Bacillus sonorensis sp. nov., a close relative of *Bacillus licheniformis*, isolated from soil in the Sonoran Desert, Arizona

Margaret M. Palmisano,[1] L. K. Nakamura,[2] Kathleen E. Duncan,[3] Conrad A. Istock[4,5] and Frederick M. Cohan[1]

[1] Department of Biology, Wesleyan University, Middletown, CT 06459-0170, USA

[2] Microbial Properties Research Unit, National Center for Agricultural Utilization Research, Agricultural Research Service, US Department of Agriculture, Peoria, IL 61604, USA

[3] Department of Botany and Microbiology, University of Oklahoma, Norman, OK 73019, USA

[4] Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14850, USA

[5] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA

Author for correspondence: Frederick M. Cohan. Tel: +1 860 685 3482. Fax: +1 860 685 3279. e-mail: FCohan@Wesleyan.edu

Eight *Bacillus* strains isolated from Sonoran Desert soil were shown to belong to a previously unidentified species, for which the name *Bacillus sonorensis* sp. nov. is proposed. The type strain is strain L87-10ᵀ (= NRRL B-23154ᵀ). On the basis of phenotypic and genetic data, *B. sonorensis* is most closely related to *Bacillus licheniformis*. *B. sonorensis* can be distinguished from *B. licheniformis* by salt tolerance, pigmentation, multilocus enzyme electrophoresis, reassociation of genomic DNA and sequence differences in protein-coding genes and 16S rRNA.

Keywords: new species, taxonomy, *Bacillus*

INTRODUCTION

The species most closely related to *Bacillus subtilis* are unusually similar at the phenotypic level. For example, fatty acid composition is the only known phenotypic character that distinguishes *Bacillus mojavensis* and *Bacillus vallismortis* from one another or from *B. subtilis* (Roberts *et al*., 1994, 1996) and *Bacillus atrophaeus* is distinguishable from *B. subtilis* only by differences in pigmentation (Nakamura, 1989). Molecular methods have proven more effective in distinguishing close relatives of *B. subtilis* and, indeed, three taxa related to *B. subtilis* were discovered through restriction site analysis of PCR-amplified genes [*B. subtilis* subsp. *spizizenii*, formerly the W23 group of *B. subtilis* (Nakamura *et al*., 1999), *B. mojavensis* (Roberts *et al*., 1994) and *B. vallismortis* (Roberts *et al*., 1996)]. The paucity of diagnostic phenotypic characteristics within this group suggests that relatives of *B. subtilis* may still be discovered by molecular means.

A study by Duncan *et al*. (1994) suggested that some strains classified phenotypically as *Bacillus licheniformis* may belong to an unidentified species. Although identified originally as *B. licheniformis* on the basis of 61 metabolic and tolerance tests, these strains formed a cluster distinct from the rest of *B. licheniformis* on the basis of multilocus enzyme electrophoresis (MLEE). The two clusters of *B. licheniformis*-like strains were designated Group 1 (*B. licheniformis sensu stricto*) and Group 2 (newly identified). Groups 1 and 2 differed in the frequency of strains that tested positive for several metabolic characters by the API method, but these differences were not diagnostic (Duncan *et al*., 1994). A preliminary survey has shown that strains from Groups 1 and 2 fall into separate clusters on the basis of RAPD characters (C. A. Istock, unpublished data).

The present study provides further phenotypic and molecular evidence that the Group 2 isolates form a group distinct from *B. licheniformis sensu stricto* and merit status as a separate species. Group 2 is distinguishable from *B. licheniformis* by several newly found phenotypic characters, by DNA–DNA reassociation assays and by sequence differences in protein-coding genes and 16S rRNA.

METHODS

Bacterial strains. All strains and their sources are listed in Table 1. Eighteen strains were previously isolated from 200 cm^3 of soil at one site on Tumamoc Hill in the Sonoran Desert, AZ, USA. These strains were identified as *B.*

The GenBank accession numbers for the *rpoB* and 16S rRNA gene sequences reported in this paper are AF226065–AF226075 and AF302118–AF302125, respectively.

Table 1. *Bacillus* strains compared in this study

Culture collections: ATCC, American Type Culture Collection, Manassas, VA, USA; BGSC, *Bacillus* Genetic Stock Center, Ohio State University, Columbus, OH, USA; NRRL, National Center for Agricultural Research (NCAUR), Peoria, IL, USA. In the case of NRRL accession numbers, the prefix B indicates that the organisms were obtained directly from individuals or were isolated at NCAUR and the prefix NRS indicates that the strains were obtained from the N. R. Smith *Bacillus* Collection. *B. licheniformis* strains NRRL B-2179, NRRL B-14262, NRRL B-14268, NRRL NRS-1114, NRRL NRS-1217, NRRL NRS-1292 and NRRL NRS-700 were also included.

Strain	Other designation(s)	Source*
Group 2		
L87-10^T†	NRRL B-23154^T; TG-8-8^T	1
L89-15†	NRRL B-23155; T89-39	1
L89-9†	NRRL B-23156; T89-33	1
L89-14†	NRRL B-23157; T89-38	1
L89-16†	NRRL B-23158; T89-40	1
L89-18†	NRRL B-23159; Rf-1	1
L87-2†	NRRL B-23160; Te-11	1
L87-4†	NRRL B-23161; Te-45	1
B. licheniformis		
L87-3†	NRRL B-23317; Te-12	1
L87-7†	NRRL B-23318; TG1-15	1
L88-2†	NRRL B-23319; T88-15	1
L89-1†	NRRL B-23320; T89-11	1
L89-2†	NRRL B-23321; T89-26	1
L89-3†	NRRL B-23322; T89-27	1
L89-6†	NRRL B-23323; T89-30	1
L89-7†	NRRL B-23324; T89-31	1
L89-17†	NRRL B-23325; T89-54	1
L89-19†	NRRL B-23326; Rk1	1
DV5-A-1-10†	NRRL B-23327	2
DV5-A-2-2	NRRL B-14950	2
DV5-A-3-1	NRRL B-14951	2
DV5-A-4-4	NRRL B-14952	2
DV5-A-4-5	NRRL B-14953	2
DV5-A-4-9†	NRRL B-23328	2
DV5-A-5-5	NRRL B-14954	2
DV6-B-2†	NRRL B-14955	2
D6-50-2†	NRRL B-23329	2
DV7-B-2†	NRRL B-23330	2
DV8-3-1†	NRRL B-14956	2
DV8-3-8†	NRRL B-23331	2
DV8-3-9†	NRRL B-23332	2
DV8-3-10†	NRRL B-23333	2
DV3-B-5†	NRRL B-23334	2
D5-0-3	NRRL B-23335	2
NRRL NRS-1264^T	ATCC 24580^T	–
B. mojavensis RO-H-1^T	NRRL B-14698^T	2
B. atrophaeus NRRL NRS-213^T	–	–
B. subtilis subsp. *subtilis*		
BGSC 1A2	NRRL B-14819	–
NRRL NRS-744^T	–	–
B. subtilis subsp. *spizizenii* BGSC 2A2	–	–
B. amyloliquefaciens NRRL B-14393^T	ATCC 23350^T	–
B. vallismortis DV1-F-3^T	NRRL B-14890^T	2

* Sources are indicated as: 1, strain isolated from the Sonoran Desert, AZ, USA, by Duncan *et al.* (1994) from a site on Tumamoc Hill (the second alternative designation for each such strain was used by Istock *et al.*, 1996); 2, strains collected from Death Valley, CA, USA, by F. M. Cohan.

† Strain used in pigmentation assays.

licheniformis on the basis of metabolic characters; eight were classified as Group 2 and ten as *B. licheniformis sensu stricto* based on MLEE. The MLEE analysis distinguished the eight Group 2 strains into seven genotypes (Duncan *et al.*, 1994). From soil in Death Valley, CA, USA, 16 strains were isolated previously and were classified as *B. licheniformis* on the basis of five metabolic characters and restriction digest analysis of three protein-coding genes (Roberts & Cohan, 1995; M. Palmisano and F. Cohan, unpublished results). Also included in this study were the type strains of *B. mojavensis*, *B. vallismortis*, *B. atrophaeus* and *B. licheniformis* and two laboratory strains of *B. subtilis* representing the subspecies *B. subtilis* subsp. *subtilis* and *B. subtilis* subsp. *spizizenii*, as well as additional *B. licheniformis* strains from the Agricultural Research Service Culture Collection (NRRL).

Phenotypic characterization. Physiological, morphological and metabolic characteristics (except pigmentation and fatty acid composition) were determined as described previously (Gordon *et al.*, 1973; Nakamura & Swezey, 1983). In order to evaluate pigmentation, cells were grown at 30 °C for 24 h and then incubated on one of the following media at 30 °C: pH 5·6 agar of Gordon *et al.* (1973), tyrosine agar (Gordon *et al.*, 1973) or glycerol/glutamate agar (Arai & Mikami, 1972). Cultures were observed daily for 2 weeks. *B. atrophaeus* was cultured on glycerol/glutamate agar as a positive control. Strains used in the pigmentation evaluation are indicated in Table 1.

For assays of salt tolerance, cultures were incubated in nutrient broth containing 0, 3, 5, 7 and 10% (w/v) NaCl. Duplicate culture tubes containing 6 ml medium were inoculated with a loopful of 18–20 h culture grown in nutrient broth at 30 °C. The inoculated tubes were incubated at 30 °C and monitored for growth at 2, 4, 7 and 14 d. Growth, if it was to occur, was usually observable by 4 d. The eight Group 2 strains and the type strain of *B. licheniformis* were tested.

The whole-cell fatty acid composition was determined by using the MIDI system of Sasser (1990) on organisms that had been grown for 24 h at 28 °C on Trypticase soy agar. We determined the fatty-acid content of the eight Group 2 strains and the 14 *B. licheniformis* strains listed in the legend to Table 3.

DNA–DNA reassociation and G+C content determination. The renaturation rates of genomic fragments from pairs of strains were determined spectrophotometrically with a model 250 UV spectrophotometer (Gilford Systems) equipped with a model 2527 thermoprogrammer (Nakamura & Swezey, 1983). The equation of De Ley *et al.* (1970) was used to calculate the extent of DNA–DNA reassociation. The G+C content was estimated by the thermal melting procedure described by Mandel & Marmur (1968). *Escherichia coli* DNA with a G+C content of 51 mol% was used for comparison.

PCR amplification. Genomic DNA was extracted and purified as described by Cohan *et al.* (1991). Purified preparations were used as template DNA for amplification of the *secY*, *rpoB* and 16S rRNA genes. The genes *secY* and *rpoB* were chosen because they are highly conserved and had previously been sequenced in *B. licheniformis* (Tschauder *et al.*, 1992; Majewski & Cohan, 1999). Moreover, previous studies of *Bacillus* systematics (Roberts & Cohan, 1995) and recombination (Cohan *et al.*, 1991; Roberts & Cohan, 1993; Zawadzki *et al.*, 1995; Zawadzki & Cohan, 1995; Majewski & Cohan, 1998, 1999; Majewski *et al.*, 2000) were based on sequence variation at *rpoB*. PCR was performed as described by Roberts & Cohan (1995).

Oligonucleotide primers for *secY* were derived from the *B. licheniformis* coding sequence (Tschauder *et al.*, 1992; strain DSM 13) and extended from bp 448 to 465 (5′-TTACATCACAGCTTCTAT-3′) and from bp 1497 to 1479 (5′-CGATAGTTTCGTTTTACCA-3′), yielding a 1050 bp PCR product. Primers based on the homologous *secY* sequence of *B. subtilis* (Suh *et al.*, 1990; Marburg strain) extended from bp 391 to 408 (5′-TTATATCACGGCTTCGAT-3′) and from bp 1441 to 1423 (5′-CGGTAGTTTCGTTTCACCA-3′), also yielding a 1050 bp product. All members of *B. licheniformis* and Group 2 were amplified using the *B. licheniformis* primers and all other strains were amplified using the *B. subtilis* primers. For *secY*, the PCR profile for *B. licheniformis* and Group 2 was 40 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 45 s and extension at 72 °C for 2·5 min; the PCR profile for the other species was 40 cycles of denaturation at 94 °C for 30 s, annealing at 37 °C for 45 s and extension at 72 °C for 2·5 min. Primers for *B. licheniformis secY* were synthesized by Oligos, Etc. (Guilford, CT, USA) and the *B. subtilis*-derived primers were synthesized by Integrated DNA Technologies (Coralville, IA, USA).

The *rpoB* primers were based on the *B. subtilis* coding sequence (Boor *et al.*, 1995; GenBank accession no. L24376). The primers, synthesized by Integrated DNA Technologies, extended from bp 1094 to 1117 (5′-AGGTCAACTAGTTCAGTATGGACG-3′) and from bp 1669 to 1651 (5′-ACCGTAACCGGCAACTTAC-3′), yielding a 576 bp product. The PCR profile was 40 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 45 s and extension at 72 °C for 2 min.

Primers 27f and 1492r were used to amplify the 16S rRNA gene of the eight Group 2 strains, following the protocol of Lane (1991).

Sequencing. A 402 bp segment of the *rpoB* gene was sequenced from the eight strains of Group 2 and *B. licheniformis* strains L89-17, D5-0-3 and L87-7. Sequencing reactions were performed by the DNA sequencing facility at the University of Pennsylvania School of Medicine, using the PCR-amplified *rpoB* segment (576 bp) as the template for all of the reactions. The sequencing reaction used the leading-strand primer from PCR, extending from bp 1094 to 1117. This yielded a single-stranded sequence of the sense strand. The 16S rRNA gene was sequenced following Nakamura *et al.* (1999).

Restriction digest analysis. Each PCR-amplified segment of *secY* from the eight Group 2 strains was treated with the four-cutter restriction endonucleases *Alu*I, *Dpn*II, *Fnu*4HI, *Hae*III, *Hin*fI, *Hpa*II and *Rsa*I (New England Biolabs), as described by Roberts & Cohan (1995). The PCR products from the type strains of *B. licheniformis*, *B. atrophaeus* and *B. vallismortis* and strains of the two *B. subtilis* subspecies (*B. subtilis* subsp. *subtilis* strain 1A2 and *B. subtilis* subsp.

spizizenii strain 2A2) were digested with *Dpn*II, *Fnu*4HI, *Hae*III, *Hin*fI, *Hpa*II and *Rsa*I. Restriction digest products were observed on 5% polyacrylamide gels that were stained with 0·5 μg ethidium bromide ml^{-1}.

Fragment lengths were estimated using Bio-ladder: 50–1000 bp (Bio-Synthesis). Restriction site differences among strains were inferred from restriction-fragment length variation using a computer algorithm developed by Cohan *et al.* (1991). In cases where the exact size and identity of a digest fragment were ambiguous, we performed supplementary restriction digests using multiple endonucleases.

RESULTS

Phenotypic characteristics

Metabolic tests with API strips were shown previously to provide no diagnostic characters that can distinguish Group 2 from *B. licheniformis* (Duncan *et al.*, 1994). However, we found these groups to be distinguishable by several other phenotypic tests (Table 2). Group 2 and *B. licheniformis* showed different pigmentation characteristics: on pH 5·6 agar, Group 2 strains appeared bright yellow, while *B. licheniformis* strains appeared pale-cream; on tyrosine agar, the Group 2 strains appeared brown while *B. licheniformis* appeared pale cream; and on glycerol/glutamate medium, *B. licheniformis* strains appeared reddish-brown while Group 2 strains appeared pale yellowish-cream. Group 2 and *B. licheniformis* could also be distinguished by their salt tolerances. Group 2 strains grew in 3% NaCl but not in 5, 7 or 10% NaCl, whereas the type strain of *B. licheniformis* grew in 3, 5, 7 and 10% NaCl.

Group 2 was phenotypically most similar to *B. licheniformis* (Claus & Berkeley, 1986) and these two could be distinguished from *B. mojavensis* (Roberts *et al.*, 1994), *Bacillus amyloliquefaciens* (Nakamura, 1987), *B. subtilis* (Claus & Berkeley, 1986) and *B. atrophaeus* (Nakamura, 1989) on the basis of their ability to grow anaerobically, to use propionate and to grow at a slightly higher range of temperatures (Table 2).

Group 2 was significantly different from *B. licheniformis* and all other taxa tested in cellular fatty acid composition (Table 3). Nevertheless, fatty acid characters could not distinguish Group 2 diagnostically from *B. licheniformis*.

Table 2. Comparisons of the phenotypes and G+C contents of Group 2 strains and related *Bacillus* species

Taxa are identified as: 1, Group 2 strains; 2, *B. licheniformis* NRRL NRS-1264[T]; 3, *B. mojavensis* NRRL B-14698[T]; 4, *B. amyloliquefaciens* NRRL B-14393[T]; 5, *B. subtilis* NRRL NRS-744[T]; 6, *B. atrophaeus* NRRL NRS-213[T]. Characteristics are scored as: −, negative reaction; +, positive reaction; W, weak reaction; ND, not determined. For all growth assays: +, growth; −, no growth. All taxa grow at pH 5·7 and in the presence of 3% NaCl but not in the presence of 0·001% lysozyme, show catalase activity and acetylmethylcarbinol, produce acid from glucose, arabinose, xylose and mannitol, hydrolyse starch, utilize citrate, reduce nitrate to nitrite and decompose casein. Data for taxa other than Group 2 were taken from Roberts *et al.* (1996).

Characteristic	1*	2	3	4	5	6
Anaerobic growth	+	+	−	−	−	−
Growth in the presence of:						
5% NaCl	−	+	+	+	+	+
7% NaCl	−	+	+	+	+	+
10% NaCl	−	+	+	+	+	+
Oxidase activity	ND	+	+	+	+	−
Temperature for growth (°C):						
Maximum	55	55	50	50	50	50
Minimum	15	15	10	10	10	10
Acid produced from lactose	ND	−	−	+	−	−
Utilization of propionate	+	+	−	−	−	−
Tween 80 decomposition	ND	ND	+	ND	W	+
Pigmentation assays:						
pH 5·6 agar	Bright yellow	Cream†	ND	ND	ND	ND
Tyrosine agar	Brown	Cream†	ND	ND	ND	ND
Glycerol/glutamate agar	Pale yellowish-cream	Reddish-brown†	ND	ND	ND	Brown‡
DNA G+C content (mol%)	46	46	43	43	43	42

* Eight Group 2 strains were used in all assays except for the G+C content assay, which is a mean value based on strains NRRL B-23154[T], NRRL B-23156, NRRL B-23157, NRRL B-23160 and NRRL B-23161.

† Twenty natural isolates of *B. licheniformis sensu stricto*, including ten from the Sonoran Desert and ten from Death Valley, were assayed for pigmentation and are indicated in Table 1. All 20 isolates of *B. licheniformis* as well as the *B. licheniformis* type strain gave identical pigmentation results for all pigmentation assays.

‡ Data from Nakamura (1989).

Table 3. Cellular fatty acid compositions of Group 2 strains and related *Bacillus* species

Taxa are identified (with the number of strains analysed indicated in parentheses) as: 1, Group 2 ($n = 8$); 2, *B. licheniformis* (14); 3, *B. subtilis* (5); 4, *B. amyloliquefaciens* (3); 5, *B. atrophaeus* (5); 6, *B. mojavensis* (22); 7, *B. vallismortis* (5). Results are expressed as mean percentages of total cellular fatty acids ± SD. The 14 *B. licheniformis* strains included the following NRRL strains: B-2179, B-14262, B-14268, B-14950, B-14951, B-14952, B-14953, B-14954, B-14955, B-14956, NRS-1114, NRS-1292, NRS-1217 and NRS-1264ᵀ. Values in bold were significantly different from those of Group 2 in a *t*-test accounting for comparisons with multiple (6) taxa. Data for taxa other than Group 2 were taken from Table 5 (*B. vallismortis*) or Table 7 (remaining species) of Roberts *et al*. (1994).

Fatty acid	1	2	3	4	5	6	7
14:0 iso	0·15 ± 0·42	**1·09 ± 0·17**	**1·13 ± 0·24**	**2·46 ± 0·69**	**1·44 ± 0·14**	**0·98 ± 0·24**	**1·07 ± 0·02**
15:0 iso	30·01 ± 2·99	32·18 ± 3·76	29·27 ± 4·64	30·50 ± 5·93	**15·02 ± 2·55**	**22·33 ± 3·19**	**24·60 ± 1·37**
15:0 anteiso	37·31 ± 1·64	39·89 ± 2·98	40·19 ± 3·98	36·48 ± 7·88	**51·36 ± 1·08**	**42·51 ± 1·67**	37·50 ± 0·77
16:0 iso	3·49 ± 0·50	3·34 ± 0·62	**2·36 ± 0·34**	4·40 ± 0·75	3·10 ± 0·61	**2·56 ± 0·41**	4·06 ± 0·39
16:1 *cis*5	1·28 ± 0·58	1·83 ± 0·50	1·52 ± 0·45	2·14 ± 0·11	1·72 ± 0·11	1·74 ± 0·40	0·64 ± 0·03
16:0	4·82 ± 0·63	**3·43 ± 0·73**	**3·14 ± 0·40**	4·52 ± 0·50	**1·99 ± 0·32**	**2·05 ± 0·41**	**2·71 ± 0·42**
17:1 *cis*7 iso	0·65 ± 0·70	**1·45 ± 0·30**	**1·72 ± 0·42**	1·67 ± 0·61	**1·99 ± 0·44**	**3·45 ± 0·62**	1·55 ± 0·40
17:0 iso	8·64 ± 0·79	**6·26 ± 0·91**	9·59 ± 1·56	9·01 ± 1·34	**4·97 ± 0·59**	8·92 ± 1·46	**14·43 ± 0·66**
17:0 anteiso	12·37 ± 1·08	**9·89 ± 1·61**	**9·38 ± 0·95**	**7·06 ± 2·69**	14·83 ± 2·36	12·53 ± 1·89	12·07 ± 1·09

Table 4. Levels of DNA–DNA reassociation between genomes of Group 2 isolates

Values are mean percentages of reassociation from two determinations; the maximum difference noted between determinations was 6%. Values in parentheses indicate that the reassociation values were 100% by definition. The reassociation values between the Group 2 strains and strains of other taxa were: 30–46% for *B. licheniformis* strains NRS-1264ᵀ and NRS-700, 21–42% for *B. mojavensis* B-14698ᵀ, 22–39% for *B. amyloliquefaciens* B-14393ᵀ, 18–39% for *B. subtilis* subsp. *subtilis* NRS-744ᵀ, 23–39% for *B. vallismortis* B-14890ᵀ and 18–35% for *B. atrophaeus* NRS-213ᵀ.

Strain	1	2	3	4	5	6	7	8
1. NRRL B-23154ᵀ	(100)							
2. NRRL B-23155	94	(100)						
3. NRRL B-23156	94	100	(100)					
4. NRRL B-23157	101	102	101	(100)				
5. NRRL B-23158	89	81	101	98	(100)			
6. NRRL B-23159	101	100	101	80	80	(100)		
7. NRRL B-23160	95	101	100	103	103	100	(100)	
8. NRRL B-23161	92	80	100	99	101	101	98	(100)

DNA G+C content

The mean G+C content of Group 2 strains was 46 mol% (Table 2). This is not significantly different from *B. licheniformis*, the G+C content of which ranges from 43 to 46 mol% (Claus & Berkeley, 1986).

DNA–DNA reassociation

Table 4 shows the levels of DNA–DNA reassociation among Group 2 strains and between Group 2 strains and the type strains of *B. licheniformis*, *B. mojavensis*, *B. amyloliquefaciens*, *B. subtilis*, *B. vallismortis* and *B. atrophaeus*. This list includes all of the species of the clade that includes *B. licheniformis* and *B. subtilis*, based on 16S rRNA sequence data (Ash *et al*., 1991; Pettersson *et al*., 1999) and restriction digest data for *gyrA*, *polC* and *rpoB* (Roberts & Cohan, 1995). The eight strains of Group 2 exhibited high levels of DNA–DNA reassociation with one another (80–100%) but very low levels of reassociation with the type strains of other taxa (18–46%).

Restriction site variation at *secY*

All eight Group 2 strains showed the same set of 31 restriction sites at *secY* (Fig. 1) and therefore our estimate of the mean within-Group-2 sequence divergence is zero. A previous restriction analysis showed the mean sequence divergence (π) within *B. licheniformis* in the homologous segment of *secY* to be 0·005 ± 0·0001 (M. Palmisano and F. Cohan, unpub-

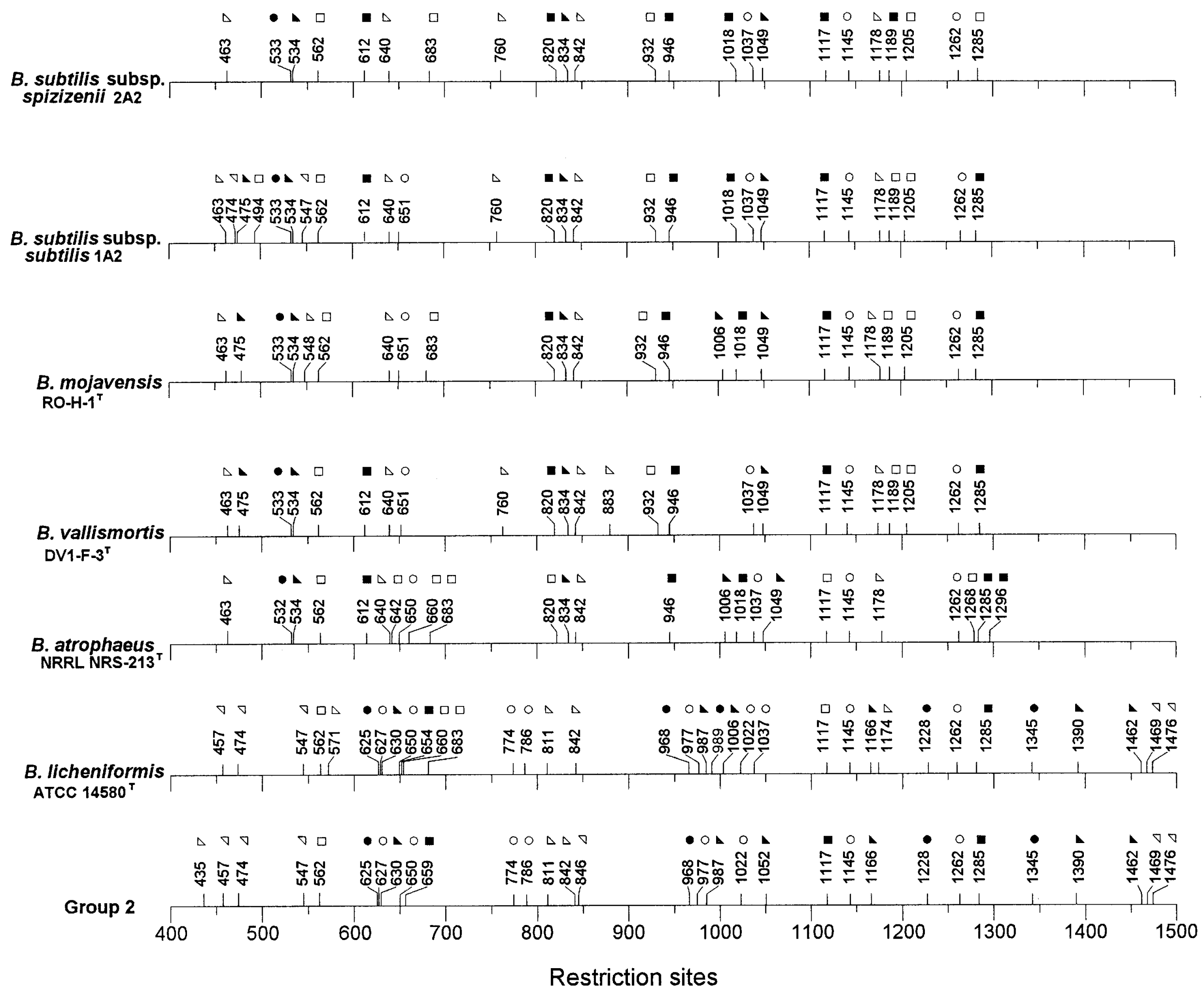

Fig. 1. Diagrammatic representation of restriction sites at *secY* for eight Group 2 strains and one strain from each related taxon. All Group 2 strains showed identical restriction digest patterns. All sites were inferred from restriction fragment patterns observed on 5% polyacrylamide gels using a computer algorithm developed by Cohan *et al.* (1991). Symbols for restriction enzymes: □, *Rsa*I; ■, *Hin*fI; ○, *Hpa*II; ●, *Hae*III; ◺, *Dpn*II; ◣, *Fnu*4HI; ◹, *Alu*I. Restriction site analysis using *Alu*I was not performed on *B. atrophaeus*, *B. vallismortis*, *B. mojavensis* or *B. subtilis* subsp. *spizizenii*.

lished). The mean sequence divergence between Group 2 and *B. licheniformis* was much greater than that within either group ($\pi = 0{\cdot}040 \pm 0{\cdot}0002$). The *secY* sequence of Group 2 was much more similar to that of *B. licheniformis* than to those of all other taxa in the *B. subtilis–B. licheniformis* clade: Group 2 showed divergence levels of 0·29 or greater from taxa other than *B. licheniformis* (divergence levels of 0·30, 0·29, 0·30, 0·29, 0·30 between Group 2 and *B. atrophaeus*, *B. subtilis* subsp. *spizizenii*, *B. subtilis* subsp. *subtilis*, *B. mojavensis* and *B. vallismortis*, respectively).

Nucleotide site variation at the *rpoB* gene

Based on the 402 bp (bases 1136–1537) sequenced from the *rpoB* gene of eight Group 2 strains and four *B. licheniformis* strains, Group 2 and *B. licheniformis* appear as distinct groups. The mean sequence divergence levels within Group 2 ($\pi = 0{\cdot}0013 \pm 0{\cdot}0006$) and within *B. licheniformis* ($\pi = 0{\cdot}015 \pm 0{\cdot}014$) are each much less than the mean sequence divergence between them ($\pi = 0{\cdot}060 \pm 0{\cdot}014$). Group 2 and *B. licheniformis* are each more divergent from *B. subtilis* ($\pi = 0{\cdot}123 \pm 0{\cdot}041$ and $\pi = 0{\cdot}126 \pm 0{\cdot}055$, respectively) than they are from one another. Group 2 strains can be distinguished from *B. licheniformis* on the basis of the following signature nucleotides: G (position 1220), T (1304), A (1337), A (1340), T (1349), A (1358), C (1370), G (1403), T (1406), C (1412), T (1424), T (1448), T (1469), G (1481), T (1484), T (1487), C (1505), A (1525) and C (1526).

Nucleotide site variation at the 16S rRNA gene

Complete and partial 16S rRNA sequences were obtained from GenBank and *SubtiList* and were aligned with sequences from Group 2 (using Omiga 2.0; Oxford Molecular); analysis was based

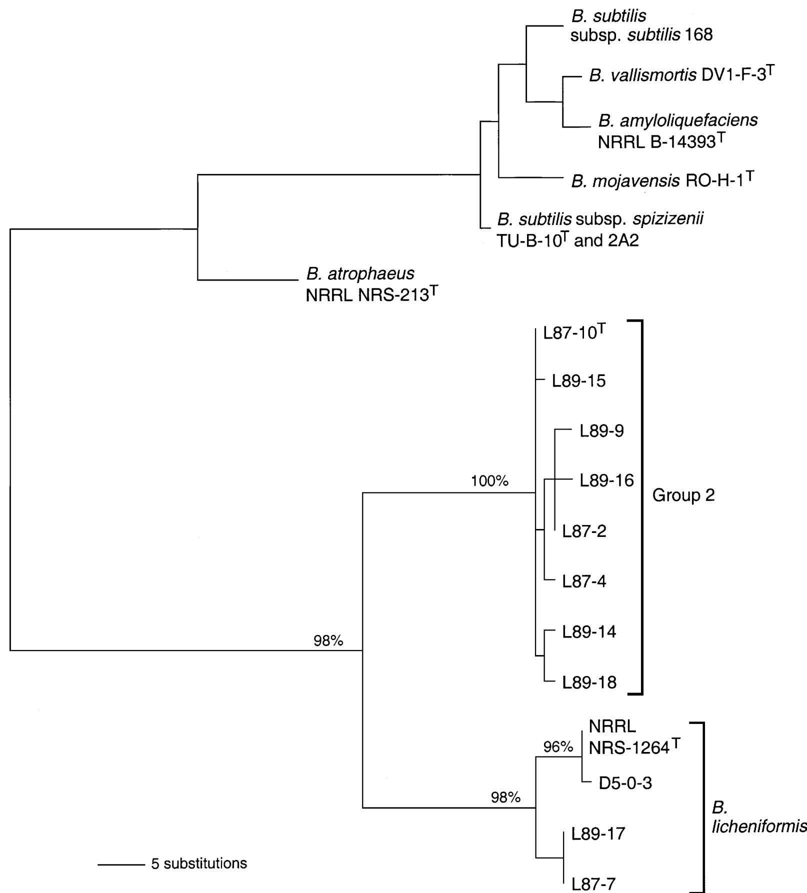

Fig. 2. A maximum-parsimony phylogeny of *Bacillus* species based on a composite of sequence data for *rpoB* and 16S rRNA, as well as restriction site data for *secY*. Numbers indicate bootstrap support values from 100 replicates. The phylogeny was based on: (i) partial (1410 bp) 16S rRNA sequences of the eight *B. sonorensis* (Group 2) strains, *B. subtilis* subsp. *subtilis* strain 168 and the type strains of the other taxa; (ii) partial (402 bp) *rpoB* sequences of the *B. sonorensis* strains, *B. subtilis* strain 168 and *B. licheniformis* strains NRRL NRS-1264^T, D5-0-3, L89-17 and L87-7; and (iii) restriction digests of *secY* for the *B. sonorensis* strains, *B. subtilis* subsp. *subtilis* strain 1A2 (a laboratory strain derived from strain 168), *B. subtilis* subsp. *spizizenii* strain 2A2 and the type strains of the other taxa. Nucleotide substitutions in *rpoB* and 16S rRNA and restriction-site changes in *secY* were counted equally in the phylogeny construction.

on 1410 bp shared by all the available sequences. The eight Group 2 strains were nearly identical in their 16S rRNA sequences, with a mean ± SD of 2·89 ± 1·20 substitutions (0·08 %) between strains. The sequences of Group 2 strains were most similar to that of *B. licheniformis* (X60623), with a mean of 8·88 ± 1·25 substitutions (0·63 %). Divergence levels between Group 2 and the other taxa of the *B. subtilis–B. licheniformis* clade were 2·19 % (*B. amyloliquefaciens*; GenBank X60605), 2·05 % (*B. atrophaeus*; GenBank AB021181), 2·19 % (*B. mojavensis*; GenBank AB021191), 2·12 % (*B. subtilis* subsp. *subtilis*; *SubtiList* BG00005), 2·12 % (*B. subtilis* subsp. *spizizenii*; GenBank AB074970) and 2·33 % (*B. vallismortis*; GenBank AB021198).

Phylogeny

A phylogeny was constructed based on a composite of sequence data for *rpoB* and 16S rRNA, as well as the restriction site data for *secY*. Fig. 2 shows the most parsimonious tree, based on the branch-and-bound algorithm of PAUP* (Swofford, 2000). The tree is of length 185 substitutions, with a consistency index of 0·886 and a retention index of 0·943.

Group 2 is supported as a monophyletic group, with a bootstrap value of 100 % (Fig. 2). Also, a bootstrap value of 98 % indicates that Group 2 and *B. licheniformis* together form a clade distinct from *B. subtilis* and the other taxa of the *B. subtilis–B. licheniformis* clade. Separate phylogenetic analyses of *rpoB*, *secY* and 16S rRNA also indicated the monophyly of Group 2, as well as the monophyly of the group containing Group 2 and *B. licheniformis* (data not shown).

DISCUSSION

A previous survey of strains with *B. licheniformis* phenotype revealed two highly distinct groups, *B. licheniformis* Group 1 (*B. licheniformis sensu stricto*) and *B. licheniformis* Group 2 (Duncan *et al.*, 1994). This earlier study could not distinguish these groups phenotypically but, based on their high divergence in allozyme frequencies, Duncan *et al.* (1994) suggested that Group 2 and *B. licheniformis* were most likely separate species. The present analysis of Group 2 strains has provided additional evidence that these strains represent a newly identified species of *Bacillus*.

DNA–DNA reassociation experiments showed Group 2 to be a taxon distinct from *B. licheniformis* and all other species in the *B. subtilis–B. licheniformis* clade. The Group 2 strains formed a cohesive group with 80–100 % reassociation. In contrast, when Group 2 isolates were compared with the type strains of other species of *Bacillus*, the reassociation values were low

and ranged from 18 to 46 %. These data pass the DNA reassociation criterion of Wayne *et al.* (1987) for naming Group 2 as a separate species.

Group 2 forms a sequence cluster distinct from other taxa of the clade that includes *B. subtilis* and *B. licheniformis*. Bootstrap analysis of the phylogeny based on *secY*, *rpoB* and 16S rRNA shows Group 2 to be monophyletic. The present study has also revealed diagnostic differences between Group 2 and *B. licheniformis* in their salt tolerances and pigmentation patterns. Because Group 2 is distinguishable from other closely related taxa by phenotypic characters, the extent of DNA–DNA relatedness and sequence divergence in three genes, this group merits recognition as a separate species, for which we propose the name *Bacillus sonorensis* sp. nov.

B. sonorensis appears to be related most closely to *B. licheniformis*. These two taxa form a monophyletic group on the basis of the sequences of *secY*, *rpoB* and 16S rRNA (Fig. 2). Moreover, *B. sonorensis* shares more phenotypic traits with *B. licheniformis* than with any other taxon (Table 2).

Two lines of evidence suggest that *B. sonorensis* and *B. licheniformis* are ecologically distinct. Firstly, these species are sympatric in at least part of their ranges (having been sampled from the same 200 cm^3 of soil) and they form distinct sequence clusters. Palys *et al.* (1997) and Dykhuizen (1998) have argued that highly distinct sequence clusters could not co-exist in sympatry unless they were ecologically distinct. Further evidence of ecological distinctness comes from the geographical distributions of these taxa. While *B. licheniformis* occurs with high frequency at both Tumamoc Hill and in Death Valley, Group 2 appears only at Tumamoc Hill. Because species of the *B. subtilis*–*B. licheniformis* clade migrate long distances at extremely high frequencies (Roberts & Cohan, 1995; M. Palmisano and F. Cohan, unpublished data), the geographical differences observed are likely due to ecological differences between taxa. It is not clear what might be the nature of the ecological differences between these species, but one possibility is that the greater salt tolerance of *B. licheniformis* adapts this species to more saline soils.

Description of *Bacillus sonorensis* sp. nov.

Bacillus sonorensis (so.no.ren′sis. N.L. adj. *sonorensis* of the Sonoran, named after the Sonoran Desert, where the organism was collected).

Vegetative cells are rod-shaped and are 1·0 μm wide by 2–5 μm long, as determined by measurements of photomicrographs. Cells often occur singly but a few chains of two to four cells are also seen. Gram-positive, motile, catalase-positive, facultatively anaerobic. Colonies on TBAB agar are a yellowish-cream colour, form mounds or lobes of amorphous slime and are approximately 2–4 mm in diameter after 2 d at 30 °C. Colonies are bright yellow on pH 5·6 agar, brown on tyrosine agar and pale yellowish-cream on glycerol/glutamate agar. Spores are formed in unswollen sporangia and are slightly less than 1·0 μm wide and 1·5–2·0 μm long. They are ellipsoidal in shape and their position within the sporangia appears to be generally subterminal. The maximum growth temperature is about 55 °C and the minimum is about 15 °C. Growth occurs at pH 5·7. Growth occurs in 3 % NaCl but not in 5, 7 or 10 % NaCl. Growth is inhibited in 0·001 % lysozyme. Acid is produced from glucose, arabinose, xylose and mannitol. The Voges–Proskauer reaction is positive and the pH after 2 d at 28 °C ranges from 5·1 to 6·6. The egg yolk reaction is negative. Hydrolyses starch and decomposes casein. Citrate and propionate are utilized. Reduces nitrate to nitrite. Tyrosine is not degraded. All of the characteristics described above are identical to those of *B. licheniformis* with these exceptions: growth of the *B. licheniformis* type strain occurs in 5, 7 and 10 % NaCl; colonies of *B. licheniformis* are cream coloured on pH 5·6 and tyrosine agar and reddish-brown on glycerol/glutamate agar. The thermal denaturation temperature for DNA of the type strain of *B. sonorensis* is 88·1 °C and the G + C content determined from this value is 46 mol %, which does not distinguish this species from *B. licheniformis*. A number of traits can be used to distinguish *B. sonorensis* from *B. licheniformis*. Diagnostic phenotypic traits include pigmentation when cells are grown on tyrosine agar, pH 5·6 agar and glycerol/glutamate agar and differences in growth in 5, 7 and 10 % NaCl. Distinguishing genetic traits include DNA–DNA reassociation values, sequences of *secY*, *rpoB* and 16S rRNA genes and enzyme electrophoresis analysis of ten loci.

Isolated from desert soil. The type strain is L87-10^T, which has been deposited in the Agricultural Research Service Culture Collection as strain NRRL B-23154^T and in the DSMZ as strain DSM 13779^T.

ACKNOWLEDGEMENTS

We are indebted to Ms Nancy Ferguson for laboratory assistance throughout the original isolation and analysis of wild isolates of *B. licheniformis* and *B. sonorensis* from Tumamoc Hill, Tucson, Arizona. This research was supported in part by National Science Foundation grant DEB 9214040 to C.A.I. and by research grants to F.M.C. from Wesleyan University.

REFERENCES

Arai, T. & Mikami, Y. (1972). Chromogenicity of *Streptomyces*. *Appl Microbiol* **23**, 402–406.

Ash, C., Farrow, J. A. E., Wallbanks, S. & Collins, M. D. (1991). Phylogenetic heterogeneity of the genus *Bacillus* revealed by comparative analysis of small-subunit-ribosomal RNA sequences. *Lett Appl Microbiol* **13**, 202–206.

Boor, K. J., Duncan, M. L. & Price, C. W. (1995). Genetic and transcriptional organization of the region encoding the β subunit of *Bacillus subtilis* RNA polymerase. *J Biol Chem* **270**, 20329–20336.

Claus, D. & Berkeley, R. C. W. (1986). Genus *Bacillus* Cohn 1872, 174^{AL}. In *Bergey's Manual of Systematic Bacteriology*, vol. 2,

pp. 1105–1139. Edited by P. H. A. Sneath, N. S. Mair, M. E. Sharpe & J. G. Holt. Baltimore: Williams & Wilkins.

Cohan, F. M., Roberts, M. S. & King, E. C. (1991). The potential for genetic exchange by transformation within a natural population of *Bacillus subtilis*. *Evolution* **45**, 1393–1421.

De Ley, J., Cattoir, H. & Reynaerts, A. (1970). The quantitative measurement of DNA hybridization from renaturation rates. *Eur J Biochem* **12**, 133–142.

Duncan, K. E., Ferguson, N., Kimura, K., Zhou, X. & Istock, C. A. (1994). Fine-scale genetic and phenotypic structures in natural populations of *Bacillus subtilis* and *Bacillus licheniformis*: important implications for bacterial evolution and speciation. *Evolution* **48**, 2002–2025.

Dykhuizen, D. E. (1998). Santa Rosalia revisited: why are there so many species of bacteria? *Antonie Leeuwenhoek* **73**, 25–33.

Gordon, R. E., Haynes, W. C. & Pang, C. H. (1973). *The Genus Bacillus*. Agriculture Handbook no. 427. Washington, DC: US Department of Agriculture.

Istock, C. A., Bell, J. A., Ferguson, N. & Istock, N. L. (1996). Bacterial species and evolution: theoretical and practical perspectives. *J Ind Microbiol* **17**, 137–150.

Lane, D. L. (1991). 16S/23S rRNA sequencing. In *Nucleic Acid Techniques in Bacterial Systematics*, pp. 115–175. Edited by E. Stackebrandt & M. Goodfellow. New York: Wiley.

Majewski, J. & Cohan, F. M. (1998). The effect of mismatch repair and heteroduplex formation on sexual isolation in *Bacillus*. *Genetics* **148**, 13–18.

Majewski, J. & Cohan, F. M. (1999). DNA sequence similarity requirements for interspecific recombination in *Bacillus*. *Genetics* **153**, 1525–1533.

Majewski, J., Zawadzki, P., Pickerill, P., Cohan, F. M. & Dowson, C. G. (2000). Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. *J Bacteriol* **182**, 1016–1023.

Mandel, M. & Marmur, J. (1968). Use of ultraviolet absorbance-temperature profile for determining the guanine plus cytosine content of DNA. *Methods Enzymol* **12B**, 195–206.

Nakamura, L. K. (1987). Deoxyribonucleic acid relatedness of lactose-positive *Bacillus subtilis* strains and *Bacillus amyloliquefaciens*. *Int J Syst Bacteriol* **37**, 444–445.

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

Nakamura, L. K. & Swezey, J. (1983). Taxonomy of *Bacillus circulans* Jordan 1890: base composition and reassociation of deoxyribonucleic acid. *Int J Syst Bacteriol* **33**, 46–52.

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

Palys, T., Nakamura, L. K. & Cohan, F. M. (1997). Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *Int J Syst Bacteriol* **47**, 1145–1156.

Pettersson, B., Rippere, K. E., Yousten, A. A. & Priest, F. G. (1999). Transfer of *Bacillus lentimorbus* and *Bacillus popilliae* to the genus *Paenibacillus* with emended descriptions of *Paenibacillus lentimorbus* comb. nov. and *Paenibacillus popilliae* comb. nov. *Int J Syst Bacteriol* **49**, 531–540.

Roberts, M. S. & Cohan, F. M. (1993). The effect of DNA sequence divergence on sexual isolation in *Bacillus*. *Genetics* **134**, 401–408.

Roberts, M. S. & Cohan, F. M. (1995). Recombination and migration rates in natural populations of *Bacillus subtilis* and *Bacillus mojavensis*. *Evolution* **49**, 1081–1094.

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

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

Sasser, M. (1990). *Identification of Bacteria by Gas Chromatography of Cellular Fatty Acids*. Technical Note 101. Newark, DE: Microbial ID, Inc.

Suh, J.-W., Boylan, S. A., Thomas, S. M., Dolan, K. M., Oliver, D. B. & Price, C. W. (1990). Isolation of a *secY* homologue from *Bacillus subtilis*: evidence for a common protein export pathway in eubacteria. *Mol Microbiol* **4**, 305–314.

Swofford, D. L. (2000). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4. Sunderland, MA: Sinauer Associates.

Tschauder, S., Driessen, A. J. M. & Freudl, R. (1992). Cloning and molecular characterization of the *secY* genes from *Bacillus licheniformis* and *Staphylococcus carnosus*: comparative analysis of nine members of the SecY family. *Mol Gen Genet* **235**, 147–152.

Wayne, L. G., Brenner, D. J., Colwell, R. R. & 9 other authors (1987). Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* **37**, 463–464.

Zawadzki, P. & Cohan, F. M. (1995). The size and continuity of DNA segments integrated in *Bacillus* transformation. *Genetics* **141**, 1231–1243.

Zawadzki, P., Roberts, M. S. & Cohan, F. M. (1995). The log-linear relationship between sexual isolation and sequence divergence in *Bacillus* transformation is robust. *Genetics* **140**, 917–932.