

## Non-contiguous finished genome sequence and description of *Oceanobacillus massiliensis* sp. nov.

Véronique Roux<sup>1\*</sup>, Matthieu Million<sup>1</sup>, Catherine Robert<sup>1</sup>, Alix Magne<sup>1</sup> and Didier Raoult<sup>1</sup>

<sup>1</sup>Aix Marseille Université, URMITE, Faculté de médecine, Aix-Marseille Université, Marseille, France

\*Correspondence: Véronique Roux (veronique.roux@univ-amu.fr)

Keywords: *Oceanobacillus massiliensis*, Firmicutes, Polar flagellum

---

*Oceanobacillus massiliensis* strain N'Diop<sup>T</sup> sp. nov. is the type strain of *O. massiliensis* sp. nov., a new species within the genus *Oceanobacillus*. This strain, whose genome is described here, was isolated from the fecal flora of a healthy patient. *O. massiliensis* is an aerobic rod. Here we describe the features of this organism, together with the complete genome sequence and annotation. The 3,532,675 bp long genome contains 3,519 protein-coding genes and 72 RNA genes, including between 6 and 8 rRNA operons.

---

### Introduction

*Oceanobacillus massiliensis* strain N'diop<sup>T</sup> (CSUR 132<sup>T</sup> = DSM 24644<sup>T</sup>) is the type strain of *O. massiliensis* sp. nov. This bacterium is a Gram-positive strictly aerobic rod, motile by a polar flagellum and was isolated from the stool of a healthy Senegalese patient as part of a "culturomics" study aiming at cultivating individually all species within human feces [1].

Presently, "the gold standard method" to define a bacterial species is DNA-DNA hybridization [2]. But this method is time-consuming and the inter-laboratory reproducibility is poor. So, with the development of PCR and sequencing methods, 16S rRNA gene sequence comparison with internationally-validated cutoff values constitutes, for most bacterial genera, a reliable, reproducible and comparable tool that enables the taxonomic classification of new bacterial species [3]. Recently, it was proposed that new bacterial taxa be described using a polyphasic approach [4] that would include the genome sequence, MALDI-TOF spectrum and main phenotypic characteristics (habitat, Gram-stain reaction, culture, cell wall structure and metabolic characteristics).

Here we present a summary classification and a set of features for *O. massiliensis* sp. nov. strain N'diop<sup>T</sup> together with the description of the complete genomic sequencing and annotation. These characteristics support the circumscription of the species *O. massiliensis*.

The genus *Oceanobacillus* was first described by Lu *et al.* [5] and was emended by Yumoto *et al.* [6]. The genus comprises 12 recognized species and two subspecies. These bacteria are motile Gram-

positive rods, obligately aerobic or facultatively anaerobic, obligately or facultatively alkaliphilic. These species were isolated from deep-sea sediment core [5,7], deteriorated mural paintings [8], salt field [9], freshwater fish [10], algal mat [11], insects in freshwater [12], *Bacillus*-dominated wastewater treatment system in Korea [13], fermented shrimp paste samples [14], soy sauce production equipment [15], a marine solar saltern [16], activated sludge in a bioreactor [17], traditional Korean fermented food [18] and a fermented *Polygonum indigo* liquor sample [19]. These bacteria belong to the phylum *Firmicutes*, in the family *Bacillaceae*. There is no evidence of pathogenicity of these bacteria.

### Classification and features

A stool sample was collected from a healthy 16-year-old male Senegalese volunteer patient living in N'diop (a rural village in the Guinean-Sudanian zone in Senegal), who was included in a research protocol. The patient gave an informed and signed consent, and the agreements of the National Ethics Committee of Senegal and the local ethics committee of the IFR48 (Marseille, France) were obtained under agreement 09-022 and 11-017. The fecal specimen was preserved at -80°C after collection and sent to Marseille. Strain N'diop (Table 1) was isolated in January 2011 by aerobic cultivation on 5% sheep blood-enriched Columbia agar (BioMerieux).

The bacterial DNA was extracted using the MagNA Pure LC DNA isolation kit III (Roche, Mannheim, Germany) with the MagNA Pure LC instrument as described by the manufacturer.

PCR amplification of the 16S rRNA gene was performed using the universal primer pair fD1 and rp2 [35]. PCR products were purified using MultiScreen PCR (Millipore) and sequencing reactions were carried out using a DNA sequencing kit (BigDye Terminator Cycle Sequencing v1.1 Ready Reactions, PE biosystems) according to the manufacturer's instructions. Sequencing products were purified and electrophoresis was performed with the 3130 Genetic Analyzer (Applied Biosystems). Base insertions were noted at the beginning and the end of the gene sequence. So, we tried to clone PCR products in pGEM-T Easy Vector (Promega) as described by the manufacturer. But the sequence was too long and no result was obtained. A new primer Omassr (GCCTGCAATCCGAACTGAGA) was chosen to reduce the length of fragment when combined with fD1 to perform PCR amplification. Seventeen clones were PCR amplified using the primers M13d (CGCCAGGGTTTTCCAGTCACGAC) and M13r (TCACACAGGAAACAGCTATGAC) and the obtained products were sequenced. The obtained sequences were compared with sequences deposited in the GenBank database by using the BLAST program through the NCBI server. This strain exhibited a 96.4-97% nucleotide sequence similarity with *O. profundus*, the phylogenetically closest validated *Oceanobacillus* species. Gene sequences were aligned using the multisequence alignment program CLUSTAL X (1.8) [36]. Phylogenetic relationships with closely related species were determined by using MEGA version 5.0 [37]. Distance matrices were determined following the assumptions described by Kimura and were used to elaborate a dendrogram using the neighbor-joining method (Figure 1). The maximum-parsimony algorithm was also used to infer phylogenetic analysis. A bootstrap analysis (bootstrap values were obtained for a consensus tree based on 1,000 randomly generated trees) was performed to investigate the stability of the trees obtained. The clustering of the new isolate was the same with the two methods.

For DNA-DNA hybridization, cells were disrupted by using a French pressure cell (Thermo Spectronic) and the DNA in the crude lysate was purified by chromatography on hydroxyapatite as described by Cashion *et al.* [38]. DNA-DNA hybridization was carried out as described by De Ley *et al.* [39] under consideration of the modifications described by Huss *et al.* [40] using a model Cary 100 Bio UV/VIS-spectrophotometer equipped with a Peltier-thermostatted 6x6 multicell changer and

temperature controller with in-situ temperature probe (Varian). DNA-DNA reassociation rate was 25.5% between strain NdiopT and *Oceanobacillus profundus*. So, they did not belong to the same species when the recommendations of a threshold value of 70% DNA-DNA similarity for the definition of bacterial species by the *ad hoc* committee are considered [41].

Surface colonies were observed on sheep blood agar (bioMérieux) after 24 h aerobic incubation at 37°C. The colonies of the strain N'diopT were circular, greyish, shiny and smooth, 2-5 mm in diameter. Gram staining showed Gram-positive bacilli (Figure 2). Different growth temperatures (25, 30, 37, 45 and 50°C) were tested. Growth occurred between 25°C and 45°C, and optimal growth was observed between 30°C and 37°C. Growth of the strain was tested under aerobic atmosphere, in the presence of 5% CO<sub>2</sub>, and also in anaerobic and microaerophilic atmospheres which were created using GENbag anaer and GENbag microaer (bioMérieux), respectively. The strain was aerobic and also grew in microaerophilia and in the presence of 5% CO<sub>2</sub> but did not grow in an anaerobic atmosphere. The NaCl concentrations allowing growth of strain N'DiopT, were determined on DifcoTMBrain Heart Infusion Agar plates (Becton Dickinson). The powder was supplemented with NaCl (Euromedex) to obtain the tested concentrations (0.5, 1, 2, 3, 5, 10, 15%, w/v). Growth occurred between 0.5-10% NaCl but the optimum growth was between 0.5-5% NaCl. Growth in the range of pH 5.0-10.0 was tested using BBLTM Brain Heart Infusion (Becton Dickinson) supplemented with 5% NaCl. The final pH was adjusted with HCl or NaOH solution. Growth occurred between pH 7-9 but optimum pH was 7-8.

The size and ultrastructure of cells were determined by negative staining transmission electron microscopy. Cells were grown on Columbia agar with 5% sheep blood for 24 h at 37°C. The bacteria were suspended in phosphate buffer (Gibco; KCl 200 mg/L, KH<sub>2</sub>PO<sub>4</sub> 200 mg/L, NaCl 8 g/L, Na<sub>2</sub>HPO<sub>4</sub>-7H<sub>2</sub>O 2.16 g/L) and pre-fixed in 5% (v/v) glutaraldehyde in phosphate buffer for at least 1 h at room temperature, washed in the same buffer and stained with 1% (w/v) phosphotungstic acid. The samples were examined on a Morgagni 268D (Philips) electron microscope at an operating voltage of 60 kV. The rods were 1.2-1.9 µm long and 0.4-0.7 µm wide (Figure 3). Polar flagella were observed.

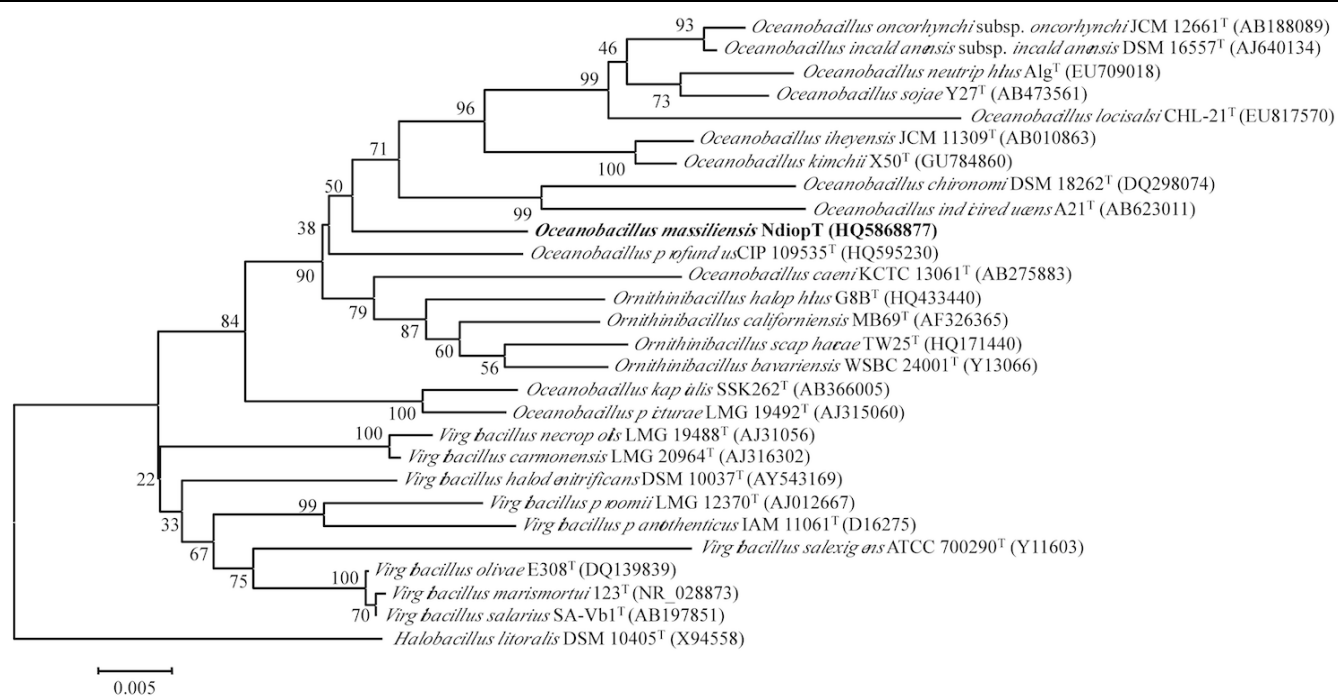
Strain N'Diop<sup>T</sup> exhibited catalase activity and oxidase activity. The commercially available Api ZYM, Api 20NE (bioMérieux) were used to characterize the biochemical properties of the strain according to the manufacturer's instructions and incubation was performed at 37 °C for 4h and 24h, respectively. Api 50 CH strips were inoculated with a heavy bacterial suspension in Api 50CHB/E medium supplemented

with 5% NaCl (w/v) and incubation was performed at 37°C for 48h. Phenotypic characteristics were compared to those of species that were most closely related in terms of 16S rRNA gene sequences. Characteristic traits are presented in Table 2.

**Table 1.** Classification and general features of *Oceanobacillus massiliensis* strain N'Diop<sup>T</sup> according to the MIGS recommendations [20]

MIGS ID	Property	Term	Evidence code <sup>a</sup>
		Domain <i>Bacteria</i>	TAS [21]
		Phylum <i>Firmicutes</i>	TAS [22-24]
		Class <i>Bacilli</i>	TAS [25,26]
	Current classification	Order <i>Bacillales</i>	TAS [27,28]
		Family <i>Bacillaceae</i>	TAS [28,29]
		Genus <i>Oceanobacillus</i>	TAS [5, 30-33]
		Species <i>Oceanobacillus massiliensis</i>	IDA
		Type strain N'Diop <sup>T</sup>	IDA
	Gram stain	Positive	IDA
	Cell shape	Bacilli	IDA
	Motility	Motile by polar flagellum	IDA
	Sporulation	Nonsporulating	IDA
	Temperature range	Mesophile	IDA
	Optimum temperature	37°C	IDA
MIGS-6.3	Salinity	Growth in BHI medium + 5% NaCl	IDA
MIGS-22	Oxygen requirement	Aerobic	IDA
	Carbon source	Unknown	NAS
	Energy source	Unknown	NAS
MIGS-6	Habitat	Human gut	IDA
MIGS-15	Biotic relationship	Free living	IDA
	Pathogenicity	Unknown	NAS
	Biosafety level	2	
MIGS-14	Isolation	Human feces	
MIGS-4	Geographic location	Senegal	IDA
MIGS-5	Sample collection time	September 2009	IDA
MIGS-4.1	Latitude	13.41	IDA
MIGS-4.1	Longitude	-16.22	IDA
MIGS-4.3	Depth	Surface	IDA
MIGS-4.4	Altitude	<100 m above sea level	IDA

Evidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable 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 [34]. If the evidence is IDA, then the property was directly observed for a live isolate by one of the authors or an expert mentioned in the acknowledgements.

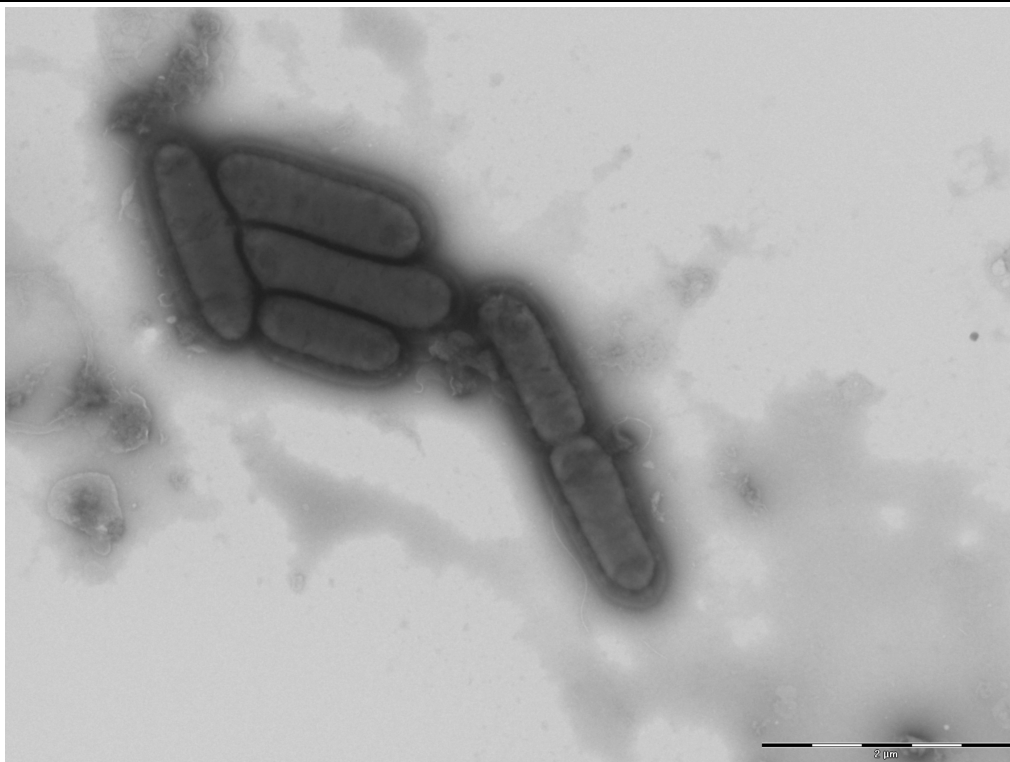


**Figure 1.** Phylogenetic tree highlighting the position of *Oceanobacillus massiliensis* strain N'Diop<sup>T</sup> relative to other type strains within the *Oceanobacillus*, *Ornithinibacillus*, and *Virgibacillus* genera. GenBank accession numbers are indicated in parentheses. Sequences were aligned using CLUSTALX, and phylogenetic inferences obtained using the neighbor joining method within the MEGA 5 software [18]. Numbers at the nodes are percentages of bootstrap values obtained by repeating the analysis 1,000 times to generate a majority consensus tree. *Halobacillus litoralis* was used as outgroup. The scale bar represents 0.005 nucleotide change per nucleotide position.



**Figure 2.** Gram staining of *O. massiliensis* strain N'Diop<sup>T</sup>





**Figure 3.** Transmission electron microscopy of *O. massiliensis* strain N'Diop<sup>T</sup>, using a Morgani 268D (Philips) at an operating voltage of 60kV. The scale bar represents 1  $\mu$ m.

Analysis of respiratory quinones by HPLC was carried out by the Identification Service and Dr Brian Tindall, DSMZ, Braunschweig, Germany. Respiratory lipoquinones were extracted from 100 mg of freeze dried cell material as described by Tindall [42,43]. Respiratory lipoquinones were separated into their different classes (menaquinones and ubiquinones) by thin layer chromatography on silica gel, using hexane:terbutylmethylether (9:1 v/v) as solvent. UV absorbing bands corresponding to menaquinones or ubiquinones were removed from the plate and further analyzed by HPLC at 269 nm. The respiratory quinones were MK-7 (97%) and MK6 (7%) for strain N'Diop<sup>T</sup>. Preparation and determination of cellular fatty acids were carried out by following the procedures given for the Sherlock Microbial identification System (MIDI). The major fatty acids were C15:0 anteiso 63.24% and C17:0 anteiso 26.86%. The DNA base composition was determined by using the HPLC method of Mesbah *et al.* [44]. The value found for strain N'diop<sup>T</sup> was 40.8%. Polar lipids were extracted from 100 mg of freeze dried cell material using a chloroform:methanol:0.3% aqueous NaCl mixture 1:2:0.8 (v/v/v) (modified after [45]). The extraction solvent was stirred overnight and the cell

debris pelleted by centrifugation. Polar lipids were recovered into the chloroform phase by adjusting the chloroform:methanol:0.3% aqueous NaCl mixture to a ratio of 1:1:0.9 (v/v/v). Polar lipids were separated as previously described [46]. The polar lipids present were diphosphatidylglycerol, phosphatidylglycerol, phospholipid 1, an unidentified aminolipid and sulfoquinovosyl diacylglycerol. The whole cells of *O. massiliensis* were hydrolyzed (4N HCl, 100 °C, 16 hours) and the hydrolysates were subjected to thin-layer chromatography on cellulose plates using the solvent system of Rhuland *et al.* [47]. Meso-diaminopimelic acid (*meso*-Dpm) was found as diagnostic diamino acid of the peptidoglycan. The occurrence of *meso*-Dpm has been reported up to now only for the peptidoglycan type A1 $\gamma$  (and A1 $\gamma'$  with glycine instead of L-alanine) and for three variations of peptidoglycan type A4 $\gamma$  (A31.1, A31.2 and A31.3, see [48]). The variations of peptidoglycan type A4 $\gamma$  based on *meso*-Dpm have been found so far exclusively in members of the genera *Brachybacterium*, *Devriesea* and *Dermabacter* (family *Dermabacteraceae*). Because a close relationship between *O. massiliensis* and the family *Dermabacteraceae* can be ruled out, we conclude the peptidoglycan type to be A1 $\gamma$ .

**Table 2.** Diagnostic trait differentiation.

Characteristic	1	2	3	4	5	6	7	8	9	10
<u>Assimilation of</u>										
Glycerol	w	w	+	+	w	-	w	-	-	-
L-Arabinose	-	-	-	+	-	-	w	-	-	-
Ribose	-	+	-	+	-	-	-	-	-	-
D-Xylose	+	+	-	+	-	-	w	-	-	-
Methyl- $\beta$ -D-Xylopyranoside	-	-	-	+	-	-	-	-	-	-
D-Galactose	-	-	+	+	-	-	-	-	-	-
D-Glucose	+	+	+	+	w	-	w	-	-	-
D-Fructose	+	+	+	+	w	-	w	-	-	-
D-Mannose	+	+	+	+	w	-	w	-	-	-
L-Sorbose	-	-	+	-	-	-	-	-	-	-
Dulcitol	-	-	+	+	-	-	-	-	-	-
Inositol	-	-	+	+	-	-	-	-	-	-
D-Mannitol	+	+	+	+	-	-	-	-	-	-
Methyl- $\alpha$ -D-Mannopyranoside	-	-	-	-	-	-	w	-	-	-
Methyl- $\alpha$ -D-Glucopyranoside	-	-	+	+	-	-	w	-	-	-
N-Acetylglucosamine	+	+	+	+	w	-	w	-	-	-
Amygdalin	-	+	+	+	-	-	-	-	-	-
Arbutin	-	-	+	+	-	-	-	-	-	-
Salicin	+	+	+	+	-	-	w	-	-	-
D-Cellobiose	+	+	+	+	-	-	w	-	-	-
D-Maltose	+	+	+	+	-	-	w	-	-	-
D-Saccharose	-	+	+	+	-	-	-	-	-	-
D-Trehalose	+	+	+	+	-	-	-	-	-	-
D-Raffinose	-	-	+	+	-	-	-	-	-	-
Gentiobiose	-	-	+	+	-	-	-	-	-	-
D-Turanose	-	-	+	+	-	-	-	-	-	-
D-Tagatose	-	+	+	+	-	-	-	-	-	-
D-Arabitol	-	w	w	w	-	-	-	-	-	-
potassium Gluconate	-	-	w	w	-	+	-	-	-	-
Reduction of nitrates to nitrites	-	+	+	+	-					
Urease	-	+	-	-	-	-	-	-	-	-
Esculin hydrolysis	+	+	+	+	-	+	-	-	-	-
Para-nitrophenyl- $\beta$ -D-galactopyranosidase	-	w	-	w	-	+	w	-	+	+
Esterase (C4)	+	+	+	+	+	+	+	-	-	-
Esterase lipase (C8)	+	+	+	+	w	+	w	w	w	+
Leucine arylamidase	-	-	-	-	w	w	-	+	w	-
$\alpha$ -chemotrypsin	-	-	w	-	-	-	w	-	-	-
$\beta$ -glucuronidase	-	+	-	-	-	-	-	+	+	-
$\beta$ -glucosidase	+	-	+	-	-	-	-	-	-	-

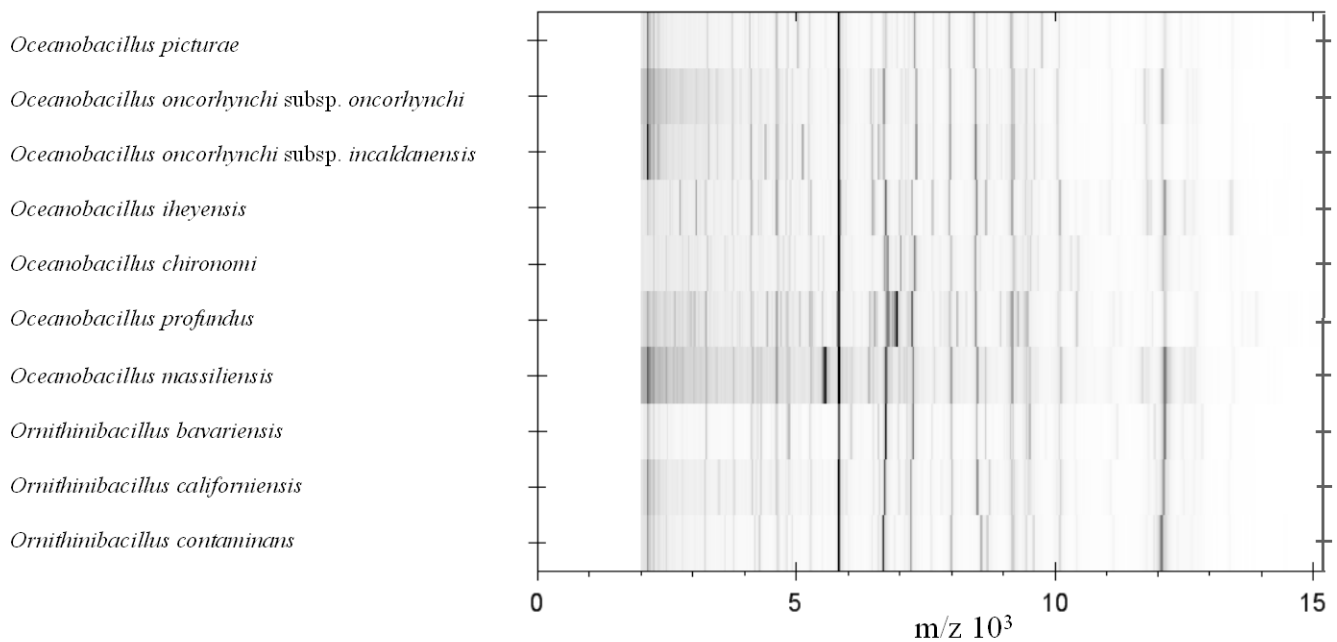
Strains: 1, *O. massiliensis* sp. nov. NdiopT; 2, *O. profundus* CIP 109535T ; 3, *O. oncorhynchi* subsp. *oncorhynchi* CIP 108867T; 4, *O. oncorhynchi* subsp. *incaldanensis* CIP 109235T; 5, *O. iheyensis* CIP 107618T; 6, *O. chironomi* CIP109536T ; 7, *O. picturae* CIP 108264T; 8, *Ornithinibacillus bavariensis* DSM 15681T; 9, *Ornithinibacillus californiensis* DSM 16628T; 10, *Ornithinibacillus contaminans* DSM 22953T.

+: positive result, -: negative result, w: weak positive result

Antimicrobial susceptibility was determined according to the National Committee for Clinical Laboratory Standards (NCCLS) criteria. Strain N'diop<sup>T</sup> was found to be susceptible to doxycycline, rifampicine, vancomycine, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamycine, penicilline, imipenem. But it was resistant to trimethoprim/sulfamethoxazole and metronidazole.

Matrix-assisted laser-desorption/ionization time-of-flight (MALDI-TOF) MS protein analysis was carried out. Briefly, a pipette tip was used to pick one isolated bacterial colony from a culture agar plate, and to spread it as a thin film on a MALDI-TOF target plate (Bruker Daltonics). Twelve distinct deposits were done for strain N'Diop<sup>T</sup> from twelve isolated colonies and the manipulation was repeated another day. After air-drying, 1.5 µl matrix solution (saturated solution of α-cyanohydroxycinnamic acid in 50% aqueous acetonitrile containing 2.5% trifluoroacetic acid) per spot was applied. MALDI-TOF MS was conducted using the Microflex LT spectrometer (Bruker Daltonics). All spectra were recorded in linear, positive ion mode. The acceleration voltage was 20 kV. Spectra were collected as a sum of 240 shots across a spot. Preprocessing and identification steps were performed using the manufacturer's parameters. The N'Diop<sup>T</sup> spectra

were imported into the MALDI BioTyper software (version 3.0, Bruker) and analyzed by standard pattern matching (with default parameter settings) against the main spectra of 4,108 bacteria including the spectra from *Oceanobacillus profundus* CIP 109535<sup>T</sup>, *Oceanobacillus picturae* CIP 108264<sup>T</sup>, *Oceanobacillus chironomi* CIP 109536<sup>T</sup>, *Oceanobacillus iheyensis* CIP 107618<sup>T</sup>, *Oceanobacillus oncorhynchi* subsp. *oncorhynchi* CIP 108867<sup>T</sup> and *Oceanobacillus oncorhynchi* subsp. *incaldanensis* CIP 109235<sup>T</sup> which were the most closely related species when 16S rRNA gene sequences were compared and *Ornithinibacillus bavariensis* DSM 15681<sup>T</sup>, *Ornithinibacillus californiensis* DSM 16628<sup>T</sup> and *Ornithinibacillus contaminans* DSM 22953<sup>T</sup>, used as reference data, in the BioTyper database. A score enabled the identification, or not, from the tested species: a score > 2.3 with a validated species enabled the identification at the species level, a score > 1.7 but < 2 enabled the identification at the genus level; and a score < 1.7 did not enable any identification. For strain N'Diop<sup>T</sup>, none of the obtained scores were > 1.5, thus suggesting that our isolate was not a member of a known species. We incremented our database with the spectrum from strain N'Diop<sup>T</sup> (Figure 4). The spectrum was made available online in our free-access URMS database [49].



**Figure 4.** Reference mass spectra from *O. massiliensis* strain N'Diop<sup>T</sup>, other representatives of the genus *Oceanobacillus* and type strains of the genus *Ornithinibacillus*. Spectra from 24 individual colonies were compared and a reference spectrum was generated.

## Genome sequencing information

### Genome project history

The organism was selected for sequencing on the basis of its phylogenetic position and 16S rRNA similarity to other members of the genus *Oceanobacillus*, and is part of a “culturomics” study of the human digestive flora aiming at isolating all bacterial species within human feces. It was the second sequenced genome of an *Oceanobacillus* species, *Oceanobacillus massiliensis* sp. nov. A summary of the project information is shown in Table 3. The EMBL accession number is CAER01000000 and consists of 95 contigs ( $\geq 500$  bp) and 9 scaffold ( $> 2,500$  bp). Table 2 shows the project information and its association with MIGS version 2.0 compliance.

### Growth conditions and DNA isolation

*O. massiliensis* sp. nov. strain N'Diop<sup>T</sup>, CSUR P132<sup>T</sup>, DSM 24644<sup>T</sup>, was grown aerobically on 5% sheep blood-enriched Columbia agar at 37°C. Three petri dishes were spread and resuspended in 3x100 $\mu$ l of G2 buffer. A first mechanical lysis was performed by glass powder on the Fastprep-24 device (Sample Preparation system) from MP Biomedicals, USA during 2x20 seconds. DNA was then incubated for a lysozyme treatment (30 minutes at 37°C) and extracted through the BioRobot EZ 1 Advanced XL (Qiagen). The DNA was then concentrated and purified on a Qiam kit

(Qiagen). The yield and the concentration were measured by the Quant-it Picogreen kit (Invitrogen) on the Genios Tecan fluorometer at 139 ng/ $\mu$ l.

### Genome sequencing and assembly

Shotgun and 3-kb paired-end sequencing strategies were performed. The shotgun library was constructed with 500 ng of DNA with the GS Rapid library Prep kit (Roche). For the paired-end sequencing, 5  $\mu$ g of DNA was mechanically fragmented on a Hydroshear device (Digilab) with an enrichment size at 3-4 kb. The DNA fragmentation was visualized using the 2100 BioAnalyzer (Agilent) on a DNA labchip 7500 with an optimal size of 3.1 kb. The library was constructed according to the 454 GS FLX Titanium paired-end protocol. Circularization and nebulization were performed and generated a pattern with an optimal size of 450 bp. After PCR amplification through 17 cycles followed by double size selection, the single stranded paired-end library was then quantified using the Genios fluorometer (Tecan) at 220 pg/ $\mu$ L. The library concentration equivalence was calculated as 8.97 E+08 molecules/ $\mu$ L. The library was stored at -20°C until further use.

**Table 3.** Project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	High-quality draft
MIGS-28	Libraries used	One paired end 3-kb library and one Shotgun library
MIGS-29	Sequencing platforms	454 GS FLX Titanium
MIGS-31.2	Fold coverage	23x
MIGS-30	Assemblers	Newbler version 2.5.3
MIGS-32	Gene calling method	Prodigal
	EMBL ID	CAER01000000
	EMBL Date of Release	February 28,2013
	Project relevance	Study of the human gut microbiome



The shotgun and paired-end libraries were clonally-amplified with 3 cpb and 1 cpb in 3SV-emPCR reactions respectively on the GS Titanium SV emPCR Kit (Lib-L) v2 (Roche). The yields of the emPCR were 15.1% and 12.1% respectively. Approximately 790,000 beads for the shotgun application and for the 3kb paired end were loaded on the GS Titanium PicoTiterPlate PTP Kit 70x75 and sequenced with the GS FLX Titanium Sequencing Kit XLR70 (Roche). The run was performed overnight and then analyzed on the cluster through the gsRunBrowser and Newbler assembler (Roche). A total of 279,999 passed filter wells were obtained and generated 81 Mb with a length average of 289 bp. The passed filter sequences were assembled using Newbler with 90% identity and 40 bp as overlap. The final assembly identified 9 scaffolds and large 82 contigs (>1,500 bp).

### 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 (COG) databases [52] using BLASTP. The tRNAscan-SE tool [53] was used to find tRNA genes, whereas ribosomal RNAs were found by using RNAmmer [54].

Transmembrane domains and signal peptides were predicted using TMHMM [55] and SignalP [56], respectively. ORFans were identified if their BLASTp *E*-value was lower than 1e-03 for alignment length greater than 80 amino acids. If

alignment lengths were smaller than 80 amino acids, we used an *E*-value of 1e-05. Such parameter thresholds have been used in previous works to define ORFans.

To estimate the mean level of nucleotide sequence similarity at the genome level between *O. massiliensis* and *O. iheyensis* (GenBank accession number PRJNA57867), the only available *Oceanobacillus* genome to date, we compared the ORFs only using comparison sequence based in the server RAST [57] at a query coverage of  $\geq 70\%$  and a minimum nucleotide length of 100 bp.

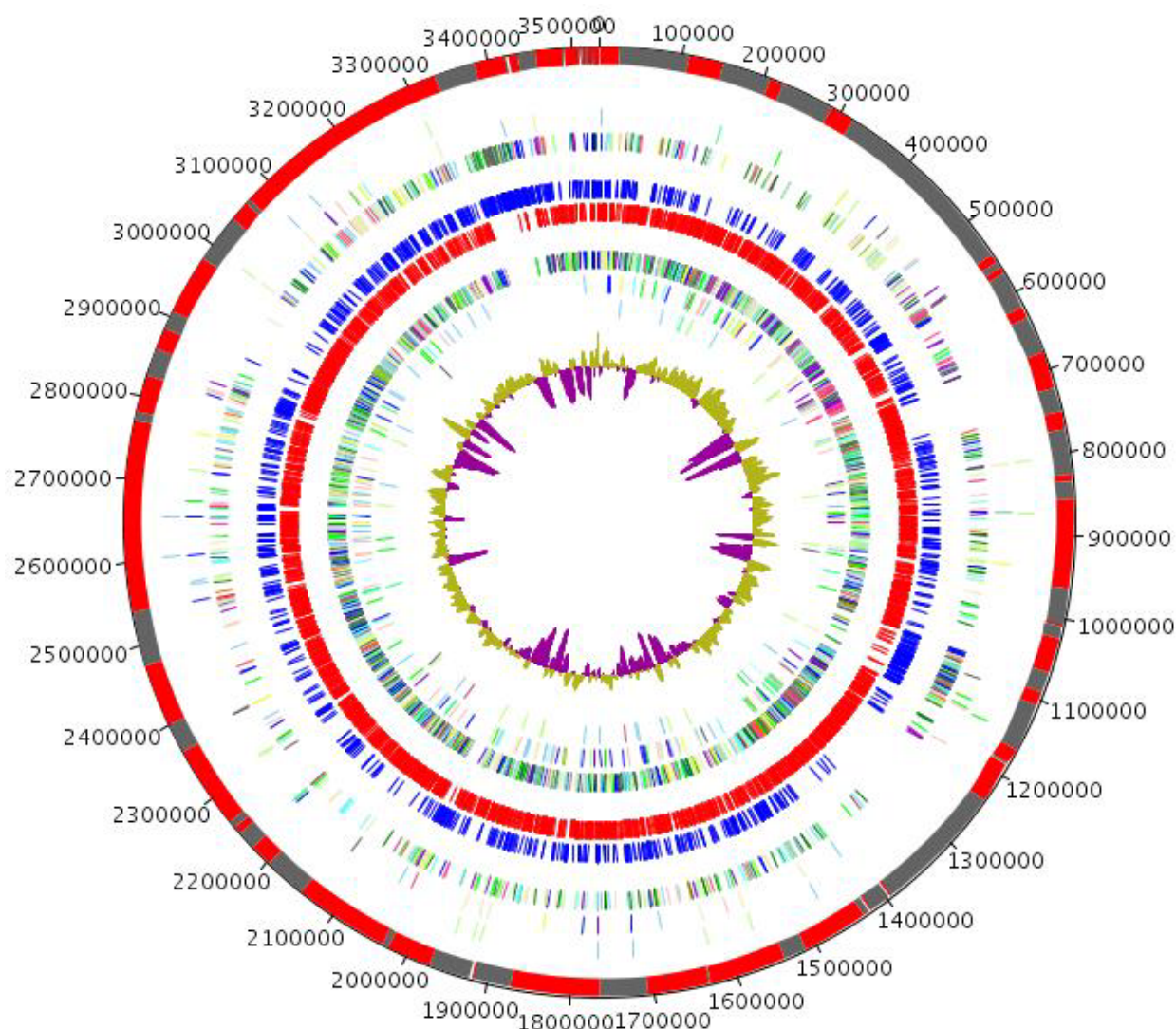
### Genome properties

The genome is 3,532,675 bp long with 40.35% GC content (Table 4 and Figure 5). It is composed of 95 Contigs (9 Scaffolds). Of the 3,589 predicted genes, 3,519 were protein-coding genes, and 72 were RNAs (1 gene is 16S rRNA, 1 gene is 23S rRNA, 9 genes are 5S rRNA, and 61 are tRNA genes). A total of 2,536 genes (72.07%) were assigned a putative function (by cogs or by NR blast). In addition, 84 genes were identified as ORFans (2.39%). The remaining genes were annotated as hypothetical proteins (618 genes (17.56%)). The distribution of genes into COGs functional categories is presented in Table 5. The properties and the statistics of the genome are summarized in Tables 4 and 5. Two CRISPRs were found using CRISPERfinder program online [58] which included at least 48 predicted spacer regions (contigs 39-41) and 13 predicted spacer regions (contig 92).

**Table 4.** Nucleotide content and gene count levels of the genome

Attribute	Value	% of total <sup>a</sup>
Genome size (bp)	3,532,675	100
DNA coding region (bp)	2,971,565	84.12
DNA G+C content (bp)	1,425,470	40.35
Total genes	3,589	100
RNA genes	72	1.95
Protein-coding genes	3,519	98.05
Genes with function prediction	2,536	72.07
Genes assigned to COGs	2,673	75.96
Genes with peptide signals	348	9.89
Genes with transmembrane helices	885	25.15

a) The total is based on either the size of the genome in base pairs or the total number of protein coding genes in the annotated genome.



**Figure 5.** Graphical circular map of the *O. massiliensis* strain N'Diop genome. From outside to the center: Contigs (red / grey), COG category of genes on the forward strand (three circles), genes on forward strand (blue circle), genes on the reverse strand (red circle), COG category on the reverse strand (three circles), G+C content.

### Comparison with other *Oceanobacillus* genomes

To date, only one genome of bacteria belonging to the genus *Oceanobacillus* was sequenced, *Oceanobacillus iheyensis*. By comparison with the genome of *O. iheyensis*, *O. massiliensis* had a smaller genome (3.53 vs 3.63 Mb, respectively), a higher G+C content (40.35 vs 35.7%) but a smaller number of predicted genes (3,589 vs 3,593 genes). In addition, *O. massiliensis* and *O. iheyensis* shared a mean nucleotide sequence similarity of 80.49% at the genome level (range 70.03-98.32%).

### Prophage genome properties

Prophage Finder [59] was used to identify potential proviruses in *O. massiliensis* strain N'Diop<sup>T</sup> genome. The genome contains at least one genetic

element of around 48.9 kb (with a GC content of 38.9%), we named OM1, on contig 65. A total of 51 open reading frames (ORFs) were recovered from OM1, larger than 160 nucleotides and most of them (42) encode proteins sharing a high identity with proteins found in *Clostridiales* genus viruses. The preliminary annotation of OM1 was performed and the majority of the putative genes (31) encode hypothetical proteins. The ORFs with an attributed function (21) encode proteins involved in DNA packaging, cell lysis, tail structural components and assembly, head structural components and assembly, lysogeny control, DNA replication, recombination and modification. Fifty one of the ORFs are located on one strand and 1 on the opposite strand.

**Table 5.** Number of genes associated with the 25 general COG functional categories

Code	Value	%age	Description
J	170	4.83	Translation
A	0	0	RNA processing and modification
K	232	6.59	Transcription
L	184	5.23	Replication, recombination and repair
B	1	0.03	Chromatin structure and dynamics
D	37	1.05	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	48	1.36	Defense mechanisms
T	142	4.04	Signal transduction mechanisms
M	156	4.43	Cell wall/membrane biogenesis
N	64	1.82	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	52	1.48	Intracellular trafficking and secretion
O	104	2.96	Posttranslational modification, protein turnover, chaperones
C	149	4.23	Energy production and conversion
G	194	5.51	Carbohydrate transport and metabolism
E	283	8.04	Amino acid transport and metabolism
F	84	2.39	Nucleotide transport and metabolism
H	100	2.84	Coenzyme transport and metabolism
I	84	2.39	Lipid transport and metabolism
P	224	6.37	Inorganic ion transport and metabolism
Q	54	1.53	Secondary metabolites biosynthesis, transport and catabolism
R	424	12.05	General function prediction only
S	281	7.99	Function unknown
X	846	24.04	Not in COGs

a) The total is based on the total number of protein coding genes in the annotated genome.

## Conclusion

Bacteria included in genera *Oceanobacillus* and *Ornithinibacillus* are closely related. The inclusion of these genera is based only on the iso-<sub>C15:0</sub>/anteiso-<sub>C15:0</sub> ratio (>1 for the *Ornithinibacillus* genus and <1 for the *Oceanobacillus* genus) and the peptidoglycan type (A4β for the *Ornithinibacillus* genus and A1γ for the genus *Oceanobacillus* genus). All the other criteria proposed are not valid for all the representatives of the genus, or can be discordant when manipulations are performed in different laboratories.

On the basis of phenotypic, phylogenetic and genomic analyses, we propose the creation of *Oceanobacillus massiliensis* sp. nov. that contains the strain N'Diop<sup>T</sup>. This bacterium has been found in Senegal.

In the future, sequencing and comparison of several genomes of *Oceanobacillus* and *Ornithinibacillus* representatives will allow better understanding of these taxa.

**Description of *Oceanobacillus massiliensis* sp. nov.**

*Oceanobacillus massiliensis* (mas.si.li.en'sis. L. masc. adj. *massiliensis* of Massilia, the old Roman name for Marseille, where the type strain was isolated). Cells are aerobic, Gram-positive, straight, and motile rods. Catalase-positive and oxidase-positive. Growth occurred in aerobic conditions. Optimal growth occurs at 30-37 °C, pH 7-8 and NaCl range of 0,5-5%. After 24 hours growth on sheep blood agar, surface colonies are up to 2-5 mm in diameter, smooth, circular, greyish and shiny. The rods measure 1.2-1.9 µm in length and 0.4-0.7 µm in diameter (as determined by electron microscopy). Using Api ZYM, activities of esterase (C4), esterase lipase (C8) and α-glucosidase are detected. Activities of alkaline phosphatase, lipase (C14), leucine arylaminidase, valine arylaminidase, cystine arylaminidase, trypsin, α-chemotrypsin, acid phosphatase, naphthol-AS-BI-phosphohydrolase, β-galactosidase, β-glucuronidase, β-glucosidase, N-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase are not detected. Using Api 50CH strips after an incubation time of 48 h acid is produced from glycerol (weakly), D-xylose, D-glucose, D-fructose, D-mannose, D-mannitol, N-acetylglucosamine, esculin ferric citrate, salicin, D-cellobiose, D-maltose and D-trehalose. Acid is not produced from

erythritol, D-arabinose, L-arabinose, D-ribose, L-xylose, D-adonitol, methyl-βD-xylopyranoside, D-galactose, L-sorbose, L-rhamnose, dulcitol, inositol, D-sorbitol, methyl-αD-mannopyranoside, methyl-αD-glucopyranoside, amygdalin, arbutin, D-lactose, D-melibiose, D-saccharose, inulin, D-melezitose, D-raffinose, amidon, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabinol, L-arabinol, potassium gluconate, potassium 2-ketogluconate and potassium 5-ketogluconate. Using Api 20 NE, hydrolysis of esculine is positive. Nitrate reduction, indole production, glucose fermentation, arginine dihydrolase, urease, gelatin hydrolysis and β-galactosidase are negative.

The respiratory quinones are MK-7 (93%) and MK-6 (7%). The major fatty acids are C15:0 anteiso (63.24%) and C17:0 anteiso (26.86%). The G+C content of the genomic DNA is 40.35-40.8%.

The type strain, which was isolated from human feces, is N'diop<sup>T</sup>. It has been deposited in the Collection de Souches de l'Unité des Rickettsies, Marseille, France, as CSUR P132<sup>T</sup> and in the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, as DSM 24644<sup>T</sup>.

**Acknowledgements**

The authors thank Julien Paganini at Xegen Company ([www.xegen.fr](http://www.xegen.fr)) for automating the genomic annotation process.

**References**

1. Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, Bittar F, Fournous G, Gimenez G, Maraninchi M, *et al.* Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect* 2012; **18**:1185-1193. [PubMed](http://pubmed.ncbi.nlm.nih.gov/22811193/)
2. Rossello-Mora R. DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation. In: Stackebrandt E (ed), *Molecular Identification, Systematics, and population Structure of Prokaryotes*, Springer, Berlin, 2006, p. 23-50.
3. Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, *et al.* Report of the *ad hoc* committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* 1987; **37**:463-464. <http://dx.doi.org/10.1099/00207713-37-4-463>
4. Tindall BJ, Rossello-Mora R, Busse HJ, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol* 2010; **60**:249-266. [PubMed](http://pubmed.ncbi.nlm.nih.gov/20169490/) <http://dx.doi.org/10.1099/ijs.0.016949-0>
5. Lu J, Nogi Y, Takami H. *Oceanobacillus iheyensis* gen. nov., sp. nov., a deep-sea extremely halotolerant and alkaliphilic species isolated from a depth of 1050 m on the Iheya Ridge. *FEMS Microbiol Lett* 2001; **205**:291-297. [PubMed](http://pubmed.ncbi.nlm.nih.gov/11574696/) <http://dx.doi.org/10.1111/j.1574-6968.2001.tb10963.x>
6. Yumoto I, Hirota K, Nodasaka Y, Nakajima K. *Oceanobacillus oncorhynchi* sp. nov., a halotolerant obligate alkaliphile isolated from the



- skin of a rainbow trout (*Oncorhynchus mykiss*), and emended description of the genus *Oceanobacillus*. *Int J Syst Evol Microbiol* 2005; **55**:1521-1524. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.63483-0>
7. Kim YG, Choi DH, Hyun S, Cho BC. *Oceanobacillus profundus* sp. nov., isolated from a deep-sea sediment core. *Int J Syst Evol Microbiol* 2007; **57**:409-413. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.64375-0>
8. Heyrman J, Logan NA, Busse HJ, Balcaen A, Lebbe L, Rodriguez-Diaz M, Swings J, De Vos P. *Virgibacillus carmonensis* sp. nov., *Virgibacillus necropolis* sp. nov. and *Virgibacillus picturae* sp. nov., three novel species isolated from deteriorated mural paintings, transfer of the species of the genus *Salibacillus* to *Virgibacillus*, as *Virgibacillus marismortui* comb. nov. and *Virgibacillus salexigens* comb. nov., and emended description of the genus *Virgibacillus*. *Int J Syst Evol Microbiol* 2003; **53**:501-511. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.02371-0>
9. Lee JS, Lim JM, Lee KC, Lee JC, Park YH, Kim CJ. *Virgibacillus koreensis* sp. nov., a novel bacterium from salt field, and transfer of *Virgibacillus picturae* to the genus *Oceanobacillus* as *Oceanobacillus picturae* comb. nov. with emended descriptions. *Int J Syst Evol Microbiol* 2006; **56**:251-257. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.63734-0>
10. Yumoto I, Hirota K, Nodasaka Y, Nakajima K. *Oceanobacillus oncorhynchi* sp. nov., a halotolerant obligate alkaliphile isolated from the skin of a rainbow trout (*Oncorhynchus mykiss*), and emended description of the genus *Oceanobacillus*. *Int J Syst Evol Microbiol* 2005; **55**:1521-1524. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.63483-0>
11. Romano I, Lama L, Nicolaus B, Poli A, Gambacorta A, Giordano A. *Oceanobacillus oncorhynchi* subsp. *incaldanensis* subsp. nov, an alkalitolerant halophile isolated from an algal mat collected from a sulfurous spring in Campania (Italy), and emended description of *Oceanobacillus oncorhynchi*. *Int J Syst Evol Microbiol* 2006; **56**:805-810. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.63575-0>
12. Raats D, Halpern M. *Oceanobacillus chironomi* sp. nov., a halotolerant and facultatively alkaliphilic species isolated from a chironomid egg mass. *Int J Syst Evol Microbiol* 2007; **57**:255-259. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.64502-0>
13. Nam JH, Bae W, Lee DH. *Oceanobacillus caeni* sp. nov., isolated from a *Bacillus*-dominated wastewater treatment system in Korea. *Int J Syst Evol Microbiol* 2008; **58**:1109-1113. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.65335-0>
14. Namwong S, Tanasupat S, Lee KC, Lee JS. *Oceanobacillus kapialis* sp. nov., from fermented shrimp paste in Thailand. *Int J Syst Evol Microbiol* 2009; **59**:2254-2259. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.007161-0>
15. Tominaga T, An SY, Oyaizu H, Yokota A. *Oceanobacillus soja* sp. nov. isolated from soy sauce production equipment in Japan. *Int J Syst Evol Microbiol* 2009; **55**:225-232. [PubMed](#)
16. Lee SY, Oh TK, Kim W, Yoon JH. *Oceanobacillus locisalsi* sp. nov., isolated from a marine solar saltern of the Yellow Sea, Korea. *Int J Syst Evol Microbiol* 2010; **60**:2758-2762. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.021907-0>
17. Whon TW, Jung MJ, Roh SW, Nam YD, Park EJ, Shin KS, Bae JW. *Oceanobacillus kimchii* sp. nov. isolated from a traditional Korean fermented food. *J Microbiol* 2010; **48**:862-866. [PubMed](#) <http://dx.doi.org/10.1007/s12275-010-0214-7>
18. Hirota K, Aino K, Nodasaka Y, Yumoto I. *Oceanobacillus indicireducens* sp. nov., a facultative alkaliphile that reduces an indigo dye. *Int J Syst Evol Microbiol* 2013; **63**:1437-1442. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.034579-0>
19. Yang JY, Huo YY, Xu XW, Meng FX, Wu M, Wang CS. *Oceanobacillus neutriphilus* sp. nov., isolated from activated sludge in a bioreactor. *Int J Syst Evol Microbiol* 2010; **60**:2409-2414. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.016295-0>
20. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 2008; **26**:541-547. [PubMed](#) <http://dx.doi.org/10.1038/nbt1360>
21. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc Natl Acad Sci USA* 1990; **87**:4576-4579. [PubMed](#) <http://dx.doi.org/10.1073/pnas.87.12.4576>
22. Murray RGE. The Higher Taxa, or, a Place for Everything...? In: Holt JG (ed), *Bergey's Manual of Systematic Bacteriology*, First Edition, Volume 1, The Williams and Wilkins Co., Baltimore, 1984, p. 31-34.

23. Garrity GM, Holt JG. The Road Map to the Manual. In: Garrity GM, Boone DR, Castenholz RW (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 1, Springer, New York, 2001, p. 119-169.
24. Gibbons NE, Murray RGE. Proposals Concerning the Higher Taxa of Bacteria. *Int J Syst Bacteriol* 1978; **28**:1-6. <http://dx.doi.org/10.1099/00207713-28-1-1>
25. Ludwig W, Schleifer KH, Whitman WB. Class I. Bacilli class nov. In: De Vos P, Garrity G, Jones D, Krieg NR, Ludwig W, Rainey FA, Schleifer KH, Whitman WB (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 3, Springer-Verlag, New York, 2009, p. 19-20.
26. Euzéby J. List of new names and new combinations previously effectively, but not validly, published. List no. 132. *Int J Syst Evol Microbiol* 2010; **60**:469-472. <http://dx.doi.org/10.1099/ijs.0.022855-0>
27. Prévot AR. In: Hauderoy P, Ehringer G, Guillot G, Magrou. J., Prévot AR, Rosset D, Urbain A (eds), *Dictionnaire des Bactéries Pathogènes*, Second Edition, Masson et Cie, Paris, 1953, p. 1-692.
28. Skerman VBD, McGowan V, Sneath PHA. Approved Lists of Bacterial Names. *Int J Syst Bacteriol* 1980; **30**:225-420. <http://dx.doi.org/10.1099/00207713-30-1-225>
29. Fischer A. Untersuchungen über bakterien. *Jahrbücher für Wissenschaftliche Botanik* 1895; **27**:1-163.
30. List Editor. Validation List no. 85. Validation of publication of new names and new combinations previously effectively published outside the IJSEM. *Int J Syst Evol Microbiol* 2002; **52**:685-690. [PubMed http://dx.doi.org/10.1099/ijs.0.02358-0](http://dx.doi.org/10.1099/ijs.0.02358-0)
31. Hirota K, Aino K, Nodasaka Y, Yumoto I. *Oceanobacillus indicireducens* sp. nov., a facultative alkaliphile that reduces an indigo dye. *Int J Syst Evol Microbiol* 2013; **63**:1437-1442. [PubMed http://dx.doi.org/10.1099/ijs.0.034579-0](http://dx.doi.org/10.1099/ijs.0.034579-0)
32. Yumoto I, Hirota K, Nodasaka Y, Nakajima K. *Oceanobacillus oncorhynchi* sp. nov., a halotolerant obligate alkaliphile isolated from the skin of a rainbow trout (*Oncorhynchus mykiss*), and emended description of the genus *Oceanobacillus*. *Int J Syst Evol Microbiol* 2005; **55**:1521-1524. [PubMed http://dx.doi.org/10.1099/ijs.0.63483-0](http://dx.doi.org/10.1099/ijs.0.63483-0)
33. Lee JS, Lim JM, Lee KC, Lee JC, Park YH, Kim CJ. *Virgibacillus koreensis* sp. nov., a novel bacterium from a salt field, and transfer of *Virgibacillus picturae* to the genus *Oceanobacillus* as *Oceanobacillus picturae* comb. nov. with emended descriptions. *Int J Syst Evol Microbiol* 2006; **56**:251-257. [PubMed http://dx.doi.org/10.1099/ijs.0.63734-0](http://dx.doi.org/10.1099/ijs.0.63734-0)
34. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; **25**:25-29. [PubMed http://dx.doi.org/10.1038/75556](http://dx.doi.org/10.1038/75556)
35. Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 1991; **173**:697-703. [PubMed http://dx.doi.org/10.1093/nar/25.24.4876](http://dx.doi.org/10.1093/nar/25.24.4876)
36. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997; **25**:4876-4882. [PubMed http://dx.doi.org/10.1093/molbev/msr121](http://dx.doi.org/10.1093/molbev/msr121)
37. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* 2011; **28**:2731-2739. [PubMed http://dx.doi.org/10.1016/0003-2697\(77\)90720-5](http://dx.doi.org/10.1016/0003-2697(77)90720-5)
38. Cashion P, Hodler-Franklin MA, McCully J, Franklin M. A rapid method for base ratio determination of bacterial DNA. *Anal Biochem* 1977; **81**:461-466. [PubMed http://dx.doi.org/10.1111/j.1432-1033.1970.tb00830.x](http://dx.doi.org/10.1111/j.1432-1033.1970.tb00830.x)
39. De Ley J, Cattoir H, Reynaerts A. The quantitative measurement of DNA hybridization from renaturation rates. *Eur J Biochem* 1970; **12**:133-142. [PubMed http://dx.doi.org/10.1016/S0723-2020\(83\)80048-4](http://dx.doi.org/10.1016/S0723-2020(83)80048-4)
40. Huss VAR, Festi H, Schleifer KH. Studies on the spectrophotometric determination of DNA hybridization from the renaturation rates. *Syst Appl Microbiol* 1983; **4**:184-192. [PubMed http://dx.doi.org/10.1099/00207713-37-4-463](http://dx.doi.org/10.1099/00207713-37-4-463)
41. Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, *et al.* Report of the *ad hoc* committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* 1987; **37**:463-464. <http://dx.doi.org/10.1099/00207713-37-4-463>

42. Tindall BJ. A comparative study of the lipid composition of *Halobacterium saccharovorum* from various sources. *Syst Appl Microbiol* 1990; **13**:128-130. [http://dx.doi.org/10.1016/S0723-2020\(11\)80158-X](http://dx.doi.org/10.1016/S0723-2020(11)80158-X)
43. Tindall BJ. Lipid composition of *Halobacterium lacusprofundi*. *FEMS Microbiol Lett* 1990; **66**:199-202. <http://dx.doi.org/10.1111/j.1574-6968.1990.tb03996.x>
44. Mesbah M, Premachandran U, Whitman W. Precise measurement of the G+C content of deoxyribonucleic acid by high performance liquid chromatography. *Int J Syst Bact* 1989; **39**:159-167. <http://dx.doi.org/10.1099/00207713-39-2-159>
45. Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* 1959; **37**:911-917. [PubMed](http://dx.doi.org/10.1139/o59-099)  
<http://dx.doi.org/10.1139/o59-099>
46. Tindall BJ, Sikorski J, Smibert RM, Kreig NR. Phenotypic characterization and the principles of comparative systematics. In: Reddy CA, Beveridge TJ, Breznak JA, Marzluf G, Schmidt TM, Snyder LR (eds), *Methods for General and Molecular Microbiology*, 3<sup>rd</sup> edition, ASM Press, Washington, 2007, p. 330-393.
47. Rhuland LE, Work E, Denman RF, Hoare DS. The behaviour of the isomers of  $\alpha$ , $\epsilon$ -diaminopimelic acid on paper chromatograms. *J Am Chem Soc* 1955; **77**:4844-4846. <http://dx.doi.org/10.1021/ja01623a047>
48. [http://www.dsmz.de/microorganisms/main.php?content\\_id=35](http://www.dsmz.de/microorganisms/main.php?content_id=35)
49. [http://ifr48.timone.univ-mrs.fr/portail2/index.php?option=com\\_content&task=view&id=97&Itemid=54](http://ifr48.timone.univ-mrs.fr/portail2/index.php?option=com_content&task=view&id=97&Itemid=54).
50. Prodigal <http://prodigal.ornl.gov/>
51. GenBank database. <http://www.ncbi.nlm.nih.gov/genbank>
52. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 2000; **28**:33-36. [PubMed](http://dx.doi.org/10.1093/nar/28.1.33)  
<http://dx.doi.org/10.1093/nar/28.1.33>
53. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 1997; **25**:955-964. [PubMed](http://dx.doi.org/10.1093/nar/25.5.955)
54. 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**:3100-3108. [PubMed](http://dx.doi.org/10.1093/nar/gkm160)  
<http://dx.doi.org/10.1093/nar/gkm160>
55. Krogh A, Larsson B, von Heijni G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001; **305**:567-580. [PubMed](http://dx.doi.org/10.1006/jmbi.2000.4315)  
<http://dx.doi.org/10.1006/jmbi.2000.4315>
56. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004; **340**:783-795. [PubMed](http://dx.doi.org/10.1016/j.jmb.2004.05.028)  
<http://dx.doi.org/10.1016/j.jmb.2004.05.028>
57. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, et al. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 2008; **9**:75-89. [PubMed](http://dx.doi.org/10.1186/1471-2164-9-75)  
<http://dx.doi.org/10.1186/1471-2164-9-75>
58. <http://crispr.u-psud.fr/Server>
59. Bose M, Barber RD. Prophage Finder: a prophage loci prediction tool for prokaryotic genome sequences. *In Silico Biol* 2006; **6**:223-227. [PubMed](http://dx.doi.org/10.1007/s10264-006-0001-1)