

Comparison of *gyrB* gene sequences, 16S rRNA gene sequences and DNA–DNA hybridization in the *Bacillus subtilis* group

Li-Ting Wang,¹ Fwu-Ling Lee,¹ Chun-Ju Tai¹ and Hiroaki Kasai²

Correspondence
Fwu-Ling Lee
fll@firdi.org.tw

¹Bioresource Collection and Research Center, Food Industry Research and Development Institute, PO Box 246, Hsinchu 30099, Taiwan

²Marine Biotechnology Institute, Heita, Kamaishi, Iwate, 026-0001, Japan

The *Bacillus subtilis* group comprises eight closely related species that are indistinguishable from one another by 16S rRNA gene sequence analysis. Therefore, the *gyrB* gene, which encodes the subunit B protein of DNA gyrase, was selected as an alternative phylogenetic marker. To determine whether *gyrB* gene sequence analysis could be used for phylogenetic analysis and species identification of members of the *B. subtilis* group, the congruence of *gyrB* grouping with both 16S rRNA gene sequencing and DNA–DNA hybridization data was evaluated.

Ranges of *gyrB* nucleotide and translated amino acid sequence similarities among the eight type strains were 75.4–95.0% and 88.5–99.2%, respectively, whereas 16S rRNA gene sequence similarities were 98.1–99.8%. Results showed that *gyrB* gene sequences provide higher resolution than 16S rRNA gene sequences. The classification achieved by *gyrB* sequence analysis was in agreement with results obtained with DNA–DNA hybridization. It is concluded that the *gyrB* gene may be an efficient alternative target for identification and taxonomic analysis of members of the *B. subtilis* group.

Bacillus subtilis is a Gram-positive, spore-forming, fermentative, aerobic, rod-shaped bacterium. The *Bacillus subtilis* group contains the closely related taxa *Bacillus subtilis* subsp. *subtilis* (Smith *et al.*, 1964; Nakamura *et al.*, 1999), *Bacillus licheniformis* (Skerman *et al.*, 1980), *Bacillus amyloliquefaciens* (Priest *et al.*, 1987), *Bacillus atrophaeus* (Nakamura, 1989), *Bacillus mojavensis* (Roberts *et al.*, 1994), *Bacillus vallismortis* (Roberts *et al.*, 1996), *Bacillus subtilis* subsp. *spizizenii* (Nakamura *et al.*, 1999) and *Bacillus sonorensis* (Palmisano *et al.*, 2001). These taxa can be differentiated from one another by fatty acid composition analysis, restriction digest analysis and DNA–DNA hybridization analysis, but are quite difficult to differentiate by phenotypic characteristics (Roberts *et al.*, 1994; Nakamura *et al.*, 1999).

16S rRNA gene sequence analysis is the most commonly used method for identifying bacteria or for constructing bacterial phylogenetic relationships (Woese, 1987;

The GenBank/EMBL/DDBJ accession numbers for the *gyrB* gene sequences of 32 strains used in this study are DQ309293–DQ309325 and those of the 16S rRNA gene sequences are DQ993674–DQ993679, EF423592–EF423609 and EF433402–EF433411, as detailed in Supplementary Table S1.

Details of strains and sequence accession numbers and a table of DNA–DNA reassociation values and *gyrB* and 16S rRNA gene sequences similarities are available as supplementary material with the online version of this paper.

Vandamme *et al.*, 1996; Joung & Cote, 2002); however, its usefulness is limited because of the high percentage of sequence similarity between closely related species (Ash *et al.*, 1991; Martínez-Murcia *et al.*, 1992; Christensen *et al.*, 1998). The use of protein-encoding genes as phylogenetic markers is now a common approach (Yamamoto & Harayama, 1998; Ko *et al.*, 2004; Chelo *et al.*, 2007). Detailed investigations have demonstrated that sequences from protein-encoding genes can accurately predict genome relatedness and may replace DNA–DNA hybridization for species identification and delineation in the future (Stackebrandt *et al.*, 2002; Zeigler, 2003).

The *gyrB* gene encodes the subunit B protein of DNA gyrase, a type II DNA topoisomerase, which plays an essential role in DNA replication and is distributed universally among bacterial species (Watt & Hickson, 1994; Huang, 1996). The rate of molecular evolution inferred from *gyrB* gene sequences is faster than that inferred from 16S rRNA gene sequences (Yamamoto & Harayama, 1995). *gyrB* gene sequences have been used in phylogenetic studies of *Pseudomonas* (Yamamoto & Harayama, 1998), *Acinetobacter* (Yamamoto & Harayama, 1996; Yamamoto *et al.*, 1999), *Mycobacterium* (Kasai *et al.*, 2000; Niemann *et al.*, 2000), *Salmonella*, *Shigella* and *Escherichia coli* (Fukushima *et al.*, 2002), *Aeromonas* (Yáñez *et al.*, 2003) and the *Bacillus anthracis*–*cereus*–*thuringiensis* group (La Duc *et al.*, 2004); results from these studies have indicated that *gyrB* is a

suitable phylogenetic marker for the study of phylogenetic and taxonomic relationships at the species level. In the present study, it has been shown that direct sequencing of the *gyrB* gene could be used for identification and phylogenetic analysis of species of the *B. subtilis* group.

A total of eight *Bacillus* type strains and 24 *Bacillus* reference strains were used in this study (see Supplementary Table S1 available in IJSEM Online). They were obtained from the Bioresource Collection and Research Center (BCRC; <http://wdcm.nig.ac.jp/CCINFO/CCINFO.xml?59>). All strains were cultivated on nutrient agar or in nutrient broth (Difco) at 30 °C for 24 h under aerobic conditions.

Genomic DNA was extracted using the Qiagen Blood & Cell Culture DNA kit. DNA–DNA relatedness values were determined using the fluorometric hybridization method in microdilution wells as described previously (Ezaki *et al.*, 1989; Chern *et al.*, 2004; Tai *et al.*, 2006).

The *gyrB* gene was amplified by PCR as described previously (Yamamoto & Harayama, 1995). PCR was performed using the Takara *Ex Taq* kit. PCR products were purified with the PCR-M clean up system (Viogene) and sequenced with a BigDye Terminator v3.1 cycle-sequencing kit on a 3730 DNA sequencer (Applied Biosystems and Hitachi). DNA sequencing was determined using *gyrB* degenerate primers UP-1S and UP-2Sr (Yamamoto & Harayama, 1995) and BS-F (5'-GAAGGCGGNACNCAYG-AAG-3') and BS-R (5'-CTTCRTGNGTNC CGCCTTC-3') (designed from conserved regions of *gyrB* nucleotide sequences of members of the *B. subtilis* group) at 3.2 µM concentration. The DNA sequence was double-checked by sequencing both strands. Approximately 1.5 kb of the 16S rRNA gene was determined using the MicroSeq Full Gene 16S rDNA Bacterial Identification kit (Applied Biosystems).

Sequence similarities were calculated using programs of the Wisconsin Package, version 10.1 (Accelrys). Multiple sequences were aligned using the program CLUSTAL_X, version 1.8 (Thompson *et al.*, 1997). Phylogenetic analysis was performed using PHYLIP (Felsenstein, 1993) and MEGA (Kumar *et al.*, 2004). Evolutionary distances were calculated by Kimura's two-parameter model (Kimura, 1980). Phylogenetic trees were constructed by the neighbour-joining (Saitou & Nei, 1987) and maximum-parsimony (Fitch, 1971) methods with bootstrap values based on 1000 replications.

Approximately 1.2 kb of the *gyrB* gene was successfully amplified using universal primers UP-1 and UP-2r for all *Bacillus* species examined in this study. After direct sequencing, 1171 bp *gyrB* gene sequences, corresponding to nt 316–1480 of the *Escherichia coli* K-12 sequence with gaps, and 1468 bp 16S rRNA gene sequences, corresponding to nt 30–1489 of the *E. coli* K-12 sequence with gaps, were used for analysis and the resulting data were deposited in GenBank/EMBL/DDBJ. Accession numbers of *gyrB* and 16S rRNA gene sequences are listed in Supplementary Table S1.

Species identification and phylogenetic analysis of *B. subtilis* and related taxa based on *gyrB* gene sequence analysis

The *gyrB* gene sequence similarities between the eight type strains were 75.4–95.0 % (mean 81.6 %) (see Supplementary Table S2). The *gyrB* translated amino acid sequence similarity values were 88.5–99.2 % (mean 92.5 %) (data not shown). In contrast, 16S rRNA gene sequence similarities between the same strains were 98.1–99.8 % (mean 98.9 %). It is clear from comparative sequence analysis that the base substitution rate of the *gyrB* gene sequence was much faster than that of the 16S rRNA gene sequence and that the *gyrB* nucleotide sequence showed significantly higher genetic variation than the translated amino acid sequence.

The *gyrB* gene sequence showed remarkable discrimination (75.4–91.9 %). At the intraspecies level, the nucleotide substitution rates were 0–5 % and were <2 % for most *Bacillus* species. At the interspecies level, the nucleotide substitution rates were usually >7 %, except for *B. vallismortis* and *B. subtilis* subsp. *spizizenii*, the most similar pair (6.1 % sequence divergence). At the subspecies level, the *gyrB* gene sequence could be used to discriminate *B. subtilis* subsp. *subtilis* from *B. subtilis* subsp. *spizizenii* (5 % sequence divergence). This result was consistent with that of *gyrA* gene sequence analysis (4.8 % sequence divergence) (Chun & Bae, 2000). The 16S rRNA gene sequences of all *Bacillus* strains tested revealed more than 98 % similarity. These results indicate that the *gyrB* gene is more useful than the 16S rRNA gene for species and subspecies identification in the *B. subtilis* group.

The phylogenetic trees constructed from the 16S rRNA and *gyrB* gene sequences of the 32 *Bacillus* strains are shown in Fig. 1. The *gyrB*-based tree clearly delineated four distinct clusters with high bootstrap values (100 %): cluster 1 contained *B. subtilis*, *B. vallismortis* and *B. mojavensis* strains; clusters 2 and 3 contained strains of *B. atrophaeus* and *B. amyloliquefaciens*, respectively; and cluster 4 contained *B. sonorensis* and *B. licheniformis* strains (Fig. 1). Of the *B. subtilis* strains, all strains of *B. subtilis* subsp. *subtilis* formed a monophyletic clade with 100 % bootstrap support; the *gyrB* gene sequence similarities among them were 98.1–99.1 % (mean 98.7 %) (data not shown) and the sequence divergence between the *B. subtilis* subsp. *subtilis* and *B. subtilis* subsp. *spizizenii* strains was 5.0 %. Comparatively, the 16S rRNA gene-based tree yielded two clusters with a bootstrap value of 79 %: cluster 1 contained strains of *B. subtilis*, *B. mojavensis*, *B. vallismortis*, *B. atrophaeus* and *B. amyloliquefaciens* and cluster 2 contained strains of *B. sonorensis* and *B. licheniformis*. All *Bacillus* strains among these two cluster groups showed more than 99 % 16S rRNA gene sequence similarity, indicating that the *gyrB* gene is a better molecular marker than the 16S rRNA gene for the study of phylogenetic and taxonomic relationships at the species level in the *B. subtilis* group.

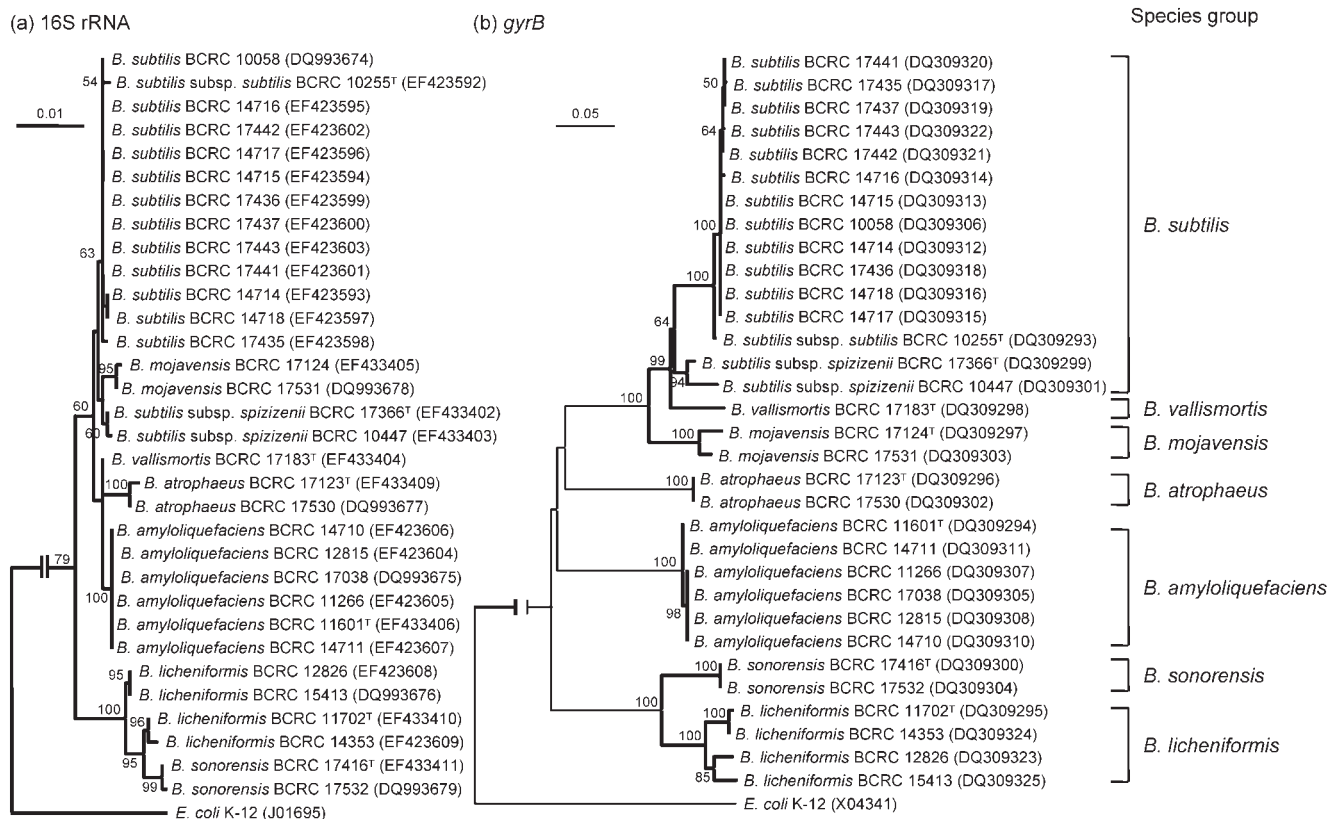


Fig. 1. Phylogenetic trees of 32 *Bacillus* strains based on 16S rRNA (a) and *gyrB* (b) gene sequences. The trees were constructed with the neighbour-joining method. Genetic distances were computed by Kimura's two-parameter model. *E. coli* K-12 was included as an outgroup. Only bootstrap percentages above 50% are shown (based on 1000 replications). Bars, 0.01 (a) or 0.05 (b) substitutions per nucleotide position.

Correlation between *gyrB* gene sequence similarity and DNA–DNA relatedness

In the modern taxonomy of bacteria, direct genomic DNA–DNA hybridization comparisons have been deemed the 'gold standard' in bacterial characterization and identification (Wayne *et al.*, 1987; Stackebrandt *et al.*, 2002). Reassociation values obtained from DNA–DNA hybridization analyses among the type strains and other *Bacillus* strains examined in this study were determined (data not shown). The DNA–DNA reassociation values between the eight type strains were low, 11–67% (Supplementary Table S2). Nakamura *et al.* (1999) divided *B. subtilis* into *B. subtilis* subsp. *subtilis* and *Bacillus subtilis* subsp. *spizizenii* mainly using DNA–DNA reassociation data. DNA–DNA relatedness (67%) and 16S rRNA gene sequence similarity (99.8%) indicated that these two taxa are closely related; however, the *gyrB* gene sequence can be used to discriminate between them (95% sequence similarity). Pairwise analyses indicated that *gyrB* gene sequences (75.4–95.0% similarity) are more discriminatory than 16S rRNA gene sequences (98.1–99.8% similarity) for species differentiation. The *gyrB* gene sequence-based phylogenetic analysis was consistent with DNA–DNA reassociation data

and a linear correlation was observed between *gyrB* gene sequence similarities and levels of DNA–DNA relatedness (Fig. 2). All strains that exhibited 95–100% *gyrB* gene sequence similarity showed high DNA–DNA relatedness (70–100%), suggesting that these strains are conspecific. Strains with approximately 95% or higher *gyrB* gene sequence similarity in the cluster exhibited DNA–DNA relatedness of >70%, an acceptable value for proposal of a single species. Strains with 93–95% *gyrB* gene sequence similarity exhibited DNA–DNA relatedness of 60–70% without exception, indicating that they are grouped at the subspecies level. Exceptionally, *B. subtilis* subsp. *spizizenii* BCRC 17366^T and *B. vallismortis* BCRC 17183^T showed 93.9% *gyrB* gene sequence similarity and 52% DNA–DNA relatedness. This finding indicated that it might be necessary to use several gene sequences and DNA–DNA hybridization to discriminate species relationships. Nevertheless, based on data obtained from the present study, the *gyrB* gene sequences have been shown to be a more efficient phylogenetic tool than the 16S rRNA gene sequences for discriminating between species of this group.

In conclusion, the 16S rRNA gene sequence is extremely limiting in the discrimination of species in the *B. subtilis*

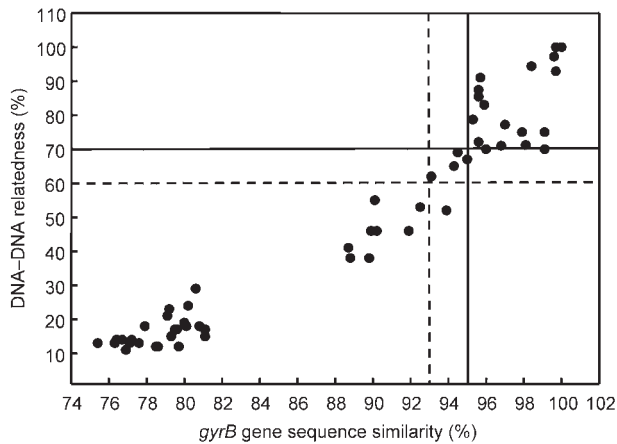


Fig. 2. Correlation between *gyrB* gene sequence similarity and DNA–DNA relatedness of *Bacillus* strains used in this study. Solid lines represent 70 % DNA–DNA relatedness and 95 % *gyrB* gene sequence similarity; dashed lines represent 60 % DNA–DNA relatedness and 93 % *gyrB* gene sequence similarity.

group. It is suggested that the *gyrB* gene may be a useful alternative to DNA–DNA hybridization for the identification and phylogenetic analysis of members of the *B. subtilis* group at the species and subspecies level, with some exceptions.

Acknowledgements

We thank T. Y. Liu, C. C. Liao and G. F. Yuan (Food Industry Research and Development Institute, Taiwan) for their encouragement. This research was supported by the Taiwanese Ministry of Economic Affairs (project no. 95-EC-17 A-17-R7-0525).

References

Ash, C., Farrow, J. A., Dorsch, M., Stackebrandt, E. & Collins, M. D. (1991). Comparative analysis of *Bacillus anthracis*, *Bacillus cereus*, and related species on the basis of reverse transcriptase sequencing of 16S rRNA. *Int J Syst Bacteriol* **41**, 343–346.

Chelo, I. M., Zé-Zé, L. & Tenreiro, R. (2007). Congruence of evolutionary relationships inside the *Leuconostoc–Oenococcus–Weissella* clade assessed by phylogenetic analysis of the 16S rRNA gene, *dnaA*, *gyrB*, *rpoC* and *dnaK*. *Int J Syst Evol Microbiol* **57**, 276–286.

Chern, L.-L., Stackebrandt, E., Lee, S.-F., Lee, F.-L., Chen, J.-K. & Fu, H.-M. (2004). *Chitinibacter tainanensis* gen. nov., sp. nov., a chitin-degrading aerobe from soil in Taiwan. *Int J Syst Evol Microbiol* **54**, 1387–1391.

Christensen, H., Nordentoft, S. & Olsen, J. E. (1998). Phylogenetic relationships of *Salmonella* based on rRNA sequences. *Int J Syst Bacteriol* **48**, 605–610.

Chun, J. & Bae, K. S. (2000). Phylogenetic analysis of *Bacillus subtilis* and related taxa based on partial *gyrA* gene sequences. *Antonie Van Leeuwenhoek* **78**, 123–127.

Ezaki, T., Hashimoto, Y. & Yabuuchi, E. (1989). Fluorometric deoxy-ribonucleic acid–deoxyribonucleic acid hybridization in microdilution wells as an alternative to membrane filter hybridization in which

radioisotopes are used to determine genetic relatedness among bacterial strains. *Int J Syst Bacteriol* **39**, 224–229.

Felsenstein, J. (1993). PHYLIP (phylogeny inference package), version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, USA.

Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specific tree topology. *Syst Zool* **20**, 406–416.

Fukushima, M., Kakinuma, K. & Kawaguchi, R. (2002). Phylogenetic analysis of *Salmonella*, *Shigella*, and *Escherichia coli* strains on the basis of the *gyrB* gene sequence. *J Clin Microbiol* **40**, 2779–2785.

Huang, W. M. (1996). Bacterial diversity based on type II DNA topoisomerase genes. *Annu Rev Genet* **30**, 79–107.

Joung, K. B. & Cote, J. C. (2002). Evaluation of ribosomal RNA gene restriction patterns for the classification of *Bacillus* species and related genera. *J Appl Microbiol* **92**, 97–108.

Kasai, H., Ezaki, T. & Harayama, S. (2000). Differentiation of phylogenetically related slowly growing mycobacteria by their *gyrB* sequences. *J Clin Microbiol* **38**, 301–308.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.

Ko, K. S., Kim, J. W., Kim, J. M., Kim, W., Chung, S. I., Kim, I. J. & Kook, Y. H. (2004). Population structure of the *Bacillus cereus* group as determined by sequence analysis of six housekeeping genes and the *plcR* gene. *Infect Immun* **72**, 5253–5261.

Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.

La Duc, M. T., Satomi, M., Agata, N. & Venkateswaran, K. (2004). *gyrB* as a phylogenetic discriminator for members of the *Bacillus anthracis–cereus–thuringiensis* group. *J Microbiol Methods* **56**, 383–394.

Martinez-Murcia, A. J., Benlloch, S. & Collins, M. D. (1992). Phylogenetic interrelationships of members of the genera *Aeromonas* and *Plesiomonas* as determined by 16S ribosomal DNA sequencing: lack of congruence with results of DNA–DNA hybridization. *Int J Syst Bacteriol* **42**, 412–421.

Nakamura, L. K. (1989). Taxonomic relationship of black-pigmented *Bacillus subtilis* strains and a proposal for *Bacillus atrophaeus* sp. nov. *Int J Syst Bacteriol* **39**, 295–300.

Nakamura, L. K., Roberts, M. S. & Cohan, F. M. (1999). Relationship of *Bacillus subtilis* clades associated with strains 168 and W23: a proposal for *Bacillus subtilis* subsp. *subtilis* subsp. nov. and *Bacillus subtilis* subsp. *spizizenii* subsp. nov. *Int J Syst Bacteriol* **49**, 1211–1215.

Niemann, S., Harmsen, D., Rusch-Gerdes, S. & Richter, E. (2000). Differentiation of clinical *Mycobacterium tuberculosis* complex isolates by *gyrB* DNA sequence polymorphism analysis. *J Clin Microbiol* **38**, 3231–3234.

Palmisano, M. M., Nakamura, L. K., Duncan, K. E., Istock, C. A. & Cohan, F. M. (2001). *Bacillus sonorensis* sp. nov., a close relative of *Bacillus licheniformis*, isolated from soil in the Sonoran Desert, Arizona. *Int J Syst Evol Microbiol* **51**, 1671–1679.

Priest, F. G., Goodfellow, M., Shute, L. A. & Berkeley, R. C. W. (1987). *Bacillus amyloliquefaciens* sp. nov., nom. rev. *Int J Syst Bacteriol* **37**, 69–71.

Roberts, M. S., Nakamura, L. K. & Cohan, F. M. (1994). *Bacillus mojavensis* sp. nov., distinguishable from *Bacillus subtilis* by sexual isolation, divergence in DNA sequence, and differences in fatty acid composition. *Int J Syst Bacteriol* **44**, 256–264.

Roberts, M. S., Nakamura, L. K. & Cohan, F. M. (1996). *Bacillus vallismortis* sp. nov., a close relative of *Bacillus subtilis*, isolated from soil in Death Valley, California. *Int J Syst Bacteriol* **46**, 470–475.

- Saitou, N. & Nei, M. (1987).** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
- Skerman, V. B. D., McGowan, V. & Sneath, P. H. A. (editors) (1980).** Approved lists of bacterial names. *Int J Syst Bacteriol* **30**, 225–420.
- Smith, N. R., Gibson, T., Gordon, R. E. & Sneath, P. H. (1964).** Type cultures and proposed neotype cultures of some species in the genus *Bacillus*. *J Gen Microbiol* **34**, 269–272.
- Stackebrandt, E., Frederiksen, W., Garrity, G. M., Grimont, P. A. D., Kämpfer, P., Maiden, M. C. J., Nesme, X., Rosselló-Mora, R., Swings, J. & other authors (2002).** Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* **52**, 1043–1047.
- Tai, C. J., Kuo, H. P., Lee, F. L., Chen, H. K., Yokota, A. & Lo, C. C. (2006).** *Chryseobacterium taiwanense* sp. nov., isolated from soil in Taiwan. *Int J Syst Evol Microbiol* **56**, 1771–1776.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997).** The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.
- Vandamme, P., Pot, B., Gillis, M., De Vos, P., Kersters, K. & Swings, J. (1996).** Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* **60**, 407–438.
- Watt, P. M. & Hickson, I. D. (1994).** Structure and function of type II DNA topoisomerases. *Biochem J* **303**, 681–695.
- Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E. & other authors (1987).** International Committee on Systematic Bacteriology. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* **37**, 463–464.
- Woese, C. R. (1987).** Bacterial evolution. *Microbiol Rev* **51**, 221–271.
- Yamamoto, S. & Harayama, S. (1995).** PCR amplification and direct sequencing of *gyrB* genes with universal primers and their application to the detection and taxonomic analysis of *Pseudomonas putida* strains. *Appl Environ Microbiol* **61**, 1104–1109.
- Yamamoto, S. & Harayama, S. (1996).** Phylogenetic analysis of *Acinetobacter* strains based on the nucleotide sequences of *gyrB* genes and on the amino acid sequences of their products. *Int J Syst Bacteriol* **46**, 506–511.
- Yamamoto, S. & Harayama, S. (1998).** Phylogenetic relationships of *Pseudomonas putida* strains deduced from the nucleotide sequences of *gyrB*, *rpoD* and 16S rRNA genes. *Int J Syst Bacteriol* **48**, 813–819.
- Yamamoto, S., Bouvet, P. J. & Harayama, S. (1999).** Phylogenetic structures of the genus *Acinetobacter* based on *gyrB* sequences: comparison with the grouping by DNA–DNA hybridization. *Int J Syst Bacteriol* **49**, 87–95.
- Yáñez, M. A., Catalán, V., Apráiz, D., Figueras, M. J. & Martínez-Murcia, A. J. (2003).** Phylogenetic analysis of members of the genus *Aeromonas* based on *gyrB* gene sequences. *Int J Syst Evol Microbiol* **53**, 875–883.
- Zeigler, D. R. (2003).** Gene sequences useful for predicting relatedness of whole genomes in bacteria. *Int J Syst Evol Microbiol* **53**, 1893–1900.