rationalize the mechanism of the intriguing dual specificity, possibly related to probiotic functions within bifidobacteria.

Results

Identification and sequence analysis of a RFO/IMO utilization locus in *B. lactis* Bl-04

Gene landscape analysis of *B. lactis* Bl-04 identified a genomic locus encoding an oligosaccharide specific ABC transporter (Balac_1597–1599), two putative glycoside hydrolase family (GH, www.cazy.org) 36 α-galactosidases (Balac_1596 and Balac_1601) and a putative GH13 oligo-α-1,6-glucosidase (Balac_1593), with the three hydrolases predicted to be intracellular (33) and the locus under transcriptional regulation by a NagC type regulator (Balac_1600). Sequence analysis of the GH36 α-galactosidase Balac_1601 revealed the presence of the $[CSSGGGR]_{514-520}$ active site motif (Balac_1601 numbering) thus assigning this enzyme into subfamily I of GH36 harboring α-galactosidases specific for raffinose (α-D-Gal*p*-(1–6)- D-Glc*p*-(α1,β2)-D-Fru*f*) and RFO (34). This together with the 67% amino acid sequence identity with the previously characterized α-galactosidase from *B. bifidum* (35) supports the specificity of this enzyme toward this class of α-1,6-galactosides abundant in human diet. The alter α-galactosidase Balac_1596 was assigned to subfamily II harboring dominantly plant raffinose synthases. The putative oligo-α-1,6-glucosidase showed 66–72% amino acid sequence identity to two α-1,6-glucosidases from *B. breve* shown to catalyze the hydrolysis of isomaltose (α-D-Glc*p*-(1–6)-D-Glc*p*), isomaltotriose (α-D-Glc*p*-(1–6)-α-D-Glc*p*-(1–6)-D-Glc*p)* and panose (α-D-Glc*p*(1–6)-α-D-Glc*p*-(1–4)-D-Glc*p*) (36). The sequence analysis thus suggested that this locus encodes the uptake of RFO and IMO via a single ABC transporter and their subsequent degradation by the mentioned glycoside hydrolases.

Dual substrate affinity characterization of recombinant *Bl*16GBP

To confirm the specificity of *Bl*16GBP, the recombinant protein was produced and purified to yield a 44 kDa protein corresponding to the theoretical 43.7 kDa mature polypeptide comprising residues 46–437, N-terminally flanked by four residues [GSHM] introduced form the cloning vector following the cleavage of the N-terminal His-tag. The structural integrity of the recombinant protein was assessed using

differential scanning calorimetry (DSC) (Figure S1). The calorimetric trace of *Bl*16GBP showed a well defined single thermal transition with a T_m of 68.2 °C, thus confirming the structural integrity and the thermostability of the protein. Interestingly a modest increase in T_m (0.9 °C) was observed in the presence 2 mM raffinose suggesting ligand mediated stabilization.

Ligand preference of *Bl*16GBP was screened using surface plasmon resonance (SPR) with *Bl*16GBP immobilized to a CM5 chip to 3900 response units (RU) for measuring the steady state binding of carbohydrate binding to *Bl*16GBP (Figure S2). No binding to the tested monosaccharides (fructose, galactose and glucose) was observed in agreement with the reported divergence of mono- and oligosaccharide binding proteins (26). Furthermore, the solute binding protein was specific for IMO and RFO as no binding was detected the β-glycosides cellobiose (β-D-Glc*p*-(1–4)-D-Glc*p)*, β-1,4-xylo-oligosaccharides ([β- D-xyl*f*-(1–4)]$_{1-6}$-D-xyl*f*), β-galactooligosaccharides ([β-D-Gal*p*-(1–4)]$_{1-5}$-D-Glc*p*), β-fructo-oligosaccharides ([β-D-Fru*f*-(1–2)]$_{1-5}$-(β2,α1)-D-Glc*p*) or to α-1,4-glucooligosaccharides ([α-D-Glc*p*-(1–6)]$_{1-6}$-D-Glc*p*) was measured. The dissociation constants for the tested RFO and IMO substrates were determined from a one binding site model to the steady state response as a function of concentration (Table 1 and Figure 1A). The highest affinity was measured towards the trisaccharides panose and raffinose with K_D values of 8.7 μM and 22.7 μM, respectively (Table 1). Notably, about a 100 fold reduction in affinity was measured for the disaccharides melibiose and isomaltose as compared to panose.

Distinguishable differences in the affinity for longer oligosaccharides were observed between RFO and IMO with the affinity dropping 16 fold for stachyose ([α-D-Gal*p*-(1–6)]$_2$-D-Glc*p*-(α1,β2)-D-Fru*f*) as compared to raffinose and no measurable binding of verbascose ([α-D-Gal*p*-(1–6)]$_3$-D-Glc*p*-(α1,β2)-D-Fru*f*), whereas the affinity for IMO with degree of polymerization (DP) 3–7 was very similar and only about 10–16 fold lower than for panose (Table 1). The binding affinity of the solute binding protein for raffinose was essentially unchanged in the pH range 5.5–8.0 (Table 2).

The binding affinities of panose and raffinose were also measured by isothermal titration calorimetry (ITC) (Figure 2), and the experimental binding stoichiometries determined were consistent with a 1:1 binding model and enthalpically dominated binding of ligands was measured (Table 3) and the binding affinities were comparable.

The temperature dependence of raffinose binding to *Bl*16GBP was measured by SPR and the raffinose dissociation (Figure 2B) was modeled to a van't Hoff equation of the binding data to yield an estimated enthalpy (ΔH) of -67 kJ/mol and entropy ($T\Delta S$) of -40 kJ/mol. Thus the substrate binding was found to be energetic favorable by enthalpic contributions and negative entropy confirming the above temperature increase of protein unfolding with substrate bound.

Sequence comparison for structure-function insight in the *Bl*16GBP binding cleft

Multiple sequence alignment of representative bifidobacteria homologs of *Bl*16GBP, together with a previously identified and structure determined lacto-*N*-biose (Gal-β-(1–3)-GlcNAc) specific solute binding protein, referred to as GL-BP, (PDB: 2Z8F) from *B. longum* (27), was done to deduce both conserved and different functional amino acids in proximity of the substrate binding cleft (Figure 4). Notably, from the global alignment a significant part of the conserved residues were related to secondary structural elements indicating an overall similar structural fold of *Bl*16GBP although a loop deletion of two amino acids $GG_{289\text{-}290}$ (GL-BP numbering) was found and corresponded to a flexible hinge region (37), thus the deletion identified in *Bl*16GBP (and homologs) may result in a more rigid conformation as possibly reflected by the low difference in unfolding temperature increase in the presence of raffinose. The residues potentially lining the substrate binding cleft of *Bl*16GBP were compared to the corresponding 21 residues found within 4 Å of the bound lacto-*N*-tetraose (Gal-β-(1–3)-GlcNAc-β-(1–3)-Gal-β-(1–4)-Glc) in GL-BP (residues highlighted in Figure 4). The corresponding 21 residues in *Bl*16GBP indicated how the side chains were overall less bulky thus displaying a structurally broader

binding cleft required for the dual specificity observed. Two tryptophans in the lacto-*N*-biose solute binding protein (W_{231} and W_{252}) were found to be the main residues involved with aromatic stacking to the ligand and hence proposed to exhibit a commonly found motif for carbohydrate binding (38). In *Bl*16GBP only a single tryptophan were functionally conserved by a tyrosine substitution, whereas the other tryptophan was lacking and the corresponding residues was positioned in a putative variable loop region. These observations indicate a certain plasticity of the *Bl*16GBP binding cleft and the apparent wide binding cleft with the loss of aromatic stacking for substrate binding may explain the weaker binding in the μM range and the dual specificity of both RFO and IMO.

Discussion

The significance of ABC transporters in probiotics for uptake of prebiotic oligosaccharides has become evident through gene and protein identification (39, 40). To date however, characterization of protein structure-function relationships is limited for understanding the diversity of substrate specificities proposed for ABC transporters and their contribution to selective prebiotic utilization by probiotics.

The present work identified a putative locus encoding an ABC transporter for potential prebiotic utilization in *B. lactis* Bl-04 and presents biophysical SPR and ITC characterization of the dual specificity solute binding protein conferring the initial step of uptake of the candidate type of prebiotic RFO and IMO. To support the selective mechanism of RFO and IMO utilization as a key attribute of prebiotics (12, 41), we mapped the phylogenetic comparison of *Bl*16GBP homologs (Figure 4) together with putative raffinose solute binding proteins previously identified by transcriptional analysis in *B. longum* NCC2705 {252 Parche 2007;}} and *Streptococcus mutant* UA159, which also was upregulated by isomaltose (42). The taxonomical distribution was confined to the Actinobacteria and Firmicutes phylae with the dominant species being *Bifidobacterium* or *Streptococcus* and *Lactobacillus*, respectively, leading to species prevailingly found in the gastrointestinal tract (*Cellulomonas* excluded albeit being a plant biomass

degrading organism) strengthening the potential selective utilization of RFO and IMO. Gene landscape analysis of the gene clusters encoding the above putative RFO/IMO solute binding proteins (Figure 5) revealed a consistent co-encoding of a putative oligosaccharide ABC transporter permease and the identified solute binding protein with a GH36 α-galactosidase. Notably, only gene clusters from Firmicutes encoded a GH13_18 sucrose phosphorylase, suggesting divergent catabolism of simpler carbohydrates within the two phylae. No distinct pattern was found for the occurrence of GH13 oligo-α-1,6-glucosidases, although not encoded in the lactobacilli gene clusters, where an alternative pathway for utilization of IMO was proposed (Møller et al., J. Bacteriol. 2012, in press.; Andersen et al., in review at PLoS ONE). Interestingly, the *B. animalis* gene clusters encoded an additional GH36 α-galactosidase sub-classified into GH36_II (34) with no predicted signal peptide indicating that the *B. lactis* Bl-04 gene may have a potential novel specificity for α-1,6-glycosides beyond the substrates binding to *Bl*16GBP.

The SPR characterization of the dual substrate specificity *Bl*16GBP revealed strongest affinity for the trisaccharides panose and raffinose while binding of longer RFO was weaker and confined to melibiose and stachyose, as no binding of the longer verbascose was measured. This apparently restricted binding mode was not observed for the IMO, found to bind all the tested DP 2–7 IMO, indicating that the ligand binding cleft recognizes the non-reducing ends of either α-1,6-glucosides or α-1,6-galactosides where the terminal hexose-glycosyl defines the end-point of RFO binding. No binding of sucrose (α-D-Glc*p*-(α1,β2)-D-Fru*f*) or isomaltulose (α-D-Glc*p*-α-(1–6)-D-Fru*f*) was observed, emphasizing that *Bl*16GBP recognizes the reducing end of an α-1,6 linked hexose as a key determinant of substrate binding.

The higher μM range of binding affinities was in striking contrast to previous characterizations of oligosaccharide binding proteins were sub μM binding affinities were reported (Table 4), indicating that the dual substrate specificity may have evolved with the cost of reduced affinity. Sequence comparison to the structural determined lacto-*N*-biose solute binding protein identified key residues in the ligand binding cleft, where a key tryptophan was lacking among the corresponding *Bl*16GBP residues and with an

overall display of a seemingly open binding cleft assist to rationalize the decrease in affinity by increase plasticity of the *Bl*16GBP binding cleft.

The enthalpic and entropic finger-prints of ligand binding compared to previously reported oligosaccharide solute binding proteins, confirms the ligand binding to be driven by enthalpy with a negative entropic contributing (Table 4). An overall, albeit weak, tendency of the entropic contribution being one third the enthalpic, which was largely unchanged, for trisaccharide binding, was also observed for solute binding proteins being mono-specific, however as for panose binding a larger entropy was measured. This implies the apparent lower affinity of *Bl*16GBP to be enforced by an increase in entropy indicating how the dual substrate specificity is affecting the mechanism of oligosaccharide binding.

The identification of a putative RFO/IMO specific ABC transporter from the probiotic *B. lactis* Bl-04 has been further pursued by functional analysis. Screening of proposed carbohydrate ligands by SPR revealed novel dual substrate specificity by binding of RFO (DP2–4) and IMO (DP2–7) with binding affinities in the higher μM range markedly weaker than previous observations of oligosaccharide solute binding proteins in the sub μM range. The thermodynamics of ligand binding by ITC proposed the reduced affinity to be linked to greater entropic contributions. This was supported by sequence comparison to a structurally determined solute binding protein where corresponding identified key residues in the substrate binding cleft indicated plasticity of the *Bl*16GBP binding cleft. Comparative Gene landscape analysis revealed the substrate specific solute binding protein to be co-encoded with both GH36 and GH13 enzymes and taxonomical predominantly restricted to species known to harbor probiotic strains thus enforcing the selective utilization of RFO and ISO as a requirement for prebiotic use by the strains of bifidobacteria and lactobacilli.

Materials and methods

Bioinformatics analysis

Phylogenetic analysis was done using ClustalW (43) and visualized using Dendroscope (44). All homology searching was done using BLAST (45).

Cloning of the *Bl*16GBP coding open reading frame

B. animalis subsp. *lactis* Bl-04 genomic DNA was used as template for the PCR amplification of the Balac_1599 open reading frame (Genbank gene ID: 8009526) with the following primer pair:

Forward: 5' GAATTC**CATATG**GGCAGCGGGCAGGTCACGCTC (Nde1 restriction site in bold)

Reverse: 5' CGC**GGATCC**CTACTTGCGGAAGTCACGAGCC (BamHi restriction site in bold)

The PCR amplicon (1201 basepairs), flanked by Nde1 and BamH1 restriction sites was constructed to include the natural stop-codon but excluding the signal peptide as predicted by SignalP (33), was ligated into the pET28a(+) vector (Novagen, Darmstadt, Germany) and transformed by heat shock into *E. coli* DH5α (Invitrogen, Carlsbad, CA, USA). Single colony clones were selected on LB agar plates with 50 µg/ml kanamycin and colonies harboring the engineered pET28a(+) with the *Bl*16GBP insert were confirmed by restriction analysis and sequencing using the T7 primer pair (Eurofins MWG, Ebersberg, Germany). Verified plasmid was transformed in *E. coli* BL21(DE3) for protein expression.

Production and purification of recombinant *Bl*16GBP

Recombinant *Bl*16GBP was produced in a 5 L Biostat B bioreactor (B. Braun Biotech International, Melsungen, Germany) according to a fed-batch protocol developed for the production of other proteins

(34), with the following modifications: heterologous expression was induced at 37 ºC when OD_{600} reached a value of 8 by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG; Sigma-Aldrich, St. Louis, MO, USA) to a final concentration of 100 µM. The cells were harvested after 18 h of induction by centrifugation (10.000 *g* at 4 °C for 15 min) and stored at -20 °C.

Cells were resuspended in Bugbuster (Novagen, Darmstadt, Germany) with Benzonase Nuclease treatment (Novagen) and incubated at 4 ºC for 1 hour, hereafter the suspension was centrifuged (43,000 g, 90 min) and sterile filtered (0.22 µm) before loading to a 5 ml HisTrap HP column (GE Healthcare, Uppsala, Sweden) as described elsewhere (34). The affinity purification was followed by anion exchange chromatography using an 8 ml Mono Q 10/100 GL column (GE Healthcare) equilibrated in 20 mM phosphate buffer, pH 7.0 (buffer A) and installed on an ÄKTA Advant™ chromatograph (GE Healthcare). Protein was elution by an increasing gradient of buffer A with 500 mM NaCl added over 10 column volumes.

The purity of eluted protein was validated by SDS-PAGE and fractions containing pure protein were pooled and buffer exchanged into 20 mM phosphate, pH 7.0. The pET28(a)+ encoded N-terminal hexa-his-tag was removed by incubation for 24 h at room temperature with thrombin, 1 U/100 µg *Bl*16GBP protein (Novagen). The reaction mixture was spun down and the supernatant was affinity purified as above with a one-step gradient over 5 column volumes where the flow-through contained the cleaved *Bl*16GBP, which was buffer exchanged into 20 mM phosphate, pH 7.0.

The protein concentration was determined by measuring A_{280} using a molar extinction coefficient ε_{280nm} = 51750 $M^{-1}cm^{-1}$ determined experimentally by aid of amino acid analysis (46), comparable to the predicted value of ε_{280nm} = 45380 $M^{-1}cm^{-1}$.

Carbohydrate ligands

Fructose, galactose, glucose, sucrose, maltose, maltotriose, maltotetraose, maltopentaose, maltohexaose, maltoheptaose were all purchased from Sigma. XOS was obtained from Shandong Longlive Bio-technology Co., Ltd, (China). Cellobiose was from Fluka AG (Switzerland). GOS and FOS were kindly supplied from DuPont Health and Nutrition as custom preparations.

Differential scanning calorimetry

*Bl*16GBP (0.5 mg/mL) was dialyzed against 1000 volumes of either 20 mM MES, pH 6.5 or in the same buffer including 2 mM raffinose to assess possible stabilization of protein in the ligand bound form, and degassed for 10 min at 20 °C. DSC analysis was performed using a VP-DSC calorimeter (MicroCal, Northampton, MA, USA) with a cell volume of 0.52061 mL at a scan rate of 1 °C·min^{-1}. Baseline scans collected with buffer in the reference and sample cells were subtracted from sample scans. Origin v7.038 software with a DSC add-on module was used for assigning T_m.

Surface plasmon resonance

Surface plasmin resonance (SPR) analysis was performed using a Biacore T100 (GE Healthcare) and 100 µg/ml *Bl*16GBP in 10 mM sodiumacetate pH 4.5 was immobilized onto a CM5 chip (GE Healthcare) using a standard amine coupling protocol before aiming for an immobilization level of 4000 response units. Binding studies were carried out at 25 °C in a 20 mM phosphate pH 7.0, 150 mM sodium chloride, 0.005% (v/v) P20 (GE Healthcare) running buffer unless otherwise stated and all solutions were filtered prior to analysis (0.22 µm). In the initial screening of binding activity, carbohydrate ligands were dissolved in the running buffer to a final concentration of 1 mM and injected over the chip surface at a flow of 30 µL/min with association times of 90 and 180 s, respectively.

Carbohydrate ligands, which displayed binding affinity towards the immobilized protein, were further analyzed (Table S1) using the above flow rate, contact and dissociation time. Binding of IMO ligands were tested at 10 concentrations (3.9–1000 μM), whereas melibiose and stachyose was tested at 10 concentrations (3–1600 μM) and (8–4000 μM), respectively. Raffinose binding was measured at 13 concentrations (0.24–1000 μM). Binding of verbascose was done using the same conditions as for stachyose but no saturation was observed. Binding affinities (K_D) were fitted to a one-binding site model (Biacore Evaluation software, GE Healthcare) to binding levels measured for each carbohydrate in triplicates.

The pH dependence of binding was measured using raffinose at four concentrations (4–250 μM) in 20 mM sodium acetate pH 4.0–5.5) or 20 mM sodium citrate-phosphate (5.5–8.0) and a similar NaCl and surfactant concentration as above in both buffers. The temperature dependence of raffinose binding (13 concentrations as in range above) to *Bl*16GBP was analyzed by measuring the K_D at eight temperatures in the range 15–43 °C and the energetic of binding were determined using linear van't Hoff analysis using the Biacore Evaluation software.

Isothermal titration calorimetry

The affinity of raffinose and panose to *Bl*16GBP was determined by isothermal titration calorimetry (iTC200, GE healthcare). Titrations were conducted in triplicates at 25°C by injection of 2 mM raffinose or panose into 108 μM *Bl*16GBP in 20 mM sodium citrate-phosphate pH 7.0. ITC heat trace profiles are shown in figure 2 for raffinose and panose. Injection of panose and raffinose into buffer was performed for measuring the heat of dilution which was subtracted in the data analysis. The ITC experiments include a pre-injection of 0.5 μL, which was discarded from the analysis, followed by 38·1μL injections into an ITC cell volume of 204 μL. The ITC heat trace was processed as described previously (47) and fitted to a single site binding model governed by the equilibrium association constant, K_D, the molar enthalpy for binding, ΔH, and the average number of binding sites on the protein n.

References:

1. Reid G, *et al* (2003) New scientific paradigms for probiotics and prebiotics. *J Clin Gastroenterol* 37: 105–118.

2. de Vrese M, Kristen H, Rautenberg P, Laue C & Schrezenmeir J (2011) Probiotic lactobacilli and bifidobacteria in a fermented milk product with added fruit preparation reduce antibiotic associated diarrhea and *Helicobacter pylori* activity. *J Dairy Res* 78: 396–403.

3. Guandalini S (2011) Probiotics for prevention and treatment of diarrhea. *J Clin Gastroenterol* 45: S149–153.

4. Ringel-Kulka T, *et al* (2011) Probiotic bacteria *Lactobacillus acidophilus* NCFM and *Bifidobacterium lactis* bi-07 versus placebo for the symptoms of bloating in patients with functional bowel disorders: A double-blind study. *J Clin Gastroenterol* 45: 518–525.

5. Szajewska H, Guandalini S, Morelli L, Van Goudoever JB & Walker A (2010) Effect of *Bifidobacterium animalis* subsp. *lactis* supplementation in preterm infants: A systematic review of randomized controlled trials. *J Pediatr Gastroenterol Nutr* 51: 203–209.

6. Ouwehand AC, *et al* (2008) *Bifidobacterium* microbiota and parameters of immune function in elderly subjects. *FEMS Immunol Med Microbiol* 53: 18–25.

7. Turroni F, van Sinderen D & Ventura M (2011) Genomics and ecological overview of the genus *Bifidobacterium*. *Int J Food Microbiol* 149: 37–44.

8. Lee J & O'Sullivan D,J. (2010) Genomic insights into *Bifidobacteria*. *Microbiology and Molecular Biology Reviews* 74: 378–416.

9. Locascio RG, Desai P, Sela DA, Weimer B & Mills DA (2010) Comparative genomic hybridization of *Bifidobacterium longum* strains reveals broad conservation of milk utilization genes in subsp. *infantis*. *Appl Environ Microbiol* 76:7373–7381

10. Sela DA, *et al* (2008) The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. *Proc Natl Acad Sci U S A* 105: 18964–18969.

11. Schell Ma, *et al* (2002) The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc Natl Acad Sci U S A* 99: 14422–14427.

12. Roberfroid M (2007) Prebiotics: The concept revisited. *J Nutr* 137: S830–S837.

13. Davis LM, Martinez I, Walter J, Goin C & Hutkins RW (2011) Barcoded pyrosequencing reveals that consumption of galactooligosaccharides results in a highly specific bifidogenic response in humans. *PLoS One* 6: e25200.

14. Costabile A, *et al* (2010) A double-blind, placebo-controlled, cross-over study to establish the bifidogenic effect of a very-long-chain inulin extracted from globe artichoke (*Cynara scolymus*) in healthy human subjects. *Br J Nutr* 104: 1007–1017.

15. Kukkonen K, *et al* (2007) Probiotics and prebiotic galacto-oligosaccharides in the prevention of allergic diseases: A randomized, double-blind, placebo-controlled trial. *J Allergy Clin Immunol* 119: 192–198.

16. Pokusaeva K, Fitzgerald GF & van Sinderen D (2011) Carbohydrate metabolism in bifidobacteria. *Genes & Nutrition* 6: 285–306.

17. Ventura M, *et al* (2009) Genome-scale analyses of health-promoting bacteria: Probiogenomics. *Nat Rev Microbiol* 7: 61–71.

18. Flint HJ & Bayer EA (2008) Plant cell wall breakdown by anaerobic microorganisms from the mammalian digestive tract. *Ann N Y Acad Sci* 1125: 280–288.

19. Flint HJ, Duncan SH, Scott KP & Louis P (2007) Interactions and competition within the microbial community of the human colon: Links between diet and health. *Environ Microbiol* 9: 1101–1111.

20. Gilad O, *et al* (2010) Combined transcriptome and proteome analysis of *Bifidobacterium animalis* subsp. *lactis* BB-12 grown on xylo-oligosaccharides and a model of their utilization. *Appl Environ Microbiol* 76: 7285–7291.

21. Parche S, *et al* (2007) Sugar transport systems of *Bifidobacterium longum* NCC2705. *J Mol Microbiol Biotechnol* 12: 9–19.

22. LoCascio RG, Desai P, Sela Da, Weimer B & Mills Da (2010) Broad conservation of milk utilization genes in *Bifidobacterium longum* subsp. *infantis* as revealed by comparative genomic hybridization. *Appl Environ Microbiol* 76: 7373–7381.

23. Broekaert WF, *et al* (2011) Prebiotic and other health-related effects of cereal-derived arabinoxylans, arabinoxylan-oligosaccharides, and xylooligosaccharides. *Crit Rev Food Sci Nutr* 51: 178–194.

24. Davidson AL, Dassa E, Orelle C & Chen J (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev* 72: 317–364.

25. Eitinger T, Rodionov Da, Grote M & Schneider E (2011) Canonical and ECF-type ATP-binding cassette importers in prokaryotes: Diversity in modular organization and cellular functions. *FEMS Microbiol Rev* 35: 3–67.

26. Berntsson RP-, Smits SHJ, Schmitt L, Slotboom D & Poolman B (2010) A structural classification of substrate-binding proteins. *FEBS Lett* 584: 2606–2617.

27. Suzuki R, *et al* (2008) Structural and thermodynamic analyses of solute-binding protein from *Bifidobacterium longum* specific for core 1 disaccharide and lacto-N-biose I. *J Biol Chem* 283: 13165–13173.

28. Barrangou R, *et al* (2009) Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and Bl-04. *J Bacteriol* 191: 4144–4151.

29. Hayakawa K, *et al* (1990) Effects of soybean oligosaccharides on human faecal flora. *Microb Ecol Health Dis* 3: 293–303.

30. Tenorio MD, Espinosa-Martos I, Préstamo G, Rupérez P (2010) Soybean whey enhance mineral balance and caecal fermentation in rats. *Eur J Nutr* 49: 155–163.

31. Yen CH, Tseng YH, Kuo YW, Lee MC & Chen HL (2011) Long-term supplementation of isomalto-oligosaccharides improved colonic microflora profile, bowel function, and blood cholesterol levels in constipated elderly people - a placebo-controlled, diet-controlled trial. *Nutrition* 27: 445–450.

32. Goffin D, *et al* (2011) Will isomalto-oligosaccharides, a well-established functional food in Asia, break through the European and American market? The status of knowledge on these prebiotics. *Crit Rev Food Sci Nutr* 51: 394–409.

33. Petersen TN, Brunak S, von Heijne G & Nielsen H (2011) SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat Methods* 8: 785–786.

34. Fredslund F, *et al* (2011) Crystal structure of α-galactosidase from *Lactobacillus acidophilus* NCFM: Insight into tetramer formation and substrate binding. *J Mol Biol* 412: 466–480.

35. Goulas T, Goulas A, Tzortzis G & Gibson GR (2009) A novel α-galactosidase from *Bifidobacterium bifidum* with transgalactosylating properties: Gene molecular cloning and heterologous expression. *Appl Microbiol Biotechnol* 82: 471–477.

36. Pokusaeva K, O'Connell-Motherway M, Zomer A, Fitzgerald GF & van Sinderen D (2009) Characterization of two novel α-glucosidases from *Bifidobacterium breve* UCC2003. *Appl Environ Microbiol* 75: 1135–1143.

37. Chan S, *et al* (2010) Apo and ligand-bound structures of ModA from the archaeon *Methanosarcina acetivorans*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 66: 242–250.

38. Dam TK & Brewer CF (2002) Thermodynamic studies of lectin-carbohydrate interactions by Isothermal Titration Calorimetry. *Chem Rev* 102: 387–429.

39. Barrangou R, Altermann E, Hutkins R, Cano R & Klaenhammer TR (2003) Functional and comparative genomic analyses of an operon involved in fructooligosaccharide utilization by *Lactobacillus acidophilus*. *Proc Natl Acad Sci U S A* 100: 8957–8962.

40. Andersen JM, *et al* (2011) Transcriptional and functional analysis of galactooligosaccharide uptake by *lacS* in *Lactobacillus acidophilus*. *Proc Natl Acad Sci U S A* 108: 17785–17790.

41. Gibson GR, Probert HM, Loo JV, Rastall RA & Roberfroid MB (2004) Dietary modulation of the human colonic microbiota: Updating the concept of prebiotics. *Nutr Res Rev* 17: 259–275.

42. Ajdić D & Pham VT (2007) Global transcriptional analysis of *Streptococcus mutans* sugar transporters using microarrays. *J Bacteriol* 189: 5049–5059.

43. Larkin MA, *et al* (2007) Clustal W and clustal X version 2.0. *Bioinformatics* 23: 2947–2948.

44. Huson DH, *et al* (2007) Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 8: 460.

45. Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410.

46. Barkholt V & Jensen AL (1989) Amino acid analysis: Determination of cysteine plus half-cystine in proteins after hydrochloric acid hydrolysis with a disulfide compound as additive. *Anal Biochem* 177: 318–322.

47. Henriksen JR, Andresen TL, Feldborg LN, Duelund L & Ipsen JH (2010) Understanding detergent effects on lipid membranes: A model study of lysolipids. *Biophys J* 98: 2199–2205.

48. Wei X, *et al* (2012) Fructose uptake in bifidobacterium longum NCC2705 is mediated by an ATP-binding cassette transporter. *The Journal of Biological Chemistry* 287: 357–367.

49. Stam MR, Danchin EGJ, Rancurel C, Coutinho PM & Henrissat B (2006) Dividing the large glycoside hydrolase family 13 into subfamilies: Towards improved functional annotations of α-amylase-related proteins. *Protein Engineering, Design & Selection: PEDS* 19: 555–562.

50. Abbott DW, *et al* (2010) The molecular basis of glycogen breakdown and transport in *Streptococcus pneumoniae*. *Mol Microbiol* 77: 183–199.

51. Shulami S, *et al* (2011) The L-arabinan utilization system of *Geobacillus stearothermophilus*. *J Bacteriol* 193: 2838–2850.

52. Higgins Ma, Abbott DW, Boulanger MJ & Boraston AB (2009) Blood group antigen recognition by a solute-binding protein from a serotype 3 strain of *Streptococcus pneumoniae*. *J Mol Biol* 388: 299–309.

Figures:

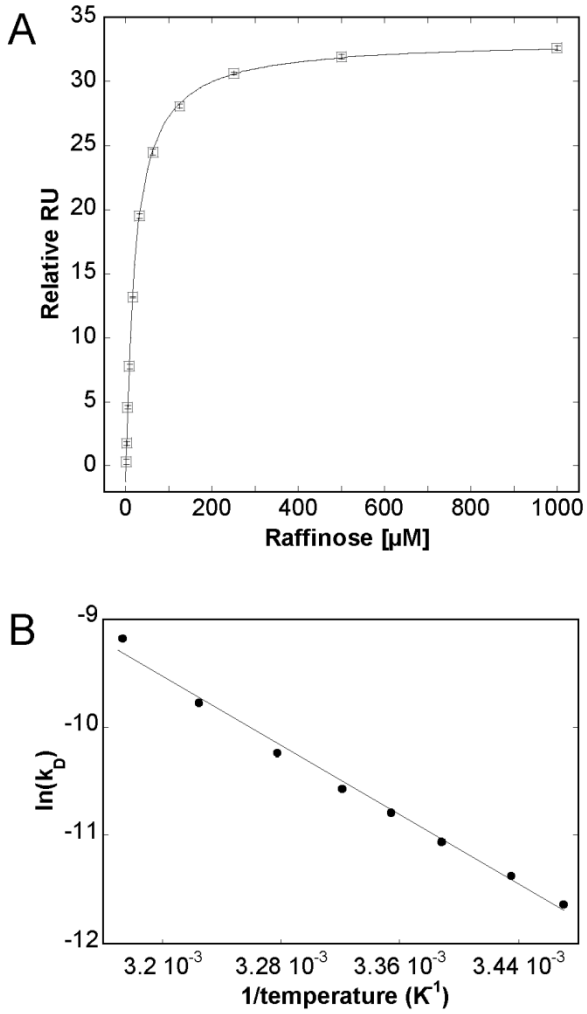


Figure 1: SPR quantified binding of raffinose to *Bl*16GBP. A: The relative response units as function of raffinose concentration (□) fitted to a 1:1 binding model shown with error bars. B: The temperature dependence of raffinose binding to *Bl*16GBP depicted by a van't Hoff plot.

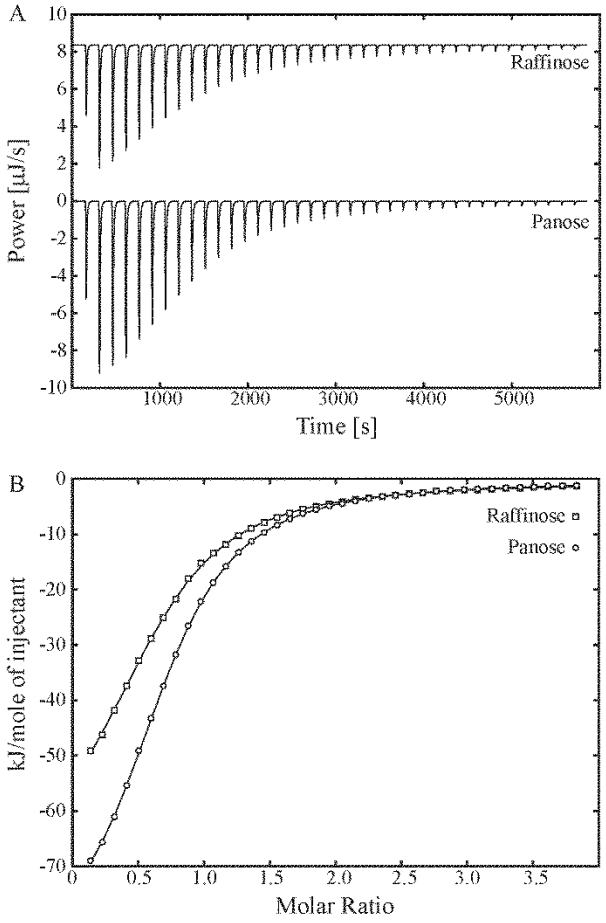


Figure 2: ITC study of the binding affinity of panose and raffinose to *Bl*16GBP. A: Representative ITC heat traces of 2mM panose and raffinose titrated into 108 µM *Bl*16GBP. B: Integrated corresponding heat of reaction to a single site binding model where the best fits are shown by solid lines. The experiments were performed at 25°C by 38 injection of 1 µL.

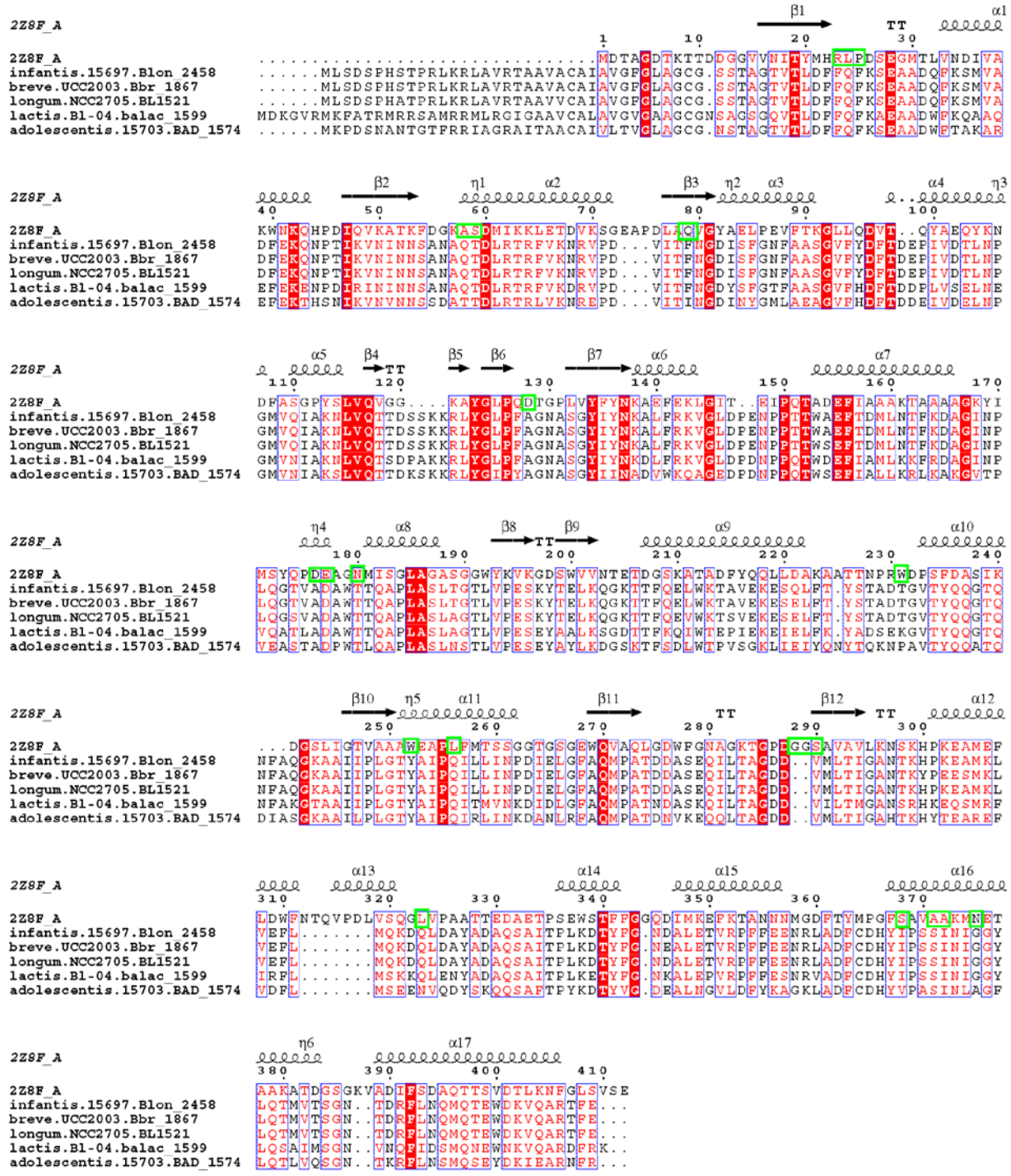


Figure 3: Multiple sequence alignment of putative functional homologs of *Bl*16GBP and the lacto-*N*-biose binding protein (GL-BP) from *B. longum* (PDB: 2Z8F). Conserved residues are highlighted in red and all residues within 4 Å of the bound ligand (lacto-*N*-tetraose) have been highlighted in green boxes in the in

GL-BP (2Z8F) sequence. Secondary structure extracted from the GL-BP structure has been represented using ESPript (http://espript.ibcp.fr/ESPript/ESPript) and the alignment was constructed using ClustalW (43). The *Bl*16GBP homologs are identified by their locus tags: Blon_2458, Bbr_1867, BL1521 and BAD_1574.

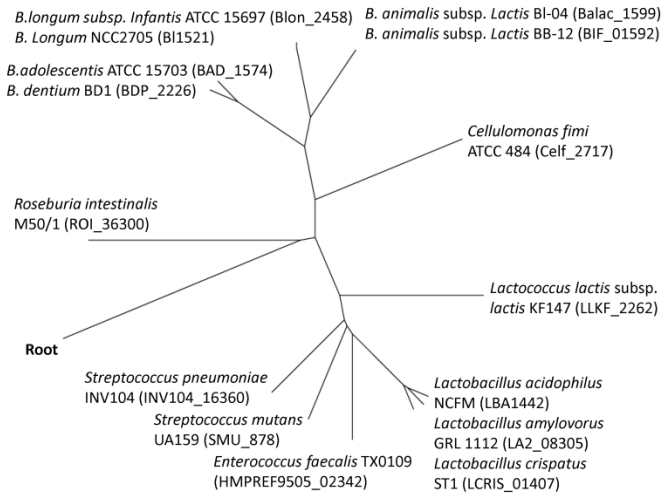


Figure 4: Phylogenetic map of putative RFO/IMO specific solute binding proteins representing taxonomical clusters harboring gene clusters for RFO/IMO utilization. All entries are shown by the full strain name and the corresponding solute binding protein locus tag in brackets. The tree has been rooted using a *B. longum* fructose solute binding protein (48).

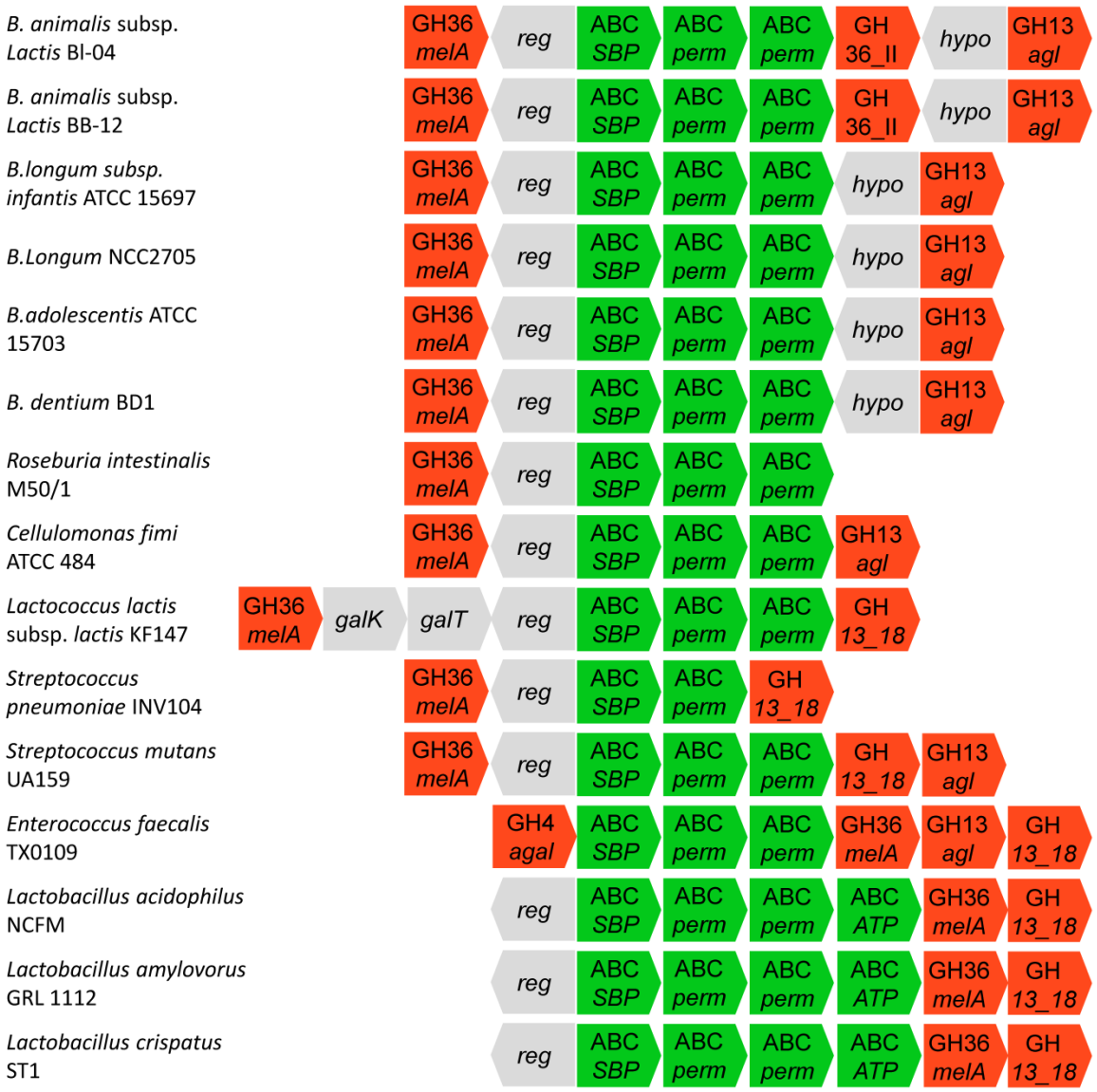


Figure 5: Gene landscape analysis of representative gene clusters encoding RFO/IMO solute binding proteins. The gene clusters have been aligned according to their solute binding proteins. Gene functions are colored as glycoside hydrolases in red where GH36 *melA* refer to Balac_1601 homologs, GH36_II refers to Balac_1596 homologs, GH4 *agal* refers to α-galactosidase homolog, GH13 *ag1* refers to Balac_1593 homologs and GH13_18 refers to subgroup 18 of GH13 encoding sucrose phosphorylases (49). Genes in green represents the components of ABC transporter by ABC SPB being the solute binding protein, ABC perm being the permeases and ABC ATP being the ATP binding kinases. Genes in grey refer to transcriptional regulators (reg), hypothetical proteins (hypo) and genes of the Leloir Pathway:

galactose kinase (*galK*) and galactose-1-phosphate uridylyltransferase (*galT*). Putative ORF less than 100 amino acids were not shown.

Tables:

Table 1: Substrate specificities of *Bl*16GBP measured by SPR. Dissociation constants (K_D) were measured as triplicates and listed with the standard errors together with the modeled maximal binding and Chi^2-estimate for the fitted one-binding site model. Isomaltooligosaccharides are denoted as IM where the number state the degree of polymerization.

Raffinose-like carbohydrates				
Carbohydrate	**K_D (µM)**	**SE (µM)**	**R_{max}**	**Chi^2**
Melibiose	729	72	24.21	0.219
Raffinose	20.7	0.36	34.81	0.115
Stachyose	327	11	37.2	0.268
Verbascose	>4000	-	54.7	0.382
Isomaltooligosaccharides				
Carbohydrate	**K_D (µM)**	**SE (µM)**	**R_{max}**	**Chi^2**
IM2	1059	73.00	19.11	0.0578
IM3	126.4	2.80	32.97	0.699
Panose	8.7	0.15	34.39	0.11
IM4	93.9	1.10	41.87	0.05
IM5	102.7	1.30	50.96	0.0734
IM6	103.9	1.10	56.38	0.0628
IM7	142.5	0.96	68.53	0.0383

Table 2: The influence of pH upon binding of raffinose to *Bl*16GBP.

pH	K_D (µM)	R_{max}	SE (µM)
4.5	62.2	14.4	2.30
5.5	21.2	25	0.08
6.0	19.7	25.1	0.07
6.5	23.5	25.1	1.46
7.0	20.2	22.5	2.40
7.5	16.5	25.2	0.71
8.0	17.9	24.98	0.73

Table 3: Dissociation constants and thermodynamics of panose and raffinose binding to *Bl*16GBP measured by ITC.

	K_D [µM]	ΔG[kJ/mol]	ΔH [kJ/mol]	-TΔS[kJ/mol]	n
Panose	17.5 ± 0.3	-27.1	-63.7 ± 0.5	36.5	0.70 ± 0.01
Raffinose	27 ± 2	-26.1	-46.3 ± 0.1	20.2	0.60 ± 0.02

Table 4: Comparison of thermodynamic finger-prints for trisaccharide binding by oligosaccharide solute binding proteins.

Strain	**Ligand**	K_D **[µM]**	**ΔG [kcal/mol]**	**ΔH [kcal/mol]**	**-TΔS [kcal/mol]**	**Reference**
Bifidobacterium animalis subsp. *lactis*	Raffinose	27.0	-6.49	-11.33	4.84	This study
	Panose	17.5	-6.23	-14.98	8.75	This study
Streptococcus pneumonia	Maltotriose	0.51	-7.65	-6.32	1.33	(50)
Geobacillus stearothermophilus	Arabinotriose	0.22	-9.2	-14.3	5.1	(51)
Streptococcus pneumonia	Blood group antigen A [GalNac-α(1,3)-Fuc-α(1,2)-Gal]	1.01	-8.21	-9.2	0.99	(52)
	Blood group antigen B [Gal-α(1,3)-Fuc-α(1,2)-Gal]	1.15	-8.26	-8.4	0.14	(52)

Supplementary material:

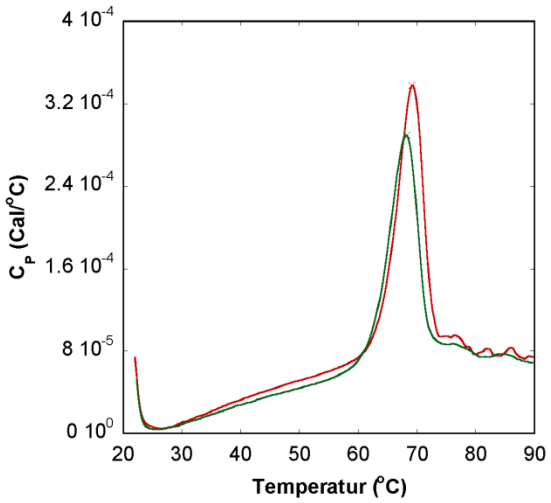


Figure S1: Differential scanning calorimetry of *Bl*16GBP. Unbound *Bl*16GBP is shown in green and *Bl*16GBP in the presence of 2 mM raffinose is shown in red.

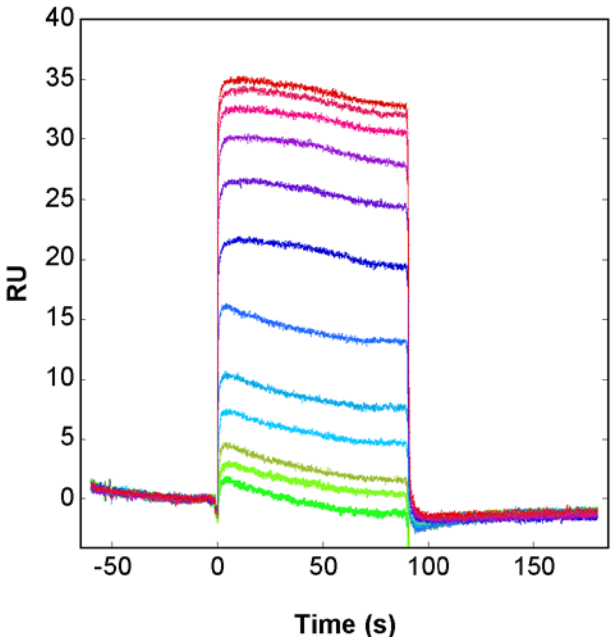


Figure S2: SPR binding curves raffinose binding to *Bl*16GBP at 25 °C. The sensorgram is shown with baseline stabilization (-60:0 s), measurement of response unit by raffinose concentrations (0.49–1000 µM) (0:90 s) and raffinose wash-off with baseline stabilization (90–180 s). The increase in raffinose concentration is illustrated by the color gradient leading from green (0.49 µM) to red (1000 µM).

Table S1: Carbohydrates used for SPR and ITC measurements. All carbohydrates are listed with chemical structure, Manufacturer and purity.

Carbohydrate	**Structure**	**Manufacturer or supplier**	**Purity (as given by Manufacturer or supplier)**
Melibiose	α-D-Gal*p*-(1–6)-D-Glc*p*	Sigma	> 98%
Raffinose family oligosaccharides	[α-D-Gal*p*-(1–6)]$_a$-D-Glc*p*-(α1,β2)-D-Fru*f*	Sigma	> 99%
Panose	α-D-Glc*p*(1–6)-α-D-Glc*p*-(1–4)-D-Glc*p*	Carbosynth Ltd. (UK)	> 98%
Isomaltooligo-saccharides	[α-D-Glc*p*-(1–6)]$_b$-D-Glc*p*	Kind gift from Professor Atsuo Kimura, Hokkaido University (Japan)	-

[a] Lists the three main types of raffinose family oligosaccharides. For a=1: Raffinose, for a=2: stachyose and for a=3: verbascose.

[b] Lists the distribution of isomaltooligosaccharides through isomaltose (b=1) till isomaltoheptaose (b=6).

6.5 Ongoing collaborative work

This appendix outlines a small activity of the experimental work initiated as part of this Ph.D. project, which is not completed at the present time. This is a collaboration with Professor Hanne Frøkiær, Department of Basic Sciences and Environment, Faculty of Life Sciences, University of Copenhagen.

Carbohydrate dependent immuno-modulation by *Lactobacillus acidophilus* NCFM stimulated with oligosaccharides

Probiotic microorganisms interact with the host immune system by modulating the immune response when degraded by dendritic cells (1). The immune modulation is largely screened by the interleukins (IL) 10 and 12 produced by the dendritic cells as part of the innate immune system, where the ratio signals a pro-inflammatory response by increased IL-12 and anti-inflammatory response by increased IL-10. The dendritic response to probiotic lactobacilli and bifidobacteria differ on the strain level (2) and functional studies have proposed lipoteichoic acid (3) and S-layer proteins (4) to affect the IL-profile and interactions with dendritic cells, respectively.

From the current project, it was indicated how the different utilized glycosides affected the applied probiotics *Lactobacillus acidophilus* NCFM and *Bifidobacterium animalis* subsp. *lactis* Bl-04. Building on this, it was hypothesized that the changed catabolism could lead to metabolic changes of factors such as cell membrane lipid or cell wall glycoside compositions, which could potentially change the innate immune response.

For the experimental design *L. acidophilus* NCFM was selected for dendritic stimulation based on prior knowledge of the bacteria and the induced immune response pathway through Toll-like receptors (5).

L. acidophilus NCFM cultures (harvested in the stationary phase) were prepared in semi-defined media (6) supplemented with either glucose as a control carbohydrate, cellobiose (β-D-Glc*p*-(1–4)-D-Glc*p*) as a β-glucoside representative), lactose (β-D-Gal*p*-(1–4)-D-Glc) as a β-galactoside representative and raffinose (α-D-Gal*p*-(1–6)-D-Glc*p*-(α1,β2)-D-Fru*f*) as a α-galactoside representative.

Dose-response experiments have been performed using murine dendritic cells measuring the IL-10 and IL-12 productions, to determine the concentration of *L. acidophilus* NCFM applied based on OD_{600} measurements and confirmed by cell counts. Cultures have been grown at DTU and all dendritic and enterocyte experiments have been performed by technician Anni Mehlsen at Copenhagen University in our collaboration with Professor Hanne Frøkiær.

Preliminary results showed a reduced IL-12 profile for *L. acidophilus* NCFM when grown on cellobiose and lactose. Currently these results are being further analyzed by realtime quantitative-PCR to deduce the intracellular pathway changes in dendritic cells that reflect the mechanism of IL-12 changes compared to earlier work (5).

In conclusion, this on-going work will add to the understanding of factors underlying the immune response of probiotics upon the host and may further highlight the importance of prebiotic induced changes in carbohydrate metabolism of probiotic microorganisms.

References:

1. Bron PA, van Baarlen P & Kleerebezem M (2011) Emerging molecular insights into the interaction between probiotics and the host intestinal mucosa. *Nat Rev Microbiol* 10: 66-78.

2. Weiss G, *et al* (2011) Lactobacilli and bifidobacteria induce differential interferon-β profiles in dendritic cells. *Cytokine* 56: 520-530.

3. Mohamadzadeh M, *et al* (2011) Regulation of induced colonic inflammation by *Lactobacillus acidophilus* deficient in lipoteichoic acid. *Proc Natl Acad Sci U S A* 108 Suppl 1: 4623-4630.

4. Konstantinov SR, *et al* (2008) S layer protein A of *Lactobacillus acidophilus* NCFM regulates immature dendritic cell and T cell functions. *Proc Natl Acad Sci U S A* 105: 19474-19479.

5. Weiss G, *et al* (2010) *Lactobacillus acidophilus* induces virus immune defence genes in murine dendritic cells by a toll-like receptor-2-dependent mechanism. *Immunology* 131: 268-281.

6. Barrangou R, Altermann E, Hutkins R, Cano R & Klaenhammer TR (2003) Functional and comparative genomic analyses of an operon involved in fructooligosaccharide utilization by *Lactobacillus acidophilus*. *Proc Natl Acad Sci U S A* 100: 8957-8962.

6.6 Posters contributions

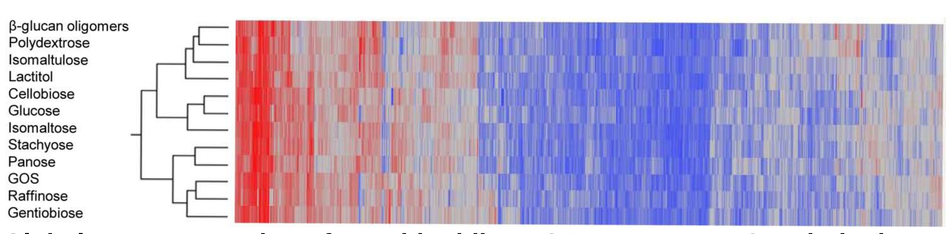
Prebiotic galacto-oligosaccharide utilization by *Lactobacillus acidophilus* NCFM. Establishment of a methodological platform for protein discovery.

Andersen, J.M.[1,4*]; Majumder, A.[1]; Fredslund, F.[1]; Ejby, M. [1]; van Zanten, G.C.[1,2] ;Barrangou, R.[3]; Goh, Y.J.[4]; Lahtinen, S.J.[5]; Lo Leggio, L.[6]; Coutinho, P.M.[7] Jacobsen, S.[1]; Abou Hachem, M.[1]; Klaenhammer, T.[4]; Svensson, B.[1]

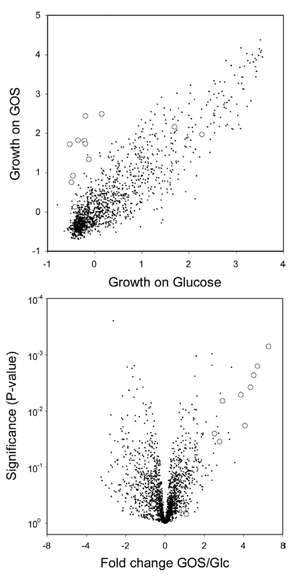
[1]Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark. [2] Department of Food Science, Faculty of Life Sciences, University of Copenhagen, Denmark. [3]DuPont Nutrition and Health, Madison, WI. USA [4]North Carolina State University, Raleigh, NC, USA. [5]DuPont Nutrition and Health, Kantvik, Finland. [6]Department of Chemistry, University of Copenhagen, Denmark. [7]CNRS, Université d'Aix Marseille, France.

CASE*:* *Lactobacillus acidophilus* NCFM is a documented probiotic able to utilize the prebiotics β-galacto-oligosaccharides and raffinose family oligosaccharides. The specific pathways for potential prebiotics remains to be characterized to advance the understanding of selective metabolism of probiotics.
AIM: Identify single genes and their protein products involved with prebiotic utilization and characterize key proteins for prebiotic uptake and catabolism.
METHODS: Global transcriptional analysis, DIGE-proteomics, *in silico* pathway reconstruction, functional genomics, recombinant protein characterization.
OUTPUT: Mapping of transporters and full catabolic pathways and validation by gene deletion mutants. Key protein molecular architectural understanding.

Transcriptional screening[1]



Global gene expression of *L. acidophilus* NCFM grown on 12 carbohydrates



GOS transcriptome

The global transcriptome influenced by various potential prebiotic (**top**) was measured.

A mixed model ANOVA was applied for data analysis of the global transcriptome, resulting in defined gene clusters putatively involved with carbohydrate uptake and catabolism being upregulated.

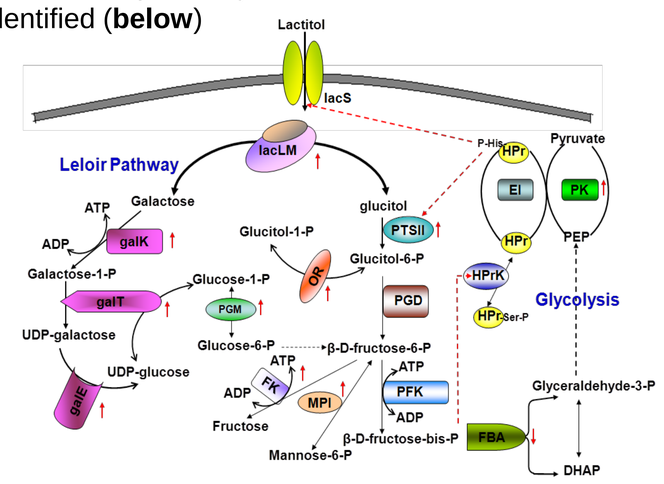
Differential transcriptomics (**left**) of β-galacto-oligo-saccharides (GOS) *versus* glucose identified a lactose permease of the glycoside-pentoside-hexuronide type and a GH2 and GH42 β-galactoside.

By the same approach, raffinose and stachyose induced an ATP binding cassette (ABC) transporter and GH36 a α-galactosidase, a key hydrolytic enzyme as presented below.

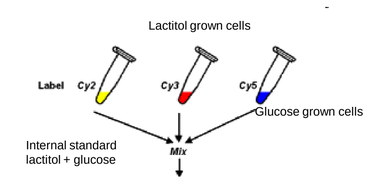
Differential proteomics[2]

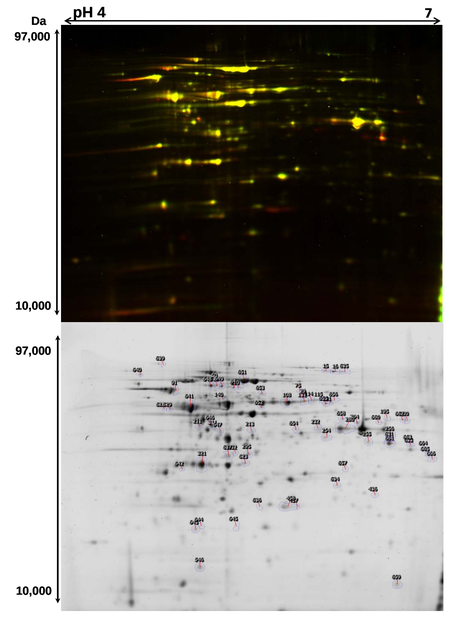
Initially, we established the reference proteome of *L. acidophilus* NCFM with 625 proteins identified, yielding knowledge of the main intracellular processes and formed the basis for differential in gel electrophoresis (DIGE) proteomics.

L. acidophilus NCFM were grown on the prebiotic lactitol and harvested in the late log phase. DIGE-proteomics (**left**) identified 62 proteins to be differentially expressed. The total intracellular path-way for lactitol catabolism were identified (**below**)



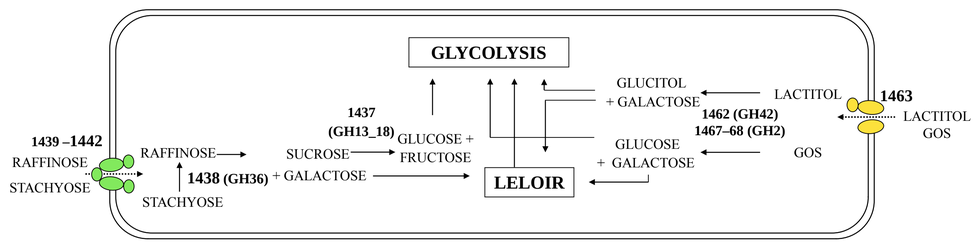
DIGE-principle:



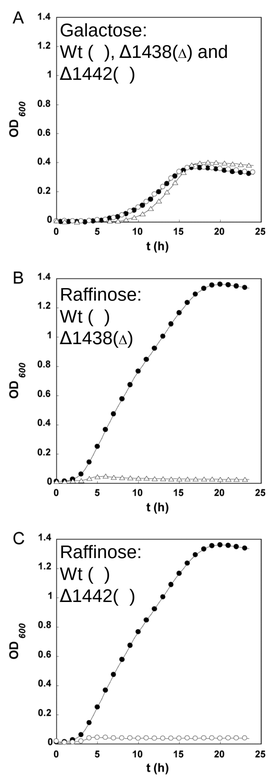


Labeled DIG proteome (top)
Reference lactitol map (below)

Functional genomics validation



***In silico* pathway reconstruction from transcriptional and proteomics findings**



By pathway analysis, key genes of interest were selected for gene deletion, with the resulting mutant phenotypes in the table **below**.

This validated the omics-based findings and revealed a broad substrate uptake profile of the GOS transporter, being the first identified *Lactobacillus* GOS transporter. Gene deletion within the raffinose pathway showed how a single gene can impact a ABC transporter (**left**)

Phenotypic characterization of key metabolic genes

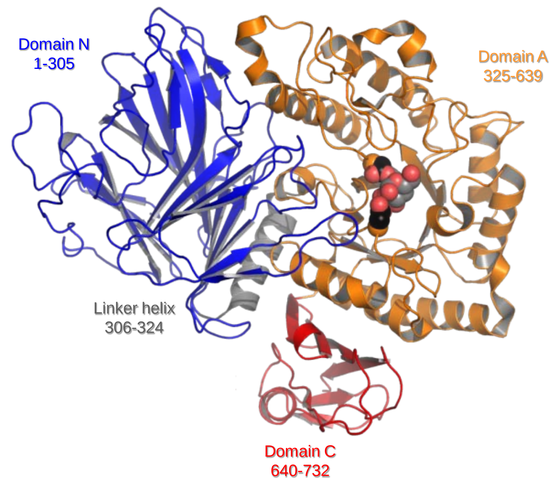
Gene locus	Function	No growth Phenotype	Chemical structure
LBA1438	α-galactosidase	Melibiose Raffinose Stachyose	α-D-Gal*p*-(1–6)-D-Glc*p* α-D-Gal*p*-(1–6)-D-Glc*p*-(α1,β2)-D-Fru*f* [α-D-Gal*p*-(1–6)]$_2$- D-Glc*p*-(α1,β2)-D-Fru*f*
LBA1442	Solute binding protein of ABC transporter	Melibiose Raffinose Stachyose	- - -
LBA1463	GPH permease	Lactose Lactitol GOS	β-D-Gal*p*-(1–4)-D-Glc β-D-Gal*p*-(1–4)-D-Glc-ol [β-D-Gal*p*-(1–4)]$_{1-5}$-D-Glc*p*

Protein structure-function relationship[3]

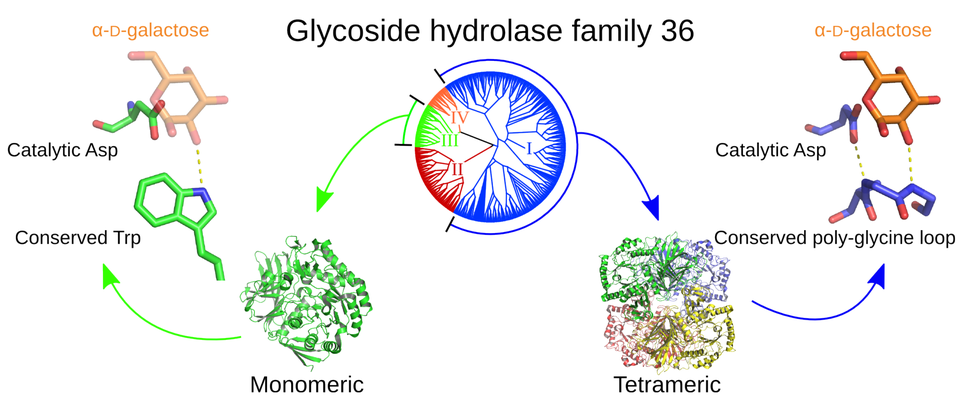
The key α-galactosidase (LBA1438, *La*MelA36A) was produced recombinant. The native enzymes was found as a tetramer and the structure was determined (**right,** only one monomer shown).

The active site topology revealed a tight pocket maintained through the tetramer interactions supporting the substrate specify found through gene deletion of LBA1438.

Sequence comparison within the glycoside hydrolase family (GH) 36, based on the *La*MelA36A structure, differentiated the family based on structural motifs relating to putative specificities (**below**).



Structure and domain organization of *La*MelA36A



Sequence clustering in GH36. Protein oligomerization change the active site accessibility.

Data integration and summary

The combination of the above methods and obtained results lead to in-depth molecular understanding of prebiotic utilization by *L. acidophilus* NCFM through identification of key proteins and their characterization.

This poster represents a methodological platform to generate data of commercial value and relate the results into a systems biological perspective through functional data and comparative sequence analysis.

References:

[1] Andersen et al. PNAS (2011) 108: 17785-17790 [2] Majumder et al. Proteomics (2011) 17: 3470–3481 [3] Fredslund et al. JMB (2011) 412: 466–480

Acknowledgements and funding

This work is supported by Danish Council for Strategic Research, Committee for Health, Food and Welfare, Danish Council for Independent Research Natural Sciences, Danisco USA and North Carolina Dairy Foundation, a HC-Ørsted Postdoctoral Fellowship (to AM), PhD stipends from the Technical University of Denmark (to JMA and ME) and Danisco A/S (to JMA).

Transcriptional Analysis of Prebiotic Utilization by *Lactobacillus acidophilus* NCFM

J. M. Andersen[1,3], R. Barrangou[2], M. Abou Hachem[1], B. Svensson[1], Y. Goh[3], T. R. Klaenhammer[3]

[1]Technical University of Denmark, Kgs. Lyngby, Denmark, [2]Danisco Inc., Madison, WI - USA, [3]North Carolina State Univ., Raleigh, NC - USA

Abstract

Probiotics microbes depend on their ability to survive in the gastrointestinal tract, adhere to mucosal surfaces, and metabolize available energy sources from non-digestible dietary compounds. We identified genetic loci in *Lactobacillus acidophilus* NCFM responsible for utilization of complex carbohydrates that may function as prebiotic substrates, *in vivo*. Whole genome oligonucleotide microarrays were used to survey the global transcriptome during logarithmic growth of *L. acidophilus* NCFM in the presence of 11 different carbohydrates (glucose, raffinose, cellobiose, panose, stachyose, isomaltose, gentiobiose, lactitol, β-glucan oligomers, polydextrose®, isomaltulose). The data were analyzed in JMP Genomics, using a mixed-model ANOVA. Specific transporters of the ATP-binding cassette (ABC), phosphotransferase system (PTS) and galactoside-pentose hexuronide (GPH) families were identified for the uptake of stachyose, cellobiose and lactitol, respectively. The identified genes were functionally validated by targeted gene deletion within an *in silico* reconstructed stachyose operon. We identified a series of genes that are responsible for the uptake and catabolism of a variety of potential prebiotic di- and oligo-saccharides in *L. acidophilus* NCFM.

Experimental starting point

Lactobacillus acidophilus NCFM is a proven probiotic bacterium commercially used in dietary supplements and fermented dairy products. Extensive Work has been done to understand the underlying molecular mechanisms of the probiotic effects, among others: bile tolerance, adherence to mucosal surfaces, mammalian host interactions and prebiotic utilization for pathogen exclusion.

The genome of *L. acidophilus NCFM* [1] encodes a significant part of transport systems and enzymatic machinery to process a wide array of complex carbohydrates, as summarized in table 1 and figure 1.

Yet the specific metabolic pathways for potential prebiotics remains to be identified and characterized.

Lactobacillus acidophilus NCFM	
Genome size	1.99 Mb
GC content	34,70%
ORFS	1,864
Carbohydrate metabolism	
ABC transporters	3
PEP-PTS systems	20
GPH permeases	2
Glycoside hydrolases	37
… of which are extracellular	2

Table 1: Summary of genes annotated with carbohydrate metabolism

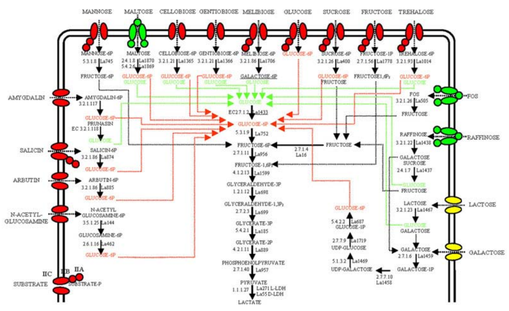


Figure 1: Predicted carbohydrate transport and intracellular metabolism by *L. acidophilus* NCFM

This study goes in depth with identification of genetic loci involved with prebiotic metabolism.

11 potential prebiotic carbohydrates (shown in table 2 with predicted glycoside hydrolase for intracellular breakdown) were selected for measuring the total transcriptome in response to each carbohydrate. Genes of interest were assessed by targeted gene deletions to confirm their role in carbohydrate metabolism.

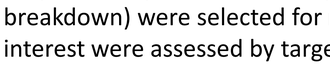
Carbohydrate	Structure	Predicted degrading enzymes
Glucose	**Glc**	Glycolytic pathway
Raffinose	**Gal-** *α1,6*-**Glc-** *α1,2*-**Fru**	α-galactosidase + sucrose phosphorylase
Gentiobiose	**Glc**-*β1,6*-**Glc**	6-phospho-β-glucosidases
Panose	**Glc**-*α1,6*-**Glc**-*α1,4*-**Glc**	α-1,6-glucosidases, maltose phosphorylase
Isomaltose	**Glc**-*α1,6*-**Glc**	α-1,6-glucosidases
Stachyose	**Gal**-*α1,6*-**Gal**-*α1,6*-**Glc**-*α1,2*-**Fru**	α-galactosidase + sucrose phosphorylase
Cellobiose	**Glc**- *β1,4*-**Glc**	phospho-β-glucosidases
Polydextrose	*All glucose linkages*	Various α-glucosidases
β-glucan	…**Glc**-*β1,3*-**Glc** (*β1,4*)…	β-glucanase + β-glucosidase
Isomaltulose	**Glc**-*α1,6*-**Fru**	α-1,6-glucosidases
Lactitol	*Sugar alcohol*	β-galactosidase or phospho-β-galactosidase

Table 2: Carbohydrates used in this study together with structural glycoside composition, O-linkages and the enzymes predicted to facilitate intracellular hydrolysis

In silico operon reconstruction

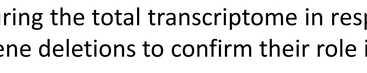
Gene clusters involved with carbohydrate metabolism were reconstructed from the identified up-regulated genes. This pictures the structure of genetic loci for potential prebiotic utilization in the Lactobacilli genus.

β-glucoside operon



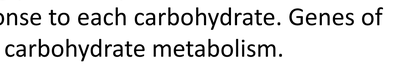
Locus tag	Putative function
LBA0724	Regulator
LBA0725	PTS CII Component
LBA0726	6-P-β-glucosidase

Stachyose operon



Locus tag	Putative function
LBA1443	Regulator
LBA1442	Solute binding, ABC
LBA1441	Permease, ABC
LBA1440	Permease, ABC
LBA1439	Kinase, ABC
LBA1438	α-galactosidase, GH36
LBA1437	Sucrose phos-phorylase, GH13_18

Lactose operon



Locus tag	Putative function
LBA1460	Surface protein
LBA1461	Regulator
LBA1462	β-galactosidase GH42
LBA1463	GPH permease
LBA1464	Transposase
LBA1465	Regulator
LBA1467	β-galactosidase, large subunit GH2
LBA1468	β-galactosidase, small subunit GH2
LBA1469	UDP-galactose epimerase

Figure 2: Reconstruction of putative prebiotic genetic operons based on DNA microarray data. Regulatory genes are shown in red, Transporter complexes in blue and glycoside hydrolases in green

References

[1] Altermann, E., Russell, M. W., Azcarate-Peril, A., Barrangou, R., Buck, L. B., McAuliffe, O., Souther, N., Dobson, A., Duong. Tri., Callanan, M., Lick, S., Hamrick, A., Cano, R. and Klaenhammer T.R. 2005 Complete genome sequence of the probiotic lactic acid bacterium Lactobacillus acidophilus NCFM. Proc.Natl. Acad. Sci. **102**: 3906-3912. **[2]** Goh, Y. J., Azcárate-Peril, A., O'Flaherty S., Durmaz E., Valence F., Jardin J., Lortal S., and Klaenhammer, T. R. 2009. Development and Application of a *upp*-Based Counterselective Gene Replacement System for the Study of the S-Layer Protein SlpX of *Lactobacillus acidophilus* NCFM. Appl. Environ. Microbiol. **75**: 3093-3105.

Acknowledgements and funding

This work is financially supported by Danisco USA and North Carolina Dairy Foundation. Joakim Mark Andersen is supported by a FøSu grant from the Danish Strategic Research Council to the project "Gene discovery and molecular interactions in prebiotics/probiotics systems. Focus on carbohydrate prebiotics".

Global transcriptome analysis

Total RNA were isolated, reverse transcribed and labeled with Cyanine3 and Cyanine5, two technical replicates for each condition. Hybridized probe intensities were background corrected and normalized before ANOVA modeling using JMP genomics 4.1. The global expression pattern was visualized by hierarchical clustering (figure 3) for the 11 growth conditions. Overall low variance between each condition was observed, correlating with regulation of few genetic loci in response to specific carbohydrate metabolism.

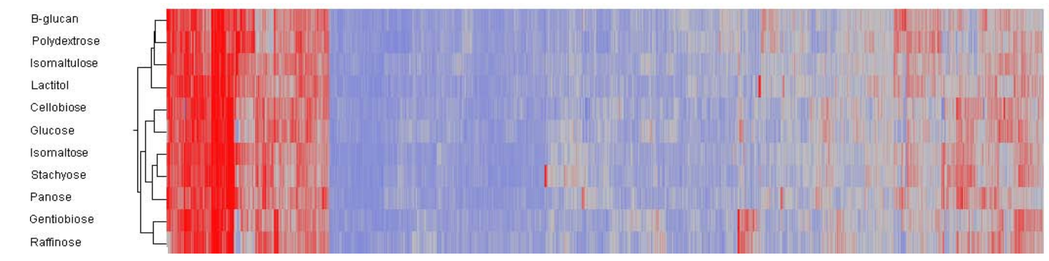


Figure 3: Rrepresentation of the hierarchical clustering of the global gene expression of *L. acidophilus* NCFM by carbohydrate. Red coloring indicates up regulation and blue coloring indicates down regulation of genes

Carbohydrate metabolism

Significant, differentially expressed genes were identified ($P < 10^{-2.75}$) and visualized by volcano plots (figure 3). Selected up-regulated genes involved with carbohydrate metabolism are shown as an expression heat map in figure 4 and listed with gene annotation and fold up-regulation in table 3. Several operon-like sets of genes were discovered, all including both a membrane transporter and at least one glycoside hydrolase.

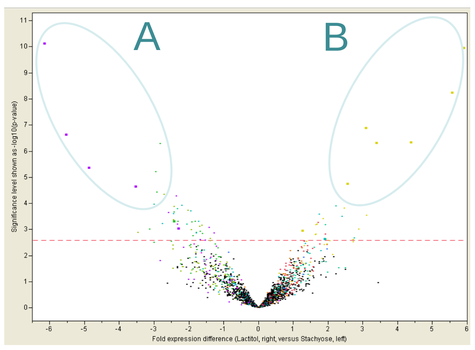


Figure 3: Volcano plot of the differential gene expression by carbohydrate stimulation. Up-regulated genes by stachyose, are circled at A while genes up-regulated by lactitol are circled at B

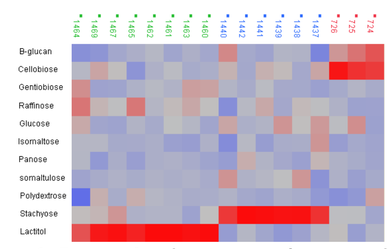


Figure 4: Expression heat map of genes related to metabolism of potential prebiotics.

Locus tag	Carbohydrate	Annotation	Fold induction
724	Cellobiose	Regulator	4,9
725	Cellobiose	PTS system IIC component	23,3
726	Cellobiose	6-P-β-glucosidase	12,1
1437	Stachyose	Sucrose phosphorylase	4,7
1438	Stachyose	α-galactosidase	15,1
1439	Stachyose	ABC transporter ATP-binding protein	18,1
1440	Stachyose	ABC transporter permease	3,2
1441	Stachyose	ABC transporter permease	7,6
1442	Stachyose	ABC transporter Solute-binding protein	53,1
1460	Lactitol	mucus binding protein	8,5
1461	Lactitol	Transcriptional regulator	16,0
1462	Lactitol	β-D-galactosidase	42,0
1463	Lactitol	GPH permease	22,7
1465	Lactitol	Regulator	2,2
1467	Lactitol	β-D-galactosidase Large subunit	22,4
1469	Lactitol	UDP-galactose 4-epimerase	6,7

Table 3: Identified genes involved with carbohydrate metabolism for cellobiose, stachyose and lactitol.

These findings show how *L. acidophilus* NCFM processes potential prebiotics, by a diverse set of transport systems and glycoside hydrolases.

Functional genomics

The functionality of the stachyose induced operon was validated using the *upp* gene deletion system [2], as illustrated in figure 6. The GH36 α-galactosidase (LBA1438) and substrate recognizing, solute binding protein of the ABC transporter (LBA1442) were deleted. Phenotypes were assessed by the introduced growth limitations, figure 7. **Both LBA1438 and LBA1442 were found to be essential for metabolism of raffinose and melibiose.** It is also highly likely that the other raffinose family oligosaccharides are catabolyzed via this pathway.

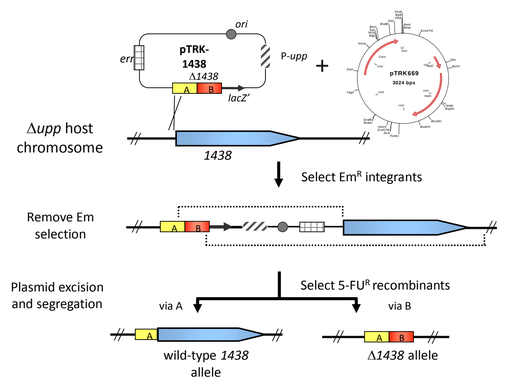


Figure 6: Overview mechanism of the *upp* gene replacement system.

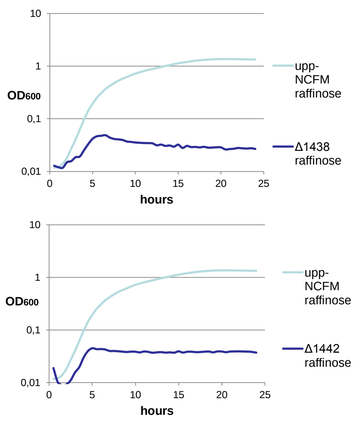


Figure 7: ΔLBA1438 (top) and ΔLBA1442 (below) grown on 1 % (w/v) raffinose in semi defined media compared to *upp*-wildtype.

Study summary

Identification of specific metabolic pathways allows future pre/pro-biotic health claims to both organism and carbohydrates for novel food and medical products

- Differential gene expression show specific regulatory patterns in response to carbohydrate stimulations
- Potential prebiotics in *L. acidophilus* NCFM are metabolized by a range of glycoside hydrolases, ABC transporters, PTS systems and a GPH permease
- Genes LBA1438 and LBA1442 are essential for hydrolysis and transport, respectively, of raffinose and melibiose in *L. acidophilus* NCFM

Transcriptional Analysis of Prebiotic Utilization by *Lactobacillus acidophilus* NCFM

J. M. Andersen[1,3], R. Barrangou[2], M. Abou Hachem[1], B. Svensson[1], Y. Goh[3], T. R. Klaenhammer[3]

[1]Technical University of Denmark, Kgs. Lyngby, Denmark, [2]Danisco Inc., Madison, WI, [3]North Carolina State Univ., Raleigh, NC

Abstract

Microbes delivered as functional probiotics depend on their ability to survive in the gastrointestinal tract, adhere to mucosal surfaces, and metabolize available energy sources from non-digestible dietary compounds. The objective of this study was to identify genetic loci in *Lactobacillus acidophilus* NCFM that are responsible for the utilization of complex carbohydrates that may function as prebiotic substrates, *in vivo*.

Whole genome oligonucleotide microarrays (1,823 ORFs, 97% coverage) were used to survey the global transcriptome during logarithmic growth of *L. acidophilus* NCFM in the presence of 11 different carbohydrates (glucose, raffinose, cellobiose, panose, stachyose, isomaltose, gentiobiose, lactitol, β-glucan oligomers, polydextrose®, isomaltulose). The data was compiled and analyzed in JMP Genomics, using a mixed-model ANOVA. Specific transporters of the ATP-binding cassette (ABC), phosphotransferase system (PTS) and galactoside-pentose hexuronide (GPH) families were identified for the uptake of stachyose, isomaltose and lactitol, respectively. The predicted roles of these transporters and carbohydrates were functionally assessed by targeted gene deletion.

The study has identified a series of genes that are responsible for the uptake and catabolism of a variety of prebiotic oligosaccharides in *L. acidophilus* NCFM.

Introduction

Lactobacillus acidophilus NCFM, isolated from human gut, is a proven probiotic bacterium commercially used in dietary supplements and fermented dairy products. Extensive Work has been done to understand the underlying molecular mechanisms of the probiotic effects, among others: bile tolerance, adherence to mucosal surfaces, mammalian host interactions and prebiotic utilization. The genome sequence of *L. acidophilus NCFM* [1] showed how a significant part of the total genome encodes transport systems and enzymatic machinery, to process a wide array of oligomeric carbohydrates , as summarized in table 1 and illustrated in figure 1.

Lactobacillus acidophilus NCFM	
genome size	1.99 Mb
GC content	34,70%
ORFS	1,864
Encoded carbohydrate metabolism	
ABC transporters	3
PEP-PTS systems	20
GPH permeases	2
Glycoside hydrolases	37
… of which are extracallular	2

Table 1: Summary of *L. acidophilus* NCFM predicted genes involved with carbohydrate metabolism

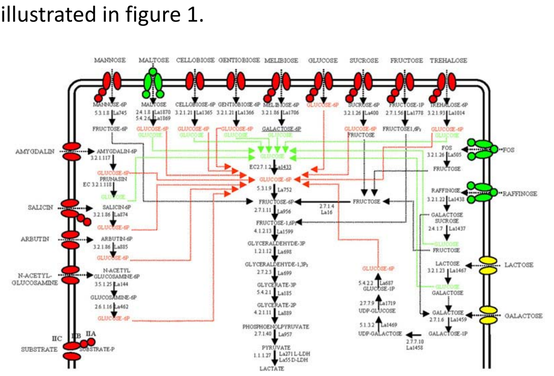


Figure 1: Predicted carbohydrate transport and intracellular metabolism in *L. acidophilus* NCFM

DNA microarray technology was previously used to map the gene expression in *L. acidophilus* NCFM in the presence of mono- and disaccharides carbohydrates [2]. The current study goes in depth with identification of genetic loci potentially involved with prebiotic catabolism. A series of 11 potential prebiotic carbohydrates (shown in table 2 together with predicted responsible glycoside hydrolase for intracellular breakdown) were selected for measurement of the total transcriptome in response to growth on each carbohydrate. Genes of interest were assessed by targeted gene deletions in *L. acidophilus* NCFM to confirm the role in the specific carbohydrate metabolism.

Carbohydrate	Structure	Predicted degrading enzymes
Glucose	**Glc**	Glycolytic pathway
Raffinose	**Gal-** *α1,6*-**Glc-** *α1,2*-**Fru**	α-galactosidase + sucrose phosphorylase
Gentiobiose	**Glc-***β1,6*-**Glc**	6-P-β-glucosidases
Panose	**Glc-***α1,6*-**Glc-***α1,4*-**Glc**	α-1,6-glucosidases, maltose phosphorylase
Isomaltose	**Glc-***α1,6*-**Glc**	α-1,6-glucosidases
Stachyose	**Gal-***α1,6*-**Gal-***α1,6*-**Glc-***α1,2*-**Fru**	α-galactosidase + sucrose phosphorylase
Cellobiose	**Glc-** *β1,4*-**Glc**	P-β-glucosidases
Polydextrose	*All glucose linkages*	Various α-glucosidases
Beta-glucan	…**Glc-***β1,3*-**Glc** (*β1,4*)…	β-glucanase + β-glucosidase
Isomaltulose	**Glc-***α1,6*-**Fru**	α-1,6-glucosidases
Lactitol	*Sugar alcohol*	β-galactosidases

Table 2: Carbohydrates used in this study listed together with structural glycoside composition O-linkages and the enzyme class predicted to facilitate intracellular hydrolysis of the carbohydrates

References

[1] Altermann, E., Russell, M. W., Azcarate-Peril, A., Barrangou, R., Buck, L. B., McAuliffe, O., Souther, N., Dobson, A., Duong. Tri., Callanan, M., Lick, S., Hamrick, A., Cano, R. and Klaenhammer T.R. 2005 Complete genome sequence of the probiotic lactic acid bacterium Lactobacillus acidophilus NCFM. Proc.Natl. Acad. Sci. **102**: 3906-3912

[2] Barrangou, R., Azcárate-Peril, A., Duong, T., Conners, S., Kelly, R. and Klaenhammer, T. R. 2006. Global analysis of carbohydrate utilization by *Lactobacillus acidophilus* using cDNA microarrays. Proc.Natl. Acad. Sci. **103**: 3816-3821

[3] Goh, Y. J., Azcárate-Peril, A., O'Flaherty S., Durmaz E., Valence F., Jardin J., Lortal S., and Klaenhammer, T. R. 2009. Development and Application of a *upp*-Based Counterselective Gene Replacement System for the Study of the S-Layer Protein SlpX of *Lactobacillus acidophilus* NCFM. Appl. Environ. Microbiol. **75**: 3093-3105.

Global transcriptome analysis

In order to analyze the global transcriptome of *L. acidophilus* NCFM, total RNA was isolated from cell cultures harvested at the logarithmic phase while growing on each of the single carbohydrates listed in table 2. The RNA was reverse transcribed and labeled with Cyanine3 and Cyanine5, with two technical replicates to each condition, prior to chip hybridization and scan. Acquired scanned images were background corrected and spot intensities were log2 transformed before being compiled into JMP genomics 4.1 for quantile normalization and ANOVA modeling.

The global transcriptome expression pattern of *L. acidophilus* NCFM was visualized by hierarchical clustering for the 11 carbohydrate growth conditions. Overall low variance between each condition was observed, correlating with regulation of few genetic loci in response to specific carbohydrate metabolism.

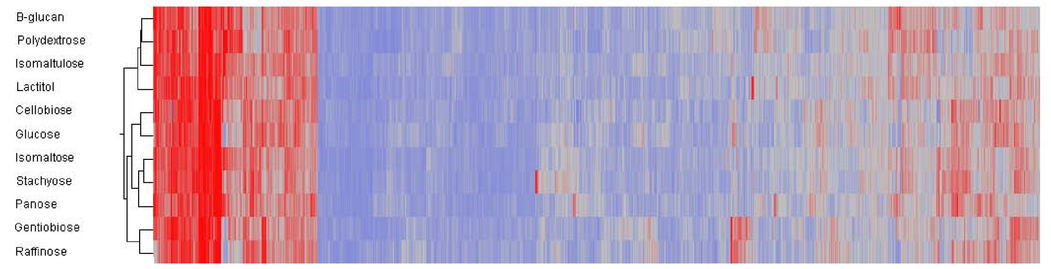


Figure 2: Rrepresentation of the hierarchical clustering of the global gene expression of *L. acidophilus* NCFM by carbohydrate. Red coloring indicates up regulation and blue coloring indicates down regulation of genes

Carbohydrate metabolism

Significant, differential expressed genes were identified from ANOVA modeling ($P < 10^{-2.75}$) and visualized by volcano plots (figure 3). Upregulated genes involved with carbohydrate metabolism to each growth condition was listed and are shown as a expression heat map in figure 4. A of genes involved with carbohydrate metabolism and their fold up regulation are shown in table 3. Several sets of operon-like sets of genes were discovered, all including both a membrane transporter and at least one glycoside hydrolase. These findings show how *L. acidophilus* NCFM process oligosaccharides by a diverse set of transport systems and glycoside hydrolases.

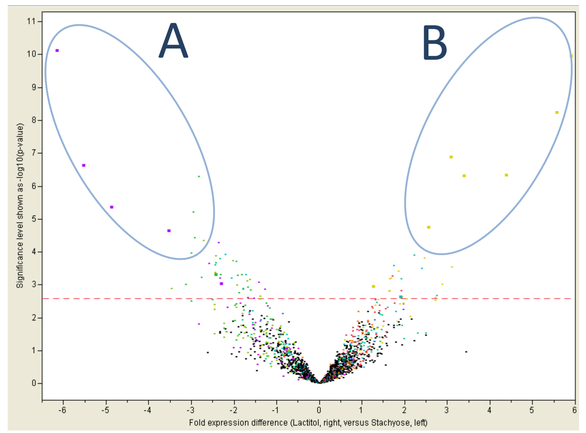


Figure 3: Volcano plot illustrating the differential gene expression in *L. acidophilus* NCFM grown on lactitol compared to stachyose as an example. Significantly upregulated genes of interest, in the presense of stachyose, are circled at A while genes upregulated by lactitol are circled at B

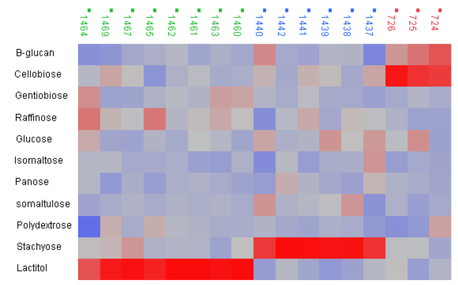


Figure 4: Expression heat map of genes related to metabolism of potential prebiotics in *L. acidophilus* NCFM.

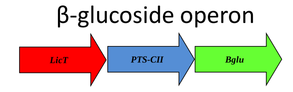
Locus tag	Carbohydrate	Putative function	Fold induction range
724	Cellobiose, β-glucan	Transcriptional regulator	3,7 – 4,9
725	Cellobiose, β-glucan	PTS system IIC component	5,4 – 23,3
726	Cellobiose, β-glucan	6-P-β-glucosidase	1,8 – 12,1
1437	Stachyose	Sucrose phosphorylase	4,7
1438	Stachyose	α-galactosidase	15,1
1439	Stachyose	ABC transporter ATP-binding protein	18,1
1440	Stachyose	ABC transporter Transmembrane permease	3,2
1441	Stachyose	ABC transporter Transmembrane permease	7,6
1442	Stachyose	ABC transporter Substrate-binding protein	53,1
1460	Lactitol	mucus bindingprotein precursor	8,5
1461	Lactitol	Transcriptional regulator	16,0
1462	Lactitol	β-D-galactosidase	42,0
1463	Lactitol	GPH permease	22,7
1465	Lactitol	Lactose transcriptional regulator	2,2
1467	Lactitol	β-D-galactosidase Large subunit	22,4
1469	Lactitol	UDP-galactose 4-epimerase	6,7

Table 3: All identified upregulated genes involved with carbohydrate metabolism. For each gene (numbered with Locus tag) the inducing carbohydrates are shown with annotation and the fold induction span. (The highest inducing carbohydrate for each gene is underlined)

In silico operon reconstruction

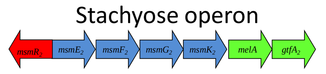
Putative operons encoding genes involved with carbohydrate catabolism were reconstructed from the identified upregulated genes. This pictures the structure of genetic loci for potential prebiotic utilization in the Lactobacilli genus.

β-glucoside operon



Locus tag	Putative function
LBA0724	Regulator
LBA0725	PTS CII Component
LBA0726	6-P-β-glucosidase

Stachyose operon



Locus tag	Putative function
LBA1443	Regulator
LBA1442	Sugar binding protein, ABC
LBA1441	Permease, ABC
LBA1440	Permease, ABC
LBA1439	Kinase, ABC
LBA1438	α-galactosidase, GH36
LBA1437	Sucrose phosphorylase, GH13_18

Lactose operon

Locus tag	Putative function
LBA1460	Surface protein
LBA1461	tetR Regulator
LBA1462	β-galactosidase GH42
LBA1463	GPH permease
LBA1464	Transposase
LBA1465	Repressor of β-galactosidase
LBA1467	β-galactosidase, large subunit GH2
LBA1468	β-galactosidase, small subunit GH2
LBA1469	UDP-glucose epimerase

Figure 5: Reconstruction of putative prebiotic genetic operons based on DNA microarray data. Regulator genes are shown in red, Transporter complexes in blue and glycoside hydrolases in green

Functional genomics

The functionality of selected upregulated genes were analyzed using the *upp* gene deletion system [3], which allows high efficiency in-frame deletion of coding DNA as illustrated in figure 6. The stachyose induced operon was chosen as target and the glycoside hydrolase family 36 α-galactosidase (LBA1438) and substrate recognizing solute binding protein of the ABC transporter (LBA1442) were deleted. Mutant phenotypes were assessed by growth/no growth on selected carbohydrates. *L. acidophilus* NCFM wild type and mutant growth curves are shown in figure 7. Both LBA1438 and LBA1442 were found to be essential for metabolism of raffinose and melibiose. It is also highly likely that the other raffinose family oligosaccharides are catabolyzed via this pathway.

upp-based counterselectable gene replacement system

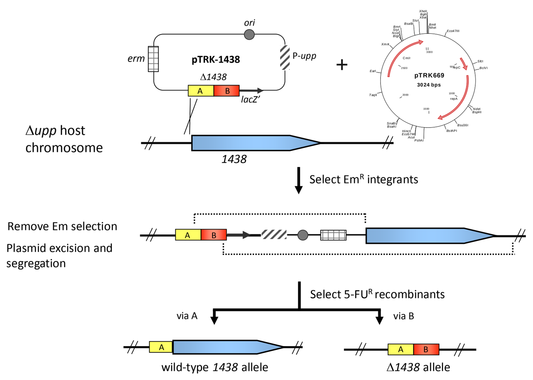


Figure 6: Overview mechanism of the *upp* gene replacement system: Transformation of target gene flanking region, integration into the chromosome and excision to generate the gene deletion mutant.

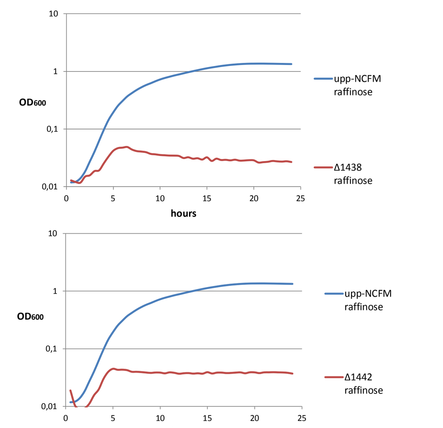


Figure 7: Growth curves of ΔLBA1438 (top) and ΔLBA1442 (button) grown on 1 % raffinose in semi defined media compared to *upp*-wildtype control.

Study summary

- Differential gene expression showed specific regulatory patterns in response to carbohydrate growth conditions
- Potential prebiotics in *L. acidophilus* NCFM are metabolized by a range of glycoside hydrolases, ABC transporters, PTS systems and a GPH permease
- Genes LBA1438 and LBA1442 are essential for hydrolysis of raffinose and transport of melibiose in *L. acidophilus* NCFM

Acknowledgements

This work is supported by Danisco USA and North Carolina Dairy Foundation.

Joakim Mark Andersen is supported by a FøSu grant from the Danish Strategic Research Council to the project "*Gene discovery and molecular interactions in prebiotics/probiotics systems. Focus on carbohydrate prebiotics*".