

ORIGINAL ARTICLE

Open Access



# Characterization of GH2 and GH42 $\beta$ -galactosidases derived from bifidobacterial infant isolates

Valentina Ambrogi<sup>1,2</sup>, Francesca Bottacini<sup>1</sup> , Joyce O'Sullivan<sup>1</sup>, Mary O'Connell Motherway<sup>1</sup>, Cao Linqiu<sup>3</sup>, Barry Schoemaker<sup>3</sup>, Margriet Schoterman<sup>3</sup> and Douwe van Sinderen<sup>1,2\*</sup>

## Abstract

Bifidobacteria are among the first and most abundant bacterial colonizers of the gastrointestinal tract of (breast-fed) healthy infants. Their success of colonising the infant gut is believed to be, at least partly, due to their ability to metabolize available carbon sources by means of secreted or intracellular glycosyl hydrolases (GHs). Among these,  $\beta$ -galactosidases are particularly relevant as they allow bifidobacteria to grow on  $\beta$ -galactosyl-linked saccharidic substrates, which are present in copious amounts in the milk-based diet of their infant host (e.g. lactose and human milk oligosaccharides). In the present study we employed an *in silico* analysis to identify GH family 2 and 42  $\beta$ -galactosidases encoded by typical infant-associated bifidobacteria. Comparative genome analysis followed by characterisation of selected  $\beta$ -galactosidases revealed how these GH2 and GH42 members are distributed among these infant-associated bifidobacteria, while their hydrolytic activity towards growth substrates commonly available in the infant gut were also assessed.

**Keywords:**  $\beta$ -Galactosidases, Bifidobacteria, Lactose, HMOs, Infant gut microbiota

## Introduction

Bifidobacteria are common members of the human gut microbiota, and are especially abundant in the gastrointestinal tract (GIT) of healthy, breast-fed infants (Milani et al. 2017). The presence of bifidobacteria in the gut is believed to provide a number of health benefits for the human host, although the mechanisms by which such benefits are delivered are currently not fully understood (Arboleya et al. 2016; Milani et al. 2017). Bifidobacteria are saccharolytic microorganisms and their ability to utilize complex dietary glycans or host-derived mucins is an important property to assist in their establishment and persistence into the GIT (Koropatkin et al. 2012; Milani et al. 2015; Riviere et al. 2016). The saccharolytic metabolism of bifidobacteria is facilitated by a large array of carbohydrate degrading enzymes, in particular glycosyl

hydrolases (GHs), which provide the capacity to directly or indirectly (through syntrophy) utilize a range of glycan substrates available in the gut (Milani et al. 2015; Turroni et al. 2018a). Bifidobacteria commonly represent the dominant component of the gut microbiota of healthy, breast-fed infants (Roger et al. 2010), a phenomenon that is believed to be due in part by their ability to metabolize human milk oligosaccharides (HMOs), which contain one or more  $\beta$ -linked galactose moieties (Fuhrer et al. 2010). Among the glycosyl hydrolases encoded by bifidobacteria (Milani et al. 2015)  $\beta$ -galactosidases (which are members of GH families 2 and 42) (van den Broek et al. 2008) have been described to participate to the utilization of (human) milk and milk-based substrates (i.e. lactose, HMOs as well as GOS) in *Bifidobacterium bifidum*, *Bifidobacterium longum* subsp. *infantis* and *Bifidobacterium breve* (Garrido et al. 2013; Goulas et al. 2007; James et al. 2016; Miwa et al. 2010; Moller et al. 2001; O'Connell Motherway et al. 2013; Yoshida et al. 2012). In addition,  $\beta$ -galactosidases in *B. bifidum* can also participate to the degradation of mucin (Turroni et al. 2010), while in *B.*

\*Correspondence: d.vansinderen@ucc.ie

<sup>1</sup> APC Microbiome Ireland, University College Cork, Western Road, Cork, Ireland

Full list of author information is available at the end of the article

*breve* such enzymes are involved in the degradation of the plant polymer galactan (O'Connell Motherway et al. 2011).

Human milk glycans can be quite diverse and they are composed of 13 core structures generated through the elongation of lactose at the reducing end with one or more  $\beta$ 1,3-linked lacto-*N*-biose (type-I chain) and/or  $\beta$ 1,3/6-linked *N*-acetylglucosamine units (type-II chain) (Urashima et al. 2012). These core structures (including lactose itself) can in turn be substituted at terminal positions by fucose connected via  $\alpha$ 1,2/3/4 links, and/or sialic acid residues attached by  $\alpha$ 2,3/6 links (Smilowitz et al. 2014). Notably, HMOs are especially rich in type-I chains and constitute a characteristic feature of human milk (Urashima et al. 2012).

Some of the glycosidic linkages and monosaccharides found in HMOs are also present in mucins found in the intestine, for example the presence of  $\beta$ -linked galactose and *N*-acetyl-glucosamine, as well as fucose and sialic acid residues which are substituted to mucin via  $\alpha$ 1,2/3/4 and  $\alpha$ 2,3/6 linkages (Tailford et al. 2015).

Recent genome analyses have highlighted metabolic pathways responsible for the degradation of HMOs and mucin-type *O*-glycans in certain infant-derived bifidobacteria. With regards to HMO utilization it has been shown that *B. bifidum* and *B. longum* subsp. *infantis* can utilize many different HMOs directly (Asakuma et al. 2011; Garrido et al. 2016; Sela 2011; Ward et al. 2007). In contrast, *B. breve* and *B. longum* subsp. *longum* can only metabolize a small number of these directly, thus relying on cross-feeding strategies (Asakuma et al. 2011; Egan et al. 2014; LoCasio et al. 2007; Ward et al. 2007). *B. bifidum* is the only bifidobacterial species described so far to be able to utilize mucin-type *O*-glycans (Turroni et al. 2010), and the products of (extracellular) mucin degradation constitute growth substrates for other (bifido)bacterial species, revealing that bifidobacteria can adopt a syntrophic strategy to access different substrates available in the gut (Egan et al. 2014; Turroni et al. 2018b).

The complex structural and functional heterogeneity of HMOs in mother's milk is presumed to be essential in conferring the bifidogenic activity and associated health benefits to the host, yet it also represents the main limitation for the deliberate incorporation of HMOs in infant formulations (Akkerman et al. 2018). For this reason alternative ways of recreating the beneficial effects of human milk have focused on the enzymatic production of other non-digestible, lactose-derived prebiotics, such as GOS produced by  $\beta$ -galactosidases that have transgalactosylation activities (Macfarlane et al. 2008).

Based on the above, the relevance of diet-based, lactose-derived oligosaccharides (e.g. HMOs and GOS) in supporting bifidobacterial establishment and persistence

in the gut is obvious, especially during early life when bifidobacteria are particularly abundant (Milani et al. 2017). The current study focused on a survey and subsequent hydrolytic characterization of (selected)  $\beta$ -galactosidases identified from sequenced strains of bifidobacterial species which have been previously shown to be members of the infant gut microbiota (i.e. *B. breve*, *B. bifidum*, *B. longum* subsp. *longum* and *B. longum* subsp. *infantis*) (Lewis and Mills 2017; Milani et al. 2017; Turroni et al. 2012). Following comparative genome analysis, our analysis focused on a number of selected  $\beta$ -galactosidases of which hydrolytic capabilities were assessed for various  $\beta$ -galactoside-containing substrates.

## Materials and methods

### In silico genome analysis to identify putative $\beta$ -galactosidase-encoding genes

For comparative purposes, genomic datasets, consisting of 34 genome sequences (Table 1) from four representative infant-derived bifidobacterial (sub)species (i.e. *B. breve*, *B. bifidum*, *B. longum* subsp. *longum* and *B. longum* subsp. *infantis*) were retrieved from GenBank (<https://www.ncbi.nlm.nih.gov/genbank>). Of the corresponding 34 bifidobacterial strains, we selected six strains as representative reference strains for our analysis as their genome had been sequenced to completion and because these strains are available within our culture collection [i.e. *B. breve* UCC2003 (GenBank: CP000303), *B. breve* JCM 7017 (GenBank: CP006712), *B. breve* JCM 7019 (GenBank: CP006713), *B. bifidum* LMG 13195 (WGS: AMPL01000000), *B. longum* subsp. *longum* NCIMB 8809 (GenBank: CP011964) and *B. longum* subsp. *infantis* ATCC 15697 (GenBank: CP001095)]. The selected genomes were searched for predicted  $\beta$ -galactosidase-encoding genes on the basis of a functional annotation supported by PFAM analysis (<http://pfam.xfam.org/>), combined with information retrieved from the Cazy database (<http://www.cazy.org>). Predictions were refined using comparative genome analysis with detection of orthologous  $\beta$ -galactosidase-encoding genes across genomes using BLASTP searches (Altschul et al. 1990). BLASTP alignments were performed using a stringency of 50% of identity across 50% of the length of the examined proteins, with a cut-off E-value of < 0.0001. Predicted  $\beta$ -galactosidase-encoding genes were then grouped in clusters of orthologs using the cd-hit pipeline (<http://weizhongli-lab.org/cd-hit/>) with a clustering threshold of 70% identity across 90% of the corresponding protein length (Bottacini et al. 2014).

Phylogenetic inference was conducted using the MEGA7 suite (<http://www.megasoftware.net/>). Protein sequence alignments were performed using the Muscle module available within MEGA7 and the resulting

**Table 1 Bifidobacterial strains used for comparative analyses**

Species	Strain	Accession number WGS
<i>B. bifidum</i>	JCM 1254	BBBT00000000
<i>B. bifidum</i>	DSM 20215	–
<i>B. bifidum</i>	DSM 20456	JDUM00000000
<i>B. bifidum</i>	NCIMB 41171	AKCA01000000
<i>B. bifidum</i>	PRL2010	CP001840
<i>B. bifidum</i>	S17	CP002220
<i>B. bifidum</i>	BGN4	CP001361
<i>B. bifidum</i>	IPLA20015	AMPM01000000
<i>B. bifidum</i>	LMG 11041	JGYO01000000
<i>B. bifidum</i>	LMG 13195	AMPL01000000
<i>B. breve</i>	2L	AWUG00000000
<i>B. breve</i>	JCM 7019	CP006713
<i>B. breve</i>	UCC2003	CP000303
<i>B. breve</i>	31L	AWUF01000000
<i>B. breve</i>	NCFB 2258	CP006714
<i>B. breve</i>	S27	CP006716
<i>B. breve</i>	12L	CP006711
<i>B. breve</i>	JCM 7017	CP006712
<i>B. breve</i>	JCP 7499	AWSX01000000
<i>B. longum</i> subsp. <i>infantis</i>	ATCC 15697	CP001095
<i>B. longum</i> subsp. <i>infantis</i>	EK3	JNWB01000000
<i>B. longum</i> subsp. <i>infantis</i>	157F	AP010890
<i>B. longum</i> subsp. <i>infantis</i>	CCUG 52486	ABQQ01000000
<i>B. longum</i> subsp. <i>longum</i>	NCIMB 8809	CP011964
<i>B. longum</i> subsp. <i>longum</i>	EK13	JNWD01000000
<i>B. longum</i> subsp. <i>longum</i>	F8	FP929034
<i>B. longum</i> subsp. <i>longum</i>	KACC 91563	CP002794
<i>B. longum</i> subsp. <i>longum</i>	VMKB44	JRWN01000000
<i>B. longum</i> subsp. <i>longum</i>	ATCC 55813	ACHI01000000
<i>B. longum</i> subsp. <i>longum</i>	BBMN68	CP002286
<i>B. longum</i> subsp. <i>longum</i>	JCM 1217	AP010888
<i>B. longum</i> subsp. <i>longum</i>	LMG 13197	JGYZ01000000
<i>B. longum</i> subsp. <i>longum</i>	JDM301	CP002010
<i>B. longum</i> subsp. <i>longum</i>	CECT 7347	CALH01000000

phylogenetic tree was built using the neighbour joining approach with statistical assessment based on 1000 bootstrap replicates.

#### Growth conditions and strain manipulation

Bifidobacterial strains were cultured in de Man Rogosa and Sharpe (MRS) medium supplemented with cysteine-HCl (0.05% w/v) and incubated at 37 °C degrees under anaerobic conditions using an anaerobic chamber (Davidson and Hardy, Belfast, Ireland). *L. lactis* strain NZ9000 (University of Groningen, The Netherlands) was selected for cloning (de Ruyter et al. 1996), overproduction and

purification purposes as several bifidobacterial GH-encoding genes had previously been successfully cloned and expressed in this host (James et al. 2016; O'Connell et al. 2013; O'Connell Motherway et al. 2013). *L. lactis* NZ9000 was grown in M17 broth (Oxoid, UK) supplemented with 0.5% glucose (GM17) at 30 °C degrees (Terzaghi and Sandine 1975). *L. lactis* NZ9000 cells carrying (suspected) recombinant plasmids were selected on GM17 agar containing chloramphenicol (Cm 5 µg ml<sup>-1</sup>), and supplemented with X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (40 µg ml<sup>-1</sup>) to visually identify recombinant clones expressing β-galactosidase activity.

Chromosomal DNA was isolated from *B. breve* UCC2003, *B. bifidum* LMG 13195, *B. longum* subsp. *longum* NCIMB 8809 and *B. longum* subsp. *infantis* ATCC 15697 as previously described (O'Riordan and Fitzgerald 1998). Plasmid extraction from *L. lactis* was achieved using the Roche High Pure Plasmid Isolation Kit (Roche Diagnostics GmbH, Mannheim, Germany). For *L. lactis* an initial lysis step was included where cells are incubated in lysis buffer containing 30 mg ml<sup>-1</sup> of lysozyme for 30 min at 37 °C. Oligonucleotide PCR primers used in the study were synthesized by Eurofins (Ebersberg, Germany; Table 2). Standard PCRs were performed using Thermo Scientific Extensor Hi-Fidelity PCR Master mix (Thermo Scientific), while high fidelity PCR was achieved using Q5 High-Fidelity DNA Polymerase (New England BioLabs, New Brunswick). PCR fragments were purified with Roche High Pure PCR Product Purification Kit (Roche Diagnostics GmbH). Plasmid pNZ8150 was used as cloning and nisin-inducible expression vector. DNA fragments encompassing predicted β-galactosidase-encoding genes and the pNZ8150 vector were restricted with specific restriction enzymes (Table 2) according to the supplier's instructions (Roche Diagnostics, East Sussex, United Kingdom). Ligation was achieved with T4 DNA ligase according to the manufacturer's instructions (Promega, Wisconsin, USA, worldwide.promega.com). Electroporation of (ligated) plasmid DNA into *L. lactis* was performed as described previously (Wells et al. 1993).

#### Protein overproduction and purification

For bifidobacterial protein overproduction 0.4 ml of M17 broth supplemented with 0.5% glucose was inoculated with a 2% inoculum of *L. lactis* strain harbouring the pNZ8150 cloning vector containing a (predicted) β-galactosidase-encoding gene, and cultivated at 30 °C until the culture reached an Optical Density (OD<sub>600 nm</sub>) of 0.5. At this point targeted protein expression was induced by the addition of filter sterilised cell free supernatant of a nisin-producing strain *L. lactis* NZ9700 (0.2% v/v)

**Table 2 Oligonucleotide primers used in this study**

Gene	Primer	Sequence	Restriction enzymes
Bbr_0010	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatgcatcaccatcaccatcaccatcacc	<i>EcoRV</i>
	Rv	tgcgcaTCTAGAtcagatgagttcgagtgac	<i>XbaI</i>
Bbr_0285	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatggagcgaatccaatacccc	<i>EcoRV</i>
	Rv	tgcgcaTCTAGAtcacacctgcagtagccg	<i>XbaI</i>
Bbr_0310	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatggggacgacaggacacagc	<i>EcoRV</i>
	Rv	tgcgcaTCTAGAtcaactcttttcgattgcg	<i>XbaI</i>
Bbr_0420	Fw	tgcacCCCGGGcatcaccatcaccatcaccatcaccatcacatgactactcgtagagc	<i>XmaI</i>
	Rv	ctcgaTCTAGActagcaggacgttttagcg	<i>XbaI</i>
Bbr_0529	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatggaacatcgcaattcaag	<i>EcoRV</i>
	Rv	tgcgcaTCTAGAttacagctttaccaccagcac	<i>XbaI</i>
Bbr_1552	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatgaacacacaccgacgatcag	<i>EcoRV</i>
	Rv	tgcgcaTCTAGAtcagatgagttcgaggtcac	<i>XbaI</i>
Bbr_1689-1690	Fw	tgcacCAGCTGcatcaccatcaccatcaccatcaccatcacatgagcaagcagaacgattg	<i>PvuII</i>
	Rv	tgcgcaTCTAGAgctggcatcttctgaacg	<i>XbaI</i>
B7017_2031	Fw	tgcacGAATTCcatcaccatcaccatcaccatcaccatcacatgaccgacacatggcacacacacaca	<i>EcoRI</i>
	Rv	tgcacTCTAGActgatgatgaaggatgactgaagccg	<i>XbaI</i>
B216_06500	Fw	tgcac ATTTAAAT atg catcaccatcaccatcaccatcaccatcac gtaataaccgttagggttgt	<i>SwaI</i>
	Rv	tgcac TCTAGAcccggggagactcgcgagagt	<i>XbaI</i>
B216_08266	Fw	tgcacATTTAAATcatcaccatcaccatcaccatcaccatcacatgagtaaacgcagaagcacag	<i>SwaI</i>
	Rv	tgcgcaTCTAGAgatgtcgcgtgtcaccg	<i>XbaI</i>
B216_08730	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatgctgctgctgctgctgctgctt	<i>EcoRV</i>
	Rv	tgcac TCTAGA aacgttgaaatagagccggaac	<i>XbaI</i>
B216_09411	Fw	tgcac ATTTAAATatg catcaccatcaccatcaccatcaccatcac ttcattccccgactacg	<i>SwaI</i>
	Rv	tgcac TCTAGA atccgataaccgtaccctg	<i>XbaI</i>
B216_09623	Fw	tgcacATTTAAATcatcaccatcaccatcaccatcaccatcacatgaacacacaccgacgatcagc	<i>SwaI</i>
	Rv	tgcgcaTCTAGAAatgagcagagaggacctggcg	<i>XbaI</i>
B8809_0321	Fw	tgcacATTTAAATcatcaccatcaccatcaccatcaccatcacatgactactcatagagcatttag	<i>SwaI</i>
	Rv	tgcac TCTAGAcattctagcgcgttttag	<i>XbaI</i>
B8809_0415	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatggaacgtaagagttcaagtgg	<i>EcoRV</i>
	Rv	tgcacTCTAGAccgttgggtaattagcgcct	<i>XbaI</i>
B8809_0611	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatgacagacgctcacacatgctg	<i>EcoRV</i>
	Rv	tgcac TCTAGAtgcacggtggactatcggatc	<i>XbaI</i>
B8809_1361	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatgcagatccccaccaccac	<i>EcoRV</i>
	Rv	tgcac TCTAGAcagcagcagataaagaagctccctcg	<i>XbaI</i>
Blon_2016	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatggaacatagagcgttcaagt	<i>EcoRV</i>
	Rv	tgcacTCTAGAccgctccctgctgctgatga	<i>XbaI</i>
Blon_2123	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatggtgctgctgctgctgctgctt	<i>EcoRV</i>
	Rv	tgcacTCTAGAccatgtacgtcggcaccgt	<i>XbaI</i>
Blon_2416	Fw	tgcacGATATCcatcaccatcaccatcaccatcaccatcacatgaccgacacatggcaca	<i>EcoRV</i>
	Rv	tgcacTCTAGAccggttctgacttgggat	<i>XbaI</i>

(de Ruyter et al. 1996) and incubation was continued at 30 °C for 90 min. Cells were harvested by centrifugation (5000 rpm for 10 min), the obtained pellet was recovered in lysis buffer (50 mM sodium phosphate buffer, pH 8; 300 mM NaCl; 10 mM Imidazole) and cells were disrupted by bead beating (1 min, three times). Cell debris was removed by centrifugation (15,000 rpm for 20 min

at 4 °C) and the resulting supernatant, representing the crude cell extract, was used for (His-tagged) protein purification using a nickel-nitrilotriacetic acid column (Qiagen GmbH) according to the manufacturer's instructions (QIAexpressionist, June 2003). Elution fractions were analysed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), as described previously (Laemmli 1970), on

a 12.5% polyacrylamide (PAA) gel. After electrophoresis, PAA gels were fixed and stained with Coomassie brilliant blue to identify fractions containing the purified protein. For molecular weight estimation of (purified) proteins the Prestained Protein Marker, Broad Range (7–175 kDa) was used (New England BioLabs, Hertfordshire, United Kingdom).

#### Beta-galactosidase assay: crude cell extract

Colorimetric and spectrophotometric assay was performed to evaluate the hydrolytic activity of predicted  $\beta$ -galactosidases towards the artificial substrate *O*-nitrophenyl  $\beta$ -D-galactopyranose (ONPG assay). This assay was performed at 30 °C according to a previously published protocol (Miller, Cold Spring Harbor, 1972). Following the development of a yellow color the reaction was terminated by the addition of 250  $\mu$ l 1 M sodium carbonate. This was followed by measurement of absorbance at a wave length of 420 nm. Calculation of  $\beta$ -galactosidase activity was performed according to the following formula:  $\beta$ -gal unit =  $1000 \cdot \text{OD}_{420 \text{ nm}} / \text{Time} \cdot \text{V} \cdot \text{OD}_{600 \text{ nm}}$ , with incubation time of 5 min (or longer if required) and volume of resuspended pellet (V) of 0.05 ml.

#### Hydrolysis assay and product analysis

Hydrolysis assays of  $\beta$ -galactosidase activity were performed on the following commercially available substrates: lactose, lactulose, two different galactotrioses (Gal $\beta$ 1-4Gal $\beta$ 1-4Gal and Gal $\alpha$ 1-3Gal $\beta$ 1-4Gal), two different galactobioses (Gal $\beta$ 1-6Gal and Gal $\beta$ 1-4Gal), lacto-*N*-tetraose (LNT), lacto-*N*-neotetraose (LNnT), 2'-fucosyllactose (2'-FL) and 3-fucosyllactose (3-FL). The hydrolysis assays were performed using the following conditions: 250  $\mu$ l enzyme solution was incubated with 0.1 mg ml<sup>-1</sup> of a particular carbohydrate in MOPS [3-(*N*-morpholino) propanesulfonic acid] buffer pH 6.5 in a total volume of 1 ml at 37 °C. Aliquots of 250  $\mu$ l were collected after 1, 3 and 24 h from the start of the reaction and following collection samples were immediately heated at 80 °C for 5 min in order to terminate the reaction, after which they stored at -20 °C until analysis by high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) employing a CarboPac PA1 (Thermo Scientific) analytical-anion exchange column (dimensions, 250 mm by 4 mm) with a CarboPac PA1 (Thermo Scientific) guard column (dimensions, 50 mm by 4 mm) and a detector (ED40) in the pulsed amperometric detection PAD mode (Dionex, Thermo Scientific). Elution was performed at a constant flow rate of 1.0 ml min<sup>-1</sup> at 30 °C using the following eluents for the analysis: (A) 200 mmol NaOH, (B) 100 mmol NaOH, 550 mmol sodium acetate (NaAC), and (C) Milli-Q water. The elution gradient is reported in Table 3.

**Table 3** Elution gradient used in HPAEC-PAD

Minute	Concentration NaOH (mmol l <sup>-1</sup> )	Concentration NaAC (mmol l <sup>-1</sup> )
0–50	100	0
50–51	0	0
51–56	0	16
56–61	0	100

The obtained data sets were then analysed employing CHROMELEON software Ver.6.70 (Dionex, Thermo Scientific).

## Results

### Comparative analysis of $\beta$ -galactosidase-encoding genes

In order to identify putative  $\beta$ -galactosidase-encoding genes of infant-associated bifidobacterial strains, we initially selected the genomes of six strains to act as representatives for four bifidobacterial (sub) species (see “Materials and methods” section) that are typically prevalent and abundant in the (healthy) infant gut microbiota.

Putative  $\beta$ -galactosidase-encoding genes were identified on the genome sequence of each of these six reference strains (representing four species) based on manual searches and cross-validation of the obtained result with PFAM searches and Cazy database (see “Materials and methods”). The obtained list of predicted  $\beta$ -galactosidase-encoding genes was then used for further comparative analyses, aimed at identifying homologous genes across a more extensive set of 34 bifidobacterial genomes, belonging to the same (sub)species as the reference set (Table 1). Employing this comparative search approach we established that the total number of predicted GH2 and GH42  $\beta$ -galactosidase-encoding genes encoded by (34 representatives of) these four bifidobacterial species is 137, of which 44, 45 and 48 were assigned to *B. breve*, *B. bifidum* and *B. longum* subsp. *longum* spp., respectively (Additional file 1: Table S1).

### Comparative sequence analysis and phylogeny of predicted $\beta$ -galactosidase genes

To establish if a putative  $\beta$ -galactosidase-encoding gene is represented by homologs across other strains or (sub) species, we organized the 137 identified  $\beta$ -galactosidases in 18 orthologous families (here referred as clusters based on a cut-off value of 70% of similarity over 90% of sequence length; see “Materials and methods”) (Additional file 1: Figure S1). This analysis revealed that members of each of the 18 orthologous clusters are variably distributed across a given species, where certain  $\beta$ -galactosidases seem to be more species-specific (e.g.

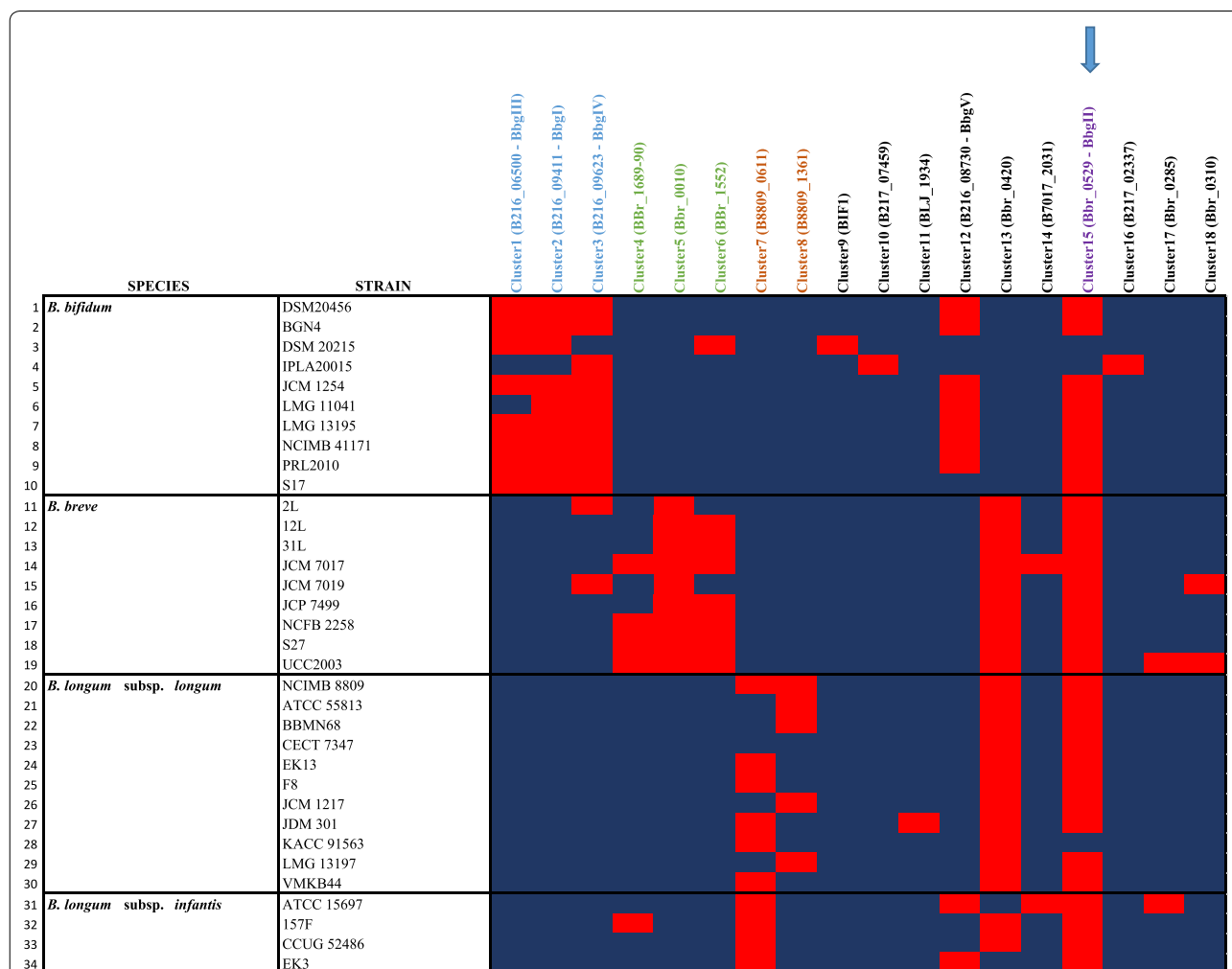
Clusters 1–3 for *B. bifidum*, Clusters 4–6 for *B. breve* and Clusters 7–8 for *B. longum* subsp. *longum*), while others are present among all members of the four (sub)species (e.g. Cluster 15) (Fig. 1 and Additional file 1: Table S1).

As mentioned above comparative sequence analysis organized the 137 identified  $\beta$ -galactosidases into 18 clusters of orthology (Fig. 1). A subsequent analysis was carried out in order to establish the phylogenetic relationships between bifidobacterial  $\beta$ -galactosidases. This analysis resulted in the generation of a phylogenetic tree which subdivided these 137 genes into two distinct clades, corresponding to members that either belong to GH family 2 or GH family 42 (Fig. 2). Within these GH2 and GH42 clades, our analysis revealed that the  $\beta$ -galactosidase-encoding genes are further distributed across 14 distinct phylogenetic groups (or subclades

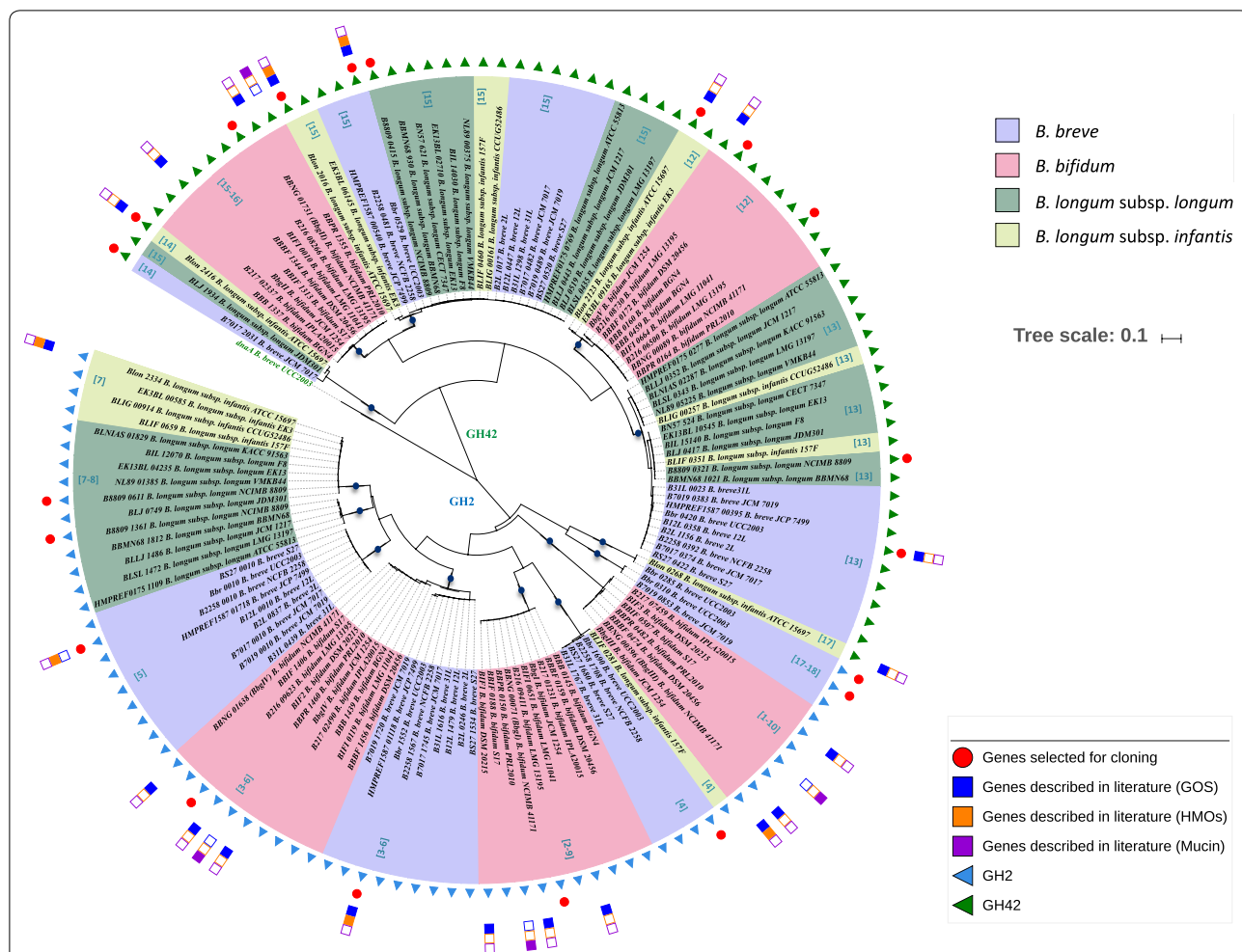
defined by nodes of bootstrap value above 90%), where genes from strains belonging to the same species often cluster together (Fig. 2).

**Cloning of bifidobacterial  $\beta$ -galactosidase-encoding genes**

Based on the outcome of the phylogenetic analyses, a total of 20 most diverse  $\beta$ -galactosidases (present in the six representative genomes available within our culture collection, while also considering practical limitations of cloning this large number of genes) were selected as candidates for further experimental validation, which in most cases locate in separate branches of the phylogenetic tree (Fig. 3a). This selection therefore captures much of the observed diversity among the identified (putative)  $\beta$ -galactosidases (Fig. 3). Notably, of these 20 selected  $\beta$ -galactosidases, four (corresponding to locus



**Fig. 1** Comparative analysis of  $\beta$ -galactosidases in infant-derived bifidobacterial (sub)species. Heatmap showing presence (red) and absence (blue) of 18 clusters of  $\beta$ -galactosidase-encoding genes across the 34 bifidobacterial representatives of *B. breve*, *B. bifidum*, *B. longum* subsp. *longum* and *B. longum* subsp. *infantis*. A blue arrow highlights Cluster 15, which is present across members of all the four (sub)species, while blue, green and brown colours identify clusters observed only in *B. bifidum*, *B. breve* and *B. longum* species, respectively



**Fig. 2** Phylogenomics of  $\beta$ -galactosidases. Neighbour-joining tree based on the alignment of 137 putative  $\beta$ -galactosidases identified across 34 members of infant-derived bifidobacterial (sub)species. A red circle highlights the genes used for experimental assessment in this study, while blue and orange squares identify those genes of which involvement in the assimilation of GOS and HMOs have been reported in literature. Circles on the phylogenetic tree highlight nodes with bootstrap values above 90% and dark blue circles identify the 14 phylogenetic groups defined by our analysis

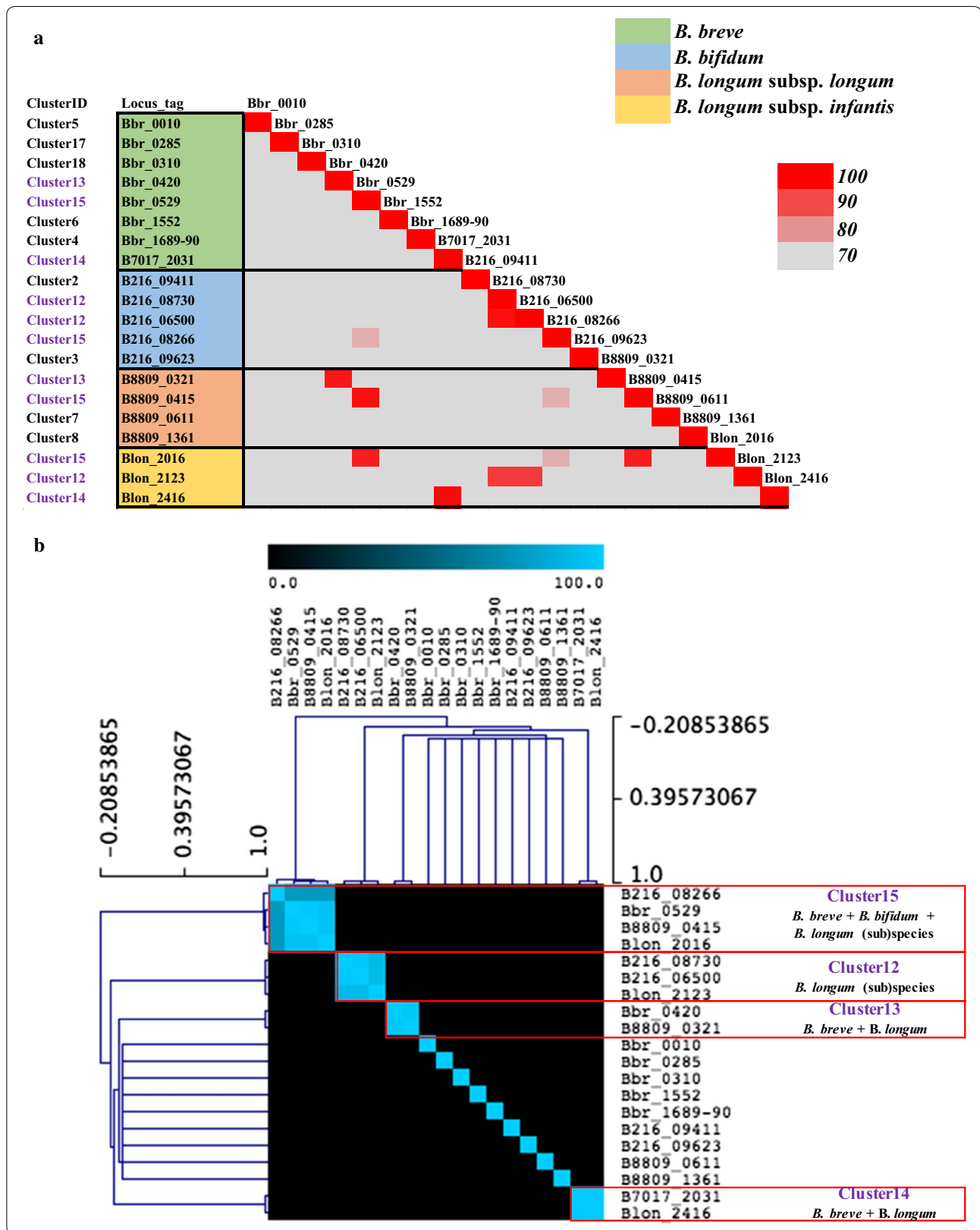
tags B216\_08266, Bbr\_0529, B8809\_0415 and Blon\_2016) are orthologous genes within Cluster 15, of which members are present across all four infant-derived bifidobacterial (sub)species, while the remaining 16 selected genes are present only within certain (sub)species (Fig. 3b).

In order to investigate if the 20 putative  $\beta$ -galactosidase-encoding genes as selected above do in fact encode this enzymatic activity, cloning and (over)expression of each of these bifidobacterial genes in *Lactococcus lactis* was

performed. The 20  $\beta$ -galactosidase-encoding genes were therefore amplified by PCR using primer pairs, one of which included a six His-tag-encoding sequence (to aid in subsequent purification), and cloned into a nisin-inducible expression vector in *L. lactis* (see “Materials and methods”). Two out of the 20 predicted  $\beta$ -galactosidase-encoding genes (corresponding to locus tags B216\_06500, B216\_09411) appeared to be unclonable in *L. lactis* (as based on multiple failed attempts). One explanation

(See figure on next page.)

**Fig. 3** Distribution of  $\beta$ -galactosidases selected for functional assessment across the five bifidobacterial reference strains. **a** Similarity plot showing the percentage of similarity of the 20 selected  $\beta$ -galactosidases across the reference strains (cut-off of 70% of identity over 90% of protein length). For each gene the relative cluster of appartenance from comparative genomics is also indicated. **b** Hierarchical clustering analysis showing the co-occurrence of the 20 selected  $\beta$ -galactosidases and their cluster of orthology across the reference genomes and bifidobacterial (sub)species





could be the plasmid instability due to the (size of the) inserted gene, or toxicity of their products in the cloning host. In one case (i.e. Blon\_2123 from *B. longum* subsp. *infantis* ATCC 15697), we found that the cloned gene contained a stop codon located 100 bp downstream of the start of the gene, thus making this a pseudogene. It is worth mentioning that BLAST analysis of this particular sequence against the deposited genome sequence of *B. longum* subsp. *infantis* ATCC 15697 showed that the publicly available sequence does not contain a stop codon at this point, suggesting that this mutation may be unique to our stock of *B. longum* ATCC 15697 (NB. the chromosomal sequence of this stock was verified by sequence analysis of a PCR product that encompassed the relevant section of Blon\_2123). The obtained *L. lactis* transformants (or at least a proportion of the colonies corresponding to a given transformation) from the remaining 17 cloning efforts were in 14 cases shown to form easily identifiable blue colonies on plates including the chromogenic substrate X-gal (yet in the absence of the inducer nisin) (Table 4). This result clearly indicates that the majority of the cloned genes indeed encodes a  $\beta$ -galactosidases active on this artificial substrate. In three cases (i.e. corresponding to the cloning of genes with locus tags Bbr\_0285, Bbr\_0310 and B7017\_2031) transformants only formed white colonies (though they contained the expected recombinant plasmid as verified by sequencing), suggesting that either the encoded product does not encode a  $\beta$ -galactosidase capable of cleaving

X-Gal, or that the cloned gene was not (highly) expressed or its product not properly folded (Table 4). The plasmid content of a selection of positively identified clones for each cloning was sent for sequencing to validate sequence integrity of the cloned gene. In summary, of the total of 20 selected (putative)  $\beta$ -galactosidase-encoding genes, 17 were successfully cloned (as an intact gene), of which 14 turned out to encode  $\beta$ -galactosidase activity as based on the applied X-gal plate assay.

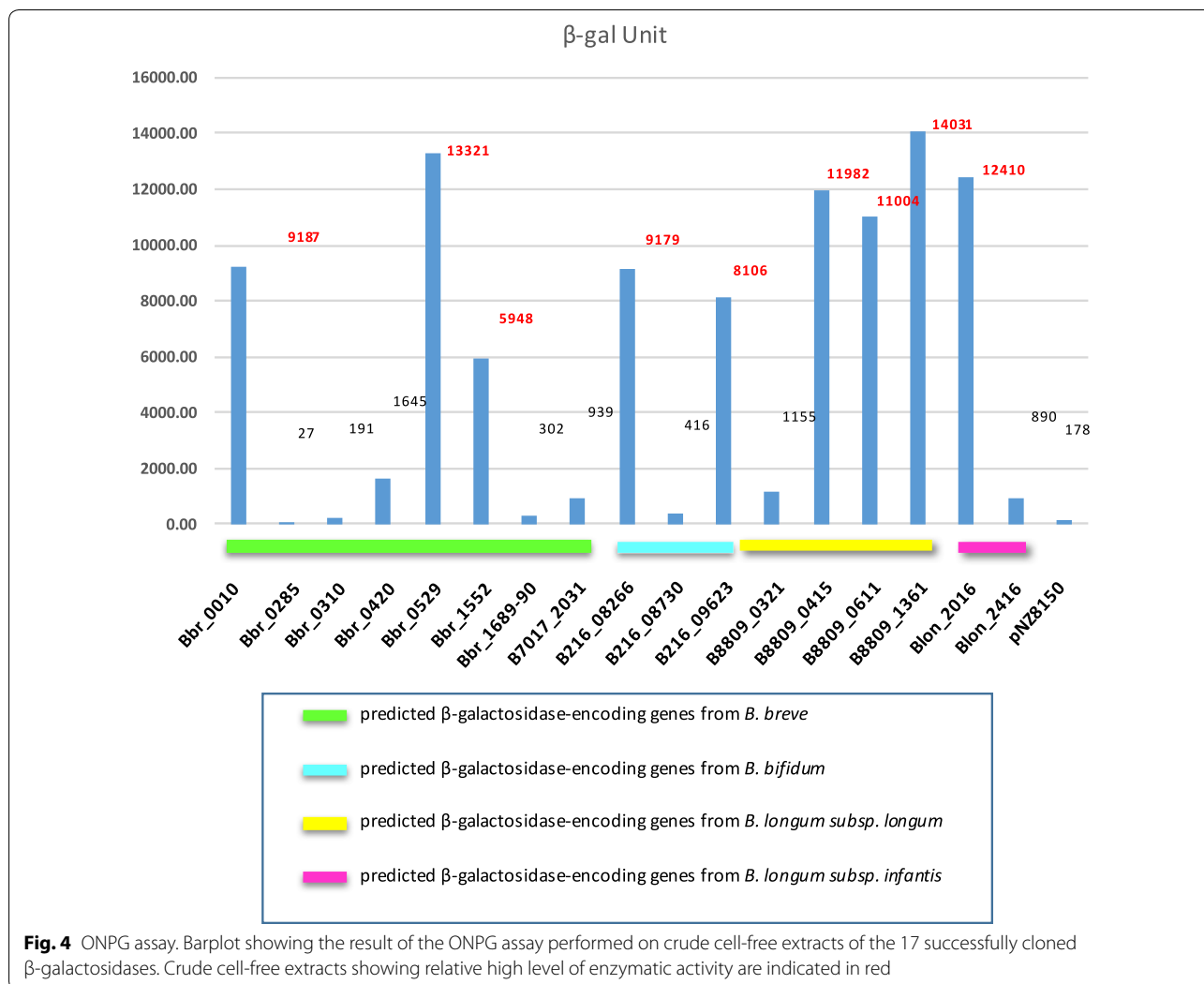
**$\beta$ -galactosidase assays: crude cell extract**

In order to determine if the products of the 17 cloned genes were active on substrates other than or in addition to X-gal,  $\beta$ -galactosidase assays were conducted using crude cell extracts obtained from individual, nisin-induced cultures of *L. lactis* in which a predicted  $\beta$ -galactosidase-encoding gene had been cloned (Additional file 1: Figure S2). Out of the 17 different crude cell extracts tested, nine were shown to exhibit high activity against the artificial substrate ONPG (between 5000 and 14,000 Miller units) (Fig. 4). In addition, five crude cell-free extracts were demonstrated to display a relative low level of enzymatic activity (between 400 and 2000 Miller units) to cleave this ONPG substrate, while the remaining three crude cell-free extracts (corresponding to the expressed proteins specified by Bbr\_0285, Bbr\_0310 and Bbr\_1689/1690) failed to show hydrolytic activity against this substrate (baseline set at 400 Miller units) (Fig. 4). It is worth mentioning that Bbr\_0285 and Bbr\_0310

**Table 4 Hydrolysis performance of the 17 cloned  $\beta$ -galactosidases tested on substrates lactose, lactulose and D-galactotriose**

ClusterID	Locus tag	X-GAL assay	ONPG crude extracts	Lactose			Lactulose			D-Galactotriose (Gal $\beta$ 1-4Gal $\beta$ 1-4Gal)		
				1h	3h	24h	1h	3h	24h	1h	3h	24h
<i>B. breve</i> UCC2003	Cluster 5	Bbr_0010	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
	Cluster 17	Bbr_0285	White	White	White	White	White	White	White	White	White	White
	Cluster 18	Bbr_0310	White	White	White	White	White	White	White	White	White	White
	Cluster 13	Bbr_0420	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
	Cluster 15	Bbr_0529	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
	Cluster 6	Bbr_1552	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
	Cluster 4	Bbr_1689-90	White	White	White	White	White	White	White	White	White	White
<i>B. breve</i> JCM 7017	Cluster 14	B7017_2031	White	White	White	White	White	White	White	White	White	White
<i>B. bifidum</i> LMG 13195	Cluster 15	B216_08266	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
	Cluster 12	B216_08730	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
	Cluster 3	B216_09623	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
<i>B. longum</i> subsp. <i>longum</i> 8809	Cluster 13	B8809_0321	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
	Cluster 15	B8809_0415	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
	Cluster 7	B8809_0611	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
	Cluster 8	B8809_1361	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
<i>B. longum</i> subsp. <i>infantis</i> ATCC 15697	Cluster 15	Blon_2016	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
	Cluster 14	Blon_2416	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green

Key to colour coding: Substrate hydrolysis observed (green), no hydrolysis (white) and substrate not assayed (grey). Results of X-gal and ONPG assays are also reported for comparative purposes



also did not form blue colonies on plates, thus providing further confirmation of their presumed inactivity as β-galactosidases (at least on the substrates tested, though lack of activity may also be due to lack of expression or protein misfolding).

**Protein overproduction, purification and hydrolysis assay**

All cloned (putative) β-galactosidase-encoding genes were then utilized for protein overproduction and subsequent His6-tagged protein purification using the *L. lactis* nisin-inducible protein production system (see “Materials and methods”). In order to verify the success of these steps, individually expressed and purified proteins were visualised by SDS-PAGE analysis followed by staining with Coomassie Brilliant Blue (Additional file 1: Figure S2). In order to assess substrate specificity of each of the 17 cloned (predicted) bifidobacterial β-galactosidases, qualitative hydrolysis assays by

HPAEC-PAD analysis were performed using a number of (β-)galactosidic linkage-containing carbohydrates. The (putative) β-galactosidases were first tested for hydrolytic activity towards lactose, lactulose and β-D-galactotriose (Galβ1-4Galβ1-4Gal) (Table 4). The obtained results demonstrate that the purified proteins exhibited varying ability to cleave these different sugars. In the reaction with lactose the best activity (i.e. complete or partial hydrolysis of the substrate observed within the first hour of reaction) was exhibited by the proteins corresponding to Bbr\_0010, Bbr\_0529 and Bbr\_1552 (all from *B. breve*), B216\_08266 (from *B. bifidum*), B8809\_0415, B8809\_0611 and B8809\_1361 (all from *B. longum subsp. longum*), and Blon\_2016 (from *B. longum subsp. infantis*) (Table 4). When the reaction was performed with lactulose the most active proteins were those corresponding to locus tags Bbr\_0529, Bbr\_1552, B216\_08266, B8809\_0415, B8809\_1361, Blon\_2016 and Blon\_2416 (Table 4). When

the hydrolysis reaction was performed with the substrate  $\beta$ -D-galactotriose (Gal $\beta$ 1-4Gal $\beta$ 1-4Gal), the proteins corresponding to locus tags Bbr\_0420, Bbr\_0529, B216\_08266, B216\_08730, B8809\_0321, B8809\_0415, Blon\_2016 and Blon\_2416 were shown to hydrolyse the substrate within the first hour of the applied reaction conditions (Table 4). Following these three tests it was decided to eliminate the proteins corresponding to locus tags Bbr\_0285, Bbr\_0310, Bbr\_1689/90, B216\_8730, B216\_09623 and B7017\_2031 from further analyses because they were shown to be unable to hydrolyse any of the sugars tested. Notably, among these negatives are Bbr\_0285 and Bbr\_0310, which also did not show activity towards X-gal or ONPG substrates (Table 4; Fig. 4). In addition Bbr\_1689/90, B216\_09623 and B7017\_2031 showed discordant activity towards hydrolytic activity towards X-gal, ONPG and lactose substrates. For this reason we decided to test their hydrolytic activity also on crude cell extract (which represents an unpurified enzyme; our purification method could have caused inactivation of the enzyme). Absence of protein expression was not considered to be the reason for the lack of activity of these enzymes because all these proteins were successfully overproduced and purified (Additional file 1: Figure S2). Based on the results of the hydrolysis assay on

crude cell extracts, B216\_09623 and B7017\_2031 were shown to be capable of lactose hydrolysis, thus suggesting that the protein purification may have negatively affected their activity. Hydrolysis assays employing the remaining eleven  $\beta$ -galactosidases were then performed utilizing commercially available  $\beta$ -galactosyl-containing structures (including HMOs and GOS) as potential substrates. In the case of galactobiose (Gal $\beta$ 1-6Gal), clear hydrolysis activity was obtained when the reaction was performed with the proteins corresponding to locus tags Bbr\_0529, Bbr\_1552, B216\_08266, B8809\_0415 and Blon\_2016. When the substrate galactotriose Gal $\alpha$ 1-3Gal $\beta$ 1-4Gal was used, no activity was detected within 24 h for any of the proteins tested, whereas four proteins (corresponding to Bbr\_0420, Bbr\_1552, B8809\_0321, Blon\_2416) were shown to be able to cleave galactobiose Gal $\beta$ 1-4Gal (Table 5). In order to further characterize the identified predicted  $\beta$ -galactosidases, their ability to hydrolyse the following carbohydrates, representing a selection of ( $\beta$ -linked galactose-containing) HMOs: lacto-*N*-tetraose (LNT), lacto-*N*-neotetraose (LNnT), 2'-fucosyllactose (2'-FL) and 3-fucosyllactose (3-FL) (Table 5). In the reaction with LNT the highest activity (positive within the first hour of reaction) was exhibited by the proteins corresponding to locus tags

**Table 5 Hydrolysis performance of the 11 selected  $\beta$ -galactosidases tested on galactobiose, galactotriose, LNT, LNnT, 2'-FL and 3-FL**

ClusterID	Locus tag	Galactobiose Gal $\beta$ 1-6Gal			Galactotriose Gal $\alpha$ 1-3Gal $\beta$ 1-4Gal			Galactobiose Gal $\beta$ 1-4Gal			Lacto- <i>N</i> -Tetraose LNT		
		1h	3h	24h	1h	3h	24h	1h	3h	24h	1h	3h	24h
<i>B. breve</i> UCC2003	Cluster 5	Bbr_0010											
	Cluster 13	Bbr_0420											
	Cluster 15	Bbr_0529											
	Cluster 6	Bbr_1552											
<i>B. bifidum</i> LMG 13195	Cluster 15	B216_08266											
<i>B. longum</i> subsp. <i>longum</i> 8809	Cluster 13	B8809_0321											
	Cluster 13	B8809_0415											
	Cluster 7	B8809_0611											
	Cluster 8	B8809_1361											
<i>B. longum</i> subsp. <i>infantis</i> ATCC 15697	Cluster 15	Blon_2016											
	Cluster 14	Blon_2416											

ClusterID	Locus tag	Lacto- <i>N</i> -Neotetraose LNnT			2'-Fucosyllactose 2'-FL			3'-Fucosyllactose 3'-FL		
		1h	3h	24h	1h	3h	24h	1h	3h	24h
<i>B. breve</i> UCC2003	Cluster 5	Bbr_0010								
	Cluster 13	Bbr_0420								
	Cluster 15	Bbr_0529								
	Cluster 6	Bbr_1552								
<i>B. bifidum</i> LMG 13195	Cluster 15	B216_08266								
<i>B. longum</i> subsp. <i>longum</i> 8809	Cluster 13	B8809_0321								
	Cluster 13	B8809_0415								
	Cluster 7	B8809_0611								
	Cluster 8	B8809_1361								
<i>B. longum</i> subsp. <i>infantis</i> ATCC 15697	Cluster 15	Blon_2016								
	Cluster 14	Blon_2416								

Key to colour coding: Substrate hydrolysis observed (green), no hydrolysis (white)

Bbr\_0529 and Bbr\_1552 (from *B. breve*), B216\_08266 (from *B. bifidum*), B8809\_0415, (from *B. longum* subsp. *longum*), and Blon\_2016 (from *B. longum* subsp. *infantis*) (Table 5). When the reaction was performed with LNnT as a substrate, the most active proteins were those corresponding to locus tags Bbr\_0010, Bbr\_0529, Bbr\_1552, B216\_08266, B8809\_0415, B8809\_0611, B8809\_1361 and Blon\_2016 (Table 5). When the hydrolysis reaction was performed with the substrate 2'-FL, it was found that none of the proteins is active on this substrate (Table 5). The last sugar tested was 3-FL, and in this case the most active proteins were Bbr\_0420, Bbr\_0529, B216\_08266, B8809\_0321, B8809\_0415, B8809\_0611, B8809\_1361 and Blon\_2016 (Table 5).

## Discussion

In the current work we investigated  $\beta$ -galactosidase enzymes belonging to four bifidobacterial species *B. breve*, *B. bifidum*, *B. longum* subsp. *longum* and *B. longum* subsp. *infantis*, all of which are commonly found in healthy, breast-fed infants. Our analysis clearly shows that multiple distinct  $\beta$ -galactosidases are encoded by these infant-derived bifidobacterial species (Additional file 1: Table S1) with a partial overlap between their hydrolytic activities (Tables 4 and 5). Notably, for some of the identified  $\beta$ -galactosidases the involvement in utilization of galacto-oligosaccharides and/or host-derived glycans has previously been reported (Additional file 1: Table S1). Of note, some of the clusters from our comparative analysis appeared to be species specific, suggesting that certain  $\beta$ -galactosidase encoding genes are only found within certain species. For example the enzymes belonging to Clusters 1, 2, 3, 6 and 9 contain a wide range of *B. bifidum*  $\beta$ -galactosidases which represent homologs of BbgIII (NCIMB 41171), BbgI (NCIMB 41171), BbgIV (NCIMB 41171), BIF2 (DSM 2015) and BIF1 (DSM 20215), respectively. Notably, these enzymes have previously been reported to be capable of GOS synthesis (Additional file 1: Table S1) (Goulas et al. 2007; Miwa et al. 2010; Moller et al. 2001).

Our comparative analysis was used as a guide to select 20 candidates to be used for further assessment of  $\beta$ -galactosidase activity towards a number of substrates including a range of selected  $\beta$ -galactosyl-containing structures. Cloning efforts of the selected genes resulted in the successful overexpression and purification of 17 presumed  $\beta$ -galactosidases, of which hydrolytic activity was further assessed on a variety of substrates. The obtained findings revealed that the majority of the expressed enzymes indeed exhibit  $\beta$ -galactosidase activity. However, these  $\beta$ -galactosidases clearly differ from

each other in their substrate specificity, since some of the enzymes are active on nearly all tested substrates, while others appear to be quite specific and merely hydrolyse one or two substrates.

Of note, the enzymes that are capable of hydrolysing the broadest range of selected substrates are members of Cluster 15 and are also widespread across the four (sub) species (i.e. Bbr\_0529, B216\_08266, B8809\_0415 and Blon\_2016), suggesting that these genes encode conserved  $\beta$ -galactosidase activities. Besides, enzymes of Cluster 15 appear to be homologous to the previously identified BbgII (from *B. bifidum* NCIMB 41171), which is capable of both GOS hydrolysis and synthesis (Goulas et al. 2007; Miwa et al. 2010; Moller et al. 2001) (Fig. 1). Notably, some of the  $\beta$ -galactosidases represented by this cluster have previously been shown to be involved in the utilization of either GOS and HMOs (Additional file 1: Table S1). In particular, it has been shown that Blon\_2016 from *B. longum* subsp. *infantis* ATCC 15697 represents a  $\beta$ -galactosidase possessing hydrolytic activity towards LNT and LNnT, as well as GOS (Garrido et al. 2013), which is indeed confirmed by our observations (Tables 4 and 5).

Another member of Cluster 15 is the  $\beta$ -galactosidase encoded by Bbr\_0529, which is required for the utilization of GOS and certain HMOs by *B. breve* UCC2003 (James et al. 2016; O'Connell Motherway et al. 2013), being consistent with our observed hydrolytic activities towards the majority of tested substrates including galactobiose, galactotriose, but also the central moieties of type I and type II HMOs (Tables 3, 4). As type I chains represent the most abundant HMO core structure in human milk (Urashima et al. 2012), the prevalence of Cluster 15 members across all four infant-derived bifidobacterial (sub)species (Fig. 3b) and observed ability to cleave LNT (but also LNnT) (Table 4) supports their relevance to bifidobacteria in obtaining access to substrates derived from a human milk-based diet.

Surprisingly, one of the  $\beta$ -galactosidases whose product was shown to be incapable of hydrolysing lactose or lactulose is Bbr\_0420, which was instead shown to be active towards D-galactotriose Gal $\beta$ 1-4Gal $\beta$ 1-4Gal (Table 4). Previous studies established that Bbr\_0420, a  $\beta$ -galactosidase which is upregulated when *B. breve* UCC2003 is grown on purified GOS (O'Connell Motherway et al. 2013), seems to be dedicated to the hydrolysis of D-galactotriose, which in turn is produced from the extracellular degradation of potato galactan in this strain (O'Connell Motherway et al. 2011). The transcription of this gene is, however, not induced when *B. breve* UCC2003 is grown on lactose and certain HMOs

(e.g. LNT or LNTnT) (James et al. 2016). Bbr\_0420 from *B. breve* UCC2003 and its homologue B8809\_0321 from *B. longum* subsp. *longum* NCIMB 8809 (both belonging to Cluster 13) were shown here to hydrolyse galactobiose (Gal $\beta$ 1-4Gal) and galactotriose (Gal $\beta$ 1-4Gal- $\beta$ 1-4Gal), but incapable of hydrolysing either LNT or LNTnT. This suggests that they represent  $\beta$ -galactosidases with a narrow substrate specificity, which is directed towards  $\beta$ 1,4 galactosidic links (Tables 4, 5). Interestingly, a recent study has associated a homolog corresponding to this  $\beta$ -galactosidase and located within the galactan cluster of *B. longum* AH1206 (BL1206\_0411) with the persistence of this strain in the GIT tract (Maldonado-Gomez et al. 2016). This finding represents another indication of the role played by  $\beta$ -galactosidases and their dietary galactose-containing substrates in promoting bifidobacterial colonization.

Notably, purified Bbr\_0010 from *B. breve* UCC2003 has previously been reported to hydrolyze lactose and LNTnT (James et al. 2016), while Bbr\_1552 represents a  $\beta$ -galactosidase with a broader substrate specificity, being involved in the hydrolysis of GOS, as well as lactose, LNT and LNTnT in this strain (James et al. 2016; O'Connell Motherway et al. 2013), which is also consistent with our observations. Interestingly, a previous study employing a Tn5-based random mutagenesis system in *B. breve* UCC2003 identified Bbr\_0010 as the main  $\beta$ -galactosidase responsible for lactose utilization in this strain (Ruiz et al. 2013). Based on our analysis this enzyme does indeed hydrolyze lactose (as well as LNTnT) and the corresponding gene is conserved across *B. breve* (as Cluster 5 members).

Two interesting cases are constituted by B8809\_0611 (Cluster 7) and B8809\_1361 (Cluster 8) from *B. longum* subsp. *longum* NCIMB 8809 as these two  $\beta$ -galactosidases show similar hydrolytic profile and are particularly active towards LNTnT (type II chain) despite being member of two different clusters (Tables 4 and 5). Cluster 7 member Blon\_2334 from *B. longum* subsp. *infantis* ATCC 15697 has been described as a  $\beta$ -galactosidase responsible for degradation of lactose and type II HMOs, but not type I HMOs (Yoshida et al. 2012), being consistent with what we observed for B8809\_0611. In contrast, B8809\_1361, which shows a hydrolytic profile similar to Blon\_2334 and B8809\_0611, belongs to a different cluster (Cluster 8). Finally, B8809\_1361 is only present in certain strains of *B. longum* subsp. *longum*, and perhaps constitutes an auxiliary  $\beta$ -galactosidase responsible for degradation of lactose and type II HMOs in this subspecies (Additional file 1: Table S1).

In conclusion, the information collected in this study highlights the importance of GHs in bifidobacterial

saccharolytic metabolism, in particular  $\beta$ -galactosidases which are involved in the utilization of a range of substrates such as lactose, HMOs, and GOS, found as part of the milk-based infant diet. The qualitative assay presented in this study provides a clear insight on the diversity of  $\beta$ -galactosidases in terms of substrate specificity. In fact, some appear to be more specialized towards milk-based substrates, while others are specific for plant-derived substrates. Therefore, the findings presented here constitute a solid foundation for future studies on bifidobacterial  $\beta$ -galactosidases and further investigation on the role of milk-derived substrates in establishing bifidobacterial predominance in the infant GIT.

## Additional file

**Additional file 1: Table S1.** List of bifidobacterial genes used for comparative purposes. **Figure S1.** *In silico* analysis and functional characterization of  $\beta$ -galactosidases. **Figure S2.** SDS-PAGE of purified proteins.

### Authors' contributions

DvS, MS and MOM conceived the study. VA, JOS and MOM performed the experiments with the help of BS and CL. FB performed the bioinformatics analyses. DvS, FB and VA wrote the manuscript. All authors read and approved the final manuscript.

### Author details

<sup>1</sup> APC Microbiome Ireland, University College Cork, Western Road, Cork, Ireland. <sup>2</sup> School of Microbiology, National University of Ireland, Western Road, Cork, Ireland. <sup>3</sup> FrieslandCampina, Amersfoort, The Netherlands.

### Acknowledgements

We thank Dr. Bas Kuiper and Dr. Eric Benjamins for their support in performing this study.

### Competing interests

AV, FB, JOS, MOM and DvS have no competing interests. MS, BS, and CL are employees of Friesland Campina.

### Availability of data and materials

The data supporting the conclusions of this article are available and included within the article.

### Consent for publication

Not applicable.

### Ethics approval and consent to participate

Not applicable.

### Funding

This work was financially supported by FrieslandCampina. DvS, VA, FB and MOM are members of APC Microbiome Ireland, which is a research centre funded by Science Foundation Ireland (SFI), through the Irish Government's National Development Plan. The authors and their work were supported by SFI (Grant SFI/12/RC/2273), FEMS Research Grant FEMS-RG-2016-0103 and HRB (Grant No. 513 PDTM/20011/9).

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 22 December 2018 Accepted: 11 January 2019

Published online: 19 January 2019

## References

- Akkerman R, Faas MM, de Vos P (2018) Non-digestible carbohydrates in infant formula as substitution for human milk oligosaccharide functions: effects on microbiota and gut maturation. *Crit Rev Food Sci Nutr*. <https://doi.org/10.1080/10408398.2017.1414030>
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410
- Arbolea S, Watkins C, Stanton C, Ross RP (2016) Gut bifidobacteria populations in human health and aging. *Front Microbiol* 7:1204
- Asakuma S, Hatakeyama E, Urashima T, Yoshida E, Katayama T, Yamamoto K, Kumagai H, Ashida H, Hirose J, Kitaoka M (2011) Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria. *J Biol Chem* 286(40):34583–34592
- Bottacini F, O'Connell Motherway M, Kuczynski J, O'Connell KJ, Serafini F, Duranti S, Milani C, Turrioni F, Lugli GA, Zomer A, Zhurina D, Riedel C, Ventura M, van Sinderen D (2014) Comparative genomics of the *Bifidobacterium breve* taxon. *BMC Genomics* 15:170
- de Ruyter PG, Kuipers OP, de Vos WM (1996) Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. *Appl Environ Microbiol* 62(10):3662–3667
- Egan M, Motherway MO, Kilcoyne M, Kane M, Joshi L, Ventura M, van Sinderen D (2014) Cross-feeding by *Bifidobacterium breve* UCC2003 during co-cultivation with *Bifidobacterium bifidum* PRL2010 in a mucin-based medium. *BMC Microbiol* 14:282
- Fuhrer A, Sprenger N, Kurakevich E, Borsig L, Chassard C, Hennet T (2010) Milk sialyllactose influences colitis in mice through selective intestinal bacterial colonization. *J Exp Med* 207(13):2843–2854
- Garrido D, Dallas DC, Mills DA (2013) Consumption of human milk glycoconjugates by infant-associated bifidobacteria: mechanisms and implications. *Microbiology* 159(Pt 4):649–664
- Garrido D, Ruiz-Moyano S, Kirmiz N, Davis JC, Totten SM, Lemay DG, Ugalde JA, German JB, Lebrilla CB, Mills DA (2016) A novel gene cluster allows preferential utilization of fucosylated milk oligosaccharides in *Bifidobacterium longum* subsp. *longum* SC596. *Sci Rep* 6:35045
- Goulas TK, Goulas AK, Tzortzis G, Gibson GR (2007) Molecular cloning and comparative analysis of four beta-galactosidase genes from *Bifidobacterium bifidum* NCIMB41171. *Appl Microbiol Biotechnol* 76(6):1365–1372
- James K, Motherway MO, Bottacini F, van Sinderen D (2016) *Bifidobacterium breve* UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways. *Sci Rep* 6:38560
- Koropatkin NM, Cameron EA, Martens EC (2012) How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol* 10(5):323–335
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227(5259):680–685
- Lewis ZT, Mills DA (2017) Differential Establishment of bifidobacteria in the breastfed infant gut. *Nestle Nutr Inst Workshop Ser* 88:149–159
- LoCasco RG, Ninonuevo MR, Freeman SL, Sela DA, Grimm R, Lebrilla CB, Mills DA, German JB (2007) Glycoprofiling of bifidobacterial consumption of human milk oligosaccharides demonstrates strain specific, preferential consumption of small chain glycans secreted in early human lactation. *J Agric Food Chem* 55(22):8914–8919
- Macfarlane GT, Steed H, Macfarlane S (2008) Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics. *J Appl Microbiol* 104(2):305–344
- Maldonado-Gomez MX, Martinez I, Bottacini F, O'Callaghan A, Ventura M, van Sinderen D, Hillmann B, Vangay P, Knights D, Hutkins RW, Walter J (2016) Stable engraftment of *Bifidobacterium longum* AH1206 in the human gut depends on individualized features of the resident microbiome. *Cell Host Microbe* 20(4):515–526
- Milani C, Lugli GA, Duranti S, Turrioni F, Mancabelli L, Ferrario C, Mangifesta M, Hevia A, Viappiani A, Scholz M, Arioli S, Sanchez B, Lane J, Ward DV, Hickey R, Mora D, Segata N, Margolles A, van Sinderen D, Ventura M (2015) Bifidobacteria exhibit social behavior through carbohydrate resource sharing in the gut. *Sci Rep* 5:15782
- Milani C, Duranti S, Bottacini F, Casey E, Turrioni F, Mahony J, Belzer C, Delgado Palacio S, Arbolea Montes S, Mancabelli L, Lugli GA, Rodriguez JM, Bode L, de Vos W, Gueimonde M, Margolles A, van Sinderen D, Ventura M (2017) The first microbial colonizers of the human gut: composition, activities, and health implications of the infant gut microbiota. *Microbiol Mol Biol Rev*. <https://doi.org/10.1128/MMBR.00036-17>
- Miwa M, Horimoto T, Kiyohara M, Katayama T, Kitaoka M, Ashida H, Yamamoto K (2010) Cooperation of beta-galactosidase and beta-N-acetylhexosaminidase from bifidobacteria in assimilation of human milk oligosaccharides with type 2 structure. *Glycobiology* 20(11):1402–1409
- Moller PL, Jorgensen F, Hansen OC, Madsen SM, Stougaard P (2001) Intra- and extracellular beta-galactosidases from *Bifidobacterium bifidum* and *B. infantis*: molecular cloning, heterologous expression, and comparative characterization. *Appl Environ Microbiol* 67(5):2276–2283
- O'Connell KJ, O'Connell Motherway M, O'Callaghan J, Fitzgerald GF, Ross RP, Ventura M, Stanton C, van Sinderen D (2013) Metabolism of four alpha-glycosidic linkage-containing oligosaccharides by *Bifidobacterium breve* UCC2003. *Appl Environ Microbiol* 79(20):6280–6292
- O'Connell Motherway M, Fitzgerald GF, van Sinderen D (2011) Metabolism of a plant derived galactose-containing polysaccharide by *Bifidobacterium breve* UCC2003. *Microb Biotechnol* 4(3):403–416
- O'Connell Motherway M, Kinsella M, Fitzgerald GF, van Sinderen D (2013) Transcriptional and functional characterization of genetic elements involved in galacto-oligosaccharide utilization by *Bifidobacterium breve* UCC2003. *Microb Biotechnol* 6(1):67–79
- O'Riordan K, Fitzgerald GF (1998) Evaluation of bifidobacteria for the production of antimicrobial compounds and assessment of performance in cottage cheese at refrigeration temperature. *J Appl Microbiol* 85(1):103–114
- Riviere A, Selak M, Lantin D, Leroy F, De Vuyst L (2016) Bifidobacteria and butyrate-producing colon bacteria: importance and strategies for their stimulation in the human gut. *Front Microbiol* 7:979
- Roger LC, Costabile A, Holland DT, Hoyles L, McCartney AL (2010) Examination of faecal *Bifidobacterium* populations in breast- and formula-fed infants during the first 18 months of life. *Microbiology* 156(Pt 11):3329–3341
- Ruiz L, Motherway MO, Lanigan N, van Sinderen D (2013) Transposon mutagenesis in *Bifidobacterium breve*: construction and characterization of a Tn5 transposon mutant library for *Bifidobacterium breve* UCC2003. *PLoS ONE* 8(5):e64699
- Sela DA (2011) Bifidobacterial utilization of human milk oligosaccharides. *Int J Food Microbiol* 149(1):58–64
- Smilowitz JT, Lebrilla CB, Mills DA, German JB, Freeman SL (2014) Breast milk oligosaccharides: structure-function relationships in the neonate. *Annu Rev Nutr* 34:143–169
- Tailford LE, Crost EH, Kavanaugh D, Juge N (2015) Mucin glycan foraging in the human gut microbiome. *Front Genet* 6:81
- Terzaghi BE, Sandine WE (1975) Improved medium for lactic streptococci and their bacteriophages. *Appl Microbiol* 29(6):807–813
- Turrioni F, Bottacini F, Foroni E, Mulder I, Kim JH, Zomer A, Sanchez B, Bidossi A, Ferrarini A, Giubellini V, Delledonne M, Henrissat B, Coutinho P, Oggioni M, Fitzgerald GF, Mills D, Margolles A, Kelly D, van Sinderen D, Ventura M (2010) Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-derived glycan foraging. *Proc Natl Acad Sci USA* 107(45):19514–19519
- Turrioni F, Peano C, Pass DA, Foroni E, Severgnini M, Claesson MJ, Kerr C, Hourihane J, Murray D, Fuligni F, Gueimonde M, Margolles A, De Bellis G, O'Toole PW, van Sinderen D, Marchesi JR, Ventura M (2012) Diversity of bifidobacteria within the infant gut microbiota. *PLoS ONE* 7(5):e36957
- Turrioni F, Milani C, Duranti S, Ferrario C, Lugli GA, Mancabelli L, van Sinderen D, Ventura M (2018a) Bifidobacteria and the infant gut: an example of co-evolution and natural selection. *Cell Mol Life Sci* 75(1):103–118
- Turrioni F, Milani C, Duranti S, Mahony J, van Sinderen D, Ventura M (2018b) Glycan utilization and cross-feeding activities by bifidobacteria. *Trends Microbiol* 26(4):339–350
- Urashima T, Asakuma S, Leo F, Fukuda K, Messer M, Oftedal OT (2012) The predominance of type I oligosaccharides is a feature specific to human breast milk. *Adv Nutr* 3(3):473S–482S
- van den Broek LA, Hinze SW, Beldman G, Vincken JP, Voragen AG (2008) *Bifidobacterium* carbohydrases-their role in breakdown and synthesis of (potential) prebiotics. *Mol Nutr Food Res* 52(1):146–163

Ward RE, Ninonuevo M, Mills DA, Lebrilla CB, German JB (2007) In vitro fermentability of human milk oligosaccharides by several strains of bifidobacteria. *Mol Nutr Food Res* 51(11):1398–1405

Wells JM, Wilson PW, Le Page RW (1993) Improved cloning vectors and transformation procedure for *Lactococcus lactis*. *J Appl Bacteriol* 74(6):629–636

Yoshida E, Sakurama H, Kiyohara M, Nakajima M, Kitaoka M, Ashida H, Hirose J, Katayama T, Yamamoto K, Kumagai H (2012) *Bifidobacterium longum*

subsp. *infantis* uses two different beta-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides. *Glycobiology* 22(3):361–368

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)

---