Draft Genome Sequence of the Symbiotic Frankia Sp. Strain KB5 Isolated from Root Nodules of Casuarina equisetifolia

Celine Pesce
*University of New Hampshire, Durham*, Celine.Pesce@unh.edu

Erik Swanson
*University of New Hampshire, Durham*

Stephen Simpson
*University of New Hampshire, Durham*

Krystalynne Morris
*University of New Hampshire, Durham*

W. Kelley Thomas
*University of New Hampshire, Durham*, Kelley.Thomas@unh.edu

See next page for additional authors

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

**Recommended Citation**

Pesce, C. E. Swanson, S. Simpson, K. Morris, W.K. Thomas, L.S. Tisa, and A. Sellstedt. 2017. Draft Genome Sequence for Frankia sp. Strain KB5 Isolated from the Root Nodules of Casuarina equisetifolia. *J. Genomics* 5: 64-67. doi: 10.7150/jgen.20887

This Article is brought to you for free and open access by the Molecular, Cellular and Biomedical Sciences at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Molecular, Cellular and Biomedical Sciences Scholarship by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact Scholarly.Communication@unh.edu.
Authors
Celine Pesce, Erik Swanson, Stephen Simpson, Krystalynne Morris, W. Kelley Thomas, Louis S. Tisa, and Anita Sellstedt
Abstract

Frankia sp. strain KB5 was isolated from Casuarina equisetifolia and previous studies have shown both nitrogenase and uptake hydrogenase activities under free-living conditions. Here, we report 5.5-Mbp draft genome sequence with a G+C content of 70.03 %, 4,958 candidate protein-encoding genes, and 2 rRNA operons.

Key words: actinobacteria, actinorhizal symbiosis, hydrogenase, nitrogen fixation, natural products, host microbe interactions, genomes.

Introduction

Biological nitrogen fixation, which is carried out by prokaryotes, converts atmospheric dinitrogen gas into a reduced biologically useful form. One group of these nitrogen-fixing microbes is the genus Frankia, which is found in association with actinorhizal plants or free-living in the soil [1, 2]. Based on molecular studies, there are four distinct Frankia lineages [3-5]. Three of the Clusters (I, II, and III) are infective on actinorhizal host plants, while members of Cluster IV are considered “atypical” strains, which are unable to re-infect the plants or form ineffective nodules that unable to fix nitrogen.

Frankia sp. strain KB5 was isolated from the root nodules of Casuarina equisetifolia ssp. incana in Australia [6]. This strain is a member of Frankia subcluster Ic, which has its host range limited to Casuarina and Allocasuarina plants. The physiology of Frankia sp. strain KB5 has been well studied especially the close relationship between hydrogenase and nitrogenase activities [7-9]. Interestingly, the addition of nickel results in increased uptake hydrogenase activity [7]. Frankia sp. strain KB5 genome was chosen to be sequenced for several reasons including an interesting physiology. Furthermore, this strain represents an isolate from Australia, the native biogeographical region of Frankia subcluster Ic [10]. This database could provide more information on the evolution of this Frankia subcluster.

Sequencing of the draft genome of Frankia sp. strain KB5 was performed at the Hubbard Center for Genome Studies (University of New Hampshire, Durham, NH) using Illumina technology techniques [11]. A standard Illumina shotgun library was constructed and sequenced using the Illumina HiSeq2500 platform, using a pair-end library with an average size of 600 bp obtaining 10,539,470 reads of 250 bp in length. The Illumina sequence data were trimmed by Trimmonatic version 0.32 [12], assembled
using Spades version 3.5 [12] and ALLPaths-LG version r52488 [13]. The final draft assembly for Frankia sp. strain KB5 consisted of 420 contigs with an N50 contig size of 24.2 kb and 236.8X coverage of the genome. The final assembled genome contained a total sequence length of 5,453,564 bp with a G+C content of 70.03%.

The assembled Frankia sp. strain KB5 genome was annotated via the NCBI Prokaryotic Genome Annotation Pipeline (PGAP), and resulted in 4,958 candidate protein-encoding genes and 2 rRNA operons. The genome features of Frankia sp. strain KB5 fall with the realm of the other Casuarina genomes, including Frankia casuarinae strain (CcI3) [14] the type strain (Table 1).

The genome also contained a nif and 2 hup operons encoding the nitrogenase and uptake hydrogenase enzymes, respectively. The operons were organized similar to those reported for Frankia cluster I genomes [15]. The 2 hup operons have been shown to be expressed differently, with hup 1 being mainly expressed in free-living conditions and hup 2 in symbiotic [16]. Bioinformatic analysis of this genomes by the use of the AntiSMASH program [17, 18] revealed the presence of high numbers of secondary metabolic biosynthetic gene clusters, which is consistent with previous results with other Frankia genomes including subcluster Ic [15, 19]. Table 2 shows a comparison of the various profiles of different Casuarina isolates for these secondary metabolic biosynthetic gene clusters. Although the majority of these secondary metabolic biosynthetic gene clusters were shared among the Casuarina genomes, the Frankia sp. strain KB5 genome contained several unique clusters that had homologues in other bacteria or were completely novel. For example, cluster KB-24 (location: KBl5_16785-KBl5_16785 genes) involved in terpene biosynthesis had no homology with any of the Frankia genomes, but was homologous to the phenalinolactoneA biosynthetic gene cluster of Streptomyces pactum KLBMP 5084, including the surrounding gene neighborhood.

Cluster KB-26 (location: KBl5_17660-KBl5_17705 genes) containing a nonribosomal peptide synthase (NRPS) is only found in the Frankia sp. strain KB5 genome and predicted to produce a core structure (Fig. 1). Further bioinformatics analysis revealed that the Frankia sp. strain KB5 genome contained 715 unique genes that had no homologues in any of the other Casuarina genomes. Although some of these genes like cluster KB-24 genes have predicted functions, the majority of the genes code for hypothetical proteins without a known function. The hypothesis that these genes play a role in the biogeographical distribution of this strain remains to be tested.

Table 1. Genome features of Frankia sp. strain KB5 and other Frankia strains isolated from Casuarina root nodules.

| Strain | Source | Location | Size (Mb) | No. of Contig(s) | G+C (%) | No. of CDS | No. of rRNA | No. of tRNA |
|--------|--------|----------|-----------|-----------------|---------|------------|-------------|-------------|
| KB5    | This study | Australia | 5.46 | 420 | 70.0 | 4,958 | 6 | 45 |
| CcI3   | [10] | USA | 5.43 | 1 | 70.1 | 4,598 | 6 | 46 |
| CeD    | [20] | Senegal | 5.00 | 120 | 70.1 | 4,403 | 7 | 45 |
| Allo2  | [21] | Uruguay | 5.33 | 110 | 69.8 | 4,838 | 7 | 46 |
| Thr    | [22] | Egypt | 5.31 | 171 | 70.0 | 4,805 | 5 | 46 |
| BMG5.23 | [23] | Tunisia | 5.27 | 167 | 70.0 | 4,747 | 9 | 47 |
| CcI6   | [24] | Egypt | 5.39 | 138 | 67.6 | 4,902 | 9 | 46 |
| BR     | [25] | Brazil | 5.23 | 180 | 70.0 | 4,777 | 5 | 46 |

1 The source of the isolate.

Table 2. Biosynthetic gene clusters for natural products found in the genomes from Casuarina Frankia strains.

| Strain | No. of Biosynthetic gene clusters 1 | NRPS 2 | PKS 3 | Terpene | Siderophore | Bacteriocin | Lantipeptide |
|--------|---------------------------------|--------|--------|---------|-------------|-------------|-------------|
| KB5    | 34                              | 4      | 9      | 6       | 1           | 1           | 4           |
| CcI3   | 29                              | 3      | 5      | 4       | 1           | 3           | 6           |
| CeD    | 30                              | 7      | 7      | 4       | 1           | 1           | 4           |
| Allo2  | 32                              | 7      | 9      | 4       | 1           | 3           | 5           |
| Thr    | 33                              | 6      | 7      | 4       | 1           | 1           | 6           |
| BMG5.23 | 31                            | 8      | 6      | 4       | 1           | 2           | 4           |
| CcI6   | 33                              | 8      | 8      | 4       | 1           | 3           | 5           |
| BR     | 29                              | 5      | 5      | 4       | 1           | 2           | 5           |

1 Biosynthetic gene clusters were identified by the use of the AntiSMASH software [17, 18]
2 NRPS: Nonribosomal peptide synthase
3 PKS: polyketide synthase including Type I, II, III, Trans-AT, and other types
Fixing Actinobacterium Isolated from the Root Nodules of Casuarina equistifolia and related genera and emendation of the family Frankiaceae. Int J Syst Bacteriol. 1996; 46: 1-9.

Normand P, Ghodhibane-Gtari F, Nouioui I, Chair M, Boudabous A, Gtari M. Phylogeny of members of the Frankia genus based on gyrB, mflB and glnII sequences. Anton Leeuw Int J G. 2011; 100: 579-87.

Ghodhibane-Gtari F, Nouioui I, Chair M, Boudabous A, Gtari M. 16S-23S rRNA Intergenic Spacer Region Variability in the Genus Frankia. Microb Ecol. 2010; 60: 487-95.

Figure 1. The predicted Frankia chemical structure for secondary metabolic biosynthetic gene cluster KB-24.

In summary, the Frankia sp. strain KB5 genome has revealed an interesting potential natural product profile and serves as a representative of Frankia subcluster Ic from its native environment. Further analysis of this genome and experimental evidence will be needed to support the predicted natural product profile and to provide insight on the evolution of this Frankia subcluster.

Nucleotide sequence accession numbers

This whole-genome shotgun sequence has been deposited at DDBJ/EMBL/GenBank under the accession number MRUJ0000000. The version described in this paper is the first version, MRUJ01000000.

Acknowledgement

Partial funding was provided by the New Hampshire Agricultural Experiment Station. This is Scientific Contribution Number 2725. This work was also supported by the Swedish Energy Agency 38239-1 (AS), USDA National Institute of Food and Agriculture Hatch 022821 (LST), Agriculture and Food Research Initiative Grant 2015-67014-22849 from the USDA National Institute of Food and Agriculture (LST), and the College of Life Science and Agriculture at the University of New Hampshire-Durham. Sequencing was performed on an Illumina HiSeq2500 purchased with an NSF MRI Grant: DBI-1229361 to WK Thomas.

Competing Interests

The authors have declared that no competing interest exists.

References

1. Normand P, Benson DR, Berry AM, Tisa LS. Family Frankiaceae. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The Prokaryotes – Actinobacteria Berlin Heidelberg: Springer-Verlag; 2014:339-56

2. Chait EE, Wall LG, Huss-Danell K. Life in soil by the actinorhizal root nodule endophyte Frankia. A review. Symbiosis. 2010; 51: 201-26.

3. Normand P, Orso S, Cournoyer B, Jeannin P, Chapelon C, Dawson J, et al. Molecular phylogeny of the genus Frankia and related genera and emendation of the family Frankiaceae. Int J Syst Bacteriol. 1996; 46: 1-9.

4. Nouioui I, Ghodhibane-Gtari F, Beauchemin NJ, Tisa LS, Gtari M. 16S-23S rRNA Intergenic Spacer Region Variability in the Genus Frankia. Microb Ecol. 2010; 60: 487-95.

5. Ghodhibane-Gtari F, Nouioui I, Chair M, Boudabous A, Gtari M. 16S-23S rRNA Intergenic Spacer Region Variability in the Genus Frankia. Microb Ecol. 2010; 60: 487-95.

6. Rosbrook PA, Burggraaf AJ, Reddell P. A Comparison of 2 Methods and Different Media for Isolating Frankia from Casuarina Root-Nodules. Plant Soil. 1989; 120: 187-93.

7. Mattsson U, Sellstedt A, Nickel affects activity more than expression of hydrogenase protein in Frankia. Curr Microbiol. 2002; 44: 88-93.

8. Mattsson U, Johansson I, Sandstrom G, Sellstedt A. Frankia KB5 possesses a hydrogenase immunologically related to membrane-bound. Curr Microbiol. 2001; 42: 438-41.

9. Mattsson U, Sellstedt A. Hydrogenase in Frankia KB5: Expression of and relation to nitrogenase. Can J Microbiol. 2000; 46: 1091-5.

10. Normand P, Lapiere P, Tisa LS, Alloisio N, Bagnol E, et al. Genome characteristics of facultatively symbiotic Frankia sp strains reflect host range and host plant biogeography. Genome research. 2007; 17: 7-15.

11. Bennett S. Solexa Ltd. Pharmacogenomics. 2004; 5: 433-8.

12. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114-20.

13. Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. P Natl Acad Sci USA. 2011; 108: 1513-8.

14. Nouioui I, Ghodhibane -Gtari F, Monterro-Calasanz MD, Goker M, Meyer-Kolhoff JP, Schumann P, et al. Proposal of a type strain for Frankia alni (Woronin 1866) Von Tubeuf 1895, emended description of Frankia alni, and recognition of Frankia casuarinae sp. nov. and Frankia elaeagni sp. nov. Int J Syst Evol Microbiol. 2016; 66: 5201-10.

15. Tisa LS, Oshone R, Sarkar L, Ktari A, Sen A, Gtari M. Genomic approaches toward understanding the actinorhizal symbiosis: an update on the status of the Frankia genomes. Symbiosis. 2016; 70: 5-16.

16. Leul M, Normand P, Sellstedt A. The organization, regulation and phylogeny of uptake hydrogenase genes in Frankia. Physiol Plantarum. 2007; 130: 464-70.

17. Blin K, Medema Ml, Kazempour E, Fischbach MA, Breitling R, Takano E, et al. antiSMASH 2.0-a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res. 2013; 41: W204-W12.

18. Weber T, Blin K, Duivetera D, Krug D, Kim HU, Bruccoleri R, et al. antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res. 2015; 43: W237-W43.

19. Udwary DW, Gontang EA, Jones AC, Jones CS, Schultz AW, Winter JM, et al. Significant Natural Product Biosynthetic Potential of Actinorhizal Symbionts of the Genus Frankia, as Revealed by Comparative Genomic and Proteomic Analyses. Appl Environ Microb. 2011; 77: 3617-25.

20. Ngom M, Oshone R, Hurst SG, Abebe-Akele F, Simpson S, Morris K, et al. Permanent Draft Genome Sequence for Frankia sp. Strain CeD, a Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of Casuarina equistifolia Grown in Senegal. Genome announcements. 2016; 4.

21. Oshone R, Ngom M, Abebe-Akele F, Simpson S, Morris K, Sy MO, et al. Permanent Draft Genome Sequence of Frankia sp. Strain Allo2, a Salt-Tolerant Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of Allocasuarina Genome announcements. 2016; 4.

22. Hurst SG, Oshone R, Ghodhibane-Gtari F, Morris K, Abebe-Akele F, Thomas WK, et al. Draft Genome Sequence of Frankia sp. Strain Thr, a Salt-Tolerant Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of Allocasuarina Genome announcements. 2016; 4.

23. Ghodhibane-Gtari F, Hurst SG, Oshone R, Morris K, Abebe-Akele F, Thomas WK, et al. Draft Genome Sequence of Frankia sp. Strain BMG5.23, a Salt-Tolerant Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of Allocasuarina. Genome announcements. 2014; 2: e00493-14.

http://www.jgenomics.com
from the Root Nodules of *Casuarina glauca* Grown in Tunisia. Genome announcements. 2014; 2: e00520-14.
24. Mansour SR, Oshone R, Hurst SGt, Morris K, Thomas WK, Tisa LS. Draft Genome Sequence of *Frankia* sp. Strain CcI6, a Salt-Tolerant Nitrogen-Fixing Actinobacterium Isolated from the Root Nodule of *Casuarina cunninghamiana*. Genome announcements. 2014; 2.

25. D’Angelo T, Oshone R, Abebe-Akele F, Simpson S, Morris K, Thomas WK, et al. Permanent Draft Genome Sequence of *Frankia* sp. Strain BR, a Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of *Casuarina equisetifolia*. Genome announcements. 2016; 4.