



Article

Revisiting the Metabolic Capabilities of *Bifidobacterium longum* subsp. *longum* and *Bifidobacterium longum* subsp. *infantis* from a Glycoside Hydrolase Perspective

Guillermo Blanco^{1,2}, Lorena Ruiz^{2,3,*}, Hector Tamés^{2,3}, Patricia Ruas-Madiedo^{2,3} , Florentino Fdez-Riverola^{1,4,5} , Borja Sánchez^{2,3}, Anália Lourenço^{1,4,5,6,*} and Abelardo Margolles^{2,3}

- ¹ Escuela Superior de Ingeniería Informática, Edificio Politécnico, Campus Universitario As Lagoas s/n, University of Vigo, 32004 Ourense, Spain; guillermo@guillermoblancos.es (G.B.); riverola@uvigo.es (F.F.-R.)
 - ² Department of Microbiology and Biochemistry of Dairy Products, Instituto de Productos Lácteos de Asturias (IPLA), Consejo Superior de Investigaciones Científicas (CSIC), Paseo Río Linares S/N, Villaviciosa, 33300 Asturias, Spain; hector.tames@ipla.csic.es (H.T.); ruas-madiedo@ipla.csic.es (P.R.-M.); borja.sanchez@csic.es (B.S.); amargolles@ipla.csic.es (A.M.)
 - ³ Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Oviedo, 33011 Asturias, Spain
 - ⁴ CINBIO-Centro de Investigaciones Biomédicas, University of Vigo, Campus Universitario Lagoas-Marcosende, 36310 Vigo, Spain
 - ⁵ SING Research Group, Galicia Sur Health Research Institute (IIS Galicia Sur), SERGAS-UVIGO, Hospital Álvaro Cunqueiro, 36312 Vigo, Spain
 - ⁶ CEB-Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal
- * Correspondence: lorena.ruiz@ipla.csic.es (L.R.); analia@uvigo.es (A.L.)

Received: 16 March 2020; Accepted: 11 May 2020; Published: 13 May 2020



Abstract: Bifidobacteria are among the most abundant microorganisms inhabiting the intestine of humans and many animals. Within the genus *Bifidobacterium*, several beneficial effects have been attributed to strains belonging to the subspecies *Bifidobacterium longum* subsp. *longum* and *Bifidobacterium longum* subsp. *infantis*, which are often found in infants and adults. The increasing numbers of sequenced genomes belonging to these two subspecies, and the availability of novel computational tools focused on predicting glycolytic abilities, with the aim of understanding the capabilities of degrading specific carbohydrates, allowed us to depict the potential glycoside hydrolases (GH) of these bacteria, with a focus on those GH profiles that differ in the two subspecies. We performed an in silico examination of 188 sequenced *B. longum* genomes and depicted the commonly present and strain-specific GHs and GH families among representatives of this species. Additionally, GH profiling, genome-based and 16S rRNA-based clustering analyses showed that the subspecies assignment of some strains does not properly match with their genetic background. Furthermore, the analysis of the potential GH component allowed the distinction of clear GH patterns. Some of the GH activities, and their link with the two subspecies under study, are further discussed. Overall, our in silico analysis poses some questions about the suitability of considering the GH activities of *B. longum* subsp. *longum* and *B. longum* subsp. *infantis* to gain insight into the characterization and classification of these two subspecies with probiotic interest.

Keywords: *Bifidobacterium*; *longum*; *infantis*; carbohydrates; glycoside hydrolases; computational screening

1. Introduction

Bifidobacteria are a group of ubiquitous microorganisms, whose main habitat includes different animals, including mammals, some insects and birds, although they have also been isolated from food and environmental samples, among other sources [1]. Several species can be found in humans and are considered among the first colonizers of the gut, constituting one of the most abundant genera in the fecal microbiota of healthy individuals [2]. Among them, *Bifidobacterium longum*, *Bifidobacterium breve*, *Bifidobacterium bifidum*, *Bifidobacterium animalis*, and *Bifidobacterium adolescentis* have received a QPS (Qualified Presumption of Safety) designation according to EFSA (European Food Safety Authority) [3]. For this reason, these QPS bifidobacterial species have been widely studied as potential probiotics and some strains are included in a large variety of food or food supplements due to the proven health benefits attributed to specific strains [4]. It is widely accepted that *B. longum* is currently subdivided into three subspecies: *B. longum* subsp. *longum*, *B. longum* subsp. *infantis*, and *B. longum* subsp. *suis* [5], although a fourth subspecies (*suillum*) has recently been proposed [6]. While the subspecies *suis* and *suillum* are not common in humans, they are mainly found in piglet feces, subspecies *longum* and *infantis* are very common in children, especially in breast-fed infants, the subspecies *longum* also being one of the most frequently detected bifidobacteria in the microbiota of healthy adult individuals [7,8].

Representatives of the subspecies *longum* and *infantis* were the first bifidobacteria to be fully sequenced [9,10]. *B. longum* is currently the species for which the highest number of fully sequenced genomes is available in public databases, the large majority of them belonging to the subspecies *longum* and, to a lesser extent, *infantis*. The extensive genetic information currently available for these two subspecies has allowed studies of comparative functional genomics that indicate some strategies used for these microorganisms to colonize the gastrointestinal tract. Among them, it is worth highlighting its ability to use a large amount of glycan rich sources. It is noteworthy that a large percentage of gene families in the *Bifidobacterium* pangenome are associated with carbohydrate metabolism, which allows these bacteria to metabolize host and diet derived carbohydrates that are not digested by the host and give them an adaptive advantage in the intestine. In general, genetic and functional studies seem to indicate that the subspecies *infantis* is better adapted to the intestinal environment of nursing children because it is capable of metabolizing human milk oligosaccharides (HMOs) [10–12]. Nonetheless, the carbohydrate metabolic capabilities in the subspecies *longum* seem to be more oriented to metabolize complex vegetable carbohydrates, and do not have the same ability as the subspecies *infantis* to degrade HMOs [13,14]. Within the enzymatic arsenal responsible for carbohydrate metabolism, glycoside hydrolases (GHs) constitute one of the main players and are responsible for the cleavage of complex carbohydrates, i.e., cleaving glycosidic bonds in poly- and oligosaccharides, and releasing shorter metabolizable products. However, the genetic information available in the databases does not always establish clearly different genetic patterns between the two subspecies.

In this work, we took advantage of novel computational tools to depict the GH capabilities of *B. longum* subsp. *longum* and *B. longum* subsp. *infantis*, with a specific focus on those activities that differentiate the two subspecies, with the aim of contributing to understanding metabolic traits that partially explain their differences in carbohydrate utilization.

2. Materials and Methods:

2.1. Phylogenetic Trees Based on 16S rRNA, *ldh*, *recA*, and *tuf* Genes

B. longum 16S rRNA gene sequences with a length >1500 bp and a sequence quality >90% were downloaded from the SILVA database in FASTA format [15]. Only the strains with assigned *longum* or *infantis* subspecies were used in the analysis. Additionally, complete and partial sequences of the genes *ldh*, *recA*, and *tuf* were analyzed. For multiple sequence comparison, Clustal Omega (version 1.2.4) was used for alignment [16]. Basically, a single plain text file containing all sequences was uploaded using HHaligh package, via pairwise comparisons. Progressive alignment was carried out using a hidden Markov model (HMM) and iterative clustering steps. Finally, multiple reference phylogenetic

trees were built up (via neighbor joining) and using data from pairwise comparison, trees were refined until finally one reference tree met criteria. Phylogenetic tree information was pasted into a plain text file again, and visualized in iTOL v5 (interactive Tree of Life) [17].

2.2. Whole Genome Phylogenetic Tree

From all *B. longum* completed and closed genomes available in the National Center of Biotechnology Information (NCBI) database, only the genomes with assigned *longum* or *infantis* subspecies were considered for phylogenetic tree building. Complete genome sequences were retrieved from the NCBI database without annotations. In order to enable a consistent comparative analysis, first open reading frame predictions and gene annotations were performed using PROKKA v1.12 [18]. Then, comparative genomics and phylogenetic tree construction were created using a ROARY pipeline [19]. Basically, GFF3 sequence outputs from Prokka were preclustered all against all using BLASTP and MCL (Markov Cluster), and after the preclustering step, further iterative comparisons between sequences with identity lower than 100% were performed using BLAST and clustering using CDhit algorithm. After grouping similar sequences in clustering steps, paralogs were split using conserved gene neighborhood (CGN) to finally distinguish between core, soft-core, shell, and cloud genes based on prevalence. All this information was used to create the phylogenetic tree using FastTree. Finally, iTOL was used for visualization of phylogenetic tree [17].

2.3. Genome Sequences of *B. longum*

To obtain protein files of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum*, the Gleukos tool was used [20]. Gleukos automatized the process of extracting and decompressing protein files of *Bifidobacterium* strains. With this tool we were able to retrieve genomic information from the NCBI genomic database and collect 188 genomes of different strains of *B. longum*, 171 of them had translated protein files (147 genomes of subspecies *longum*; 24 genomes of subspecies *infantis*). Supplementary File 1 includes a summary of the genomic information of each strain.

2.4. In Silico Analysis of *B. longum* GH Capabilities

The first step in this analysis entailed the compilation of the GHs of *B. longum* subsp. *longum* and *B. longum* subsp. *infantis*. The completed list of GHs is displayed in Supplementary File 2. Different data sources were used to gather together gene and protein information, namely the Carbohydrate-Active Enzymes database (CAZy) [21], the Kyoto Encyclopedia of Genes and Genomes (KEGG) [22], and the Genome database of NCBI [23]. To automatize this process, the Gleukos tool was used [20].

The glycolytic profiles of the reference strains were retrieved from the available genomes in KEGG. In particular, reference strains include three genomes of *B. longum* subsp. *infantis* (two of them belong to the same strain that has been sequenced in two different sequencing projects), i.e., *B. longum* subsp. *infantis* ATCC 15697 JGI (identified as “bln”), *B. longum* subsp. *infantis* ATCC 15697 Tokyo (identified as “blon”), and *B. longum* subsp. *infantis* 157F (identified as “blf”). Additionally, six strains of *B. longum* subsp. *longum* were analyzed as reference strains, i.e., *B. longum* subsp. *longum* JDM301 (identified as “bll”), *B. longum* subsp. *longum* BBM68 (identified as “blb”), *B. longum* subsp. *longum* JCM 1217 (identified as “blm”), *B. longum* subsp. *longum* KACC 91563 (identified as “blk”), *B. longum* subsp. *longum* F8 (identified as “blg”), and *B. longum* subsp. *longum* GT15 (identified as “blz”). For each of these strains, all annotated proteins were screened, keeping only GHs for further analyses. In total, this screening resulted in 417 GH protein matches.

The previously described 417 proteins were matched against all the genomes of *B. longum* strains with associated protein files. The tool BLASTp (version 2.7.1) [24] enabled the large-scale screening of glycolytic activities. Specifically, the BLAST database, containing the previously described 417 proteins, was built by executing the command “makeblastdb” with the parameter “dbtype prot”. The command “blastp”, with the parameters “outfmt 6” and “max_target_seqs 1”, enabled the matching of all the genomes of *B. longum* strains against this database (i.e., 171 strains with associated protein files). These

results were represented in a heatmap created with the Gleukos using InCHlib library [25], in which the Y axis represents the GH proteins and the X axis represents the collection of *B. longum* strains. Furthermore, each axis included a dendrogram with the corresponding hierarchical clustering of its data. These hierarchical clusters were built considering the Euclidean distance as intra-cluster proximity metric and average linkage as inter-cluster distance. Heatmap coloring represented the scoring of protein homologies, i.e., red indicates <50% identity and green >50% identity (with at least 50% identity across at least 50% of either protein sequence) [26].

3. Results and Discussion

In order to discern the global genetic and evolutionary variability among annotated *B. longum* subsp. *longum* and *B. longum* subsp. *infantis*, a clustering analysis was performed based on different genetic traits, including 16S rRNA gene and whole genome sequences. Phylogenetic analysis allowed us to construct two trees with strains of *B. longum* subsp. *longum* and *B. longum* subsp. *infantis*, whose complete genomes or 16S rRNA genes are available in public databases (Figure 1). First, a tree was built based on the comparison of complete genomes that had been annotated in the databases with the corresponding subspecies assignment (Figure 1A). Subsequently, a tree was constructed taking into account the sequences of the 16S rRNA gene, using only the sequences available in the SILVA database of those strains for which a subspecies has been assigned (Figure 1B). Furthermore, a clustering analysis considering full or partial sequences of *recA*, *tuf*, and *ldh* genes, previously used for subspecies identification in *B. longum* [27–30], was also carried out with the strains included in Figure 1A (Supplementary File 3). Neither of the trees have included strains for which the subspecies is not specified. That is, all the strains deposited as “*B. longum*”, without specifying the subspecies, were not taken into account in our analysis.

When comparing the complete genomes, two clearly defined clusters are observed, the most populated including 11 strains mainly belonging to the subspecies *longum*. The second one, consisting of three strains, includes two strains of *infantis* and one of *longum*. This result seems to indicate that there is no good correlation between the comparative genomic analysis and the corresponding subspecies assignment, since in the group dominated by strains belonging to the subspecies *longum*, strains deposited as *B. longum* subsp. *infantis* are also present (see CECT 7210 and 157F). Meanwhile, one *B. longum* subsp. *longum* strain (JDM301) is also found in the group with the majority of *infantis* strains. The *B. longum* subsp. *infantis* strains CECT 7210 and 157F are also included in branches dominated by *B. longum* subsp. *longum* strains when complete and partial sequences of the genes *recA* and *tuf* are used for the clustering analysis (Supplementary File 3). Therefore, this highlights that there could be certain discrepancies when assigning subspecies to strains that are genetically closer to the other subspecies. This lack of harmony between the genetic profile and the subspecies assignment in some strains is also evident when *Bifidobacterium* strains were clustered according to their rRNA 16S ribosomal gene sequence (Figure 1B).

Previous comparative genomics studies have already reported that some *B. longum* subsp. *infantis* strains might have been taxonomically misassigned [13,31]. O’Callaghan and coworkers carried out a phylogenetic analysis of 33 complete and incomplete *B. longum* strains and showed that *B. longum* subsp. *infantis* 157F is a member of the subspecies *longum* [13]. In our work, we also showed that the strain *B. longum* subsp. *infantis* CECT 7210 clusters together with other strains of the subspecies *longum*, and it is most likely a member of the *longum* subspecies. This strain, isolated from infant feces, provided in vivo protection against rotavirus infection in a mouse model [32]. Thus, our data suggest that care must be taken when comparing *B. longum* genomes deposited in public databases, since our results and the results of others show that subspecies assignment might have not been accurate for some strains. However, to avoid confusion, the probable misclassifications were not corrected in the subsequent comparative analysis of the GH families.

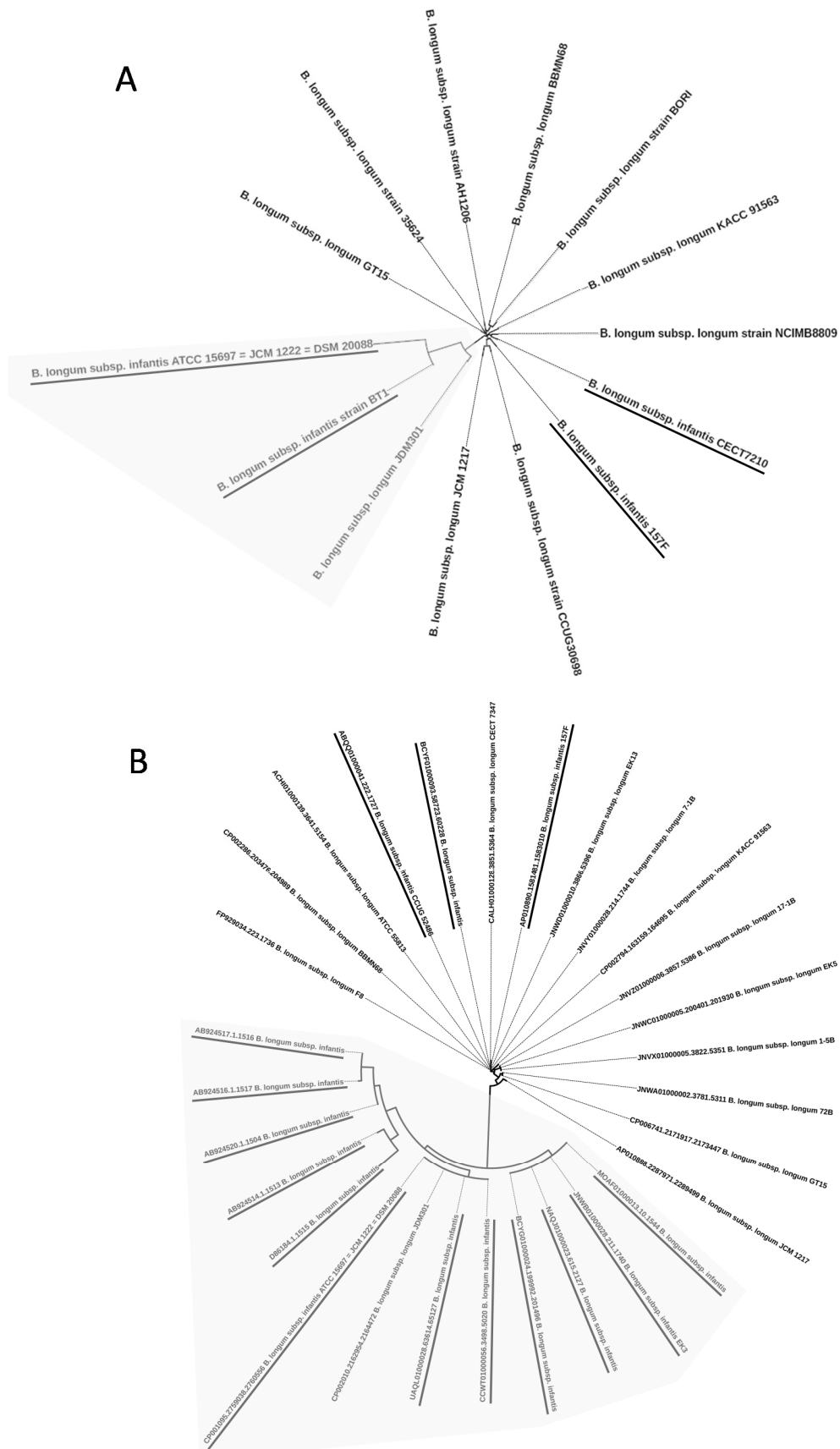


Figure 1. Genome-based (**A**) and 16S rRNA gene-based (**B**) clustering analysis. *Bifidobacterium longum* subsp. *infantis* strains are underlined. The clusters rich in *B. longum* subsp. *infantis* strains are represented with grey background.

In this context, we hypothesized that genetic traits relevant for bifidobacterial adaptation to the gastrointestinal tract ecosystem of different population groups, might add relevant information for *B. longum* subspecies classification. Indeed, traditionally, it has been documented that the metabolism of complex sugars is a differential trait of these two subspecies. In this regard, the metabolism of *infantis* seems to be oriented to the degradation of the oligosaccharides found in human milk, whereas the subspecies *longum* lacks this ability and focuses its carbohydrate metabolic potential in the utilization of complex vegetable carbohydrates obtained from diet [10,11,13,14,33,34]. Glycoside hydrolases (GHs), enzymes responsible for cleaving glycosidic bonds in sugar polymers, releasing shorter oligo- and polymers or sugar moieties, play a key role in many of the steps of the carbohydrate metabolic processes in bacteria [20]. A comparative analysis of the presence of GHs in the reference strains of these two subspecies (the nine reference *B. longum* strains previously defined and for which KEGG and functional annotations were publicly available), enabled us to detect 46 different GHs, 38 of them were found in both subspecies, and the results are shown in Figure 2 and Supplementary File 2. Furthermore, these 46 different GHs were distributed across 64 GH families, 57 of them being present in both subspecies. Figure 2 shows Venn diagrams displaying the distribution of different GHs and their family allocation by subspecies.

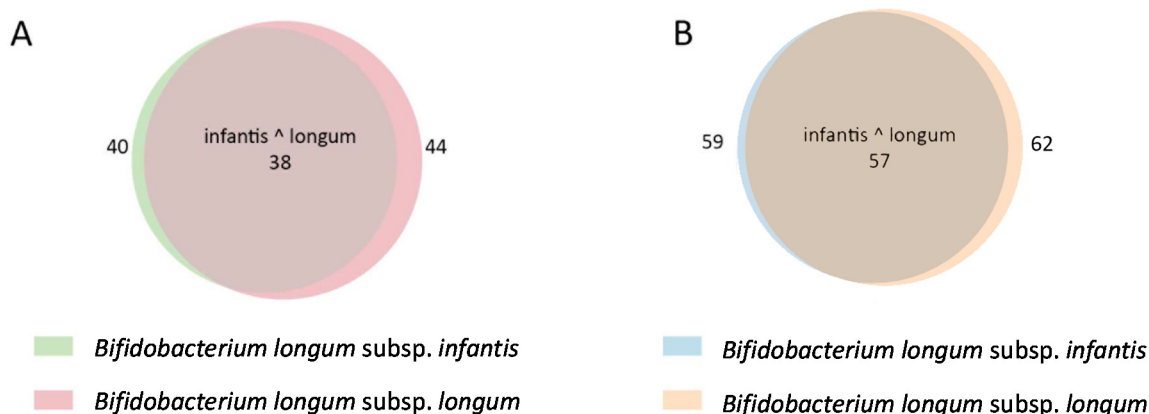


Figure 2. Venn diagram representing the distribution of glycoside hydrolases (GHs) in the *Bifidobacterium longum* reference strains (A) and the distribution of GH families (B). The numbers in the overlapping areas correspond to shared GHs (A) or GH families (B).

According to the Glycoside Hydrolase Family Classification (CAZY), and considering the reference strains, GHs from the families GH8, GH70, GH72, GH79, and GH94 are only present in the subspecies *longum*, and GH34 and GH83 have only been found in the subspecies *infantis*. Supplementary File 2 details the matching of these families and the different subspecies *longum* and *infantis*. Regarding specific GHs, the presence of EC 2.4.1.230 (kajibiose phosphorylase) and EC 3.2.1.18 (sialidase) was unique for *infantis*, and EC 2.4.1.4 (amylsucrase), EC 3.2.1.31 (beta-glucuronidase), EC 3.2.1.41 (pullulanase), EC 3.2.1.45 (glucocerebridase), EC 3.2.1.99 (arabinanase), and EC 3.2.1.156 (reducing end xylose-releasing exo-oligoxylanase) were only found in *longum*. These enzymes represent a GH spectrum that can allow the selection of specific carbohydrates preferentially metabolized by each subspecies. Indeed, some GH substrates containing the specific bonds hydrolyzed by these enzymes, have already been reported as preferential carbon sources for *longum* or *infantis*. This is the case of sialidases, since sialic acid decorates the structure of HMOs, whose specific metabolism has been associated with *infantis* [11]. Additionally, vegetable oligo and polysaccharides containing arabinose, xylose, or both, have been previously used to promote the growth of *B. longum* subsp. *longum*, as well as other *Bifidobacterium* species [35,36]. However, our analysis also showed kojibiose as a potential novel substrate specific for *infantis*. Kojibiose is a disaccharide product of glucose caramelization, composed of two glucose molecules linked through an alpha 1–2 bond, which is present in honey, beer, and other foods, normally in low amounts. Interestingly, the prebiotic and bifidogenic activity

of kojibiose and some kojibiose-derivatives have been previously studied [37], but their potential to specifically favor the growth of *infantis* and not that of *longum*, to the best of our knowledge, has not been explored.

In an attempt to perform a more rational analysis of the GH capabilities of *B. longum*, we used Gleukos, a new bioinformatics tool that may be of help in a wide range of biotechnological applications [20]. This new public Web portal integrates the CAZY and KEGG information and contributes with insights by extending the search scope of potential glycolytic sources in a custom set of publicly available genomes, without the need to conduct extensive genome annotations and comparative whole genomic analysis. Additionally, Gleukos enables the identification of the GH arsenal of intestinal bacteria at different taxonomic levels. Expanding the array of bacterial glycoside-acting enzymes capable of metabolizing specific bifidogenic substrates was our aim in this analysis. To do this, the Interactive Cluster Heatmap library (InChlib) was applied to display and cluster heatmaps [25]. As an example that illustrates the outcome of Gleukos, Figure 3 shows the heatmap for the 171 *B. longum* selected strains and the GHs belonging to the GH3 (Figure 3A), GH5 (Figure 3B), and GH10 (Figure 3C) families, three of the families containing GHs involved in the degradation of prebiotic products and widely found in *B. longum* (Supplementary File 2) and contain many enzymes. A global heatmap, considering all the GH families and the 171 *B. longum* strains, is depicted in Supplementary File 4. Additionally, an interactive heatmap showing the presence/absence of GH homologs in the 171 genomes is available in the URL <http://193.147.87.218:8888/gleukos/task?id=3279>. Remarkably, some of the GH enzymes associated with the subspecies *infantis* include sialidases, fucosidases, and *N*-acetyl-galactosaminidases, all of them involved in the degradation of complex HMOs [12,38,39]. In addition, EC.3.2.1.55 was the GH enzyme more frequently associated with *longum*. EC 3.2.1.55 are alpha-L-arabinofuranosidases that catalyze the hydrolysis of terminal non-reducing alpha-L-arabinofuranoside residues in alpha-L-arabinosides, frequently found as a component of hemicelluloses and pectins, and that cannot be absorbed by intestinal cells. A search within the six genomes of the subspecies *longum* strains included in the KEGG database (strains JDM 301, BBMN68, JCM 1217, KACC 91563, F8, and GT15), showed that the genes encoding these enzymes are present in the genomes organized in clusters of carbohydrate utilization. The most typical cluster structure is composed of nine genes, including genes coding for transcriptional regulators, ABC transporters, sugar binding proteins, a beta-galactosidase, and an alpha-L-arabinofuranosidase. This cluster structure is present in the six genomes analyzed, although the genome of the strain F8 also contains other alpha-L-arabinofuranosidase enzymes, organized in different clusters including other sugar-actin enzymes nearby, such as beta-xylosidases and alpha-glucosidases (Figure 4). Indeed, arabinose and xylose can often be together in arabinoxylan structures [40,41], suggesting that those bifidobacteria possessing arabinose and xylose acting enzymes would have an additional advantage to metabolize this substrate.

Several papers have reported the involvement of *Bifidobacterium* alpha-L-arabinofuranosidases in the hydrolysis of arabinans, arabinoxylans, and arabinogalactans [40–43], and the prebiotic and/or bifidogenic activity of these substrates [44,45]. Since cereals are very rich in arabinan-containing sugar polymers, the fact that alpha-L-arabinosidases are over-represented in the subspecies *longum*, but are absent or scarcely present in *infantis*, could indicate the adaptation of the subspecies *longum* to fiber-rich vegetable substrates in an adult-like diet. In this regard, it is worth noting that several authors have already demonstrated that *B. longum* subsp. *longum* representative strains are capable to degrade either arabinan or arabinoxylans [34,41,46]. Remarkably, among the bifidobacterial species with QPS status, L-arabinosidases and L-xylosidases appear to be restricted to *B. longum* subsp. *longum*, suggesting that it represents a specific trait related to this bifidobacterial group that could be exploited in future biotechnological applications aimed at developing novel subspecies specific bifidogenic substrates and symbiotic formulations.

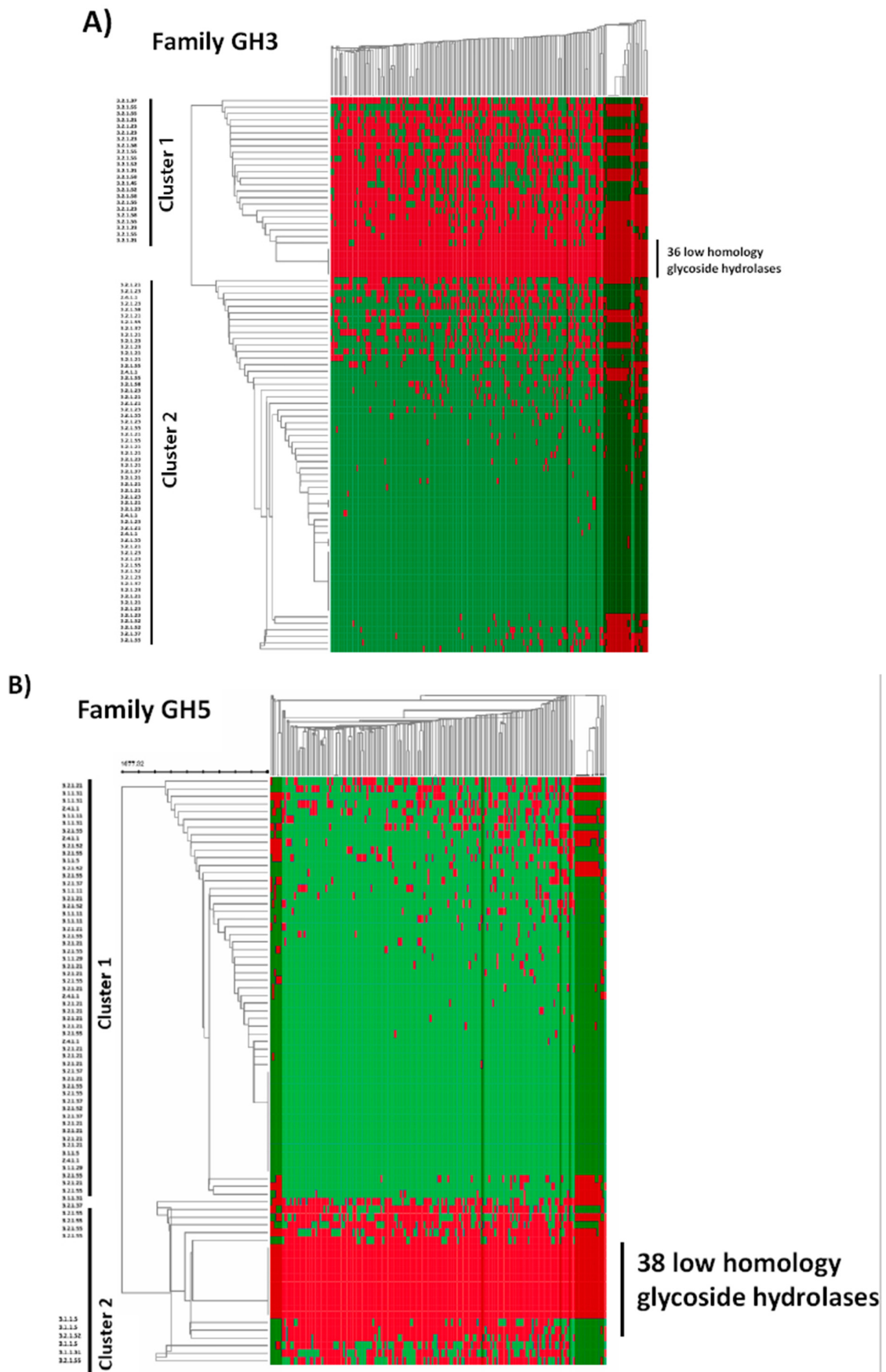


Figure 3. Cont.

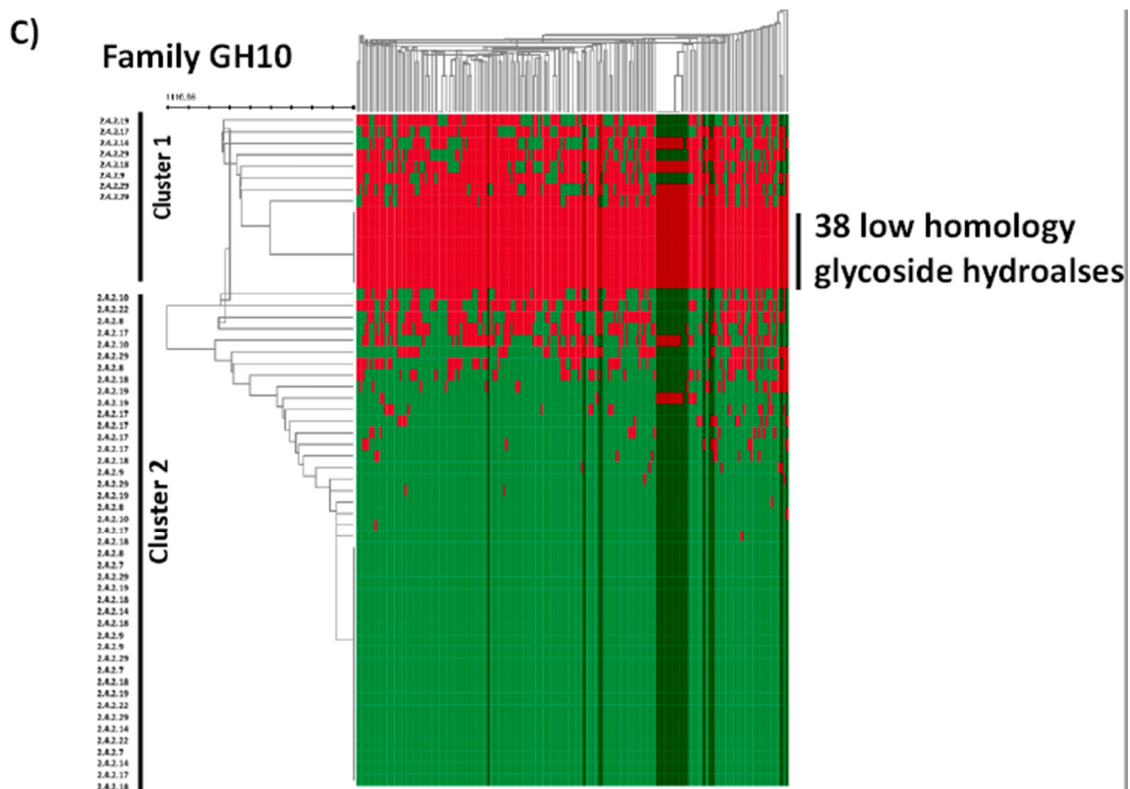


Figure 3. Heatmaps of the GH3 (A), GH5 (B), and GH10 families (C). The X axes represent 171 *B. longum* strains, and in the Y axes the different GH proteins are included. The red colors indicate homologies (protein level) below 50%, and the green color indicates that the homology is higher than 50%. The *B. longum* subsp. *infantis* strains are represented in darker columns.

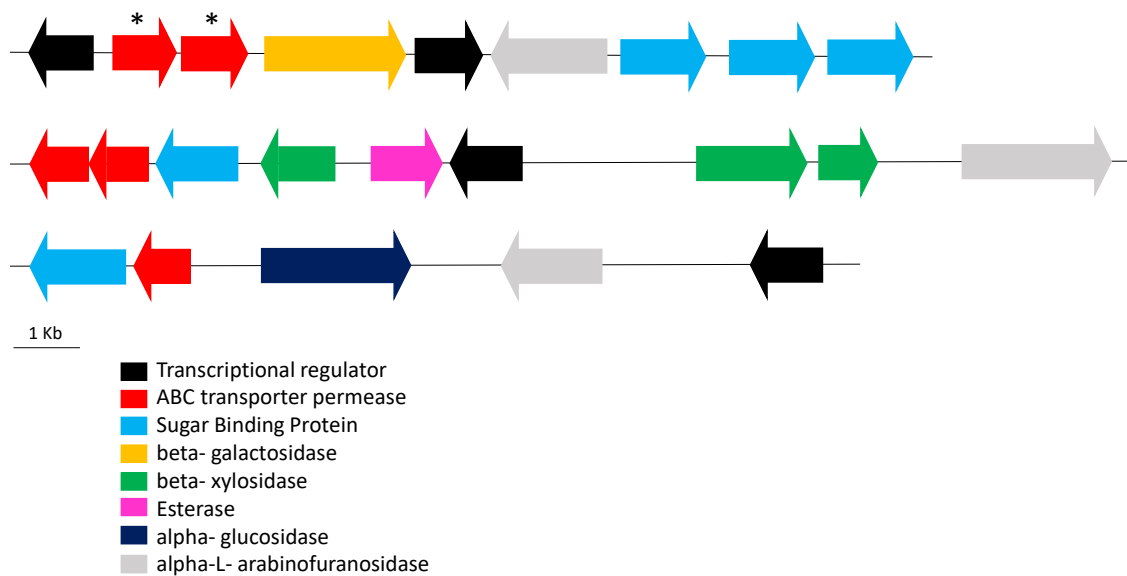


Figure 4. Representation of the different gene clusters containing alpha-L-arabinofuranosidases from family E.C. 3.2.1.55 in the genome of *B. longum* subsp. *longum* F8. *, genes not annotated in the genome of the F8 strain, but annotated as “ABC transporter permease” in the genomes of the strains JDM 301, BBMN68, JCM 1217, KACC 91563, and GT15.

4. Conclusions

The work presented in this manuscript highlights that the publicly available genomes of some strains of *B. longum* subsp. *longum* and *B. longum* subsp. *infantis* clearly depicted GH patterns associated with the two subspecies. Some of the GH activities discussed in this work, such as the preference of *B. longum* subsp. *infantis* for HMOs, were already reported in literature, but other observations open new possibilities to explore the GH activities of these two subspecies, in order to favor the representativeness of one or the other subspecies in the human gut microbiota, depending on the substrate composition of the surrounding environment, as well as the design of “a la carte” oligo and/or polysaccharides specifically targeted to fulfil the metabolic capabilities of these subspecies. In addition, our findings also showed the potential of the Gleukos tool to explore GH maps in *Bifidobacterium*, a process that could be extrapolated to other bacterial groups.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2076-2607/8/5/723/s1>, Supplementary File 1: Information of the *Bifidobacterium* genomes retrieved from the NCBI database for the present study; Supplementary File 2: GH and GH families considered in the current work. Green boxes indicate the presence of GH and GH families. Supplementary File 3: Clustering analysis of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* strains based on *ldh* (A,D), *recA* (B,E), and *tuf* (C,F) complete and partial sequences. Supplementary File 4: Heatmap of the different GH found in the 171 genomes of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum*. The X axes represent 171 *B. longum*, and in the Y axes the different GH proteins are included. The red colors indicate homologies (protein level) below 50%, and the green color indicates that the homology is higher than 50%. The *B. longum* subsp. *infantis* strains are represented in darker columns.

Author Contributions: Conceptualization, L.R., F.F.R., P.R.-M., B.S., A.L., A.M.; Methodology, G.B., L.R., H.T., A.L., A.M.; Software, G.B., A.L., B.S., F.F.-R.; Formal Analysis, G.B., H.T.; Writing—Original Draft Preparation, G.B., A.M.; Writing—Review and Editing, all authors. All authors have read and agreed to the published version of the manuscript.

Funding: The research included in the present work was funded by the grants AGL2016-78311-R (funded by AEI/FEDER, UE), RTI2018-095021-J-I00 (funded by MCIU/AEI/FEDER, UE), and IDI/2018/000236 (funded by PCTI Gobierno del Principado de Asturias/FEDER, UE). This study was supported by the Consellería de Educación, Universidades e Formación Profesional (Xunta de Galicia) under the scope of the strategic funding of ED431C2018/55-GRC Competitive Reference Group, the “Centro singular de investigación de Galicia” (accreditation 2019-2022) and the European Union (European Regional Development Fund—ERDF)-Ref. ED431G2019/06, and the Portuguese Foundation for Science and Technology (FCT) under the scope of the strategic funding of UIDB/04469/2020 unit and BioTecNorte operation (NORTE-01-0145-FEDER-000004) funded by the European Regional Development Fund under the scope of Norte2020—Programa Operacional Regional do Norte.

Acknowledgments: SING group thanks CITI (Centro de Investigación, Transferencia e Innovación) from the University of Vigo for hosting its IT infrastructure.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bottacini, F.; Ventura, M.; van Sinderen, D.; O’Connell Motherway, M. Diversity, ecology and intestinal function of bifidobacteria. *Microb. Cell Fact.* **2014**, *13* (Suppl. 1), S4. [[CrossRef](#)]
2. Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K.S.; Manichanh, C.; Nielsen, T.; Pons, N.; Levenez, F.; Yamada, T.; et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **2010**, *464*, 59–65. [[CrossRef](#)]
3. Koutsoumanis, K.; Allende, A.; Alvarez-Ordoñez, A.; Bolton, D.; Bover-Cid, S.; Chemaly, M.; Davies, R.; De Cesare, A.; Hilbert, F.; Lindqvist, R.; et al. Update of the list of QPS-recommended biological agents intentionally added to food or feed as notified to EFSA. *EFSA J.* **2019**, *15*, 5753.
4. Hidalgo-Cantabrana, C.; Delgado, S.; Ruiz, L.; Ruas-Madiedo, P.; Sánchez, B.; Margolles, A. Bifidobacteria and Their Health-Promoting Effects. *Microbiol. Spectr.* **2018**. [[CrossRef](#)]
5. Mattarelli, P.; Bonaparte, C.; Pot, B.; Biavati, B. Proposal to reclassify the three biotypes of *Bifidobacterium longum* as three subspecies: *Bifidobacterium longum* subsp. *longum* subsp. nov., *Bifidobacterium longum* subsp. *infantis* comb. nov. and *Bifidobacterium longum* subsp. *suis* comb. nov. *Int. J. Syst. Evol. Microbiol.* **2008**, *58*, 767–772. [[CrossRef](#)]

6. Yanokura, E.; Oki, K.; Makino, H.; Modesto, M.; Pot, B.; Mattarelli, P.; Biavati, B.; Watanabe, K. Subspeciation of *Bifidobacterium longum* by multilocus approaches and amplified fragment length polymorphism: Description of *B. longum* subsp. *suillum* subsp. nov.; isolated from the faeces of piglets. *Syst. Appl. Microbiol.* **2015**, *38*, 305–314. [[CrossRef](#)] [[PubMed](#)]
7. Milani, C.; Duranti, S.; Bottacini, F.; Casey, E.; Turrone, F.; Mahony, J.; Belzer, C.; Delgado-Palacio, S.; Arbolea-Montes, S. The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota. *Microbiol. Mol. Biol. Rev.* **2017**, *81*, e00036-17. [[CrossRef](#)] [[PubMed](#)]
8. Turrone, F.; Milani, C.; Duranti, S.; Mancabelli, L.; Mangifesta, M.; Viappiani, A.; Lugli, G.A.; Ferrario, C.; Gioiosa, L.; Ferrarini, A.; et al. Deciphering bifidobacterial-mediated metabolic interactions and their impact on gut microbiota by a multi-omics approach. *ISME J.* **2016**, *10*, 1656–1668. [[CrossRef](#)] [[PubMed](#)]
9. Schell, M.A.; Karmirantzou, M.; Snel, B.; Vilanova, D.; Berger, B.; Pessi, G.; Zwahlen, M.C.; Desiere, F.; Bork, P.; Delley, M.; et al. The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14422–14427. [[CrossRef](#)]
10. Sela, D.A.; Chapman, J.; Adeuya, A.; Kim, J.H.; Chen, F.; Whitehead, T.R.; Lapidus, A.; Rokhsar, D.S.; Lebrilla, C.B.; German, J.B.; et al. The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 18964–18969. [[CrossRef](#)]
11. Sela, D.A.; Li, Y.; Lerno, L.; Wu, S.; Marcobal, A.M.; German, J.B.; Chen, X.; Lebrilla, C.B.; Mills, D.A. An infant-associated bacterial commensal utilizes breast milk sialyloligosaccharides. *J. Biol. Chem.* **2011**, *286*, 11909–11918. [[CrossRef](#)] [[PubMed](#)]
12. Zabel, B.; Yde, C.C.; Roos, P.; Marcussen, J.; Jensen, H.M.; Salli, K.; Hirvonen, J.; Ouwehand, A.C.; Morovic, W. Novel Genes and Metabolite Trends in *Bifidobacterium longum* subsp. *infantis* Bi-26 Metabolism of Human Milk Oligosaccharide 2'-fucosyllactose. *Sci. Rep.* **2019**, *9*, 7983. [[CrossRef](#)] [[PubMed](#)]
13. O'Callaghan, A.; Bottacini, F.; O'Connell Motherway, M.; van Sinderen, D. Pangenome analysis of *Bifidobacterium longum* and site-directed mutagenesis through by-pass of restriction-modification systems. *BMC Genom.* **2015**, *16*, 832. [[CrossRef](#)] [[PubMed](#)]
14. Milani, C.; Turrone, F.; Duranti, S.; Lugli, G.A.; Mancabelli, L.; Ferrario, C.; van Sinderen, D.; Ventura, M. Genomics of the Genus *Bifidobacterium* Reveals Species-Specific Adaptation to the Glycan-Rich Gut Environment. *Appl. Environ. Microbiol.* **2015**, *82*, 980–991. [[CrossRef](#)]
15. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, D590–D596. [[CrossRef](#)]
16. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [[CrossRef](#)]
17. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [[CrossRef](#)]
18. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [[CrossRef](#)]
19. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T.; Fookes, M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **2015**, *31*, 3691–3693. [[CrossRef](#)]
20. Blanco, G.; Sánchez, B.; Ruiz, L.; Fdez-Riverola, F.; Margolles, A.; Lourenco, A. Computational approach to the systematic prediction of glycolytic abilities: Looking into human microbiota. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2020**. [[CrossRef](#)]
21. Cantarel, B.L.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for Glycogenomics. *Nucleic Acids Res.* **2009**, *37*, D233–D238. [[CrossRef](#)] [[PubMed](#)]
22. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [[CrossRef](#)] [[PubMed](#)]
23. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **2018**, *46*, D8–D13. [[CrossRef](#)] [[PubMed](#)]

24. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
25. Skuta, C.; Bartůněk, P.; Svozil, D. InChIlib-interactive cluster heatmap for web applications. *J. Cheminform.* **2014**, *6*, 44. [[CrossRef](#)]
26. Bottacini, F.; Morrissey, R.; Esteban-Torres, M.; James, K.; van Breen, J.; Dikareva, E.; Egan, M.; Lambert, J.; van Limpt, K.; Knol, J.; et al. Comparative genomics and genotype-phenotype associations in *Bifidobacterium breve*. *Sci. Rep.* **2018**, *8*, 10633. [[CrossRef](#)]
27. Kullen, M.J.; Brady, L.J.; O’Sullivan, D.J. Evaluation of using a short region of the *recA* gene for rapid and sensitive speciation of dominant bifidobacteria in the human large intestine. *FEMS Microbiol. Lett.* **1997**, *154*, 377–383. [[CrossRef](#)]
28. Roy, D.; Sirois, S. Molecular differentiation of *Bifidobacterium* species with amplified ribosomal DNA restriction analysis and alignment of short regions of the *ldh* gene. *FEMS Microbiol. Lett.* **2000**, *191*, 17–24. [[CrossRef](#)]
29. Sakanaka, M.; Nakakawaji, S.; Nakajima, S.; Fukiya, S.; Abe, A.; Saburi, W.; Mori, H.; Yokota, A. A Transposon Mutagenesis System for *Bifidobacterium longum* subsp. *longum* Based on an IS3 Family Insertion Sequence, ISBlo11. *Appl. Environ. Microbiol.* **2018**, *84*, e00824-18. [[CrossRef](#)]
30. Ventura, M.; Canchaya, C.; Meylan, V.; Klaenhammer, T.R.; Zink, R. Analysis, characterization, and loci of the *tuf* genes in *Lactobacillus* and *Bifidobacterium* species and their direct application for species identification. *Appl. Environ. Microbiol.* **2003**, *69*, 6908–6922. [[CrossRef](#)]
31. Chaplin, A.V.; Efimov, B.A.; Smeianov, V.V.; Kafarskaia, L.I.; Pikina, A.P.; Shkoporov, A.N. Intraspecies Genomic Diversity and Long-Term Persistence of *Bifidobacterium longum*. *PLoS ONE* **2015**, *10*, e0135658. [[CrossRef](#)]
32. Muñoz, J.A.; Chenoll, E.; Casinos, B.; Bataller, E.; Ramón, D.; Genovés, S.; Montava, R.; Ribes, J.M.; Buesa, J.; Fàbrega, J.; et al. Novel probiotic *Bifidobacterium longum* subsp. *infantis* CECT 7210 strain active against rotavirus infections. *Appl. Environ. Microbiol.* **2011**, *77*, 8775–8783.
33. Milani, C.; Lugli, G.A.; Duranti, S.; Turrone, F.; Mancabelli, L.; Ferrario, C.; Mangifesta, M.; Hevia, A.; Viappiani, A.; Scholz, M.; et al. Bifidobacteria exhibit social behavior through carbohydrate resource sharing in the gut. *Sci. Rep.* **2015**, *5*, 5782. [[CrossRef](#)]
34. Arboleya, S.; Bottacini, F.; O’Connell-Motherway, M.; Ryan, C.A.; Ross, R.P.; van Sinderen, D.; Stanton, C. Gene-trait matching across the *Bifidobacterium longum* pan-genome reveals considerable diversity in carbohydrate catabolism among human infant strains. *BMC Genom.* **2018**, *19*, 33. [[CrossRef](#)] [[PubMed](#)]
35. Rivière, A.; Selak, M.; Geirnaert, A.; Van den Abbeele, P.; De Vuyst, L. Complementary Mechanisms for Degradation of Inulin-Type Fructans and Arabinoxylan Oligosaccharides among Bifidobacterial Strains Suggest Bacterial Cooperation. *Appl. Environ. Microbiol.* **2018**, *84*, e02893-17. [[CrossRef](#)] [[PubMed](#)]
36. Pastell, H.; Westermann, P.; Meyer, A.S.; Tuomainen, P.; Tenkanen, M. In vitro fermentation of arabinoxylan-derived carbohydrates by bifidobacteria and mixed fecal microbiota. *J. Agric. Food Chem.* **2009**, *57*, 8598–8606. [[CrossRef](#)] [[PubMed](#)]
37. Ruiz-Aceituno, L.; Esteban-Torres, M.; James, K.; Moreno, F.J.; van Sinderen, D. Metabolism of biosynthetic oligosaccharides by human-derived *Bifidobacterium breve* UCC2003 and *Bifidobacterium longum* NCIMB 8809. *Int. J. Food Microbiol.* **2020**, *316*, 108476. [[CrossRef](#)]
38. Garrido, D.; Ruiz-Moyano, S.; Kirmiz, N.; Davis, J.C.; Totten, S.M.; Lemay, D.G.; Ugalde, J.A.; German, J.B.; Lebrilla, C.B.; Mills, D.A. A novel gene cluster allows preferential utilization of fucosylated milk oligosaccharides in *Bifidobacterium longum* subsp. *longum* SC596. *Sci. Rep.* **2016**, *6*, 35045. [[CrossRef](#)]
39. Sakanaka, M.; Hansen, M.E.; Gotoh, A.; Katoh, T.; Yoshida, K.; Odamaki, T.; Yachi, H.; Sugiyama, Y.; Kurihara, S.; Hirose, J.; et al. Evolutionary adaptation in fucosyllactose uptake systems supports bifidobacteria-infant symbiosis. *Sci. Adv.* **2019**, *5*, eaaw7696. [[CrossRef](#)]
40. Margolles, A.; de los Reyes-Gavilán, C.G. Purification and functional characterization of a novel alpha-L-arabinofuranosidase from *Bifidobacterium longum* B667. *Appl. Environ. Microbiol.* **2003**, *69*, 5096–5103. [[CrossRef](#)]
41. Komeno, M.; Hayamizu, H.; Fujita, K.; Ashida, H. Two Novel α -L-Arabinofuranosidases from *Bifidobacterium longum* subsp. *longum* Belonging to Glycoside Hydrolase Family 43 Cooperatively Degrade Arabinan. *Appl. Environ. Microbiol.* **2019**, *85*, e02582-18. [[CrossRef](#)] [[PubMed](#)]

42. Fujita, K.; Sasaki, Y.; Kitahara, K. Degradation of plant arabinogalactan proteins by intestinal bacteria: Characteristics and functions of the enzymes involved. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 7451–7457. [[CrossRef](#)] [[PubMed](#)]
43. Van Laere, K.M.; Beldman, G.; Voragen, A.G. A new arabinofuranohydrolase from *Bifidobacterium adolescentis* able to remove arabinosyl residues from double-substituted xylose units in arabinoxylan. *Appl. Microbiol. Biotechnol.* **1997**, *47*, 231–235. [[CrossRef](#)]
44. Broekaert, W.F.; Courtin, C.M.; Verbeke, K.; Van de Wiele, T.; Verstraete, W.; Delcour, J.A. Prebiotic and other health-related effects of cereal-derived arabinoxylans, arabinoxylan-oligosaccharides, and xylooligosaccharides. *Crit. Rev. Food Sci. Nutr.* **2011**, *51*, 178–194. [[CrossRef](#)]
45. Walton, G.E.; Lu, C.; Trogh, I.; Arnaut, F.; Gibson, G.R. A randomised, double-blind, placebo controlled cross-over study to determine the gastrointestinal effects of consumption of arabinoxylan-oligosaccharides enriched bread in healthy volunteers. *Nutr. J.* **2012**, *11*, 36. [[CrossRef](#)]
46. Rivière, A.; Gagnon, M.; Weckx, S.; Roy, D.; De Vuyst, L. Mutual Cross-Feeding Interactions between *Bifidobacterium longum* subsp. *longum* NCC2705 and *Eubacterium rectale* ATCC 33656 Explain the Bifidogenic and Butyrogenic Effects of Arabinoxylan Oligosaccharides. *Appl. Environ. Microbiol.* **2015**, *81*, 7767–7781.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).