



Complete Genome Sequence of *Staphylococcus epidermidis* PH1-28, Isolated from the Forehead of a Hyperseborrheic Donor

Pascal Hilaire,^a Leslie Landemaine,^b Sandy Contreras,^c H  l  ne Blanquart-Goudezeune,^c Patricia Siguier,^d Fran  ois Cornet,^d H  l  ne Chiapello,^e Valentin Loux,^e C  cile Clavaud,^b  Stanislas C. Morand^b

^aL'Or  al R&I, Tours, France

^bL'Or  al R&I, Aulnay-sous-Bois, France

^cGenoScreen, Lille, France

^dCBI, LMGM, Universit   de Toulouse, UPS, CNRS, Toulouse, France

^eUniversit   Paris-Saclay, INRAE, Bioinformatics, MIGALE Bioinformatics Facility, Jouy-en-Josas, France

ABSTRACT We report the complete genome sequence of *Staphylococcus epidermidis* commensal strain PH1-28, isolated from the forehead of a healthy donor. The assembled 2.6-Mbp genome consisted of one chromosome and five plasmids. These data will provide valuable information and important insights into the physiology and metabolism of this skin flora microorganism.

Staphylococcus epidermidis is a Gram-positive commensal *Firmicutes* bacterium ubiquitously present on human skin and mucous membranes (1, 2). *S. epidermidis* strains exhibit ambivalent roles; as an opportunistic pathogen, *S. epidermidis* is a common cause of nosocomial infections due to its capability to form biofilms on medical implants (3). In contrast, some commensal strains protect against skin colonization by pathogens or modulate the immune system (4, 5). We present here the complete *S. epidermidis* strain PH1-28 genome sequence, isolated after swabbing the forehead of a healthy hyperseborrheic donor.

S. epidermidis PH1-28 was isolated during a research protocol (study number ARC/COPEG/1125) on a tryptic soy agar plate and cultured with tryptic soy broth (TSB) medium at 37  C. Initial *S. epidermidis* taxonomic identification was deciphered by a BLASTn search on the nonredundant database v2.2.29+ (6) after 16S PCR followed by Sanger sequencing. High-molecular-weight genomic DNA was extracted from an overnight culture grown in TSB at 37  C using the Genra Puregene kit (Qiagen). Sequencing was first performed on a MinION device with an R9.4.1 flow cell (Oxford Nanopore Technologies; rapid sequencing SQK-RAD004 library with 400 ng of DNA and 50-s tagmentation; base calling with Guppy v4.2.2 in 450_bps fast configuration). Long reads (12.9 Gbp; 1,265,187 reads; N_{50} , 17,655 bp) were controlled with MinIONQC v1.4.1 (7) and filtered and trimmed using Nanofilt v2.6.0 (-q 7 -l 1,000), and adaptors were removed with Porechop v0.2.4 (8, 9). Then, sequencing on a HiSeq 2500 system (Illumina; 2-kb Nextera XT library) generated short reads, which were assessed with FastQC v0.11.5 (10) and cleaned with cutadapt v1.18 (11) and Prinseq v0.20.4 (12) (with the parameters -trim_qual_right 30 -trim_qual_type min -trim_qual_rule lt -trim_qual_window 7 -ns_max_n 0 -noniupac -min_qual_mean 30 -trim_left 15 -min_len 60), resulting in 11,094,664 reads (936 Mbp). Both short and long cleaned reads were *de novo* assembled using the Canu v2.0 (13), SPAdes v3.10.1 (14), and MaSuRCA v3.3.0 (PE = pe 600 50) (15) tools. The consensus assembly was manually curated and carefully verified using Bowtie 2 v2.1.0 (16), minimap2 v2.17 (17), and Geneious Assembler (Biomatters) by calculating the Reads Mapped Back to Contigs (RMBC) index (99.6 to 99.9% rate for the three tools). After assembly polishing with Pilon v1.23 (18), the

Citation Hilaire P, Landemaine L, Contreras S, Blanquart-Goudezeune H, Siguier P, Cornet F, Chiapello H, Loux V, Clavaud C, Morand SC. 2021. Complete genome sequence of *Staphylococcus epidermidis* PH1-28, isolated from the forehead of a hyperseborrheic donor. *Microbiol Resour Annot* 10:e00165-21. <https://doi.org/10.1128/MRA.00165-21>.

Editor Steven R. Gill, University of Rochester School of Medicine and Dentistry

Copyright    2021 Hilaire et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Stanislas C. Morand, stanislas.morand@rd.loreal.com.

Received 12 February 2021

Accepted 15 February 2021

Published 4 March 2021

chromosomal and plasmid coverage and circularity were validated with GeniousPrime software v2020.2.5. Finally, the chromosome was rotated to start at *dnaA*. The whole-genome median coverage depth reaches 4,828-fold. The assembled ungapped genome is 2,598,183 bp long (GC content, 32.05%), comprising one chromosome (2,468,349 bp) and five plasmids, pSE1 (51,568 bp), pSE2 (45,209 bp), pSE3 (28,218 bp), pSE4 (2,491 bp), and pSE5 (2,348 bp). pSE5 is identical to *Staphylococcus aureus* plasmid CP047868. Structural and functional annotations, carried out with comparisons using the Prokka v1.14.6 (19), RAST v0.11 (20), and PGAP v4.12 (21) tools, identified 19 rRNAs (6 operons [16S-23S-5S] with an extra 5S copy) and 61 tRNAs. This process pointed out 2,332 protein-coding sequences, 248 of which (11%) encode hypothetical proteins.

GC skew analysis using BRIG v0.95 (22) showed a clear chromosomal GC+/- region segmentation. Comparison with other *S. epidermidis* complete genome sequences using dRep v2.5.4 (23) based on average nucleotide identity (ANI) showed that the closest strain is ATCC 14990 (ANI, 0.994). PHASTER analysis (24) revealed that the chromosome comprises a ca. 40-kb locus corresponding to the *Siphoviridae* class I StB20 prophage (25, 26). Searches for mobile elements using ISsaga v2.0 (27) and ISEScan v1.5.4 (28) showed the presence of 43 transposases, 21 of which belong to the identical IS110 family. The PH1-28 strain accumulation-associated protein aap does not contain the classical C-terminal collagen triple helix domain, and both massive *ebh* and *fmtB* protein-coding genes enclose internal stop codons. According to TADB v2.0 (29), plasmids pSE1 and pSE2 contain two and one type-II toxin-antitoxin clusters, respectively. Finally, the Abricate v1.0.1 tool showed that the PH1-28 strain does not carry the *blaZ* beta-lactamase gene present in the closely related ATCC 14990 strain (30). Default parameters were used for all software unless otherwise specified.

Data availability. The complete genome sequence is available in GenBank under the accession numbers CP066371 to CP066376, BioProject PRJNA662445, and BioSample SAMN16085533. The version described in this paper is the first version. Next-generation sequencing (NGS) raw reads have been deposited at the Sequence Read Archive under the numbers SRR13274404 and SRR13274405.

ACKNOWLEDGMENT

P.H., L.L., H.B.-G., and S.C.M. performed the research experiments; S.C., P.S., H.C., V.L., C.C., and S.C.M. analyzed the data; S.C.M. wrote the first draft of the manuscript, and all the authors contributed to and approved the final version.

REFERENCES

- Oh J, Byrd AL, Deming C, Conlan S, Kong HH, Segre JA, NISC Comparative Sequencing Program. 2014. Biogeography and individuality shape function in the human skin metagenome. *Nature* 514:59–64. <https://doi.org/10.1038/nature13786>.
- Oh J, Byrd AL, Park M, Kong HH, Segre JA, NISC Comparative Sequencing Program. 2016. Temporal stability of the human skin microbiome. *Cell* 165:854–866. <https://doi.org/10.1016/j.cell.2016.04.008>.
- Fey PD, Olson ME. 2010. Current concepts in biofilm formation of *Staphylococcus epidermidis*. *Future Microbiol* 5:917–933. <https://doi.org/10.2217/fmb.10.56>.
- Cogen AL, Yamasaki K, Muto J, Sanchez KM, Crotty Alexander L, Tanios J, Lai Y, Kim JE, Nizet V, Gallo RL. 2010. *Staphylococcus epidermidis* antimicrobial delta-toxin (phenol-soluble modulin-gamma) cooperates with host antimicrobial peptides to kill group A *Streptococcus*. *PLoS One* 5: e8557. <https://doi.org/10.1371/journal.pone.0008557>.
- Linehan JL, Harrison OJ, Han S-J, Byrd AL, Vujkovic-Cvijin I, Villarino AV, Sen SK, Shaik J, Smelkinson M, Tamoutounour S, Collins N, Bouladoux N, Dzutsev A, Rosshart SP, Ar buckle JH, Wang C-R, Kristie TM, Rehmann B, Trinchieri G, Brenchley JM, O'Shea JJ, Belkaid Y. 2018. Non-classical immunity controls microbiota impact on skin immunity and tissue repair. *Cell* 172:784–796.e18. <https://doi.org/10.1016/j.cell.2017.12.033>.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Lanfear R, Schalamun M, Kainer D, Wang W, Schwessinger B. 2019. MinIONQC: fast and simple quality control for MinION sequencing data. *Bioinformatics* 35:523–525. <https://doi.org/10.1093/bioinformatics/bty654>.
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>.
- Wick RR. 2018. Porechop: adapter trimmer for Oxford Nanopore reads. <https://github.com/rwick/Porechop>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17:10–12. <https://doi.org/10.14806/ej.17.1.200>.
- Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>.
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
- Bankovich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.

15. Zimin A, Marçais G, Puiu D, Roberts M, Salzberg S, Yorke J. 2013. The MaSuRCA genome assembler. *Bioinformatics* 29:2669–2677. <https://doi.org/10.1093/bioinformatics/btt476>.
16. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
17. Li H. 2018. minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
18. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
19. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
20. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.
21. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
22. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402. <https://doi.org/10.1186/1471-2164-12-402>.
23. Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* 11:2864–2868. <https://doi.org/10.1038/ismej.2017.126>.
24. Arndt D, Grant J, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–W21. <https://doi.org/10.1093/nar/gkw387>.
25. Deghorain M, Bobay L-M, Smeesters PR, Bousbata S, Vermeersch M, Perez-Morga D, Drèze P-A, Rocha EPC, Touchon M, Van Melderen L. 2012. Characterization of novel phages isolated in coagulase-negative staphylococci reveals evolutionary relationships with *Staphylococcus aureus* phages. *J Bacteriol* 194:5829–5839. <https://doi.org/10.1128/JB.01085-12>.
26. van Zyl LJ, Abrahams Y, Stander EA, Kirby-McCollough B, Jourdain R, Clavaud C, Breton L, Trindade M. 2018. Novel phages of healthy skin metaviromes from South Africa. *Sci Rep* 8:12265. <https://doi.org/10.1038/s41598-018-30705-1>.
27. Varani AM, Siguier P, Gourbeyre E, Charneau V, Chandler M. 2011. ISsaga is an ensemble of Web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome Biol* 12:R30. <https://doi.org/10.1186/gb-2011-12-3-r30>.
28. Zhiquan X, Haixu T. 2017. ISEScan: automated identification of insertion sequence elements in prokaryotic genomes. *Bioinformatics* 33:3340–3347. <https://doi.org/10.1093/bioinformatics/btx433>.
29. Xie Y, Wei Y, Shen Y, Li X, Zhou H, Tai C, Deng Z, Ou HY. 2018. TADB 2.0: an updated database of bacterial type II toxin-antitoxin loci. *Nucleic Acids Res* 46:D749–D753. <https://doi.org/10.1093/nar/gkx1033>.
30. Seemann T. 2015. Abricate: mass screening of contigs for antimicrobial resistance or virulence genes. <https://github.com/tseemann/abricate>.