

# Draft Genome Sequence of *Frankia* sp. Strain CN3, an Atypical, Noninfective (Nod<sup>-</sup>) Ineffective (Fix<sup>-</sup>) Isolate from *Coriaria nepalensis*

Faten Ghodhbane-Gtari,<sup>a,b</sup> Nicholas Beauchemin,<sup>a</sup> David Bruce,<sup>c</sup> Patrick Chain,<sup>c</sup> Amy Chen,<sup>d</sup> Karen Walston Davenport,<sup>c</sup> Shweta Deshpande,<sup>d</sup> Chris Detter,<sup>c</sup> Teal Furnholm,<sup>a</sup> Lynne Goodwin,<sup>c</sup> Maher Gtari,<sup>a,b</sup> Cliff Han,<sup>c</sup> James Han,<sup>d</sup> Marcel Huntemann,<sup>d</sup> Natalia Ivanova,<sup>d</sup> Nikos Kyrpides,<sup>d</sup> Miriam L. Land,<sup>e</sup> Victor Markowitz,<sup>d</sup> Kostas Mavrommatis,<sup>d</sup> Matt Nolan,<sup>d</sup> Imen Nouioui,<sup>b</sup> Ioanna Pagani,<sup>d</sup> Amrita Pati,<sup>d</sup> Sam Pitluck,<sup>d</sup> Catarina L. Santos,<sup>f</sup> Arnab Sen,<sup>g</sup> Saubashya Sur,<sup>g</sup> Ernest Szeto,<sup>d</sup> Fernando Tavares,<sup>h</sup> Hazuki Teshima,<sup>c</sup> Subarna Thakur,<sup>g</sup> Luis Wall,<sup>i</sup> Tanja Woyke,<sup>d</sup> Louis S. Tisa<sup>a</sup>

University of New Hampshire, Durham, New Hampshire, USA<sup>a</sup>; University of Tunis El Manar, Tunis, Tunisia<sup>b</sup>; Los Alamos National Laboratory, Los Alamos, New Mexico, USA<sup>c</sup>; DOE, Joint Genome Institute, Walnut Creek, California, USA<sup>d</sup>; Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA<sup>e</sup>; Centro de Investigação em Biodiversidade e Recursos Genéticos, Porto, Portugal<sup>f</sup>; University of North Bengal, Siliguri, India<sup>g</sup>; Instituto de Biologia Celular e Molecular, Porto, Portugal<sup>h</sup>; University of Quilmes, Bernal, Argentina<sup>i</sup>

**We report here the genome sequence of *Frankia* sp. strain CN3, which was isolated from *Coriaria nepalensis*. This genome sequence is the first from the fourth lineage of *Frankia*, strains of which are unable to reinfect actinorhizal plants. At 10 Mb, it represents the largest *Frankia* genome sequenced to date.**

Received 6 February 2013 Accepted 14 February 2013 Published 14 March 2013

**Citation** Ghodhbane-Gtari F, Beauchemin N, Bruce D, Chain P, Chen A, Walston Davenport K, Deshpande S, Detter C, Furnholm T, Goodwin L, Gtari M, Han C, Han J, Huntemann M, Ivanova N, Kyrpides N, Land ML, Markowitz V, Mavrommatis K, Nolan M, Nouioui I, Pagani I, Pati A, Pitluck S, Santos CL, Sen A, Sur S, Szeto E, Tavares F, Teshima H, Thakur S, Wall L, Woyke T, Tisa LS. 2013. Draft genome sequence of *Frankia* sp. strain CN3, an atypical, noninfective (Nod<sup>-</sup>) ineffective (Fix<sup>-</sup>) isolate from *Coriaria nepalensis*. *Genome Announc.* 1(2):e00085-13. doi:10.1128/genomeA.00085-13.

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

Address correspondence to Louis S. Tisa, louis.tisa@unh.edu.

Among the *Actinobacteria*, the genus *Frankia* is well known for its facultative lifestyle as a plant symbiont of dicotyledonous plants and as a free-living soil dweller (1–3). Our understanding of this genus has been greatly enhanced by the sequencing of several *Frankia* genomes representing three different *Frankia* lineages (4–6). Besides these three major lineages of *Frankia* that are known to reinfect their host plants, a fourth *Frankia* lineage (termed atypical *Frankia* isolates) is recognized that represents isolates that are unable to reinfect actinorhizal plants. Although this group forms a clade within *Frankia*, it is phylogenetically distinct from the other three lineages based on several criteria, including the 16S rRNA gene (7), *glmII* (8, 9), and a 16S–23S rRNA intergenic spacer region (10).

*Frankia* sp. strain CN3 was chosen for sequencing and is a member of the fourth *Frankia* lineage and an atypical *Frankia* strain that is noninfective (Nod<sup>-</sup>) and non-nitrogen fixing (Fix<sup>-</sup>). The bacterium was isolated from root nodules obtained on soil samples collected from Muree, Northern Pakistan, using *Coriaria nepalensis* seedlings (11). *Frankia* sp. strain CN3 is also resistant to elevated levels of toxic heavy metals, including Cu<sup>+2</sup> (12), and has potential applications in bioremediation. Strain CN3 was sequenced to provide information about the potential ecological roles of the atypical *Frankia* strains and their interaction with actinorhizal plants. Since symbiotic interactions between *Frankia* and the host plant are not well understood at a molecular level, comparative genomics of atypical and infective *Frankia* strains should provide clues on their symbiotic lifestyle.

The draft genome of *Frankia* sp. strain CN3 was generated at the Department of Energy (DOE) Joint Genome Institute (JGI) using a combination of 454-GS-FLX-Titanium (13) and Illumina

GAii (14) techniques. A standard 454 Titanium library, which generated 918,729 reads, a paired-end 454 library with an average insert size of 8 kb, which generated 756,574 reads totaling 511.5 Mb of 454 data, and an Illumina GAii shotgun library, which generated 65,924,070 reads totaling 5,010.2 Mb, were created. All techniques for DNA isolation, library construction, and sequencing were performed according to JGI standards and protocols (<http://www.jgi.doe.gov>). The 454 Titanium standard data and the 454 paired-end data were assembled together with Newbler, version 2.3 (13), and resulted in 379.9 Mb of 454 draft data, which provided an average 38× coverage of the genome. The Illumina sequencing data were assembled with Velvet, version 1.0.13 (15), and the resulting 4,476.3 Mb of Illumina draft data provided an average 447.6× coverage of the genome. For finishing, the gaps and misassemblies were resolved by editing in Consed, by PCR, and by bubble PCR primer walks.

The draft genome of *Frankia* strain CN3 was resolved to 2 scaffolds consisting of 9,978,592 bp with a G+C content of 71.81%, 8,333 candidate protein-encoding genes, 79 tRNA genes, and 3 rRNA regions. Strain CN3 represents the largest *Frankia* genome sequenced to date.

**Nucleotide sequence accession numbers.** The *Frankia* sp. strain CN3 genome sequence has been deposited at DDBJ/EMBL/GenBank under the accession number [AGJN000000000](https://www.ncbi.nlm.nih.gov/nuccore/AGJN000000000). The version described in this paper is the first version, [AGJN010000000](https://www.ncbi.nlm.nih.gov/nuccore/AGJN010000000).

## ACKNOWLEDGMENTS

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231. This project (L.S.T.)

was supported in part by Agriculture and Food Research Initiative grant 2010-65108-20581 from the USDA National Institute of Food and Agriculture, Hatch grant NH530, and the College of Life Sciences and Agriculture at the University of New Hampshire, Durham. M.G. and F.G.-G. were supported in part by a Visiting Scientist and Postdoctoral Scientist Program administered by the NH AES at the University of New Hampshire.

## REFERENCES

- Benson DR, Silvester WB. 1993. Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. *Microbiol. Rev.* 57:293–319.
- Schwencke J, Caru M. 2001. Advances in actinorhizal symbiosis: Host plant-*Frankia* interactions, biology, and applications in arid land reclamation. A review. *Arid Land Res. Manag.* 15:285–327.
- Chaia EE, Wall LG, Huss-Danell K. 2010. Life in soil by actinorhizal root nodule endophyte *Frankia*. A review. *Symbiosis* 51:201–226.
- Normand P, Lapierre P, Tisa LS, Gogarten JP, Alloisio N, Bagnarol E, Bassi CA, Berry AM, Bickhart DM, Choisine N, Couloux A, Cournoyer B, Cruveiller S, Daubin V, Demange N, Francino MP, Goltsman E, Huang Y, Kopp OR, Labarre L, Lapidus A, Lavire C, Marechal J, Martinez M, Mastrorunzio JE, Mullin BC, Niemann J, Pujic P, Rawnsley T, Rouy Z, Schenowitz C, Sellstedt A, Tavares F, Tomkins JP, Vallet D, Valverde C, Wall LG, Wang Y, Medigue C, Benson DR. 2007. Genome characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant biogeography. *Genome Res.* 17:7–15.
- Normand P, Queiroux C, Tisa LS, Benson DR, Rouy Z, Cruveiller S, Medigue C. 2007. Exploring the genomes of *Frankia*. *Physiol. Plant.* 130:331–343.
- Persson T, Benson DR, Normand P, Vanden Heuvel B, Pujic P, Chertkov O, Teshima H, Bruce DC, Detter C, Tapia R, Han S, Han J, Woyke T, Pitlock S, Pennacchio L, Nolan M, Ivanova N, Pati A, Land ML, Pawlowski K, Berry AM. 2011. Genome sequence of “*Candidatus* *Frankia* *datiscacae*” Dg1, the uncultured Microsymbiont from nitrogen-fixing root nodules of the dicot *Datisca glomerata*. *J. Bacteriol.* 193:7017–7018.
- Normand P, Orso S, Cournoyer B, Jeannin P, Chapelon C, Dawson JO, Evtushenko L, Mirsra AK. 1996. Molecular phylogeny of the genus *Frankia* and related genera and emendation of the family *Frankiaceae*. *Int. J. Syst. Bacteriol.* 46:1–9.
- Cournoyer B, Lavire C. 1999. Analysis of *Frankia* evolution radiation using *glnII* sequences. *FEMS Microbiol. Lett.* 177:29–34.
- Nouioui I, Ghodhbane-Gtari F, Beauchemin NJ, Tisa LS, Gtari M. 2011. Phylogeny of members of the *Frankia* genus based on *gyrB*, *nifH* and *glnII* sequences. *Antonie Van Leeuwenhoek* 100:579–587.
- Ghodhbane-Gtari F, Nouioui I, Chair M, Boudabous A, Gtari M. 2010. 16S-23S rRNA intergenic spacer region variability in the genus *Frankia*. *Microb. Ecol.* 60:487–495.
- Mirza MS, Hahn D, Akkermans ADL. 1992. Isolation and characterization of *Frankia* strains from *Coriaria nepalensis*. *System. Appl. Microbiol.* 15:289–295.
- Richards JW, Krumholz GD, Chval MS, Tisa LS. 2002. Heavy metal resistance patterns of *Frankia* strains. *Appl. Environ. Microbiol.* 68:923–927.
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLI, Jarvie TP, Jirage KB, Kim J-B, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Mc Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasza A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–380.
- Bennett S. 2004. Solexa Ltd. *Pharmacogenomics* 5:433–438.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829.