Noncontiguous finished genome sequence and description of *Paenibacillus antibioticophila* sp. nov. GD11[T], the type strain of *Paenibacillus antibioticophila*

G. Dubourg[1,2], T. Cimmino[1], S. a. Senkar[1], J.-C. Lagier[1,2], C. Robert[1], C. Flaudrops[1], P. Brouqui[3], D. Raoult[1,2,4], P.-E. Fournier[1,2] and J.-M. Rolain[1,2]

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, Aix-Marseille Université,* 2) *Pôle des Maladies Infectieuses et Tropicales Clinique et Biologique, Fédération de Bactériologie−Hygiène−Virologie, University, Hospital Centre Timone, Institut Hospitalo-Universitaire (IHU) Méditerranée Infection,* 3) *Service des Maladies Infectieuses et Tropicales. Hôpital Nord, Assistance Publique-Hôpitaux de Marseille, France and* 4) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdul Aziz University, Jeddah, Saudi Arabia*

Abstract

Paenibacillus antibioticophila strain GD11[T] sp. nov. is the type strain of a new species within the genus *Paenibacillus*. This strain, whose genome is described here, was isolated from human faeces of a 63-year-old woman with multidrug-resistant tuberculosis who was receiving numerous antibiotics at the time of stool collection. *P. antibioticophila* is a Gram-positive aerobic bacterium. We describe here the features of this bacterium, together with the complete genome sequence and annotation. The 5 562 631 bp long genome contains 5084 protein-coding and 71 RNA genes.
New Microbes and New Infections © 2015 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases.

Corresponding author: J.-M. Rolain, 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, Aix-Marseille Université, 27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France

E-mail: jean-marc.rolain@univ-amu.fr

Introduction

Paenibacillus antibioticophila strain GD11[T] (= DSM 28228 = CSUR P1358) is the type strain of *Paenibacillus antibioticophila* sp. nov. It is a Gram-positive, aerobic, indole-negative rod-shaped bacterium isolated as part of a culturomics study [1,2]. This bacterium was isolated from a 63-year-old woman with multidrug-resistant tuberculosis who was receiving a broad-spectrum antibiotic regimen at the time of stool collection, as recently reported [2].

The current classification of prokaryotes is based on a combination of phenotypic and genotypic characteristics [3,4] which includes 16S rRNA gene phylogeny, G+C content and DNA−DNA hybridization. Although these tools are considered to be the reference standard, they have several pitfalls [5,6]. As a result of the declining cost of sequencing, there has been rapid growth in the number of bacterial genomes being sequenced [7], and we thus recently proposed to add genomic information to phenotypic criteria for the description of new bacterial species [7−36].

The genus *Paenibacillus* initially included Gram-variable, facultative anaerobic endospore-forming bacteria that were reclassified as a separate genus in 1993 [37] after being grouped with the genus *Bacillus*. To date, the genus consists of 170 described species and four subspecies that have been isolated from a variety of environments such as soil, water, rhizosphere, vegetable matter, forage and insect larvae [38−40]. Isolation from human specimens has also been described [41,42]. One *Paenibacillus* species was cultured for the first time from human faeces as a part of a culturomics study [14].

Here we present a summary of the classification and set of features for *Paenibacillus antibioticophila* sp. nov. strain GD11[T], together with the description of the complete genomic sequencing and annotation. These characteristics support the creation of the *Paenibacillus antibioticophila* species.

Organism information

Classification and features

A stool sample was collected from a 63-year-old woman with a pulmonary form of multidrug-resistant tuberculosis [2]. The study was approved by the ethics committee of the Institut Fédératif de Recherche IFR48, Faculty of Medicine, Marseille,

France, under agreement 09-002. The faecal specimen was preserved at −80°C after collection. Strain GD11[T] (Table 1) was isolated in March 2012 by cultivation on 5% sheep's blood agar in aerobic conditions at 37°C after a 21-day preincubation in a blood culture bottle with sterile cow rumen fluid and sheep's blood.

Strain GD11[T] exhibited a 97.6% 16S rRNA sequence identity with *P. puldeungensis* (GenBank accession no. NR117451), the phylogenetically closest bacterial species with standing in nomenclature (Fig. 1). Its 16S rRNA sequence was deposited in GenBank under accession number KC158472. This value was lower than the 98.7% 16S rRNA gene sequence threshold recommended by Stackebrandt and Ebers [4] to delineate a new species without carrying out DNA-DNA hybridization.

Growth at different temperatures (25, 30, 37, 45 and 56°C) was tested; no growth was observed at 45°C or 56°C. Growth occurred between 25°C and 37°C, after 24 to 48 hours of incubation. Colonies were 0.5 μm in diameter on blood-enriched Columbia agar. Growth of the strain was tested in 5% sheep's blood−enriched Columbia agar (bioMérieux, Marcy l'Etoile, France) under anaerobic and microaerophilic conditions using the GENbag anaer and GENbag microaer systems, respectively (bioMérieux), and under aerobic conditions, with or without 5% CO_2. Growth was achieved only both aerobically and anaerobically. Gram staining showed Gram-positive bacilli (Fig. 2). A motility test was positive. Cells grown on agar were soft and translucent after 24 hours and had a mean width of 0.49 μm and mean length of 2.67 μm (Fig. 3).

Strain GD11[T] exhibits neither catalase nor oxidase activity. Using an API ZYM strip (bioMérieux), positive reactions were observed for esterase (C4), esterase lipase (C8), naphthol-AS-BI-phosphohydrolase, β-galactosidase, α-galactosidase and α-glucosidase. Using rapid ID32A, positive reactions were observed for α-glucosidase, α-arabinosidase, β-glucuronidase, N-acetyl-β-glucosaminidase, nitrate reduction, glutamic acid decarboxylase, fermentation of mannose and raffinose.

Using an API 50 CH strip (bioMérieux), positive reactions were recorded for esculin hydrolysis and fermentation of L-arabinose, D-ribose, D-xylose, methyl-βD-xylopranoside, D-galactose, D-glucose, D-fructose, D-mannose, L-rhamnose, D-mannitol, N-acetylglucosamine, amygdalin, arbutin, salicin, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-saccharose, D-trehalose, inulin, D-melezitose, D-raffinose, starch, glycogen and D-lyxose.

Using an API ZYM strip (bioMérieux), negative reactions were observed for acid phosphatase, alkaline phosphatase, leucine arylamidase, valine arylamidase, cystine arylamidase, lipase (C14), trypsin, α-chymotrypsin, β-glucosidase, α-mannosidase, and α-fucosidase. Using rapid API 32A, negative reactions were observed for arginine dihydrolase, urease, production of indole, leucine arylamidase, histidine arylamidase, phenylalanine arylamidase, tyrosin arylamidase, alanine arylamidase α-mannosidase, β-glucosidase, α-fucosidase. an API 50 CH strip (bioMérieux), negative reactions were recorded for fermentation of erythritol, D-arabinose, L-xylose, D-adonitol, L-sorbose, dulcitol, inositol, D-sorbitol, xylitol, D-turanose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate, potassium 5-ketogluconate and potassium-5-ketogluconate.

Cells are susceptible to penicillin G, amoxicillin, amoxicillin−clavulanic acid, ceftriaxone, imipenem, vancomycin, rifampicin, erythromycin, gentamicin, ciprofloxacin and trimethoprim−sulfamethoxazole, but resistant to metronidazole.

By comparison with *P. puldeungensis* strain CAU 9324[T], its phylogenetically closest neighbor, *P. antibioticophila* differed in alkaline phosphates, acid phosphatase, oxidase and β-glucosidase (Table 2).

Extended features descriptions

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) protein analysis was carried out as previously described [48]. Briefly, using a pipette tip, one isolated bacterial colony from a culture on agar plate was transferred and spread as a thin film on a MSP 96 MALDI-TOF target plate (Bruker Daltonics, Leipzig, Germany). Twelve distinct deposits from 20 isolated colonies were performed for strain GD11[T]. Each smear was overlaid with 2 μL of matrix solution (saturated solution of alpha-cyano-4-hydroxycinnamic acid) in 50% acetonitrile, 2.5% trifluoroacetic acid, and dried for 5 minutes. Microflex spectrometer (Bruker) was used for measurements, and spectra were then recorded in the positive linear mode for the mass range of 2000 to 20 000 Da (parameter settings: ion source 1 (ISI), 20 kV; IS2, 18.5 kV; lens, 7 kV). A spectrum was obtained after 240 shots with variable laser power. The 20 GD11[T] spectra were imported into MALDI BioTyper 3.0 software (Bruker) and analysed by standard pattern matching (with default parameter settings) against the main spectra of 7379 bacteria, including 129 spectra from 70 *Paenibacillus* species. The method of identification included the *m*/z from 3000 to 15 000 Da. A maximum of 100 peaks were compared with spectra in the database for every spectrum. The resulting score enabled the identification of tested species (or not): a score of ≥2 with a validly published species enabled identification at the species level; a score of ≥1.7 but <2 enabled identification at the genus level; and a score of <1.7 did not enable any identification. No significant MALDI-TOF score was obtained for strain GD11[T] against the Bruker

MIGS ID	Property	Term	Evidence code[a]
	Current classification	Domain: *Bacteria*	TAS [43]
		Phylum: *Firmicutes*	TAS [44]
		Class: *Bacillus*	TAS [45]
		Order: *Bacillales*	TAS [46]
		Family: *Paenibacillaceae*	TAS [45]
		Genus: *Paenibacillus*	TAS [37]
		Species: *Paenibacillus antibioticophila*	IDA
		Type strain: GD11[T]	IDA
	Gram stain	Positive	IDA
	Cell shape	Bacilli	IDA
	Motility	Motile	IDA
	Sporulation	Nonsporulating	IDA
	Temperature range	Mesophile	IDA
	Optimum temperature	37°C	IDA
MIGS-6.3	Salinity	Unknown	
MIGS-22	Oxygen requirement	Aerobic	IDA
	Carbon source	Unknown	IDA
	Energy source	Unknown	IDA
MIGS-6	Habitat	Human gut	IDA
MIGS-15	Biotic relationship	Free-living	IDA
MIGS-14	Pathogenicity	Unknown	
	Biosafety level	2	
	Isolation	Human faeces	
MIGS-4	Geographic location	Marseille, France	IDA
MIGS-5	Sample collection time	March 2012	IDA
MIGS-4.1	Latitude	43.296482	IDA
MIGS-4.1	Longitude	5.36978	IDA
MIGS-4.3	Depth	Surface	IDA
MIGS-4.4	Altitude	0 m above sea level	IDA

MIGS, minimum information about a genome sequence.
[a]Evidence codes are as follows: IDA, inferred from direct assay; TAS, traceable author statement (i.e., a direct report exists in the literature); NAS, nontraceable author statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species or anecdotal evidence). These evidence codes are from the Gene Ontology project (http://www.geneontology.org/GO.evidence.shtml) [47]. If the evidence code is IDA, then the property should have been directly observed, for the purpose of this specific publication, for a live isolate by one of the authors, or by an expert or reputable institution mentioned in the acknowledgements.

database, suggesting that our isolate was not a member of a known species. We added the spectrum from strain GD11[T] to our database (Fig. 4). Finally, the spectral differences with other members of the genus *Paenibacillus* were highlighted by gel view (Fig. 5).

Genome sequencing information

Genome project history

As part of a culturomics study isolating all bacterial species from the human digestive flora from patients with multidrug-resistant tuberculosis and treated with broad-spectrum antibiotics, this organism was isolated and selected for sequencing on the basis of its phenotypic differences, phylogenetic position and 16S rRNA sequence similarity to other members of the genus *Paenibacillus*. It is the first sequenced genome of *P. antibioticophila* sp. nov. The GenBank Bioproject number is PRJEB1962 and consists of 131 large contigs in nine scaffolds. Table 3 shows the

project information and its association with minimum information about a genome sequence (MIGS) 2.0 compliance [49].

Growth conditions and DNA isolation

Paenibacillus antibioticophila strain GD11[T] (= DSM 28228 = CSUR P1358) was cultured aerobically on Columbia agar (bioMérieux). A total of 200 μL of bacterial suspension was diluted in 1 mL Tris-EDTA (TE) buffer for lysis treatment. After a lysozyme incubation of 30 minutes at 37°C, the lysis was performed with laurylsarcosylby 1% final and RNAseA treatment at 50 μG/μL final concentration during 1 hour at 37°C, followed by an overnight proteinase K incubation at 37°C. The DNA was purified three times by phenol−chloroform extractions, and ethanolic precipitation was performed at −20°C overnight. After centrifugation, the DNA was resuspended in 150 μL TE buffer. The concentration was measured by the Quant-it Picogreen kit (Invitrogen; Life Technologies, Carlsbad, CA, USA) on the Genios_Tecan fluorometer at 48 ng/μL.
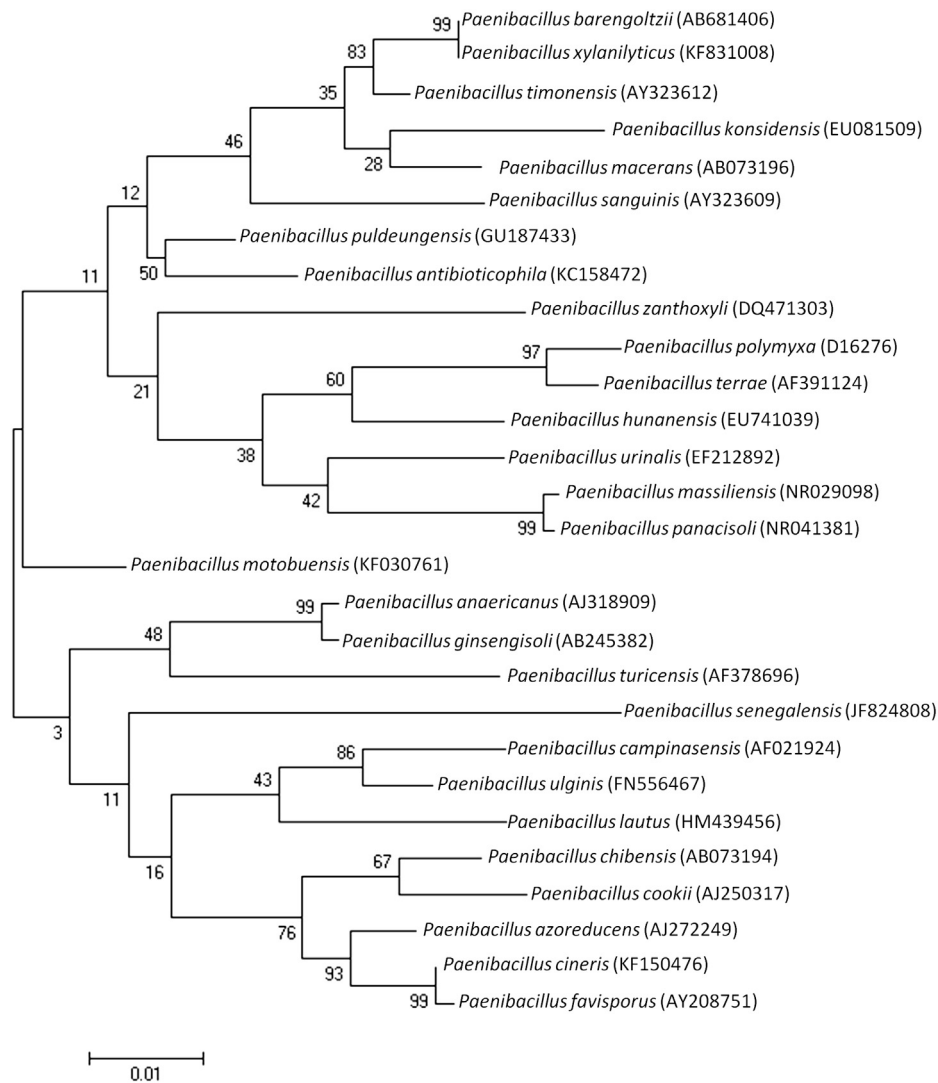
Genome sequencing and assembly

Genomic DNA of *Paenibacillus antibioticophila* strain GD11[T] was sequenced on 454_Roche_Titanium (Roche, Basel, Switzerland).

A paired end library was pyrosequenced on the 454_Roche_Titanium. This project was loaded twice on a $^1/_4$ region for the 5 kb insert libraries on PTP PicoTiterPlates. The library was constructed with 2.5 μg of DNA according to the 454_Roche_Titanium paired end protocol and manufacturer. It was mechanically fragmented on the Covaris device (KBio-Science−LGC Genomics, Teddington, UK) through miniTUBE-Red 5 kb. The DNA fragmentation was visualized through the Agilent 2100 BioAnalyzer on a DNA labchip 7500, with an optimal size of 4.7 kb. The library was constructed according to the 454_Roche_Titanium paired end protocol and manufacturer. Circularization and nebulization were performed and generated a pattern with optimal at 490 bp. After PCR amplification through 17 cycles followed by double size selection, the library was then quantified on the Quant-it Ribogreen kit (Invitrogen) on the Genios_Tecan fluorometer at 759 pg/μL. The library concentration equivalence was calculated as 2.84×10^9 molecules/μL. The library was stocked at −20°C until use.

The 5 kb paired end library was clonal amplified with 0.25 and 0.5 cpb in 4 emPCR reactions per condition, with the GS Titanium SV emPCR Kit (Lib-L) v2. The yield of the emPCR was 5.21% and 5.69%, respectively, according to the quality expected by the range of 5 to 20% from the Roche procedure.

Twice, 790 000 beads were loaded on the GS Titanium PicoTiterPlates PTP Kit 70 × 75 and sequenced with the GS Titanium Sequencing Kit XLR70.
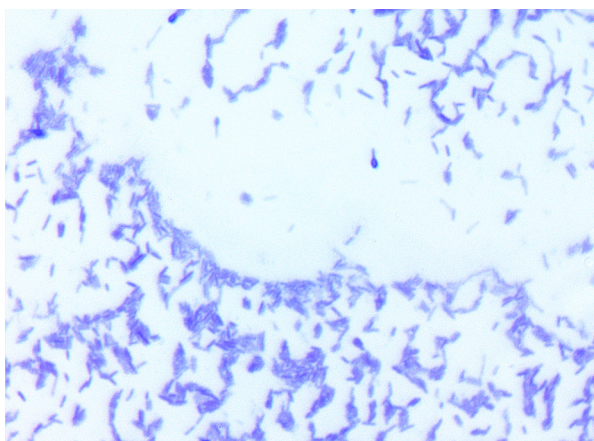
FIG. 1. Phylogenetic tree showing position of *Paenibacillus antibioticophila* strain GD11[T] relative to other type strains within *Paenibacillaceae*. The strains and their corresponding GenBank accession numbers for 16S rRNA genes are (type = [T]): *P. sanguinis* strain 2301083[T], AY323612; *P. macerans* strain ATCC 8244[T], AB073196; *P. timonensis* strain 2301032[T], AY323612; *P. barengoltzii* strain SAFN-016[T], AY167814; *P. antibioticophila* strain GD11[T], KC158472; *P. puldeungensis* strain CAU 9324[T], GU187433; *P. motobuensis* strain MC10[T], AY741810; *P. senegalensis* strain JC66[T], JF824808; *P. zanthoxyli* strain JH29[T], DQ471303; *P. sabinae* strain T27[T], DQ338444; *P. lautus* strain ATCC 43898[T], AB073188; *P. terrae* strain AM141[T], AF391124; *P. polymyxa* strain ATCC 842[T], D16276; *P. xylanilyticus* strain XIL14[T], AY427832; *P. massiliensis* strain 2301065[T], AY323608; *P. panacisoli* strain Gsoil 1411[T], AB245384; *Bacillus subtilis* ATCC 6051[T], AJ276351. Sequences were aligned using CLUSTALW, and phylogenetic inferences were obtained using the maximum likelihood method within MEGA6. Numbers at nodes are percentages of bootstrap values obtained by repeating analysis 1000 times to generate majority consensus tree. *Bacillus subtilis* ATCC 6051[T] (AJ276351) was used as outgroup. Scale bar = 1% nucleotide sequence divergence.

The runs were performed overnight and then analysed through the gsRunBrowser and Newbler assembler _Roche. The global 411 382 passed filter sequences generated 150.40 Mb with a length average of 354 bp. These sequences were assembled on the gsAssembler from Roche with 90% identity and 40 bp as overlap. It led to nine scaffolds and 131 large contigs (>1500 bp), generating a genome size of 5.6 Mb, which corresponds to a coverage of 19.96 genome equivalent.

Genome annotation

Open reading frames (ORFs) were predicted using Prodigal [50] with default parameters, but the predicted ORFs were excluded if they were spanning a sequencing gap region. The predicted bacterial protein sequences were searched against the GenBank database [51] and the Clusters of Orthologous Groups (COGs) database using BLASTP. The tRNAScanSE tool [52] was used to find tRNA genes, whereas ribosomal RNAs were found by using

FIG. 2. Gram staining of *Paenibacillus antibioticophila* strain GD11[T] (= DSM 28228 = CSUR P1358).

RNAmmer [53] and BLASTn against the GenBank database. Lipoprotein signal peptides and the number of transmembrane helices were predicted using SignalP [54] and TMHMM [55], 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. Here, we compared the genome sequence of *P. antibioticophila* strain GD11[T] with those of *Paenibacillus barengoltzii* strain G22 (GenBank accession no. ASSZ00000000.1), *Paenibacillus massiliensis* strain 2301065[T]

FIG. 3. Transmission electron microscopy of *P. antibioticophila* strain GD11[T] with Morgani 268D device (Philips, Amsterdam, The Netherlands) at operating voltage of 60 kV. Scale bar = 200 nm.

(GenBank accession no. ARIL00000000.1), *Paenibacillus panacisoli* strain DSM 21345 (GenBank accession no. AUFO00000000.1), *Paenibacillus polymyxa* strain ATCC 842[T] (GenBank accession no. AFOX00000000.1), *Paenibacillus sanguinis* strain 2301083[T] (GenBank accession no. ARGO00000000.1), *Paenibacillus senegalensis* strain JC66[T] (GenBank accession no. CAES00000000.1), *Paenibacillus terrae* strain HP L-003 (GenBank accession no. CP003107.1) and *Paenibacillus zanthoxyli* strain JH29[T] (GenBank accession no. ASSD00000000.1), which were identified using the Proteinortho software (version 1.4) [56] using a 30% protein identity and 1e-05 *E* value. The average percentage of nucleotide sequence identity (Table 4) between corresponding orthologous sets was determined using the Needleman-Wunsch algorithm global alignment technique. Artemis [57] was used for data management, and DNA Plotter [58] was used for visualization of genomic features. The Mauve alignment tool was used for multiple genomic sequence alignment and visualization [59]. PHAST (PHAge search Tool) was employed to identify phage sequences [60].

Genome properties

The genome of *P. antibioticophila* strain GD11[T] is 5 562 631 bp long with a 49.1% G+C content (Fig. 6). Of the 5155 predicted genes, 5084 were protein-coding genes and three were RNAs. Three rRNA genes (one 16S rRNA, one 23S rRNA and one 5S rRNA) and 68 predicted tRNA genes were identified in the genome. A total of 3814 genes (73.98%) were assigned a putative function. One hundred forty-three genes (2.77%) were identified as ORFans. The remaining genes were annotated as hypothetical proteins. The properties and the statistics of the genome are summarized in Tables 5 and 6. The distribution of genes into COGs functional categories is presented in Table 5.

Insights from genome sequence

Extended insights

The genome of *P. antibioticophila* strain GD11[T] was compared to each of *Paenibacillus barengoltzii* strain G22 (GenBank accession no. ASSZ00000000.1), *Paenibacillus massiliensis* strain 2301065[T] (GenBank accession no. ARIL00000000.1), *Paenibacillus panacisoli* strain DSM 21345 (GenBank accession no. AUFO00000000.1), *Paenibacillus polymyxa* strain ATCC 842[T] (GenBank accession no. AFOX00000000.1), *Paenibacillus sanguinis* strain 2301083[T] (GenBank accession no. ARGO00000000.1), *Paenibacillus senegalensis* strain JC66[T] (GenBank accession no. CAES00000000.1), *Paenibacillus terrae* strain HP L-003 (GenBank accession no. CP003107.1) and
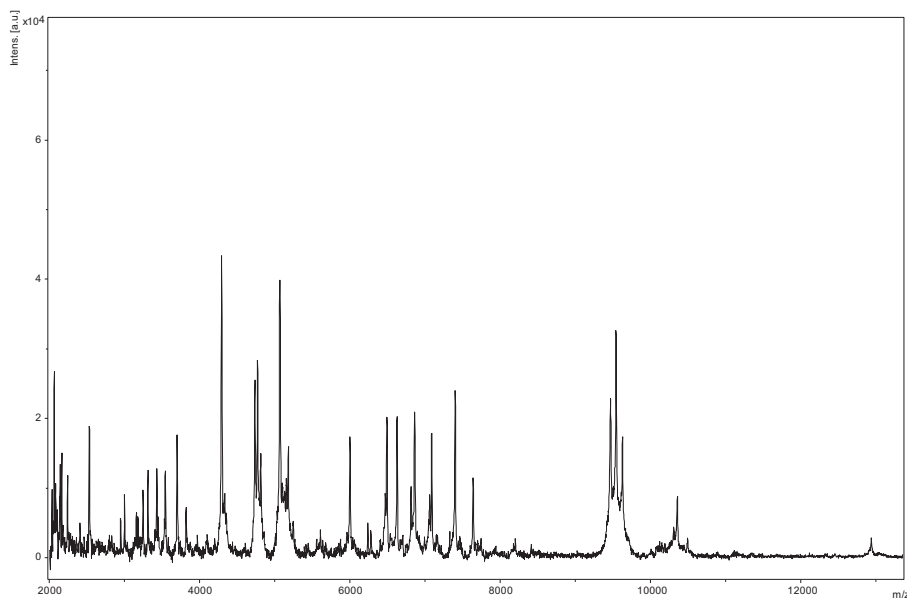
TABLE 2. Differential characteristics of *Paenibacillus antibioticophila* strain GD11[T] (data from this study) with *P. sanguinis* strain 2301083[T], *P. zanthoxyli* strain JH29[T], *P. puldeungensis* strain CAU 9324[T], *P. terrae* strain AM141[T]

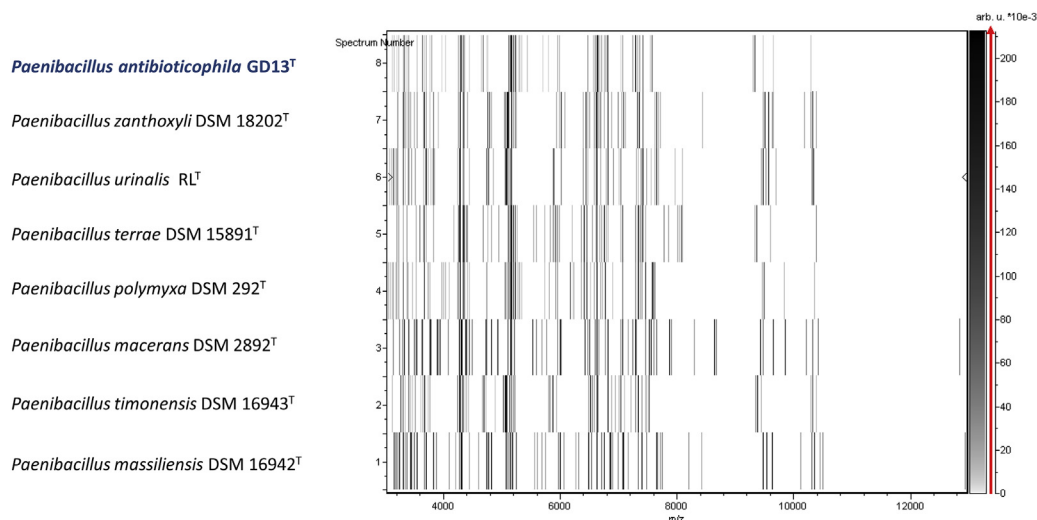| Property	*P. antibioticophila*	*P. sanguinis*	*P. zanthoxyli*	*P. puldeungensis*	*P. terrae*
Cell diameter (μm)	0.49 × 2.67	0.5 × 2–3	0.4 × 4–4.8	0.3 × 1.3–2.3	1.5 × 4–7
Oxygen requirement	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic
Gram stain	+	+	+	+	+
Salt requirement	No	No	No	No	
Motility	+	+	+	−	−
Endospore formation	NA	+	+	+	+
Production of:					
Alkaline phosphatase	−	+	NA	+	−
Acid phosphatase	−	+	NA	+	−
Catalase	−	−	+	−	+
Oxidase	−	−	−	+	+
Nitrate reductase	+	+	NA	+	+
Urease	−	−	NA	NA	NA
α-Galactosidase	+	NA	NA	+	NA
β-Galactosidase	+	NA	NA	+	NA
β-Glucuronidase	+	NA	NA	+	NA
α-Glucosidase	+	NA	NA	+	NA
β-Glucosidase	−	NA	NA	+	NA
Esterase	+	NA	NA	+	NA
Esterase lipase	+	NA	NA	−	NA
Naphthol-AS-BI-phosphohydrolase	+	NA	NA	NA	NA
N-acetyl-β-glucosaminidase	−	NA	NA	−	NA
Pyrazinamidase	NA	NA	NA	NA	NA
α-Mannosidase	−	NA	NA	−	NA
α-Fucosidase	−	NA	NA	−	NA
Leucine arylamidase	+	NA	NA	+	NA
Valine arylamidase	−	NA	NA	+	NA
Cystine arylamidase	−	NA	NA	+	NA
α-Chymotrypsin	−	NA	NA	+	NA
Trypsin	−	NA	NA	+	NA
Utilization of:					
5-Keto-gluconate	−	NA	NA	NA	NA
D-Xylose	−	−	−	NA	+
D-Fructose	−	+	−	NA	+
D-Glucose	−	−	−	NA	−
D-Mannose	−	−	NA	NA	+
Habitat	Human gut	Human	Soil	Soil	Algae

+, positive result; −, negative result; NA, data not available.

Paenibacillus zanthoxyli strain JH29[T] (GenBank accession no. ASSD00000000.1). The draft genome of *P. antibioticophila* has a larger size than that of *P. barengoltzii, P. sanguinis* and *P. zanthoxyli* (5.56, 4.75, 4.8 and 5.05 Mb, respectively). The G+C content of *P. antibioticophila* is higher than those of *P. massiliensis, P. panacisoli, P. polymyxa, P. senegalensis* and *P. terrae* (49.1, 48.5, 48.3, 44.9, 48.2 and 46.8%, respectively) but less than that of *P. barengoltzii, P. sanguinis* and *P. zanthoxyli*

FIG. 4. Reference mass spectrum from *Paenibacillus antibioticophila* strain GD11[T] (= DSM 28228 = CSUR P1358). Spectra from 16 individual colonies were compared and reference spectrum generated.

FIG. 5. Gel view comparing *Paenibacillus antibioticophila* strain GD11[T]. Gel view displays raw spectra of loaded spectrum files arranged in pseudo gel-like look. The *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 indicating relation between color peak are displayed, with peak intensity in arbitrary units. Displayed species are indicated at left, with GD11[T] highlighted in blue.

(51.9, 49.3 and 50.9%, respectively). The gene content of *P. antibioticophila* is larger than those of *P. barengoltzii*, *P. sanguinis*, *P. senegalensis* and *P. zanthoxyli* (5155, 4394, 4209, 4422 and 4878, respectively). However, the distribution of genes into COGs categories was similar in all nine compared genomes (Fig. 7). In addition, *P. antibioticophila* shared 5084, 4307, 5055, 5059, 5068, 4093, 4278, 5525 and 4676 ortholo-gous genes with *P. barengoltzii, P. massiliensis, P. panacisoli, P. polymyxa, P. sanguinis, P. senegalensis, P. terrae* and *P. zanthoxyli*, respectively (Table 6). The average nucleotide sequence iden-tity ranged from 96.47 to 66.74% between the species. Finally, no sequences coding for nonribosomal peptide synthetases or polyketide synthases were found within the *P. antibioticophila* genome. The analysis of virome revealed the presence of one

intact phage of 20.8 kb, and other three incomplete phages, of, respectively, 17, 15.4 and 14.3 kb with 43.96, 50.60 and 49.84% G+C content (Table 7).

Conclusions

On the basis of phenotypic, phylogenetic and genomic analyses, we formally propose the creation of *Paenibacillus antibioticophila* sp. nov. The strain has been isolated from the stool sample of a 63-year-old woman with multidrug-resistant tuberculosis in Marseille, France. Several other previously undescribed bacte-rial species were also cultivated from different faecal samples through diversification of culture conditions [1,8–36], thus suggesting that the human faecal flora of humans remains partially unknown.
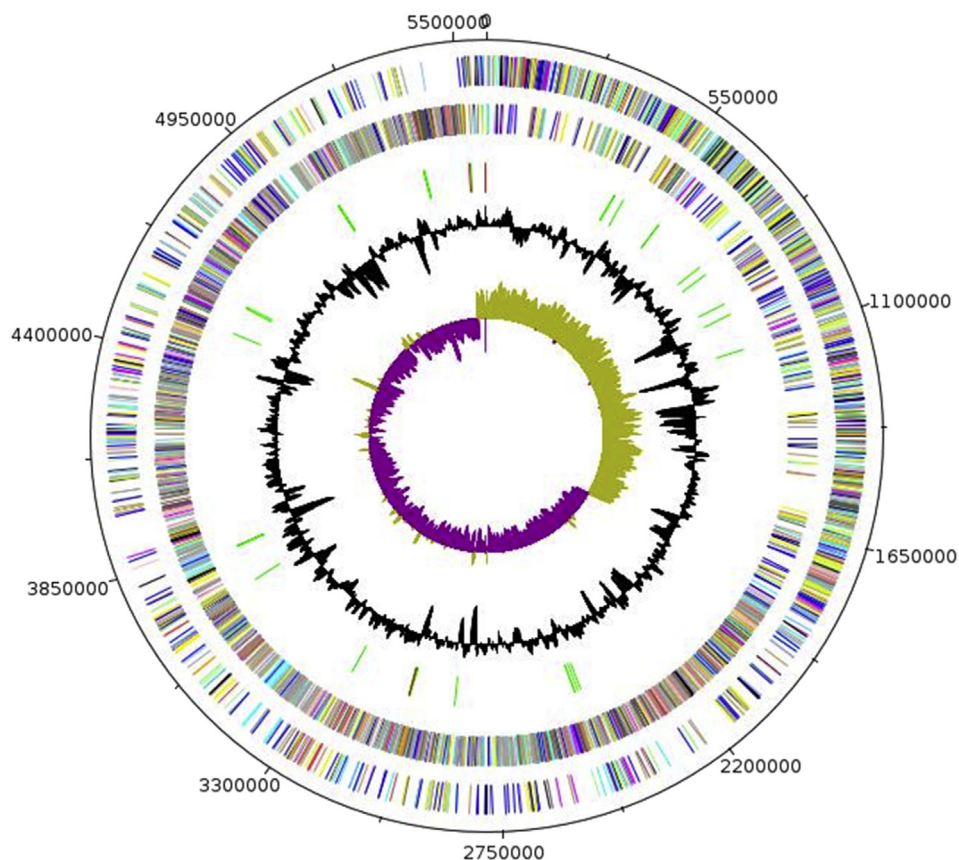
TABLE 3. Project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	High-quality draft
MIGS-28	Libraries used	454 paired-end 5 kb library
MIGS-29	Sequencing platform	454 Roche Titanium
MIGS-31.2	Fold coverage	19.96×
MIGS-30	Assemblers	Newbler
MIGS-32	Gene calling method	Prodigal
	GenBank date of release (04-06-2013)	
	NCBI project ID (PRJEB1771)	
	EMBL accession (CBLK000000000)	
MIGS-13	Project relevance	Study of human gut microbiome

EMBL, European Molecular Biology Laboratory; MIGS, minimum information about a genome sequence; NCBI, National Center for Biotechnology Information.

TABLE 4. Nucleotide content and gene count levels of genome

Attribute	Value	% of total[a]
Size (bp)	5 562 631	100
G+C content (bp)	2 731 251	49.1
Coding region (bp)	4 803 279	86.34
Total genes	5155	100
RNA genes	71	1.37
Protein-coding genes	5084	98.62
Genes assigned to COGs	3814	73.98
Genes with peptide signals	415	8.05
Genes with transmembrane helices	1508	29.25

COGs, Clusters of Orthologous Groups database.
[a]Total is based on either size of genome (in base pairs) or total number of protein-coding genes in annotated genome.

FIG. 6. Graphical circular map of chromosome. From outside in, outer two circles show ORFs oriented clockwise (colored by COGs categories) and anticlockwise (colored by COGs categories) direction, respectively. Third circle marks rRNA gene operon (red) and tRNA genes (green). Fourth circle shows GC% content plot. Innermost circle shows GC skew, with purple indicating negative values and olive positive values.

TABLE 5. Number of genes associated with 25 general COGs functional categories

Code	Value	% of total[a]	Description
J	184	3.61	Translation
A	0	0	RNA processing and modification
K	486	9.55	Transcription
L	154	3.02	Replication, recombination and repair
B	1	0.01	Chromatin structure and dynamics
D	35	0.68	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	132	2.59	Defense mechanisms
T	314	6.17	Signal transduction mechanisms
M	179	3.52	Cell wall/membrane biogenesis
N	70	1.37	Cell motility
Z	2	0.03	Cytoskeleton
W	0	0	Extracellular structures
U	50	0.98	Intracellular trafficking and secretion
O	110	2.16	Posttranslational modification, protein turnover, chaperones
C	148	2.91	Energy production and conversion
G	637	12.52	Carbohydrate transport and metabolism
E	296	5.82	Amino acid transport and metabolism
F	89	1.75	Nucleotide transport and metabolism
H	122	2.39	Coenzyme transport and metabolism
I	88	1.73	Lipid transport and metabolism
P	266	5.23	Inorganic ion transport and metabolism
Q	77	1.51	Secondary metabolites biosynthesis, transport and catabolism
R	576	11.32	General function prediction only
S	323	6.35	Function unknown
—	1,27	24.98	Not in COGs

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

Description of *Paenibacillus antibioticophila* strain GD11[T] sp. nov.

Paenibacillus antibioticophila strain GD11[T] (= DSM 28228 = CSUR P1358) is the type strain of the genus *Paenibacillus*. It was isolated from the stool samples of a 63-year-old woman with a
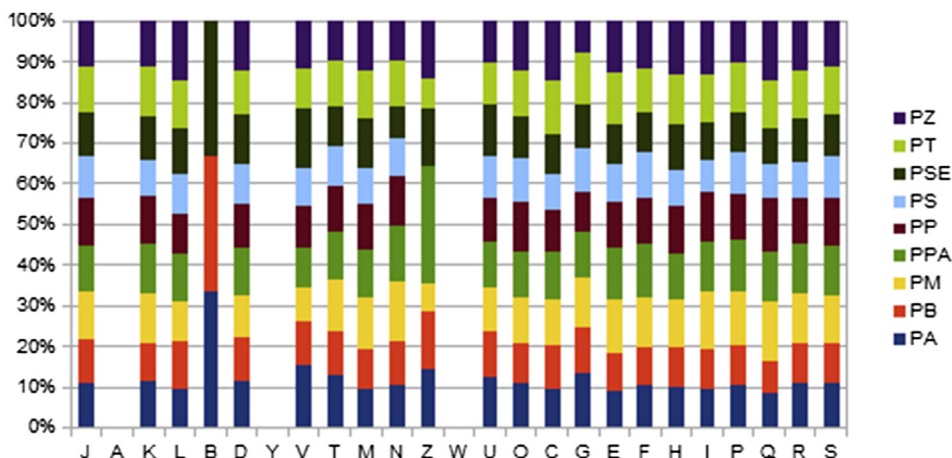
TABLE 6. Numbers of orthologous proteins shared between genomes

	PN	PB	PM	PPA	PP	PS	PSE	PT	PZ
PN	**5084**	2553	2421	2381	2242	2562	1883	2318	2078
PB	72.85	**4307**	2357	2304	2249	2456	1833	2292	2063
PM	69.61	69.12	**5055**	3881	3084	2339	1900	3142	2283
PPA	69.60	69.14	96.47	**5059**	3038	2302	1894	3104	2267
PP	69.11	69.12	71.29	71.23	**5068**	2167	1798	3279	2210
PS	73.96	75.72	69.47	69.42	69.29	**4093**	1802	2228	1948
PSE	67.11	67.15	66.78	66.75	66.74	67.17	**4278**	1797	1651
PT	69.62	69.55	71.63	71.53	86.20	69.68	66.99	**5525**	2240
PZ	69.97	70.46	69.28	69.26	69.32	69.99	67.16	69.85	**4676**

Average percentage similarity of nucleotides corresponding to orthologous protein shared between genomes (below diagonal) and numbers of proteins per genome (bold).
PA, *Paenibacillus antibioticophila*; PB, *P. barengoltzii*; PM, *P. massiliensis*; PP, *P. polymyxa*; PPA, *P. panacisoli*; PS, *P. sanguinis*; PSE, *P. senegalensis*; PT, *P. terrae*; PZ, *P. zanthoxyli*.

FIG. 7. Distribution of functional classes of predicted genes of *Paenibacillus* chromosomes according to clusters of orthologous groups of proteins. PA, *P. antibioticophila*; PB, *P. barengoltzii*; PM, *P. massiliensis*; PP, *P. polymyxa*; PPA, *P. panacisoli*; PS, *P. sanguinis*; PSE, *P. senegalensis*; PT, *P. terrae*; PZ, *P. zanthoxyli*.

pulmonary form of multidrug-resistant tuberculosis hospitalized in an infectious diseases ward in Marseille, France. The main scope of the culturomics study is to cultivate all the species within the human faeces. *P. antibioticophila* is a Gram-positive bacilli that does exhibits neither catalase nor oxidase activity. Colonies were 0.5 mm in diameter, and cells have a mean width of 0.49 μm and a mean length of 2.67 μm. Esterase (C4), esterase lipase (C8), naphthol-AS-BI-phosphohydrolase, β-galactosidase, α-galactosidase and α-glucosidase, α-arabinosidase, β-glucuronidase, N-acetyl-β-glucosaminidase, nitrate reduction, glutamic acid decarboxylase, fermentation L-arabinose, D-ribose, D-xylose, methyl-βD-xylopranoside, D-galactose, D-glucose, D-fructose, D-mannose, L-rhamnose, D-mannitol, N-acetylglucosamine, amygdalin, arbutin, salicin, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-saccharose, D-trehalose, inulin, D-melezitose, D-raffinose, starch, glycogen and D-lyxose were positive. Acid phosphatase, alkaline phosphatase, leucine arylamidase, valine arylamidase, cystine arylamidase, lipase (C14), trypsin, α-chymotrypsin, β-glucosidase, α-mannosidase, α-fucosidase, arginine dihydrolase, urease, production of indole, leucine arylamidase, histidine arylamidase, phenylalanine arylamidase, tyrosin arylamidase, alanine arylamidase α-mannosidase, fermentation erythritol, D-arabinose, L-xylose, D-adonitol, L-

sorbose, dulcitol, inositol, D-sorbitol, xylitol, D-turanose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate, potassium 5-ketogluconate and potassium-5-ketogluconate were negative. Cells are susceptible to penicillin G, amoxicillin, amoxicillin–clavulanic acid, ceftriaxone, imipenem, vancomycin, rifampicin, erythromycin, gentamicin, ciprofloxacin and trimethoprim–sulfamethoxazole and resistant to metronidazole.

The G+C content of the genome is 49.1%. The 16S rRNA gene sequence and whole-genome shotgun sequence of *P. antibioticophila* strain GD11^T is deposited in GenBank under accession number KC158472. The type strain is GD11^T (= DSM 28228 = CSUR P1358).

Acknowledgements

Conflict of Interest

None declared.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.nmni.2015.10.006.

TABLE 7. **Characteristics associated to phages in *Paenibacillus antibioticophila* strain GD11^T**

Region length (bp)	Completeness	GC%
20.8	Intact	48.31
17	Incomplete	43.96
15.4	Incomplete	50.60
14.3	Incomplete	49.84

References

[1] 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.

[2] Dubourg G, Lagier JC, Armougom F, Robert C, Hamad I, Brouqui P, et al. The gut microbiota of a patient with resistant tuberculosis is more comprehensively studied by culturomics than by metagenomics. Eur J Clin Microbiol Infect Dis 2013;32:637–45.

[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] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.

[5] Wayne LG, Brenner DJ, Colwell PR, Grimont PAD, Kandler O, Krichevsky MI, et al. Report of the ad hoc committee on reconciliation of approaches to bacterial systematic. Int J Syst Bacteriol 1987;37: 463–4.

[6] Rosselló-Móra R. DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation. In: Molecular identification, systematics, and population structure of prokaryotes. Heidelberg: Springer Berlin; 2006. p. 23–250.

[7] Genomes OnLine Database (GOLD). https://gold.jgi.doe.gov/index.

[8] Lagier JC, El Karkouri K, Nguyen TT, Armougom F, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Bacillus timonensis* sp. nov. Stand Genomic Sci 2012;6:346–55.

[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] Mishra AK, Gimenez G, Lagier JC, Robert C, Raoult D, Fournier PE. Genome sequence and description of *Alistipes senegalensis* sp. nov. Stand Genomic Sci 2012;6:1–16.

[11] 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.

[12] Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Clostridium senegalense* sp. nov. Stand Genomic Sci 2012;6:386–95.

[13] Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Peptoniphilus timonensis* sp. nov. Stand Genomic Sci 2012;7:1–11.

[14] Mishra AK, Lagier JC, Rivet R, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Paenibacillus senegalensis* sp. nov. Stand Genomic Sci 2012;7:70–81.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] Lagier JC, Ramasamy D, Rivet R, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Cellulomonas massiliensis* sp. nov. Stand Genomic Sci 2012;7:258–70.

[20] 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.

[21] Mishra AK, Hugon P, Lagier JC, Nguyen TT, Robert C, Couderc C, et al. Non contiguous-finished genome sequence and description of *Peptoniphilus obesi* sp. nov. Stand Genomic Sci 2013;7:357–69.

[22] Mishra AK, Lagier JC, Nguyen TT, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Peptoniphilus senegalensis* sp. nov. Stand Genomic Sci 2013;7:370–81.

[23] 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.

[24] Hugon P, Ramasamy D, Lagier JC, Rivet R, Couderc C, Raoult D, et al. Non contiguous-finished genome sequence and description of *Alistipes obesi* sp. nov. Stand Genomic Sci 2013;7:427–39.

[25] Mishra AK, Hugon P, Robert C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Peptoniphilus grossensis* sp. nov. Stand Genomic Sci 2012;7:320–30.

[26] Mishra AK, Hugon P, Lagier JC, Nguyen TT, Couderc C, Raoult D, et al. Non contiguous-finished genome sequence and description of *Enorma massiliensis* gen. nov., sp. nov., a new member of the family *Coriobacteriaceae*. Stand Genomic Sci 2013;8:290–305.

[27] Ramasamy D, Lagier JC, Gorlas A, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Bacillus massiliosenegalensis* sp. nov. Stand Genomic Sci 2013;8:264–78.

[28] Ramasamy D, Lagier JC, Nguyen TT, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Dielma fastidiosa* gen. nov., sp. nov., a new member of the family *Erysipelotrichaceae*. Stand Genomic Sci 2013;8:336–51.

[29] Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Genome sequence and description of *Timonella senegalensis* gen. nov., sp. nov., a new member of the suborder *Micrococcinae*. Stand Genomic Sci 2013;8:318–35.

[30] Roux V, Lagier JC, Gorlas A, Robert C, Raoult D. Non-contiguous finished genome sequence and description of *Kurthia senegalensis* sp. nov. Stand Genomic Sci 2014;9(3).

[31] Padmanabhan R, Dubourg G, Lagier JC, Nguyen TT, Couderc C, Rossi-Tamisier M, et al. Non-contiguous finished genome sequence and description of *Collinsella massiliensis* sp. nov. Stand Genomic Sci 2014;9(3).

[32] Ramasamy D, Dubourg G, Robert C, Caputo A, Papazian L, Raoult D, et al. Non contiguous-finished genome sequence and description of *Enorma timonensis* sp. nov. Stand Genomic Sci 2014;9(3).

[33] Edouard S, Bibi F, Ramasamy D, Lagier JC, Azhar EI, Robert C, et al. Non-contiguous finished genome sequence and description of *Corynebacterium jeddahense* sp. nov. Stand Genomic Sci 2014;9(3).

[34] Pagnier I, Croce O, Robert C, Raoult D, La Scola B. Non-contiguous finished genome sequence and description of *Anaerococcus provencensis* sp. nov. Stand Genomic Sci 2014;9(3).

[35] Pagnier I, Croce O, Robert C, Raoult D, La Scola B. Non-contiguous finished genome sequence and description of *Fenollaria massiliensis* gen. nov., sp. nov., a new genus of anaerobic bacterium. Stand Genomic Sci 2014;9(3).

[36] Pfeiderer A, Robert C, Caputo A, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Alistipes massilioanorexius* sp. nov. Stand Genomic Sci 2014;9(3).

[37] Ash C, Priest FG, Collins MD. Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test. Proposal for the creation of a new genus *Paenibacillus*. Antonie Van Leeuwenhoek 1993;64:253–60.

[38] Lal S, Tabacchioni S. Ecology and biotechnological potential of *Paenibacillus polymyxa*: a minireview. Indian J Microbiol 2009;49:2–10.

[39] McSpadden Gardener BB. Ecology of *Bacillus* and *Paenibacillus* spp. in agricultural systems. Phytopathology 2004;94:1252–8.

[40] Montes MJ, Mercade E, Bozal N, Guinea J. *Paenibacillus antarcticus* sp. nov., a novel psychrotolerant organism from the Antarctic environment. Int J Syst Evol Microbiol 2004;54:1521–6.

[41] Ouyang J, Pei Z, Lutwick L, Dalal S, Yang L, Cassai N, et al. Case report: *Paenibacillus thiaminolyticus*: a new cause of human infection, inducing bacteremia in a patient on hemodialysis. Ann Clin Lab Sci 2008;38:393−400.

[42] Glaeser SP, Falsen E, Busse HJ, Kämpfer P. *Paenibacillus* vulneris sp. nov., isolated from a necrotic wound. Int J Syst Bacteriol 2013;63(pt 2): 777−82.

[43] Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archae, Bacteria*, and *Eukarya*. Proc Natl Acad Sci U S A 1990;87:4576−9.

[44] Skerman VBD, Sneath PHA. Approved list of bacterial names. Int J Syst Bact 1980;30:225−420.

[45] Garrity GM, Holt J. Taxonomic outline of the *Archae* and *Bacteria*. In: Garrity GM, Boone DR, Castenholz RW, editors. Bergey's manual of systematic bacteriology. New York: Springer Verlag; 2001. p. 155−66.

[46] Prevot AR. Dictionnaire des bactéries pathogènes. In: Hauduroy P, Ehringer G, Guillot G, Magrou J, Prevot AR, Rosset A, et al., editors. Paris: Masson; 1953. p. 1−692.

[47] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25−9.

[48] Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, et al. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Infect Dis 2009;49:543−51.

[49] Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, et al. The minimum information about a genome sequence (MIGS) specification. Nat Biotechnol 2008;26:541−7.

[50] Prodigal. http://prodigal.ornl.gov/.

[51] Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res 2012;40:D48−53.

[52] 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.

[53] 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.

[54] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004;340:783−95.

[55] TMHMM. http://www.cbs.dtu.dk/services/TMHMM/.

[56] 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.

[57] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. Bioinformatics 2000;16:944−5.

[58] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 2009;25:119−20.

[59] Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 2004;14:1394−403.

[60] Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: a fast phage search tool. Nucleic Acids Res 2011;39(Web server issue): W347−52.