



HAL
open science

**Noncontiguous finished genome sequence and
description of *Streptococcus timonensis* sp. nov.
isolated from the human stomach**

D. Ricaboni, M. Mailhe, C. Lagier, C. Michelle, N. Armstrong, F. Bittar, V.
Vitton, A. Benezech, Didier Raoult, M. Million

► **To cite this version:**

D. Ricaboni, M. Mailhe, C. Lagier, C. Michelle, N. Armstrong, et al.. Noncontiguous finished genome sequence and description of *Streptococcus timonensis* sp. nov. isolated from the human stomach. *New Microbes and New Infections*, 2016, 15, pp.77 - 88. 10.1016/j.nmni.2016.11.013 . hal-01478508

HAL Id: hal-01478508

<https://amu.hal.science/hal-01478508>

Submitted on 9 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

Noncontiguous finished genome sequence and description of *Streptococcus timonensis* sp. nov. isolated from the human stomach

D. Ricaboni^{1,3}, M. Mailhe¹, J.-C. Lagier¹, C. Michelle¹, N. Armstrong¹, F. Bittar¹, V. Vitton², A. Benezech², D. Raoult^{1,4} and M. Million¹

1) *Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, CNRS (UMR 7278), IRD (198), INSERM (U1095), AMU (UM63), Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine, Aix-Marseille Université*, 2) *Service de Gastroentérologie, Hôpital Nord, Assistance Publique-Hôpitaux de Marseille, Marseille, France*, 3) *Department of Biomedical and Clinical Sciences, 3rd Division of Clinical Infectious Disease, University of Milan, Luigi Sacco Hospital, Milan, Italy* and 4) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdul Aziz University, Jeddah, Saudi Arabia*

Abstract

Strain Marseille-P2915^T, a Gram-positive, facultative anaerobic and nonmotile coccus, was isolated from the gastric lavage of a patient with severe anaemia. The 16S rRNA and *rpoB* gene comparison exhibited a sequence identity of 98.7 and 92.6% with *Streptococcus infantis* strain JCM 10157^T, respectively, collocating it within the '*Streptococcus mitis*' group. On the basis of phenotypic and genomic analysis, we propose the validation of the type strain *Streptococcus timonensis* sp. nov. Marseille-P2915^T (= DSM 103349 = CSUR P2915).

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases.

Keywords: Culturomics, human gut microbiota, stomach, *Streptococcus timonensis*, taxonogenomics

Original Submission: 23 September 2016; **Revised Submission:** 26 October 2016; **Accepted:** 4 November 2016

Article published online: 18 November 2016

Corresponding author: M. Million, Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, INSERM 1095, Marseille, France

E-mail: matthieumillion@gmail.com

Introduction

The genus *Streptococcus* comprises 116 officially recognized species (<http://www.bacterio.net/>) partially clustered into six species group (*pyogenic*, *anginosus*, *mitis*, *salivarius*, *bovis* and *mutans*) on the basis of systematic 16S rRNA sequence study [1]. The exact taxonomic classification of each species within this genus remains challenging, especially in the *mitis* group, as a result of a high genetic and phenotypic similarity shared within different species [2], in particular with different species sharing a 16S rRNA sequence identity greater than 98.7% [3]. Because species belonging to this group gather together highly virulent species involved in pathologies as meningitis, endocarditis, pneumonia and low-pathogenic commensal species, the rapid

identification of clinical isolates is mandatory. To reach this goal, several different molecular targets have been tested, including *sodA* [4], *rnpB* [5], *tuf* [6] and *groEL* [7]. The use of *rpoB* was initially proposed in our laboratory [8]. Since then, it has been validated by its use in the classification of different new streptococcal species [9–11].

Nowadays, the widespread use of matrix-assisted desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) in the clinical and research setting, coupled with the rapid development of next-generation sequencing technology, gives us a new and more complete insight into the taxonomy of the *Streptococcus* genus. In this context, a polyphasic approach to describe new bacterial species, termed taxonogenomics, was proposed in our laboratory in 2004, combining common phenotypic tests such as API strips, whole genomic sequencing and MALDI-TOF MS spectra [12].

Here we present a phenotypic and genetic comparison of 16S rRNA genes, *rpoB* genes as well as whole-genomic-level analysis which led us to propose the validation of *Streptococcus timonensis* strain Marseille-P2915^T (= DSM 103349 = CSUR P2915) as a new member within the genus *Streptococcus*.

Materials and Methods

Sample collection

In the context of a study aimed at the description of different gut microflora levels by culturomics [13], a gastric lavage sample from a 60-year-old patient was collected during a gastroscopy performed for medical reasons (severe anaemia). The patient had been receiving long-term proton pump inhibitor therapy. Informed and signed consent, approved by the Institut Fédératif de Recherche IFR48 (Faculty of Medicine, Marseille, France), under agreement 09-022, was obtained from the patient. After collection, the sample was immediately placed in an antioxidant transport medium [14] and plated within 2 hours.

Isolation and identification of strain

Strain Marseille-P2915^T's first growth was obtained in April 2016 on 5% sheep's blood-enriched Columbia agar medium (bioMérieux, Marcy l'Etoile, France) under aerobic conditions at 37°C. Once isolated on pure culture, proteomic analysis was carried out with MALDI-TOF MS as previously described [15,16] using a Microflex spectrometer (Bruker Daltonics, Bremen, Germany). Strain Marseille-P2915^T spectra was thus obtained, imported into the MALDI BioTyper software (version 3.0, Bruker) and processed by standard pattern matching (with default parameter settings) against the main spectra of 7567 bacteria. The comparison with the BioTyper database spectra enabled the rapid matching of the analysed species with those present in the database. The resulting score, if >2, enabled the identification at the species level, while a score of <1.7 did not enable any identification.

To obtain the 16S rRNA sequence of strain Marseille-P2915^T, PCR analysis was performed using a GeneAmp PCR System 2720 thermal cycler (Applied Biosystems, Bedford, MA, USA) as previously described [17]. The sequencing step was performed on an ABI Prism 3130xl Genetic Analyser capillary sequencer (Applied Biosystems). Chromas Pro 1.34 software (Technelysium, Tewantin, Australia) was used to fix the sequences, and a BLASTn (Basic Local Alignment Search Tool) search [18] was performed against the National Center for Biotechnology Information (NCBI) database (default parameters, uncultured/environmental sample sequences excluded).

The *rpoB* gene sequence of strain Marseille-P2915^T was obtained using the 1730_F and 3700_R primers as previously described [8–19]. Then a BLAST search was performed against the NCBI database using default parameters.

Phylogenetic analysis

On the basis of the BLAST research results, 16S rRNA reference sequences (<http://www.bacterio.net/streptococcus.html>)

of the closest species with standing in nomenclature were downloaded from NCBI. Because there is no officially recognized reference for *rpoB*, gene sequences were downloaded directly from the NCBI database after the BLASTn search. Sequences were then aligned using Muscle v3.8.31 with default parameters, and phylogenetic inferences were obtained using the maximum-likelihood method with 1000 bootstrap replicates within MEGA6 software.

Phenotypic and biochemical characterization

Growth conditions. Growth of the strain was tested on sheep's blood-enriched Columbia agar (bioMérieux) under anaerobic conditions using anaeroGEN (Oxoid, Basingstoke, UK), microaerophilic (CampyGen, Oxoid) and aerobic conditions. Different growth temperatures (20, 28, 37, 45 and 55°C) were also tested. The acceptance limit of salinity by strain Marseille-P2915^T was tested on Columbia agar using 10, 15 and 20% of NaCl concentrations. Moreover, seven pHs were tested: 5.0; 5.5; 6.0; 6.5; 7.0; 7.5 and 8.

Microscopy. The Gram coloration was performed using color Gram 2 kit (bioMérieux) and observed using the DMI1000 photonic microscope (Leica Microsystems, Wetzlar, Germany) with a 100× oil-immersion objective lens. The ability to produce spores was studied by thermal shock (80°C during 20 minutes). The motility test was performed by observing fresh colonies between blades and slats using a DMI1000 photonic microscope (Leica Microsystems) with a 100× oil-immersion objective lens. Transmission electron microscopic images were obtained using a Tecnai G20 (FEI Company, Limeil-Brevannes, France) at an operating voltage of 200 keV. Briefly, cells were fixed with 2.5% glutaraldehyde in 0.1 M cacodylate buffer for at least 1 hour at 4°C. A drop of cell suspension was deposited for approximately 5 minutes on glow-discharged formvar carbon film on 400 mesh nickel grids (FCF400-Ni, EMS, Hatfield, PA, USA). The grids were dried on blotting paper, and cells were negatively stained for 10 seconds with 1% ammonium molybdate solution in filtered water at room temperature.

Biochemical assays. A basic biochemical study was performed using the API Gallery systems: API ZYM, API strep and API 50CH (L medium) according to the manufacturer's instructions (bioMérieux). Oxidase (Becton Dickinson, Franklin Lakes, NJ, USA) and catalase assays (bioMérieux) were done separately.

Antibiotic susceptibility. The antibiogram profile of strain Marseille-P2915^T was obtained with the disk diffusion method following European Committee on Antimicrobial Susceptibility Testing (EUCAST) 2016 recommendations. Tested antibiotics included: penicillin G 10 IU, amoxicillin 25 µg, ceftriaxone 30 µg, erythromycin 15 IU, rifampicin 30 µg, sulfamethoxazole

23.75 µg, trimethoprim 1.25 µg, imipenem 10 µg and vancomycin 30 µg.

Fatty acids analysis. Cellular fatty acid methyl ester (FAME) analysis was performed by gas chromatography mass spectrometry (GC/MS). Two samples were prepared with approximately 40 mg of bacterial biomass per tube collected from several 5% sheep's blood-enriched Columbia agar plates. FAMES were prepared as described by Sasser [20]. GC/MS analyses were carried out as previously described [21]. FAMES were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500-SQ 8 S, PerkinElmer, Waltham, MA, USA). A spectral database search was performed using MS Search 2.0 operated with the Standard Reference Database 1A (National Institute of Standards and Technology, Gaithersburg, MD, USA) and the FAME mass spectral database (Wiley, Chichester, UK).

Genomic DNA extraction and genome sequencing and assembly

After a pretreatment by a lysozyme incubation at 37°C, DNA was extracted on the EZ1 biorobot (Qiagen, Germantown, MD, USA) with a EZ1 DNA tissues kit. The elution volume was 50 µL. Genomic DNA (gDNA) was quantified by a Qubit assay with the high-sensitivity kit (Life Technologies, Carlsbad, CA, USA) to 46 ng/µL. gDNA of strain Marseille-P2915^T was sequenced with the MiSeq Technology apparatus (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 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 the 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 of 6.920 kb. No size selection was performed, and 678.5 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with optima on a bimodal curve at 675 and 1445 bp on the Covaris S2 device in T6 tubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies), and the final concentration library was measured at 42.2 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 39-hour

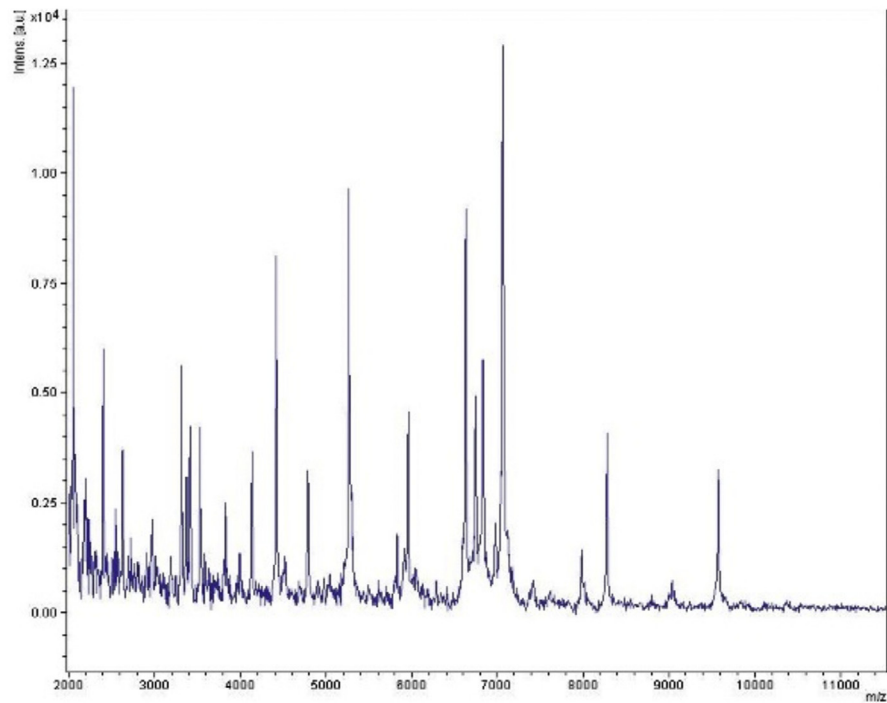
run at a 2 × 151 bp read length. Total information of 7.9 Gb was obtained from a 863K/mm² cluster density with a cluster passing quality control filters of 94% (15 627 000 passing filter paired reads). Within this run, the index representation for *Streptococcus timonensis* strain Marseille-P 2915^T was determined to be 11.94%. The 1 865 795 paired reads were trimmed, then assembled in four scaffolds.

Genome annotation and comparison

Open reading frames (ORFs) were predicted using Prodigal (<http://prodigal.ornl.gov/>) 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 [22] and Clusters of Orthologous Groups (COGs) databases using BLASTP. The tRNAs and rRNAs were predicted using the tRNAScan-SE [23] and RNAMmer [24] tools, respectively. Signal peptides and number of transmembrane helices were predicted using SignalP [25] and TMHMM [26], respectively. Mobile genetic elements were predicted using PHAST [27] and RAST [28]. ORFans were identified if their BLASTP E-value was lower than 1e-03 for an alignment length greater than 80 amino acids. If alignment lengths were smaller than 80 amino acids, an E value of 1e-05 was used. Such parameter thresholds have already been used in previous works to define ORFans. Artemis [29] and DNA Plotter [30] were used for data management and visualization of genomic features, respectively. The Mauve alignment tool (version 2.3.1) was used for multiple genomic sequence alignment [31].

The mean level of nucleotide sequence similarity at the genome level between strain Marseille-P2915^T and other bacteria (*Streptococcus oralis* strain ATCC 35037 (ADMV00000000.1); *Streptococcus infantis* ATCC 700779 (AEVD00000000.1); *Streptococcus pseudopneumoniae* ATCC BAA960 (AICS00000000.1); *Streptococcus sanguinis* SK1 (CP000387.1); *Streptococcus mitis* ATCC 49456 (AEDX00000000.1); *Streptococcus parasanguinis* ATCC 15912 (CP002843.1); *Streptococcus tigurinus* strain AZ_3a (AORU00000000.1); *Streptococcus pneumoniae* strain R6 (AE007317.1); *Streptococcus dentisani* strain 7747 (CAUK00000000.1)) was estimated by average genomic identity of orthologous gene sequences (AGIOS) software [12]. Overall, this software combines two other software packages: Proteinortho [32] (to detect orthologous proteins between genomes compared two by two, then to retrieve the corresponding genes) and the Needleman-Wunsch global alignment algorithm (to determine the mean percentage of nucleotide sequence identity among orthologous ORFs). To evaluate genomic similarity within the strains, we also performed digital DNA-DNA hybridization (<http://ggdc.dsmz.de/>), which exhibits high correlation with DNA-DNA hybridization [33].

a.



b.

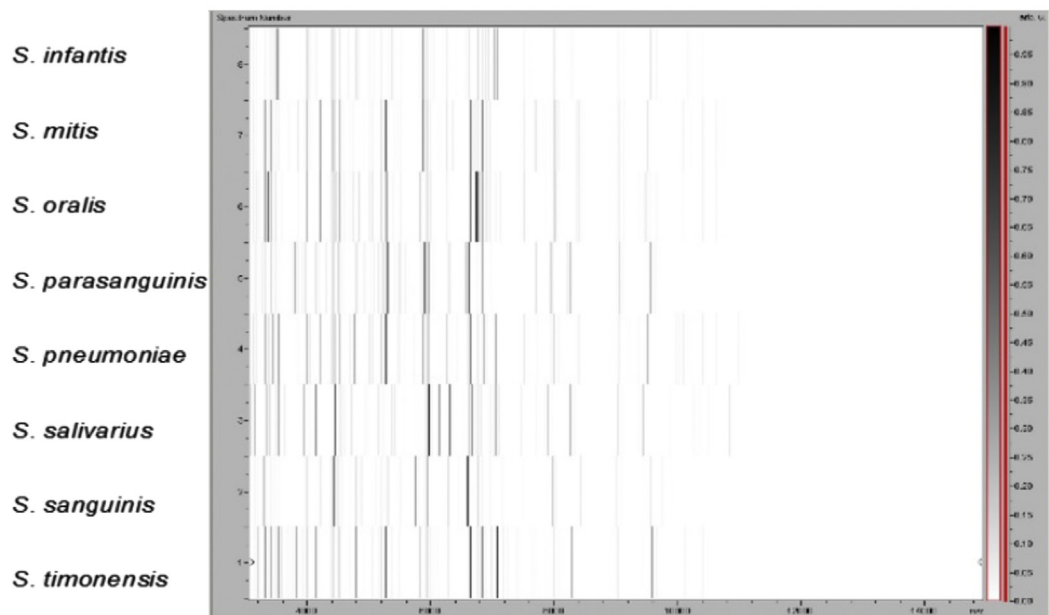


FIG. 1. Matrix-assisted Laser desorption ionization–time of flight mass spectrometry analysis of *Streptococcus timonensis* strain Marseille-P2915^T. (a) Reference mass spectrum from *S. timonensis* strain Marseille-P2915^T. (b) Gel view comparing *S. timonensis* sp. nov. strain Marseille-P2915^T spectra with other members of *Streptococcus* genus. Gel view displays raw spectra of loaded spectrum files arranged in pseudo-gel-like look. *x*-axis records *m/z* value. Left *y*-axis displays running spectrum number originating from subsequent spectra loading. Peak intensity is expressed by greyscale scheme code. Colour bar and right *y*-axis indicate relationship between colour in which peak is displayed and peak intensity in arbitrary units. Displayed species are indicated at left.

Results

Phylogenetic analysis

After three failed identification attempts with MALDI-TOF MS, the reference spectrum of strain Marseille-P2915^T (Fig. 1) was entered in our database and its 16S rRNA sequenced. The 16S rRNA comparison showed a sequence identity of 99.0% with *Streptococcus dentisani* strain 7747^T and 98.8% with *Streptococcus tigurinus* strain AZ-3a^T (Fig. 2). Because the 16S rRNA phylogenetic analysis did not distinguish within different species belonging to the *Streptococcus mitis* group, the *rpoB* gene was sequenced, resulting in a 92.6% sequence identity with the closest *Streptococcus* with an available *rpoB* sequence, *Streptococcus infantis* strain ATCC 700779^T (Fig. 3). A gel view comparing the spectrum of strain Marseille-P2915^T with other *Streptococcaceae* species is shown in Fig. 1(b). The *Streptococcus timonensis* strain Marseille-P2915^T 16S rRNA accession number from European Molecular Biology Laboratory (EMBL)-European Bioinformatics Institute (EBI) is LT576411. The MALDI-TOF MS reference spectrum is available online and in the public domain (<http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database>).

Phenotypic and biochemical characterization

On sheep's blood-enriched Columbia agar, after 48 hours of aerobic incubation, cells formed 0.5 to 1 mm punctiform,

greyish, α -hemolytic colonies with undulated edges. Growth was achieved at all tested pHs (5.0; 5.5; 6.0; 6.5 and 7.0; and 7.5 and 8), whereas no growth was registered in any salt-containing medium (>5 g/L of NaCl). Growth was obtained from 20 to 37°C under aerobic, anaerobic, and microaerophilic conditions, whereas no growth was obtained at 45 or 55°C. Cells were Gram-positive cocci (Fig. 4(a)) with a mean diameter of 0.6 μ m (range, 0.4–0.8 μ m) (Fig. 4(b)). Catalase and oxidase tested negative. The sporulation and motility tests were negative. Table 1 summarizes the classification and main features of strain Marseille-P2915^T.

Using an API 20 strip gallery, positive reactions were recorded only for leucyl-aminopeptidase and acid production from starch. Negative reactions were recorded for: Voges-Proskauer reaction, hippuric acid hydrolysis, esculin hydrolysis, pyrrolidiny arylamidase, α -galactosidase, β -glucuronidase, β -galactosidase, alkaline phosphatase and arginine hydrolase; fermentation reactions were negative for D-ribose, L-arabinose, D-mannitol, D-sorbitol, D-lactose, D-trehalose, inuline and D-raffinose. Using an API ZYM strip, positive reactions were observed for esterase (C4) and leucine arylamidase. Negative reactions were recorded for alkaline phosphatase, acid phosphatase, α -galactosidase, β -galactosidase, β -glucosidase, N-acetyl- β -glucosaminidase and naphthol-AS-BI-phosphohydrolase, lipase (C14), esterase lipase (C8), valine arylamidase, cystine arylamidase, trypsin, α -chymotrypsin, β -glucuronidase, α -glucosidase, α -fucosidase and α -mannosidase. Using an API 50 CH strip, positive reactions were observed for D-

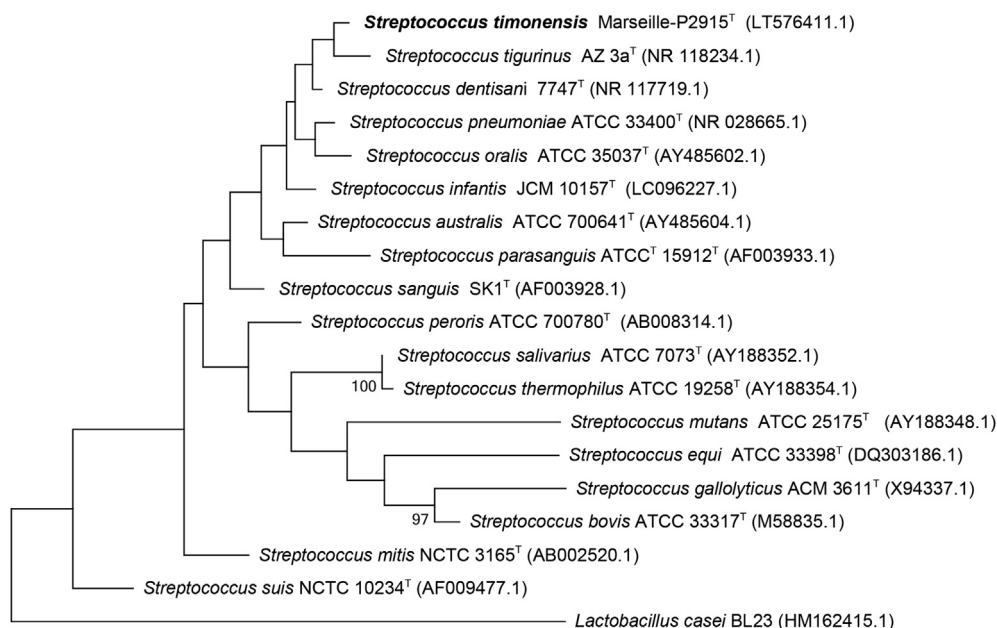


FIG. 2. Phylogenetic tree based on 16S rRNA showing position of *Streptococcus timonensis* strain Marseille-P2915^T relative to other type strains within genus *Streptococcus*. Strains and their corresponding GenBank accession numbers for 16S rRNA genes are in parentheses. Tree was constructed by maximum likelihood method with Kimura two-parameter model and 1000 bootstrap replications using MEGA6 software and rooted by using *Lactobacillus casei* strain BL 23 (HMI162415.1) as outgroup. Only bootstrap values of >95% are shown. Scale bar represents 0.02% nucleotide sequence divergence.

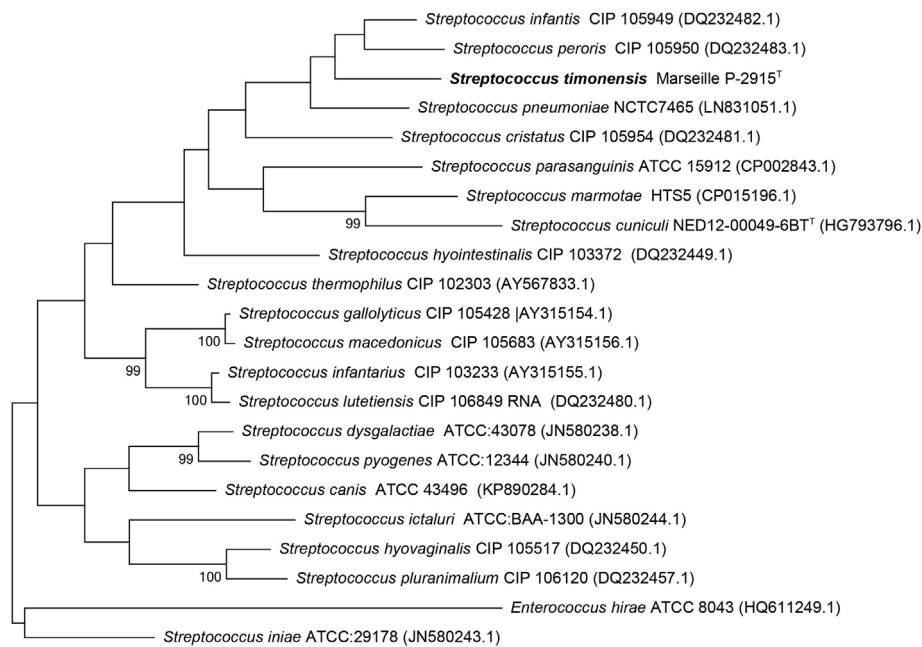


FIG. 3. Phylogenetic tree based on *rpoB* gene sequence showing position of *Streptococcus timonensis* strain Marseille-P2915^T relative to other strains within genus *Streptococcus*. Strains and their corresponding GenBank accession numbers for *rpoB* genes are in parentheses. Tree was constructed using maximum likelihood method with Kimura two-parameter model and 1000 bootstrap replications using MEGA6 software and rooted by using *Enterococcus hirae* strain ATCC 29187 (HQ611249.1) as outgroup. Only bootstrap values of >95% are shown. Scale bar represents 0.02% nucleotide sequence divergence.

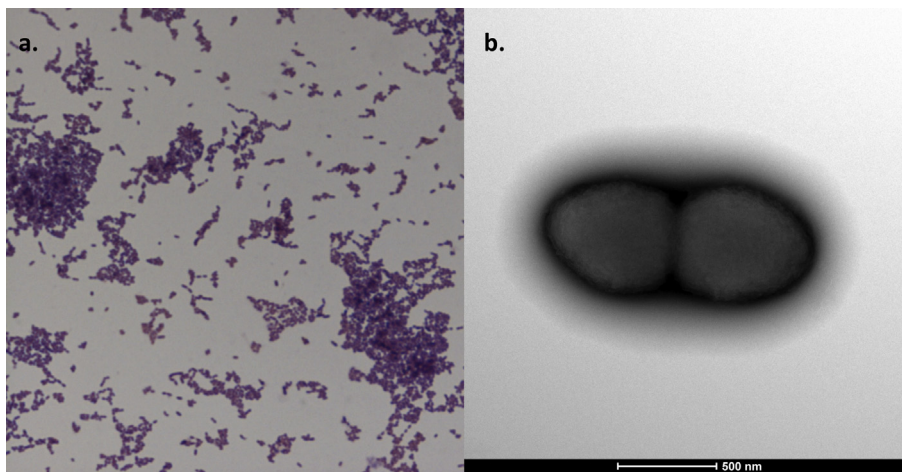


FIG. 4. Phenotypic features of *Streptococcus timonensis* strain Marseille-P2915^T. (a) Gram staining of *S. timonensis* strain Marseille-P2915^T. (b) Transmission electron microscopy of *S. timonensis* strain Marseille-P2915^T using Tecnai G20 (FEI Company) at operating voltage of 200 keV. Scale bar = 200 nm.

glucose, D-fructose, D-maltose and D-saccharose. Negative reactions were recorded for D-ribose, L-arabinose, D-xylose, D-mannose, inositol, glycerol, arbutin, D-tagatose, D-arabinose, D-fucose, N-acetylglucosamine, salicin, D-cellobiose, D-trehalose, amidone, erythritol, L-xylose, D-adonitol, methyl-βD-xylopyranoside, D-galactose, L-sorbose, L-rhamnose, dulcitol, D-sorbitol, methyl-αD-mannopyranoside, methyl-αD-glucopyranoside, esculin, D-lactose, D-melibiose, inulin, D-melezitose, D-mannitol, amygdalin, D-raffinose, glycogen, xylitol, gentiobiose, D-turanose,

D-lyxose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate and potassium 5-ketogluconate. Strain Marseille-P2915^T resulted susceptible to penicillin, ceftriaxone, imipenem, vancomycin and rifampicin. Susceptibility to erythromycin was classified as intermediate. Strain Marseille-P2915^T was tested resistant to trimethoprim-sulfamethoxazole. Biochemical characteristics that differentiate strain Marseille-P2915^T from other related species within the family *Streptococcaceae* are summarized in Table 2. The major fatty acid is hexadecanoic acid

TABLE 1. Classification and general features of strain *Streptococcus timonensis* Marseille-P2915^T

Property	Term
Classification	Domain <i>Bacteria</i> Phylum <i>Firmicutes</i> Class <i>Bacilli</i> Family <i>Streptococcaceae</i> Genus <i>Streptococcus</i> Species <i>Streptococcus timonensis</i> Type strain Marseille-P2915 ^T
Gram stain	Positive
Cell shape	Cocci
Motility	Nonmotile
Sporulation	Non-spore forming
Temperature range	Mesophile
Optimum temperature	37°C
Habitat	Human gut
Oxygen requirement	Facultative anaerobe
Biotic relationship	Free-living
Pathogenicity	Unknown
Geographical location	Marseille
Sample collection	April 2016
Latitude	43.296346° N
Longitude	5.369889° E

(43%). Abundant fatty acids were either saturated or unsaturated. Moreover, several branched fatty acids were also described at lower abundances. A complete report of fatty acid analysis of strain Marseille-P2915^T is reported in Table 3.

Genome properties

The draft genome of the strain Marseille-P2915^T is 1 925 331 bp long (Fig. 5) with a 38.56% G+C content. It is composed of four scaffolds (comprising four contigs). Among

TABLE 2. Differential characteristics of *Streptococcus timonensis* strain Marseille-P2915^T compared to closely related strains

Trait	1	2	3	4	5	6	7	8
Amygdalin	-	-	-	-	-	-	+	-
Arbutin	-	-	-	-	-	-	+	-
Inulin	-	-	-	(+)	-	-	-	-
Lactose	-	+	+	+	(+)	+	-	+
Maltose	+	+	+	+	+	(+)	+	+
Raffinose	-	(+)	+	-	-	-	(+)	+
Ribose	-	-	+	-	-	-	-	-
Starch	(+)	-	-	-	-	-	(+)	+
Trehalose	-	-	-	+	(+)	-	-	v
α-Galactosidase	-	(+)	-	-	-	-	-	-
Alkaline phosphatase	-	(+)	-	+	-	-	-	-
VP-reaction	-	-	-	-	-	-	+	-
β-Glucosidase	-	-	-	-	(+)	-	+	-
Leucine-aminopeptidase	+	NA	NA	NA	NA	NA	NA	+
Cellobiose	-	-	-	-	-	-	+	-
Fructose	+	(+)	+	+	+	+	-	+
Galactose	-	+	+	+	+	+	-	+
Gentiobiose	-	-	-	-	-	-	+	-
Glucose	+	-	+	+	+	+	-	+
Hippuric acid	-	-	-	-	-	-	+	-
Mannose	-	(+)	(+)	+	+	+	-	+
Methyl β-D-xylopyranoside	-	-	+	+	+	+	-	-
Sucrose	+	(+)	+	+	+	+	-	+

Strains: 1, *Streptococcus timonensis* strain Marseille-P2915^T; 2, *S. dentisani* strain 7747^T; 3, *S. mitis* ATCC 49465^T; 4, *S. oralis* ATCC35037^T; 5, *S. sanguinis* ATCC 10556^T; 6, *S. infantis* JCM 10157^T; 7, *S. salivarius* ATCC 7073^T; 8, *S. tigurinus* Az_3a^T; data for strains 2 to 7 are from [34], while data for strain 8 are derived from [35]. +, positive reaction; -, negative reaction; (+), weakly positive reaction; NA, not available; v, variable.

TABLE 3. Fatty acid profiles of *Streptococcus timonensis* strain Marseille-P2915^T

Fatty acid	Name	Mean relative %
16:0	Hexadecanoic acid	43.3 ± 1.5
18:1n9	9-Octadecenoic acid	19.5 ± 0.4
18:0	Octadecanoic acid	14.1 ± 0.1
14:0	Tetradecanoic acid	8.5 ± 0.1
18:2n6	9,12-Octadecadienoic acid	6.0 ± 0.3
18:1n7	11-Octadecenoic acid	2.7 ± 0.4
18:1n5	13-Octadecenoic acid	1.5 ± 0.3
17:0	Heptadecanoic acid	1.0 ± 0.2
12:0	Dodecanoic acid	1.0 ± 0.1
15:0	Pentadecanoic acid	TR
16:1n7	9-Hexadecenoic acid	TR
13:0	Tridecanoic acid	TR
17:1n7	10-Heptadecenoic acid	TR
13:0 iso	11-Methyl-dodecanoic acid	TR
16:1n5	11-Hexadecenoic acid	TR
16:1n9	7-Hexadecenoic acid	TR
17:0 anteiso	14-Methyl-hexadecanoic acid	TR
15:0 iso	13-Methyl-tetradecanoic acid	TR
17:0 iso	15-Methyl-Hexadecanoic acid	TR

TR, trace amounts <1%.

the 2056 predicted genes, 1974 were protein-coding genes and 82 were RNAs (six genes are 5S rRNA, six genes are 16S rRNA, six genes are 23S rRNA and 64 genes are tRNA genes). A total of 1436 genes (72.75%) were assigned a putative function (by COGs or by NR BLAST). Twenty-two genes were identified as ORFans (1.11%). The remaining genes were annotated as hypothetical proteins (426 genes ≥21.58%). Genomic statistics are reported in Table 4, while Table 5 lists the distribution of genes into COGs functional categories (Fig. 5). The genome sequence has been deposited in EMBL-EBI under accession number FMIX01000000.

Comparison with other genomes

The draft genome sequence of strain Marseille-P2915^T (1925 Mb) is smaller than that of *Streptococcus pseudopneumoniae* (2086 Mb), *Streptococcus tigurinus* (2185 Mb), *Streptococcus pneumoniae* (2161 Mb), *Streptococcus parasanguinis* (2154 Mb) and *Streptococcus sanguinis* (2303 Mb), but larger than that of *Streptococcus dentisani* (1884 Mb), *Streptococcus infantis* (1857 Mb), *Streptococcus mitis* (1916 Mb) and *Streptococcus oralis* (1914 Mb). The G+C content of *S. timonensis* (38.5%) is smaller than that of *S. dentisani* (41.1%), *S. pseudopneumoniae* (39.8%), *S. infantis* (38.9%), *S. mitis* (41.1%), *S. tigurinus* (40.3%), *S. pneumoniae* (39.7%), *S. parasanguinis* (41.7%), *S. sanguinis* (43.2%) and *S. oralis* (41.4%).

The gene content of strain Marseille-P2915^T (1.974 genes) is smaller than that of *S. pseudopneumoniae* (2.113), *S. tigurinus* (2.114), *S. pneumoniae* (2.125), *S. parasanguinis* (2.022) and *S. sanguinis* (2.268), but larger than that of *S. dentisani* (1.797), *S. infantis* (1.876), *S. mitis* (1.793) and *S. oralis* (1.847). Within species with standing in nomenclature, AGIOS values ranged from 82.23% between *S. mitis* and *S. oralis* to 58.20% between

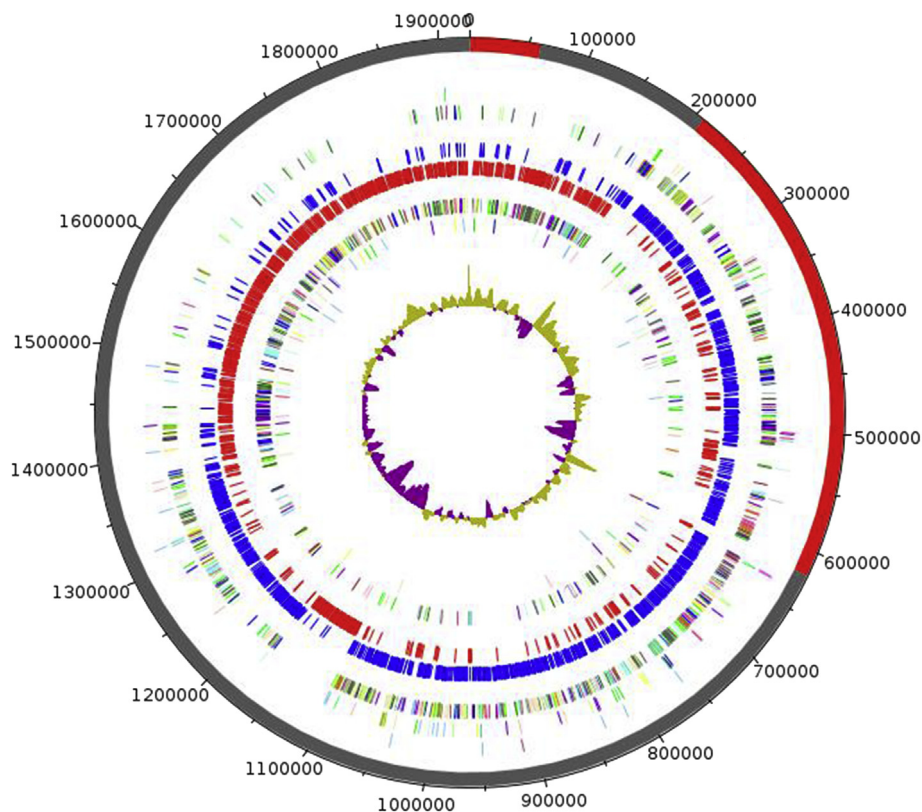


FIG. 5. Graphical circular map of genome of strain *Streptococcus timonensis* Marseille-P2915^T. From outside to centre: contigs (red/grey), COGs category of genes on forward strand (three circles), genes on forward strand (blue circle), genes on reverse strand (red circle), COGs category on reverse strand (three circles), GC content.

TABLE 4. Nucleotide content and gene count levels of the genome of *Streptococcus timonensis* strain Marseille-P2915^T

Attribute	Genome (total)	
	Value	% of total
Size (bp)	1 925 331	100
G+C content (%)	742 420	38.56
Coding region (bp)	1 731 074	89.9
Total genes	2056	100
RNA genes	82	3.98
Protein-coding genes	1974	96.01
Genes with function prediction	1436	72.75
Genes assigned to COGs	1221	61.85
Genes with peptide signals	182	9.21
Genes with transmembrane helices	445	22.54
Genes associated to PKS or NRPS	5	0.25
Genes associated to ORF	22	1.11
Genes associated to mobilome	887	44.93
Genes associated to toxin/antitoxin	72	3.64
Genes associated to resistance genes	0	0
Genes associated to virulence	422	21.37
Genes associated to bacteriocin	22	1.11
Genes with paralogues (<i>E</i> value: 1e-10)	289	14.64
Genes with paralogues (<i>E</i> value: 1e-25)	170	8.61
Genes associated to hypothetical proteins	426	21.58
Genes larger than 5000 nucleotides	6	0

COGs, Clusters of Orthologous Groups database; NRPS, nonribosomal peptide synthase; ORF, open reading frame; PKS, polyketide synthase.

TABLE 5. Number of genes associated with 26 general COGs functional categories

Code	Value	% of total ^a	Description
J	189	9.57	Translation
A	0	0	RNA processing and modification
K	93	4.71	Transcription
L	76	3.85	Replication, recombination and repair
B	0	0	Chromatin structure and dynamics
D	26	1.31	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	42	2.12	Defense mechanisms
T	54	2.73	Signal transduction mechanisms
M	68	3.44	Cell wall/membrane biogenesis
N	9	0.45	Cell motility
Z	0	0	Cytoskeleton
W	2	0.10	Extracellular structures
U	17	0.86	Intracellular trafficking and secretion
O	55	2.78	Posttranslational modification, protein turnover, chaperones
C	34	1.72	Energy production and conversion
X	15	0.75	Mobilome: prophages, transposons
G	99	5.01	Carbohydrate transport and metabolism
E	122	6.18	Amino acid transport and metabolism
F	72	3.64	Nucleotide transport and metabolism
H	57	2.88	Coenzyme transport and metabolism
I	50	2.53	Lipid transport and metabolism
P	64	3.24	Inorganic ion transport and metabolism
Q	18	0.91	Secondary metabolites biosynthesis, transport and catabolism
R	105	5.31	General function prediction only
S	90	4.55	Function unknown
—	753	38.14	Not in COGs

COGs, Clusters of Orthologous Groups database.

^aTotal is based on the total number of protein-coding genes in the annotated genome.

TABLE 6. Number of orthologous proteins shared between *Streptococcus* genomes (upper right), average percentage similarity of nucleotides 502 corresponding to orthologous proteins shared between genomes (lower left) and number of proteins per genome (**bold**)

	<i>S. oralis</i>	<i>S. infantis</i>	<i>S. pseudopneumoniae</i>	<i>S. sanguinis</i>	<i>S. mitis</i>	<i>S. parasanguinis</i>	<i>S. tigurinus</i>	<i>S. pneumoniae</i>	<i>S. timonensis</i>	<i>S. dentisani</i>
<i>S. oralis</i>	1847	1280	1290	1246	1449	1241	1377	1290	1285	1401
<i>S. infantis</i>	67.65	1876	1214	1145	1272	1199	1260	1203	1297	1278
<i>S. pseudopneumoniae</i>	67.39	67.27	2113	1163	1299	1200	1277	1325	1220	1292
<i>S. sanguinis</i>	6078	58.20	61.42	2268	1245	1235	1203	1176	1158	1238
<i>S. mitis</i>	82.23	65.24	65.83	62.81	1793	1254	1355	1307	1285	1392
<i>S. parasanguinis</i>	62.65	62.01	61.30	62.62	62.75	2022	1231	1213	1212	1248
<i>S. tigurinus</i>	75.57	68.32	67.28	59.43	69.27	62.57	2114	1268	1263	1369
<i>S. pneumoniae</i>	71.68	66.61	68.25	62.63	72.49	63.40	70.54	2125	1210	1285
<i>S. timonensis</i>	65.09	64.76	65.47	73.09	67.07	63.17	62.27	67.09	1974	1274
<i>S. dentisani</i>	69.44	61.94	67.78	73.88	71.69	63.24	66.11	69.08	81.41	1797

TABLE 7. Pairwise comparison of *Streptococcus timonensis* Marseille-P2915^T with nine other *Streptococcus* species using GGDC, formula 2 (DDH estimates based on identities/HSP length)

	<i>S. timonensis</i>	<i>S. oralis</i>	<i>S. infantis</i>	<i>S. pseudopneumoniae</i>	<i>S. sanguinis</i>	<i>S. mitis</i>	<i>S. parasanguinis</i>	<i>S. tigurinus</i>	<i>S. pneumoniae</i>	<i>S. dentisani</i>
<i>S. timonensis</i>	100% ± 00	25.3% ± 2.8	37.4% ± 2.5	26.2% ± 2.6	25.2% ± 2.0	26.0% ± 2.4	26.0% ± 2.4	25.2% ± 2.4	25.8% ± 2.4	25.2% ± 2.4
<i>S. oralis</i>		100% ± 00	25.8% ± 2.4	31.8% ± 2.4	23.7% ± 1.9	31.9% ± 2.4	24.8% ± 2.3	49.8% ± 2.6	31.6% ± 2.4	44.8% ± 2.5
<i>S. infantis</i>			100% ± 00	26.4% ± 2.4	24.5% ± 2.4	26.2% ± 2.4	26.5% ± 2.4	25.7% ± 2.4	25.7% ± 2.4	25.5% ± 2.4
<i>S. pseudopneumoniae</i>				100% ± 00	24.2% ± 2.4	48.5% ± 2.6	25.1% ± 2.4	31.6% ± 2.4	59.1% ± 2.8	31.5% ± 2.4
<i>S. sanguinis</i>					100% ± 00	23.5% ± 2.4	23.1% ± 2.3	23.1% ± 2.3	23.3% ± 2.3	23.4% ± 2.4
<i>S. mitis</i>						100% ± 00	24.2% ± 2.4	31.7% ± 2.6	46.7% ± 2.5	31.7% ± 2.5
<i>S. parasanguinis</i>							100% ± 00	24.5% ± 2.4	25.6% ± 2.4	23.6% ± 2.4
<i>S. tigurinus</i>								100% ± 00	31.8% ± 2.4	43.1% ± 2.5
<i>S. pneumoniae</i>									100% ± 00	31.3% ± 2.4
<i>S. dentisani</i>										100% ± 00

DDH, DNA-DNA hybridization; GGDC, Genome-to-Genome Distance Calculator; HSP, high-scoring segment pairs.

S. sanguinis and *S. infantis* (Table 6). Genome-to-Genome Distance Calculator (GGDC) values [36] ranged from 59.1 ± 2.8 between *S. pneumoniae* and *S. pseudopneumoniae* to 23.1 ± 2.3 between *S. sanguinis* and *S. parasanguinis* (Table 7). Compared

to the closest phylogenetic neighbour according to *rpoB* tree (*Streptococcus infantis* ATCC 700779T, Fig. 3), the probability that both are same species is lower than 5% (1.35%, GGDC 2.1 formula 2, <http://ggdc.dsmz.de/>).

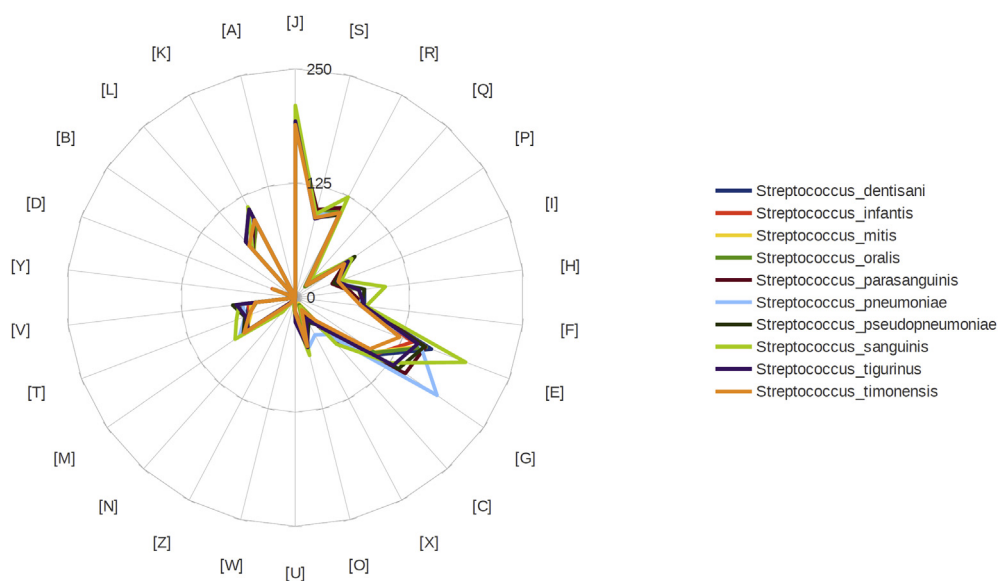


FIG. 6. Distribution of functional classes of predicted genes according to clusters of orthologous groups of proteins from *Streptococcus timonensis* strain Marseille-P2915^T.

TABLE 8. 16S rRNA similarity within: 1, *Streptococcus timonensis* strain Marseille-P2915^T; 2, *Streptococcus dentisani* strain 7747^T; 3, *Streptococcus pneumoniae* strain ATCC 33400^T; 4, *Streptococcus infantis* strain JCM 10157^T; 5, *Streptococcus tigurinus* strain AZ 3a^T; 6, *Streptococcus mitis* strain NCTC 3165^T; 7, *Streptococcus oralis* 35037^T; 8, *Streptococcus parasanguinis* strain ATCC 15912^T; 9, *Streptococcus sanguinis* strain SK1^T; 10, *Streptococcus salivarius* strain ATCC 7073^T

	1	2	3	4	5	6	7	8	9
1									
2	99.0								
3	98.8	98.9							
4	98.7	98.9	98.5						
5	99.0	98.7	98.4	98.6					
6	97.0	96.9	96.3	96.6	97.0				
7	98.3	98.4	98.9	98.6	98.3	97.0			
8	96.8	97.0	97.0	97.0	97.3	96.5	97.1		
9	98.1	97.9	97.8	97.8	98.5	97.4	97.7	97.2	
10	95.4	95.2	95.3	95.6	95.3	95.5	95.3	95.3	95.7

Distribution of functional classes of predicted genes according to the COGs categories was realized for the closest species of strain Marseille-P2915^T (Fig. 6) and showed a similar profile.

Discussion

A strong argumentation to support the recognition of strain Marseille-P2915^T as a new species within the *Streptococcus* genus is brought by the low value of GGDC (formula 2) obtained with other close species (highest value 37.4% with *Streptococcus infantis*, probability that both are the same species

<5%), particularly when compared to the high value obtained within different streptococcal species with standing in nomenclature (59.1% between *Streptococcus pneumoniae* and *Streptococcus pseudopneumoniae* and 48.5% between *Streptococcus mitis* and *Streptococcus pseudopneumoniae*) (Table 7). Other evidence comes from the phenotypic analysis obtained by the API strips, which showed a unique enzymatic profile (Table 2) compared to the eight most closely related species. In the family *Streptococcaceae* and in particular within the '*Streptococcus mitis*' group, it is a common feature to share a high 16S rRNA similarity [3] with other streptococcal species (Table 8). Another important point that differentiates strain Marseille-P2915^T from *Streptococcus dentisani* and *Streptococcus tigurinus* (the 16S rRNA closest species with standing in nomenclature) is the difference in G+C content (2.6 and 1.8%, respectively). To our knowledge, there is no officially recognized threshold for *rpoB* gene sequence similarity to delineate a new species, but the 92.6% value obtained with *Streptococcus infantis* is well below the 96.7% of the *rpoB*'s gene similarity between *Streptococcus oralis* (strain KCCM 41567) and *Streptococcus cristatus* (strain CIP 56.62) or the 96.4% gene similarity between *Streptococcus salivarius* (strain 735-09) and *Streptococcus thermophilus* (strain CIP 105446) (Table 9).

Conclusion

On the basis of the phenotypic, phylogenetic and genomic analyses, we propose the validation of *Streptococcus timonensis* sp. nov. within the family *Streptococcaceae*. Strain Marseille-P2915^T is the type strain, and it was isolated from human stomach.

TABLE 9. *rpoB*'s gene similarity within: 1, *Streptococcus infantis* strain ChDC B 194; 2, *Streptococcus peroris* strain ChDC B648; 3, *Streptococcus mitis* strain ChDC B183; 4, *Streptococcus pneumoniae* strain NCTC 7465; 5, *Streptococcus cristatus* strain CIP 56.62; 6, *Streptococcus oligofermentans* strain ChDC B685; 7, *Streptococcus oralis* strain KCCM 41567; 8, *Streptococcus pseudopneumoniae* strain CIP 1086; 9, *Streptococcus oligofermentans* strain ChCD B689; 10, *Streptococcus infantarius* strain GMRS55; 11, *Streptococcus thermophilus* strain CIP 105446; 12, *Streptococcus salivarius* strain 735-09; 13, *Streptococcus timonensis* strain Marseille-P2915^T

	1	2	3	4	5	6	7	8	9	10	11	12
1												
2	94.1											
3	94.0	93.7										
4	92.5	91.9	95.8									
5	94.0	93	95.5	94.3								
6	94.0	92.7	94.0	94.0	94.9							
7	93.5	93.7	94.1	93.7	96.7	94.9						
8	92.8	92.2	96.4	95.3	95.0	93.8	94.9					
9	92.5	92.2	94.1	93.5	94.3	95.6	95.2	94.0				
10	90.0	89.1	89.1	87.9	88.6	88.0	88.6	88.6	88.5			
11	88.6	89.2	88.6	88.2	88.6	88.2	88.9	87.9	87.6	90.9		
12	30.1	90.3	90.1	88.8	89.2	88.5	89.4	89.2	88.3	91.2	96.4	
13	93.7	92.5	92.5	91.2	92.1	91.5	91.6	90.3	90.1	88.0	87.9	88.6

Description of *Streptococcus timonensis* strain Marseille-P2915^T sp. nov.

Streptococcus timonensis (ti.mon.e'n.sis, L. gen. masc. 'originating from La Timone,' the hospital where the sample was collected). *S. timonensis* is a nonmotile, non-spore-forming, facultative anaerobe and Gram-positive coccus. Growth is achieved under aerobic, anaerobic and microaerophilic atmospheres in a temperature range of 20 to 37°C and at an optimum temperature of 37°C. After 48 hours of aerobic incubation on 5% sheep's blood-enriched Columbia agar, colonies are pinpoint, greyish and α-haemolytic, with undulated edges and with a diameter of 0.5 to 1 mm. Cells are roughly round with a 0.6 μm diameter. Catalase and oxidase are negative. *Streptococcus timonensis* strain Marseille-P2915^T exhibits positive reactions for leucyl-aminopeptidase, esterase C4 and leucine arylamidase. It is able to ferment starch, D-glucose, D-fructose, D-maltose and D-saccharose. Strain Marseille-P2915^T was found to be susceptible to penicillin, ceftriaxone, imipenem, vancomycin and rifampicin. The major fatty acid is hexadecanoic acid (43%). The G+C content of the genome is of 38.56%. The 16S rRNA sequence and whole-genome shotgun sequence are deposited in EMBL-EBI under accession numbers LT576411 and FMIX01000000, respectively. The type strain is Marseille-P2915^T (= CSUR P2915 = DSM 103349), and it was isolated from a human stomach sample in Marseille, France.

Acknowledgements

The authors thank the Xegen Company (www.xegen.fr/) for automating the genomic annotation process. This study was funded by the Fondation Méditerranée Infection. We also thank C. Andrieu for administrative assistance and M. Lardière for English-language review.

Conflict of Interest

None declared.

References

- [1] Kawamura Y, Hou XG, Sultana F, Miura H, Ezaki T. Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus *Streptococcus*. *Int J Syst Bacteriol* 1995;45:406–8.
- [2] Kilian M, Poulsen K, Blomqvist T, Håvarstein LS, Bek-Thomsen M, Tettelin H, et al. Evolution of *Streptococcus pneumoniae* and its close commensal relatives. *PLoS One* 2008;3:e2683.
- [3] Arbiqwe JC, Poyart C, Trieu-Cuot P, Quesne G, Carvalho M da GS, Steigerwalt AG, et al. Accuracy of phenotypic and genotypic testing for identification of *Streptococcus pneumoniae* and description of *Streptococcus pseudopneumoniae* sp. nov. *J Clin Microbiol* 2004;42:4686–96.
- [4] Poyart C, Quesne G, Coulon S, Berche P, Trieu-Cuot P. Identification of streptococci to species level by sequencing the gene encoding the manganese-dependent superoxide dismutase. *J Clin Microbiol* 1998;36:41–7.
- [5] Tapp J, Tholleson M, Herrmann B. Phylogenetic relationships and genotyping of the genus *Streptococcus* by sequence determination of the RNase P RNA gene, *rnpB*. *Int J Syst Evol Microbiol* 2003;53:1861–71.
- [6] Picard FJ, Ke D, Boudreau DK, Boissinot M, Huletsky A, Richard D, et al. Use of *tuf* sequences for genus-specific PCR detection and phylogenetic analysis of 28 streptococcal species. *J Clin Microbiol* 2004;42:3686–95.
- [7] Glazunova OO, Raoult D, Roux V. Partial sequence comparison of the *rpoB*, *sodA*, *groEL* and *gyrB* genes within the genus *Streptococcus*. *Int J Syst Evol Microbiol* 2009;59:2317–22.
- [8] Drancourt M, Roux V, Fournier PE, Raoult D. *rpoB* gene sequence-based identification of aerobic Gram-positive cocci of the genera *Streptococcus*, *Enterococcus*, *Gemella*, *Abiotrophia*, and *Granulicatella*. *J Clin Microbiol* 2004;42:497–504.
- [9] Vela AI, Mentaberre G, Lavín S, Domínguez L, Fernández-Garayzábal JF. *Streptococcus caprae* sp. nov., isolated from Iberian ibex (*Capra pyrenaica hispanica*). *Int J Syst Evol Microbiol* 2016;66:196–200.
- [10] Okamoto M, Imai S, Miyahara M, Saito W, Momoi Y, Nomura Y, et al. *Streptococcus panodentis* sp. nov. from the oral cavities of chimpanzees. *Microbiol Immunol* 2015;59:526–32.
- [11] Vela AI, Casas-Díaz E, Lavín S, Domínguez L, Fernández-Garayzábal JF. *Streptococcus pharyngis* sp. nov., a novel streptococcal species isolated from the respiratory tract of wild rabbits. *Int J Syst Evol Microbiol* 2015;65:2903–7.
- [12] 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.
- [13] Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect* 2012;18:1185–93.
- [14] La Scola B, Khelaifia S, Lagier JC, Raoult D. Aerobic culture of anaerobic bacteria using antioxidants: a preliminary report. *Eur J Clin Microbiol Infect Dis* 2014;33:1781–3.
- [15] 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.
- [16] Seng P, Abat C, Rolain JM, Colson P, Lagier JC, Gouriet F, et al. Identification of rare pathogenic bacteria in a clinical microbiology laboratory: impact of matrix-assisted laser desorption ionization–time of flight mass spectrometry. *J Clin Microbiol* 2013;51:2182–94.
- [17] Drancourt M, Bollet C, Carlioz 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.
- [18] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;215:403–10.
- [19] Renesto P, Lorvellec-Guillon K, Drancourt M, Raoult D. *rpoB* gene analysis as a novel strategy for identification of spirochetes from the genera *Borrelia*, *Treponema*, and *Leptospira*. *J Clin Microbiol* 2000;38:2200–3.
- [20] Sasser M. Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME). MIDI Technical Note 101. 2006.
- [21] Dione N, Sankar SA, Lagier JC, Khelaifia S, Michele C, Armstrong N, Richez M, et al. Genome sequence and description of *Anaerobaculum massiliensis* sp. nov. *New Microb New Infect* 2016;10:66–76.

- [22] Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucl Acids Res* 2012;40:D48–53.
- [23] Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucl Acids Res* 1997;25:955–64.
- [24] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucl Acids Res* 2007;35:3100–8.
- [25] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004;340:783–95.
- [26] 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.
- [27] 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.
- [28] Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: a fast phage search tool. *Nucl Acids Res* 2011;39:W347–52.
- [29] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. *Bioinformatics* 2000;16:944–5.
- [30] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* 2009;25:119–20.
- [31] Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 2004;14:1394–403.
- [32] Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co)orthologs in large-scale analysis. *BMC Bioinformatics* 2011;12:124.
- [33] Auch AF, von Jan M, Klenk H-P, Göker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2010;2:117–34.
- [34] Camelo-Castillo A, Benítez-Páez A, Belda-Ferre P, Cabrera-Rubio R, Mira A. *Streptococcus dentisani* sp. nov., a novel member of the mitis group. *Int J Syst Evol Microbiol* 2014;64:60–5.
- [35] Zbinden A, Mueller NJ, Tarr PE, Eich G, Schulthess B, Bahlmann AS, et al. *Streptococcus tigurinus*, a novel member of the *Streptococcus mitis* group, causes invasive infections. *J Clin Microbiol* 2012;50:2969–73.
- [36] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.