Non-contiguous finished genome sequence and description of *Streptococcus varani* sp. nov.

S. Bakour¹, J. Rathored¹, C. I. Lo¹,², O. Mediannikov¹,², M. Beye¹, C. B. Ehounoud¹, P. Biagini³, D. Raoult¹,²,⁴, P.-E. Fournier¹ and F. Fenollar¹,²

1) Aix-Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Faculté de médecine, Marseille cedex 05, France, 2) Campus international UCAD-IRD, Dakar, Senegal, 3) Equipe Emergence et Co-évolution Virale, UMR 7268 – Aix-Marseille Université/ EFS / CNRS, IHU Méditerranée Infection, Marseille, France and 4) Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdul Aziz University, Jeddah, Saudi Arabia

Abstract

Strain FF10ᵀ (= CSUR P1489 = DSM 100884) was isolated from the oral cavity of a lizard (*Varanus niloticus*) in Dakar, Senegal. Here we used a polyphasic study including phenotypic and genomic analyses to describe the strain FF10ᵀ. Results support strain FF10ᵀ being a Gram-positive coccus, facultative anaerobic bacterium, catalase-negative, non-motile and non-spore forming. The sequenced genome counts 2.46 Mb with one chromosome but no plasmid. It exhibits a G+C content of 40.4% and contains 2471 protein-coding and 45 RNA genes. On the basis of these data, we propose the creation of *Streptococcus varani* sp. nov.

Keywords: Culturomics, genome, *Streptococcus varani*, taxonogenomics, *Varanus niloticus*

Original Submission: 5 February 2016; Accepted: 14 March 2016

Article published online: 19 March 2016

Introduction

The genus *Streptococcus* contains 113 recognized species and 22 subspecies ([http://www.bacterio.net/streptococcus.html](http://www.bacterio.net/streptococcus.html)) as of 5 February 2016. These species are Gram-positive cocci, chain-forming, facultative anaerobes and catalase negative [1]. Many of them are associated with human or animal hosts. Previous studies have reported that a large number of *Streptococcus* species colonize the oral cavities of humans and animals [2]. Among oral streptococci isolated from animals, we note: *Streptococcus ursoris* isolated from the oral cavities of bears [1], *Streptococcus orisratti* isolated from the surface of the lower molars of Sprague-Dawley rats [3], *Streptococcus oriloxodontae* isolated from the oral cavities of elephants [4], and *Streptococcus mutans* isolated from dental plaque [5].

Recently, with next-generation sequencing technology able to sequence whole genomes in a short time, and mass spectrometric analysis of bacteria, we have had easy access to genetic and proteomic information [6]. Therefore, we propose a polyphasic approach combining genomic properties in combination with matrix-assisted laser-desorption/ionization time-of-flight (MALDI-TOF) spectra and phenotypic characteristics to describe new bacterial species. The strain FF10ᵀ (= CSUR P1489 = DSM 100884) was isolated from the oral cavity of the Nile monitor (*Varanus niloticus*) in Dakar, Senegal.

Here, we present a summary classification and a set of features for *Streptococcus varani* sp. nov. strain FF10ᵀ (= CSUR P1489 = DSM 100884), including the description of its complete genome and annotation. These characteristics support the circumscription of the species *S. varani*.
Classification and Features of the Strain

Strain identification
In 2014, a sample was collected from the oral cavity of a lizard reptile named ‘Nile monitor’ (Varanus niloticus) in Dakar, Senegal, and stored at −80°C. In October 2014, the strain FF10T (Table 1) was isolated from this sample by cultivation on 5% sheep blood-enriched Columbia agar (BioMérieux, Marcy l’Étoile, France) at 37°C with 5% CO₂.

To identify the strain, MALDI-TOF MS protein analysis was performed as previously described [7] using a Microflex spectrometer (Bruker Daltonics, Leipzig, Germany). The 12 FF10T spectra were imported into the MALDI Biotyper software (version 2.0, Bruker) and analysed by standard pattern matching (with default parameter settings) against the main spectra of 6252 bacteria. From the resulting scores, the tested species may or may not be identified compared with the instrument’s database; a score ≥ 2 with a validly published species enables identification at the species level; a score ≥ 1.7 and < 2 allows identification at the genus level; and a score < 1.7 does not enable any identification. For strain FF10T, the scores obtained were lower than 1.306, suggesting that our strain was not a member of any known species. The reference mass spectrum of Streptococcus minor strain DSM 17118 was identified using 16S rRNA PCR coupled with sequencing, as previously described [8]. Strain FF10T exhibited 96% 16S rRNA sequence similarity with Streptococcus varani strain FF10T (Table 1) was isolated from this sample by cultivation on 5% sheep blood-enriched Columbia agar (BioMérieux, Marcy l’Étoile, France) at 37°C with 5% CO₂.

TABLE 1. Classification and general features of Streptococcus varani strain FF10T

| Property                  | Term                          | References  |
|---------------------------|-------------------------------|-------------|
| Classification            | Domain Bacteria               | [23]        |
|                           | Phylum Firmicutes             | [26,27]     |
|                           | Class Bacilli                 | [28]        |
|                           | Order Lactobacillales         | [29,30]     |
|                           | Family Streptococcaceae       | [31,32]     |
|                           | Genus Streptococcus           | [31,33,5]   |
|                           | Species Streptococcus varani  |             |
|                           | Type strain FF10T             |             |

Gram stain: Positive
Cell shape: Cocci
Motility: Non-motile
Sporulation: Non-spore forming
Temperature range: Mesophile
Optimum temperature: 37°C
Carbon source: Unknown
Habitat: Lizard
Salinity: Unknown
Oxygen requirement: Facultative anaerobe
Biotic relationship: Free living
Pathogenicity: Unknown
Geographic location: Dakar
Sample collection: October 22, 2014
Latitude: 13.7167
Longitude: −16.4167
Altitude: 51 m above sea level

Different growth temperatures (25°C, 30°C, 37°C, 45°C and 56°C) were tested. Growth was obtained between 25 and 37°C, with optimal growth at 37°C. Growth of the strain was tested also under anaerobic and microaerophilic conditions using GENbag anaer and GENbag microaer systems, respectively (bioMérieux), and under aerobic conditions, with or without 5% CO₂. Strain growth was observed under anaerobic and microaerophilic conditions but optimal growth was observed under aerobic conditions. Colonies were translucent and yellow with a regular surface, haemolytic on 5% sheep blood-enriched Columbia agar (bioMérieux), and approximately 1 mm in diameter. A motility test was negative. Cells were Gram-positive cocci, unable to form spores (Fig. 4), and with mean diameter of 0.6 μm (range 0.4–0.8 μm) and mean length of 1.3 μm (range 0.7–1.9 μm) (Fig. 5).

Biochemical characterization and antibiotic susceptibility
This bacterium, FF10T, exhibits neither catalase nor oxidase activities. Using an API ZYM strip (bioMérieux), positive reactions were observed for esterases, β-galactosidase, β-glucosidase, β-glucuronidase, naphthol-AS-BI-phosphorylase, and N-acetyl-β-glucosaminidase. Using an API 20 NE strip (bioMérieux), positive reactions were observed only for the esculin hydrolysis test whereas negative reactions were observed for nitrate reduction, urease, indole production, arginine dihydrolase, glucose fermentation, arabinose, mannose, mannitol, N-acetyl-glucosamine, maltose, glucuronate, caprate, adipate, malate, citrate, phenyl-acetate assimilation, and gelatin hydrolysis.

Using API 50 CH strip (bioMérieux), negative reactions were observed for the fermentation of glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xyllose, L-xyllose, D-adenosine, methyl-β-D-d-xylopyranoside, D-galactose, D-glucose, D-fructose, D-mannose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-β-D-d-xylopyranoside, methyl-β-D-glucopyranoside, N-acetylglucosamine, amygdalin, arbutin,
FIG. 1. Reference mass spectrum from *Streptococcus varani* sp. nov. strain FF10T. Spectra from 12 individual colonies were compared and a reference spectrum was generated.

FIG. 2. Gel view comparing *Streptococcus varani* sp. nov. strain FF10T spectra with other members of the *Streptococcus* genus. The gel view displays the raw spectra of loaded spectrum files arranged in a pseudo-gel like look. The x-axis records the m/z value. The left y-axis displays the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by a Grey-scale scheme code. The colour bar and the right y-axis indicate the relation between the colour in which a peak is displayed and the peak intensity in arbitrary units. Displayed species are indicated on the left.
salicin, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-saccharose, D-trehalose, inulin, D-melezitose, D-raf fines, amidon, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate and potassium-5-ketogluconate.

Streptococcus varani strain FF10T was susceptible to penicillin, amoxicillin, amoxicillin/clavulanic acid, ticarcillin, ceftriaxone, cefalotin, imipenem, gentamicin, kanamycin, trimethoprim/sulfamethoxazol, erythromycin, doxycycline, metronidazole, vancomycin and rifampicin, but resistant to ciprofloxacin, nitrofurantoin and colistin.

The differential phenotypic characteristics with other Streptococcus species (S. minor [10], Streptococcus plurextorum [11], S. ursoris [1] and S. orisratti [3]) are summarized in Table 2.

### Genome Description

#### Genome sequencing and assembly

The DNA was extracted using the phenol/chloroform method as previously described [12]. Genomic DNA (gDNA) of S. varani FF10T was sequenced on the MiSeq Technology (Illumina Inc., San Diego, CA, USA) using the mate-pair strategy. The gDNA was bar-coded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). The gDNA was quantified by a Qubit assay with the high sensitivity kit (Thermo Fisher Scientific, Waltham, MA, USA) to 62.4 mg/L. The mate-pair library was prepared with 1.5 μg of genomic DNA using the Nextera mate-pair Illumina guide. The gDNA sample was simultaneously fragmented and tagged with a mate-pair junction adapter. The fragmentation pattern was validated on an Agilent 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA, USA) with a DNA 7500 LabChip. The
DNA fragments ranged from 1.5 kb up to 11 kb with optimal size of 5.8 kb. No size selection was performed and 591 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments, optimally at 697 bp, on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc.) and the final concentration library was measured at 75.52 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 were performed in a single 39-h run in a 2 × 251-bp.

Total information of 5.9 Gb was obtained from a 624 K/mm² cluster density with a cluster passing quality control filters of 96.33% (12 040 000 clusters). Within this run, the index representation for S. varani FF10T was determined to 5.26%. The 610 160 paired reads were filtered according to the read qualities. These reads were trimmed then assembled.

**Genome Annotation and Genome Analysis**

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 GenBank[13] and Clusters of Orthologous Groups (COG) databases using BLASTP. The tRNAs and rRNAs were predicted using the tRNAScan-SE[14] and RNAmmer[15] tools, respectively. Signal peptides and numbers of transmembrane helices were predicted using

| Characteristics | S. varani | S. minor | S. plurextorum | S. ursoris | S. orisratti |
|-----------------|-----------|---------|---------------|-----------|-------------|
| Cell diameter (μm) | 0.4–0.8 | <1 | na | 0.5–0.7 | na |
| Oxygen requirement | Facultative anaerobic | Microaerobic | Facultative anaerobic | Facultative anaerobic | Facultative anaerobic |
| Gram stain | + | + | + | + | + |
| Motility | – | – | – | – | – |
| Endospore formation | – | – | – | – | – |
| Production of: | | | | | |
| Alkaline phosphatase | + | – | – | – | – |
| Acid phosphatase | + | + | + | + | + |
| Catalase | – | na | – | – | – |
| Oxidase | – | na | – | na | na |
| Nitrate reductase | – | na | – | na | na |
| Urease | – | – | – | – | – |
| α-galactosidase | – | W | + | + | + |
| β-galactosidase | + | – | + | + | + |
| β-glucuronidase | – | – | – | – | – |
| α-glucosidase | – | na | – | + | na |
| β-glucosidase | + | na | – | + | na |
| Esterase | + | na | – | na | na |
| Esterase lipase | + | na | – | na | na |
| Naphthol-AS-BI-phosphohydrolase | – | na | – | + | + |
| N-acetyl-β-glucosaminidase | – | na | – | – | na |
| Utilization of: | | | | | |
| Gluconate | – | – | DP | na | na |
| α-ribose | – | – | – | + | – |
| d-ribose | – | – | – | + | – |
| α-xylene | – | – | – | + | – |
| d-xylene | – | – | – | + | – |
| α-fructose | – | + | + | + | + |
| d-fructose | – | + | + | + | + |
| α-glucose | – | + | + | + | + |
| d-mannose | – | + | + | + | + |
| β-mannose | – | + | + | + | + |
| α-lactose | – | + | + | + | + |
| β-lactose | – | + | + | + | + |
| α-maltose | – | + | + | + | + |
| d-maltose | – | + | + | + | + |
| α-arabinose | – | + | + | + | + |
| α-arabinose | – | + | + | + | + |
| d-arabinose | – | + | + | + | + |
| d-arabinose | – | + | + | + | + |
| Habitats | Lizard | Dog, cat and calf | Pigs | Bears | Rats |

*+, positive; –, negative; DP, delayed production; W, weak reaction; na, not available.
SignalP [16], and TMHMM [17], respectively. Mobile genetic elements were predicted using PHAST [14] and RAST [18]. ORFans were identified if their BLASTP E-value was lower than 1e-03 for alignment length >80 amino acids. If alignment lengths were <80 amino acids, we used an E-value of 1e-05.

Such parameter thresholds have already been used in previous work to define ORFans. Artemis [19] and DNA Plotter [20] 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 [21].

The mean level of nucleotide sequence similarity at the genome level between S. varani and other bacteria (S. orisratti (ARCG01000001.1), S. tigurinus (AORU01000001.1), Streptococcus parasanguinis (CP002843.1), Streptococcus agalactiae (AEQQ01000001.1), S. plurextorum (AUIO01000001.1), S. minor (AQYB01000001.1), Streptococcus oralis (ADMV01000001.1), Streptococcus pyogenes (AE004092.2), Streptococcus vestibularis (AEVI01000001.1), Streptococcus peroris (AEVF01000001.1) and Lactococcus lactis (NC_002662.1)) was estimated using the Average Genomic Identity of gene Sequences (AGIOS) home-made software [22]. Overall, this software combines the functionality of other software programs: PROTEINORTHO [23] (detects orthologous proteins between genomes compared two by two, then retrieves the corresponding genes) and the Needleman–Wunsch global

**TABLE 3. Nucleotide content and gene count levels of the genome**

| Attribute                  | Value          | % of total |
|----------------------------|----------------|------------|
| Genome size (bp)           | 2,460,376      | 100        |
| DNA coding (bp)            | 2,170,614      | 88.2       |
| DNA G+C (bp)               | 993,991        | 40.4       |
| Total genes                | 2,516          | 100        |
| Protein coding genes       | 2,471          | 98.2       |
| RNA genes                  | 45             | 1.78       |
| Genes with function prediction | 1,185        | 73.60      |
| Genes assigned to COGs     | 1,746          | 69.39      |
| Genes with signal peptides | 118            | 4.68       |
| Genes with transmembrane helices | 565         | 22.45      |
| Genes associated with resistant genes | 5         | 0.19       |
| ORFans genes               | 0              | 0          |
| CRISPR repeats             | 0              | 0          |
| Genes with Pfam-A domains  | 224            | 89.98      |

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

**FIG. 6.** Graphical circular map of the chromosome. From outside to the centre. Genes on the forward strand coloured by COG categories (only genes assigned to COG), genes on the reverse strand coloured by COG categories (only gene assigned to COG), RNA genes (tRNAs green, rRNAs red), GC content, and GC skew.

New Microbes and New Infections © 2016 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, NMNI, 11, 93–102

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
TABLE 4. Number of genes associated with the 25 general COG functional categories*

| Code | Value | %   | Description                                      |
|------|-------|-----|--------------------------------------------------|
| J    | 144   | 8.18| Translation                                      |
| A    | 0     | 0.00| RNA processing and modification                  |
| K    | 101   | 5.74| Transcription                                    |
| L    | 91    | 5.17| Replication, recombination and repair            |
| B    | 0     | 0.00| Chromatin structure and dynamics                 |
| D    | 23    | 1.31| Cell cycle control, mitosis and meiosis          |
| Y    | 0     | 0.00| Nuclear structure                                |
| V    | 37    | 2.10| Defence mechanisms                              |
| T    | 41    | 2.33| Signal transduction mechanisms                   |
| M    | 83    | 4.71| Cell wall/membrane biogenesis                    |
| N    | 5     | 0.28| Cell motility                                   |
| Z    | 0     | 0.00| Cytoskeleton                                    |
| W    | 0     | 0.00| Extracellular structures                         |
| U    | 20    | 1.13| Intracellular trafficking and secretion          |
| O    | 54    | 3.07| Post-translational modification, protein turnover, chaperones |
| C    | 37    | 2.10| Energy production and conversion                 |
| G    | 111   | 6.30| Carbohydrate transport and metabolism            |
| E    | 145   | 8.24| Amino acid transport and metabolism              |
| F    | 69    | 3.92| Nucleotide transport and metabolism              |
| H    | 42    | 2.39| Coenzyme transport and metabolism                |
| I    | 35    | 1.99| Lipid transport and metabolism                   |
| P    | 71    | 4.03| Inorganic ion transport and metabolism           |
| Q    | 13    | 0.74| Secondary metabolites biosynthesis, transport and catabolism |
| R    | 156   | 8.86| General function prediction only                 |
| S    | 154   | 8.75| Function unknown                                 |
| C21  | 328   | 18.64| Not in COGs                                     |

*The total is based on the total number of protein-coding genes in the annotated genome.

alignment algorithm (determines the mean percentage of nucleotide sequence identity among orthologous ORFs).

Genome properties

The genome is 2,460,376 bp long with 40.4% GC content (Table 3). It is composed of 16 contigs (Fig. 6). Of the 2,516 predicted genes, 2,471 were protein-coding genes, and 45 were RNAs (one gene is 55 rRNA, one gene is 16S rRNA, one gene is 23S rRNA, 42 genes are tRNA genes). No genes were identified as ORFans (0.0%). The remaining genes were annotated as hypothetical proteins (293 genes ≥ 11.64%). Table 4 distributes the genes into COG functional categories. The genome sequence has been deposited in GenBank under accession number GCA_001375655.

Comparison with other genomes

The draft genome of *S. varani* (2.46 Mb) is larger than those of *S. agalactiae* (2.11 Mb), *S. minor* (1.93 Mb), *S. oralis* (1.90 Mb), *S. orisratti* (2.41 Mb), *S. parasangunis* (2.15 Mb), *S. peroris* (1.63 Mb), *S. plurextorum* (2.10 Mb), *S. pyogenes* (1.85 Mb), *S. tigurinus* (2.18 Mb), *S. vestibularis* (1.84 Mb). The G+C content of *S. varani* (40.40%) is lower than those of *S. minor* (41.10%), *S. oralis* (44.40%), *S. parasangunis* (41.70%), *S. plurextorum* (41.10) but higher than those of *S. agalactiae* (35.30%), *S. orisratti* (38.50%), *S. peroris* (39.10%), *S. pyogenes* (38.50%), *S. tigurinus* (40.30%) and *S. vestibularis* (39.60%).

Similarly, the protein-coding genes of *S. varani* (2,471) are larger than those of *S. agalactiae*, *S. minor*, *S. oralis*, *S. orisratti*, *S. parasangunis*, *S. peroris*, *S. plurextorum*, *S. pyogenes*, *S. tigurinus* and *S. vestibularis* (2139, 1903, 1795, 2319, 1982, 1608, 2081, 1776, 2146 and 1812, respectively). The gene content of *S. varani* (2516) is larger than that of *S. agalactiae* (2348), *S. minor* (2055), *S. oralis* (1776), *S. orisratti* (2116), *S. parasangunis* (1854), *S. plurextorum* (1838), *S. pyogenes* (2197), *S. tigurinus* (1838), *S. vestibularis* (1776), *S. varani* (1812) and *S. vestibularis* (1838), respectively.

However, the COG category gene distribution is similar in all genomes compared (Fig. 7). Among species with standing in nomenclature, AGIOS values ranged from 81.54 between *S. varani* and *S. peroris* (20%) but higher than those of *S. agalactiae* (38.50%)

**FIG. 7.** Distribution of predicted genes of Streptococcus varani strain FF10T and 11 other closely related species into COG categories.
colonies are translucent and yellow with a regular surface, between 25°C and 37°C, with optimal growth at 37°C. The anaerobe, and Gram-positive coccus. Growth is obtained isolated) is a non-motile, non-spore forming, facultative anaerobe, and Gram-positive coccus. Growth is obtained between 25°C and 37°C, with optimal growth at 37°C. The colonies are translucent and yellow with a regular surface,

### Conclusion

On the basis of phenotypic, phylogenetic and genomic analyses, we formally propose the creation of Streptococcus varani sp. nov. that contains the strain FF10T. This bacterial strain has been isolated from the oral cavities of a lizard (Varanus niloticus) in Dakar, Senegal.

### Description of Streptococcus varani sp. nov. strain FF10T

Streptococcus varani (va.ra’ni. NL. gen. n. varani, from Varanus, the Latin name from the lizard which the type strain was isolated) is a non-motile, non-spore forming, facultative anaerobe, and Gram-positive coccus. Growth is obtained between 25°C and 37°C, with optimal growth at 37°C. The colonies are translucent and yellow with a regular surface,

### Table 5. The numbers of orthologous protein shared between genomes (upper right)a

|          | S. varani | S. minor | S. agalactiae | S. oralis | S. orisratti | S. parasanguinis | S. peroris | S. pluerextorum | S. pyogenes | S. tigurinus | S. vestibularis | L. lactis |
|----------|-----------|----------|---------------|-----------|--------------|-----------------|------------|----------------|-------------|--------------|----------------|----------|
| Streptococcus varani | 2471 | 476 | 1079 | 1016 | 1133 | 1067 | 975 | 1044 | 972 | 1089 | 983 | 952 |
| Streptococcus minor | 76.99 | 1903 | 997 | 1000 | 1089 | 1045 | 962 | 1024 | 930 | 1023 | 989 | 902 |
| Streptococcus agalactiae | 69.62 | 69.74 | 2139 | 962 | 1082 | 1030 | 936 | 1020 | 1022 | 1036 | 1008 | 915 |
| Streptococcus oralis | 70.91 | 70.91 | 69.62 | 1795 | 1025 | 1149 | 1171 | 994 | 911 | 1261 | 1031 | 887 |
| Streptococcus orisratti | 71.00 | 70.83 | 73.03 | 70.46 | 2319 | 1110 | 1001 | 1059 | 976 | 1092 | 1066 | 930 |
| Streptococcus parasanguinis | 70.27 | 70.37 | 69.26 | 74.95 | 69.97 | 1982 | 1101 | 1061 | 959 | 1156 | 1068 | 915 |
| Streptococcus peroris | 70.79 | 70.69 | 70.31 | 81.54 | 70.73 | 75.43 | 1608 | 950 | 870 | 1170 | 978 | 845 |
| Streptococcus pluerextorum | 70.19 | 70.31 | 70.95 | 70.04 | 72.20 | 69.49 | 70.12 | 2081 | 951 | 1031 | 946 | 856 |
| Streptococcus pyogenes | 70.19 | 70.00 | 74.23 | 70.10 | 73.04 | 69.97 | 70.59 | 71.26 | 1176 | 932 | 903 | 840 |
| Streptococcus tigurinus | 70.38 | 70.52 | 69.82 | 93.09 | 70.19 | 74.91 | 81.42 | 69.96 | 69.92 | 2146 | 1051 | 892 |
| Streptococcus vestibularis | 70.47 | 70.45 | 71.67 | 71.82 | 72.94 | 72.11 | 72.02 | 71.56 | 72.05 | 71.69 | 1812 | 892 |
| Lactococcus lactis | 65.60 | 65.36 | 66.90 | 66.33 | 66.46 | 66.03 | 66.70 | 65.63 | 66.68 | 66.09 | 66.63 | 2650 |

*aAverage percentage similarity of nucleotides corresponding to orthologous protein shared between genomes (lower left) and numbers of proteins per genome (bold).

### Table 6. Pairwise comparison of Streptococcus varani with eight other species using GGDC, formula 2 (DDH estimates based on identities / HSP length)a

|          | S. varani | S. minor | S. agalactiae | S. oralis | S. orisratti | S. parasanguinis | S. peroris | S. pluerextorum | S. pyogenes | S. tigurinus | S. vestibularis | L. lactis |
|----------|-----------|----------|---------------|-----------|--------------|-----------------|------------|----------------|-------------|--------------|----------------|----------|
| Streptococcus varani | 100% ± 00 | 24.7% ± 2.57 | 25.7% ± 2.53 | 25.0% ± 2.57 | 22.4% ± 2.56 | 22.7% ± 2.57 | 25.2% ± 2.60 | 26.3% ± 2.72 | 23.3% ± 2.53 | 25.9% ± 2.25 | 1.1% ± 2.59 |
| Streptococcus minor | 100% ± 00 | 23.4% ± 2.53 | 25.0% ± 2.57 | 23.2% ± 2.58 | 24.8% ± 2.73 | 25.0% ± 3.01 | 22.7% ± 2.58 | 27.4% ± 2.59 | 49.8% ± 3.14 | 3.1% ± 2.54 | 4.6% ± 2.64 |
| Streptococcus agalactiae | 100% ± 00 | 25.0% ± 2.57 | 22.4% ± 2.56 | 25.6% ± 2.56 | 22.7% ± 2.57 | 25.4% ± 2.63 | 26.3% ± 2.72 | 23.3% ± 2.53 | 25.9% ± 2.25 | 1.1% ± 2.59 |
| Streptococcus oralis | 100% ± 00 | 24.9% ± 2.57 | 23.5% ± 2.59 | 25.6% ± 2.56 | 22.3% ± 2.56 | 22.3% ± 2.56 | 23.2% ± 2.56 | 23.5% ± 2.60 | 25.3% ± 2.60 | 1.1% ± 2.59 |
| Streptococcus orisratti | 100% ± 00 | 24.9% ± 2.57 | 23.5% ± 2.59 | 25.6% ± 2.56 | 22.3% ± 2.56 | 23.2% ± 2.56 | 23.5% ± 2.60 | 25.3% ± 2.60 | 1.1% ± 2.59 |
| Streptococcus parasanguinis | 100% ± 00 | 26.3% ± 2.77 | 23.1% ± 2.58 | 25.8% ± 2.74 | 27.6% ± 2.59 | 24.5% ± 2.72 | 27.4% ± 2.68 |
| Streptococcus peroris | 100% ± 00 | 24.9% ± 2.57 | 25.8% ± 2.59 | 24.9% ± 3.01 | 25.9% ± 2.25 | 1.1% ± 2.59 |
| Streptococcus pluerextorum | 100% ± 00 | 23.8% ± 2.60 | 21.9% ± 2.58 | 23.5% ± 2.58 | 25.3% ± 2.58 | 25.3% ± 2.58 | 25.3% ± 2.58 |
| Streptococcus pyogenes | 100% ± 00 | 25.9% ± 2.59 | 24.7% ± 2.61 |
| Streptococcus tigurinus | 100% ± 00 | 26.9% ± 2.64 |
| Streptococcus vestibularis | 100% ± 00 | 26.9% ± 2.64 |

*aThe confidence intervals indicate the inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical test data sets (which are always limited in size). These results are in accordance with the 16S rRNA (Fig. 3) and phylogenomic analyses as well as the GGDC results.

New Microbes and New Infections © 2016 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, NMNI, 11, 93–102

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Authors’ Contributions

SB, JR, and MB performed the genomic analyses and drafted the manuscript. CIL and CBE performed the phenotypic characterization of the bacterium and drafted the manuscript. OM participated in its design and helped to draft the manuscript. PB helped to draft the manuscript. DR conceived the study and helped to draft the manuscript. PEF and FF conceived the study, participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Transparency Declaration

The authors declare no conflict of interest.

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 Claudia Andrieu for administrative assistance.

References

[1] Shinozaki-Kuwahara N, Takada K, Hirasawa M. Streptococcus urorii sp. nov., isolated from the oral cavities of bears. Int J Syst Evol Microbiol 2015;65:1:40–4.
[2] Facklam R. What happened to the streptococci: overview of taxonomic and nomenclature changes. Clin Microbiol Rev 2002;15:613–30.
[3] Zhu H, Willcox MDP, Knox KW. A new species of oral Streptococcus isolated from Sprague-Dawley rats, Streptococcus orrisrati sp. nov. Int J Syst Evol Microbiol 2000;50:55–61.
[4] Shinozaki-Kuwahara N, Saito M, Hirasawa M, Takada K. Streptococcus orisrati sp. nov., isolated from the oral cavities of elephants. Int J Syst Evol Microbiol 2014;64:3755–9.
[5] Hamada S, Slade H. Biology, immunology, and cariogenicity of Streptococcus mutans. Microbiol Rev 1980;44:331–84.
[6] Welker M, Moore ER. Applications of whole-cell matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst Appl Microbiol 2011;34:2–11.
[7] 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 In-fect Dis 2009;49:543–51.
[8] Mourembou G, Rathore J, Lekana-Douki JB, Ndjojyi-Mbiguino A, Fenollar F, Michelle C, et al. Non-contiguous finished genome sequence and description of Kalliopyga gabonensis sp. nov. New Microbes New Infect 2015;9:15–23.
[9] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.
[10] Vancanneyt M, Devriese LA, Haesebrouck F, Baele M, Lefebvre K, Snaauwaert C, et al. Streptococcus minor sp. nov., from fecal samples and tonsils of domestic animals. Int J Syst Evol Microbiol 2004;54:449–52.
[11] Vela A, Casamayor A, Sanchez Del Rey V, Dominguez L, Fernandez-Garayzabal J. Streptococcus pluteumtor sp. nov., isolated from pigs. Int J Syst Evol Microbiol 2009;59:504–8.
[12] Lagier JC, Bibi F, Ramasamy D, Azhar EI, Robert C, Yasir M, et al. Non contiguous-finished genome sequence and description of Clostridium jeddahense sp. nov. Base, ESBGHost, 2014 (accessed November 24, 2015).
[13] Benson DA, Karchs-Mizrachi I, Clark K, Lipman DJ, Ostell J. Sayers EW. GenBank. Nucl Acids Res 2012;40:D48–53.
[14] Lowe TM, Eddy SR. RNA-Scan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucl Acids Res 1997;25:955–64.
[15] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RnaNner: consistent and rapid annotation of ribosomal RNA genes. Nucl Acids Res 2007;35:3100–8.
[16] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004;340:783–95.
[17] 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.
[18] Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: a fast genome search tool. Nucl Acids Res 2011;39:W437–52.
[19] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. Bioinformatics 2000;16:944–5.
[20] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 2009;25:119–20.
[21] Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 2004;14:1394–403.
[22] Ramasamy D, Mishra AK, Lagier JC, Padmanaban R, Rossi-Tamisier M, Sentausa E, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of new bacterial species. Int J Syst Evol Microbiol 2014;64:384–91.
[23] Lechner M, Findeis S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co)orthologs in large-scale analysis. BMC Bioinformatics 2011;12:124.
[24] Auch AF, Von Jan M, Klenk HP, Goker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2010;2:117–34.
[25] Woese CR, Kandler O, Woese ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 1990;87:4576–9.
[26] Gibbons NE, Murray RGE. Proposals Concerning the Higher Taxa of Bacteria. Int J Syst Bacteriol 1978;28:1–6.
[27] Murray RGE. The Higher Taxa, or, a Place for Everything...? In: Holt JG, editor. Bergey's Manual of Systematic Bacteriology. First EditionVolume 1. Baltimore, MD: The Williams and Wilkins Co.; 1984. p. 31–4.
[28] Euzéby J. List of new names and new combinations previously effectively, but not validly, published. List no. 132. Int J Syst Evol Microbiol 2010;60:469–72.
[29] Ludwig W, Schleifer KH, Whitten WM, Class I. Bacilli class nov. In: De Vos P, Garrity G, Jones D, Krieg NR, Ludwig W, Rainey FA, Schleifer KH, Whitten WM, editors. Bergey’s Manual of Systematic Bacteriology. Second EditionVolume 3. New York: Springer-Verlag; 2009. p. 19–20.
[30] Ludwig W, Schleifer KH, Whitten WM, Order II. Lactobacilloidea ord. nov. In: De Vos P, Garrity G, Jones D, Krieg NR, Ludwig W, Rainey FA, Schleifer KH, Whitten WM, editors. Bergey’s Manual of Systematic Bacteriology. Second EditionVolume 1. Baltimore, MD: The Williams and Wilkins Co.; 1984. p. 31–4.
[31] Skerman VBD, McGowan V, Sneath PHA. Approved Lists of Bacterial Names. Int J Syst Bacteriol 1980;30:225–420.

[32] Deibel RH, Seeley HW. Family II. Streptococcaceae. In: Buchanan RE, Gibbons NE, editors. Bergey’s Manual of Determinative Bacteriology. Eighth Edition. Baltimore, MD: The Williams and Wilkins Co.; 1974. p. 490–515.

[33] Rosenbach FJ. In: Bergmann JF (ed), Microorganismen bei den Wund-Infections-Krankheiten des Menschen., Wiesbaden; 1884, p. 1–122.

[34] Deibel RH, Seeley HW. Genus I. Streptococcus Rosenbach 1884, 22. In: Buchanan RE, Gibbons NE, editors. Bergey’s Manual of Determinative Bacteriology, Eighth Edition. Baltimore: The Williams and Wilkins Co.; 1974. p. 490–509.