An update on the taxonomy of the genus *Frankia* Brunchorst, 1886, 174^{AL}

Maher Gtari^{1*}, Imen Nouiou², Indrani Sarkar³, Faten Ghodhbane-Gtari^{1,4}, Louis S Tisa⁵,
Arnab Sen³, Hans-Peter Klenk²

¹ Institut National des Sciences Appliquées et de Technologie, Université Carthage, Centre Urbain Nord, BP 676-1080 Tunis Cedex, Tunisia.

² School of Natural and Environmental Sciences, Newcastle University, Ridley Building 2, Newcastle upon Tyne, NE1 7RU, United Kingdom.

³ NBU Bioinformatics Facility, Department of Botany, University of North Bengal, Siliguri 734013, India

⁴ Laboratoire Microorganismes et Biomolécules Actives, Université Tunis El Manar, 2092 Tunis, Tunisia.

⁵ Department of Molecular, Cellular & Biomedical Sciences, University of New Hampshire, 46 College Road, Durham, NH 03824-2617, USA

Corresponding author: Maher Gtari maher.gtari@insat.rnu.tn

Abstract

Since the recognition of the name *Frankia* in the approved lists of bacterial names (1980), few amendments were given to the genus description. Successive editions of Bergey's Manual of Systematics of Archaea and Bacteria have broadly conflicting suprageneric treatments of the genus without any advance for subgenera classification. This review focuses on the recent results from taxongenomics and phenoarray approaches to the positioning and the structuring the genus *Frankia*. Based on phylogenomic treeing, *Frankia* is no longer hampered within the artificial assembly and is now the single member of the monophyletic order, *Frankiales*, within *Frankiaceae* family. Polyphasic strategy incorporating genome to genome data, omniLog® phenoarray together with classical approaches, has allowed the designation of a type strain and an amended description of the type species *Frankia alni* and the recognition of up to 10 novel species covering symbiotic and non symbiotic taxa within the genus. Genome to phenome data will be shortly incorporated for proposing novel species including those recalcitrant to isolation in axenic culture.

Introduction

In the last edition of the Bergey's Manual of Systematic Bacteriology, *Frankia* was treated as the sole genus found in the family *Frankiaceae* and within the order *Frankiales* a polyphenetic assemblage that was marginally supported by 16S rRNA gene sequences (Normand and Benson 2015). The genus descriptions was consistently the same as it appeared in previous editions in particular first reported by Lechevalier and Lechevalier (1989). While they denoted that 16S rRNA gene sequence based phylogeny of the genus *Frankia* (Normand et al. 1996) together with several DNA:DNA relatedness studies (Akimov and Dobritsa 1992; Akimov et al. 1991; Fernandez et al. 1989; Lumini et al. 1996) have been shown to marginally overlap phenotypic-based grouping based on morphology, chemotaxonomic properties (Lechevalier 1994), host specificity (Baker 1987) and mode of infection (Berry and Sunell 1990), physiological traits (Dobritsa 1998; Lechevalier 1994). Even so they clearly indicated the heterogeneity nature of the genus *Frankia*. The limited number of strains studied and the variability of strains from study to study have not moved forward these groupings to species delineation (Gtari et al. 2013).

From 2016 onwards, the availability of many genomes (Tisa et al. 2016) covering almost up-to-date observed diversity (Gtari et al. 2013), and the incorporation of taxongenomics and phenoarray, have helped long-expected species definition according to conventional nomenclature. Several classical and time-consuming bacteriological techniques for polyphasic taxonomy have been revolutionized. Genome-to-genome alignments allow phylogenomic treeing, digital calculation of whole-genome average nucleotide identity (ANI) (Konstantinidis and Tiedje 2005) and digital DNA:DNA hybridization (dDDH) (Auch et al. 2010). The threshold for 16S rRNA gene similarity for delineating a bacterial species was also revised in light of taxongenomics (Meier-Kolthoff et al. 2013). Thus, a threshold of 99.0 % (with a maximum probability of error of 1.0 %) for 16S rRNA sequence similarities and DDH values above 70 % cut-off point are recommended for species delineation (Wayne et al. 1987). The omniLog® phenotype microarray, allows the assessment of cultural and physiological properties of cells based on their ability to reduce an indicator dye rather than estimation of their growth (Miller et al. 1991). Polyphasic strategy incorporating these two approaches together with chemotaxonomy and host specificity, has allowed speciation of the genus *Frankia*.

Frankia Brunchorst 1886, 174^{AL} represents the type genus that belongs to the monogeneric family *Frankiaceae* (Becking 1970) which is the only member of the order *Frankiales* as amended by Sen et al. (2014). It encompasses heterogeneous assemblage of actinobacteria that may be asymbiotic or facultative symbiotic of taxonomically disparate host plant species

designed actinorhizal plants (Lechevalier 1993). The root structures that result are designated; nodules or actinorhiza where the actinobacteria are housed and fed, fix nitrogen and make it accessible to the host plants (Fig. 1). After the successful isolation of *Frankia* strains from 1978 onwards (Callaham et al. 1978), microsymbionts of several actinorhizal plants have been readily isolated in axenic condition (Lechevalier and Lechevalier 1990) and fewer strains have been directly isolated from soil without passing through plant trapping assay and following a complex protocol (Baker and O’Keefe 1984). Actually *Frankia* strains are, virtually, all cultivable in axenic condition including those yet considered as uncultivable or obligate microsymbionts. Genome analysis and physiological bioassays has recently resulted in the isolation of previously recalcitrant strains (Gtari et al. 2015; Gueddou et al. 2018) opening the route to extend the approach for the isolation of several other *Frankia* strains.

Frankia genus is Gram-positive, aerobic, chemoorganoheterotroph and may fix dinitrogen in culture and in association with the host plant. Colonies (Fig. 2a) are ovoid or circular with a network of highly branched hyphae that can be compact or diffuse in liquid and solid media. Morphologically, most strains are developmentally complex and are characterized by three morphological cell structures. Rudimentary to highly branched hyphae (0.3–0.5 µm in diameter) which are the primary vegetative state that may bear with different extend two additional unique developmental structures either terminally or in an intercalary position on the hyphae sporangia and vesicles (Fig. 2b-c). Sporangia are multilocular, ovate to spherical (5–100 µm in diameter), containing non-motile spores and that may be abundant for some species and very large size up to 50 µm in diameter (Normand and Lalonde 1983; Normand et al. 2018) or completely suppressed for others (Gtari et al. 2015; Nouioui et al. 2017a; Gueddou et al. 2018). Vesicles called also diazovesicles are walled swollen (0.6–2.0 µm diameter) represent the site of nitrogen fixation produced under aerobic conditions of low nitrogen availability. They may be present in the presence of fixed nitrogen source for some strains or totally absent in some asymbiotic strains (Baker et al. 1980; Nouioui et al. 2017c).

Current chemotaxonomy is in good agreement with previously reports (for a review see Lechevalier 1994). Cell wall composition is of type III containing *meso*-diaminopimelic acid, alanine, glucosamine, glutamic acid and muramic acid (Lechevalier et al. 1982; Lechevalier and Lechevalier 1990; Nouioui et al. 2016a). Glucose, galactose, mannose and ribose with other species-specific sugars are present in whole-organism hydrolysates (Lechevalier and Lechevalier 1979; Lechevalier et al. 1983; Lechevalier and Ruan 1984; Nouioui et al. 2016a-2018c, Normand et al. 2018). Phospholipid pattern PI consists of phosphatidylinositol (PI), phosphatidylglycerol (PG), diphosphatidylglycerol (DPG) (Lechevalier et al. 1983 ; Lechevalier and Lechevalier 1979 ; Nouioui et al. 2016a-2018c, Normand et al. 2018). Fatty acids include *iso*-C_{16:0}, C_{17:1} ω8c (Simon et al. 1989; Nouioui et al. 2016a-2018c; Normand et al. 2018). Menaquinones are predominantly MK-9(H₈), MK-9(H₄) and MK-9(H₆)

(Lechevalier 1994; Nouioui et al. 2016a-2018c, Normand et al. 2018). Genome sizes range from 4.9 to 11.2 Mb with 67.6 to 72.8 % CG content (Tisa et al. 2016; Nouioui et al. 2016a-2018c, Normand et al. 2018).

The two physiological groups A and B previously proposed (for a review see Lechevalier, 1994) are recoverable with recent cultural and metabolic peculiarities revealed by omniLog® microarray phenotyping (Nouioui et al. 2016a-2018c). Group A strains lack one or more morphological and/or symbiotic features (asymbiotic), relatively fast growing and metabolically diverse with an array of hydrolytic activities and carbon source uses. Group B species fulfill Koch's postulates and are able to re-infect their original host and/or other actinorhizal plants (symbiotic), are slow-growing and physiologically more restricted than group A.

Molecular methods have been extensively used to characterize cultured but also uncultured strains in root nodules and soil (Hahn et al. 1999) and have magnificently helped our understanding of the genus diversity. Molecular phylogeny consistently identifies three clusters among symbiotic strains and one cluster of asymbiotic strains (Ghodhbane-Gtari et al. 2010; Gtari et al. 2004; 2015; Normand et al. 1996; Nouioui et al. 2011; Tisa et al. 2016) (Fig. 3). Cluster 1 subdivided into three cluster 1a-c. Cluster 1a includes *Frankia alni*, *Frankia torreyi* and relatives and are infective on *Alnus* and *Myricaceae*. Cluster 1b contains *Frankia canadensis* and relatives that are infective and effective on *Alnus* and *Myricaceae*. *Frankia casuarinae* (Cluster 1c) effectively infects members of *Allocasuarina*, *Casuarina*, and *Myricaceae* species. Cluster 2 associates the infective and effective species *Frankia coriariae* together with other yet uncultured microsymbionts of *Coriariaceae*, *Datisceae*, *Dryadoideae* and *Ceanothus*. *Frankia elaeagni*, *Frankia discariae*, *Frankia irregularis* and other related strains are members of cluster 3 and effectively nodulate *Colletieae*, *Elaeagnaceae*, *Gymnostoma* and *Myricaceae* species. The fourth cluster includes asymbiotic actinorhizal isolates which lack one or more morphological and/or symbiotic features (Lechevalier 1996).

A restructuration of the order *Frankiales* based on phylogenomic treeing using 54 concatenated conserved proteins retrieved from genomes of 35 families and 17 orders of the class *Actinobacteria* permitted to amend the order *Frankiales* as monophyletic assemblage to only contain *Frankiaceae* as type family (Sen et al. 2014). The other hampered taxa were reclassified into three new orders *Geodermatophilales* (type family *Geodermatophilaceae*), *Acidotherrmales* (type family *Acidotherrmaceae*) and *Nakamurellales* (type family *Nakamurellaceae*).

At present, up to 10 species with effective and/or validly published names have been proposed within *Frankia* genus (Table 1) covering species with facultative symbiotic to asymbiotic lifestyle. Cultures of type strains have been made available in public culture collections primarily in DSMZ -

German collection of microorganisms and cell cultures and then in CECT-Spanish type culture collection for taxonomic and applied purposes.,

Type species *Frankia alni* (Woronin 1866) Von Tubeuf 1895 Emended Nouioui et al. 2016a

The species name refers to alder plant. The original description in roots of alder occurred well before the isolation of the availability of any cultivated strain of the genus. It is a diazotrophic actinobacterium with three types of cell structures: substrate hyphae, multilocular sporangia and vesicles. Grows optimally at 28 °C and from pH 6.3 to 6.8. It is able to metabolize d-fructose 6-phosphate, d-glucose 6-phosphate, fusidic acid, malic acid and methyl pyruvate and grows in the presence of guanidine hydrochloride, lithium chloride, potassium tellurite, propionate, sodium bromate, 1 % sodium lactate and Tween 40. Resists to lincomycin, minocycline, vancomycin and nalidixic acid and at up to 5 mM of biphenyl and polychlorinated biphenyl. The cell sugars are galactose, glucose, mannose, rhamnose, ribose and xylose. Predominant menaquinone is MK-9(H 8). The major fatty acids are iso-C_{16:0} and C_{17:1} ω8 c, whereas the polar lipids are PI, DPG, GPL1–3, PG and unknown lipid (L). The G+C content of type strain is 72.8 mol% and its total genome size is 7.50 Mb (Normand et al. 2007).

The type strain is ACN14a^T (=DSM 45986^T =CECT 9034^T) was isolate from *Alnus viridis* ssp. *crispa* (Tadoussac, Canada; Normand and Lalonde 1982) and is able to infect *Alnus* and *Myricaceae* species.

Based on dDDH (94.3 %) and ANI (99.49%) with type strain, strain AvcII, an *Alnus viridis* subsp. *crispa* isolate (Baker and Torrey 1980) is a member of *Frankia alni* species. The strain has the same host range being infective *Alnus*, *Comptonia* and *Myrica* (Baker and Torrey 1980). Its genome size is 7.7-Mbp with a G+C content of 72.4 mol% (Swanson et al. 2015b).

***Frankia torreyi* Nouioui et al. 2018c**

The species was named for John G. Torrey, who worked in Harvard University, Petersham, MA USA, in recognition of his great contributions to *Frankia* research including the isolation of the first *Frankia* strains.

Nitrogen-fixing, with three types of cell structures: substrate hyphae, multilocular sporangia and vesicles. Able to metabolise propionate, α-keto-butyric acid, D-gluconic acid, p-hydroxy-phenylacetic acid, citric acid, bromo-succinic acid, and α-keto-glutaric acid and to grow in presence of aztreonam, lincomycin, minocycline, nalidixic acid, troleandomycin and vancomycin. Cell walls contain *meso*-Diaminopimelic acid, galactose, glucose, mannose, rhamnose, ribose and xylose. Polar lipid profile consisted of PI, DPG, GPL1–3, PG, aminophospholipid (APL) and unknown phospholipids (PL1-2) and lipids (L). Predominant menaquinones is MK-9 (H8), as polar lipid, while major fatty acids are

composed of iso-C_{16:0} and C_{17:1} ω 8c. The genome size of the type strain is 7.6 Mb and digital DNA G+C content is 72.4%.

The type strain CpII^T (=DSM 44263^T = CECT 9035^T) was isolated from the root nodules of *Comptonia peregrina* (Callaham et al. 1978) and is able to nodulate *Alnus*, *Comptonia* and *Myricac* species.

Based on dDDH value (94.2%) and ANI score (99.5 %) with CpII^T, strain ACN1ag is a member of *Frankia torreyi*. The strain was reported as a re-isolate from root nodules of *Alnus glutinosa* (Lalonde et al. 1981) inoculated with a strain originally isolated from *Alnus viridis* spp. *crispa* collected from Atikokan, Ontario, Canada (Baker and Torrey 1979). Its genome is of 7.5-Mbp with a G+C content of 72.3 mol% (Swanson et al. 2015a).

Strain CpI-P was reported by Tisa et al. (1983) as derivative of the original strain CpII (Callaham et al. 1978). CpI-P will not uses succinate as the original isolate that used instead propionate. The genome of the original strain CpIIT and its derivative have 100% homology of 16S rRNA gene sequences, 99.6% dDDH value and 99.9% of ANI score indicating that are very closely related and members of the same species *Frankia torreyi*.

It is worth noting that by analyzing the two genomes (both of 7.6 Mb with G+C content of 72.4 mol%) reported by Oshone et al. (2016), revealed that they contain exactly the same genes (in term of number and sequence homology= 100%) that are related to propionate and succinate metabolisms (results not shown). Moreover 16S rRNA gene sequence, MLSA and pan-genome phylogenies (Fig. 3) together with other comparative genomic (data not shown) indicate that they undeniably represent the same strain.

***Frankia canadensis* Normand et al. 2018**

The species was named based on the country, Canada, from which the type strain was collected. Nitrogen-fixing actinobacterium growing as three cell structures: septate hyphae, vesicles and multilocular sporangia. Optimum growth of the strain was observed after 3-4 weeks of incubation at 28°C in BAP medium. Whole cell hydrolysates are rich in *meso*-diaminopimelic acid, glucose, mannose, rhamnose (trace), ribose and xylose; polar lipid profile contains PI, DPG, and GPL. The predominant menaquinone (>25%) was MK-9(H8) and major fatty acid (>30%) consists of *iso*-C_{16:0} and C_{17:1}ω8c. The DNA G+C content is 72.4 mol%.

The type strain ARgP5^T (=DSM 45898^T = CECT 9033^T) was isolated from *Alnus incana* ssp *rugosa* tree growing on a roadside ditch bank in Quebec city, Province of Quebec, Canada (Normand and Lalonde, 1982) and able to nodulate *Alnus* and *Myricaceae* species.

***Frankia casuarinae* Nouiou et al. 2016a**

224 The species was named for *Casuarina*, referring to the source of the isolate. Diazotrophic
 225 actinobacterium. Multilocular sporangia are formed either terminally or intercalary positions on
 226 substrate hyphae while vesicles are formed terminally. Metabolizes acetic acid, acetoacetic acid, butyric
 227 acid, fusidic acid, β -gentiobiose, methyl pyruvate, propionic acid, sodium lactate, α -hydroxybutyric
 228 acid, p-hydroxyphenylacetic acid, α -ketobutyric acid, potassium tellurite and Tween 40. A positive
 229 reaction is observed in the presence of tetrazolium violet and tetrazolium blue. It resists to nalidixic acid,
 230 minocycline, vancomycin and to polychlorinated biphenyl at concentrations up to 5 mM. The major
 231 fatty acids are iso-C 16 : 0 and C 17 : 1 ω 8 c. The predominant menaquinone is MK-9(H 6). The polar
 232 lipids are PI, DPG, PG, three GPL and an unknown lipid. The cell sugars are galactose, glucose,
 233 mannose, rhamnose, ribose and xylose. The G+C content of the strain is 70.1 mol% and its genome size
 234 is 5.3 Mb (Normand et al. 2007).

235 The type strain CcI3^T (=DSM 45818^T =CECT 9043^T) isolated from *Casuarina cunninghamiana*
 236 (Florida, USA; Zhang et al. 1984) and the host plants include *Casuarinaceae* (except *Gymnostoma*) and
 237 *Myricaceae*.

238 As previously expected (Dobritsa 1998), strains world widely isolated from *Allocasuarina* and
 239 *Casuarina* species (Australia; KB5 (Rosbrook et al. 1989), India; Cg70.9, Senegal; CeD (Diem et al.
 240 1982), Brazil; BR, Tunisia; BMG5.23 (Ghodhbane-Gtari et al. 2010), and several strains from Egypt;
 241 CgIM4, CgIS1, CcI156 and CcI6 (Mansour and Moussa 2005), Allo2 and Thr (Girgis and Schwencke
 242 1993) fall into *Frankia casuarinae* species based on dDDH and ANI values that range from 94.6 to
 243 95.7% and 99.3-99.9% respectively. The genome size of these strains ranges from 4.9 to 5.6 with G+C
 244 content of 69.3 to 70.1 mol% (Ghodhbane-Gtari et al. 2014 ; Hurst et al. 2014; Mansour et al. 2014;
 245 D'Angelo et al. 2016; Ngom et al. 2016; Pesce et al. 2017).

246 ***Frankia coriariae* Nouioui et al. 2017a**

247 The name of species is referring to the host origin of isolation, *Coriaria japonica*, of the type strain.
 248 Mildly alkaliphilic, and slow growing actinobacterium that form two unique cell structures: dense
 249 mycelium forming brownish colonies and spherical or oval vesicles produced under nitrogen-fixing
 250 conditions. No sporangia are detected. Grows well at 25–30 °C and pH 7–9.5. Optimal growth
 251 temperature and pH are 28 °C and 9.5, respectively. Metabolize acetic acid, acetoacetic acid, citric acid,
 252 cellobiose, D-fructose 6-phosphate, glucuronamide, L-lactic acid, methyl pyruvate, p-
 253 hydroxyphenylacetic acid, propionate, pyruvate, Tween 40, α -hydroxybutyric acid, α -ketoglutaric acid
 254 and 1 % sodium lactate. Tolerates up to 2 % (w/v) NaCl. The whole-cell hydrolysate contains *meso*-
 255 diaminopimelic acid, galactose, glucose, mannose and a trace of ribose. The polar lipid pattern is formed
 256 by PI, PG, DPG, GPL1–3 and unknown lipids (L1–3). The major fatty acid (>30 %) and the predominant

isoprenolog (>30 %) are C_{18:1} ω₉c and MK-9(H6) (44.7 %) respectively. The G+C content of the type strain is 70.1 mol % (Gtari et al. 2015).

The type strain BMG5.1^T (=CECT 9032^T =DSM 100624^T), was isolated from the root nodule of *Coriaria japonica* (Gtari et al. 2015) and was able to re-establish effective symbiosis with *Coriaria* spp. and *Datisca* spp.

Based on 16S rRNA gene homology (100%), dDDH value (95.6%) and ANI score (99.71%) with BMG5.1^T, strain BMG5.30 is a new member of *Frankia coriariae* species. This strain was isolated from soil collected in Tunisia using *Coriaria myrtifolia* as plant trapping bioassay (Gueddou et al. 2018). Its genome size is 5.8-Mbp with a G + C content of 70. mol%.

***Frankia elaeagni* Nouioui et al. 2016a**

The name of species is in reference to *Elaeagnus*, the host plant source of the isolation of the type strain. Forms vegetative hyphae with a septate mycelium differentiate into sporangia and vesicles. The strain fixes nitrogen and grows well on BAP medium producing a red pigment. Metabolize acetic acid, butyric acid, cellobiose, d-fructose 6-phosphate, d-glucose 6-phosphate, l-lactic acid, β-hydroxybutyric acid, Tween 40, α-ketobutyric acid, propionic acid, tetrazolium blue and tetrazolium violet. It resists to lincomycin. The cell sugars are galactose, glucose, mannose, rhamnose, ribose and xylose while the predominant menaquinone is MK-9(H4). The iso-C_{16:0}, C_{16:0} and C_{17:1} ω₈c as major fatty acid content and polar lipids are PI, DPG, GPL1–3, PG and unknown lipid (L). The G+C content is 71.7 mol% and the genome size is 7.6 Mb (Nouioui et al. 2013).

The type strain BMG5.12^T (=DSM 46783^T=CECT 9031^T) *Elaeagnus angustifolia* (Gafsa, Tunisia; Gtari et al. 2004) and the host plant range includes members of *Elaeagnaceae* and *Myricaceae*.

***Frankia discariae* Nouioui et al. 2017b**

The species was named for *Discaria*, the host plant origin of isolation of the type strain. It is diazotrophic actinobacterium that grow as substrate hyphae, vesicle and multilocular sporangia appeared in terminally or intercalary position on the hyphae. Uses D-cellobiose, d-glucose-6-phosphate, d-fructose-6-phosphate, α-hydroxy-butyric acid, l-lactic acid, citric acid, bromosuccinic acid, fusidic acid, acetic acid, lithium chloride, potassium tellurite and sodium lactate as carbon source and guanidine hydrochloride as nitrogen compound. It resists to phenotype to lincomycin, nalidixic acid, minocycline and vancomycin. Cell walls contain meso-diamino acid is the principal diamino acid of peptidoglycan, galactose, glucose, mannose, xylose and ribose. The fatty acid profile (>5%) is composed of iso-C_{16:0}, C_{17:1} ω₈c, C_{15:0}, C_{17:0}, C_{16:0}, C_{16:1} ω₇c C_{18:1} ω₉c and C_{18:0}. The predominant quinone (>50%) is MK-9(H4). Polar lipids

pattern consists of PI, DPG, PG, GPL1–3 and an unknown lipid (L). The genome size is 7.89 Mb and the GC content is 72.4% (Wall et al. 2013).

Type strain BCU110501^T (=DSM 46785^T=CECT 9042^T) was isolated from the nodule of *Discaria trinervis* (Chaia 1998) and its host plants include *Rhamnaceae* and *Elaeagnaceae* species.

***Frankia irregularis* Nouioui et al. 2018a**

The species name refers to irregular ie; inability of the species to infect its original host plant and to infect taxonomically disparate host plants. Diazotrophic actinobacterium that grow as red pigmented colonies forming the three cell structures: substrate hyphae, multilocular sporangia and vesicles. It is able to oxidise D-cellobiose, α -keto-butyric acid, methyl pyruvate, L-lactic acid, bromo-succinic acid, acetic acid, guanidine hydrochloride; growth in presence of sodium lactate, potassium tellurite, lincomycin, minocycline and vancomycin. Whole cell hydrolysates are formed by meso diaminopimelic acid, galactose, glucose, mannose, rhamnose, ribose and xylose. The predominant menaquinones (>20%) are MK-9 (H₄) and MK-9 (H₆) while the major fatty acids (>15%) are composed of *iso*-C_{16:0}, C_{17:1} ω 8c and C_{15:0}. Polar lipid pattern consisted of PI, DPG, PG, GPL1–3, APL and unknown lipids. (L). The size of the genome is 9.5 Mb and digital DNA G+C content is 70.9% (Nouioui et al. 2016b).

The type strain G2^T (=DSM 45899^T = CECT 9038^T) was isolated from *Casuarina equisetifolia* (Diem et al. 1982) while the host range include *Elaeagnaceae*, *Rhamnaceae* and *Myricaceae* species.

***Frankia inefficax* Nouioui et al. 2017c**

The species name refers to inefficient ie; the inability of the bacterium to form effective nitrogen-fixing symbiosis with its plant host. Grows as extremely branched substrate hyphae with multilocular sporangia while vesicles are not produced accordingly to its inability to fix nitrogen. Metabolizes D-dextrin, D-glucose-6-phosphate, D-fructose-6-phosphate, D-malic acid, citric acid, bromo-succinic acid, glucuronamide and aketo-glutaric acid. Tolerates fusidic acid, lithium chloride, potassium tellurite, sodium bromate, 1% sodium lactate and niaproof. Guanidine hydrochloride and D-serine are usable nitrogen sources. The cell sugars are galactose, glucose, mannose, ribose, rhamnose and also fucose being a diagnostic sugar of this species. The major fatty acids include *iso*-C_{16:0}, C_{17:1} ω 8c and C_{15:0} and C_{17:0}. The major menaquinones are MK-9(H₆) and MK-9(H₄) while polar lipids are composed of PI, DPG, two

320 GPL, PG and an unidentified lipid. The GC% content is 72.3% and the total size of its genome
321 is 8.815 Mb (Tisa et al. 2016).

322 The type strain EuI1c^T (= DSM 45817^T = CECT 9037^T) was isolated from *Elaeagnus*
323 *umbellata* (Baker et al. 1980) and forms ineffective nodules on *Elaeagnaceae* and *Morella*
324 species.

325 ***Frankia asymbiotica* Nouioui et al. 2017d**

326 The naming of the species is for asymbiotic lifestyle of the strain. It is a diazotrophic
327 actinobacterium which grows as branched substrate hyphae, vesicles and multilocular
328 sporangia. Metabolizes acetic acid, butyric acid, cellobiose, a- and b-hydroxybutyric acid, D-
329 glucose, a-ketobutyric acid, methyl pyruvate, maltose, trehalose, sucrose, turanose, D-mannose,
330 L-rhamnose, D-gluconic acid and propionic acid. Tolerates troleandomycin, rifamycin SV and
331 aztreonam. Whole-cell contain meso-diaminopimelic acid isomers; galactose, glucose,
332 mannose, rhamnose and ribose as sugars. Majors menaquinones are MK-9(H₄) and MK-9(H₆).
333 Polar lipids are composed of PI, PG, DPG, phosphoglycolipid and phospholipid. Fatty acids
334 composed of iso-C_{16:0} and C_{17:1ω8c}. The DNA G+C content of the M16386^T genome was
335 71.8% (Gueddou et al. 2017).

336 The type strain M16386^T (=DSM 100626^T =CECT 9040^T) was isolated from *Morella*
337 *californica* located in Westport-Legget, CA (Lechevalier 1986) is able to fix dinitrogen but
338 unable to nodulate its host plant of origin or any of the tested actinorhizal plant species.

339 ***Frankia saprophytica* Nouioui et al. 2018b**

340 The name saprophytica refers here to the asymbiotic lifestyle of the type strain). Grown as
341 branched substrate hyphae and multilocular sporangia. No vesicle is formed accordingly to its
342 inability to fix nitrogen. Metabolizes propionate, D-fucose, b-gentiobiose (disaccharide),
343 acetoacetic acid, butyric acid and L-malic acid (organic acids) and D-serine (nitrogen source)
344 and to grow in the presence of sodium bromate (salt), tetrazolium blue and tetrazolium violet
345 (redox indicators), rifamycin SV, minocycline, nalidixic acid and potassium tellurite. Whole-
346 cell hydrolysates contain meso-diaminopimelic acid, galactose, glucose, mannose, rhamnose
347 and ribose (in traces) as cell-wall sugars. The predominant menaquinone is MK-9(H₆). The
348 polar lipid profile contains PI, PG, DPG, two phosphoglycolipids (PGL1-2), phospholipid, six
349 uncharacterized glycolipids (GL1-6) and two uncharacterized lipids (L1-2). The major fatty

acids are iso-C_{16:0}, C_{17:1} ω8c and C_{15:0}. The NA G+C content of the type strain is 71.8 mol% (Ghodhbane-Gtari et al. 2013).

The type strain is CN3^T (=DSM 105290^T=CECT 9314^T), isolated from *Coriaria nepalensis* by Mirza et al. (1991) and is unable to nodulate any actinorhizal plant species.

Based on 16S rRNA gene homology (100%), dDDH value (99.7%) and ANI score (99.9%) with CN^T, strain EUN1h isolated from soil collected in Tunisia using *Elaeagnus umbellata* as plant trapping assay (Gueddou et al. 2017) is a member of *Frankia saprophytica* species.

Conclusion

Polyphasic taxonomy of actinobacteria from the genus *Frankia* has been long-standing hampered due to persistent problem of studying them with traditional bacteriological techniques (Beson et al. 2011). Recently the strategy had incorporated genome to genome data, omniLog® phenoarray together with morphology, chemotaxonomy and host plant range and had permitted to describe several species with effective and/or validly published names. The heyday of *Frankia* taxonomy will certainly continue with further incorporation of genome to phenome data for species delineation purpose. Prediction of chemotaxonomic and cultural features is at present being explored through annotated genomes (Riesco et al. 2018; Pujalte et al. 2018). This approach may be applied in the next future to yet uncultured microsymbionts (Torrey 1990) or *Candidatus*; *Candidatus* *Frankia discarea* (Persson et al. 2011) and *Candidatus* *Frankia californiensis* (Normand et al. 2017; Nguyen et al. 2016).

Competing Interests

The authors have declared that they have no competing interest exists.

Ethical approval:

This article does not contain any studies with human participants or animals performed by any of the authors.

Authors Contribution: MG conceived the study. IS and AS performed phylogenomic analysis. MG, IN, FGG, LST and HPK wrote the manuscript.

References

Akimov VN, Dobritsa SV (1992) Grouping of *Frankia* strains on the basis of DNA relatedness. Syst Appl Microbiol 15 372–379.

Akimov VN, Dobritsa SV, Stupar OS (1991) Grouping of *Frankia* strains by DNA:DNA homology: How many genospecies are in the genus *Frankia*? Nitrogen fixation. In Developments in Plant and Soil Sciences, pp. 635–636. Edited by Polsinelli M., Materassi R., Vincenzini M. Dordrecht, The Netherlands: Kluwer Academic Publisher.

- 384 Auch AF, Jan M, Klenk HP, Göker M (2010) Digital DNA-DNA hybridization for microbial species
385 delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2(1),117.
- 386 Baker D, Newcomb W, Torrey JG (1980) Characterization of an ineffective actinorhizal microsymbiont,
387 *Frankia* sp. EuI1 (*Actinomycetales*). *Can J Microbiol* 26: 1072-1089.
- 388 Baker D, O'Keefe D (1984) A modified sucrose fractionation procedure for the isolation of frankiae
389 from actinorhizal root nodules and soil samples. *Plant Soil* 78:23–8.
- 390 Baker D, Torrey JG (1979) The isolation and cultivation of actinomycetes root nodule
391 endophytes. *Symbiotic nitrogen fixation in the management of temperate forests*. Eds. Gordon
392 JC, Wheeler CT and Perry DA. Oregon State University Press, Corvallis, OR 38-56.
- 393 Baker D, Torrey JG (1980) Characterization of an effective actinorhizal microsymbiont, *Frankia* sp.
394 AvcI1 (*Actinomycetales*). *Can J Microbiol* 26:1066-71.
- 395 Baker D, Torrey JG, Kidd GH (1979) Isolation by sucrose-density fractionation and cultivation in vitro
396 of actinomycetes from nitrogen-fixing root nodules. *Nature* 281:76–78.
- 397 Baker DD (1987) Relationships among pure cultured strains of *Frankia* based on host specificity.
398 *Physiol Plant* 70 245–248.
- 399 Becking JH (1970) *Frankiaceae* fam. nov. (*Actinomycetales*) with one new combination and six new
400 species of the genus *Frankia* Brunchorst 1886, 174. *Int J Syst Evol Microbiol* 20: 201-220.
- 401 Benson DR, Brooks JM, Huang Y, Bickhart DM, Mastronunzio JE (2011) The biology of *Frankia* sp.
402 strains in the post-genome era. *Mol Plant Microbe Interact* 24:1310-1316.
- 403 Berry A, Sunell L (1990) The infection process and nodule development. In *The biology of Frankia and*
404 *actinorhizal plants*. Edited by Schwintzer C. R., Tjepkema J. D. San Diego: Academic Press.
- 405 Brunchorst J (1886) Über einige Wurzelanschwellungen, besonders diejenigen von *Alnus* und den
406 *Elaegnaceen*. *Botanische Institut Tübingen* 2:151-77.

407 Callaham D, Deltredici P, Torrey JG (1978) Isolation and cultivation in vitro of the actinomycete
408 causing root nodulation in *Comptonia*. Science 199: 899-902.

409 Chaia E (1998) Isolation of an effective strain of *Frankia* from nodules of *Discaria trinervis*
410 (*Rhamnaceae*). Plant and Soil, 205(2), 99-102.

411 D'Angelo T, Oshone R, Abebe-Akele F, Simpson S, Morris K, Thomas WK et al. (2016) Permanent
412 draft genome sequence of *Frankia* sp. strain BR, a nitrogen-fixing actinobacterium isolated from
413 the root nodules of *Casuarina equisetifolia*. Genome Announc, 4(5), e01000-16.

414 Diem HG, Gauthier D, Dommergues YR. (1982) Isolation of *Frankia* from nodules of *Casuarina*
415 *equisetifolia*. Can J Microbiol 28:526–530.

416 Dobritsa SV (1998) Grouping of *Frankia* strains on the basis of susceptibility to antibiotics, pigment
417 production and host specificity. Int J Syst Evol Microbiol 48:1265–1275

418 Fernandez MP, Meugnier H, Grimont PAD, Bardin R (1989). Deoxyribonucleic acid relatedness among
419 members of the genus *Frankia* . Int J Syst Bacteriol 39 424–429.

420 Ghodhbane-Gtari F, Beauchemin N, Bruce D, Chain P, Chen A, Davenport KW et al. (2013) Draft
421 genome sequence of *Frankia* sp. strain CN3, an atypical, noninfective (Nod–) ineffective (Fix–)
422 isolate from *Coriaria nepalensis*. Genome announc 1(2), e00085-13.

423 Ghodhbane-Gtari F, Hurst SG, Oshone R, Morris K, Abebe-Akele F, Thomas WK et al. (2014) Draft
424 genome sequence of *Frankia* sp. strain BMG5. 23, a salt-tolerant nitrogen-fixing actinobacterium
425 isolated from the root nodules of *Casuarina glauca* grown in Tunisia. Genome Announc, 2(3),
426 e00520-14.

427 Ghodhbane-Gtari F, Nouioui I, Chair M, Boudabous A, Gtari M (2010) 16S-23S rRNA intergenic spacer
428 region variability in the genus *Frankia*. Microb Ecol60:487-495.

429 Girgis MG, Schwencke J (1993) Differentiation of *Frankia* strains by their electrophoretic patterns of
430 intracellular esterases and aminopeptidases. J Gen Microbiol 139:2225–2232

431 Gtari M, Ghodhbane-Gtari F, Nouioui I, Ktari A, Hezbri K, Mimouni W, Sbissi I et al. (2015)
 432 Cultivating the uncultured: growing the recalcitrant cluster-2 *Frankia* strains. Scientific reports 5:
 433 13112.

434 Gtari M, Tisa LS, Normand P (2013) Diversity of *Frankia* strains, actinobacterial symbionts of
 435 actinorhizal plants. In: Aroca R (editor). Symbiotic Endophytes. Berlin, Heidelberg: Springer-
 436 Verlag; 2013.pp. 123–148.

437 Gtari M, Brusetti L, Hassen A, Mora D, Daffonchio D, Boudabous A (2007) Genetic diversity among
 438 *Elaeagnus* compatible *Frankia* strains and sympatric-related nitrogen-fixing actinobacteria
 439 revealed by *nifH* sequence analysis. Soil Biol Biochem 39:372–377

440 Gtari M, Brusetti L, Skander G, Mora D, Boudabous A, Daffonchio D (2004) Isolation of *Elaeagnus*-
 441 compatible *Frankia* from soils collected in Tunisia. FEMS Microbiol Lett 15;234:349-55.

442 Gueddou A, Swanson E, Hezbri K, Nouioui I, Ktari I, Simpson S, Morris K, Thomas WK, Ghodhbane-
 443 Gtari F, Gtari M, Tisa LS (2018) Draft genome sequence of the symbiotic *Frankia* sp. strain
 444 BMG5.30 isolated from root nodules of *Coriaria myrtifolia* in Tunisia. Antonie van Leeuwenhoek
 445 doi:10.1007/s10482-018-1138-1

446 Gueddou A, Swanson E, Ktari A, Nouioui I, Hezbri K, Ghodhbane-Gtari F et al. (2017) Permanent draft
 447 genome sequences of three *Frankia* sp. strains that are atypical, noninfective, ineffective isolates.
 448 Genome announc, 5(15), e00174-17.

449 Hahn D, Nickel A, Dawson J (1999) Assessing *Frankia* populations in plants and soil using molecular
 450 methods. FEMS Microbiol Ecol, 29: 215-227.

451 Hurst SG, Oshone R, Ghodhbane-Gtari F, Morris K, Abebe-Akele F, Thomas WK et al. (2014) Draft
 452 genome sequence of *Frankia* sp. strain Thr, a nitrogen-fixing actinobacterium isolated from the
 453 root nodules of *Casuarina cunninghamiana* grown in Egypt. Genome Announc, 2(3), e00493-14.

454 Konstantinidis KT, Tiedje JM (2005) Towards a genome-based taxonomy for prokaryotes. Journal of
 455 bacteriology 187: 6258-6264.

- 456 Lechevalier MP (1994) Taxonomy of the genus *Frankia* (*Actinomycetales*). Int J Syst Evolution
457 Microbiol 44: 1-8.
- 458 Lechevalier MP, Baker D, Horrière F (1983) Physiology, chemistry, serology, and infectivity of two
459 *Frankia* isolates from *Alnus incana* subsp. *rugosa* . Can J Bot 61 2826–2833.
- 460 Lechevalier MP, Horriere F, Lechevalier HA (1982) The biology of *Frankia* and related organisms. Dev
461 Ind Microbiol 23 51–60.
- 462 Lechevalier MP, Lechevalier HA (1979) The taxonomic position of the actinomycetic endophytes. In
463 Symbiotic Nitrogen Fixation in the Management of Temperate Forests, pp. 111–121. Edited by
464 Gordon J. C., Wheeler C. T., Perry D. A. Forest Research Laboratory, Corvallis, OR: Oregon
465 State University.
- 466 Lechevalier MP, Lechevalier HA (1990) Systematics, isolation and culture of *Frankia* . In The Biology
467 of *Frankia* and Actinorhizal Plants, pp. 35–60. Edited by Schwintzer C. R., Tjepkema J. D. San
468 Diego: Academic Press.
- 469 Lechevalier MP, Ruan JS (1984) Physiology and chemical diversity of *Frankia* spp. isolated from
470 nodules of *Comptonia peregrina* (L.) Coult and *Ceanothus americanus* L. Plant Soil 78 15–22.
- 471 Lechevalier MP (1986) Catalog of *Frankia* strains. Actinomycete 19:131–162
- 472 Lumini E, Bosco M, Fernandez MP (1996) PCR-RFLP and total DNA homology revealed three related
473 genomic species among broad-host-range *Frankia* strains. FEMS microbiology ecology, 21(4),
474 303-311.
- 475 Mansour SR, Moussa LAA (2005) Role of Gamma-radiation on spore germination and infectivity of
476 *Frankia* strains CeI523 and CcI6 isolated from Egyptian *Casuarina*. Isot Radiat Res 37:1023–38.
- 477 Mansour SR, Oshone R, Hurst SG, Morris K, Thomas WK, Tisa LS (2014) Draft genome sequence of
478 *Frankia* sp. strain CcI6, a salt-tolerant nitrogen-fixing actinobacterium isolated from the root
479 nodule of *Casuarina cunninghamiana*. Genome Announc, 2(1), e01205-13.

480 Meier-Kolthoff JP, Göker M, Spröer C, Klenk HP (2013) When should a DDH experiment be mandatory
481 in microbial taxonomy? Archives of Microbiology 195: 413-418.

482 Miller JM, Rhoden DL (1991) Preliminary evaluation of Biolog, a carbon source utilization method for
483 bacterial identification. J Clin Microbiol 29, 1143–1147.

484 Mirza MS, Hameed S, Akkermans AD (1994) Genetic diversity of *Datisca cannabina*-compatible
485 *Frankia* strains as determined by sequence analysis of the PCR-amplified 16S rRNA
486 gene. Applied and environmental microbiology 60: 2371-2376.

487 Ngom M, Oshone R, Hurst SG, Abebe-Akele F, Simpson S, Morris K et al. (2016) Permanent draft
488 genome sequence for *Frankia* sp. strain CeD, a nitrogen-fixing actinobacterium isolated from the
489 root nodules of *Casuarina equisetifolia* grown in Senegal. Genome Announc, 4(2), e00265-16.

490 Nguyen TV, Wibberg D, Battenberg K, Blom J, Heuvel BV, Berry AM, Kalinowski J et al. (2016) An
491 assemblage of *Frankia* Cluster II strains from California contains the canonical nod genes and
492 also the sulfotransferase gene *nodH*. BMC genomics 17: 796.

493 Normand N, Nouioui I, Pujic P, Fournier P, Dubost A, Klenk H-P, Nguyen A et al. (2018) *Frankia*
494 *canadensis* sp. nov., isolated from root nodules of *Alnus incana* ssp *rugosa* growing in Canada.
495 Int J Syst Evol Microbiol 10.1099/ijsem.0.002939

496 Normand P, Lalonde M. (1982) Evaluation of *Frankia* strains isolated from provenances of two *Alnus*
497 species. Can J Microbiol 28:1133-1142.

498 Normand P, Nguyen TV, Battenberg K, Berry AM, Heuvel BV, Fernandez MP, Pawlowski K (2017)
499 Proposal of '*Candidatus Frankia californiensis*', the uncultured symbiont in nitrogen-fixing root
500 nodules of a phylogenetically broad group of hosts endemic to western North America. Int J Syst
501 Evol Microbiol 67: 3706-3715.

502 Normand P, Orso S, Cournoyer B, Jeannin P, Chapelon C, Dawson J et al. (1996). Molecular phylogeny
503 of the genus *Frankia* and related genera and emendation of the family *Frankiaceae*. International
504 Journal of Systematic and Evolutionary Microbiology, 46(1), 1-9.

505 Normand P, Lapierre P, Tisa LS, Gogarten JP, Alloisio N, Bagnarol E et al. (2007) Genome
506 characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant
507 biogeography. *Genome Res* 17(1):7-15.

508 Nouioui I, Beauchemin N, Cantor MN, Chen A, Detter JC, Furnholm T et al. (2013). Draft genome
509 sequence of *Frankia* sp. strain BMG5. 12, a nitrogen-fixing actinobacterium isolated from
510 Tunisian soils. *Genome Announc* 1(4), e00468-13.

511 Nouioui I, Ghodhbane-Gtari F, Beauchemin NJ, Tisa LS, Gtari M (2011) Phylogeny of members of the
512 *Frankia* genus based on *gyrB*, *nifH* and *glnII* sequences. *Antonie Van Leeuwenhoek* 100:579-587

513 Nouioui I, Ghodhbane-Gtari F, Jando M, Rhode M, Klenk H-P, Gtari M (2018c) *Frankia torreyi* sp.
514 nov., *Frankia torreyi* sp. nov., the first actinobacterium of the genus *Frankia* Brunchorst 1886,
515 174^{AL} isolated in axenic culture. *Antonie Van Leeuwenhoek* doi: 10.1007/s10482-018-1131-8.
516

517 Nouioui I, Ghodhbane-Gtari F, Klenk HP, Gtari M (2018b) *Frankia saprophytica* sp. nov. an atypical
518 non-infective (Nod-) and non-nitrogen fixing (Fix-) actinobacterium isolated from *Coriaria*
519 *nepalensis* root nodules. *Int J Syst Evol Microbiol* 68: 1090-1095.

520 Nouioui I, Ghodhbane-Gtari F, Montero-Calasanz MD, Göker M, Meier-Kolthoff JP, Schumann P et al.
521 (2016a) Proposal of a type strain for *Frankia alni* (Woronin 1866) Von Tubeuf 1895, emended
522 description of *Frankia alni*, and recognition of *Frankia casuarinae* sp. nov. and *Frankia elaeagni*
523 sp. nov. *Int J Syst Evol Microbiol* 66: 5201-5210.

524 Nouioui I, Ghodhbane-Gtari F, Montero-Calasanz MDC, Rohde M, Tisa LS, Gtari M et al. (2017c)
525 *Frankia inefficax* sp. nov., an actinobacterial endophyte inducing ineffective, non nitrogen-fixing,
526 root nodules on its actinorhizal host plants. *Antonie Van Leeuwenhoek* 110: 313-320.

527 Nouioui I, Ghodhbane-Gtari F, Rhode M, Sangal V, Klenk HP, Gtari M (2018a) *Frankia irregularis* sp.
528 nov., an actinobacterium unable to nodulate its original host, *Casuarina equisetifolia*, but
529 effectively nodulate members of the actinorhizal *Rhamnales*. *Int J Syst Evol Microbiol* doi:
530 10.1099/ijsem.0.002914.

531 Nouioui I, Ghodhbane-Gtari F, Rohde M, Klenk HP, Gtari M (2017a) *Frankia coriariae* sp. nov., an
532 infective and effective microsymbiont isolated from *Coriaria japonica*. Int J Syst Evol Microbiol
533 67: 1266-1270.

534 Nouioui I, Gtari M, Göker M, Ghodhbane-Gtari F, Tisa LS, Fernandez MP et al. (2016b) Draft genome
535 sequence of Frankia strain G2, a nitrogen-fixing actinobacterium isolated from *Casuarina*
536 *equisetifolia* and able to nodulate actinorhizal plants of the order Rhamnales. Genome announc
537 4(3): e00437-16.

538 Nouioui I, Gueddou A, Ghodhbane-Gtari F, Rhode M, Gtari M, Klenk H-P (2017d) *Frankia asymbiotica*
539 sp. nov., a non infective actinobacterium isolated from *Morella californica* root nodule. Int J Syst
540 Evol Microbiol 67: 4897-4901.

541 Nouioui I, Montero-Calasanz MDC, Ghodhbane-Gtari F, Rohde M, Tisa LS, Klenk HP et al. (2017b)
542 *Frankia discariae* sp. nov.: an infective and effective microsymbiont isolated from the root nodule
543 of *Discaria trinervis*. Arch Microbiol 199: 641-647.

544 Oshone R, Hurst SG, Abebe-Akele F, Simpson S, Morris K, Thomas WK et al. (2016) Permanent draft
545 genome sequences for two variants of *Frankia* sp. strain CpII, the first *Frankia* strain isolated
546 from root nodules of *Comptonia peregrina*. Genome Announc, 4(1): e01588-15.

547 Persson T, Benson DR, Normand P, vanden Heuvel B, Pujic P, Chertkov O et al. (2011) Genome
548 sequence of "*Candidatus Frankia datiscae*" Dg1, the uncultured microsymbiont from nitrogen-
549 fixing root nodules of the dicot *Datisca glomerata*. J Bacteriol 193:7017–7018.

550 Pesce C, Swanson E, Simpson S, Morris K, Thomas WK, Tisa LS et al. (2017) Draft Genome Sequence
551 of the Symbiotic *Frankia* sp. Strain KB5 Isolated from Root Nodules of *Casuarina equisetifolia*.
552 J Genomics 5:64.

553 Pujalte MJ, Lucena T, Rodrigo-Torres L, Arahal DR (2018). Comparative genomics of *Thalassobius*
554 including the description of *Thalassobius activus* sp. nov., and *Thalassobius autumnalis* sp. nov.
555 Front Microbiol 8:2645.

556 Riesco R, Carro L, Román-Ponce B, Prieto C, Blom J, Klenk H-P et al. (2018). Defining the species
557 *Micromonospora saelicesensis* and *Micromonospora noduli* under the framework of genomics.
558 *Frontiers in Microbiology* 9:1360.

559 Rosbrook PA, Burggraaf AJP, Reddell P (1989) A Comparison of 2 Methods and Different Media for
560 Isolating *Frankia* from *Casuarina* Root-Nodules. *Plant Soil* 120:187–93.

561 Sen A, Daubin V, Abrouk D, Gifford I, Berry AM, Normand P (2014) The phylogeny of actinobacteria
562 revisited in the light of complete genomes, the orders *Frankiales* and *Micrococcales* should be
563 split into coherent entities. Proposal of *Frankiales* ord. nov., *Geodermatophilales* ord. nov.,
564 *Acidothermates* ord. nov. and *Nakamurellales* ord. nov. *Int J Syst Evol Microbiol* 64: 3821-3832.

565 Simon L, Jabaji-Hare S, Bousquet J, Lalonde M (1989) Confirmation of *Frankia* species using cellular
566 fatty acids analysis. *Syst Appl Microbiol* 11 229–235.

567 Swanson E, Oshone R, Simpson S, Morris K, Abebe-Akele F, Thomas WK et al. (2015a) Permanent
568 draft genome sequence of *Frankia* sp. strain ACN1ag, a nitrogen-fixing actinobacterium isolated
569 from the root nodules of *Alnus glutinosa*. *Genome Announc* 3(6):e01483-15.

570 Swanson E, Oshone R, Simpson S, Morris K, Abebe-Akele F, Thomas WK et al. (2015b) Permanent
571 draft genome sequence of *Frankia* sp. strain AvcI1, a nitrogen-fixing actinobacterium isolated
572 from the root nodules of *Alnus viridis* subsp. *crispa* grown in Canada. *Genome Announc* 3(6):
573 e01511-15.

574 Tisa LS, Oshone R, Sarkar I, Ktari A, Sen A, Gtari M (2016) Genomic approaches toward understanding
575 the actinorhizal symbiosis: an update on the status of the *Frankia* genomes. *Symbiosis* 70:5-16.

576 Von Tubeuf K (1895) In *Pflanzenkrankheiten durch Kryptogame Parasiten verursacht*. Springer, Berlin,
577 pp 1–599

578 Wall LG, Beauchemin N, Cantor MN, Chaia E, Chen A, Detter JC et al. (2013). Draft genome sequence
579 of *Frankia* sp. strain BCU110501, a nitrogen-fixing actinobacterium isolated from nodules of
580 *Discaria trinevis*. *Genome announc* 1(4): e00503-13.

581 Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI et al. (1987) Report of
582 the ad hoc committee on reconciliation of approaches to bacterial systematics. Int J Syst Bacteriol
583 37:463-464.

584 Woronin MS (1866) Über die bei der Schwarzerle (*Alnus glutinosa*) und bei der gewöhnlichen Garten-
585 Lupine (*Lupinus mutabilis*) auftretenden Wurzelschwellungen. Mem Acad Imp Sci St
586 Petersburg VII Ser 10:1–13.

587 Zhang Z, Lopez MF, Torrey JG (1984) A comparison of cultural characteristics and infectivity of
588 *Frankia* isolates from root nodules of *Casuarina* species. Plant Soil 78:79–90.

589

590

591

592

593

594

595

596

597

598

599

600

601

602 **Figure legends**

603 **Figure 1:** Root nodules of *Casuarina* (A). Cross section through *Coriaria myrtifolia* nodule
604 showing infected cells (IC) (B).

Figure 2: Colonies of *Frankia coriariae* BMG5.1^T grown in BAP medium for 3 weeks (A). Cell structures of *Frankia casuarinae* CcI3^T; hyphae (h), vesicle and sporangia under scanning electron microscopy (B) and under transmission electron microscopy (C).

Figure 3: (A); Maximum-likelihood phylogenetic tree based on 16S rRNA gene sequences. The numbers above the branches are bootstrap support values greater than 60% for ML (left) and MP (right). (B); Maximum-likelihood phylogeny based on concatenated amino acid sequences from *atpI*; *gyrA*; *ftsZ*; *secA* and *dnaK* proteins. The numbers above the branches are bootstrap support values greater than 50% for ML (left) and MP (right). (C) Maximum-likelihood phylogenomic tree based on core genome sequences.

626 Table 1. Distinguishable phenotypic and chemotaxonomic properties among *Frankia* species (Nouioui et al. 2016a-2018c, Normand et al. 2018).

	<i>Frankia alni</i> ACN14A ^T	<i>Frankia torreyi</i> CpII ^T	<i>Frankia canadensis</i> ARgP5 ^T	<i>Frankia casuarinae</i> CcI3 ^T	<i>Frankia coriariae</i> BMG5.1 ^T	<i>Frankia elaeagni</i> BMG5.12 ^T	<i>Frankia discariae</i> BCU110501 ^T	<i>Frankia irregularis</i> G2 ^T	<i>Frankia inefficax</i> EuI1c ^T	<i>Frankia asymbiotica</i> M16386 ^T	<i>Frankia saprophytica</i> CN3 ^T
Colony colour	white	white	white	white	brown	red	yellow	red	white	white	white greyish
Vesicles/N ₂ -fixation	+	+	+	+	+	+	+	+	-	+	-
Sporangia	+	+	+	+	-	+	+	+	+	+	+
Carbon source											
Dextrin	-	-	+	-	-	-	-	-	+	-	-
D-cellobiose	-	-	-	-	+	+	+	+	-	+	+
β-gentiobiose	-	-	-	+	-	-	-	-	-	-	+
D-glucose-6-phosphate	+	w	+	-	-	+	+	-	-	-	+
D-fructose-6-phosphate	+	w	+	-	+	+	+	-	+	-	+
α-hydroxy-butyricacid	-	-	-	+	+	-	+	w	-	+	-
β-hydroxy-butyricacid	-	-	-	-	-	+	-	-	-	+	-
α-keto-butyricacid	-	+	+	+	-	+	-	+	-	+	+
Aceto-aceticacid	-	-	+	+	+	-	-	w	-	-	+
Methyl pyruvate	+	-	-	+	+	-	-	+	-	+	+
D-gluconicacid	-	+	-	-	-	-	-	+	-	+	-
L-lacticacid	-	-	+	-	+	+	+	+	-	-	-
L-malicacid	+	+	-	-	-	-	-	w	-	-	+
D-malicacid	-	w	-	-	-	-	-	w	+	-	+
Citricacid	-	+	+	-	+	-	+	-	+	-	+
Bromo-succinicacid	-	+	-	-	-	-	+	+	+	-	+
p-hydroxy-phenyl acetic acid	-	+	-	+	+	-	-	-	-	-	-
L-pyroglutamicacid	-	+	-	-	-	-	-	-	-	-	-
α-keto-glutaricacid	-	+	-	-	+	-	-	w	+	-	+
Glucuronamide	-	-	-	-	+	-	-	w	+	-	+

Grow in presence of											
Acetic acid	-	+	+	+	+	+	+	+	-	+	+
1% sodium lactate	+	+	+	+	+	-	+	+	+	-	+
Fusidic acid	+	+	+	+	-	-	+	w	+	-	+
Lithium chloride	+	w	+	-	-	-	+	w	+	-	+
Potassium tellurite	+	+	+	+	-	-	+	+	+	+	+
Sodium bromate	+	+	+	-	-	-	-	w	+	+	+
Nitrogen sources											
Guanidine hydrochloride	+	w	+	-	-	-	+	+	+	-	-
D-serine	+	+	+	-	-	-	-	w	+	+	+
Antibiotic resistance to[#]											
Lincomycin	R	R	R	S	S	R	R	R	R	S	S
Nalidixic acid	R	R	R	R	S	S	R	R	R	R	R
Minocycline and vancomycin	R	R	R	R	S	S	R	R	R	S	R
Major fatty acids (>15%)	<i>iso</i> -C _{16:0} , C _{17:1} ω8c	<i>iso</i> -C _{16:0} , C _{17:1} ω8c	<i>iso</i> -C _{16:0} , C _{17:1} ω8c	<i>iso</i> -C _{16:0} , C _{17:1} ω8c	C _{18:1} ω9c, C _{16:0}	<i>iso</i> -C _{16:0} , C _{16:0} , C _{17:1} ω8c	<i>iso</i> -C _{16:0} , C _{17:1} ω8c, C _{16:0}	<i>iso</i> -C _{16:0} , C _{17:1} ω8c, C _{15:0}	<i>iso</i> -C _{16:0} , C _{17:1} ω8c, C _{17:0}	<i>iso</i> -C _{16:0} , C _{17:1} ω8c	<i>iso</i> -C _{16:0} , C _{17:1} ω8c, C _{15:0}
Predominant menaquinones (>20%)	MK-9(H ₈), MK-9(H ₄)	MK9(H ₈)	MK-9(H ₈)	MK-9(H ₆), MK-9(H ₈)	MK9(H ₆), MK9(H ₄)	MK-9(H ₄), MK-9(H ₆)	MK-9(H ₄)	MK-9(H ₄); MK-9(H ₆)	MK-9(H ₆), MK-9(H ₄)	MK-9(H ₄), MK-9(H ₆)	MK-9(H ₆)
Host plant origin	<i>Alnus viridis</i> ssp. <i>crispa</i>	<i>Comptonia</i> <i>peregrina</i>	<i>Alnus incana</i> ssp. <i>rugosa</i>	<i>Casuarina</i> <i>cunninghamiana</i>	<i>Coriaria japonica</i>	<i>Elaeagnus</i> <i>sangustifolia</i>	<i>Discaria</i> <i>nervosa</i>	<i>casuarina</i> <i>equisetifolia</i>	<i>Elaeagnus</i> <i>nusumbelata</i>	<i>Morella californica</i>	<i>Coriaria nepalensis</i>

Host plant range	<i>Alnus,</i> <i>Comptonia,</i> <i>Myrica</i>	<i>Alnus,</i> <i>Comptonia,</i> <i>Myrica</i>	<i>Alnus,</i> <i>Comptonia,</i> <i>Myrica</i>	<i>Alloea</i> <i>suarina,</i> <i>Casuarina,</i> <i>Myrica</i> <i>ceae</i>	<i>Coriaria</i> <i>ceae,</i> <i>Datisca</i> <i>ceae</i>	<i>Elaeagnaceae,</i> <i>Colletieae</i> <i>, Morella</i>	<i>Colletieae,</i> <i>Elaeagnaceae</i> <i>Morella</i>	<i>Elaeagnaceae,</i> <i>Colletieae,</i> <i>Morella</i>	<i>Elaeagnaceae,</i> <i>Colletieae,</i> <i>Morella</i>	-	-
Genomic G+C content (%)	72.8	72.4	72.4	70.1	70.2	71.7	72.3	70.9	72.3	72.0	71.8

627 +, positive reaction; -, negative reaction; w, weak reaction; R, resistance; S, sensitive

628