9-12-2017

Permanent Draft Genome sequence for Frankia sp. strain CcI49, a Nitrogen-Fixing Bacterium Isolated from Casuarina cunninghamiana that Infects Elaeagnaceae

Samira R. Mansour
Suez Canal University

Erik Swanson
University of New Hampshire, Durham

Zakkary McNutt
University of New Hampshire, Durham

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

Kelsey Harrington
University of New Hampshire, Durham

Comments
See next page for additional authors

This is an article published by Ivyspring Interantional Publisher in Journal of Genomics in 2017, available online:

https://dx.doi.org/10.7150/jgen.22138

IVYSPRING INTERNATIONAL PUBLISHER

Journal of Genomics

2017; 5: 119-123. doi: 10.7150/jgen.22138

Short Research Paper

Permanent Draft Genome sequence for *Frankia* sp. strain CcI49, a Nitrogen-Fixing Bacterium Isolated from *Casuarina cunninghamiana* that Infects *Elaeagnaceae*

Samira Mansour[1]✉, Erik Swanson[2], Zakkary McNutt[2], Céline Pesce[2], Kelsey Harrington[2], Feseha Abebe-Alele[2], Stephen Simpson[2], Krystalynne Morris[2], W. Kelley Thomas[2], and Louis S. Tisa[2]✉

1. Suez Canal University, Ismailia, Egypt,
2. University of New Hampshire, Durham, New Hampshire, USA.

✉ Corresponding authors: Louis S. Tisa, Mailing address: Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire, 46 College Rd., Durham, NH 03824-2617. Email: louis.tisa@unh.edu Telephone: 1-603-862-2442 Fax: 1-603-862-2621. Samira Mansour, Mailing address: Botany department, Faculty of Science, Suez Canal University, Ismailia, Egypt. Email: samirarmansour@hotmail.com

Received: 2017.07.28; Accepted: 2017.08.31; Published: 2017.09.12

Abstract

Frankia sp. strain CcI49 was isolated from *Casuarina cunninghamiana* nodules. However the strain was unable to re-infect *Casuarina*, but was able to infect other actinorhizal plants including *Elaeagnaceae*. Here, we report the 9.8-Mbp draft genome sequence of *Frankia* sp. strain CcI49 with a G+C content of 70.5 % and 7,441 candidate protein-encoding genes. Analysis of the genome revealed the presence of a *bph* operon involved in the degradation of biphenyls and polychlorinated biphenyls.

Key words: Actinorhizal symbiosis, bioremediation, nitrogen fixation, natural products, host microbe interactions, genomes.

Soil dwelling actinobacteria of the genus *Frankia* form an endophytic symbiosis with actinorhizal plants, which are comprised of over 200 species from 8 angiosperm families [1, 2]. Actinorhizal plants in symbiosis with *Frankia* play important ecological roles as pioneer species and are used in agroforestry, land reclamation, crop protection, and soil stabilization projects [3]. Molecular phylogenetic approaches have identified four major clusters of *Frankia* that also follow host plant specificity groups [4-7]. Members of cluster 1 are divided into sub-cluster 1a that are infective on *Alnus* and *Myricaceae* and sub-cluster 1b strains which are infective on *Allocasuarina*, *Casuarina* and *Myricacaeae*. Cluster 2 represents strains infective on *Coriariaceae*, *Datiscaceae*, *Dryadoideae* and *Ceanothus*, while cluster 3 comprises strains that are infective on *Colletieae*, *Elaeagnaceae*, *Gymnostoma* and *Myricaceae*. Finally, cluster 4 groups *Frankia* strains isolated from actinorhizal nodules that are unable to undertake the nitrogen-fixation process (Fix^-) and/or re-infect their host plant causing nodulation (Nod^-) and are classified as "atypical *Frankia*". Genomes for representatives from each cluster have been sequenced [8]. The availability of these *Frankia* genome databases has opened up the use of "omics" approaches. Analysis of *Frankia* genomes has revealed new potential in respect to metabolic diversity, natural product biosynthesis, and stress tolerance, which may aid the cosmopolitan nature of the actinorhizal symbiosis.

Several *Frankia* strains isolated from *Casuarina* nodules are unable to re-infect it, but are able to infect other actinorhizal plant genera like *Elaeagnus* [9, 10]. Although isolated from *Casuarina* nodules, these *Frankia* strains are classified as members of cluster 3 based on molecular phylogeny and genomes for two members of this group have been sequenced [11, 12]. *Frankia* sp. strain CcI49 was isolated from root nodules

of *Casuarina cunninghamiana* grown on the edge of a cultivated field on side of the highway in Ismailia-Port Said, Egypt. The fresh nodules were washed, dissected into individual lobes, and surface-sterilized as described previously [13]. Each lobe was checked for sterility in sterile nutrient-rich medium. Nodules that were free from contamination were selected, dissected and homogenized, the homogenates were transferred to 100-ml screw caped bottle containing modified BAP medium for outgrowth. Hyphal outgrowth was homogenized and plated onto solid medium. After 3-4 weeks, colonies picked from the plates were homogenized and incubated in liquid medium. Surprisingly, *Frankia* sp. strain CcI49 produced reddish colonies, while other *Frankia* isolates from *Casuarina* do not. *Frankia* sp. strain CcI49 produced sporangia and spores that were smaller and narrower than normal *Frankia* sporangia and spores (Figure 1). Spores from *Frankia* sp. strain CcI49 had a high germination rate similar to *Frankia* strain CeI5 [14, 15]. We tested the ability of *Frankia* sp. strain CcI49 to re-infect actinorhizal plants. Four different actinorhizal plant species were tested to assay the plant host range and ten plants of each species were inoculated. *Frankia* sp. strain CcI49 was unable to infect *C. cunninghamiana* and *Alnus glutinosa,* but formed nodules on *Elaeagnus angustifolia* and *Hippophäe rhamnoides*. All ten of the *E, angustifolia* and *H. rhamnoides* plants tested formed nodules. Thus, *Frankia* sp. strain CcI49 had a plant-host-specificity pattern similar to *Frankia* sp. strains G2 and R43 [9, 10] from cluster 3 also isolated from *Casuarina* root nodules. *Frankia* sp. strain CcI49 genome was chosen to be sequenced for several reasons including an interesting physiology including the production of a reddish pigment, development of smaller sporangia and spores than are typical found with *Frankia*, and providing more information on this *Frankia* subcluster.

Sequencing of the draft genome of *Frankia* sp. strain CcI49 was performed at the Hubbard Center for Genome Studies (University of New Hampshire, Durham, NH) using Illumina technology techniques [16]. A standard Illumina shotgun library was constructed and sequenced using the Illumina HiSeq2500 platform, which generated 7,939,466 reads (260-bp insert size) totaling 1,921 MBp. The Illumina sequence data were trimmed by Trimmonatic version 0.32 [17], assembled using Spades version 3.5 [18], and ALLPaths-LG version r52488 [19]. The final draft assembly for *Frankia* sp. strain CcI49 consisted of 78 contigs with an N_{50} contig size of 282.1 kb and 167X coverage of the genome. The final assembled genome contained a total sequence length of 9,758,130 bp with a G+C content of 70.5%.

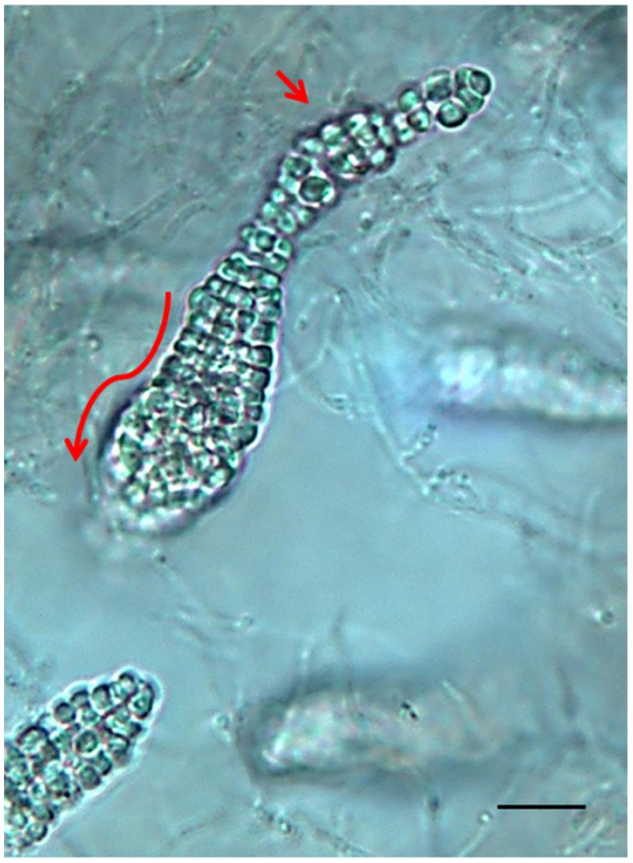

Figure 1. Photomicrograph of *Franakia* sp. strain CcI49 grown in liquid culture. The elongated arrow shows the presence of a long, narrow sessile sporangium containing differentiated mature spores at distal end (short arrow). Size bar represents 32 µm.

The assembled *Frankia* sp. strain CcI49 genome was annotated via the NCBI Prokaryotic Genome Annotation Pipeline (PGAP), and resulted in 7,411 candidate protein-encoding genes, 46 tRNA and 2 rRNA regions. The genome features of *Frankia* sp. strain CcI49 fall outside the realm of the other cluster 1b genomes, but similar to other cluster 3 isolates from *Casuarina* (Table 1). Phylogenetic analysis of the 23S rDNA shows that *Frankia* sp. strain CcI49 groups with the cluster 3 strains (Figure S1). The genome size and corresponding number of CDSs were larger than the typical cluster 1b, but fit within those values reported for cluster 3 genomes [8]. The genome also contained a *nif,* 2 *hup,* and 1 *shc* operons encoding the nitrogenase and uptake hydrogenase enzymes and the hopanoid biosynthetic pathway, respectively. The operons were organized similar to those reported for *Frankia* cluster 3 genomes [8].

Table 1. Genome features of *Frankia* sp. strain CcI49 and other *Frankia* strains isolated from *Casuarina* root nodules.

Strain	Source	Location[1]	Size (Mb)	No. of Contigs	*Frankia* cluster	No. of CDS	Host Plants[2]
CcI49	This study	Egypt	9.76	78	3	7,441	*Elaeagnaceae*
R43	[12]	USA	10.45	46	3	7,644	*Elaeagnaceae*
G2	[11]	Guadeloupe	9.54	90	3	7,790	*Elaeagnaceae*
KB5	[23]	Australia	5.46	420	1b	4,958	*Casuarinaceae*
CcI3	[24]	USA	5.43	1	1b	4,598	*Casuarinaceae*
CeD	[25]	Senegal	5.00	120	1b	4,403	*Casuarinaceae*
Allo2	[26]	Uruguay	5.33	110	1b	4,838	*Casuarinaceae*
Thr	[27]	Egypt	5.31	171	1b	4,805	*Casuarinaceae*
BMG5.23	[28]	Tunisia	5.27	167	1b	4,747	*Casuarinaceae*
CcI6	[29]	Egypt	5.39	138	1b	4,902	*Casuarinaceae*
BR	[30]	Brazil	5.23	180	1b	4,777	*Casuarinaceae*

[1] The source of the isolate

2 Re-infection plant host range

Analysis of the *Frankia* sp. strain CcI49 revealed the presence of the *bph* operon coding for a potential metabolic pathway involved in the degradation of biphenyl and polychlorinated biphenyls (Figure 2). The *bph* operon is also present in *Frankia* sp. strains EuI1c and EUN1f genomes [20] and was also found in the genomes of *Frankia* sp. strains G2 and R43 [11, 12], cluster 3 strains isolated from *Casuarina* root nodules. Both *Frankia* sp. strains CcI49 and EUN1f contained the entire *bph* operon, while two genes (*bphA3* and *bphH*) are missing in *Frankia* sp. EuI1c (Figure 2). The presence of the complete *bph* operon suggests that *Frankia* sp. strain CcI49 may be capable of degrading these recalcitrant xenobiotics.

Bioinformatic analysis of this genomes by the use of the AntiSMASH program [21] revealed the presence of high numbers of secondary metabolic biosynthetic gene clusters, which is consistent with previous results with other *Frankia* genomes including cluster 3 [8, 22]. Table 2 shows a comparison of the various profiles of different *Frankia* strains isolated from *Casuarina* for these secondary metabolic biosynthetic gene clusters. The profile of *Frankia* sp. strain CcI49 differed from those shown by *Frankia* strains that are able re-infect *Casuarina* and was similar to the pattern exhibited by the other two cluster 3 strains (R43 and G2) isolated from *Casuarina* nodules. These cluster 3 genomes contained more polyketide synthase (PKS) biosynthetic clusters than the cluster 1b genomes. The *Frankia* sp. strain CcI49 genome contained several unique clusters that had homologues in other bacteria or were completely novel.

In summary, the *Frankia* sp. strain CcI49 genome has revealed an interesting potential metabolic pathways and natural product profile, and serves as a representative of *Frankia* cluster 3. Further analysis of this genome and experimental evidence will be needed to support the predicted natural product profile and metabolic potential of *Frankia* sp. strain CcI49.

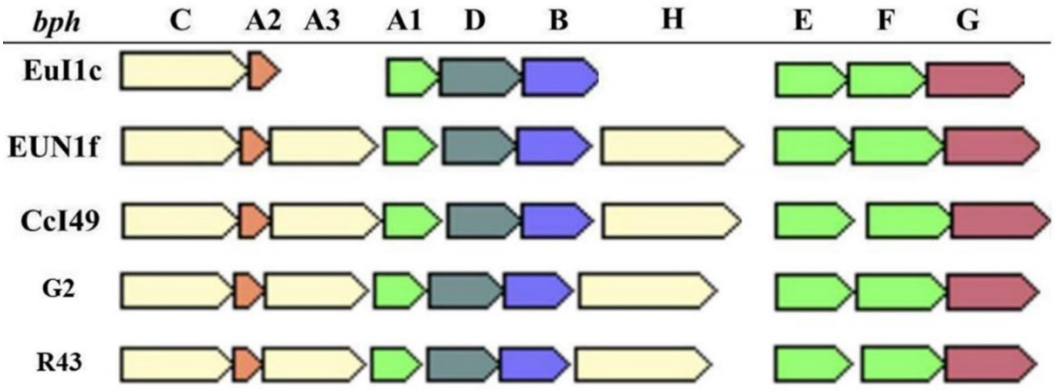

Figure 2. The *bph* operon is present in *Frankia* strains CcI49, EuI1c, G2, R43, and EUN1f. *bph* genes encode the following : BphA, A2 and A3, Biphenyl 2,3-dioxygenase; BphB, *cis* -2,3-dihydro-2,3-dihydroxybiphenyl dehydrogenase; BphC, 2,3-dihydroxybiphenyl 1,2-dioxygenase; BphD, 2-hydroxy-6-phenyl-6-oxohexa-2,4-dieneoate (HOPDA) hydrolase; BphE, 2-hydroxypenta-2,4-dienoate hydratase; BphF, acylating acetaldehyde dehydrogenase; and BphG, 4-hydroxy-2-oxovalerate aldolase.

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

Strain	*Frankia* Cluster	No. of Biosynthetic gene clusters [1]	NRPS [2]	PKS [3]	Terpene	Siderophore	Bacteriocin	Lantipeptide
CcI49	3	42	6	17	3	1	2	5
R43	3	38	4	14	3	1	2	4
G2	3	35	8	13	3	1	2	2
KB5	1b	34	4	9	6	1	1	4
CcI3	1b	29	3	5	4	1	3	6
CeD	1b	30	7	7	4	1	1	4
Allo2	1b	32	7	9	4	1	3	5
Thr	1b	33	6	7	4	1	1	6
BMG5.23	1b	31	8	6	4	1	2	4
CcI6	1b	33	8	8	4	1	3	5
BR	1b	29	5	5	4	1	2	5

[1] Biosynthetic gene clusters were identified by the use of the AntiSMASH software [21]

[2] NRPS: Nonribosomal peptide synthase

[3] PKS: polyketide synthase including Type I, II, III, Trans-AT, and other types

Nucleotide sequence accession numbers

This whole-genome shotgun sequence has been deposited at DDBJ/EMBL/GenBank under the accession number MOWP00000000.1. The version described in this paper is the first version, MOWP01000000.

Supplementary Material

Figure S1.
http://www.jgenomics.com/v05p0119s1.pdf

Acknowledgement

Partial funding was provided by the New Hampshire Agricultural Experiment Station. This is Scientific Contribution Number 2471. This work was supported by the 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), the College of Life Science and Agriculture at the University of New Hampshire-Durham and the Science and Technology Development Fund (STDF), grant No. 900, as International Cooperation Project (SM and LST). 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 Prokaryote – Actinobacteria Berlin Heidelberg: Springer-Verlag; 2014: 339-56
2. Chaia EE, Wall LG, Huss-Danell K. Life in soil by the actinorhizal root nodule endophyte *Frankia*. A review. Symbiosis. 2010; 51: 201-26.
3. Benson DR, Dawson JO. Recent advances in the biogeography and genecology of symbiotic Frankia and its host plants. Physiol Plantarum. 2007; 130: 318-30.
4. 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.
5. Ghodhbane-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. Nouioui I, Ghodhbane-Gtari F, Beauchemin NJ, Tisa LS, Gtari M. Phylogeny of members of the *Frankia* genus based on *gyrB*, *nifH* and *glnII* sequences. Anton Leeuw Int J G. 2011; 100: 579-87.
7. Gtari M, Tisa LS, Normand P. Diversity of Frankia Strains, Actinobacterial Symbionts of Actinorhizal Plants. In: Aroca R, editor. Symbiotic Endophytes. 2013: 123-48.
8. Tisa LS, Oshone R, Sarkar I, 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.
9. Zhang Z, Lopez MF, Torrey JG. A Comparison of Cultural-Characteristics and Infectivity of *Frankia* Isolates from Root-Nodules of *Casuarina* Species. Plant Soil. 1984; 78: 79-90.
10. Diem HG, Gauthier D, Dommergues YR. Isolation of *Frankia* from Nodules of *Casuarina equisetifolia*. Can J Microbiol. 1982; 28: 526-30.
11. Nouioui I, Gtari M, Goker M, Ghodhbane-Gtari F, Tisa LS, Fernandez MP, et al. Draft Genome Sequence of Frankia Strain G2, a Nitrogen-Fixing Actinobacterium Isolated from Casuarina equisetifolia and Able To Nodulate Actinorhizal Plants of the Order Rhamnales. Genome Announc. 2016; 4: e00437-16.
12. Pujic P, Bolotin A, Fournier P, Sorokin A, Lapidus A, Richau KH, et al. Genome Sequence of the Atypical Symbiotic *Frankia* R43 Strain, a Nitrogen-Fixing and Hydrogen-Producing Actinobacterium. Genome Announc. 2015; 3: e01387-15.
13. Mansour SR, Dewedar A, Torrey JG. Isolation, Culture, and Behavior of *Frankia* Strain HFPCgI4 from Root Nodules of *Casuarina glauca*. Bot Gaz. 1990; 151: 490-6.
14. Tzean SS, Torrey JG. Spore Germination and the Life Cycle of *Frankia Invitro*. Can J Microbiol. 1989; 35: 801-6.
15. Mansour SR, Torrey JG. *Frankia* Spores of Strain Hfpcgi4 as Inoculum for Seedlings of *Casuarina glauca*. Can J Bot. 1991; 69: 1251-6.
16. Bennett S. Solexa Ltd. Pharmacogenomics. 2004; 5: 433-8.
17. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114-20.
18. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, et al. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. Journal of computational biology : a journal of computational molecular cell biology. 2013; 20: 714-37.
19. 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.
20. Rehan M, Swanson E, Tisa LS. *Frankia* as a Biodegrading Agent. In: Dhanasekaran D, Jiang Y, editors. Actinobacteria - Basics and Biotechnological Applications. Croatia: In Tech Open Access 2016: 271-90.
21. Blin K, Wolf T, Chevrette MG, Lu XW, Schwalen CJ, Kautsar SA, et al. antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Res. 2017; 45: W36-W41.
22. 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.
23. Pesce C, Swanson E, Simpson S, Morris K, Thomas WK, Tisa LS, et al. Draft Genome Sequence for *Frankia* sp. Strain KB5 Isolated from the Root Nodules of *Casuarina equisetifolia*. . J Genomics. 2017; 5: 64-7
24. Normand P, Lapierre P, Tisa LS, Gogarten JP, Alloisio N, Bagnarol E, et al. Genome characteristics of facultatively symbiotic *Frankia* sp strains reflect host range and host plant biogeography. Genome research. 2007; 17: 7-15.
25. Ngom M, Oshone R, Hurst SGt, 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: e00265-16.

26. 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: e00388-16.
27. Hurst SGt, Oshone R, Ghodhbane-Gtari F, Morris K, Abebe-Akele F, Thomas WK, et al. Draft Genome Sequence of *Frankia* sp. Strain Thr, a Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of *Casuarina cunninghamiana* Grown in Egypt. Genome announcements. 2014; 2: e00493-14.
28. Ghodhbane-Gtari F, Hurst SGt, 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 *Casuarina glauca* Grown in Tunisia. Genome announcements. 2014; 2: e00520-14.
29. 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: e01205-13.
30. 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: e01000-16.