Differentiation of *Bifidobacterium* species using partial RNA polymerase \(\beta\)-subunit (*rpoB*) gene sequences

Byoung Jun Kim, Hee-Youn Kim, Yeo-Jun Yun, Bum-Joon Kim and Yoon-Hoh Kook

Department of Microbiology, Seoul National University College of Medicine, Seoul 110-799, Republic of Korea

Partial RNA polymerase \(\beta\)-subunit gene (*rpoB*) sequences (315 bp) were determined and used to differentiate the type strains of 23 species of the genus *Bifidobacterium*. The sequences were compared with those of the partial *hsp60* (604 bp) and 16S rRNA genes (1475 or 1495 bp). The *rpoB* gene sequences showed nucleotide sequence similarities ranging from 84.1 % to 99.0 %, while the similarities of the *hsp60* sequences ranged from 78.5 % to 99.7 % and the 16S rRNA gene sequence similarities ranged from 89.4 % to 99.2 %. The phylogenetic trees constructed from the sequences of these three genes showed similar clustering patterns, with the exception of several species. The *Bifidobacterium catenulatum–Bifidobacterium pseudocatenulatum, Bifidobacterium pseudolongum* subsp. *pseudolongum–Bifidobacterium pseudolongum* subsp. *globosum* and *Bifidobacterium gallinarum–Bifidobacterium pullorum–Bifidobacterium saeculare* groups were more clearly differentiated in the partial *rpoB* and *hsp60* gene sequence trees than they were in the 16S rRNA gene tree. Based on sequence similarities and tree topologies, the newly determined *rpoB* gene sequences are suitable molecular markers for the differentiation of species of the genus *Bifidobacterium* and support various other molecular tools used to determine the relationships among species of this genus.

Bifidobacteria are Gram-positive, non-spore-forming, non-motile, non-gas-producing, anaerobic, bifid-shaped rods that colonize the gastrointestinal tracts of humans, as well as other mammals and some insects. There has been increased interest in these organisms recently due to their beneficial effects on gastrointestinal health. Indeed, they have been used as probiotics and evaluated for their ability to reduce several gastrointestinal tract diseases and protect against pathogens (Guarner & Malagelada, 2003; Mitsuoka & Kaneuchi, 1977; Ventura et al., 2008).

*Bacillus bifidus* was first isolated from a healthy breast-fed infant by Tissier in 1899. The genus *Bifidobacterium* was later designated as an independent taxon by Orla-Jensen (1924) (Ventura et al., 2008), after which many species belonging to this genus were identified and classified. Traditionally, phenotypic analyses based on morphology and the patterns of various carbohydrate fermentation tests have been used to differentiate species of the genus *Bifidobacterium*. However, it has been suggested that these classification criteria may be unclear and may provide questionable results (Mitsuoka & Kaneuchi, 1977). Therefore, improved methods of identification and differentiation are necessary. Accordingly, various molecular tools that are currently available may be useful for the reliable identification of such organisms.

The recent rapid developments in molecular biology have resulted in the use of various new methods for the identification and differentiation of species of the genus *Bifidobacterium*. These include PCR-linked methods such as nucleotide-dependent phylogenetic analysis targeting housekeeping genes (Jian et al., 2001; Ventura & Zink, 2003; Ventura et al., 2004a, 2007), as well as the 16S rRNA gene (Leblond-Bourget et al., 1996; Miyake et al., 1998), and other forms of molecular typing such as pulsed-field gel electrophoresis (PFGE) (Roy et al., 1996), automated ribotyping (Sakata et al., 2006), random amplified polymorphic DNA (RAPD) assays (Vincent et al., 1998), enterobacterial repetitive intergenic consensus sequences (ERIC) PCR (Ventura et al., 2003) and PCR restriction length polymorphism analysis (Ventura & Zink, 2003). Additionally, whole genome sequencing of some species of...
the genus *Bifidobacterium* has provided useful information (Kim *et al*., 2009; Lee *et al*., 2008; Schell *et al*., 2002; Sela *et al*., 2008).

Although these molecular approaches are useful for the detection, identification and characterization of *Bifidobacterium* species, they can be somewhat laborious and time-consuming (Ventura *et al*., 2004b). In addition, even though analysis of the 16S rRNA gene is widely used to identify and differentiate various bacterial species, it is limited by its highly conserved nature (Fox *et al*., 1992). Moreover, the presence of multiple copies of the 16S rRNA gene may affect the phylogenetic information for *Bifidobacterium* species (Bourget *et al*., 1993; Satokari *et al*., 2001). Thus, it has been recommended that analyses be conducted using multiple gene sequences (Palys *et al*., 1997) to evaluate precise bacterial genetic relationships.

Recently, a multilocus approach that employed housekeeping genes such as *clpC*, *dnaA*, *dnaG*, *dnaJ1*, *purF*, *rpoC* and *xfp* for an evolutionary study of the genus *Bifidobacterium* was reported (Ventura *et al*., 2006). In addition, useful gene sequence information regarding other bacteria is available in databases such as GenBank (Holmes *et al*., 1999; Ventura *et al*., 2006).

The housekeeping *hsp60* gene encodes a heat-shock protein (about 60 kDa), which plays a crucial role in stress responses, especially those due to changes in temperature. The *hsp60* gene has been used to identify and infer phylogenetic relationships among species of the genus *Bifidobacterium* (De Dea Lindner *et al*., 2007; Jian *et al*., 2001) in addition to the 16S rRNA, *hsp20*, *tsf* and *recA* genes and the internal transcribed spacer (ITS) (Leblond-Bourget *et al*., 1996; Miyake *et al*., 1998; Ventura & Zink, 2003; Ventura *et al*., 2004a, 2007).

The *rpoB* gene, which encodes the RNA polymerase β-subunit, is associated with the rifampicin resistance region of *Mycobacterium tuberculosis* (Kim *et al*., 1999) and *Escherichia coli* (Severinov *et al*., 1996). Recently, the *rpoB* gene has been shown to be related to the rifamycin resistance of *Bifidobacterium* species (Vitali *et al*., 2008). As a molecular marker, the *rpoB* gene sequence has been successfully used to identify and determine the phylogenetic relationships of species within the genera *Mycobacterium* (Kim *et al*., 1999), *Borrelia* (Lee *et al*., 2000), *Bartonella* (Renesto *et al*., 2000), *Bacteroides* (Ko *et al*., 2007) and *Legionella* (Ko *et al*., 2002). In addition, it has been reported that the *rpoB* gene is useful for determining the relationships among various bacterial species (Adekambi *et al*., 2009; Dahllof *et al*., 2000). However, *rpoB* gene sequences that might be useful for the differentiation of *Bifidobacterium* species by means of a single marker gene or as a target gene for multilocus sequence analysis (MLSA) have not been reported to date.

In this study, we used partial *rpoB* gene sequences to differentiate species of the genus *Bifidobacterium* and then compared the results with those conducted using the 16S rRNA and partial *hsp60* genes, which have been widely used previously.

Twenty-three type strains of species of the genus *Bifidobacterium* were used in this study. The strains were obtained from the Korean Collection for Type Cultures (KCTC) (Table 1). All *Bifidobacterium* strains were cultured anaerobically in fluid thioglycollate broth medium (Difco) at 37 °C for 2 days, after which they were subcultured on blood agar plates in an anaerobic jar at 37 °C under 5 % CO₂ for 2–3 days. Total DNA was extracted from each cultured *Bifidobacterium* strain using the bead beater–phenol extraction method (Kim *et al*., 1999; Yun *et al*., 2005). The DNA was then used as a template for the PCR.

The partial *rpoB*, *hsp60* and 16S rRNA gene sequences of the *Bifidobacterium* strains were amplified by PCR using the primers listed in Table 2. Template DNA (50 ng µl⁻¹) and each primer pair (20 pmol) were added to PCR tubes (containing AccuPower PCR PreMix; Bioneer) and PCR was conducted by subjecting the samples to 5 min at 95 °C, followed by 40 cycles of 95 °C for 30 s, 58–60 °C for 30 s, and 72 °C for 1 min; and a final extension at 72 °C for 5 min, which was performed in a model 9700 Thermocycler (Perkin-Elmer Cetus). The PCR products were detected and then purified using QIAEX II gel extraction kits (Qiagen) for direct sequencing. The purified 16S rRNA genes were transformed into TOP10 *E. coli* cells (Invitrogen Corp.) using a TOPO TA cloning kit (Invitrogen Corp.) for sequencing.

Sequencing reactions were performed using a MJ Research PTC-225 Peltier Thermal Cycler and ABI PRISM BigDye Terminator Cycle Sequencing kits with AmpliTaq DNA polymerase (FS enzyme; Applied Biosystems) following the manufacturers’ protocols.

The determined *rpoB* (315 bp), *hsp60* (604 bp) and 16S rRNA gene sequences (1475 or 1495 bp) of the 23 species of the genus *Bifidobacterium* were aligned using CLUSTAL W. The pairwise distance values of each gene were calculated using MEGA4 software (Tamura *et al*., 2007) and then plotted to determine the variances among genes. The nucleotide sequence similarities between the *Bifidobacterium* species and the deduced amino acid sequences were determined using the MEGALIGN software package (DNASTAR). Phylogenetic trees were inferred from the *rpoB*, *hsp60* and 16S rRNA gene sequences using the neighbour-joining (NJ) method and the bootstrap values were calculated from 1000 replicates using MEGA4 software. In addition, a concatenated tree was constructed using the combined *rpoB*, *hsp60*, and 16S rRNA gene sequences.

The incongruence length difference test (ILD; partition homogeneity test in the PAUP* software package) was conducted (Cunningham, 1997; Swoford, 2003) to determine whether the three gene trees were correlated. The maximum-likelihood (ML) tree produced for each gene was constructed using the HK85 + Γ substitution model. The log-likelihood values were calculated from three datasets.
under each ML topology constraint using the Shimodaira-Hasegawa (SH) test (Shimodaira & Hasegawa, 1999). The test was performed in the PAUP* program (Swofford, 2003) with 1000 bootstrap replicates and the full optimization (one-tailed test) option.

The newly determined partial rpoB sequences for the type strains of the 23 Bifidobacterium species used in this study were deposited in the GenBank database under accession numbers EU780709–EU780731 (Table 1).

The partial rpoB sequences (315 bp) of the 23 Bifidobacterium species were compared. The G+C contents determined for these nucleotide sequences were higher (61.27 % to 69.52 %) than those of whole genome sequences of Bifidobacterium longum and Bifidobacterium adolescentis (59 % to 60 %) (GenBank accession nos AE014295, CP000605 and AP009256) (Lim et al., 2003; Schell et al., 2002).

The rpoB nucleotide sequence similarities were determined using the MEGALIGN software in the DNASTAR package and were then compared with those of the hsp60 and 16S rRNA gene sequences. There were no gaps in the sequenced rpoB and hsp60 regions, whereas gaps resulting in variations in sequence length were observed in the 16S rRNA gene sequences. The divergences within the rpoB, hsp60, and 16S rRNA genes were 1–17.4 %, 0.6–8.4 % and 0.3–23.5 %, respectively. In addition, the inter-species gene sequence similarities for the rpoB gene ranged from 84.1 % to 99.0 %.

Table 1. Reference strains of the Bifidobacterium species used in this study

All gene sequences were determined in this study.

<table>
<thead>
<tr>
<th>Species</th>
<th>Strain no.</th>
<th>Origin</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. animalis subsp. animalis</td>
<td>KCTC 3219&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Rat faeces</td>
<td>EU780709</td>
</tr>
<tr>
<td>B. bifidum</td>
<td>KCTC 3202&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Stool of breast-fed infant</td>
<td>EU780710</td>
</tr>
<tr>
<td>B. breve</td>
<td>KCTC 3227&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Rumens of cattle</td>
<td>EU780711</td>
</tr>
<tr>
<td>B. catenulatum</td>
<td>KCTC 3220&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Intestine of infant</td>
<td>EU780712</td>
</tr>
<tr>
<td>B. choerinum</td>
<td>KCTC 3221&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Human faeces</td>
<td>EU780713</td>
</tr>
<tr>
<td>B. cuniculi</td>
<td>KCTC 3275&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Piglet faeces</td>
<td>EU780714</td>
</tr>
<tr>
<td>B. dentium</td>
<td>KCTC 3276&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Rabbit faeces</td>
<td>EU780715</td>
</tr>
<tr>
<td>B. gallinarum</td>
<td>KCTC 3222&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Dental caries</td>
<td>EU780716</td>
</tr>
<tr>
<td>B. gallicum</td>
<td>KCTC 3277&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Intestine of adult</td>
<td>EU780717</td>
</tr>
<tr>
<td>B. gallinarum</td>
<td>KCTC 3235&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Chicken caecum</td>
<td>EU780718</td>
</tr>
<tr>
<td>B. indicum</td>
<td>KCTC 3230&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Hindgut of honeybee, Apis cerana indica</td>
<td>EU780719</td>
</tr>
<tr>
<td>B. longum subsp. longum</td>
<td>KCTC 3128&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Intestine of adult</td>
<td>EU780720</td>
</tr>
<tr>
<td>B. magnus</td>
<td>KCTC 3228&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Rabbit faeces</td>
<td>EU780721</td>
</tr>
<tr>
<td>B. merycicum</td>
<td>KCTC 3369&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Bovine rumen</td>
<td>EU780731</td>
</tr>
<tr>
<td>B. minimum</td>
<td>KCTC 3273&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Sewage</td>
<td>EU780722</td>
</tr>
<tr>
<td>B. pseudocatenulatum</td>
<td>KCTC 3223&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Faeces of infant</td>
<td>EU780723</td>
</tr>
<tr>
<td>B. pseudolongum subsp. globosum</td>
<td>KCTC 3234&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Bovine rumen</td>
<td>EU780724</td>
</tr>
<tr>
<td>B. pseudolongum subsp. pseudolongum</td>
<td>KCTC 3224&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Swine faeces</td>
<td>EU780725</td>
</tr>
<tr>
<td>B. pullorum</td>
<td>KCTC 3274&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Faeces of chicken</td>
<td>EU780726</td>
</tr>
<tr>
<td>B. ruminantium</td>
<td>KCTC 3425&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Bovine rumen</td>
<td>EU780727</td>
</tr>
<tr>
<td>B. sacculare</td>
<td>KCTC 3427&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Rabbit faeces</td>
<td>EU780728</td>
</tr>
<tr>
<td>B. subtilis</td>
<td>KCTC 3272&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Sewage</td>
<td>EU780729</td>
</tr>
<tr>
<td>B. thermophilum</td>
<td>KCTC 3225&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Swine faeces</td>
<td>EU780730</td>
</tr>
</tbody>
</table>

Table 2. PCR primers used in this study

<table>
<thead>
<tr>
<th>Target</th>
<th>Primer</th>
<th>Oligonucleotide sequence</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>rpoB</td>
<td>BifF</td>
<td>5’-TCGATCGGGGCACTACGG-3’ (BifF–BifR) 351 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BifR</td>
<td>5’-CGACCACCTTCGGCAACCCG-3’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BifR2</td>
<td>5’-ACAAGATCAACCGCAACGC-3’ (BifF–BifR2) 523 bp</td>
<td></td>
</tr>
<tr>
<td>hsp60</td>
<td>HspF3</td>
<td>5’-ATCGCCAAAGGAGATCGAGCT-3’ (HspF3–HspR4) 644 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HspR4</td>
<td>5’-AAGGTGCCGGATCTTGTT-3’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HspBF3</td>
<td>5’-CCGACCATACCAACAGATGG-3’ (HspBF3–HspBR4) 680 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HspBR4</td>
<td>5’-CAGAACTTGAACGTGCCACG-3’</td>
<td></td>
</tr>
<tr>
<td>16S rRNA</td>
<td>Bif 285</td>
<td>5’-GGAGGTTGATTTGCCACAGCTCAG-3’ (Bif 285–261) 1516–1536 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>261</td>
<td>5’-AAGGAGCTGATCTCCAGCCGCA-3’</td>
<td></td>
</tr>
</tbody>
</table>
The rpoB-based sequence similarity range fell between that of the partial hsp60 (from 78.5 % to 99.7 %) and the 16S rRNA genes (from 89.4 % to 99.2 %). When the rpoB gene sequences were evaluated, the lowest similarities were between Bifidobacterium indicum and Bifidobacterium magnun (84.1 %), while the highest similarities were observed between Bifidobacterium gallinarum and Bifidobacterium saeculare (99.0 %). When the 16S rRNA sequences were evaluated, the lowest similarities were between Bifidobacterium animalis and Bifidobacterium bifidum and Bifidobacterium ruminantium (89.4 %) and the highest similarity was found between B. gallinarum and B. saeculare (99.2 %). Evaluation of the hsp60 sequences revealed that the lowest and highest similarities were between B. magnum and Bifidobacterium minimum (78.5 %) and B. gallinarum and B. saeculare (99.7 %), respectively. For all three cases, the highest gene sequence similarity was observed between B. gallinarum and B. saeculare. In addition, the similarities between these two species obtained using the partial rpoB gene sequences were relatively lower than the similarities obtained using the other genes.

The pairwise distance values of each gene were calculated using the MEGA4 program and then compared (see Supplementary Fig. S1 in IJSEM Online). The plotted graphs showed that the distance values for the hsp60 and rpoB genes were higher than that of the 16S rRNA gene (see Supplementary Fig. S1a and S1b). However, when the hsp60 and rpoB genes were compared, the results revealed that the hsp60 gene sequences were more diverse than the rpoB sequences (Supplementary Fig. S1c and S1d).

A total of 104 amino acids (R452 to F555) were deduced from the partial rpoB gene sequences. Only 1–3 substitutions were observed between the Bifidobacterium species analysed. The rpoB amino acid sequence similarities were 97.1–100 %, whereas those of the 201-amino-acid sequences (E or D62 to L262) deduced from the hsp60 gene were 88.6–100 %.

Phylogenetic trees were constructed based on the partial rpoB, hsp60 and 16S rRNA gene sequences of the Bifidobacterium species using the NJ method in conjunction with bootstrap analysis using the MEGA4 program (Fig. 1). The trees were rooted using the gene sequence of Mycobacterium tuberculosis H37Rv T, which is well known as the causative agent of tuberculosis. The topology of the rpoB gene tree was similar to the topologies of the hsp60 and 16S rRNA gene trees and to previously published reports (Jian et al., 2001; Sakata et al., 2006; Ventura et al., 2008) (Fig. 1). Each tree had similar clusters; however, several species were positioned at different branches in each tree. In the hsp60 and 16S rRNA gene trees, B. minimum and B. indicum were clustered together (Fig. 1a, c), but they were not clustered together in the rpoB tree (Fig. 1b). In the 16S rRNA gene tree, B. animalis and Bifidobacterium choerinum were clustered together, but in the rpoB and hsp60 trees these organisms formed independent branches in the same group (Bifidobacterium pseudolongum group) (Ventura et al., 2008). In the 16S rRNA gene tree, Bifidobacterium mericycum and B. ruminantum were clustered together, but this was not the case in the rpoB and hsp60 gene trees. In the rpoB and hsp60 trees, B. gallicum and B. magnun were clustered together; however, they formed independent branches in the same group in the 16S rRNA gene tree (B. pseudolongum group) (Ventura et al., 2008).

Although B. pseudolongum subsp. pseudolongum—B. pseudolongum subsp. globosum, B. pseudocatenulatum—Bifidobacterium catenulatum and B. saeculare—B. gallinarum—Bifidobacterium pullorum appeared as closely related groups in the 16S rRNA gene tree, these species were resolved by longer branches in the rpoB and hsp60 gene trees (Fig. 1). The combined sequences of the three genes were also used to construct a concatenated tree that was better at resolving closely related species than any of the single-gene trees. In addition, each node was supported by higher bootstrap values than the trees that were based on each single gene (Fig. 1d).

The ILD test was conducted to test the concordancy of the three gene trees. The results revealed that the three datasets were not congruent (P≤0.001). The results of the SH test also demonstrated that their topologies could not be combined (P<0.05).

Despite having shorter sequences than those of the 16S rRNA genes (1475 or 1495 bp), the nucleotide sequence similarities among species when the partial rpoB gene sequences were used (from 84.1 % to 99.0 %) were slightly lower than those obtained from the 16S rRNA gene sequences (from 89.4 % to 99.2 %). In addition, the pairwise distance values obtained when the rpoB gene was used were greater than those obtained with the 16S rRNA gene sequences, indicating greater sequence variance within the rpoB gene (see Supplementary Fig. S1). These results suggested that the short rpoB sequences were more informative than the 16S rRNA gene sequences. However, the hsp60 sequences showed very broad range similarities (from 78.5 % to 99.7 %) and were more divergent than the rpoB gene sequences (see Supplementary Fig. S1). This difference in rpoB and hsp60 gene sequences was also found in their deduced amino acid sequences. Most of the base substitutions in the rpoB gene sequences were synonymous, but this was not the case in the hsp60 gene sequences. These results suggest that the hsp60 gene is under weaker selective pressure than the rpoB gene in Bifidobacterium species.

The rpoB and hsp60 gene trees showed more distinctive patterns than the 16S rRNA gene tree. This finding was especially clear among related species such as B. pseudocatenulatum—B. catenulatum, B. saeculare—B. gallinarum—B. pullorum and B. pseudolongum subsp. pseudolongum—B. pseudolongum subsp. globosum (Fig. 1), which could be separated more easily using the rpoB gene sequences than with the 16S rRNA gene sequences. In addition, the hsp60-based results showed clear differentiation between these
Fig. 1. Phylogenetic trees based on the 16S rRNA gene (a), rpoB (b), hsp60 (c), and the combined 16S rRNA, rpoB and hsp60 gene sequences (d) of 23 type strains of species of the genus *Bifidobacterium*. The trees were constructed using the NJ method. The bootstrap values were calculated from 1000 replications; those with values of <50% are not shown. *Mycobacterium tuberculosis* H37Rv was used as an outgroup. Bars, number of substitutions per nucleotide position.
closely related groups, which was in agreement with the results of a previous report (Jian et al., 2001). These closely related species were reported to have almost the same ecological niches (Table 1) (Gavini et al., 1991) and to have highly similar 16S rRNA gene sequences (98.1 %, 98.7–99.2 % and 98.8 %, respectively). Additionally, there have been some efforts to differentiate species within the B. pseudocatenulatum–B. catenulatum and B. saeculare–B. gallinarum–B. pullorum groups (Jian et al., 2001; Requena et al., 2002; Ventura et al., 2003). In the present study, these closely related species were clearly differentiated using the rpoB gene, which supported the finding that the rpoB gene can be used to infer the genetic relationships of Bifidobacterium species.

In the 16S rRNA gene tree, B. minimum–B. indicum, B. animalis–B. choerinum and B. merycicum–B. ruminantium were clustered together. However, in the rpoB, hsp60 and concatenated trees, they were not grouped or were grouped with low bootstrap values (Fig. 1b, c and d). The B. gallicum–B. magnum group was supported with a high bootstrap value (>80), except in the 16S rRNA gene tree (Fig. 1). These results suggest that the clustering patterns in the 16S rRNA gene tree may not reflect the genetic relationships of members of the genus Bifidobacterium well. With these few exceptions, the rpoB gene tree showed similar clustering patterns to the partial hsp60 and 16S rRNA gene trees (Fig. 1) and to trees generated using previously reported data (Jian et al., 2001; Sakata et al., 2006; Ventura et al., 2008).

Topological differences were analysed by the ILD and SH tests. The results showed that the trees based on the three gene sequences were statistically incongruent. These incongruencies may be due to the different evolution rates for each gene. Additionally, these incongruencies were supported by different variances within each gene (Supplementary Fig. S1) and the similarities of their amino acid sequences. Moreover, parasexual events such as horizontal gene transfer (HGT), may have contributed to the phylogenetic incongruence. It is likely that there have been few HGT events in the rpoB region due to its role as a housekeeping gene. However, because there is evidence of HGT events in the rpoB region of other bacteria (Iyer et al., 2004) and prophage-like elements have been detected in several Bifidobacterium species (Ventura et al., 2005), HGT events in the rpoB region could also affect the phylogenetic relationships among members of the genus Bifidobacterium.

Although single gene analysis using the rpoB gene can provide useful resolving power to differentiate species of the genus Bifidobacterium, there may be limitations on its usefulness due to the different topologies between the 16S rRNA and hsp60 gene sequence trees. Indeed, these differences suggest that none of these trees can precisely explain the genetic relationships among Bifidobacterium species. Thus, a multi-gene-based analysis of housekeeping genes (at least five loci) (Stackebrandt et al., 2002) may be necessary to provide more detailed molecular analysis and to reveal the precise relationships between Bifidobacterium species. In such a situation, the rpoB gene would be a good candidate for use as a housekeeping gene for multilocus sequence typing, which is a portable, easily accessible and highly resolving nucleotide-based molecular typing system. In this study, the rpoB, hsp60, and 16S rRNA gene sequences were combined to infer the phylogenetic relationships of species of the genus Bifidobacterium (Fig. 1d). Although only three gene loci were used to construct the concatenated tree, the resolution of this tree was better than any other single gene tree (Fig. 1a, b and c).

In conclusion, the newly determined rpoB sequences in this study provided sufficient information to provide good discriminatory power for the differentiation of members of the genus Bifidobacterium. Accordingly, a new rpoB gene sequence database may be used for gene-based analysis alone or as a supplement for other molecular methods such as 16S rRNA gene-based analysis. However, we suggest that, due to the incongruent topological information provided by the rpoB, hsp60 and 16S rRNA gene trees, multi-gene based analysis may provide better genetic information regarding Bifidobacterium species than the use of any of these genes alone.

Acknowledgements

We would like to thank Dr Soon Gyu Hong of the Polar Research Institute (KOPRI) and Professor Kwan Soo Ko of the Asian-Pacific Research Foundation for Infectious Diseases (ARFID) and Sungkyunkwan University School of Medicine for advice on how to use the PAUP* program.

References


Differentiation of Bifidobacterium species by rpoB gene sequences


