Check for updates

Whole-Genome Shotgun (WGS) Sequence of *cis*-Isoprene Polymer-Degrading *Nocardia* sp. strain BSTN01

Biraj Sarkar,[a] ◎ Amit Kumar Mandal,[b] Amit Ghati,[c] Pranab Ghosh,[d] Sukhendu Mandal,[a] ◎ Ahmet Kati[e,f]

[a]Laboratory of Molecular Bacteriology, Department of Microbiology, University of Calcutta, Kolkata, India
[b]Chemical Biology Laboratory, Department of Sericulture, Raiganj University, North Dinajpur, West Bengal, India
[c]Department of Microbiology, Barrackpore Rastraguru Surendranath College, Kolkata, India
[d]Department of Chemistry, University of North Bengal, Darjeeling, West Bengal, India
[e]Experimental Medicine Research and Application Center, University of Health Sciences Turkey, Uskudar, Istanbul, Turkey
[f]Department of Biotechnology, Institution of Health Sciences, University of Health Sciences Turkey, Uskudar, Istanbul, Turkey

Biraj Sarkar and Amit Kumar Mandal contributed equally. The order of authors were determined on the basis of seniority.

ABSTRACT Species belonging to the genus *Nocardia* are known to be facultative human pathogens. There are also reports of *Nocardia* species capable of degrading various forms of rubber. Here, we report the whole-genome shotgun (WGS) sequence of *Nocardia* sp. strain BSTN01, isolated from stored water in latex-collecting cups thrown away near a local rubber processing unit in Tripura, India.

Species belonging to the genus *Nocardia* are aerobic, Gram-positive, nonmotile, filamentous actinomycetes and have been reported to be isolated from a wide range of environments (1–4). Numerous species belonging to the genus *Nocardia* have also been reported to be facultative intracellular human pathogens (1, 4–6). Studies have shown that some *Nocardia* spp. have the ability to degrade numerous toxic and harmful hydrocarbon-based materials. Different studies revealed that some *Nocardia* spp. harbor *lcp* genes, which are responsible for the initiation of the biodegradation of polymeric materials such as rubber (4, 7–10). The isolated strain BSTN01 was obtained from the stored water in latex-collecting cups (used to collect the freshly tapped latex from the *Hevea brasiliensis* trees) thrown away near a local rubber processing unit in Tripura, India. Initially, the collected water sample was serially diluted in sterile NaCl solution (0.9%), and 100 μL of the final dilution was spread on a sterile ISP-2 (International *Streptomyces* Project-2) agar plate (composition: malt extract, 10 g; yeast extract, 4 g; dextrose, 4 g; agar, 20 g, in 1 L of Milli-Q-grade autoclaved water, pH 7.0) supplemented with 50 μg/mL of both filter-sterilized cycloheximide and nystatin to check fungal contamination. The inoculated agar plate was then incubated for 4 to 5 days at 30°C. Thereafter, single colonies were individually picked and transferred to fresh ISP-2 plates using a sterile inoculating loop to obtain pure colonies using the streak plate method. The genome sequencing of the isolate *Nocardia* sp. strain BSTN01 was performed to explore the putative genes responsible for the rubber biodegradation and its possible pathogenicity (4).

The strain BSTN01 was grown in ISP-2 broth medium (composition: malt extract, 10 g; yeast extract, 4 g; dextrose, 4 g, in 1 L of Milli-Q-grade autoclaved water, pH 7.0) at 30°C under shaking conditions at 400 rpm for 4 to 5 days. The genomic DNA was isolated from freshly grown cells of BSTN01 as per the standard phenol:chloroform method (11, 12). The paired-end libraries were prepared and sequenced using the Illumina HiSeq X10 platform (AgriGenome Labs Pvt. Ltd., Kochi, Kerala, India), producing a total of 14,888,026 reads with 2 × 150-bp paired-end read length. The DNA library was prepared using the NEBNext Ultra DNA library prep kit according to the manufacturer's manual. To accomplish data preprocessing, unique reads were first fetched using BBTools v38.57 (https://sourceforge.net/projects/bbmap) (13). Adapter

Subsystem Coverage

Subsystem Category Distribution

Subsystem Feature Counts

- Cofactors, Vitamins, Prosthetic groups, Pigments (217)
- Cell wall and Capsule (29)
- Virulence, Disease and Defense (56)
- Potassium metabolism (9)
- Photosynthesis (0)
- Miscellaneous (42)
- Phages, Prophages, Transposable elements, Plasmids (17)
- Membrane transport (44)
- Iron acquisition and metabolism (10)
- RNA metabolism (63)
- Nucleosides and Nucleotides (94)
- Protein metabolism (190)
- Cell division and Cell Cycle (0)
- Motility and Chemotaxis (0)
- Regulation and Cell signaling (28)
- Secondary metabolism (5)
- DNA metabolism (108)
- Fatty acids, Lipids, and Isoprenoids (306)
- Nitrogen metabolism (34)
- Dormancy and Sporulation (4)
- Respiration (105)
- Stress response (63)
- Metabolism of Aromatic compounds (34)
- Amino acids and derivatives (403)
- Sulfur metabolism (33)
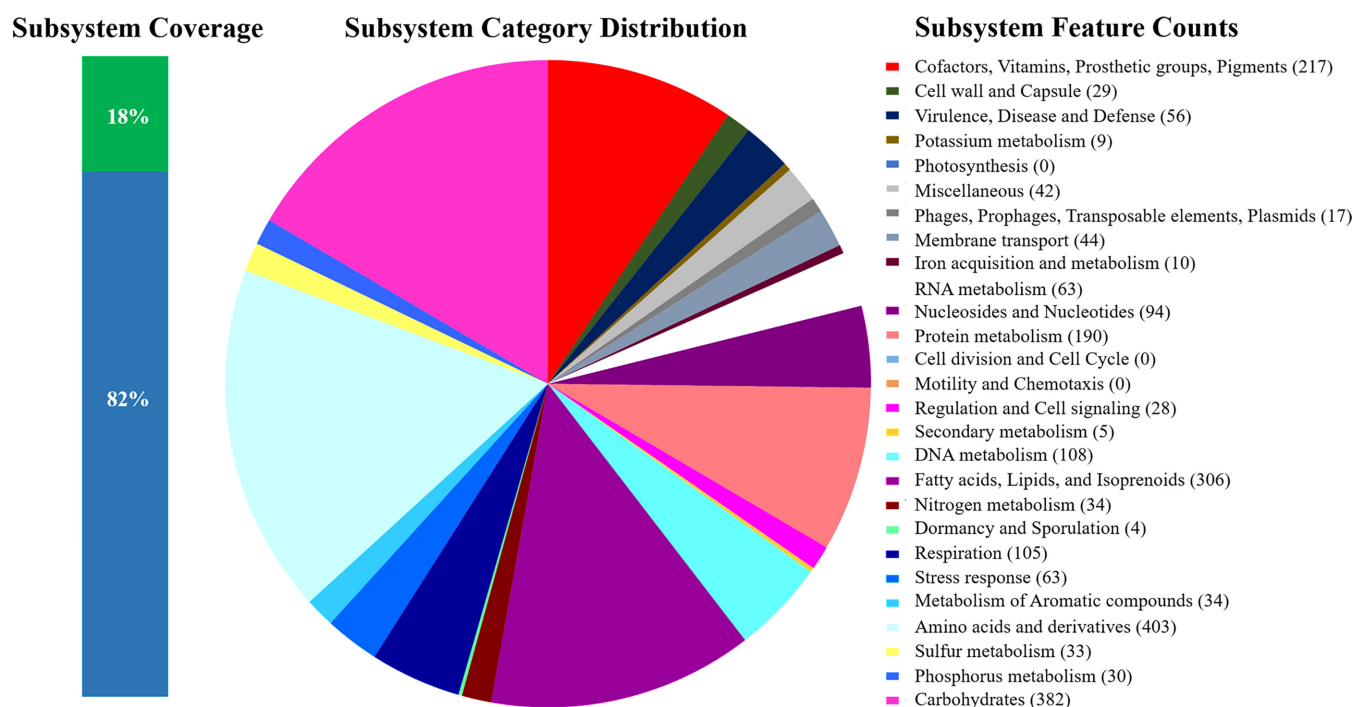- Phosphorus metabolism (30)
- Carbohydrates (382)

FIG 1 An overview of the subsystem categories assigned to the genome of *Nocardia* sp. BSTN01. The WGS sequence of the strain BSTN01 was annotated using the RAST server.

removal and low-quality end trimming was done using the Adapter Removal-v2 tool v2.3.1 (https://github.com/MikkelSchubert/adapterremoval) (14). The preprocessed reads were aligned with a plasmid sequence database (in-house database curated from NCBI RefSeq by AgriGenome Labs Pvt. Ltd.), and the unaligned reads were used for performing assembly. The *de novo* assembly was completed using the Unicycler v0.4.8 (https://github .com/rrwick/Unicycler) assembler (15). The annotation was carried out employing the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.13 with the methods "best-placed reference protein set" and "GeneMarkS-2+" (16). The assembly produced a draft genome sequence encompassing 208 contigs. The N_{50} length is 108,911 bp, and the L_{50} count is 20. The estimated genome size is 7,795,358 bp with a 67.7% G+C content and 96.3× coverage. A total of 7,253 coding sequences were annotated, including 6 rRNA genes (2 16S and 4 23S) and 49 tRNAs. The WGS sequence of BSTN01 was also later uploaded for rapid annotations in the Rapid Annotations using Subsystems Technology (RAST) server (17). For all the software tools used, default parameters were applied unless otherwise specified. The total number of coding sequences predicted was 7,702, and the number of predicted RNAs was 51. An overview of the subsystem categories assigned to the genome of *Nocardia* sp. BSTN01 is shown in Fig. 1. The WGS analysis revealed the presence of the *lcp* gene (responsible for the initiation of the biodegradation process of rubber) along with other virulence genes (e.g., *mce*, *SodA*, *KatG*, *ahpC*, etc.) (4). Further insight into the genome of BSTN01 will undoubtedly contribute toward the molecular basis of rubber biodegradation for solid-polymeric-rubber waste management.

Data availability. This whole-genome shotgun project has been deposited at NCBI under the accession number JADKYP000000000. The version described in this paper is the first version, JADKYP010000000. The BioSample and BioProject accession numbers are SAMN16604275 and PRJNA673320, respectively. The raw data are available from the Sequence Read Archive (SRA) under the accession number SRR17194475.

REFERENCES

1. Bafghi MF, Eshraghi SS, Heidarieh P, Habibnia S, Nasab MR. 2014. Nocardiosis in immune disorder disease. Malays J Med Sci 21:75–76.

2. Dhakal D, Kumar Jha A, Pokhrel A, Shrestha A, Sohng JK. 2016. Genetic manipulation of *Nocardia* species. Curr Protoc Microbiol 40:10F.2.1–10F.2.18. https://doi.org/10.1002/9780471729259.mc10f02s40.

3. Shariff M, Gunasekaran J. 2016. Pulmonary nocardiosis: review of cases and an update. Can Respir J 2016:7494202. https://doi.org/10.1155/2016/7494202.

4. Sarkar B, Gupta AM, Mandal S. 2021. Insights from the comparative genome analysis of natural rubber degrading *Nocardia* species. Bioinformation 17:880–890. https://doi.org/10.6026/97320630017880.

5. Hashemi-Shahraki A, Heidarieh P, Bostanabad SZ, Hashemzadeh M, Feizabadi MM, Schraufnagel D, Mirsaeidi M. 2015. Genetic diversity and antimicrobial susceptibility of *Nocardia* species among patients with nocardiosis. Sci Rep 5:17862. https://doi.org/10.1038/srep17862.

6. Fellag M, Levasseur A, Delerce J, Bittar F, Marié JL, Davoust B, Drancourt M. 2018. Draft genome sequence of "*Nocardia suismassiliense*" strain S-137 (CSUR P4007). Genome Announc 6:e00212-18. https://doi.org/10.1128/genomeA.00212-18.

7. Tsuchii A, Takeda K, Tokiwa Y. 1996. Degradation of the rubber in truck tires by a strain of *Nocardia*. Biodegradation 7:405–413. https://doi.org/10.1007/BF00056424.

8. Luo Q, Hiessl S, Poehlein A, Daniel R, Steinbüchel A. 2014. Insights into the microbial degradation of rubber and gutta-percha by analysis of the complete genome of *Nocardia nova* SH22a. Appl Environ Microbiol 80:3895–3907. https://doi.org/10.1128/AEM.00473-14.

9. Linh DV, Huong NL, Tabata M, Imai S, Iijima S, Kasai D, Anh TK, Fukuda M. 2017. Characterization and functional expression of a rubber degradation gene of a *Nocardia* degrader from a rubber-processing factory. J Biosci Bioeng 123:412–418. https://doi.org/10.1016/j.jbiosc.2016.11.012.

10. Rodrigues EM, Vidigal P, Pylro VS, Morais DK, Leite LR, Roesch L, Tótola MR. 2017. Draft genome of *Nocardia farcinica* TRH1, a linear and polycyclic aromatic hydrocarbon-degrading bacterium isolated from the coast of Trindade Island, Brazil. Braz J Microbiol 48:391–392. https://doi.org/10.1016/j.bjm.2016.09.014.

11. 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.

12. Maiti PK, Mandal S. 2020. *Lentzea indica* sp. nov., a novel actinobacteria isolated from Indian Himalayan-soil. Antonie Van Leeuwenhoek 113:1411–1423. https://doi.org/10.1007/s10482-020-01449-8.

13. Bushnell B. BBMap. sourceforge.net/projects/bbmap/.

14. Schubert M, Lindgreen S, Orlando L. 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res Notes 9:88. https://doi.org/10.1186/s13104-016-1900-2.

15. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.

16. 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.

17. 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.