Genomic and phenotypic description of *Streptococcus resistens* sp. nov., *Streptococcus buccae* sp. nov. and *Streptococcus mediterraneus* sp. nov., three new members of the *Streptococcus* genus isolated from the human oral cavity

Sabrina Naud
Aix Marseille Univ
Issam Hasni
Institut Hospitalo-Universitaire Méditerranée Infection
Sara Bellali
Institut Hospitalo-Universitaire Méditerranée Infection
Hoang Thong Kieu
Institut Hospitalo-Universitaire Méditerranée Infection
Cheikh Ibrahima Lo
Institut Hospitalo-Universitaire Méditerranée Infection
Hussein Anani
Institut Hospitalo-Universitaire Méditerranée Infection
Marion Bonnet
Institut Hospitalo-Universitaire Méditerranée Infection
Nicholas Armstrong
Institut Hospitalo-Universitaire Méditerranée Infection
Magali Richez
Institut Hospitalo-Universitaire Méditerranée Infection
Simon Robinne
Institut Hospitalo-Universitaire Méditerranée Infection
Didier Raoult
Institut Hospitalo-Universitaire Méditerranée Infection
Jean-Christophe Lagier
Institut Hospitalo-Universitaire Méditerranée Infection
Maryam Tidjani Alou (tidjani_maryam@hotmail.com)
Aix-Marseille Université IHU Méditerranée Infection: IHU Mediterranee Infection https://orcid.org/0000-0002-0167-6235

Research Article

Keywords: Culturomics, taxono-genomics, Streptococcus sp. nov., Genome, Bacteria, Human

Posted Date: February 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-238204/v1

Abstract

Phenotypic, phylogenetic and genomic studies were carried out on three unidentified Gram-stain positive, facultative anaerobic, and cocci-shaped bacteria isolated from the human oral cavity. The 16S rRNA gene of strains Marseille-P5794 T , Marseille-P6264 T and Marseille-P7376 T exhibited a sequence identity of 99,41%, 99.67% and 97.88%, respectively with Streptococcus cristatus, their closest phylogenetic relative with standing in nomenclature. Moreover, the rpoB gene sequence of strains Marseille-P5794 T and Marseille-P6264 T shared a similarity level with 96.1%, and 95.9% with Streptococcus cristatus whereas strain Marseille-P7376 T shared a 93.98% identity with Streptococcus sanguinis. Whole genome comparison of strains Marseille-P5794 T , Marseille-P6264 T and Marseille-P7376 T with their phylogenetic neighbours were under the threshold values set to define new species using digital DNA-DNA hybridization and Orthologous Average Nucleotide Identity. The taxonogenomics analysis thus allowed the classification of these strains as new species within the Streptococcus genus named Streptoccocus resistens sp. nov. Strain Marseille-P5794 T (=CSUR P5794 = CECT9902), Streptococcus buccae sp. nov. Strain Marseille-P6264 T (=CSUR P6264 = CECT9910) and Streptococcus mediterraneus sp. nov. Strain Marseille-P7376 (=CSUR P7376 = CECT30035).

Introduction

The genus *Streptococcus* currently consists of 108 species and 27 subspecies according the list of prokaryotic names with standing in nomenclature (Parte et al. 2020). Members of this genus are Gram-positive cocci, forming relatively long chains, facultative anaerobes that never possess a catalase (Facklam 2002). *Streptococcus* species are mostly found in the oral cavities of humans and animals (Facklam 2002; Shinozaki-Kuwahara et al. 2011; Bakour et al. 2016). The taxonomic classification of bacterial species belonging to the *Streptococcus* genus remains a challenge, as a result of a high genetic and phenotypic similarity between several *Streptococcus* species particularly when solely considering the 16S rRNA gene sequence which is the gold standard to discriminate bacterial species (Arbique et al. 2004; Kilian et al. 2008).

Although *Streptococci* are commensals of the human oral cavity, members of this genus have been associated to different diseases. For example, *Streptococcus agalactiae* have been associated with the development of pneumonia, septicemia or, more rarely, meningitis in neonates (Krzyściak et al. 2013). In addition, studies showed that *Streptococcus canis, Streptococcus equinus* or *Streptococcus bovis* can be responsible for epizootic opportunistic infections (Cole et al. 2008). Despite the wide diversity within the *Streptococcus* genus, using the culturomics method, our laboratory previously isolated new members of this genus (Ricaboni et al. 2016). This high throughput culture technique consists in using various culture media and physico-chemical conditions to mimic the original environment of microorganisms and explore as exhaustively as possible a given ecosystem (Lagier et al. 2012). It relies on the Matrix Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-TOF MS) technology to identify bacterial isolates expeditiously as well as on the taxonogenomics strategy, which combines phylogenetic analysis, phenotypic criteria and genomic characterization to describe new species (Ramasamy et al. 2014; Ngom

et al. 2020). Here, we present a detailed description according to the taxonogenomics concept (Ramasamy et al. 2014; Fournier et al. 2015) of three putative new species for which we suggest the names Streptoccocus resistens sp. nov. strain Marseille-P5794^T, *Streptococcus buccae* sp. nov. strain Marseille-P6264^T and *Streptococcus mediterraneus* sp. nov. strain Marseille-P7376^T.

Methods

Isolation and identification

Three oral samples were collected from a 39-year-old French female who spontaneously recovered from HIV (Colson et al. 2020) as well as from two healthy French males aged 27- and 28-year-old. Each subject gave an informed and written consent for this study which was approved by the ethics committee of the Institut Federatif de Recherche IFR48 under agreement number 2016-011. Samples were cultured in aerobic and anaerobic blood culture bottles (bioMérieux, Marcy l'Etoile, France) supplemented with 5% of sterile rumen fluid and 5% of sterile defibrinated sheep blood (bioMérieux). At different incubation timepoints, ten-fold serial dilutions of the culture were inoculated on 5% defibrinated sheep blood enriched Colombia agar (COS, bioMérieux, Marcy l'Etoile, France) incubated at 37°C in aerobic or anaerobic atmosphere. After an incubation for 24-48h, bacterial colonies were purified by subculture and identified using MALDI-TOF MS (Bruker Daltonics, Bremen, Germany) as previously described (Lo et al. 2015a). The unidentified spectra were imported into MALDI Biotyper 3.0 software (Bruker Daltonics) and analyzed against two databases, that of Bruker and the constantly updated in lab URMS database (https://www.mediterranee-infection.com/acces-ressources/base-de-donnees/urms-data-base/).

16S rRNA gene sequencing

In order to classify the isolated bacteria, the 16S rRNA gene sequencing was performed. The primer pair fD1 and rP2 (Eurogentec, Angers, France) was used to amplify the 16S rRNA gene prior to sequencing using the Big Dye® Terminator v1.1 Cycle Sequencing Kit and a 3500xL sequencer (Thermofisher, Saint-Aubin, France) as previously described (Drancourt et al. 2000). The 16S rRNA nucleotide sequences were assembled using CodonCode Aligner software (http://www.codoncode.com). The obtained sequence was matched against the GenBank database using BLASTn (2015). A sequence similarity under 98.7% and 95% with the phylogenetically closest species was used to define a new species and a new genus respectively (Kim et al. 2014). A phylogenetic tree based on 16S rRNA gene was built with the MEGA 7 software (Edgar 2004) with sequences aligned using the Muscle software (Kumar et al. 2016).

Phenotypic and biochemical characterization

Optimal growth condition. The optimal growth of the studied strains was determined using COS agar. Several atmospheres namely aerobic, anaerobic and microaerophilic, were tested and generated using anaerobic and microaerophilic generators respectively (Thermo Fisher Scientific, Dardilly, France). Different growth temperatures (25, 28, 37, 45 and 56°C) were tested in each atmosphere. Salt tolerance was tested for these strains using Columbia agar (Thermo Fisher Scientific Dardilly France)

supplemented with 5, 10 and 15% of NaCl as well. The pH range tolerated by the described isolates was determined by testing different pH levels (6, 6.5, 7, 7.5, 8 and 8.5) in the optimal growth condition.

Biochemical properties. Oxidase (Becton Dickinson) and catalase assays (bioMérieux) were carried out on each strain according to the manufacturer's instructions. Biochemical characteristics of these strains were further explored using API 50CH, API ZYM and API 20 Strep strips (bioMérieux) according to the manufacturer's instructions. Spore-forming capacity was explored by subjecting the described strains to a thermal shock through heating at 80°C for 20 minutes before inoculation on COS agar.

Morphological properties. Scanning microscopy electron (Hitachi TM4000) allowed the morphologic observation of the bacterial cells, as previously described (Belkacemi et al. 2019). The motility of each bacteria was studied by observing a suspension of fresh colonies in a slide and coverslip under a DM1000 photonic microscope (Leica Microsystems, Wetzlar, Germany) at a 100x magnification. Moreover, the cellular fatty acid methyl ester (FAME) profile was established using gas chromatography/ mass spectrometry (GC/MS). Briefly, two samples were prepared with approximately 70 mg of bacterial biomass per tube harvested from several culture plates. FAMEs were prepared as described by Sasser (Sasser). GC/MS analyses were performed as previously described (Dione et al. 2016). FAMEs were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500 - SQ 8 S, Perkin Elmer, Courtaboeuf, France). A spectral database search was then performed using MS Search 2.0 operated with the Standard Reference Database 1A (NIST, Gaithersburg, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK).

Antibiotic susceptibility profile. The antibiotic susceptibility profile of the strains was assessed using the disc diffusion method, specifically E-test strips (Citron et al. 1991; Balouiri et al. 2016) to determine the minimal inhibitory concentration (MIC) of the following antibiotics: amoxicillin, erythromycin, clindamycin, rifampicin, oxacillin, penicillin, vancomycin, ceftriaxone, linezolid, gentamicin, fosfomycin and doxycycline.

Genome sequencing and analysis

DNA extraction and sequencing. To study the genome properties of the studied strains, the genomic DNA (gDNA) was extracted using the EZ1 biorobot (Qiagen, Courtaboeuf, France) with the EZ1 DNA tissue kit prior to sequencing on the MiSeq technology (Illumina, San Diego, CA, USA) with the Nextera XT Paired end (Illumina), as previously described (Lo et al. 2015b). Then, the quality control of raw data from DNA sequencing was performed using the FastQC software (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

Assembly and annotation. The assembly of the genomes of the strains was performed with multiple softwares (Trimmomatic software (Bolger et al. 2014), GapCloser (Llorens et al. 2011) and SPAdes (Bankevich et al. 2012)) as previously described (Boxberger et al. 2019). Assembled genomes were annotated using Dfast (DDBJ Fast Annotation and Submission Tool) (https://dfast.nig.ac.jp/) and/or the PROKKA software (Seemann 2014).

Open Reading Frames (ORF) were predicted using Prodigal (Hyatt et al. 2010) whereas protein-coding sequences were predicted by matching ORFs to the NR database using BLASTP with the following parameters: E-value of $1e^{-03}$ ($1e^{-05}$ for a sequence shorter than 80 amino acids, coverage 0.7 and identity percentage of 30%). The resulting protein sequences were matched against the Clutster of Orthologous Groups (COG) database (Tatusov et al. 2000, 2015). RNA sequences, specifically ribosomal RNA (rRNA) and tranfer RNA (tRNA) were predicted using the RNAmmer (Lagesen et al. 2007) and tRNAScanSE (Lowe and Eddy 1997) tools respectively.

Genomic comparison. RpoB gene sequences extracted from the genomes of the studied strains were used to further classify the aforementioned isolates within the *Streptococcus* genus though the construction of a phylogenetic tree based built with the Muscle (Kumar et al. 2016) and MEGA 7 softwares (Edgar 2004).To assess the level of genomic similarity of the described strains with closely related species, genomic data from the following species were used: *Enterococcus hirae* strain R17 (CP015516.1), *Streptococcus constellatus* subsp. pharyngis C1050 (NC_022238.1), *Streptococcus cristatus* ATCC 51100 (NZ_CP050133.1) and *Streptococcus gordonii* strain FDAARGOS_257 (CP020450.2). The Type Strain Genome Server web server available online (https://tygs.dsmz.de/) (Meier-Kolthoff and Göker 2019) was used for genomic comparison through *in silico* digital DNA-DNA hybridization (dDDH) (Meier-Kolthoff et al. 2013) which allowed the estimation of the percentage of nucleotide identity between the compared genomes. In addition, the *Orthologous* Average Nucleotide Identity (OrthoANI) was assessed using the OAT software (Lee et al. 2016).

Results

Strain identification

Strain isolation and growth conditions. Strain Marseille-P5794^T was isolated from the oral sample of a 39-year-old French female who spontaneously recovered from HIV while strain Marseille-P6264^T was isolated from the oral sample of a 28-year-old healthy French male. Similarly, strain Marseille-P7376^T was isolated from the oral sample of a 27-year-old healthy French male. Strains Marseille-P6264^T and Marseille-P7376^T were isolated after a 7-day pre-incubation in an aerobic blood culture bottle (bioMérieux) while strain Marseille-P5794^T was isolated after two days of pre-incubation in an anaerobic blood culture bottle. An identification attempt using MALDI-TOF MS allowed to the creation of the reference spectrum of each strain (Figure 1) which was subsequently incremented in our database that did not contain any matching spectrum.

Strain identification. The 16S rRNA gene of strains Marseille-P5794^T, Marseille-P6264^T and Marseille-P7376^T exhibited a 99.47%, 99.77% and 97.9 % sequence identity respectively with *Streptococcus cristatus* strain ATCC 51100 (Genbank accession number EU156757), the phylogenetically closest species with standing in nomenclature (Figure 2). As the use of the 16S rRNA gene was not conclusive for the classification of species within the *Streptococcus* genus, the complete genome sequencing was

performed and the *rpoB* gene sequences were extracted *in silico* as this gene was previously described as a useful marker to delineate species within this genus (Drancourt et al. 2004). Using the *rpoB* gene sequence, strains Marseille-P5794^T and Marseille-P6264^T shared a 96.1% and 95.9% nucleotide sequence similarity respectively with *Streptococcus cristatus* ATCC 51100 (NZ_AFUE00000000.1), the phylogenetically closest species with a validly published name. As for strain Marseille-P7376^T, it showed a 93.98% identity with *Streptococcus sanguinis* strain CGMH010 (CP040556.1, Figure 3).

Phenotypic characteristics

Optimal growth condition. The optimal growth of the three strains was recorded in aerobic condition at 37°C after a 24-hour incubation although growth was recorded in aerobic, microaerophilic and anaerobic conditions at temperatures ranging from 25 to 45°C classifying as facultative anaerobes. Growth was observed at pH 6.0, 6.5, 7.0 and 7.5 for strains Marseille-P5794^T and Marseille-P6264^T whereas strain Marseille-P7376^T grew at pH levels ranging from 7 to 9. However, no salt tolerance was exhibited by any of the strains.

Morphological characteristics. Cells from the three strains were Gram-stain positive, non-spore-forming and non-motile cocci which formed cream-colored colonies on COS agar. Scanning electron microscopy revealed that strains Marseille-P5794^T, Marseille-P6264 ^{T} and Marseille-P7376^T had a mean diameter of 550nm, 736nm and 500nm respectively (Figure 4). Fatty acids with chain saturated structures were the most commonly found in the three strains of interest. In fact, the main cellular fatty acids found were hexadecanoic acid, tetradecanoic acid and dodecanoic acid. Several unsaturated structures were detected with minor amounts (Table 3).

Biochemical characteristics. Strains Marseille-P5794^T, Marseille-P6264^T and Marseille-P7376^T possessed neither catalase nor oxidase activities. Using an API 50CH strip, all strains exhibited a positive reaction for D-sorbitol, N-acetylglucosamine, D-saccharose and D-maltose (Table S1). Using an API ZYM strip, the three strains presented high enzymatic activity for leucine arylamidase (Table S2). Finally, using an API Strep, all strains were also able to ferment for esculin iron citrate, L-arginine and D-ribose (Table S3). The main phenotypic characteristics of these strains are summarized in Table 1.

Antibiotic susceptibility. The MIC obtained for strain Marseille-P5794^T were 0.38 µg/ml for amoxicillin, 0.0132 µg/ml for rifampicin, 1.5 µg/ml for oxacillin, 0.125 µg/ml for benzylpenicillin, 0.50 µg/ml for vancomycin, 0.25 µg/ml for ceftriaxone, 0.5 µg/ml for linezolid, 2 µg/ml for gentamicin, 32 µg/ml for fosfomycin and 1.0 µg/ml for doxycycline. In addition, strain Marseille-P5794 had a MIC>256 µg/mL for erythromycin and clindamycin (Table 2). The MIC obtained for strain Marseille-P6264^T for amoxicillin, erythromycin, clindamycin, rifampicin, oxacillin, benzylpenicillin, vancomycin, ceftriaxone, linezolid, gentamicin, fosfomycin, doxycycline, amikacin and ciprofloxacin were 0.047 µg/ml, 6 µg/ml, 12 µg/ml, 0.023 µg/ml, 0.125 µg/ml, 0.032 µg/ml, 0.5 µg/ml, 0.025 µg/ml, 0.75 µg/ml, 6 µg/ml, 24 µg/ml and 24 µg/ml respectively (Table 2). As for strain Marseille-P7376^T, the observed MIC were 0.064 µg/mL for amoxicillin and penicillin, 0.08 µg/mL for rifampicin, 0.095 µg/mL for ceftriaxone, 0.16 µg/mL for

erythromycin and clindamycin, 0.25 µg/mL for linezolid, 0.5 µg/mL for vancomycin, 1.0 µg/mL for Gentamycin, 16 µg/mL for doxycycline, 48 µg/mL for fosfomycin and 125 µg/mL for oxacillin (Table 2).

Genome analysis and comparison

Genome properties. The genome of strain Marseille-P5794^T was 2,073,737 bp long with 42.3 mol% G+C content and consisted of 12 contigs (Figure 5). Strain Marseille-P6264^T had a genome size 2,043,666 bp with 42.5 mol% G+C content (Figure 5). As for strain Marseille-P7376^T, it had a genome size of 2,292,862 bp consisting of 56 contigs with a G+C content of 41.9 mol% (Figure 5). The genome of Strain Marseille-P5794^T included 1,973 coding DNA sequences (CDS) among which 40 transfer ribonucleic acids (tRNAs), 4 ribosomal ribonucleic acids (rRNA) and 2 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) while that of strain Marseille-P6264^T contained 1,972 CDS including 44 tRNAs, 6 rRNAs and 1 CRISPRs. As for strain Marseille-P7376, the genome annotation predicted 2,180 CDS, 3 rRNA, 36 tRNA and 5 CRISPRs. Using Dfast annotation (https://dfast.nig.ac.jp/), the average protein length was 306.4, 299.4 and 304.4 amino acids for strains Marseille-P5794^T, Marseille-P6264^T and Marseille-P7376^T respectively. In addition, similar proportions of annotated genes were associated with COG functions in the genomes of the three strains: 79.3% (1564/1973) for strain Marseille-P5794, 79.7% (1572/1972 for strain Marseille-P6264^T) and 77.1% (1681/2180) for strain Marseille-P6264^T (Figure 6). Moreover, the COG annotation showed a distribution of protein-coding genes into 19 COG categories with the cluster of “General function prediction only”, “Function unknown”, “Translation” and “Amino acid transport and metabolism” being the most represented in the genome of the three described strains (Figure 6).

Genomic comparison. The genomic similarity of the studied *Streptococcus* strains with their closely related species was estimated using OrthoANI and dDDH. OrthoANI values among compared genomes ranged from 65.7% between strain Marseille-P5794^T and *Enterococcus hirae* to 94.97% between strain Marseille-P5794^T and *S. cristatus.* The OrthoANI heat-map showed a cluster consisting of strains and *S. cristatus* (Figure 7) thus confirming their membership to the *Streptococcus* genus. The highest similarity for strains Marseille-P5794^T, Marseille-P6264^T and Marseille-P7376^T was observed with *S. cristatus* (94.97%, 93.92% and 87.22% respectively). Furthermore, the dDDH yielded values ranging from 21.9 ±4.7% between *S. gordonii* and *S. constellatus* to 59.2 ±5.6% between *S. cristatus* and strain Marseille-P5794^T (Table 4). The highest similarity observed using dDDH were all under 70%, precisely 59.2%, 53.4%, and 33.9% with *S. cristatus* for strains Marseille- P5794^T, Marseille-P6264^T and Marseille-P7376^T respectively. Finally, a phylogenetic tree based on the whole genome sequences highlighted strains Marseille-P5794^T, Marseille-P6264^T and Marseille-P7376^T as a part of the Mitis group of the *Streptococcus* genus alongside *S. cristatus.*

Discussion

The taxonogenomics analysis performed on these putative new species revealed a close taxonomic distance from the Mitis group of the *Streptococcus* genus (Kawamura et al. 1995; Kilian et al. 2014).

Consequently, the high similarity observed between the 16S rRNA gene sequences warranted the use of the *rpoB* housekeeping gene to discriminate these strains (Drancourt et al. 2004). Based on their phylogenetic position as well as their respective percentage of similarity with the *rpoB* gene of the closest validly published species, these three strains were classified as new members within the *Streptococcus* genus. In addition, the highest dDDH value obtained by comparing the genomes of these strains was 59.2% shared between *S. cristatus* and strain Marseille-P5794^T (Table 4). This value was below the recommended threshold of 70% to describe a new bacterial species [28] demonstrating that all the genomic sequences compared here were taxonomically distinct. Similar dDDH values were observed in the description of other novel *Streptococcus* spp. (Ricaboni et al. 2016; Martínez-Lamas et al. 2020). Consequently, based on their phylogenetic, phenotypic, biochemical, and genomic characteristics, we classify strains Marseille-P6264^T, Marseille-P5794^T and Marseille-P7376^T within the *Streptococcus* genus, as novel species distinct from their closest phylogenetic neighbours. Therefore, we formally propose the creation of Streptoccocus resistens sp. nov., (strain Marseille-P5794^T), *Streptococcus buccae* sp. nov. (strain Marseille-P6264^T) and *Streptococcus mediterraneus* sp. nov. (strain Marseille-P7376^T) as new species within the *Streptococcus* genus from the *Streptococcaceae* family, *Lactobacillales* order, *Bacilli* class of the *Firmicutes* phylum.

Description

Description of *Streptococcus resistens* sp. nov.

Streptococcus resistens (re.sis'tens L. part. adj. *resistens*, referring to the HIV resistance exhibited by the patient from whom the strain was isolated). Cells are facultative anaerobes and Gram-stain positive cocci with a mean diameter of 550 nm. They are non-motile and non-spore-forming, catalase and oxidase negative. Strain Marseille-P5794^T is mesophilic with an optimal growth at 37°C. Colonies are circular and cream-colored. *S. resistens* is neither halotolerant nor halophilic but grows on agar between pH 6 and pH 7.5. Using an API ZYM strip, a positive reaction is observed for esterase (C4), leucine arylamidase, valine arylamidase, cystine arylamidase, acid phosphatase, naphtalo-AS-BI-phosphohydrolase and β-galactosodase. Using an API 50 CH strip, reaction is positive for salicine, D-maltose, D-lactose (bovine origin), D-saccharose (sucrose), D-trehalose, D-sorbitol, N-acetylglucosamine and arbutine. Finallys, using API 20 Strep, S.resistens can ferment esculin iron citrate, L-arginine, D-ribose, D-lactose (bovin origin), D-tréhalose, Inuline and Glycogen. $C_{16:0}$ (39.8%), $C_{14:0}$ (29.5%) and $C_{12:0}$ (16.9%) were the major fatty acids. The genome of *S. resistens* was 2.07 Mbp long with 42.3 % of G+C content. Strain Marseille-P5794 was isolated from an oral sample provided by a 39-year-old French female who spontaneously recovered from HIV. The 16S rRNA gene and genome sequences are available in GenBank under accession numbers LR597666 and CABPTQ010000000 respectively. This strain is deposited in the "*Collection de Souches* de l'Unité des Rickettsies" (CSUR) and "Spanish Type Culture Collection" (STCC) under number CSURP5794 and CECT 9902 respectively.

Description of *Streptococcus buccae* sp. nov.

Streptococcus buccae sp. nov. (buc.ca'e L. gen. fem. n. *buccae*, of the mouth, referring to the habitat where the strain was isolated). *Streptococcus buccae* consists of Gram-stain positive cocci which are facultative anaerobes. Cells are non-motile and non-spore-forming with a mean diameter of 736 nm. Catalase and oxidase activities are negative. *S. buccae* was mesophilic with an optimal growth temperature at 37°C and was able to tolerate pH levels ranging from 6 to 7.5. Using an API ZYM strip, a positive reaction was observed for leucine arylamidase, valine arylamidase, cystine arylamidase, naphthol-AS-BI-phosphohydrolase, α-galactosidase and β-galactosodase. Using an API 50 CH strip, positive reactions were obtained for D-maltose, D-saccharose (sucrose), D-galactose, D-raffinose, D-glucose, D-fructose, D-mannose, D-sorbitol and N-acetylglucosamine. Finally, using API 20 Strep, *S.buccae* can also ferment Sodium pyruvate, esculin iron citrate, 6-bromo-2-naphthyl- α D-galactopyranoside, L-arginine, D-ribose, D-lactose (bovin origin), starch and glycogen. The major fatty acids were $C_{16:0}$ (45.1%), $C_{14:0}$ (25.7%) and $C_{12:0}$ (14.7%). Its genome was 2.04 Mbp long with 42.5 % of G+C content. The accession numbers of 16S rRNA and genome sequences are LR699782 and CABPTO010000000 respectively in the Genbank database. The type strain Marseille-P6264 was isolated from the oral sample of a 28-year-old healthy French male and is available in the CSUR and STCC collections under number CSURP6264 and CECT 9910 respectively.

Description of *Streptococcus mediterraneus* sp. nov.*Streptococcus mediterraneus* sp. nov. (me.di.ter.ra'ne.us L. masc. adj. *mediterraneus*, from the Mediterranean Sea, which borders Marseille, where the strain was isolated). Cells are Gram-stain positive cocci which were facultatively anaerobic, non motile and non-spore-forming with a mean average diameter of 500 nm. Catalase and oxidase activities are negative. *S. mediterraneus* grows optimally on COS agar after 24 hours in aerobic condition at 37°C at pH 7.5. Using an API ZYM strip, a positive reaction is observed only for leucine arylamidase. Using an API 50 CH strip, positive reactions were obtained with N-acetylglucosamine, D-galactose, D-glucose, D-fructose, D-mannose, D-mannitol, D-sorbitol, arbutin, esculin, salicine, D-cellobiose, D-maltose, D-lactose, D-saccharose and D-trehalose. In addition, using API 20 Strep, *S.mediterraneus* can also ferment esculin iron citrate, 6-bromo-2-naphthyl- α D-galactopyranoside, L-arginine and D-ribose. The major fatty acids were $C_{16:0}$ (56 %), $C_{14:0}$ (19 %) and $C_{12:0}$ (7.2%). Its genome was 2.29 Mbp long with 41.9 % of G+C content. The 16S rRNA gene and genome sequences are available in the Genbank database under number LR699798 and CABPTS010000000, respectively. The type strain Marseille-P7376 was isolated from the oral sample of a 27-year-old healthy French male and it is deposited in the CSUR and STCC collections under number CSURP7376 and CECT 30035 respectively.

Declarations

Funding

This study was supported by the Méditerranée Infection Foundation and the National Research Agency under the program « Investissements d'avenir », reference ANR-10-IAHU-03. This work was supported by the Région Provence-Alpes-Côte d'Azur and European funding FEDER PRIMI

Conflicts of interest/Competing interests

None to declare

Availability of data and material

The datasets presented in this study can be found in online repositories. Associated accession numbers are cited in the manuscript.

Code availability

The accession numbers can be found in this article or in the supplementary Material.

Authors' contributions

Investigation: SN, SB, HTK, NA, MR, MB. Data curation: SR. Formal analysis: SN, NA. Conceptualization: DR. Methodology: DR, JCL, MTA. Visualization: SN, MTA. Supervision: DR, JCL, MTA. Writing – original draft: SN, NA, MTA. Writing – review & editing: MTA, JCL and DR.

Ethics approval

This study was approved by the ethics committee of the Institut Federatif de Recherche IFR48 under agreement number 2016-011.

Consent to participate

Each subject gave an informed and written consent for this study.

Consent for publication

All authors had access to this publication and approved its submission.

Acknowledgements

The authors thank Aurelia Caputo for submitting the 16S rRNA and complete genome sequences to GenBank as well as Stéphane Alibar for the deposition of the strains in the CSUR and STCC collections. In addition, we thank the genomics team for their technical contribution.

References

Arbique JC, Poyart C, Trieu-Cuot P, et al (2004) Accuracy of phenotypic and genotypic testing for identification of Streptococcus pneumoniae and description of Streptococcus pseudopneumoniae sp. nov. J Clin Microbiol 42:4686–4696. https://doi.org/10.1128/JCM.42.10.4686-4696.2004

Bakour S, Rathored J, Lo CI, et al (2016) Non-contiguous finished genome sequence and description of Streptococcus varani sp. nov. New Microbes New Infect 11:93–102.

https://doi.org/10.1016/j.nmni.2016.03.004

Balouiri M, Sadiki M, Ibnsouda SK (2016) Methods for in vitro evaluating antimicrobial activity: A review. Journal of Pharmaceutical Analysis 6:71–79. https://doi.org/10.1016/j.jpha.2015.11.005

Bankevich A, Nurk S, Antipov D, et al (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021

Belkacemi S, Bou Khalil J, Ominami Y, et al (2019) Passive Filtration, Rapid Scanning Electron Microscopy, and Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry for Treponema Culture and Identification from the Oral Cavity. J Clin Microbiol 57:. https://doi.org/10.1128/JCM.00517-19

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

Boxberger M, Anani H, La Scola B (2019) Genome sequence and description of Alterileibacterium massiliense gen. nov., sp. nov., a new bacterium isolated from human ileum of a patient with Crohn's disease. New Microbes New Infect 30:. https://doi.org/10.1016/j.nmni.2019.100533

Citron DM, Ostovari MI, Karlsson A, Goldstein EJ (1991) Evaluation of the E test for susceptibility testing of anaerobic bacteria. J Clin Microbiol 29:2197–2203

Cole JN, Henningham A, Gillen CM, et al (2008) Human pathogenic streptococcal proteomics and vaccine development. Proteomics Clin Appl 2:387–410. https://doi.org/10.1002/prca.200780048

Colson P, Dhiver C, Tamalet C, et al (2020) Full-length title: Dramatic HIV DNA degradation associated with spontaneous HIV suppression and disease-free outcome in a young seropositive woman following her infection. Sci Rep 10:. https://doi.org/10.1038/s41598-020-58969-6

Dione N, Sankar SA, Lagier J-C, et al (2016) Genome sequence and description of Anaerosalibacter massiliensis sp. nov. New Microbes New Infect 10:66–76. https://doi.org/10.1016/j.nmni.2016.01.002

Drancourt M, Bollet C, Carlioz A, et al (2000) 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol 38:3623–3630

Drancourt M, Roux V, Fournier P-E, Raoult D (2004) rpoB gene sequence-based identification of aerobic Gram-positive cocci of the genera Streptococcus, Enterococcus, Gemella, Abiotrophia, and Granulicatella. J Clin Microbiol 42:497–504. https://doi.org/10.1128/jcm.42.2.497-504.2004

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340

Facklam R (2002) What happened to the streptococci: overview of taxonomic and nomenclature changes. Clin Microbiol Rev 15:613–630. https://doi.org/10.1128/cmr.15.4.613-630.2002

Fournier P-E, Lagier J-C, Dubourg G, Raoult D (2015) From culturomics to taxonomogenomics: A need to change the taxonomy of prokaryotes in clinical microbiology. Anaerobe 36:73–78. https://doi.org/10.1016/j.anaerobe.2015.10.011

Hyatt D, Chen G-L, Locascio PF, et al (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119

Kawamura Y, Hou XG, Sultana F, et al (1995) Determination of 16S rRNA sequences of Streptococcus mitis and Streptococcus gordonii and phylogenetic relationships among members of the genus Streptococcus. Int J Syst Bacteriol 45:406–408. https://doi.org/10.1099/00207713-45-2-406

Kilian M, Poulsen K, Blomqvist T, et al (2008) Evolution of Streptococcus pneumoniae and its close commensal relatives. PLoS ONE 3:e2683. https://doi.org/10.1371/journal.pone.0002683

Kilian M, Riley DR, Jensen A, et al (2014) Parallel evolution of Streptococcus pneumoniae and Streptococcus mitis to pathogenic and mutualistic lifestyles. mBio 5:e01490-01414. https://doi.org/10.1128/mBio.01490-14

Kim M, Oh H-S, Park S-C, Chun J (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY 64:346–351. https://doi.org/10.1099/ijs.0.059774-0

Krzyściak W, Pluskwa KK, Jurczak A, Kościelniak D (2013) The pathogenicity of the Streptococcus genus. Eur J Clin Microbiol Infect Dis 32:1361–1376. https://doi.org/10.1007/s10096-013-1914-9

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054

Lagesen K, Hallin P, Rødland EA, et al (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160

Lagier J-C, Armougom F, Million M, et al (2012) Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 18:1185–1193. https://doi.org/10.1111/1469-0691.12023

Lee I, Ouk Kim Y, Park S-C, Chun J (2016) OrthoANI: An improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 66:1100–1103. https://doi.org/10.1099/ijsem.0.000760

Llorens C, Futami R, Covelli L, et al (2011) The Gypsy Database (GyDB) of mobile genetic elements: release 2.0. Nucleic Acids Res 39:D70–D74. https://doi.org/10.1093/nar/gkq1061

Lo CI, Fall B, Sambe-Ba B, et al (2015a) MALDI-TOF Mass Spectrometry: A Powerful Tool for Clinical Microbiology at Hôpital Principal de Dakar, Senegal (West Africa). PLoS ONE 10:e0145889. https://doi.org/10.1371/journal.pone.0145889

Lo CI, Padhmanabhan R, Mediannikov O, et al (2015b) High-quality genome sequence and description of Bacillus dielmoensis strain FF4(T) sp. nov. Stand Genomic Sci 10:41. https://doi.org/10.1186/s40793-015-0019-8

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964

Martínez-Lamas L, Limeres-Posse J, Diz-Dios P, Álvarez-Fernández M (2020) Streptococcus downii sp. nov., isolated from the oral cavity of a teenager with Down syndrome. Int J Syst Evol Microbiol 70:4098–4104. https://doi.org/10.1099/ijsem.0.004180

Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60. https://doi.org/10.1186/1471-2105-14-60

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

Ngom II, Hasni I, Lo CI, et al (2020) Taxono-genomics and description of Gordonibacter massiliensis sp. nov., a new bacterium isolated from stool of healthy patient. New Microbes New Infect 33:100624. https://doi.org/10.1016/j.nmni.2019.100624

Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, et al (2020) List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. Int J Syst Evol Microbiol. https://doi.org/10.1099/ijsem.0.004332

Ramasamy D, Mishra AK, Lagier J-C, et al (2014) A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol 64:384–391. https://doi.org/10.1099/ijs.0.057091-0

Ricaboni D, Mailhe M, Lagier J-C, et al (2016) Noncontiguous finished genome sequence and description of Streptococcus timonensis sp. nov. isolated from the human stomach. New Microbes New Infect 15:77–88. https://doi.org/10.1016/j.nmni.2016.11.013

Sasser M Bacterial Identification by Gas Chromatographic Analysis of Fatty Acid Methyl Esters (GC-FAME). 6

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

Shinozaki-Kuwahara N, Takada K, Hirasawa M (2011) Streptococcus ursoris sp. nov., isolated from the oral cavities of bears. Int J Syst Evol Microbiol 61:40–44. https://doi.org/10.1099/ijs.0.019638-0

Tatusov RL, Galperin MY, Natale DA, Koonin EV (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 28:33–36

(2015) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 43:D6–D17. https://doi.org/10.1093/nar/gku1130

Tables

Table 1 - General information regarding strains Marseille-P5794^{T}, Marseille-P6264^{T} and Marseille-P7376^{T}.

Property	Marseille-P5794^T	Marseille-P6264^T	Marseille-P7376^T
Species name proposed	*Streptococcus resistens*	*Streptococcus buccae*	*Streptococcus mediterraneus*
Genus name	*Streptococcus*	*Streptococcus*	*Streptococcus*
Specific epithet	*resistens*	*buccae*	*mediterraneus*
Species status	sp. nov.	sp. nov.	sp. nov.
Designation of the type strain	Marseille-P5794	Marseille-P6264	Marseille-P7376
Strain collection numbers	CSUR P5794	CSUR P6264	CSUR P7376
16S rRNA gene accession number	LR597666	LR699782	LR699798
Genome accession number	CABPTQ010000000	CABPTO010000000	CABPTS010000000
Genome size (bp)	2,073,737 bp	2,043,666 bp	2,292,862 bp
GC (mol %)	42.3%	42.5%	41.9 mol%.
Origin	Marseille, France	Marseille, France	Marseille, France
Years of isolation	2016	2016	2017
Source of isolation	Human oral cavity	Human oral cavity	Human oral cavity
Conditions used for standard cultivation	COS agar for 24h	COS agar for 24h	COS agar for 24h
Gram stain	Positive	Positive	Positive
Cell shape	Coccus	Coccus	Coccus
Cell size average (diameter)	550nm	736nm	500nm
Motility	Non-motile	Non-motile	Non-motile
Sporulation	Negative	Negative	Negative
Colony color	Cream	Cream	Cream
Relationship to O_2	Facultative anaerobe	Facultative anaerobe	Facultative anaerobe
Oxidase	Negative	Negative	Negative
Catalase	Negative	Negative	Negative
pH range	From 6 to 7.5	From 6 to 7.5	From 7 to 9

Temperature growth	Room temperature, 28°C, 37°C and 45°C.	Room temperature, 28°C, 37°C and 45°C.	Room temperature, 28°C, 37°C and 45°C.
Salinity test	Non-halophilic	Non-halophilic	Non-halophilic
Pathogenicity	Unknown	Unknown	Unknown
Biotic relationship	Free-living	Free-living	Free-living

Table 2. Summary table of the minimal inhibitory concentration in µg/ml for strain Marseille-P5794^T, strain Marseille-P6264^T and strain Marseille-P7376^T.

Antibiotics	Marseille-P5794^T	Marseille-P6264^T	Marseille-P7376^T
Amoxicillin	0.38	0.047	0.064
Erythromycin	>256	6	0.16
Clindamycin	>256	12	0.16
Rifampicin	0.0132	0.023	0.08
Oxacillin	1.5	0.125	125
Benzylpenicillin	0.125	0.032	0.064
Vancomycin	0.50	0.5	0.5
Ceftriaxone	0.25	0.025	0.095
Linezolid	0.5	0.75	0.25
Gentamicin	2	6	1.0
Fosfomycin	32	24	48

Table 3 - Cellular fatty acid composition (%) of strain Marseille-P5794^T, strain Marseille-P6264^T and strain P7376^T. [a] Mean peak area percentage. NA: not applicable.

Fatty acids	Name	Marseille-P5794^T [a]	Marseille-P6264^T [a]	Marseille-P7376^T [a]
$C_{16:0}$	Hexadecanoic acid	39.8	45.1	56.0
$C_{14:0}$	Tetradecanoic acid	29.5	25.7	19.0
$C_{12:0}$	Dodecanoic acid	16.9	14.7	7.2
$C_{18:1n9}$	9-Octadecenoic acid	4.2	4.9	4.7
$C_{18:0}$	Octadecanoic acid	4.0	5.1	4.3
$C_{18:2n6}$	9,12-Octadecadienoic acid	2.6	2.8	2.7
$C_{18:1n7}$	11-Octadecenoic acid	1.7	1.6	2.1
$C_{16:1n7}$	9-Hexadecenoic acid	1.4	NA	1.5

Table 4 – dDDH values obtained by comparison of all studied genomes which are *Streptococcus gordonii* (**St. go**), *Streptococcus constellatus* (**St. co**) and *Streptococcus cristatus* (**St. cr**), strain Marseille-P5794^T (**P5794^T**), strain Marseille-P6264^T (**P6264^T**) and strain Marseille-P7376^T (**P7376^T**).

	P5794^T	**P6264^T**	**St.go**	**St.co**	**St.cr**	**P7376^T**
P5794^T	100%	53.8 ±5.4%	27.3 ±4.9%	23.7 ±4.7%	59.2 ±5.6%	34.2 ±5%
P6264^T		100%	27.0 ±4.8%	23.3 ±4.8%	53.4 ±5.3%	34.7 ± 4.9%
St.go			100%	21.9 ±4.7%	27.6 ±4.8%	27 ±4.9%
St.co				100%	22.8 ±4.8%	22.9 ±4.8%
St.cr					100%	33.9 ±4.9%
P7376^T						100%

Figures

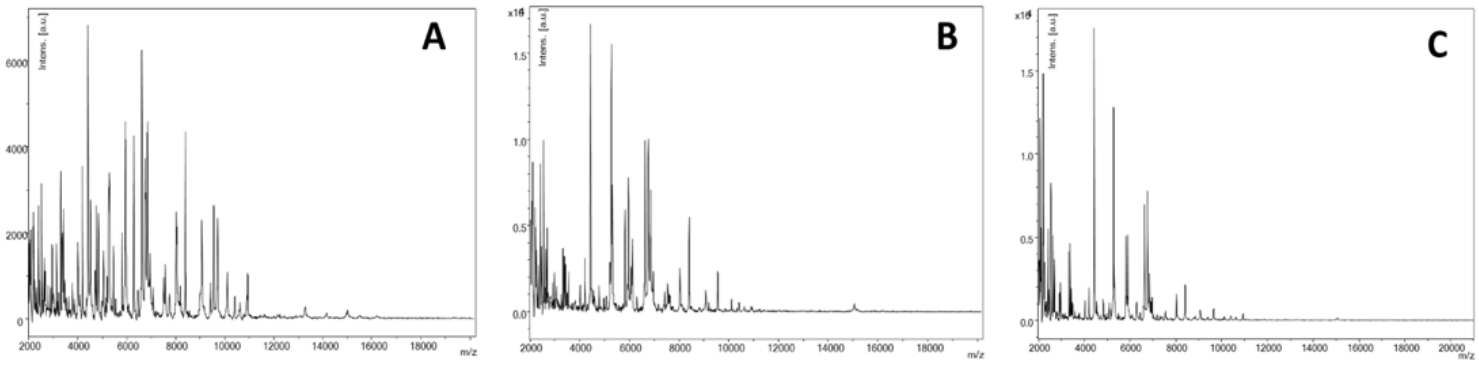

Figure 1

Reference mass spectra from of strain Marseille-P5794T (A), strain Marseille-P6264T (B) and strain Marseille-P7376T (C) obtained with MALDI-TOF MS.

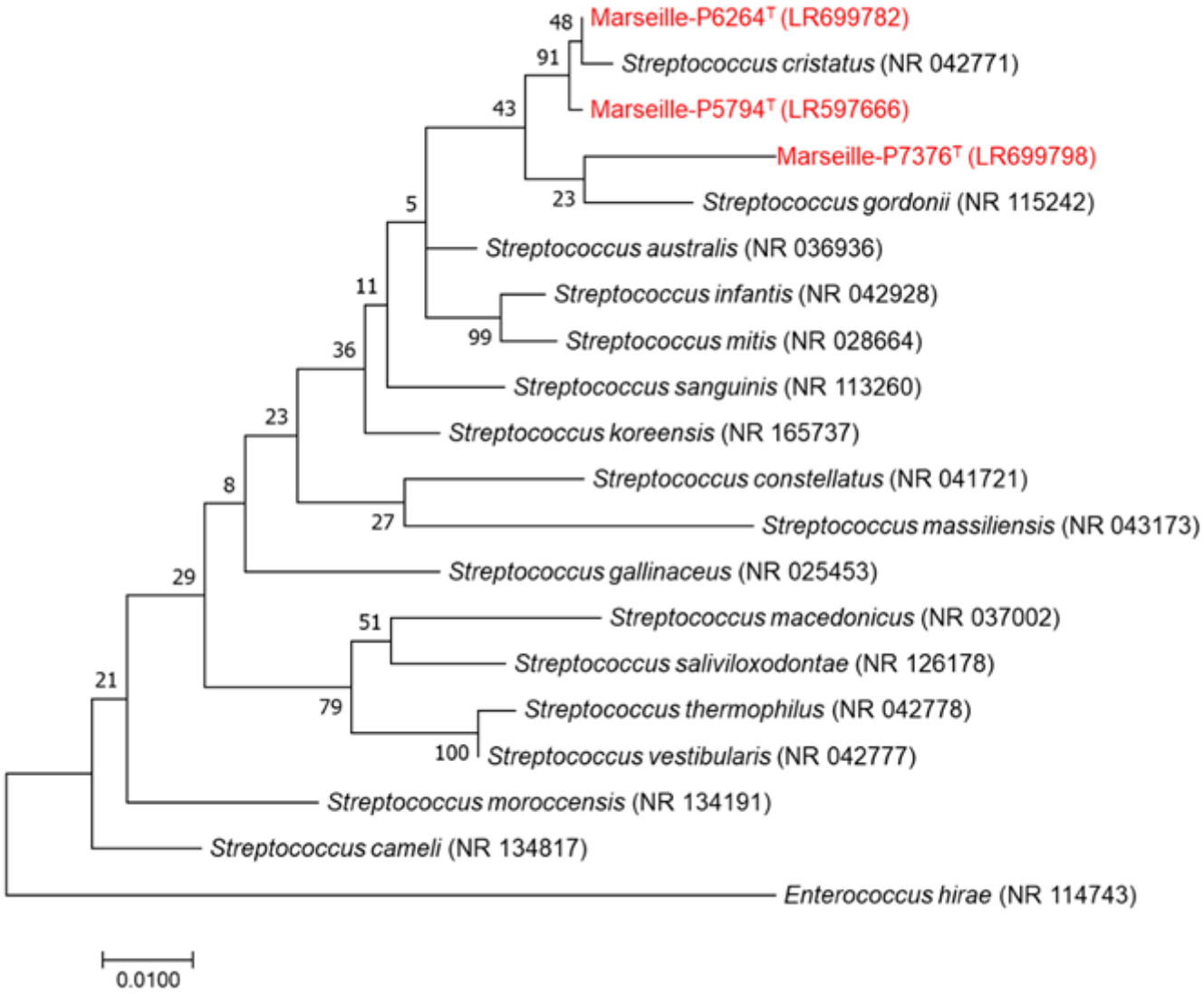

Figure 2

16S rRNA-based phylogenetic tree showing the position of strain Marseille-P5794T, strain Marseille-P6264T and strain Marseille-P7376T relative to other phylogenetically close species with Enterococcus hirae used as as outgroup. The described isolates are highlighted in red. The scale bar indicates a 1 % nucleotide sequence divergence.

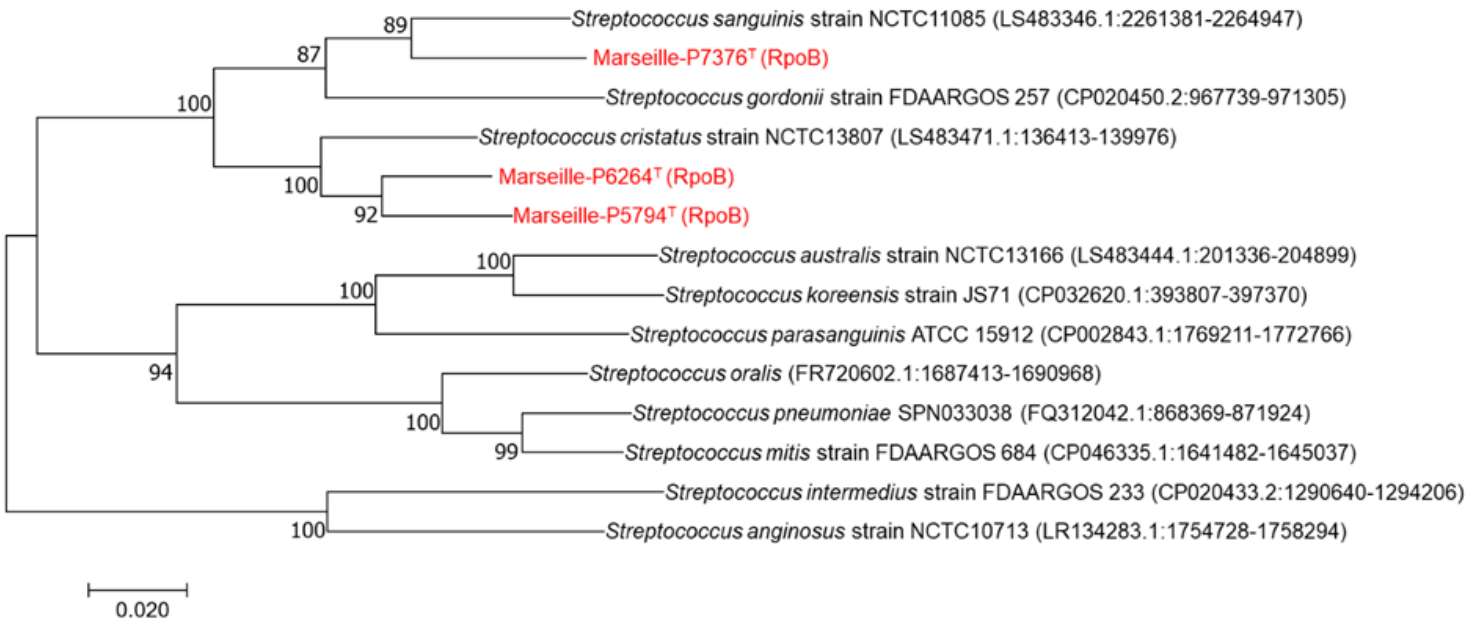

Figure 3

rpoB based-phylogenetic tree showing the position of strain Marseille-P5794T, strain Marseille-P6264T and strain Marseille-P7376T relatively to phylogenetically close species with Enterococcus gallinarium used as an outgroup. The isolates of interest are highlighted in red with rpoB gene sequences were extracted in silico from the annotation genome using Prokka online software (https://usegalaxy.org.au/). The scale bar indicates a 5 % nucleotide sequence divergence.

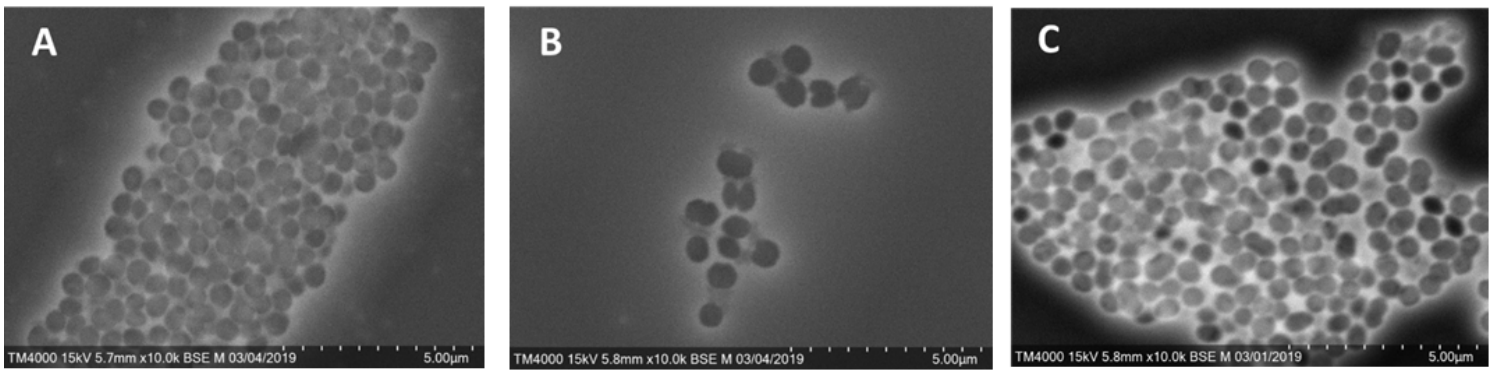

Figure 4

Scanning electron micrograph of strain Marseille-P5794T (A), strain Marseille-P6264T (B) and strain Marseille-P7376T (C) using TM4000 microscope from Hitachi (Japan). Scale bar and acquisition settings are shown on the original micrograph and image parameters were indicated at the bottom.

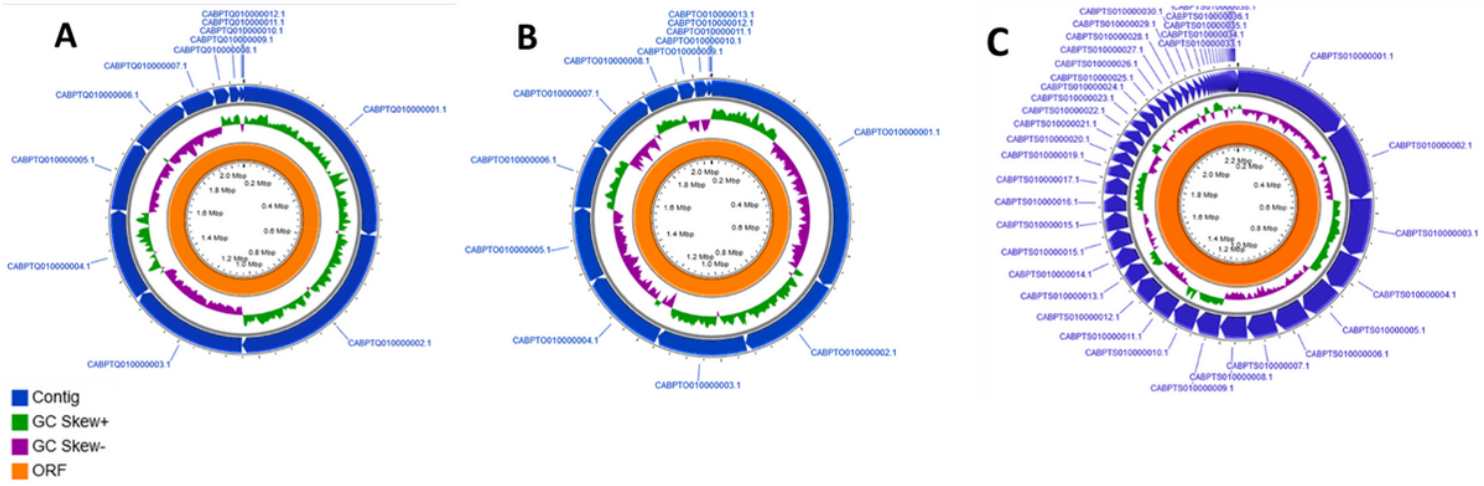

Figure 5

The graphical maps of genomes of (A) strain Marseille- P5794T, strain Marseille-P6264T (B) and strain Marseille-P7376T (C). These maps were obtained using CG view server (http://cgview.ca/).

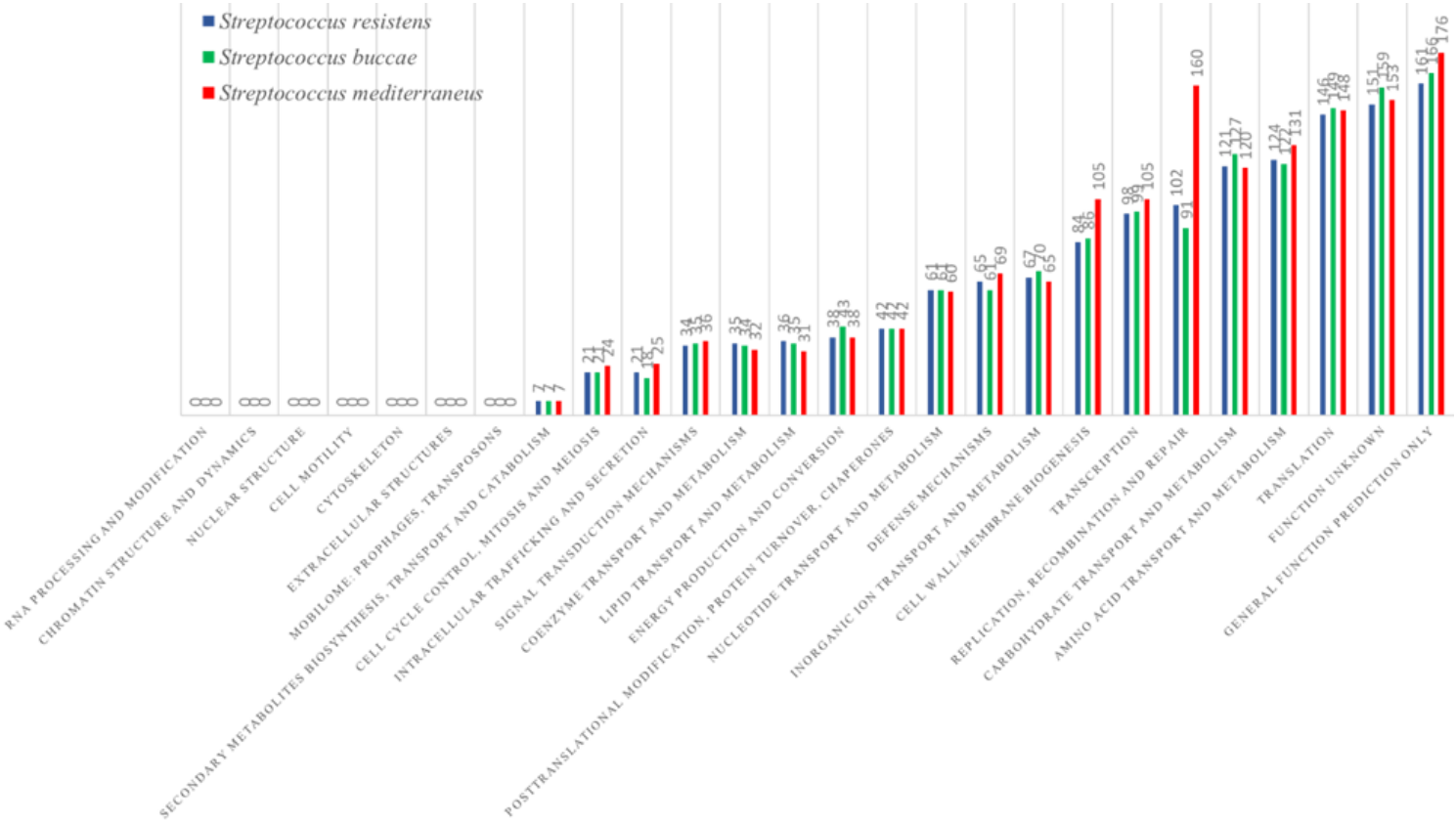

Figure 6

Numbers of genes of strain Marseille- P5794T (in blue), strain Marseille-P6264T (in green) and strain Marseille-P7376T (in orange) associated with general COG functional categories.

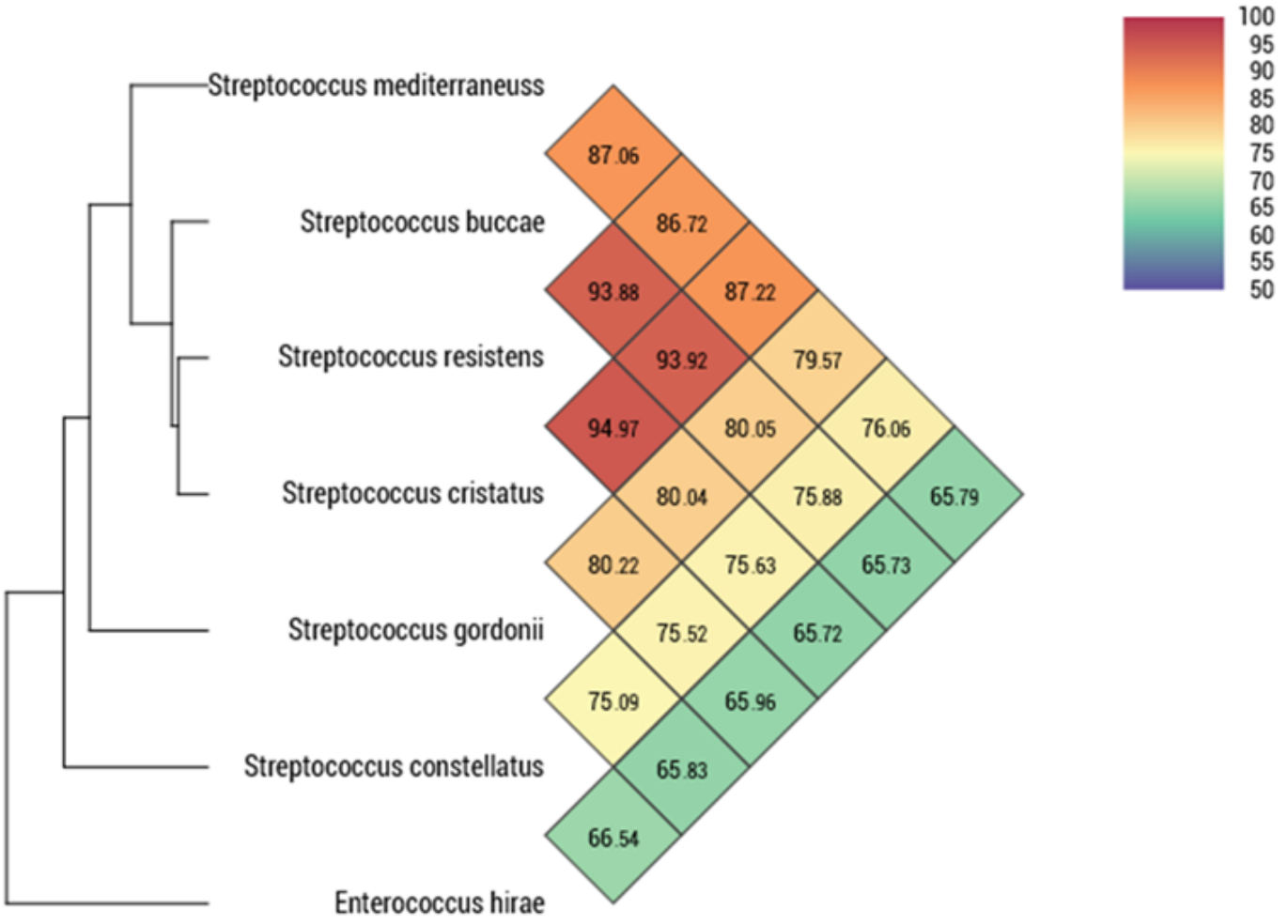

Figure 7

Heatmap generated with OrthoANI values calculated using the OAT software between strain Marseille-P5794T, strain Marseille-P6264T and strain Marseille-P7376T and other closely related species with standing in nomenclature.

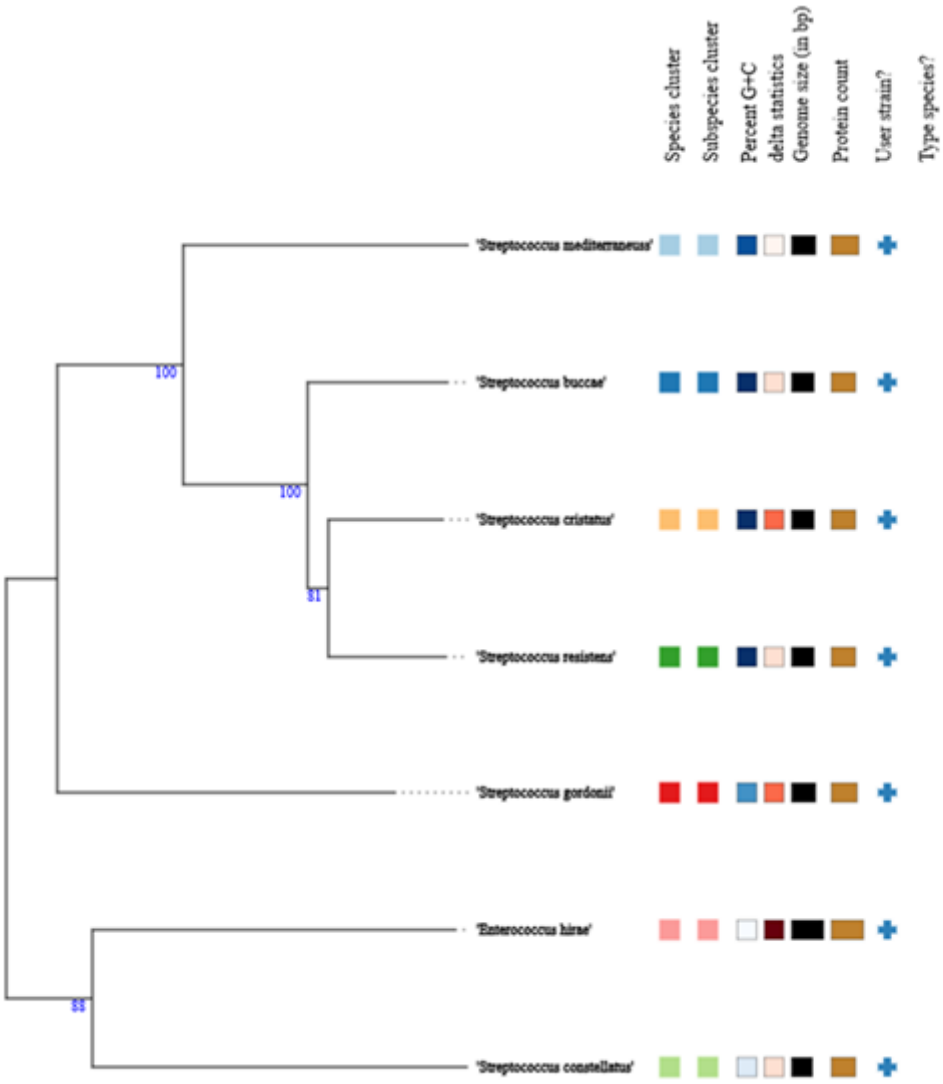

Figure 8

Tree based on Genome BLAST Distance Phylogeny (GBDP) distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d5 (https://tygs.dsmz.de/). The numbers above branches are GBDP pseudo-bootstrap support values > 60 % from 100 replications, with an average branch support of 92.3 %. This tree was obtained using TYGS server online (https://tygs.dsmz.de/) (Meier-Kolthoff and Göker 2019).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarydata.docx