

Application of the partial 16S rDNA sequence as an index for rapid identification of species in the genus *Bacillus*

Keiichi Goto,* Tomoko Omura, Yukihiko Hara, and Yoshito Sadaie^{1,†}

Laboratory of QC and Food Analysis, Food Research Institute, Mitsui Norin Co., Ltd., Shizuoka 426–0133, Japan

¹Radioisotope Center, National Institute of Genetics, Shizuoka 411–0801, Japan

(Received April 21, 1999; Accepted November 26, 1999)

A comparison of 16S rRNA gene (rDNA) sequences was made among type strains of 69 *Bacillus* species approved in the International Journal of Systematic Bacteriology (IJSB) until 1998. The results suggested that 5' end region (approx. 275 bp) was the hypervariant region (HV region) in the gene and was highly specific for each type strain. Furthermore, a sequence analysis of the HV region of *Bacillus* strains revealed that this region was highly conserved within the species. These results indicate that the HV region is a useful index for the identification or grouping of *Bacillus* species.

Key Words—*Bacillaceae*; genus *Bacillus*; grouping; identification; 16S rRNA gene

Introduction

The genus *Bacillus* is a large and heterogeneous collection of aerobic or facultatively anaerobic, rod-shaped, endospore-forming bacteria that are widely distributed in the environment. Many kinds of species belong to this genus, which have acidophilic, alkalophilic, thermophilic, or other properties. The reclassification of genus *Bacillus*, which began in 1991, revealed at least eight genera: *Alicyclobacillus* (Wisotzkey et al., 1992), *Aneurinibacillus* (Shida et al., 1996), *Bacillus*, *Brevibacillus* (Shida et al., 1995, 1996), *Gracilibacillus* (Waino et al., 1999), *Paenibacillus* (Ash et al., 1991, 1993), *Salibacillus* (Waino et al., 1999), and *Virgibacillus* (Heyndrickx et al., 1998). Since these eight genera consist of more than 100 species that have similar characteristics, identifying

them is difficult.

The identification of *Bacillus* species has been performed mainly with morphological and physiological criteria, and this method is widely employed in various fields. However, the process used requires skillful techniques and is very complex and time-consuming. With the advance of genetic engineering, the randomly amplified polymorphic DNA (RAPD) method (Yamazaki et al., 1997), the hybridization method (Giffel et al., 1997; Timothy et al., 1994), or restriction mapping (Ilan et al., 1995) were adapted for the identification of *Bacillus* species. These methods are effective for identification or detection among a small number of species, but they are not suitable for identification among a large number of species.

Over the years, a sizable database of 16S rRNA gene (rDNA) has been built, and this has been successfully applied in determining phylogenetic relationships or in identifying bacteria. Moreover, it has been reported that a partial region of 16S rDNA is effective for the classification and identification of acetic acid bacteria (Yamada et al., 1997) and *Streptomyces* (Kataoka et al., 1997).

In this work, we investigated the identification

* Address reprint requests to: Dr. Keiichi Goto, Laboratory of QC and Food Analysis, Food Research Institute, 223–1 Miyahara, Fujieda, Shizuoka 426–0133, Japan.

E-mail: kgoto@mitsui-norin.co.jp

[†] Present address: Department of Molecular Biology, Faculty of Science, Saitama University, 255, Shimo-Okubo, Urawa, Saitama 338–8570, Japan.

method of *Bacillus* species by using the partial 16S rDNA sequence. First, to search the most informative region of the 16S rDNA, the type strains of 69 *Bacillus* species were subjected to a comparison of the 16S rDNA sequences. The results revealed that the 5' end region was the hypervariant region (HV region) and highly specific for each type strain. Furthermore, sequence analyses of the HV region from 51 strains belonging to four clusters suggested that the HV region was highly conserved within species. These results showed that the HV region is a very efficient index for the rapid identification or grouping of *Bacillus* species.

Materials and Methods

Strains and culture conditions. The 107 bacterial strains used in this study were obtained from ATCC (American Type Culture Collection, Manassas, VA, USA), DSM (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany), IAM (Institute of Molecular and Cellular Biosciences, The University of Tokyo, Tokyo, Japan), IFO (Institute for Fermentation, Osaka, Japan), JCM (Japan Collection of Microorganisms, Saitama, Japan), NCIMB (The National Collections of Industrial and Marine Bacteria, Aberdeen, Scotland, UK), NIG (National Institute of Genetics, Mishima, Shizuoka, Japan), and NRRL (Agricultural Research Service Culture Collection, Peoria, IL, USA). These bacteria were cultivated by the methods recommended in the ATCC, DSM, IAM, IFO, JCM, NCIMB, NIG, and NRRL strain catalogues. The type strain of *B. anthracis* was not obtained because of its pathogenicity, so the 16S rDNA sequence of *B. anthracis*^T (accession number X55059) was used in this work. A bacterial cell of *B. thermamylovorans* CNCM I-1378^T (Collection Nationale de Cultures de Microorganismes) was not obtained, but chromosomal DNA of the strain was kindly provided by Dr. Bharat Patel (Griffith University, Brisbane, Australia).

Preparation of chromosomal DNA. The methods used for the preparation of chromosomal DNA were carried out by using a technique described previously (Sadaie et al., 1997).

Complete 16S rDNA sequences. Complete 16S rDNA sequences of type strains of 19 *Bacillus* species (*B. atrophaeus*, *B. carboniphilus*, *B. chitinolyticus*, *B. ehimensis*, *B. flexus*, *B. halodenitrificans*, *B. halodurans*, *B. halophilus*, *B. lentus*, *B. marinus*, *B. mojavensis*, *B. mycoides*, *B. naganoensis*, *B. niacini*, *B. psychrosaccharolyticus*, *B. stearothermophilus*, *B. thermoglucosidasius*, *B. vallismortis*, and *B. weihenstephanensis*) were determined by using a 16S rRNA Gene Kit following the protocols of the manufacturer (Perkin-Elmer Co.). These sequences which we determined have been deposited in the DDBJ under the accession numbers AB021181 to AB021199. The sequence data of *B. horikoshii* DSM 8719^T was obtained from Dr. F. A. Rainey. Other 16S rDNA sequences were obtained from the DDBJ-EMBL-GenBank database.

Sequencing of the HV region. For amplification and sequencing of the HV region, two primers were constructed based on the result of the multiple alignment of 16S rDNA sequences from 69 *Bacillus* species (type strain) by using CLUSTAL W version 1.7 (Fig. 1). A forward primer: 5'-TGT AAA ACG ACG GCC AGT GCC TAA TAC ATG CAA GTC GAG CG-3' (-21 M13 forward primer [5'-TGT AAA ACG AAC GGC CAG T-3']) was bound to an oligonucleotide [5'-GCC TAA TAC ATG CAA GTC GAG CG-3'] corresponding to nucleotide positions 47 to 69), and a reverse primer: 5'-CAG GAA ACA GCT ATG ACC ACT GCT GCC TCC CGT AGG AGT-3' (M13 reverse primer [5'-CAG GAA ACA GCT ATG ACC-3']) was bound to an oligonucleotide [5'-ACT GCT GCC TCC CGT AGG AGT-3'] corresponding to nucleotide positions 345 to 365). The HV region was amplified with the primer pair according to a method of Sadaie et al. (1997). The sequence of PCR product was determined by using a Dye Primer Cycle Sequencing Kit (Perkin-Elmer Co.) following the protocols of the manufacturer, and a model ABI 373 automatic DNA sequencer (Perkin-Elmer Co.).

Computer analysis and construction of phylogenetic tree. A sequence analysis was performed by using Gene Works (version 2.0, IntelliGenetics, Inc.) and the EMBL, GenBank, and DDBJ database. Multiple sequence alignment, calculation of nucleotide substitution rates (K_{nuc} values), construction of a neighbor-joining phylogenetic tree, and bootstrap analysis with 1,000 replicates for evaluation of phylogenetic tree topology were carried out with the Clustal W version 1.7 program (Kimura, 1980; Saitou and Nei, 1987; Thompson et al., 1994). Alignment gaps and unidentified base positions were not taken into account for the calculations.

DNA-DNA hybridization. For an analysis of chromosomal DNA homologies between different strains,

```

1 -----GCCT AATACATGCA
61 AGTCGAGCG-----A-- GGCG-A-GGG TGAG--AC-C
121 GT-G-----C -----G ---GG-A-AA C-----GAAA -----GCT-A --CC---T--
181 -----AAA -----C-- ---G-----
241 C-GCG--GC- TTAG-T--TT GG---GGTA- -GGC---CCA AG-C---GAT GC-TAGCCG-
301 C--GAGAGGG T---CG-CCA CA--GG-ACT -AGA-A-GG- C---ACTCCT ACGGGAGGCA
361 GCAGT-GGGA -----C--CA A--G-CG-AA G-C-GA--G- GC-A--C-G- -TG-----G
421 A-GG-----G G-T-G--AA- ---TG----- -G-G--GA-- -----
481 -----GA CGG-A-C--- --AG-AAGCC -CG---A--T ACGTGCCAGC AGCCGCGGT-
541 A-ACGT--G- -GC-AG-GT- -TCC-GA-T- --TGGGCGT- AAG-G--CG- --G-GG---
601 ----GTC--- -GT--AA--- ---G-T--A C----- -C---GGA-A C-G-----CT
661 --AG---G- AG-G---G- -GAATTCC-- GTGT-GCGGT -AAATGCCGA GA-AT---GA
721 GGAA-ACC-- T--CGAA-GC --C---T-G -C-----CT- AC-CTGAGG- -CG-AA----
781 ----AGC-AA CAGGATTAGA TACCCT-GTA G-CCA-GC-- -AAACG-TG- -TGCT---TG
841 T---G-G-- -----C- -----
901 -----TT- ACGGGG-CC- GCAC-AGC-G T-G----TGT
961 -GTTTA-TTC GA----ACGC GA--AACCT- ACCAGG--TT GA--TC--- G-G-----
1021 GAGA----- -----G- -----GAC -GGTG--GC- T-G--GTCG- CAGCT-G-GT
1081 CGTGAGATGT TGGGT--AGT CCCG-AACGA GCGC--CCC- -----GT TG-CA-C---
1141 --G--GGG-A CTC----- ACT-CCG--- ---A--CGGA GGAA-G--GG GA-GACGT-A
1201 A-TCATCATG CCCC--ATG- -CT-GG-T-C ACACGT-CTA C-ATGG---- -AC-----G-
1261 -G--A---G --A-----G C---C--- AA-----C --AGT-CGG- T-G--G--TG
1321 -----G-CT ---T-A-G-- GGA-T-G-T- GT---CG--G --CAG--TGC ----G-GA-T
1381 -CGT-CC-GG G-CT-GT--A -ACCGC--G- ---A---G- -AG-----C-----
1441 -----
1501 -----

```

Fig. 1. Consensus nucleotide sequence derives from *B. subtilis* (*rrnE*) (Sadaie et al., 1997) based on the result of the multiple sequence alignment of 16S rDNA from type strains of 69 *Bacillus* species.

Nonconserved nucleotides and gaps are represented '-'. Underlines indicate a part of sequencing primer used in this work.

bacterial DNA was isolated and purified by using a QIAGEN Blood & Cell Culture DNA Midi Kit (QIAGEN Inc.), according to the QIAGEN Genomic DNA Handbook 09/97. The DNA concentration and purity were determined spectrometrically at 230, 260, and 280 nm. Levels of DNA relatedness were determined fluorometrically by the method of Ezaki et al. (1989), using photobiotin-labeled DNA probes and microplates.

Results

Selection of variable region in the 16S rDNA for identification of Bacillus species

To find the most informative region of 16S rDNA for the rapid identification of *Bacillus* species based on sequence differences, multiple sequence alignment was carried out by using the 16S rDNA sequences of 69 type strains of *Bacillus* species. We were able to select a hypervariable region (HV region; nucleotide positions 70–344) in common with the 69 type strains. A sequence similarity of the HV region among the type strains was shown to be mostly lower than 90.0%; thus

the HV region was specific for the individual type strains. The only exception was that the HV region of *B. atrophaeus*^T coincided with that of *B. mojavensis*^T in spite of low hybridization value (25%) (Nakamura, 1989; Robert et al., 1994). Other variable regions were not common to all strains because of lack of enough sequence information and large nucleotide deletions of *B. schlegelii* or *B. tusciae*.

Using sequences of the selected HV region, the phylogenetic tree was constructed by the neighbor-joining method. This HV region tree was not the same as a tree obtained from complete 16S rDNA sequences; however, major clusters were strictly conserved, and the genetic distance was more emphasized in the HV region tree (Figs. 2 and 3). Thus we considered that the HV region was sufficient for identification of *Bacillus* strains at species level.

Amplification of the HV region

To examine the specificity of the designated primers to the HV region, PCR was performed with chromosomal DNA of 68 type strains (*B. anthracis*^T was not

Complete 16S rDNA

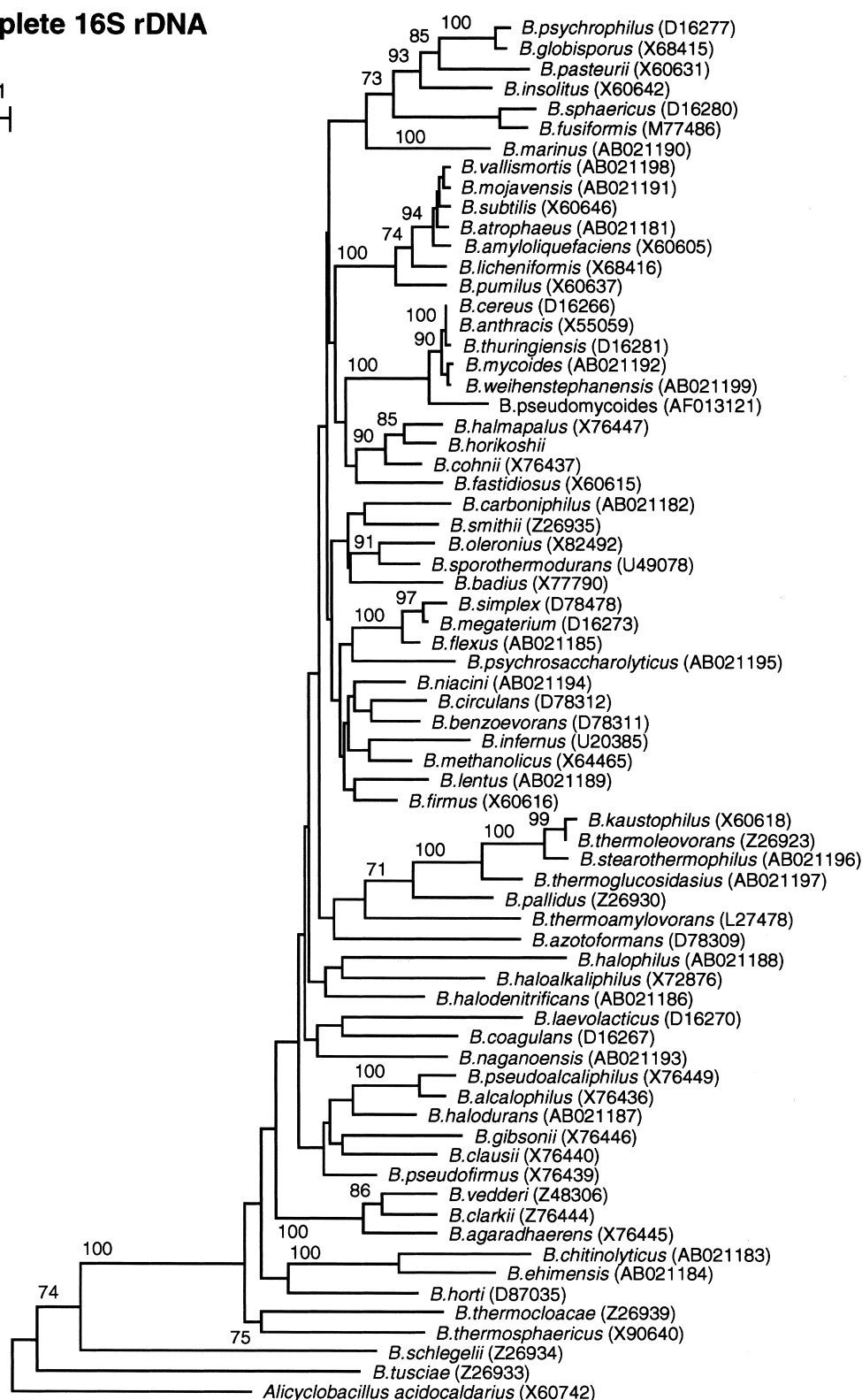
0.01
|

Fig. 2. Phylogenetic relationships of *Bacillus* species based on the 16S rDNA sequences.

The branching pattern, which is rooted by using *A. acidocaldarius* as the outgroup, was generated by the neighbor-joining method based on an alignment of 1,057 nucleotides of the 16S rDNA. Bootstrap values are given for each node having 70% or greater support. Bar=0.01 nucleotide substitution per site. The accession number is shown in parentheses. The sequence data of *B. horikoshii* DSM 8719^T were obtained from Dr. F. A. Rainey.

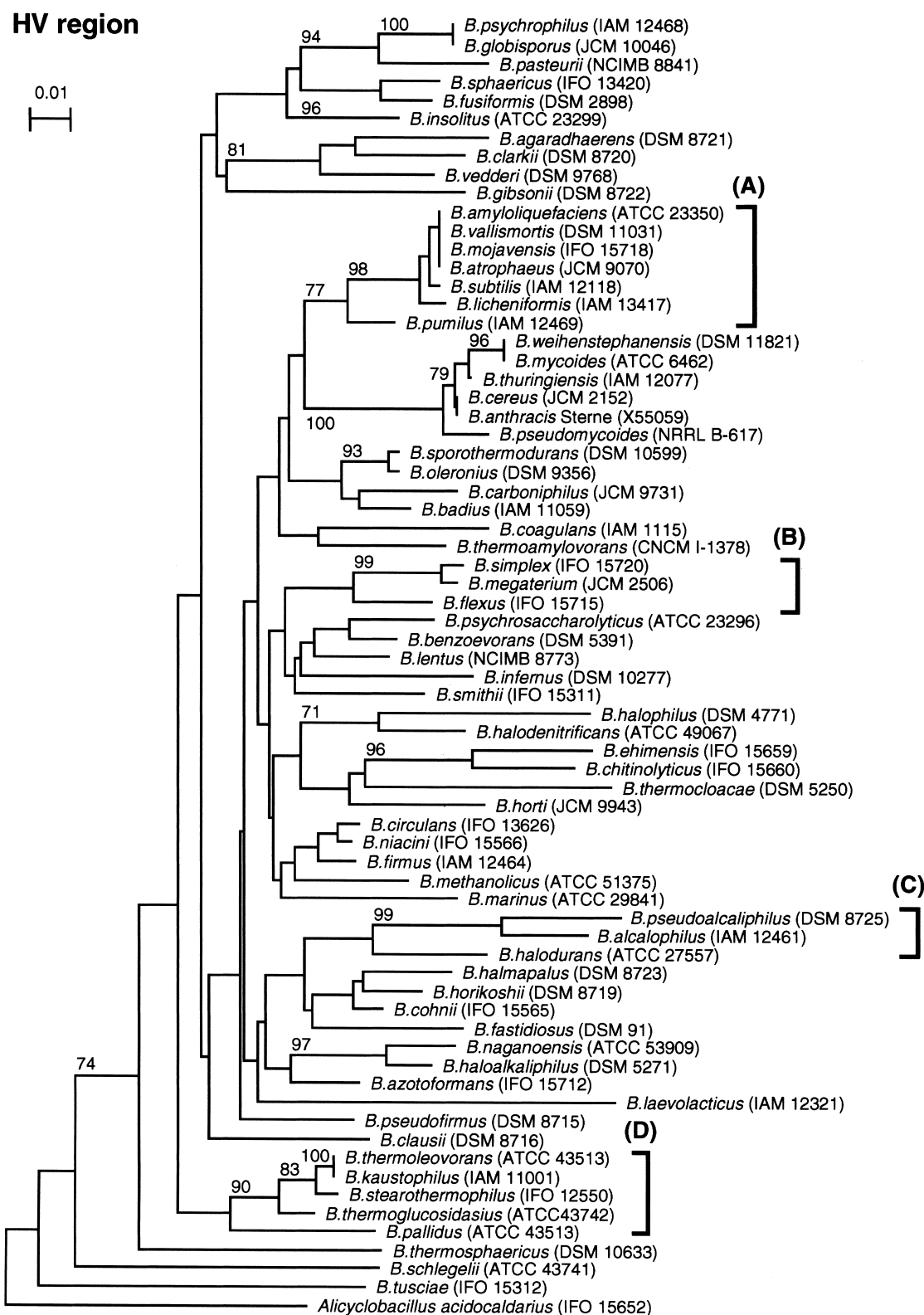


Fig. 3. Phylogenetic relationships of *Bacillus* species based on sequences of the HV region.

The branching pattern, which is rooted by using *A. acidocaldarius* as the outgroup, was generated by the neighbor-joining method based on an alignment of 219 nucleotides of the HV region. Bootstrap values are given for each node having 70% or greater support. Bar=0.01 nucleotide substitution per site. The accession number is shown in parentheses.

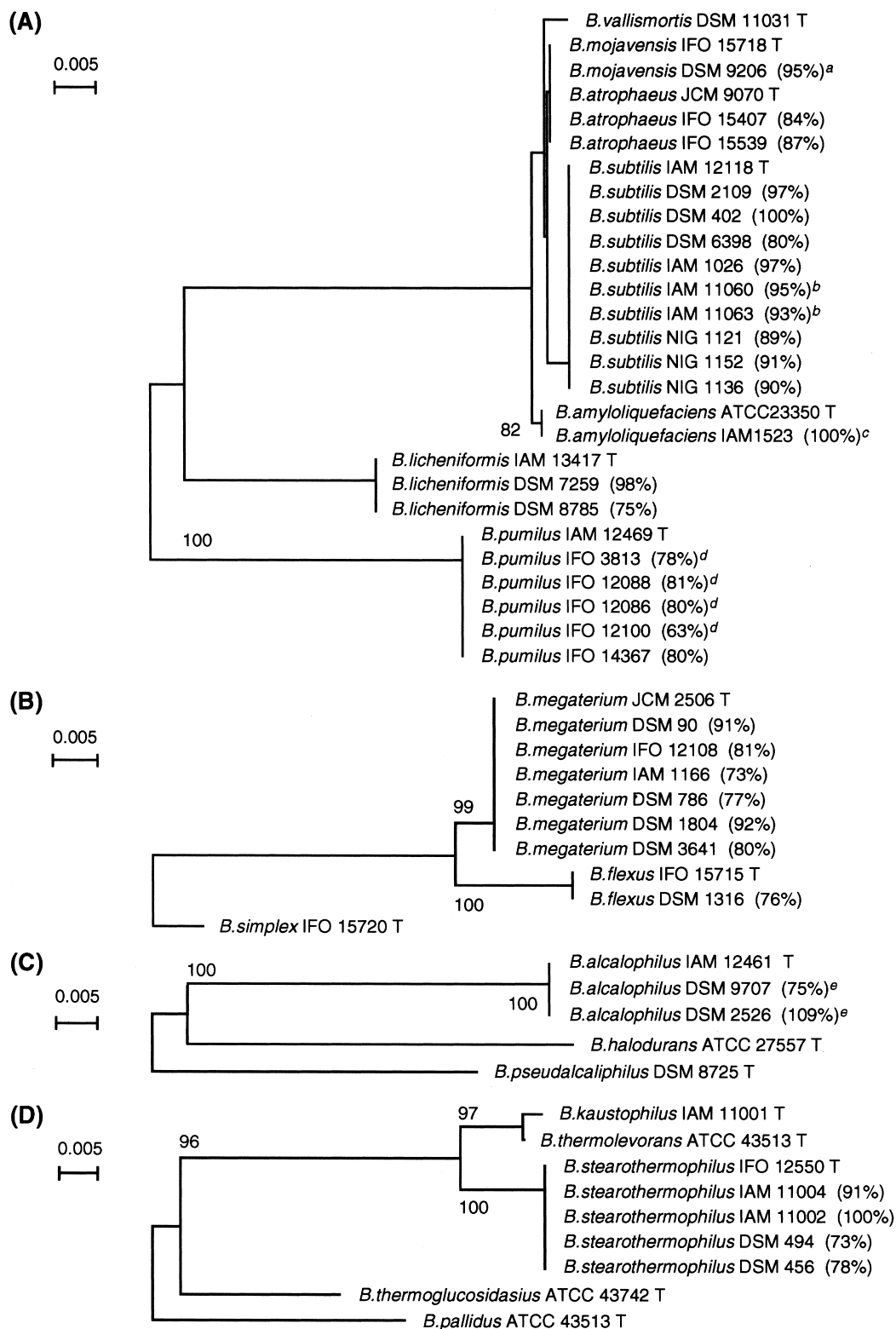


Fig. 4. Similarity pattern of *Bacillus* strains based on the sequence analysis of the HV region.

(A) *B. subtilis* cluster, (B) *B. megaterium* cluster, (C) *B. alcalophilus* cluster, (D) *B. stearothermophilus* cluster. The number indicates bootstrap values. Bar=0.005 nucleotide substitution per site. The values (%) in parentheses indicate the results of DNA-DNA hybridization to identical type strain, and the values are obtained from the average of five determinations. ^a Nakamura (1989); ^b Robert et al. (1994); ^c O'Donnell et al. (1980); ^d Seki et al. (1978); ^e Nielsen et al. (1995).

tested). Agarose gel electrophoresis of the PCR products showed that only one predicted fragment (approx. 320 bp) was amplified in all the strains (data not shown). The sequences of PCR products were in good agreement with the HV region of the identical strain. It was therefore confirmed that the designated primers amplified the HV region specifically.

Conservation of the HV region within species

To corroborate conservation of the HV region within species, we compared 51 sequences of the HV region from *Bacillus* strains belonging to the clusters A, B, C, and D in Figs. 3 and 4. The sequences of the HV regions of the strains tested were demonstrated to be in good agreement with those of their corresponding type strain. Moreover, the HV region of each species was not generally superimposable on that of other species. Thus it was revealed that the HV region was highly conserved within species.

As described above, it was clearly demonstrated that the HV region was different among species and was conservative within species. Consequently, the HV region will be applicable for facile identification or grouping of species in the genus *Bacillus*.

Discussion

In this paper, we proposed a facile method for the identification or grouping of species in the genus *Bacillus* based on the sequence differences of the HV region. The identification of an unknown *Bacillus* strain will be performed by selecting already known *Bacillus* species of which the sequence of the HV region has the highest homology with the unknown strain, using the BLAST database search system.

Exceptionally, no sequence differences were found between *B. atrophaeus* and *B. mojavensis* in either the HV region (sequence similarity 100%) or the complete 16S rDNA gene (sequence similarity 99.3%). Accordingly, it is necessary for the identification of these species that oxidase activity, which is the only key distinctive characteristic (Nakamura, 1989; Robert et al., 1994) be examined besides the analysis of the HV region. Although these two species were not clearly differentiated by the HV region, these strains will be dealt with one group, and this group can be distinguished from other species by the sequence differences of the HV region. Furthermore, *B. anthracis*, *B. cereus*, and *B. thuringiensis* belonging to the *B. cereus* cluster (Fig.

2) were not distinguished from each other, but they formed one group with the same sequence of the HV region (data not shown). This group was distinguishable from other species similarly to the case of *B. mojavensis* and *B. atrophaeus*.

We also attempted to apply these methods to other genera (*Alicyclobacillus acidocaldarius* IFO 15652^T, *Aneurinibacillus aneurinilyticus* IAM 1077^T, *Brevibacillus brevis* IFO 15304^T, *Gracilibacillus dipsosauri* DSM 11125^T, *Paenibacillus polymyxa* IAM 13419^T, *Salibacillus salexigens* DSM 11483^T, and *Virgibacillus panthothenticus* IAM 11061^T). We found that the amplified fragments were in agreement with the 5' end region of their 16S rDNA, which corresponded to the HV region of genus *Bacillus*. These seven sequences were different from each other; moreover, they were not found to be identical to any of the *Bacillus* strains. From these observations, it may be possible to identify or to group these bacteria in the same way as genus *Bacillus*. Further investigations are in progress, using many strains to ascertain whether these methods are applicable to the identification or grouping of species belonging to *Bacillaceae* besides genus *Bacillus*.

Acknowledgments

We wish to thank Dr. A. Yokota (University of Tokyo, Japan) for his valuable advice, and R. Katakawa and C. Yagoshi for their kind assistance. We also wish to thank Dr. L. K. Nakamura (NRRL, USA) for providing a strain of *B. pseudomycoides* (NRRL B-617), Dr. B. Patel (Griffith University, Australia) for providing chromosomal DNA of *B. thermoamylovorans* (CNCM I-1378), and Dr. F. A. Rainey for permitting the utilization of sequence data used in this study.

References

- Ash, C., Farrow, J. A. E., Wallbanks, S., and Collins, M. D. (1991) Phylogenetic heterogeneity of the genus *Bacillus* revealed by comparative analysis of small-subunit-ribosomal RNA sequence. *Lett. Appl. Microbiol.*, **13**, 202–206.
- Ash, C., Priest, F. G., and Collins, M. D. (1993) Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks, and Collins) using a PCR probe test. *Antonie Leeuwenhoek*, **64**, 253–260.
- Ezaki, T., Hashimoto, Y., and Yabuuchi, E. (1989) Fluorometric deoxyribonucleic 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.
- Giffel, M. C., Beumer, R. R., Klijn, N., Wagendrop, A., and Rom-

- bouts, F. M. (1997) Discrimination between *Bacillus cereus* and *Bacillus thuringiensis* using specific DNA probes based on variable regions of 16S rRNA. *FEMS Microbiol. Lett.*, **146**, 47–51.
- Heyndrickx, M., Lebbe, L., Kersters, K., De Vos, P., Forsyth, G., and Logan, N.A. (1998) *Virgibacillus*: A new genus to accommodate *Bacillus pantothenicus* (Proom and Knight 1950). Emended description of *Virgibacillus pantothenicus*. *Int. J. Syst. Bacteriol.*, **48**, 99–106.
- Ian, H., Yu, D., and Peter, C. B. T. (1995) Differentiation of *Bacillus anthracis* and other 'Bacillus cereus group' bacteria using IS231-derived sequences. *FEMS Microbiol. Lett.*, **128**, 113–118.
- Kataoka, M., Ueda, K., Kudo, T., Seki, T., and Yoshida, T. (1997) Application of the variable region in 16S rDNA to create an index for rapid species identification in the genus *Streptomyces*. *FEMS Microbiol. Lett.*, **151**, 249–255.
- Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.*, **16**, 111–120.
- 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.
- Nielsen, P., Fritze, D., and Priest, F. G. (1995) Phenetic diversity of alkaliphilic *Bacillus* strains: Proposal for nine new species. *Microbiology*, **141**, 1745–1761.
- O'Donnell, A. G., Norris, J. R. W., Claus, D., Kaneko, T., Logan, N. A., and Nozaki, R. (1980) Characterization of *Bacillus subtilis*, *Bacillus pumilus*, *Bacillus licheniformis*, and *Bacillus amyloliquefaciens* by pyrolysis gas-liquid chromatography, deoxyribonucleic acid-deoxyribonucleic acid hybridization, biochemical tests, and API system. *Int. J. Syst. Bacteriol.*, **30**, 448–459.
- Robert, M. S., Nakamura, L. K., and Cohan, F. M. (1994) *Bacillus mojaviensis* 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.
- Sadaie, Y., Yata, K., Fujita, M., Sagai, H., Itaya, M., Kasahara, Y., and Ogasawara, N. (1997) Nucleotide sequence and analysis of the *phoB-rrnE-groESL* region of the *Bacillus subtilis* chromosome. *Microbiology*, **143**, 1861–1866.
- Saitou, N. and Nei, M. (1987) The neighbor-joining method: A new method for reconstruction phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.
- Seki, T., Chung, C., Mikami, H., and Oshima, Y. (1978) Deoxyribonucleic acid homology and taxonomy of the genus *Bacillus*. *Int. J. Syst. Bacteriol.*, **28**, 182–189.
- Shida, O., Takagi, H., Kadowaki, K., and Komagata, K. (1996) Proposal for two new genera, *Brevibacillus* gen. nov. and *Aneurinibacillus* gen. nov. *Int. J. Syst. Bacteriol.*, **46**, 939–946.
- Shida, O., Takagi, H., Kadowaki, K., Uda, S., Nakamura, L. K., and Komagata, K. (1995) Proposal of *Bacillus reuszeri* sp. nov., *Bacillus formosus* sp. nov., nom. rev., and *Bacillus borstelensis* sp. nov., nom. rev. *Int. J. Syst. Bacteriol.*, **45**, 93–100.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence weighing, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4682.
- Timothy, C. R., Malcolm, J., Suresh, D. P., and Mitchell, C. (1994) Identification of capsule-forming *Bacillus anthracis* spores with the PCR and a novel dual-probe hybridization format. *Appl. Environ. Microbiol.*, **60**, 1622–1625.
- Waino, M., Tindall, B. J., Schumann, P., and Ingvorsen, K. (1999) *Gracilibacillus* gen. nov., with description of *Gracilibacillus halotolerans* gen. nov., sp. nov.; transfer of *Bacillus dipsosauri* to *Gracilibacillus dipsosauri* comb. nov., and *Bacillus salexigens* to the genus *Salibacillus* gen. nov., as *Salibacillus salexigens* comb. nov. *Int. J. Syst. Bacteriol.*, **49**, 821–831.
- Wisotzkey, J. D., Jurtshuk, P., Jr., Fox, G. E., Deinhard, G., and Poralla, K. (1992) Comparative sequence analyses on the 16S rRNA (rDNA) of *Bacillus acidocaldarius*, *Bacillus acidoterrestris*, and *Bacillus cycloheptanicus* and proposal for creation of new genus, *Alicyclobacillus* gen. nov. *Int. J. Syst. Bacteriol.*, **42**, 263–269.
- Yamada, Y., Hoshino, K., and Ishikawa, T. (1997) The phylogeny of acetic acid bacteria based on the partial sequences of 16S ribosomal RNA: The elevation of the subgenus *Gluconoacetobacter* to the genetic level. *Biosci. Biotech. Biochem.*, **61**, 1244–1251.
- Yamazaki, K., Okubo, T., Inoue, N., and Shinano, H. (1997) Randomly Amplified Polymorphic DNA (RAPD) for rapid identification of spoilage bacterium *Alicyclobacillus acidoterrestris*. *Biosci. Biotech. Biochem.*, **61**, 1016–1018.