AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Permanent Draft Genome Sequences of Three *Frankia* sp. Strains That Are Atypical, Noninfective, Ineffective Isolates

Abdellatif Gueddou,[a] Erik Swanson,[b] Amir Ktari,[a] Imen Nouioui,[a] Karima Hezbri,[a] Faten Ghodhbane-Gtari,[a] Stephen Simpson,[b] Krystalynne Morris,[b] W. Kelley Thomas,[b] Arnab Sen,[c] Maher Gtari,[a] Louis S. Tisa[b]

Université de Tunis El Manar, Tunis, Tunisia[a]; University of New Hampshire, Durham, New Hampshire, USA[b]; University of North Bengal, Siliguri, India[c]

ABSTRACT Here, we present draft genome sequences for three atypical *Frankia* strains (lineage 4) that were isolated from root nodules but are unable to reinfect actinorhizal plants. The genome sizes of *Frankia* sp. strains EUN1h, BMG5.36, and NRRL B16386 were 9.91, 11.20, and 9.43 Mbp, respectively.

Received 14 February 2017 **Accepted** 16 February 2017 **Published** 13 April 2017

Citation Gueddou A, Swanson E, Ktari A, Nouioui I, Hezbri K, Ghodhbane-Gtari F, Simpson S, Morris K, Thomas WK, Sen A, Gtari M, Tisa LS. 2017. Permanent draft genome sequences of three *Frankia* sp. strains that are atypical, noninfective, ineffective isolates. Genome Announc 5:e00174-17. https://doi.org/10.1128/genomeA.00174-17.

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

Endosymbiotic plant-bacterium associations are contributors to terrestrial biological nitrogen fixation and include actinorhizal symbiosis. This mutually beneficial symbiotic relationship between actinobacterial *Frankia* spp. and actinorhizal plants results in the formation of plant root nodule structure. This relationship allows proliferation of the plant through the bacterium, obtaining nutrients from the host plant in exchange for a source of fixed nitrogen that is assimilated by the host plant (1). Mutualistic infective *Frankia* strains are systematically classified based on their morphology, behavior in culture, and mode of infection within one of three major phylogenetic clusters (2). Another *Frankia* group isolated from actinorhizal nodules that are unable to undertake the nitrogen fixation process (Fix^-) and/or reinfect their host plant causing nodulation (Nod^-) are classified as "atypical *Frankia*" spp. and form a fourth phylogenetic cluster within the genus *Frankia*. The phenomena of how these atypical *Frankia* spp. enter inside nodule and the host metabolic cost of their presence as parasitic cheaters remain unclear (3). Although genomes for representatives for all four clusters have been sequenced (4), only two genomes are available for atypical *Frankia* spp. from cluster 4. The purpose of this study was to expand the number of genomes sequenced from cluster 4 to provide insight on these questions.

Frankia sp. strains EUN1h, BMG5.36, and NRRL B16386 were isolated from *Elaeagnus umbellata* (Tunisia), *Coriaria myrifolia* (Algeria), and *Morella californica* (United States; A. Gueddou, M. Gtari, M. Lechevalier, unpublished data), respectively. All three strains have failed to reinfect and nodulate their respective original host and any other actinorhizal host plant tested.

Sequencing of the draft genomes of *Frankia* sp. strains EUN1h, BMG5.36, and NRRL B16386 was performed at the Hubbard Center for Genome Studies (University of New Hampshire, Durham, NH) using Illumina technology techniques (5). A standard Illumina shotgun library was constructed and sequenced using the Illumina HiSeq 2500 platform with paired-end reads of 2 × 250 bp, which generated 2,121,668 to 15,077,492 reads (Table 1). The Illumina sequence data were trimmed by Trimmonatic version 0.32 (6) and assembled using SPAdes version 3.5 (7) and ALLPaths-LG version r52488 (8). Data on the final draft assemblies for *Frankia* sp. strains EUN1h, BMG5.36, and NRRL B16386 are presented in Table 1. The final assembled genomes for *Frankia* sp. strains EUN1h,

TABLE 1 Genome statistics

Frankia strain	No. of reads	N_{50} contig size (kb)	Assembly size (Mb)	No. of contigs	Sequencing depth (×)	No. of CDSs[a]	G+C content (%)	No. of biosynthetic gene clusters[b]	Accession no.
BMG5.36	2,121,668	84.9	11.20	280	28.0	8,952	71.26	33	MBLO00000000
NRRL B-16386	10,384,450	117.5	9.43	174	161.4	7,562	71.93	27	MOMC00000000
EUN1h	15,077,492	194.6	9.91	129	305.1	7,928	71.83	30	MBLN00000000

[a]CDSs, coding sequences.
[b]Biosynthetic gene clusters for natural products were identified by the use of the antiSMASH software (9, 10).

BMG5.36, and NRRL B16386 contained total sequence lengths of 9,910,952, 11,203,906, and 9,435,764 bp, respectively, with an average G+C content of 71% (Table 1). The assembled *Frankia* sp. strains EUN1h, BMG5.36, and NRRL B16386 genomes were annotated via the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) and resulted in 7,928, 8,952, and 7,562 candidate protein-coding genes, respectively. Bioinformatic analysis of these three genomes by use of the antiSMASH program (9, 10) revealed that these genomes again provided high numbers of secondary metabolic biosynthetic gene clusters (Table 1), similar to previous findings (4, 11), and including potential compounds, like frankiamicin (12).

Accession number(s). The draft genome sequences have been deposited in GenBank under the accession numbers in Table 1.

ACKNOWLEDGMENTS

Partial funding was provided by the New Hampshire Agricultural Experiment Station. This work was supported by the USDA National Institute of Food and Agriculture Hatch 022821 (to L.S.T.), Agriculture and Food Research Initiative Grant 2015-67014-22849 from the USDA National Institute of Food and Agriculture (to L.S.T.), and the College of Life Science and Agriculture at the University of New Hampshire-Durham. Partial funding was supported by the Laboratoire Microorganismes & Biomolécules Actives, Université Tunis El-Manar, Tunisia (grant LR03ES03). This is Scientific Contribution no. 2710.

Sequencing was performed on an Illumina HiSeq 2500 purchased with an NSF MRI grant DBI-1229361 to W.K.T.

REFERENCES

1. 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 prokaryotes: *Actinobacteria*. Springer-Verlag, Berlin Heidelberg, Heidelberg, Germany.
2. Clawson ML, Bourret A, Benson DR. 2004. Assessing the phylogeny of *Frankia*-actinorhizal plant nitrogen-fixing root nodule symbioses with *Frankia* 16S rRNA and glutamine synthetase gene sequences. Mol Phylogenet Evol 31:131–138. https://doi.org/10.1016/j.ympev.2003.08.001.
3. Fujita H, Aoki S, Kawaguchi M. 2014. Evolutionary dynamics of nitrogen fixation in the legume-rhizobia symbiosis. PLoS One 9:e93670. https://doi.org/10.1371/journal.pone.0093670.
4. Tisa LS, Oshone R, Sarkar I, Ktari A, Sen A, Gtari M. 2016. Genomic approaches toward understanding the actinorhizal symbiosis: an update on the status of the *Frankia* genomes. Symbiosis 70:5–16. https://doi.org/10.1007/s13199-016-0390-2.
5. Bennett S. 2004. Solexa Ltd. Pharmacogenomics 5:433–438. https://doi.org/10.1517/14622416.5.4.433.
6. 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.
7. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol 20:714–737. https://doi.org/10.1089/cmb.2013.0084.
8. 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. https://doi.org/10.1073/pnas.1017351108.
9. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R. 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res 39:W339–W346. https://doi.org/10.1093/nar/gkr466.
10. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015. antiSMASH 3.0–a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43:W237–W243. https://doi.org/10.1093/nar/gkv437.
11. Udwary DW, Gontang EA, Jones AC, Jones CS, Schultz AW, Winter JM, Yang JY, Beauchemin N, Capson TL, Clark BR, Esquenazi E, Eustáquio AS, Freel K, Gerwick L, Gerwick WH, Gonzalez D, Liu WT, Malloy KL, Maloney KN, Nett M, Nunnery JK, Penn K, Prieto-Davo A, Simmons TL, Weitz S, Wilson MC, Tisa LS, Dorrestein PC, Moore BS. 2011. Significant natural product biosynthetic potential of actinorhizal symbionts of the genus *Frankia*, as revealed by comparative genomic and proteomic analyses. Appl Environ Microbiol 77:3617–3625. https://doi.org/10.1128/AEM.00038-11.
12. Ogasawara Y, Yackley BJ, Greenberg JA, Rogelj S, Melançon CE. 2015. Expanding our understanding of sequence-function relationships of type II polyketide biosynthetic gene clusters: bioinformatics-guided identification of frankiamicin A from *Frankia* sp. EAN1pec. PLoS One 10:e0121505. https://doi.org/10.1371/journal.pone.0121505.