Antonie van Leeuwenhoek (2026) 119:147
https://doi.org/10.1007/s10482-026-02359-x

ORIGINAL PAPER

Neobacillus camarae sp. nov. a new bacterium isolated from the breast milk of a Senegalese mother breast-feeding a healthy child and genomic description of *Neobacillus dielmonensis* DSM 27844 FF4^T and *Neobacillus drentensis* LMG 21831^T

Salematou Sarr · Ousmane Ndour · Adama Mbow · Mamadou Beye · Anissa Idrissa Abdoulaye · Sara Bellali · Stephane Alibar · Rayan Semmar · Nicolas Orain · George Diatta · Camille Valles · Clemence Thiakane · Maryam Tidjani Alou · Cheikh Sokhna · Matthieu Million

Received: 28 January 2026 / Accepted: 8 June 2026

Abstract As part of a comprehensive study investigating the breast milk microbiota of healthy and malnourished children through a culturomic approach, a novel bacterial strain belonging to the genus *Neobacillus* was isolated, identified, and characterized using a polyphasic strategy. This isolate, designated as strain *Neobacillus camarae* sp. nov. Marseille-QA0830^T, was recovered from the breast milk of a mother nursing a healthy infant. Strain Marseille-QA0830^T (=CSUR QA0830^T=CECT 31295^T) is a Gram-stain-positive, non-motile, non-spore-forming bacillus, facultatively anaerobic showing optimal growth under aerobic conditions but also capable of developing in anaerobic environment at temperatures ranging from 20 to 45 °C. With a genome size of 5.83 Mbp, strain Marseille-QA0830^T exhibits a G+C content of 42.3%. Phylogenetic analysis based on the 16S rRNA gene revealed a high sequence similarity (98.10%) between strain Marseille-QA0830^T and *Neobacillus niacini* NBRC 15566^T (=IFO15566). A significant genomic divergence was observed through digital DNA-DNA hybridization (dDDH), against *Neobacillus dielmonensis* with 26.2% similarity well below the 70% threshold for species delineation. At the same time, *rpoB* gene analysis indicated a distant relationship with *Neobacillus drentensis* LMG 21831^T (83.5% similarity). Collectively, these genomic data demonstrate that although strain Marseille-QA0830^T is phylogenetically closely related to *N. niacini* NBRC 15566^T based on 16S rRNA gene sequences, it remains distinct at the whole-genome

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10482-026-02359-x.

S. Sarr · O. Ndour · A. Mbow · M. Beye · A. Idrissa Abdoulaye · S. Bellali · S. Alibar · R. Semmar · N. Orain · C. Valles · M. Tidjani Alou · C. Sokhna · M. Million (✉)
IHU Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13385 Marseille cedex 05, France
e-mail: matthieumillion@gmail.com

S. Sarr · O. Ndour · A. Idrissa Abdoulaye · R. Semmar · C. Valles · M. Tidjani Alou · M. Million
Aix Marseille Univ, MEPHI, 19-21 Boulevard Jean Moulin, 13385 Marseille cedex 05, France

M. Beye · N. Orain · C. Valles · M. Million
Assistance Publique–Hôpitaux de Marseille (APHM), 19-21 Boulevard Jean Moulin, 13385 Marseille cedex 05, France

G. Diatta · C. Sokhna
Aix Marseille Univ, IRD, MINES, Campus International IRD-UCAD, Dakar, Senegal

C. Thiakane
ICP Saint Martin Health Centre, 99000 Dakar, Senegal

C. Sokhna
Aix Marseille Univ, RITMES, 19-21 Boulevard Jean Moulin, 13385 Marseille cedex 05, France

Published online: 13 June 2026

and *rpoB* levels from previously described species within the *Neobacillus* nomenclature. These findings confirm that this isolate represents a unique and previously undescribed taxon. Consequently, we propose the description of the type strain Marseille-QA0830^T as a new species named *Neobacillus camarae* sp. nov.

Keywords Breast milk · *Neobacillus camarae* sp. nov. · New species · Senegal · Severe acute malnutrition · Taxonogenomics

Introduction

The genus *Neobacillus* was recently proposed to resolve the polyphyletic nature of the original *Bacillus* genus, which was long considered a phylogenetically incoherent group. Originally established in 1872 by Ferdinand Julius Cohn (Cohn 1872), *Bacillus* had expanded to include nearly 293 species and subspecies by 2020 (Patel and Gupta 2020). Members of this group exert a profound impact on human activity, ranging from notorious pathogens such as *Bacillus anthracis* and *Bacillus cereus* (Koehler 2009; Bottone 2010; Schoeni and Wong 2005) to beneficial species for agriculture (*B. thuringiensis*, *B. velezensis*) (Sanchis and Bourguet 2009), industrial enzyme production (*B. subtilis*, *B. clausii*, *B. licheniformis* and *B. wakoensis*) (Logan and De Vos 2009; Logan 2012; Pignatelli and Moya 2009) and as a natural resource for plant health and nutrition (Saxena et al. 2020).

In 1991, Ash and al., revealed substantial phylogenetic heterogeneity within *Bacillus* through comparative 16S rRNA analysis, highlighting the need for a major taxonomic revision (Ash et al. 1991). Subsequently, in 2020, Patel and Gupta identified 36 molecular markers, specifically conserved signature indels (CSIs), which supported the reclassification of *Bacillus* into six new genera: *Peribacillus*, *Alkalihalobacillus*, *Cytobacillus*, *Mesobacillus*, *Metabacillus*, and *Neobacillus* (Berkeley et al. 1984; Patel and Gupta 2020). As of January 2026, the genus *Neobacillus* comprises 28 validly published species (https://lpsn.dsmz.de/genus/neobacillus) according to LPSN, 2026, represented by the type species *Neobacillus niacini* (Nagel 1991).

While members of this genus are typically isolated from soil, water, and human skin, the present report describes an isolate from human breast milk in Senegal using a culturomics approach, expending our understanding of the biogeographical diversity of host-associated *Bacillaceae*. Phylogenetic and whole-genome analyses (ANI: 83.0%, dDDH: 26.2% against *Neobacillus dielmonensis*) firmly place this non-motile, non-spore-forming isolate as a novel species within the genus *Neobacillus*. Members of this genus are ecologically renowned for their resilience in extreme niches and their distinct capacity to degrade recalcitrant organic compounds. This novel lineage underscores a highly restricted, specialized carbohydrate-metabolic profile, offering a new genomic resource to explore the latent biotechnological and enzymatic potential within the human microbiota. Consequently, we propose the name *Neobacillus camarae* sp. nov.

Materials and methods

Ethical considerations and study design

This study was initiated in 2022 following administrative authorization from the National Ethics and Scientific Committee of the Senegalese Ministry of Health and Social Action (MSAS; No. 00000233/MSAS/CNERS/SP, issued on September 11, 2022). The project aimed to characterise the breast milk microbiota of Senegalese women across four regions (Dakar, Podor, Matam, and Niakhar) to identify potential probiotic strains. All procedures were conducted in strict accordance with the ethical principles of the Declaration of Helsinki.

Sample collection and logistics

Following written informed consent, breast milk samples (10–15 mL each) were collected from participants who met established inclusion criteria. These mothers were nursing either healthy or malnourished infants and resided in Niakhar, Senegal. Strain Marseille-QA0830^T was isolated during this process. Initial processing was performed at the Institut de Recherche pour le Développement (IRD) laboratory in Dakar, where samples were transported at +4 °C, aliquoted into cryovials, and stored at −80 °C. Subsequently, the samples were shipped on dry ice to the Institut Hospitalo-Universitaire (IHU) Méditerranée

Infection in Marseille, France, for comprehensive culture-based analysis, genomic sequencing, and biological characterisation.

Isolation and identification of the strain

Strain Marseille-QA0830^T was first isolated from a breast milk sample of a mother nursing a healthy infant using a culturomic approach. The isolate was recovered following a 15-day enrichment period in BSM Broth (Nutriselect, Basic, Sigma-Aldrich, Merck, France, Catalogue N° 90,273) prepared according to the manufacturer's instructions, with the pH adjusted to 7.0 ± 0.2 at 25 °C and sterilized by autoclave at 121 °C for 15 min; within a blood culture flask under aerobic conditions. Subsequently, the enrichment was subcultured onto Columbia agar supplemented with 5% sheep blood.

Resulting colonies (5–8 per plate) were streaked onto fresh blood agar plates and incubated at 37 °C for 24 h to ensure purity. Multiple identification attempts were performed using Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) with a Biotyper® Sirius system (Bruker, Bremen, Germany). In general, a score ≥ 2 with a validly published species name indicates a correct identification, while a score between ≥ 1.7 and < 2 ensures genus identification. On the other hand, a score < 1.7 is characteristic of poor identification. Spectra were acquired and analyzed using the MBT Compass software. These spectra were compared against the BDAL 9607 MSPs (version 12) database, which contains 12,780 species and is continuously optimized with the MEPHI repository (https://www/mediterraneeinfection/com/acces-ressources/base-de-donnees/urm-data-base/).

Phenotypic and biochemical characteristics

Phenotypic characterization of strain Marseille-QA0830^T, including Gram staining, spore-forming ability, and catalase and oxidase activities, was conducted following standard microbiological procedures (Atlas and Snyder 2011). To determine optimal growth conditions, the strain was inoculated onto Columbia agar supplemented with 5% sheep blood (bioMérieux, Craponne, France) and incubated for 72 h across a temperature range of 20 °C to 56 °C (20 °C, 28 °C, 37 °C, 45 °C, and 56 °C). Growth was further assessed under various atmospheric conditions: aerobic, anaerobic (using GasPak EZ gas generators; Becton Dickinson, New Jersey, USA), and microaerophilic. Optimal pH for growth and NaCl tolerance (at concentrations of 5%, 7.5%, 10%, 15%, and 20%) were evaluated using specific media with pH values ranging from 5.0 to 8.5.

Various enzymatic activities and metabolic pathways were determined using API ZYM, API 50CH, and API 20NE strips according to the manufacturer's instructions (bioMérieux, Marcy-l'Étoile, France). These standardized phenotypic and biochemical assays were supplemented with modern descriptive procedures combining MALDI-TOF mass spectrometry (Bruker, Bremen, Germany), Next-Generation Sequencing (Illumina Miseq), gas chromatography (using fatty acid profiling) and morphological description via electron microscopy (Hitachi, Tokyo, Japan), in order to ensure a comprehensive characterization.

Antibiotic susceptibility was evaluated using the disk diffusion method in accordance with EUCAST 2025 guidelines (http://www.eucast.org/). The following antibiotics were tested: amoxicillin, amikacin, ceftazidime, ceftriaxone, ciprofloxacin, clindamycin, daptomycin, doxycycline, gentamicin, imipenem, linezolid, nitrofurantoin, oxacillin, rifampicin, teicoplanin, tobramycin, sulfamethoxazole/trimethoprim, and vancomycin. Minimum inhibitory concentrations (MICs) were subsequently determined following established recommendations (Matuschek et al. 2018; CLSI 2018).

Fatty acid analysis

Cellular fatty acid methyl esters (FAMEs) were analyzed by Gas Chromatography-Mass Spectrometry (GC–MS). Samples were prepared using approximately 50 mg of bacterial biomass per tube, harvested from multiple culture plates. FAMEs were prepared according to the protocol described by Sasser (2006), and GC–MS analyses were conducted following previously established methods (Dione et al. 2016). Briefly, fatty acid methyl esters were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500–SQ 8 S, PerkinElmer, Courtaboeuf, France). The spectral database search was performed using MS Search 2.0, utilizing the NIST

standard reference 1A (Gaithersburg, USA) and the FAME mass spectral reference library (Wiley, Chichester, UK).

Morphological characterisation by scanning electron microscopy

The morphological characteristics of strain Marseille-QA0830^T were investigated using scanning electron microscopy (SEM) with a SU5000 instrument (Hitachi, Tokyo, Japan). Fresh colonies were initially fixed in 2.5% glutaraldehyde (Electron Microscopy Sciences, USA) prepared in 0.1 M sodium cacodylate buffer (Thermo Fisher Scientific, USA). Following fixation, cell suspensions were deposited onto glass slides via cytocentrifugation at 800 rpm for 7 min using a Cytospin 4 centrifuge (Thermo Electron Corporation—Shandon, UK).

To enhance cellular contrast and optimize image resolution, the samples were stained with 10% phosphotungstic acid (PTA), pH 7.4 (Sigma-Aldrich, St. Louis, MO, USA) for 5 min. Subsequently, the specimens were sputter-coated with platinum-palladium using an MC1000 ion sputter coater (Hitachi, Tokyo, Japan). SEM micrographs were acquired at magnifications ranging from 1000× to 20,000×. Detailed acquisition parameters, including instrument model, acceleration voltage, magnification, working distance, and detection mode, are provided for each individual micrograph.

Genome extraction, sequencing and assembly

To extract genomic DNA (gDNA), the strain was first suspended in 160 µL of G2 buffer (EZ1 DNA Tissue Kit, Qiagen) and subjected to mechanical lysis with glass beads (G4649-500 g, Sigma) using a FastPrep-24™ 5G homogenizer (MP Biomedicals) at a speed of 6.5 m/s for 90 s. This step was followed by a 30-min incubation after the addition of 40 µL of lysozyme (Sigma). DNA was then extracted into a 50 µL eluate using the EZ1 DNA Tissue Kit (Qiagen) on the EZ1 Advanced XL automated workstation. gDNA was quantified using the Qubit dsDNA High Sensitivity Assay Kit (Thermo Fisher Scientific).

Sequencing was performed using MiSeq technology (Illumina Inc., San Diego, CA, USA) with a paired-end strategy. Libraries were prepared using the Nextera XT DNA Library Preparation Kit (Illumina). Briefly, gDNA was fragmented and tagged with adapters ('tagmented') before being amplified by limited-cycle PCR (12 cycles) to incorporate dual-index barcodes. The libraries were purified using AMPure XP beads (Beckman Coulter Inc., Fullerton, CA, USA) and normalized according to the Nextera XT protocol. Sequencing was conducted on a MiSeq instrument using the MiSeq Reagent Kit v2 (500 cycles), performing 2 × 250 bp paired-end reads.

To improve assembly quality, long-read sequencing was performed using GridION technology (Oxford Nanopore Technologies [ONT], UK). According to the manufacturer's instructions (NBE_9065_v109_revAP_14Aug2019), Nanopore libraries were prepared from 1000 ng of gDNA using the Native Barcoding Kit (EXP-NBD104/114) and the Ligation Sequencing Kit (SQK-LSK109). DNA repair and end-prep were performed using the NEBNext Ultra II End Repair/dA-Tailing Module and the NEBNext FFPE DNA Repair Mix (New England Biolabs [NEB]). Following purification and barcode ligation with the Blunt/TA Ligase Master Mix (NEB), adapters were ligated using the NEBNext Quick Ligation Module (NEB). After final purification, 12 µL of the library (maximum 430 ng) was loaded onto an R9.4.1 flow cell and sequenced on the GridION platform.

Genomic analysis

The quality of the raw sequencing data generated by the MiSeq platform was assessed using FastQC v0.11.9 (Andrews 2010). To improve overall read quality, low-quality bases were removed using Trimmomatic v0.39 (Bolger et al. 2014). For Oxford Nanopore Technologies (ONT) data, NanoPlot was utilized to monitor read quality and length distribution, while Filtlong (https://github.com/rrwick/Filtlong) facilitated the selection of high-quality ONT reads based on length and quality scores. Hybrid assembly of the filtered reads was performed using Unicycler v0.4.8 (Wick et al. 2017). Contig sequences shorter than 800 bp were removed to eliminate potential contaminants (Ndongo et al. 2020).

Genome annotation was conducted using Prokka v1.13 (Seemann 2014). To assign putative functions, predicted bacterial protein sequences were compared against the GenBank (Benson et al.

2018) and Clusters of Orthologous Groups (COG) (Galperin et al. 2015) databases using BLASTP (Altschul et al. 1990). Signal peptides were predicted using SignalP (Teufel et al. 2022), and transmembrane helices were detected with TMHMM (Krogh et al. 2001). Antibiotic resistance genes were screened using Abricate (https://github.com/tseemann/abricate), and CRISPR-Cas systems were identified via CRISPRCasFinder (Couvin et al. 2018).

For species delineation, digital DNA-DNA hybridization (dDDH) values were calculated using the Type (Strain) Genome Server (TYGS) online tool (Meier-Kolthoff et al. 2019), applying the standard 70% similarity threshold (Stackebrandt et al. 2002; Wayne et al. 1987). Furthermore, Average Nucleotide Identity (ANI) between genomes was estimated using Pyani to assess pairwise genomic relatedness between isolates (Pritchard et al. 2016). Finally, the genome was visualized using CGView (Grant et al. 2012). Taxonomic classification of genomes was conducted using GTDB-Tk (Genome Taxonomy Database Toolkit) (Chaumeil et al. 2019) with the GTDB database (release 226 (2026-02-25)), following default parameters (Parks et al. 2018).

Phylogenetic analysis

The phylogenetic distribution of the *Neobacillus* genomes in this study was assessed by aligning sequences with Mugsy version 1.2.3 (Angly et al. 2012). Evolutionary analyses were performed using IQ-TREE2 version 2.2.0 with the Maximum Likelihood method, applying the General Time Reversible (GTR) model (Minh et al. 2020). Bootstrap testing (1000 replicates) was used to evaluate branch support, and the percentage of replicate trees in which the associated taxa clustered together is indicated alongside the branches (Felsenstein 1985). The resulting phylogenetic tree was visualized using iTOL (Letunic and Bork 2019).

Strain analysis by integrated microbial NGS platform (IMNGS)

To assess the prevalence and abundance of strain Marseille-QA0830^T across diverse human and animal metagenomes, its 16S rRNA gene sequence was submitted to the Integrated Microbial Next-Generation Sequencing (IMNGS) platform (https://www.imngs.org; last accessed in December 2025). The screening parameters were set at a 99% similarity threshold for the 16S rRNA gene sequence, with a minimum requirement of 100 base pairs for the metagenomic sequences. Metagenomes in which the sequence was detected at least once were selected, and the relative frequency across human metagenomes was subsequently calculated (Lagkouvardos et al. 2016).

Results and discussion

Growth conditions and phenotypic and biochemical characteristics

Macroscopic observation during microbial diagnosis revealed well-isolated, circular, convex, and translucent colonies with regular margins, a smooth surface, and mucoid consistency. Microscopic examination after Gram staining showed that strain Marseille-QA0830^T consists of Gram-stain-positive, non-motile, non-spore-forming, rod-shaped cells with average dimensions of 5.4 µm in length and 2.9 µm in diameter (Fig. 1, Table 1). The strain is a facultative anaerobe, capable of growth under aerobic, microaerophilic (at 28 °C and 37 °C), and anaerobic conditions, with optimal growth observed at 37 °C under aerobic conditions on Columbia agar supplemented with 5% sheep blood (COS; bioMérieux, Marcy-l'Étoile, France). Strain Marseille-QA0830^T grew in media with a pH range of 5.0–7.5; however, no growth was observed in saline culture media (at concentration 5%, 7.5%, 10%, 15% and 20%). The isolate tested positive for catalase and negative for oxidase (Table 2).

Biochemical characterization using API 50 CHB strips revealed that strain Marseille-QA0830^T was unable to metabolize carbohydrates such as D-glucose, D-galactose, D-trehalose, gentiobiose, D-xylose, L-arabinose, D-mannose, arbutin, inulin, starch, and salicin. Furthermore, tests for

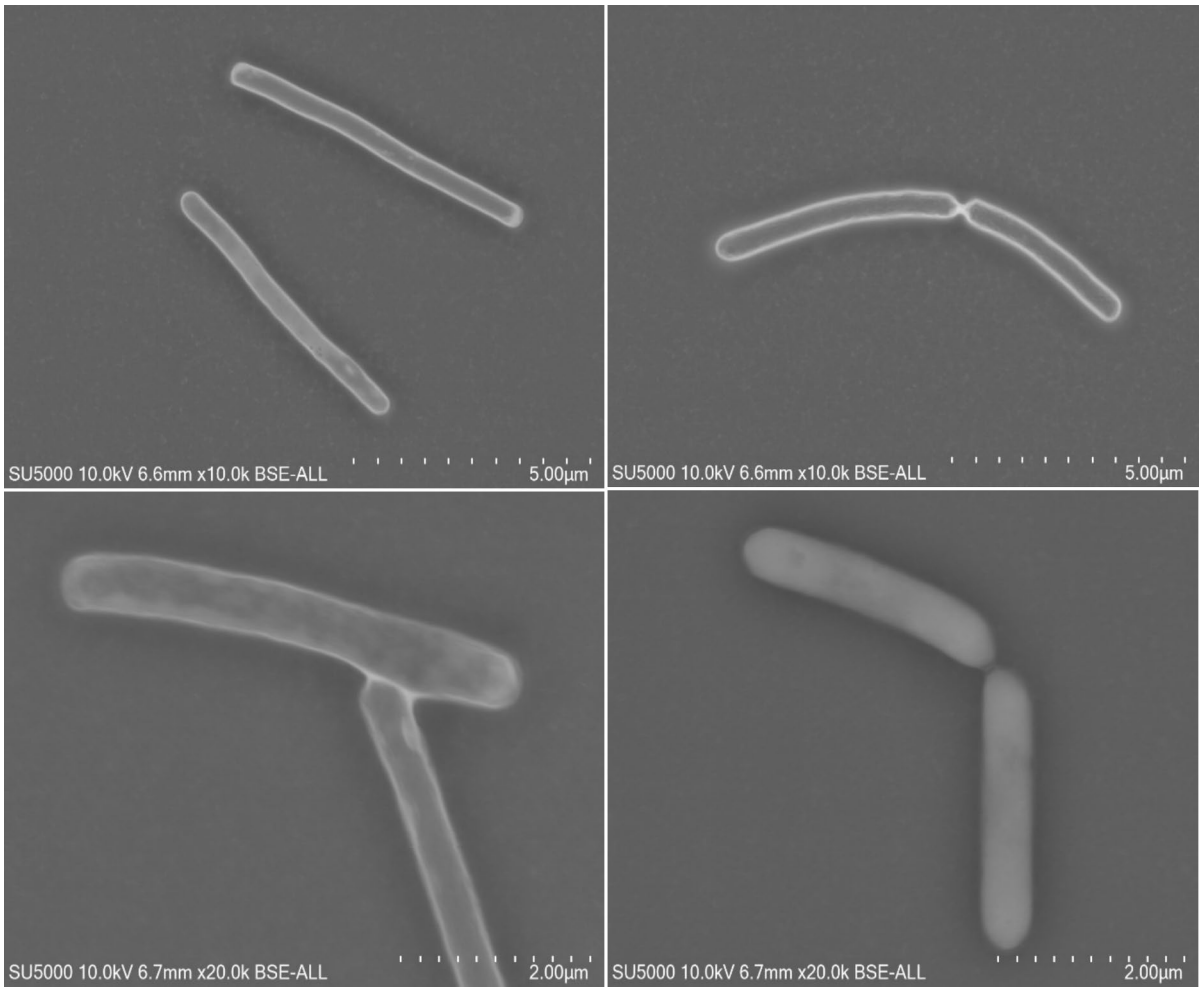

Fig. 1 Morphological structure of the *Neobacillus camarae* sp. Marseille-QA0830^T strain, revealed by the SU5000 scanning electron microscope (Hitachi High-Tech, Japan). Parametric details of the cells are shown in the figures

arginine dihydrolase, urease, and esculin hydrolysis (β-glucosidase) were negative (Table 2). This restricted metabolic profile reflects the low enzymatic activity characteristic of this genus, as previously documented in the literature (Lo et al. 2015; Mbaye et al. 2021). Regarding the API ZYM strips, positive enzymatic reactions were observed for alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, acid phosphatase, α-glucosidase, β-glucosidase, β-galactosidase, β-glucuronidase, and naphthol-AS-BI-phosphohydrolase (Table 3).

Fatty acids analysis

The cellular fatty acid composition analysis revealed that branched-chain fatty acids (BCFAs) accounted for more than 80% of the total composition. The predominant structures identified were the anteiso and iso forms of pentadecanoic acid ($C_{15:0}$), representing 28.0% and 20.4% of the total, respectively. Hexadecanoic acid ($C_{16:0}$) was the primary unbranched saturated fatty acid identified, at 12.4%. Additionally, trace amounts of other saturated and unsaturated fatty acids were detected (Table 4).

Table 1 Characteristic between (1) *Neobacillus camarae* sp. nov. Marseille-QA0830^T and strains of closely related (2) *Neobacillus dielmonensis* sp. nov. DSM 27844 FF4^T; *Neobacillus niacini* strain NBRC 15566^T

Characteristic	(1)	(2)	(3)
Cell shape	Rod-shaped	Rod-shaped	Rod-shaped
Cell diameter (µm)	3.8–6.7	2.6–5.8	2.0–4.0
Color	Translucent	White	White cream
Oxygen requirement	Aerobic	Aerobic	Aerobic
Gram stain	+	+	+
Flagellation	−	+	+
Endospore formation	−	−	+
Growth at/on			
25°	+	−	+
28°	+	−	+
37°	+	+	+
45°	+	+	+
56°	−	−	−
pH 5	−	−	−
pH 6	+	−	−
pH 7	+	+	+
pH 8,5	−	−	−
NaCl 5%	−	−	−
NaCl 7,5%	−	NA	−
NaCl 10%	−	NA	−
NaCl 15%	−	NA	−
NaCl 20%	−	NA	−
Pathogenicity	Unknown	Unknown	Unknown
Habitat	Milk	Human skin	Soil

Data taken from description of *Neobacillus dielmonensis* sp. nov

DSM 27844 FF4^T (Lo et al. 2015)

Nagel and Andreesen (1991); Bergey's Manual of Systematic Bacteriology

(+) correspond to positive reaction, (−) correspond to negative reaction, (NA) correspond to data not available

Table 2 Biochemical tests performed observed between (1) strains *Neobacillus camarae* sp. Marseille-QA0830^T and strains of closely related (2) *Neobacillus dielmonensis* sp. nov. DSM 27844 FF4^T; *Neobacillus niacini* strain NBRC 15566^T

Phenotypic characteristics (API 50 CHB)	1	2	3
Utilisation of			
D-glucose	–	–	NA
D-xylose	–	–	NA
L-arabinose	–	–	NA
D-galactose	–	–	NA
D-mannose	–	–	NA
Arbutin	–	–	NA
Citrate	–	–	NA
Salicin	–	–	NA
D-tréhalose	–	–	NA
Amidon	–	–	NA
Gentiobiose	–	–	NA
Production of			
Dihydrolase arginine	–	–	NA
N-acetyl-D-glucosamine	–	–	NA
Urease	–	–	NA
D-fructose	–	–	NA
Maltose	–	–	NA
Orthinine décarboxylase	–	–	NA
Lysine decarboxylase	–	–	NA
Hydrogen Sulfide	–	–	NA
Tryptophan desaminase	–	–	NA
Indole	–	–	NA
D-arabinose	–	–	NA
D-arabitol	–	–	NA
L-arabitol	–	–	NA
Dulcitol	–	–	NA
erythritol	–	–	NA
L-xylose	–	–	NA
L-sorbose	–	–	NA
Xylitol	–	–	NA
2-keto-D-gluconate	–	–	NA
D-tagatose	–	–	NA
Methyl D-xyloside	–	–	NA
Hydrolyse β-glucosidase (ESC)	–	+	NA

(+) means positive reaction; (−) means negative reaction; NA means data not available

Data taken from description of *Neobacillus dielmonensis* sp. nov. DSM 27844 FF4^T (Lo and al. 2015)

Antibiotic tests analysis

The antibiotic susceptibility profiles for strain Marseille-QA0830^T, determined using the disk diffusion method (Gaur et al. 2023) in accordance with EUCAST 2025 guidelines, are summarized in Table 5. In the absence of specific clinical breakpoints for the recently described genus *Neobacillus*, we compared our MIC values against the recommendations applicable to *Bacillus* spp. (CLSI M45). However, these results should be interpreted with caution, establishing reliable clinical breakpoints and distinguishing between intrinsic and acquired resistance in the Marseille-QA0830^T

Table 3 Comparison of enzymatic diversity between strain *Neobacillus camarae* sp. Marseille-QA0830^T (1) and strains of closely related (2) *Neobacillus dielmonensis* DSM 27844 FF4^T and *Neobacillus niacini* NBRC 15566^T

Enzyme activity using API ZYM	(1)	(2)	(3)
Catalase	+	+	+
Cytochrome oxydase	–	–	+
Phosphatase alcaline	+	+	NA
Esterase	+	+	NA
Esterase lipase	+	+	NA
Lipase	–	–	NA
Leucine arylamidase	+	–	NA
Valine arylamidase	+	–	NA
Cystine arylamidase	–	–	NA
Trypsine	–	–	NA
α-chymotrypsine	–	–	NA
Phosphatase acid	+	+	NA
Naphtol-AS-BI-phosphohydrolase	+	+	NA
α-galactosidase	–	–	NA
β-galactosidase	+	+	NA
β-glucuronidase	+	+	NA
α-glucosidase	+	+	NA
β-glucosidase	+	+	NA
N-acétyl-β-glucosaminidase	–	–	NA
α-mannosidase	–	–	NA
α-fucosidase	–	–	NA

+means positive reaction;– means negative reaction; NA means data Not Available

Data taken from description of *Neobacillus dielmonensis* sp. DSM 27844 FF4^T (Lo and al. 2015)

Table 4 Analysis of cellular fatty acid components (%) in the (1) *Neobacillus camarae* sp. Marseille-QA0830^T strain compared to the closest species related (2) *Neobacillus dielmonensis* DSM 27844 FF4^T

Fatty acid	Names	*Mean relative %* [a]	
		1	*2*
$C_{15:0\ anteiso}$	12-methyl-tetradecanoic acid	28.0 ± 0.2	24.0 ± 0.2
$C_{15:0\ iso}$	13-methyl-tetradecanoic acid	20.4 ± 2.9	51.2 ± 1.0
$C_{14:0\ iso}$	12-methyl-tridecanoic acid	18.4 ± 0.4	2.9 ± 0.1
$C_{16:0\ iso}$	14-methyl-pentadecanoic acid	14.1 ± 0.9	3.7 ± 0.4
$C_{16:0}$	Hexadecanoic acid	12.4 ± 1.8	4.5 ± 0.2
$C_{5:0\ iso}$	3-methyl-Butanoic acid	-	3.0 ± 0.2
$C_{18:0}$	Octadecanoic acid	1.7 ± 0.1	1.5 ± 0.2
$C_{17:0\ anteiso}$	14-methyl-hexadecanoic acid	1.3 ± 0.1	TR
$C_{18:2n6}$	9,12-Octadecadienoic acid	TR	TR
$C_{18:1n12}$	6-Octadecenoic acid	TR	1.3 ± 0.1
$C_{17:0\ iso}$	15-methyl-hexadecanoic acid	TR	2.5 ± 0.1
$C_{14:0}$	Tetradecanoic acid	TR	TR
$C_{17:0}$	Heptadecanoic acid	TR	TR
$C_{16:0\ anteiso}$	13-methyl-pentadecanoic acid	TR	–
$C_{15:0}$	Pentadecanoic acid	TR	TR

[a]Mean peak area percentage, TR = trace amounts < 1%, All data taken from cellular fatty acid methyl ester (FAMEs) analysis of this study

strain will require analysis of a larger cohort in future studies.

Strain identification

Following multiple unsuccessful identification attempts using MALDI-TOF MS, which yielded scores consistently below 1.39, assuming that the spectrum of the isolate was determined to be absent from the reference database, suggesting a potentially novel or undocumented species. Consequently, whole-genome sequencing of strain Marseille-QA0830^T was performed to establish its definitive taxonomic status.

First, the 16S ribosomal RNA (rRNA) gene sequence was extracted and analyzed. A BlastN search against the GenBank database revealed that strain Marseille-QA0830^T exhibits a 98.10% sequence similarity to *Neobacillus niacini* NBRC 15566^T (NR_113777.1) (Table 6). Phylogenetic analysis based on the 16S rRNA confirmed that the isolate clusters robustly within the genus *Neobacillus*, distinct from its closest validly published neighbors, with *Cytobacillus formosensis* CC-LY275^T serving as a stable outgroup (Fig. 2).

To complete the 16S rRNA analysis, which can exhibit high sequence conservation within the *Bacillaceae*, an evaluation of the *rpoB* gene sequence was carried out as a secondary phylogenetic marker. The *rpoB* gene sequence analysis showed a low similarity of 83% with *Neobacillus drentensis,* further verifying a significant evolutionary distance at the species level (Table 6).

Table 5 Antibiotic susceptibility profile of *Neobacillus camarae* sp. strain Marseille-QA0830^T

Antibiotics (µg/mL)	*Neobacillus camarae* sp. QA0830^T	CLSI M45 Breakpoint* (S) ≤	CLSI M45 Breakpoint* (R) ≥
Amikacin (AK)	0.5	8	32
Amoxicillin (AC)	0.064	0.25	0.5
Benzylpenicillin (PG)	0.032	0.12	0.25
Ceftazidim (TZ)	> 256	8	32
Ceftriaxon (TX)	0.5	8	64
Ciprofloxacin (CI)	0.064	1	4
Clindamycin (CM)	< 0,002	0.5	4
Daptomycin (DPC)	0.125	1	–
Doxycyclin (DC)	< 0,004	4	16
Gentamycin (GM)	< 0,0016	4	16
Imipenem (IP)	< 0,05	4	16
Linezolid (LZ)	0.125	4	–
Nitrofurantoin (NI)	0.25	32	128
Oxacillin (OX)	0.125	2	4
Rifampicin (RI)	0.32	1	4
Teicoplanin (TP)	< 0,001	8	32
Tobramycin (TM)	> 256	4	16
Trimethoprimsulmétoxazole (TS)	0.006	2/38	4/76
Vancomycin (VA)	> 256	4	–

*Reference ranges are based on CLSI M45 criteria for *Bacillus* spp. Other than *B. anthracis*. These values are provided for informative purposes only and should not be considered formal S/I/R clinical breakpoints for genus *Neobacillus*

Table 6 Summary of 16S rRNA Similarity Results and LPSN Validation

Strains	Analysis	Best BLASTN hit	% Identity	% Coverage	Accession number	LPSN Status
Neobacillus camarae sp. Marseille_QA0830^T	16S rRNA	*Neobacillus niacini* CIP 104585^T(= IFO 15556^T)	98.10	100	NR_024695	Valid name
Neobacillus camarae sp. Marseille_QA0830^T	16S rRNA	*Neobacillus dielmonensis DSM* 27844 FF4^T	97.33	100	NR_178429	Valid name
Neobacillus camarae sp. Marseille_QA0830^T	*rpoB*	*Neobacillus drentensis* LMG 21831^T	83.53	100	CP040523.1	Valid name
Neobacillus camarae sp. Marseille_QA0830^T	dDDH	*Neobacillus dielmonensis* DSM 27844 FF4^T	26.2	NA	NR_178429	Valid name

**rpoB N. drentensis* sequence was extracted from the whole-genome assembly

Genomic comparison

The genome of strain Marseille-QA0830^T is 5.83 Mbp long with a G + C content of 42.35% (Table S1). The genome comprises a total of 5451 genes, including 136 tRNAs and 44 rRNAs (Table S1). To resolve the taxonomic status of the isolate within the genus *Neobacillus*, comprehensive pairwise genomic comparisons were performed against twenty-four closely related validly published strains.

The 16S rRNA gene sequence analysis showed a high similarity of 98.10% between strain Marseille-QA0830^T and *Neobacillus niacini* strain NBRC 15566^T (Table 6). At the same time, the resolution using the *rpoB* marker revealed a significantly lower similarity of 83.53% with *Neobacillis drentensis* LMG 21831^T highlighting a phylogenetically distant relationship (Table S3). The divergence was definitively confirmed at the whole-genome level using Overall Genome Relatedness Indices (OGRI). Genomic comparisons revealed a digital DNA-DNA

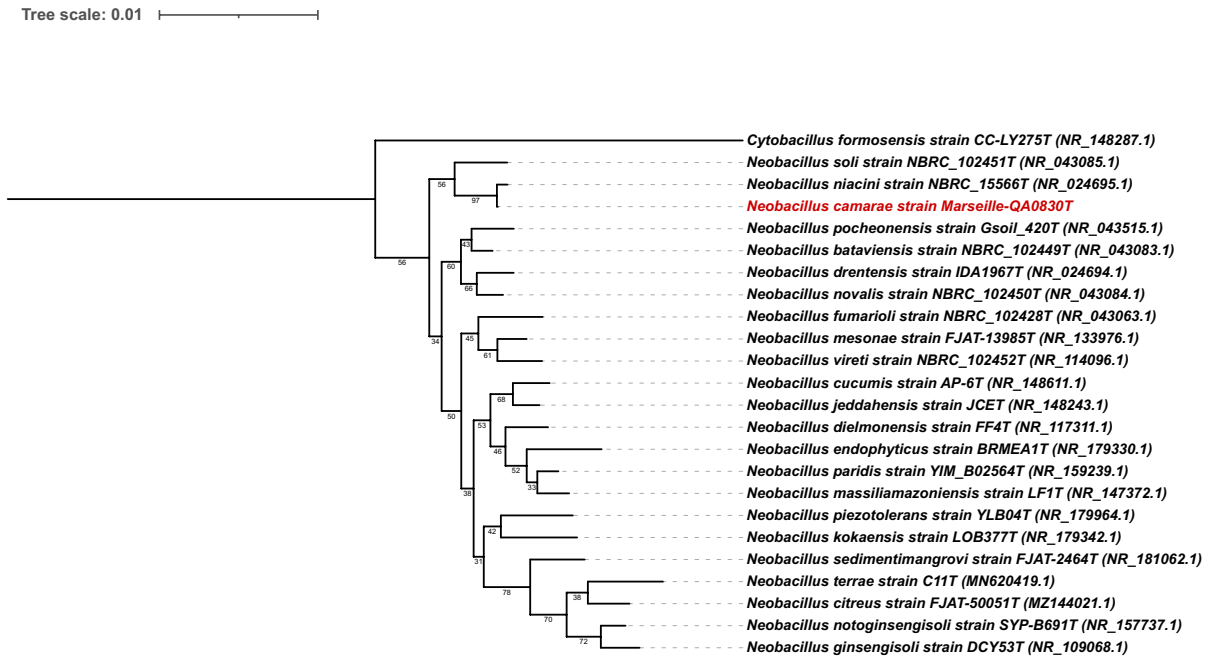

Fig. 2 Phylogenetic tree based on 16S rRNA gene sequences showing the position of *Neobacillus camarae* sp. Marseille-QA0830^T within the genus *Neobacillus*. GenBank access numbers for 16S rRNA were in brackets

hybridization (dDDH) value of only 26.2% (Fig. 3) and an Average Nucleotide Identity (ANI) value of 83.0% (Fig. 4) against its closest genomic neighbor, *Neobacillus dielmonensis* FF4^T. Both values fall drastically short of the universally recognized species delineation thresholds (70% for dDDH and 95–96%

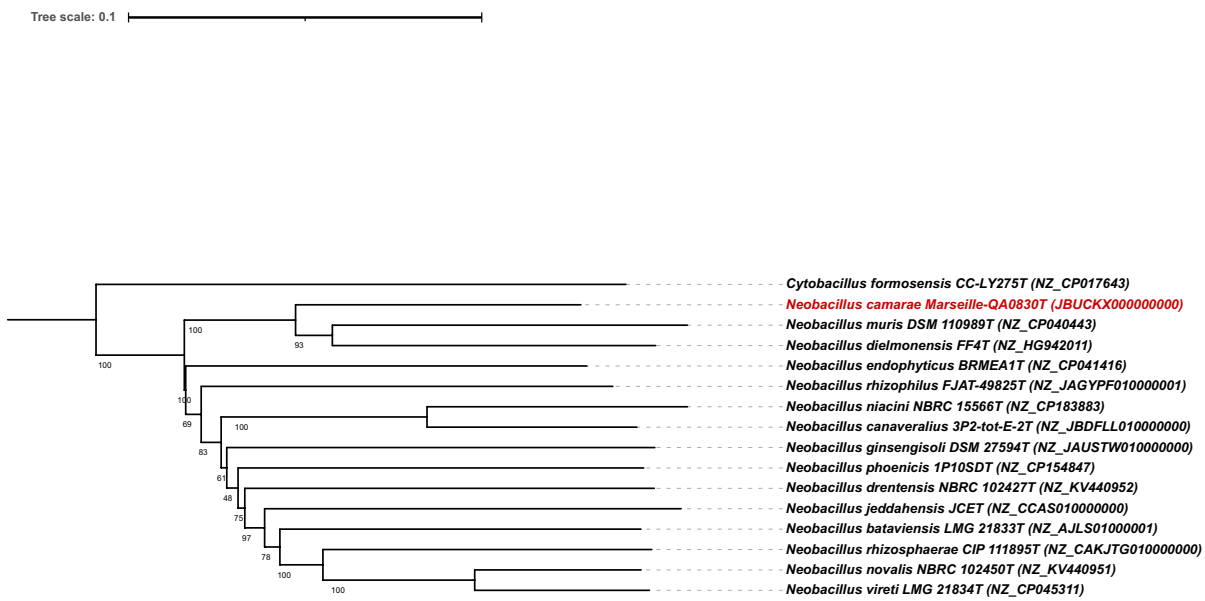

Fig. 3 Maximum-likelihood phylogenomic tree based on core genome sequences, illustrating the relationship between strain *Neobacillus camarae* sp. Marseille-QA0830^T and related species of the genus *Neobacillus*. GenBank access numbers were in brackets

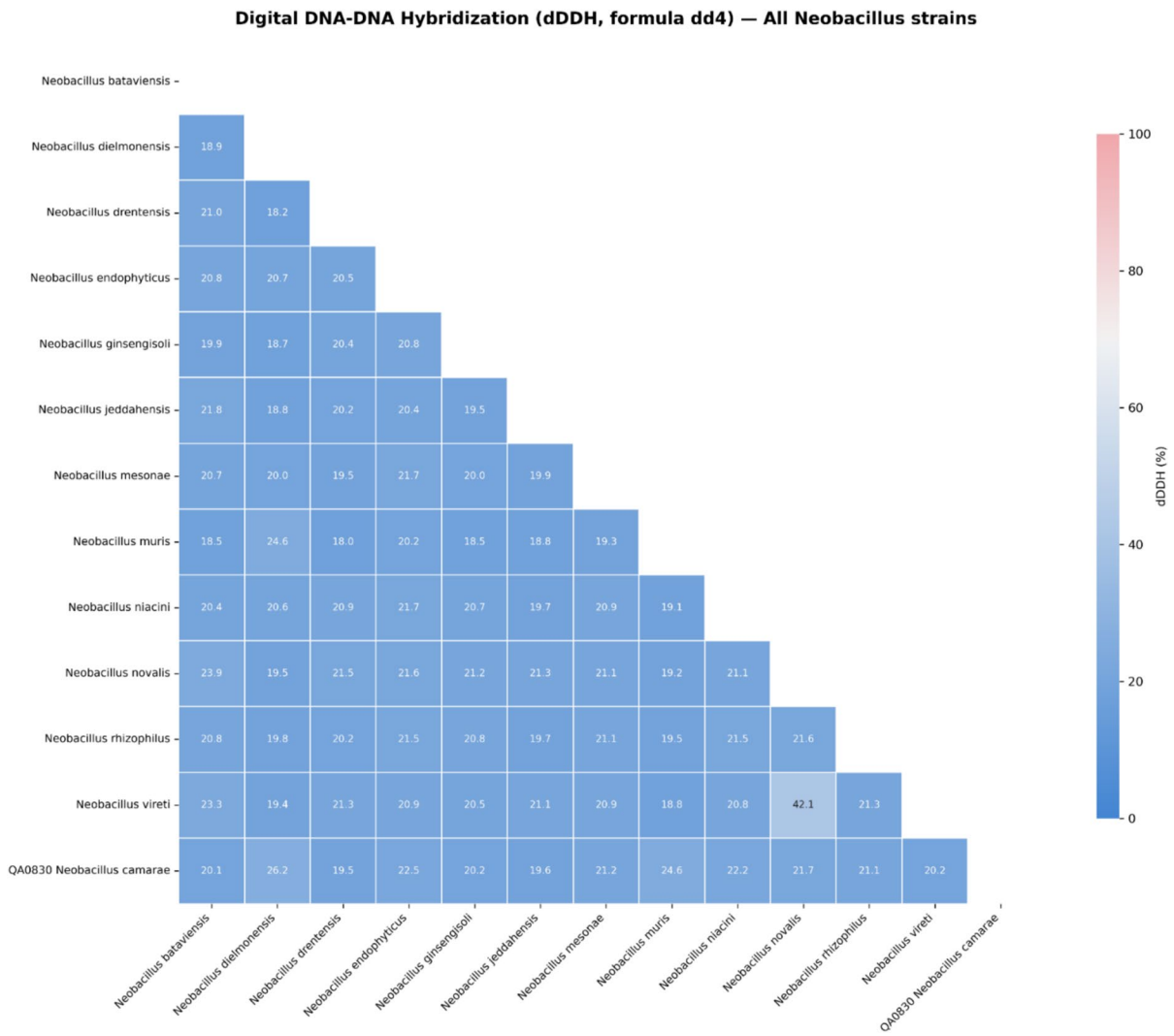

Fig. 4 Heatmap of digital DNA–DNA hybridization (dDDH, formula d4) among *Neobacillus camarae* sp. nov. Marseille-QA0830^T, and closely *Neobacillus* reference species

for ANI), confirming that strain Marseille-QA0830^T represents a genomic distinct novel taxon.

The functional distribution of Coding Sequences (CDS) into clusters of Orthologous Groups (COG) categories was analyzed to evaluate specific ecological adaptations (Fig. 5**,** Table S3). Strain Marseille-QA0830^T exhibits a robust metabolic backbone, particularly enriched in amino acid transport and metabolism [**E**] (9.2%) and carbohydrate metabolism [**G**] (8.5%). Interestingly, comparative functional analysis with its closest phylogenetic neighbours demonstrated a distinct enrichment in transcription and regulatory mechanisms [**K**], representing 7.95% of its genome compared to 6.51% in *N. drentensis* LMG 21831^T. This functional divergence, although subtle, supports its ability to adapt to fluctuating environmental or hort-associated conditions and correlates with its unique specialized phenotypic profile (Fig. 6).

IMNGS analysis of the strain *Neobacillus camarae* sp. nov. Marseille-QA0830^T

IMNGS analysis demonstrated the presence of the isolate across various human metagenomes.

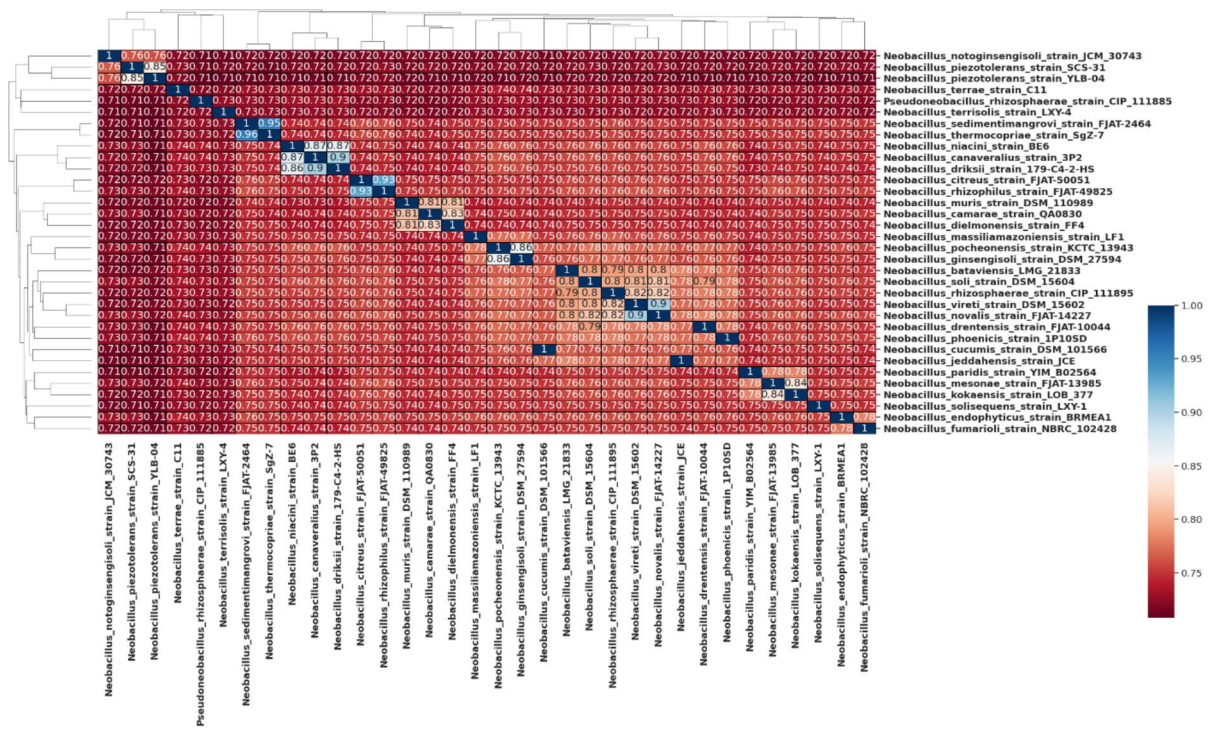

Fig. 5 Comparative heatmap of Average Nucleotide Identity (ANI) among *Neobacillus* strains and closely reference species

Specifically, detection frequencies were 0.054% in human-associated samples (3,222/6,012,454), 0.053% in skin (480/911,950), 0.01% in the oral cavity (48/499,706), 0.11% in the gut (6,991/6,088,359), 0.013% in the nasopharynx (498/3,957,055), 0.14% in milk (592/416,515), and 0.14% in the lungs (Table S4). The presence of strain Marseille-QA0830^T in these diverse metagenomes suggests a potential, albeit minimal, role in the human microbiome. Furthermore, these data indicate a possible vertical transmission route or dissemination from the oral cavity or the mammary glands to the infant's digestive tract (Rodriguez 2014). However, a geographical specificity with higher prevalence in Senegal and West Africa can't be excluded.

Discussion

The phenotypic, phylogenetic, and taxonomic analyses conducted on strain Marseille-QA0830^T allow for its clear differentiation from all other validly published species within the genus *Neobacillus*. Isolating this novel taxon from a human sample in Senegal is highly consistent with the local history of microbial culturomics. Notably, the very first novel species described via this high-throughput method was a member of the *Bacillaceae* family (*Oceanobacillus massiliensis*) discovered in the fecal flora of a healthy individual from the Dielmo and N'Diop regions (Roux et al. 2013), close to where *Neobacillus dielmonensis* was later isolated (Lo et al 2015). Recent biogeographical cohorts confirm a distinct bacterial "fingerprint" or regional enrichment of *Bacillus*-related genera in human breast milk and gut samples across West African regions, including Senegal and the Gambia, when compared to European or American cohorts (Lackey et al. 2019; Sun et al. 2025).

Beyond its biogeographical interest, a striking feature of strain Marseille-QA0830^T is its highly restricted biochemical and carbohydrate-fermentation profile. While the genus type species *Neobacillus niacini* NBRC 15566^T is also known to be relatively asaccharolytic, strain Marseille-QA0830^T is entirely inert across the vast majority of standard API 50 CHB substrates, including key structural hexoses like D-glucose and D-galactose. This phenotype strongly correlates with its genomic infrastructure.

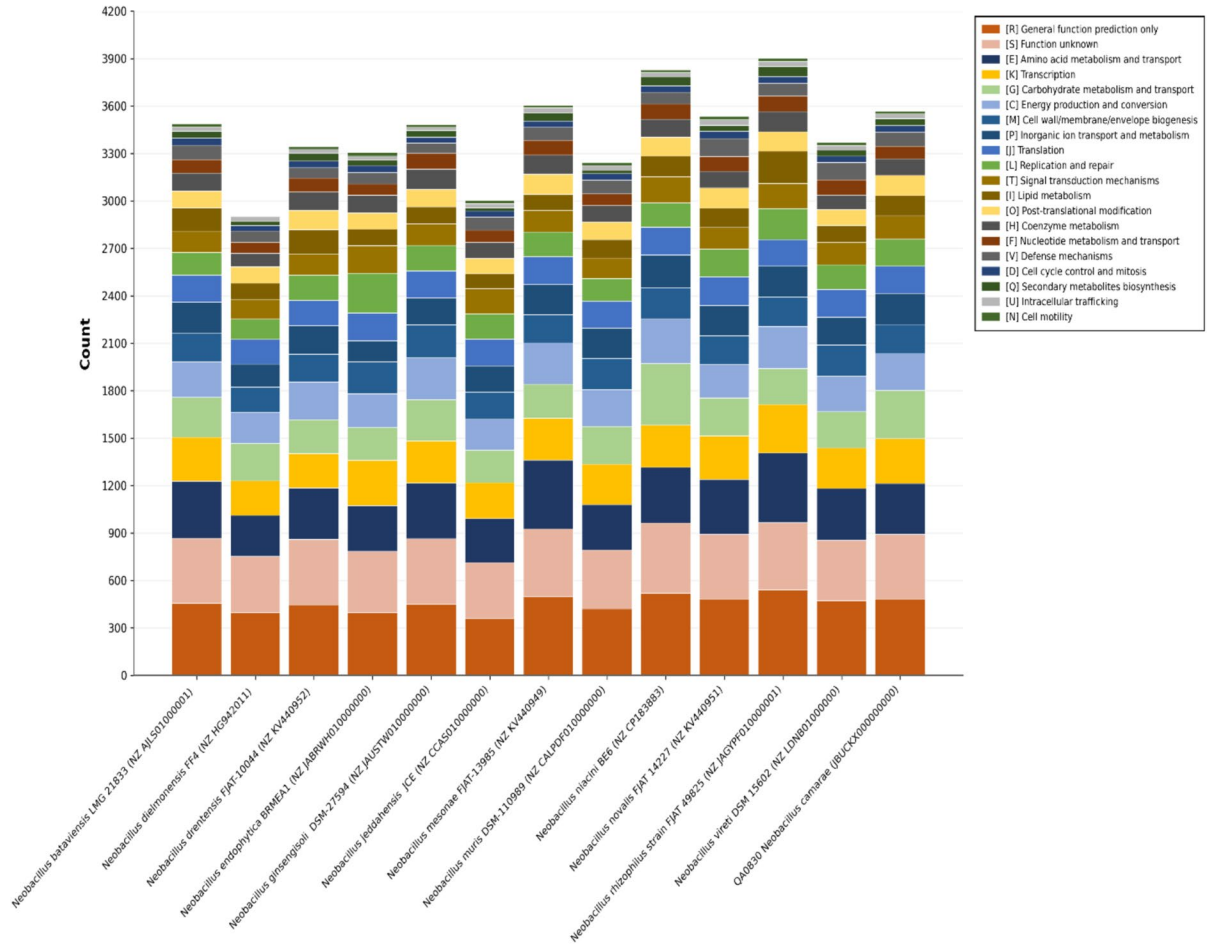

Fig. 6 Functional distribution of coding sequences (CDSs) among COG categories in *Neobacillus camarae* sp. Marseille-QA0830^T and closely strains

Genus-wide genomic comparative metrics (ANI at 83.0% and dDDH at 26.2% with *N. dielmonensis*) reveal a significant reduction or divergence in standard carbohydrate transport systems. This contrasts with its large genome (5.8 Mbp). Instead, the survival of strain Marseille-QA0830^T in highly competitive niches like human milk or the gut likely relies on alternative metabolic strategies, such as the catabolism of specific amino acids or organic acids, which is a common trait among specialized endospore-forming *Bacillaceae*.

From an ecological and biotechnological standpoint, members of the genus *Neobacillus* (formerly part of the *Bacillus* rRNA group 5) are well-documented for their exceptional environmental robustess and enzymatic potential. Many *Neobacillus* species harbor specialized genetic pathways for the degradation of recalcitrant aromatic compounds, xenobiotics, or complex environmental polymers, alongside the secretion of highly stable extracellular hydrolases. Although strain Marseille-QA0830^T displayed a restricted response to stress, oxidase stress tolerance, and nutrient scavenging. Given that related *Neobacillus* lineages are actively screened for the production of thermostable enzymes or secondary metabolites with antimicrobial properties, the isolation and genomic sequencing of *Neobacillus camarae* strain Marseille-QA0830^T opens insteresting avenues. Future functional multi-omics studies under specific inductive substrates will be essenstial to fully resolve its accurate ecological

role within the human-associated microbiota and to unveil its latent biotechnological applications.

Conclusion

In this study, a comprehensive polyphasic taxonomic approach was successfully employed to define the evolutionary and taxonomic status of strain Marseille-QA0830^T. The integration of single-gene phylogeny, core-genome phylogenomics, and comparative phenotypic profiles provides unequivocal evidence that this isolate represents a distinct maternal lineage within the genus *Neobacillus*. While 16S rRNA gene sequence analysis firmly anchored the strain within the *Neobacillus* cluster, exhibiting a 98.10% similarity to the genus type *Neobacillus niacini* NBRC 15566^T whole-genome relatedness indices provided definitive resolution regarding its status as a novel taxon. The dDDH value (26.2%) and ANI score (83.0%) obtained against its closest genomic neighbor, *Neobacillus dielmonensis* FF4^T, stand drastically below the universally accepted thresholds for bacterial species delineation (70% and 95–96% respectively).

Crucially, this clear genomic divergence is perfectly mirrored by unique physiological and biochemical traits, most notably the strain's lack of motility, lack of endospore formation under standard conditions, and an unreactive carbohydrate profile that clearly distinguishes it from both its phylogenomic neighbors and the genus type species.

In conclusion, the robust congruence between phenotypic limitations, single-gene markers (*rpoB*), and definitive whole-genome restrictions fully justify the circumscription of this isolate as a novel species. Consequently, we formally propose the name *Neobacillus camarae* sp. nov., with Marseille-QA0830^T (=CSUR QA0830^T=CECT 31295^T) as the designated type strain.

Description of ***Neobacillus camarae*** sp. nov. Marseille-QA0830^T

Named *Neobacillus camarae* (ca.ma'rae. N.L. gen. n. camarae, in honor of Professor Makhtar Camara, for his significant contributions to microbiology and to the study of human microbiota and antimicrobial resistance surveillance in Senegal), cells of the type strain are Gram-stain-positive, non-motile, and non-spore-forming rods, measuring approximately 5.4 µm in length and 2.9 µm in diameter. Flagella are absent. It is a facultative anaerobe, capable of growth on Columbia agar supplemented with 5% sheep blood under aerobic, microaerophilic, and anaerobic conditions. Growth occurs between room temperature and 45 °C, with an optimal temperature of 37 °C after 72 h of incubation. Colonies are well-isolated, circular, convex, and translucent, characterised by regular margins, a smooth surface, and a mucoid consistency. The strain grows at pH levels ranging from 5.5 to 7.5 but shows no tolerance for NaCl.

Using the API 50 CHB system, the strain does not ferment carbohydrates, including D-glucose, D-mannose, D-trehalose, D-xylose, D-galactose, L-arabinose, arbutin, salicin, inulin, starch, or gentiobiose. Tests for arginine dihydrolase, urease, and esculin hydrolysis (β-glucosidase) are negative. However, high enzymatic diversity is observed via API ZYM, with positive results for esterase (C4), esterase lipase (C8), leucine arylamidase, naphthol-AS-BI-phosphohydrolase, and trypsin. It is catalase-positive and oxidase-negative. The cell wall fatty acid composition is dominated by branched-chain structures (> 80%), primarily anteiso-$C_{15:0}$ (28.0%) and iso- $C_{15:0}$ (20.4%). The major unbranched structure is hexadecanoic acid ($C_{16:0}$, 12.4%). The type strain, *Neobacillus camarae* sp. nov. Marseille-QA0830^T (=CECT 31295^T), was isolated from the breast milk of a healthy breastfeeding mother in Senegal. The 16S rRNA gene and whole-genome sequences are deposited in GenBank under accession numbers (JBUCKX000000000).

Acknowledgements The authors would like to thank Amael Fadlane and Dion Del Rio from Collection des Souches de l'Unité des Richettsies, (CSUR) for their availability in response to numerous requests during the completion of this descriptive work.

Author contributions All authors participated in the design and implementation of the study. M.M. and C. S. have designed the study, S.S. and O.N. have collected samples, A.M. and M.B. carried out the genomic part of this article, S.S. and A.I.A. have performed phenotypic description tests, S.B. carried out all the microscopic work, S.A. and C.V. coordinated all administrative aspects of obtaining second collection numbers and verifying the etymology of strain names,R. S. carried out the IMNGS analysis of strain,N.O. analysed the fatty acids in the cell walls of strain, M.M. have coordinated the financial acquisition of this study, S.S. and A.M. have realised investigation, M.T.A. monitored the descriptive methodology of new species, M.M. and C.S. have coordinated the research project, C.S. and G.D. participated in the drafting of the protocol

and in obtaining ethical approval for this study, C.T. authorised the collection of samples at his health center, M.M., M.T.A., C.S. have supervised the work, S.S. and A.M. have validated the draft, wrote the original draft; and viewed the draft, M.M. and S.S. and did the writing, review and editing of the manuscrit. All authors have read and approved the final version of the manuscript.

Funding This work was supported by a grant from the French Government managed by the National Research Agency under the "Investissements d'avenir (Investments for the Future)" programme with the reference ANR-10-IAHU-03 (Méditerranée Infection), by the Contrat Plan Etat-Région and the European funding FEDER IHUPERF.

Data availability All data relating to this study are available from the corresponding author. The strain have been deposited in the microorganism collection of the IHU Méditerranée Infection. (Collection des Souches de l'Unité des Richettsies, CSUR [https://csur.eu/] (https:/csur.eu)) under number CSUR QA0830^{T} and in the Spanish type culture collection "Colección Española de Cultivos Tipo" (CECT [https://www.uv.es/cect] (https:/www.uv.es/cect)) under the number CECT 31295^{T}. The genome sequences have been deposited in Genbank under accession number (ongoing).

Declarations

Conflict of interest The authors declare no conflict of interest.

Ethical approval This collection was initiated as part of the study of the microbiota of breast milk in malnutrition following the granting of administrative authorisation from the Ethics and Scientific Committee of the Senegalese Ministry of Health and Social Action (MSAS) under no. 00000233/MSAS/CNERS/SP on 11 September 2022.

References

Altschul SF, Gish W, Miller W et al (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. babraham bioinformatics. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Angly C, Chan JSM, Henshaw CDG et al (2012) Mugsy: a multiple genome alignment of microbial genomes. Bioinformatics 28(18):2477–2483. https://doi.org/10.1093/bioinformatics/bts491

Ash C, Farrow AE, Wallbanks S, Collins MD (1991) Phylogenetic heterogeneity of the genus *Bacillus* revealed by comparative analysis of small subunit ribosomal RNA sequences. Lett Appl Microbiol 13(4):202–206

Atlas RM, Snyder JW (2011) Reagents, stains and media: bacteriology*. Human Clin Microbiol 10:272–303. https://doi.org/10.1128/9781555816728.ch17

Benson DA, Cavanaugh M, Clark K et al (2018) GenBank. Nucleic Acids Res 46(D1):D41–D47. https://doi.org/10.1093/nar/gkx1094

Berkeley RCW, Logan NA, Shute LA, Capey AG (1984) Identification of *Bacillus* species. Methods Microbiol 16:291–328

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bottone EJ (2010) *Bacillus cereus*, a volatile human pathogen. Clin Microb Rev 23:382–398. https://doi.org/10.1128/CMR.00073-09

Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH (2019) GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics 36(6):1925–1927. https://doi.org/10.1093/bioinformatics/btz848

Clinical and Laboratory Standards Institute (CLSI) (2018) Performance standards for antimicrobial susceptibility testing, 28th edition (CLSI supplement M100). Clinical and Laboratory Standards Institute, Wayne, PA

Cohn F (1875) Untersuchungen über Bakterien. In: Cohn F (ed) Beiträge zur Biologie der Pflanzen 1 (Heft 2), 1872, Max Müller, Breslau, pp 127–224

Couvin D, Bernheim A, Toffano-Nioche C et al (2018) CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res 46(W1):W246–W251. https://doi.org/10.1093/nar/gky425

De Coster W, D'Hert S, Schultz DT et al (2018) NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34(15):2666–2669. https://doi.org/10.1093/bioinformatics/bty149

Dione N, Sankar SA, Lagier JC et al (2016) Genome sequence and description of *Anaerosalibacter massiliensis* sp. nov. New Microbes Nex Infect 10:66–76

Ensign JC (1963) PhD thesis. University of Southern California; Los Angeles

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39(4):783–791. https://doi.org/10.2307/2408678

Galperin MY, Makarova KS, Wolf YI, Koonin EV (2015) Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res 43(Database issue):D261–D269

Gaur P, Hada V, Rath RS, Mohanty A, Singh P, Rukadikar A (2023) Interpretation of antimicrobial susceptibility testing using european committee on antimicrobial susceptibility testing (EUCAST) and clinical and laboratory standards institute (CLSI) breakpoints analysis of agreement. Cureus 15(3):e36977

Grant JR, Arantes AS, Stothard, P (2012) Comparing thousands of circular genomes using the CGView Comparison Tool. BMC Genomics 13 202. https://doi.org/10.1186/1471-2164-13-202

Hendrickson DA (1985) Reagents and stains. 1093–1107. In: Lenette EH, Balows A, Hausler WJ, Shadomy HJ (eds)

Manual of clinical microbiology, 4th. American Society for Microbiology, Washington D.C., pp 1093–1107

Heyrman J, Vanparys B, Logan NA, Balcaen A, Rodriguez-Diaz M, Felske A et al (2004) *Bacillus novalis* sp. nov., *Bacillus vireti* sp. nov., *Bacillus soli* sp. nov. *Bacillus bataviensis* sp. nov. and *Bacillus drentensis* sp. nov., from the Drentse A grasslands. Int J Syst Evol Microbiol 54(1):47–57

Koehler TM (2009) *Bacillus anthracis* physiology and genetics. Mol Aspects Med 30:386–396. https://doi.org/10.1016/j.mam.2009.07.004

Krogh A, Larsson B, von Heijne G, Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305(3):567–580. https://doi.org/10.1006/jmbi.2000.4315

Lackey KA, Williams JE, Meehan CL, Zachek JA, Benda ED, Price WJ et al (2019) What's Normal? Microbiomes in human milk and infant feces are related to each other but vary geographically: the INSPIRE study. Front Nutr 2019(6):45. https://doi.org/10.3389/fnut.2019.00045

Lagkouvardos I, Joseph D, Kapfhammer M (2016) IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. Sci Rep. 6:33721. Published 2016 Sep 23. https://doi.org/10.1038/srep33721

Letunic I, Bork P (2019) Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 47(W1):W256–W259. https://doi.org/10.1093/nar/gkz239

Lo CI, Padhmanabhan R, Mediannikov O et al (2015) High-quality genome sequencing and description of *Bacillus dielmonensis* strain FF4 T sp. nov. Stand in Genomic Sci 10:41

Logan NA (2012) *Bacillus* and relatives in foodborne diseases. J Appl Microbiol 112:417–429. https://doi.org/10.1111/j.1365-2672.2011.05204.x

Logan NA, De Vos P (2009) Bergey's manual of systematic bacteriology. In: De Vos P, Garrity M, Jones D et al (eds) Springer, New York, pp 21–128

Matuschek E, Åhman J, Kahlmeter G, Yagupsky P (2018) Antimicrobial susceptibility testing of Kingella kingae with broth microdilution and disk diffusion using EUCAST recommended media. Clin Microbiol Infect 24:396–401. https://doi.org/10.1016/j.cmi.2017.07.019

Mbaye B, Tidjani Alou M, Fadlane A, Fregiere L, Alibar S, Million M et al (2021) *Neobacillus massiliamazoniensis* sp. nov., a new bacterial species isolated from a stool sample of an inhabitant of the amazon region. New Microbes New Infect. 42:100900

Meier-Kolthoff JP, Göker M (2019) TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun. 10(1):2182. Published 2019 May 16. https://doi.org/10.1038/s41467-019-10210-3

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Haeseler AV (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37(5):1530–1534. https://doi.org/10.1093/molbev/msaa015

Nagel M, Andreesen JR (1991) *Bacillus niacini* sp. nov., a nicotinate-metabolizing mesophile isolated from soil. Int J Syst Bacteriol 41:134–139. https://doi.org/10.1099/00207713-41-1-134

Ndongo S, Beye M, Laba N et al (2020) *Gorillibacterium timonense* sp. nov., isolated from an obese patient. Arch Microbiol 202:1223–1229. https://doi.org/10.1007/s00203-020-01827-0

Pal D, Kumar RM, Kaur N, Kumar N, Kaur G, Singh NK et al (2017) *Bacillus maritimus* sp. nov., a new member of the genus *Bacillus* isolated from marine sediments. Int J Syst Evol Microbiol 67:60–66. https://doi.org/10.1099/ijsem.0.001569

Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36(10):996–1004. https://doi.org/10.1038/nbt.4229

Patel S, Gupta RS (2020) A phylogenomic and comparative genomic framework for resolving the polyphyly of the genus *Bacillus*: proposal for six new genera of *Bacillus* species, *Peribacillus* gen. nov., *Cytobacillus* gen. nov., *Mesobacillus* gen. nov., *Neobacillus* gen. nov., *Metabacillus* gen. nov. and *Alkalihalobacillus* gen. nov. Int J Syst Evol Microbiol 70:406–438. https://doi.org/10.1099/ijsem.0.003775

Pignatelli M, Moya A (2009) Judge Tamames, EnvDB, a database for describing the environmental distribution of prokaryotic taxa. Environ Microbiol Rep 1:191–197. https://doi.org/10.1111/j.1758-2229.2009.00030.x

Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK (2016)Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Analytical Methods 8(1):12–24. https://doi.org/10.1039/C5AY02550H

Rodríguez JM (2014) The origin of human milk bacteria: is there a bacterial entero-mammary pathway during late pregnancy and lactation? Adv Nutr 5(6):779–784. https://doi.org/10.3945/an.114.007229

Roux V, Million M, Robert C, Magne A, Raoult D (2013) Noncontiguous finished genome sequence and description of *Oceanobacillus massiliensis* sp. nov. Stand Genomic Sci 9(2):370–384. https://doi.org/10.4056/sigs.4267953

Sanchis C, Bourguet D (2009) Applications in agriculture and insect resistance management—a review. In: Lichtfouse E, Navarrete M, Debaeke P, et al (eds) EDP Sciences, Springer, pp 243–55 https://doi.org/10.1007/978-90-481-2666-8_16

Sasser M (2006) Bacterial identification by gas chromatography analysis of fatty acids methyl esther (GC-FAME). MIDI, Technical Note #101, p 6. MIDI, Newark, DE

Saxena AK, Kumar M, Chakdar H et al (2020) *Bacillus* species in soil as a natural resource for plant health and nutrition. J Appl Microbiol 128:1583–1594. https://doi.org/10.1111/jam.14506

Schoeni JL, Wong ACL (2005) Food poisoning caused by *Bacillus cereus* and its toxins. Pal J Food Prot 68:636–648. https://doi.org/10.4315/0362-028x-68.3.636

Seemann T (2014) Prokka: rapid prokaryotic genome annotation. Bioinformatics 30(14):2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Stackebrandt E, Frederiksen W, Garrity GM (2002) Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. Int J Syst Evol Microbiol.

52(Pt 3):1043-1047. https://doi.org/10.1099/00207713-52-3-1043

Stein T (2005) *Bacillus subtilis* antibiotics: structures, synthesis and specific functions. Mol Microbiol 56:845–857. https://doi.org/10.1111/j.1365-2958.2005.04587.x

Sun H, Finlay B, Azad MB, Cuomo CA, Cowen LE, Berdy B et al (2025) The human milk bacteriome and mycobiome and their inter-kingdom interactions viewed across geography. Front Nutr [Internet]. Jul 7 [cited 2025 Nov 29]. Available from: https://www.frontiersin.org/journals/nutrition/articles/https://doi.org/10.3389/fnut.2025.1610346/full

Teufel F, Almagro Armenteros JJ, Johansen AR et al (2022) SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 40(7):1023–1025. https://doi.org/10.1038/s41587-021-01156-3

Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP Trüper HG (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. International Journal of Systematic Bacteriology 37(4):463–464. https://doi.org/10.1099/00207713-37-4-463

Wick RR, Judd LM, Holt KE (2017) Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13(6):e1005595. https://doi.org/10.1371/journal.pcbi.1005595