Research Paper

Draft genomes of non-nitrogen-fixing Frankia strains

Camila Carlos-Shanley, Trina Guerra, Dittmar Hahn

Texas State University, Department of Biology, 601 University Drive, San Marcos, TX 78666, USA.

Corresponding author: Camila Carlos-Shanley, Department of Biology, Texas State University, 601 University Drive, San Marcos, TX 78666. (512) 245-5283 (voice); (512) 245-8713 (fax); Email: carlos-shanley@txstate.edu

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). See http://ivyspring.com/terms for full terms and conditions.

Received: 2021.07.27; Accepted: 2021.09.14; Published: 2021.10.22

Abstract

In this study, we describe the genomes of two novel candidate species of non-nitrogen fixing Frankia that were isolated from the root nodules of Coriaria nepalensis and Alnus glutinosa, genospecies CN and Ag, respectively. Comparative genomic analyses revealed that both genospecies lack genes essential for nitrogen-fixation and possess genes involved in the degradation of plant cell walls. Additionally, we found distinct biosynthetic gene clusters in each genospecies. The availability of these genomes will contribute to the study of the taxonomy and evolution of actinorhizal symbioses.

Key words: Frankia, Actinorhizal symbiosis, genome, non-nitrogen-fixing frankiae, biosynthetic gene clusters

Introduction

The actinobacterial genus Frankia consists of a diverse group of filamentous, gram-positive soil bacteria that are able to form root nodules with members of eight plant families representing about 25 genera of woody, dicotyledonous, perennial angiosperms [1–3]. Root nodule formation is host plant-specific, and largely correlates with assignments of strains to specific clusters derived from comparative sequence analyses of amplicons of genes such as 16S or 23S rRNA genes [4], nifH genes [5–7], gyrB or glnII [8] and concatenated sequences of some housekeeping genes (dnaA, ftsZ, and pgk) [9, 10]. Clusters 1, 2, and 3 represent Frankia strains that are capable of fixing atmospheric nitrogen (N2) in pure culture and in symbiosis, while members of cluster 4 are typically unable to fix N2, with one exception, and are often not able to form root nodules [4, 11].

While comparative sequence analyses of specific gene amplicons have provided evidence for group-specific assignments of isolates, whole genome sequencing has opened a new avenue to more sophisticated classification of species within the genus Frankia. These analyses include both isolates deposited as type strains in culture collections, and uncultured Frankia populations in root nodules of specific host plants described as candidate species. Frankia alni ACN14aT [12] and Frankia torreyi CpII1T [13], as well as Frankia casuarinae CcI3T [12] and Frankia canadensis ARgP5T [14] have been identified in cluster 1, representing cluster 1a, 1c, and 1d, respectively. Related, but distinct to F. canadensis is Candidatus Frankia nodulisporulans AgTrS, and AgUmAS1 and AgUmASH1 [15] as well as Candidatus Frankia alpina AiOr, and AvVan [15]. Frankia coriariae BMG5.1T [16], Candidatus Frankia datiscae DgI [17], Candidatus Frankia californiensis Dg2 [18] and Candidatus Frankia meridionalis Ccppg1 [19] represent species in cluster 2, while Frankia elaeagni BMG5.12T [12], Frankia discariae BCU110501T [20], Frankia soli NRRL B-16219T [21] and Frankia irregularris G2T [22] are species in cluster 3. Cluster 4 frankiae are represented by three species, Frankia inefficax EuIIcT [23], Frankia asymbiotica M16386T [24] and Frankia saprophytica CN3T [25].

Using comparative sequence analyses of amplicons of an actinobacteria-specific insertion in the 23S rRNA genes of additional cluster 4 frankiae, we identified several strains clustering together but distinct from type strains of cluster 4 [26]. These strains included CN4, CN6, CN7, and CNM7 isolated from Coriaria nepalensis [27], AgW1.1 and AgB1.9 isolated from Alnus glutinosa [28], and a yet
undescribed strain AgB1.8 obtained from the same location as strain AgB1.9. The goal of this study was to use whole genome sequence analyses to assess the viability of our previous amplicon-based analysis, and thus affirm the potential of these strains for new species descriptions.

Materials and Methods

Sample preparation

Seven *Frankia* strains previously identified to represent members of cluster 4 (AgW1.1, AgB1.9, AgB1.8, CN4, CN6, CN7, CNM7) were grown in Defined Propionate Medium (DPM) containing propionate and NH₄Cl as C and N source, respectively [29], at 30°C for two weeks. Cells were harvested by centrifugation (15,000 x g, 5 min), and cell aggregates homogenized by brief sonication (10s at 20% output in a S-450 sonifier, Branson Ultrasonics, Danbury, CT) [30]. After centrifugation, cell pellets were used for DNA extraction using the Sureprep™ Soil DNA Isolation Kit (Fisher Scientific, Houston, TX) with small modifications as described before [31]. Extractions of all samples were done in triplicate, and DNA concentrations measured with a Qubit® 2.0 Fluorometer (Life Technologies, Carlsbad, USA). Library preparation and sequencing was performed at the Microbial Genomics Sequencing Center, Pittsburgh, PA, USA using the Illumina tagmentation protocol and the NextSeq Illumina platform (2 x 150 bp).

Genome assembling

Sequence reads were filtered and trimmed using the default settings of fastp [32], and bbduk (https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/) was used to remove reads with average % GC < 54. Genomes were assembled using SPAdes 3.13.0 [33]. Quast was used to check the quality of the assembled genomes. Comparative genomic analysis

All *Frankia* genomes available in NCBI Genomes (https://www.ncbi.nlm.nih.gov/genome/) in November, 14, 2019 were downloaded for comparative genomic analyses. Anvi’o [34] was used to compare the shared COG functions among *Frankia* genomes and generate a core genome phylogenetic tree. The core genome phylogenetic tree was generated using anvi-gen-phylogenomic-tree with the concatenated alignment of 113 single-copy core genes (present in all genomes) that resulted in a total alignment length of 40,557 amino acids. Fragments of the 16S rRNA gene were recovered from the genomes using the HMM profile Ribosomal RNAs implemented in anvi’o. 16S rRNA gene sequences were aligned using SINA [35] and sequence identity was calculated with MEGAX 3 [36]. Pairwise Average Nucleotide Identity (ANI) was calculated with the fastANI software [37]. Identification of open reading frames (ORFs) was performed using Prokka using the default settings [38] and KEGG orthologous annotation was performed with kofamKOALA using the default settings [39]. We used antiSMASH 5.0 using the default settings [40] to investigate the presence of biosynthetic gene clusters (BGCs) categories across cluster 4 *Frankia* strains.

Results

Phylogenetic analysis of *Frankia* spp. isolates

Assembled genomes ranged from 9.8 Mb and 10.7 Mb and the GC content of the genomes was between 71.5 and 72.7% (Table 1). A phylogenetic tree generated from the concatenated alignment of single-copy orthologous proteins present in all 58 *Frankia* genomes revealed that the closest strains to AgB1.8, AgB1.9, AgW1.1, CNM7, CN4, CN6, CN7 were members of cluster 4, i.e. *Frankia saprophytica* CN3, *Frankia* sp. EUN1h, *Frankia asymbiotica* NRRL B-16386, *Frankia* sp. BMG5.36, *Frankia* sp. DC12 and *Frankia inefficax* EuI1c (Figure 1).

| Strain | RefSeq assembly accession | Number of reads a | Number of scaffolds | N50 scaffold b | Total length | GC (%) c | Protein coding genes |
|--------|--------------------------|-------------------|--------------------|---------------|-------------|---------|---------------------|
| AgB1.8 | GCF_016792105.1          | 4,651,340         | 383                | 46,557        | 9,788,199   | 71.67   | 9,155               |
| AgB1.9 | GCF_016786355.1          | 6,564,124         | 278                | 66,606        | 9,788,632   | 71.65   | 9,075               |
| AgW1.1 | GCF_016786345.1          | 3,546,242         | 558                | 35,357        | 9,961,786   | 71.53   | 9,540               |
| CNM7   | GCF_016786385.1          | 3,665,004         | 1,106              | 15,926        | 10,670,374  | 72.68   | 10,336              |
| CN4    | GCF_016786305.1          | 5,509,032         | 607                | 31,737        | 10,707,582  | 72.68   | 9,888               |
| CN6    | GCF_016792085.1          | 6,112,674         | 397                | 48,718        | 10,191,543  | 72.69   | 9,882               |
| CN7    | GCF_016786325.1          | 7,180,944         | 398                | 51,919        | 10,729,937  | 72.68   | 9,843               |

a Number of reads used in the genome assembling after quality control.

b N50 scaffold is the largest length L of which 50% of all nucleotides are in scaffolds of size at least L.

c Total number of G and C nucleotides in the genome, divided by the total length of the genome.
Figure 1. Maximum likelihood tree from concatenated alignments of 113 single-copy protein (total alignment length of 40,557 amino acids) generated with Anvi’o. Strains sequenced in this study are highlighted. The tree was re-rooted using Cryptoperangium aurantiacum as outgroup. Circles represent branch support values (maximum likelihood bootstrap) ≥ 0.8.
In order to determine if the strains described in this work were potential novel species the average nucleotide identity (ANI) was calculated among all the cluster 4 strains (Figure 2). AgB1.8, AgB1.9, AgW1.1, CNM7, CN4, CN6, CN7 did not show ANI ≥ 95% to any of the reference strains. Strains AgB1.8, AgB1.9, AgW1.1 showed ANI ≥ 99% to each other, which indicates they belong to the same genospecies (genospecies Ag). Strains CNM7, CN4, CN6, CN7 showed ANI ≥ 99% to each other, indicating that they belong to a single genospecies (genospecies CN). Analysis of nearly full length 16S rRNA gene fragment (alignment length 1,202) showed that the identity of cluster 4 frankiae ranged from 99.57% to 100% (Figure 2).

**Functional analysis of Frankia spp. isolates**

We compared the presence of KEGG orthologues (KO) of the two novel genospecies of *Frankia* Ag and CN with other *Frankia* strains (Table S2). As expected, nitrogenase complex genes (*nif*) were absent in all the seven strains sequenced in this study (Table S2). KEGG orthologues common to all cluster 4 genomes and absent in other frankiae included genes involved in the urea transporter system (K11959 to K11963), and the complete tryptophan metabolism pathway (KEGG module M00038). Additionally, KEGG orthologues involved in the biodegradation of xenobiotics were enriched in cluster 4 *Frankia* genomes when compared to other frankiae (Table S2). 415 KEGG orthologues were found in all 58 *Frankia* genomes analyzed in this study, including: alkaline phosphatases (K01113), which are involved in solubilization of soil phosphate (Nautilyal et al., 2000, Hahn et al., 2003); adherence proteins (K12510 and K12511), described to be involved in *Actinobacillus actinomycetemcomitans* host colonization (Kachalany et al. 2001) (Table S2).

Unique KOs of genomes of *Frankia* Ag included feruloyl esterase (K09252) and beta-mannosidase (K01192). KOs found in the genomes of *Frankia* CN genospecies and not found in other cluster 4 genomes included a protein involved nitrate transport (K15577), and a flavohemoglobin (K05916). Additionally, genomes of *Frankia* Ag genospecies had more Type 1 polyketide synthase (PKS) and beta-lactone clusters than other cluster 4 *Frankia*, while *Frankia* CN genospecies had more non-ribosomal peptide synthetase clusters (NRPS) than other cluster 4 frankiae (Figure 3).

**Discussion**

In this study, we compared the genomes of three *Frankia* strains isolated from *Alnus glutinosa* and four strains isolated from *Coriaria nepalensis* with 51 publicly available genomes of *Frankia*. Phylogenomic and ANI analyses confirmed that these strains are related to cluster 4 frankiae, and indicate that the three strains isolated from *Alnus glutinosa* (AgB1.8, AgB1.9 and AgW1.1; Ag genospecies) belong to a novel species, as do the four strains from *Coriaria nepalensis* (CN4, CN6, CN7 and CNM7; CN genospecies). Genome sizes of all strains assembled as Ag and CN genospecies were about 9.8 Mb and 10.7 Mb, respectively. Sizes were comparable to those of other strains representing cluster 4 frankiae, i.e. *F. inefficax* and *F. saprophytica* with 8.8 Mb and 10 Mb, respectively, and much larger than genomes of most cluster 1 and 3 frankiae (7.5 Mb to 7.9 Mb), including *F. casuarinae* (4.9 to 5.6 Mb) and *A. nodulisporulans* (4.9 Mb), as well as *F. coriariae* as cluster 2 representative (5.8 Mb) (see [11] for review). An exception was the genome size of *F. irregularis*, a cluster 3 representative with a genome size similar to our cluster frankiae (9.5 Mb) [22], and of related cluster 3 strain R43 (10.44 Mb) [41] and EAN1pec (9.04 Mb) [42].
Smaller genome sizes in frankiae have been related to genome reductions are associated to reduced saprotrophic potential, but not symbiotic potential, while larger genomes created through genome expansions, often through duplication of genes involved in introducing substrates into central metabolic pathways, allowed frankiae to exploit a large variety of environments [17, 42]. The comparatively large genomes of cluster 3 and cluster 4 Frankia strains therefore suggest an increased saprotrophic potential compared to strains assigned to other clusters. In support of this hypothesis, indigenous cluster 3 frankiae as well as introduced strain EAN1pec have been shown to persist and grow under a broader range of environmental conditions (e.g. plant species, carbon resource or matric
potential) than indigenous cluster 1a frankiae or introduced strains ArI3 (cluster 1a) and CcI3 (cluster 1c) [43, 44]. The results of these studies have been used to suggest that cluster 3 frankiae represent a group of generalists [44]. In addition, cluster 1a frankiae were characteristic populations of later stages in a succession of Frankia populations in soil associated with plant growth and succession. Frankiae of cluster 1b represented by Frankia strain Ag45/Mut15 were characteristic of soils in early stages of plant-mediated organic matter accumulation. Cluster 1c frankiae represented a group of highly-specialized frankiae that depended on the presence of their host plant to grow saprotrophically in soil [44].

Diversity of frankiae in soils vegetated with host plant species such as Alnus glutinosa or non-host plant species such as Betula nigra was found to be restricted to cluster 1 and 3 frankiae, and generally members of cluster 1b dominating [26]. Cluster 1b frankiae have been detected as major populations in several studies, with absolute numbers depending on the sampling depth, physicochemical conditions and the vegetation [43, 45, 46]. So far, cluster 4 frankiae have been detected as major Frankia population only in prairie soils where they represented about 50% of all frankiae, that included clusters 1b, 2 and 3, in native, restored and cultivated black prairie soil [26]. Native and restored prairie were dominated by Sorghastrum nutans (L.) (Indiangrass) and Andropogon gerardii Vitman (Big bluestem), while the cultivated prairie site was adjacent to the native prairie site, but cultivated continuously with corn. These results suggest that grasses are supporting a diverse community of frankiae in prairie soils, with significant positive effects on the abundance of cluster 4 frankiae. Although it has been shown that frankiae can grow in the rhizosphere of grasses [47], it is yet unclear for prairie environments whether these effects are related to rhizosphere effects of dense grass root environments or to organic matter accumulation in top soils.

Isolates representing cluster 4 frankiae have been obtained from a variety of different host plant species, i.e. Coriaria nepalensis, Alnus glutinosa, Elaeagnus umbellata, Datisca cannabina, and Ceanothus americanus (CN3, AgW1.1, Eu1c, DC12 and Ca1I) [27, 28, 48-50]. Cluster 4 frankiae have also been detected in bioassays highlighting that they form an important fraction of all frankiae in wet soils under A. glutinosa [51, 52]. Some progenies of A. glutinosa, however, exhibited natural resistance to infection by these populations with no or only very small nodules formed [28, 53]. The failure to detect nodules has therefore resulted in the assumption that many cluster 4 frankiae are not only ineffective with respect to nitrogen-fixation, but also non-infective with respect to root nodule formation. Cluster 4 frankiae could therefore be characterized as largely or even entirely saprotrophic, with isolates obtained from root nodules as surfaces contaminants [54].

Both genospecies CN and Ag lack genes essential for nitrogen-fixation, as observed for cluster 4 species F. inefficax EU1cT [23] and F. saprophytica CN3T [25], but not for F. asymbiotica M16386T [24]. While F. inefficax EU1cT was observed to form ineffective nodules, none of the other species were able to form nodules on their original host plant. Members of genospecies Ag have been shown to induce small ineffective nodules on their respective host plant A. glutinosa [28] while members of genospecies CN failed to induce nodules on their host plant Coriaria nepalensis [27]. Genomes of genospecies Ag strains harbor genes encoding feruloyl esterases and beta-mannosidases that potentially participate in the degradation of plant cell walls as indicated for some plant fungal pathogens, and could therefore be involved in the plant infection process or in nutrient acquisition [55, 56]. Genes for both flavohemoglobin and nitrite/nitrate transport proteins are unique for genospecies CN strains potentially involved in the conversion of nitric oxide to nitrate [57]. Our analysis supports previous studies that found significant biosynthetic potential of Frankia spp. It has been hypothesized that specialized metabolites, such as the molecules produced by PKS and NRPS gene clusters, might participate in plant-microbial signaling and interactions [58]. While our genome studies confirm our previous expectations and affirm the potential of these strains for new species descriptions, the assessment of ecological differences between our genospecies CN and Ag, and described cluster 4 species requires additional experimental data.

Data Summary

1. Genomes of the strains sequenced in this study from Dr. Dittmar Hahn culture collection and were deposited in the National Center for Biotechnology Information (NCBI), under BioProject Number PRJNA680372. Individual RefSeq assembly accession numbers can be found in Table S1.

2. A list of other Frankia genomes utilized in this study can be found in Table S1. All sequences were downloaded from the NCBI Assembly database.

Supplementary Material

Supplementary tables. http://www.jgenomics.com/v09p0068s1.xlsx
Acknowledgements

The authors acknowledge Quinlan J. Carlos-Shanley for writing a script to parse the genome annotation data.

Competing Interests

The authors have declared that no competing interest exists.

References

1. Schwintzer CR, Tjepkema JD. The biology of Frankia and actinorhizal plants. San Diego: Academic Press; 1990.
2. Huss-Danell K. Actinorhizal symbioses and their N2 fixation. New Phytol. 1997; 136: 375-405.
3. Benson DR, Silvester WB. Biology of Frankia strains, actinomycete symbionts of actinorhizal plants. Microb Rev. 1993; 57: 293-319.
4. Normand P, Osos P, 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. Welsh A, Mirza BS, Rieder JP, Paschke MW, Hahn D. Diversity of Frankia in root nodules of Morella pennsylvanica grown in soils from five continents. Syst Appl Microbiol. 2009; 32: 201-210.
6. Welsh AK, Dawson JO, Gottfried GJ, Hahn D. Diversity of Frankia in root nodules of geographically isolated Arizona alders in central Arizona (USA). Appl Environ Microbiol. 2009; 75: 6913-6918.
7. Mirza BS, Welsh A, Rasul G, Rieder JP, Paschke MW, Hahn D. Variation in Frankia populations of the Elaeagnus host infection group in nodules of six host plant species after inoculation with soil. Microb Ecol. 2009; 58: 384-393.
8. 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. Antonie Leeuw. 2011; 100: 579-587.
9. Pozzi AC, Bautista-Guerrero HH, Nouioui I, Cotin-Galvan L, Pepin R, Fournier P, et al. In planta sporulation phenotype: a major life history trait to understand the evolution of Frankia infective strains. Environ Microbiol. 2015; 17: 3125-3138.
10. Pozzi AC, Bautista-Guerrero HH, Abbs S, Herrera-Belaroussi A, Abrout D, Normand P, et al. Robust Frankia phylogeny, species delineation and intraspecies diversity based on Multi-Locus Sequence Analysis (MLSA) and Single-Locus Strain Typing (SLST) adapted to a Large sample size. Syst Appl Microbiol. 2018; 41: 311-323.
11. Gtari M, Nouioui I, Sarkar I, Ghodhbane-Gtari F, Tisa LS, Sen A, et al. An update on the taxonomy of the genus Frankia Bruchousch, 1886, 174(Al). Anton Leeuw. 2019; 112: 5-21.
12. Nouioui I, Ghodhbane-Gtari F, Montero-Calasanz MD, Gtari M, Vander Heul B, Klenk HP, et al. Candidatus Frankia sp nov., an actinobacterium unable to nodulate its host plant. FEMS Microbiol Ecol. 2017; 92: 137-142.
13. Mirza BS, Welsh A, Hahn D. Saprophytic growth of inoculated Frankia sp. in soil microcosms. FEMS Microbiol Ecol. 2007; 62: 280-289.
14. Eren AM, Kiefl E, Shaber A, Veseli I, Miller SE, Schechter MS, et al. Community-led, integrated, reproducible multi-omics with anvio. Nature Microbiology. 2021; 6: e54-64.
15. Pujic P, Bolotin A, Fournier P, Sorokin A, Lapidus A, Richau KH, et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucl Acids Res. 2020; 48: W58-W7.
16. Smit, Smit, and Grindley, 2013; Ensembl, 2013; Maier, 2013; Smith, 2013; Sturrock, 2013; Tatusova, 2013; Venter, 2013; Wang, 2013; Wallis, 2013; Wendl, 2013; Yandell, 2013; Zhang, 2013.
43. Samant SS, Dawson JO, Hahn D. Growth responses of indigenous *Frankia* populations to edaphic factors in actinorhizal rhizospheres. System Appl Microbiol. 2015; 38: 501-505.

44. Samant S, Dawson JO, Hahn D. Growth responses of introduced *Frankia* strains to edaphic factors. Plant Soil. 2016; 400: 123-132.

45. Samant S, Huo T, Dawson JO, Hahn D. Abundance and relative distribution of *Frankia* host infection groups under actinorhizal *Alnus glutinosa* and non-actinorhizal *Betula nigra* trees. Microb Ecol. 2016; 71: 473-481.

46. Samant S, Amann RI, Hahn D. Evaluation of the 23S rRNA gene as target for qPCR based quantification of *Frankia* in soils. Syst Appl Microbiol. 2014; 37: 229-234.

47. Rönkkö R, Smolander A, Nurmiha-Lassila EL, Haaheta K. *Frankia* in the rhizosphere of nonhost plants: A comparison with root-associated nitrogen-fixing *Enterobacter*, *Klebsiella* and *Pseudomonas*. Plant Soil. 1993; 153: 85-95.

48. Lechevalier MP, Ruan JS. Physiology and chemical diversity of *Frankia* spp. isolated from nodules of *Comptonia peregrina* (L.) Coul. and *Ceanothus americanus* L. Plant Soil. 1984; 78: 15-22.

49. Mirza MS, Janse JD, Hahn D, Akkermans ADL. Identification of atypical *Frankia* strains by fatty acid analysis. FEMS Microbiol Lett. 1991; 83: 91-97.

50. Baker D, Newcomb W, Torrey JG. Characterization of an atypical actinorhizal microsymbiont, *Frankia* sp. EuI1 (Actinomycetales). Can J Microbiol. 1980; 26: 1072-1089.

51. Wolters DJ, Akkermans ADL, Van Dijk C. Ineffective *Frankia* strains in wet stands of *Alnus glutinosa* L. Gaertn. in the Netherlands. Soil Biol Biochem. 1997; 29: 1707-1712.

52. Wolters DJ, Van Dijk C, Zoetendal EG, Akkermans AD. Phylogenetic characterization of ineffective *Frankia* in *Alnus glutinosa* (L.) Gaertn. nodules from wetland soil inoculants. Mol Ecol. 1997; 6: 971-981.

53. Van Dijk C, Sluimer A. Resistance to an ineffective *Frankia* strain type in *Alnus glutinosa* (L.) Gaertn. New Phytol. 1994; 128: 497-504.

54. Vemulapally S, Guerra T, Hahn D. Localization of typical and atypical *Frankia* isolates from *Casuarina* sp. in nodules formed on *Casuarina equisetifolia*. Plant Soil. 2019; 435: 385-393.

55. Xu M, Gao X, Chen J, Yin Z, Feng H, Huang L. The feruloyl esterase genes are required for full pathogenicity of the apple tree canker pathogen Valsa mali. Mol Plant Pathol. 2018; 19: 1353-1363.

56. Kubicek CP, Starr TL, Glass NL. Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. Annu Rev Phytopathol. 2014; 52: 427-451.

57. Frey AD, Kallio PT. Bacterial hemoglobins and flavohemoglobins: versatile proteins and their impact on microbiology and biotechnology. FEMS Microbiol Rev. 2003; 27: 525-545.

58. 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 Microbiol. 2011; 77: 3617-3625.