Title: Degenerate PCR primers for detecting putative priming glycosyltransferase genes in Bifidobacterium strains.

Running title: Priming-GTF detection in Bifidobacterium.

Authors: Claudio Hidalgo-Cantabrana*, Irene Ordoñez, Patricia Ruas-Madiedo, and Abelardo Margolles

Addresses:
1 Department of Microbiology and Biochemistry of Dairy Products, Instituto de Productos Lácteos de Asturias – Consejo Superior de Investigaciones Científicas (IPLA-CSIC), Paseo Río Linares s/n, 33300 Villaviciosa, Asturias, Spain.

* Corresponding author: Instituto de Productos Lácteos de Asturias (IPLA-CSIC), Paseo Río Linares s/n, 33300 Villaviciosa, Asturias, Spain. Tel.: +34 985892131, e-mail: claudio@ipla.csic.es
Abstract

A new PCR-based method to detect putative exopolysaccharide (EPS) producers from the genus *Bifidobacterium* was developed based on the detection of two priming glycosyltransferase genes: *rfbP* (undecaprenyl-phosphate sugar phospho-transferase) and *cpsD* (galactosyl-transferase). An *in silico* analysis of the genomes of 28 bifidobacterial strains, belonging to 8 different species, allowed us to detect *rfbP*, *cpsD*, or both, in the large majority of the genomes. Based on DNA sequence homology studies, twenty four degenerated primers were synthesized in order to select the primer pairs with the broadest capacity to detect the presence of these genes. Four primer pairs targeting internal regions of *rfbP* and *cpsD* were selected, allowing the detection of at least one of the two genes in 63 out of 99 bifidobacterial strains analyzed, whereas control strains from other genera yielded negative results, suggesting that these genes are widely spread in this genus. The use of these primers is recommended to screen for the potential of *Bifidobacterium* strains to produce EPS.

**Keywords:** Bifidobacteria, exopolysaccharide, priming-glycosyltransferase, PCR detection.
1. Introduction

Exopolysaccharides (EPS) are carbohydrate polymers present as an extracellular layer in Gram positive and Gram negative bacteria that could form a capsule or could be secreted displaying a network-like structure. One of the main interests in EPS producing strains is related to their application in the industrial sector (De Philippis et al., 2011, Rehm, 2010) and their positive impact on human health (Round JL, et al., 2009, 2011). The health benefits promoted in the host by the probiotic strains, i.e. “live microorganisms which when administered in adequate amounts confer a health benefit on the host” (FAO/WHO, 2006), have been related, on some occasions, with the production of EPS (Fanning et al., 2012; Nikolic et al., 2012). Most common bacteria used as probiotics belong to Lactobacillus and Bifidobacterium genera. Both are Gram positive, natural inhabitants of the gastrointestinal tract of animals (Turroni et al., 2013, Ventura et al., 2012) and some species also inhabit the oral cavity (Ventura et al., 2009). The composition and structure of the repeating unit of some EPS polymers has been described using nuclear magnetic resonance (NMR) for members of Lactobacillus and other lactic acid bacteria (LAB) (Behare et al., 2009, Laws et al., 2001) and for bifidobacteria (Hidalgo-Cantabrana et al., 2014, Leivers et al., 2011).

To date, the mechanism of EPS synthesis in these bacterial groups is still not clear, although, in recent years some information has been elucidated and several reports have characterized the genes involved in the EPS synthesis in Streptococcus thermophilus (Broadbent et al., 2003) and other LAB (Boels et al., 2001; Denou et al., 2008; Dertli et al., 2013). In some species of LAB, the genes involved in EPS synthesis are located in clusters with a highly conserved structural-functional organization. These genes can be categorized into four groups: regulatory genes, proteins involved in polymerization and chain length determination, glycosyltransferases (GTF), and the last group is genes involved in transport and polymerization (Hidalgo-Cantabrana et al., 2014). In bifidobacteria there is not such a consensus in the genetic organization but the genes present have similar functions based on homology studies, although the organization within the clusters is largely different (Hidalgo-Cantabrana et al., 2014). One of these genes is the priming-GTF, which is a specific GTF that transfers a sugar-1-phosphate to a lipophilic carrier molecule that is anchored in the bacterial membrane. This is the first step of EPS synthesis and some authors have described that the deletion, or inactivation, of the priming-GTF dramatically reduces EPS synthesis (Low et al., 1998; Stingele et al., 1996; van Kranenburg et al., 1997). In Lactobacillus jonschonii FI9785 the deletion of the priming-GTF gene, epsE, reduces the two different EPS usually produced to only one (Dertli et al., 2013). In bifidobacteria, the priming-GTF genes could be annotated in genomic databases as “undecaprenyl-phosphate sugar phospho-transferase” (rfbP) or “galactosyl-transferase” (cpsD). Both, rfbP and cpsD, are highly conserved between Bifidobacterium species and most bifidobacterial genomes harbor one of these priming-gtf, some strains harboring both genes (Hidalgo-Cantabrana et al., 2014).

A molecular approach to detect cpsD and rfbP has been used before to identify EPS-producing strains for LAB (Provencher et al., 2003) and bifidobacteria (Ruas Madiedo et al., 2007). In this study, after in silico analyzing the presence of priming-GTF genes in publically available bifidobacterial genomes, we designed degenerated oligonucleotide primers to detect, by PCR, the genes cpsD and rfbP specifically in Bifidobacterium genus.

2 Material and methods

2.1 In silico analysis

To analyse the ubiquity of the priming-GTF genes in the genus Bifidobacterium an in silico analysis was performed with the 28 bifidobacterial complete genomes available at the time of writing in the GenBank database (http://www.ncbi.nlm.nih.gov/genbank). The two priming-GTF genes, rfbP (accession number: NP_695455) and cpsD (NP_695447), from
Bifidobacterium longum subsp. longum NCC2705 (AE014295, Schell et al., 2002), were used as a template to screen and identify the priming-GTF genes in all complete bifidobacterial genomes. For this purpose, the Basic Local Alignment Search Tool, BLAST (Zhang et al., 2000) available at the National Center for Biotechnology Information site, NCBI (http://www.ncbi.nlm.nih.gov/) was used. Then, a phylogenetic tree was depicted based on the amino acid sequence alignment of both priming-GTF using the ClustalX 2.1 software (http://www.clustal.org, Larkin et al., 2007); for clustering, the Neighbour-Joining method with a total of 100 bootstrap replications was applied, as implemented in the free software SplitsTree 4.12.3 (http://www.splitstree.org, Huson and Bryant, 2006).

2.2 Bacterial strains and growth conditions

A collection of 99 Bifidobacterium strains held in the “Instituto de Productos Lácteos de Asturias” (IPLA) collection, from different origins, such as breast milk, infant and adult faeces, commercial dairy products and culture collection strains, were used in this study. Another 12 strains belonging to other genera were used as a control. Bifidobacterial strains were routinely cultivated in MRSC [MRS Difco (BD, Biosciences, San Diego, CA) containing 0.25% L-cysteine-HCl (Sigma-Chemical Co., St. Louis, MO)] at 37°C under anaerobic conditions (80% N₂, 10% CO₂, 10% H₂) in an MG500 chamber (Don Whitley Scientific, West Yorkshire, UK). The media and growth temperature used for negative control microorganisms was: MRS at 37°C for Lactobacillus rhamnosus LMG 18243 (GG), Lb. plantarum H2 and Lb. casei BA61, M17 (Oxoid, Hampshire, England) at 32°C for Lactococcus lactis subsp. cremoris V31, Streptococcus thermophilus St.-Body and Enterococcus faecalis IF3/1, MRS at 32°C for Leuconostoc mesenteroides CECT 394, LB at 37°C for Escherichia coli LMG 2092. GAMS [GAM (Nissui Seiyaku Co., Ltd., Tokyo) containing 0.25% L-cysteine] was used for Bacteroides thetaiotaomicron DSMZ 2079 and Blautia cocoides DSMZ 935, RCM (Oxoid) for Clostridium difficile LMG 21717 and MRSC for Parascardovia denticolles BM7/11. Bacteroides, Blautia, Clostridium and Parascardovia were cultivated at 37°C under anaerobic conditions.

Strains from stocks (stored at -80°C in broth with 20% glycerol) were plated on the surface of agar media, after 48 h incubation a single colony per strain was picked to inoculate 10 ml broth that was incubated overnight. This overnight culture was used to inoculate 1% (v/v) 10 ml broth and cultures were grown for 18 h at optimal conditions for DNA extraction.

2.3 DNA isolation and purification

Bacterial cells were harvested from 1.5 ml cultures by centrifugation at 6,000 xg for 10 min. Cells were washed once with phosphate-buffered saline (PBS) (Amresco, Ohio, USA) and centrifuged as before. Chromosomal DNA of bacterial cells was isolated using the “GenElute Bacterial Genomic DNA” kit (Sigma-Aldrich) according to the manufacturer’s recommendations. For Gram positive strains the first step was modified with the addition of lysozyme (10 mg/ml) (Merck, Darmstadt, Germany) and mutanolysin (5 U) (Sigma-Aldrich) and incubation at 37°C for 1 h. The isolated DNA was stored at -20°C. Samples were run on 1% agarose gel electrophoresis and concentration of DNA was measured in Gen5™ Teck3 Module (BioTek, Vermont, USA).

2.4 Design of degenerate PCR primers

Degenerate primers (Table 1) were designed for each priming-GTF genes, cpsD and rfbP, to detect potential EPS producing bifidobacteria by PCR amplification. Firstly, each priming-GTF from the strain B. longum subsp. longum NCC2705 was used as a template to search the priming-GTF in other bifidobacterial genomes available in the GenBank database, as described above. Then, sequences encoding the priming-GTF of several bifidobacteria
were selected from the GenBank database. For the *cpsD*, 7 gene sequences were selected:
BL0237 (accession no. NP_695447) from *B. longum* subsp. *longum* NCC2705, BAD_1389
(YP_910252) from *Bifidobacterium adolescentis* ATCC 15703, BIFANG_03177
(ZP_04448174) from *Bifidobacterium angulatum* DSM 20098, BLIG_00247 (ZP_04665465)
from *B. longum* subsp. *infantis* CCUG 52486, BIFGAL_04272 (ZP_05966491) from
*Bifidobacterium gallicum* DSM 20093, BIFPSEUDO_0324 (ZP_03742673) from
*Bifidobacterium pseudocatenulatum* DSM 20438 and BLA_0595 (YP_002469468) from
*Bifidobacterium animalis* subsp. *lactis* AD011. For the *rfbP*, 7 gene sequences were selected:
BL0249 (NP_695455) from *B. longum* subsp. *longum* NCC2705, BIFBRE_03332
(ZP_06595517) from *Bifidobacterium breve* DSM 20213, Blon_2114 (YP_002323554) from
*B. longum* subsp. *infantis* ATCC 15697, BIFADO_01843 (ZP_02029386) from *B.
adolescentis* L2-32, HMPREF9003_1455 (ZP_07696327) from *Bifidobacterium dentium*
JCVIHMP022, BIFCAT_01253 (ZP_03324464) from *Bifidobacterium catenulatum* DSM
16992 and BIFPSEUDO_0324 (ZP_03742673) from *B. pseudocatenulatum* DSM 20438. The
ClustalX algorithm was used to align the amino acid and nucleotide sequences of each
priming-GTF (Supplementary material 1) to find the most conserved part of the gene. Finally,
degenerated primers were designed in this conserved part and were synthesized by Sigma-
Aldrich.

Five forward (F) and six reverse (R) primers were designed for the *cpsD* and seven
forward and six reverse for the *rfbP* (Table 1). The possible combinations between forward
and reverse primers for each priming-GTF, yielding amplicons between 120 bp to 1.2 kb
(Table 2), were tested in the *Bifidobacterium* strains and in the 12 strains of other genera,
selected as negative controls.

Table 1. Degenerated primers designed in this work to amplify by PCR the priming-GTF
genes, *cpsD* and *rfbP*, specifically in *Bifidobacterium* genus. The nucleotide annealing
position of each primer is indicated (5' → 3') based on the nucleotide sequence of *cpsD* and
*rfbP* from *B. longum* subsp. *longum* NCC2705.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5'→ 3')</th>
<th>Annealing position (5'→ 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>cpsD_F1</td>
<td>GGCRTKTAYCAYCGBCATGT</td>
<td>271 - 290</td>
</tr>
<tr>
<td>cpsD_F2</td>
<td>CATGTKATGGGTGAHGYTAY</td>
<td>286 - 307</td>
</tr>
<tr>
<td>cpsD_F3</td>
<td>YBRTGGCSWTBGCATCAA</td>
<td>1022 - 1040</td>
</tr>
<tr>
<td>cpsD_F4</td>
<td>YGGCAAGTCTGGYWKATS</td>
<td>1092 - 1110</td>
</tr>
<tr>
<td>cpsD_F5</td>
<td>TWCAAGMTSAGVAYGAYCC</td>
<td>1195 - 1214</td>
</tr>
<tr>
<td>cpsD_R1</td>
<td>TTRWSDATMTWRAACGCTTG</td>
<td>1115 - 1095</td>
</tr>
<tr>
<td>cpsD_R2</td>
<td>KCGGRTCTRTBCTTSAAKCTT</td>
<td>1216 - 1197</td>
</tr>
<tr>
<td>cpsD_R3</td>
<td>ARGAACTGSGGVARYTACCT</td>
<td>1277 - 1258</td>
</tr>
<tr>
<td>cpsD_R4</td>
<td>CRGTRAIRRCRGGCTTVAC</td>
<td>1393 - 1375</td>
</tr>
<tr>
<td>cpsD_R5</td>
<td>GTRATRCCRGGCTTVAC</td>
<td>1391 - 1371</td>
</tr>
<tr>
<td>cpsD_R6</td>
<td>CCDDGANACYTGCGCAVGGRC</td>
<td>1412 - 1391</td>
</tr>
<tr>
<td>rfbP_F1</td>
<td>TYGVMCGTTGGMWKATGCG</td>
<td>620 - 638</td>
</tr>
<tr>
<td>rfbP_F2</td>
<td>TYGGVTCWYCSGAAGYATC</td>
<td>695 - 714</td>
</tr>
<tr>
<td>rfbP_F3</td>
<td>CTSAAGGTGWTKCCRTSAA</td>
<td>871 - 890</td>
</tr>
<tr>
<td>rfbP_F4</td>
<td>TCBYTGGCIGAYKTSACGG</td>
<td>1030 - 1049</td>
</tr>
<tr>
<td>rfbP_F5</td>
<td>ATGTAYAARTTCGCTCATG</td>
<td>1285 - 1305</td>
</tr>
<tr>
<td>rfbP_F6</td>
<td>TTCAARMTSAARGAYGATCC</td>
<td>1372 - 1391</td>
</tr>
<tr>
<td>rfbP_F7</td>
<td>GARKTYCCARTTCTWYAA</td>
<td>1438 - 1457</td>
</tr>
<tr>
<td>rfbP_R1</td>
<td>TTSARYGMMWACACCTTSAG</td>
<td>890 - 871</td>
</tr>
<tr>
<td>rfbP_R2</td>
<td>GTTTACGRATRAARYKACC</td>
<td>1427 - 1408</td>
</tr>
</tbody>
</table>
Amersham Bioscience, The annealing...

181
180
179
178
177
176
175
174
173
172
171
170
169
168
167
166
165
164
163

2.5 PCR amplification

After the optimization of PCR conditions, the final reaction mixture and amplification steps were: 1 µl DNA, 0.2 µM of each primer, 200 µM dNTPs (Amersham Bioscience, Upsala, Sweden), and 2.5 U Taq DNA-polymerase (Eppendorf, Hamburg, Germany) in a 25 µl reaction volume. PCR reactions were performed in duplicate in an UnoCycler thermal cycler (VWR International Eurolab S.L., Barcelona, Spain). The PCR thermal conditions were an initial incubation step of 95°C for 5 min, 30 amplification cycles of a denaturation step at 95°C for 1 min, annealing step at 50-60°C (depending on the primer) for 50 s and elongation step of 68°C for 1-1.3 min, and a final elongation step of 68°C for 10 min. PCR products were stored at 4°C and visualized under UV after electrophoresis in 1-1.5% agarose gels and staining with ethidium bromide.

Table 2. Possible combinations between forward and reverse primers used to amplify by PCR each priming-GTF, with expected amplicons between 120 bp to 1.2 kb. The annealing temperature and the elongation time, as well as amplicon size (base pair) for each combination are indicated. The grey shadow indicates pair combinations selected for further checking steps and those shadowed in dark-grey were the pair combinations finally used for the genetic screening in a collection of 118 bifidobacteria. NP indicates not performed.

<table>
<thead>
<tr>
<th>Primers for priming-GTF cpsD</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cpsD_R1</td>
</tr>
<tr>
<td>cpsD_F1</td>
<td>56°C, 60 s (824 bp)</td>
</tr>
<tr>
<td>cpsD_F2</td>
<td>56°C, 60 s (809 bp)</td>
</tr>
<tr>
<td>cpsD_F3</td>
<td>NP</td>
</tr>
<tr>
<td>cpsD_F4</td>
<td>NP</td>
</tr>
<tr>
<td>cpsD_F5</td>
<td>NP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primers for priming-GTF rfbP</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rfbP_R1</td>
</tr>
<tr>
<td>rfbP_F1</td>
<td>56°C, 60 s (251 bp)</td>
</tr>
<tr>
<td>rfbP_F2</td>
<td>56°C, 60 s (177 bp)</td>
</tr>
<tr>
<td>rfbP_F3</td>
<td>NP</td>
</tr>
<tr>
<td>primer</td>
<td>50°C, 60 s (159 bp)</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------</td>
</tr>
<tr>
<td>rfbP_F4</td>
<td>60°C, 60 s (537 bp)</td>
</tr>
<tr>
<td>rfbP_F5</td>
<td>60°C, 60 s (415 bp)</td>
</tr>
<tr>
<td>rfbP_F6</td>
<td>50°C, 60 s (501 bp)</td>
</tr>
<tr>
<td>rfbP_F7</td>
<td>50°C, 60 s (567 bp)</td>
</tr>
</tbody>
</table>

2.6 DNA sequencing and analysis

PCR products of the required size were sent to Macrogen (Seoul, Korea) for purification and sequencing of each DNA strand using the forward and reverse primers. Sequences obtained were processed with the free software Chromas 1.45 (Technelysium Pty Ltd., Australia) and then compared with those held in the GenBank database using the BLASTn tool.

3. Results and Discussion

3.1 In silico analysis

An in silico analysis was performed with the priming-GTF genes present in the 28 bifidobacterial complete genomes available in the GenBank database. The priming-GTF genes were detected in most of the bifidobacterial genomes analysed, reflecting the ubiquity of these genetic determinants and the wide-spread putative EPS production capability in the genus Bifidobacterium. Remarkably, some strains harboured both priming-GTF genes, cpsD and rfbP, whereas other strains from the same species only one. The phylogenetic tree performed with the amino acid sequences of the priming-GTF found in the bifidobacterial genomes showed a clear separation of the two enzymes, RfbP and CpsD, in two phylogenetic branches (Figure 1). Even this division was kept when a given strain harboured both genes in its eps cluster, such as is the case of B. animalis subsp. lactis DSM10140 and B. longum subsp. longum NCC2705. This clearly showed that there are two different genes conserved between species, and not two copies of the same gene. The homology in the amino acid sequences of the priming-GTF could be related to the presence of conserved regions involved in the interaction with the lipid carrier. Whereas, the divergences between species could be due to sequence domains involved in the recognition of specific sugar moieties (Provencher et al., 2003; Ruas-Madiedo et al., 2007).

Figure 1. Phylogenetic tree based on the amino acid sequences of the priming-GTF genes, cpsD and rfbP, detected in the bifidobacterial complete genomes available in the GenBank data base

3.2 Priming-GTF genes detection by PCR using specific degenerated primers

Twenty four degenerated primers targeted to the priming-GTF genes, cpsD and rfbP, present in bifidobacteria were designed as described above. Initially, all possible combinations (24 for cpsD and 37 for rfbP) between forward and reverse degenerated primers were tested in B. longum subsp. longum NCC2705 to check positive amplification and to standardize the PCR protocols (Table 2). Some of the primer pairs (17 out of 24 for cpsD and 30 out of 37 for rfbP) were discarded due to the absence of, or unspecific, amplification. The PCR products of
the remaining primers were sequenced, analyzed and compared to sequences held in the GenBank database to confirm the specific amplification of each priming-GTF. All amplicons displayed DNA sequences that matched 100% with the corresponding *B. longum* genes; therefore 14 primer pair combinations (7 for *cpsD* and 7 for *rfbP*) (grey and dark-grey shadow, Table 2) were selected to continue with the study. These primers were tested in one representative strain from 9 species of bifidobacteria, putatively harboring at least one of the priming-GTF genes according to the in silico analysis: *B. adolescentis* CECT 5781, *B. animalis* subsp. *lactis* A1dOxR, *B. bifidum* A8, *B. breve* BM12/11, *B. dentium* LMG 110445, *B. longum* subsp. *infantis* CECT 4551, *B. longum* subsp. *longum* NCC2705, *B. pseudocatenulatum* A102 and *B. pseudolongum* LMG 11569. Some primers were discarded due to the absence of amplification. Then the remaining primer combinations (4 for *cpsD* and 3 for *rfbP*) were tested in 12 representative bacteria, not belonging to *Bifidobacterium* genus, from food and intestinal origin, which were selected as negative controls to check the specificity for bifidobacteria of the selected primers. The primers which amplified any of the negative controls were discarded. After these procedures, only two pairs of degenerated primers for each priming-GTF gene were selected (dark-grey shadow, Table 2) as potential candidates to detect EPS-producing strains specifically in *Bifidobacterium* genus.

These 4 primer pairs were tested in another 90 bifidobacteria strains from the IPLA collection isolated from different sources (Table 3). From a total of 99 strains, a maximum of 88 amplicons were obtained (51 from *cpsD* and 37 from *rfbP*). Additionally, the PCR products of some representative strains for each species were sequenced and all of them displayed the expected DNA sequence (data not shown). The BLASTn analysis of sequences confirmed that the primers *cpsD* F1R2, *cpsD* F4R6 for *cpsD* and *rfbP* F3R4, *rfbP* F5R5 for *rfbP* amplified each priming-GTF specifically in bifidobacteria. Both primers for *cpsD* gene were able to amplify this priming-GTF in a wide range of *Bifidobacterium* species (Table 3); this result gives an idea of how spread the *cpsD* gene in *Bifidobacterium* genus is. A higher number of strains was amplified using the primer *cpsD* F4R6 (49.5%) with respect to that of *cpsD* F1R2 (40.4%); however, the *cpsD* gene of *B. pseudolongum* and *B. dentium* was exclusively detected with the primer *cpsD* F1R2. Therefore, the use of both pairs of primers will allow the successful detection of this priming-GTF in a wider bifidobacterial spectrum.

**Table 3.** Number of positive amplifications obtained for the final primers selected to detect putative EPS-producing strains in *Bifidobacterium* genera.

<table>
<thead>
<tr>
<th>Species</th>
<th>Nº strains</th>
<th><em>cpsD</em></th>
<th></th>
<th><em>rfbP</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>F1R2</td>
<td>F4R6</td>
</tr>
<tr>
<td><em>B. adolescentis</em></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td><em>B. animalis</em> subsp. <em>lactis</em></td>
<td>21</td>
<td>18</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td><em>B. bifidum</em></td>
<td>19</td>
<td>2</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td><em>B. breve</em></td>
<td>8</td>
<td>-</td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td><em>B. dentium</em></td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>B. longum</em> subsp. <em>longum</em></td>
<td>37</td>
<td>16</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td><em>B. longum</em> subsp. <em>infantis</em></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><em>B. pseudocatenulatum</em></td>
<td>9</td>
<td>-</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td><em>B. pseudolongum</em></td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>99</strong></td>
<td><strong>40</strong></td>
<td><strong>49</strong></td>
<td><strong>36</strong></td>
</tr>
</tbody>
</table>
The primers rfbP_F3R4 and rfbP_F5R5 were able to amplify the rfbP in five species, but it was detected in a lower number of species than the cpsD. The number of positive strains was also considerably reduced (36.4% and 23.2% for rfbP_F3R4 and rfbP_F5R5, respectively). With these results, the rfbP gene seems to be less wide-spread in bifidobacteria than cpsD. However, after the in silico analysis we found other species, such as B. asteroides and B. thermophilum (Figure 1), that harbor this gene, which were not screened by PCR in this study. Therefore, in light of the current study, we cannot ascertain that the priming-GTF rfbP is present in a lower percentage in bifidobacteria taxon.

No amplification was obtained for B. breve strains with the primers targeting cpsD, whereas all the strains of this species were positive with rfbP primers; therefore it seems that the only gene coding for the priming-GTF in the B. breve strains analyzed in this work is the rfbP. Moreover, the presence of rfbP is not a trait particular to the B. breve strains analyzed in this study, as B. breve UCC2003 also contains this gene (Fanning et al., 2012) and it was found in the in-silico analyses of the B. breve genomes available in the Genbank (Figure 1).

On the other hand, no amplification was obtained for the rfbP gene in the strains of B. adolescentis, B. animalis subsp. lactis, B. bifidum and B. pseudolongum. First thoughts are that this species doesn’t harbor the rfbP gene and only harbors the cpsD; but, as was shown in the in silico analysis (Figure 1), some strains from B. animalis subsp. lactis and B. bifidum S17 also harbor the rfbP gene. The absence of amplification in this species could be related to the high genetic variability of this gene in these two species (data not shown).

Then, the four primers proposed in this work could be used by PCR as a screening method to detect putative EPS-producing bifidobacteria strains. These primers are not general for all species, neither for the two priming-GTF genes. Therefore, the best approach to search for these genes in a collection of bifidobacteria will be the use of the four primer pairs.

Finally, a homology study was done to analyze the similarity between the proteins encoded by rfbP and cpsD. For this purpose, the primers cpsD_F4R6 and rfbP_F5R5 were selected as representative of each gene, since both primer pairs amplified DNA regions of a similar size in the 3’-end of the gene (coding for the C-terminal region of each protein). From these PCR sequences, 65 amino acids from the C-terminal region were selected out of the 100 amino acids obtained. The multiple alignments performed, using the ClustalW algorithm, showed that both primers amplify a similar, conserved region in each gene (Figure 2). This region, which is the most conserved part, has been described as involving a domain implicated in the interaction of the priming-GTF with the lipid carrier; the other domain is implicated in the recognition of the sugar specificity and therefore seems to be more variable (Wang et al., 1996; Ruas-Madiedo et al., 2007).

**Figure 2.** Alignment of the amino acid sequences (65 amino acids each) deduced from the PCR amplifications with the degenerated primers cpsD_F4R6 and rfbP_F5R5

The data obtained in this work supports previous analyses carried out in our group and suggests that the EPS production capability is widely distributed among bifidobacteria (Hidalgo-Cantabrana et al., 2014). Here we have confirmed that the priming-GTF genes are present in 88 strains of the 99 analyzed. In fact, the EPS production and the structure of the repeating unit of the EPS-polymer have been characterized for some of the strains used in this study: B. infantis ATCC15697 (Tone-Shimokawa et al. 1996) and B. animalis subsp. lactis A1dOxR (Leivers et al., 2011). In other strains, although the EPS structure has not been characterized, the production has been proven as in the EPS-producing B. animalis subsp. lactis E43 (Salazar et al., 2008), A1, A1dOx, PBT, Mn6, IPLA4549 and the strains of B. longum H67, H73, L55 (Salazar et al., 2008) and NB667 (Salazar et al., 2012). In these strains positive amplification was obtained for cpsD but not for rfbP. The absence of amplification of
rfbP, which is present in the eps cluster from B. animalis subsp. lactis (Hidalgo-Cantabrana et al., 2014), could be related to the nucleotide divergence of this gene compared with other bifidobacteria (Supplementary material 1). On the other hand, in the EPS-producing strains of B. pseudocatenulatum E515, A102 and H34G (Salazar et al., 2008) the rfbP gene was amplified but not cpsD, probably indicating that these strains only harbour one of the priming-GTF. The molar mass distribution and the monosaccharide ratio of the isolated EPS of these B. longum and B. pseudocatenulatum strains have also been determined (Salazar et al., 2009), proving the EPS production by these strains.

4. Conclusion

In conclusion, we have analyzed 15 Bifidobacterium strains in which EPS production has been described in previous studies, and positive amplification was obtained for all of them with the selected primers, supporting the use of degenerate primers to detect putative EPS-producing strains. The developed degenerated primers of the priming-GTF genes are useful to detect putative EPS-producing strains in bifidobacteria. This method will facilitate the screening and selection of bifidobacteria harboring genes involved in the synthesis of EPS from different origins, as a prior step to the tedious and time-consuming task of polymer isolation and purification.

Acknowledgements

This work was funded by the Spanish “Plan Nacional I+D+I” from the “Ministerio de Ciencia e Innovación” (MICINN, FEDER funds from European Union) through the project AGL2012-33278 and C. Hidalgo-Cantabrana acknowledges his FPI fellowship to MICINN.

References


403 Bifidobacterium strains of human origin, and metabolic activity of the producing bacteria in
405 Salazar, N., Ruas-Madiedo, P., Prieto, A., Calle, L. P., de los Reyes-Gavilán, C. G.,
406 2012. Characterization of Exopolysaccharides Produced by Bifidobacterium longum NB667
407 and Its Cholate-Resistant Derivative Strain IPLA B667dCo. Journal of Agricultural and Food
408 Chemistry 60:1028–1035.
409 Schell, M.A., Karmirantzou, M., Snel, B., Vilanova, D., Berger, B., Pessi, G., Zwahlen,
410 M.C., Desiere, F., Bork, P., Delley, M., Pridmore, R.D., Arigoni, F. 2002. The genome of
411 Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract. Proceedings
412 of the National Academy of Sciences of the United States of America 99:14422-14427.
413 Stingele, F., Neeser, J.R., Mollet, B., 1996. Identification and characterization of the eps
414 (exopolysaccharide) gene cluster of Streptococcus thermophilus Sfi6. Journal of Bacteriology
415 178:1680–1690.
416 Tone-Shimokawa, Y., Toida, T., Kawashima, T., 1996. Isolation and structural analysis of
417 polysaccharide containing galactofuranose from the cell walls of Bifidobacterium infantis.
419 Turroni, F., Ventura, M., Buttó, L.F., Duranti, S., O’Toole, P.W., O’Connell-
420 Motherway, M., van Sinderen, D., 2013. Molecular dialogue between the human gut
421 microbiota and the host: a Lactobacillus and Bifidobacterium perspective. Cellular and
423 van Kranenburg, R., van Swam, I. I., Marugg, J. D., Kleerebezem, M., de Vos, W. M.,
424 1997. Molecular characterization of the plasmid-encoded eps gene cluster essential for
425 exopolysaccharide biosynthesis in Lactococcus lactis. Molecular Microbiology 24:387–397.
427 The Bifidobacterium dentium Bd1 genome sequence reflects its genetic adaptation to the
428 human oral cavity. PLOS Genetics. 5:e1000785.
429 Ventura, M., Turroni, F., O’Connell-Motherway, M., MacSharry, J., van Sinderen, D.,
430 2012. Host-microbe interactions that facilitate gut colonization by commensal bifidobacteria.
431 Trends in Microbiology 20:467-476.
432 Wang, L., Liu, D., Reeves, P. R., 1996. C-terminal half of Salmonella enterica WbaP (RfbP)
433 is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen
436 DNA sequences. Journal of Computational Biology, 7 (1–2), 203–214
Figure 1. Phylogenetic tree based on the amino acid sequences of the priming-GTF genes, *cpsD* and *rfbP*, detected in the bifidobacterial complete genomes available in the GenBank data base. The annotation code of each gene is indicated between brackets. The clustering algorithm used was the Neighbour-joining method and the number associated with the branches indicates the bootstrap (statistical confidence of the phylogenetic relationships) for a total of 100 replicates. The tree was rooted using the *dnaA* of *B. animalis* subsp. *lactis* as an outgroup. Bar scale indicates the number of substitutions per site.

Figure 2. Alignment of the amino acid sequences (65 amino acids each) deduced from the PCR amplifications with the degenerated primers *cpsD* F4R6 and *rfbP* F5R5. The picture shows residues from D^{404} to Q^{568} in *cpsD* and from D^{465} to Q^{527} in *rfbP*, according to the genome annotation of *B. longum* NCC2705 (Acc. No. AE014295). For each strain, the corresponding priming-GTF gene is indicated. Dark shading indicates the amino acids that are identical in all the samples, and light shading indicates conserved amino acid residues (only two different amino acids between strains). Consensus sequence was depicted taking into consideration the identical amino acids.
**Figure 1**

- *B. adolescentis* ATCC15703 (BAD_1389)
- *B. longum* subsp. *longum* NCC2705 (BL0237)
- *B. longum* subsp. *longum* BBMN68 (BBMN68_1012)
- *B. longum* subsp. *longum* JCM1217 (BLJ_0364)
- *B. animalis* subsp. *animalis* ATCC25527 (BANAN_06765)
  - BB12 (BIF_00944)
  - V9 (BalV_1349)
  - Bi-07 (W91_1429)
  - BLC1 (BLC1_1349)
  - CNCMI 2494 (BALAC2494_01344)
    - *B. animalis* subsp. *lactis*
      - B420 (W7y_1394)
      - DSM10140 (Balat_1392)
      - B1-04 (Balac_1392)
      - AD011 (BLA_0595)
  - B. longum subsp. infantis ATCC15697 (Blon_2114)
  - B. breve ACS-071-V-Sch8b (HMPREF9228_0447)
  - B. longum subsp. *longum* NCC2205 (BL0249)
  - B. longum subsp. *infantis* BBMN68 (BBMN68_1012)
    - *B. animalis* subsp. *lactis*
      - B420 (W7y_1394)
      - BSC104 (Balat_1392)
      - AD011 (BLA_0595)
  - B. therophilum RBL67 (D805_0348)
  - *B. bifidum* S17 (BBIF_0393)
    - *B. longum* subsp. *longum* F8 (BIL_15040)
    - B. breve ACS-071-V-Sch8b (HMPREF9228_0447)
      - B. longum subsp. *longum* NCC2205 (BL0249)
      - B. longum subsp. *infantis* ATCC15697 (Blon_2114)
      - B. breve UCC2003 (Bbr_0430)
      - B. longum subsp. *infantis* 157F (BLIF_0362)
      - B. longum subsp. *longum* KACC91563 (BLNIAS_02272)
  - *B. dentium* Bd1 (BDP_1857)
    - B1-04 (Balac1371)
    - BLC1 (BLC1_1328)
    - DSM10140 (Balat_1371)
    - B420 (W7y_1374)
    - Bi-07 (W91_1409)
    - V9 (BalV_1328)
    - AD011 (BLA_0576)
    - BB12 (BIF_00983)
    - CNCMI2494 (BALAC2494_01362)

---

**Figure 1**
<table>
<thead>
<tr>
<th>Strain</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. breve LMG13208_rfbP</td>
<td>DPRVTKIGHFIRKTSDEFPQFFNMFKGDMSLVGPRPPLPEEVARYDMLYSTRLLLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. longum IF6/1_cpsD</td>
<td>DPRVTKIGHFIRKTSDEFPQFFNVFKGDMSLVGPRPPLPEEVARYDMLYSTRLLLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. breve IF23/10_rfbP</td>
<td>DPRVTKIGHFIRKTSDEFPQFFNVFKGDMSLVGPRPPLPEEVARYDMLYSTRLLLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. longum NCIMB8809_rfbP</td>
<td>DPRVTKIGHFIRKTSDEFPQFFNVFKGDMSLVGPRPPLPEEVARYDMLYSTRLLLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. longum NCC2705_cpsD</td>
<td>DPRVTKIGHFIRKTSDEFPQFFNVFKGDMSLVGPRPPLPEEVARYDMLYSTRLLLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. longum NCC2705_rfbP</td>
<td>DPRVTKIGHFIRKTSDEFPQFFNVFKGDMSLVGPRPPLPEEVARYDMLYSTRLLLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. bifidum D119_cpsD</td>
<td>DPRVTRIGHFIRKTSIDEIPQFFNVFKGDMSLVGPRPPLPEEVARYGTLYSARLLLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. bifidum A8_cpsD</td>
<td>DPRVTRIGHFIRKTSIDEIPQFFNVFKGDMSLVGPRPPLPEEVARYGTLYSARLLLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. pseudocatenolatum IF12/5_cpsD</td>
<td>DPRITKVGHFIRKFSIDEFPQFLNVKGDMSVVGPRPPLPEEVARYGNRLYATRMLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. dentium LMG11045_rfbP</td>
<td>DPRITKVGHFIRKFSIDEFPQFLNVKGDMSVVGPRPPLPEEVARYGNRLYATRMLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. longum NB667_cpsD</td>
<td>DPRITKVGHFIRRFSIDELPQFLNVWMGDMSVVGPRPPLPEEYARNRLYATRMLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. longum BM18/6_cpsD</td>
<td>DPRITKVGHFIRRFSIDELPQFLNVWMGDMSVVGPRPPLPEEYARNRLYATRMLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. longum IF6/1_rfbP</td>
<td>DPRITKVGHFIRRFSIDELPQFLNVWMGDMSVVGPRPPLPEEYARNRLYATRMLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. longum D12_cpsD</td>
<td>DPRITKVGHFIRRFSIDELPQFLNVWMGDMSVVGPRPPLPEEYARNRLYATRMLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. adolescentis CECT5781_cpsD</td>
<td>DPRVTTVGRFIRKFSIDEFPQFLNVLRGDMSVVGPRPPLPEEYARNRLYATRMLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. breve NCIMB8807_rfbP</td>
<td>DPRVTTVGRFIRKFSIDEFPQFLNVLRGDMSVVGPRPPLPEEYARNRLYATRMLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. animalis subsp. lactis A1dOxR_cpsD</td>
<td>DPRITKVGFIRKFSIDEFPQFFDVFMGSMSVVGPRPQRFEVAEYDPVYATRMLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. animalis IF20/1_cpsD</td>
<td>DPRITKVGFIRKFSIDEFPQFFDVFMGSMSVVGPRPQRFEVAEYDPVYATRMLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. animalis subsp. lactis A1_cpsD</td>
<td>DPRITKVGFIRKFSIDEFPQFFDVFMGSMSVVGPRPQRFEVAEYDPVYATRMLKVPGITGPWQ</td>
</tr>
<tr>
<td>B. animalis subsp. lactis LMG18314_cpsD</td>
<td>DPRITKVGFIRKFSIDEFPQFFDVFMGSMSVVGPRPQRFEVAEYDPVYATRMLKVPGITGPWQ</td>
</tr>
</tbody>
</table>

**Figure 2**

Consensus: DPR T G FIR S DE PQF G MS VGPRP E Y Y LVKPGITGPWQ