



First Draft Genome Sequence of a Clinical Strain of *Nocardia cerradoensis*

Gema Carrasco,^a Sara Monzón,^b Pilar Jiménez,^c Isabel Cuesta,^b Joaquín Bartolomé-Alvarez,^d  Sylvia Valdezate^a

Reference and Research Laboratory for Taxonomy, Department of Bacteriology, National Centre of Microbiology, Instituto de Salud Carlos III, Majadahonda, Madrid, Spain^a; Laboratory of Bioinformatics, Common Scientific Technical Units, Instituto de Salud Carlos III, Majadahonda, Madrid, Spain^b; Genomics Laboratory, S. G. for Applied Services, Training and Research, Instituto de Salud Carlos III, Majadahonda, Madrid, Spain^c; Microbiology Department, University Hospital of Albacete, Albacete, Spain^d

ABSTRACT This paper reports the first draft genome sequence for a strain of *Nocardia cerradoensis* obtained from an immunocompetent patient with a knee infection. The 8.2-Mb genome has 8,329 coding sequences, including intrinsic resistance genes, biosynthetic gene clusters for polyketide synthase and nonribosomal peptide synthase, virulence genes, and prophages.

Nocardia species are environmental bacteria that cause severe opportunistic infections of the lung, skin, and central nervous system, although mainly in immunocompromised patients. Of 110 recognized species of this genus (<http://www.bacterio.net/>), more than half are described as human pathogens. *Nocardia cerradoensis* was initially described in Brazilian soil (1); as a pathogen it was first noticed causing disseminated infection in a patient who had received a kidney graft (2). Although *N. cerradoensis* has been proposed as an emerging global pathogen (3), no genome of an *N. cerradoensis* clinical strain has previously been sequenced. Here, we describe the draft genome sequence of *N. cerradoensis* CNM20130759, isolated from a postinfiltration cutaneous infection on the knee of a 52-year-old immunocompetent woman (4). The antimicrobial profile of this isolate fitted that described for the *Nocardia nova* complex pattern (5). The 16S rRNA and *gyrB* genes, and the multilocus sequence type (6), were most similar to those of *N. cerradoensis* W9747^T (99.9%, 99.5%, and 99.6% similarities, respectively).

A paired-end library (Nextera-XT DNA library preparation kit, Illumina 1.9) was adapted and sequenced at 2 × 150 in the Illumina NextSeq 500 platform. After quality control analysis using fastQC v.0.11.3 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and Trimmomatic v.0.36 software (7), ~4.9 million paired-end reads of 50 to 151 bp were retrieved. Assembly of the reads was performed using SPAdes v.3.8.0 (8), and the quality of the assembly was evaluated using QUAST software (9).

The genome was found to be 8,193,765-bp long, with a G+C content of 68.16% (150 contigs of ≥500 bp; maximum, 1,035,523 bp; *N*₅₀ value, 394,043 bp). The assembled sequences were annotated using Prokka 1.12-beta software (10), which predicted 8,406 genes (8,329 protein-coding sequences, 5 rRNA genes, 1 transfer-messenger RNA [tmRNA] gene, and 71 tRNA genes) and 543 signal peptides. Among the protein-coding genes, 2,060 were unique and 704 repeated 2 to 41 times. Multiple antibiotic resistance and efflux pump genes [*acc(6')*-I, *blal*, *rrrA*, *rrrB*, *rrrC*, *emrY*, *inhA*, *mdtG*, *mdtK*, *mdtH*, *mdtL*, *mmr*, NCRCNM_01156, NCRCNM_01157, NCRCNM_01690, NCRCNM_01779, NCR CNM_06233, NCRCNM_06526, *norM*, *stp*, *tetR*, and *yfmO*, among others], antiseptic genes (*sugE* and *qacA*), metalloproteinase genes (*tldD*, *yflN*, *ycnJ*, and *merR*), and integrase genes (Int-Tn 1 to 3) were detected, but no clustered regularly interspaced

Received 28 April 2017 Accepted 11 August 2017 Published 28 September 2017

Citation Carrasco G, Monzón S, Jiménez P, Cuesta I, Bartolomé-Alvarez J, Valdezate S. 2017. First draft genome sequence of a clinical strain of *Nocardia cerradoensis*. *Genome Announc* 5:e00551-17. <https://doi.org/10.1128/genomeA.00551-17>.

Copyright © 2017 Carrasco 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 Sylvia Valdezate, svaldezate@isciii.es.

short palindromic repeats (CRISPR) were found. One intact prophage (with *attL/attR* sites) and five incomplete prophages were identified by PHAST software (11), representing 0.84% of the genome. Neither SRST2 (12) nor ARGannot (13) software detected any plasmid or acquired resistance genes.

Virulence genes for *mce2* (mammalian cell invasion), hemolysin, superoxide dismutase, catalase-peroxidase, cutinase, esterases, hemagglutinin, the antigen 85 complex family, ESX-1, and sortase A (among others) were all detected; some of these have been previously described for *Nocardia farcinica* (14). Biosynthetic gene clusters such as 12 nonribosomal peptide synthases (NRPS), 8 polyketide synthases (PKS-I-III), 4 terpene cyclases, 6 NRPS, and 2 unique PKS-I genes were found using antiSMASH software (15).

This clinical strain showed 99.4% similarity, by ANI software (16), to the *N. cerra-doensis* W9747^T draft genome (a soil bacterium), which has a length of 7.6 Mbp, a G+C content of 68.16%, 7,216 protein-coding genes (8 rRNAs), and 1 CRISPR. Increased understanding of the differences between the human and environmental strains expands our knowledge of *N. cerra-doensis*.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number [NGAF00000000](https://doi.org/10.1093/nar/gkr485). The version described in this paper is version NGAF01000000.

ACKNOWLEDGMENTS

We are grateful to N. Garrido, J. A. Saéz-Nieto, M. D. Pérez-Vázquez, and A. Zaballos for their help in the different stages of the genome sequencing and to Adrian Burton for language and editing assistance.

This work was funded by project MPY1278/15 of the Instituto de Salud Carlos III.

REFERENCES

- Albuquerque de Barros EV, Manfio GP, Ribiero Maitan V, Mendes Bataus LA, Kim SB, Maldonado LA, Goodfellow M. 2003. *Nocardia cerra-doensis* sp. nov., a novel isolate from Cerrado soil in Brazil. *Int J Syst Evol Microbiol* 53:29–33. <https://doi.org/10.1099/ijs.0.02020-0>.
- Piau C, Kerjouan M, Le Mouel M, Patrat-Delon S, Henaux PL, Brun V, Morin MP, Gautier P, Rodriguez-Nava V, Kayal S. 2015. First case of disseminated infection with *Nocardia cerra-doensis* in a human. *J Clin Microbiol* 53:1034–1037. <https://doi.org/10.1128/JCM.02979-14>.
- Ercibengoa M, Pérez-Trallero E, Marimón JM. 2016. Autochthonous *Nocardia cerra-doensis* infection in humans, Spain, 2011 and 2014. *Emerg Infect Dis* 22:109–111. <https://doi.org/10.3201/eid2201.150771>.
- Escudero-Jiménez Á, Saéz-Nieto JA, Turowicz M, Bartolomé-Álvarez J. 2016. Infección cutánea por *Nocardia cerra-doensis* tras infiltración. *Enferm Infecc Microbiol Clin* 34:325–326. <https://doi.org/10.1016/j.eimc.2015.07.015>.
- CLSI. 2011. Susceptibility testing of *Mycobacteria*, *Nocardia*, and other aerobic actinomycetes, 2nd ed, vol 31, no. 5. Approved standard M24-A2. National Committee for Clinical Laboratory Standards, Wayne, PA.
- McTaggart LR, Richardson SE, Witkowska M, Zhang SX. 2010. Phylogeny and identification of *Nocardia* species on the basis of multilocus sequence analysis. *J Clin Microbiol* 48:4525–4533. <https://doi.org/10.1128/JCM.00883-10>.
- 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>.
- Bankovich 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>.
- 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>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. *Nucleic Acids Res* 39:W347–W352. <https://doi.org/10.1093/nar/gkr485>.
- Inouye M, Dashnow H, Raven LA, Schultz MB, Pope BJ, Tomita T, Zolber J, Holt KE. 2014. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med* 6:90. <https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-014-0090-6>.
- Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, Rolain JM. 2014. ARG-Annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother* 58:212–220. <https://doi.org/10.1128/AAC.01310-13>.
- Ishikawa J, Yamashita A, Mikami Y, Hoshino Y, Kurita H, Hotta K, Shiba T, Hattori M. 2004. The complete genomic sequence of *Nocardia farcinica* IFM 10152. *Proc Natl Acad Sci U S A* 101:14925–14930. <https://doi.org/10.1073/pnas.0406410101>.
- Blin K, Medema MH, Kottmann R, Lee SY, Weber T. 2017. The antiSMASH database, a comprehensive database of microbial secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res* 45:D555–D559. <https://doi.org/10.1093/nar/gkw960>.
- Zhang W, Du P, Zheng H, Yu W, Wan L, Chen C. 2014. Whole-genome sequence comparison as a method for improving bacterial species definition. *J Gen Appl Microbiol* 60:75–78. <https://doi.org/10.2323/jgam.60.75>.