



OPEN ACCESS

EDITED BY
Benoit Pugin,
ETH Zürich,
Switzerland

REVIEWED BY
Corina-Diana Ceapa,
National Autonomous University of Mexico,
Mexico
Li Weicheng,
Inner Mongolia Agricultural University,
China

*CORRESPONDENCE
Marco Ventura
✉ marco.ventura@unipr.it

SPECIALTY SECTION
This article was submitted to
Microbial Physiology and Metabolism,
a section of the journal
Frontiers in Microbiology

RECEIVED 23 December 2022
ACCEPTED 17 January 2023
PUBLISHED 08 February 2023

CITATION
Alessandri G, Fontana F, Tarracchini C,
Rizzo SM, Bianchi MG, Taurino G, Chiu M,
Lugli GA, Mancabelli L, Argentini C, Longhi G,
Anzalone R, Viappiani A, Milani C, Turroni F,
Bussolati O, van Sinderen D and
Ventura M (2023) Identification of a prototype
human gut *Bifidobacterium longum* subsp.
longum strain based on comparative and
functional genomic approaches.
Front. Microbiol. 14:1130592.
doi: 10.3389/fmicb.2023.1130592

COPYRIGHT
© 2023 Alessandri, Fontana, Tarracchini, Rizzo,
Bianchi, Taurino, Chiu, Lugli, Mancabelli,
Argentini, Longhi, Anzalone, Viappiani, Milani,
Turroni, Bussolati, van Sinderen and Ventura.
This is an open-access article distributed under
the terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Identification of a prototype human gut *Bifidobacterium longum* subsp. *longum* strain based on comparative and functional genomic approaches

Giulia Alessandri¹, Federico Fontana^{1,2}, Chiara Tarracchini¹,
Sonia Mirjam Rizzo¹, Massimiliano G. Bianchi^{3,4}, Giuseppe Taurino^{3,4},
Martina Chiu³, Gabriele Andrea Lugli¹, Leonardo Mancabelli^{3,4},
Chiara Argentini¹, Giulia Longhi^{1,2}, Rosaria Anzalone²,
Alice Viappiani², Christian Milani^{1,4}, Francesca Turroni^{1,4},
Ovidio Bussolati^{3,4}, Douwe van Sinderen⁵ and Marco Ventura^{1,4*}

¹Laboratory of Probiogenomics, Department of Chemistry, Life Sciences, and Environmental Sustainability, University of Parma, Parma, Italy, ²GenProbio srl, Parma, Italy, ³Department of Medicine and Surgery, University of Parma, Parma, Italy, ⁴Microbiome Research Hub, University of Parma, Parma, Italy, ⁵APC Microbiome Institute and School of Microbiology, Bioscience Institute, National University of Ireland, Cork, Ireland

Bifidobacteria are extensively exploited for the formulation of probiotic food supplements due to their claimed ability to exert health-beneficial effects upon their host. However, most commercialized probiotics are tested and selected for their safety features rather than for their effective abilities to interact with the host and/or other intestinal microbial players. In this study, we applied an ecological and phylogenomic-driven selection to identify novel *B. longum* subsp. *longum* strains with a presumed high fitness in the human gut. Such analyses allowed the identification of a prototype microorganism to investigate the genetic traits encompassed by the autochthonous bifidobacterial human gut communities. *B. longum* subsp. *longum* PRL2022 was selected due to its close genomic relationship with the calculated model representative of the adult human-gut associated *B. longum* subsp. *longum* taxon. The interactomic features of PRL2022 with the human host as well as with key representative intestinal microbial members were assayed using *in vitro* models, revealing how this bifidobacterial gut strain is able to establish extensive cross-talk with both the host and other microbial residents of the human intestine.

KEYWORDS

probiotics, bifidobacteria, host-microbe interactions, microbe-microbe interactions, RNA sequencing, intestinal barrier, inflammation

Introduction

Members of the genus *Bifidobacterium* are abundant symbiotic microorganisms inhabiting the gastrointestinal tract of mammals (Alessandri et al., 2021). This bacterial genus currently counts more than 100 species (Lugli et al., 2021; Li et al., 2022), and in recent decades particular scientific focus has been addressed on *Bifidobacterium longum* subsp. *longum*, due to its purported health benefits and because it is the most prevalent and abundant bifidobacterial species colonizing the human intestine throughout the entire lifespan (Santos et al., 2021; Kim et al., 2022; Quintanilha et al., 2022). In this

context, convincing scientific evidence has accumulated demonstrating that several *B. longum* subsp. *longum* strains are able to support intestinal barrier integrity, provide protection against pathogen proliferation, limit the onset/exacerbation of inflammatory gut diseases and metabolic disorders, and modulate host immunity to promote gut homeostasis (Schiavi et al., 2016, 2018; McCarville et al., 2017; Alessandri et al., 2019; Yan et al., 2019; Hao et al., 2022; Quintanilha et al., 2022). Furthermore, *B. longum* subsp. *longum* strains not only significantly contribute to host metabolism by degrading, through saccharolytic activities, a wide variety of otherwise non-digestible glycans, but also produce various metabolites, including polyphenols, vitamins, conjugated linoleic acids and short-chain fatty acids (SCFAs) that are believed to beneficially impact on both the gut microbial ecosystem and epithelial host cells (Tanno et al., 2019; Blanco et al., 2020; Alessandri et al., 2021; Sasaki et al., 2021; Komeno et al., 2022).

However, despite clear experimental evidence supporting the purported probiotic features of certain *B. longum* subsp. *longum* strains, the molecular mechanisms behind such health-promoting activities are still far from being fully understood (Turroni et al., 2014; Hidalgo-Cantabrana et al., 2017; Kim et al., 2018). Certainly, the rapid evolution of systems biology toward high-throughput “omics” technologies allowed mechanistic investigation of probiotic actions, an approach sometimes also referred to as probiogenomics (Ventura et al., 2009, 2012; Turroni et al., 2014). The latter has, in particular due to genome decoding and functional analysis of bacterial strains, contributed to improve insights into the diversity and evolution of health-promoting bacteria, while in some cases it has also shed light on the mechanisms underlying their beneficial effects (Ventura et al., 2012; Turroni et al., 2018a; Choi et al., 2021). However, despite our deepening insights into probiotic mechanisms, most of the probiotic strains that are currently on the market have been tested and selected principally for their survival and safety features, thus limiting knowledge about the ability of these strains to interact with their host and/or with other microbial players of the human gut environment.

In this context, to identify a prototype human gut *B. longum* subsp. *longum* strain with a presumed high fitness in the human gut, we applied an ecological and phylogenomic-driven selection approach. These analyses allowed identification of *B. longum* subsp. *longum* PRL2022 as the genetically and functionally closest strain to the most representative strains of this bifidobacterial species present in the human intestinal microbiota, thus highlighting this microorganism as a model strain for this (sub)species. Genomic insights of PRL2022 were obtained, and a comparative genomic analysis was performed that included the genome of PRL2022 and those of *B. longum* subsp. *longum* strains that are currently included in commercially available probiotic products, leading to the identification of PRL2022-associated unique genes putatively involved in host- and/or microbe-microbe interactions, suggesting its adaptation to the human gut ecological niche. In order to validate such *in silico*-based data, functional genomic experiments using *in vitro* models of the human gut were carried out, showing how this putative model microorganism is able to establish extensive cross-talks with both the host and other microorganisms of the intestinal ecosystem.

Materials and methods

Dataset selection

All publicly available datasets corresponding to fecal samples of healthy adults sequenced through a shotgun metagenomics approach were selected and downloaded from the National Center for

Biotechnology Information (NCBI) SRA repository using NCBI SRA Toolkit 2.11.0 faster-dump (Leinonen et al., 2011). Specifically, to achieve high quality and coverage data, only shotgun metagenomic data sets based on Illumina sequencing technology were retained (Supplementary Table S1).

Bifidobacterium longum subsp. *longum* genome selection

To select a model candidate of the *B. longum* subsp. *longum* species, all publicly available *B. longum* subsp. *longum* genomes from the NCBI genome list were downloaded and then subjected to genome quality assessment. In detail, only genome sequences with a genome coverage higher than 30-fold and containing less than 100 contigs were considered. Subsequently, the retained high-quality genomes were processed through the DRep and CheckM tools to cluster identical genomes, i.e., those with an Average Nucleotide Identity (ANI) \geq 99%, to eliminate redundant genomes, and to obtain a high-quality non-redundant *B. longum* subsp. *longum* genome database (Supplementary Table S2).

Identification of novel model candidates among the *Bifidobacterium longum* subsp. *longum* genomic database

The retained high-quality non-redundant *B. longum* subsp. *longum* genomes together with the.fastq corresponding to the selected shotgun metagenomics sequencing-based publicly available dataset from adult fecal samples were submitted to and processed by the open-source Strain Genome Explorer (StrainGE) software¹ (Supplementary Tables S1, S2). Specifically, Strain Genome Search Tool (StrainGST), i.e., an integrated tool of the StrainGE suite, allowed detection of bifidobacterial strains included in the abovementioned *B. longum* subsp. *longum* genomic database in the selected shotgun metagenomics dataset of fecal samples from adult (van Dijk et al., 2022). StrainGE allows assessment of the prevalence of each genome contained in the *B. longum* subsp. *longum* genomic database within the analyzed shotgun metagenomics samples. Subsequently, to identify the model candidate among the *B. longum* subsp. *longum* strains included in the genomic database, we assessed the AxP index, defined as [average ANI value calculated for each genome] * [prevalence score of the strain in the shotgun metagenomics dataset] * [100] (Fontana et al., 2022).

Application of the RefBifSelector

Since the identified optimal candidate among the *B. longum* subsp. *longum* genomic database is not available in our laboratory, to identify, among those strains available in our local bacterial repository, a *B. longum* subsp. *longum* model strain genomically and functionally closest to the above-identified optimal candidate, the RefBifSelector tool was used (Fontana et al., 2022). In detail, the tool enables the identification of the strain, within a local biobank repository, that is

1 <https://github.com/broadinstitute/strainge>

genomically the most closely related to the identified optimal model strain, as based on a genomic comparison. Specifically, this selection is driven by ANI value in conjunction with the use of the average Percentage of Positive Scoring matches (PPOS). While ANI analysis investigates the nucleotide identity between genome pairs, the PPOS score is calculated from the comparison of the translated amino acid sequences obtained through Blastp analysis following the [(number of identical matches) + (number of similar matches)]/(alignment length) formula (Fontana et al., 2022).

Genome sequencing and assembly

To obtain a fully representative sequence of *B. longum* subsp. *longum* PRL2022 genome, the previously obtained sequencing data from Illumina platform were combined with long reads obtained from genome sequencing on a MinION sequencer (Oxford Nanopore, United Kingdom). Specifically, the DNA extracted from the selected strain was subjected to whole-genome sequencing using a MinION (Oxford Nanopore, United Kingdom). DNA library preparation was performed using the Native Barcoding Genomic DNA kit according to the manufacturer's instructions. One microgram input DNA from each sample was used for library preparation. The samples were pooled and diluted to a final concentration of 5–50 fmol. Sequencing was performed on MinION device (Oxford Nanopore, United Kingdom) using a R9.4.1 flow cell. Long reads were filtered by quality using the Filtlong tool,² while short reads were filtered through the fastq-mcf script.³ Then, filtered fastq files of MinION long reads obtained from genome sequencing efforts were used as input for genome assembly through CANU software (Koren et al., 2017). The resulting genome sequences have been polished through Polypolish (Wick and Holt, 2022) using Illumina paired-end reads (250 bp). Polished contigs were then employed by MEGAnnotator (Lugli et al., 2016) for the prediction of protein-encoding open reading frames (ORFs) using Prodigal (Hyatt et al., 2010).

Gene annotation and prediction of genes involved in host- and microbe-microbe interaction in PRL2022 genome

Open reading frames (ORFs) in the PRL2022 genome were predicted and annotated using the most recent release of the MEGAnnotator pipeline (Lugli et al., 2016). In detail, contigs longer than 1,000 bp were employed to predict protein encoding ORFs through Prodigal v2.0 (Hyatt et al., 2010). Predicted ORFs were then functionally annotated using RAPSearch2 (cutoff *e*-value of 1×10^{-5} and minimum alignment length 20) using the NCBI reference sequences (RefSeq) database (Zhao et al., 2012a) coupled with hidden Markov model profile (HMM) searches against the manually curated Pfam-A database (cutoff *e*-value of 1×10^{-10}). The PRL2022 genome sequence was subjected to prediction of genes encoding glycosyl hydrolases (GHs), glycosyl transferases (GTs), carbohydrate binding modules (CBMs), and carbohydrate esterases (CEs) through

sequence similarity search in the carbohydrate-active enzyme (CAZy) database (Lombard et al., 2014) using HMMER v3.3 (Wheeler and Eddy, 2013) (cutoff *e*-value of 1×10^{-15}) and BLASTP analysis (Altschul et al., 1990) (cutoff *e*-value of 1×10^{-10}). The presence of pilus gene clusters was scrutinized through homology search tool RAPSearch2 (cutoff *e*-value of 1×10^{-5} and minimum alignment length 20) (Zhao et al., 2012a) exploiting the custom pilus gene database previously described (Milani et al., 2017). In addition, to identify loci encoding exopolysaccharides, protein sequences of priming glycosyltransferases (pGTFs) were retrieved from the NCBI database and were compared against the predicted PRL2022 proteome. Once the putative pGTF was identified, the genomic regions flanking the corresponding gene were investigated to identify genes involved in EPS biosynthesis, i.e., glycosyl transferases, flippases, ABC transporters, and carbohydrate precursor biosynthesis/modification enzymes. Finally, the presence of putative bile salt hydrolase- and serpin-encoding genes was investigated through sequence similarity searches in the protein sequence NCBI database (cutoff *e*-value of 1×10^{-5}).

Comparative genome analysis

The PRL2022 genome together with various genomes of *B. longum* subsp. *longum* strains that are currently included in a variety of commercially available probiotic products (Tarracchini et al., 2022; Supplementary Table S3), were subjected to a pangenome analysis pipeline (PGAP) (Zhao et al., 2012b). Predicted proteome of a specific *B. longum* subsp. *longum* strain was screened for orthologous encoding genes against the proteome of the other considered *B. longum* subsp. *longum* strains employing BLAST analysis (cutoff *e*-value of 1×10^{-10} and exhibiting at least 50% identity across at least 80% of both protein sequences) (Altschul et al., 1990). The obtained data were then clustered into protein families, i.e., clusters of orthologous genes (COGs) employing the Markov clustering algorithm (Enright et al., 2002), by means of the method gene family (GF). Based on the presence/absence matrix encompassing all COGs identified in the analyzed genomes, unique genes present in PRL2022 genome and not in the other 10 considered genomic sequences were identified. Functional annotation of each unique gene was accomplished using the EggNog database (Huerta-Cepas et al., 2016).

Cultivation conditions

PRL2022 strain was cultivated in the de Man-Rogosa-Sharpe (MRS) medium (Sharlau Chemie, Spain) supplemented with 0.05% (wt/vol) L-cysteine hydrochloride (Merk, Germany) and incubated at 37°C in a chamber (Concept 400, Ruskinn) with an anaerobic atmosphere (2.99% H₂, 17.01% CO₂, and 80% N₂).

pH, sodium chloride, and bile salts tolerance tests

To evaluate the ability of the selected strain to tolerate various pH levels, *B. longum* subsp. *longum* PRL2022 was cultivated in 10 mL of MRS broth at 37°C under anaerobic conditions to reach a final concentration of 10^8 cells/mL. Subsequently, cells were centrifuged at

2 <https://github.com/rwwick/Filtlong>

3 <https://github.com/ExpressionAnalysis/ea-utils>

3,000 rpm for 8 min, washed with a saline solution (PBS, pH 6.5) and resuspended in 10 mL of MRS broth whose pH was adjusted to pH 2.0, pH 3.0, or pH 4.0 with the addition of HCl. Cells were incubated under anaerobic conditions at 37°C for 2 h, as previously described (Yasmin et al., 2020). The same procedure was performed to assess the ability of PRL2022 to tolerate different NaCl (2, 6, and 10%) or bile salt (0.5, 1, and 2%) concentrations with an exposure of 3 h to these stressful conditions, as previously reported (Yasmin et al., 2020). All experiments were carried out in triplicate and a control sample was obtained by inoculating PRL2022 cells in MRS broth. After incubation, cell viability was evaluated by means of the LIVE/DEAD BacLight Bacterial Viability kit (ThermoFisher Scientific, United States) and an Attune NxT flow cytometer (ThermoFisher Scientific, United States).

Flow cytometry bacterial viability assay

Following exposure to acidic environment, or various bile salt or NaCl concentrations, a 10-fold serial dilution in Phosphate Buffered Saline (PBS) was obtained from each tested condition. The diluted cells were then used for a flow cytometry cell viability assay using the fluorescent dyes SYTO9 (3.34 mM) and PI (20 mM) of the LIVE/DEAD BacLight Bacterial Viability kit (ThermoFisher Scientific, United States), following the manufacturer's protocol (Manual of the LIVE/DEAD BacLight Bacterial Viability and counting kit, ThermoFisher Scientific, United States). Briefly, two aliquots of 1 ml of bacterial cell dilution (1:1000) were harvested by centrifugation at 3,000 rpm for 8 min and washed with PBS. Subsequently, one of the two aliquots of bacterial suspension was exposed to 70% isopropyl alcohol and kept on ice for 1 h to permeabilize cell membranes and cause cell death, while the other 1 mL aliquot was maintained in PBS to preserve cell viability. Subsequently, 1.5 μ L of a specific dye was added to samples for single staining assay, while 1.5 μ L of both dyes were added for the double staining assay. Once stained, samples were incubated in the dark for 15 min at room temperature. Furthermore, while single-stained controls were used for instrument parameter adjustment, non-stained cells were used as a background control. Cell viability assay was performed with the Attune NxT flow cytometer (ThermoFisher Scientific, United States), and all data were analyzed with the Attune NxT flow cytometer software.

Antibiotic susceptibility assay

Strain PRL2022 susceptibility to ampicillin, vancomycin, gentamicin, kanamycin, streptomycin, erythromycin, clindamycin, tetracycline, and chloramphenicol was investigated by Minimum Inhibitory Concentration (MIC) assays using the broth microdilution method (MDIL) following the European Food Safety Authority (EFSA) recommendations (Masco et al., 2006; EFSA, 2012). Briefly, from each antibiotic stock solution, two-fold dilution series were obtained and aliquoted in a 96-well microtiter plate. Subsequently, an overnight culture of strain PRL2022 was diluted to obtain an Optical Density at 600 nm (OD_{600nm}) = 1, and 15 μ L of the diluted cells were inoculated in 135 μ L of culture medium supplemented with a specific antibiotic concentration. Microplates were incubated under anaerobic conditions at 37°C for 48 h. The MIC breakpoints expressed in μ g/mL represent the highest concentration of a given antibiotic to which the strain was shown to be resistant. All experiments were performed in triplicate. The same procedure was used for the quality control strain *B. longum* subsp. *longum* ATCC 15707^T (ES ISO 10932:2010).

Human cell line culture

Human colorectal adenocarcinoma-derived Caco-2 cells (purchased from ATCC) and human colon carcinoma-derived mucin-secreting goblet cells HT29-MTX (kindly provided by Prof. Antonietta Baldi, University of Milan) were cultured in Minimum Essential medium (MEM) and Dulbecco's Modified Eagle's medium (DMEM) with high glucose (4.5 g/L) and 10 mM of sodium pyruvate, respectively, as previously described (Bianchi et al., 2019). Both media were supplemented with 10% Fetal Bovine Serum (FBS), 2 mM glutamine, 100 g/mL streptomycin, and 100 U/mL penicillin. Cultures were maintained at 5% CO₂ at 37°C and passaged three times a week. Subsequently, a mixed suspension of Caco-2 and HT29-MTX cells (7:3) was seeded in DMEM + FBS at a density of $\approx 10^5$ cells/cm² into cell culture inserts with membrane filters (pore size 0.4 μ m) for Falcon 24-well-multitrays (Becton, Dickinson & Company, Franklin Lakes, NJ, United States), and cultured for 21 days with a medium replacement every 3 days until a tight monolayer was formed (TEER >600 $\Omega \cdot$ cm²).

Human cell line monolayer in contact with PRL2022

After 21 days of seeding, the culture medium of the 24-well plates was replaced with fresh antibiotic-free DMEM. Subsequently, PRL2022 cells (final concentration 10⁸ cells/mL) were added to Caco-2/HT29-MTX cell monolayer, as previously described (Serafini et al., 2013; Fontana et al., 2022). The 24-well plates were subsequently incubated at 5% CO₂ at 37°C. After 4 h of incubation, bacterial cells and human cell lines were separately recovered in RNA later and preserved at -80°C until processing.

Specifically, for this trial, *B. longum* subsp. *longum* PRL2022 was grown in MRS broth under anaerobic conditions at 37°C. Once the exponential growth phase ($0.6 < OD_{600nm} < 0.8$) was reached, bifidobacterial cells were enumerated by using the Thoma cell counting chamber (Herka), possibly diluted to reach a final concentration of 1×10^8 cells/mL, washed in PBS, resuspended in 400 μ L of antibiotic-free DMEM, and seeded on Caco-2/HT29-MTX cell monolayers. The strain PRL2022 resuspended in DMEM and maintained under the same incubation conditions of the 24-well plates without any contact with human cell lines was used as bacterial control sample, while Caco-2/HT29-MTX cell monolayers without any bifidobacterial seeding were used as human cell line control sample. All experiments were carried out in triplicate.

Prokaryotic and eukaryotic RNA extraction

Total RNA from bacterial cells was isolated using a previously described method (Turrone et al., 2016; Milani et al., 2020). Briefly, bifidobacterial cell pellets were resuspended in 1 mL of QIAzol lysis reagent (Qiagen, Germany) in a sterile tube containing glass beads. Cells were lysed by alternating 2 min of stirring the mix on a bead beater with 2 min of static cooling on ice. These steps were repeated three times. Lysed cells were centrifuged at 12,000 rpm for 15 min, and the upper phase was recovered. Bacterial RNA was subsequently purified using the RNeasy Mini Kit (Qiagen, Germany) following the manufacturer's instruction. Total RNA from human cell lines was extracted by adding 350 μ L of RLT buffer from the RNeasy Mini Kit (Qiagen, Germany) and

following the manufacturer's instructions. In both cases, the RNA concentration and purity were evaluated using a spectrophotometer (Eppendorf, Germany).

Prokaryotic and eukaryotic mRNA sequencing analysis

Total bacterial RNA (from 100 ng to 1 µg) was treated to remove rRNA by means of the QIAseq FastSelect – 5S/16S/23S following the manufacturer's instructions (Qiagen, Germany). The yield of rRNA depletion was checked with a 2,200 TapeStation (Agilent Technologies, United States). Then, a whole transcriptome library for both prokaryotic and eukaryotic RNA was constructed using the TruSeq Stranded mRNA Sample preparation kit (Illumina, San Diego, USA). Samples were then loaded onto a NextSeq high-output v2 kit (150 cycles) (Illumina) as indicated by the technical support guide. The obtained prokaryotic reads were filtered to remove low-quality reads (minimum mean quality 20, minimum length 150 bp) as well as any remaining ribosomal locus-encompassing reads using the METAnnotator X2 pipeline (Milani et al., 2021). Subsequently, the retained reads were aligned to the specific reference genome through Bowtie2 software (Langdon, 2015), while high quality .fastq were aligned to the Human reference genome sequence (GRCh38.p13) by the use of the splice-aware STAR algorithm (version 2.7.10a) (Dobin et al., 2013), and quality of the alignments was evaluated using Picard software tool (version 2.26.11) (<https://broadinstitute.github.io/picard/>). Subsequently, quantification of reads mapped to individual transcripts was achieved through htseq-counts script of HTSeq software in “union” mode (Anders et al., 2015). Raw counts were then normalized using CPM (Counts per million mapped reads) for filtering genes with low counts (CPM <1) and TMM (Trimmed Mean of *M*-Values) for statistically robust differential gene expression analysis through the EdgeR package (Robinson et al., 2010). Evaluation of expression differences was calculated for each gene as log₂ fold change (log₂FC) of average expression between the control (no contact between human cell lines and strain PRL2022) and “treated” samples (contact between human cell lines and strain PRL2022). Additionally, for each comparison, a Volcano plot was created to simultaneously visualize expression changes (log fold change) and their statistical significance (*value of p*).

In vitro evaluation of PRL2022 response to human gut microbiota exposure

To evaluate how PRL2022 interacts with the gut microbial community, batch cultures were set up to co-cultivate the selected strain with two different bacterial communities, previously stabilized through a bioreactor system (Solaris Biotech Solutions, Italy) in IGS medium (Alessandri et al., 2022), dominated by species of either *Bacteroides* or *Prevotella*, i.e., two of the most abundant and representative genera of the human gut microbiota (Arumugam et al., 2011; Costea et al., 2018).

Batch cultures were obtained by inoculating 0.1% (vol/vol) of a stabilized intestinal microbial community and 1% (vol/vol) of an overnight culture of the strain PRL2022, as previously described (Mancabelli et al., 2021), in 30 mL of IGS medium adjusted to 6.8 ± 0.2 pH to mimicking the human intestinal environment (Alessandri et al., 2022). In addition, a batch culture with 1% (vol/vol) of strain PRL2022 was obtained as a control sample. All microbial cultures were performed

in triplicate and incubated under anaerobic conditions at 37°C. After 8 h of incubation, cultures were centrifuged at 7,000 rpm for 5 min, the supernatants were discarded, while the obtained bacterial pellets were used for RNA extraction. RNA extraction and sequencing, as well as RNA sequencing analysis were performed as described above.

Statistical analysis

For differential gene expression analysis, EdgeR package was used to estimate the statistical significance of differences between fold changes as the False Discovery Rate (FDR).

Data availability statement

Raw sequences of the RNAseq experiments are accessible under BioProject accession number PRJNA914637 and PRJNA833139. The updated genome sequence of *B. longum* subsp. *longum* PRL2022 is available under the accession number PRJNA692178.

Results and discussion

Ecological and phylogenomic-driven identification of model/prototype candidates belonging to the *Bifidobacterium longum* subsp. *longum* species

Bacterial strains that naturally colonize the human intestine are expected to possess high ecological fitness and to be equipped with a genomic assembly that supports their persistence in the intestinal environment. To identify a candidate prototype with a presumed high fitness in the human intestinal ecosystem among the *B. longum* subsp. *longum* strains whose genome sequences are publicly available, an ecological and phylogenomic-driven selection approach was performed, which is based on a recent published protocol with some modifications (Fontana et al., 2022). Specifically, all publicly available *B. longum* subsp. *longum* genomes were first checked for genome completeness, presence of contaminant genomic sequences and genome heterogeneity. Genomes that passed this initial quality check were then assessed for genome redundancy resulting in the retention of 306 high-quality, non-redundant *B. longum* subsp. *longum* genomes (Supplementary Table S2). Then, a strain tracking analysis was performed to predict the ecological distribution or prevalence, of these high-quality non-redundant genomes in the intestinal microbiota of adults. For this purpose, a total of 4,020 publicly available Illumina shotgun metagenomics-based fecal samples of healthy adults, covering 81 different cohorts from various geographical areas, were selected, and used for such a strain tracking exercise (Supplementary Table S1). Subsequently, strain prevalence and average ANI data were integrated into a specific index score, the AxP, as previously described (Fontana et al., 2022). This analysis revealed that among the selected publicly available genomes, *B. longum* subsp. *longum* strain GCA_015551245.1 showed the highest AxP score, thus highlighting the latter as the ecologically and genomically most representative reference strain for the *B. longum* subsp. *longum* species (Supplementary Table S4). However, since this strain is not publicly available, all genomes of *B. longum* subsp. *longum* strains from our in-house microbial

repository were screened to identify the closest related strain to the above-defined optimal model microorganism by using the previously assessed RefBifSelector, which is a tool specifically established for biobank screening based on genomic and functional similarity to an identified model strain (Fontana et al., 2022). This screening assay revealed that the highest score was obtained for *B. longum* subsp. *longum* PRL2022, indicating this strain as a very closely related strain to the above identified reference strain for the *B. longum* subsp. *longum* species (Supplementary Table S5). However, to validate this *in silico*-driven strain selection, genomic insights and *in vitro* analyses were performed to assess PRL2022 safety as well as its ability to cope with human gastrointestinal challenges and to interact with both host cells and other intestinal microorganisms (see also Supplementary Text for further information).

Unique genomic features of PRL2022 with respect to other *Bifidobacterium longum* subsp. *longum* probiotic strains

To assess unique features encoded by the PRL2022 genome that may for example enhance its fitness or gut-associated microbe-host/microbe-microbe interactions, a comparative genomics analysis was performed by considering the genomes of 10 *B. longum* subsp. *longum* strains currently included in commercially probiotic products (Tarracchini et al., 2022). Specifically, to provide a fully representative PRL2022 genome sequence, the PRL2022 genome generated by combining sequencing data obtained from both Illumina sequencing (sequencing output of 556,485 short reads) and Oxford Nanopore MinION (sequencing output of 57,198 long reads) was used. This analysis revealed the presence of 279 genes in the PRL2022 genome that are not present in the genomes of other probiotic *B. longum* subsp. *longum* strains. Insights into these genes highlighted that PRL2022 is provided with “unique genes” involved in the breakdown of complex polysaccharides (Supplementary Table S6), including, genes encoding a β -galactosidase and α -L-arabinofuranosidase (GH39 and GH43) (Supplementary Table S6). These enzymes are required for the degradation of a complex polysaccharide commonly present in plant cell walls, i.e., arabinogalactan, whose degradation has been demonstrated to be strain-dependent within the taxon *Bifidobacterium longum* subsp. *longum* (Fujita et al., 2019; Wang and LaPointe, 2020; Sasaki et al., 2022). Furthermore, PRL2022 was shown to possess various genes involved in the transport/uptake of glycans as well as various glycosyl transferases (Supplementary Table S6). These findings suggest that PRL2022 can activate specific (poly)saccharide-related degradative and transport strategies to compete with other intestinal microorganisms for nutrients. In addition, various genetic sequences corresponding to proteins or protein domains involved in adhesion to/interaction with human intestinal epithelial cells, including Ig-like domain, fibronectin type III domain, von Willebrand-domain, and collagen-binding surface protein, were identified as unique genes in PRL2022 (Supplementary Table S6). This indicates that PRL2022 possesses specific genetic features to facilitate interaction with intestinal cells to ensure its colonization and persistence in the human intestine. Finally, the PRL2022 genome harbors “unique genes” that may have a role in limiting the growth of other microorganisms, including various cysteine/histidine-dependent aminohydrolases/peptidases (CHAP) domain-containing protein as well as a type II toxin-antitoxin system

(Supplementary Table S6). Specifically, while CHAP domain-containing proteins exhibit lytic activity toward peptidoglycans (Guglielmetti et al., 2014), the toxin-antitoxin system is involved in bacterial persistence, biofilm formation, and antibiotic tolerance, with the toxin causing cell death or suppression of cell proliferation and antitoxin acting as an antagonist of the toxin action (Klimina et al., 2019, 2020). Therefore, PRL2022 appears to possess specific strategies to not only adhere to/interact with the host and compete for nutrients, but also to limit growth of other microorganisms ensuring its colonization and survival in the competitive human intestinal environment.

Evaluation of the ability of PRL2022 strain to interact/communicate with the human host through *in vitro* assays

A key biological feature that may be exerted by a bacterium belonging to the autochthonous human gut microbiota, such as PRL2022, is represented by its ability to establish cross-talk with the host intestinal epithelium. Therefore, to investigate possible strategies with which PRL2022 may interact with the host, thereby promoting its fitness and survival in the human intestinal environment, *B. longum* subsp. *longum* PRL2022 was seeded on a Caco-2/HT29-MTX cell monolayer for 4 h. Subsequently, changes in gene expression between PRL2022 cells in contact with the human cell line and those not exposed to this human cell monolayer, i.e., bacterial control cells, were evaluated through RNA sequencing, as previously described (Fontana et al., 2022), generating a total of 7,653,473 quality-filtered reads with an average of 1,275,579 reads per sample. In this context, only genes showing a fold-change of ≥ 2 in combination with a value of $p \leq 0.05$ calculated through correction for multiple comparisons using the False Discovery Rate (FDR) procedure were considered as significantly differentially expressed between the two conditions. Specifically, a total of 449 genes were found to be differentially expressed by comparing the transcriptome of PRL2022 following contact with human cell lines to that of the control (Supplementary Table S7). Interestingly, in depth functional scrutiny of 258 PRL2022 genes whose transcription was up-regulated upon being in contact with human cell lines revealed the presence of a gene encoding a O-antigen ligase (67B_0395), three genes encoding polysaccharide biosynthesis/chain length determinant proteins (67B_0392, 67B_0396, 67B_0398), and two genes coding for glycosyltransferases (GT2 and GT4) (67B_0393 and 67B_0394), all predicted to belong to the genetic locus of PRL2022 genome responsible for exopolysaccharide (EPS) production (Figure 1; Supplementary Table S7). In this context, since EPS have been described as complex glycans exposed on bacterial cell surface directly involved in promoting both bacterial adhesion to the epithelium and microbe-host interactions (Alessandri et al., 2019, 2021), it can be postulated that PRL2022 enhances EPS production as a strategy to adhere to and interact with the human intestinal epithelial cells. In a similar manner, when PRL2022 was placed in contact with the human cell line monolayer, specific transcriptional induction of particular genes occurred, corresponding to proteins implied in the establishment of a microbe-host interaction, i.e., genes encoding a transaldolase (67B_1138) as well as a fibronectin type III domain-containing protein (67B_0918) and a von Willebrand domain-containing protein (67B_1562) (Figure 1;

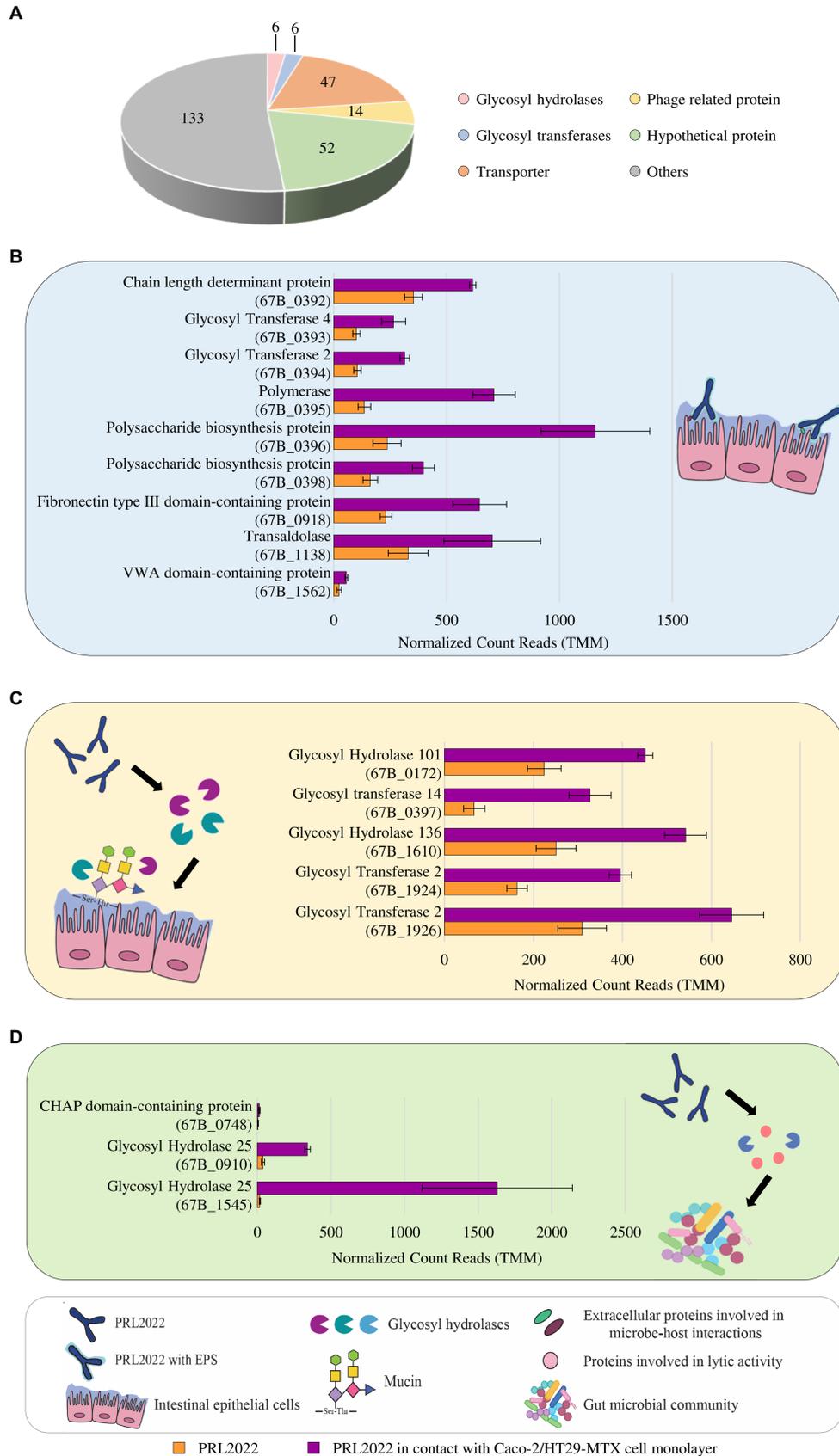


FIGURE 1
 Statistically significant gene expression modulation of *B. longum* subsp. *longum* PRL2022 after 4h of contact with Caco-2/HT29-MTX monolayer. Panel (A) summarizes the number of genes per functional category among the transcriptionally up-regulated genes of PRL2022. The image reports transcriptional modulation of genes, expressed as average of the normalized count reads obtained from each independent biological triplicate, involved in adhesion to the intestinal epithelium [panel (B)], mucin degradation and uptake of the derived saccharides [panel (C)], and cell-to-cell communications [panel (D)]. Each bar plot shows the average of the normalized count reads obtained.

Supplementary Table S7). Specifically, although only demonstrated under *in vitro* experimental conditions, bifidobacterial transaldolases have been described to act as extracellular appendages promoting adherence to the host mucosal surface, therefore acting as adhesive moonlighting proteins with high binding affinities to mucin (Nishiyama et al., 2020; Shang et al., 2022). Proteins harbouring a fibronectin type III- or von Willebrand domain have been shown to promote anchoring to the extracellular matrix protein (Heidler et al., 2021; Nezametdinova et al., 2021; van Leeuwen et al., 2021; Webster et al., 2022). These results indicate that PRL2022 can employ different strategies, including strain-specific features, to adhere to the human intestinal epithelium and, therefore, persist in the human intestinal environment.

Furthermore, beyond genes predicted to be involved in host adhesion, PRL2022 cells exposed to human cell lines showed several upregulated genes predicted to be responsible for the degradation of host-derived glycan secreted by the HT29-MTX cells, i.e., mucin, and for the uptake of the resulting oligosaccharides (Figure 1). Specifically, the transcriptome of PRL2022 exposed to Caco-2/HT29-MTX cell monolayer showed a statistically significant transcriptional upregulation of genes encoding GH101, acting as an endo- α -N-acetylgalactosaminidase, GH136 specifically identified as an extracellular lacto-N-biosidase but also capable of degrading mucin glycan core structure, as previously described (Yamada et al., 2017), as well as GT2, and GT14 both involved in transfer of mucin-related saccharide building blocks, i.e., N-acetylgalactosamine and N-acetylglucosamine (Figure 1). In this context, since mucin is a host-derived carbon source accessible only to a limited number of bacterial strains colonizing the human gut (Tailford et al., 2015), these findings suggest that, when exposed to a human intestine-simulating environment, PRL2022 activates specific carbohydrate degradation strategies to ensure its survival and persistence in the extremely competitive intestinal environment.

Finally, exposure to human cell lines caused a statistically significant higher transcription of PRL2022 genes involved in bacterial cell-to-cell communication, including two genes (67B_0910 and 67B_1545) predicted to encode lysozyme (GH25) and an above-identified unique gene (67B_0748) whose translation produces a cysteine/histidine-dependent aminohydrolases/peptidases (CHAP) domain-containing protein (Figure 1). Interestingly, similar proteins have been shown to exhibit lytic activity toward peptidoglycans (Guglielmetti et al., 2014), suggesting that PRL2022 not only induces mucin degradation genes as a strategy to survive in the competitive intestinal environment, but also produces enzymes potentially able to actively limit growth of other microorganisms.

Altogether, these observations not only suggest the ability of PRL2022 to interact with the host to ensure its survival and persistence in the intestinal microbial ecosystem, but also the multifactorial nature of this process involving microbial surface components and specific carbohydrate-degrading enzymes.

Assessment of host-activated genes in response to PRL2022 cell exposure

To investigate if and how the contact between PRL2022 and the Caco-2/HT29-MTX cell monolayer causes modulation of gene expression in the eukaryotic cells, RNA extracted from the human

cell line following exposure to PRL2022 was subjected to RNA sequencing, generating a total of 133,961,309 reads with an average of 22,326,885 reads per sample. As described above, only genes showing a fold-change of ≥ 2 and a *value of p* ≤ 0.05 calculated through the FDR correction for multiple comparisons procedure were considered as significantly differentially expressed in the Caco-2/HT29-MTX cells in the presence and absence (control samples) of the selected bifidobacterial strain. Interestingly, 875 out of the 1,254 genes differentially expressed in the two conditions were shown to be upregulated in the eukaryotic cells exposed to PRL2022 when compared to the control (Supplementary Table S8). Among these, a gene responsible for the production of a glucosaminyl (N-acetyl) transferase that catalyzes the formation of core 2 and 4 O-glycans on mucin-type glycoprotein, together with three genes, i.e., MUC3A, MUC17, and MUC5B, encoding two mucin-related epithelial glycoproteins and for one of the major gel-forming mucin protein, respectively, were identified (Figure 2). These findings suggest that PRL2022, perhaps because of its ability to utilize mucin, acts as a stimulus for eukaryotic cells to produce/secrete components of the mucous layer covering the human intestinal epithelium, potentially contributing to the reinforcement of the intestinal barrier integrity. Furthermore, various other host genes such as those encoding for the tight junction protein, claudin-1, syndecan, adrenomedullin, amphiregulin, and gasdermin directly or indirectly involved in preserving the integrity/homeostasis of the intestinal epithelial barrier, were significantly more expressed in the human cell monolayer exposed to PRL2022 when compared to the unexposed control (Supplementary Table S8; Bode et al., 2008; Martinez-Herrero and Martinez, 2016; Chen et al., 2018; Frohling et al., 2018; Nishii et al., 2020; Hindson, 2022). These observations reinforce the notion that PRL2022 plays a role in fortifying the intestinal barrier and its ability to stimulate health-promoting effects to the host.

In addition to the genes responsible for maintaining epithelial barrier integrity, various host genes encoding extracellular matrix proteins as well as proteins associated to cilia formation showed a significantly higher transcription level in Caco-2/HT29-MTX cells following PRL2022 exposure (Figure 2). Remarkably, these proteins are known to mediate cell-to-cell communication/adhesion favoring the establishment of an intimate dialog between prokaryotic and eukaryotic cells (Milani et al., 2017; Jeffery, 2019; Vaca et al., 2020), indicating that the presence of PRL2022 induces the production of host proteins involved in host-microbe interaction/adhesion ensuring its persistence in a human intestine-simulating environment. At the same time, a plethora of host genes encoding pattern recognition receptors and cell signaling molecules both involved in eliciting an immune response by the host were shown to be significantly upregulated in eukaryotic cells when placed in contact with PRL2022 cells (Supplementary Table S8). Interestingly, while several host genes associated with pro-inflammatory protein/molecule production were induced in the cell monolayer as a natural result of exposure to bifidobacterial cells, genes coding for anti-inflammatory molecules were also over-expressed. This suggests that PRL2022, as observed for several other bifidobacterial strains, can engage in cross-communication with the host immune system, yet without causing a detrimental inflammatory cascade response, but rather alerting the immune system to promptly react to the possible presence of pathogens (Fanning et al., 2012; Turrone et al., 2013; Longhi et al., 2020).

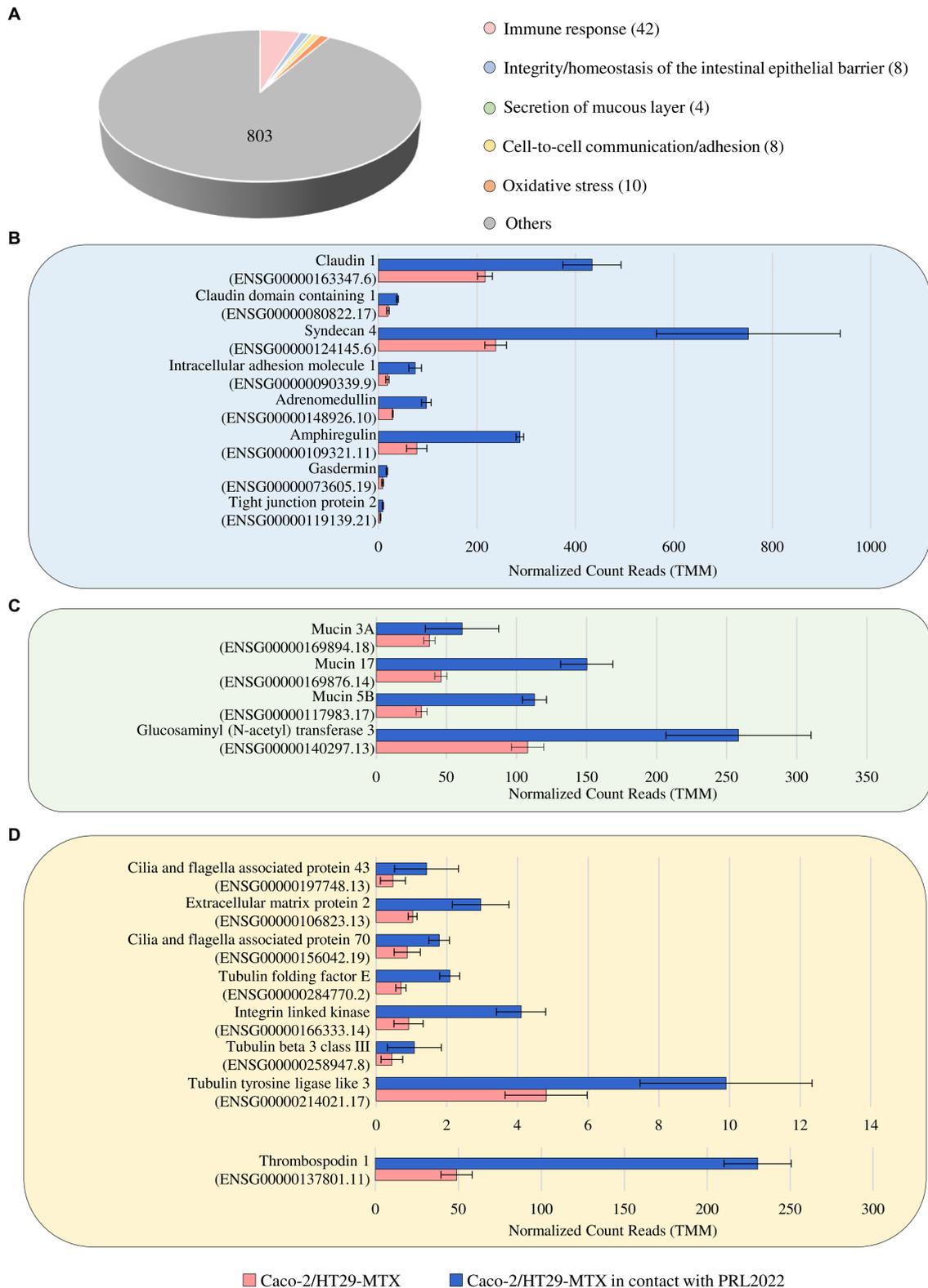


FIGURE 2

Statistically significant up-regulated genes of Caco-2/HT29-MTX cell monolayer after 4h of contact with PRL2022. Panel (A) reports the number of genes per functional category among the up-regulated genes of the human cell lines when in contact with PRL2022. Panels (B–D) display, respectively, transcriptional modulation of those genes involved in integrity/homeostasis of the intestinal epithelial barrier, secretion of mucous layer, and cell-to-cell communication/adhesion. Transcriptional modulation of genes was reported as average of the normalized count reads obtained from each independent biological triplicate.

Overall, these observations suggest that PRL2022 can induce not only the production of extracellular proteins by human intestinal

epithelial cells to favor host cell adhesion and, therefore, its persistence in the intestinal environment, but also the overexpression of genes

involved in guaranteeing intestinal barrier integrity and priming the immune system.

Disentangling the molecular interaction of PRL2022 cells with other human gut microorganisms

To assess possible cross-talk of PRL2022 cells with other bacterial players of the human intestinal microbiota, co-cultivation assays of this strain with two synthetic microbial communities previously stabilized through a bioreactor-based fermentation system were performed to generate an environment that mimics the human intestine (Alessandri et al., 2022). Specifically, the two synthetic microbial communities were selected to allow co-cultivation of PRL2022 in the presence of a high abundance of species of the genus *Bacteroides* (C1) or *Prevotella* (C2) to mimic an intestinal environment typical of the two most widespread enterotypes of the adult human gut microbiota (Arumugam et al., 2011; Costea et al., 2018). To identify PRL2022 genes whose transcription is modulated following 8 h of co-cultivation, RNA extracted from co-cultures as well as from PRL2022 cells grown under the same conditions yet not exposed to other microbial players was sequenced. RNA sequencing generated a total of 31,369,940 quality-filtered reads with an average of 3,485,549 reads per sample. Specifically, only genes with a fold-change of ≥ 2 coupled with a value of $p \leq 0.05$ based on FDR correction were considered as differentially expressed between PRL2022 cells in the presence or absence of the C1 or C2 microbial ecosystem. In detail, of the 1,035 and 1,117 genes that were determined to be significantly differentially transcribed, a total of 511 and 543 genes were shown to be significantly up-regulated in PRL2022 cells when co-cultivated with either C1 or C2 microbial communities, respectively (Figure 3; Supplementary Table S9). Interestingly, most of the PRL2022 genes showing increased transcription when co-cultivated with the C1 microbial community (compared to the PRL2022 culture without co-cultivation) were shown to correspond to PRL2022 genes that also exhibited increased transcription in the presence of the C2 microbial ecosystem, suggesting that PRL2022 adopts similar response strategies to human intestinal microbiota exposure regardless of the taxonomic composition of the latter. Particularly, PRL2022 cells were shown to increase transcription of a wide variety of genes involved in the metabolism and transport of amino acids/proteins and carbohydrates when co-cultivated with other intestinal microbes (Figure 3; Supplementary Table S9). Probably, the presence of other microorganisms may not only have stimulated PRL2022 to activate multiple strategies to compete for nutrients, but it may also have made accessible, through cross-feeding or resource-sharing activities, specific nutrients that PRL2022 alone would not be able to use (Bunesova et al., 2018; Turrioni et al., 2018b). Directly linked to the higher abundance of over-expressed genes involved in carbohydrate degradation, the interaction with other microorganisms combined with the exposure to an intestinal-simulating environment, seem to favor also the increased transcription of PRL2022 genes involved in acetate production, i.e., glucose-6-phosphate isomerase, pyruvate kinase, acetate kinase, phosphoketolase, transketolase, and transaldolase (Figure 3; Supplementary Table S9; Lanigan et al., 2019). These observations emphasize the important interactomic features of this *B. longum* subsp. *longum* strain with the human host microbiota.

Indeed, since acetate can be consumed by certain members of the gut microbiota to generate butyrate, a short chain fatty acid known for its beneficial activity to the host promoting intestinal barrier integrity and anti-inflammatory response, the enhancement of the expression of genes involved in the production of acetate by PRL2022 may have an *in vivo* butyrogenic effect and, consequently, exerting health-promoting effects upon the host (Riviere et al., 2016; Alessandri et al., 2019).

Furthermore, and as also observed when PRL2022 interacts with a Caco-2/HT29-MTX cell line monolayer, when exposed to a human intestinal-simulating environment with the presence of a microbial community, the selected strain was shown significantly increase transcription of eight genes of the PRL2022 *eps* locus, specifying four GTs, the polymerase, and three genes coding for polysaccharide biosynthesis protein/chain length determinant protein (Supplementary Table S9; Figure 3). This result indicates that (increased) EPS production is the elective strategy adopted by PRL2022 to interact with both the host and other microbial players ensuring its persistence in the human gut.

Finally, although PRL2022 displayed a similar behavior in response to the interaction with two different microbial communities, only the exposure to C2 induced in the selected bifidobacterial strain a significantly higher number of transcripts of genes predicted to encode S-ribosylhomocysteine lyase (LuxS), i.e., the enzyme responsible for production of autoinducer-2 (AI-2) (67B_0488), and autoinducer-2 ABC transporter (67B_1938) (Supplementary Table S9; Figure 3). These two proteins are involved in production and transport of AI-2, allowing an intricate cell-to-cell communication system, known as quorum sensing, which is able to regulate diverse phenomena, including virulence, adherence to the host cells or biofilm formation (Sun et al., 2014). This suggests that PRL2022 activates this specific interspecies communication system particularly in presence of certain bacterial species.

Overall, these observations highlight the ability of PRL2022 to activate multiple strategies to survive and persist in the competitive gut environment and to interact with other intestinal microbial players.

Conclusion

B. longum subsp. *longum* strains are highly prevalent in the human gut and several strains of this species have been associated with exerting health-promoting effects upon their host. In this context, to identify a *B. longum* subsp. *longum* strain that could act as a valuable model for high fitness in the human intestine, an ecological and phylogenomic approach was applied. This approach allowed the identification of *B. longum* subsp. *longum* PRL2022 as a strain that is, from a genetic and functional point of view, a very close relative of the most representative strains of this bifidobacterial species in the human intestinal microbiota. In-depth insights into PRL2022 genomic determinants coupled with a comparative genome analysis involving PRL2022 and genomes of those *B. longum* subsp. *longum* strains earmarked as model microorganisms of this species and included in commercially available probiotic products revealed the presence of a plethora of genes, some of which identified as PRL2022 unique genes, involved in host- and microbe-microbe interactions. This indicates that PRL2022 possesses a genetic makeup that could favor the survival, colonization, and persistence of this strain in the competitive human intestinal ecological niche. A notion that was corroborated through

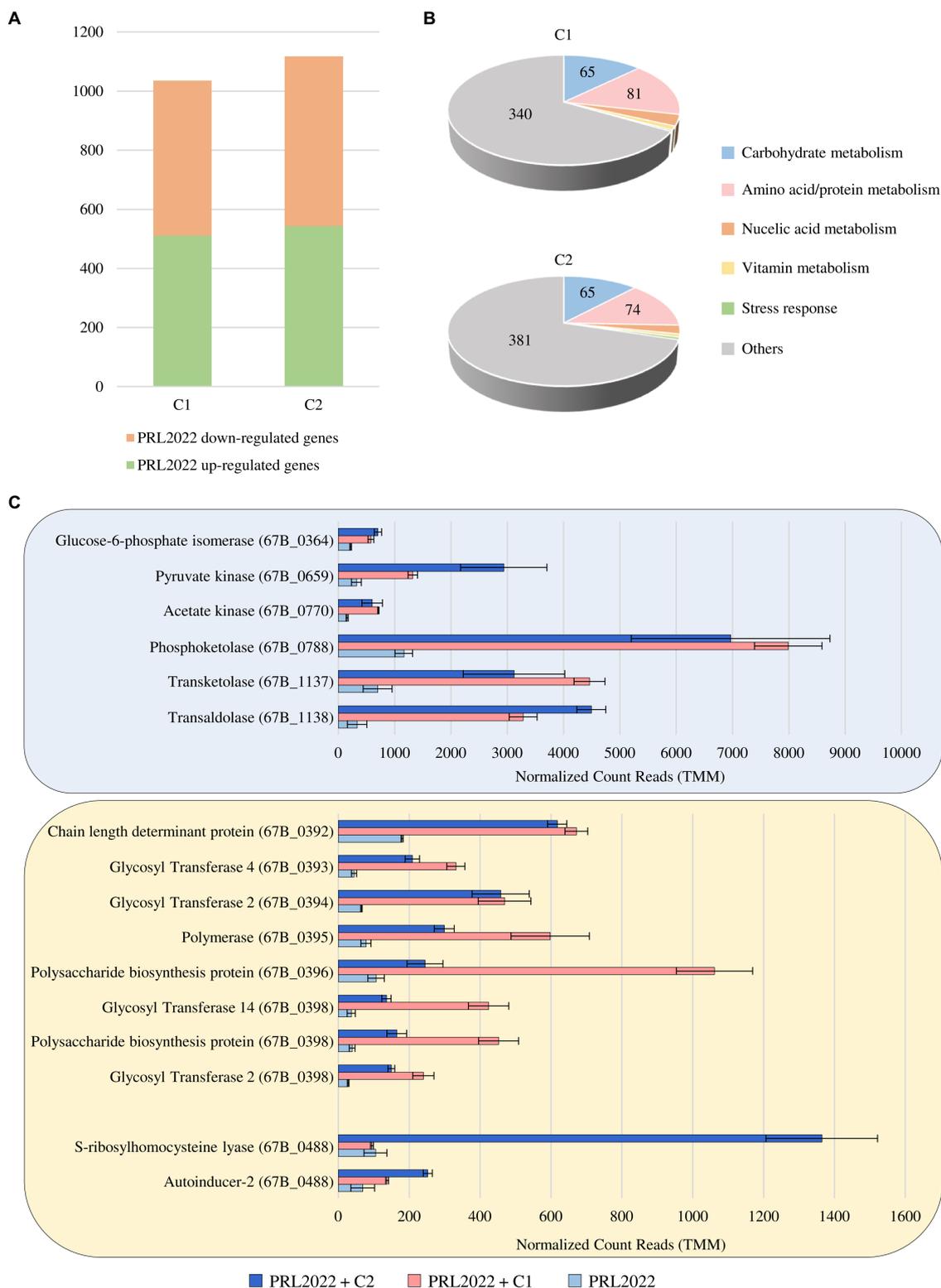


FIGURE 3

Transcriptional modulation of PRL2022 genes when exposed to either of two stabilized intestinal bacterial communities C1 and C2. Panel (A) reports the number of up-regulated and down-regulated genes in PRL2022 when exposed to C1 and C2 when compared to the control. Panel (B) displays the number of genes per functional category among the up-regulated genes of PRL2022 in the presence of C1 or C2. Panel (C) shows the transcriptional modulation of genes predicted to be involved in acetate and EPS production as well as in quorum sensing-based cell-to-cell communication. Transcriptional modulation of genes was reported as average of the normalized count reads obtained from each independent biological triplicate.

functional genomic experiments using *in vitro* models mimicking the human intestinal environment and microbial biodiversity, highlighting the

ability of PRL2022 to establish extensive cross-talk with both the host and other intestinal microbial players. Therefore, the proposed *in silico* method

based on ecological- and phylogenomic-driven selection of a potential *B. longum* subsp. *longum* model strain candidate with a presumed high fitness in the human gut should be considered as a valuable tool to be applied to other (bifido) bacterial species to identify model strains among the autochthonous strains of the human gut.

Data availability statement

The data presented in the study are deposited in the NCBI SRA repository, accession number PRJNA914637 and PRJNA833139. Furthermore, the updated genome sequence of *B. longum* subsp. *longum* PRL2022 is available under the accession number PRJNA692178.

Author contributions

GA wrote the manuscript and performed *in vitro* analyses. SR, MB, GT, MC, and CA performed *in vitro* analysis. GA, FF, and CT performed bioinformatics analyses. GAL and LM validated the bioinformatic analyses. GL, RA, and AV performed the sequencing. CM, FT, OB, and DS supervised the project and edited the manuscript. MV supervised the project and designed the study. All authors contributed to the article and approved the submitted version.

Funding

The PhD fellowship of CA is financially supported by Fondazione Cariparma, Parma, Italy.

References

- Alessandri, G., Fontana, F., Mancabelli, L., Lugli, G. A., Tarracchini, C., Argentini, C., et al. (2022). Exploring species-level infant gut bacterial biodiversity by meta-analysis and formulation of an optimized cultivation medium. *NPJ Biofilms Microbiomes* 8:88. doi: 10.1038/s41522-022-00349-1
- Alessandri, G., Ossiprandi, M. C., Macsharry, J., Van Sinderen, D., and Ventura, M. (2019). Bifidobacterial dialogue with its human host and consequent modulation of the immune system. *Front. Immunol.* 10:2348. doi: 10.3389/fimmu.2019.02348
- Alessandri, G., Van Sinderen, D., and Ventura, M. (2021). The genus bifidobacterium: from genomics to functionality of an important component of the mammalian gut microbiota running title: Bifidobacterial adaptation to and interaction with the host. *Comput. Struct. Biotechnol. J.* 19, 1472–1487. doi: 10.1016/j.csbj.2021.03.006
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638
- Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180. doi: 10.1038/nature09944
- Bianchi, M. G., Chiu, M., Taurino, G., Brighenti, F., Del Rio, D., Mena, P., et al. (2019). Catechin and Procyanidin B(2) modulate the expression of tight junction proteins but do not protect from inflammation-induced changes in permeability in human intestinal cell monolayers. *Nutrients* 11:2271. doi: 10.3390/nu11102271
- Blanco, G., Ruiz, L., Tames, H., Ruas-Madiedo, P., Fdez-Riverola, F., et al. (2020). Revisiting the metabolic capabilities of *Bifidobacterium longum* subsp. *Longum* and *Bifidobacterium longum* subsp. *infantis* from a glycoside hydrolase perspective. *Microorganisms* 8:723. doi: 10.3390/microorganisms8050723
- Bode, L., Salvestrini, C., Park, P. W., Li, J. P., Esko, J. D., Yamaguchi, Y., et al. (2008). Heparan sulfate and syndecan-1 are essential in maintaining murine and human intestinal epithelial barrier function 118, 229–238. doi: 10.1172/JCI32335
- Bunesova, V., Lacroix, C., and Schwab, C. (2018). Mucin cross-feeding of infant *Bifidobacteria* and *Eubacterium hallii*. *Microb. Ecol.* 75, 228–238. doi: 10.1007/s00248-017-1s037-4
- Chen, F., Yang, W., Huang, X., Cao, A. T., Bilotta, A. J., Xiao, Y., et al. (2018). Neutrophils promote Amphiregulin production in intestinal epithelial cells through TGF-beta and Contribute to intestinal homeostasis. *J. Immunol.* 201, 2492–2501. doi: 10.4049/jimmunol.1800003
- Choi, I. Y., Kim, J., Kim, S. H., Ban, O. H., Yang, J., and Park, M. K. (2021). Safety evaluation of *Bifidobacterium breve* IDCC4401 isolated from infant feces for use as a commercial probiotic. *J. Microbiol. Biotechnol.* 31, 949–955. doi: 10.4014/jmb.2103.03041
- Costea, P. I., Hildebrand, F., Arumugam, M., Bäckhed, F., Blaser, M. J., Bushman, F. D., et al. (2018). Enterotypes in the landscape of gut microbial community composition. *Nat. Microbiol.* 3, 8–16. doi: 10.1038/s41564-017-0072-8
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635
- EFSA (2012). Guidance on the assessment of bacterial susceptibility to antimicrobials of human and veterinary importance. *EFSA J.* 10:2740. doi: 10.2903/j.efsa.2012.2740
- Enright, A. J., Van Dongen, S., and Ouzounis, C. A. (2002). An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* 30, 1575–1584. doi: 10.1093/nar/30.7.1575
- Fanning, S., Hall, L. J., and Van Sinderen, D. (2012). *Bifidobacterium breve* UCC2003 surface exopolysaccharide production is a beneficial trait mediating commensal-host interaction through immune modulation and pathogen protection. *Gut Microbes* 3, 420–425. doi: 10.4161/gmic.20630
- Fontana, F., Alessandri, G., Tarracchini, C., Bianchi, M. G., Rizzo, S. M., Mancabelli, L., et al. (2022). Designation of optimal reference strains representing the infant gut bifidobacterial species through a comprehensive multi-omics approach. *Environ. Microbiol.* 24, 5825–5839. doi: 10.1111/1462-2920.16205
- Frohling, M., Tepasse, P., Intemann, J., Sambale, M., Sherwood, J., Paruzel, P., et al. (2018). Syndecan-4 modulates epithelial gut barrier function and epithelial regeneration in experimental colitis. *Inflamm. Bowel Dis.* 24, 2579–2589. doi: 10.1093/ibd/izy248
- Fujita, K., Sakamoto, A., Kaneko, S., Kotake, T., Tsumuraya, Y., and Kitahara, K. (2019). Degradative enzymes for type II arabinogalactan side chains in *Bifidobacterium longum* subsp. *longum*. *Appl. Microbiol. Biotechnol.* 103, 1299–1310. doi: 10.1007/s00253-018-9566-4

Acknowledgments

We thank GenProbio srl for financial support of the Laboratory of Probiogenomics. Part of this research has been conducted using the High Performance Computing (HPC) facility of the University of Parma.

Conflict of interest

FF, GL, RA, and AV were employed by GenProbio Srl.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationship that could be construed as potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2023.1130592/full#supplementary-material>

- Guglielmetti, S., Balzaretto, S., Taverniti, V., Miriani, M., Milani, C., Scarafoni, A., et al. (2014). TgaA, a VirB1-like component belonging to a putative type IV secretion system of *Bifidobacterium bifidum* MIMBb75. *Appl. Environ. Microbiol.* 80, 5161–5169. doi: 10.1128/AEM.01413-14
- Hao, J., Zhang, Y., Wu, T., Liu, R., Sui, W., Zhu, J., et al. (2022). The antidiabetic effects of *Bifidobacterium longum* subsp. *longum* BL21 through regulating gut microbiota structure in type 2 diabetic mice. *Food Funct.* 13, 9947–9958. doi: 10.1039/D2FO01109C
- Heidler, T. V., Ernits, K., Ziolkowska, A., Claesson, R., and Persson, K. (2021). *Porphyromonas gingivalis* fimbrial protein Mfa5 contains a von Willebrand factor domain and an intramolecular isopeptide. *Commun. Biol.* 4:106. doi: 10.1038/s42003-020-01621-w
- Hidalgo-Cantabrana, C., Delgado, S., Ruiz, L., Ruas-Madiedo, P., Sanchez, B., and Margolles, A. (2017). *Bifidobacterium* and their health-promoting effects. *Microbiol. Spectr.* 5. doi: 10.1128/microbiolspec.BAD-0010-2016
- Hindson, J. (2022). Gasdermin B in IBD and epithelial barrier repair. *Nat. Rev. Gastroenterol. Hepatol.* 19:216. doi: 10.1038/s41575-022-00589-8
- Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., et al. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44, D286–D293. doi: 10.1093/nar/gkv1248
- Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* 11:119. doi: 10.1186/1471-2105-11-119
- Jeffery, C. J. (2019). Intracellular/surface moonlighting proteins that aid in the attachment of gut microbiota to the host. *AIMS Microbiol.* 5, 77–86. doi: 10.3934/microbiol.2019.1.77
- Kim, Y. T., Kim, C. H., Kwon, J. G., Cho, J. H., Shin, Y. S., Kim, H. B., et al. (2022). *In vivo* trial of *Bifidobacterium longum* revealed the complex network correlations between gut microbiota and health promotional effects. *Front. Microbiol.* 13:886934. doi: 10.3389/fmicb.2022.886934
- Kim, M. J., Ku, S., Kim, S. Y., Lee, H. H., Jin, H., Kang, S., et al. (2018). Safety evaluations of *Bifidobacterium bifidum* BGN4 and *Bifidobacterium longum* BORI. *Int. J. Mol. Sci.* 19:1422. doi: 10.3390/ijms19051422
- Klimina, K. M., Kasianov, A. S., Poluektova, E. U., Emelyanov, K. V., Voroshilova, V. N., Zakharevich, N. V., et al. (2019). Employing toxin-antitoxin genome markers for identification of *Bifidobacterium* and *Lactobacillus* strains in human metagenomes. *PeerJ* 7:e6554. doi: 10.7717/peerj.6554
- Klimina, K. M., Voroshilova, V. N., Poluektova, E. U., Veselovsky, V. A., Yunes, R. A., Kovtun, A. S., et al. (2020). Toxin-antitoxin systems: a tool for taxonomic analysis of human intestinal microbiota. *Toxins (Basel)* 12:388. doi: 10.3390/toxins12060388
- Komeno, M., Yoshihara, Y., Kawasaki, J., Nabeshima, W., Maeda, K., Sasaki, Y., et al. (2022). Two alpha-L-arabinofuranosidases from *Bifidobacterium longum* subsp. *longum* are involved in arabinoxylan utilization. *Appl. Microbiol. Biotechnol.* 106, 1957–1965. doi: 10.1007/s00253-022-11845-x
- Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116
- Langdon, W. B. (2015). Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. *BioData Min.* 8:1. doi: 10.1186/s13040-014-0034-0
- Lanigan, N., Kelly, E., Arzamasov, A. A., Stanton, C., Rodionov, D. A., and Van Sinderen, D. (2019). Transcriptional control of central carbon metabolic flux in *Bifidobacteria* by two functionally similar, yet distinct LacI-type regulators. *Sci. Rep.* 9:17851. doi: 10.1038/s41598-019-54229-4
- Leinonen, R., Sugawara, H., and Shumway, M. International Nucleotide Sequence Database, C (2011). The sequence read archive. *Nucleic Acids Res.* 39, D19–D21. doi: 10.1093/nar/gkq1019
- Li, T. T., Zhang, H. X., and Gu, C. T. (2022). *Bifidobacterium mizhiense* sp. nov., isolated from the gut of honeybee (*Apis mellifera*). *Int. J. Syst. Evol. Microbiol.* 72. doi: 10.1099/ijsem.0.005390
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178
- Longhi, G., Van Sinderen, D., Ventura, M., and Turrone, F. (2020). Microbiota and cancer: the emerging beneficial role of *Bifidobacteria* in cancer immunotherapy. *Front. Microbiol.* 11:575072. doi: 10.3389/fmicb.2020.575072
- Lugli, G. A., Calvete-Torre, I., Alessandri, G., Milani, C., Turrone, F., Laiolo, P., et al. (2021). Phylogenetic classification of ten novel species belonging to the genus *Bifidobacterium* comprising *B. phasiani* sp. nov., *B. pongonis* sp. nov., *B. saguiniicoloris* sp. nov., *B. colobi* sp. nov., *B. simiiventris* sp. nov., *B. santillanense* sp. nov., *B. miconis* sp. nov., *B. amazonense* sp. nov., *B. pluvialisilvae* sp. nov., and *B. miconisargentati* sp. nov. *Syst. Appl. Microbiol.* 44:126273. doi: 10.1016/j.syapm.2021.126273
- Lugli, G. A., Milani, C., Mancabelli, L., Van Sinderen, D., and Ventura, M. (2016). MEGAnnotator: a user-friendly pipeline for microbial genomes assembly and annotation. *FEMS Microbiol. Lett.* 363. doi: 10.1093/femsle/fnw049
- Mancabelli, L., Mancino, W., Lugli, G. A., Argentini, C., Longhi, G., Milani, C., et al. (2021). Amoxicillin-Clavulanic acid resistance in the genus *Bifidobacterium*. *Appl. Environ. Microbiol.* 87:e03137-20. doi: 10.1128/AEM.03137-20
- Martinez-Herrero, S., and Martinez, A. (2016). Adrenomedullin regulates intestinal physiology and pathophysiology. *Domest. Anim. Endocrinol.* 56, S66–S83. doi: 10.1016/j.domaniend.2016.02.004
- Masco, L., Van Hoorde, K., De Brandt, E., Swings, J., and Huys, G. (2006). Antimicrobial susceptibility of *Bifidobacterium* strains from humans, animals and probiotic products. *J. Antimicrob. Chemother.* 58, 85–94. doi: 10.1093/jac/dkl197
- Mccarville, J. L., Dong, J., Caminero, A., Bermudez-Brito, M., Jury, J., Murray, J. A., et al. (2017). A commensal *Bifidobacterium longum* strain prevents gluten-related immunopathology in mice through expression of a serine protease inhibitor. *Appl. Environ. Microbiol.* 83:17. doi: 10.1128/AEM.01323-17
- Milani, C., Alessandri, G., Mancabelli, L., Mangifesta, M., Lugli, G. A., Viappiani, A., et al. (2020). Multi-omics approaches to decipher the impact of diet and host physiology on the mammalian gut microbiome. *Appl. Environ. Microbiol.* 86:e01864-20. doi: 10.1128/AEM.01864-20
- Milani, C., Lugli, G. A., Fontana, F., Mancabelli, L., Alessandri, G., Longhi, G., et al. (2021). METAnnotatorX2: a comprehensive tool for deep and shallow metagenomic data set analyses. *mSystems* 6:e0058321. doi: 10.1128/mSystems.00583-21
- Milani, C., Mangifesta, M., Mancabelli, L., Lugli, G. A., Mancino, W., and Viappiani, A. (2017). The Sortase-dependent Fimbriome of the genus *Bifidobacterium*: extracellular structures with potential to modulate microbe-host dialogue. *Appl. Environ. Microbiol.* 83:e01295-17. doi: 10.1128/AEM.01295-17
- Nezametinova, V. Z., Yunes, R. A., Dukhinova, M. S., Alekseeva, M. G., and Danilenko, V. N. (2021). The role of the PFNA operon of *Bifidobacterium* in the recognition of Host's immune signals: prospects for the use of the FN3 protein in the treatment of COVID-19. *Int. J. Mol. Sci.* 22:9219. doi: 10.3390/ijms22179219
- Nishii, N., Oshima, T., Li, M., Eda, H., Nakamura, K., Tamura, A., et al. (2020). Lubiprostone induces Claudin-1 and protects intestinal barrier function. *Pharmacology* 105, 102–108. doi: 10.1159/000503054
- Nishiyama, K., Takaki, T., Sugiyama, M., Fukuda, I., Aiso, M., Mukai, T., et al. (2020). Extracellular vesicles produced by *Bifidobacterium longum* export Mucin-binding proteins. *Appl. Environ. Microbiol.* 86:e01464-20. doi: 10.1128/AEM.01464-20
- Quintanilha, M. F., Miranda, V. C., Souza, R. O., Gallotti, B., Cruz, C., Santos, E. A., et al. (2022). *Bifidobacterium longum* subsp. *longum* 5(1A) attenuates intestinal injury against irinotecan-induced mucositis in mice. *Life Sci.* 289:120243. doi: 10.1016/j.lfs.2021.120243
- Riviere, A., Selak, M., Lantin, D., Leroy, F., and De Vuyst, L. (2016). *Bifidobacteria* and butyrate-producing colon bacteria: importance and strategies for their stimulation in the human gut. *Front. Microbiol.* 7:979. doi: 10.3389/fmicb.2016.00979
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Santos, S. S., Miranda, V. C., Trindade, L. M., Reis, D. C., Cassali, G. D., Nicoli, J. R., et al. (2021). *Bifidobacterium longum* subsp. *longum* 5(1A) attenuates signs of inflammation in a murine model of food allergy. *Probiotics Antimicrob. Proteins.* doi: 10.1007/s12602-021-09846-9
- Sasaki, Y., Horigome, A., Odamaki, T., Xiao, J. Z., Ishiwata, A., Ito, Y., et al. (2021). Novel 3-O-alpha-d-Galactosyl-alpha-L-Arabinofuranosidase for the assimilation of gum Arabic Arabinogalactan protein in *Bifidobacterium longum* subsp. *longum*. *Appl. Environ. Microbiol.* 87:e02690-20. doi: 10.1128/AEM.02690-20
- Sasaki, Y., Komeno, M., Ishiwata, A., Horigome, A., Odamaki, T., Xiao, J. Z., et al. (2022). Mechanism of cooperative degradation of gum Arabic Arabinogalactan protein by *Bifidobacterium longum* surface enzymes. *Appl. Environ. Microbiol.* 88:e0218721. doi: 10.1128/aem.02187-21
- Schiavi, E., Gleinser, M., Molloy, E., Groeger, D., Frei, R., Ferstl, R., et al. (2016). The surface-associated exopolysaccharide of *Bifidobacterium longum* 35624 plays an essential role in dampening host Proinflammatory responses and repressing local TH17 responses. *Appl. Environ. Microbiol.* 82, 7185–7196. doi: 10.1128/AEM.02238-16
- Schiavi, E., Plattner, S., Rodriguez-Perez, N., Barcik, W., Frei, R., Ferstl, R., et al. (2018). Exopolysaccharide from *Bifidobacterium longum* subsp. *longum* 35624 modulates murine allergic airway responses. *Benef. Microb.* 9, 761–773. doi: 10.3920/BM2017.0180
- Serafini, F., Strati, F., Ruas-Madiedo, P., Turrone, F., Foroni, E., Duranti, S., et al. (2013). Evaluation of adhesion properties and antibacterial activities of the infant gut commensal *Bifidobacterium bifidum* PRL2010. *Anaerobe* 21, 9–17. doi: 10.1016/j.anaerobe.2013.03.003
- Shang, J., Yang, S., Tang, Z., Chen, Y., Duan, B., and Meng, X. (2022). *Bifidobacterium bifidum* H3-R2 and its molecular communication within the context of ulcerative colitis. *J. Agric. Food Chem.* 70, 11678–11688. doi: 10.1021/acs.jafc.2c02909
- Sun, Z., He, X., Brancaccio, V. F., Yuan, J., and Riedel, C. U. (2014). *Bifidobacteria* exhibit LuxS-dependent autoinducer 2 activity and biofilm formation. *PLoS One* 9:e88260. doi: 10.1371/journal.pone.0088260
- Tailford, L. E., Crost, E. H., Kavanaugh, D., and Juge, N. (2015). Mucin glycan foraging in the human gut microbiome. *Front. Genet.* 6:81. doi: 10.3389/fgene.2015.00081
- Tanno, H., Fujii, T., Ose, R., Hirano, K., Tochio, T., and Endo, A. (2019). Characterization of fructooligosaccharide-degrading enzymes in human commensal *Bifidobacterium longum* and *Anaerostipes caccae*. *Biochem. Biophys. Res. Commun.* 518, 294–298. doi: 10.1016/j.bbrc.2019.08.049
- Tarracchini, C., Viglioli, M., Lugli, G. A., Mancabelli, L., Fontana, F., Alessandri, G., et al. (2022). The integrated probiotic database: a genomic compendium of *Bifidobacterium* health-promoting strains. *Microb. Res. Rep.* doi: 10.20517/mrr.2021.13

- Turroni, F., Milani, C., Duranti, S., Ferrario, C., Lugli, G. A., Mancabelli, L., et al. (2018a). Bifidobacteria and the infant gut: an example of co-evolution and natural selection. *Cell. Mol. Life Sci.* 75, 103–118. doi: 10.1007/s00018-017-2672-0
- Turroni, F., Milani, C., Duranti, S., Mahony, J., Van Sinderen, D., and Ventura, M. (2018b). Glycan utilization and cross-feeding activities by Bifidobacteria. *Trends Microbiol.* 26, 339–350. doi: 10.1016/j.tim.2017.10.001
- Turroni, F., Milani, C., Duranti, S., Mancabelli, L., Mangifesta, M., Viappiani, A., et al. (2016). Deciphering bifidobacterial-mediated metabolic interactions and their impact on gut microbiota by a multi-omics approach. *ISME J.* 10, 1656–1668. doi: 10.1038/ismej.2015.236
- Turroni, F., Serafini, F., Foroni, E., Duranti, S., O'connell Motherway, M., Taverniti, V., et al. (2013). Role of sortase-dependent pili of *Bifidobacterium bifidum* PRL2010 in modulating bacterium-host interactions. *Proc. Natl. Acad. Sci. U. S. A.* 110, 11151–11156. doi: 10.1073/pnas.1303897110
- Turroni, F., Taverniti, V., Ruas-Madiedo, P., Duranti, S., Guglielmetti, S., Lugli, G. A., et al. (2014). *Bifidobacterium bifidum* PRL2010 modulates the host innate immune response. *Appl. Environ. Microbiol.* 80, 730–740. doi: 10.1128/AEM.03313-13
- Vaca, D. J., Thibau, A., Schutz, M., Kraiczky, P., Happonen, L., Malmstrom, J., et al. (2020). Interaction with the host: the role of fibronectin and extracellular matrix proteins in the adhesion of gram-negative bacteria. *Med. Microbiol. Immunol.* 209, 277–299. doi: 10.1007/s00430-019-00644-3
- Van Dijk, L. R., Walker, B. J., Straub, T. J., Worby, C. J., Grote, A., Henry, L. S., et al. (2022). StrainGE: a toolkit to track and characterize low-abundance strains in complex microbial communities. *Genome Biol.* 23:74. doi: 10.1186/s13059-022-02630-0
- Van Leeuwen, H. C., Roelofs, D., Corver, J., and Hensbergen, P. (2021). Phylogenetic analysis of the bacterial pro-pro-endopeptidase domain reveals a diverse family including secreted and membrane anchored proteins. *Curr. Res. Microb. Sci.* 2:100024. doi: 10.1016/j.crmicr.2021.100024
- Ventura, M., Oflaherty, S., Claesson, M. J., Turroni, F., Klaenhammer, T. R., Van Sinderen, D., et al. (2009). Genome-scale analyses of health-promoting bacteria: probio-genomics. *Nat. Rev. Microbiol.* 7, 61–71. doi: 10.1038/nrmicro2047
- Ventura, M., Turroni, F., and Van Sinderen, D. (2012). Probiogenomics as a tool to obtain genetic insights into adaptation of probiotic bacteria to the human gut. *Bioeng. Bugs* 3, 73–79. doi: 10.4161/bbug.18540
- Wang, Y., and Lapointe, G. (2020). Arabinogalactan utilization by *Bifidobacterium longum* subsp. *longum* NCC 2705 and *Bacteroides caccae* ATCC 43185 in monoculture and Coculture. *Microorganisms* 8:1703. doi: 10.3390/microorganisms8111703
- Webster, S. S., Wong, G. C. L., and Otoole, G. A. (2022). The power of touch: type 4 Pili, the von Willebrand domain, and surface sensing by *Pseudomonas aeruginosa*. *J. Bacteriol.* 204:e0008422. doi: 10.1128/jb.00084-22
- Wheeler, T. J., and Eddy, S. R. (2013). Nhmmer: DNA homology search with profile HMMs. *Bioinformatics* 29, 2487–2489. doi: 10.1093/bioinformatics/btt403
- Wick, R. R., and Holt, K. E. (2022). Polypolish: short-read polishing of long-read bacterial genome assemblies. *PLoS Comput. Biol.* 18:e1009802. doi: 10.1371/journal.pcbi.1009802
- Yamada, C., Gotoh, A., Sakanaka, M., Hattie, M., Stubbs, K. A., Katayama-Ikegami, A., et al. (2017). Molecular insight into evolution of Symbiosis between breast-fed infants and a member of the human gut microbiome *Bifidobacterium longum*. *Cell Chem. Biol.* 24:e515, 515–524.e5. doi: 10.1016/j.chembiol.2017.03.012
- Yan, S., Yang, B., Zhao, J., Zhao, J., Stanton, C., Ross, R. P., et al. (2019). A ropy exopolysaccharide producing strain *Bifidobacterium longum* subsp. *longum* YS108R alleviates DSS-induced colitis by maintenance of the mucosal barrier and gut microbiota modulation. *Food Funct.* 10, 1595–1608. doi: 10.1039/C9FO00014C
- Yasmin, I., Saeed, M., Khan, W. A., Khaliq, A., Chughtai, M. F. J., Iqbal, R., et al. (2020). In vitro probiotic potential and safety evaluation (hemolytic, cytotoxic activity) of *Bifidobacterium* strains isolated from raw camel Milk. *Microorganisms* 8:354. doi: 10.3390/microorganisms8030354
- Zhao, Y., Tang, H., and Ye, Y. (2012a). RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinformatics* 28, 125–126. doi: 10.1093/bioinformatics/btr595
- Zhao, Y., Wu, J., Yang, J., Sun, S., Xiao, J., and Yu, J. (2012b). PGAP: pan-genomes analysis pipeline. *Bioinformatics* 28, 416–418. doi: 10.1093/bioinformatics/btr655