

Genome Sequence of *Staphylococcus massiliensis* Strain S46, Isolated from the Surface of Healthy Human Skin

Rajpal Srivastav,^{a,b} Ajit Singh,^a Pramod Kumar Jangir,^a Chanchal Kumari,^a Suchismita Muduli,^a Rakesh Sharma^{a,c}

Institute of Genomics and Integrative Biology, Council of Scientific and Industrial Research (CSIR), Delhi, India^a; Department of Biotechnology, University of Pune, Pune, India^b; National Chemical Laboratory, Council of Scientific and Industrial Research (CSIR), Pune, India^c

***Staphylococcus massiliensis* strain S46 was isolated from the surface of healthy human skin. Here, we report the draft genome sequence of *S. massiliensis* S46 (2,447,110 bp, with a G+C content of 36.3%).**

Received 28 June 2013 Accepted 8 July 2013 Published 8 August 2013

Citation Srivastav R, Singh A, Jangir PK, Kumari C, Muduli S, Sharma R. 2013. Genome sequence of *Staphylococcus massiliensis* strain S46, isolated from the surface of healthy human skin. *Genome Announc.* 1(4):e00553-13. doi:10.1128/genomeA.00553-13.

Copyright © 2013 Srivastav et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Rakesh Sharma, rsharma@igib.res.in.

Staphylococcus massiliensis strain S46 was isolated from the surface of healthy human skin and was identified as *S. massiliensis* based on the *rrs* (16S rRNA) gene sequence identity of 100% with type strain *S. massiliensis* 5402776^T. *S. massiliensis* is a coagulase-negative, nonmotile, non-spore-forming bacterium that was first isolated from a human brain abscess (1) and later was proposed as a component of normal human skin microflora (2). The genome sequence of strain 5402776^T was reported recently (3). Here, we report the draft genome sequence of *S. massiliensis* S46.

The genomic DNA of *S. massiliensis* S46 was extracted from culture grown on Luria-Bertani (LB) broth. The genome of *S. massiliensis* S46 was sequenced by a whole-genome shotgun sequencing strategy using 454 GS FLX, which produced a total of 157,855 reads. The reads were assembled into 64 large (>500 bp) contigs by Newbler (version 2.6). The contigs were annotated using the prokaryotic genome automatic annotation pipeline at NCBI. The automated pipeline uses GeneMark and Glimmer (4, 5, 6) for gene prediction and compares the translated proteins with the nonredundant protein database at GenBank, Entrez Protein Clusters (7), the Conserved Domain Database (8), and COGs (9) for annotation. The genome sequence was also submitted to the RAST (10) server for annotation and identification of metabolic pathways.

The *S. massiliensis* S46 unclosed draft genome is 2,447,110 bp in length, with a G+C content of 36.3%. The genome contains 2,324 coding sequences, 56 tRNA genes, and 6 rRNA genes. The strain S46 genome has predicted genes for glycolysis, the tricarboxylic acid (TCA) cycle, fatty acid metabolism, phosphorus metabolism, iron acquisition, and organic sulfur assimilation.

A comparison with the genome sequences available at RAST showed that *Staphylococcus epidermidis* strain RP62A (score, 523) was the closest neighbor of *S. massiliensis* S46. The genes for N-linked glycosylation, sialic acid metabolism, thiamin biosynthesis, and trehalose uptake and utilization were present in *S. massiliensis* S46 but not in *S. epidermidis* RP62A. However, the genes for biofilm formation, quorum sensing, late competence, arginine biosynthesis, inorganic sulfur assimilation, and urea decomposition were absent in *S. massiliensis* S46 but were present in *S. epidermidis* RP62A. A detailed comparative genome analysis

between the strains of *S. massiliensis* and *S. epidermidis* will reveal their metabolic differences and strategies to survive as commensals on the human skin surface.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [AMSQ00000000](https://www.ncbi.nlm.nih.gov/nuccore/AMSQ00000000). The version described in this paper is the first version, [AMSQ01000000](https://www.ncbi.nlm.nih.gov/nuccore/AMSQ01000000). The 64 contigs have been deposited under the accession no. [AMSQ01000001](https://www.ncbi.nlm.nih.gov/nuccore/AMSQ01000001) to [AMSQ01000064](https://www.ncbi.nlm.nih.gov/nuccore/AMSQ01000064).

ACKNOWLEDGMENTS

We thank Rajesh S. Gokhale, director of the IGIB, for encouragement and support.

This work was supported by the Council of Scientific and Industrial Research (CSIR) grant “Toward understanding of skin cell homeostasis” (TOUCH), grant no. BSC0302. A.S. acknowledges CSIR for a senior research fellowship. P.K.J. acknowledges CSIR grant no. BSC0124.

REFERENCES

1. Al Masalma M, Raoult D, Roux V. 2010. *Staphylococcus massiliensis* sp. nov., isolated from a human brain abscess. *Int. J. Syst. Evol. Microbiol.* 60:1066–1072.
2. Zong Z. 2012. The newly-recognized species *Staphylococcus massiliensis* is likely to be part of the human skin microflora. *Antonie Van Leeuwenhoek* 101:449–451.
3. Roux V, Robert C, Gimenez G, Raoult D. 2012. Draft genome sequence of *Staphylococcus massiliensis* strain 5402776^T. *J. Bacteriol.* 194:6984–6985.
4. Borodovsky M, McIninch J. 1993. GeneMark: parallel gene recognition for both DNA strands. *Comput. Chem.* 17:123–133.
5. Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 26:1107–1115.
6. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with Glimmer. *Nucleic Acids Res.* 27:4636–4641.
7. Klimke W, Agarwala R, Badrettin A, Chetvernin S, Ciufu S, Fedorov B, Kiryutin B, O'Neill K, Resch W, Resenchuk S, Schafer S, Tolstoy I, Tatusova T. 2009. The National Center for Biotechnology Information's Protein Clusters Database. *Nucleic Acids Res.* 37:D216–D223. doi:10.1093/nar/gkn734.
8. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH. 2011. CDD: a

- Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 39:D225–D229. doi:[10.1093/nar/gkq1189](https://doi.org/10.1093/nar/gkq1189).
9. Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28:33–36.
 10. 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. doi:[10.1186/1471-2164-9-75](https://doi.org/10.1186/1471-2164-9-75).