

Cellular and Molecular Biology

E-ISSN: 1165-158X / P-ISSN: 0145-5680

www.cellmolbiol.org



Sequence comparison of *phoR*, *gyrB*, *groEL*, and *cheA* genes as phylogenetic markers for distinguishing *Bacillus amyloliquefaciens* and *B. subtilis* and for identifying *Bacillus* strain B29

C. Yu^{1,2#}, J. Jin^{3#}, L-Q. Meng^{1,2}, H-H. Xia^{1,2}, H-F. Yuan⁴, J. Wang^{1,2}, D-S. Yu^{1,2}, X-Y. Zhao^{1,2*}, C-Q. Sha^{4*}

¹ Institute of Microbiology, Heilongjiang Academy of Sciences, Harbin , China

² Institute of Advanced Technology, Heilongjiang Academy of Sciences, Harbin, China

³ Department of biochemistry and molecular biology, Harbin Medical University, Harbin, China

⁴ Heilongjiang Academy of Sciences, Heilongjiang Province, China

Correspondence to: zhao1612@hotmail.com, shachangqing@163.com

Received October 19, 2016; Accepted May 7, 2017; Published May 20, 2017

[#] These authors contributed equally to this work

Doi: http://dx.doi.org/10.14715/cmb/2017.63.5.4

Copyright: © 2017 by the C.M.B. Association. All rights reserved.

Abstract: Given the close genetic relationship between Bacillus amyloliquefaciens and B. subtilis, distinguishing the two solely based on their physiological and biochemical characteristics and 16S rRNA sequences is difficult. Molecular identification was used to discover suitable genes for distinguishing the two bacteria, and to identify the bio-controlling strain B29, due to molecular identification has been paid more and more attention. The similarity of four genes, cheA, gyrB, groEL and phoR, of the two species was compared by the software BLASTN and MAGA, and phylogenetic tree was constructed. The B29 strain was re-identified by using the screened genes. The similarities of the four genes, gyrB, groEL, cheA and phoR, of the two species were 93-95%, 82-84%, 76-78% and 76-77%, respectively. The homologies of the four genes of the strain B29 and the strains of B. amyloliquefaciens strains were more than 95%. We determined how well the phoR and cheA genes could be used to differentiate B. amyloliquefacien and B. subtilis. The previously isolated biological control strain B29, initially classified as B. subtilis, was re-classified as B. amyloliquefaciens. Our data indicate that other than the phoR gene, the cheA gene might be a useful phylogenetic marker for differentiating B. subtilis and B. amyloliquefaciens.

Key words: Biological control; Bacillus subtilis; Bacillus amyloliquefaciens; Phylogeny; Molecular classification.

Introduction

The genus Bacillus contains gram-positive bacteria and it is widely used in biological control given its convenient cultivation and storage (1). Two Bacillus species, B. amyloliquefaciens and B. subtilis, produce gibberellin, indole acetic acid (2), extracellular phytase (3), chitinases (4,5), and antifungal peptides (6-9), biologically active molecules that promote plant growth and thus improve crop production (10-13). B. amyloliquefaciens and B. subtilis share similar phenotypic characteristics and are genetically closely related. Previously, B. amyloliquefaciens was considered a subspecies of B. subtilis before it was identified as a separate species in 1967.

Given the close genetic relationship between B. amyloliquefaciens and B. subtilis, distinguishing the two based solely on their physiological and biochemical characteristics is very difficult; incorrect classification of related strains might affect their use in biological control. Currently, there are no sensitive methods/gene sequences for distinguishing B. amyloliquefaciens and B. subtilis, which hampers the classification of new isolated strains.

Analysis of the 16S rRNA gene sequence is the most frequently used method for constructing bacterial phy-

logenetic relationships and for distinguishing species with up to 98.7% sequence similarity (15). However, its use is limited in closely related species due to their high sequence similarity (99–100%) (16). Thus, many researchers are discovering more phylogenetic genes to distinguish bacteria at the species level. Generally, housekeeping genes encoding proteins involved in the central metabolism of organisms evolve faster than the 16S rRNA gene does, and can be used to distinguish closely related species (17). Presently, several genes are used to identify bacterial species or for identification within specific bacterial genera. The gyrB genes encode the subunit B protein of DNA gyrase, a type II DNA topoisomerase that plays an important role in DNA replication and prokaryotic transcription that is widely distributed in bacteria (18,19). The gyrB gene sequence has been used in the phylogenetic analysis of many bacterial genera, including Bacillus (20,21). In bacteria, the groEL gene encodes a 60-kD heat shock protein that is involved in maintaining normal physiological functions, and is a well-established phylogenetic marker for detecting many bacteria (22,23). The phoR gene encodes a histidine kinase belonging to the histidine kinase protein superfamily expressed in many bacteria (24); phoR is universally distributed in bacteria and is a single-copy gene (25) A recent study indicated

that the phoR gene is a useful marker for differentiating B. subtilis subspecies (25).

The cheA gene encodes a histidine kinase that couples environmental stimuli to bacterial swimming motions (26). CheA converts the sensory signal detected by transmembrane chemoreceptors into a cytosolic chemical signal via autophosphorylation. The phosphorylated CheA induces CheY phosphorylation; phosphorylated CheY is distributed in the cytoplasm and it interacts with flagellar motors, resulting in reversed flagellar rotation and cell chemotactic movement. Six proteins are involved in this action; CheA is the pivotal protein for chemotactic movement.

The biological control Bacillus strain B29 was isolated from major soybean varieties in Heilongjiang Province, China. B29 has biological activity against many soil-borne fungal diseases and exhibits >90% efficacy against Fusarium oxysporum (27). Based on its physiological and biochemical characteristics and 16S rRNA sequence, B29 was previously identified as B. subtilis. However, following whole-genome sequencing, we found that more genes in B29 exhibited high similarity with that of B. amyloliquefaciens. To classify the species for B29 accurately, we first compared the usefulness of several candidate markers for differentiating B. amyloliquefaciens and B. subtilis, and then determined the phylogenetic taxonomy for B29. As one gene may provide insufficient information on the genomic differences between closely related species, we used four gene sequences to distinguish B. amyloliquefaciens and B. subtilis.

Materials and Methods

Strains

We used two type strains: B. subtilis subspecies (subsp.) subtilis strain (str.) 168 (GenBank Accession No. GI: 728882887, CP010052.1) and B. amyloliquefaciens subsp. plantarum str. FZB42 (GenBank Accession No. GI: 154350369, CP000560.1), for the sequence comparison. The B29 experimental strain was isolated from major soybean varieties and preserved in our laboratory (preservation no.: CGMCC 0752). B29 is a highly effective broad-spectrum biological control strain that inhibits the growth of many fungi that infect plants, including F. oxysporum f. cucumerinum, Rhizoctonia solani, and Pythium spp. We sequenced the entire B29 genome and labeled its genes.

The control strains were the strains containing the whole-genome sequence of, which are available from the National Center for Biotechnology Information (NCBI) GenBank database, and specific genes had been labeled (Table 1 lists the NBCI accession numbers).

Comparison between genes

The 16S rRNA, phoR, gyrB, cheA, and groEL gene sequences of the above-mentioned strains were obtained from the NCBI GenBank database. We compared the sequence similarity among the five genes between B. subtilis subsp. subtilis str. 168 and B. amyloliquefaciens subsp. plantarum str. FZB42 and among B. subtilis or B. amyloliquefaciens using BLASTN 2.2.30+ (NCBI). Then, we compared the sequence similarity of the five genes between B. subtilis subsp. subtilis str. 168

and other B. amyloliquefaciens strains and between B. amyloliquefaciens subsp. plantarum str. FZB42 and B. subtilis strains (from NCBI) using BLASTN 2.2.30+.

Species identification and phylogenetic analysis of B29 based on cheA and phoR gene sequences

The sequence comparison between B29 and the B. amyloliquefaciens and B. subtilis strains was aligned using BLAST (NCBI). Phylogenetic analysis was performed using MEGA 6.0 (http://www.megasoftware.net/); the genetic distances were calculated and phylogenetic trees were constructed using the neighborjoining method with bootstrap values based on 1000 replications.

Results

Comparison between B. subtilis subsp. subtilis str. 168 and B. amyloliquefaciens subsp. plantarum str. FZB42

There was only 80% genome-level similarity between B. subtilis subsp. subtilis str. 168 and B. amyloliquefaciens subsp. plantarum str. FZB42 (Figure 1), further indication that B. subtilis and B. amyloliquefaciens are different species. Table 2 lists the analysis results of the similarities between the five genes in the two type strains. There was 99% similarity between the 16S rRNA genes in the two type strains, followed by 93% similarity between the groEL genes and 82% similarity between the gyrB genes; the cheA and phoR genes had the least similarity (both 77% similarity).

Comparison of genes among species and between type strains and control strains

Next, we compared the five genes among one species and between one type strain and other control strains. The genes all had 99% similarity in B. subtilis and 98–99% similarity in B. amyloliquefaciens. There was lower similarity between the phoR, gyrB, cheA, and groEL genes in B. subtilis subsp. spizizenii str. **Table 2.** Similarity between five genes in B. subtilis str. 168 and B. amyloliquefaciens str. FZB42.

| Gene | Similarity (%) | |
|--------|----------------|--|
| 16sRNA | 99 | |
| phoR | 77 | |
| gyrB | 82 | |
| cheA | 77 | |
| groEL | 93 | |



Figure 1. Dot matrix comparison of the sequences of B. subtilis subsp. subtilis str. 168 and B. amyloliquefaciens subsp. plantarum str. FZB42.

| ÷ | ž |
|---|---|
| - | - |
| ÷ | Ę |
| 5 | |
| 5 | 5 |
| 5 | 1 |
| 10 | 3 |
| 0 | 2 |
| | aL |
| 40 | 10 |
| 4 | į |
| 5 | 2 |
| 4 | 1 |
| 100 | 2 |
| 9/11 | |
| £ | ž |
| 4 | |
| 6 | Ì |
| 0 | 3 |
| ĕ. | ž |
| ć | 3 |
| Ē | 2 |
| έ | ŋ |
| Ę | 5 |
| Ę | Ę |
| ē | 5 |
| 00 | ž |
| đ | 2 |
| 0 |) (I) |
| 170 | 2 |
| £ | - |
| Ċ | 5 |
| Ę | 5 |
| .10 | |
| 10 | 1 |
| 5 | 1 |
| 5 | 5 |
| C |) |
| 2 | |
| e | |
| Q | ċ |
| . 5 | 3 |
| | |
| 0.4 | 3 |
| ctra | 3 TI C 1 |
| ol etra | 101 2010 |
| ntrol etra | 2116 10 111 |
| ontrol etra | |
| d control etra | The TOTITION IN |
| and control etra | |
| e and control etra | |
| ins and control stra | זווני מוות אוומי מוווי |
| trains and control stra | |
| strains and control stra | |
| me strains and control stra | be summe and control suc |
| tyme strains and control stra- | type suams and control suc |
| trine strains and control stra | on type suams and common suc |
| teen tyme strains and control stra | 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 |
| tween twne strains and control stra | itw con type sutating and control suc |
| between type strains and control stra | ociw con type suams and control suc |
| d between type strains and control stra | in octween type suams and control suc |
| and hetween type strains and control stra | and octword type suams and control suc |
| ss and between type strains and control stra | c_{2} and octword type submits and control suc |
| ries and hetween type strains and control stra | $c_{1}c_{2}$ and $c_{1}c_{2}$ $c_{2}c_{3}$ and $c_{1}c_{1}c_{2}$ |
| secies and between type strains and control stra | protes and octween type suants and control suc |
| snecies and between type strains and control stra | , speeres and oetween type suants and control suc |
| ng species and between type strains and control stra | ing spectra and octween type analitis and control and |
| tong species and between type strains and control stra | |
| among species and between type strains and control stra | among speets and octween type suams and common suc |
| s among species and between type strains and control stra | es annous speetes and octween type suams and connot suc |
| nes among species and henveen type strains and control stra | - 1103 allouig speetes and octween type suams and control suc |
| renes among species and between type strains and control stra | בעווכם מוווטווע סףכטכס מוות טכוא ככוו ואףכ סתמוווס מוות כטוונוטו סוונ |
| e denes among species and between type strains and control stra | יד באוואס מוווטווד אוויט מווע טרוא כרוו גארי איז אוויט מווע באוואס אוויט אוויט אוויט אוויט אוויט אוויט אוויט אי |
| five genes among species and between type strains and control stra | 11 v genues among speeds and occaven type summer and control suc |
| of five genes among species and between type strains and control stra | 11 11 VC genues allioning spectres ally occurrent type suballits ally collition sate |
| s of five genes among species and between type strains and control stra | If of inverse allious species and occurrent type submits and control suc |
| on of five genes among species and between type strains and control stra | 2011 OT 1110 SOUCE MILIOUS SPECIES AND OCIMENT 12 PC SUMINIS AND COURTS 1 2010 |
| rison of five genes among species and between type strains and control stra | נוסטון טו ווער צעונט מוווטווצ סעיכונט מווע טכואיכטו ואףר שנומוש מווע נטוונטו שנינ |
| ustrison of five genes smong species and between type strains and control stra | המוזטטוו טו ווידי צירוונט מוווטווצ סףיכיונט מווע טינויינענו ואיין אין אימוווט מווע כטוועטו סווג |
| muaricon of five acres among energies and between type strains and control stra | נווף אנוסטוו טו ווזיך צרוונים אווטווצ סףיכטנים אווע טינוייעיטו ואיך שעמוווס אווע כטוועטו סווג |
| Tommarison of five cenes among energies and hetwaen type strains and control stra | ϕ |
| Communicon of five remes among energies and hetwaen type strains and control stra | • Comparison of its genes anong species and octaven type stands and control suc |
| Comparison of five genes among species and between type strains and control strails | cy. Comparison of nye genes among speeces and octaven type suams and control suc |
| (A) Communism of five remes among species and between type strains and control stra | (x). Comparison of the genes among speece and octaven type shares and control suc |
| 1 (A) Communicon of five menes among energies and hetwaen type strains and control stra | . (A). Companison of new genes annous species and octween type suams and control suc |

| Domilalizations | NCDI accession number | Compar | ison among I | 3. amyloliq | uefaciens (| (%) | Compariso | n with B. sı | ubtilis str. 1 | 168 (%) | |
|--|-----------------------|--------------|--------------|--------------|-------------|-------------|--------------|-----------------------|----------------------------|-----------|------------|
| b. amyronqueracrens | NCBI accession number | 16S rRN | A gyrB | groEl | cheA | phoR | 16S rRNA | gyrB | groEL | cheA | phoR |
| B. a. str. FZB42 | CP000560.1 | | | | | | 66 | 82 | 92 | LL | 77 |
| B. a. str. L-H15 | CP010556.1 | 100 | 66 | 66 | 98 | 98 | 66 | 82 | 93% | 76 | 77 |
| B. a. DSM7 | FN597644.1 | 66 | 66 | 97 | 95 | 94 | 66 | 81 | 93% | 76 | 77 |
| B. a. subsp. amyloliquefaciens KHG19 | CP007242.1 | 66 | 66 | 66 | 66 | 98 | 66 | 81 | 93 | 77 | 77 |
| B. a.CC178, complete genome | CP006845.1 | 100 | 100 | 100 | 100 | 100 | 66 | 82 | 93 | LL | 77 |
| B. a. XH7, complete genome | CP002927.1 | 66 | 96 | 97 | 95 | 94 | 66 | 81 | 93 | 76 | 77 |
| B. a. subsp. plantarum TrigoCor1448 | CP007244.1 | 66 | 66 | 98 | 66 | 66 | 66 | 82 | 93 | LL | 77 |
| B. a. LFB112 | CP006952.1 | 66 | 66 | 66 | 98 | 98 | 66 | 82 | 93 | 76 | 77 |
| B. a. subsp. plantarum NAU-B3 | HG514499.1 | 66 | 66 | 66 | 98 | 66 | 66 | 82 | 94 | 76 | 76 |
| B. a. subsp. plantarum UCMB5033 | HG328253.1 | 66 | 66 | 66 | 66 | 98 | 66 | 82 | 63 | | 76 |
| B. a. subsn. nlantariim UCMB 5036 | HF563562_1 | 66 | 66 | 66 | 66 | 66 | 66 | 28 | 63 | | |
| B a subso plantariim YAU B9601-Y2 | HE774679.1 | 66 | 66 | 66 | 98 | 66 | 66 | 2 2 7 8 | 66 | 76 | 76 |
| B a subsn nlantariim CAU B946 | HE6171591 | 66 | 66 | 98 | 98 | 98 | 66 | 8 | 63 | 26 | |
| B a IT-45 | CP004065 1 | 66 | 66 | 66 | 80 | 86 | 66 | 28 | 66 | 76 | |
| B a subsn nlantariim AS43 3 | CP003838 1 | 00 | 66 | 66 | 66 | 66 | 66 | 20 | 6 | | |
| D. a. suber vlantani ICMR 5112 | HC22825A 1 | 100 | 00 | 00 | 00 | 00 | 00 | 70 70 | 03 | | 76 |
| B. a. SORQ complete genome | | 00 | 00 | 00 | 00 | 00 | 00 | 20 20 20 | 00 | 76 | |
| D. a. TADA | | 00 | 00 | 10 | 05 | 070 | 00 | 81 8 | 03 | 26 | |
| D. a. 17200 R a 113 chromosoma | CD002627.1 | 00 | 06 | 10 | 50 | 10 | 00 | 81 81 | 02 | 0/ | |
| | | 00 | 00 | 00 | 00 | 1 00 | 00 | 10 | 00 | 91 | |
| Average | | 66 | 66 | 66 | 06 | 20 | 66 | 70 | C | 0/ | |
| | | Comnarison : | առոց Ք շոհ | tilie 168 (% | | | Comnarison v | ith R a F | ZR42 (%) | | |
| B. subtilis | NCBI accession number | 16S FRNA | ovrR | oroF.L. | ched | nhoR | 16S rRNA | ovrR | oroF.I. | And | nhoR |
| R s sulten sultifie str 168 | CP010052 1 | | â | â | | | 00 | 2072 2020 | 03 | 17 | 77 |
| B s str PS837 | CP010053 1 | 100 | 100 | 100 | 100 | 100 | 66 | 22 | 63 | <i>LL</i> | |
| B s subsn subtilis 6051-HGW | CP003379 1 | 100 | 100 | 100 | 100 | 100 | 00 | 200 | 93 | <i>LL</i> | |
| B & etr CG6 | | 00 | 00 | 00 | 00 | 00 | 00 | 200 | 60 | | |
| $\mathbf{D} = \mathbf{D} \mathbf{E} \mathbf{C} \mathbf{T} \mathbf{C} 1 2 \mathbf{D} \mathbf{N} 1 \mathbf{A}$ | A DO17405 1 | 100 | 100 | | 100 | 100 | 00 | 10 | 50 | | |
| D, S. D Let I (11) D (14) D | | 100 | 100 | 00 | 100 | 100 | 00 | 70 | | - 10 | - [|
| B. S. suosp. spizizenii sur. W 23 | CFU02183.1 | 99 00 | 4 r | 90 | 50 | 4 4 5 | 99 | 70 | <i>c c c c c c c c c c</i> | 0/0 | - C - C |
| B. S. subsp. spizizenii 10-B-10 | CP002905.1 | 99 | 56 100 | 80 20 | с <u>е</u> | 93 100 | 99 00 | 78 | 95 07 | 8/ | |
| B. S. BESI /003 DNA | AP012496.1 | 100 | 100 | 91 | 100 | 100 | 99 0.0 | 78 | с <u>к</u> 0 | | |
| B. s. subsp. subtilis str. 3NA | CP010314.1 | 100 | 100 | 100 | 100 | 100 | 99 00 | 7.8 | 93 20 | | |
| B. s. PY79 | CP006881.1 | 66 | 100 | 100 | 100 | 100 | 66 | 82 | 93 | 77 | 77 |
| B. s. BSn5 | CP002468.1 | 66 | 66 | 66 | 66 | 66 | 66 | 82 | 93 | LL | 76 |
| B. s. subsp. subtilis str. BAB-1 | CP004405.1 | 66 | 66 | 98 | 66 | 66 | 66 | 82 | 94 | LL | 76 |
| B. s. subsp. subtilis RO-NN-1 | CP002906.1 | 66 | 66 | 98 | 66 | 98 | 66 | 82 | 93 | LL | 77 |
| B. s. subsp. subtilis str. OH 131.1 | CP007409.1 | 66 | 66 | 66 | 66 | 66 | 66 | 82 | 93 | LL | 77 |
| B. s. XF-1 | CP004019.1 | 100 | 66 | 98 | 66 | 66 | 66 | 82 | 94 | LL | 76 |
| B. s. subsp. natto BEST195 DNA | AP011541.2 | 66 | 66 | 98 | 66 | 66 | 66 | 82 | 94 | LL | <i>LL</i> |
| B. s. subsp. subtilis str. AG1839 | CP008698.1 | 100 | 100 | 100 | 100 | 100 | 66 | 82 | 93 | LL | 77 |
| B. s. subsp. subtilis str. BSP1 | CP003695.1 | 100 | 66 | 98 | 66 | 66 | 66 | 82 | 93 | LL | 77 |
| B. s. subsp. spizizenii str. NRS 231 | CP010434.1 | 66 | 94 | 98 | 95 | 94 | 66 | 82 | 93 | 78 | 77 |
| B. s. subsp. str. JH642 | CP007800.1 | 100 | 100 | 100 | 100 | 100 | 66 | 82 | 93 | LL | 77 |
| B. s. OB928 | CP003783.1 | 66 | 100 | 100 | 100 | 100 | 66 | 82 | 93 | LL | 77 |
| Average | | 66 | 66 | 66 | 66 | 66 | 66 | 82 | 93 | 77 | 77 |



V22 D subtilie suber anizizanii TU D 10 and D sub

W23, B. subtilis subsp. spizizenii TU-B-10, and B. subtilis subsp. spizizenii strain NRS 231 than in the other strains. There was lower similarity between the phoR, gyrB, cheA, and groEL genes of B. amyloliquefaciens XH7, B. amyloliquefaciens TA208, and B. amyloliquefaciens LL3 chromosome as compared to that of the other B. amyloliquefaciens strains.

Comparison of the control strains with B. subtilis subsp. subtilis str. 168 and B. amyloliquefaciens subsp. plantarum str. FZB42 revealed that the 16S rRNA gene in both type strains shared 99% similarity with the other B. subtilis and B. amyloliquefaciens strains. There was 76–77%, 82%, 77–78%, and 93–94% similarity between the phoR, gyrB, cheA, and groEL genes, respectively, of B. subtilis subsp. subtilis str. 168 and the other B. amyloliquefaciens strains.

There was 76–77%, 81–82%, 76–77%, and 93–95% similarity between the phoR, gyrB, cheA, and groEL genes, respectively, of B. amyloliquefaciens subsp. plantarum str. FZB42 and the other B. subtilis strains.

Phylogenetic tree construction

The above data suggest that the 16S rRNA gene cannot be used to differentiate B. amyloliquefaciens and B. subtilis, nor is the groEL gene a good gene for phylogenetic tree construction, given the high level of similarity of the gene between the two species. Although there was 81–82% similarity between the gyrB genes of the two species, the 80% similarity between B. amyloliquefaciens subsp. plantarum str. FZB42 and B. subtilis subsp. subtilis str. 168 indicate that the gene is not ideal for distinguishing the two species. Therefore, we selected the cheA and phoR genes for the phylogenetic tree construction.

We constructed phylogenetic trees based on the cheA (Figure 2A) and phoR genes (Figure 2B). B. subtilis strains and B. amyloliquefaciens were located in two clades in both trees. The cheA gene phylogenetic tree placed B. pumilus, B. licheniformis, and B. atrophaeus in three different clades; the phoR gene phylogenetic tree placed B. atrophaeus, B. mojavensis, and B. subtilis subsp. spizizenii in three distinct clades. Three B. subtilis strains: BS-916, ATCC13592, and ATCC19217, were in the same clade as B. amyloliquefaciens, and the phoR, gyrB, groEL, and cheA genes from these strains shared close similarity with that of B. amyloliquefaciens subsp. plantarum str. FZB42. Thus, these three strains may warrant further re-classification.

Table 3 lists the comparison analysis of the phoR, gyrB, groEL, and cheA genes between B29 and the B. subtilis or B. amyloliquefaciens strains. There was >95% similarity among the four genes between B29 and the B. amyloliquefaciens strains, which was about 20% greater similarity than that between B29 and the B. subtilis strains. The cheA gene phylogenetic tree indicated a close genetic distance between B29 and B. amyloliquefaciens subsp. plantarum str. FZB42, a wide genetic

| Table 3. Similarities between four genes from B29 and other |
|---|
| strains available in the NCBI database. |

| Gene | B. subtilis (%) | B. amyloliquefaciens (%) |
|-------|-----------------|--------------------------|
| phoR | 70–77 | 95–99 |
| gyrB | 81 | 98–99 |
| groEL | 93–97 | 98–99 |
| cheA | 70–78 | 95–99 |

distance between B29 and B. subtilis subsp. subtilis str. 168, and that B29 belonged to a different sub-cluster. Thus, B29 should be classified as B. amyloliquefaciens.

Discussion

To better distinguish B. subtilis and B. amyloliquefaciens, we compared the 16S rRNA, phoR, gyrB, groEL, and cheA genes from strains of the two species. Phylogenetic trees based on the cheA and phoR genes were constructed, enabling successful differentiation of B. subtilis and B. amyloliquefaciens. Using the cheA and phoR phylogenetic trees, we determined that B29 is a strain of B. amyloliquefaciens.

Generally, a phylogenetic gene is widely distributed in most microbes as a single copy of a certain length, can be cloned easily, and predicts whole-genome relationships accurately (28). The 16S rRNA gene is the most commonly used gene for such studies; however, it is not suitable for differentiating species within the same genus because of high inter-species similarity. Consequently, more sensitive genes should be used for distinguishing closely related species.

The groEL gene is a well-established and widely used phylogenetic marker in many bacteria (22,23). However, our data revealed 93-95% similarity between the groEL genes of B. subtilis and B. amyloliquefaciens, indicating it is not a good phylogenetic marker.

The gyrB gene is a 1.2-1.4-kb single-copy housekeeping gene that has an average base substitution rate of 0.7-0.8% every 100 million years; it evolves more quickly than the 16S rRNA gene, which has an average base substitution rate of 1% every 5000 million years (18,21,29). The gyrB gene is commonly distributed in bacteria without horizontal transfer. Moreover, it contains conserved sequences and variable regions, enabling amplification from different bacterial species and intra-species differentiation, respectively. Given these advantages, the gyrB gene is better suited for differentiating closely related species than are rRNA genes. In recent years, the gyrB gene has been widely used as a phylogenetic gene in different arenas (21,30). The phoR gene encodes a histidine kinase belonging to the superfamily of histidine kinase proteins expressed in many bacteria (24) and is a phylogenetic marker that differentiates B. subtilis subspecies (25). Our data revealed 82–84% similarity between the gyrB genes and 76–78% similarity between the phoR genes of B. subtilis and B. amyloliquefaciens, suggesting these two genes might be good phylogenetic markers for differentiating the two. The cheA gene is a single-copy housekeeping gene and has not been reported as a phylogenetic gene. However our data showed that there was 76-77% similarity between the cheA genes of B. subtilis and B. amyloliquefaciens, suggesting it may be a good phylogenetic condicate for differentiating the two.

The International Committee on Systematic Bacteriology considers individuals with >70% similarity to be one species (31). However, a single gene may be insufficient for molecular differentiation. Combining ≥ 2 genes may be a future trend in genetic classification. Therefore, we used the cheA and phoR genes to differentiate B. subtilis and B. amyloliquefaciens.

groEL, and cheA genes of B. subtilis and B. amyloliquefaciens. The groEL genes shared >93% similarity and were not suitable for distinguishing the two. There was 81-82% similarity between the gyrB genes of the two species, indicating it is a potential phylogenetic gene. However, the inter-species similarities of the groEL and gyrB genes closely resemble the intra-species similarities, and they cannot be used to distinguish between species effectively. There was <70% inter-species similarity and >95% intra-species similarity among the cheA and phoR genes, indicating they are potential phylogenetic genes. Further inter-species comparisons confirmed their potential as phylogenetic genes.

CheA gene: a novel phylogenetic markers for distinguishing Bacillus affinis.

We know of no previous report involving the use of the cheA gene in taxonomy. In this study, we differentiated the closely related B. subtilis and B. amyloliquefaciens using a cheA gene-based phylogenetic tree. However, further studies are required before the cheA gene is confirmed as a phylogenetic marker for differentiating B. subtilis and B. amyloliquefaciens.

We spliced the phoR, gyrB, groEL, and cheA genes of the B29 strain based on their genome sequences, and identified B29 using cheA and phoR gene-based phylogenetic trees. We found that B29 belongs to B. amyloliquefaciens and that its previous classification was incorrect.

Microbial taxonomy is a rigorous subject: incorrectly classifying one strain can adversely affect its application. B29 is a biological control strain with high economic value and social benefit, thus it would be of great importance to classify it correctly. In China, many isolated biological control strains have been classified based only on their physiological and biochemical characteristics or the 16S rRNA gene (32), which might be inaccurate, thus classification might be incorrect. The rapid development of high-throughput sequencing techniques and bioinformatics analysis as facilitated the identification of bacterial strains, thus gene-based species identification is becoming increasingly important. Hence, discovering more housekeeping genes for taxonomy will provide a solid foundation for the further study of microorganisms.

In conclusion, we compared the similarities among five genes from B. subtilis and B. amyloliquefaciens and found that the cheA and phoR genes are potential phylogenetic genes. Phylogenetic trees based on the two genes successfully distinguished B. subtilis and B. amyloliquefaciens. Based on the phylogenetic trees, the B29 strain was classified as B. amyloliquefaciens. Other than the phoR gene, the cheA gene might be a useful phylogenetic gene for differentiating B. subtilis and B. amyloliquefaciens, although confirmation of this requires further study.

Acknowledgements

This study was supported by the Science and Technology ResearchProjects of Heilongjiang province (GZ13B017), the Basic Application Technology Research Projects of Heilongjiang province (CZ14BZXY14)., the Natural Science Foundation of Heilongjiang (H2015010).

References

There were similarities between the phoR, gyrB,

1. Reva ON, Dixelius C, Meijer J, Priest FG. Taxonomic charac-

terization and plant colonizing abilities of some bacteria related to Bacillus amyloliquefaciens and Bacillus subtilis. FEMS Microbiol Ecol, 2004; 48(2):249-59.

2. Turner JT. and Backman PA. Factors relating to peanut yield increases after seed treatment with Bacillus subtilis. J Plant Disease, 1991; 75(4): 347-353.

3. Idriss EE, Makarewicz O, Farouk A, Rosner K, Greiner R, Bochow H, et al. Extracellular phytase activity of Bacillus amyloliquefaciens FZB45 contributes to its plant-growth-promoting effect. Microbiology, 2002;148: 2097-109.

4. Manjula K. and Podile AR. Chitin-supplemented formulations improve biocontrol and plant growth promoting efficiency of Bacillus subtilis AF 1. Can J Microbiol, 2001; 47(7): 618-25.

5. Wang SL, Shih IL, Liang TW, Wang CH. Purification and characterization of two antifungal chitinases extracellularly produced by Bacillus amyloliquefaciens V656 in a shrimp and crab shell powder medium. J Agric Food Chem, 2002; 50(8): 2241-8.

6. Leifert C, Li H, Chidburee S, Hampson S, Workman S, Sigee D, et al. Antibiotic production and biocontrol activity by Bacillus subtilis CL27 and Bacillus pumilus CL45. J Appl Bacteriol, 1995; 78(2): 97-108.

7. Pinchuk IV, Bressollier P, Sorokulova IB, Verneuil B, Urdaci MC. Amicoumacin antibiotic production and genetic diversity of Bacillus subtilis strains isolated from different habitats. Res Microbiol, 2002; 153(5): 269-76.

8. Raaijmakers JM., Vlami M., and de Souza JT.. Antonie Van Leeuwenhoek, 2002; 81(1-4): 537-47.

9. Walker R, Powell AA, Seddon B. Bacillus isolates from the spermosphere of peas and dwarf French beans with antifungal activity against Botrytis cinerea and Pythium species. J Appl Microbiol, 1998; 84(5): 791-801.

10. Yang C , Hamel C, Gan Y, Vujanovic V. Bacterial endophytes mediate positive feedback effects of early legume termination times on the yield of subsequent durum wheat crops. Can J Microbiol. 2012; 58(12):1368-77

11. M Alemayehu. Use of Bacillus subtilis as Biocontrol Agent. I. Activities and Characterization of Bacillus subtilis Strains. J Plant Disease Protection, 1998; 105(2):181-197

12. Wipat A, Harwood CR., The Bacillus subtilis genome sequence: the molecular blueprint of a soil bacterium. FEMS Microbiology Ecology, 1999; 28(1): 1-9.

13. Chanway CP. Bacterial endophytes Ecological and practical implications. Sydowia, 1998; 50(2): 149-170

14. Welker NE. and. Campbell LL. Unrelatedness of Bacillus amyloliquefaciens and Bacillus subtilis. J Bacteriol, 1967; 94(4): 1124-30.
15. Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiology Today, 2006; 33(4): 152-155.

16. Nakamura LK, Roberts MS, Cohan FM. 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, 1999; 49:.1211-5. 17. Maiden, MC. Multilocus sequence typing of bacteria. Annu Rev Microbiol, 2006; 60: 561-88.

18. Huang WM. Bacterial diversity based on type II DNA topoisomerase genes. Annu Rev Genet, 1996; 30: 79-107.

19. Yamada S, Ohashi E, Agata N, Venkateswaran K. Cloning and nucleotide sequence analysis of gyrB of Bacillus cereus, B. thuringiensis, B. mycoides, and B. anthracis and their application to the detection of B. cereus in rice. Appl Environ Microbiol, 1999; 65(4): 1483-90.

20. Wang, LT., Lee FL, Tai CJ, Kasai H. Comparison of gyrB gene sequences, 16S rRNA gene sequences and DNA-DNA hybridization in the Bacillus subtilis group. Int J Syst Evol Microbiol, 2007; 57: 846-50.

21. La Duc MT., Satomi M, Agata N, Venkateswaran K. gyrB as a phylogenetic discriminator for members of the Bacillus anthraciscereus-thuringiensis group. J Microbiol Methods, 2004; 56(3): 383-94.

22.Hossain MT, Kim EY, Kim YR, Kim DG, Kong IS. Development of a groEL gene-based species-specific multiplex polymerase chain reaction assay for simultaneous detection of Vibrio cholerae, Vibrio parahaemolyticus and Vibrio vulnificus. J Appl Microbiol, 2013; 114(2): 448-56.

23. Hossain MT., Kim YO, Kong IS. Multiplex PCR for the detection and differentiation of Vibrio parahaemolyticus strains using the groEL, tdh and trh genes. Mol Cell Probes, 2013; 27(5-6): 171-5.

24. Parkinson JS, Kofoid EC. Communication modules in bacterial signaling proteins. Annu Rev Genet, 1992; 26:71-112.

25. Guo Q ,Li S,Lu X, Li B, Stummer B, Dong W, Ma P. phoR sequences as a phylogenetic marker to differentiate the species in the Bacillus subtilis group. Can J Microbiol, 2012; 58(11): 1295-305.

26. Ellefson DD, Weber U, Wolfe AJ. Genetic analysis of the catalytic domain of the chemotaxis-associated histidine kinase CheA. J Bacteriol.1997; 179(3): 825-30.

27. Li J, Yang Q, Zhao LH, Zhang SM, Wang YX, Zhao XY. Purification and characterization of a novel antifungal protein from Bacillus subtilis strain B29. J. Zhejiang Univ. Sci. B, 2009; 10(4): 264-272.

28. Zeigler DR. Gene sequences useful for predicting relatedness of whole genomes in bacteria. Int J Syst Evol Microbiol, 2003; 53:1893-900.

29. Ochman, H, Wilson AC. Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. J Mol Evol, 1987; 26(1-2): 74-86.

30. Alam SI. Bansod S, Goel AK, Singh L. Characterization of an environmental strain of Bacillus thuringiensis from a hot spring in Western Himalayas. Curr Microbiol, 2011; 62(2): 547-56.

31. Graham PH. Proposed Minimal Standards for the Description of New Genera and Species of Root-and Stem-Nodulating Bacteria. International Journal of Systematic Bacteriology, 1991; 41(4): 582-587.

32. Sun Z. and L. Yang. Screening and identification of antagonistic bacteria against Panax ginseng pathogeny fungus. Acta Phytophylacica Sinica, 2015; 42(1): 79-86.