NOTE

Comparative sequence analysis of a *recA* gene fragment brings new evidence for a change in the taxonomy of the *Lactobacillus casei* groupGiovanna E. Felis,¹ Franco Dellaglio,¹ Luca Mizzi² and Sandra Torriani^{1,3}¹ Dipartimento Scientifico e Tecnologico, Facoltà di Scienze MM. FF. NN., Università degli Studi di Verona, Strada le Grazie, 15, 37134 Verona, Italy² Dipartimento di Genetica e di Biologia dei Microrganismi, Via Celoria 26, 20133 Milano, Italy³ DOFATA, Università degli Studi di Catania, 95124 Catania, Italy

Author for correspondence: Franco Dellaglio. Tel: +39 045 8027917. Fax: +39 045 8027928. e-mail: dellaglio@sci.univr.it

The taxonomic positions of species of the *Lactobacillus casei* group have been evaluated by sequencing and phylogenetic analysis of a 277 bp *recA* gene fragment. High sequence similarity between strain ATCC 393^T, currently designated as the type strain of *L. casei*, and the type strain of *Lactobacillus zeae*, LMG 17315^T, has been established, while *L. casei* ATCC 334 and *Lactobacillus paracasei* NCDO 151^T form a single phylogenetic group. The taxonomic status of species and strains at issue is discussed.**Keywords:** *Lactobacillus casei*, *recA*, taxonomy

Lactobacillus casei and related species include phenotypically and genetically heterogeneous strains able to colonize various natural and man-made environments (Vescovo *et al.*, 1995). Many strains are currently employed as probiotics or mixed cultures for dairy products. Despite its importance for the food industry, the taxonomy of the *L. casei* group is still unclear. Historically, the systematics of the *L. casei* group have been problematic and adjustments have been introduced based on new taxonomic methods that have gradually become available. Initially, Hansen & Lessel (1971) designated strain ATCC 393^T as the type strain of *L. casei* subsp. *casei* on the basis of a few phenotypic traits. However, Dellaglio *et al.* (1975) found little homology between authentic strains belonging to the subspecies of *L. casei* and ATCC 393^T in DNA–DNA hybridization experiments. Based on such evidence, Collins *et al.* (1989) reclassified the majority of reference strains of *L. casei* subsp. *casei*, *L. casei* subsp. *alactosus* and *L. casei* subsp. *pseudoplatantarum* as *Lactobacillus paracasei*. Subsequent DNA–DNA hybridization experiments and numerical analysis of total soluble cell protein patterns led to the proposal of designating strain ATCC 334 as the neotype for the species *L. casei*, with rejection of the name *L. paracasei* (Dellaglio *et al.*, 1991). The proposal was denied (Wayne, 1994). Dicks *et al.* (1996) provided new indications in favour of the reclassification of strain

ATCC 334 as the neotype of the species *L. casei*; moreover, they proposed the reclassification of *L. casei* subsp. *casei* ATCC 393^T and *Lactobacillus rhamnosus* ATCC 15820 as strains of *Lactobacillus zeae*. Mori *et al.* (1997) performed a comprehensive study of the phylogenetic relationships existing between *L. casei*, *L. zeae*, *L. rhamnosus* and *L. paracasei* on the basis of a comparative sequence analysis of 16S rRNA genes (16S rDNA). High similarity was found between the sequences analysed, with a maximum of 19 variable nucleotide positions out of more than 1520. Thirteen nucleotides at different positions were indicated as sequence signatures and the results supported the classification of *L. paracasei* and *L. casei* ATCC 334 in one species. About 1 % maximum divergence and only two signatures differed between *L. casei* ATCC 393^T and *L. zeae* LMG 17315^T, fewer than the seven different signatures found between *L. casei* ATCC 393^T and *L. casei* ATCC 334 that were not a sufficiently strong argument to show clear interspecific differentiation between *L. casei* ATCC 393^T and *L. casei* ATCC 334.

Nour (1998) proposed that 16S–23S and 23S–5S intergenic spacer region sequences might represent a useful alternative to 16S and 23S rDNA for establishing the phylogeny of closely related bacterial species, because of the higher sequence and fragment length polymorphism. Comparative analysis of the 23S–5S rRNA intergenic spacer regions of *L. casei*-related strains was performed by Chen *et al.* (2000). Their observations supported the request of Dellaglio *et al.* (1991) to replace strain ATCC 393^T with strain ATCC 334 as the neotype of *L. casei* subsp. *casei*.

The EMBL/GenBank/DBJ accession numbers for the partial *recA* sequences of strains LMG 9191^T, LMG 17315^T, LMG 6400^T, ATCC 393^T, ATCC 334, LMG 9438, NCDO 151^T and DF1 are AJ286115–AJ286117, AJ286121–AJ286124 and AJ297821, respectively.

However, the nomenclature of *L. casei*-group species is still ambiguous, as demonstrated by recently published articles that indicate either ATCC 334 (Yoon *et al.*, 2000) or NCDO 161^T (= ATCC 393^T) (Swezey *et al.*, 2000) as the *L. casei* type strain.

The *recA* gene has been proposed as a useful marker in inferring bacterial phylogeny (Lloyd & Sharp, 1993; Eisen, 1995) and has recently been used successfully to differentiate species of some bacterial genera (Kullen *et al.*, 1997; Gruber *et al.*, 1998; Smith *et al.*, 1999; Maréchal *et al.*, 2000; van Waasbergen *et al.*, 2000). The aim of the present study was to elucidate the phylogenetic relationships between species and strains of the *L. casei* group as delineated by the *recA* gene.

The bacterial strains used in this study were *L. casei* ATCC 334, *L. casei* ATCC 393^T, *L. paracasei* subsp. *paracasei* NCDO 151^T, *L. paracasei* subsp. *paracasei* LMG 9438, *L. paracasei* subsp. *tolerans* LMG 9191^T, *L. rhamnosus* LMG 6400^T, *L. zeae* LMG 17315^T and *Lactobacillus* sp. DF1, a strain isolated from a commercial probiotic product. Strains obtained from different culture collections were checked for purity. Strain DF1 was identified biochemically as *L. casei* by the API 50 CHL kit (bioMérieux).

The strains were grown aerobically in MRS broth for 16 h at 37 °C. Genomic DNA of lactic acid bacteria was extracted following the procedure described by Marmur (1961). Three hundred ng chromosomal DNA from each strain was added to a 50 µl PCR mixture composed of 2.5 U RED *Taq* DNA polymerase (Sigma-Aldrich) in its buffer, 50 pmol each primer and 100 µM dNTPs. The PCR program was made up of initial denaturation at 94 °C for 5 min and 30 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 90 s and extension at 72 °C for 1 min. Primers used for the amplification had the following sequences: 5'-TTYATHGAYGCNGARCAAYGC-3' (forward primer, corresponding to amino acids 92–98 of the *Escherichia coli* RecA protein; Dybvig *et al.*, 1992) and 5'-CCWCCWGKWGTHGTYTCNGG-3' (reverse primer, corresponding to amino acids 206–211 of *E. coli* RecA; Duwat *et al.*, 1992). The expected size of the amplicon was 360 bp. Because of the lack of amplification with this primer pair from *L. rhamnosus* LMG 6400^T and strain DF1, we designed an internal reverse primer (5'-ACTTTYTCMCGMAWCTGATT-3') based on the sequences obtained for the other strains tested. The amplification program applied with this primer coupled with the above forward primer differed from the former program in the annealing temperature (50 °C). The expected size of the amplicon was 318 bp. Fragments of the expected sizes were purified from 2% agarose gel by using the GeneClean kit (Bio 101), ligated in pGEM-T vector (Promega) and electroporated into *E. coli* XL-1 Blue. Plasmids were extracted from white colonies by the plasmid miniprep kit (QIAGEN) and the inserted fragment length was verified by agarose gel electrophoresis after digestion with *AfaI* (Takara). Sequencing of inserts

was performed at Centro Genoma Vegetale-ENEA CR Casaccia (Rome).

The similarity of the sequences obtained to previously reported *recA* sequences was determined by a BLASTX search (Altschul *et al.*, 1997) against major molecular databases. Moreover, the presence, in the putative translations of fragments, of Asp-100, Tyr-103, Asp-144 and Ser-145 further confirmed their identity because those residues are considered functionally important (Story & Steitz, 1992). Regions of primer annealing were excluded from the analysis because point mutations were observed due to primer degeneration. Only homologous regions complete in all sequences were included in the analysis. The CLUSTAL W program (Thompson *et al.*, 1994) was used to align both nucleotide and putative amino acid sequences of the *recA* gene fragments.

In order to avoid misleading phylogenetic conclusions, different reference organisms were selected to be used as outgroups from those for which *recA* gene sequences have already been determined (Stackebrandt & Ludwig, 1994): *Acholeplasma laidlawii* (accession number M81465), *Acidothermus cellulolyticus* (AJ006705), *Anabaena variabilis* (M29680), *Arthrobacter globiformis* (AF214780), *Bacillus subtilis* (X52132), *Bifidobacterium breve* (AF094756), *Clostridium perfringens* (U61497), *Frankia alni* (AJ006707), *Geodermatophilus obscurus* (AJ006706), *Mycoplasma mycoides* (L22073), *Ruminococcus albus* (U30293), *Spirulina platensis* (U33924), *Staphylococcus aureus* (L25893), *Streptococcus pyogenes* (U21934), *Synechococcus* sp. (M29495) and *Lactococcus lactis* (M88106). Trees were constructed by calculating distance matrices with the DNADIST program in the PHYLIP software package (Felsenstein, 1993) on nucleotide sequences and PROTDIST on putative amino acid sequences with the default models. Trees were inferred from the matrices using the NEIGHBOR program in PHYLIP, which implements the neighbour-joining method of Saitou & Nei (1987). In order to improve the accuracy of phylogenetic estimation with this outgroup, we produced trees with different methods (Kim, 1993). A phylogenetic tree from gene sequences was also constructed with the DNAML (maximum-likelihood) and DNAPARS (parsimony) programs in PHYLIP (Felsenstein, 1993). Results obtained with all the tree methods were in agreement. Moreover, since it is known that the order of sequence alignment can influence the tree topology (Lake, 1991), we always selected the 'jumble' option in program menus to test different sequence orders. Finally, in order to prove the reliability of tree topologies, a bootstrap confidence analysis on 1000 replicates was performed (Felsenstein, 1985), generating a consensus tree. All the different methods tested produced the same topology of the tree. The tree obtained with *Bacillus subtilis* as the outgroup is reported in Fig. 1. The same classification scheme was obtained with nucleotide and amino acid sequences, but phylogenetic distances for the amino acid tree were shorter (data not shown), due to the

Fig. 1. Neighbour-joining tree obtained with *recA* gene sequences. Bootstrap values are reported for a total of 1000 replicates. Bootstrap values for the *L. paracasei* cluster range between 853 and 1000. Trees generated by the maximum-likelihood and parsimony methods have the same topology (not shown). Bar, 10% sequence divergence. In this tree topology, the phylogenetic distance between organisms is the sum of the horizontal segments.

masking of synonymous nucleotide substitution after translation.

Fig. 1 shows that the phylogenetic distance between *L. casei* ATCC 393^T and *L. casei* ATCC 334 is greater than that between ATCC 393^T and the type strain of *L. zeae*. In contrast, the distances between strains of *L. paracasei* and *L. casei* ATCC 334 are virtually zero. These data are strongly supported by the bootstrap values reported.

Because of the similar G + C content of the organisms considered, the differences in sequences are likely due

to real evolutionary divergence; we therefore focused our interest on *recA* nucleotide sequences.

K_{nuc} values (Hori & Osawa, 1979) calculated for *recA* sequences and K_{nuc} values for 16S rDNA as calculated by Mori *et al.* (1997) for the type strains are reported in Table 1. It can be noted that, except for the K_{nuc} values between *L. casei* ATCC 393^T and *L. zeae* LMG 17315^T, the K_{nuc} values calculated for *recA* sequences are at least 30-fold greater than those calculated for 16S rDNA sequences. From these data, it appears that intraspecific K_{nuc} values are at least one logarithmic unit smaller than those between species: for example, *L. casei* ATCC 334 and *L. paracasei* strains share K_{nuc} values between 0 and 0.021, while they are clearly distinguished from *L. rhamnosus* by K_{nuc} values between 0.265 and 0.279. Similarly, they are distinct from *L. zeae* LMG 17315^T, with values between 0.372 and 0.390. From this point of view, it is evident that *L. casei* ATCC 393^T, sharing a K_{nuc} value of 0.0358 with strain LMG 17315^T, forms an intraspecific group with *L. zeae*, while the values it shares with *L. casei* ATCC 334 and related strains (0.367–0.390) indicate an inter-specific relationship. We cannot yet propose a numerical value for *recA* homology to be used as a taxonomic discriminator among species, however, since *L. paracasei* and *L. rhamnosus*, sharing K_{nuc} values between 0.265 and 0.279, are validly considered distinct species; the same status should be accepted for *L. casei* ATCC 334 and *L. casei* ATCC 393^T, which show higher K_{nuc} values.

Strain DF1, identified biochemically as *L. casei*, is characterized by a particular position in the phylogenetic tree: it groups together with *L. zeae* LMG 17315^T and *L. casei* ATCC 393^T, but their K_{nuc} values are greater than those obtained for the strains of the *L. paracasei* group. Therefore, we cannot unequivocally assign this strain to *L. zeae*.

Table 1. K_{nuc} values for *recA* internal sequences of *Lactobacillus* strains

Values in bold are K_{nuc} values for 16S rDNA for the same species, as reported by Mori *et al.* (1997).

Strain	1	2	3	4	5	6	7	8
1. <i>L. casei</i> ATCC 393 ^T	—	0.0358	0.367	0.367	0.378	0.390	0.378	0.162
		0.0040	0.0073	0.0073	0.0073	0.0073	0.0086	
2. <i>L. zeae</i> LMG 17315 ^T		—	0.372	0.372	0.390	0.390	0.367	0.161
			0.0086	0.0086	0.0086	0.0086	0.0126	
3. <i>L. paracasei</i> subsp. <i>paracasei</i> LMG 9438			—	0	0.012	0.012	0.265	0.383
				0	0	0	0.0113	
4. <i>L. paracasei</i> subsp. <i>paracasei</i> NCDO 151 ^T				—	0.012	0.012	0.265	0.383
					0	0	0.0113	
5. <i>L. paracasei</i> subsp. <i>tolerans</i> LMG 9191 ^T					—	0.021	0.279	0.395
						0	0.0113	
6. <i>L. casei</i> ATCC 334						—	0.275	0.407
							0.0113	
7. <i>L. rhamnosus</i> LMG 6400 ^T							—	0.367
8. <i>Lactobacillus</i> sp. DF1								—

Analyses based on the *recA* gene or RecA protein were in general agreement with those performed with 16S rRNA (e.g. Eisen, 1995), thus validating the use of *recA* as a taxonomic marker; on the other hand, it is plausible that the *recA* gene, which unlike 16S rRNA encodes a protein, can accumulate a greater degree of variability at the nucleotide level without modifying the product of translation substantially. Therefore, the *recA* gene can be proposed as a reliable marker for inferring phylogenetic/taxonomic relationships between closely related species because it is not completely free of mutation; though the RecA protein has a fundamental role in the cell and mutations must not alter its function, point mutations that do not alter gene product activity are tolerated. Moreover, while the 16S rRNA is a unique analysable sequence, *recA* has a double potential: the amino acid sequence can be used to infer relationships between distant taxa without the disadvantages of rRNAs (e.g. Collins *et al.*, 1991; Fox *et al.*, 1992; Hasegawa & Hashimoto, 1993; Sogin *et al.*, 1993; Cilia *et al.*, 1996; Ueda *et al.*, 1999; Gupta, 2000) and the nucleotide sequences can be analysed to separate closely related species.

The observations on *recA* K_{nuc} values and phylogenetic inference complement the data on 23S–5S rRNA intergenic spacer regions (Chen *et al.*, 2000), 16S rDNA K_{nuc} values and sequence signatures (Mori *et al.*, 1997), DNA–DNA relatedness (Dellaglio *et al.*, 1975) and numerical analysis of total soluble cell proteins patterns (Dellaglio *et al.*, 1991) in demonstrating that *L. casei* ATCC 393^T is not representative of the *L. casei* species.

The high genetic similarity between *L. casei* ATCC 393^T and *L. zeae* LMG 17315^T indicates that they are members of the same species. Moreover, *L. casei* ATCC 334 and the type strain of the species *L. paracasei* form a single phylogenetic group.

In an accompanying Request for an Opinion (Dellaglio *et al.*, 2002), we propose that the Judicial Commission consider that the strain ATCC 393^T is scientifically unsuitable as the type strain of the species *L. casei*. For strains belonging to the same taxon as strain ATCC 334 and NCDO 151, the species name *L. casei* has to be retained on the basis of its priority over *L. paracasei*. Consequently, the latter species, phylogenetically indistinguishable from many representative strains of *L. casei*, should be rejected.

References

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
- Chen, H., Lim, C. K., Lee, Y. K. & Chan, Y. N. (2000). Comparative analysis of the genes encoding 23S–5S rRNA intergenic spacer regions of *Lactobacillus casei*-related strains. *Int J Syst Evol Microbiol* **50**, 471–478.
- Cilia, V., Lafay, B. & Christen, R. (1996). Sequence heterogeneities among 16S ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. *Mol Biol Evol* **13**, 451–461.
- Collins, M. D., Phillips, B. A. & Zanoni, P. (1989). Deoxyribonucleic acid homology studies of *Lactobacillus casei*, *Lactobacillus paracasei* sp. nov., subsp. *paracasei* and subsp. *tolerans*, and *Lactobacillus rhamnosus* sp. nov., comb. nov. *Int J Syst Bacteriol* **39**, 105–108.
- Collins, M. D., Rodrigues, U. M., Ash, C., Aguirre, M., Farrow, J. A. E., Martinez-Murcia, A., Phillips, B. A., Williams, A. M. & Wallbanks, S. (1991). Phylogenetic analysis of the genus *Lactobacillus* and related lactic acid bacteria as determined by reverse transcriptase sequencing of 16S rRNA. *FEMS Microbiol Lett* **77**, 5–12.
- Dellaglio, F., Bottazzi, V. & Vescovo, M. (1975). Deoxyribonucleic acid homology among *Lactobacillus* species of the subgenus *Streptobacterium* Orla-Jensen. *Int J Syst Bacteriol* **25**, 160–172.
- Dellaglio, F., Dicks, L. M. T., Du Toit, M. & Torriani, S. (1991). Designation of ATCC 334 in place of ATCC 393 (NCDO 161) as the neotype strain of *Lactobacillus casei* subsp. *casei*, and rejection of the name *Lactobacillus paracasei* (Collins *et al.*, 1989). Request for an Opinion. *Int J Syst Bacteriol* **41**, 340–342.
- Dellaglio, F., Felis, G. E. & Torriani, S. (2002). The status of the species *Lactobacillus casei* (Orla-Jensen 1916) Hansen and Lessel 1971 and *Lactobacillus paracasei* Collins *et al.* 1989. Request for an Opinion. *Int J Syst Evol Microbiol* **52**, (in press).
- Dicks, L. M. T., Du Plessis, E. M., Dellaglio, F. & Lauer, E. (1996). Reclassification of *Lactobacillus casei* subsp. *casei* ATCC 393 and *Lactobacillus rhamnosus* ATCC 15820 as *Lactobacillus zeae* nom. rev., designation of ATCC 334 as the neotype of *L. casei* subsp. *casei*, and rejection of the name *Lactobacillus paracasei*. *Int J Syst Bacteriol* **46**, 337–340.
- Duwat, P., Ehrlich, S. D. & Gruss, A. (1992). A general method for cloning *recA* genes of gram-positive bacteria by polymerase chain reaction. *J Bacteriol* **174**, 5171–5175.
- Dybvig, K., Hollingshead, S. K., Heath, D. G., Clewell, D. B., Sun, F. & Woodard, A. (1992). Degenerate oligonucleotide primers for enzymatic amplification of *recA* sequences from gram-positive bacteria and mycoplasmas. *J Bacteriol* **174**, 2729–2732.
- Eisen, J. A. (1995). The RecA protein as a model molecule for the molecular systematic studies of bacteria: comparison of trees of RecAs and 16S RNAs from the same species. *J Mol Evol* **41**, 1105–1123.
- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.
- Felsenstein, J. (1993). PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, WA, USA.
- Fox, G. E., Wisotzkey, J. D. & Jurtshuk, P., Jr (1992). How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int J Syst Bacteriol* **42**, 166–170.
- Gruber, T. M., Eisen, J. A., Gish, K. & Bryant, D. A. (1998). The phylogenetic relationships of *Chlorobium tepidum* and *Chloroflexus aurantiacus* based upon their RecA sequences. *FEMS Microbiol Lett* **162**, 53–60.
- Gupta, R. S. (2000). Evolutionary relationships among bacteria: does 16S rRNA provide all the answers? *ASM News* **66**, 189–190.
- Hansen, P. A. & Lessel, E. F. (1971). *Lactobacillus casei* (Orla-Jensen) comb. nov. *Int J Syst Bacteriol* **21**, 69–71.
- Hasegawa, M. & Hashimoto, T. (1993). Ribosomal RNA trees misleading? *Nature* **361**, 23.

- Hori, H. & Osawa, S. (1979). Evolutionary change in 5S RNA secondary structure and a phylogenetic tree of 54 5S RNA species. *Proc Natl Acad Sci U S A* **76**, 381–385.
- Kim, J. (1993). Improving the accuracy of phylogenetic estimation by combining different methods. *Syst Biol* **42**, 331–340.
- Kullen, M. J., Brady, L. J. & O'Sullivan, D. J. (1997). Evaluation of using a short region of the *recA* gene for the rapid and sensitive speciation of dominant bifidobacteria in the human large intestine. *FEMS Microbiol Lett* **154**, 377–383.
- Lake, J. A. (1991). The order of sequence alignment can bias the selection of tree topology. *Mol Biol Evol* **8**, 378–385.
- Lloyd, A. T. & Sharp, P. M. (1993). Evolution of the *recA* gene and the molecular phylogeny of bacteria. *J Mol Evol* **37**, 399–407.
- Maréchal, J., Clement, B., Nalin, R., Gandon, C., Orso, S., Cvejic, J. H., Bruneteau, M., Berry, A. & Normand, P. (2000). A *recA* gene phylogenetic analysis confirms the close proximity of *Frankia* to *Acidothermus*. *Int J Syst Evol Microbiol* **50**, 781–785.
- Marmur, J. (1961). A procedure for the isolation of DNA from microorganisms. *J Mol Biol* **3**, 208–218.
- Mori, K., Yamazaki, K., Ishiyama, T., Katsumata, M., Kobayashi, K., Kawai, Y., Inoue, N. & Shinano, H. (1997). Comparative sequence analyses of the genes coding for 16S rRNA of *Lactobacillus casei*-related taxa. *Int J Syst Bacteriol* **47**, 54–57.
- Nour, M. (1998). 16S–23S and 23S–5S intergenic spacer regions of lactobacilli: nucleotide sequence, secondary structure and comparative analysis. *Res Microbiol* **149**, 433–448.
- Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
- Smith, N. H., Holmes, E. C., Donovan, G. M., Carpenter, G. A. & Spratt, B. G. (1999). Networks and groups within the genus *Neisseria*: analysis of *argF*, *recA*, *rho*, and 16S rRNA sequences from human *Neisseria* species. *Mol Biol Evol* **16**, 773–783.
- Sogin, M. L., Hinkle, G. & Leipe, D. D. (1993). Universal tree of life. *Nature* **362**, 795.
- Stackebrandt, E. & Ludwig, W. (1994). The importance of using outgroup reference organisms in phylogenetic studies: the *Atopobium* case. *Syst Appl Microbiol* **17**, 39–43.
- Story, R. M. & Steitz, T. A. (1992). Structure of the *recA* protein–ADP complex. *Nature* **355**, 374–376.
- Swezey, J. L., Nakamura, L. K., Abbott, T. P. & Peterson, R. E. (2000). *Lactobacillus arizonensis* sp. nov., isolated from jojoba meal. *Int J Syst Evol Microbiol* **50**, 1803–1809.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.
- Ueda, K., Seki, T., Kudo, T., Yoshida, T. & Kataoka, M. (1999). Two distinct mechanisms cause heterogeneity of 16S rRNA. *J Bacteriol* **181**, 78–82.
- Vescovo, M., Torriani, S., Scolari, G. & Dicks, L. M. T. (1995). *Lactobacillus casei* and related species: a review. *Ann Microbiol Enzimol* **45**, 51–83.
- van Waasbergen, L. G., Balkwill, D. L., Crocker, F. H., Bjornstad, B. N. & Miller, R. V. (2000). Genetic diversity among *Arthrobacter* species collected across a heterogeneous series of terrestrial deep-subsurface sediments as determined on the basis of 16S rRNA and *recA* gene sequences. *Appl Environ Microbiol* **66**, 3454–3463.
- Wayne, L. G. (1994). Action of the Judicial Commission of the International Committee on Systematic Bacteriology on Requests for Opinions published between January 1985 and July 1993. *Int J Syst Bacteriol* **44**, 177–178.
- Yoon, J.-H., Kang, S.-S. & Mheen, T.-I. & 7 other authors (2000). *Lactobacillus kimchii* sp. nov., a new species from kimchi. *Int J Syst Evol Microbiol* **50**, 1789–1795.