Staphylococcus borealis sp. nov. -A novel species isolated from skin and blood in humans

Maria Pain[1*], Runa Wolden[1], Daniel Jaén-Luchoro[2,3], Francisco Salvà-Serra[3,4,5,6], Beatriz Piñeiro Iglesias[3,4], Roger Karlsson[3,4], Claus Klingenberg[1,7] Jorunn Pauline Cavanagh[1,7]

[1] Pediatric Infection Group, Department of Clinical Medicine, UiT The Arctic University of Norway, Tromsø, Norway

[2] Department of Infectious Diseases, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[3] Centre for Antibiotic Resistance Research (CARe), University of Gothenburg, Gothenburg, Sweden

[4] Department of Clinical Microbiology, Sahlgrenska University Hospital, Region Västra Götaland, Gothenburg, Sweden

[5] Culture Collection University of Gothenburg (CCUG), Sahlgrenska Academy, University of Gothenburg, Sweden

[6] Microbiology, Department of Biology, University of the Balearic Islands, Palma de Mallorca, Spain

[7] Department of Paediatrics, University Hospital of North Norway, Tromsø, Norway

*Corresponding email: maria.pain@uit.no

Non-standard abbreviations

ANI: Average Nucleotide Identity

CFA-FAME: Cell Fatty Acid-Fatty Acid Methyl Ester

CoNS: Coagulase-Negative Staphylococci

dDDH: digital DNA-DNA Hybridisation

GGDC: Genome-to-Genome Distance Calculator

HSP: High-scoring Segment Pairs

ML: Maximum Likelihood

MSA: Multi Sequence Alignment

OGRI: Overall Genome Related Index

UBCG: Up-to-date Bacterial Core Gene set

The 16S rRNA sequence of *Staphylococcus borealis* 51-48^{T} is available under the accession number MT586030.

The genome sequence data from *Staphylococcus borealis* 51-48^{T} is available under these accessions: BioSample number; SAMN15197055 and assembly accession number; GCA_013345165.1.

The genome sequence data from *Staphylococcus borealis* strains 57-14, 57-74, 58-22 and 58-52 is available under these accessions; BioSample number; SAMN15197056, SAMN15197057, SAMN15197058 and SAMN15197059, and assembly accession number; GCA_013345185.1., GCA_013345175.1., GCA_013345195.1. and GCA_013345205.1, respectively.

Abstract

When analysing a large cohort of *Staphylococcus haemolyticus*, using whole-genome sequencing, five human isolates (four from the skin and one from a blood culture) with aberrant phenotypic and genotypic traits were identified. They were phenotypically similar with yellow colonies, nearly identical 16S rRNA gene sequences and initially speciated as *S. haemolyticus* based on 16S rRNA gene sequence and MALDI-TOF MS. However, compared to *S. haemolyticus,* these five strains demonstrate i) considerable phylogenetic distance with an average nucleotide identity < 95 % and inferred DNA-DNA hybridization < 70 %, ii) a pigmented phenotype, iii) urease production and iv) different fatty acid composition. Based on the phenotypic and genotypic results, we conclude that these strains represent a novel species, for which the name *Staphylococcus borealis* sp. nov. is proposed. The novel species belong to the *Staphylococcus* genus and is coagulase- and oxidase negative and catalase positive. The type strain 51-48^{T} is deposited in the Culture Collection University of Gothenburg (CCUG 73747^{T}) and in the Spanish Type Culture Collection (CECT 30011^{T}).

Introduction

Members of the genus *Staphylococcus*, currently consisting of 54 species and 22 subspecies[1] with validly published names, are most often found on the skin and mucus membranes of mammals and birds (1). Staphylococci, and particularly the coagulase-positive *Staphylococcus aureus*, are a major cause of clinical disease in both humans and animals (2–4). The coagulase-negative staphylococci (CoNS) colonise different niches of the human skin (5) and are part of the commensal human host microbiota. However, over the last decades some CoNS-species such as *Staphylococcus epidermidis, Staphylococcus hominis* and *Staphylococcus haemolyticus* have emerged as important opportunistic pathogens primarily causing disease in patients with foreign body implants or impaired immunity (2).

As part of a previous study analysing a large cohort of *S. haemolyticus* (6), we detected five bacterial strains with aberrant phenotypic and genotypic traits. All five strains originated from the same geographic location, Tromsø, in North Norway. Four strains were isolated from skin swabs from the groin and armpit of healthy volunteers (5), and one strain was isolated from blood culture in 1997 at the University Hospital of North Norway (7). The five strains were all initially identified as *S. haemolyticus* based on 16S rRNA gene sequencing and matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) using a Microflex LT instrument (Bruker Daltonics, Bremen, Germany), Flex Control software and the MALDI Biotyper 3.1 software (Bruker Daltonics). The genome of the blood culture strain was published as *S. haemolyticus* under the accession number GCA_001224225.1 in 2015 (7).

We decided to perform further analyses of these five strains due to the differences observed in both genotypic and phenotypic features compared to other *S. haemolyticus* strains. Based on results from established phenotypic and genotypic methods for species identification (8,9), we propose that these five strains belong to a new staphylococcal species hereafter designated *Staphylococcus borealis* sp. nov.

[1] Based on validly published *Staphylococcus* species from the List of Prokaryotic Names with Standing in Nomenclature (https://lpsn.dsmz.de).

Whole genome sequencing (WGS) and phylogenetic analysis

Genomic DNA from the blood culture strain was isolated according to Chachaty and Saulnier (10) with the addition of RNase A (10 mg/mL; Qiagen, Hilden, Germany) for Illumina sequencing, and the MasterPure™ Gram Positive DNA Purification Kit (Lucigen, Middleton, WI, USA) for isolating genomic DNA for PacBio sequencing. Genomic DNA from the four commensal isolates was isolated by using the Wizard® Genomic DNA kit (Promega, Durham, NC, USA). Whole genome sequencing (WGS) was performed using Illumina Genome Analyzer II (for the blood culture strain) and Illumina MiSeq (for the four commensal strains), as described previously (6,7). Strain 51-48^T was additionally sequenced with a PacBio RS II instrument (Pacific Biosciences, Menlo Park, CA, USA). Assembly of Illumina reads was done using Shovill version 0.8.0 (https://github.com/tseemann/shovill). For the PacBio long reads, consensus sequences were generated and assembled with HGAP v3 (Pacific Biosciences, SMRT Analysis Software v2.3.0) (11). The assembly was polished with Pilon v1.23 (12), using the Illumina sequences (BioSample: SAMEA1035138; SRA: ERS066311) generated previously by Cavanagh *et al.* (7). Mapping of Illumina sequences and the PacBio assembly were done using BWA-MEM (version 0.715-r1140) (13). The resultant draft genome sequences were deposited in GenBank under the BioProject PRJNA638539.

As these strains were initially identified as *S. haemolyticus* based on both 16S rRNA gene sequence similarity and MALDI-TOF, we performed a core-genome phylogeny analysis on all available *S. haemolyticus* genomes deposited in National Center for Biotechnology Information (NCBI) at the time to see whether there were additional isolates related to *S. borealis*. Two draft genomes isolated from cattle in Canada (SNUC119, assembly accession no: GCA_003580835.1 and SNUC 1342, assembly accession no: GCA_003042555.1) (14) clustering with the five proposed *S. borealis* strains were identified. Additionally, a recent publication presented the draft genome of a novel *Staphylococcus* sp. isolated from human skin in Denmark (*Staphylococcus* sp. strain 170179, accession no: GCA_009735325.1); the authors reported that the closest related genome was the clinical *S. borealis* strain (51-48^T, GCA_001224225.1) (15). These three additional draft genomes were included in all the comparative genomic analyses. All eight genomes were annotated with PROKKA (v 1.13) (16) for downstream analysis.

The genome size range was 2,521,961 - 2,797,948 bp, with 2,288 - 2,529 coding sequences (CDSs). The G+C content of the novel species ranged from 33.54% to 33.80% (**Table 1**), which

is in the range of 33-40% expected for species of the genus *Staphylococcus* (1). The G+C content was 0.64-0.9% higher than for the *S. haemolyticus* type strain NCTC 11042^T.

The draft genome of strain 51-48^T was 2,797,948 bp long and had 292.2x depth of sequencing coverage. The draft genome had 33.75 mol% G+C content and contained a total of 2,529 CDS, 22 rRNA[2] genes (8 copies of each 5S rRNA, seven copies of 16S rRNA and seven copies of 23S rRNA), 1 tmRNA and 67 tRNAs.

The 16S rRNA gene sequence of *S. borealis* 51-48^T, 57-14, 57-74, 58-22 and 58-52 was determined by Sanger-sequencing (forward primer; 5'-TACATGCAAGTCGAGCGAAC-3' and reverse primer 5'-ACCTTCCGATACGGCTACCT-3'). The 16S rRNA sequence of SNUC119, SNUC1342 and 170179 was retrieved from the genome assemblies. The 16S rRNA gene sequences from the *S. borealis* strains were analysed using the EzBioCloud online tool 16S-based ID (17). These results showed that the highest similarities were found with *S. haemolyticus* NCTC 11042^T (99.79 % for strain SNUC119, 99.86% for strains 51-48^T, 57-14, 57-74 and SNUC1342; 99.93% for 58-22 and 58-52) and *Staphylococcus petrasii* subsp. *jettensis* (99.5 % for isolates 51-48^T, 57-14, 57-74 and SNUC1342; 99.4 % for 58-22, 58-52 and SNUC119) (**Table 2**). The full length 16S rRNA gene was detected in all strains but SNUC119 (96.7 % of full length). For the Danish strain 170179 the 16S rRNA gene was fragmented with some parts missing, and was omitted from the analysis. In total 16 staphylococcal species and subspecies showed 16S rRNA gene identity > 98.7 %.

The complete 16S rRNA sequences of all staphylococcal type strains were downloaded from the EzBioCloud database (17) and a multi sequence alignment (MSA) was created in MEGA7 (18) using the software MUSCLE (MUltiple Sequence Comparison by Log-Expectation) (19). The maximum likelihood method was used and bootstrapping was set to 500 replicates, using the RAxML software (20). **Figure S1** shows the reconstructed phylogenetic tree generated from the MSA of the 16S rRNA gene sequence of the staphylococci type strains, which accession numbers are listed in **Table S1**.

We also retrieved 16S rRNA sequences from a diverse collection of *S. haemolyticus* (5 representative strains from each of six phylogenetic groups (6)) and compared them to *S.*

[2] *Among the 22 rRNA, one 5S rRNA and one 16S rRNA were partial sequences.*

borealis. We identified intraspecies variations in 16S rRNA among the diverse groups of *S. haemolyticus,* with some strains having identical 16S rRNA sequence with *S. borealis*, emphasising that 16S rRNA cannot be used to distinguish between the two species.

As the 16S rRNA gene was unable to discriminate *S. haemolyticus* from *S. borealis* we investigated whether other single housekeeping genes could distinguish the two species. We analysed the sequence identity between *S. haemolyticus* NCTC 11042^{T} and *S. borealis* 51-48^{T} of the following housekeeping genes; *gap* (21), *tuf* (22), *sodA* (23), *rpoB* (24), *dnaJ* (25), *hsp60* (26), *recA* (27) and *gyrB* (28). Additionally, we looked at intraspecies variations within the eight *S. borealis* strains and within a diverse group of 30 *S. haemolyticus* (the same strains as used in 16S rRNA comparison). Even though only the *hsp60* gene meet the cut-off criteria for different species we believe that the genes *rpoB*, *dnaJ*, *hsp60*, *recA* and *gyrB* can be used to discriminate between *S. haemolyticus* and *S. borealis* as *S. borealis* contains many unique signature bases (found in all eight *S. borealis* strains and not identified in any of the tested *S. haemolyticus*). These results are summarised in **Table 3**.

We also performed multilocus sequence typing (MLST) using the *S. haemolyticus* specific MLST-scheme (29). All *S. borealis* strains were non-typeable following this scheme, and we observed variations ranging between 23 to 79 SNPs for each of the seven genes to the closest allele in the MLST database, including gaps and insertions, further supporting the identification of a new species.

Genome-based phylogeny plays a central role in taxonomy and phylogeny of bacteria, and provides higher resolution than 16S rRNA/single gene phylogeny (30). WGS comparisons were performed according to the recommended minimal standards for description of new staphylococcal species (8). We used the up-to-date bacterial core gene set (UBCG) (30), which produced an alignment based on 92 single-copy core genes extracted from WGS of staphylococcal type strains available in GenBank (accession numbers listed in **Table S1**). From the concatenated gene sequences extracted by UBCG, a maximum likelihood phylogenetic tree was inferred using RAxML (raxmlGUI2 beta) (20) using the GTRGAMMA model and 100 bootstrap replications (31). The results from the phylogenomic tree confirmed that the eight *S. borealis* strains belong to a novel CoNS-species forming their own well-supported branch (**Figure 1**) most closely related to *S. haemolyticus, S. hominis, S. petrasii* and *Staphylococcus devriesei*.

We used the overall genome related index (OGRI) methods to calculate average nucleotide identity (ANI) and tetra-nucleotide analysis using the online tool JSpeciesWS (32). The *in silico* DNA-DNA hybridization (dDDH) was calculated using the genome-to-genome distance calculator (GGDC) version 2.1 (33). The GGDC results were based on the recommended formula 2 (sum of all identities found in high-scoring segment pairs (HSPs), divided by the overall HSP length), which is independent of genome size. Both the ANI and dDDH values between the *S. borealis* strains and related staphylococci were much lower than those considered to represent the same species (9). The results from all OGRI methods confirm that the eight *S. borealis* strains belong to a novel species, which is related to, but distinctly different from *S. haemolyticus.* The OGRI values between the closest related staphylococcal type strains and the *S. borealis* strains are summarised in **Table 2**. The OGRI values between the individual *S. borealis* strains confirmed that these were different strains and that they belonged to the same species (**Table S2 and S3**).

Based on the WGS data we constructed core-genome SNP-based maximum likelihood (ML) tree using the KSNP3 package (34) including 169 *S. haemolyticus* strains, our five *S. borealis* strains and the three draft genomes similar to *S. borealis* found in NCBI. The resultant ML tree clearly demonstrates that the five *S. borealis* strains and the strains included from NCBI, form a distinct cluster separated from *S. haemolyticus* (**Figure 2**).

Based on the OGRI analysis we clearly see a significant difference in the genomes between *S. borealis* and the closest related staphylococcal species. To what extent this represents *S. borealis* specific genes or SNP variability within genes of similar functions the OGRI tools does not answer. Thus, in order to identify genes specific for *S. borealis*, a genome comparison between the eight *S. borealis* strains and 169 whole genome sequences of *S. haemolyticus* strains (6) was performed. We used the pan-genome tool Roary v3.11.2 (35) at default settings but changed the parameter for minimum percentage identity for blastp to 70% in order to identify genes significantly different between the two species. We then extracted the genes found in all eight *S. borealis* which were found in 1% or less of the 169 *S. haemolyticus* strains. We identified 74 *S. borealis* specific genes (**Table S4**), and among these genes we found a urease operon (*ureABCEFGD*). Performing pan-genome analysis with the default minimum percentage for BLAST at 95% produced a *S. borealis* pan-genome of 3,267 genes, of which 1480 were *S. borealis* specific and not shared with any *S. haemolyticus.* The common core

genome between the two species comprised 861 genes, which totals 34.6% of the average *S. borealis* gene content.

All eight *S. borealis* genomes contained capsule-like genes, similar to both the *S. haemolyticus* described capsule (*capA-capG*, and *capK-capM*) and to *S. aureus* capsule genes (c*ap5H-cap5J* and *cap5/8L-cap5/8P*). The presence of capsule-like genes was initially identified by performing a local BLAST of the *S. borealis* genomes against the virulence factor database (VFDB) (36), and was subsequently manually inspected in all eight strains. The six human associated strains contained the same capsule-like operon, while the two Canadian animal associated strains had a slightly different type. Whether these are functional genes, and their role in this species remains to be investigated.

Antibiotic resistance genes were identified using the following databases; Comprehensive Antibiotic Resistance Database (CARD) (37), NCBI AMRFinderPlus (38) and MEGARes (39). The identified antibiotic resistance genes (ARGs) are summarised in **Table 4**. All eight strains contained a *vga(A)* gene variant (98.48 % identity). *Vga(A)* variants confer different levels of resistance to streptogramin A, lincosamides and/or pleuromutilins (40–42). All eight strains also harboured *mgrA*, a global regulator shown to play a role in regulation of virulence factors and contributing to decreased susceptibility to antibiotics like quinolones and β-lactams (43).

Phenotypic tests and metabolic profiling

The phenotype and metabolic profiling were only performed on the five *S. borealis* strains identified in our own collection. Coagulase activity was determined using the Staphaurex™ Plus Latex Agglutination Test (Thermo Scientific ™, Waltham, MA, USA). DNAse activity was tested on DNAse agar with methyl green (Oxoid, Basingstoke; UK). Catalase production was determined by the slide catalase test using hydrogenperoxide, and oxidase activity was determined using the filter paper spot method with 1 % Kovács oxidase reagent. All five *S. borealis* strains were Gram-stain positive cocci growing in clusters. They were non-motile on motility agar. All were oxidase-, DNAse-, coagulase- and clumping factor negative, and catalase positive. All five strains were facultative anaerobic, as determined by using the Brewer thioglycollate medium.

Scanning electron microscopy analyses were performed using a Zeiss Sigma scanning electron microscope (SEM) (Zeiss, Oberkochen, Germany). Whole cells were fixed with 2.5% glutaraldehyde and 4% formaldehyde in PHEM-buffer, before sedimentation onto poly-L-lysin

coated coverslips. Samples were further processed according to the protocol of Cocchiaro using the Pelco Biowave (Ted Pella, Redding, CA, USA)(44). Samples were dried in a Leica EM CPD300 (Leica, Wetzlar, Germany) and mounted on SEM-stubs; gold/palladium was applied with a Polaron Range Sputter Coater (Newhaven, UK). The bacterial cells had a round coccoid shape, appeared in clusters and had a diameter of 650 nm to 1.23 µm.

We tested temperature (4, 15, 30, 37, 42 and 45 ºC) and NaCl tolerance (0, 0.5, 1.5, 3, 5, 7.5, 10 and 15%) according to the protocol by Freney *et al*. (8), using P-agar plates (1). The haemolysis assay was performed on blood agar plates (Oxoid, Basingstoke, UK). *S. haemolyticus* CCUG 7323^{T} was included as a reference strain in all tests. All five *S. borealis* strains were able to grow at 30 to 42 ºC, showed tolerance to NaCl up to 15% and displayed yellow pigmentation on P-agar plates (**Figure 3**). After 24 h of aerobic incubation on horse blood agar at 37°C, the *S. borealis* strains formed smooth, circular, raised or slightly convex colonies reaching 3-5 mm in diameter. A clear β-haemolysis (2 mm) was observed in the *S. borealis* strains and *S. haemolyticus* CCUG 7323^{T} (1.5 mm) on horse blood agar plates.

Metabolic profiling of the five *S. borealis* strains and *S. haemolyticus* CCUG 7323^{T} was performed. The CCUG STX phenotypic worksheet was followed using the API bacterial identification systems API®Staph, API® 32 Staph and API® Coryne test (bioMérieux, Marcy-l'Étoile, France) (www.ccug.se/identification/worksheets), following the instructions of the manufacturer. The metabolic profiles are summarized in **Table 5.** Biochemically, the five *S. borealis* strains differed in three tests when compared with the *S. haemolyticus* CCUG 7323^{T} type strain. Briefly, all *S. borealis* strains were urease positive, where *S. haemolyticus* is considered urease negative. Urease production is one of the main phenotypic factor differentiating *S. haemolyticus* from *Staphylococcus warneri* and *Staphylococcus saprophyticus* (45). All five *S. borealis* strains were positive for fermentation of fructose and mannitol, which is a variable trait in *S. haemolyticus* (46).

Cell fatty acid-fatty acid methyl ester (CFA-FAME) analysis was performed for the five *S. borealis* strains and the resultant profiles were contrasted with the type strain of *S. aureus* (type species of the genus *Staphylococcus*) and two closely related species (*S. haemolyticus* and *S. devriesei*) (**Table 6**). The strains were cultivated on Columbia Blood Agar Base plus 5% defibrinated horse blood, at 37°C, aerobically, overnight (18 to 24 h). An approximate biomass of 100 mg from each strain was harvested in early stationary phase to carry out fatty acid methylester extraction. FAMEs were extracted and washed with alkaline solution after

saponification and methylation of the cell biomass, following the protocol detailed by MIDI (Microbial IDentification Inc, Newark, DE, USA) (47). The CFA-FAME profile was determined, using a HP 5890 gas chromatograph (Hewlett-Packard, Palo Alto, CA, USA) and a standardized protocol similar to the MIDI Sherlock MIS system (47) as described previously (48). CFA-FAME analysis of the five *S. borealis* strains determined the major CFAs to be long chain saturated fatty acids, C15:0 ISO (11%), C15:0 ANTEISO (63%) and C17:0 ANTEISO (13%), while other CFAs observed included C17:0 ISO (5%) and C18:0 (2.5%) (**Table 5**). Preponderance of uneven branched-chain fatty acid pairs with a difference of two carbons, ISO/ANTEISO C15:0 and ISO/ANTEISO C17:0, were present in all samples; of which the ANTEISO fatty acids, C15:0 ANTEISO (mean of approximately 63%) and C17:0 ANTEISO (mean close to 13%) had higher relative proportions than the ISO fatty acids from the same pair, C15:0 ISO (mean near to 11%) and C17:0 ISO (mean of approximately 5%). Only a few straight chain fatty acids were identified, dominated by stearic acid, C18:0 (mean close to 2.5%). As fatty acids are highly preserved in the bacterial membrane, due to their role in the cellular structure, they are useful markers for bacterial differentiation (49). Staphylococci exhibit specific CFA-FAME patterns at genus level (50), although characteristic fatty acids at species level (51). The relative proportion of C15:0 ANTEISO is remarkably higher in *S. borealis* type strain CCUG 73747^T compared to *S. haemolyticus* CCUG 7323^T, showing 61.8% and 47% respectively. The increased amount of C15:0 ANTEISO may regulate membrane fluidity in lower temperatures (52,53).

The determination of peptidoglycan structure was carried out by the German Collection of Microorganisms and Cell Culture GmbH (DSZM) identification service, as described by Schumann (54). The total hydrolysate (100 °C, 4N HCl, 16 h) of the peptidoglycan contained muramic acid (Mur) and the amino acids lysine (Lys), alanine (Ala), serine (Ser), glycine (Gly) and glutamic acid (Glu). Quantification of amino acids by GC/MS of N-heptafluorobutyric amino acid isobutylesters resulted in the following molar ratio: 0.9 Lys : 1.7 Ala : 0.7 Ser : 1.0 Glu : 2.4 Gly : 1.1 Mur. The identity of all amino acids was confirmed by agreement in the gas-chromatographic retention time with those of authentic standards and by characteristic mass spectrometric fragment ions of the derivatives. After hydrolysis under milder conditions (100 °C, 4 N HCl, 0.75 h), the hydrolysate contained (in addition to the amino acids) the peptides Lys-Ala-Ala (backbone isomer), Lys-Ala (backbone isomer), Mur-Ala, Ala-Glu, Ala-Lys-Gly, Ala-Glu-Lys-Gly, Gly-Gly and Gly-Ser but no Gly-Gly-Gly or Lys-Ser peptide. From these data it was concluded that the strain *S. borealis* 51-48^T (CCUG 737547^T) displayed the

peptidoglycan type A3α L-Lys–Gly–Gly–L-Ser–Gly (type A11.3, www.peptidoglycan-types.info).

Antimicrobial resistance testing was performed using the disc diffusion method and microbroth dilution test according to the EUCAST guidelines (55). Briefly, a 0.5 McFarland bacterial cell suspension was inoculated on Mueller Hinton agar plates (Oxoid, Basingstoke; UK). Discs or MIC gradient strips were place on the inoculated agar plates and were incubated at 35 ±1 ºC for 16-18 hours, and zones of inhibition were measured. The five *S. borealis* strains were susceptible to the antimicrobial agents cefoxitin (30 µg), ciprofloxacin (5 µg), clindamycin (2 µg), gentamicin (10 µg), linezolid (10 µg), rifampicin, tetracycline (30 µg), trimethoprim-sulfamethoxazole (1.25-23.75 µg), vancomycin (0.015-256 µg), clindamycin (0.125-256 µg) and lincomycin (0.125-256 µg). 51-48^{T} and 58-52 were resistant to erythromycin (both of which harboured the *ermC* gene) while 58-52 was also resistant to fusidic acid (harbouring the *fusC* gene). All strains showed resistance to the pleuromutilin antibiotic tiamulin according to the MIC breakpoints given by Frey et al., (56), which could be conferred by the *vga(A)* gene.

In conclusion, although the eight *S. borealis* strains share near identical 16S rRNA gene sequences to *S. haemolyticus* NCTC 11042^{T}, and are phylogenetically closely related to *S. haemolyticus,* there are strong phenotypic and genomic justifications for assigning the strains to a novel species of the genus *Staphylococcus,* for which the name *Staphylococcus borealis* sp. nov. is proposed.

These justifications are:

1) Phylogenetic distance, ANI < 95% and inferred DDH < 70 %.
2) Genome comparisons
3) Pigmented phenotype
4) Production of urease
5) Different cell fatty acid composition

Description of *Staphylococcus borealis* sp. nov.

Staphylococcus borealis (bo.re.a'lis. L. masc. adj. *borealis* related to the North, boreal)

Colonies are 3-5 mm in diameter, round, smooth and have a yellow tint. The difference in pigmentation between typical *S. haemolyticus* and *S. borealis* is particularly evident on different supplemented P-agars (non-supplemented, full fat milk and horse blood) after 48 h at 37 ºC.

Cells are Gram-stain positive, coccoid, 650 nm to 1.23 µm in diameter and form clusters. They are facultative anaerobic and coagulase negative and catalase positive. They are biochemically negative for fructose, mannitol and positive for production of urease. The major fatty acids are branched fatty acid pairs C15:0 ISO, C15:0 ANTEISO and C17:0 ANTEISO, C17:0 ISO while the straight chain fatty acid C18:0 is present in much lower amount. The peptidoglycan type is type A11.3.

The five *S. borealis* strains are deposited to the Culture Collection University of Gothenburg (CCUG), with the following numbers: type strain, 51-48^T = CCUG 73747^T, 57-14 = CCUG 73748, 57-74 = CCUG 73749, 58-22 = CCUG 73750 and 58-52 = CCUG 73751. Two strains are deposited to the Spanish Type Culture Collection (CECT) with the following identifiers: 51-48^T = CECT 30011^T and 57-14 = CECT 30010.

Author Statements

Funding Information

This study was funded by UiT - The Arctic University of Norway, and funds from the Northern Norway Regional Health Authority (HNF1344-17).

Acknowledgements

The Culture Collection of the University of Gothenburg (CCUG) and the staff are acknowledged for providing reference strains and expert characterization analyses. The CCUG is supported by the Department of Clinical Microbiology, Sahlgrenska University Hospital.

Conflicts of Interest

The authors declare no competing interests

References

1. Götz F, Bannerman T, Schleifer K. The genera Staphylococcus and Macrococcus. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K, Stackebrandt E, editors. The Prokayotes, vol 4: Bacteria: Firmicutes, Cyanobacteria. New York, NY: Springer US; 2006. p. 5–75.
2. Becker K, Heilmann C, Peters G. Coagulase-negative staphylococci. Clin Microbiol Rev. 2014;27(4):870–926.
3. Tong SYC, Davis JS, Eichenberger E, Holland TL, Fowler VG. Staphylococcus aureus infections: Epidemiology, pathophysiology, clinical manifestations, and management. Clin Microbiol Rev. 2015;28(3):603–61.
4. Fluit AC. Livestock-associated Staphylococcus aureus. Clin Microbiol Infect [Internet]. 2012;18(8):735–44. Available from: http://dx.doi.org/10.1111/j.1469-0691.2012.03846.x
5. Cavanagh JP, Wolden R, Heise P, Esaiassen E, Klingenberg C, Aarag Fredheim EG. Antimicrobial susceptibility and body site distribution of community isolates of coagulase-negative staphylococci. Apmis. 2016;124:973–8.
6. Pain M, Hjerde E, Klingenberg C, Cavanagh JP. Comparative genomic analysis of Staphylococcus haemolyticus reveals key to hospital adaptation and pathogenicity. Front Microbiol. 2019;10(September):1–13.
7. Cavanagh JP, Hjerde E, Holden MTG, Kahlke T, Klingenberg C, Flægstad T, et al. Whole-genome sequencing reveals clonal expansion of multiresistant Staphylococcus haemolyticus in European hospitals. J Antimicrob Chemother. 2014;69:2920–7.
8. Freney J, Kloos WE, Hajek V, Webster JA, Bes M, Brun Y, et al. Recommended minimal standards for description of new staphylococcal species. Int J Syst Bacteriol. 2009;49(2):489–502.
9. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, et al. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol Microbiol. 2018;68(1):461–6.
10. Chachaty E, Saulnier P. Isolating chromosomal DNA from. In: Rapley R, editor. The Nucleic Acid Protocols Handbook. New Jersey: Humana Press; 2000.
11. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods. 2013;10(6):563–9.
12. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: An

integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One. 2014;9(11).

13. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint [Internet]. 2013;00(00):1–3. Available from: http://arxiv.org/abs/1303.3997
14. Naushad S, Barkema HW, Luby C, Condas LAZ, Nobrega DB, Carson DA, et al. Comprehensive phylogenetic analysis of bovine non-aureus staphylococci species based on whole-genome sequencing. Front Microbiol. 2016;7(DEC).
15. Stødkilde K, Poehlein A, Brüggemann H. Draft genome sequence of a new staphylococcal species isolated from human skin. Microbiol Resour Announc. 2020;9:e01499-19.
16. Seemann T. Prokka: Rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068–9.
17. Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, et al. Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol. 2017;67(5):1613–7.
18. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33(7):1870–4.
19. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7.
20. Edler D, Klein J, Antonelli A, Silvestro D. raxmlGUI 2.0 beta: a graphical interface and toolkit for phylogenetic analyses using RAxML. bioRxiv. 2019;(October):800912.
21. Yugueros J, Temprano A, Berzal B, Sánchez M, Hernanz C, Luengo JM, et al. Glyceraldehyde-3-phosphate dehydrogenase-encoding gene as a useful taxonomic tool for Staphylococcus spp. J Clin Microbiol. 2000;38(12):4351–5.
22. Martineau F, Picard FJ, Ke D, Paradis S, Roy PH, Ouellette M, et al. Development of a PCR Assay for Identification of Staphylococci.pdf. 2001;39(7):2541–7.
23. Poyart C, Quesne G, Boumaila C, Trieu-Cuot P. Rapid and accurate species-level identification of coagulase-negative staphylococci by using the sodA gene as a target. J Clin Microbiol. 2001;39(12):4296–301.
24. Drancourt M, Raoult D. rpoB gene sequence-based identification of Staphylococcus species. J Clin Microbiol. 2002;40(4):1333–8.
25. Shah MM, Iihara H, Noda M, Song SX, Nhung PH, Ohkusu K, et al. dnaJ gene sequence-based assay for species identification and phylogenetic grouping in the genus

Staphylococcus. Int J Syst Evol Microbiol. 2007;57(1):25–30.

26. Goh SH, Potter S, Wood JO, Hemmingsen SM, Reynolds RP, Chow AW. HSP60 Gene Sequences as Universal Targets for Microbial Species Identification. J Clin Microbiol [Internet]. 1996;34(4):818–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8815090
27. Landeta G, Reverón I, Carrascosa A V., Rivas B de las, Muñoz R. Use of recA gene sequence analysis for the identification of Staphylococcus equorum strains predominant on dry-cured hams. Food Microbiol [Internet]. 2011;28(6):1205–10. Available from: http://dx.doi.org/10.1016/j.fm.2011.04.006
28. Poirier S, Rué O, Peguilhan R, Coeuret G, Zagorec M, Champomier-Vergès MC, et al. Deciphering intra-species bacterial diversity of meat and seafood spoilage microbiota using gyrB amplicon sequencing: A comparative analysis with 16S rDNA V3-V4 amplicon sequencing. PLoS One. 2018;13(9):1–26.
29. Cavanagh JP, Klingenberg C, Hanssen AM, Fredheim EA, Francois P, Schrenzel J, et al. Core genome conservation of Staphylococcus haemolyticus limits sequence based population structure analysis. J Microbiol Methods [Internet]. 2012;89(3):159–66. Available from: http://dx.doi.org/10.1016/j.mimet.2012.03.014
30. Na SI, Kim YO, Yoon SH, Ha S min, Baek I, Chun J. UBCG: Up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. J Microbiol. 2018;56(4):281–5.
31. Felsenstein J. Confidence Limits on Phylogenies : An Approach Using the Bootstrap. Evolution (N Y). 1985;39(4):783–91.
32. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics. 2015;32(6):929–31.
33. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics. 2013;14.
34. Gardner SN, Slezak T, Hall BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. Bioinformatics. 2015;31(17):2877–8.
35. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691–3.

36. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: A comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. 2019;47(D1):D687–92.
37. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, et al. CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res. 2017;45:D566–73.
38. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, et al. Validating the AMRFINder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob Agents Chemother. 2019;63(11):1–20.
39. Doster E, Lakin SM, Dean CJ, Wolfe C, Young JG, Boucher C, et al. MEGARes 2.0: a database for classification of antimicrobial drug, biocide and metal resistance determinants in metagenomic sequence data. Nucleic Acids Res. 2020;48(D1):D561–9.
40. Novotna G, Janata J. A new evolutionary variant of the streptogramin A resistance protein, Vga(A)LC, from Staphylococcus haemolyticus with shifted substrate specificity towards lincosamides. Antimicrob Agents Chemother. 2006;50(12):4070–6.
41. Tessé S, Trueba F, Berthet N, Hot C, Chesneau O. Resistance genes underlying the LSA phenotype of staphylococcal isolates from France. Antimicrob Agents Chemother. 2013;57(9):4543–6.
42. Gentry DR, McCloskey L, Gwynn MN, Rittenhouse SF, Scangarella N, Shawar R, et al. Genetic characterization of Vga ABC proteins conferring reduced susceptibility to pleuromutilins in Staphylococcus aureus. Antimicrob Agents Chemother. 2008;52(12):4507–9.
43. Truong-Bolduc QC, Hooper DC. The transcriptional regulators NorG and MgrA modulate resistance to both quinolones and β-lactams in Staphylococcus aureus. J Bacteriol. 2007;189(8):2996–3005.
44. Cocchiaro JL, Kumar Y, Fischer ER, Hackstadt T, Valdivia RH. Cytoplasmic lipid droplets are translocated into the lumen of the Chlamydia trachomatis parasitophorous vacuole. Proc Natl Acad Sci U S A. 2008;105(27):9379–84.
45. Cunha MDLRS, Sinzato YK, Silveira LVA. Comparison of methods for the identification of coagulase-negative staphylococci. Mem Inst Oswaldo Cruz. 2004;99(8):855–60.
46. Schleifer KH, Kloos WE. Isolation and Characterization of Staphylococci from Human Skin I. Amended Descriptions of Staphylococcus epidermidis and Staphylococcus saprophyticus and Descriptions of Three New Species: Staphylococcus cohnii,

Staphylococcus haemolyticus, and Staphyloc. Int J Syst Bacteriol. 1975;25(1):50–61.
47. Sasser M. Identification of bacteria by gas chromatography of cellular fatty acids. MIDI - Tech note #101 [Internet]. 1990;(February):1–6. Available from: http://natasha.eng.usf.edu/gilbert/courses/Biotransport Phenomena/pdf/bacteria_gc_1.pdf
48. Zamora L, Fernández-Garayzábal JF, Svensson-Stadler LA, Palacios MA, Domínguez L, Moore ERB, et al. Flavobacterium oncorhynchi sp. nov., a new species isolated from rainbow trout (Oncorhynchus mykiss). Syst Appl Microbiol [Internet]. 2012;35(2):86–91. Available from: http://dx.doi.org/10.1016/j.syapm.2011.11.007
49. De Carvalho CCCR, Caramujo MJ. The various roles of fatty acids. Molecules. 2018;23(10).
50. Welch DF. Applications of cellular fatty acid analysis. Clin Microbiol Rev. 1991;4(4):422–38.
51. Kotilainen P, Huovinen P, Eerola E. Application of gas-liquid chromatographic analysis of cellular fatty acids for species identification and typing of coagulase-negative staphylococci. J Clin Microbiol. 1991;29(2):315–22.
52. Kaneda T. Iso- and anteiso-fatty acids in bacteria: Biosynthesis, function, and taxonomic significance. Microbiol Rev. 1991;55(2):288–302.
53. Onyango LA, Alreshidi MM. Adaptive Metabolism in Staphylococci: Survival and Persistence in Environmental and Clinical Settings. J Pathog. 2018;2018(Cm):1–11.
54. Schumann P. Peptidoglycan structure. Methods Microbiol. 2011;38:101–29.
55. The European Committee on Antimicrobial Suceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters, version 9.0, 2019.
56. Frey Y, Rodriguez JP, Thomann A, Schwendener S, Perreten V. Genetic characterization of antimicrobial resistance in coagulase-negative staphylococci from bovine mastitis milk. J Dairy Sci. 2013;96(4):2247–57.

Figure legends

Figure 1. Phylogenetic relationship of Staphylococcal type strains and the eight *S. borealis* strains based on core genes. The maximum likelihood method was used and bootstrapping was set to 100 replicates, using the RAxML software. *Macrococcus caseolyticus* was used for rooting the tree.

Figure 2. SNP-based core-genome phylogenetic tree using the kSNP3 suite, of 169 *S. haemolyticus* strains and the eight *S. borealis* strains.

Figure 3. Yellow pigmentation of the five *S. borealis* isolates from this study, in comparison to *S. haemolyticus* CCUG 7323^{T} shown on P-agar.

Table 1. Overview of genomic information for all eight *S. borealis* strains and *S. haemolyticus* type strain.

Isolate ID	Genome size	Contigs	CDS	N50	GC%	Coverage	Accession
51-48^T **(=CCUG 73747^T = CECT 30011^T)**	2,797,948 bp	5	2,529	2,689,815	33.75%	292.2x	GCA_013345165.1
57-14 **(=CCUG 73748 = CECT 30010)**	2,626,230 bp	36	2,403	645,817	33.66%	337.7x	GCA_013345185.1
57-74 **(=CCUG 73749)**	2,615,713 bp	41	2,398	390,616	33.66%	463.8x	GCA_013345175.1
58-22 **(=CCUG 73750)**	2,666,192 bp	40	2,475	391,465	33.69%	319.4x	GCA_013345195.1
58-52 **(=CCUG 73751)**	2,664,706 bp	30	2,420	805,534	33.54%	387.0x	GCA_013345205.1
Staphylococcus sp. 170179	2,629,435 bp	48	2,324	212,499	33.58%	334.0x	GCA_009735325.1
SNUC119	2,521,961 bp	166	2,288	42,538	33.80%	50x	GCA_003580835.1
SNUC1342	2,522,218 bp	99	2,290	119,419	33.80%	93x	GCA_003042555.1
S. haemolyticus NCTC 11042^T	2,569,468 bp	4	2,323	2,515,409	32.90%	100x	GCA_900458595.1

Table 2. Overview of results for species identity of the closest related *Staphylococcus* species and subspecies, compared to the proposed type strain of *S. borealis* 51-48^T. The numbers in brackets are the threshold values for species delineation.

Staphylococcal type strains	16S rRNA (98.7%)	ANIb (<95%)	ANIm (<95%)	Tetra (<0.989)	dDDH (<70 %)
***S. devriesei* NCTC 13828^T**	99.25	79.82	84.92	0.95612	23.8
S. petrasii subsp. ***petrasii* CCM8418^T**	99.39	80.10	85.49	0.96349	23.3
S. petrasii subsp ***croceolyticus* CCM8421^T**	99.39	80.35	85.72	0.96578	23.5
S. petrasii subsp***. jettensis* SEQ110^T**	99.51	80.28	85.65	0.96301	23.6
S. petrasii subsp***. pragensis* NRL/St 12/356^T**	99.46	80.55	85.74	0.96981	23.6
S. hominis subsp***. hominis* DSM 20328^T**	99.25	78.54	85.11	0.95310	22.6
S. hominis subsp***. novobiosepticus* GTC 1228^T**	98.83	78.67	85.38	0.95569	23.0
***S. haemolyticus* NCTC 11042^T**	99.86	87.40	88.66	0.98571	34.2
***S. haemolyticus* SNUC 119**	99.93	97.67	98.14	0.99807	82.1
***S. haemolyticus* SNUC 1342**	100	97.65	98.18	0.99727	82.4
***Staphylococcus sp.* 170179**	-	99.54	99.80	0.99903	98.0
S. borealis 57-14	100	99.58	99.82	0.99895	97.2
S. borealis 57-74	100	99.56	99.78	0.99879	97.4
S. borealis 58-22	99.93	99.56	99.83	0.99910	98.0
S. borealis 58-52	99.93	99.74	99.81	0.99905	98.1

Table 3: Percent identity between housekeeping genes of *S. borealis* 51-48^T and *S. haemolyticus* NCTC 11042^T and the intraspecies variations for each gene within each species. Intraspecies variation within *S. haemolyticus* was based on representatives from each phylogenetic group of a diverse collection of *S. haemolyticus* (6). For the *S. borealis* strains of human origin (51-48, 57-14, 57-74, 58-22, 58-52 and 170179) all housekeeping genes with the exception of 16S rRNA were identical. Within all housekeeping genes (except 16S rRNA) *S. borealis* specific conserved bases/ were observed (specific bases/SNPs found in all *S. borealis* and in no *S. haemolyticus)*.

	Housekeeping genes (identity cut-off)								
	tuf (98%) (22)	***gap*** (96 %) (21)	***sodA*** (97 %) (23)	***rpoB*** (93.6 %) (24)	***hsp60*** (93 %) (26)	***dnaJ*** (88.8 %) (25)	***recA*** (27)	***gyrB*** (28)	**16S** (98.7 %)
S. haemolyticus NCTC 11042^T versus *S. borealis* 51-48^T	99.2 %	99.4 %	97.2 %	96.1 %	91.4 %	93.1 %	91.2 %	94.7 %	99.86 %
S. borealis intraspecies variation Conserved SNPs only in *S. borealis*	100 % 4	99.7-100 % 9	99.8 -100 % 12	99.3-100 % 113	99.1-100 % 117	99.4-100 % 61	99-100 % 70	99.5-100 % 83	99.93-100% 0
S. haemolyticus intraspecies variation	99-100 %	99.3.100%	98.2-100 %	98.9-100 %	98.3-100 %	97.7-100 %	97.0-100 %	98.7-100%	99.67-100%

Table 4: Antibiotic resistance genes[3] identified in the different *S. borealis* strains, based on the antibiotic databases CARD, MegaRes and NCBI. For each resistance gene the percentage identity with the genes identified in the *S. borealis* strain is presented. For numbers marked in bold the resistance phenotype was also confirmed.

	ANT4	***ble***	***ermC***	***fusC***	***mgrA***	***tetK***	***qacC***	***vga(A)***
51-48^T			**100%**		93%			98.5%
57-14					93%			98.5%
57-74					93%			98.5%
58-22					93%			98.5%
58-52			**100%**	**100%**	93%		100%	98.5%
170179	100%	100%		100%	93%			98.5%
SNUC119					93%	100%		98.5%
SNUC1342					93%			98.5%

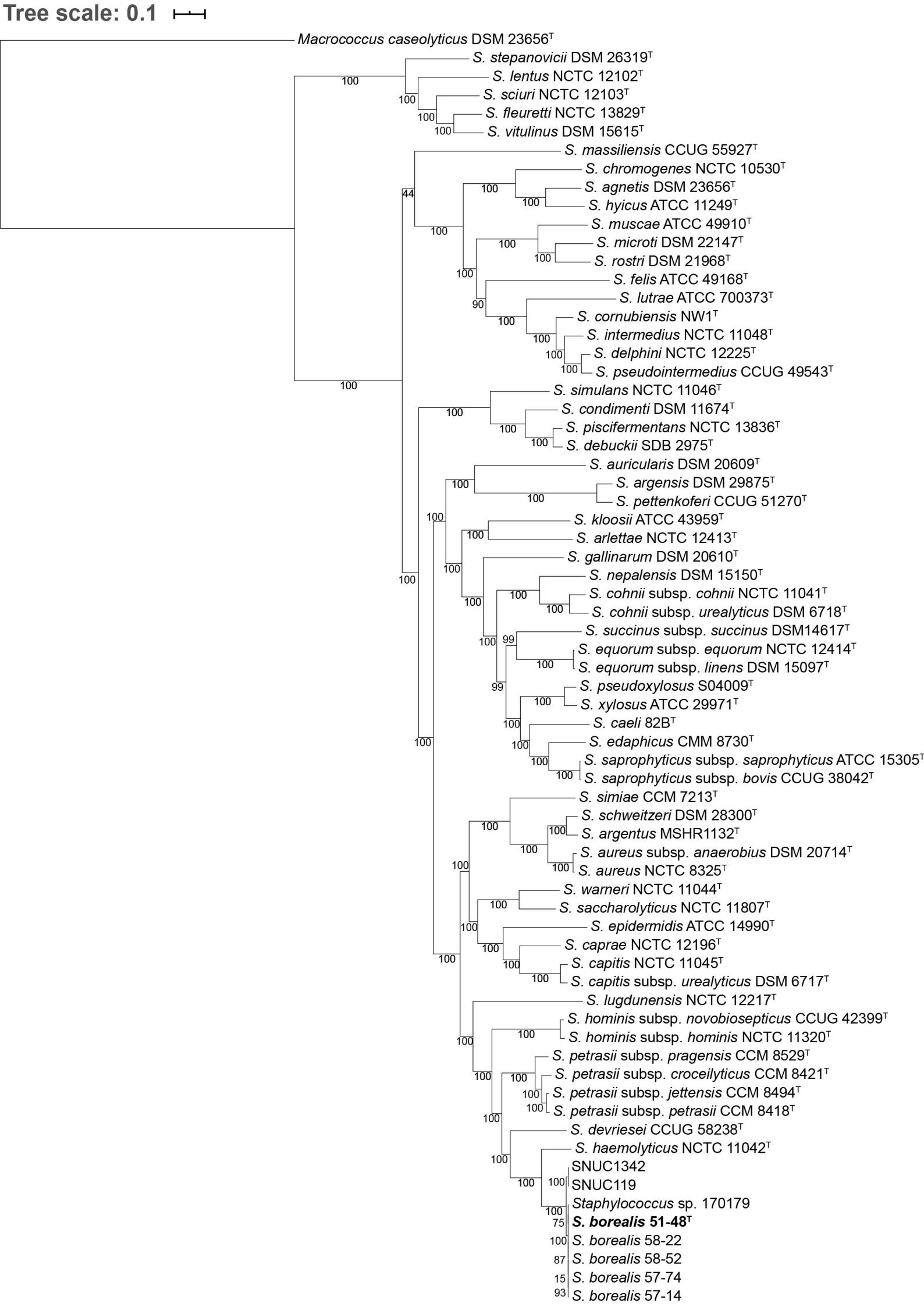
[3] The resistance genes listed in the table can confer resistance to the following antimicrobials; *ANT4* – aminoglycoside, *ble* – bleomycin, *ermC* – erythromycin, *fusC* – fusidic acid, *mgrA* – global regulator (B-lactams and quinolones), *qacC* – quaternary ammonium compounds, *vga(A)* - streptogramin A, lincosamides and/or pleuromutilins.

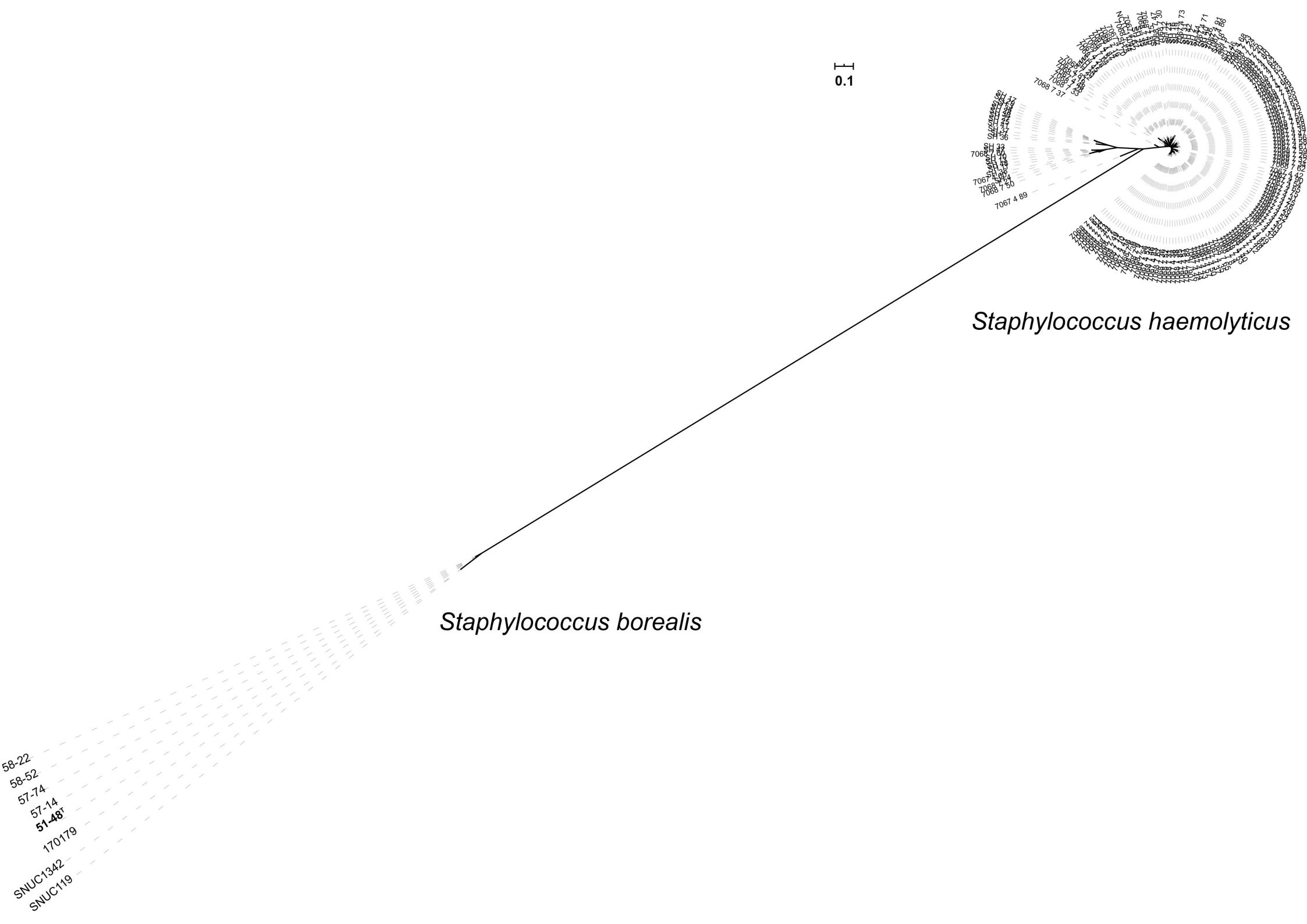
Table 5. Biochemical tests, based on API®Staph, API® 32 Staph and API® Coryne, of the five *S. borealis* isolates and the type strain of *S. haemolyticus*

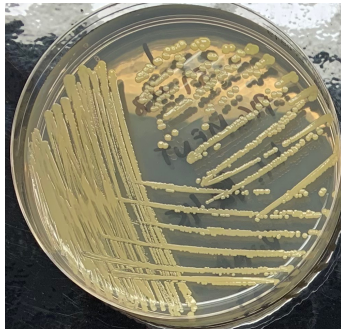
		S. borealis					*S. haemolyticus*
Culture Collection University of Gothenburg (CCUG)		**73747^T**	**73748**	**73749**	**73750**	**73751**	**7323^T**
Local strain identification number		**51-48**	**57-14**	**57-74**	**58-22**	**58-52**	**63-42**
Glucose	**GLU**	+	+	+	+	+	+
Fructose	**FRU**	+	+	+	+	+	-
Arabinose	**ARA**	-	-	-	-	-	-
Ribose	**RIB**	+	+	+	+	+	+
Mannose	**MNE**	-	-	-	-	+	-
Xylose	**XYL**	-	-	-	-	-	-
Sucrose	**SAC**	+	+	+	+	+	+
Lactose	**LAC**	-	-	-	-	-	-
Turanose	**TUR**	-	+	+	-	-	-
Cellobiose	**CEL**	-	-	-	-	-	-
Maltose	**MAL**	+	+	+	+	+	+
Trehalose	**TRE**	+	+	+	+	+	+
Melibiose	**MEL**	-	-	-	-	-	-
Raffinose	**RAF**	-	-	-	-	-	-
Glycogen	**GLYG**	-	-	-	-	-	-
N-acetyl-glucosamine	**NAG**	+	-	+	+	-	+
MethyL-α-D-glucopiranoside	**MDG**	-	+	+	+	+	-
Mannitol	**MAN**	+	+	+	+	+	-
Xylitol	**XLT**	-	-	-	-	-	-
Nitrate	**NIT**	+	+	+	+	+	+
Acetoin Production	**VP**	-	-	+	-	+	+
Novobiocin	**NOVO**	-	-	-	-	-	-
Gelatine	**GEL**	-	-	-	-	-	-
Esculin	**ESC**	+	-	+	-	+	-
Catalase	**CAT**	+	+	+	+	+	+
Urease	**URE**	+	+	+	+	+	-
N-acetyl-β-Glucosaminidase	**βNAG**	-	-	-	-	-	-
α-glucosidase	**αGLU**	-	-	-	-	-	-
β-galactosidase	**βGAL**	-	-	-	-	-	-
β-glucuronidase	**βGUR**	+	-	-	+	-	+
Alkaline phosphatase	**PAL**	+	-	-	+	-	+
Pyrazinamidase	**PYZ**	+	+	+	+	+	+
Arginine arylamidase	**ArgA**	-	-	-	-	-	-
Pyrrolidonyl arylamidase	**PyrA**	+	+	+	+	+	+
Ornithine decarboxilase	**ODC**	-	-	-	-	-	-
Arginine dihydrolase	**ADH**	+	+	+	+	+	+

Table 6: Cell fatty acid-fatty acid methyl ester (CFA-FAME) analysis of *S. borealis* strains (CCUG 73747^{T}, CCUG 73748, CCUG 73749, CCUG 73750, CCUG 73751), showing the ECL (equivalent chain length), name of CFAs and the Area per peak (%). Also shown are the corresponding CFA profiles of *S. aureus* (CCUG 1800^{T}), *S. devriesei* (CCUG 58238^{T}) and *S. haemolyticus* (CCUG 7323^{T}). tr denotes "trace" which means a peak has been recorded, but too small to be integrated.

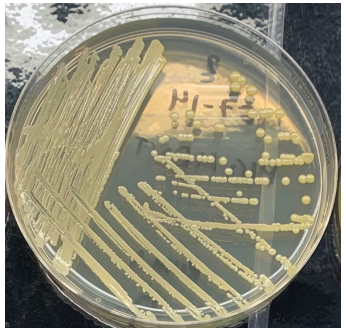
ECL	**13.618**	**14.621**	**14.711**	**15.626**	**16.000**	**16.629**	**16.722**	**17.724**	**17.769**	**18.000**	**18.633**	**18.729**	**20.000**
Peak name of CFA	**C14:0 ISO**	**C15:0 ISO**	**C15:0 ANTESIO**	**C16:0 ISO**	**C16:0**	**C17:0 ISO**	**C17:0 ANTESIO**	**Summed feature***	**18:1 ω9c**	**C18:0**	**C19:0 IS0**	**C19:0 ANTESIO**	**C20:0**
Spp													
***S. aureus* CCUG 1800^{T}**	tr	7.3	55.1	1.6	2,0	3.9	19,5	3.3	1.4	2.6	tr	1.3	tr
***S. devriesei* CCUG 58238^{T}**	0,0	3.2	52.1	tr	1,3	5.1	25.8	2.6	1,5	2,5	1,0	3.8	tr
***S. haemolyticus* CCUG 7323^{T}**	1.1	6.9	42,7	1.3	1,8	6,8	18.8	1.2	0,0	8,5	2.8	5,7	2,5
***S. borealis* 51-48 CCUG 73747^{T}**	1.1	10,0	61.8	tr	1.1	4,9	12.8	tr	1,0	2.3	1,0	1.9	tr
***S. borealis* 57-14 CCUG 73748**	tr	12,0	64,3	tr	1,2	5,6	13,1	0,0	tr	2,4	tr	1,6	0,0
***S. borealis* 57-74 CCUG 73749**	1,1	10,9	65,6	tr	tr	4,6	12,2	tr	tr	2,1	1,3	2,3	tr
***S. borealis* 58-22 CCUG 73750**	1,2	10.8	61,1	1,0	1,2	5,5	13,3	tr	tr	2.6	1.1	1,7	tr
***S. borealis* 58-52 CCUG 73751**	1,3	11,7	59.2	1.1	1,7	5,3	13,8	tr	tr	3.9	0,0	1,2	tr


Tree scale: 0.1
Macrococcus caseolyticus DSM 23656T
S. stepanovicii DSM 26319T
S. lentus NCTC 12102T
S. sciuri NCTC 12103T
S. fleuretti NCTC 13829T
S. vitulinus DSM 15615T
S. massiliensis CCUG 55927T
S. chromogenes NCTC 10530T
S. agnetis DSM 23656T
S. hyicus ATCC 11249T
S. muscae ATCC 49910T
S. microti DSM 22147T
S. rostri DSM 21968T
S. felis ATCC 49168T
S. lutrae ATCC 700373T
S. cornubiensis NW1T
S. intermedius NCTC 11048T
S. delphini NCTC 12225T
S. pseudointermedius CCUG 49543T
S. simulans NCTC 11046T
S. condimenti DSM 11674T
S. piscifermentans NCTC 13836T
S. debuckii SDB 2975T
S. auricularis DSM 20609T
S. argensis DSM 29875T
S. pettenkoferi CCUG 51270T
S. kloosii ATCC 43959T
S. arlettae NCTC 12413T
S. gallinarum DSM 20610T
S. nepalensis DSM 15150T
S. cohnii subsp. cohnii NCTC 11041T
S. cohnii subsp. urealyticus DSM 6718T
S. succinus subsp. succinus DSM14617T
S. equorum subsp. equorum NCTC 12414T
S. equorum subsp. linens DSM 15097T
S. pseudoxylosus S04009T
S. xylosus ATCC 29971T
S. caeli 82BT
S. edaphicus CMM 8730T
S. saprophyticus subsp. saprophyticus ATCC 15305T
S. saprophyticus subsp. bovis CCUG 38042T
S. simiae CCM 7213T
S. schweitzeri DSM 28300T
S. argentus MSHR1132T
S. aureus subsp. anaerobius DSM 20714T
S. aureus NCTC 8325T
S. warneri NCTC 11044T
S. saccharolyticus NCTC 11807T
S. epidermidis ATCC 14990T
S. caprae NCTC 12196T
S. capitis NCTC 11045T
S. capitis subsp. urealyticus DSM 6717T
S. lugdunensis NCTC 12217T
S. hominis subsp. novobiosepticus CCUG 42399T
S. hominis subsp. hominis NCTC 11320T
S. petrasii subsp. pragensis CCM 8529T
S. petrasii subsp. croceilyticus CCM 8421T
S. petrasii subsp. jettensis CCM 8494T
S. petrasii subsp. petrasii CCM 8418T
S. devriesei CCUG 58238T
S. haemolyticus NCTC 11042T
SNUC1342
SNUC119
Staphylococcus sp. 170179
S. borealis 51-48T
S. borealis 58-22
S. borealis 58-52
S. borealis 57-74
S. borealis 57-14

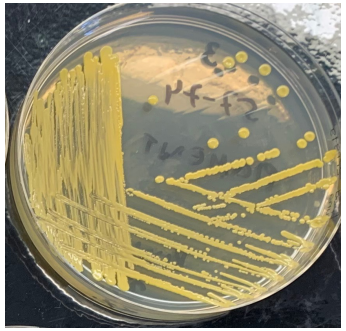

0.1
Staphylococcus haemolyticus
Staphylococcus borealis
58-22
58-52
57-74
57-14
51-48T
170179
SNUC1342
SNUC119



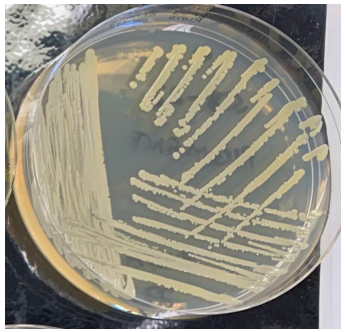
S. borealis 51-48^{T}



S. borealis 57-14



S. borealis 57-74



S. borealis 58-22

S. borealis 58-52

S. haemolyticus
CCUG 7323^{T}

Supplementary material

Staphylococcus borealis sp. nov. -A novel species isolated from skin and blood in humans

Maria Pain, Runa Wolden, Daniel Jaén-Luchoro, Francisco Salvà-Serra, Beatriz Piñeiro Iglesias, Roger Karlsson, Claus Klingenberg, Jorunn Pauline Cavanagh·

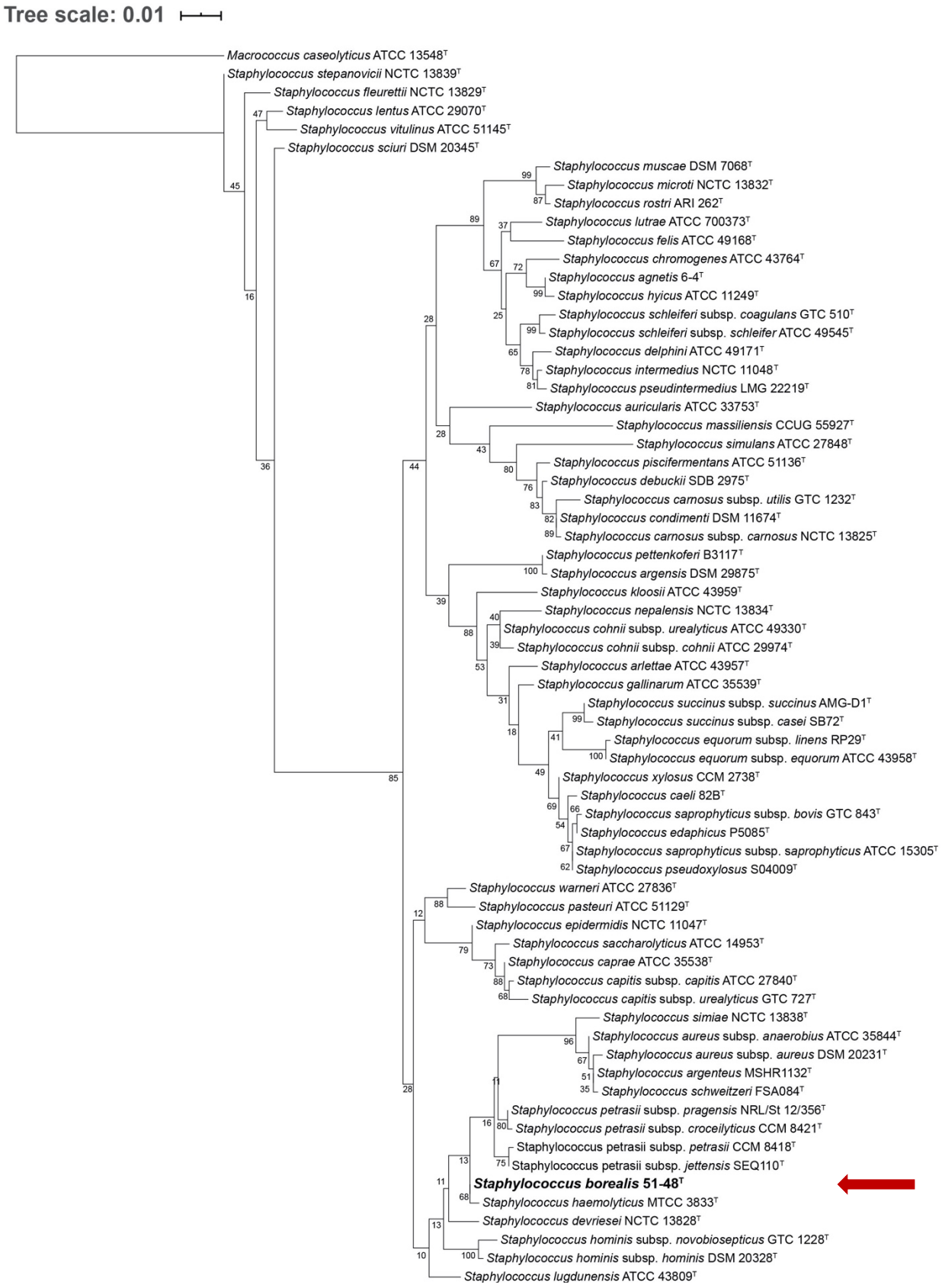


Figure S1. Phylogenetic relationship of Staphylococcal type strains based on 16S rRNA gene sequence. The phylogenetic tree was generated with MUSCLE using 16S rRNA sequences of Staphylococcal type strains and the type strain of *Macrococcus caseolyticus* as outgroup. The maximum likelihood method was used and bootstrapping was set to 500 replicates, using the RAxML software. rRNA sequences were obtained from EzBioCloud (1) (Type strain ID displayed in figure, accession numbers listed in Supplementary **Table S1**).

Table S1: Accession numbers of staphylococcal type strains used in the genomic analysis.

Species	Subspecies	WGS assembly accession	16S accession
S. agnetis		GCA_002901865.1	HM484980
S. argensis		GCA_002902305.1	PPPX01000013
S. argenteus		GCA_000236925.1	FR821777
S. arlettae		GCA_002902345.1	AB009933
S. aureus	*aureus*	GCA_000013425.1	AMYL01000007
	anaerobius	GCA_002902425.1	D83355
S. auricularis		GCA_001500315.1	L37598
S. caeli		GCA_900097965.1	MH431939
S. capitis	*capitis*	GCA_002902325.1	L37599
	urealyticus	GCA_002901925.1	AB233325
S. caprae		GCA_002902725.1	AB009935
S. carnosus	*carnosus*	-	UHCT01000001
	utilis	-	AB233329
S. chromogenes		GCA_900458195.1	D83360
S. cohnii	*cohnii*	GCA_002902365.1	D83361
	urealyticus	GCA_002902235.1	AB009936
S. condimenti		GCA_001618885.1	CP015114
S. cornubiensis		GCA_900183575.1	-
S. debuckii		-	MK121623
S. delphini		GCA_900636325.1	AB009938
S. devriesei		GCA_002902625.1	UHCZ01000002
S. edaphicus		GCA_002614725.1	KY315825
S. epidermidis		GCA_006094375.1	UHDF01000003
S. equorum	*equorum*	GCA_900458565.1	AB009939
	linens	GCA_002901955.1	AF527483
S. felis		GCA_003012915.1	D83364
S. fleuretti		GCA_900458505.1	UHDL01000001
S. gallinarum		GCA_000875895.1	D83366
S. haemolyticus		GCA_900458595.1	LILF01000056
S. hominis	*hominis*	GCA_002901845.1	X66101
	novobiosepticus	GCA_002902465.1	AB233326
S. hyicus		GCA_000816085.1	CP008747
S. intermedius		GCA_000308095.1	CAIB01000045
S. kloosi		GCA_003019255.1	AB009940
S. lentus		GCA_900458735.1	D83370
S. lugdunensis		GCA_900478255.1	AB009941
S. lutrae		GCA_002101335.1	CP020773
S. massiliensis		GCA_000298075.1	EU707796
S. microti		GCA_000934465.1	UHDT01000001
S. muscae		GCA_003019275.1	FR733703
S. nepalensis		GCA_002902745.1	UHDS01000001
S. pasteuri		-	AF041361
S. petrasii	*petrasii*	GCA_004684865.1	JX139845
	croceilyticus	GCA_004684875.1	AY953148
	jettensis	GCA_004745855.1	JN092118
	pragensis	GCA_004785665.1	KM873669
S. pettenkoferi		GCA_002902685.1	AF322002
S. piscifermentans		GCA_900156985.1	AB009943
S. pseudintermedius		GCA_002902125.1	AJ780976
S. pseudoxylosus		-	MH643903
S. rostri		GCA_002902145.1	FM242137
S. saccharolyticus		GCA_003970495.1	L37602
S. saprophyticus	*saprophyticus*	GCA_000010125.1	AP008934
	bovis	GCA_002902545.1	AB233327
S. schleiferi	*schleiferi*	-	AB009945
	coagulans	-	AB233334
S. schweitzeri		GCA_002902405.1	CCEL01000025
S. sciuri		GCA_900474615.1	AJ421446
S. simiae		GCA_000235645.1	LT906460
S. simulans		GCA_900474685.1	D83373
S. stepanovicii		GCA_002902525.1	LT906462
S. succinus	*succinus*	GCA_001006765.1	AF004220
	casei	-	AJ320272
S. vitulinus		GCA_002902265.1	AB009946
S. warneri		GCA_900636385.1	L37603
S. xylosus		GCA_900240095.1	MRZO01000018
Macrococcus caseolyticus		GCA_002902665.1	D83359

Table S2: Result for species delineation of *S. borealis* against *S. haemolyticus* NCTC 11042^T using 16S RNA gene sequence, ANI, tetra and dDDH comparisons. Values in brackets indicate percent alignment between *S. haemolyticus* NCTC 11042^T and the different *S. borealis* isolates.

Isolate	16S rRNA gene (<98.7%)	ANIb (<95%)	ANIm (<95%)	Tetra (<0.99)	dDDH (<70 %)
51-48^T	99.86%	87.74% [83.78]	88.67% [86.04]	0.98571	34.2%
57-14	99.86%	87.68% [82.94]	88.59% [85.37]	0.98435	34.2%
57-74	99.86%	87.66% [83.32]	88.59% [85.81]	0.98445	34.2%
58-22	99.93%	87.66% [82.87]	88.58% [85.44]	0.98413	34.2%
58-52	99.93%	87.89% [83.00]	88.61% [85.75]	0.98455	34.3%
SNUC119	99.79%	87.96% [81.42]	88.61% [85.34]	0.98289	34.5%
SNUC1342	99.86%	87.88% [81.96]	88.61% [85.04]	0.98197	34.5%
170179	-	87.65% [82.78]	88.58% [85.67]	0.98492	34.2%

Table S3: Results for species delineation between the *S. borealis* strains. Percent ANIb (based on blast+) scores and [aligned nucleotides].

	S. borealis 51-48^T	***S. borealis 57-14***	***S. borealis 57-74***	***S. borealis 58-22***	***S. borealis 58-52***	***S. borealis 179170***	***S. borealis SNUC119***
S. borealis 57-14	99.58 [91.58]	-					
S. borealis 57-74	99.56 [92.35]	99.82 [95.97]	-				
S. borealis 58-22	99.65 [93.49]	99.89 [97.60]	99.89 [96.45]	-			
S. borealis 58-52	99.74 [92.63]	99.87 [96.02]	99.92 [96.20]	99.94 [94.96]	-		
S. borealis 179170	99.54 [90.53]	99.79 [95.52]	99.77 [95.13]	99.79 [94.14]	99.59 [94.97]	-	
S. borealis SNUC119	97.67 [81.85]	97.64 [85.67]	97.73 [85.61]	97.61 [83.67]	97.63 [84.28]	97.67 [85.00]	-
S. borealis SNUC1342	97.65 [82.41]	97.63 [86.34]	97.73 [84.88]	97.59 [84.37]	97.58 [84.78]	97.68 [85.54]	99.84 [96.70]

1. Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, et al. Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol. 2017;67(5):1613–7.