Virgibacillus senegalensis sp. nov., a new moderately halophilic bacterium isolated from human gut

E. Seck[1], J. Rathored[1], S. Khelaifia[1], O. Croce[1], C. Robert[1], C. Couderc[1], F. Di Pinto[1], C. Sokhna[2,3], D. Raoult[1,4] and J.-C. Lagier[1]

1) *Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM 63, CNRS 7278, IRD 198, Inserm 1095, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine,* 2) *Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes IRD 198, CNRS 7278, Aix-Marseille Université, Marseille, France,* 3) *Campus Commun UCAD-IRD of Hann, Dakar, Senegal and* 4) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

Virgibacillus senegalensis SK-1^T (= CSUR P1101 = DSM 28585) is the type strain of *V. senegalensis* sp. nov. It is an aerobic, Gram positive, moderately halophilic, motile bipolar flagellum isolated from a healthy Senegalese man. Here we describe the genomic and phenotypic characteristics of this isolate. The 3 755 098 bp long genome (one chromosome, no plasmid) exhibits a G + C content of 42.9% and contains 3738 protein-coding and 95 RNA genes.

Keywords: Culturomics, genome, human gut, moderately halophilic bacteria, taxonogenomics, *Virgibacillus senegalensis*
Original Submission: 12 August 2015; **Revised Submission:** 23 September 2015; **Accepted:** 24 September 2015
Article published online: 3 October 2015

Corresponding author: J.-C. Lagier, URMITE, UMR CNRS 7278, L'Institut de Recherche pour le Développement 198, INSERM U1095, Faculté de Médecine, Aix-Marseille Université, 27 Boulevard Jean Moulin, 13385 Marseille Cedex 5, France
E-mail: jclagier@yahoo.fr

Introduction

The concept of microbial culturomics is based on the variation of physicochemical parameters of the culture conditions so as to express the maximum of microbial diversity. It is based on rapid methods for identification, such as matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) and 16S rRNA amplification and sequencing for unidentified colonies. This concept considerably enriches the gut microbiota repertoire, including new species not previously isolated from humans [1,2].

This isolation was part of a culturomics study we undertook using high-salt-containing culture conditions to grow halophilic bacteria from human stool [1].

The typical parameters used to define bacterial species comprise 16S rRNA sequencing and phylogeny, G + C content genomic diversity and DNA-DNA hybridization (DDH). However, some limitations have been noted [3–6]. By using the availability of data in genomics through the development of new tools for sequencing DNA, we introduced a new taxonomic method for the description of new bacterial species. This concept, which we named taxonogenomics, includes their genomic features [7] and proteomic information obtained by MALDI-TOF analysis [8–17].

The genus *Virgibacillus* was first proposed by Heyndrickx in 1998 with the transfer of *Bacillus pantothenticus* to *Virgibacillus pantothenticus* [18]. To date, there are more than 25 recognized species [19]. These bacteria are positive, Gram-variable rods which are ellipsoidal to oval endospores and have DNA G + C content ranging from 36% to 43% [20]. These species were isolated from sediments of a salt lake [20–23], fermented seafood in traditional salt [24], a permafrost core collected from the Canadian high Arctic [25], a navy solar salt marsh [26,27], soil [28], seawater [29], field soil, a dairy product [30], residual wash water produced during processing wastewater, Spanish-style green table olives [31], saline sample of mud, salt

http://dx.doi.org/10.1016/j.nmni.2015.09.014

crust [32] and Thai fermented fish [33]. Here we present a brief classification and a set of features for strain SK-1T (= CSUR P1101 = DSM 28585), with a description of the complete genome sequence and annotation. We named this new isolate *Virgibacillus senegalensis*.

Materials and Methods

Sample and culture condition

The stool sample was collected from a healthy male Senegalese volunteer patient living in N'diop, a rural village in the Guinean–Sudanian zone in Senegal. After the patient provided signed informed consent, the sample was collected in a sterile pot and transported to our laboratory. The study and the assent procedure were approved by the National Ethics Committee of Senegal and by the ethics committees of the Institut Fédératif de Recherche 48, Faculty of Medicine, Marseille, France (agreement 09-022). The salt concentration of the stool specimen was determined by a digital refractometer (Fisher Scientific, Illkirch, France) and the pH with a pH meter (Cyberscan 510PH; Eutech Instruments, Singapore).

Strain SK-1T was isolated in February 2014 by aerobic culture on a homemade culture medium consisting of a Columbia agar culture medium (Sigma-Aldrich, Saint-Quentin Fallavier, France) modified by adding (per liter): $MgCl_2$ $6H_2O$, 5 g; $MgSO_4$ $7H_2O$, 5 g; KCl, 2 g; $CaCl_2$ $2H_2O$, 1 g; NaBr, 0.5 g; $NaHCO_3$, 0.5 g; glucose, 2 g; and 100 g/L of NaCl. The pH was adjusted to 7.5 with 10 M NaOH before autoclaving.

MALDI-TOF identification

An isolated colony was deposited in duplicate on a MALDI-TOF target to be analysed. A matrix of 1.5 µL (saturated solution of α-cyano-4-hydroxycinnamic acid diluted in 500 µL acetonitrile, 250 µL of acid tri-fluoro-acetic to 10%, and 250 µL of HPLC water) was used on each spot. This solution enables ionization and desorption of the homogeneous biological sample with which it crystallizes. The analysis was performed by a Microflex (Bruker Daltonics, Leipzig, Germany) device, and protein spectra were compared with those of the hospital database. A score was assigned indicating the reliability of the identification of the bacteria; above 1.9 was considered proper identification. Conversely, if the bacterium was not referenced in the database, sequencing the 16S rRNA was used to achieve the correct identification [34].

Identification by sequencing of 16S rRNA

Colonies not identified by the MALDI-TOF after three tests were suspended in 200 µL of distilled water for DNA extraction by EZ1 DNA Tissue Kit (Qiagen, Venlo, The Netherlands). The amplification of the 16S rRNA was performed by standard PCR in a thermocycler using the universal primer pair FD1 and rp2 according to the following amplification program: activation of the polymerase (95°C for 5 minutes), followed by 40 cycles (95°C 30 seconds, 52°C 45 seconds, 72°C 2 minutes), followed by 5 minutes at 72°C. The DNA amplified by this reaction was revealed by electrophoresis on 1.5% agarose gel. Once validated, the PCR product was purified and sequenced using the Big Dye Terminator Sequencing Kit using the internal primers 536F, 536R, 800F, 800R, 1050F and 1050R, as previously described [2].

Phylogenetic analysis

Phylogenetic analysis based on 16S rRNA of our isolates was performed to identify its phylogenetic affiliations with other near isolates, including other members of the genus *Virgibacillus*. MEGA 6 software (http://www.megasoftware.net/mega.php) allowed us to construct a phylogenetic tree. Sequence alignment of the different species was performed using Clustal W (http://www.clustal.org/clustal2/), and the evolutionary distance was calculated with the Kimura two-parameter model [35].

Biochemical, atmospheric and antimicrobial susceptibility tests

Biochemical tests were performed using the commercially available Api ZYM (bioMérieux, Marcy l'Étoile, France), API 50CH (bioMérieux) and 20 NE (bioMérieux) strips. The incubation time was 48 hours for the API 50CH and 20 NE, and 4 hours for Api ZYM. Growth of strain SK-1T was tested in aerobic atmosphere, in the presence of 5% CO_2 and also in anaerobic and microaerophilic atmospheres, created using AnaeroGen (Atmosphere Generation Systems, Dardily, France). Antibiotic susceptibility was determined by Müller-Hinton agar in a petri dish (bioMérieux). The following antibiotics were tested: doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamicin, penicillin, trimethoprim/sulfamethoxazole, imipenem and metronidazole.

Genome Sequencing Information

Genomic DNA preparation

We cultured our strain in the homemade culture. After 48 hours, bacteria grown on four petri dishes were resuspended in sterile water and centrifuged at 4°C at 2000 × *g* for 20 minutes. Cell pellets were resuspended in 1 mL Tris/EDTA/NaCl (10 mM Tris/HCl (pH7.0), 10 mM EDTA (pH8.0) and 300 mM

(NaCl)) and recentrifuged under the same conditions. The pellets were then resuspended in 200 µL Tris-EDTA buffer and proteinase K and kept overnight at 37°C for cell lysis. DNA was purified with phenol/chloroform/isoamyl alcohol (25:24:1), followed by an overnight precipitation with ethanol at −20°C. The DNA was resuspended in 205 µL Tris-EDTA buffer. DNA concentration was 155 ng/µL as measured by a Qubit fluorometer using the high-sensitivity kit (Life Technologies, Carlsbad, CA, USA).

Genome sequencing and assembly

Genomic DNA of *Virgibacillus senegalensis* was sequenced on the MiSeq Technology (Illumina, San Diego, CA, USA) with the mate pair strategy. The gDNA was bar coded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). gDNA was quantified by a Qubit assay with the high sensitivity kit (Life Technologies) to 155 ng/µL. The mate pair library was prepared with 1 µg of genomic DNA using the Nextera mate pair Illumina guide. The genomic DNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The pattern of the fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 lab chip. The DNA fragments ranged in size from 1 to 11 kb, with an optimal size at 4.008 kb. No size selection was performed, and 388.3 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal at 634 bp on the Covaris device S2 in microtubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies), and the final concentration library was measured at 35.59 nmol/L. The libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and sequencing run were performed in a single 39-hour run in a 2 × 251 bp. Total information of 10.6 Gb was obtained from a 1326K/mm^2 cluster density with a cluster passing quality control filters of 99.1% (24 492 260 clusters). Within this run, the index representation for *Virgibacillus senegalensis* was determined to be 7.06%. The 1 481 197 paired reads were filtered according to the read qualities. These reads were trimmed, then assembled using CLC genomicsWB4 software.

Genome annotation and comparison

Open reading frames (ORFs) were predicted using Prodigal [36] with default parameters, but the predicted ORFs were excluded if they spanned a sequencing gap region. The predicted bacterial protein sequences were searched against the GenBank database [37] and the Clusters of Orthologous Groups (COGs) database using BLASTP. The tRNAScanSE tool [38] was used to find tRNA genes, whereas ribosomal RNAs were found using RNAmmer [39] and BLASTn against the GenBank database. Lipoprotein signal peptides and the number of transmembrane helices were predicted using SignalP [40] and TMHMM [41], respectively. ORFans were identified if their BLASTP E value was lower than 1e^{-03} for alignment length greater than 80 aa. If alignment lengths were smaller than 80 aa, we used an E-value of 1e^{-05}. Such parameter thresholds have already been used in previous works to define ORFans. Artemis [42] was used for data management and DNA Plotter [43] for visualization of genomic features. Mauve alignment tool (version 2.3.1) was used for multiple genomic sequence alignment [44]. To estimate the mean level of nucleotide sequence similarity at the genome level, we used an in-lab pipeline software named Marseille Average Genomic Identity (MAGi) to calculate the average genomic identity of gene sequences (AGIOS) among compared genomes [45]. Briefly, this software combines the Proteinortho software [45] for detecting orthologous proteins in pairwise genomic comparisons, then retrieves the corresponding genes and determines the mean percentage of nucleotide sequence identity among orthologous ORFs using the Needleman-Wunsch global alignment algorithm. Genomes from the genus *Virgibacillus* and closely related genera were used for the calculation of AGIOS values. Here we compared the genome sequences of *Virgibacillus senegalensis* strain SK-1^T (GenBank accession number PRJEB1962) with those of *Virgibacillus kekensis* strain YIM kkny9 (NR_042744.1), *Virgibacillus albus* strain YIM 93624 (NR_109613.1), *Aquibacillus salifodinae* WSY08-1 (AB859945.1), *Virgibacillus halodenitrificans* DSM 10037 (AY543169), *Thalassobacillus devorans* MSP14 (JX518269.1), *Halobacillus dabanensis* HD 02 (HG931924.2), *Halobacillus kuroshimensis* DSM 18393 (AB195680.1), *Thalassobacillus devorans strain* XJSL7-8 (GQ903447.1), *Bacillaceae bacterium* EFN-4 (EU817569.1), *Virgibacillus marismortui* strain M3-23 (GQ282501.1), *Halobacillus salinus* strain GSP59 (AY505517.1), *Virgibacillus alimentarius* J18 (GU202420), *Pseudomonas aeruginosa* PAO1 (NR_074828.1) and *Virgibacillus massiliensis* (CCDP010000001).

Results

Phenotypic description

Strain SK-1T was isolated in February 2014 (Table 1) by aerobic culture on a homemade culture medium at 37°C after 48 hours. No significant MALDI-TOF result for the strain SK-1^T against

TABLE 1. Classification of *Virgibacillus senegalensis* strain SK-1

Property	Term
Current classification	Domain: *Bacteria* Phylum: *Firmicutes* Class: *Bacilli* Order: *Bacillales* Family: *Bacillaceae* Genus: *Virgibacillus* Species: *Virgibacillus senegalensis* Type strain: SK-1
Gram stain	Positive
Cell shape	Rod shaped
Motility	Motile by polar flagellum
Sporulation	Endospore forming
Temperature range	Mesophile
Optimum temperature	37°C
pH	pH 5 to 9
Optimum pH	7.5
Salinity	0.5–10%
Optimum salinity	7.5%
Oxygen requirement	Aerobic

our MALDI-TOF database was obtained, suggesting that our isolate was not a previously known species. We added the specter of SK-1T strain to our database (Fig. 1). The gel view allowed us to observe the spectral differences with other members of *Virgibacillus* genus (Fig. 2). PCR-based identification of the 16S rRNA of our new isolate (GenBank accession number LK021111) yielded 96.3% 16S rRNA sequence similarity with the reference *Virgibacillus kekensis* (GenBank accession number NR042744), the phylogenetically closest validated *Virgibacillus* species (Fig. 3).

After growth for 24 hours on our homemade culture medium at 37°C, the surface colonies were circular, greyish, shiny and smooth, with a diameter of 1 to 2 mm. *V. senegalensis* is Gram positive (Fig. 4).

Growth was observed at temperatures ranging from 25 to 40°C, with an optimum at 37°C. The growth required a salinity ranging from 5 to 200 g/L of NaCl (optimum at 75 g/L). The optimum pH for growth was 7.5 (pH range 5 to 9). Growth of the strain SK-1^T was tested in an aerobic atmosphere, in the presence of 5% CO_2 and also in anaerobic and microaerophilic atmospheres created using AnaeroGen (Atmosphere Generation Systems), respectively. The strain was strictly aerobic and also grew in the presence of 5% CO_2 but did not grow in an anaerobic atmosphere. The size and ultrastructure of cells were determined by negative staining transmission electron microscopy 2 to 6 µm in length and 0.5 µm in diameter (Fig. 5). Using the commercially available Api ZYM, Api 20NE (bioMérieux), to characterize the biochemical *V. senegaiensis* strain SK-1^T, positive reactions were observed for urease, β-glucosidase,

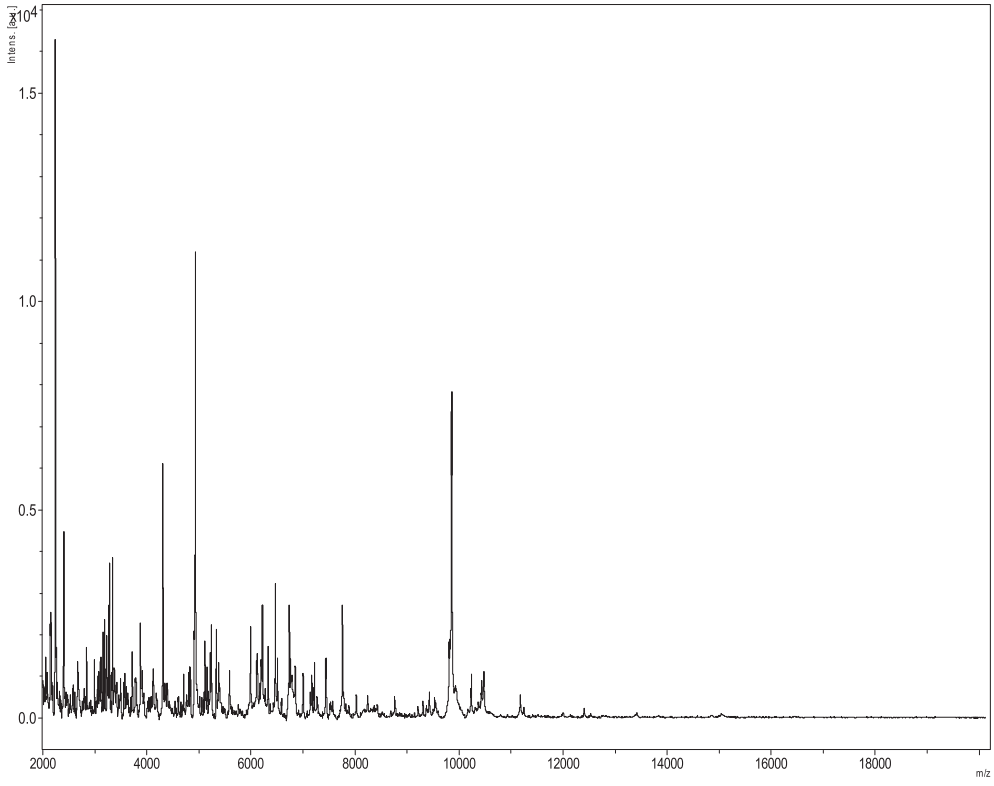

FIG. 1. Reference mass spectrum from *Virgibacillus senegalensis* sp. nov. SK-1^T. Spectra from 12 individual colonies were compared and reference spectrum generated.

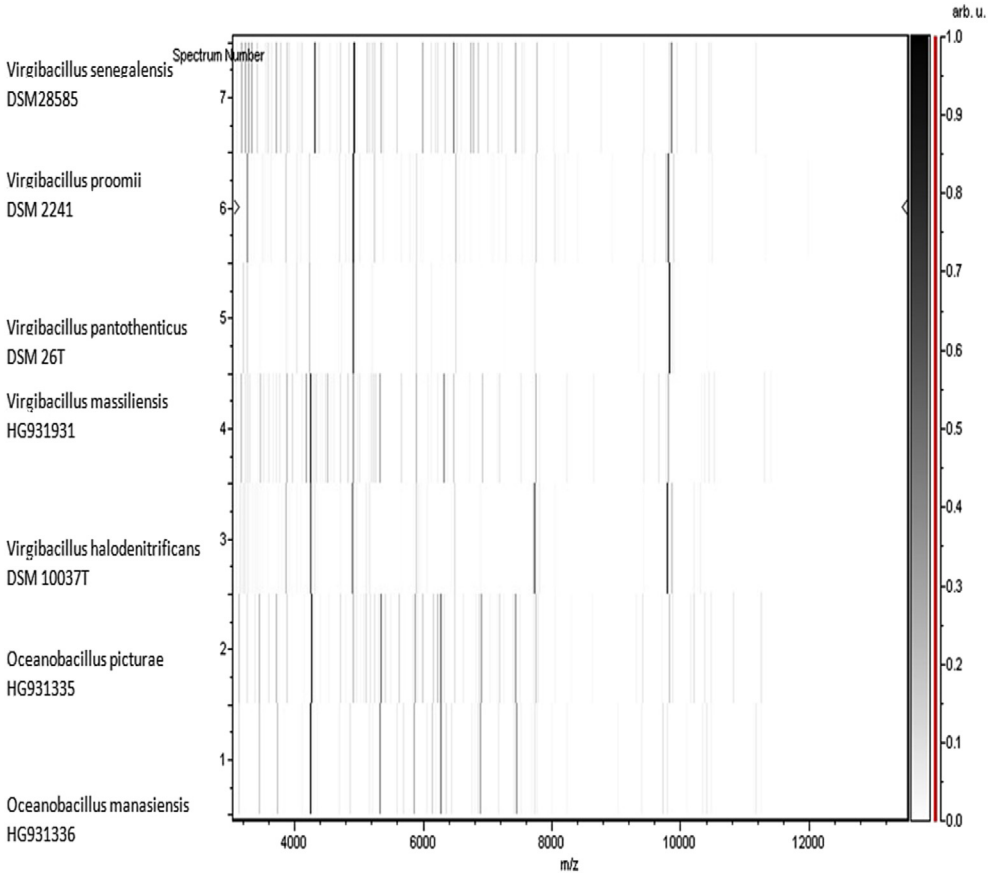

FIG. 2. Gel view comparing *Virgibacillus senegalensis* sp. nov. SK-1^{T}to members of family *Virgibacillus* and *Oceanobacillus*. Gel view displays raw spectra of all loaded spectrum files arranged in pseudo-gel-like look. X-axis records m/z value. Left y-axis displays running spectrum number originating from subsequent spectra loading. Peak intensity is expressed by greyscale scheme code. Color bar and right y-axis indicate relation between color peak; peak intensity expressed in arbitrary units. Displayed species are indicated at left.

protease, β-galactosidase and arginine hydrolase. All other tested reactions were negative, notably nitrate reduction alkaline phosphatase and N-acetyl-β-glucosaminidase. The stain was also catalase and oxidase negative. Substrate oxidation and assimilation were examined with an API 50CH strip (bioMérieux) at 37°C. Negative reactions were obtained for D-mannose, D-lactose, L-arabinose, D-galactose, D-ribose, D-sucrose, D-fructose, D-glucose, D-mannitol and D-maltose. Phenotypic characteristics were compared to those of the most closely related species (Table 2).

Finally, antimicrobial susceptibility testing demonstrate that the strain SK-1^T was susceptible to doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamicin, penicillin, trimethoprim/sulfamethoxazole and imipenem, but resistant to metronidazole.

Genome Sequencing Information

Genome properties

The draft genome of *V. senegalensis* consists of nine scaffolds with 59 contigs.

The genome is, 755 098 bp long with a 42.9% G + C content (Table 3, Fig. 6). Of the 3833 predicted genes, 3738 (96.46%) were protein-coding genes and 95 (2.44%) were RNAs (14 genes were 5S rRNA, five genes 16S rRNA, seven genes 23S rRNA, 69 genes tRNA). A total of 2773 genes (62.82%) were assigned a putative function, 2427 genes (65.43%) were assigned to COGs and 155 genes (4.04%) contained peptide signals, whereas 980 (25.56%) genes had transmembrane helices. A total of 245 genes were identified as ORFans (6.39%). The remaining genes were annotated as hypothetical proteins. The properties and statistics of the genome are summarized in

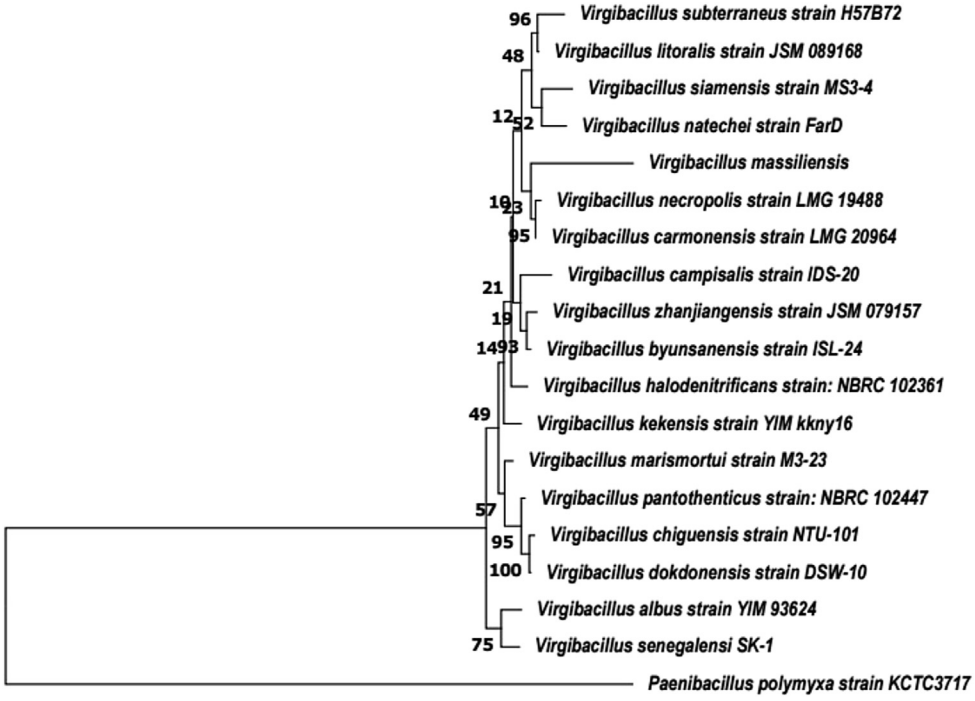

FIG. 3. Phylogenetic tree highlighting position of *Virgibacillus senegalensis* strain SK-1T (= CSUR P1101 = DSM 28585) relative to other type strains of *Virgibacillus albus* strain YIM 93624 (NR_109613.1), *Virgibacillus kekensis* strain YIM kkny16 (NR_042744.1), *Virgibacillus alimentarius* strain J18 (NR_108710.1), *Virgibacillus marismortui* strain M3-23 (GQ282501.1), *Virgibacillus necropolis* strain LMG 19488 (NR_025472.1), *Virgibacillus carmonensis* strain LMG 20964 (NR_025481.1), *Virgibacillus subterraneus* strain H57B72 (FJ746573.1), *Virgibacillus zhanjiangensis* strain JSM 079157 (FJ425904.1), *Virgibacillus litoralis* strain JSM 089168 (FJ425909.1), *Virgibacillus dokdonensis* strain DSW-10 (NR_043206.1), *Virgibacillus siamensis* strain MS3-4 (AB365482.1), *Virgibacillus salarius* strain SA-Vb1 (NR_041270.1), *Virgibacillus halophilus* strain 5B73C (NR_041358.1), *Virgibacillus natechei* strain FarD (NR_132721.1), *Virgibacillus chiguensis* strain NTU-101 (NR_044086.1), *Virgibacillus dokdonensis* strain DSW-10 (NR_043206.1), *Virgibacillus campisalis* strain IDS-20 (GU586225.1), *Virgibacillus pantothenticus* strain NBRC 102447 (AB681789.1), *Virgibacillus halodenitrificans* strain NBRC 102361 (AB681753.1), *Virgibacillus byunsanensis* strain ISL-24 (FJ357159.1), *Virgibacillus massiliensis* strain Vm-5 (HG931931.1) and *Paenibacillus polymyxa* strain KCTC3717 (AY359637.1). GenBank accession numbers are indicated in parentheses. Sequences were aligned using Clustal W (http://www.clustal.org/clustal2/), and phylogenetic inferences were obtained using maximum-likelihood method within MEGA 6 (http://www.megasoftware.net/mega.php). *Paenibacillus polymyxa* was used as outgroup. Scale bar = 0.005% nucleotide sequence divergence.

Tables 3 and 4. The distribution of genes into COGs functional categories is presented in Table 5.

Genome comparison

The draft genome of V. senegalensis SK-1^T is smaller than those of Halobacillus kuroshimensis DSM 18393, Virgibacillus halodenitrificans DSM10037, Thalassobacillus devorans XJSL7-8, Thalassobacillus devorans XJSL7-8, Halobacillus dabanensis HD 02 and Pseudomonas aeruginosa PAO1 (3.85, 3.92, 3.94, 4.1 and 6.26 Mb, respectively) but larger than that of Virgibacillus alimentarius J18 (3.05 Mb). The G + C content of V. senegalensis SK-1^T is smaller than those of Halobacillus kuroshimensis DSM 18393 and Pseudomonas aeruginosa PAO1 (47.0% and 66.60%, respectively) and larger than those of Virgibacillus alimentarius J18, Virgibacillus halodenitrificans DSM10037 and Halobacillus dabanensis HD 02 (37.1%, 37.4% and 41.5% respectively) but equal to Thalassobacillus devorans XJSL7-8 (42.9%). Protein-coding genes of V. senegalensis SK-1^T were smaller than those of Virgibacillus halodenitrificans DSM10037, Thalassobacillus devorans XJSL7-8, Halobacillus kuroshimensis DSM 18393, Halobacillus dabanensis HD 02 and Pseudomonas aeruginosa PAO1 (3748, 3752, 3832, 3835 and 5572 Mb, respectively) but larger than those of Virgibacillus alimentarius J18 (2889 Mb). Total gene content of V. senegalensis SK-1^T (3883) is smaller than those of Halobacillus kuroshimensis DSM 18393, Halobacillus dabanensis HD 02 and Pseudomonas aeruginosa PAO1 (3915, 4011, and 5697 respectively) but larger than those of Thalassobacillus devorans XJSL7-8, Virgibacillus halodenitrificans DSM10037 and Virgibacillus alimentarius J18 (3840, 3822 and 3022, respectively).

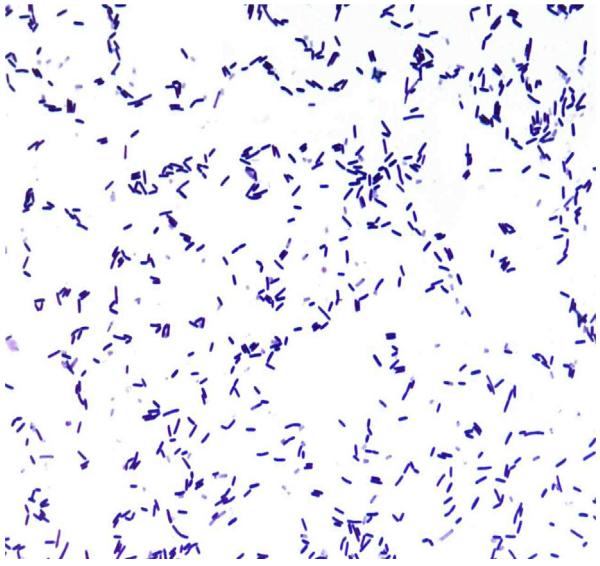

FIG. 4. Gram staining of *Virgibacillus senegalensis* sp. nov. SK-1^T.

Among species with standing in nomenclature, AGIOS values ranged from 66.41% between *V. senegalensis* SK-1^T and *Halobacillus kuroshimensis* DSM 18393 to 73.39% between *Halobacillus dabanensis* HD 02 and *Halobacillus kuroshimensis* DSM 18393. To evaluate the genomic similarity among studied strains, in addition to AGIOS [7], which was designed to be independent from DDH, we determined a digital DDH that exhibited a high correlation with DDH [46,47]. Digital DDH ranged from 18.4% to 27.2% between the different species tested (Table 6, Fig. 7).

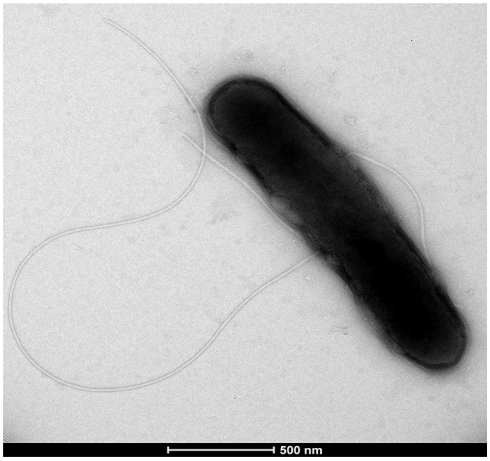

FIG. 5. Transmission electron microscopy of *Virgibacillus senegalensis* sp. nov. SK-1^T. Cells were observed on Tecnai G20 transmission electron microscope operated at 200 keV. Scale bar = 500 nm.

TABLE 2. **Differential characteristics of *Virgibacillus* species**

Property	*V. senegalensis*	*V. massiliensis*	*V. olivae*	*V. salarius*	*V. marismortui*	*V. sediminis*	*V. xinjiangensis*	*V. kekensis*	*V. halodenitrificans*	*V. proomii*	*V. dokdonensis*
Cell diameter (µm)	0.6–0.9	0.5–0.8	0.4–0.6	0.6–0.9		0.4–0.7	1.4–2.4	0.3–0.5	0.6–0.8	0.5–0.7	
Oxygen requirement	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic
Gram stain	+	+	+	+	+	+	+	+	+	+	+
Salt requirement	+	+	+	+	−	+	−	−	+	NA	+
Motility	+	+	+	+	+	+	+	+	+	+	+
Endospore formation	+	+	+	+	+	+	+	+	+	+	+
Indole	−	−	−	−	−	−	−	−	−	−	−
Production of											
Alkaline phosphatase	−	−	NA	NA	NA	−	NA	−	NA	NA	−
Catalase	−	+	+	+	+	+	+	+	+	+	+
Oxidase	−	+	+	+	+	+	−	+	+	NA	+
Nitrate reductase	−	+	+	−	+	+	+	−	+	+	−
Urease	+	−	NA	−	NA	−	+	−	+	NA	−
β-Galactosidase	+	−	NA	−	−	−	−	−	+	+	−
N-acetyl-glucosamine	−	−	NA	+	+	−	NA	−	NA	+	−
Acid from:											
L-Arabinose	−	−	NA	−	−	−	−	−	−	−	−
Ribose	−	−	NA	NA	NA	+	−	−	NA	+	+
D-Mannose	−	+	−	+	+	−	−	+	+	+	+
D-Manitol	−	+	NA	−	−	−	−	w	+	+	−
D-Sucrose	−	+	−	−	−	−	+	−	−	+	+
D-Glucose	−	+	−	+	+	+	+	+	+	+	+
D-Fructose	−	+	+	+	+	+	+	−	+	+	+
D-Maltose	−	+	−	+	−	+	−	+	+	+	−
D-Lactose	−	−	−	−	−	−	−	−	+	+	+
Habitat	Human gut	Human gut	Waste wash water	Salt lake	Mural paintings	Salt lake	Salt lake	Salt lake	Solar saltern	Soil	Soil

NA, data not available; w, weak reaction.

TABLE 3. Nucleotide content and gene count levels of genome

Attribute	Value	% of total[a]
Size (bp)	3 755 098	100
G + C content (%)	1 610 937	42.9
Coding region (bp)	3 129 675	83.34
Total genes	3883	100
RNA genes	95	2.44
Protein-coding genes	3738	98.46
Genes with function prediction	2773	62.82
Genes assigned to COGs	2421	65.43
Genes with peptide signals	155	4.04
Genes with transmembrane helices	980	25.56
CRISPRs	2	0.05
Genes with Pfam domains	2011	52.46

COGs, Clusters of Orthologous Groups database; CRISPR, clustered regularly interspaced short palindromic repeat.
[a]Total is based on either size of genome (bp) or total number of protein-coding genes in annotated genome.

TABLE 4. Number of genes associated with the 25 general COGs functional categories

Code	Value	% value[a]	Description
J	172	4.60	Translation
A	0	0	RNA processing and modification
K	262	7.01	Transcription
L	196	5.24	Replication, recombination and repair
B	1	0.03	Chromatin structure and dynamics
D	30	0.80	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	49	1.31	Defense mechanisms
T	153	4.09	Signal transduction mechanisms
M	156	4.17	Cell wall/membrane biogenesis
N	64	1.71	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	51	1.36	Intracellular trafficking and secretion
O	92	2.46	Posttranslational modification, protein turnover, chaperones
C	137	3.67	Energy production and conversion
G	275	7.36	Carbohydrate transport and metabolism
E	305	8.16	Amino acid transport and metabolism
F	82	2.19	Nucleotide transport and metabolism
H	93	2.49	Coenzyme transport and metabolism
I	102	2.73	Lipid transport and metabolism
P	202	5.40	Inorganic ion transport and metabolism
Q	74	1.98	Secondary metabolites biosynthesis, transport and catabolism
R	448	11.99	General function prediction only
S	287	7.68	Function unknown
—	352	9.06	Not in COGs

COGs, Clusters of Orthologous Groups database.
[a]Total is based on total number of protein-coding genes in annotated genome.

Conclusion

On the basis of phenotypic, genomic and phylogenetic analyses, we formally propose the creation of *Virgibacillus senegalensis* sp. nov., represented here by the SK-1T strain. The strain was isolated from a stool sample of a Senegalese healthy individual.

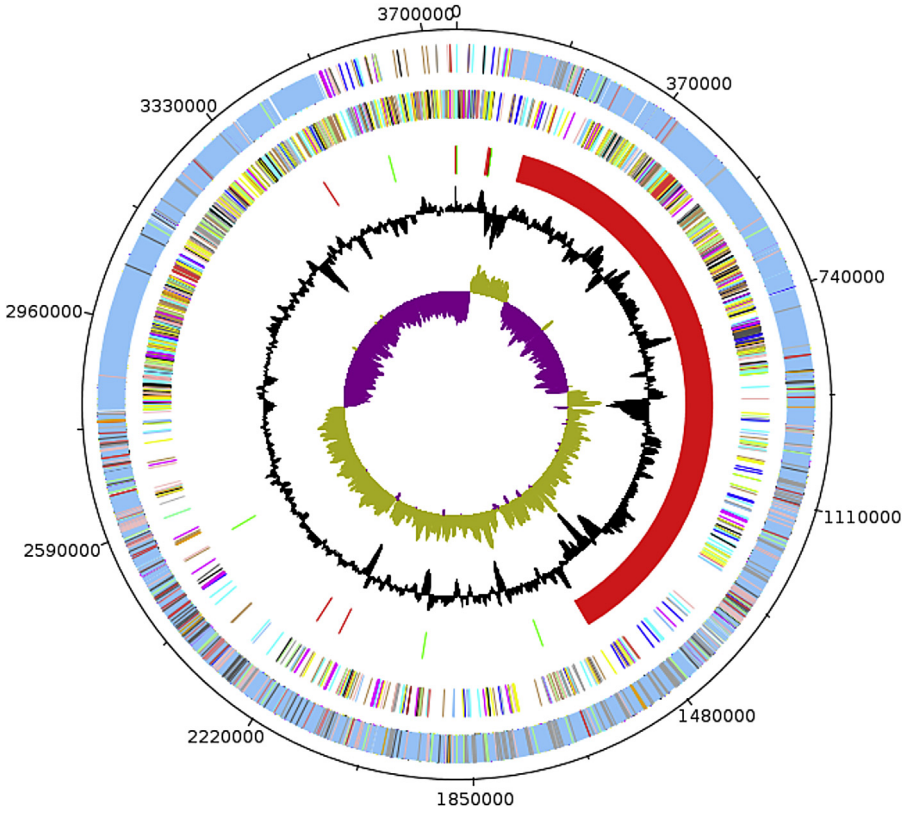

FIG. 6. Graphical circular map of *Virgibacillus senegalensis* sp. nov. SK-1^T chromosome. From outside in, outer two circles show open reading frames oriented in forward (colored by COGs categories) and reverse (colored by COGs categories) directions, respectively. Third circle marks tRNA genes (green). Fourth circle shows percentage G + C content plot. Innermost circle shows GC skew, purple indicating negative values and olive positive values. COGs, Clusters of Orthologous Groups database.

TABLE 5. **Numbers of orthologous proteins shared between genomes (upper right), average percentage similarity of nucleotides corresponding to orthologous proteins shared between genomes (lower left) and numbers of proteins per genome (bold)**

	Virgibacillus senegalensis	*Halobacillus dabanensis*	*Halobacillus kuroshimensis*	*Thalassobacillus devorans*	*Virgibacillus alimentarius*	*Virgibacillus halodenitrificans*	*Pseudomonas aeruginosa*	*Virgibacillus massiliensis*
V. senegalensis	**3378**	1791	1786	1776	1446	1741	601	588
H. dabanensis	66.82	**4063**	2218	2122	1581	1866	697	642
H. kuroshimensis	66.41	73.39	**3926**	2119	1554	1861	695	630
T. devorans	67.76	69.22	68.46	**3880**	1596	1923	687	618
V. alimentarius	67.16	66.25	64.45	66.55	**3119**	1663	607	604
V. halodenitrificans	67.19	66.41	64.71	66.67	71.39	**3876**	661	676
P. aeruginosa	53.94	52.73	55.86	53.96	49.86	50.02	**5681**	244
V. massiliensis	66.21	65.24	63.48	65.57	69.67	70.87	50.45	**1768**

TABLE 6. **Pairwise comparison of *Virgibacillus senegalensis* with eight other species**[a]

	Virgibacillus senegalensis	*Halobacillus dabanensis*	*Halobacillus kuroshimensis*	*Thalassobacillus devorans*	*Virgibacillus alimentarius*	*Virgibacillus halodenitrificans*	*Pseudomonas aeruginosa*	*Virgibacillus massiliensis*
V. senegalensis	100% ± 00	24.3% ± 2.55	26.7% ± 2.54	22.8% ± 2.54	26.4% ± 2.54	21.8% ± 2.56	20.8% ± 2.53	22.8% ± 2.53
H. dabanensis		100% ± 00	26.3% ± 2.55	24.4% ± 2.56	27.7% ± 2.55	27% ± 2.55	21.2% ± 2.57	20.0% ± 2.53
H. kuroshimensis			100% ± 00	24.6% ± 2.55	27.2% ± 2.55	24.7% ± 2.57	21.8% ± 2.54	19.5% ± 2.53
T. devorans				100% ± 00	22.1% ± 2.55	21.3% ± 2.57	18.4% ± 2.56	20.1% ± 2.53
V. alimentarius					100% ± 00	23% ± 2.56	21.1% ± 2.53	20.0% ± 2.54
V. halodenitrificans						100% ± 00	22.5% ± 2.54	21.3% ± 2.55
P. aeruginosa							100% ± 00	20.3% ± 2.52
V. massiliensis								100% ± 00

[a]Comparison made using GGDC, formula 2 (DDH estimates based on identities/HSP length). Confidence intervals indicate inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical test data sets (which are always limited in size). These results are in accordance with 16S rRNA (Fig. 1) and phylogenomic analyses as well as GGDC results. DDH, DNA-DNA hybridization; HSP, high-scoring pair.

Taxonomic and nomenclatural proposals

Description of *Virgibacillus senegalensis* sp. nov.

Virgibacillus senegalensis (se.ne.ga.len'sis. L. masc. adj. *senegalensis* of Senegalia, the Roman name for Senegal, where the type strain was isolated). Growth occurred between 15°C and 45°C on a homemade culture medium (described above), with optimal growth observed at 37°C in an aerobic atmosphere. Strain SK-1T required a salinity ranging from 5 to 200 g/L of NaCl (optimum at 100 g/L). The optimum pH for growth was 7.5 (range, 5 to 9). The strain SK-1T was strictly aerobic and also grew in the presence of 5% CO_2, but no growth was observed under anaerobic and microaerophilic conditions. The colonies of the strain SK-1T were circular, greyish, shiny and smooth, with a diameter of 2 to 6 mm. Cells stained Gram positive. They were motile by polar flagella, spore forming (2 to 6 µm in length and 0.5 µm in diameter) and generally occurred individually or in pairs. Strain SK-1T is catalase and oxidase negative. Using API 50 CH and API20 NE (bioMérieux), strain SK-1T was positive

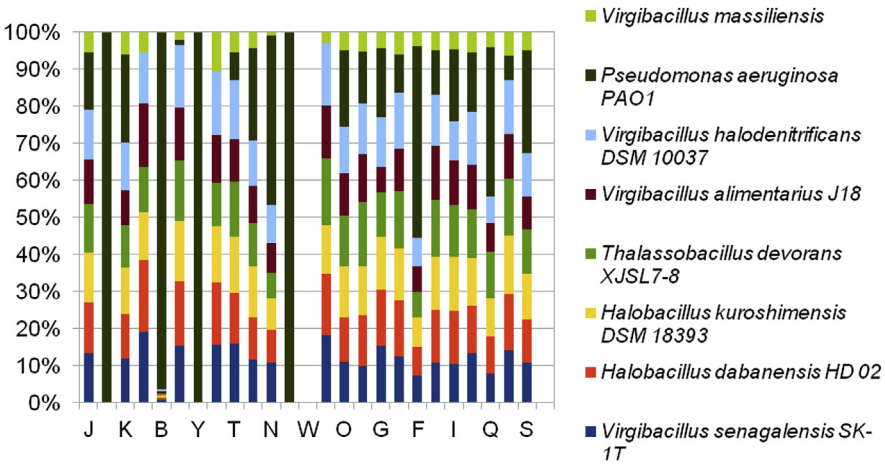

FIG. 7. Distribution of functional classes of predicted genes according to COGs proteins. COGs, Clusters of Orthologous Groups (COGs) database.

for reduction of nitrates but negative for phosphatase alkaline activity, â-galactosidase, áN-acetyl-â-glucosaminidase and urease. Strain SK-1T was negative for ribose, L-arabinose and D-lactose assimilation and positive for D-glucose, D-fructose, D-mannose, D-mannitol, D-maltose and D-sucrose. The strain SK-1T was susceptible to doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamicin, penicillin, trimethoprim/sulfamethoxazole and imipenem, but resistant to metronidazole.

The percentage of G + C content of the genome is 42.9%. The 16S rRNA and genome sequences are deposited in GenBank under accession numbers LK021111. The habitat of the microorganism is the human digestive tract. The type strain SK-1T (= CSUR P1101, = DSM 28585) was isolated from a stool specimen of a healthy Senegalese man.

Acknowledgements

The authors thank the Xegen Company (http://www.xegen.fr/) for automating the genomic annotation process. This study was funded by the Méditerranée Infection Foundation.

Conflict of Interest

None declared.

References

[1] Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, Raoult D. The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. Clin Microbiol Rev 2015;1:237–64.
[2] Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 2012;18:1185–93.
[3] Tindall BJ, Rosselló-Móra R, Busse HJ, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. Int J Syst Evol Microbiol 2010;60:249–66.
[4] Wayne LG, Brenner DJ, Colwell PR, Grimont PAD, Kandler O, Krichevsky, et al. Report of the *ad hoc* committee on reconciliation of approaches to bacterial systematic. Int J Syst Bacteriol 1987;37:463–4.
[5] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.
[6] Rosselló-Móra R. DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation. In: Stackebrandt E, editor. Molecular identification, systematics, and population structure of prokaryotes. Berlin: Springer; 2006. p. 23–50.
[7] Ramasamy D, Mishra AK, Lagier JC, Padhmanabhan R, Rossi M, Sentausa E, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of new bacterial species. Int J Syst Evol Microbiol 2014;64:384–91.
[8] Welker M, Moore ER. Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst Appl Microbiol 2011;34:2–11.
[9] Lagier JC, El Karkouri K, Nguyen TT, Armougom F, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Anaerococcus senegalensis* sp. nov. Stand Genomic Sci 2012;6:116–25.
[10] Lagier JC, Armougom F, Mishra AK, Nguyen TT, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Alistipes timonensis* sp. nov. Stand Genomic Sci 2012;6:315–24.
[11] Roux V, El Karkouri K, Lagier JC, Robert C, Raoult D. Non-contiguous finished genome sequence and description of *Kurthia massiliensis* sp. nov. Stand Genomic Sci 2012;7:221–32.
[12] Kokcha S, Ramasamy D, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Brevibacterium senegalense* sp. nov. Stand Genomic Sci 2012;7:233–45.
[13] Ramasamy D, Kokcha S, Lagier JC, Nguyen TT, Raoult D, Fournier PE. Genome sequence and description of *Aeromicrobium massiliense* sp. nov. Stand Genomic Sci 2012;7:246–57.
[14] Lagier JC, Elkarkouri K, Rivet R, Couderc C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Senegalemassilia anaerobia* gen. nov., sp. nov. Stand Genomic Sci 2013;7:343–56.
[15] Lagier JC, El Karkouri K, Mishra AK, Robert C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Enterobacter massiliensis* sp. nov. Stand Genomic Sci 2013;7:399–412.
[16] Hugon P, Mishra AK, Lagier JC, Nguyen TT, Couderc C, Raoult D, et al. Non-contiguous finished genome sequence and description of *Brevibacillus massiliensis* sp. nov. Stand Genomic Sci 2013;8:1–14.
[17] Lagier JC, Gimenez G, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Herbaspirillum massiliense* sp. nov. Stand Genomic Sci 2012;7:200–9.
[18] Heyndrickx ML, Kersters K, DeVos P, Forsyth G, Logan NA. Virgibacillus: a new genus to accommodate *Bacillus pantothenticus* (Proom and Knight 1950). Emended description of *Virgibacillus pantothenticus*. Int J Syst Bacteriol 1998;48:99–106.
[19] Heyrman J, Logan NA, Busse HJ, Balcaen A, Lebbe L, Rodriguez-Diaz M, et al. *Virgibacillus carmonensis* sp. nov., *Virgibacillus necropolis* sp. nov. and *Virgibacillus picturae* sp. nov., three novel species isolated from deteriorated mural paintings, transfer of the species of the genus *Salibacillus* to *Virgibacillus*, as *Virgibacillus marismortui* comb. nov. and *Virgibacillus salexigens* comb. nov., and emended description of the genus *Virgibacillus*. Int J Syst Evol Microbiol 2003;53:501–11.
[20] Carrasco IJ, Marquez MC, Ventosa A. *Virgibacillus salinus* sp. nov., a moderately halophilic bacterium from sediment of a saline lake. Int J Syst Evol Microbiol 2009;59:3068–73.
[21] Chen YG, Cui XL, Wang YX, Zhang YQ, Tang SK, Li WJ, et al. *Virgibacillus sediminis* sp. nov., a moderately halophilic bacterium isolated from a salt lake in China. Int J Syst Evol Microbiol 2009;59:2058–63.
[22] Jeon C, Kim J, Park DJ, Xu LH, Jiang CL, Kim CJ. *Virgibacillus xinjiangensis* sp. nov., isolated from a Salt Lake of Xin-jiang Province in China. J Microbiol 2009;47:705–9.
[23] Zhang YJ, Zhou Y, Ja M, Shi R, Chun-Yu WX, Yang LL, et al. *Virgibacillus albus* sp. nov., a novel moderately halophilic bacterium isolated from Lop Nur salt lake in Xinjiang Province, China. Antonie Van Leeuwenhoek 2012;102:553–60.
[24] Kim J, Jung MJ, Roh SW, Nam YD, Shin KS, Bae JW. *Virgibacillus alimentarius* sp. nov., isolated from a traditional Korean food. Int J Syst Evol Microbiol 2011;61:2851–5.
[25] Niederberger TD, Steven B, Charvet S, Barbier B, Whyte LG. *Virgibacillus arcticus* sp. nov., a moderately halophilic, endospore-forming bacterium from permafrost in the Canadian high Arctic. Int J Syst Evol Microbiol 2009;59:2219–25.

[26] Yoon JH, Kang SJ, Jung YT, Lee KC, Oh HW, Oh TK. *Virgibacillus byunsanensis* sp. nov., isolated from a marine solar saltern. Int J Syst Evol Microbiol 2010;60:291–5.
[27] Lee SY, Kang CH, Oh TK, Yoon JH. *Virgibacillus campisalis* sp. nov., from a marine solar saltern. Int J Syst Evol Microbiol 2012;62:347–51.
[28] An SY, Asahara M, Goto K, Kasai H, Yokota A. *Virgibacillus halophilus* sp. nov., spore-forming bacteria isolated from soil in Japan. Int J Syst Evol Microbiol 2007;57:1607–11.
[29] Yoon JH, Kang SJ, Lee SY, Lee MH, Oh TK. *Virgibacillus dokdonensis* sp. nov. isolated from a Korean island, Dokdo, located at the edge of the East Sea in Korea. Int J Syst Evol Microbiol 2005;55:1833–7.
[30] Seiler H, Wenning M. *Virgibacillus halotolerans* sp. nov., isolated from a dairy product. Int J Syst Evol Microbiol 2013;63:3358–63.
[31] Quesada T, Aguilera M, Morillo JA, Ramos-Cormenzana A, Monteoliva-Sanchez M. *Virgibacillus olivae* sp. nov., isolated from waste wash-water from processing of Spanish-style green olives. Int J Syst Evol Microbiol 2007;57:906–10.
[32] Chen YG, Cui XL, Fritze D, Chai LH, Schumann P, Wen ML, et al. *Virgibacillus kekensis* sp. nov., a moderately halophilic bacterium isolated from a salt lake in China. Int J Syst Evol Microbiol 2008;58:647–53.
[33] Tanasupawat S, Chamroensaksri N, Kudo T, Itoh T. Identification of moderately halophilic bacteria from Thai fermented fish and proposal of *Virgibacillus siamensis* sp. nov. J Gen Appl Microbiol 2010;56:369–79.
[34] Seng P, Abat C, Rolain M, Colson P, Lagier J, Gouriet F. Laboratory: impact of matrix-assisted laser desorption ionization–time of flight mass spectrometry. J Clin Microbiol 2013;51:2182–94.
[35] Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 1980;16:111–20.
[36] Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 2010;11:119.
[37] Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res 2012;40:D48–53.
[38] Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 1997;25:955–64.
[39] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007;35:3100–8.
[40] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004;340:783–95.
[41] Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 2001;305:567–80.
[42] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. Bioinformatics 2000;16:944–5.
[43] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 2009;25:119–20.
[44] Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 2004;14:1394–403.
[45] Lechner M, Findeib S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics 2011;12:124.
[46] Auch AF, Von Jan M, Klenk HP, Goker M. DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2010;2:117–34.
[47] Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 2013;14:60.