J. Microbiol. Biotechnol. (2018), 28(10), 1760–1768
https://doi.org/10.4014/jmb.1712.12006

jmb

Complete Genome of *Bacillus subtilis* subsp. *subtilis* KCTC 3135^T and Variation in Cell Wall Genes of *B. subtilis* Strains S

Seonjoo Ahn[1,2,3], Sangmi Jun[3,4,5], Hyun-Joo Ro[3,4,5], Ju Han Kim[1,6*], and Seil Kim[2,3,5*]

[1]*Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul 03080, Republic of Korea*
[2]*Division of Chemical and Medical Metrology, Center for Bioanalysis, Korea Research Institute of Standards and Science, Daejeon 34113, Republic of Korea*
[3]*Convergent Research Center for Emerging Virus Infection, Korea Research Institute of Chemical Technology, Daejeon 34114, Republic of Korea*
[4]*Drug and Disease Target Team, Korea Basic Science Institute, Chungcheongbuk-do 28119, Republic of Korea*
[5]*Department of Bio-Analysis Science, University of Science & Technology (UST), Daejeon 34113, Republic of Korea*
[6]*Seoul National University Biomedical Informatics (SNUBI), Division of Biomedical Informatics, Seoul National University College of Medicine, Seoul 03080, Republic of Korea*

Received: December 6, 2017
Revised: August 16, 2018
Accepted: August 27, 2018

First published online
September 3, 2018

**Corresponding authors*
S.K.
Phone: +82-42-868-5619;
Fax: +82-42-868-5801;
E-mail: stapler@kriss.re.kr
J.H.K.
Phone: +82-2-740-8320;
Fax: +82-2-3673-2167;
E-mail: juhan@snu.ac.kr

Supplementary data for this paper are available on-line only at http://jmb.or.kr.

pISSN 1017-7825, eISSN 1738-8872

Copyright© 2018 by
The Korean Society for Microbiology and Biotechnology

The type strain *Bacillus subtilis* subsp. *subtilis* KCTC 3135^T was deeply sequenced and annotated, replacing a previous draft genome in this study. The *tar* and *tag* genes were involved in synthesizing wall teichoic acids (WTAs), and these genes and their products were previously regarded as the distinguishing difference between *B. s. subtilis* and *B. s. spizizenii*. However, a comparative genomic analysis of *B. subtilis* spp. revealed that both *B. s. subtilis* and *B. s. spizizenii* had various types of cell walls. These *tar* and *tag* operons were mutually exclusive and the *tar* genes from *B. s. spizizenii* were very similar to the genes from non-*Bacillus* bacteria, unlike the *tag* genes from *B. s. subtilis*. The results and previous studies suggest that the *tar* genes and the *tag* genes are not inherited after subspecies speciation. The phylogenetic tree based on whole genome sequences showed that each subspecies clearly formed a monophyletic group, while the tree based on *tar* genes showed that monophyletic groups were formed according to the cell wall type rather than the subspecies. These findings indicate that the *tar* genes and the presence of ribitol as a cell-wall constituent were not the distinguishing difference between the subspecies of *B. subtilis* and that the description of subspecies *B. s. spizizenii* should be updated.

Keywords: *Bacillus subtilis* subsp. *subtilis* KCTC 3135^T, complete genome sequence, illumina sequencing, cell-wall teichoic acid-related genes

Introduction

Bacillus subtilis has been one of the best-known and most extensively studied Gram-positive bacteria for decades [1–3]. *B. subtilis* is highly ubiquitous and widely used in both industry and laboratory [4, 5]. Genetically engineered *B. subtilis* strains have been used to produce various biomolecules which can be used as enzymes [6], drugs [7] and raw chemicals [8]. Currently, *B. subtilis* has three subspecies [9, 10]. According to subspecies descriptions and previous studies, the presence of ribitol as a cell-wall constituent and the genetic structure of wall teichoic acid-(WTAs) related genes are regarded as the major differences among *B. s. subtilis* and *B. s. spizizenii* [9, 11, 12]. Previous studies have shown that WTAs can be synthesized by *tar* (*B. s. spizizenii* strain W23) or *tag* genes (*B. s. subtilis* strain 168) and that these genes synthesize different repeat units [11–13]. Although most *tar* and *tag* genes are homologous, the genes have different genetic structures. The *tar* genes have a divergon, *tarABIJKL-tarDF*, and produce a repeat unit with ribitol, while the *tag* genes have *tagABC-tagDEF* and produce a repeat unit with glycerol [13, 14]. In this

study, a comparative genomic analysis of the complete genome sequence of *B. s. subtilis* KCTC 3135^T and other *B. subtilis* genomes revealed that both *B. s. subtilis* and *B. s. spizizenii* have both genotypes of WTA-related genes, suggesting that *B. s. subtilis* and *B. s. spizizenii* cannot be distinguished by the presence of ribitol as a principal cell-wall constituent.

Materials and Methods

DNA Extraction and Complete Genome Sequencing

The strain of *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* KCTC 3705^T used here were obtained from the Korean Collection for Type Cultures (KCTC) and cultured using Luria-Bertani (LB) medium (Difco) at 37°C. The genomic DNA was extracted using a GenElute Bacterial Genomic DNA extraction kit (Sigma-Aldrich, USA). The library construction was performed by a TruSeq Nano DNA sample preparation kit and the genomic DNA was sequenced using HiSeq 2500 (Illumina, USA) and MiSeq (Illumina, USA) by a sequencing company (Chunlab, Korea).

Genome Assembly and Annotation

The raw sequences from HiSeq (150 bp PE, 9.5 Gbp) and MiSeq (300 bp PE, 870 Mbp) were assembled using the SPAdes assembler 3.6.0 [15]. The ncRNA and rRNA genes were searched using Infernal 1.02 [16] based on the Rfam database 12.0 [17]. CRISPR repeats were identified by PilerCR 1.06 [18] and the CRISPR recognition tool 1.2 [19]. Gene prediction was carried out using Prodigal 2.6.2 [20] and the predicted genes were annotated using KEGG [21], Swissprot [22], EggNOG [23], SEED [24] databases and the NCBI prokaryotic genome annotation pipeline. Antibiotic resistant genes were identified using HMMER3.1b2 [25] with the core Resfams HMM database v1.2 [26]. HMMER3 was run using the hmmscan command with --tblout and --cut_ga parameters. A comparative genomic analysis was carried out using BIOCYC [27]. The presence of specific genes in the pathways was confirmed manually by blastp.

Reference Genome Sequences

The complete genome sequence of *B. s. subtilis* KCTC 3135^T was deposited at GenBank (CP015375). Other genomes of *B. subtilis* were retrieved from EzBioCloud [28]. The complete genomes of forty-one other *B. subtilis* subsp. *subtilis* strains were selected. Five complete and nine incomplete genomes of *B. subtilis* subsp. *spizizenii* strains were also selected. Three genomes of *B. subtilis* subsp. *inaquosorum* strains were chosen as reference genomes. The criteria for selecting the reference genomes were completeness and availability. All available complete genomes of *B. subtilis* subsp. *subtilis* strains were retrieved from EzBioCloud DB [28]. However only a few complete genome sequences of *B. s. spizizenii* and *B. s. inaquosorum* strains were available, while some incomplete genome of *B. s. spizizenii* and *B. s. inaquosorum* strains were also retrieved. The genomes used in this study are listed in Table S1. The genome of *Staphylococcus aureus* subsp. *aureus* DSM 20231^T (CP011526) was selected as an outgroup sequence.

Detection of Putative Horizontal Gene Transfer (HGT)

The putative HGT discovery processes of *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* TU-B-10^T were carried out using HGTector [29]. The protein sequences of each strain were used as a blastp query against the GenBank nr database. The blastp results were used as input files for HGTector. The self group and close group were defined as *B. subtilis* and *Bacillus*, respectively. All other taxa were regarded as a distal group. The cutoff values of the close group and the distal group were 35.27 (KCTC 3135^T)/39.78 (TU-B-10^T) and 15.06 (KCTC 3135^T)/15.05 (TU-B-10^T), respectively.

Phylogenetic Analysis

The genome sequences of 60 *B. subtilis* strains were obtained from EzBioCloud [28]. Most of the obtained genome sequences were complete genome sequences of *B. s. subtilis* and *B. s. spizizenii* while the rest were genomes of three *B. s. inaquosorum* strains. The genome of the *S. aureus* strain DSM 20231^T was used as an outgroup (Table S1). The conserved genes (*glpF*, *ilvD*, *pta*, *purH*, *pycA*, *rpoD*, *tpiA*) of *B. subtilis* strains were designated as previous studies [30]. The *tar/tagABDF* gene sequences and conserved genes were retrieved from 60 genome sequences using the tblastn method of BLAST 2.2.30+ [31]. Among the retrieved genes, *tarABD* genes and/or the conserved genes were sequentially concatenated and used for a multi-locus sequence analysis (MLSA) [32]. The alignment was carried out using MUSCLE [33] and Neighbor-joining (NJ) tree was constructed using the Maximum Composite Likelihood method [34]. The Tar proteins from *B. s. spizizenii* W23 and the Tag proteins from *B. s. subtilis* 168 were used as reference queries. The accession numbers of the query proteins (Tar and Tag) and tblastn identities (similarities) of *tar/tagABDF* genes to the query proteins are shown in Table S2.

Among the 60 genomes, 25 unique genomes were selected to construct concise phylogenetic trees and a heatmap. The phylogenetic trees were constructed using the neighbor-joining (NJ) algorithm [35] with MEGA 7 [36]. The distance matrix was generated based on the whole genome similarity by OrthoANI [37] (Fig. 1) or on the distance based on MLSA using *tarABD* genes (Fig. 2A). Heatmaps (Figs. 2B and S3B) were generated based on the tblastn similarities of the *tar/tagABDF* target genes using R 3.4.2. The extensive trees and a heatmap are shown in Figs. S2 and S3. The distance matrix was constructed and sequence parsing was performed using an in-house script with Python 2.7.5 (on Linux) or Python 2.7.12 (on Windows).

Antibiotic Susceptibility Test

An antibiotic susceptibility test was carried out as previously described [38] with some modifications. The strains were

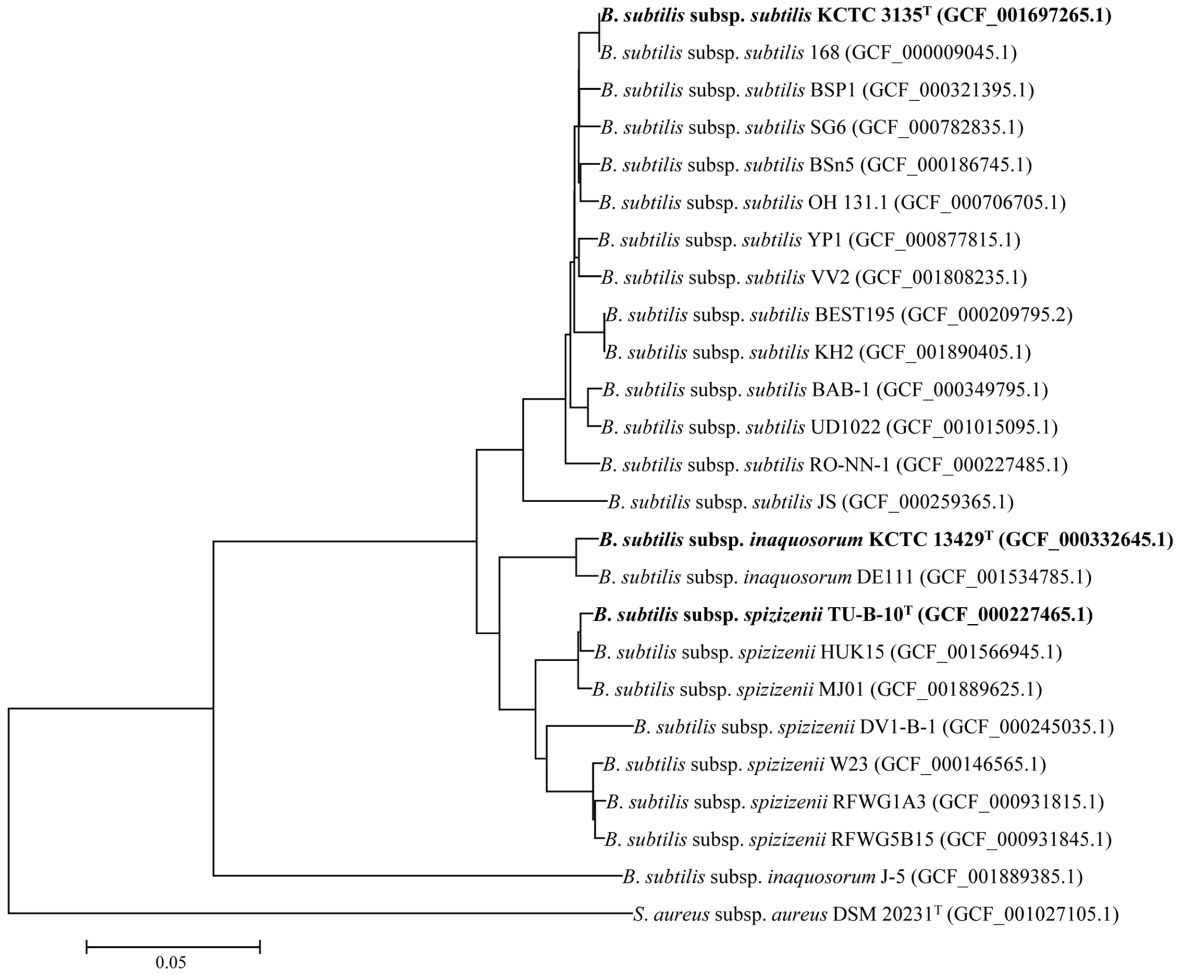

Fig. 1. A neighbor-joining tree based on the OrthoANI distance matrix.

B. subtilis with three subspecies contains 24 genomes, and *S. aureus* subsp. *aureus* DSM 20231^T was used as an outgroup. The scale bar indicates the sequence divergence.

inoculated into the LB medium at 37°C for one day and then spread onto LB agar plates. The plates were incubated at 37°C for three days.

Transmission Electron Microscopy (TEM)

Samples for TEM were prepared by a method described previously [39]. The ultrathin sections (80 nm) were viewed under ZEISS LEO 912 and FEI Tecnai G2 Spirit Twin TEM at 120kV using Digital Micrograph software.

Results and Discussion

General Features of *Bacillus subtilis* subsp. *subtilis* KCTC 3135^T

The sequences were assembled into a single contig with approximately 2,485X coverage. The genome size of *B. s. subtilis* KCTC 3135^T is 4,211,343bp with a G+C content of 43.51%. The numbers of rRNA operons, tRNAs and

Table 1. Genomic features of *Bacillus subtilis* subsp. *subtilis* KCTC 3135^T and *B. subtilis* subsp. *spizizenii* TU-B-10^T.

Features	*B. subtilis* subsp. *subtilis* KCTC 3135^T (CP015375)	*B. subtilis* subsp. *spizizenii* TU-B-10^T (CP002905)
Genomic size (bp)	4,211,343	4,207,222
G + C content (%)	43.51	43.82
Total number of genes	4,436	4,315
Protein-coding genes (CDS)	4,315	4,188
rRNA operons	10	10
tRNA genes	86	92

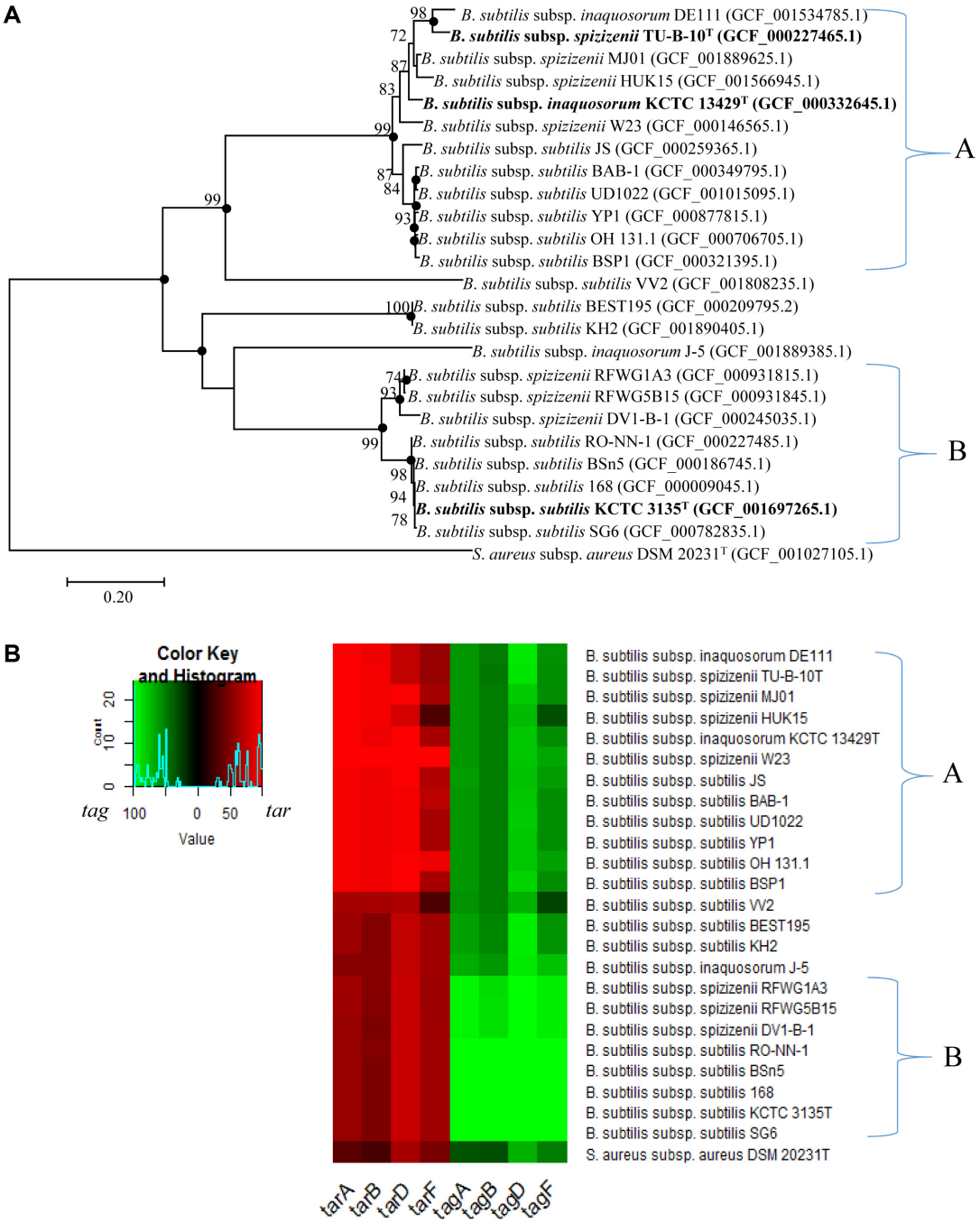

Fig. 2. (**A**) Phylogenetic tree using the *tarABD* MLSA with the neighbor-joining method and (**B**) heatmap based on the nucleotides similarities of the *tarABD* from *B. subtilis* genome to the reference *tar* or *tag* genes.

B. subtilis with three subspecies contains 24 genomes, and *S. aureus* subsp. *aureus* DSM 20231^{T} was used as an outgroup. Bootstrap values (>70%) based on 1000 replicates are shown at the branch nodes. The branch nodes also recovered in maximum-likelihood and maximum-parsimony trees are marked by black dots. The bar indicates the nucleotide substitution rate at the given length of the scale. The color indicates the tblastn nucleotide similarities of the *tarABD* from *B. subtilis* genome to the reference *tar* (red) or *tag* (green) genes.

ncRNAs were 10, 86, and 5, respectively. The total number of genes predicted is 4,436, including 4,315 coding sequence (CDS). The genomic features of *B. s. subtilis* KCTC 3135^{T} and *B. s. spizizenii* TU-B-10^{T} are listed in Table 1 and Fig. S1

describes the genome of *B. s. subtilis* KCTC 3135^T.

Based on EzBioCloud [28], 3,775 genes (87.49% out of all 4,315 total CDS) were classified according to the COG category. Excluding 1,245 genes (32.98%) in category S (Function unknown), the most abundant COG group is as follows; 301 genes (7.97% of COG-assigned genes) in Category E (Amino acid transport and metabolism); 288 genes (7.63% of COG-assigned genes) in Category K (Transcription); 271 genes (7.18% of COG assigned-genes) in Category G (Carbohydrate transport and metabolism); 209 genes (5.54% of COG-assigned genes) in Category M (Cell wall/membrane/envelope biogenesis); and 207 genes (5.48% of COG-assigned genes) in Category P (Inorganic ion transport and metabolism). The overall COG distribution of strain KCTC 3135^T was similar to the COG distributions of other related *B. subtilis* spp. The COG distributions of these genomes are listed in Table S3. The result of the comparative genomic analysis showed that most of the pathways were conserved in both *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* TU-B-10^T. The pathways of the cell structure biosynthesis processes including peptidoglycan maturation (*meso*-diaminopimelate containing), lipoteichoic acid biosynthesis and UDP-*N*-acetylmuramoyl-pentapeptide biosynthesis (*meso*-diaminopimelate containing) were conserved in both strains. The pathway of the poly-(ribitol phosphate) wall teichoic acid biosynthesis was conserved only in *B. s. spizizenii* TU-B-10^T while poly-(glycerol phosphate) wall teichoic acid biosynthesis processes were only in *B. s. subtilis* KCTC 3135^T. The biodegradation pathways including amines, amino acids, aromatic compounds, carbohydrate, carboxylate, fatty acids and protein degradation were also well conserved in both strains and did not show significant differences.

Genetic Variation between *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* TU-B-10^T

The putative horizontal gene transfer events of *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* TU-B-10^T are shown in Tables S4 and S5. There were 395 (9.2%) and 453 (10.5%) putative HGT events, respectively. The putative horizontal gene transfer percentages were higher than those of other prokaryotic genomes [40]. The majority of putative HGT events were classified into the transcription (K) and carbohydrate transport and metabolism (G) categories (Table S6). Most genes classified into category G were phosphotransferase system (PTS) transporters of hexose and major facilitator superfamily (MFS) transporters. Previous studies showed that these genes in various bacteria can be acquired by HGT [41, 42].

According to the HGTector results, both *tarABDF* and *tagABDF* were not regarded as putative HGTs. However, both genes were mutually exclusive (Table S2). Previous study suggested 'pseudoallelism' of these genes in *B. subtilis* strains [43]. The interstrain hybrid of 168/W23 showed that these genes can be substituted between *B. s. subtilis* and *B. s. spizizenii*, indicating that the presence of the ribitol as a cell-wall constituent cannot be solid phenotype of *B. s. spizizenii*. The presence of poly-ribitol teichoic acid production genes (*tarIJKL*) in *B. s. subtilis* genomes also showed that the presence of ribitol as a cell-wall constituent was not a distinguishing characteristic between *B. s. subtilis* and *B. s. spizizenii*. (Table S8). Both genomes of *B. s. subtilis* and *B. s. spizizenii* strains had one of the operons. The only exceptions were some *B. s. subtilis* strains. These strains did not show high similarity (approximately 60%) to either *tarABDF* or *tagABDF* genes (Table S8). Among these strains, strain VV2 had *tarIJKL* genes. These indicated that *tag* operons can be substituted with *tar* operons or vice versa in the genome of *B. subtilis* strains regardless of subspecies, showing the presence of ribitol is not a distinguishing characteristic of specific subspecies.

Phylogenetic Analysis of the *B. subtilis* Strains

Each *B. subtilis* subspecies was clearly distinguished and formed separate groups in OrthoANI trees and MLSA trees with conservative genes (Figs. 1, S2, and S4). This is also clearly congruent with previous study [44]. However, most *B. subtilis* were clustered into two clades according to the cell wall genotype in the *tarABD* MLSA NJ tree and the combined MLSA tree (Figs. 2, S3, and S5). The genes in clade A were *tar*-like genes while the genes in clade B were *tag*-like genes. However, some strains were not clustered into clade A or clade B, and the genes of these strains did not show high levels of similarity to either *tar* or *tag* genes (Tables S2 and S8). Both clade A and clade B had both *B. s. subtilis* and *B. s. spizizenii* strains as members, suggesting that the cell wall difference was independent of speciation and could not be used to distinguish between *B. s. subtilis* and *B. s. spizizenii*, as previously described [11, 12].

The OrthoANI tree showed that the overall degree of genome similarity between the strains coincided with the subspecies classification. However, the *tarABD* MLSA NJ tree and combined MLSA tree did not coincide with the OrthoANI tree. The MLSA trees showed that each subspecies had multiple cell wall types, also indicating that the *B. subtilis* subspecies cannot be distinguished by the

presence of ribitol as a cell-wall constituent. It also indicated that *tar* genes were not inherited after the speciation of *B. s. spizizenii*.

Antibiotic Susceptibility and Electron Microscopy

B. s. subtilis KCTC 3135^T and *B. s. spizizenii* KCTC 3705^T showed different antibiotic susceptibilities with rifampicin, streptomycin, ampicillin and tetracycline (Table 2). Compared to *B. s. subtilis* KCTC 3135^T, *B. s. spizizenii* KCTC 3705^T was more susceptible to rifampicin, streptomycin and tetracycline, and was resistant to ampicillin. However, the genetic structures of specific antibiotic resistance genes or related genes did not contribute to the difference in the antibiotic phenotypes. The distribution of antibiotic genes was not significantly different among *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* TU-B-10^T (Table S7). Rifampicin is an antibiotic that inhibits RNA synthesis by binding *rpoB* [45]. The amino acid sequence difference in the *rpoB* genes from *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* TU-B-10^T was three residues, and these residues are not known as rifampicin binding sites [46]. These findings indicate that the difference in the antibiotic susceptibility levels did not arise from the variation in the *rpoB* genes. It was reported that the inhibition of cell wall synthesis increased the sensitivity to rifampicin though the antibiotic did not target the bacterial cell wall [47]. The mechanism of action of streptomycin is the inhibition of protein synthesis by binding the 30S subunit of ribosome, and streptomycin resistance is related to mutation of the S12 protein [48]. Although the levels of susceptibility of the two strains were different, the sequences of S12 proteins from both strains were identical. Ampicillin, in the family of β-lactam antibiotics, inhibits the formation of peptidoglycan by binding DD-transpeptidase [49, 50]. The modes of ampicillin resistance were the degradation of the β-lactam ring by β-lactamase and the alternation of the binding sites [51, 52]. Both *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* TU-B-10^T had class A and class D β-lactamase genes. The similarity rates of the class A and class D β-lactamase protein sequences of *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* TU-B-10^T were 91.2% and 94.3%, respectively. It was unclear whether the sequence differences in the lactamase genes were related to the different levels of ampicillin susceptibility of the strains. However, there were no nonsense mutations in the lactamase genes. It is known that the inhibition of WTA synthesis can sensitize MRSA to β-lactam [53]. The inhibition of WTA synthesis led to the destabilization of penicillin-binding proteins, which had a crucial role in β-lactam resistance [53]. Tetracycline also binds to the 30S subunit of ribosome, and tetracycline resistance depends on various tetracycline-resistant genes [54]. *B. s. subtilis* KCTC 3135^T was more susceptible to tetracycline however despite the presence of tetracycline-resistant genes. It is unclear whether the differences in the ABC efflux pump were related to the differences in the antibiotic susceptibility levels of *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* KCTC 3705^T. Due to the extreme diversity of ABC efflux pump genes, it is difficult to discover all of the genes.

Table 2. Antibiotic susceptibility.

	KCTC 3135^T (mm)[b]	KCTC 3705^T (mm)
RD (5)[a]	15	30
S (10)	10	20
AMP (10)	10	R
K (30)	25	25
NA (30)	20	18
NV (30)	20	20
C (30)	27	25
CN (10)	25	22
TE (30)	10	24

[a]Amount of antibiotics (μg).

[b]Diameter of clear zone (mm).

RD: Rifampicin, S: Streptomycin, AMP: Ampicillin, K: Kanamycin, NA: Nalidixic acid, NV: Novobiocin, C: Chloramphenicol, CN: Gentamycin, TE: Tetracycline R: Resistance

The bacterial cell walls of the two strains, *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* KCTC 3705^T, were examined by TEM. Although the ultrastructure of the bacterial cell walls was the visibly dissimilar (Figs. 3A and 3E), no significant morphological differences were observed. *B. s. subtilis* KCTC 3135^T has a smooth outer membrane, whereas *B. s. spizizenii* KCTC 3705^T has a rough outer membrane with electron-dense surface coats and spikes (Figs. 3F and 3H). The width of the outermost cell wall in *B. s. spizizenii* KCTC 3705^T (Fig. 3H) is wider than that in *B. s. subtilis* KCTC 3135^T (Fig. 3D), which provides evidence of the presence of electron-dense surface spikes in *B. s. spizizenii* KCTC 3705^T. Electron microscopy of each strain's cell wall revealed different cell wall structures. Based on previous reports, the morphology of *B. subtilis* cells was known to be affected by WTA [55, 56]. It is assumed that the different cell-wall structures were related to the genetic structures of WTA biosynthesis genes. The initial steps of WTA synthesis by both *tar* and *tag* genes were conserved, but the extension of the WTA polymer diverged [12]. The

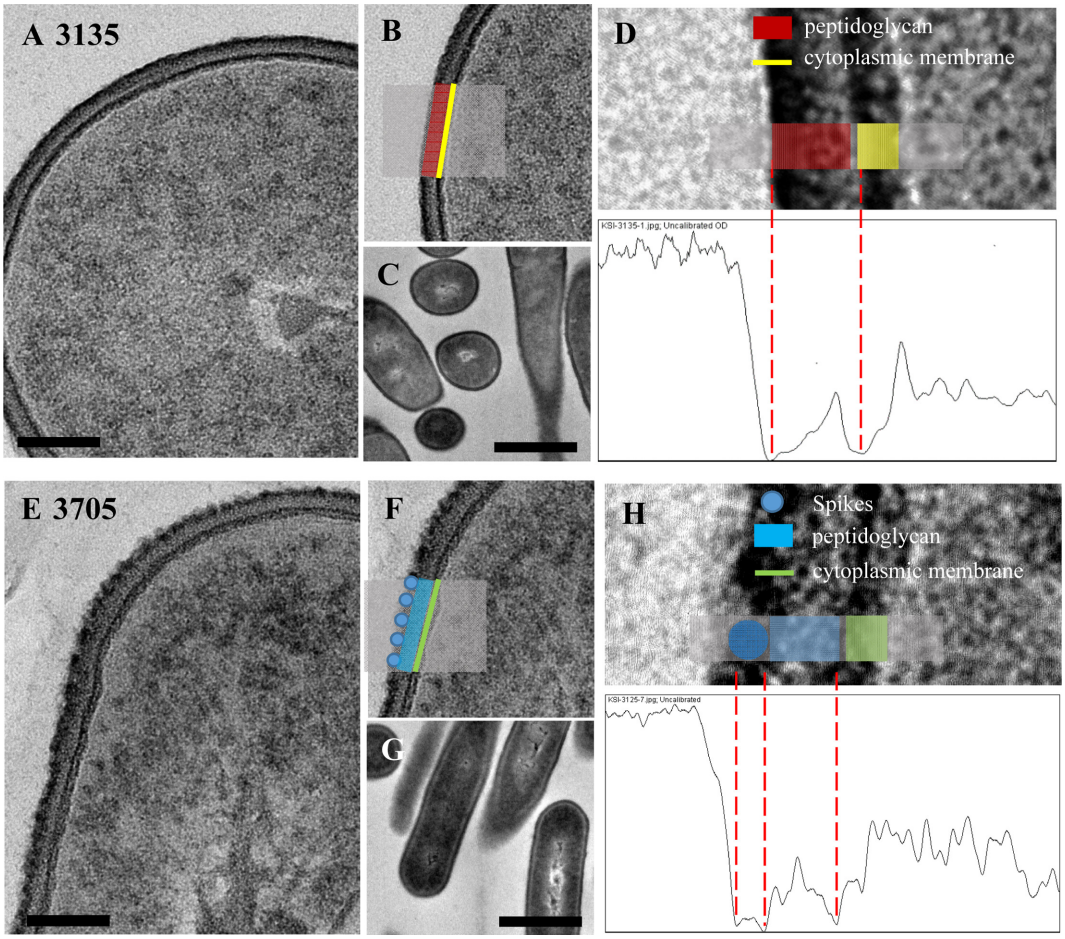

Fig. 3. Bacterial cell walls of the strains *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* KCTC 3705^T were examined by TEM.
In Figs. 3D and 3H, the x-axis and y-axis of the plot profile refer to the distance from the outer space to the bacterial cytoplasm and the average pixel intensity, respectively.

tar genes extended the polymer with ribitol while the *tag* genes extended the polymer with glycerol [12].

In conclusion, the overall genomic structures of *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* KCTC 3705^T were similar apart from the WTA synthesis genes, a difference which was previously regarded as a characteristic distinguishing *B. s. spizizenii* from *B. s. subtilis*. However, a comparative genomic analysis of multiple *B. subtilis* genomes showed that WTA with ribitol can be synthesized from both *B. s. spizizenii* and *B. s. subtilis*, indicating that the subspecies description of *B. s. spizizenii* should be corrected. The cell wall morphologies of *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* KCTC 3705^T were examined using TEM and showed clear differences. Considering the overall genomic structures of *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* KCTC 3705^T, the differences in the WTA-synthesizing genes may contribute to the differences in the cell-wall morphologies. The antibiotic susceptibility levels of *B. s. subtilis* KCTC 3135^T and *B. s. spizizenii* KCTC 3705^T were different, though the distribution of antibiotic genes was similar, suggesting that the difference in antibiotic resistance was not due to the presence of specific antibiotic genes but rather to other factors like differences in the cell-wall structures or the ABC efflux pumps, etc.

Acknowledgments

This research was supported by a project entitled "Development of Bio and Clinical Measurement Standards" funded by the Korea Research Institute of Standards and Science (KRISS - GP2017 – 0009) and by a grant from the National Research Council of Science & Technology (NST) under the Korean government (MSIP) (No. CRC-16-01-KRICT).

Conflict of Interest

The authors have no financial conflicts of interest to declare.

References

1. Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V, *et al.* 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390:** 249-256.
2. Nazina TN, Tourova TP, Poltaraus AB, Novikova EV, Grigoryan AA, Ivanova AE, *et al.* 2001. Taxonomic study of aerobic thermophilic bacilli: descriptions of *Geobacillus subterraneus* gen. nov., sp. nov. and *Geobacillus uzenensis* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus*, *Bacillus thermocatenulatus*, *Bacillus thermoleovorans*, *Bacillus kaustophilus*, *Bacillus thermoglucosidasius* and *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus*, *G. thermocatenulatus*, *G. thermoleovorans*, *G. kaustophilus*, *G. thermoglucosidasius* and *G. thermodenitrificans*. *Int. J. Syst. Evol. Microbiol.* **51:** 433-446.
3. Porwal S, Lal S, Cheema S, Kalia VC. 2009. Phylogeny in aid of the present and novel microbial lineages: diversity in *Bacillus*. *PLoS One* **4:** e4438.
4. Harwood CR. 1992. *Bacillus subtilis* and its relatives: molecular biological and industrial workhorses. *Trends Biotechnol.* **10:** 247-256.
5. Earl AM, Losick R, Kolter R. 2008. Ecology and genomics of *Bacillus subtilis*. *Trends Microbiol.* **16:** 269-275.
6. Schallmey M, Singh A, Ward OP. 2004. Developments in the use of *Bacillus* species for industrial production. *Can. J. Microbiol.* **50:** 1-17.
7. Jin P, Zhang L, Yuan P, Kang Z, Du G, Chen J. 2016. Efficient biosynthesis of polysaccharides chondroitin and heparosan by metabolically engineered *Bacillus subtilis*. *Carbohydr. Polym.* **140:** 424-432.
8. Deshmukh AN, Nipanikar-Gokhale P, Jain R. 2016. Engineering of *Bacillus subtilis* for the Production of 2,3-Butanediol from Sugarcane Molasses. *Appl. Biochem. Biotechnol.* **179:** 321-331.
9. Nakamura LK, Roberts MS, Cohan FM. 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.
10. Rooney AP, Price NP, Ehrhardt C, Swezey JL, Bannan JD. 2009. Phylogeny and molecular taxonomy of the *Bacillus subtilis* species complex and description of *Bacillus subtilis* subsp. *inaquosorum* subsp. nov. *Int. J. Syst. Evol. Microbiol.* **59:** 2429-2436.
11. Swoboda JG, Campbell J, Meredith TC, Walker S. 2010. Wall teichoic acid function, biosynthesis, and inhibition. *ChemBioChem.* **11:** 35-45.
12. Brown S, Santa Maria JP Jr, Walker S. 2013. Wall teichoic acids of gram-positive bacteria. *Annu. Rev. Microbiol.* **67:** 313-336.
13. Lazarevic V, Abellan FX, Moller SB, Karamata D, Mauel C. 2002. Comparison of ribitol and glycerol teichoic acid genes in *Bacillus subtilis* W23 and 168: identical function, similar divergent organization, but different regulation. *Microbiology* **148:** 815-824.
14. Mauel C, Young M, Karamata D. 1991. Genes concerned with synthesis of poly(glycerol phosphate), the essential teichoic acid in *Bacillus subtilis* strain 168, are organized in two divergent transcription units. *J. Gen. Microbiol.* **137:** 929-941.
15. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, *et al.* 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19:** 455-477.
16. Nawrocki EP, Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29:** 2933-2935.
17. Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, *et al.* 2015. Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res.* **43:** D130-D137.
18. Edgar RC. 2007. PILER-CR: Fast and accurate identification of CRISPR repeats. *BMC Bioinformatics* **8:** 18.
19. Bland C, Ramsey TL, Sabree F, Lowe M, Brown K, Kyrpides NC, *et al.* 2007. CRISPR Recognition Tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics* **8:** 209.
20. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11:** 119.
21. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* **42:** D199-D205.
22. Chen C, Huang H, Wu CH. 2017. Protein bioinformatics databases and resources. *Methods Mol. Biol.* **1558:** 3-39.
23. Powell S, Forslund K, Szklarczyk D, Trachana K, Roth A, Huerta-Cepas J, *et al.* 2014. eggNOG v4.0: nested orthology inference across 3686 organisms. *Nucleic Acids Res.* **42:** D231-D239.
24. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, *et al.* 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* **33:** 5691-5702.
25. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. 2013. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* **41:** e121-e121.
26. Gibson MK, Forsberg KJ, Dantas G. 2015. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J.* **9:** 207-216.
27. Caspi R, Billington R, Ferrer L, Foerster H, Fulcher CA, Keseler IM, *et al.* 2016. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* **44:** D471-480.

28. Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, *et al.* 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* **67:** 1613-1617.
29. Zhu Q, Kosoy M, Dittmar K. 2014. HGTector: an automated method facilitating genome-wide discovery of putative horizontal gene transfers. *BMC Genomics* **15:** 717.
30. Jolley KA, Maiden MC. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* **11:** 595.
31. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, *et al.* 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10:** 421.
32. Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, *et al.* 1998. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. USA* **95:** 3140-3145.
33. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32:** 1792-1797.
34. Tamura K, Nei M, Kumar S. 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. USA* **101:** 11030-11035.
35. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4:** 406-425.
36. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33:** 1870-1874.
37. Lee I, Kim YO, Park SC, Chun J. 2016. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int. J. Syst. Evol. Microbiol.* **66:** 1100-1103.
38. Kim S, Gong G, Woo HM, Kim Y, Um Y. 2016. *Burkholderia jirisanensis* sp. nov., isolated from forest soil. *Int. J. Syst. Evol. Microbiol.* **66:** 1260-1267.
39. Yun SH, Lee SY, Choi CW, Lee H, Ro HJ, Jun S, *et al.* 2017. Proteomic characterization of the outer membrane vesicle of the halophilic marine bacterium *Novosphingobium pentaromativorans* US6-1. *J. Microbiol.* **55:** 56-62.
40. Garcia-Vallvé S, Romeu A, Palau J. 2000. Horizontal gene transfer in bacterial and archaeal complete genomes. *Genome Res.* **10:** 1719-1725.
41. Zúñiga M, Comas I, Linaje R, Monedero V, Yebra MJ, Esteban CD, *et al.* 2005. Horizontal gene transfer in the molecular evolution of mannose PTS transporters. *Mol. Biol. Evol.* **22:** 1673-1685.
42. Comas I, González-Candelas F, Zúñiga M. 2008. Unraveling the evolutionary history of the phosphoryl-transfer chain of the phosphoenolpyruvate: phosphotransferase system through phylogenetic analyses and genome context. *BMC Evol. Biol.* **8:** 147.
43. Young M, Mauel C, Margot P, Karamata D. 1989. Pseudo-allelic relationship between non-homologous genes concerned with biosynthesis of polyglycerol phosphate and polyglycerol phosphate teichoic acids in *Bacillus subtilis* strains 168 and W23. *Mol. Microbiol.* **3:** 1805-1812.
44. Yi H, Chun J, Cha CJ. 2014. Genomic insights into the taxonomic status of the three subspecies of *Bacillus subtilis*. *Syst. Appl. Microbiol.* **37:** 95-99.
45. Halling SM, Burtis K, Doi R. 1977. Reconstitution studies show that rifampicin resistance is determined by the largest polypeptide of *Bacillus subtilis* RNA polymerase. *J. Biol. Chem.* **252:** 9024-9031.
46. Goldstein BP. 2014. Resistance to rifampicin: a review. *J. Antibiot.* **67:** 625.
47. Sigle S, Steblau N, Wohlleben W, Muth G. 2016. Polydiglycosylphosphate transferase PdtA (SCO2578) of *Streptomyces coelicolor* A3 (2) is crucial for proper sporulation and apical tip extension under stress conditions. *Appl. Environ. Microbiol.* **82:** 5661-5672.
48. Sharma D, Cukras AR, Rogers EJ, Southworth DR, Green R. 2007. Mutational analysis of S12 protein and implications for the accuracy of decoding by the ribosome. *J. Mol. Biol.* **374:** 1065-1076.
49. Nguyen-Distèche M, Leyh-Bouille M, Ghuysen J-M. 1982. Isolation of the membrane-bound 26 000-Mr penicillin-binding protein of Streptomyces strain K15 in the form of a penicillin-sensitive D-alanyl-D-alanine-cleaving transpeptidase. *Biochem. J.* **207:** 109-115.
50. Gordon E, Mouz N, Duee E, Dideberg O. 2000. The crystal structure of the penicillin-binding protein 2x from *Streptococcus pneumoniae* and its acyl-enzyme form: implication in drug resistance. *J. Mol. Biol.* **299:** 477-485.
51. Lowy FD. 2003. Antimicrobial resistance: the example of *Staphylococcus aureus*. *J. Clin. Invest.* **111:** 1265-1273.
52. Drawz SM, Bonomo RA. 2010. Three decades of β-lactamase inhibitors. *Clin. Microbiol. Rev.* **23:** 160-201.
53. Farha MA, Leung A, Sewell EW, D'Elia MA, Allison SE, Ejim L, *et al.* 2012. Inhibition of WTA synthesis blocks the cooperative action of PBPs and sensitizes MRSA to β-lactams. *ACS Chem. Biol.* **8:** 226-233.
54. Roberts MC. 2005. Update on acquired tetracycline resistance genes. *FEMS Microbiol. Lett.* **245:** 195-203.
55. Bhavsar AP, Beveridge TJ, Brown ED. 2001. Precise deletion of *tagD* and controlled depletion of its product, glycerol 3-phosphate cytidylyltransferase, leads to irregular morphology and lysis of *Bacillus subtilis* grown at physiological temperature. *J. Bacteriol.* **183:** 6688-6693.
56. D'Elia MA, Millar KE, Beveridge TJ, Brown ED. 2006. Wall teichoic acid polymers are dispensable for cell viability in *Bacillus subtilis*. *J. Bacteriol.* **188:** 8313-8316.