**Title**

Draft Genome Sequence of Frankia sp. Strain DC12, an Atypical, Noninfective, Ineffective Isolate from Datisca cannabina

**Permalink**

https://escholarship.org/uc/item/8sh0q2mk

**Journal**

Microbiology Resource Announcements, 3(4)

**ISSN**

2576-098X

**Authors**

Tisa, Louis S
Beauchemin, Nicholas
Cantor, Michael N
et al.

**Publication Date**

2015-08-27

**DOI**

10.1128/genomea.00889-15

Peer reviewed

# Draft Genome Sequence of *Frankia* sp. Strain DC12, an Atypical, Noninfective, Ineffective Isolate from *Datisca cannabina*

Louis S. Tisa,[a] Nicholas Beauchemin,[a] Michael N. Cantor,[b] Teal Furnholm,[a] Faten Ghodhbane-Gtari,[a,c] Lynne Goodwin,[b] Alex Copeland,[d] Maher Gtari,[a,c] Marcel Huntemann,[d] Natalia Ivanova,[d] Nikos Kyrpides,[d] Victor Markowitz,[d] Kostas Mavrommatis,[d] Natalia Mikhailova,[d] Imen Nouioui,[c] Rediet Oshone,[a] Galina Ovchinnikova,[d] Ioanna Pagani,[d] Krishnaveni Palaniappan,[d] Amrita Pati,[d] Arnab Sen,[e] Nicole Shapiro,[d] Ernest Szeto,[d] Luis Wall,[f] Jessie Wishart,[a] Tanja Woyke[d]

University of New Hampshire, Durham, New Hampshire, USA[a]; Los Alamos National Laboratory, Los Alamos, New Mexico, USA[b]; University of Tunis-El Manar, El Manar, Tunisia[c]; DOE Joint Genome Institute, Walnut Creek, California, USA[d]; University of North Bengal, Siliguri, India[e]; University of Quilmes, Buenos Aires, Argentina[f]

This article is scientific contribution number 2612 from the NH Agricultural Experiment Station.

***Frankia* sp. strain DC12, isolated from root nodules of *Datisca cannabina*, is a member of the fourth lineage of *Frankia*, which is unable to reinfect actinorhizal plants. Here, we report its 6.88-Mbp high-quality draft genome sequence, with a G+C content of 71.92% and 5,858 candidate protein-coding genes.**

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

*F*rankia spp. are well known as plant symbionts of dicotyledonous plants and also are found as free-living soil dwellers (1–3). This genus has not yet been described to the species level, but it has become an area of recent interest. Four major *Frankia* lineages have been identified (4–7). Three of them are known to reinfect their host plant, while the fourth lineage (termed atypical *Frankia* isolates) are unable to reinfect actinorhizal plants or will reinfect the host plant but form ineffective nodules. Our understanding of this genus has been greatly enhanced by the sequencing of several *Frankia* genomes from the different *Frankia* lineages (8–18).

As a member of the fourth *Frankia* lineage, *Frankia* sp. strain DC12 was chosen for sequencing. This atypical noninfective (Nod⁻) and non-nitrogen-fixing (Fix⁻) *Frankia* strain was isolated from root nodules of *Datisca cannabina* L. collected from Swat, Pakistan (19, 20). Strain DC12 is also resistant to elevated levels of toxic heavy metals (21), and the spores germinate well under controlled conditions, enabling single genomic units to be isolated (22). Strain DC12 was sequenced to provide information about the potential ecological roles of the atypical *Frankia* strains and interaction with actinorhizal plants.

The draft genome of *Frankia* sp. strain DC12 was generated at the Department of Energy (DOE) Joint Genome Institute (JGI) using Illumina data (23). An Illumina short-insert paired-end library with an average (± standard deviation) insert size of 242 ± 59 bp, which generated 16,229,834 reads, and an Illumina long-insert paired-end library with an average insert size of 6,525 ± 1,400 bp, which generated 20,981,340 reads totaling 4,533 Mbp of Illumina data, were generated and sequenced. All techniques for DNA isolation, library construction, and sequencing were performed according to JGI standards and protocols (http://www.jgi.doe.gov). The Illumina sequencing data were assembled with Velvet version 1.0.13 (24) and AllPaths version r41043 (25). The final draft assembly contained 12 contigs in 1 scaffold. The total size of the genome is 6.88 Mbp, and the final assembly is based on 4,533 Mbp of Illumina draft data, which provides an average 657× coverage of the genome. For finishing, the gaps and misassemblies were resolved by editing in Consed, PCR, and sequencing of bridging PCR fragments with Sanger and/or PacBio technologies.

Project information is available in the Genomes Online Database (26). Genes were identified using Prodigal (27), followed by a round of manual curation using GenePRIMP (28) as part of the microbial annotation pipeline of the JGI (29). Additional gene prediction analysis and manual functional annotation were performed within the Integrated Microbial Genomes-Expert Review (IMG-ER) platform (http://img.jgi.doe.gov) developed by the Joint Genome Institute (Walnut Creek, CA, USA) (30).

The high-quality draft genome of *Frankia* sp. DC12 was resolved to 1 scaffold consisting of 6,884,336 bp, with a G+C content of 71.92%, 5,858 candidate protein-coding genes, 46 tRNA genes, and 3 rRNA regions.

**Nucleotide sequence accession numbers.** The *Frankia* sp. strain DC12 genome sequence has been deposited at DDBJ/EMBL/GenBank under the accession no. LANG00000000. The version described in this paper is the first version, LANG01000000.

## ACKNOWLEDGMENTS

## REFERENCES

1. **Schwencke J, Carú 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. http://dx.doi.org/10.1080/153249801753127615.

2. **Chaia EE, Wall LG, Huss-Danell K**. 2010. Life in soil by the actinorhizal root nodule endophyte *Frankia*. A review. Symbiosis **51**:201–226. http://dx.doi.org/10.1007/s13199-010-0086-y.

3. **Normand P, Benson DR, Berry AM, Tisa LS**. 2014. Family *Frankiaceae*, p 339–356. *In* Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (ed), The prokaryote–actinobacteria, 4th ed. Springer-Verlag, Berlin, Germany.

4. **Normand P, Orso S, Cournoyer B, Jeannin P, Chapelon C, Dawson J, Evtushenko L, Misra AK**. 1996. Molecular phylogeny of the genus *Frankia* and related genera and emendation of the family *Frankiaceae*. Int J Syst Bacteriol **46**:1–9. http://dx.doi.org/10.1099/00207713-46-1-1.

5. **Cournoyer B, Lavire C**. 1999. Analysis of *Frankia* evolution radiation using *glnII* sequences. FEMS Microbiol Lett **117**:29–34.

6. **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. http://dx.doi.org/10.1007/s10482-011-9613-y.

7. **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.

8. **Normand P, Lapierre P, Tisa LS, Gogarten JP, Alloisio N, Bagnarol E, Bassi CA, Berry AM, Bickhart DM, Choisne 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, Mastronunzio JE, Mullin BC, Niemann J, Pujic P, Rawnsley T, Rouy Z, Schenowitz C, Sellstedt A, Tavares F, Tomkins JP, Vallenet 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. http://dx.doi.org/10.1101/gr.5798407.

9. **Normand P, Queiroux C, Tisa LS, Benson DR, Rouy Z, Cruveiller S, Médigue C**. 2007. Exploring the genomes of *Frankia*. Physiol Plant **130**:331–343. http://dx.doi.org/10.1111/j.1399-3054.2007.00918.x.

10. **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, Pitluck S, Pennacchio L, Nolan M, Ivanova N, Pati A, Land ML, Pawlowski K, Berry AM**. 2011. Genome sequence of "*Candidatus* Frankia datiscae" Dg1, the uncultured microsymbiont from nitrogen-fixing root nodules of the dicot *Datisca glomerata*. J Bacteriol **193**:7017–7018. http://dx.doi.org/10.1128/JB.06208-11.

11. **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 LG, Woyke T, Tisa LS**. 2013. Draft genome sequence of *Frankia* sp. strain CN3, an atypical, non-infective (Nod⁻) ineffective (Fix⁻) isolate from *Coriaria nepalensis*. Genome Announc **1**(2):-13. http://dx.doi.org/10.1128/genomeA.00085-13.

12. **Sen A, Beauchemin N, Bruce D, Chain P, Chen A, Walston Davenport K, Deshpande S, Detter C, Furnholm T, Ghodbhane-Gtari F, 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, Sur S, Szeto E, Tavares F, Teshima H, Thakur S, Wall LG, Woyke T, Tisa LS**. 2013. Draft genome sequence of *Frankia* sp. strain QA3, a nitrogen-fixing actinobacterium isolated from the root

13. **Nouioui I, Beauchemin N, Cantor MN, Chen A, Detter C, Furnholm T, Ghodhbane-Gtari F, Goodwin L, Gtari M, Han C, Han J, Huntemann M, Hua SX, Ivanova N, Kyrpides N, Markowitz V, Mavrommatis K, Mikhailova N, Nordberg HP, Ovchinnikova G, Pagani I, Pati A, Sen A, Sur S, Szeto E, Thakur S, Wall LG, Wei C-L, Woyke T, Tisa LS**. 2013. Draft genome sequence of *Frankia* sp. strain BMG5.12, a nitrogen-fixing actinobacterium isolated from Tunisian soils. Genome Announc **1**(4):. http://dx.doi.org/10.1128/genomeA.00468-13.

14. **Wall LG, Beauchemin N, Cantor MN, Chaia E, Chen A, Detter JC, Furnholm T, Ghodhbane-Gtari F, Goodwin L, Gtari M, Han C, Han J, Huntemann M, Hua SX, Ivanova N, Kyrpides N, Markowitz V, Mavrommatis K, Mikhailova N, Nordberg HP, Nouioui I, Ovchinnikova G, Pagani I, Pati A, Sen A, Sur S, Szeto E, Thakur S, Wei C-L, Woyke T, Tisa LS**. 2013. Draft genome sequence of *Frankia* sp. strain BCU110501, a nitrogen-fixing actinobacterium isolated from nodules of *Discaria trinevis*. Genome Announc **1**(4):e00503-13 http://dx.doi.org/10.1128/genomeA.00503-13.

15. **Mansour SR, Oshone R, Hurst SG, IV, Morris K, Thomas WK, Tisa LS**. 2014. Draft genome sequence of *Frankia* sp. strain CcI6, a salt-tolerant nitrogen-fixing actinobacterium isolated from the root nodule of *Casuarina cunninghamiana*. Genome Announc **2**(1):. http://dx.doi.org/10.1128/genomeA.01205-13.

16. **Hurst SG, IV, Oshone R, Ghodhbane-Gtari F, Morris K, Abebe-Akele F, Thomas WK, Ktari A, Salem K, Mansour S, Gtari M, Tisa LS**. 2014. Draft genome sequence of *Frankia* sp. strain Thr, a nitrogen-fixing actinobacterium isolated from the root nodules of *Casuarina cunninghamiana* grown in Egypt. Genome Announc **2**(3):. http://dx.doi.org/10.1128/genomeA.00493-14.

17. **Ghodhbane-Gtari F, Hurst SG, IV, Oshone R, Morris K, Abebe-Akele F, Thomas WK, Ktari A, Salem K, Gtari M, Tisa LS**. 2014. Draft genome sequence of *Frankia* sp. strain BMG5.23, a salt-tolerant nitrogen-fixing actinobacterium isolated from the root nodules of *Casuarina glauca* grown in Tunisia. Genome Announc **2**(3):. http://dx.doi.org/10.1128/genomeA.00520-14.

18. **Tisa LS, Beauchemin N, Gtari M, Sen A, Wall LG**. 2013. What stories can the *Frankia* genomes start to tell us? J Biosci **38**:719–726. http://dx.doi.org/10.1007/s12038-013-9364-1.

19. **Hameed S, Hafeez FY, Mirza MS, Malik KA, Akkermans ADL**. 1994. Confirmation of an isolate from *Datisca cannabina* as atypical *Frankia* strain using PCR amplified 16 rRNA sequence analysis. Pak J Bot **26**:247–251.

20. **Hafeez F**. 1983. Nitrogen fixation and nodulation in *Datisca cannabina* L. and *Alnus nitida* Endl. PhD thesis. Quaid-e-Azam University, Islamabad, Pakistan.

21. **Richards JW, Krumholz GD, Chval MS, Tisa LS**. 2002. Heavy metal resistance patterns of *Frankia* strains. Appl Environ Microbiol **68**:923–927. http://dx.doi.org/10.1128/AEM.68.2.923-927.2002.

22. **Krumholz GD, Chval MS, McBride MJ, Tisa LS**. 2003. Germination and physiological properties of *Frankia* spores. Plant Soil **254**:57–67. http://dx.doi.org/10.1023/A:1024911231889.

23. **Bennett S**. 2004. Solexa Ltd. Pharmacogenomics **5**:433–438. http://dx.doi.org/10.1517/14622416.5.4.433.

24. **Zerbino DR, Birney E**. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res **18**:821–829. http://dx.doi.org/10.1101/gr.074492.107.

25. **Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB**. 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci U S A **108**:1513–1518. http://dx.doi.org/10.1073/pnas.1017351108.

26. **Reddy TB, Thomas AD, Stamatis D, Bertsch J, Isbandi M, Jansson J, Mallajosyula J, Pagani I, Lobos EA, Kyrpides NC**. 2015. The Genomes Online Database (GOLD) v.5: a metadata management system based on a four level (meta)genome project classification. Nucleic Acids Res **43**:D1099–D1106. http://dx.doi.org/10.1093/nar/gku950.

27. **Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ**.

2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11:**119. http://dx.doi.org/10.1186/1471-2105-11-119.

28. **Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC.** 2010. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. Nat Methods **7:**455–457. http://dx.doi.org/10.1038/nmeth.1457.

29. **Mavromatis K, Ivanova NN, Chen IM, Szeto E, Markowitz VM, Kyrpides NC.** 2009. The DOE-JGI standard operating procedure for the annotations of microbial genomes. Stand Genomic Sci **1:**63–67.

30. **Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC.** 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics **25:**2271–2278. http://dx.doi.org/10.1093/bioinformatics/btp393.