Description of *Gracilibacillus phocaeensis* sp. nov., a new halophilic bacterium isolated from a Senegalian human stool.

Issa Isaac NGOM[1,2], Issam HASNI[1,2,4], Bruno SENGHOR[1,2], Cheikh Ibrahima LO[2,3], Nicholas ARMSTRONG[1,2], Cheikh SOKHNA[2,3], Didier RAOULT[1,2], Pierre-Edouard FOURNIER[2,3], Jean-Christophe LAGIER[1,2*].

[1] Aix Marseille Univ, IRD, AP-HM, MEPHI, Marseille, France

[2] IHU-Méditerranée Infection, Marseille, France

[3] Aix Marseille Univ, IRD, AP-HM, SSA, VITROME, Marseille, France

[4] Amoeba R&D Department, Chassieu, France.

* Corresponding author : Jean-Christophe LAGIER, Institut Hospitalo-Universitaire Méditerranée-Infection, 19-21 Boulevard Jean Moulin, 13385, Marseille cedex 05, France ; tel +33 413 732 401 ; fax : +33 413 732 402 ; e-mail: jean-christophe.lagier@ap-hm.fr

Running title: *Gracilibacillus phocaeensis* sp. nov.

Keywords: *Gracilibacillus phocaeensis* sp. nov. ; genome; taxonogenomics; culturomics; gut

Abbreviations :

CSUR: Collection de Souches de l'Unité des Rickettsies

MALDI-TOF: Matrix-Assisted Laser-Desorption/Ionization Time-Of-Flight

FAME: Fatty Acid Methyl Ester

GC/MS: Gaz Chromatography/Mass Spectrometry

ASP: Antibiotic susceptibility profile

Abstract

Using Taxonogenomics method, we described here *Gracilibacillus phocaeensis* strain Marseille-P380, a new species previously isolated from a salty stool of a 20-year-old man from N'Diop. It is Gram-positive, aerobic and motile bacilli. The major fatty acids are $C_{15:0\ anteiso}$ (59 %), $C_{16:0}$ (16 %) and $C_{17:0\ anteiso}$ (11 %). Strain Marseille-P3801 exhibits a 98.45% sequence similarity with *Gracilibacillus thailandensis* TP2-8, the phylogenetically closest species. Its genome is 4.66 Mb with 39.6 mol% G+C content.

Introduction

The genus *Gracilibacillus* has been described as moderately halophilic, motile endospore-forming bacteria [1]. Moderately halophilic bacteria have been found in a variety of fermented foods and one of the most important food preservation methods in history has been the use of salt (NaCl) [2]. Salt is the main source of sodium in our diet. Some gut bacteria, such as lactobacilli, are very sensitive to salt [3]. However, it has been demonstrated to favor the emergence and growth of others, mainly halophilic bacteria, including *Gracilibacillus* [4]. Halophilic and halotolerant bacteria are most commonly isolated from the human gut microbiota [5,6]. The culturomics approach, based on the multiplication of culture conditions (variation of media, temperature, and atmosphere) with a more rapid bacterial identification by Matrix-Assisted Laser-Desorption/Ionization Time-Of-Flight Mass Spectrometry (MALDI-TOF MS) [7] was used to explore the human gut halophilic microbiota. Thanks to culturomics, we used high salt-containing culture media, which allowed us to isolate a new moderately halophilic bacterial strain Marseille-P3801^{T}, which belongs to the genus *Gracilibacillus* [8]. The genus *Gracilibacillus* currently compromises 13 species validly standing in nomenclature [9]. *Gracilibacillus* species were isolated from diverse salty environmental samples, including sea water, salty lakes [4,10,11], soil [12,13] and/or food [1,14,15] and gut microbiota [6].

Various parameters, including phenotypic and genotypic characteristics, such as DNA-DNA hybridization, were used to define a new species, but they present certain limitations [16]. In our study, using a taxonogenomic approach that includes phenotypic characteristics, proteomic information obtained by MALDI-TOF MS and the analysis of the complete genome sequence, we were able to achieve a complete description of the new halophilic species called *G. phocaeensis*; the strain Marseille-P3801T (= CSUR P3801) is the type strain of *Gracilibacillus phocaeensis* sp. nov.

Materials and Methods

Bacterial strains and growth conditions

Strain Marseille-P3801^{T} was isolated from stool samples of a 20-year-old man from N'Diop. The study was approved by the ethics committee of the Institut Hospitalo-Universitaire Méditerranée Infection under number 2016-011, and the patient provided a signed informed consent. The percentage of NaCl in the stool sample was determined using a salinity refractometer (Thermo Scientific, Villebon-sur-Yvette, France) by diluting 1 g in 10 mL of distilled water and centrifuging it for 10 minutes at 5000g. In a second time, 100 µL of supernatant was deposited in the refractometer; the result was in a straight line, displayed on the screen in per mille and then reported in percentage of NaCl. To culture the bacteria from stool samples, we used an aerobic blood culture bottle (Becton Dickinson, Le Pont-de-Claix, France) containing a halophilic medium prepared in a modified Columbia broth (Sigma-Aldrich, Saint-Quentin-Fallavier, France), by adding (per litre: 1% (w/v) MgSO4, 0.1% (w/v) MgCl2, 0.4% (w/v) KCl, 0.1% (w/v) CaCl2, 0.05% (w/v) NaHCO3, 0.2% (w/v) of glucose, 0.5% (w/v) of yeast extract (Becton Dickinson), and from 10 to 15% (w/v) NaCl according to the required salinity with a pH adjusted to 7.5 and incubated for 3 days in an aerobic atmosphere at 37°C [8]. All strains were first isolated in a halophilic culture medium with 15% (w/v) NaCl. The initial growth of colonies on agar were obtained after 24 hours of incubation at 37°C under aerobic conditions. The oxygen requirement was evaluated by incubating strain Marseille-P3801^{T} under aerobic, microaerophilic, and anaerobic conditions using AnaeroGenTM (Atmosphere Generation Systems, Dardily, France) at 37°C. The colonies isolated were identified using MALDI-TOF MS, as previously described [17]. As for unidentified colonies, the 16S rRNA gene was sequenced and the obtained sequence was matched against the NCBI database using the BLAST algorithm [18].

16S rRNA gene sequencing and phylogenetic analysis

The 16S rRNA gene sequence of the strain was determined for subsequent phylogenetic analysis. The genomic DNA of the strain was amplified by PCR using the primers pair fD1 and rP2 (Eurogentec, Angers, France) [19] and sequenced with the MiSeq Technology (Illumina Inc, San Diego, CA, USA) as previously described [20]. The 16S rRNA nucleotide sequences were assembled and corrected using CodonCode Aligner software (http://www.codoncode.com). A BLAST research (Basic Local Alignment Search Tool) was further performed against the GenBank nucleotide collection (http://blast.ncbi.nlm.nih.gov/Blast.cgi). If the 16S rRNA sequence similarity value was lower than 98.65% with the most closely related species with standing in nomenclature, as proposed by Stackebrandt [21], the strain was proposed as belonging to a new species [22].

Phenotypic and biochemical characteristics

The morphology of strain Marseille-P3801^{T} was revealed by a negative staining observed on a Hitachi SU5000 scanning electron microscope (Hitachi Group, Krefeld, Germany) and Gram staining observed on photonic microscope Leica DM2500 (Leica Microsystems, Nanterre, France) with a 100X oil-immersion objective. Sporulation, motility catalase and oxidase were tested, as previously reported [23,24]. To determine the optimal growth conditions, the strain Marseille-P3801 was cultivated in Mueller Hinton agar (Sigma-Aldrich) by varying the NaCl concentrations from 5 to 20% (w/v) as well as the pH values (5, 5.5, 6, 6.5, 7, 7.5 and 8). It was also seeded on 5% sheep blood-enriched Columbia agar (bioMérieux, Marcy l'Etoile, France) and incubated under different temperatures (25, 28, 37, 45 and 55°C). API 50CH, ZYM and 20 NE test strips (bioMérieux, Marcy-l'Etoile, France) were used to study the carbohydrate metabolism, enzymes activities and biochemical criteria of strain Marseille-P3801, according to the manufacturer's instructions.

Cellular fatty acid methyl ester (FAME) analysis was performed by GC/MS. Two samples were prepared with approximately 90 mg of bacterial biomass per tube, harvested from

several culture plates. Fatty acid methyl esters were prepared as described by Sasser [25]. GC/MS analyses were carried out as previously described [26]. Briefly, fatty acid methyl esters were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500 - SQ 8 S, Perkin Elmer, Courtaboeuf, France). Spectral database search was performed using MS Search 2.0 operated with the Standard Reference Database 1A (NIST, Gaithersburg, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK).

Genome: extraction, sequencing and assembly

Genomic DNA (gDNA) of *Gracilibacillus phocaeensis* was extracted in two steps : a mechanical treatment was first performed by glass beads acid washed (G4649-500g Sigma) using a FastPrep BIO 101 instrument (Qbiogene, Strasbourg, France) at maximum speed (6.5m/s) for 90 seconds. After 30 minutes incubation of the lysozyme at 37°C, the DNA was extracted on the EZ1 biorobot (Qiagen) with the EZ1 DNA tissue kit. The elution volume is 50µL. Genomic DNA was evaluated with a Qubit test (Life technologies, Carlsbad, CA, USA). The mate pair library was prepared with 1.5 µg of genomic DNA using the Nextera mate pair Illumina guide. The DNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The pattern of fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies Inc, Santa Clara, CA, USA) with a DNA 7500 labchip kit. The DNA fragments size ranged from 1.5 kb up to 11 kb, with an optimal size at 8.10 kb. No size selection was performed and 600 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal at 1086 bp on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip kit (Agilent Technologies Inc) and the final concentration library was measured at 31.31 nmol/l. The libraries were normalized at 2nM 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 2x251-bp. Total information of 8.2 Gb was obtained from a 932 K/mm2 cluster density with a cluster passing quality control filters of 91 %. Within this run, the index representation for *Gracilibacillus phocaeensis* was determined at 13.20 %. The 2 141 870 paired end reads were filtered according to the read qualities.

The assembly was performed with a pipeline incorporating different softwares, including Velvet[27], Spades [28] and Soap Denovo [29] on trimmed (Trimmomatic) [30] or raw data. To reduce assembly gap, GapCloser software was used. The scaffolds inferior to 800 bp and scaffolds with a depth value < 25% of the mean depth were discarded. The best assembly was selected using different criteria (number of scaffolds, N50, number of N).

Genome annotation and comparisons

Open reading frames (ORFs) were predicted using the Prodigal tool (http://prodigal.ornl.gov) with defaults parameters. tRNAs and rRNAs were detected using tRNAscan-SE v.1.2129 and RNAmmer v.1.230, respectively[31, 32]. The protein sequence annotation was performed on NCBI GenBank non-redundant protein sequence database (nr) using BLAST protein with an e-value of 1e-03 as significance thresholds [33]. Then, we obtained the functional classification of gene families (COG ID and letters) using EggNOG against the COG database [34]. The genome of *Gracilibacillus phocaeensis* strain Marseille-P3801^{T} (EMBL EBI accession number UZBG00000000) was compared with that of *Gracilibacillus boraciitolerans* strain JCM 21714 (BAVS00000000), *Gracilibacillus massiliensis* strain Awa-1^{T} (CZRP00000000), *Gracilibacillus lacisalsi* DSM 19029 (ARIY00000000), *Gracilibacillus ureilyticus* CGMCC (FOGL00000000), *Gracilibacillus dipsosauri* (QGTD00000000) and *Halobacillus karajensis* DSM 14948 (FNWW00000000) using OrthoANI software [35].

Results

Strain identification and phylogenetic analysis

Strain Marseille-P3801^T formed yellow colonies after 1 day of culture on agar with horse blood ranging from 2 to 20% (w/v) NaCl (optimum at 7.5 (w/v) NaCl) at 37°C. The spectrum resulting from the 8 pure colonies of strain Marseille-P3801 deposited on MALDI-TOF target plate did not allow the identification of this bacterium because there was no spectrum match with those in the Bruker database (**Supplementary Fig. S1**). Using 16S rRNA sequence of *G. phocaeensis* (LT934503.1), the phylogenetic analysis revealed that strain Marseille-P3801 exhibited a sequence similarity of 98.39% with *Gracilibacillus thailandensis* strain TP2-8 (GenBank accession no. NR_116568.1), the phylogenetically closest species with standing in nomenclature (**Fig. 1**). Therefore, we classified this strain as a member of a new species within the genus *Gracilibacillus*, family *Bacillaceae* and phylum *Firmicutes*. This value was lower than the 98.7% 16S rRNA gene sequence, threshold advised by Meier-Kolthoff *et al.* [36] to delineate a new species without carrying out DNA-DNA hybridization. Classification and general features of strain Marseille-P3801 are summurazed in **Table S1**.

Physiological and biochemical characteristics

G. phocaeensis sp. nov. strain Marseille-P3801^T (= CSUR P3801) is Gram positive. Colonies are circular, yellow with a mean diameter of 2 mm after 2-3 days of growth on 5% sheep blood-enriched Columbia agar medium (BioMérieux, Marcy l'Etoile, France). Bacterial cells of strain Marseille-P3801 were motile, rod shaped and polymorphic (**Fig. 2**).

The major fatty acids were saturated structures mainly branched: 12-methyl-tetradecanoic acid (59 %), hexadecanoic acid (16 %) and 14-methyl-hexadecanoic acid (11 %). Other saturated and branched fatty acids were also described. 7-Hexadecenoic acid was the only one unsaturated structure detected (**Supplementary Table S2**). Catalase and oxidase were positive. Using API ZYM trips, positive reaction were detected for lipases (C4, C8 and C14), leucine arylamidase, acid phosphatase, naphthol-as-bi-phosphohydrolase, β-galactosidase, α-glucosidase, β-glucosidase and N-acetyl-β-glucosaminidase, but no reaction was observed for

alkaline phosphatase, valine and cysteine arylamidase, α-chymotrypsin, α-galactosidase, trypsin, β-glucuronidase, α-mannosidase and α-fucosidase. API 20NE strip exhibited positive reactions of fermentation of glucose, urease activity, and metabolism of L-arginine and esculin. In contrast, negative reactions were observed for nitrate and indole production as well as metabolism of D-glucose, L-arabinose, D-mannose, D-maltose, D-mannitol, N-acetyl-glucosamine, potassium gluconate, capric acid, malic acid and phenylacetic acid. Using API 50CH strip , strain Marseille-P3801^{T} exhibited esculin hydrolysis and negative reactions for D-galactose, D-lactose, D-maltose, D-ribose, D-saccharose, D-lyxose, D-mannose, L-sorbose, D-tagatose, D-turanose, D-xylose, L-xylose, D-arabinose, L-arabinose, D-sorbitol, D-cellobiose, D-melezitose, D-melibiose, D-trehalose, D-raffinose, D-arabitol, L-arabitol, D-glucose, D-fructose, D-fucose, L-rhamnose, D-adonitol, D-mannitol, L-fucose, amygdalin, arbutin, erythritol, dulcitol, gentiobiose, glycerol, glycogen, inositol, inulin, salicin, starch, xylitol, αD-glucopyranoside, methyl-βD-xylopyranoside, methyl-αD-mannopyranoside, potassium gluconate and N-acetylglucosamine. The differential characteristics of *Gracilibacillus phocaeensis* with respect to other bacteria related to the genus *Gracilibacillus* are outlined in **Table 1.** Phenotypic characterization of *Gracilibacillus phocaeensis* sp. nov., based on analytical profile index (API) tests was summurized on **Table 2.**

Genome properties

The genome length of *Gracilibacillus phocaeensis* strain Marseille-P3801^{T} is 4.66 megabases encompassing 12 scaffolds (11 contigs). The guanine-cytosine (GC) content is 39.6 mol%. Among the 4,390 predicted genes, 4,255 were protein-coding genes, 67 were RNAs (11 rRNA, 52 tRNAs and 4 non-coding RNAs) and 68 pseudogenes. The BLASTp annotation of *G. phocaeensis* strain Marseille-P3801 had assigned a putative function to 3,606 genes and 649 genes annotated as hypothetical proteins. To further insight on the gene function, a

comparison of *G. phocaeensis* protein sequences with COG database was realized. Out of 4,390 protein encoding genes, 3,774 were assigned to a COG function (86%) distributed in 20 COG categories (**Supplementary Fig. S2**).

Genome comparison

To explore the genomic similarity of *G. phoceeanensis* with closely related bacteria, we performed an OrthoANI analysis. Among closely related species, we found OrthoANI values ranging from 66.64% between *Gracilibacillus phocaeensis* strain Marseille-P3801 and *Halobacillus karajensis* DSM 14948, to 78.39 % between *Gracilibacillus boraciitolerans* strain JCM 21714 and *Gracilibacillus massiliensis* strain Awa-1. When *Gracilibacillus phocaeensis* strain Marseille-P3801 was compared with these closely related species, we found values ranging from 66.64% with *Halobacillus karajensis* DSM 14948, to 72.42% with *Gracilibacillus lacisalsi* strain DSM 19029 (**Fig. 3**). The representations of strain Marseille-P3801 genome and its genes repartition into functional categories are illustrated in **Fig. 4**.

Description of *Gracilibacillus phocaeensis*

Gracilibacillus phocaeensis (pho.ca.een'sis, N.L. masc. adj., from *phocaeensis*, related to the Phocaeans, the founders of Marseille, France where the type strain was isolated and characterized like many others species). It is Gram positive, motile and aerobic bacterium. The colonies are yellow, circular with a mean diameter of 2 mm. Bacterial cells were rod-shaped and polymorphic. Strain Marseille-P3801 grows at an optimal temperature of 37°C, pH 7 with 7.5% (w/v) of NaCl. It is catalase and oxidase positive. Positive reactions were observed for esterase (C4), esterase lipase (C8), acid phosphatase, naphthol-as-bi-phosphohydrolase, leucine arylamidase, β-galactosidase, α-glucosidase, β-glucosidase and N-acetyl-β-glucosaminidase. *G. phocaeensis* strain Marseille-P3801^{T} was isolated from a stool sample of a 20-year-old man from N'Diop, Senegal. This strain exhibited a G+C content of 39.6 mol%. Its 16S rRNA sequence was deposited in GenBank under accession number

LT934503, and the whole genome shotgun sequence was deposited in GenBank under accession number UZBG00000000.

Discussion and conclusion

The concept of "microbial culturomics", based on the variation of physicochemical parameters of the culture conditions, allows the exploration of microbial diversity of different ecosystems such as gut microbiota [7]. Microbial culturomics provides culture conditions that simulate, reproduce or mimic all the selective constraints that have shaped the natural microbiota for millions of years. Many new bacterial species have been discovered, particularly those belonging to the *Bacillales* order, which is one of the most represented bacterial orders [37]. To explore the halophilic microbiota of the human gut, the use of culture media with a high salt content, allowed us to isolate a new moderately halophilic bacterial strain, Marseille-P3801^{T} which belongs to the genus *Gracilibacillus* [8]. To the best of our knowledge, this is the second *Gracilibacillus* species described in the human gut. Based on its phenotypic, phylogenetic and genomic characteristics, this strain is proposed to represent a novel species in the genus *Gracilibacillus*, for which the name *Gracilibacillus phocaeensis* sp. nov. is proposed, with Marseille-P3801^{T} as the type strain.

References

1. Oh YJ, Lee H-W, Lim SK, Kwon M-S, Lee J, Jang J-Y, et al. *Gracilibacillus kimchii* sp. nov., a halophilic bacterium isolated from kimchi. J Microbiol Seoul Korea. 2016;54:588–93.
2. Gibtan A, Park K, Woo M, Shin J-K, Lee D-W, Sohn JH, et al. Diversity of Extremely Halophilic Archaeal and Bacterial Communities from Commercial Salts. Front Microbiol [Internet]. 2017 [cited 2019 Aug 19];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5423978/
3. Miranda PM, De Palma G, Serkis V, Lu J, Louis-Auguste MP, McCarville JL, et al. High salt diet exacerbates colitis in mice by decreasing *Lactobacillus* levels and butyrate production. Microbiome. 2018; 6:57.
4. Wainø M, Tindall BJ, Schumann P, Ingvorsen K. *Gracilibacillus* gen. nov., with description of *Gracilibacillus halotolerans* gen. nov., sp. nov.; transfer of *Bacillus dipsosauri* to *Gracilibacillus dipsosauri* comb. nov., and *Bacillus salexigens* to the genus *Salibacillus* gen. nov., as *Salibacillus salexigens* comb. nov. Int J Syst Bacteriol. 1999; 49 Pt 2:821–31.
5. Khelaifia S, Lagier J-C, Bibi F, Azhar EI, Croce O, Padmanabhan R, et al. Microbial Culturomics to Map Halophilic Bacterium in Human Gut: Genome Sequence and Description of *Oceanobacillus jeddahense* sp. nov. OMICS J Integr Biol. 2016; 20:248–58.
6. Diop A, Seck E hadji, Dubourg G, Armstrong N, Blanc-Tailleur C, Raoult D, et al. Genome sequence and description of *Gracilibacillus timonensis* sp. nov. strain Marseille-P2481T, a moderate halophilic bacterium isolated from the human gut microflora. MicrobiologyOpen [Internet]. 2018 [cited 2019 Aug 8]; 8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6485780/
7. Lagier J-C, 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 Off Publ Eur Soc Clin Microbiol Infect Dis. 2012;18:1185–93.
8. Senghor B, Seck EH, Khelaifia S, Bassène H, Sokhna C, Fournier P-E, et al. Description of '*Bacillus dakarensis*' sp. nov., '*Bacillus sinesaloumensis*' sp. nov., '*Gracilibacillus timonensis*' sp. nov., '*Halobacillus massiliensis*' sp. nov., '*Lentibacillus massiliensis*' sp. nov., '*Oceanobacillus senegalensis*' sp. nov., '*Oceanobacillus timonensis*' sp. nov., '*Virgibacillus dakarensis*' sp. nov. and '*Virgibacillus marseillensis*' sp. nov., nine halophilic new species isolated from human stool. New Microbes New Infect. 2017; 17:45–51.
9. Parte AC. LPSN—list of prokaryotic names with standing in nomenclature. Nucleic Acids Res. 2014;42:D613–6.
10. Gao M, Liu Z-Z, Zhou Y-G, Liu H-C, Ma Y-C, Wang L, et al. *Gracilibacillus kekensis* sp. nov., a moderate halophile isolated from Keke Salt Lake. Int J Syst Evol Microbiol. 2012;62:1032–6.
11. Jeon CO, Lim J-M, Jang HH, Park D-J, Xu L-H, Jiang C-L, et al. *Gracilibacillus lacisalsi* sp. nov., a halophilic Gram-positive bacterium from a salt lake in China. Int J Syst Evol Microbiol. 2008;58:2282–6.

12. Huo Y-Y, Xu X-W, Cui H-L, Wu M. *Gracilibacillus ureilyticus* sp. nov., a halotolerant bacterium from a saline–alkaline soil. Int J Syst Evol Microbiol. 2010;60:1383–6.
13. Chen Y-G, Cui X-L, Zhang Y-Q, Li W-J, Wang Y-X, Xu L-H, et al. *Gracilibacillus halophilus* sp. nov., a moderately halophilic bacterium isolated from saline soil. Int J Syst Evol Microbiol. 2008;58:2403–8.
14. Chamroensaksri N, Tanasupawat S, Akaracharanya A, Visessanguan W, Kudo T, Itoh T. *Gracilibacillus thailandensis* sp. nov., from fermented fish (pla-ra). Int J Syst Evol Microbiol. 2010;60:944–8.
15. Diop A, Khelaifia S, Armstrong N, Labas N, Fournier P-E, Raoult D, et al. Microbial culturomics unravels the halophilic microbiota repertoire of table salt: description of *Gracilibacillus massiliensis* sp. nov. Microb Ecol Health Dis [Internet]. 2016 [cited 2019 Aug 18];27. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5071648/
16. Vandamme P, Pot B, Gillis M, de Vos P, Kersters K, Swings J. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol Rev. 1996;60:407–38.
17. Lo CI, Fall B, Sambe-Ba B, Diawara S, Gueye MW, Mediannikov O, Sokhna C et al. MALDI-TOF Mass Spectrometry: A Powerful Tool for Clinical Microbiology at Hôpital Principal de Dakar, Senegal (West Africa). PLoS One 2015; 10(12):e0145889. doi: 10.1371/journal.pone.0145889.
18. Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral J-P, Raoult D. 16S Ribosomal DNA Sequence Analysis of a Large Collection of Environmental and Clinical Unidentifiable Bacterial Isolates. J Clin Microbiol. 2000;38:3623–30.
19. Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol. 1991;173:697–703.
20. Morel A-S, Dubourg G, Prudent E, Edouard S, Gouriet F, Casalta J-P, et al. Complementarity between targeted real-time specific PCR and conventional broad-range 16S rDNA PCR in the syndrome-driven diagnosis of infectious diseases. Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol. 2015; 34:561–70.
21. Stackebrandt E. Taxonomic parameters revisited : tarnished gold standards. Microbiol Today. 2006;33:152–5.
22. Sentausa E, Fournier P-E. Advantages and limitations of genomics in prokaryotic taxonomy. Clin Microbiol Infect. 2013; 19:790–5.
23. Diop K, Diop A, Khelaifia S, Robert C, Pinto FD, Delerce J, Raoult D, Fournier PE, Bretelle F, Fenollar F. Characterization of a novel Gram-stain-positive anaerobic coccus isolated from the female genital tract: Genome sequence and description of *Murdochiella vaginalis* sp. nov. Microbiologyopen. 2018; 7(3):e00570. doi: 10.1002/mbo3.570.
24. Bilen M, Mbogning Fonkou MD, Nguyen TT, Richez M, Daoud Z, Fournier PE, Raoult D, Cadoret F. *Miniphocibacter massiliensis* gen. nov., sp. nov., a new species isolated from the human gut and its taxono-genomics description. Microbiologyopen. 2019; 8(5):e00735. doi: 10.1002/mbo3.735.
25. Sasser M. Bacterial Identification by Gas Chromatographic Analysis of Fatty Acids Methyl Esters (GC-FAME). :6.

26. Dione N, Sankar SA, Lagier JC, Khelaifia S, Michele C, Armstrong N, et al. Genome sequence and description of *Anaerosalibacter massiliensis* sp. nov. New Microbes New Infect. 2016; 10:66–76.
27. Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008; 18:821–9.
28. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol. 2012;19:455–77.
29. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience. 2012; 1:18.
30. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30:2114–20.
31. 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.
32. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 2007; 35:3100–8.
33. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215:403–10.
34. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, et al. Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. Mol Biol Evol. 2017; 34:2115–22.
35. Lee I, Ouk Kim Y, Park S-C, Chun J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol. 2016; 66:1100–3.
36. Meier-Kolthoff JP, Göker M, Spröer C, Klenk H-P. When should a DDH experiment be mandatory in microbial taxonomy? Arch Microbiol. 2013; 195:413–8.
37. Lagier J-C, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P, et al. Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol. 2016; 1:16203.
38. Hirota K, Hanaoka Y, Nodasaka Y, Yumoto I. *Gracilibacillus alcaliphilus* sp. nov., a facultative alkaliphile isolated from indigo fermentation liquor for dyeing. Int J Syst Evol Microbiol. 2014 Sep;64(Pt 9):3174-3180. doi: 10.1099/ijs.0.060871-0.
39. Tang SK, Wang Y, Lou K, Mao PH, Jin X, Jiang CL, Xu LH, Li WJ. *Gracilibacillus saliphilus* sp. nov., a moderately halophilic bacterium isolated from a salt lake. Int J Syst Evol Microbiol. 2009 Jul;59(Pt 7):1620-4. doi: 10.1099/ijs.0.006569-0.
40. Carrasco IJ, Márquez MC, Yanfen X, Ma Y, Cowan DA, Jones BE, Grant WD, Ventosa A. *Gracilibacillus orientalis* sp. nov., a novel moderately halophilic bacterium isolated from a salt lake in Inner Mongolia, China. Int J Syst Evol Microbiol. 2006 Mar;56(Pt 3):599-604. doi: 10.1099/ijs.0.63971-0. PMID: 16514034.

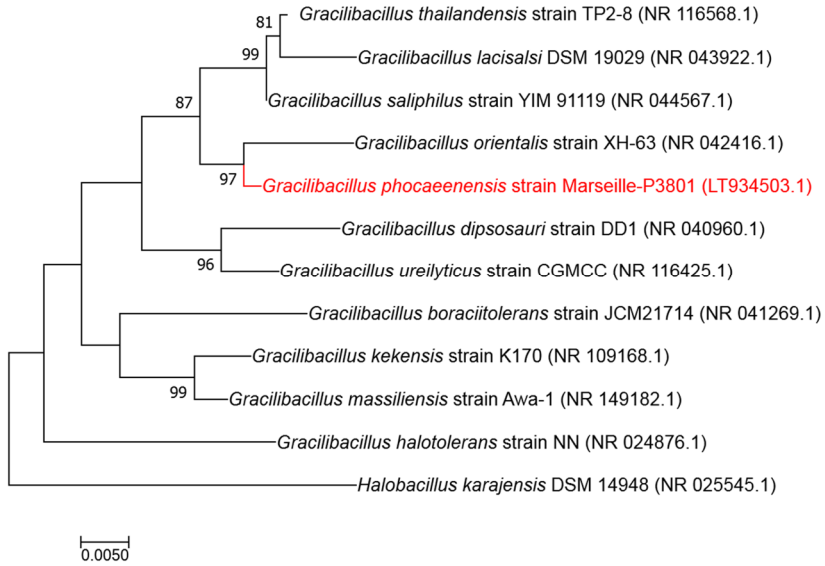
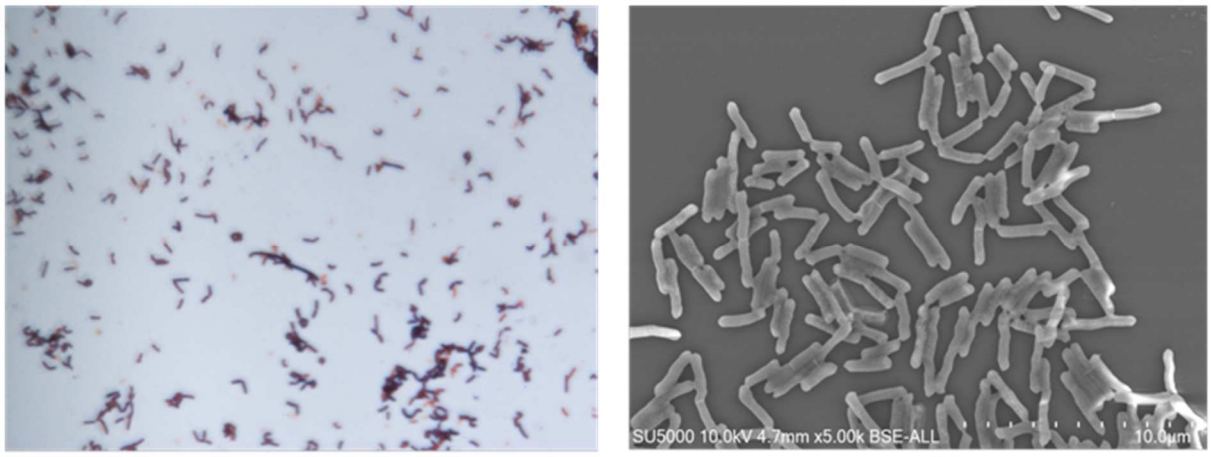


Fig. 1: Phylogenetic tree highlighting the phylogenetic position of *Gracilibacillus phocaeensis* strain Marseille-P3801T relative to other phylogenetically close members of the family *Bacillaceae*. Sequences were aligned using CLUSTALW, and phylogenetic inferences were obtained using the maximum likelihood method within the MEGA7 software. Numbers at the nodes are percentages of bootstrap values obtained by repeating the analysis 500 times to generate a majority consensus tree.



(a) (b)

Fig. 2: (a) Gram staining of *Gracilibacillus phocaeensis* strain Marseille-P3801T. (b) Morphological structure of *G. phocaeensis* strain Marseille-P3801T obtained with Hitachi SU5000 scanning electron microscope. Scale bar and acquisition settings are shown of figure.

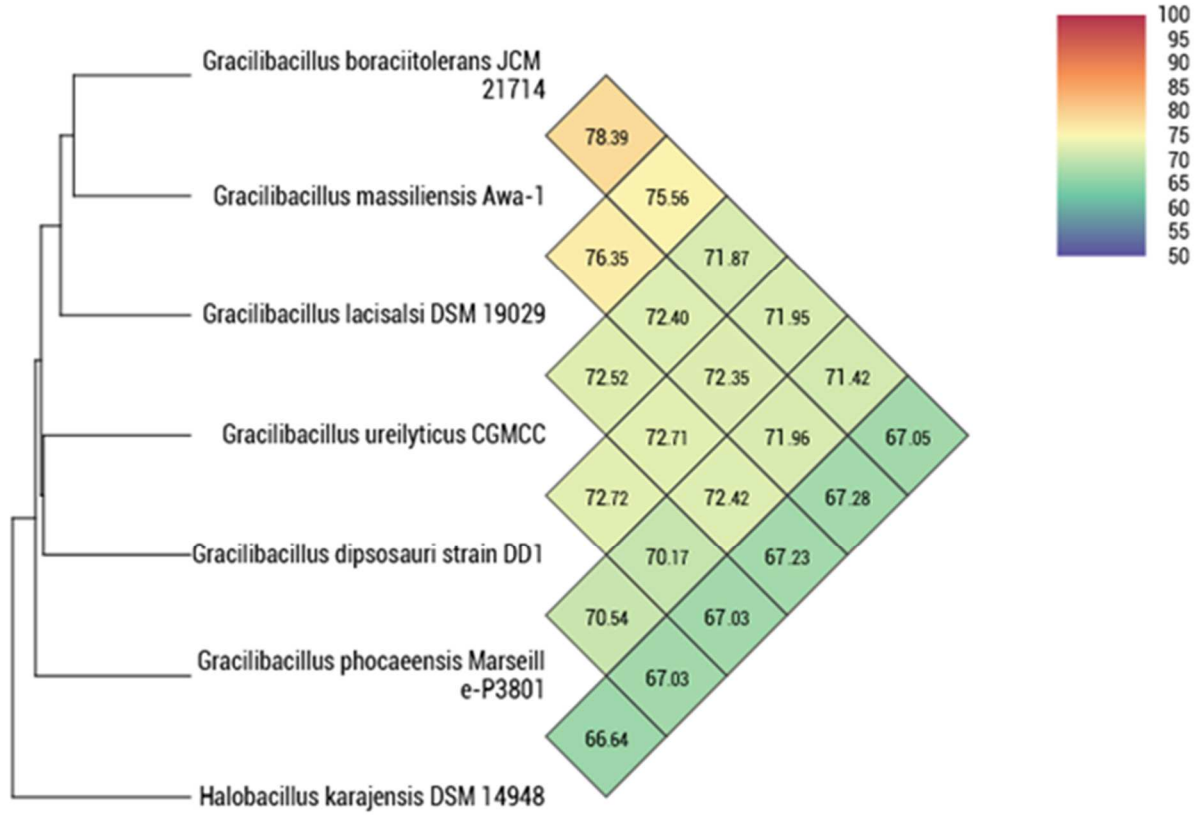


Fig. 3: Heatmap generated with OrthoANI values calculated using the OAT software between *Gracilibacillus phocaeensis* sp. nov. strain Marseille-P3801. nov. and other closely related species with standing in nomenclature.

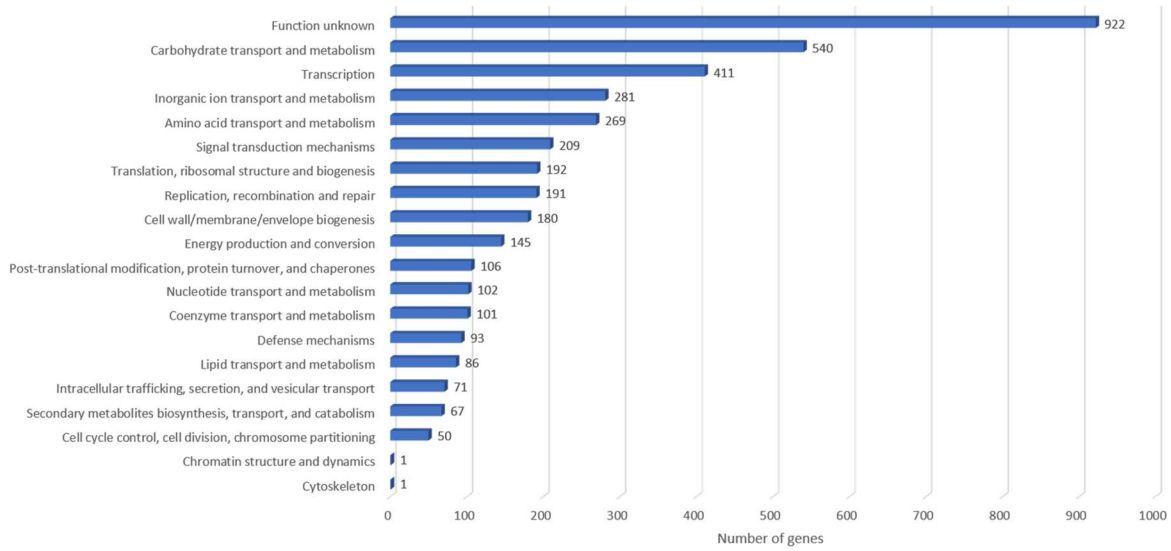


Fig. 4: Distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins of *Gracilibacillus phocaeensis* strain Marseille-P3801

Table 1: Differential characteristics of **1**, *Gracilibacillus phocaeensis* strain Marseille-P3801 compared to other close bacteria of the genus *Gracilibacillus*: **2,** *G. timonensis* strain Marseille-P2481[6]; **3**, *G. massiliensis* strain Marseille-P1441[15]; **4**, *G. alcaliphilus* strain SG103 [38]; **5,** *G. saliphilus* strain YIM 91119 [39]; **6**, *G. orientalis* strain XH-63 [40] and **7**, *G. halophilus* strain YIM-C55.5 [13].

Properties	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Cell diameter (µm)	0.3-0.6	0.5-0.8	0.3-1.8	0.5-0.7	0.7-0.9	0.7-0.9	0.3-0.5
Pigmentation	Yellow	Creamy orange	White	Creamy white	Creamy white	Creamy	White
Oxygen requirement	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic
Gram stain	+	+	+	+	+	+	+
Salt requirement	+	+	+	+	+	+	+
Motility	+	+	+	+	+	+	+
Sporulation	+	+	−	+	+	+	+
Indole	−	−	−	−	−	−	−
Production of :							
Alkaline phosphate	−	−	−	−	+	NA	+
Catalase	+	+	+	+	+	+	+
Oxidase	+	−	−	−	+	−	+
Nitrate reductase	−	−	−	+	+	−	+
Urease	+	+	+	+	+	−	−
β-galactosidase	+	+	−	NA	+	NA	+
α-galactosidase	−	−	+	−	−	NA	−
N-acetyl-glucosamine	−	−	−	+	+	NA	−
Acid from							
L-Arabinose	−	−	−	+	+	+	−
D-mannose	−	−	−	−	+	−	−
D-mannitol	−	−	−	+	+	+	+
D-glucose	−	−	−	+	+	+	+
D-fructose	−	−	−	+	+	+	+
D-maltose	−	−	−	+	+	+	−
D-lactose	−	−	−	+	+	+	−

DNA G+C content (mol%)	39.6	39.8	36.05	41.3	40.1	37.1	42.3	
Habitat	Human gut	Human gut	Cooking salt	Fermentation liquor	Salt lake	Salt lake	Salt soil	

4 +, positive reaction; -, negative reaction; NA, Not available data.

Table 2: Phenotypic characterization of *Gracilibacillus phocaeensis* sp. nov., strain Marseille-P3801 based on analytical profile index (API) tests.

Tests	Characteristics	Results	Tests	Characteristics	Results
API ZYM	Alkaline phosphatase	-	API 50 CH	Glycerol	-
	Esterase (C4)	+		Erythritol	-
	Esterase Lipase (C8)	+		D-arabinose	-
	Lipase (C14)	+		L-arabinose	-
	Leucine arylamidase	+		D-ribose	-
	Valine arylamidase	-		D-xylose	-
	Cystine arylamidase	-		L-xylose	-
	Trypsin	-		D-Adonitol	-
	α-chymotrypsin	-		Methyl βD-xylopyranoside	-
	Acid phosphatase	+		D-galactose	-
	Naphthol-AS-BI-phosphohydrolase	+		D-glucose	-
	α-galactosidase	-		D-fructose	-
	β-galactosidase	+		D-mannose	-
	β-glucuronidase	-		L-sorbose	-
	α-glucosidase	+		L-rhamnose	-
	β-glucosidase	+		Dulcitol	-
	N-acetyl-β-glucosaminidase	+		Inositol	-
	α-mannosidase	-		D-mannitol	-
	α-fucosidase	-		D-sorbitol	-
API 20 NE	Nitrates to nitrites	-		Methyl αD-mannopyranoside	-
	Indole	-		Methyl αD-glucopyranoside	-
	Glucose fermentation	+		N-acetyl-glucosamine	-
	Arginine dihydrolase	-		Amygdalin	-
	Urease	+		Arbutin	-
	β-glucosidase	-		Esculin ferric citrate	+
	Protease	-		Salicin	-
	β-galactosidase	-		D-cellobiose	-
	Glucose assimilation	-		D-maltose	-
	Arabinose	-		D-lactose	-
	Mannose	-		D-melibiose	-
	Mannitol	-		Sucrose	-
	N-acetyl-glucosamine	-		D-trehalose	-
	Maltose	-		Inulin	-
	Potassium gluconate	-		D-melezitose	-
	Capric acid	-		D-raffinose	-
	Adipic acid	-		Starch	-
	Malate	-		Glycogen	-
	Trisodium Citrate	-		Xylitol	-

Phenylacetic acid	-	Gentiobiose	-
		D-turanose	-
		D-lyxose	-
		D-tagalose	-
		D-fucose	-
		L-fucose	-
		D-arabitol	-
		L-arabitol	-
		Potassium gluconate	-
		Potassium 2-ketogluconate	-
		Potassium 5-ketogluconate	-

3 **+**, positive reaction; **-**, negative reaction.