

University of New Hampshire

University of New Hampshire Scholars' Repository

Molecular, Cellular and Biomedical Sciences
Scholarship

Molecular, Cellular and Biomedical Sciences

9-4-2017

Permanent draft genome sequence of Frankia sp. NRRL B-16219 reveals the presence of canonical nod genes, which are highly homologous to those detected in Candidatus Frankia Dg1 genome

Amir Ktari

Université de Tunis El-Manar

Imen Nouioui

Université de Tunis El-Mana

Teal Furnholm

University of New Hampshire, Durham

Erik Swanson

University of New Hampshire, Durham

Faten Ghodhbane-Gtari

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

Université de Tunis El-Mana

Comments

This is an article published by BioMed Central (BMC) in Standards in Genomic Sciences in 2017, available online:

~~See next page for additional authors~~
<https://dx.doi.org/10.1186/s40793-017-0261-3>

Recommended Citation

Ktari, A., I. Nouioui; T. Furnholm, E. Swanson, F. Ghodhbane-Gtari, L. S Tisa, and M. Gtari. 2017. Permanent draft genome sequence of Frankia sp. NRRL B-16219 reveals the presence of canonical nod genes, which are highly homologous to those detected in Candidatus Frankia Dg1 genome. Standards in Genomic Science 12:e51 doi:10.1186/s40793-017-0261-3

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

Amir Ktari, Imen Nouioui, Teal Furnholm, Erik Swanson, Faten Ghodhbane-Gtari, Louis S. Tisa, and Maher Gtari

EXTENDED GENOME REPORT

Open Access



Permanent draft genome sequence of *Frankia* sp. NRRL B-16219 reveals the presence of canonical *nod* genes, which are highly homologous to those detected in *Candidatus Frankia Dg1* genome

Amir Ktari¹, Imen Nouioui¹, Teal Furnholm², Erik Swanson², Faten Ghodhbane-Gtari¹, Louis S. Tisa² and Maher Gtari^{1*}

Abstract

Frankia sp. NRRL B-16219 was directly isolated from a soil sample obtained from the rhizosphere of *Ceanothus jepsonii* growing in the USA. Its host plant range includes members of *Elaeagnaceae* species. Phylogenetically, strain NRRL B-16219 is closely related to "*Frankia discariae*" with a 16S rRNA gene similarity of 99.78%. Because of the lack of genetic tools for *Frankia*, our understanding of the bacterial signals involved during the plant infection process and the development of actinorhizal root nodules is very limited. Since the first three *Frankia* genomes were sequenced, additional genome sequences covering more diverse strains have helped provide insight into the depth of the pangenome and attempts to identify bacterial signaling molecules like the rhizobial canonical *nod* genes. The genome sequence of *Frankia* sp. strain NRRL B-16219 was generated and assembled into 289 contigs containing 8,032,739 bp with 71.7% GC content. Annotation of the genome identified 6211 protein-coding genes, 561 pseudogenes, 1758 hypothetical proteins and 53 RNA genes including 4 rRNA genes. The NRRL B-16219 draft genome contained genes homologous to the rhizobial common nodulation genes clustered in two areas. The first cluster contains *nodACIJH* genes whereas the second has *nodAB* and *nodH* genes in the upstream region. Phylogenetic analysis shows that *Frankia nod* genes are more deeply rooted than their sister groups from rhizobia. PCR-sequencing suggested the widespread occurrence of highly homologous *nodA* and *nodB* genes in microsymbionts of field collected *Ceanothus americanus*.

Keywords: *Frankia*, Actinorhizal symbiosis, Plant-microbe interactions, Genome, Canonical nod genes, *Ceanothus*

Introduction

The symbiosis resulting from members of the genus *Frankia* interacting with the roots of 8 dicotyledonous plant families (referred to actinorhizal plants) is found worldwide and contributes to the ability of actinorhizal pioneer plants to grow in poor and marginally fertile soils [1]. This symbiotic association has drawn interest because of its higher rate of soil nitrogen input and the ability of the plants to overcome harsh environmental

conditions [2]. The molecular mechanism for the establishment of an actinorhizal nitrogen-fixing root nodule remains elusive [3]. Molecular phylogeny of the *Frankia* genus has consistently identified four main clusters regardless of the typing locus used [1]. These *Frankia* clusters also follow and support the host specificity groups proposed by Baker [4]. Cluster 1 is divided into sub-cluster 1a including *F. alni* and relatives that are infective on *Alnus* and *Myricaceae* and sub-cluster 1b strains that are infective on *Allocasuarina*, *Casuarina* and *Myricaceae* including *F. casuarinae* [5]. Cluster 2 contains *F. coriariae* [6] and uncultured microsymbionts of *Coriariaceae*, *Datisceae*, *Dryadoideae* and

* Correspondence: maher.gtari@fst.rnu.tn

¹Laboratoire Microorganismes et Biomolécules Actives, Université Tunis El Manar (FST) & Université de Carthage (INSAT), 2092 Tunis, Tunisia
Full list of author information is available at the end of the article

Ceanothus, while cluster 3, associated *F. elaeagni* [5], “*F. discariae*” [7] and closely related strains are infective on *Colletiaeae*, *Elaeagnaceae*, *Gymnostoma* and *Myricaceae*. Finally, cluster 4 groups a broad range of non-nitrogen-fixing and infective strains including *F. inefficax* species [8] together with “*F. asymbiotica*” [9] and other related strains that are unable to establish a symbiosis with actinorhizal plants. As has been established for rhizobial and arbuscular mycorrhizal symbioses, the LysM-RLKs are also involved in the perception of *Frankia* signal molecules by the actinorhizal plant [10, 11]. However, the bacterial signals triggering this symbiosis remain unknown. At present, more than 30 *Frankia* genomes from strains in pure culture have been sequenced and annotated [12–30] and two *Candidatus* genomes were generated from nodule metagenomes [31, 32]. Analysis of the *Frankia* genomes failed to reveal the presence of common canonical *nodABC* genes [33] which also appear to be missing in several photosynthetic [34] and non-

photosynthetic [35] bradyrhizobia. The only exceptions were found in the two *Candidatus Frankia* genomes, which contained the canonical *nodABC* and sulfotransferase *nodH* genes [32, 36]. This contradictory situation justifies additional sequencing of genomes from cultivated *Frankia* strains to gain insight into the depth of the pangenome pool covered. Here we report the first proof of the presence of rhizobial homologous canonical *nodABCH* genes within the draft genome of cultivated *Frankia* isolate, strain NRRL B-16219 and widespread occurrence of *nodAB* in field collected *Ceanothus americanus* microsymbionts.

Organism information

Classification and features

Strain NRRL B-16219 metabolizes short fatty acids, TCA-cycle intermediates and carbohydrates (Table 1). It is infective on members of *Elaeagnaceae* and *Morella cerifera* and produces effective root nodules [4, 37]. In

Table 1 Classification and general features of *Frankia* sp. strain NRRL B-16219 according to MIGS [45]

MIGS ID	Property	Term	Evidence code ^a
	Classification	Domain <i>Bacteria</i>	TAS [46]
		Phylum <i>Actinobacteria</i>	TAS [47]
		Class <i>Actinobacteria</i>	TAS [48]
		Order <i>Frankiales</i>	TAS [49]
		Family <i>Frankiaceae</i>	TAS [50, 51]
		Genus <i>Frankia</i>	TAS [52, 53]
		Species <i>Frankia</i> sp.	IDR
		Strain NRRL B-16219IDA	
	Gram stain	Positive	IDA
	Cell shape	Filament-shaped	IDA
	Motility	Non-motile	NAS
	Sporulation	Sporulating	NAS
	Temperature range	25–35 °C	TAS [5]
	Optimum temperature	28 °C	TAS [5]
	pH range; Optimum	pH 6.3 – pH 6.8	NAS
	Carbon source	short fatty acids, TCA-cycle intermediates and carbohydrates	IDA
MIGS-6	Habitat	Soil and Host-associated	IDA
MIGS-6.3	Salinity	Not reported	
MIGS-22	Oxygen requirement	Aerobic	NAS
MIGS-15	Biotic relationship	Free-living and Host plant-associated	NAS
MIGS-14	Pathogenicity	Non-pathogen	NAS
MIGS-4	Geographic location	Soil beneath <i>Ceanothus jepsonii</i> , USA	IDA
MIGS-5	Sample collection	1982	IDA
MIGS-4.1	Latitude	Not reported	-
MIGS-4.2	Longitude	Not reported	-
MIGS-4.4	Altitude	Not reported	-

^a Evidence codes – IDA INFERRED FROM DIRECT ASSAY, TAS traceable author statement (i.e., a direct report exists in the literature) NAS non-traceable author statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence)

coherence with its host range, strain NRRL B-16219 is phylogenetically affiliated to cluster 3, known to effectively nodulate members of *Elaeagnaceae*, *Rhamnaceae* and *Myricaceae* families. Phylogenetic analysis based on 16S rRNA gene sequence showed that strain NRRL B-16219 was most closely related to type strains of “*F. discariae*” DSM 46785^T (99.78%) and *F. elaeagni* (98.26%) (Fig. 1).

Frankia sp. strain NRRL B-16219 shows typical *Frankia* morphological structures; branched hyphae, vesicles, the site of nitrogenase activity, and multilocular sporangia containing non-motile spores (Fig. 2).

Extended feature descriptions

Strain NRRL B-16219 represents one of the rare *Frankia* strains directly isolated from soil on plate medium without passing through plant trapping assay. The strain was isolated from the rhizosphere of *Ceanothus jepsonii* [37] following a complex protocol of soil treatment with phenol (0.7%), sample fractionation through ultracentrifugation in sucrose density gradient, and plating on solid DPM without nitrogen source. Strain NRRL B-16219 developed unpigmented white colonies after 4 weeks growth on DPM medium at 28 °C without shaking. The strain was phenotyped using GENIII microplates in an Omnilog device (BIOLOG Inc., Haywood, USA) as previously described [5]. It was able to metabolize acetic acid, citric acid, D-cellobiose, dextrin, D-fructose, D-mannitol, D-mannose, fructose-6-phosphate, fusidic acid, glucose-6-phosphate, D and L malic acid, *p*-hydroxy-phenylacetic acid, propionic acid and D-serine and to grow in presence of 1% sodium lactate and up to 1% NaCl. Growth occurred between pH 5.0–6.8. The strain showed tolerant only to rifamycin.

Genome sequencing information

Genome project history

Because it is one of the rare strains isolated directly from the soil, NRRL B-16219 strain was selected as part of an effort to gain insight into the depth of the pangenome pool and to identify symbiotic signaling molecules. The sequencing project was completed in April 2016 and the generated data was submitted as draft genome to Genbank under BioProject PRJNA318440 and the accession number MAXA00000000.1.

Growth conditions and genomic DNA preparation

The studied strain was kindly provided by David Labeda, ARS USDA bacterial collection, as NRRL B-16219 strain ID. The strain was grown at 28 °C in stationary culture in 1-l bottles containing DPM medium [5], supplemented with 0.5 mM NH₄Cl as nitrogen source maintained. Biomass from 1 month-old culture was harvested by centrifugation at 9000 x g for 15 min, rinsed several times with sterile distilled water. The mycelial mats were broken by repeated passages through syringes with progressively smaller diameters (21 g to 27 g). Genomic DNA extraction was performed using Plant DNeasy kits (Qiagen, Hilden, Germany) following the recommendation of the manufacturer. Prior to genome sequencing, the quality of the isolated DNA was checked by using the prepared DNA as template for PCR and partial sequences of several housekeeping genes and the 16S rRNA gene were generated and analyzed [16].

Genome sequencing and assembly

Sequencing of the draft genome of *Frankia* sp. NRRL B-16219 was performed at the Hubbard Center for

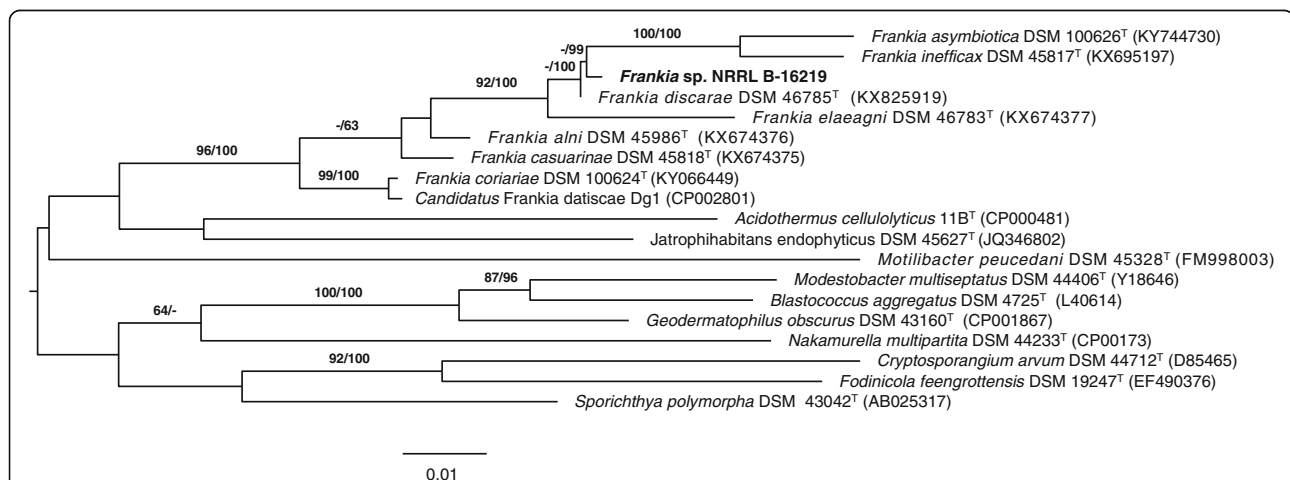
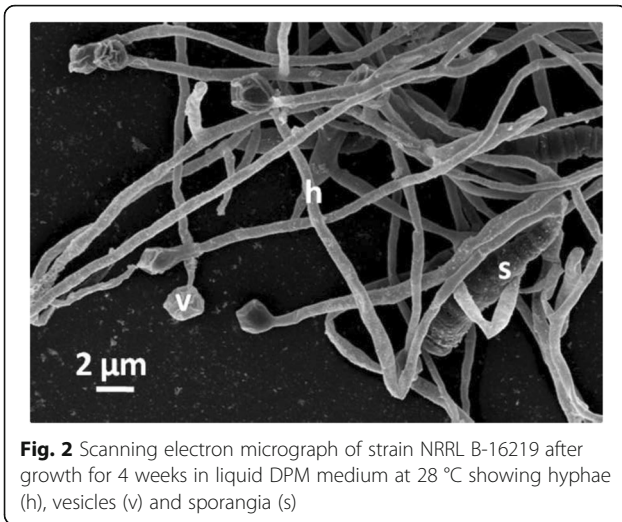


Fig. 1 Maximum likelihood (ML) phylogenetic tree based on the 16S rRNA gene sequences (1400 nt), showing the relationships between *Frankia* NRRL B-16219 and *Frankia* species. The ML tree was inferred using the GTR + GAMMA model and rooted by midpoint-rooting; the branches are scaled in terms of the expected number of substitutions per site. The numbers above the branches are support values when larger than 60% from ML (left) and MP (right) bootstrapping



Genome Studies (University of New Hampshire, Durham, NH) using Illumina technology [38]. A standard Illumina shotgun library was constructed and sequenced using the Illumina HiSeq2500 platform with pair-end reads of 2 × 250 bp. The Illumina sequence data were trimmed by Trimomatic version 0.32 [39], and assembled using Spades version 3.5 [40], and ALLPaths-LG version r52488 [41].

Genome annotation

The genome was annotated via the NCBI Prokaryotic Genome Annotation Pipeline. Additionally nod gene prediction analysis was done within the Integrated Microbial Genomes-Expert Review system developed by the Joint Genome Institute, Walnut Creek, CA, USA [42] developed by the Joint Genome Institute, Walnut

Creek, CA, USA, using similarity search tools. This whole-genome shotgun sequence has been deposited at DDBJ/EMBL/GenBank under the accession number MAXA00000000.1. The version described in this paper is the first version, MAXA00000000.1. A summary of the project information is shown in Table 2.

Genome properties

The draft genome of *Frankia* NRRL B-16219 consisted of 289 DNA contigs that correspond to estimated genome size of 8,032,739 bp and a GC content of 71.7%. The draft genome contained 6859 total genes, including 6211 protein-encoding genes (90.55%), 561 pseudo genes (8.17%) and 53 RNAs (0.76%) (Table 3). Classification of genes into the COG functional categories is shown in Table 4.

Insights from the genome sequence

Comparison of genomes from *Frankia* sp. NRRL B-16219 and other *Frankia* species

The *Frankia* sp. NRRL B-16219 genome was compared to all of the *Frankia* genomes available at NCBI genome database including seven *Frankia* species including *F. alni*, *F. casuarinae*, *F. elaeagni*, *F. coriariae*, “*F. discariae*”, *F. inefficax*, and “*F. asymbiotica*”, two *Candidatus Frankia* and other *Frankia* sp. strains. As shown for other closely related strains from cluster 3, strain NRRL B-16219 has one of the largest genome sizes (8,032,739 bp) with a high GC content of 71.72%. Genes shown or suggested to be involved in the actinorhizal symbiosis were detected. Nitrogenase genes were organized into one operon: *nifH-D-K-E-N-X-orf1-orf2-W-Z-B-U* and a non-linked *nifV* gene. Genes encoding the hydrogenase subunits were clustered into two operons.

Table 2 Project information

MIGS ID	Property	Term
MIGS 31	Finishing quality	Draft genome
MIGS-28	Libraries used	Illumina Standard library
MIGS 29	Sequencing platforms	Illumina HiSeq2500 platform
MIGS 31.2	Fold coverage	120.5x
MIGS 30	Assemblers	Spades version 3.5, ALLPaths-LG version r52488
MIGS 32	Gene calling method	GeneMarkS+ v3.3
	Locus Tag	BBK14_RS02460
	Genbank ID	MAXA00000000.1
	Genbank Date of Release	October 30, 2016
	GOLD ID	Gp0153653
	BIOPROJECT	PRJNA224116
MIGS 13	Source Material Identifier	NRRL B-16219
	Project relevance	Agricultural

Table 3 Genome statistics

Attribute	Value	% of Total ^a
Genome size (bp)	8,032,739	100.0
DNA coding (bp)	6,603,166	82.20
DNA G + C (bp)	5,760,840	71.72
DNA Contigs289100		.0
Total genes	6859	100.0
Protein coding genes	6, 211	91.01
RNA genes	53	0.77
Pseudo genes ^b	561	8.18
Genes in internal clusters	-	-
Genes with function prediction	5046	73.60
Genes assigned to COGs	3609	52.64
Genes with Pfam domains	4735	69.06
Genes with signal peptides	176	2.57
Genes with transmembrane helices	296	4.32
CRISPR repeats	2	-

^aThe total is based on either the size of the genome in base pairs or the total genes in the annotated genome

^bPseudo genes may also be counted as protein coding or RNA genes, so is not additive under total gene count

Genes for two different types of truncated hemoglobins, HbN and HbO, were also present.

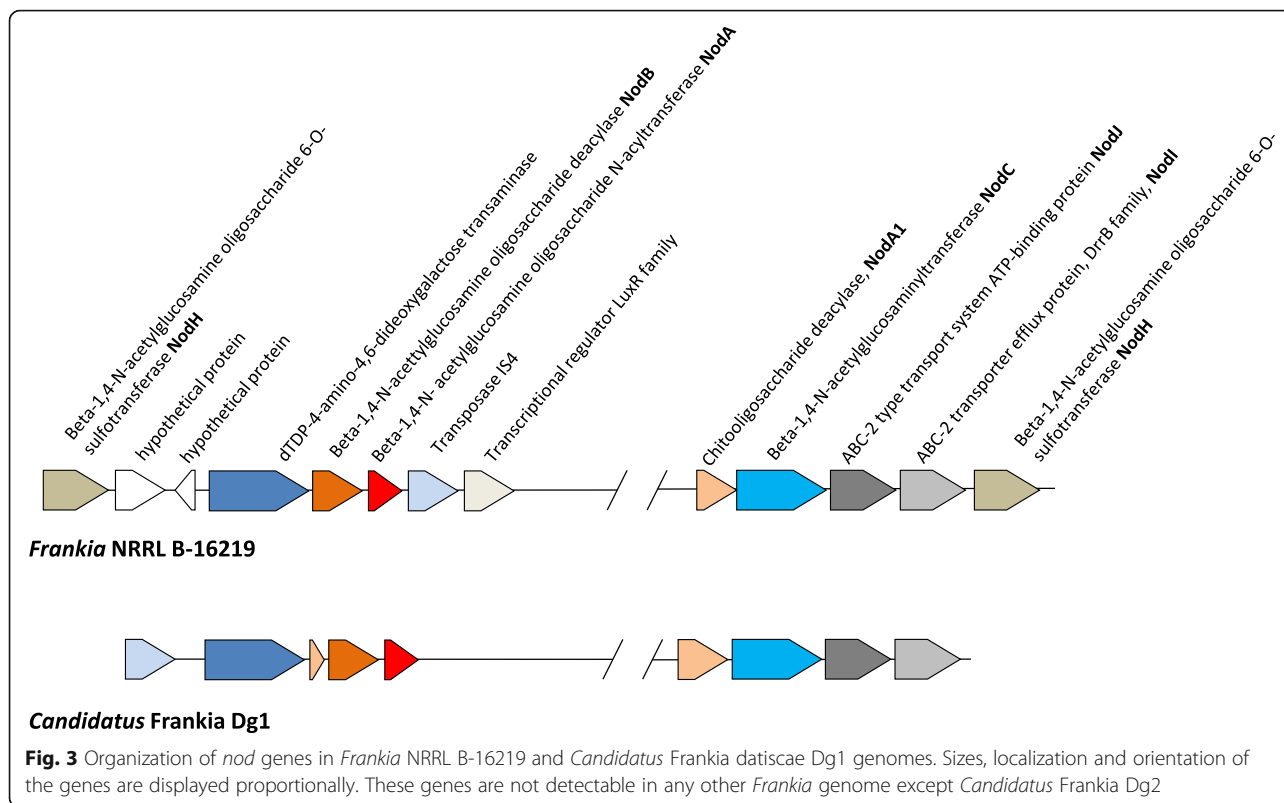
Nodulation pathway

In rhizobia, the common canonical *nodABC* genes playing a key role in triggering root nodule formation in Legumes. These signals are secreted as a reply to host-plant flavonoids perceived by the compatible rhizobial strains [43]. The Nod factors perceived by host plant through the LysM-RLKs, and the resulting signal transduction cascade triggers a bacterial invasion of root cortical cells and the genesis of functional nodules. Despite the presence of these LysM-RLKs in the actinorhizal plants [11], none of the *Frankia* genomes from cultivated strains contained any homologous *nod* genes [33], but they are present in the two *Candidatus* Frankia genomes [32, 36]. Six *nod*-like genes were detected in the NRRL B-16219 draft genome (Additional file 1: Table S1) organized into two regions (Fig. 3). The first cluster contained genes encoding the *nodA1*, *nodC*, ABC-2 type transport system ATP-binding protein (*nodI*), ABC-2 transporter efflux protein, DrrB family

Table 4 Number of genes associated with the general COG functional categories

Code	Value	% age ^a	Description
J	178	4.27	Translation, ribosomal structure and biogenesis
A	1	0.02	RNA processing and modification
K	408	9.79	Transcription
L	109	2.62	Replication, recombination and repair
B	1	0.02	Chromatin structure and dynamics
D	32	0.77	Cell cycle control, cell division, chromosome partitioning
V	135	3.24	Defense mechanisms
T	249	5.98	Signal transduction mechanisms
M	173	4.15	Cell wall/membrane biogenesis
N	21	0.5	Cell motility
U	30	0.72	Intracellular trafficking, secretion, and vesicular transport
O	140	3.36	Posttranslational modification, protein turnover, chaperones
C	250	6	Energy production and conversion
G	207	4.97	Carbohydrate transport and metabolism
E	297	7.13	Amino acid transport and metabolism
F	94	2.26	Nucleotide transport and metabolism
H	262	6.29	Coenzyme transport and metabolism
I	351	8.42	Lipid transport and metabolism
P	210	5.04	Inorganic ion transport and metabolism
Q	256	6.14	Secondary metabolites biosynthesis, transport and catabolism
R	508	12.19	General function prediction only
S	178	4.27	Function unknown
-	3247	47.36	Not in COGs

^aThe total is based on the total number of protein-coding genes in the genome



(*nodI*) and *nodH*. The second cluster contained *nodA*, *nodB* and a *nodH* genes. Amino acid sequence similarities between *Frankia* sp. strain NRRL B-16219 NodA, B, C, and H predicted proteins ranged from 86 to 93% and 57–67% with the uncultured *Frankia* (Dg1 and Dg2) and (α - and β -) rhizobia, respectively (Additional file 2: Table S2). Further phylogenetic analysis (Fig. 4) showed that the *Frankia* Nod proteins were positioned at the root of both the α - and β -rhizobial NodABC proteins as previously reported [4, 8]. They were most closely related to plant nodulating *Betaproteobacteria* of *Burkholderia* and *Paraburkholderia* genera. The GC content of *Frankia nod* genes ranged from 57.9% for *nodA* to 66.37% for *nodB* which is quite similar to that of some rhizobial species including *Methylobacterium* and *Burkoldaria*. For both *Frankia* and rhizobia, GC% of the *nod* genes was lower than that of total genome sequences.

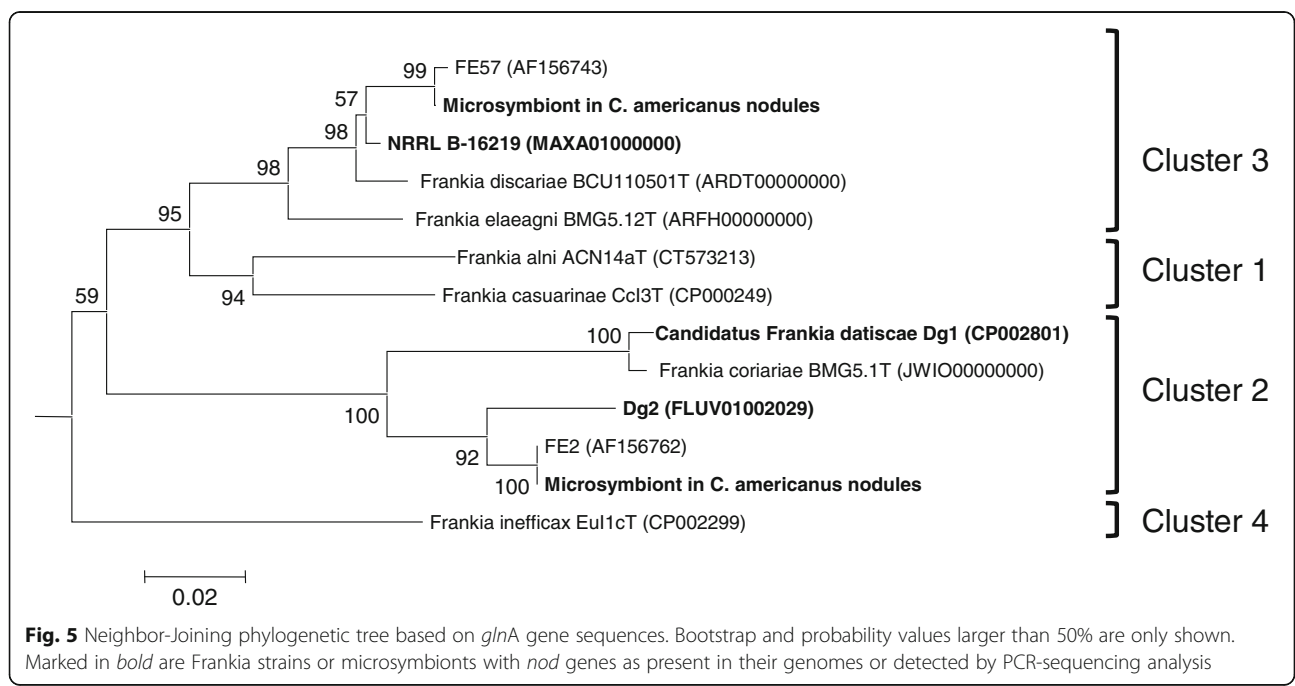
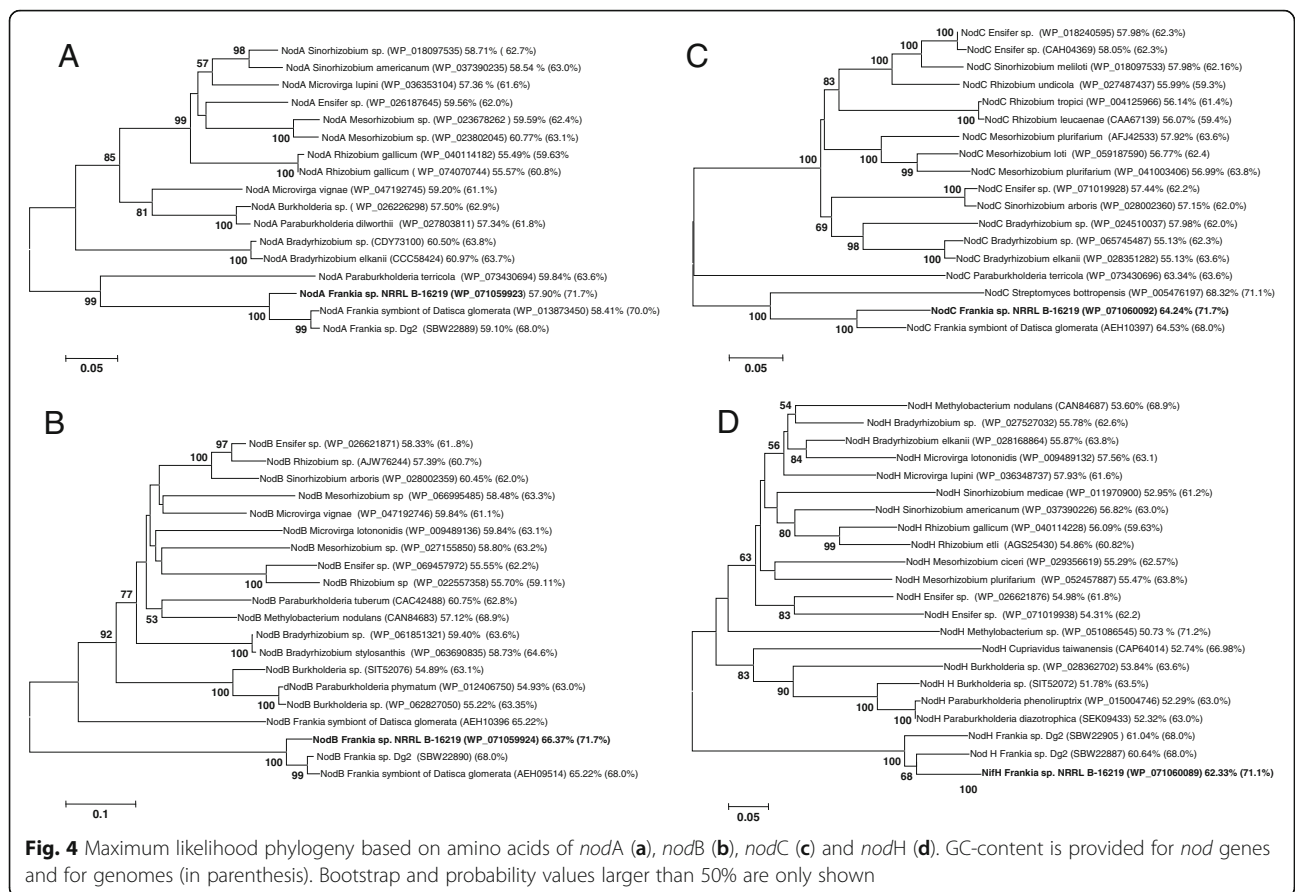
Field collected microsymbionts of *Ceanothus americanus* contain *nod* genes

Root nodules from *Alnus glutinosa*, *Casuarina glauca* and *Elaeagnus angustifolia* growing in Tunisia and *Ceanothus americanus* and *Elaeagnus umbellata* growing in Durham New Hampshire, USA, were collected. The *nodA-nodB* region from *C. americanus* nodules was PCR-amplified and sequenced. Following the alignments of the *nodA* and *nodB* gene sequences of

Dg1 and NRRL B-16219, the primer set (forward primer *nodAF* 5'-AGCGCGACCCGAGCTCAGGATA ATCG-3' and reverse *nodBF* (5'-CGATCCCACCCGG ATGGAGCTGC-3')) was designed in this study. The sequenced PCR-products were translated into amino acid sequences to permit the detection of the 23 aa sequence at the beginning of the 193 aa of the NodA, the intergenic region (160 nucleotides) and 41 aa at the end of the 230aa of the NodB. Both sequences showed 100% sequence similarities to their respective homologous region in NodA (23/193aa) and NodB (41/230aa) protein sequences for *Candidatus Frankia* Dg1. Regardless of their affiliation to cluster 2 or to cluster 3 (Fig. 5), all of the analyzed *C. americanus* microsymbionts contained the *nodAB* genes. In contrast, *A. glutinosa*, *C. glauca*, *E. umbellata* and *E. angustifolia* microsymbionts failed to amplify the expected PCR product. This result is in congruence with previous reports claiming that no homologous *nod* genes are retrievable in sequenced genomes from strains isolated from these actinorrhizal plant species [33].

Conclusions

We report here the genome sequence of a *Frankia* strain directly isolated from soil rhizosphere. The generated draft genome was assembled into 289 contigs corresponding to 8,032,739 bp, which falls within the size range of *Frankia* cluster 3 [33]. Bacterial factors triggering actinorrhizal



symbiosis remain enigmatic since many sequenced *Frankia* genomes have revealed the absence of universal nod-factors. It was hypothesized that most *Frankia* strains use a novel nod-independent pathway for the infection process of actinorhizal plants. In contrast, two *Candidatus Frankia* Dg1 and Dg2 genomes contain canonical nod genes [32, 36]. Here we provide the first proof for the presence of *nod* genes in the genome of a cultivated *Frankia* strain. In addition, a PCR-sequencing approach suggested that *nod* genes are only widespread in *C. americanus* microsymbionts. This situation is similar to legume symbionts where two nodulation pathways are described: the well-studied nod-dependent and an alternative nod-independent pathway. The majority of rhizobia use the nod-dependent pathway, while some photosynthetic [34] and non-photosynthetic [35] bradyrhizobia use the alternative nod-independent pathway. Moreover, some rhizobia use both pathways and the use of the nod-independent pathway seems to be highly dependent on host species rather than the presence or absence of *nod* genes in a given bradyrhizobial genome [44]. For *Frankia*, almost all host plants are infected through the nod-independent pathway, while the nod-dependent process may only be present in unstudied actinorhizal species such as members of the genus *Ceanothus*.

Additional files

Additional file 1: Table S1. Localizations and DNA coordinates for nod genes in NRRL B16219 and Dg1 genomes. (DOCX 12 kb)

Additional file 2: Table S2. Percent similarities based on amino acid sequence for NodA, B, C and H between *Frankia* sp. NRRL B-16219, *Candidatus Frankia* Dg1 and Dg2, and other rhizobial strains. (XLSX 308 kb)

Abbreviations

DPM: Defined Propionate Medium; *glnA*: Glutamine synthetase; LysM-RLKs: LysM-receptor-like kinases; *nifH*: Nitrogenase reductase; *nodA*: Beta-1,4-N-acetylglucosamine oligosaccharide N-acyltransferase; *nodA1*: Beta-1,4-N-acetylglucosaminyltransferase; *nodB*: Beta-1,4-N-acetylglucosamine oligosaccharide deacylase; *nodC*: Beta-1,4-N-acetylglucosaminyltransferase; *nodH*: Beta-1,4-N-acetylglucosamine oligosaccharide 6-O-sulfotransferase

Acknowledgements

This work was supported by the Laboratoire Microorganismes & Biomolécules Actives, Université Tunis El-Manar, Tunisia (grant LR03ES03) (MG), and the USDA National Institute of Food and Agriculture Hatch 022821 (LST). Sequencing was performed on an Illumina HiSeq2500 purchased with an NSF MRI Grant: DBI-1229361 to WK Thomas. The authors would like to thank Dr. David P. Labeda curator at Agricultural Research Service Culture Collection for providing *Frankia* strain and Dr. Manfred Rhode (HZI-Helmholtz Centre for Infection Research, Germany) for performing scanning electron micrograph for the studied strain.

Authors' contributions

MG conceived and designed the work. LST and ES performed the complete genome production, including genome assembly and GenBank submission. AK prepared the DNA isolation and the cultivation of *Frankia* strain NRRL B-16219. TF and MG sampled *C. americanus* and *E. umbellata* nodules. AK, IN and FGG sampled root nodules from *A. glutinosus*, *C. glauca* and *E. angustifolia*, performed DNA extraction from all sampled nodules, amplification and sequencing of *glnA*

and *nodA-B* genes. MG drafted the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Laboratoire Microorganismes et Biomolécules Actives, Université Tunis El Manar (FST) & Université de Carthage (INSAT), 2092 Tunis, Tunisia.

²Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire, 289 Rudman Hall, 46 college Road, Durham, NH 03824-2617, USA.

Received: 22 February 2017 Accepted: 22 August 2017

Published online: 04 September 2017

References

- Gtari M, Tisa LS, Normand P. Diversity of Frankia strains, actinobacterial symbionts of actinorhizal plants. In: Symbiotic Endophytes. Springer Berlin Heidelberg; 2013. p. 123–48.
- Dommergues YR. Nitrogen fixation by trees in relation to soil nitrogen economy. Fertil Res. 1995;4:215–30.
- Ktari A, Gueddou A, Nouioui I, Miotello G, Sarkar I, Ghodhbane-Gtari F, Sen A, Armengaud J, Gtari M. Host Plant Compatibility Shapes the Proteogenome of *Frankia coriariae*. Front Microbiol. 2017;8:720.
- Baker DD. Relationships among pure cultured strains of *Frankia* based on host specificity. Physiol Plant. 1987;70:245–8.
- Nouioui I, Ghodhbane-Gtari F, del Carmen M-CM, Göker M, Meier-Kolthoff JP, Schumann P, Rohde M, Goodfellow M, Fernandez MP, Normand P, Tisa LS, Klenk H-P, Gtari M. 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.
- Nouioui I, Ghodhbane-Gtari F, Rohde M, Klenk HP, Gtari M. *Frankia coriariae* sp. nov., an infective and effective microsymbiont isolated from *Coriaria japonica*. Int J Syst Evol Microbiol. 2017;67(5):1266–70.
- Nouioui I, Del Carmen M-CM, Ghodhbane-Gtari F, Rohde M, Tisa LS, Klenk HP, Gtari M. *Frankia discariae* sp. nov.: an infective and effective microsymbiont isolated from the root nodule of *Discaria trinervis*. Arch Microbiol. 2017;199(5):641–7.
- Nouioui I, Ghodhbane-Gtari F, Del Carmen M-CM, Rohde M, Tisa LS, Gtari M, Klenk HP. *Frankia inefficax* sp. nov., an actinobacterial endophyte inducing ineffective, non nitrogen-fixing, root nodules on its actinorhizal host plants. Antonie Van Leeuwenhoek. 2017;110(3):313–20.
- Nouioui I, Gueddou A, Ghodhbane-Gtari F, Rhode M, Gtari M, Klenk HP. *Frankia asymbiotica* sp. nov., a non infective actinobacterium isolated from *Morella californica* root nodule. Int J Syst Evol Microbiol. 2017. doi:10.1099/ijsem.0002153.
- Oldroyd GE. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol. 2013;11(4):252–63.
- Svistoonoff S, Benabdoun FM, Nambiar-Veetil M, Imanishi L, Vaissayre V, Cesari S, Diagne N, Hoher V, de Billy F, Bonneau J, et al. The independent acquisition of plant root nitrogen-fixing symbiosis in Fabids recruited the same genetic pathway for nodule organogenesis. PLoS One. 2013;8(5):e64515.
- D'Angelo T, Oshone R, Abebe-Akele F, Simpson S, Morris K, Thomas WK, Tisa LS. Permanent Draft Genome Sequence of *Frankia* sp. Strain BR, a Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of *Casuarina equisetifolia*. Genome Announc. 2016;4(5).
- D'Angelo T, Oshone R, Abebe-Akele F, Simpson S, Morris K, Thomas WK, Tisa LS. Permanent Draft Genome Sequence for *Frankia* sp. Strain EI5c, a Single-Spore Isolate of a Nitrogen-Fixing Actinobacterium, Isolated from the Root Nodules of *Elaeagnus angustifolia*. Genome Announc. 2016;4(4).
- Ghodhbane-Gtari F, Beauchemin N, Bruce D, Chain P, Chen A, Walston Davenport K, Deshpande S, Detter C, Furnholm T, Goodwin L, et al. Draft genome sequence of *Frankia* sp. strain CN3, an atypical, noninfective (Nod-)

- ineffective (Fix-) isolate from *Coriaria nepalensis*. *Genome Announc.* 2013; 1(2):e0008513.
15. Ghodhbane-Gtari F, Hurst SGT, Oshone R, Morris K, Abebe-Akele F, Thomas WK, Ktari A, Salem K, Gtari M, Tisa LS. 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 Announc.* 2014;2(3).
 16. Gtari M, Ghodhbane-Gtari F, Nouioui I, Ktari A, Hezbri K, Mimouni W, Sbissi I, Ayari A, Yamanaka T, Normand P, et al. Cultivating the uncultured: growing the recalcitrant cluster-2 *Frankia* strains. *Sci Rep.* 2015;5:13112. doi:10.1186/s40793-017-0261-3.
 17. Hurst SGT, Oshone R, Ghodhbane-Gtari F, Morris K, Abebe-Akele F, Thomas WK, Ktari A, Salem K, Mansour S, Gtari M, 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 Announc.* 2014;2(3).
 18. Mansour SR, Oshone R, Hurst SGT, Morris K, Thomas WK, Tisa LS. Draft Genome Sequence of *Frankia* sp. Strain Ccl6, a Salt-Tolerant Nitrogen-Fixing Actinobacterium Isolated from the Root Nodule of *Casuarina cunninghamiana*. *Genome Announc.* 2014;2(1).
 19. Ngom M, Oshone R, Hurst SGT, Abebe-Akele F, Simpson S, Morris K, Sy MO, Champion A, Thomas WK, Tisa LS. Permanent Draft Genome Sequence for *Frankia* sp. Strain CeD, a Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of *Casuarina equisetifolia* Grown in Senegal. *Genome Announc.* 2016;4(2).
 20. Normand P, Lapiere P, Tisa LS, Gogarten JP, Alloiso N, Bagnarol E, Bassi CA, Berry AM, Bickhart DM, Choise N, et al. Genome characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant biogeography. *Genome Res.* 2007;17(1):7–15.
 21. Nouioui I, Beauchemin N, Cantor MN, Chen A, Detter JC, Furnholm T, Ghodhbane-Gtari F, Goodwin L, Gtari M, Han C, et al. Draft Genome Sequence of *Frankia* sp. Strain BMG5.12, a Nitrogen-Fixing Actinobacterium Isolated from Tunisian Soils. *Genome Announc.* 2013;1(4).
 22. Nouioui I, Gtari M, Goker M, Ghodhbane-Gtari F, Tisa LS, Fernandez MP, Normand P, Huntemann M, Clum A, Pillay M, 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 Rhamales. *Genome Announc.* 2016;4(3).
 23. Oshone R, Hurst SGT, Abebe-Akele F, Simpson S, Morris K, Thomas WK, Tisa LS. Permanent Draft Genome Sequences for Two Variants of *Frankia* sp. Strain Cpl1, the First *Frankia* Strain Isolated from Root Nodules of *Comptonia peregrina*. *Genome Announc.* 2016;4(1).
 24. Oshone R, Ngom M, Abebe-Akele F, Simpson S, Morris K, Sy MO, Champion A, Thomas WK, Tisa LS. Permanent Draft Genome Sequence of *Frankia* sp. Strain Allo2, a Salt-Tolerant Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of *Allocastraria*. *Genome Announc.* 2016;4(3).
 25. Pujic P, Bolotin A, Fournier P, Sorokin A, Lapidus A, Richau KH, Briolay J, Mebarki F, Normand P, Sellstedt A. Genome Sequence of the Atypical Symbiotic *Frankia* R43 Strain, a Nitrogen-Fixing and Hydrogen-Producing Actinobacterium. *Genome Announc.* 2015;3(6).
 26. Sen A, Beauchemin N, Bruce D, Chain P, Chen A, Walston Davenport K, Deshpande S, Detter C, Furnholm T, Ghodhbane-Gtari F, et al. Draft genome sequence of *Frankia* sp. strain QA3, a nitrogen-fixing actinobacterium isolated from the root nodule of *Alnus nitida*. *Genome Announc.* 2013;1(2):e0010313.
 27. Swanson E, Oshone R, Simpson S, Morris K, Abebe-Akele F, Thomas WK, Tisa LS. Permanent Draft Genome Sequence of *Frankia* sp. Strain ACN1ag, a Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of *Alnus glutinosa*. *Genome Announc.* 2015;3(6).
 28. Swanson E, Oshone R, Simpson S, Morris K, Abebe-Akele F, Thomas WK, Tisa LS. Permanent Draft Genome Sequence of *Frankia* sp. Strain Avcl1, a Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of *Alnus viridis* subsp. *crispa* Grown in Canada. *Genome Announc.* 2015;3(6).
 29. Tisa LS, Beauchemin N, Cantor MN, Furnholm T, Ghodhbane-Gtari F, Goodwin L, Copeland A, Gtari M, Huntemann M, Ivanova N, et al. Draft Genome Sequence of *Frankia* sp. Strain DC12, an Atypical, Noninfective, Ineffective Isolate from *Datisca cannabina*. *Genome Announc.* 2015;3(4).
 30. Wall LG, Beauchemin N, Cantor MN, Chaia E, Chen A, Detter JC, Furnholm T, Ghodhbane-Gtari F, Goodwin L, Gtari M, et al. Draft Genome Sequence of *Frankia* sp. Strain BCU110501, a Nitrogen-Fixing Actinobacterium Isolated from Nodules of *Discaria trinevis*. *Genome Announc.* 2013;1(4).
 31. Persson T, Benson DR, Normand P, Vanden Heuvel B, Pujic P, Chertkov O, Teshima H, Bruce DC, Detter C, Tapia R, et al. Genome sequence of “*Candidatus Frankia daticae*” Dg1, the uncultured microsymbiont from nitrogen-fixing root nodules of the dicot *Datisca glomerata*. *J Bacteriol.* 2011;193(24):7017–8. doi:10.1186/s40793-017-0261-3.
 32. Nguyen TV, Wibberg D, Battenberg K, Blom J, Vanden Heuvel B, Berry AM, Kalinowski J, Pawlowski K. An assemblage of *Frankia* Cluster II strains from California contains the canonical nod genes and also the sulfotransferase gene nodH. *BMC Genomics.* 2016;17(1):796.
 33. 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(1-3):5–16.
 34. Giraud E, Moulin L, Vallenet D, Barbe V, Cytryn E, Avarre JC, Jaubert M, Simon D, Cartieux F, Prin Y, et al. Legumes symbioses: absence of Nod genes in photosynthetic bradyrhizobia. *Science.* 2007;316(5829):1307–12.
 35. Miche L, Moulin L, Chaintreuil C, Contreras-Jimenez JL, Munive-Hernandez JA, Del Carmen V-HM, Crozier F, Bena G. Diversity analyses of *Aeschynomene* symbionts in Tropical Africa and Central America reveal that nod-independent stem nodulation is not restricted to photosynthetic bradyrhizobia. *Environ Microbiol.* 2010;12(8):2152–64.
 36. Persson T, Battenberg K, Demina IV, Vigil-Stenman T, Vanden Heuvel B, Pujic P, Facciotti MT, Wilbanks EG, O'Brien A, Fournier P, et al. *Candidatus Frankia daticae* Dg1, the Actinobacterial Microsymbiont of *Datisca glomerata*, Expresses the Canonical nod Genes nodABC in Symbiosis with Its Host Plant. *PLoS One.* 2015;10(5):e0127630.
 37. Baker D, O'Keefe D. A modified sucrose fractionation procedure for the isolation of frankiae from actinorhizal root nodules and soil samples. *Plant Soil.* 1984;78:23–8.
 38. Bennett S. Solexa Ltd. *Pharmacogenomics.* 2004;5(4):433–8.
 39. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–20.
 40. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Pribelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, et al. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J Comput Biol.* 2013;20(10):714–37.
 41. Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *PNAS.* 2011;108(4):1513–8.
 42. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NCIMGER. A system for microbial genome annotation expert review and curation. *Bioinformatics.* 2009;25:2271–8.
 43. Lerouge P, Roche P, Faucher C, Maillat F, Truchet G, Prome JC, Denarie J. Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature.* 1990; 344(6268):781–4.
 44. Chaintreuil C, Arrighi JF, Giraud E, Miche L, Moulin L, Dreyfus B, Munive-Hernandez JA, Villegas-Hernandez Mdel C, Bena G. Evolution of symbiosis in the legume genus *Aeschynomene*. *New Phytol.* 2013;200(4):1247–59.
 45. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol.* 2008;26(5):541–7.
 46. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A.* 1990;87(12):4576–9.
 47. Goodfellow M. *Phylum XXVI. Actinobacteria phyl. nov. Bergey's Manual of Systematic Bacteriology, vol 5; 2012. p. Part A: 33. doi:10.1007/978-0-387-68233-447.*
 48. Stackebrandt E, Rainey FA, Ward-Rainey NL. Proposal for a new hierarchic classification system, *Actinobacteria* classis nov. *Int J Syst Bacteriol.* 1997;47: 479–91.
 49. Sen A, Daubin V, Abrouk D, Gifford I, Berry AM, Normand P. Phylogeny of the class Actinobacteria revisited in the light of complete genomes. The orders ‘Frankiales’ and Micrococcales should be split into coherent entities: proposal of Frankiales ord. nov., Geodermatophilales ord. nov., Acidothermales ord. nov. and Nakamurellales ord. nov. *Int J Syst Evol Microbiol.* 2014;64(Pt 11):3821–32.
 50. Becking JH. *Frankiaceae* fam. nov. (Actinomycetales) with one new combination and six new species of the genus *Frankia* Brunchorst. *Int J Syst Bacteriol.* 1970;20:201–20.
 51. Normand P, Orso S, Cournoyer B, Jeannin P, Chapelon C, Dawson J, Evtushenko L, Misra AK. Molecular phylogeny of the genus *Frankia* and

related genera and emendation of the family Frankiaceae. *Int J Syst Bacteriol.* 1996;46(1):1–9.

52. Brunchorst J. Über einige Wurzelanschwellungen, besonders diejenigen von *Alnus* und den *Elaeagnaceen*. *Unters Bot Inst Tübingen.* 1886;2:151–77.
53. Skerman VBD, McGowan V, Sneath PHA. Approved lists of bacterial names. *Int J Syst Evol Microbiol.* 1980;30:225–420.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

