Applied and Environmental Microbiology, Dec. 2011, p. 8578–8587 Vol. 77, No. 24
0099-2240/11/$12.00 doi:10.1128/AEM.06352-11

Global Genome Transcription Profiling of *Bifidobacterium bifidum* PRL2010 under *In Vitro* Conditions and Identification of Reference Genes for Quantitative Real-Time PCR▽†

Francesca Turroni,[1] Elena Foroni,[1] Barbara Montanini,[2] Alice Viappiani,[1] Francesco Strati,[1] Sabrina Duranti,[1] Alberto Ferrarini,[4] Massimo Delledonne,[4] Douwe van Sinderen,[3] and Marco Ventura[1]*

Laboratory of Probiogenomics, Department of Genetics, Biology of Microorganisms, Anthropology and Evolution, University of Parma, Parma, Italy[1]; Department of Biochemistry and Molecular Biology, University of Parma, Parma, Italy[2]; Alimentary Pharmabiotic Centre and Department of Microbiology, Bioscience Institute, National University of Ireland, Western Road, Cork, Ireland[3]; and Dipartimento di Biotecnologie, Università degli Studi di Verona, Verona, Italy[4]

Received 28 July 2011/Accepted 7 October 2011

Bifidobacteria have attracted significant scientific attention due to their perceived role as health-promoting microorganisms, although the genetics of the bacterial group is still underexplored. In this study, we investigated the transcriptome of *Bifidobacterium bifidum* PRL2010 during *in vitro* growth by microarray technology. When *B. bifidum* PRL2010 was grown in liquid broth, 425 of the 1,644 PRL2010 genes represented on the array were expressed in at least one of the three investigated growth phases, i.e., the lag, exponential, and stationary phases. These transcriptional analyses identified a core *in vitro* transcriptome encompassing 150 genes that are expressed in all phases. A proportion of these genes were further investigated as potential reference genes by quantitative real-time reverse transcription-PCR (qRT-PCR) assays. Their expression stability was evaluated under different growth conditions, which included cultivation on different carbon sources, exposure to environmental stresses (thermal, acidic, and osmotic), and growth phases. Our analyses validated six reference genes suitable for normalizing mRNA expression levels in qRT-PCR experiments applied to bifidobacteria.

Bifidobacteria represent one of the principal members of the gut microbiota of infants (12) and are thought to positively influence the health status of their hosts (9). Thus, in recent years, a very significant body of research has been directed to exploring the biology of this group of microorganisms (for recent reviews, see references 3 and 22). However, the genetic basis of these purported health-promoting effects is poorly, if at all, understood (16). Thirty-eight species of bifidobacteria have been described so far (22), and several bifidobacterial genomes have been sequenced (for reviews, see references 7, 25, 26, and 33). The advent of genomics has allowed the determination of complete genetic blueprints of many bifidobacteria and has provided genetic evidence that underpins the specific adaptation of these bacteria to the human gut (for reviews, see references 7, 10, 19, 25, 26, and 33). In this context, the decoding of the genome sequence of *Bifidobacterium bifidum* PRL2010, a microbial isolate from an infant gut, led to the identification of an arsenal of enzymes necessary for the metabolism of host glycans, such as mucins and *h*uman *m*ilk *o*ligosaccharides (HMO) (19), which opened avenues for further scientific exploration of the microorganism as a representative component of the infant gut microbiota.

Global transcriptional profiling has already been applied to explore changes in the transcriptome of *B. bifidum* PRL2010 cultivated under different growth conditions related to the available carbon source, i.e., mucin oligosaccharides, HMO, and lactose (19). Similarly, transcriptome analyses have been employed to monitor differential transcription based on carbohydrate metabolism in *Bifidobacterium dentium* Bd1 (36), and responses to various stressful conditions in *Bifidobacterium breve* UCC2003, such as heat stress and osmotic shock (39). In contrast, whole-transcriptome analyses have not yet been published to investigate changes in the transcription profile of bifidobacteria at different growth phases. Such analyses are important, not only to provide insights into the molecular mechanisms underpinning adaptation of bifidobacteria to the different growth phases, but also to generate a valuable database of housekeeping genes. The genes in this category are assumed to be constitutively and uniformly expressed (18) and are important as reference genes to normalize the level of transcription in quantitative real-time reverse transcription-PCR (qRT-PCR) protocols (23, 24). The application of the qRT-PCR technique contributes very significantly to the understanding of complex biological processes that allow adaptation of bifidobacteria to the human gut (20). However, the validity of reference genes under specific experimental conditions must be determined before the application of quantitative mRNA expression studies. To provide a genetic baseline for differential expression studies in bifidobacteria, we investigated the transcriptome of *B. bifidum* PRL2010 at different

* Corresponding author. Mailing address: Laboratory of Probiogenomics, Department of Genetics, Biology of Microorganisms, Anthropology, and Evolution, University of Parma, Parco Area delle Scienze 11a, 43100 Parma, Italy. Phone: 39-521-905666. Fax: 39-521-906014. E-mail: marco.ventura@unipr.it.

† Supplemental material for this article may be found at http://aem.asm.org/.

▽ Published ahead of print on 14 October 2011.

growth phases under *in vitro* conditions. This analysis led to the identification of a set of PRL2010 housekeeping genes confirmed by qRT-PCR.

MATERIALS AND METHODS

Growth conditions. *B. bifidum* PRL2010 was routinely cultivated in an anaerobic atmosphere (2.99% H_2, 17.01% CO_2, and 80% N_2) in a chamber (Concept 400; Ruskin) at 37°C for 32 h in Man-Rogosa-Sharp (MRS) medium (Scharlau Chemie, Barcelona, Spain), supplemented with 0.05% (wt/vol) L-cysteine hydrochloride. Growth curves were performed at least in triplicate in PRL2010 batch cultures. Samples were taken at regular intervals from the cultures to measure the optical density at 600 nm (OD_{600}) and to determine the CFU/ml and pH values until 32 h following inoculation. Aliquots of 20 ml for the lag, early exponential, late exponential, and stationary phases were centrifuged for 10 min at 10,000 × *g* and 4°C. The pellets were then immediately frozen in liquid nitrogen and subjected to RNA extraction. We had to employ this extensive centrifugation step due to the severe difficulty in pelleting *B. bifidum* PRL2010 cells. The mRNA turnover was checked by qRT-PCR using log cell fractions collected at different times of centrifugation, i.e., 2 min, 5 min, and 10 min, targeting molecular chaperone genes, such as *groEL* and *dnaK*. No obvious change in the expression level between these different centrifugation times was noticed (data not shown).

RNA isolation. RNA was isolated according to the protocol described previously (37). The quality and integrity of the RNA was checked by Experion (Bio-Rad) analysis. As shown in Fig. S2 in the supplemental material, all RNA samples appeared to be of good quality with no obvious degradation.

Microarray, description, labeling, and hybridizations. Microarray analysis was performed with an oligonucleotide array based on the *B. bifidum* PRL2010 genome: a total of 39,249 oligonucleotide probes 35 bp in length were designed on 1,644 open reading frames (ORFs) using OligoArray 2.1 software (13). The oligonucleotides were synthesized in triplicate on a 2,000 by 40,000 CombiMatrix array (Mulkiteo). Replicates were distributed on the chip at random, nonadjacent positions. A set of 74 negative-control probes designed based on phage and plant sequences were also included on the chip.

Reverse transcription and amplification of 500 ng of total RNA were performed with a MessageAmp II-Bacteria kit (Ambion, Austin, TX) according to the manufacturer's instructions. Five micrograms of RNA was then labeled with a ULS labeling kit for Combimatrix arrays with Cy5 (Kreatech, The Netherlands). Hybridization of labeled DNA to *B. bifidum* PRL2010 arrays was performed according to CombiMatrix protocols (19).

Microarray data acquisition and treatment. Fluorescence scanning was performed on an InnoScan 710 microarray scanner (Innopsys, France). Signal intensities for each spot were determined using GenePix Pro 7 software (Molecular Devices). The signal background was calculated as the mean of negative controls plus 2 times the standard deviation (1). The absolute expression pattern was determined from at least three independent biological samples. A gene was considered to be subject to active transcription when it was detected in all three experiments and, consistent with previously published studies, when its hybridization signal exceeded the background standard deviation by 3-fold (4). This definition allows comparative transcription analysis of the different growth phases, although it provides only qualitative data.

Expression profiles of actively transcribed genes were analyzed using the EPCLUST module (http://www.bioinf.ebc.ee/EP/EP/). Signal intensities (*s*) were first transformed into $\log_2$ values (*v*) [$v_{i,j} = \log_2 (s_{i,j}/S_j)$, where $s_{i,j}$ is the signal intensity of gene *j* in the experiment *i*, and S_j is the average signal intensity of gene *j* in all experiments]. The EPCLUST K-means tool (default parameters) was used to find clusters of coregulated genes. The presence of enriched *c*luster *o*rthologue *g*ene (COG) functional categories was investigated in each of the eight clusters identified. This analysis highlighted the fact that one to four COG functional categories were enriched more than 2-fold with respect to the entire genome. The EPCLUST hierarchical-clustering tool (default parameters) was used to construct a heat map of genes belonging to these functional categories.

qRT-PCR. qRT-PCR primers (see Table S1 in the supplemental material) were used to amplify the reference genes indicated in Table S1. The criteria for primer design were based on a desired melting temperature (T_m) value between 58 and 60°C and an amplicon size of approximately 100 bp. qRT-PCR was performed using the CFX96 system (Bio-Rad, CA). PCR products were detected with SYBR green fluorescent dye and amplified according to the following protocol: one cycle of 95°C for 3 min, followed by 39 cycles of 95°C for 5 s and 66°C for 20 s. The melting curve was 65°C to 95°C with increments of 0.5°C/s.

Each PCR mixture contained the following: 12.5 μl 2× SYBR SuperMix Green (Bio-Rad, CA), 1 μl of cDNA dilution, and each of the forward and reverse primers at 0.5 μM; nuclease-free water was added to obtain a final volume of 20 μl. In each run, negative controls (no cDNA) for each primer set were included.

Data on the expression levels of the housekeeping genes were obtained in the form of crossing point (CP) values based on the "second derivative maximum" method as computed by CFX96 software (Bio-Rad, CA). Further data analysis was performed with CP raw data using the Excel-based application BestKeeper tool program (11).

Microarray data accession number. The microarray data have been deposited in the Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/) with the associated accession number GSE30832.

RESULTS AND DISCUSSION

***In vitro* cultivation.** Previous studies have shown that the fermentation abilities of *B. bifidum* PRL2010 are limited to a relatively small number of carbohydrates, including complex sugars, such as human milk oligosaccharides and mucin, or disaccharides, such as lactose (19). However, here, we decided to study the fermentation behavior of PRL2010 on glucose, as it represents the carbon source present in most of the commercially available synthetic growth media. For this reason, *B. bifidum* PRL2010 was grown in MRS plus L-cysteine hydrochloride (see Materials and Methods) supplemented with 2% glucose as the unique carbon source. Growth was evaluated by constant monitoring of the OD_{600} for 32 h using a microplate reader, as well as by determining the CFU/ml and pH values at different time points. As expected, the growth curve of *B. bifidum* PRL2010 obtained in this medium consists of a lag phase, an exponential phase, and a stationary phase (Fig. 1). The viable count of the culture decreases dramatically after 24 h, and the acidification of the medium by PRL2010 rapidly increases following 6 h of incubation to reach a minimal pH value of 4.2, as previously described for other bifidobacterial cultures (6) (Fig. 1).

Analysis of the *in vitro* transcription profile. mRNA was isolated from *B. bifidum* PRL2010 cells collected from a liquid culture at four time points following inoculation (Fig. 1). In total, 12 hybridization assays were carried out, consisting of a triplicate for each of the four time points (lag, early exponential, late exponential, and stationary phases). Genes were scored as being transcribed if their signal was identified in all hybridization experiments while also fulfilling additional signal strength criteria as outlined in Materials and Methods. Overall, 425 genes, representing 26% of the total identified PRL2010 gene arsenal, were shown to meet these criteria and were thus considered to be expressed under *in vitro* conditions. This repertoire of transcriptionally active genes was shown to be differently distributed in the different growth phases (Fig. 2a). Functional classification of these PRL2010 growth phase-specific genes according to the COG families is shown in Fig. 2b. Sixteen genes were specifically transcribed in the lag phase, i.e., they were not detected in the early/late exponential and stationary phases. These genes were predicted to be involved in protein translation, lipid and nucleotide transport, and metabolism. Similarly, of the 283 genes expressed in the early exponential phase, 12 were specifically expressed in that growth phase and were predicted to be involved in protein translation; signal transduction; amino acid, nucleotide, inorganic ion, and carbohydrate transport and metabolism; and energy production (Fig. 2a). In both the lag and exponential phases,

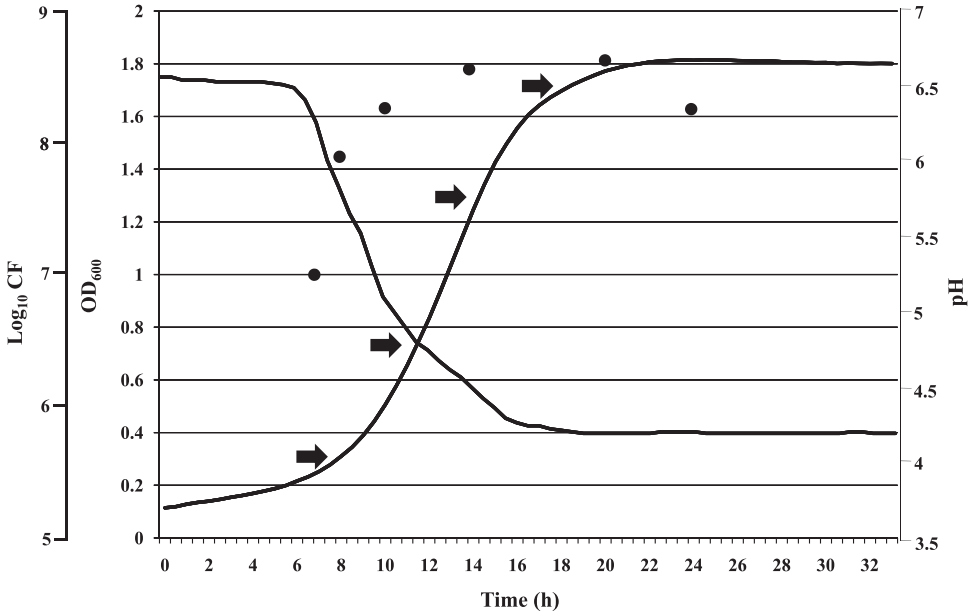

FIG. 1. Growth curve of *B. bifidum* PRL2010 in MRS broth. Growth was monitored as a function of time by measuring the OD_{600}, viable colony counts (circles), and pH values. The time points (T1 to T4) analyzed by microarray expression and qRT-PCR are indicated by arrows.

PRL2010 cells appeared to be metabolically active, as was also indicated by the rapid decrease in the pH. Finally, 21 and 62 of the 273 ORFs and 274 ORFs, respectively, were shown to be specifically transcribed during the late exponential and stationary phases. These genes are predicted to be involved in mechanisms commonly found in metabolically active cells, such as those that specify proteins involved in amino acid and carbohydrate transport and metabolism, as well as in typical processes occurring in senescent cells, like repair of misfolded proteins by molecular chaperones and of DNA damage. Moreover, among the genes whose transcription was shown to be upregulated at these time points is the F_1F_0 ATPase-encoding operon, whose function is to counteract acid stress (15, 27). Notably, during stationary phase, transcription of genes whose products are presumed glycolytic enzymes, such as glyceraldehyde-3-phosphate dehydrogenase, was increased, which would be expected to lead to increased production of reducing equivalents and energy-rich intermediates that protect cells against the deleterious consequences of this late growth phase (e.g., low pH and DNA damage). Interestingly, several genes that are known to provide protection against stressful conditions (*dnaK*, *dnaJ*, *grpE*, *clpB*, *clpC*, *hsp20*, *groEL*/*groES*, and *hspR*) in bifidobacteria (for a review, see reference 28) increase their transcription levels following the transition from lag to exponential phase (see Fig. S1 in the supplemental material). However, a large body of published data supports the notion that molecular-chaperone-encoding genes are also expressed during "normal" growth (i.e., under nonstressful conditions) of bifidobacteria in laboratory media, perhaps simply to assist the proper folding process of nascent proteins (28–32, 38, 39). Also, the increased production of certain glycolytic enzymes in response to stressful conditions has previously been reported for *Bifidobacterium longum* (14). Overall, we noticed a reduction in the transcription of genes belonging to categories that are involved in the maintenance of an active metabolic state during growth and, in particular, the transition from lag to stationary phase. Compared to their abundance as a functional category in the genome, the transcriptomes of all *in vitro* growth phases were somewhat enriched in functions related to transport and metabolism of carbohydrates (9.8%, 10%, 11%, and 10% of the transcribed genes in lag phase, early and late exponential phase, and stationary phase, respectively, compared to 9% of the genes in the genome) and translation and ribosomal biogenesis (24%, 22%, 14%, and 12% of the transcripts compared to 11% of the genome), as well as posttranslational modification, protein turnover, and chaperones (5% of the transcripts in the tested growth phases and 7% in stationary phase compared to 3% of the genes in the genome) (Fig. 2).

Accordingly, when all upregulated genes were classified based on their transcription profiles, 8 gene groups were identified, where members of each group exhibit similar expression profiles (Fig. 3). These groups encompass 83 genes, which exhibit enrichment of particular COG families based on the growth phase relative to their abundance in the PRL2010 genome (Fig. 3). The lag- and early-exponential-phase-specific gene transcription pool was enriched in genes involved in functions pertaining to an active metabolic state (e.g., ribosomal biogenesis; energy production and conversion; and nucleotide, lipid, and coenzyme transport and metabolism) (Fig. 3). The late-exponential- and early-stationary-phase transcripts displayed a higher representation of genes that perform functions in protein turnover, protein folding, energy conversion, intracellular trafficking, cell division, and signal transduction mechanisms than those of the preceding phases (Fig. 3).

Notably, in contrast to what was noticed for other bifidobacterial strains (D. van Sinderen, M. Ventura, and F. Turroni, unpublished data), 1,219 genes, corresponding to 74% of the identified ORFs of the *B. bifidum* PRL2010 genome that were present on the array, did not appear to be expressed during any

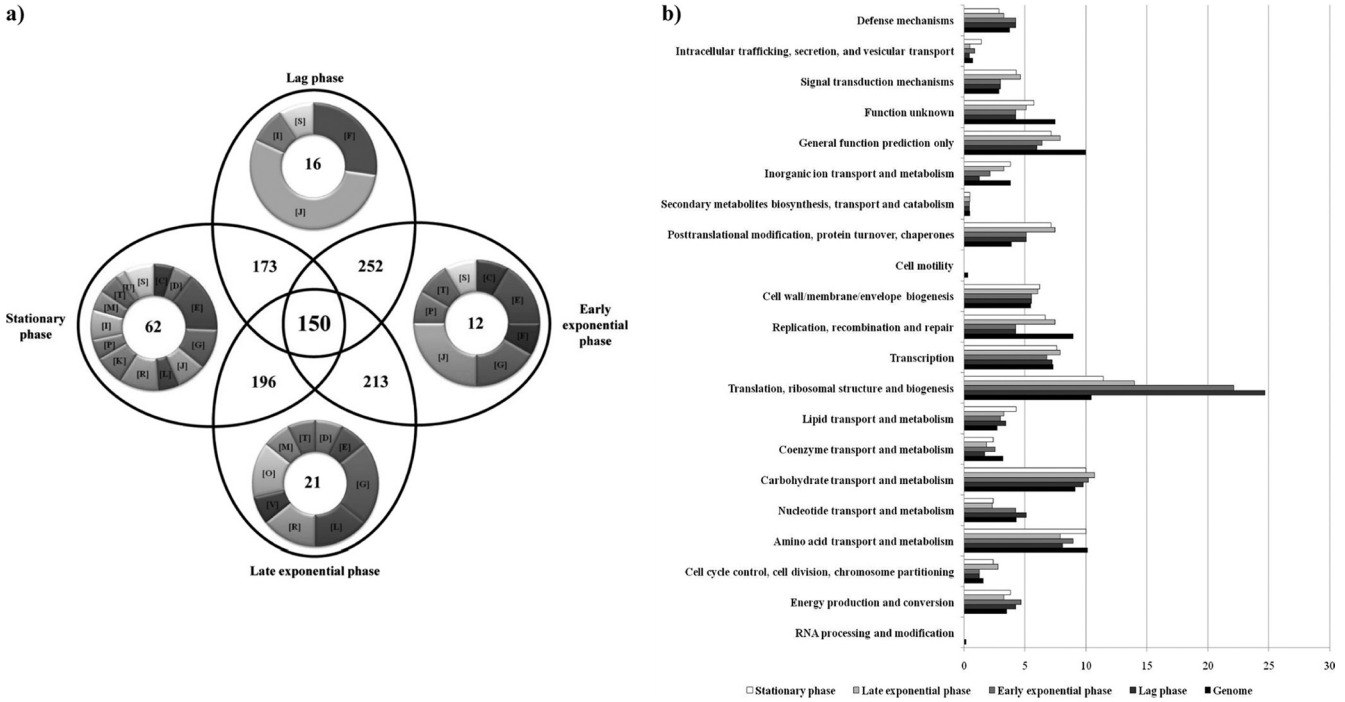

FIG. 2. Number and functional analysis according to COG categories of *B. bifidum* PRL2010 genes expressed *in vitro* by global transcription profiling. (a) Venn diagram displaying the numbers of genes expressed during the lag phase, early exponential phase, late exponential phase, and stationary phase. Inside each ring of the Venn diagram, the COG categories of the specifically upregulated genes for each of the tested growth phases are indicated. Each COG family is identified by a one-letter abbreviation: A, RNA processing and modification; B, chromatin structure and dynamics; C, energy production and conversion; D, cell cycle control and mitosis; E, amino acid metabolism and transport; F, nucleotide metabolism and transport; G, carbohydrate metabolism and transport; H, coenzyme metabolism; I, lipid metabolism; J, translation; K, transcription; L, replication and repair; M, cell wall/membrane/envelope biogenesis; N, cell motility; O, posttranslational modification, protein turnover, and chaperone functions; P, inorganic ion transport and metabolism; Q, secondary structure; T, signal transduction; U, intracellular trafficking and secretion; Y, nuclear structure; V, defense mechanisms; Z, cytoskeleton; R, general functional prediction only; S, function unknown. (b) Functional annotation of the *in vitro*-expressed genes of *B. bifidum* PRL2010 according to their COG categories. For each category, the black bar represents the percentage of genes in that category as detected in the sequenced genome of PRL2010 (19). The other bars show the percentages of genes transcribed during the lag, early exponential, late exponential, and stationary phases that belong to a particular category. The percentage was calculated as the percentage of transcribed genes belonging to the indicated COG category in respect to all transcribed genes.

of the tested growth phases. Such apparent disagreement with other microarray data might be linked to the different cutoff criteria used, as well as to the different microarray platforms used, which may have different sensitivities with regard to hybridization signal generation and detection. A large part of these genes were clustered in particular genomic loci, for example, the DNA region encompassing the prophage locus Bbif-1 (34, 35) (see Fig. S1 in the supplemental material). Nontranscribed genes were also found in PRL2010 genomic regions predicted or shown to mediate host-microbe interaction, such as those that specify sortase-dependent pili (the *pil1* and *pil3* loci) and the biosynthesis and transport system for Tad (tight-adherence) pili, which have previously been characterized for PRL2010 (5, 19) and *Bifidobacterium breve* UCC2003 (10). In contrast, other DNA regions predicted to be involved in host interaction, such as the gene encoding the putative adhesion factor BopA and the pilus-like structure specified by the *pil2* locus, are transcribed under such *in vitro* conditions (see Fig. S1 in the supplemental material).

Common genes expressed during *in vitro* growth. Transcription profiling analyses of PRL2010 cultivated under *in vitro* conditions revealed the existence of a core transcriptome, i.e., transcripts present in all three growth phases, consisting of 150 genes. Such genes are predicted to encode housekeeping functions, such as metabolism of lipids, amino acids, nucleotides, vitamins, and steroids, as well as glycolytic processes (Table 1). Furthermore, genes coding for hypothetical proteins and regulators are found among this core transcriptome of PRL2010.

Comparing the core transcriptome of PRL2010 with the core genome contents of the genus *Bifidobacterium* (2) revealed a high level of correspondence between the two sets of genes (Table 1).

Validation of the identified reference genes. Based on the PRL2010 core transcriptome, 22 genes were selected as putative reference genes that could be used for mRNA level normalization in qRT-PCR protocols (Table 1) because they have homologs in all known bifidobacterial genomes and because of their COG assignments. To verify the suitability of these genes as internal reference (IRF) genes, their expression stability was investigated and compared to that of the very commonly used reference 16S rRNA gene. The PCR efficiencies and linearity of the housekeeping-specific real-time assays were defined using standard curves based on PRL2010 genomic-DNA samples. The evaluation of efficiency is considered an essential marker in gene quantification procedure. The amplification of

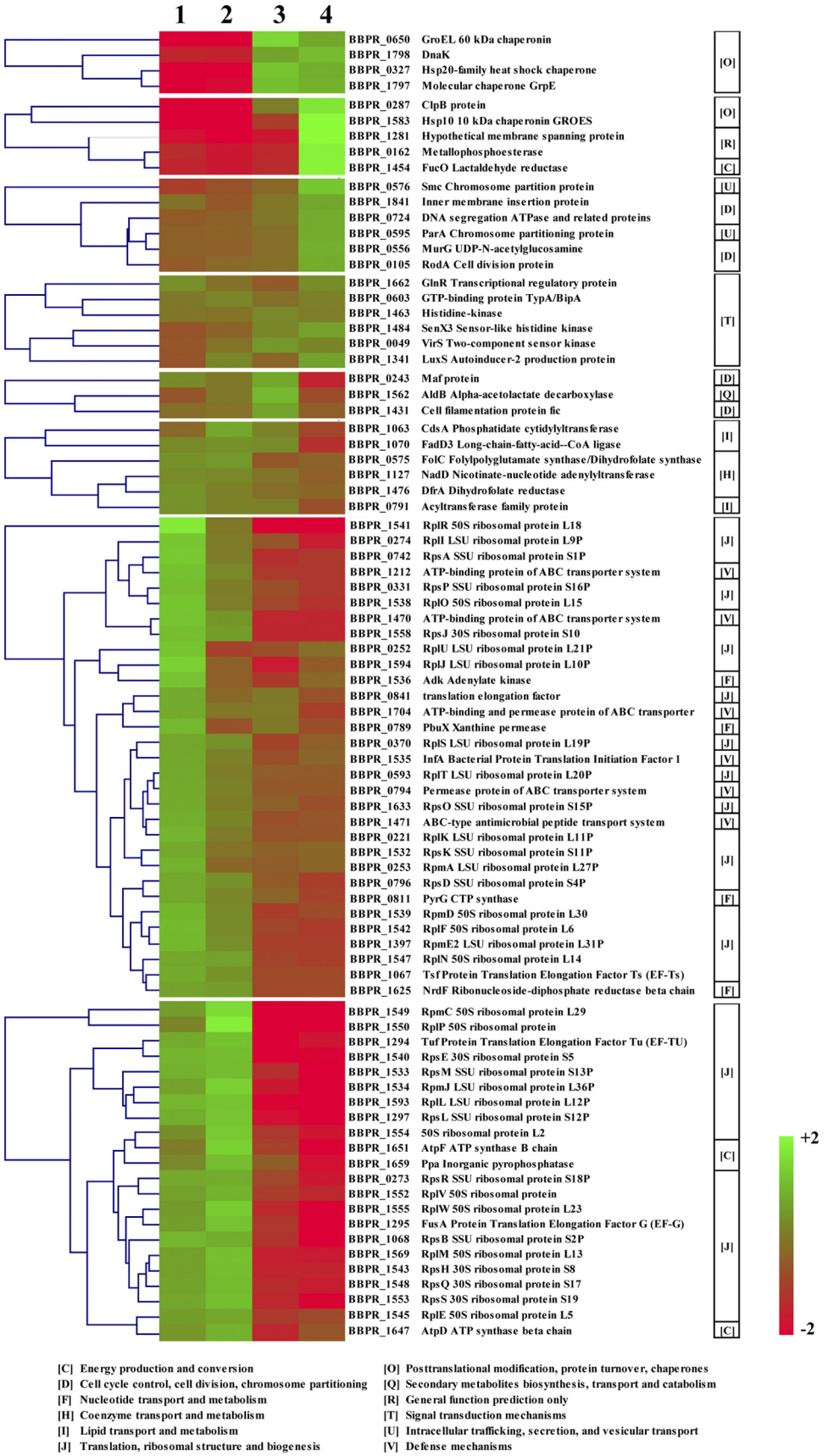

FIG. 3. Correlation between gene expression groups and COG functional categories. From left to right, the hierarchical clustering of genes belonging to cluster-enriched COG categories is shown, along with the corresponding eight clusters of coregulated genes and the description of enriched COG categories. Color legend is to the right; showing increased (green) and decreased (red) transcription levels.

TABLE 1. Core transcriptome of *B. bifidum* PRL2010 grown under *in vitro* conditions

ORF[a]	Product	COG	Presence in the minimal genome sequence[b]
BBPR_0034	Hypothetical protein in DPS family	[P]	−
BBPR_0043	Hypothetical membrane-spanning protein with DUF1212 domain	[S]	+
BBPR_0055	IS*3*/IS*911* family transposase	[L]	−
BBPR_0082	ATP-binding protein ABC transporter system for polysaccharides	[GM]	−
BBPR_0094	Narrowly hypothetical protein		−
BBPR_0107	Hypothetical secreted protein with FHA domain	[V]	+
BBPR_0115	Narrowly hypothetical membrane-spanning protein		−
BBPR_0126	DegP DO serine protease containing PDZ domain	[O]	+
BBPR_0137	XRE family transcriptional regulator	[K]	−
BBPR_0146	Transporter, MFS superfamily	[GEPR]	−
BBPR_0151	Aap amino acid permease	[E]	−
BBPR_0160	Glutaredoxin	[O]	−
BBPR_0169	Multidomain protein possibly involved in fatty acid or polyketide biosynthesis		−
BBPR_0171	IS*3*/IS*911* family transposase	[L]	−
BBPR_0172	Transcriptional regulator, TetR family	[K]	−
BBPR_0180	Mobilization protein		−
BBPR_0193	1,2-A-L-fucosidase	[N]	−
BBPR_0197	4-Oxalocrotonate tautomerase		+
BBPR_0200	Gpm phosphoglycerate mutase	[G]	+
BBPR_0201	Narrowly hypothetical protein		+
BBPR_0210	LeuC 3-isopropylmalate dehydratase large subunit	[E]	+
BBPR_0214	Hypothetical membrane-spanning protein in CAAX amino-terminal protease family		−
BBPR_0217	MurA UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase	[M]	+
BBPR_0233	LnbP lacto-*N*-biose phorylase		−
BBPR_0238	LysA diaminopimelate decarboxylase	[E]	+
BBPR_0251	Rne RNase G	[J]	+
BBPR_0252	RplU LSU ribosomal protein L21P	[J]	+
BBPR_0255	ObgE GTP-binding protein; GTP1/OBG family	[R]	+
BBPR_0258	SecE protein translocase subunit		−
BBPR_0261	GpsA glycerol-3-phosphate dehydrogenase [NAD(P)$^+$]	[C]	+
BBPR_0279	GltX glutamyl-tRNA synthetase	[J]	−
BBPR_0290	Hypothetical protein		−
BBPR_0327	Hsp20 family heat shock chaperone	[O]	−
BBPR_0329	RimM 16S rRNA-processing protein	[J]	+
BBPR_0346	RncS RNase III	[K]	+
BBPR_0370	RplS LSU ribosomal protein L19P	[J]	−
BBPR_0375	Pyridoxine biosynthesis protein	[H]	+
BBPR_0386	Hypothetical protein		−
BBPR_0412	Putative inner membrane protein	[S]	−
BBPR_0429	ATP-dependent DNA helicase; UvrD/REP family	[L]	+
BBPR_0432	Putative DNA modification methyltransferase		−
BBPR_0438	PepN membrane alanine aminopeptidase	[E]	+
BBPR_0443	Hypothetical protein		−
BBPR_0464	Phospholipase	[R]	−
BBPR_0466	DapF diaminopimelate epimerase	[E]	+
BBPR_0483	GlyS glycyl-tRNA synthetase	[J]	+
BBPR_0507	IspG 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase	[I]	+
BBPR_0519	Aldo/keto reductase family	[R]	−
BBPR_0535	HrpA ATP-dependent helicase	[L]	+
BBPR_0539	Ldh L-lactate dehydrogenase	[C]	−
BBPR_0558	FtsQ cell division protein	[M]	−
BBPR_0569	OppB oligopeptide transport system permease protein	[EP]	−
BBPR_0575	FolC folylpolyglutamate synthase/dihydrofolate synthase	[H]	+
BBPR_0581	Mur UDP-*N*-acetylmuramoyl-L-alanyl-D-glutamate lysine ligase	[M]	−
BBPR_0587	Gap glyceraldehyde 3-phosphate dehydrogenase	[G]	+
BBPR_0593	RplT LSU ribosomal protein L20P	[J]	+
BBPR_0596	ScpA segregation and condensation protein	[S]	−
BBPR_0597	ScpB segregation and condensation protein	[K]	−
BBPR_0603	GTP-binding protein TypA/BipA	[T]	+
BBPR_0635	DNA-binding protein	[L]	+
BBPR_0646	MoxR protein	[R]	−
BBPR_0650	GroEL 60-kDa chaperonin	[O]	+
BBPR_0651	Hypothetical protein with DUF909 domain		+
BBPR_0655	CspB cold shock protein	[K]	+
BBPR_0658	ClpC negative regulator of genetic competence	[O]	−
BBPR_0667	Hypothetical protein with YbaK/prolyl-tRNA synthetase-associated domain	[S]	−

Continued on following page

TABLE 1—*Continued*

ORF[a]	Product	COG	Presence in the minimal genome sequence[b]
BBPR_0670	Glutamate transport system permease protein GluC	[E]	+
BBPR_0676	Fused ATP-binding protein and permease of ABC transporter	[P]	−
BBPR_0702	IspD 2-*C*-methyl-D-erythritol 4-phosphate cytidylyltransferase	[I]	−
BBPR_0753	GlgX Glycogen operon protein	[G]	−
BBPR_0768	Xylulose-5-phosphate/Fructose-6-phosphate phosphoketolase	[G]	+
BBPR_0781	IS*3* family transposase	[L]	−
BBPR_0787	GlmS glucosamine-fructose-6-phosphate aminotransferase (isomerizing)	[M]	−
BBPR_0815	IscS/SufS cysteine desulfurase/selenocysteine lyase	[E]	+
BBPR_0827	PntA1 NAD(P) transhydrogenase subunit alpha part 1	[C]	+
BBPR_0846	Glycosyltransferase	[M]	+
BBPR_0861	Site-specific DNA methyltransferase (adenine-specific)		−
BBPR_0872	Hypothetical protein		−
BBPR_0904	Hypothetical protein		−
BBPR_0920	gp28		−
BBPR_0967	Hypothetical protein	[V]	−
BBPR_1010	Apt adenine phosphoribosyltransferase	[F]	+
BBPR_1022	Hypothetical protein		−
BBPR_1028	Tkt transketolase	[G]	+
BBPR_1029	Tal transaldolase	[G]	+
BBPR_1041	AroE shikimate 5-dehydrogenase	[E]	+
BBPR_1058	Solute-binding protein of ABC transporter system for sugars	[G]	−
BBPR_1063	CdsA phosphatidate cytidylyltransferase	[I]	+
BBPR_1065	PyrH uridylate kinase	[F]	+
BBPR_1070	FadD3 long-chain fatty acid-CoA ligase	[IQ]	−
BBPR_1081	Hypothetical protein with UPF0079 domain	[R]	−
BBPR_1109	Macrolide efflux protein		−
BBPR_1111	Hypothetical protein		−
BBPR_1127	NadD nicotinate nucleotide adenylyltransferase	[H]	+
BBPR_1130	ThrC threonine synthase	[E]	+
BBPR_1145	TyrS tyrosyl-tRNA synthetase	[J]	+
BBPR_1149	HsdR-like protein of type I restriction modification system	[V]	−
BBPR_1234	pflA pyruvate formate-lyase-activating enzyme	[O]	−
BBPR_1253	Lhr ATP-dependent helicase	[R]	+
BBPR_1263	PlsC 1-acyl-sn-glycerol-3-phosphate acyltransferase	[I]	+
BBPR_1271	Hypothetical protein		−
BBPR_1294	Tuf protein translation elongation factor Tu (EF-TU)	[J]	+
BBPR_1300	Glycosyl hydrolase family protein		−
BBPR_1302	PurC phosphoribosylamidoimidazole-succinocarboxamide synthase	[F]	+
BBPR_1306	Hypothetical protein	[S]	−
BBPR_1348	Solute-binding protein of ABC transporter system for peptides	[R]	−
BBPR_1362	IleS isoleucyl-tRNA synthetase	[J]	+
BBPR_1392	Rfe undecaprenyl-phosphate-alpha-*N*-acetylglucosaminephosphotransferase	[M]	+
BBPR_1399	Glucose uptake protein	[G]	−
BBPR_1410	Hypothetical protein		−
BBPR_1423	Narrowly hypothetical protein with DUF979 domain	[S]	−
BBPR_1441	Narrowly hypothetical membrane-spanning protein		+
BBPR_1446	HprT hypoxanthine-guanine phosphoribosyltransferase	[F]	+
BBPR_1476	DfrA dihydrofolate reductase	[H]	+
BBPR_1484	SenX3 sensor-like histidine kinase	[T]	+
BBPR_1494	Histidine kinase sensor of two-component system		+
BBPR_1508	PtsG PTS system, glucose-specific IIABC component	[G]	−
BBPR_1521	RbfA ribosome-binding factor A	[J]	+
BBPR_1531	RpoA DNA-directed RNA polymerase alpha chain	[K]	+
BBPR_1535	InfA bacterial protein translation initiation factor 1	[J]	+
BBPR_1539	RpmD 50S ribosomal protein L30	[J]	+
BBPR_1540	RpsE 30S ribosomal protein S5	[J]	+
BBPR_1545	RplE 50S ribosomal protein L5	[J]	+
BBPR_1546	RplX 50S ribosomal protein L24	[J]	+
BBPR_1551	RpsC 30S ribosomal protein S3	[J]	+
BBPR_1562	AldB alpha-acetolactate decarboxylase	[Q]	−
BBPR_1569	RplM 50S ribosomal protein L13	[J]	+
BBPR_1579	RpmG LSU ribosomal protein L33P	[J]	−
BBPR_1599	Hypothetical protein		−
BBPR_1600	Phosphate transport ATP-binding protein	[P]	−
BBPR_1604	Two-component response regulator	[TK]	−
BBPR_1606	Pfs 5′-methylthioadenosine nucleosidase/*S*-adenosylhomocysteine nucleosidase	[F]	+
BBPR_1612	PepC2 aminopeptidase C	[E]	−

Continued on following page

TABLE 1—*Continued*

ORF[a]	Product	COG	Presence in the minimal genome sequence[b]
BBPR_1649	AtpA ATP synthase alpha chain	[C]	+
BBPR_1653	AtpB ATP synthase A chain	[C]	+
BBPR_1661	Narrowly hypothetical membrane-spanning protein		+
BBPR_1662	GlnR transcriptional regulatory protein	[TK]	+
BBPR_1688	MurE UDP-*N*-acetylmuramoylalanyl-D-glutamate-2,6-diaminopimelate ligase	[M]	+
BBPR_1691	GlnD (protein-PII) uridylyltransferase	[M]	+
BBPR_1707	Fimbrial subunit FimA		−
BBPR_1738	TopA DNA topoisomerase I	[L]	+
BBPR_1771	Thiazole synthase	[H]	−
BBPR_1774	Hypothetical protein		−
BBPR_1792	Haloacid dehalogenase-like hydrolase (HAD superfamily)	[R]	−
BBPR_1795	HspR heat shock response regulator	[K]	+
BBPR_1796	DnaJ class molecular chaperone	[O]	+
BBPR_1798	DnaK	[O]	+
BBPR_1809	Sir2-type regulatory protein	[K]	+
BBPR_1836	Thioredoxin reductase	[O]	−
BBPR_1841	Inner membrane insertion protein	[U]	−

[a] The putative reference genes are shaded.
[b] +, present; −, absent. The minimal genome sequences were retrieved from reference 2.

putative IRF target genes obtained exhibited PCR efficiencies ranging from 96.4% to 100%. The expression stability of a tested housekeeping gene was analyzed as a function of different PRL2010 growth conditions, including exposure to stressful treatments, such as thermal, osmotic, and acid stress, and cultivation in the presence of different carbon sources (e.g., glucose and lactose) and different growth phases. Among the potential IRF genes tested, the most highly abundant transcript was shown to correspond to the 16S rRNA gene, whereas the *pntA1* gene was found to generate the lowest level of transcripts (Fig. 4). The achieved *c*rossing *p*oint (CP) data were further analyzed using the BestKeeper tool software (11), which allows identification of the "optimal" housekeeping reference genes (17) by calculation of CP data variation, CP data ranges, standard deviation (SD), and coefficient of variation (CV). As shown in Tables 2 and 3, the *uvrD-rep*, *gluC*, and *pdxS* genes displayed the highest level of expression stability (SD $\leq \pm 1$ CP and CV of 2.5% CP). Notably, such data show that there is only a low level of variation associated with the expression of these three candidate genes under the different conditions tested (see Table S2 in the supplemental material). In fact, inspection of the CP data variations based on all the CP values derived from the different conditions analyzed (e.g., temperature shifts, treatment with NaCl, acid exposure, various carbon sources, and change in growth phases) displayed high expression stability for each of the three genes investigated.

Conclusions. Bifidobacteria have become a prominent group of microorganisms due to their perceived role as health-promoting bacteria in the human gut (33). However, despite a

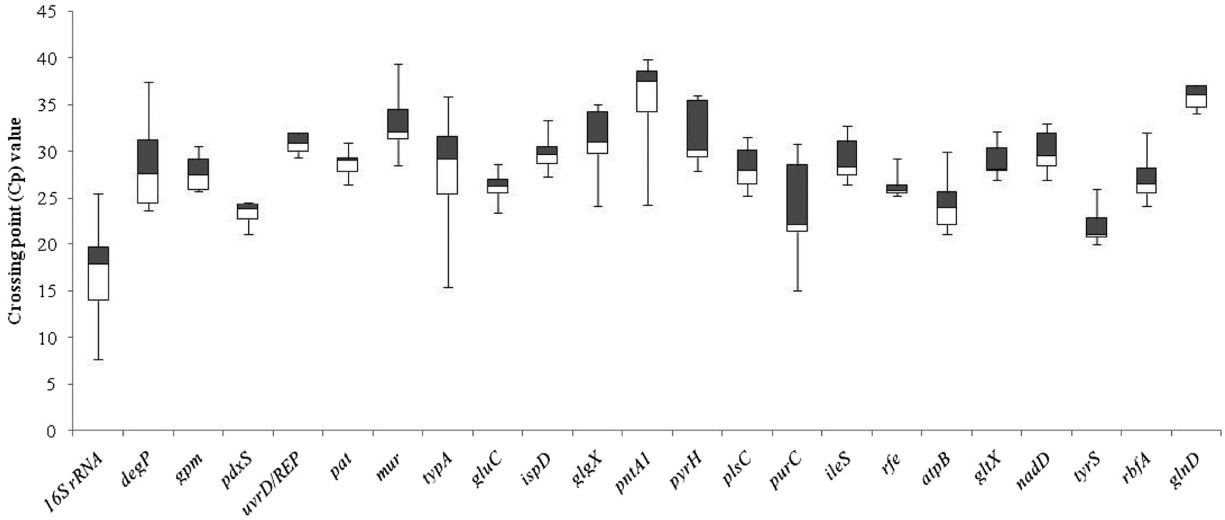

FIG. 4. Box plot overview of the CP values of *B. bifidum* PRL2010 cells cultivated under different growth conditions (growth phases; thermal, acidic, and osmotic stresses; and in the presence of different carbon sources). The box plots highlight the median 25th (white) and 75th (black) percentiles and expression range across the different reference genes.

TABLE 2. Statistical analysis related to variation of CP values obtained from transcriptional behavior of various *B. bifidum* PRL2010 genes[a]

Factor[b]	Value										
	pdxS	*uvrD*/Rep	*gluC*	*atpB*	*rfe*	*pyrH*	*gltX*	*nadD*	*tyrS*	*rbfA*	*glnD*
GM (CP)	23.73	31.56	26.12	24.19	25.68	30.67	29.12	29.52	22.3	27.31	35.71
AM (CP)	23.74	31.57	26.14	24.25	25.7	30.7	29.18	29.58	22.41	27.48	35.74
Min (CP)	22.22	30.46	23.06	21.91	24.35	28	26.87	27.35	20.05	23	33.67
Max (CP)	25.02	32.36	27.12	26.81	27.53	33.45	32.13	33.24	26.46	33.3	38.17
SD (±CP)	0.64	0.48	0.6	1.56	0.81	1.01	1.65	1.56	1.87	2.57	1.16
CV (%CP)	2.71	1.53	2.29	6.42	3.15	3.29	5.65	5.26	8.36	9.34	3.24

[a] Housekeeping gene expression was assessed on total RNA templates isolated from *B. bifidum* PRL2010 growth under different conditions.
[b] GM (CP), geometric mean of the CP; AM (CP), arithmetic mean of the CP; Min (CP) and Max (CP), extreme values of the CP; SD (±CP), standard deviation of the CP; CV (%CP), CV expressed as a percentage of the CP level.

growing interest in the genomics, molecular biology, and genetics of this group of bacteria, surprisingly little is known about genome expression under *in vivo* or *in vitro* conditions (25). Microarray-based analysis allows insights into the metabolic states of bacteria under different culture conditions. The data described in this study show that, during *in vitro* growth, the lag, exponential, and stationary phases are characterized by a specific global transcription pattern. Analysis of the functional annotations of the transcribed genes revealed a shift from cell division-focused gene expression during the lag phase to a more pronounced carbohydrate metabolism-related gene repertoire in the exponential growth phase. Later in the growth cycle of *B. bifidum* PRL2010, cells modify their transcription patterns by reducing the total number of genes that are expressed, thereby sustaining the stationary growth phase.

Surprisingly, just 26% of the total number of genes present in the genome of *B. bifidum* PRL2010 were designated as being transcribed under *in vitro* conditions using our criteria, indicating that only a modest part of the overall genetic makeup of PRL2010 is active during growth in liquid broth. However, such a finding can be explained by assuming that all gut-expressed genes (e.g., those encoding mucin breakdown, as well as host-microbe interaction) remain transcriptionally silent during broth cultivation of *B. bifidum* PRL2010. Alternatively, the applied criteria may have been too stringent for the data sets obtained here, perhaps eliminating many genes that were in fact transcriptionally active. Nevertheless, if just 26% of the total gene repertoire of the microorganism is active under the *in vitro* conditions used here, this poses new questions related to the minimal transcriptome and the global regulation of genes when a bacterium is cultivated in a "nonnatural" environment. The core transcriptome of PRL2010 contains a large number of genes that are present in all bifidobacterial genomes so far sequenced. This highlights an interesting parallel between the core transcriptome and the minimal genome content of the genus *Bifidobacterium*, suggesting that the genetic requirements for the utilization of simple carbohydrate substrates are conserved and are always expressed in bifidobacteria.

Furthermore, apart from these gene categories, there are 346 nontranscribed genes belonging to the "unknown-function" category, and we cannot even speculate under what circumstances/conditions these genes may be expressed.

Our study facilitated the identification of a set of bifidobacterial housekeeping genes suitable as internal reference genes for the quantification of mRNA targets in bifidobacterial cells by quantitative gene expression analysis techniques, such as real-time RT-PCR. This technique is frequently applied to evaluate changes in gene expression under different growth conditions involving stress treatments, fermentation on different carbon sources, and protease treatments (for examples related to bifidobacteria, see references8, 21, and 36). However, the reference genes used so far for such gene transcript quantification by qRT-PCR approaches in bifidobacteria have not been subjected to solid validation. It is worth mentioning that the proposed reference genes, as suggested in this report, represent the optimal reference conditions for relative qRT-PCR analysis of gene expression for the growth conditions used in this study. However, since it is possible that specific but as yet untested environmental conditions affect bacterial gene expression, we have to accept that universal reference genes do not exist. Thus, a certain amount of caution should be applied in the selection of reference genes in qRT-PCR assays, including those proposed here, when using different growth conditions and/or different bifidobacterial species.

This study has focused on the *B. bifidum* PRL2010 strain, which we use as a model microorganism to study the genetics and physiology of intestinal commensal bifidobacteria. However, it is worth mentioning that the genetic behavior of the various members of the genus *Bifidobacterium* is likely to be different, and thus, the findings described here might not be applicable to all bifidobacteria. For these reasons, further research activities should be carried out to facilitate the determination of global transcription patterns of bifidobacteria un-

TABLE 3. Data from BestKeeper correlation analysis[a]

Factor	Value for BestKeeper vs:										
	pdxS	*uvrD*/Rep	*gluC*	*atpB*	*rfe*	*pyrH*	*gltX*	*nadD*	*tyrS*	*rbfA*	*glnD*
Coefficient of correlation (*r*)	0.627	0.811	0.553	0.568	0.607	−0.031	0.94	0.47	0.9	0.94	−0.01
P value	0.016	0.001	0.04	0.034	0.022	0.914	0.001	0.031	0.001	0.001	0.984

[a] Measures of correlation between each candidate gene's expression and the BestKeeper index computed from the best candidate genes.

der *in vivo* conditions, allowing comparative analysis with *in vitro* transcriptome data, which will be important in order to determine if bifidobacterial transcription in different gut sections corresponds to the transcription profile observed in a specific growth phase and thus to obtain molecular clues to gut colonization by bifidobacteria.

ACKNOWLEDGMENTS

This work was financially supported by the Italian Award for Outstanding Young Researcher scheme "Incentivazione alla mobilità di studiosi stranieri e italiani residenti all'estero" and by a Marie Curie Reintegration Grant (MERG-CT-2005-03080) to M.V. and a FEMS Advanced Fellowship 2011 to F.T. D.V.S. is a member of the Alimentary Pharmabiotic Centre, which is a Centre for Science and Technology (CSET) funded by the Science Foundation Ireland (SFI) through the Irish government's National Development Plan (grant no. 02/CE/B124 and 07/CE/B1368).

We thank S. Ottonello for constructive and critical reading of the manuscript.

REFERENCES

1. **Bilban, M., L. K. Buehler, S. Head, G. Desoye, and V. Quaranta.** 2002. Defining signal thresholds in DNA microarrays: exemplary application for invasive cancer. BMC Genomics **3:**19.
2. **Bottacini, F., et al.** 2010. Comparative genomics of the genus Bifidobacterium. Microbiology **156:**3243–3254.
3. **Cronin, M., M. Ventura, G. F. Fitzgerald, and D. van Sinderen.** 2011. Progress in genomics, metabolism and biotechnology of bifidobacteria. Int. J. Food Microbiol. **149:**4–18.
4. **Denou, E., et al.** 2007. Gene expression of commensal Lactobacillus johnsonii strain NCC533 during in vitro growth and in the murine gut. J. Bacteriol. **189:**8109–8119.
5. **Foroni, E., et al.** 2011. Genetic analysis and morphological identification of pilus-like structures in members of the genus *Bifidobacterium*. Microb. Cell Fact. **10**(Suppl. 1)**:**S16.
6. **Laine, R., S. Salminen, Y. Benno, and A. C. Ouwehand.** 2003. Performance of bifidobacteria in oat-based media. Int. J. Food Microbiol. **83:**105–109.
7. **Lee, J. H., and D. J. O'Sullivan.** 2010. Genomic insights into bifidobacteria. Microbiol. Mol. Biol. Rev. **74:**378–416.
8. **Liu, D., et al.** 2011. Proteomics analysis of Bifidobacterium longum NCC2705 growing on glucose, fructose, mannose, xylose, ribose, and galactose. Proteomics **11:**2628–2638.
9. **Marco, M. L., S. Pavan, and M. Kleerebezem.** 2006. Towards understanding molecular modes of probiotic action. Curr. Opin. Biotechnol. **17:**204–210.
10. **O'Connell Motherway, M., et al.** 2011. Functional genome analysis of Bifidobacterium breve UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. Proc. Natl. Acad. Sci. U. S. A. **108:**11217–11222.
11. **Pfaffl, M. W., A. Tichopad, C. Prgomet, and T. P. Neuvians.** 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pair-wise correlations. Biotechnol. Lett. **26:**509–515.
12. **Roger, L. C., A. Costabile, D. T. Holland, L. Hoyles, and A. L. McCartney.** 2010. Examination of faecal Bifidobacterium populations in breast- and formula-fed infants during the first 18 months of life. Microbiology **156:** 3329–3341.
13. **Rouillard, J. M., M. Zuker, and E. Gulari.** 2003. OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. Nucleic Acids Res. **31:**3057–3062.
14. **Sánchez, B., et al.** 2005. Proteomic analysis of global changes in protein expression during bile salt exposure of Bifidobacterium longum NCIMB 8809. J. Bacteriol. **187:**5799–5808.
15. **Sánchez, B., C. G. de los Reyes-Gavilan, and A. Margolles.** 2006. The F1F0-ATPase of Bifidobacterium animalis is involved in bile tolerance. Environ. Microbiol. **8:**1825–1833.
16. **Servin, A. L.** 2004. Antagonistic activities of lactobacilli and bifidobacteria against microbial pathogens. FEMS Microbiol. Rev. **28:**405–440.
17. **Tasara, T., and R. Stephan.** 2007. Evaluation of housekeeping genes in Listeria monocytogenes as potential internal control references for normalizing mRNA expression levels in stress adaptation models using real-time PCR. FEMS Microbiol. Lett. **269:**265–272.
18. **Thellin, O., B. El Moualij, E. Heinen, and W. Zorzi.** 2009. A decade of improvements in quantification of gene expression and internal standard selection. Biotechnol. Adv. **27:**323–333.
19. **Turroni, F., et al.** 2010. Genome analysis of Bifidobacterium bifidum PRL2010 reveals metabolic pathways for host-derived glycan foraging. Proc. Natl. Acad. Sci. U. S. A. **107:**19514–19519.
20. **Turroni, F., et al.** 2010. Characterization of the serpin-encoding gene of Bifidobacterium breve 210B. Appl. Environ. Microbiol. **76:**3206–3219.
21. **Turroni, F., et al.** 2009. Exploring the diversity of the bifidobacterial population in the human intestinal tract. Appl. Environ. Microbiol. **75:**1534–1545.
22. **Turroni, F., D. van Sinderen, and M. Ventura.** 2011. Genomics and ecological overview of the genus Bifidobacterium. Int. J. Food Microbiol. **149:**37–44.
23. **Vandecasteele, S. J., W. E. Peetermans, R. Merckx, and J. Van Eldere.** 2001. Quantification of expression of Staphylococcus epidermidis housekeeping genes with Taqman quantitative PCR during in vitro growth and under different conditions. J. Bacteriol. **183:**7094–7101.
24. **Vandesompele, J., et al.** 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. **3:**RESEARCH0034.
25. **Ventura, M., C. Canchaya, G. F. Fitzgerald, R. S. Gupta, and D. van Sinderen.** 2007. Genomics as a means to understand bacterial phylogeny and ecological adaptation: the case of bifidobacteria. Antonie Van Leeuwenhoek **91:**351–372.
26. **Ventura, M., et al.** 2007. Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. Microbiol. Mol. Biol. Rev. **71:**495–548.
27. **Ventura, M., C. Canchaya, D. van Sinderen, G. F. Fitzgerald, and R. Zink.** 2004. Bifidobacterium lactis DSM 10140: identification of the atp (atpBEFHAGDC) operon and analysis of its genetic structure, characteristics, and phylogeny. Appl. Environ. Microbiol. **70:**3110–3121.
28. **Ventura, M., et al.** 2006. How high G+C Gram-positive bacteria and in particular bifidobacteria cope with heat stress: protein players and regulators. FEMS Microbiol. Rev. **30:**734–759.
29. **Ventura, M., C. Canchaya, Z. Zhang, G. F. Fitzgerald, and D. van Sinderen.** 2007. Molecular characterization of hsp20, encoding a small heat shock protein of bifidobacterium breve UCC2003. Appl. Environ. Microbiol. **73:** 4695–4703.
30. **Ventura, M., C. Canchaya, R. Zink, G. F. Fitzgerald, and D. van Sinderen.** 2004. Characterization of the groEL and groES loci in Bifidobacterium breve UCC 2003: genetic, transcriptional, and phylogenetic analyses. Appl. Environ. Microbiol. **70:**6197–6209.
31. **Ventura, M., G. F. Fitzgerald, and D. van Sinderen.** 2005. Genetic and transcriptional organization of the clpC locus in Bifidobacterium breve UCC 2003. Appl. Environ. Microbiol. **71:**6282–6291.
32. **Ventura, M., J. G. Kenny, Z. Zhang, G. F. Fitzgerald, and D. van Sinderen.** 2005. The clpB gene of Bifidobacterium breve UCC 2003: transcriptional analysis and first insights into stress induction. Microbiology **151:**2861–2872.
33. **Ventura, M., et al.** 2009. Genome-scale analyses of health-promoting bacteria: probiogenomics. Nat. Rev. Microbiol. **7:**61–71.
34. **Ventura, M., et al.** 2010. Analyses of bifidobacterial prophage-like sequences. Antonie Van Leeuwenhoek **98:**39–50.
35. **Ventura, M., et al.** 2009. Comparative analyses of prophage-like elements present in bifidobacterial genomes. Appl. Environ. Microbiol. **75:**6929–6936.
36. **Ventura, M., et al.** 2009. The Bifidobacterium dentium Bd1 genome sequence reflects its genetic adaptation to the human oral cavity. PLoS Genet. **5:**e1000785.
37. **Ventura, M., et al.** 2005. The ClgR protein regulates transcription of the clpP operon in Bifidobacterium breve UCC 2003. J. Bacteriol. **187:**8411–8426.
38. **Ventura, M., R. Zink, G. F. Fitzgerald, and D. van Sinderen.** 2005. Gene structure and transcriptional organization of the dnaK operon of Bifidobacterium breve UCC 2003 and application of the operon in bifidobacterial tracing. Appl. Environ. Microbiol. **71:**487–500.
39. **Zomer, A., et al.** 2009. An interactive regulatory network controls stress response in Bifidobacterium breve UCC2003. J. Bacteriol. **191:**7039–7049.