

# *Peptoniphilus lacydonensis* sp. nov., a new human-associated species isolated from a patient with chronic refractory sinusitis

M. Beye<sup>1</sup>, S. Bakour<sup>1</sup>, E. Le Dault<sup>1</sup>, J. Rathored<sup>1</sup>, C. Michelle<sup>2</sup>, F. Cadoret<sup>2</sup>, D. Raoul<sup>2,3</sup> and P.-E. Fournier<sup>1</sup>

1) UMR VITROME, Aix-Marseille Université, IRD, Service de Santé des Armées, 2) UMR MEPHI, Aix-Marseille Université, IRD, Assistance Publique-Hôpitaux de Marseille, Institut Hospitalo-Universitaire Méditerranée-Infection, Marseille, France and 3) Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdul Aziz University, Jeddah, Saudi Arabia

## Abstract

Strain ELI<sup>T</sup> was isolated from a sinus sample of an 85-year-old man with chronic refractory sinusitis complicating ethmoidal adenocarcinoma. We studied its phenotypic and genomic characteristics. This is a Gram stain-positive, anaerobic and microaerophilic coccus. Cells are catalase negative, nonmotile and non-spore forming. The major fatty acids are saturated hexadecanoic acid (34%), unsaturated 9-octadecenoic acid (32%) and 9.12-octadecadienoic acid (21%). The 1.86 Mb long genome exhibits a 29.9% G+C content and contains 1750 protein-coding and 43 RNA genes. On the basis of these data, we propose the creation of the new human-associated bacterial species *Peptoniphilus lacydonensis* sp. nov.

© 2018 The Authors. Published by Elsevier Ltd.

**Keywords:** Chronic refractory sinusitis, culturomics, genome, *Peptoniphilus lacydonensis*, sinus adenocarcinoma, taxonomy

**Original Submission:** 16 January 2018; **Revised Submission:**

1 February 2018; **Accepted:** 14 February 2018

**Article published online:** 28 February 2018

**Corresponding author:** P.-E. Fournier, UMR MEPHI, Aix-Marseille Université, IRD, AP-HM, IHU Méditerranée-Infection, 19-21 Bd Jean Moulin, 13005 Marseille, France.

**E-mail:** pierre-edouard.fournier@univ-amu.fr

## Introduction

As of 31 August 2017, the genus *Peptoniphilus* contains 15 species with validly published names (<http://www.bacterio.net/peptoniphilus.html>). Bacteria from the genus *Peptoniphilus* are Gram-positive cocci that do not form spores and are strictly anaerobic and catalase negative [1]. All species have been isolated from human clinical material except three species that were isolated from animals. *Peptoniphilus asaccharolyticus* [1], *P. olsenii* [2], *P. gorbachii* [2], *P. duerdenii* [3], *P. koenoenieniae* [3], *P. timonensis* [4] and *P. catoniae* [5] were isolated from human clinical specimens. In addition, *P. harei* [1], *P. ivorii* [1], *P. coxii* [6] and *P. tyrelliae* [6] were isolated from human clinical infections. Finally, *P. indolicus* [1], *P. methioninivorax* [7] and *P. stercorisuis* [8] were isolated from cattle, retail ground beef and a swine manure storage tank, respectively.

Recently, next-generation sequencing technology, which enables sequencing the whole genome of a prokaryote in a short time, and mass spectrometric analysis of bacteria have given easy access to genetic and proteomic information [9]. Therefore, we propose a polyphasic approach to describe new bacterial species that combines genomic properties with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) spectrum and phenotypic characteristics [10].

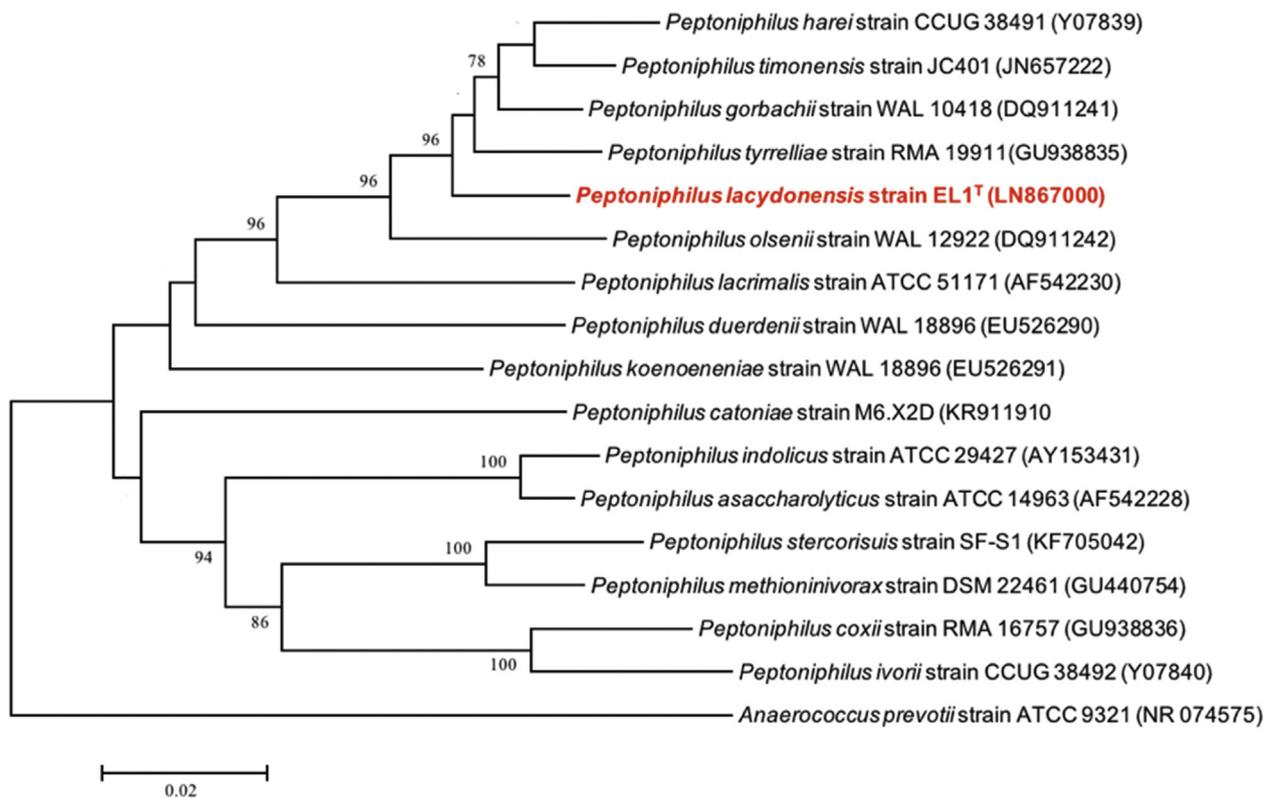
Herein we describe strain ELI<sup>T</sup> (= CSUR P2013 = DSM 100661), which was isolated from a sinus specimen of an 85-year-old patient with chronic refractory sinusitis. We present a summary classification and set of features for this strain, including the description of its complete genome sequencing and annotation that allow its description as a new *Peptoniphilus* species. These characteristics support the circumscription of the new species *Peptoniphilus lacydonensis* sp. nov.

## Results

### Strain identification by MALDI-TOF MS and 16S rRNA gene sequencing

No significant MALDI-TOF MS score was obtained for strain ELI<sup>T</sup> against the Bruker Daltonics (Leipzig, Germany) database, suggesting that our strain was not a member of any known species. The reference mass spectrum of strain ELI<sup>T</sup> was added to our database (Supplementary Fig. S1). A gel view comparing the spectrum of strain ELI<sup>T</sup> with those of other *Peptoniphilus* species is shown in Supplementary Fig. S2.

To identify strain ELI, the 16S rRNA gene was sequenced, and the obtained sequence (GenBank accession no. LN867000)



**FIG. 1.** Phylogenetic tree obtained by comparison of 16S rRNA gene sequences showing position of *Peptoniphilus lacydonensis* sp. nov. strain EL1<sup>T</sup> relative to other type strains within genus *Peptoniphilus*. GenBank accession numbers are indicated in parentheses. Sequences were aligned using default parameters in MUSCLE; phylogenetic inferences were obtained using maximum-likelihood method with 1000 replicates within MEGA6 software. Scale bar represents 2% nucleotide sequence divergence.

showed a 96.76% similarity with *P. tyrrelliae* strain RMA 19911<sup>T</sup> (accession number GU938835), the phylogenetically closest bacterial species with a validly published name (Fig. 1). The general characteristics of strain EL1<sup>T</sup> are presented in Supplementary Table S1. According to the similarity level threshold defined by Kim et al. [11], we propose strain EL1<sup>T</sup> as type strain of a new species within the genus *Peptoniphilus*, for which we suggest the name *Peptoniphilus lacydonensis* sp. nov. (Supplementary Table S1).

### Phenotypic description

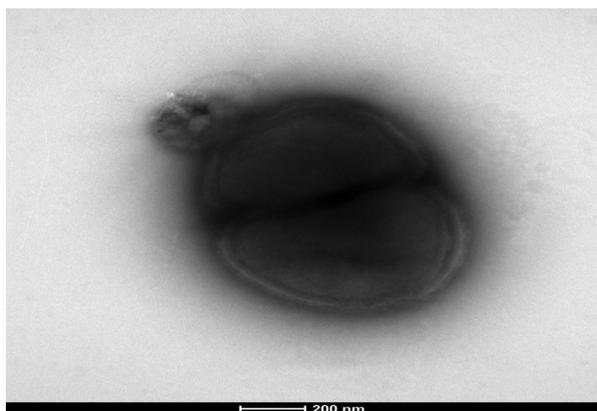
Growth was obtained at 37°C but weakly at 29°C, with optimal growth at 37°C after 48 hours' incubation. The pH range for growth was 6 to 8.5, and growth occurred in the presence of 0 to 5 g/L NaCl. Strain growth was observed in both anaerobic and microaerophilic conditions, but not in aerobic atmosphere. Colonies were translucent and grey; they exhibited a diameter of 0.5 to 1 mm after 72 hours' incubation on 5% sheep's blood-enriched Columbia agar (bioMérieux, Marcy l'Etoile, France). Bacterial cells were Gram-positive cocci (Supplementary Fig. S3) with a mean diameter of 0.73 µm, and

they were mostly grouped in pairs (Fig. 2). Cells were not motile and were non-spore forming.

### Biochemical characterization and antibiotic susceptibility

Catalase and oxidase production reactions were negative for strain EL1<sup>T</sup>. Using an API 20A strip (bioMérieux), a positive reaction was observed for indole formation. No acid production was observed from D-glucose, D-mannitol, D-lactose, D-saccharose, D-maltose, salicin, D-xylose, L-arabinose, glycerol, D-cellobiose, D-mannose, D-melezitose, D-raffinose, D-sorbitol, L-rhamnose or D-trehalose. Urease, β-glucosidase and protease activities were absent.

Using an API ZYM strip (bioMérieux), a weak reaction was observed for naphthol-AS-BI-phosphohydrolase. Negative reactions were noted for alkaline phosphatase, esterase, esterase-lipase, lipase, leucine arylamidase, valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, phosphatase acid, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, α-mannosidase, α-fucosidase and N-acetyl-β-glucosaminidase.



**FIG. 2.** Transmission electron microscopy of *Peptoniphilus lacydonensis* strain ELI<sup>T</sup> using Tecnai G20 electron microscope (FEI Company). Scale bar = 200 nm.

Using an API 50 CH strip (bioMérieux), negative reactions were observed for the fermentation of glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl-β-D-xylopyranoside, D-galactose, D-glucose, D-fructose, D-mannose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-αD-xylopyranoside, methyl-αD-glucopyranoside, N-acetylglucosamine, amygdalin, arbutin, esculin, ferric citrate, salicin, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-saccharose, D-trehalose, inulin, D-melezitose, D-raffinose, amidone, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate and potassium-5-ketogluconate.

The major fatty acid was saturated hexadecanoic acid (34%), followed by unsaturated 9-octadecenoic acid (32%) and 9,12-octadecadienoic acid (21%). Both unsaturated and saturated fatty acids were detected (Supplementary Table S2).

*Peptoniphilus lacydonensis* strain ELI<sup>T</sup> was susceptible to amoxicillin (minimum inhibitory concentration 0.125 µg/mL), amoxicillin/clavulanic acid (0.125), ticarcillin (1), cefepime (1), imipenem (0.125), gentamicin (0.5), doxycycline (0.25), tigecycline (0.125), clindamycin (0.125), fosfomycin (0.125), rifampicin (0.06), ciprofloxacin (0.125), erythromycin (0.25) and vancomycin (0.5), but resistant to trimethoprim/sulfamethoxazole (>16). Table 1 lists the antibiotic susceptibility of closely related species.

**Genome properties**

The final assembly was made of three contigs. The genome is 1 863 688 bp long (one chromosome, but no plasmid) with a 29.9% G+C content (Fig. 3, Supplementary Table S3). Of the 1793 predicted genes, 1750 were protein-coding genes and 43 were RNAs (two genes are 5S rRNA, one gene is 16S rRNA, one gene is 23S rRNA and 39 genes are tRNA genes). A total of 1321 genes (75.49%) were assigned a putative function. Twenty genes were identified as ORFans (1.14%). The remaining genes were annotated as hypothetical proteins (429 genes, 24.51%) (Supplementary Table S3).

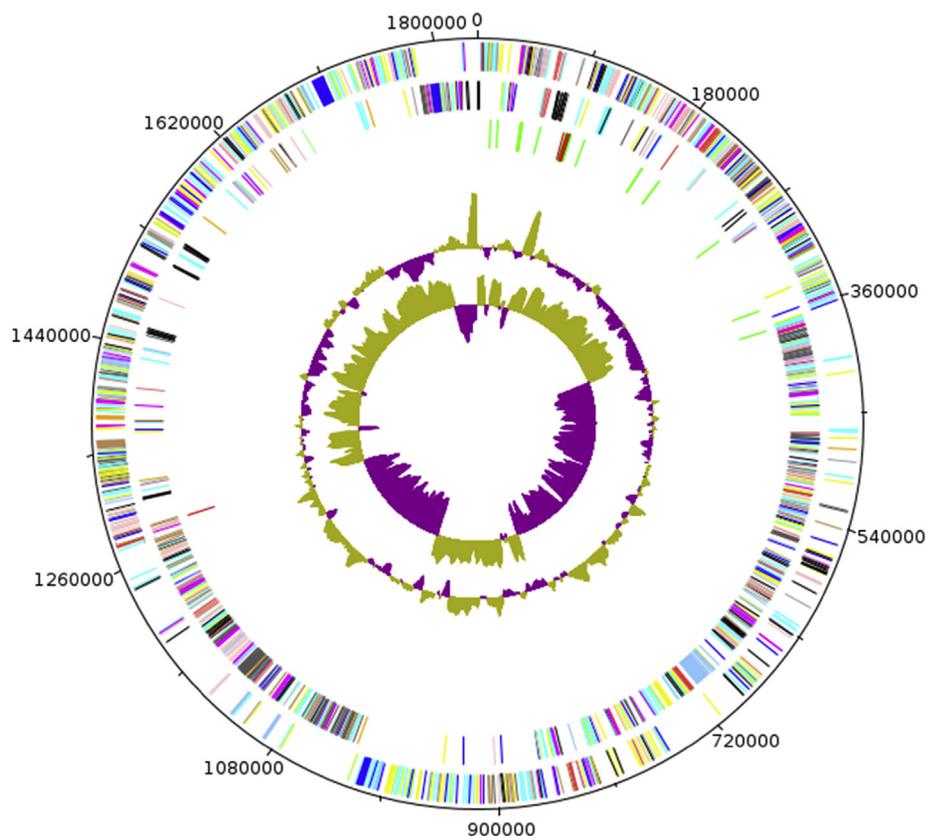
Table 2 shows the distribution of genes into Clusters of Orthologous Groups (COGs) functional categories.

The genome sequence has been deposited in GenBank under accession numbers FNWF01000001 to FNWF01000003.

**TABLE 1.** Differential phenotypic characteristics of *Peptoniphilus lacydonensis* strain ELI<sup>T</sup> and related type strains (*Peptoniphilus lacrimalis*, *Peptoniphilus asaccharolyticus*, *Peptoniphilus indolicus*, *Peptoniphilus harei*, *Peptoniphilus ivorii* and *Peptoniphilus duerdenii* [1,6])

Property	<i>P. lacydonensis</i>	<i>P. lacrimalis</i>	<i>P. asaccharolyticus</i>	<i>P. indolicus</i>	<i>P. harei</i>	<i>P. ivorii</i>	<i>P. duerdenii</i>
Gram stain	Positive	Positive	Positive	Positive	Positive	Positive	Positive
Oxygen requirement	Anaerobic and microaerophilic	Obligately anaerobic	Obligately anaerobic	Obligately anaerobic	Obligately anaerobic	Obligately anaerobic	Obligately anaerobic
Motility	-	-	-	-	-	-	-
Spore formation	-	-	-	-	-	-	-
Production of:							
Catalase	-	-	-	+	+	-	-
Indole	+	-	D	+	D	-	+
Urease	-	-	-	-	-	-	-
Alkaline phosphatase	-	-	-	+	-	-	-
Fermentation of:							
Glucose	-	-	-	-	-	-	-
Lactose	-	-	-	-	-	-	-
Raffinose	-	-	-	-	-	-	-
Mannose	-	-	-	-	-	-	-
Activity of:							
α-Galactosidase	-	-	-	-	-	-	-
β-Galactosidase	-	-	-	-	-	-	-
α-Glucosidase	-	-	-	-	-	-	-
β-Glucosidase	-	-	-	-	-	-	-
Leucine arylamidase	-	+	D	+	D	-	+
Habitat	Human	Human	Human	Animal	Human	Human	Human

+, positive result; -, negative result; D, strain dependent.



**FIG. 3.** Graphical circular map of chromosome of *Peptoniphilus lacydonensis* sp. nov. strain ELI<sup>T</sup>. From outside to centre: genes on forward strand coloured by COGs category (only genes assigned to COGs), genes on reverse strand coloured by COGs category (only genes assigned to COGs), RNA genes (tRNAs green, rRNAs red), GC content and GC skew. COGs, Clusters of Orthologous Groups database.

**TABLE 2.** Number of genes associated with 25 general COGs functional categories

Code	Value	% <sup>a</sup>	Description
J	138	7.89	Translation
A	0	0	RNA processing and modification
K	89	5.09	Transcription
L	123	7.03	Replication, recombination and repair
B	0	0	Chromatin structure and dynamics
D	17	0.97	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	61	3.49	Defense mechanisms
T	42	2.4	Signal transduction mechanisms
M	59	3.37	Cell wall/membrane biogenesis
N	3	0.17	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	13	0.74	Intracellular trafficking and secretion
O	51	2.91	Posttranslational modification, protein turnover, chaperones
X	7	0.4	Mobilome: prophages, transposons
C	85	4.86	Energy production and conversion
G	36	2.06	Carbohydrate transport and metabolism
E	111	6.34	Amino acid transport and metabolism
F	52	2.97	Nucleotide transport and metabolism
H	52	2.97	Coenzyme transport and metabolism
I	41	2.34	Lipid transport and metabolism
P	75	4.29	Inorganic ion transport and metabolism
Q	9	0.51	Secondary metabolites biosynthesis, transport and catabolism
R	158	9.03	General function prediction only
S	111	6.34	Function unknown
—	417	23.83	Not in COGs

COGs, Clusters of Orthologous Groups database.

<sup>a</sup>Total is based on total number of protein-coding genes in annotated genome.

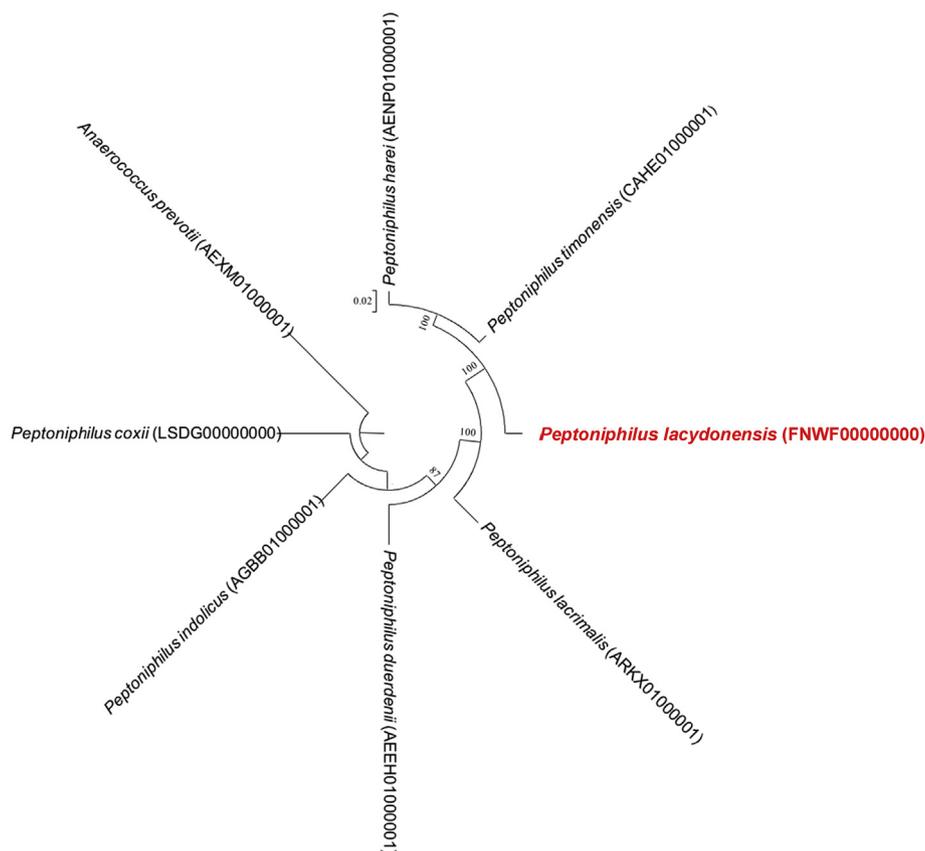
### Genome comparison

Phylogenomic analysis showed that the strain ELI<sup>T</sup> genome is closest to *Peptoniphilus harei* and *Peptoniphilus timonensis* (Fig. 4).

The draft genome of *P. lacydonensis* (1.86 Mb) has a smaller size than that of *P. lacrimalis* (1.88 Mb), *P. indolicus* (1.93 Mb), *P. harei* (1.87 Mb), and *P. duerdenii* (2.11 Mb) but is greater than those of *P. timonensis* (1.76 Mb) and *P. coxii* (1.84 Mb) (Table 3). The G+C content of *P. lacydonensis* (29.9%) was lower than that of *P. timonensis* (30.7%), *P. lacrimalis* (30.2%), *P. indolicus* (31.7%), *P. harei* (34.4%), *P. duerdenii* (34.2%) and *P. coxii* (44.6%) (Table 3).

Similarly, the gene content of *P. lacydonensis* sp. nov. (1793) is greater than those of *P. coxii*, *P. lacrimalis* and *P. harei* (1775, 1780 and 1754, respectively) but is smaller than those of *P. timonensis*, *P. indolicus* and *P. duerdenii* (1936, 2231 and 1947, respectively) (Table 3). However, the distribution of genes into COGs categories was similar in all compared genomes (Supplementary Fig. S4).

Among *Peptoniphilus* species with standing in nomenclature, average genomic identity of orthologous gene sequences (AGIOS) values ranged from 61.20% between *P. timonensis* and *P. coxii* to 84.44% between *P. timonensis* and *P. harei*. When comparing



**FIG. 4.** Phylogenomic tree obtained by comparison of whole genome sequences showing position of *Peptoniphilus lacydonensis* sp. nov. strain ELI<sup>T</sup> relative to other type strains' genomes within genus *Peptoniphilus*. GenBank accession numbers are indicated in parentheses. Sequences were aligned using default parameters in Mauve alignment programme; phylogenetic inferences were obtained using maximum-likelihood method with 1000 replicates within FastTree software.

*P. lacydonensis* sp. nov. to other species, AGIOS values were in the same range, from 62.54% with *P. coxii* to 80.54% with *P. timonensis* (Table 4). Among *Peptoniphilus* species with standing in nomenclature, digital DNA-DNA hybridization (dDDH) values ranged from 20.10% between *P. indolicus* and *P. harei* to 45.1% between *P. indolicus* and *P. coxii*. When comparing *P. lacydonensis* sp. nov. to other species, dDDH values were also in the same range, from 20.2% with *P. indolicus* to 39.5% with *P. duerdenii* (Supplementary

Table S4). Among species with standing in nomenclature, average nucleotide identity (ANI) values ranged from 64.4% between *P. timonensis* and *P. coxii* to 82.8% between *P. timonensis* and *P. harei*. When comparing *P. lacydonensis* to other species, the ANI value ranged from 66.1% between *P. lacydonensis* and *P. coxii* to 79.7% between *P. lacydonensis* and *P. timonensis* (Supplementary Table S4).

## Discussion

Strain ELI<sup>T</sup>, which was isolated from a patient with chronic refractory ethmoidal sinusitis, exhibits a close phylogenetic relatedness to *Peptoniphilus* species but with specific features, notably its ability to grow in microaerophilic atmosphere and a 16S rRNA gene sequencing similarity of 96.76% with *P. tyrrelliae* [6], which is lower than the threshold recommended by Kim et al. [11] to delineate a new species. Phenotypic characterization, MALDI-TOF MS, 16S rRNA gene sequencing and comparative genomic analyses among close phylogenetic relatives enabled the

**TABLE 3.** Genome comparison of closely related species to *Peptoniphilus lacydonensis* strain ELI

Organism name	GenBank Accession no.	Size (Mb)	G+C %	No. of proteins	Total genes
<i>Peptoniphilus lacydonensis</i>	FNWF00000000	1.86	29.9	1750	1793
<i>Peptoniphilus coxii</i>	LSDG00000000	1.84	44.6	1738	1775
<i>Peptoniphilus duerdenii</i>	AEEH00000000	2.08	34.2	1920	1947
<i>Peptoniphilus harei</i>	AENP00000000	1.84	34.4	1719	1754
<i>Peptoniphilus indolicus</i>	AGBB00000000	2.10	31.7	2205	2231
<i>Peptoniphilus lacrimalis</i>	ARKX00000000	1.85	30.2	1738	1780
<i>Peptoniphilus timonensis</i>	CAHE00000000	1.76	30.7	1914	1936

**TABLE 4.** AGIOS values (lower left), numbers of orthologous proteins shared between genomes (upper right) and numbers of proteins per genome (bold numbers)

	<i>Peptoniphilus lacydonensis</i>	<i>Peptoniphilus timonensis</i>	<i>Peptoniphilus duerdenii</i>	<i>Peptoniphilus harei</i>	<i>Peptoniphilus indolicus</i>	<i>Peptoniphilus lacrimalis</i>	<i>Peptoniphilus coxii</i>
<i>P. lacydonensis</i>	<b>1750</b>	1008	943	1138	937	1001	877
<i>P. timonensis</i>	80.54%	<b>1914</b>	780	1029	826	900	743
<i>P. duerdenii</i>	69.20%	67.78%	<b>1920</b>	965	852	941	864
<i>P. harei</i>	78.78%	84.44%	69.48%	<b>1719</b>	964	1074	879
<i>P. indolicus</i>	68.96%	68.16%	67.21%	67.25%	<b>2205</b>	878	792
<i>P. lacrimalis</i>	72.20%	71.11%	69.52%	71.21%	68.54%	<b>1738</b>	835
<i>P. coxii</i>	62.54%	61.20%	63.71%	62.19%	61.45%	62.24%	<b>1738</b>

AGIOS, average genomic identity of orthologous gene sequences.

identification of strain ELI<sup>T</sup> as a new species within the genus *Peptoniphilus*, with its closest relatives being *P. tyrrelliae*, *P. harei*, *P. timonensis* and *P. gorbachii*. ANI, AGIOS and dDDH values confirmed *P. lacydonensis* sp. nov. as a new species.

#### Taxonomic and nomenclatural proposals: description of *Peptoniphilus lacydonensis* sp. nov.

*Peptoniphilus lacydonensis* (la.cy.do.ne'n.sis, L. gen. masc. n., from *lacydonensis*, derived from Greek Λακυδών, the name of the ancient port of Marseille, where the patient from which strain ELI<sup>T</sup> was obtained, lived).

Colonies obtained after 72 hours' incubation on 5% sheep's blood-enriched Columbia agar exhibit a diameter of 0.5 to 1 mm. Bacterial cells are Gram-positive, non-spore-forming and nonmotile cocci. Strain ELI<sup>T</sup> optimally grows at 37°C after 48 hours in anaerobic atmosphere but is also able to grow in microaerophilic conditions. Catalase and oxidase production reactions are negative. The major fatty acid is the saturated hexadecanoic acid.

Using API 20A and API ZYM strips, positive reactions are observed for indole formation and naphthol-AS-BI-phosphohydrolase, respectively. Negative reactions are observed for alkaline phosphatase, esterase, esterase-lipase, lipase, leucine arylamidase, valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, phosphatase acid, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, α-mannosidase, α-fucosidase and *N*-acetyl-β-glucosaminidase using API ZYM strips. In addition, no acid production was observed from D-glucose, D-mannitol, D-lactose, D-saccharose, D-maltose, salicin, D-xylose, L-arabinose, glycerol, D-cellobiose, D-mannose, D-melezitose, D-raffinose, D-sorbitol, L-rhamnose and D-trehalose, and no activity was observed for urease, β-glucosidase or protease using API 20A strips. Negative reactions are observed for the fermentation of glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl-β-D-xylopyranoside, D-galactose, D-glucose, D-fructose, D-mannose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-α-D-xylopyranoside,

methyl-α-D-glucopyranoside, *N*-acetylglucosamine, amygdalin, arbutin, esculin ferric citrate, salicin, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-saccharose, D-trehalose, inulin, D-melezitose, D-raffinose, amidone, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate and potassium-5-ketogluconate using API 50 CH strips.

The genome is 1 863 688 bp long with a 29.9% G+C content and is accessible under GenBank accession numbers FNWF01000001 to FNWF01000003. The 16S rRNA gene sequence is also accessible in GenBank under number LN867000. The type strain ELI<sup>T</sup> (= CSUR P2013 = DSM 100661) was isolated from the ethmoidal sinus of a patient with chronic refractory sinusitis complicating an adenocarcinoma in Marseille, France.

## Material and methods

### Sample information and ELI<sup>T</sup> strain isolation

In February 2015, a sinus sample from an 85-year-old man with chronic refractory sinusitis complicating an ethmoidal adenocarcinoma was collected in Marseille, France. The patient provided written informed consent, and the study was validated by the Sud Méditerranée I ethics committee under number 15-12. The study was also registered at ClinicalTrials.gov as NCT02407275. After 10 days' culture in an anaerobic blood culture vial that contained 5% sheep's blood-enriched and sterile-filtered sheep rumen, we grew a bacterial strain that we named ELI<sup>T</sup> (= CSUR P2013 = DSM 100661). The strain was subsequently subcultured on 5% sheep's blood-enriched Columbia agar (bioMérieux) in anaerobic atmosphere at 37°C.

### Strain identification by MALDI-TOF MS

To identify the strain, MALDI-TOF MS protein analysis was performed as previously described using a Bruker Microflex spectrometer [12]. Twelve spectra obtained from 12 individual colonies were imported into MALDI BioTyper 2.0 software (Bruker) and analysed by standard pattern matching (with

default parameter settings) against the main spectra of 8500 bacteria, including those of 15 *Peptoniphilus* species. Spectra were recorded in the positive linear mode for the mass range from 2000 to 20 000 Da (parameter settings: ion source I (ISI), 20 kV; IS2, 18.5 kV; lens, 7 kV). A spectrum was obtained after 675 shots with variable laser power. The time of acquisition was between 30 seconds and 1 minute per spot. From the resulting scores, the tested species may or may not be identified when compared to the instrument's database; a score  $\geq 2$  with a validly published species enables identification at the species level; a score  $\geq 1.7$  and  $< 2$  allows identification at the genus level; and a score  $< 1.7$  does not enable any identification.

#### Strain identification by 16S rRNA gene sequencing

The 16S rRNA gene nucleotide sequence was obtained by amplification and sequencing using the fD1 and rP2 primers as previously described [13]. Kim et al. [11] determined a 98.65% similarity level threshold to define a new species without performing DNA-DNA hybridization (DDH). A BLASTN was then systematically performed (<http://blast.ncbi.nlm.nih.gov/gate1.inist.fr/Blast.cgi>). Sequences were aligned and phylogenetic inferences obtained using MUSCLE [14] and MEGA (maximum-likelihood method) [15] software, respectively.

#### Growth conditions, phenotypic tests and biochemical characterization

The bacterium was subcultured on 5% sheep's blood-enriched Columbia agar (bioMérieux) and incubated for 24 hours at 37°C in anaerobic conditions. Growth was tested at different temperatures (29, 37 and 55°C). Strain growth was tested on 5% sheep's blood-enriched Columbia agar (bioMérieux) under anaerobic and microaerophilic conditions using the GENbag anaer and GENbag microaer systems, respectively (bioMérieux), and under aerobic atmosphere, with or without 5% CO<sub>2</sub>. Growth was tested for salt tolerance, with 0 to 5%, 50% and 100% (w/v) NaCl. The pH range for growth was tested at pH 6.5 and 8.5 using trypticase soy agar. API 20 A, API ZYM and API 50CH (bioMérieux) strips were used to perform biochemical assays. Oxidase (Becton Dickinson, Franklin Lakes, NJ, USA) and catalase (bioMérieux) assays were performed separately. Cellular fatty acid methyl ester analysis was performed by gas chromatography/mass spectrometry as previously described [16]. Antibiotic susceptibility was tested using the standard disc diffusion procedure on Mueller-Hinton agar with 5% blood using SirScan Discs antibiotics (i2a, Montpellier, France) as described by the European Committee on Antimicrobial Susceptibility Testing (EUCAST) (<http://www.eucast.org>). The results were interpreted according to EUCAST breakpoints.

Electronic microscopy was performed with a formvar carbon 400-mesh nickel grid (Electron Microscopy Sciences,

Hatfield, PA, USA) on which was deposited for 60 seconds a 3.5 µL drop of bacterial suspension. After drying on filter paper, bacteria were immediately stained with 1% ammonium molybdate (ACROS Organics, Illkirch, France) for 1 second. Electron micrographs were then acquired on a Tecnai G20 transmission electron microscope (FEI Company, Hillsboro, OR, USA) at an operating voltage of 200 kV.

#### Genomic DNA preparation

Strain ELI<sup>T</sup> was cultured on 5% sheep's blood-enriched Columbia agar (bioMérieux) at 37°C in an anaerobic atmosphere. Bacteria grown on three petri dishes were resuspended in 400 µL of Tris-EDTA (TE) buffer. Then 200 µL of this suspension was diluted in 1 mL TE buffer for lysis treatment, which included a 30-minute incubation with 2.5 µg/µL lysozyme at 37°C, followed by an overnight incubation with 20 µg/µL proteinase K at 37°C. Extracted DNA was then purified using three successive phenol/chloroform extractions and ethanol precipitations at -20°C overnight. After centrifugation, the DNA was resuspended in 160 µL TE buffer.

#### Genome sequencing and assembly

Genomic DNA (gDNA) of strain ELI<sup>T</sup> was sequenced using a MiSeq sequencer (Illumina, San Diego, CA, USA) with the mate-pair strategy. The gDNA was barcoded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina).

The gDNA was quantified by a Qubit assay using the high-sensitivity kit (Life Technologies, Carlsbad, CA, USA) to 120.3 ng/µL. The mate-pair library was prepared with 1.5 µg of gDNA using the Nextera Mate Pair (Illumina) guide. The gDNA sample was simultaneously fragmented and tagged with a mate-pair junction adapter. The pattern of fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 labchip. The DNA fragments ranged in size from 1.5 to 11 kb, with an optimal size at 5.316 kb. No size selection was performed, and 640.3 ng of tagged fragments was circularized. The circularized DNA was mechanically sheared to small fragments with an optimal size of 1550 bp on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent), and the final concentration library was measured at 15.44 nmol/L. The libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and sequencing run were performed in a single 2 × 301 bp run. Total information of 7.3 Gb was obtained from a 511K/mm<sup>2</sup> cluster density, with clusters passing quality control filters of 97.0% (12 079 000 passing filter paired reads). Within

this run, the index representation for strain ELI<sup>T</sup> was determined to be 7.71%. The 931 514 paired reads were trimmed, then assembled using SPAdes [17].

### Genome annotation and comparison

Open reading frames (ORFs) were predicted using Prodigal [18] with default parameters. However, the predicted ORFs were excluded if they spanned a sequencing gap region. The predicted bacterial protein sequences were searched against the GenBank [19] and COGs databases using BLASTP. The tRNAs and rRNAs were predicted using the tRNAScan-SE [20] and RNAMmer [21] tools, respectively. Signal peptides and numbers of transmembrane helices were predicted using SignalP [22] and TMHMM [23], respectively. Mobile genetic elements were predicted using PHAST [24] and RAST [25]. The CRISPRFinder algorithm was used to identify putative CRISPR (clustered regularly interspaced short palindromic repeat) loci [26]. ORFans were identified if their BLASTP *E* value was lower than 1e-03 for an alignment length >80 amino acids. If alignment lengths were <80 aa, we used an *E* value of 1e-05. Such parameter thresholds have been used in previous work to define ORFans. Artemis [27] and DNAPlotter [28] were used for data management and visualization of genomic features, respectively.

A phylogenomic tree was constructed using an FastTree [29] based on Mauve programme genome alignment [30].

To evaluate the degree of genomic similarity at the nucleotide level among studied strains, we determined two parameters: dDDH and AGIOS. The former exhibits a high correlation with DDH, using Genome-to-Genome Distance Calculator (GGDC), formula 2, software [31]. The latter was designed to be independent from DDH. It combines Proteinortho [32] to detect orthologous proteins between genomes compared two by two, then retrieves the corresponding genes, and then determines the mean percentage of nucleotide sequence identity among orthologous ORFs using the Needleman-Wunsch global alignment algorithm [10]. ANI at the genome level between *P. lacydonensis* and the other species *P. lacrimalis* (ARKX01000001.1), *P. indolicus* (AGBB01000001), *P. duerdenii* (AEEH01000001), *P. harei* (AENP01000001.1), *P. coxii* (LSDG00000000) and *P. timonensis* (CAHE00000000) was estimated using the OrthoANI software [33].

### Acknowledgements

The authors thank the Xegen Company ([www.xegen.fr](http://www.xegen.fr)) for automating the genomic annotation process. This study was funded by the Fondation Méditerranée Infection and the French Agence Nationale de la Recherche under reference Investissements d'Avenir Méditerranée Infection ANR-10-IAHU-03.

### Conflict of Interest

None declared.

### Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.nmni.2018.02.007>.

### References

- [1] Ezaki T, Kawamura Y, Li N, Li ZY, Zhao L, Shu S. Proposal of the genera *Anaerococcus* gen. nov., *Peptoniphilus* gen. nov. and *Gallicola* gen. nov. for members of the genus *Peptostreptococcus*. *Int J Syst Evol Microbiol* 2001;51:1521–8.
- [2] Song Y, Liu C, Finegold SM. *Peptoniphilus gorbachii* sp. nov., *Peptoniphilus olsenii* sp. nov., and *Anaerococcus murdochii* sp. nov. isolated from clinical specimens of human origin. *J Clin Microbiol* 2007;45:1746–52.
- [3] Ulger-Toprak N, Lawson PA, Summanen P, O'Neal L, Finegold SM. *Peptoniphilus duerdenii* sp. nov. and *Peptoniphilus koenoeneniensis* sp. nov., isolated from human clinical specimens. *Int J Syst Evol Microbiol* 2012;62:2336–41.
- [4] Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non contiguous–finished genome sequence and description of *Peptoniphilus timonensis* sp. nov. *Stand Genomic Sci* 2012;7:1.
- [5] Patel NB, Tito RY, Obregón-Tito AJ, O'Neal L, Trujillo-Villaroel O, Marin-Reyes L, et al. *Peptoniphilus cationiae* sp. nov., isolated from a human faecal sample from a traditional Peruvian coastal community. *Int J Syst Evol Microbiol* 2016;66:2019–24.
- [6] Citron DM, Tyrrell KL, Goldstein EJ. *Peptoniphilus coxii* sp. nov. and *Peptoniphilus tyrelliae* sp. nov. isolated from human clinical infections. *Anaerobe* 2012;18:244–8.
- [7] Rooney AP, Swezey JL, Pukall R, Schumann P, Spring S. *Peptoniphilus methionivorax* sp. nov., a Gram-positive anaerobic coccus isolated from retail ground beef. *Int J Syst Evol Microbiol* 2011;61:1962–7.
- [8] Johnson CN, Whitehead TR, Cotta MA, Rhoades RE, Lawson PA. *Peptoniphilus stercorisuis* sp. nov., isolated from a swine manure storage tank and description of *Peptoniphilaceae* fam. nov. *Int J Syst Evol Microbiol* 2014;64:3538–45.
- [9] Welker M, Moore ER. Applications of whole-cell matrix-assisted laser–desorption/ionization time-of-flight mass spectrometry in systematic microbiology. *Syst Appl Microbiol* 2011;34:2–11.
- [10] Ramasamy D, Mishra AK, Lagier JC, Padhmanabhan R, Rossi M, Sentausa E, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. *Int J Syst Evol Microbiol* 2014;64:384–91.
- [11] Kim M, Oh HS, Park SC, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 2014;64:346–51.
- [12] Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, et al. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin Infect Dis* 2009;49:543–51.
- [13] Drancourt M, Bollet C, Carlloz A, Martelin R, Gayral JP, Raoult D. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. *J Clin Microbiol* 2000;38:3623–30.
- [14] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–7.

- [15] Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis, version 6.0. *Mol Biol Evol* 2013;30:2725–9.
- [16] Tidjani Alou M, Khelaifia S, Michelle C, Andrieu C, Armstrong N, Bittar F, et al. *Anaerococcus rubiinfantis* sp. nov., isolated from the gut microbiota of a Senegalese infant with severe acute malnutrition. *Anaerobe* 2016;40:85–94.
- [17] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–77.
- [18] Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform* 2010;11:119.
- [19] Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res* 2015;43:D30–5.
- [20] Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 1997;25:955–64.
- [21] Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007;35:3100–8.
- [22] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004;340:783–95.
- [23] Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001;305:567–80.
- [24] Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: a fast phage search tool. *Nucleic Acids Res* 2011;39(Web Server issue):W347–52.
- [25] Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 2008;9:75.
- [26] Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 2007;35:W52–7.
- [27] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. *Bioinforma* 2000;16:944–5.
- [28] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. *Bioinforma* 2009;25:119–20.
- [29] Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009;26:1641–50.
- [30] Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 2004;14:1394–403.
- [31] Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinform* 2013;14:60.
- [32] Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinform* 2011;12:124.
- [33] Ouk Kim Y, Chun J, Lee I, Park SC. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 2016;66:1100–3.