Draft Genome Sequence of *Deinococcus* sp. Strain S9, Isolated from Microbial Mat Deposits of Hot Springs Located atop the Himalayan Ranges at Manikaran, India

Chandni Talwar,[a] Amit Kumar Singh,[b] Durgesh Narain Singh,[a] Shekhar Nagar,[a] Yogendra Singh,[a] Mallikarjun Shakarad,[a] Ram Krishan Negi,[a] Rup Lal[a,c]

[a]Department of Zoology, University of Delhi, Delhi, India
[b]Department of Biotechnology, Jamia Millia Islamia, New Delhi, India
[c]PhiXGen Private Limited, Gurugram, Haryana, India

ABSTRACT Here, we present the draft genome sequence of *Deinococcus* sp. strain S9, a red-pigmented and moderately thermophilic bacterium isolated from microbial mat deposits around the hot springs at Manikaran, Himachal Pradesh, India. The draft genome (3.34 Mb) contains 101 contigs with an average GC content of 66.4%.

Our group has been incessantly exploring the microbial diversity at the geothermal hot springs located at Manikaran, Himachal Pradesh, India (1–7). Previously, we reported the draft genome sequence of *Deinococcus* sp. strain RL, which was isolated from sediments of this hot spring (6). In continuation of our efforts to study the extremophilic microbial diversity, we isolated a red-pigmented bacterium, *Deinococcus* sp. strain S9, from microbial mat deposits around the hot spring located at an altitude of 1,760 m at Manikaran (32°020′N, 72°210′E), Himachal Pradesh, India (surface temperatures, >57°C) (4). Briefly, 10 g of microbial mat sample was suspended in sterile normal saline solution and shaken at 45°C for 3 h. Serial dilutions from the seed sample were spread on Luria-Bertani agar (HiMedia, Mumbai, India) and incubated at 45°C. A red-pigmented bacterial colony optimally growing at 45°C was picked up after 24 h and designated S9. Based on 16S rRNA gene sequencing (1,278 bp), the isolate showed 99.06% identity with the type strain *Deinococcus geothermalis* DSM 11300.

The genomic DNA was isolated as described elsewhere (8). To determine the whole-genome sequence, library preparation was done using the TruSeq nano DNA PCR-free kit (catalog number FC-121-3001; Illumina, San Diego, CA), and sequencing was performed using the Illumina HiSeq 2500 platform. In total, 42,493,218 paired-end reads 100 bp long were generated and processed with Trimmomatic v0.39 (9) using default parameters. Assembly was performed using SPAdes v3.11.0 (10), by employing a multi-k-mer approach (k = 27, 31, 33, 35, 41, 45, 49, 53, 55, 75, 77, and 95), into 179 contigs. The assembly was validated by using QUAST v4.5 (11), and contigs <500 bp long were removed. The final assembly contains 101 contigs (N_{50}, 111,497 bp) at 107-fold coverage. The total size of the draft genome is 3,344,786 bp, with an average GC content of 66.4%. The 16S, 23S, and 5S rRNA gene sequences (1 of each) were identified within the genome by using RNAmmer (12). In addition, 49 tRNA genes and 1 transfer-messenger RNA (tmRNA) gene were also predicted using ARAGORN (13). Nine CRISPR arrays along with several CRISPR-associated (*cas*) genes were also determined within the draft genome using CRISPRFinder (14).

Annotation using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (15) predicted 3,108 protein-coding sequences. Enzymes involved in carotenoid biosynthesis, such as phytoene synthase, phytoene desaturase, and lycopene cyclase, were annotated within the draft genome (16, 17). The genomic studies of *Deinococcus* spp.

are of great interest because of their high resistance to radiation, temperatures, and heavy metals and unparalleled efficiency in repairing damaged DNA (18). Several genes responsible for the DNA repair and resistance to radiation were also annotated, such as *uvrA* (19), *recO* (20), *recA* (21), *recF*, and *recR* (22). In addition, the gene coding for ClpX, which is involved in the repair of double-stranded breaks upon gamma irradiation (23), was also annotated. In-depth analysis of this genome would provide crucial details required for the understanding of the mechanisms involved in nucleotide excision repair, carotenoid biosynthesis, and survival at extreme niches.

Data availability. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number SKCF00000000 (BioProject number PRJNA525588 and BioSample number SAMN11053996). The version described in this paper is SKCF01 and consists of sequences SKCF01000001 to SKCF01000101. The raw reads are available in NCBI under BioProject accession number PRJNA525588.

ACKNOWLEDGMENTS

REFERENCES

1. Sharma A, Schmidt M, Kiesel B, Mahato NK, Cralle L, Singh Y, Richnow HH, Gilbert JA, Arnold W, Lal R. 2018. Bacterial and archaeal viruses of Himalayan hot springs at Manikaran modulate host genomes. Front Microbiol 9:3095. https://doi.org/10.3389/fmicb.2018.03095.

2. Tripathi C, Mahato NK, Rani P, Singh Y, Kamra K, Lal R. 2016. Draft genome sequence of *Lampropedia cohaerens* strain CT6^T isolated from arsenic rich microbial mats of a Himalayan hot water spring. Stand Genomic Sci 11:64. https://doi.org/10.1186/s40793-016-0179-1.

3. Sharma A, Kohli P, Singh Y, Schumann P, Lal R. 2016. *Fictibacillus halophilus* sp. nov., from a microbial mat of a hot spring atop the Himalayan Range. Int J Syst Evol Microbiol 66:2409–2416. https://doi.org/10.1099/ijsem.0.001051.

4. Sangwan N, Lambert C, Sharma A, Gupta V, Khurana P, Khurana JP, Sockett RE, Gilbert JA, Lal R. 2015. Arsenic rich Himalayan hot spring metagenomics reveal genetically novel predator-prey genotypes. Environ Microbiol Rep 7:812–823. https://doi.org/10.1111/1758-2229.12297.

5. Sharma A, Hira P, Shakarad M, Lal R. 2014. Draft genome sequence of *Cellulosimicrobium* sp. strain MM, isolated from arsenic-rich microbial mats of a Himalayan hot spring. Genome Announc 2:e01020-14. https://doi.org/10.1128/genomeA.01020-14.

6. Mahato NK, Tripathi C, Verma H, Singh N, Lal R. 2014. Draft genome sequence of *Deinococcus* sp. strain RL isolated from sediments of a hot water spring. Genome Announc 2:e00703-14. https://doi.org/10.1128/genomeA.00703-14.

7. Dwivedi V, Sangwan N, Nigam A, Garg N, Niharika N, Khurana P, Khurana JP, Lal R. 2012. Draft genome sequence of *Thermus* sp. strain RL, isolated from a hot water spring located atop the Himalayan ranges at Manikaran, India. J Bacteriol 194:3534. https://doi.org/10.1128/JB.00604-12.

8. Marmur J. 1961. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J Mol Biol 3:208–218. https://doi.org/10.1016/S0022-2836(61)80047-8.

9. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

10. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski 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.

11. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.

12. Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160.

13. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32:11–16. https://doi.org/10.1093/nar/gkh152.

14. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a Web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res 35:W52–W57. https://doi.org/10.1093/nar/gkm360.

15. 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.

16. Lütke-Brinkhaus F, Liedvogel B, Kreuz K, Kleinig H. 1982. Phytoene synthase and phytoene dehydrogenase associated with envelope membranes from spinach chloroplasts. Planta 156:176–180. https://doi.org/10.1007/BF00395433.

17. Moreno JC, Pizarro L, Fuentes P, Handford M, Cifuentes V, Stange C. 2013. Levels of lycopene β-cyclase 1 modulate carotenoid gene expression and accumulation in *Daucus carota*. PLoS One 8:e58144. https://doi.org/10.1371/journal.pone.0058144.

18. Gerber E, Bernard R, Castang S, Chabot N, Coze F, Dreux-Zigha A, Hauser E, Hivin P, Joseph P, Lazarelli C, Letellier G, Olive J, Leonetti JP. 2015. *Deinococcus* as new chassis for industrial biotechnology: biology, physiology and tools. J Appl Microbiol 119:1–10. https://doi.org/10.1111/jam.12808.

19. Agostini HJ, Carroll JD, Minton KW. 1996. Identification and characterization of uvrA, a DNA repair gene of *Deinococcus radiodurans*. J Bacteriol 178:6759–6765. https://doi.org/10.1128/jb.178.23.6759-6765.1996.

20. Xu G, Wang L, Chen H, Lu H, Ying N, Tian B, Hua Y. 2008. RecO is essential for DNA damage repair in *Deinococcus radiodurans*. J Bacteriol 190:2624–2628. https://doi.org/10.1128/JB.01851-07.

21. Schlesinger DJ. 2007. Role of RecA in DNA damage repair in *Deinococcus radiodurans*. FEMS Microbiol Lett 274:342–347. https://doi.org/10.1111/j.1574-6968.2007.00862.x.

22. Bentchikou E, Servant P, Coste G, Sommer S. 2010. A major role of the RecFOR pathway in DNA double-strand-break repair through ESDSA in *Deinococcus radiodurans*. PLoS Genet 6:e1000774. https://doi.org/10.1371/journal.pgen.1000774.

23. Servant P, Jolivet E, Bentchikou E, Mennecier S, Bailone A, Sommer S. 2007. The ClpPX protease is required for radioresistance and regulates cell division after γ-irradiation in *Deinococcus radiodurans*. Mol Microbiol 66:1231–1239. https://doi.org/10.1111/j.1365-2958.2007.06003.x.