8-6-2015

Draft genome sequence of Frankia sp. strain DC12, an atypical, noninfective, ineffective isolate from Datisca cannabina

Louis S. Tisa
University of New Hampshire, Durham, louis.tisa@unh.edu

Nicholas Beauchemin
University of New Hampshire, Durham

Michael N. Cantor
Los Alamos National Laboratory

Teal R. Furnholm
University of New Hampshire, Durham

Faten Ghodhbane-Gtari
Université Tunis El Manar (FST) & Université Carthage (INSAT)

See next page for additional authors

Follow this and additional works at: https://scholars.unh.edu/nhaes

Recommended Citation

This Article is brought to you for free and open access by the Research Institutes, Centers and Programs at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in New Hampshire Agricultural Experiment Station by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.
Draft Genome Sequence of Frankia sp. Strain DC12, an Atypical, Noninfective, Ineffective Isolate from Datisca cannabina

ARTICLE in GENOME ANNOUNCEMENTS · AUGUST 2015
DOI: 10.1128/genomeA.00889-15

CITATIONS
2

READS
71

26 AUTHORS, INCLUDING:

Lynne Goodwin
Los Alamos National Laboratory
315 PUBLICATIONS 4,056 CITATIONS

Galina Ovchinnikova
Lawrence Berkeley National Laboratory
63 PUBLICATIONS 987 CITATIONS

Ioanna Pagani
Thermo Fisher Scientific
110 PUBLICATIONS 2,440 CITATIONS

Ernest Szeto
Lawrence Berkeley National Laboratory
58 PUBLICATIONS 3,745 CITATIONS

Available from: Louis S Tisa
Retrieved on: 14 January 2016
Draft Genome Sequence of *Frankia* sp. Strain DC12, an Atypical, Noninfective, Ineffective Isolate from *Datisca cannabina*


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

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 reinfest 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.

Frankia spp. are well known as plant symbionts of dicotyle-donous 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 reinfest their host plant, while the fourth lineage (termed atypical *Frankia* isolates) are unable to reinfest 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**

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 DE-AC02-05CH11231. This project (L.S.T.) was supported in part by Agriculture and Food Research Initiative grant 2010-
REFERENCES


