



## Genome Announcements

# Draft genome of *Nocardia farcinica* TRH1, a linear and polycyclic aromatic hydrocarbon-degrading bacterium isolated from the coast of Trindade Island, Brazil

Edmo M. Rodrigues<sup>a,\*</sup>, Pedro Marcus P. Vidigal<sup>b</sup>, Victor S. Pylro<sup>c</sup>, Daniel K. Morais<sup>c</sup>, Laura Rabelo Leite<sup>c</sup>, Luiz F.W. Roesch<sup>d</sup>, Marcos R. Tótola<sup>a</sup>

<sup>a</sup> Laboratório de Biotecnologia e Biodiversidade para o Meio Ambiente, Departamento de Microbiologia, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil

<sup>b</sup> Núcleo de Análise de Biomoléculas (NuBioMol), Centro de Ciências Biológicas, Universidade Federal de Viçosa, Minas Gerais, Brazil

<sup>c</sup> Biosystems Informatics and Genomics Group, René Rachou Research Centre – CPpRR/FIOCRUZ, Belo Horizonte, Minas Gerais, Brazil

<sup>d</sup> Universidade Federal do Pampa, Campus São Gabriel, São Gabriel, Rio Grande do Sul, Brazil

## ARTICLE INFO

## Article history:

Received 4 June 2016

Accepted 18 September 2016

Available online 10 December 2016

Associate Editor: John McCulloch

## Keywords:

Hydrocarbon bioremediation

Pristine environment

PAH biodegradation

## ABSTRACT

Here, we report the draft genome sequence and annotation of *Nocardia farcinica* TRH1, a petroleum hydrocarbons degrading Actinobacteria isolated from the coastal water of Trindade Island, Brazil.

© 2016 Sociedade Brasileira de Microbiologia. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Several species of the genus *Nocardia* are known to degrade hydrocarbons.<sup>1</sup> The strain *Nocardia farcinica* TRH1 was isolated from the coastal water of Trindade Island, a pristine oceanic island in Brazil.<sup>2</sup> It is capable of growing using several petroleum hydrocarbons as the sole source of carbon and energy, such as phenanthrene, pyrene, anthracene, eicosane, pentacontane, triacontane, tetracosane, naphthalene, hexadecane, octane, toluene and xylene.<sup>2</sup>

Genome sequencing of *N. farcinica* TRH1 was performed using the Ion Torrent PGM platform (ThermoFisher Scientific). Briefly, the genomic DNA was fragmented using the Bioruptor UCD-200. The template library was prepared with the Ion Plus fragment library kit and clonally amplified in the One Touch System with the Ion PGM template OT2 400 kit. The amplified library was sequenced using the Ion PGM sequencing 400 kit within the 318 v2 microchip. A total of

\* Corresponding author.

E-mails: [edmomontes@yahoo.com.br](mailto:edmomontes@yahoo.com.br) (E.M. Rodrigues), [totola@ufv.br](mailto:totola@ufv.br) (M.R. Tótola).

<http://dx.doi.org/10.1016/j.bjm.2016.09.014>

1517-8382/© 2016 Sociedade Brasileira de Microbiologia. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

2531733 reads were obtained with sizes ranging from 25 to 492 bp in length. The reads were filtered for length (minimum, 100 bp) and quality (minimum score, Q20) and used for *de novo* assembling using CLC Genomics Workbench version 6.5.1 (CLC bio). From assembling, we obtained 321 contigs, corresponding to 5230013 bp, with an average size ( $N_{50}$ ) of 9853 bp, longest contig size of 122221 bp, G+C content of 68.0% and genome coverage of 52.53X. Genes from the contigs were predicted using GeneMarkS,<sup>3</sup> which revealed 4946 coding sequence set (CDS). The protein sets were functionally annotated using BLAST (<http://blast.ncbi.nlm.nih.gov/>), and approximately 68.6% of the proteins were assigned to Clusters of Orthologous Groups (COG) families.<sup>4</sup>

Genome annotation was performed using BlastKOALA<sup>5</sup> and revealed 1869 protein-coding-sequences, including 106 related with xenobiotics biodegradation and metabolism, and 309 unclassified. The KEGG Automatic Annotation Server (KASS)<sup>6</sup> was used for pathways analysis, which identified 516 genes related to metabolic pathways, including 21 genes related to biodegradation of aromatic compounds.

KASS also identified the presence of *catA* and *dmpC*, which are among the genes of the catabolic pathways of ortho and meta-cleavage of catechol, respectively. Catechol is a toxic intermediate generated during the biodegradation of polycyclic aromatic hydrocarbons.<sup>7</sup> Furthermore, several genes were identified as being related to biodegradation of aliphatic hydrocarbons as the *alkB* gene. We also found *aldH*, *paaF*, *fadB* and *fadJ* genes related to biodegradation of the nylon precursor caprolactam, as well as the *atzD* gene, related to the degradation of the pesticide atrazine, showing another possibility for biotechnological applications of this strain.

The abundance of genes involved in biodegradation pathways in the genome of *N. farcinica* TRH1 implies in high metabolic plasticity of this strain, what is consistent with the results obtained during the screening of bacteria for hydrocarbon biodegradation.<sup>2</sup> The genome sequencing data from this study will support a better understanding of the metabolism and the potential applications of *N. farcinica* TRH1 in biotechnological processes, as hydrocarbons and xenobiotics bioremediation.

Nucleotide sequence accession numbers: This WGS Bio-Project has been deposited at DDBJ/EMBL/GenBank under the accession number PRJNA322144 and the sequences under the accession number LYCQ00000000. The versions described in this paper are the first versions.

---

## Conflicts of interest

The authors declare no conflicts of interest.

---

## Acknowledgments

We thank the Brazilian Navy and Rodrigo Otoch Chaves for logistic support while collecting samples and for providing the essential structure to transport and store samples. CNPq grant 405544/2012-0 (PROTRINDADE), FAPEMIG, and CAPES (PROEX) funded this work. This work is also supported by the Brazilian Microbiome Project (<http://brmicrobiome.org>).

---

## REFERENCES

1. Lara-Severino RC, Camcho-López MA, García-Macedo JM, et al. Determination of the residual anthracene concentration in cultures of haloalkalitolerant actinomycetes by excitation fluorescence, emission fluorescence, and synchronous fluorescence: comparative study. *J Anal Methods Chem.* 2016; <http://dx.doi.org/10.1155/2016/6287931>.
2. Rodrigues EM, Kalks KHM, Tótola MR. Prospect, isolation, and characterization of microorganisms for potential use in cases of oil bioremediation along the coast of Trindade Island, Brazil. *J Environ Manag.* 2015;156:15–22.
3. Besemer J, Lomsadze A, Borodovsky M. GeneMark S: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 2001;29:2607–2618.
4. Tatusov RL, Natale DA, Garkavtsev IV, et al. The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* 2001;29:22–28.
5. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol.* 2016;428:726–731.
6. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. The metabolic pathways were examined through KAAS (KEGG automatic annotation server): an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 2007;35:182–185.
7. Schweigert N, Zehnder AJ, Eggen RI. Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals. *Environ Microbiol.* 2001;3:81–91.