



HAL
open science

Species identification of staphylococci by amplification and sequencing of the *tuf* gene compared to the *gap* gene and by matrix-assisted laser desorption ionization time-of-flight mass spectrometry.

Marie Bergeron, Olivier Dauwalder, Manolo Gouy, Anne-Marie Freydiere, Michèle Bes, Hélène Meugnier, Yvonne Benito, Jérôme Etienne, Gérard Lina, François Vandenesch, et al.

► To cite this version:

Marie Bergeron, Olivier Dauwalder, Manolo Gouy, Anne-Marie Freydiere, Michèle Bes, et al.. Species identification of staphylococci by amplification and sequencing of the *tuf* gene compared to the *gap* gene and by matrix-assisted laser desorption ionization time-of-flight mass spectrometry.. *European Journal of Clinical Microbiology and Infectious Diseases*, 2011, 30 (3), pp.343-54. 10.1007/s10096-010-1091-z . hal-00698367

HAL Id: hal-00698367

<https://hal.science/hal-00698367v1>

Submitted on 31 Jul 2012

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.

1 **Title**

2 Species identification of staphylococci by amplification and sequencing of the *tuf*
3 gene compared to the *gap* gene and by matrix-assisted laser desorption ionization
4 time-of-flight mass spectrometry

5

6 **Running title:** *tuf* sequence-based identification of Staphylococci

7

8 Marie BERGERON¹, Olivier DAUWALDER^{1,2,3}, Manolo GOUY⁴, Anne-Marie
9 FREYDIERE¹, Michèle BES^{1,2,3}, Hélène MEUGNIER^{1,2,3}, Yvonne BENITO^{1,2,3},
10 Jerome ETIENNE^{1,2,3}, Gerard LINA^{1,2,3}, François VANDENESCH^{1,2,3*}, Sandrine
11 BOISSET^{1,2,3,*}

12

13 ¹Hospices Civils de Lyon, Laboratoire de Bactériologie, Centre de Biologie et
14 Pathologie Est, Bron, France

15 ²Université Lyon 1, Centre National de Référence des Staphylocoques, Lyon, France

16 ³INSERM U851, IFR128, Lyon, France

17 ⁴Laboratoire de Biométrie et Biologie Evolutive, CNRS UMR 5558, Université de
18 Lyon, F-69622 Villeurbanne, France

19

20 * François Vandenesch & Sandrine Boisset share senior authorship

21

22 Corresponding author: mailing address: Laboratoire de Bacteriologie, Centre
23 National de Référence des Staphylocoques, 59 Boulevard Pinel, 69677 BRON
24 Cedex, France. Phone: 33 4 72 12 96 64. Fax: 33 4 72 35 73 35. Email:

25 sandrine.boisset@univ-lyon.fr

26

27 **KEY-WORDS:** *Staphylococcus* identification, *tuf* sequencing, *gap* sequencing,
28 phylogeny, MALDI-TOF-MS

29 **Abstract:**

30 Staphylococcal species notably, coagulase-negative staphylococci are frequently
31 misidentified using phenotypic methods.

32 The partial nucleotide sequences of the *tuf* and *gap* genes were determined in 47
33 reference strains to assess their suitability, practicability and discriminatory power as
34 target molecules for staphylococcal identification. The partial *tuf* gene sequence was
35 selected and further assessed with a collection of 186 strains including 35 species
36 and sub-species. Then, to evaluate the efficacy of this genotyping method versus the
37 technology of matrix-assisted laser desorption ionization-time of flight mass
38 spectrometry (MALDI-TOF-MS), the 186 strains were identified using MALDI-TOF-
39 MS (Axima[®] Shimadzu) coupled to the SARAMIS[®] database (AnagnosTec). The
40 French National Reference Center for staphylococci identification method was used
41 as a reference.

42 One hundred eighty-four strains (98.9%) were correctly identified by *tuf* gene
43 sequencing. Only one strain was misidentified, and one was unidentified. MALDI-
44 TOF-MS identified properly 138 isolates (74.2%). Four strains were misidentified, 39
45 were unidentified, 5 were identified at the group (*hominis/warneri*) level and 1 strain
46 was identified at the genus level.

47 These results confirm the value of MALDI-TOF-MS identification for common species
48 in clinical laboratory practice and the value of the partial *tuf*-gene sequence for the
49 identification of all staphylococcal species as required in a reference laboratory.

50

51 Introduction

52

53 According to current knowledge, including the newly described species published in
54 2009-2010, the *Staphylococcus* genus groups together 45 species and 21
55 subspecies [1, 17, 35, 40, 48]. Staphylococcal species are widely distributed in
56 various environments: the skin and mucous membranes of humans and animals as
57 well as soil, sand, and water. Some staphylococcal species are used as starter
58 cultures for sausage manufacturing in the food industry (*Staphylococcus xylosus* and
59 *S. carnosus*) [7], whereas others are mainly associated with animal diseases such as
60 *S. pseudintermedius* in dogs. Of the 45 species and 21 subspecies, only half have
61 been cultured from human specimens. *S. aureus* is the most clinically relevant
62 staphylococcal species, but coagulase-negative staphylococci (CoNS) are
63 increasingly recognized as etiologic agents of clinical manifestations in humans.
64 CoNS have been identified as a major cause of hospital-acquired infections that
65 typically affect immunocompromised patients with implanted medical devices [52].
66 Treatment is difficult because many CoNS species carry multiple antibiotic
67 resistances, notably methicillin resistance in approximately 55-75% of nosocomial
68 isolates, as well as glycopeptide resistance, which was initially described in CoNS
69 strains [5, 36]. Identification to the species level is necessary to provide a better
70 understanding of pathogenic potential of various CoNS and could help therapeutic
71 clinical decision [18]. Furthermore, the accurate identification to the species level in
72 reference laboratories is important to establish the role of each staphylococcal
73 species as an infectious agent and to conduct epidemiologic investigations.

74 Several manual and automated phenotypic identification systems are available, such
75 as the ID32 STAPH[®] strip (bioMérieux), the VITEK 2 GP[®] identification card
76 (bioMérieux) and the PID 61 Phoenix system (Becton Dickinson), but none of these
77 systems are able to accurately identify all staphylococcal species [8, 22, 28]. These
78 methods have been designed mainly for the most frequently encountered species in
79 human clinical samples and are not able to identify rare species and atypical strains
80 such as metabolic variants of common species. More recently, peptide spectra
81 obtained by matrix-assisted laser desorption ionization-time of flight mass
82 spectrometry (MALDI-TOF-MS) have been used to identify CoNS; this technique has
83 a good performance overall for species encountered in clinical practice [6, 10, 11, 45,
84 47]. Sample preparation and analysis techniques are simple and can be performed

85 within minutes. In addition to phenotypic methods, several PCR-sequencing-based
86 methods have been developed for the identification of *Staphylococcus spp.*: the 16S
87 rRNA [4, 15], *hsp60* [13, 26], *sodA* [37], *rpoB* [9, 31], *femA* [51], *tuf* [18, 30] and *gap*
88 [27, 53, 54] genes have been used as targets. Many studies have demonstrated that
89 genotyping methods are superior to phenotypic methods [18, 28]. However, the
90 sequences of some genes are not sufficiently discriminative to differentiate closely
91 related *Staphylococcus* species, and the databases only include a limited number of
92 species. Previous studies suggest that the *tuf* and *gap* genes constitute the most
93 discriminative targets to differentiate closely related *Staphylococcus* species [12].
94 The *tuf* gene, which encodes the elongation factor (EF-Tu), is involved in peptide
95 chain formation and is a part of the core genome [44]. PCR-based assays targeting
96 the *tuf* gene have been developed for different bacterial genera such as
97 *Enterococcus* [21] *Mycobacterium* [33] and *Staphylococcus* [30]. In the latter case,
98 Martineau *et al.* used hybridization probes (and not DNA sequencing) to differentiate
99 27 species. The *gap* gene encodes a 42-kDa transferrin-binding protein (Tpn) located
100 within the bacterial cell wall that possesses a glycolytic function, converting D-
101 glyceraldehyde-3-phosphate to 1,3-bisphosphoglycerate [34]. Partial sequencing of
102 the *gap* gene has been proposed as an alternative molecular tool for the taxonomic
103 analysis of *Staphylococcus* species [12].

104 We constructed a *tuf* and *gap* gene sequence database of 47 staphylococcal-type
105 strains and evaluated the performance of this database as a molecular identification
106 tool using a 186-strain collection from the French National Reference Center for
107 staphylococci (CNRSta). Finally, to ascertain the rank of this PCR-sequencing
108 approach among the panel of newly developed techniques, the same collection of
109 strains was also tested using MALDI-TOF-MS technology.

110

111 **Materials and Methods**

112 **Bacterial strains.** Type strains representing 47 *Staphylococcus* species and
113 subspecies (Table 1) were used in this study. In addition, 186 strains collected by the
114 CNRSta (Lyon, France) from 1980 to 2008, of both human and animal origin and
115 representing 35 staphylococcal species and subspecies, were included. They were
116 distributed as follows: *S. arlettae* (n = 4), *S. aureus* (n = 9), *S. auricularis* (n = 5), *S.*
117 *capitis* subsp. *capitis* (n = 5), *S. capitis* subsp. *urealyticus* (n = 6), *S. caprae* (n = 6),
118 *S. carnosus* (n = 3), *S. chromogenes* (n = 4), *S. cohnii* subsp. *cohnii* (n = 5), *S. cohnii*
119 subsp. *urealyticus* (n = 5), *S. epidermidis* (n = 7), *S. equorum* (n = 3), *S. felis* (n = 4),
120 *S. gallinarum* (n = 4), *S. haemolyticus* (n = 8), *S. hominis* subsp. *hominis* (n = 6), *S.*
121 *hominis* subsp. *novobiosepticus* (n = 4), *S. hyicus* (n = 1), *intermedius* group with *S.*
122 *delphini* (n = 4), *S. intermedius* (n = 16) *S. pseudintermedius* (n = 4), *S. lentus* (n =
123 2), *S. lugdunensis* (n = 6), *S. pasteurii* (n = 5), *S. pettenkoferi* (n = 1), *S.*
124 *piscifermentans* (n = 3), *S. saprophyticus* (n = 7), *S. schleiferi* subsp. *coagulans* (n =
125 5), *S. schleiferi* subsp. *schleiferi* (n = 8), *S. sciuri* subsp. *carnaticus* (n = 2), *S. sciuri*
126 subsp. *sciuri* (n = 3), *S. sciuri* subsp. *rodentium* (n = 3), *S. simiae* (n = 6), *S. simulans*
127 (n = 6), *S. succinus* (n = 2), *S. warneri* (n = 7), *S. xylosus* (n = 7).

128

129 **CNR identification.** Identification of the above 186 isolates was performed using
130 phenotypic (biochemical characteristics) and genotypic methods. The genus
131 *Staphylococcus* was defined as a Gram-positive cocci with a positive catalase
132 reaction, O/129 compound resistance, bacitracin resistance and nitrofurantoin
133 susceptibility. Coagulase activity on rabbit plasma, heat-stable DNase and the
134 agglutination test (clumping factor, protein A) were used to distinguish *S. aureus* and
135 coagulase-negative Staphylococci (CoNS). In the case of negative coagulase activity
136 or discordant tests, species identification was performed using the ID32 STAPH strip
137 (bioMérieux, Marcy l'Etoile, France). In the case of incorrect identification
138 (unacceptable probability, low confidence factor or no identification), additional tests
139 suggested by the bioMérieux identification system were performed, such as
140 novobiocin susceptibility, oxidase reactions, or the deferoxamine test.

141 When phenotypic tests were not sufficient for the identification of staphylococcal
142 species, molecular methods were used. Sixty-nine strains required supplementary
143 tests for identification, among which were species not included in the ID32 STAPH
144 database (2006). Fifteen strains were identified by amplification of the 16S-23S

145 intergenic spacer regions and the restriction enzyme analysis technique as described
146 by Mendoza *et al.* [32]. In addition, other molecular methods were used, such as
147 PCR sequencing of the partial *sodA* gene [37], ribotyping [39] and DNA-DNA
148 hybridization [43]. A PCR based on the amplification of a *S. pasteurii*-specific random
149 amplified polymorphism DNA (RAPD) fragment was performed to identify *S. pasteurii*
150 species [50]. The same technique (amplification of a specific fragment generated by
151 RAPD) was used to identify *S. capitis* ([50] and unpublished). An *agr*-PCR,
152 described by Jarraud *et al.*, permitted the identification of atypical *S. aureus* strains
153 (*i.e.*, lactose negative, mannitol negative, catalase negative or coagulase negative)
154 [20].

155 Excluding *S. cohnii* subsp. *cohnii* and *S. cohnii* subsp. *urealyticus*, which are
156 discriminated by the ID32 STAPH strip, subspecies were determined using
157 phenotypic or genotypic tests according to the original description of each
158 subspecies: (i) colony pigmentation for *S. capitis* subsp. *capitis* (negative) and *S.*
159 *capitis* subsp. *urealyticus* (positive) [2], (ii) coagulase activity on rabbit plasma for *S.*
160 *schleiferi* subsp. *schleiferi* (negative) and *S. schleiferi* subsp. *coagulans* (positive)
161 [19], (iii) novobiocin susceptibility for *S. hominis* subsp. *hominis* (negative) and *S.*
162 *hominis* subsp. *novobiosepticus* (positive) [24], (iv) nitrate reduction for *S.*
163 *saprophyticus* subsp. *saprophyticus* (negative) and *S. saprophyticus* subsp. *bovis*
164 (positive) [16] and (v) ribotyping methods for the three subspecies of *S. sciuri* [23,
165 29].

166

167 **Bacterial growth and DNA isolation.** Chromosomal DNA from all staphylococcal
168 strains were obtained from overnight cultures grown on horse blood trypticase soy
169 agar plates (bioMérieux, Marcy l'Etoile, France) at 37°C. Colonies were suspended in
170 10 mM/L Tris-HCl buffer, pH 7.0. After centrifugation at 3,450 x *g* for 2 min, the
171 bacterial pellet was resuspended in 100 µL of Tris buffer (10 mM) containing 10 µL of
172 lysostaphin (1 mg/ml) (Sigma, Saint Quentin Fallavier, France), and the mixture was
173 incubated at 37°C for 30 min. DNA purification was completed on the QIAcube
174 apparatus (QIAGEN, Courtaboeuf, France).

175

176 **Bacterial identification by *tuf* and *gap* sequencing.** Based on multiple sequence
177 alignments, a region of the *tuf* gene that is highly conserved among staphylococci
178 was chosen to design the PCR primers. A 660-bp *tuf* DNA fragment was amplified

179 using the primers stat1 (TTA TCA CGT AAC GTT GGT G) and stat2 (CAT TTC WGT
180 ACC TTC TGG). The PCR program consisted of an initial denaturation step at 94°C
181 for 10 min, followed by 30 cycles of denaturation at 94°C for 30 sec, annealing at
182 53°C for 30 sec, and extension at 72°C for 40 sec., and a final extension step for 10
183 min at 72°C. Gap1-for and Gap2-rev were used to amplify a 931-bp fragment of the
184 *gap* gene as previously described [54]. PCR products were electrophoresed in a
185 0.8% agarose gel and visualized with SYBR[®] Safe DNA gel stain (Invitrogen) under
186 UV light to confirm the correct size of the amplified product. Amplicons were
187 sequenced using Genoscreen (Institut Pasteur, Lille, France). Both strands were
188 sequenced with stat1 and stat2 oligonucleotides, or with Gap1-for and Gap2-rev for
189 the reference strains. The coding strand was sequenced only for the other 186
190 strains.

191

192 **Phylogenetic analysis.** Multiple sequence alignments were performed using the
193 ClustalW Program. Phylogenetic trees were generated with the neighbor-joining
194 algorithm [41] applied to synonymous distances (Ks) using the SeaView program
195 [14]. The degree of data support for the tree topology was quantified using the
196 bootstrap method with 500 replications. The *tuf* and *gap* sequences of *Bacillus*
197 *subtilis* were obtained from GenBank (accession no. NC_000964) and used as the
198 outgroup in the phylogenetic analysis. The identification of the 186 strains of the CNR
199 collection was based on their phylogenetic position and their similarities to the
200 reference strain sequences.

201

202 **Nucleotide sequence accession numbers.** The GenBank accession numbers of
203 the staphylococcal *tuf* and *gap* sequences determined in this study are listed in Table
204 1.

205

206 **MALDI-TOF-MS**

207 Staphylococcal strains were sub-cultivated 3 times on Columbia sheep blood agar
208 plates (bioMérieux, Marcy l'Etoile, France) before MALDI-TOF-MS testing. One
209 colony was directly deposited on a MALDI-TOF-MS target plate, and each strain was
210 spotted 4 times. The preparation was overlaid with 1 µl of matrix solution (saturated
211 α-cyano-4-hydroxycinnamic acid). The matrix-sample was crystallized by air-drying at
212 room temperature. Samples were then processed in the MALDI-TOF mass

213 spectrometer Axima Assurance[®] (Shimadzu, Champs sur Marne, France) using an
214 accelerating voltage of 20 kV in linear mode. The spectra were analyzed in the mass-
215 to-charge ratio (m/z) range of 2 000-20 000. Five hundred laser shots were recorded
216 for each spectrum. Quality controls (*i.e.*, duplicate spots of the *Escherichia coli*
217 CCUG 10979 strain) were performed for each target plate. To identify the strains, the
218 spectra obtained for each isolate were compared to the SARAMIS database for
219 January 2009 (AnagnosTec, Potsdam, Germany). This database includes more than
220 2 600 SuperSpectra[™], which can be used for automatic microorganism identification,
221 over 35 000 single spectra, and notably 38 *Staphylococcus* species and subspecies.
222 The results of the matching process are expressed as percentages. Values greater
223 than 80% provide reliable identification based on a SuperSpectra[™]. Values between
224 30 and 80% allow provide identification based on a single spectrum. No identification
225 has been achieved for a score below 30%, as specified by the manufacturer. The
226 SirWeb-MALDI-TOF software (I2A, Perols, France) was used for all experiments to
227 generate the analysis and to export the results to the laboratory informatics system.

228

229

230

231

232 **Results**

233 ***tuf* and *gap* amplification and sequencing.** The utility of amplification–sequencing
234 of the *tuf* and *gap* genes for the identification of staphylococcal species was first
235 determined by analyzing 47 reference strains representing 21 staphylococcal species
236 and 23 subspecies. A partial *tuf* gene sequence (660 bp) was amplified using the
237 primers designed for this study (stat1 and stat2), sequenced and compared. An
238 amplification signal was obtained for all strains tested, and a complete reference
239 database of partial *tuf* gene sequences from the type strains was created for this
240 study. The obtained data were deposited in the GenBank database (accession
241 numbers are presented in Table 1). Similarly, amplification of the partial *gap* gene
242 (931 bp) was performed for the 47 reference strains using the primers described by
243 Yugueros *et al.* [54]. An implemented GenBank database was generated by
244 depositing the 20 missing *gap* sequences that were not deposited by Ghebremedhin
245 [12]. Overall, three species (*S. fleurettii*, *S. vitulinus* and *S. felis*) could not be
246 amplified using the *gap*-specific primers.

247

248 ***Staphylococcus* phylogeny derived from *tuf* and *gap* sequences.** Multiple
249 alignments of the partial *tuf* and *gap* DNA sequences were carried out using the
250 ClustalX[®] software, and phylogenetic trees were constructed by the neighbor-joining
251 method. Bootstrap support values are indicated at the tree nodes (Figs. 1 and 2). The
252 global topology of the *tuf* tree is in agreement with that constructed by *gap* gene
253 analysis. The two trees revealed three common major clusters (bootstrap values >
254 90): (i) the “*sciuri* group” (bootstrap value of 97 with *tuf*, 100 with *gap*) including the 3
255 subspecies of *S. sciuri*, *S. lentus*, *S. vitulinus* and *S. fleurettii*, (ii) the “*intermedius*
256 group” (bootstrap value of 93 with *tuf*, 98 with *gap*) comprising *S. intermedius*, *S.*
257 *delphini* and *S. pseudintermedius*, and (iii) the “*simulans* group” (bootstrap value of
258 91 with *tuf*, 100 with *gap*) including *S. simulans*, *S. piscifermentans*, and *S. carnosus*.
259 *S. epidermidis* and *S. saccharolyticus* formed another major cluster in the *gap* tree
260 with a bootstrap value of 95, whereas these two species were not related in the *tuf*
261 tree. In agreement with other methods, the two trees clustered *S. schleiferi*, *S.*
262 *hyicus*, *S. chromogenes*, *S. muscae*, and *S. lutrae* with the *S. intermedius* group
263 (bootstrap value of 36 with *tuf*, 67 with *gap*), *S. haemolyticus* with *S. lugdunensis* and
264 *S. hominis* and *S. warneri* with *S. pasteurii*, and finally *S. aureus* with *S. simiae*. In
265 both trees, the “*saprophyticus* group” included *S. saprophyticus*, *S. cohnii*, and *S.*

266 *xylosum*, with the addition of *S. succinus* and *S. gallinarum* in the *tuf* tree (Figs. 1 and
267 2). These analyses revealed that the two gene sequences allowed the discrimination
268 of all *Staphylococcus* species, because subspecies of the same species were always
269 clustered together with the exclusion of any other *Staphylococcus* species. Bootstrap
270 values were typically higher for *gap*. Thus, the *gap* gene had a greater discriminatory
271 power than *tuf* for the differentiation of *Staphylococcus* species. However, *tuf*
272 demonstrated greater practicability; a 660-bp amplicon of *tuf* was sufficient for the
273 analysis, versus 900 bp for the *gap* gene. In addition, *tuf* provided a more universal
274 analysis, because it resulted in the amplification of all species, in contrast with *gap*
275 (Table 1). Therefore, the *tuf* gene was selected for further analysis.

276

277 **Species identification of CNRSta laboratory collection strains by *tuf***
278 **sequencing.** The *tuf* gene-based identification matched at the species level 184/186
279 strains obtained from CNRSta (98.9%) (Table 2). Note that for the *S. intermedius*
280 group, identification was considered correct when the *tuf* sequence assigned the
281 identification to the group and not necessarily to the three recently defined species
282 constituting this group: *S. delphini*, *S. pseudintermedius* and *S. intermedius* [42]. For
283 the remaining two strains, one was identified as *S. schleiferi* by CNRSta and as *S.*
284 *warneri* by *tuf* sequencing, with the latter identification confirmed by *gap* sequencing.
285 The second strain identified as *S. warneri* by CNRSta and confirmed to be *S. warneri*
286 by *gap* sequencing could not be identified by *tuf* sequencing for reasons unknown.
287 Similarly to other molecular methods, *tuf* did not discriminate *Staphylococcus*
288 subspecies, except for *S. cohnii* subsp. *cohnii* and *S. cohnii* subsp. *urealyticus*.

289

290 **Comparison of MALDI-TOF-MS and *tuf* sequencing for species identification.**
291 Forty-four of the 47 reference strains together with the 186 strains from the CNRSta
292 collection were analyzed using the MALDI-TOF technology. The two anaerobic
293 strains (*S. aureus* subsp. *anaerobius* and *S. saccharolyticus*) and *S. fleurettii* were
294 not tested. Five species, *S. kloosii*, *S. muscae*, *S. piscifermentans*, *S. simiae* and two
295 subspecies of *S. succinus* not included in the SARAMIS[®] database, provided an
296 incorrect (*S. simiae* identified as *S. aureus*) or no identification (Table 3). Concerning
297 the species or subspecies included in the SARAMIS[®] database, seven reference
298 strains were not identified (*S. auricularis*, *S. caprae*, *S. hyicus*, *S. intermedius*, *S.*
299 *pasteuri*, *S. pettenkoferi* and *S. schleiferi* subsp. *schleiferi*), whereas for some of

300 these species, a correct identification was obtained for several isolates from the
301 CNRSta collection (Table 3). For instance, the reference strain of *S. caprae* was not
302 identified, whereas 5 out of 6 isolates from the CNRSta collection were properly
303 identified by MALDI-TOF-MS. The reference strain of *S. warneri* and 5 out of 7
304 *S. warneri* isolates of the CNRSta collection were identified only at the group
305 *hominis/warneri* level. Overall, 138 out of 186 strains (74.2%) from the CNRSta
306 collection were identified at the species level by MALDI-TOF-MS, and one *S. warneri*
307 strain was identified at the genus level. Four strains were misidentified, and 39 were
308 unidentified. Five *S. warneri* strains were assigned to the group *hominis/warneri*.
309 After exclusion of the CoNS species not included in the database (*i.e.*, *S. simiae*, *S.*
310 *kloosii*, *S. muscae*, *S. piscifermentans*, and *S. succinus*), the final percentage of
311 correct identifications in the CNRSta collection reached 81.5%.

312

313

314

315

316

317 **Discussion**

318 The use of nucleic acid targets provides an alternative technique for the accurate
319 identification of *Staphylococcus* species. Because of the large amount of 16S rDNA
320 sequence data available in public databases, this gene has been the favorite choice
321 in many studies. However due to its lack of discriminatory power, the 16S rDNA
322 sequence is not sufficient for the reliable identification of staphylococcal species [4].
323 Thus, several targets have been exploited to identify species belonging to the
324 *Staphylococcus* genus. Among these, the *sodA* [37], *rpoB* [9, 31], *hsp60* [25], *dnaJ*
325 [46], *gap* [12, 27] and *tuf* gene [30] sequences have been found to be useful for
326 staphylococcal species identification. Major interest in the use of *tuf* results from the
327 small required size of the amplicon (660 bp) together with the ability to use non-
328 degenerate oligonucleotide primers; these two conditions have not been achieved
329 simultaneously by most other targets. The *tuf* gene has thus emerged as a reliable
330 molecular tool for the accurate identification of *Staphylococcus* species [12, 18, 49].
331 However, published studies have been limited to the most common staphylococcal
332 species encountered in human diseases [30]. In the present study, we extended the
333 sequence analysis of the *tuf* gene to a total of 47 species and subspecies. Thus, the
334 present study is the most extensive *tuf*-gene sequence-based study to date on
335 staphylococcal species and sub-species.

336 Considering the phylogeny derived from the *tuf* gene, the global topology of the *tuf*
337 tree, notably the presence of three major clusters, is in agreement with trees
338 constructed based on the analysis of the other genes listed above [9, 12, 26, 30, 37,
339 46, 49]. The strains belonging to the “*sciuri* group” form an identical cluster in all
340 phylogenetic trees derived from 16S rDNA, *rpoB*, *sodA*, *hsp60*, *dnaJ*, *gap* and *tuf*
341 gene sequences. These strains are all novobiocin resistant and oxidase positive.
342 Similarly, the “*intermedius* group” clusters with *S. schleiferi*, *S. hyicus*, *S.*
343 *chromogenes*, *S. muscae* and *S. lutrae* by phylogenetic analysis of the 16S rDNA,
344 *rpoB*, *hsp60*, *dnaJ*, *gap* and *tuf* genes. In contrast, the phylogeny obtained using
345 *sodA* is slightly discordant with those phylogenies, because *S. schleiferi*, *S. hyicus*,
346 *S. muscae*, and *S. chromogenes* do not cluster with the “*intermedius* group” using
347 *sodA* [37]. The third major cluster, the “*simulans* group,” is conserved with *sodA*,
348 *rpoB*, *hsp60*, *dnaJ*, *gap* and *tuf* gene analysis but not with 16S rDNA. In addition to
349 these major clusters, the “*saprophyticus* group” appears to be partially conserved in
350 the *tuf* phylogeny (*S. saprophyticus*, *S. cohnii*, *S. xylosus*, *S. gallinarum*) with a low

351 bootstrap value, in contrast with other gene-derived phylogenies including *S. arlettae*,
352 *S. kloosi*, and *S. equorum* in the “*saprophyticus* group” [9, 25, 37] [12, 30, 46, 49].
353 The other groups with low bootstrap values in the *tuf* phylogeny appeared to be
354 poorly conserved in the phylogenies derived from other genes.

355 The present study resulted in the creation of an almost complete reference database
356 of partial *tuf* gene sequences from type strains. Indeed, the completeness of the
357 database is essential for reliable identification. Prior to the present work, numerous
358 species were either not identified or misidentified when relying on the *tuf* GenBank
359 database; for instance, *S. carnosus* could be misidentified as *S. simulans*, *S.*
360 *gallinarum* as *S. saprophyticus*, *S. lentus* as *S. sciuri*, and *S. piscifermentans* as *S.*
361 *simulans*. However, both the percentage of similarity (below 97%) and the topology
362 of the *tuf*-based phylogenetic tree should demonstrate the lack of robustness of such
363 results.

364 Considering sub-species identification, *tuf* and *gap* sequencing did not allow
365 discrimination at the subspecies level except for *S. cohnii* subsp. *cohnii* and *S. cohnii*
366 subsp. *urealyticus*, as demonstrated for other genes [12, 37]. Thus, molecular
367 methods are clearly not suitable for identification at the subspecies level, a restriction
368 with almost no consequences in clinical practice.

369 Because mass spectrometry is becoming increasingly popular for bacterial
370 identification, we wondered whether it would outcompete *tuf* sequencing in the
371 identification of the 47 species and subspecies of staphylococci. It is noteworthy that
372 no studies have yet explored such a diversity of staphylococcal species. When
373 comparing the *tuf*-based identification with the MALDI-TOF-MS technology, we
374 concluded for an overall superiority of the molecular method even though the MALDI-
375 TOF-MS based method is faster and more cost effective than the molecular method.

376 As expected the MALDI-TOF-MS with 74.2% of correct identification out-competed
377 the ID32 STAPH that identified 62.9% of isolates. When excluding species not
378 included in the databases these percentages were 81.5% versus 75% respectively.

379 The slight inferiority of the MALDI-TOF-MS versus the *tuf*-based method was rather
380 unexpected given the number of enthusiastic reports on the performance of this
381 technology for species identification [6, 10, 11, 45, 47]. Dupont *et al.* analyzed 230
382 isolates of CoNS representing 20 species. They obtained correct identifications for
383 93.2% of the isolates using MALDI-TOF-MS, and this percentage reached 97.4%
384 with exclusion of the species not included in their database [11]. Similarly, Dubois *et*

385 *al.* used the MALDI-TOF-MS Biotyper[®] to identify a collection of 156 strains
386 representing 22 different species and obtained concordant identifications for 99.3%
387 of the species [10]. There are several reasons to explain these apparent
388 discrepancies. First, the SARAMIS database is said to comprise 38 species and
389 subspecies; however, only 15 species or subspecies have a SuperSpectra[®]. It
390 appears that a reliable identification can only be obtained in the latter cases. A similar
391 limitation has been pointed out by Seng *et al.* for the Biotyper database [45]. It is
392 important to note that this drawback has limited consequences in routine clinical
393 practice, because the most frequent species encountered in humans are well
394 represented in both the SARAMIS[®] and Biotyper[®] databases. Hence, correct
395 identification scores as high as 99.3% can be reported for bloodstream isolates in
396 certain studies using MALDI-TOF-MS [47]. Expanding the database to include more
397 species and more strains tested per species would improve the performance of this
398 promising method. Second, the strain collection tested in the present study (the
399 CNRSta collection) not only contains numerous species that are exceptional in
400 clinical practice (but not necessarily never encountered) but also includes isolates of
401 rare species, which were difficult to identify using classical methods and were thus
402 referred to us as a reference laboratory. A third possible reason for the slight
403 inferiority of the MALDI-TOF-MS approach in the present study was that our strain
404 collection contained isolates that had been stored at -20°C for durations ranging from
405 months to several years. This storage period may have altered the phenotypic
406 expression of proteins and thus decreased the performance of the MALDI-TOF-MS
407 approach, which essentially depends on the expression of ribosomal proteins,
408 without affecting the efficiency of the DNA sequencing approach. In conclusion, the
409 *tuf*-based approach appears to be particularly suited for a reference laboratory in
410 which typical and atypical strains of all staphylococcal species are encountered,
411 whereas at present, MALDI-TOF remains more appropriate for routine microbiology
412 practices in clinical laboratories.

413

414 **Acknowledgments**

415 We thank Christine Gardon, Christine Courtier, Céline Spinelli, and Caroline
416 Bouveyron for technical assistance, Christian Curel, Michel Roch and Nader Baïda
417 (I2A company) for generously providing us with the Shimadzu Axima[®] mass
418 spectrometer, the SARAMIS[®] identification database and the SirWeb MALDI-TOF

419 software.

420 **References**

421

- 422 1. Al Masalma M, Raoult D, Roux V (2010) *Staphylococcus massiliensis* sp. nov.,
423 isolated from a human brain abscess. *Int J Syst Evol Microbiol* 60:1066-1072
- 424 2. Bannerman TL, Kloos WE (1991) *Staphylococcus capitis* subsp. *ureolyticus*
425 subsp. nov. from human skin. *Int J Syst Bacteriol* 41:144-147
- 426 3. Bannoehr J, Ben Zakour NL, Waller AS, Guardabassi L, Thoday KL, van den
427 Broek AH, Fitzgerald JR (2007) Population genetic structure of the
428 *Staphylococcus intermedius* group: insights into agr diversification and the
429 emergence of methicillin-resistant strains. *J Bacteriol* 189:8685-8692
- 430 4. Becker K, Harmsen D, Mellmann A, Meier C, Schumann P, Peters G, von Eiff
431 C (2004) Development and evaluation of a quality-controlled ribosomal
432 sequence database for 16S ribosomal DNA-based identification of
433 *Staphylococcus* species. *J Clin Microbiol* 42:4988-4995
- 434 5. Biavasco F, Vignaroli C, Varaldo PE (2000) Glycopeptide resistance in
435 coagulase-negative staphylococci. *Eur J Clin Microbiol Infect Dis* 19:403-417
- 436 6. Carbonnelle E, Beretti JL, Cottyn S, Quesne G, Berche P, Nassif X, Ferroni A
437 (2007) Rapid identification of *Staphylococci* isolated in clinical microbiology
438 laboratories by matrix-assisted laser desorption ionization-time of flight mass
439 spectrometry. *J Clin Microbiol* 45:2156-2161
- 440 7. Corbiere Morot-Bizot S, Leroy S, Talon R (2007) Monitoring of staphylococcal
441 starters in two French processing plants manufacturing dry fermented
442 sausages. *J Appl Microbiol* 102:238-244
- 443 8. Delmas J, Chacornac JP, Robin F, Giammarinaro P, Talon R, Bonnet R
444 (2008) Evaluation of the Vitek 2 system with a variety of *Staphylococcus*
445 species. *J Clin Microbiol* 46:311-313
- 446 9. Drancourt M, Raoult D (2002) *rpoB* gene sequence-based identification of
447 *Staphylococcus* species. *J Clin Microbiol* 40:1333-1338
- 448 10. Dubois D, Leyssene D, Chacornac JP, Kostrzewa M, Schmit PO, Talon R,
449 Bonnet R, Delmas J (2009) Identification of a variety of *Staphylococcus*
450 species by MALDI-TOF mass spectrometry. *J Clin Microbiol*
- 451 11. Dupont C, Sivadon-Tardy V, Bille E, Dauphin B, Beretti JL, Alvarez AS,
452 Degand N, Ferroni A, Rottman M, Herrmann JL, Nassif X, Ronco E,
453 Carbonnelle E (2009) Identification of clinical coagulase-negative
454 staphylococci, isolated in microbiology laboratories, by matrix-assisted laser
455 desorption/ionization-time of flight mass spectrometry and two automated
456 systems. *Clin Microbiol Infect*
- 457 12. Ghebremedhin B, Layer F, Konig W, Konig B (2008) Genetic classification and
458 distinguishing of *Staphylococcus* species based on different partial gap, 16S
459 rRNA, *hsp60*, *rpoB*, *sodA*, and *tuf* gene sequences. *J Clin Microbiol* 46:1019-
460 1025
- 461 13. Goh SH, Potter S, Wood JO, Hemmingsen SM, Reynolds RP, Chow AW
462 (1996) HSP60 gene sequences as universal targets for microbial species
463 identification: studies with coagulase-negative staphylococci. *J Clin Microbiol*
464 34:818-823
- 465 14. Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multiplatform
466 graphical user interface for sequence alignment and phylogenetic tree
467 building. *Mol Biol Evol* 27:221-224

- 468 15. Gribaldo S, Cookson B, Saunders N, Marples R, Stanley J (1997) Rapid
469 identification by specific PCR of coagulase-negative staphylococcal species
470 important in hospital infection. *J Med Microbiol* 46:45-53
- 471 16. Hajek V, Meugnier H, Bes M, Brun Y, Fiedler F, Chmela Z, Lasne Y, Fleurette
472 J, Freney J (1996) *Staphylococcus saprophyticus* subsp. *bovis* subsp. nov.,
473 isolated from bovine nostrils. *Int J Syst Bacteriol* 46:792-796
- 474 17. Hauschild T, Stepanovic S, Zakrzewska-Czerwinska J (2010) *Staphylococcus*
475 *stepanovicii* sp. nov., a novel novobiocin-resistant oxidase-positive
476 staphylococcal species isolated from wild small mammals. *Syst Appl Microbiol*
477 33:183-187
- 478 18. Heikens E, Fleer A, Paauw A, Florijn A, Fluit AC (2005) Comparison of
479 genotypic and phenotypic methods for species-level identification of clinical
480 isolates of coagulase-negative staphylococci. *J Clin Microbiol* 43:2286-2290
- 481 19. Igimi S, Takahashi E, Mitsuoka T (1990) *Staphylococcus schleiferi* subsp.
482 *coagulans* subsp. nov., isolated from the external auditory meatus of dogs with
483 external ear otitis. *Int J Syst Bacteriol* 40:409-411
- 484 20. Jarraud S, Mougel C, Thioulouse J, Lina G, Meugnier H, Forey F, Nesme X,
485 Etienne J, Vandenesch F (2002) Relationships between *Staphylococcus*
486 *aureus* genetic background, virulence factors, agr groups (alleles), and human
487 disease. *Infect Immun* 70:631-641
- 488 21. Ke D, Picard FJ, Martineau F, Menard C, Roy PH, Ouellette M, Bergeron MG
489 (1999) Development of a PCR assay for rapid detection of enterococci. *J Clin*
490 *Microbiol* 37:3497-3503
- 491 22. Kim M, Heo SR, Choi SH, Kwon H, Park JS, Seong MW, Lee DH, Park KU,
492 Song J, Kim EC (2008) Comparison of the MicroScan, VITEK 2, and Crystal
493 GP with 16S rRNA sequencing and MicroSeq 500 v2.0 analysis for coagulase-
494 negative *Staphylococci*. *BMC Microbiol* 8:233
- 495 23. Kloos WE, Ballard DN, Webster JA, Hubner RJ, Tomasz A, Couto I, Sloan GL,
496 Dehart HP, Fiedler F, Schubert K, de Lencastre H, Sanches IS, Heath HE,
497 Leblanc PA, Ljungh A (1997) Ribotype delineation and description of
498 *Staphylococcus sciuri* subspecies and their potential as reservoirs of
499 methicillin resistance and staphylolytic enzyme genes. *Int J Syst Bacteriol*
500 47:313-323
- 501 24. Kloos WE, George CG, Olgiate JS, Van Pelt L, McKinnon ML, Zimmer BL,
502 Muller E, Weinstein MP, Mirrett S (1998) *Staphylococcus hominis* subsp.
503 *novobiosepticus* subsp. nov., a novel trehalose- and N-acetyl-D-glucosamine-
504 negative, novobiocin- and multiple-antibiotic-resistant subspecies isolated
505 from human blood cultures. *Int J Syst Bacteriol* 48 Pt 3:799-812
- 506 25. Kwok AY, Chow AW (2003) Phylogenetic study of *Staphylococcus* and
507 *Macrococcus* species based on partial hsp60 gene sequences. *Int J Syst Evol*
508 *Microbiol* 53:87-92
- 509 26. Kwok AY, Su SC, Reynolds RP, Bay SJ, Av-Gay Y, Dovichi NJ, Chow AW
510 (1999) Species identification and phylogenetic relationships based on partial
511 HSP60 gene sequences within the genus *Staphylococcus*. *Int J Syst Bacteriol*
512 49 Pt 3:1181-1192
- 513 27. Layer F, Ghebremedhin B, Konig W, Konig B (2007) Differentiation of
514 *Staphylococcus* spp. by terminal-restriction fragment length polymorphism
515 analysis of glyceraldehyde-3-phosphate dehydrogenase-encoding gene. *J*
516 *Microbiol Methods* 70:542-549

- 517 28. Layer F, Ghebremedhin B, Moder KA, Konig W, Konig B (2006) Comparative
518 study using various methods for identification of *Staphylococcus* species in
519 clinical specimens. *J Clin Microbiol* 44:2824-2830
- 520 29. Marsou R, Bes M, Boudouma M, Brun Y, Meugnier H, Freney J, Vandenesch
521 F, Etienne J (1999) Distribution of *Staphylococcus sciuri* subspecies among
522 human clinical specimens, and profile of antibiotic resistance. *Res Microbiol*
523 150:531-541
- 524 30. Martineau F, Picard FJ, Ke D, Paradis S, Roy PH, Ouellette M, Bergeron MG
525 (2001) Development of a PCR assay for identification of staphylococci at
526 genus and species levels. *J Clin Microbiol* 39:2541-2547
- 527 31. Mellmann A, Becker K, von Eiff C, Keckevoet U, Schumann P, Harmsen D
528 (2006) Sequencing and staphylococci identification. *Emerg Infect Dis* 12:333-
529 336
- 530 32. Mendoza M, Meugnier H, Bes M, Etienne J, Freney J (1998) Identification of
531 *Staphylococcus* species by 16S-23S rDNA intergenic spacer PCR analysis. *Int*
532 *J Syst Bacteriol* 48 Pt 3:1049-1055
- 533 33. Mignard S, Flandrois JP (2007) Identification of *Mycobacterium* using the EF-
534 Tu encoding (*tuf*) gene and the tmRNA encoding (*ssrA*) gene. *J Med Microbiol*
535 56:1033-1041
- 536 34. Modun BJ, Cockayne A, Finch R, Williams P (1998) The *Staphylococcus*
537 *aureus* and *Staphylococcus epidermidis* transferrin-binding proteins are
538 expressed in vivo during infection. *Microbiology* 144 (Pt 4):1005-1012
- 539 35. Novakova D, Pantucek R, Hubalek Z, Falsen E, Busse HJ, Schumann P,
540 Sedlacek I (2010) *Staphylococcus microti* sp. nov., isolated from the common
541 vole (*Microtus arvalis*). *Int J Syst Evol Microbiol* 60:566-573
- 542 36. Piette A, Verschraegen G (2009) Role of coagulase-negative staphylococci in
543 human disease. *Vet Microbiol* 134:45-54
- 544 37. Poyart C, Quesne G, Boumaila C, Trieu-Cuot P (2001) Rapid and accurate
545 species-level identification of coagulase-negative staphylococci by using the
546 *sodA* gene as a target. *J Clin Microbiol* 39:4296-4301
- 547 38. Probst AJ, Hertel C, Richter L, Wassill L, Ludwig W, Hammes WP (1998)
548 *Staphylococcus condimenti* sp. nov., from soy sauce mash, and
549 *Staphylococcus carnosus* (Schleifer and Fischer 1982) subsp. *utilis* subsp.
550 nov. *Int J Syst Bacteriol* 48 Pt 3:651-658
- 551 39. Regnault B, Grimont F, Grimont PA (1997) Universal ribotyping method using
552 a chemically labelled oligonucleotide probe mixture. *Res Microbiol* 148:649-
553 659
- 554 40. Riesen A, Perreten V (2009) *Staphylococcus rostri* sp. nov., a hemolytic
555 bacterium isolated from the nose of healthy pigs. *Int J Syst Evol Microbiol*
- 556 41. Saitou N, Nei M (1987) The neighbor-joining method: a new method for
557 reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425
- 558 42. Sasaki T, Kikuchi K, Tanaka Y, Takahashi N, Kamata S, Hiramatsu K (2007)
559 Reclassification of phenotypically identified staphylococcus *intermedius*
560 strains. *J Clin Microbiol* 45:2770-2778
- 561 43. Schleifer KH, Meyer SA, Rupprecht M (1979) Relatedness among coagulase-
562 negative staphylococci: deoxyribonucleic acid reassociation and comparative
563 immunological studies. *Arch Microbiol* 122:93-101
- 564 44. Schmitt E, Guillon JM, Meinnel T, Mechulam Y, Dardel F, Blanquet S (1996)
565 Molecular recognition governing the initiation of translation in *Escherichia coli*.
566 A review. *Biochimie* 78:543-554

- 567 45. Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, Raoult
568 D (2009) Ongoing revolution in bacteriology: routine identification of bacteria
569 by matrix-assisted laser desorption ionization time-of-flight mass spectrometry.
570 Clin Infect Dis 49:543-551
- 571 46. Shah MM, Iihara H, Noda M, Song SX, Nhung PH, Ohkusu K, Kawamura Y,
572 Ezaki T (2007) dnaJ gene sequence-based assay for species identification
573 and phylogenetic grouping in the genus Staphylococcus. Int J Syst Evol
574 Microbiol 57:25-30
- 575 47. Spanu T, De Carolis E, Fiori B, Sanguinetti M, D'Inzeo T, Fadda G, Posteraro
576 B (2010) Evaluation of matrix-assisted laser desorption ionization-time-of-flight
577 mass spectrometry in comparison to rpoB gene sequencing for species
578 identification of bloodstream infection staphylococcal isolates. Clin Microbiol
579 Infect
- 580 48. Supre K, De Vlieghe S, Cleenwerck I, Engelbeen K, Van Trappen S, Piepers
581 S, Sampimon OC, Zadoks RN, De Vos P, Haesebrouck F (2010)
582 Staphylococcus devriesei sp. nov., isolated from teat apices and milk of dairy
583 cows. Int J Syst Evol Microbiol
- 584 49. Takahashi T, Satoh I, Kikuchi N (1999) Phylogenetic relationships of 38 taxa
585 of the genus Staphylococcus based on 16S rRNA gene sequence analysis. Int
586 J Syst Bacteriol 49 Pt 2:725-728
- 587 50. Vandenesch F, Perrier-Gros-Claude JD, Bes M, Fuhrmann C, Delorme V,
588 Mouren C, Etienne J (1995) Staphylococcus pasteurii-specific oligonucleotide
589 probes derived from a random amplified DNA fragment. FEMS Microbiol Lett
590 132:147-152
- 591 51. Vannuffel P, Heusterspreute M, Bouyer M, Vandercam B, Philippe M, Gala JL
592 (1999) Molecular characterization of femA from Staphylococcus hominis and
593 Staphylococcus saprophyticus, and femA-based discrimination of
594 staphylococcal species. Res Microbiol 150:129-141
- 595 52. von Eiff C, Peters G, Heilmann C (2002) Pathogenesis of infections due to
596 coagulase-negative staphylococci. Lancet Infect Dis 2:677-685
- 597 53. Yugueros J, Temprano A, Berzal B, Sanchez M, Hernanz C, Luengo JM,
598 Naharro G (2000) Glyceraldehyde-3-phosphate dehydrogenase-encoding
599 gene as a useful taxonomic tool for Staphylococcus spp. J Clin Microbiol
600 38:4351-4355
- 601 54. Yugueros J, Temprano A, Sanchez M, Luengo JM, Naharro G (2001)
602 Identification of Staphylococcus spp. by PCR-restriction fragment length
603 polymorphism of gap gene. J Clin Microbiol 39:3693-3695
604
605
606

607 Table 1: Sources and gene accession numbers of the bacterial reference strains
 608 used in this study

Strain	Source	<i>tuf</i> gene accession number	<i>gap</i> gene accession number
<i>S. aureus</i> subsp. <i>anaerobius</i>	ATCC 35844	HM352930	HM352968
<i>S. aureus</i> subsp. <i>aureus</i>	CCM 885	HM352919	HM352967
<i>S. arlettae</i>	DSM 20672	HM352954	DQ321674 ^a
<i>S. auricularis</i>	ATCC 33753	HM352956	DQ321675 ^a
<i>S. capitis</i> subsp. <i>capitis</i>	CCM 2734	HM352920	DQ321676 ^a
<i>S. capitis</i> subsp. <i>urealyticus</i>	ATCC 49326	HM352921	HM352966
<i>S. caprae</i>	CCM 3573	HM352928	DQ321677 ^a
<i>S. carnosus</i> subsp. <i>carnosus</i>	DSM 20501	HM352953	DQ321678 ^a
<i>S. chromogenes</i>	CCM 3387	HM352952	DQ321680 ^a
<i>S. cohnii</i> subsp. <i>cohnii</i>	CCM 2736	HM352938	DQ321681 ^a
<i>S. cohnii</i> subsp. <i>urealyticus</i>	ATCC 49330	HM352939	HM352971
<i>S. delphini</i>	DSM 20771	HM352940	DQ321682 ^a
<i>S. epidermidis</i>	CCM 2124	HM352922	DQ321683 ^a
<i>S. equorum</i> subsp. <i>equorum</i>	DSM 20674	HM352959	DQ321684 ^a
<i>S. equorum</i> subsp. <i>linens</i>	DSM 15097	HM352965	HM352977
<i>S. felis</i>	ATCC 49168	HM352941	Failed amplification
<i>S. fleurettii</i>	CIP 106114	HM352961	Failed amplification
<i>S. gallinarum</i>	CCM 3572	HM352942	DQ321686 ^a
<i>S. haemolyticus</i>	CCM2737	HM352923	DQ321687 ^a
<i>S. hominis</i> subsp. <i>hominis</i>	DSM 20328	HM352924	DQ321688 ^a
<i>S. hominis</i> subsp. <i>novobiosepticus</i>	ATCC 700236	HM352925	HM352973
<i>S. hyicus</i>	CCM 2368	HM352943	DQ321689 ^a
<i>S. intermedius</i>	CCM 5739	HM352933	DQ321690 ^a
<i>S. kloosii</i>	DSM 20676	HM352951	DQ321691 ^a
<i>S. lentus</i>	ATCC 29070	HM352944	DQ321692 ^a
<i>S. lugdunensis</i>	ATCC 43809	HM352926	DQ321693 ^a
<i>S. lutrae</i>	DSM 10244	HM352945	HM352978
<i>S. muscae</i>	CCM 4175	HM352957	DQ321694 ^a
<i>S. pasteurii</i>	ATCC 51129	HM352929	HM352972
<i>S. pettenkoferi</i>	CIP 107711	HM352963	HM352976
<i>S. piscifermentans</i>	JCM 6057	HM352955	HM352979
<i>S. pseudintermedius</i>	LMG 22219	HM352962	HM352982
<i>S. saccharolyticus</i>	DSM 20359	HM352932	HM352969
<i>S. saprophyticus</i> subsp. <i>bovis</i>	CCM 4410	HM352934	HM352975
<i>S. saprophyticus</i> subsp. <i>saprophyticus</i>	CCM 883	HM352935	DQ321695 ^a
<i>S. schleiferi</i> subsp. <i>schleiferi</i>	ATCC 43808	HM352936	DQ321696 ^a
<i>S. schleiferi</i> subsp. <i>coagulans</i>	JCM 7470	HM352937	HM352980
<i>S. sciuri</i> subsp. <i>carnaticus</i>	ATCC 700058	HM352946	HM352983
<i>S. sciuri</i> subsp. <i>rodentium</i>	ATCC 700061	HM352948	HM352984
<i>S. sciuri</i> subsp. <i>sciuri</i>	ATCC 29062	HM352947	HM352985
<i>S. simiae</i>	CCM 7213	HM352931	HM352970
<i>S. simulans</i>	ATCC 27848	HM352949	DQ321698 ^a
<i>S. succinus</i> subsp. <i>casei</i>	DSM 15096	HM352964	HM352981
<i>S. succinus</i> subsp. <i>succinus</i>	ATCC 700337	HM352958	HM352974
<i>S. vitulinus</i>	ATCC 51145	HM352960	Failed amplification
<i>S. warneri</i>	CCM 2730	HM352927	DQ321699 ^a
<i>S. xylosum</i>	ATCC 29971	HM352950	DQ321700 ^a

609 ATCC: American Type Culture Collection; CCM: Czech Collection of Microorganisms; DSM = DSMZ:
 610 German Collection of Microorganisms and Cell Cultures; JCM: Japan Collection of Microorganisms;
 611 LMG = BCCM/LMG = Belgian Coordinated Collections of Microorganisms; CIP: Institute Pasteur
 612 Collection

613 ^a Sequences deposited by Ghebremedhin *et al.* [12]

614 **Fig. 1** Phylogenetic tree of the reference strains based on *tuf* sequences, computed
615 by the neighbor-joining method applied to synonymous distances (Ks). Bootstrap
616 support percentages $\geq 90\%$ are indicated. The tree was rooted using *Bacillus subtilis*.

617

618

619 **Fig. 2** Phylogenetic tree of reference strains based on *gap* sequences, computed by
620 the neighbor-joining method applied to synonymous distances (Ks). Bootstrap
621 support percentages $\geq 90\%$ are indicated. The tree is rooted at its center.

622 Table 2: *Staphylococcus* species and subspecies identified by *tuf* sequencing versus
 623 CNRSta

Species and subspecies (number of strains)	CNRSta identification		<i>tuf</i> identification at the species level versus CNRSta identification
	ID32 STAPH (number of correct IDs)	Complementary tests (number of tests performed)	
<i>S. arlettae</i> (n = 4)	1	ITS-PCR (2), <i>sodA</i> (1)	4/4
<i>S. aureus</i> (n = 9)	7	Accuprobe (1) <i>agr</i> PCR (1)	9/9
<i>S. auricularis</i> (n = 5)	5		5/5
<i>S. capitis</i> subsp. <i>capitis</i> (n = 5)	4	Specific PCR (1)	5/5
<i>S. capitis</i> subsp. <i>urealyticus</i> (n = 6)	2	Specific PCR (4)	6/6
<i>S. caprae</i> (n = 6)	4	ITS-PCR (1) DNA-DNA hybridization (1)	6/6
<i>S. carnosus</i> (n = 3)	0	ITS-PCR (1) specific probes (2) [38]	3/3
<i>S. chromogenes</i> (n = 4)	2	Pigmentation (2)	4/4
<i>S. cohnii</i> subsp. <i>cohnii</i> (n = 5)	5		5/5
<i>S. cohnii</i> subsp. <i>urealyticus</i> (n = 5)	5		5/5
<i>S. delphini</i> (n = 5)	0	DNA-DNA hybridization (1) sequencing (4) [3]	5/5 ^a
<i>S. epidermidis</i> (n = 7)	6	ITS-PCR (1)	7/7
<i>S. equorum</i> (n = 3)	1	DNA-DNA hybridization (2)	3/3
<i>S. felis</i> (n = 4)	0	DNA-DNA hybridization (4)	4/4
<i>S. gallinarum</i> (n = 4)	4		4/4
<i>S. haemolyticus</i> (n = 8)	7	ITS-PCR (1)	8/8
<i>S. hominis</i> subsp. <i>hominis</i> (n = 6)	5	β-glucuronidase (1)	6/6
<i>S. hominis</i> subsp. <i>novobiosepticus</i> (n = 4)	0	ITS-PCR (4)	4/4
<i>S. hyicus</i> (n = 1)	1		1/1
<i>S. intermedius</i> group (n = 15)	13	ITS-PCR (2)	15/15
<i>S. lentus</i> (n = 2)	2		2/2
<i>S. lugdunensis</i> (n = 6)	6		6/6
<i>S. pasteurii</i> (n = 5)	0	Specific PCR (5) [50]	5/5
<i>S. pettenkoferi</i> (n = 1)	0	DNA-DNA hybridization (1)	1/1
<i>S. piscifermentans</i> (n = 3)	0	ITS-PCR (1) DNA-DNA hybridization (2)	3/3
<i>S. pseudintermedius</i> (n = 4)	0	Sequencing (4) [3]	4/4 ^a
<i>S. saprophyticus</i> subsp. <i>saprophyticus</i> (n = 5)	5		5/5
<i>S. saprophyticus</i> subsp. <i>bovis</i> (n = 2)	0	DNA-DNA hybridization (2)	2/2
<i>S. schleiferi</i> subsp. <i>schleiferi</i> (n = 9)	8		8/9
<i>S. schleiferi</i> subsp. <i>coagulans</i> (n = 5)	0	ITS-PCR (4), DNA-DNA hybridization (1)	5/5
<i>S. scirui</i> subsp. <i>carnaticus</i> (n = 2)	2		2/2
<i>S. sciuri</i> subsp. <i>sciuri</i> (n = 3)	3		3/3
<i>S. sciuri</i> subsp. <i>rodentium</i> (n = 3)	3		3/3
<i>S. simiae</i> (n = 6)	0	DNA-DNA hybridization (6)	6/6
<i>S. simulans</i> (n = 6)	6		6/6
<i>S. succinus</i> (n = 2)	0	<i>sodA</i> PCR (2)	2/2
<i>S. warneri</i> (n = 7)	7		6/7
<i>S. xylosus</i> (n = 7)	7		7/7
	117 (62.9%)	69 (37.1%)	184/186 (98.9%)

624 ^a Assigned to group *intermedius* by *tuf* sequencing

625

626 Table 3. MALDI-TOF-MS identification based on quadruplicate runs of each strain

Species	MALDI identification	
	Reference strains	CNRSta collection strains
<i>S. arlettae</i>	1/1	4/4
<i>S. aureus</i>	1/1	9/9
<i>S. auricularis</i> ^a	0/1	0/5
<i>S. capitis</i> subsp. <i>capitis</i>	1/1	3/5
<i>S. capitis</i> subsp. <i>urealyticus</i>	1/1	3/6
<i>S. caprae</i>	0/1	5/6
<i>S. carnosus</i>	1/1	2/3
<i>S. chromogenes</i>	1/1	3/4
<i>S. cohnii</i> subsp. <i>cohnii</i>	1/1	5/5
<i>S. cohnii</i> subsp. <i>urealyticus</i>	1/1	1/5
<i>S. delphini</i>	1/1	3/4
<i>S. epidermidis</i>	1/1	7/7
<i>S. equorum</i> subsp. <i>equorum</i>	1/1	3/3
<i>S. equorum</i> subsp. <i>linens</i>	1/1	-
<i>S. felis</i>	1/1	2/4
<i>S. gallinarum</i>	1/1	4/4
<i>S. haemolyticus</i>	1/1	8/8
<i>S. hominis</i> subsp. <i>hominis</i>	1/1	6/6
<i>S. hominis</i> subsp. <i>novobiosepticus</i>	1/1	4/4
<i>S. hyicus</i>	0/1	1/1
<i>S. intermedius</i>	0/1	-
<i>S. intermedius</i> group		21/24
<i>S. kloosii</i> ^a	0/1	-
<i>S. lentus</i>	1/1	1/2
<i>S. lugdunensis</i>	1/1	6/6
<i>S. lutrae</i>	1/1	-
<i>S. muscae</i> ^a	0/1	-
<i>S. pasteurii</i>	0/1	3/5
<i>S. pettenkoferi</i>	0/1	0/1
<i>S. piscifermentans</i> ^a	0/1	0/3
<i>S. pseudintermedius</i>	1/1	0/7
<i>S. saprophyticus</i> subsp. <i>bovis</i>	1/1	2/5
<i>S. saprophyticus</i> subsp. <i>saprophyticus</i>	1/1	5/5
<i>S. schleiferi</i> subsp. <i>coagulans</i>	1/1	4/7
<i>S. schleiferi</i> subsp. <i>schleiferi</i>	0/1	4/5
<i>S. sciuri</i> subsp. <i>carnaticus</i>	1/1	3/3
<i>S. sciuri</i> subsp. <i>rodentium</i>	1/1	3/3
<i>S. sciuri</i> subsp. <i>sciuri</i>	1/1	2/3
<i>S. simiae</i> ^a	0/1	4/6 ^b
<i>S. simulans</i>	1/1	6/6
<i>S. succinus</i> subsp. <i>casei</i> ^a	0/1	-
<i>S. succinus</i> subsp. <i>succcinus</i> ^a	0/1	0/2
<i>S. vitulinus</i>	1/1	-
<i>S. warneri</i>	1/1 ^c	5/7 ^c
<i>S. xylosus</i>	1/1	4/7

31/44 (70,5%)

138/186 (74,2%)

627

^aAbsent in the SARAMIS database

628

^bFalse identification: *S. aureus* instead of *S. simiae*

629

^cIdentified as a member of the *hominis/warneri* group

630

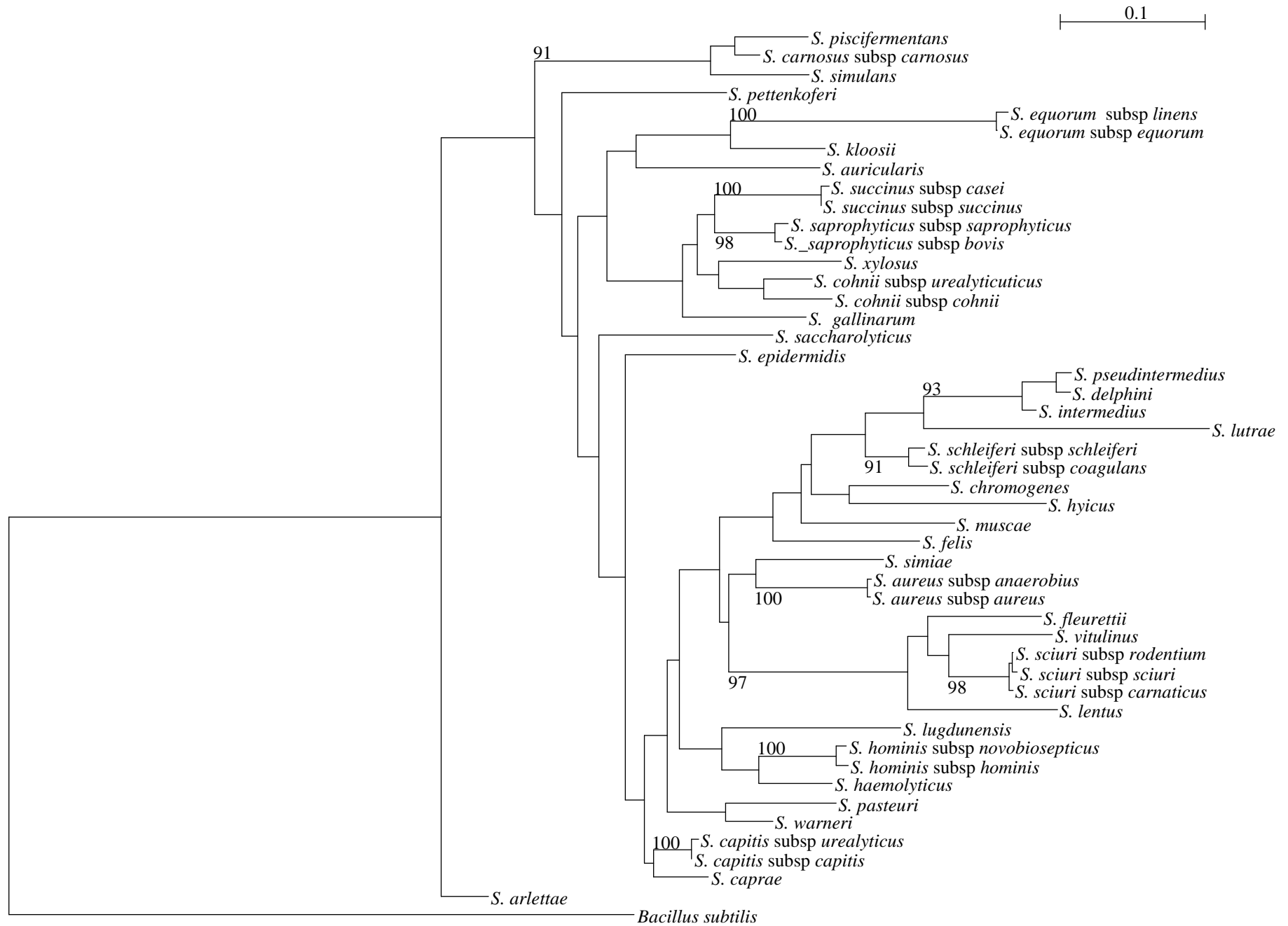


Fig. 1. Phylogenetic tree of the reference strains based on *tuf* sequences, computed by the neighbor-joining method applied to synonymous distances (Ks). Bootstrap support percentages $\geq 90\%$ are indicated. The tree was rooted using *Bacillus subtilis*.

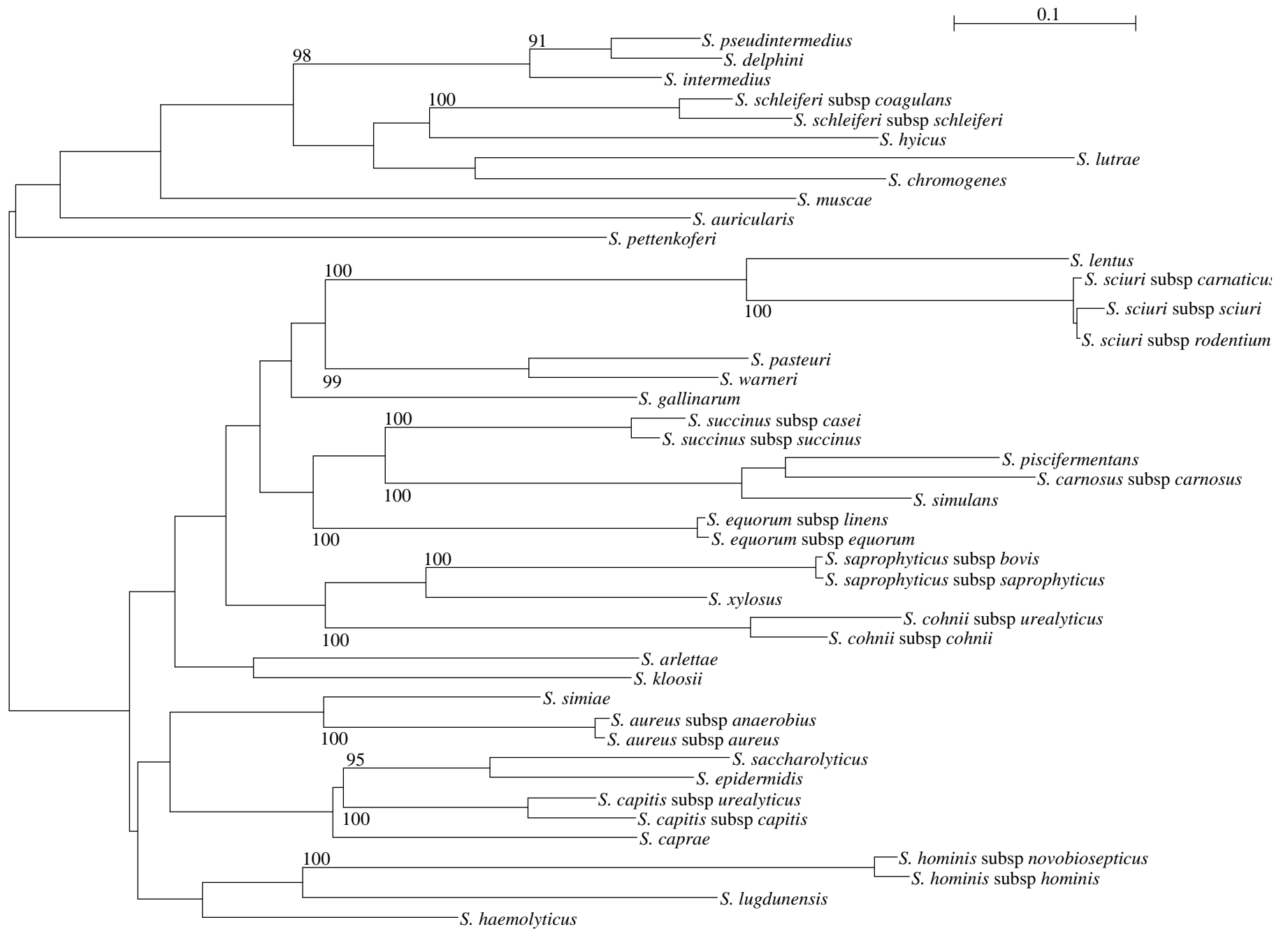


Fig. 2. Phylogenetic tree of reference strains based on *gap* sequences, computed by the neighbor-joining method applied to synonymous distances (Ks). Bootstrap support percentages $\geq 90\%$ are indicated. The tree is rooted at its center.