

# Frankia Diversity in Host Plant Root Nodules Is Independent of Abundance or Relative Diversity of Frankia Populations in Corresponding Rhizosphere Soils

Seifeddine Ben Tekaya, Trina Guerra, [...], and Dittmar Hahn

## ABSTRACT

Actinorhizal plants form nitrogen-fixing root nodules in symbiosis with soil-dwelling actinobacteria within the genus *Frankia*, and specific *Frankia* taxonomic clusters nodulate plants in corresponding host infection groups. In same-soil microcosms, we observed that some host species were nodulated (*Alnus glutinosa*, *Alnus cordata*, *Shepherdia argentea*, *Casuarina equisetifolia*) while others were not (*Alnus viridis*, *Hippophaë rhamnoides*). Nodule populations were represented by eight different sequences of *nifH* gene fragments. Two of these sequences characterized frankiae in *S. argentea* nodules, and three others characterized frankiae in *A. glutinosa* nodules. Frankiae in *A. cordata* nodules were represented by five sequences, one of which was also found in nodules from *A. glutinosa* and *C. equisetifolia*, while another was detected in nodules from *A. glutinosa*. Quantitative PCR assays showed that vegetation generally increased the abundance of frankiae in soil, independently of the target gene (i.e., *nifH* or the 23S rRNA gene). Targeted Illumina sequencing of *Frankia*-specific *nifH* gene fragments detected 24 unique sequences from rhizosphere soils, 4 of which were also found in nodules, while the remaining 4 sequences in nodules were not found in soils. Seven of the 24 sequences from soils represented >90% of the reads obtained in most samples; the 2 most abundant sequences from soils were not found in root nodules, and only 2 of the sequences from soils were detected in nodules. These results demonstrate large differences between detectable *Frankia* populations in soil and those in root nodules, suggesting that root nodule formation is not a function of the abundance or relative diversity of specific *Frankia* populations in soils.

**IMPORTANCE** The nitrogen-fixing actinobacterium *Frankia* forms root nodules on actinorhizal plants, with members of specific *Frankia* taxonomic clusters nodulating plants in corresponding host infection groups. We assessed *Frankia* diversity in root nodules of different host plant species, and we related specific populations to the abundance and relative distribution of indigenous frankiae in rhizosphere soils. Large differences were observed between detectable *Frankia* populations in soil and those in root nodules, suggesting that root nodule formation is not a function of the abundance or relative diversity of specific *Frankia* populations in soils but rather results from plants potentially selecting frankiae from the soil for root nodule formation. These data also highlight the necessity of using a combination of different assessment tools so as to adequately address methodological constraints that could produce contradictory data sets.

**KEYWORDS:** abundance, actinorhiza, *Frankia*, Illumina, *nifH*, qPCR, quantification, soil

## INTRODUCTION

The genus *Frankia* resembles nitrogen-fixing and non-nitrogen-fixing actinobacteria that live in soils (1) and can form root nodules in symbiosis with a variety of woody plants (2, 3). Root nodule formation requires interactions between members of specific *Frankia* taxonomic clusters and plants of specific host infection groups. Frankiae of cluster 1 form nodules on plants of the genera *Alnus*, *Morella*, and *Comptonia*, with a subgroup specifically nodulating the genera *Casuarina* and *Allocasuarina*. Frankiae of cluster 2 form nodules with *Ceanothus*, *Cercocarpus*, *Chamaebatia*, *Coriaria*, *Datisca*, *Dryas*, and *Purshia* species, while those of cluster 3 form nodules on members of the genera *Elaeagnus*, *Hippophaë*, *Shepherdia*, *Myrica*, *Morella*, and *Colletia* (2, 4). The last cluster, cluster 4, represents atypical, generally non-nitrogen-fixing frankiae from a variety of different host plants, including *Alnus*, *Ceanothus*, *Coriaria*, *Datisca*, *Purshia*, and *Elaeagnus* species (5–8).

The compatibility of frankiae with plants of specific host infection groups is very well documented (9–11), as is intragenetic variation in host plant compatibility with specific *Frankia* populations (12–15). Host plant species have been shown to determine the selection of *Frankia* strains from soil for potential nodule formation (14, 15), with nodules of different plant species from the same genus harboring unique *Frankia* populations, and none of them capturing the entire diversity of nodule-forming frankiae (15). While large differences in nodule-forming frankiae were observed for different soils on the same host plant (14), the level of diversity in nodule-forming frankiae from individual soils was found to be low, with generally one or two populations dominating in the nodules formed (14, 16). Comparative analyses of *Frankia*-specific *nifH* gene clone libraries from soil with root nodule populations often displayed large differences between sequences retrieved from clone libraries and those obtained from nodules; matching sequences were encountered only rarely (17) or not at all (14).

Many of these data have been retrieved from plant bioassays in which nodule-forming frankiae were identified on roots of a specific host plant after inoculation with a soil slurry (13, 16, 18–20). Since root nodules represent a natural locale of enrichment of usually one *Frankia* population, molecular tools such as PCR-assisted sequence analyses can be used to retrieve information on *Frankia* populations in root nodules. The case is different for highly complex and diverse environments, such as soil with *Frankia* populations present in small numbers within a large and complex microbial community (21), for which even molecular tools such as gene clone library analyses provide limited information (14, 17). The development of *Frankia*-specific quantitative PCR (qPCR) (22–24) or targeted Illumina sequencing (25) analyses now provides increasingly sophisticated molecular tools for the analyses of *Frankia* populations in soil.

The aim of this study was to assess *Frankia* diversity in root nodules of different host plant species growing in the same soil and to relate this diversity to the abundance and relative distribution of indigenous frankiae in rhizosphere soils. Different actinorhizal plant species, including *Alnus glutinosa*, *Alnus cordata*, *Alnus viridis*, *Casuarina equisetifolia*, *Shepherdia argentea*, and *Hippophaë rhamnoides*, as well as the nonactinorhizal plant species *Betula pendula*, were grown in soil microcosms that, together with a nonvegetated control, were analyzed by *Frankia*-specific qPCR and Illumina sequencing in order to assess the abundance and relative distribution of indigenous frankiae in rhizosphere soils. The results were related to *Frankia* diversity in root nodules analyzed by PCR-assisted sequence analyses.

## RESULTS

**Qualitative analyses of *Frankia* populations in root nodules by sequencing.** Healthy, surviving plants were growing with heights between 15 and 35 cm, depending on the species, and roots generally filling the entire microcosm. Root nodules were obtained for four of the seven plant species used for analyses: *Alnus glutinosa* ( $30 \pm 3$  nodules per microcosm; 3 microcosms), *Alnus cordata* (57 nodules; 1 microcosm), *Casuarina equisetifolia* (1 nodule; 3 microcosms), and *Shepherdia argentea* ( $23 \pm 3$  nodules; 2 microcosms). *Alnus viridis*, *Hippophaë rhamnoides*, and the nonhost plant *Betula pendula* (with 3 microcosms each) did not produce nodules (Fig. 1). Sequences were obtained from 141 lobes (51 *A. glutinosa*, 46 *A. cordata*, 1 *C. equisetifolia*, and 43 *S. argentea* lobes), with a total of eight different sequences. Two distinct sequences characterized frankiae in nodules from *S. argentea*; three sequences characterized frankiae in nodules from *A. glutinosa*; and five sequences characterized frankiae in nodules from *A. cordata*, of which two were shared with *A. glutinosa*, and one sequence was identical to that obtained from *C. equisetifolia* (Fig. 1).



FIG 1

Neighbor-joining topology showing sequence relationships for selected *Frankia* strains and uncultured endophytes in root nodules to demonstrate *Frankia* cluster assignments. Numbers above the branches represent the bootstrap values from maximum likelihood ...

Maximum likelihood (ML) analyses revealed that all sequences obtained from *S. argentea* represented frankiae of cluster 3; a minor percentage (7%) of sequences showed high similarity to *Frankia elaeagni* BMG5.12 (99.6% similarity), and the majority of sequences (93%) showed high similarity to strain EUN1f (98.3% similarity) (Fig. 1). Frankiae of cluster 3, related to strain EAN1pec (99.4% similarity), were also found in one nodule from *A. glutinosa*, corresponding to 2% of the sequences in nodules of *A. glutinosa*. Sequences present in nodules from both *A. glutinosa* and *A. cordata* were either closely related to *Frankia* strain ARgP5 (99.0% similarity), representing cluster 1d (62% of the sequences from *A. glutinosa* and 22% of the sequences from *A. cordata*), or formed a separate branch within cluster 1 characterized by sequences of uncultured frankiae without cultured relatives in the database, here tentatively named cluster 1e (36% of the sequences from *A. glutinosa* and 62% of the sequences from *A. cordata*) (Fig. 1). Two other distinct sequences from endophytes in nodules from *A. cordata* (each comprising 7% of the sequences from *A. cordata*) as well as the one from *C. equisetifolia* were assigned to this cluster as well.

**Quantitative analyses of *Frankia* sequences in soil by qPCR.** Microcosms with different vegetation harbored different numbers of *Frankia* cells; with any vegetation, numbers generally increased over those for a nonvegetated control ( $P < 0.001$ , except for *A. viridis* [ $P = 0.003$ ] and a nonsignificant value for *H. rhamnoides* [ $P = 0.1$ ]). Numbers about 4 to 5 times higher than the  $10^6$  cells (g of soil) $^{-1}$  in the nonvegetated control soil were obtained for soils with *A. glutinosa*, *A. cordata*, or *C. equisetifolia*, while those for the remaining plant species ranged between  $2 \times 10^6$  and  $3 \times 10^6$  cells (g of soil) $^{-1}$  (Fig. 2). Within treatments, results were identical independently of the target gene, i.e., the *nifH* gene allows the retrieval of information on clusters 1 and 3 together only, and the 23S rRNA gene provides information on all four clusters representing the genus *Frankia* (Fig. 2). These results suggest the absence of clusters 2 and 4, which was confirmed by specific analyses that detected frankiae only of cluster 1b, the newly assigned cluster 1e, and cluster 3; clusters 1a/d, 1c, 2, and 4 were not detected (Fig. 2). All three clusters detected were present only in nonvegetated control soil and in soil vegetated with *A. glutinosa*, while soils with the remaining plant species generally harbored cluster 3 and a smaller population of cluster 1e frankiae. Posttreatment microcosm soils with *C. equisetifolia* were an exception because they harbored cluster 1b instead of cluster 3 frankiae, and, again, a smaller population of cluster 1e frankiae (Fig. 2).

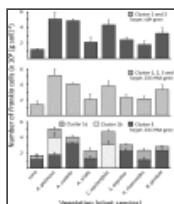


FIG 2

Sybr green-based qPCR quantification of specific *Frankia* clusters in soils vegetated with host trees (*Alnus glutinosa*, *Alnus cordata*, *Alnus viridis*, *Casuarina equisetifolia*, *Shepherdia argentea*, and *Hippophaë rhamnoides*) or nonhost trees (*Betula* ...

**Qualitative analyses of *Frankia* sequences in soil by Illumina sequencing.** Numbers of reads between 273,068 ( $\pm 14,956$ ) (*A. glutinosa*) and 425,752 (*A. cordata*) were identified as representing the genus *Frankia* (Fig. 3). A total of 24 unique sequences (LT934508 to LT934531) were retrieved from all treatments together (Table 1). Seven of these sequences were present in the majority of samples, and these reads combined represented generally >90% of all reads obtained (except for *A. cordata* [78.1%] and *S. argentea* [85.4%]). The majority of reads were represented by read 147378 (generally between 15 and 30%) and read 170193 (generally 45 to 60%), both of which identified cluster 3 frankiae as closest relatives but were not found in root nodules of our test plant species (Table 1). Only two of the seven sequences were identical to those found in nodules: read 101236 represented cluster 3 frankiae (generally 3 to 5% of the reads, except for *S. argentea*), and read 148038 represented cluster 1e frankiae (generally 4 to 9% of the reads) (Table 1). Read 101236 was present in the soil and nodules of *S. argentea*, at high percentages for both (20% of reads from soil, 93% of reads from nodules) (Fig. 1 and 3), while read 148038 was detected in the soil and nodules of *A. glutinosa* and *C. equisetifolia*, but not in soils with *A. cordata*, even though it was found in 62% of the nodules on *A. cordata* (Fig. 1 and 3). Of the reads present at minor abundances (i.e., just above the threshold of 1%), two were found to be identical to those in root nodules: one represented a cluster 3 *Frankia* population in a nodule on *A. glutinosa*, and another represented cluster 1e in nodules on *A. cordata* but, again, was not present in soil vegetated with *A. cordata* (Table 1). The remaining sequences were generally most closely related to sequences representing cluster 3 frankiae, with two exceptions that represented cluster 1a and 1b frankiae.

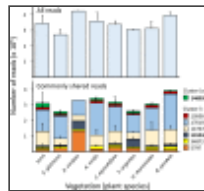


FIG 3

(Top) Numbers of reads representing *Frankia*-specific *nifH* gene fragments obtained by Illumina sequencing of DNA extracts from soils vegetated with host trees (*Alnus glutinosa*, *Alnus cordata*, *Alnus viridis*, *Casuarina equisetifolia*, *Shepherdia argentea* ...

Frankia-specific <i>nifH</i> reads from soil microcosms	Frankia-specific <i>nifH</i> reads from root nodules
Alnus glutinosa	Alnus glutinosa
Alnus cordata	Alnus cordata
Alnus viridis	Alnus viridis
Casuarina equisetifolia	Casuarina equisetifolia
Shepherdia argentea	Shepherdia argentea

TABLE 1

Similarity of *Frankia*-specific *nifH* reads from soil microcosms to sequences of uncultured frankiae in root nodules of different plant species grown in these soils and to sequences of closest *Frankia* isolates or clones

## DISCUSSION

Host plant species as well as *Frankia* populations seem to affect nodulation capacity and subsequent plant growth performance, as demonstrated in inoculation studies where spore-producing nodule homogenates produced significantly more nodules than non-spore-producing nodule homogenates on different alder species (26, 27). Spore-producing nodule homogenates produced twice as many nodules with *A. glutinosa* as with *Alnus incana* (26), while *Alnus crispa* produced more than twice as many nodules as *Alnus rubra* and six times as many nodules as *Alnus rugosa* (27). Differential root nodule formation of different alder species has been reported before, with *A. rubra* producing more nodules than *A. incana*, which ultimately produced more nodules than *A. glutinosa* (12). While the opposite sequence was found for different soils, with more nodules formed on *A. glutinosa* than on *A. incana* (13), none of these studies reported the lack of nodulation of some species (*A. viridis* and *H. rhamnoides*) while others of the same host infection group produced abundant nodules (*A. glutinosa*, *A. cordata*, *S. argentea*), as in our study. The above- and below-ground biomass of both *A. viridis* and *H. rhamnoides* was much smaller than that of the other plant species (data not shown), and *Frankia* populations were only marginally or not significantly different from those in soils that were kept nonvegetated (*A. viridis* [ $P = 0.03$ ] and *H. rhamnoides* [ $P = 0.102$ ]), while *Frankia* populations in soils vegetated by all other plant species significantly increased in abundance ( $P, < 0.001$  for all). These data indicate potential experimental effects that reduced the plant growth performance of both *A. viridis* and *H. rhamnoides*, which subsequently might have affected *Frankia* growth in the soil and rhizosphere and, finally, root nodule formation.

Although our study was not designed to assess quantitative differences in nodule formation on different host plant species, it clearly established differences in the diversity and abundance of specific *Frankia* populations in nodules and in soil. Frankiae in root nodules on each plant species were generally dominated by one or two populations, with additional populations present in a few nodules only. Two nodules were found with *Frankia* populations not typically detected on the respective plant host species. One nodule of *A. glutinosa* harbored a cluster 3 frankia instead of the usual cluster 1 populations, confirming previous reports of occasional detections of cluster 3 frankiae in nodules of the *Betulaceae* (2). A single nodule was found on *C. equisetifolia*, with a *Frankia* population representing cluster 1e. *Casuarina* species usually produce nodules with *Frankia* populations of cluster 1c (24), meant to represent a group of highly specialized frankiae that depend on cointroduction with their exotic host plant species outside their native range (28, 29). Consequently, *Casuarina*-infective *Frankia* populations were not expected to form nodules in soils that have never supported *Casuarina* spp. (21, 23, 30). Cluster 1e frankiae, which can be distinguished from other cluster 1 frankiae by highly distinctive signature sequences on a 23S rRNA gene insertion, as demonstrated in this study, have been observed in previous studies in nodules of different *Morella* species

growing on acidic soils in South Africa (31), as well as in nodules from *Morella pensylvanica*, used as capture plants in bioassays to assess *Frankia* diversity in soils from different countries throughout the world (e.g., from Japan and Rwanda) (17).

Soil *Frankia* populations showed higher abundance values in vegetated microcosms than in nonvegetated soil, indicating growth-enhancing effects of vegetation on frankiae, as demonstrated previously (21, 30, 32). This result confirms previous results documenting that all *Frankia* strains tested so far grow in the rhizosphere of their host plants (33) but also potentially in that of other plants (33–35). The diversity of *Frankia* populations, however, was relatively low, with frankiae of clusters 1b, 1e, and 3 present in nonvegetated controls, and cluster 1e and 3 frankiae generally were present in vegetated microcosms only. Previous studies reported on the low diversity of *Frankia* populations in different soil environments, e.g., the presence of cluster 1b only in wet or even water-logged soils in natural stands of *A. glutinosa* (23), clusters 1b and 3 in forest soils with *A. glutinosa* or nonhost plants, including *Betula nigra* (21, 24), clusters 1a, 1b, and 3 in forest soils with *A. glutinosa* (21), clusters 1a and 3 in microcosms with *A. glutinosa* or *C. equisetifolia* (30), or clusters 1b, 2, 3, and 4 in prairie soils with *Ceanothus* species as potential host plants (24). These studies suggest soil-specific environmental effects on *Frankia* populations in soils, with frankiae of clusters 1b and 3 being quite ubiquitous and independent of the presence of host plant species. Cluster 1b frankiae had been shown to grow with leaf litter of host and/or nonhost plants (33, 36) and thus could be adapted to carbon resources provided by the decomposition of plant material. Similarly, cluster 3 frankiae grew in bulk soil, in the rhizosphere, and with leaf litter, independently of matric potential and plant species, and thus represent a group with broad physiological adaptations (30). In contrast, cluster 2 frankiae were not detected. Prior to European settlement of the study region in the early nineteenth century, the landscape encompassing the site was tallgrass prairie and oak savanna, having the native actinorhizal shrub *Ceanothus americanus* L. as a common component. Existing native prairie remnants contain *C. americanus* and the corresponding, detectable soil frankiae specific to cluster 2 (24). It has been speculated that cluster 2 frankiae depend on their host plants for growth and establishment (24, 37), and thus, the conversion of native vegetation almost entirely to agricultural and urban uses, primarily the cultivation of maize, legume hay crops, and, in recent decades, soybeans may have resulted in the loss of cluster 2 frankiae at this site and other cultivated sites in this region.

qPCR data for soil *Frankia* populations were not exactly matched by targeted Illumina sequencing, which retrieved sequences representing clusters 1a, 1b, 1e, and 3, with all samples dominated by cluster 3 sequences, a small number of cluster 1e reads, and very similar composition of specific reads (Fig. 3). Root nodule formation does not seem to be a function of the abundance of specific *Frankia* populations in soils. More than 70% of *A. cordata* nodules harbored frankiae of cluster 1e detected by qPCR, but not by Illumina sequencing. The remaining 30% of nodules harbored cluster 1a frankiae that were not detected by either method. Similar results were obtained for nodules of *A. glutinosa* that harbored cluster 1e frankiae, as expected from detection in soils by both methods. However, most nodules contained cluster 1a frankiae that were not detected in soils by either of the two methods, while cluster 1b frankiae, detected as a major population by qPCR, did not occur in nodules at all. On a more specific scale than cluster assignments, eight distinct sequences were found in root nodules, but only four were found back in soil by Illumina sequencing. Although specific reads from soil samples were detected in nodules, the abundance of reads in soil samples did not reflect their presence in root nodules. These results corroborate previous studies where clone libraries of *Frankia*-specific *nifH* gene fragments from soils revealed large differences in cluster assignments to sequences obtained from nodules, with assignments to the same cluster only rarely encountered for individual soils (17).

These results demonstrate large differences between detectable *Frankia* populations in soil and those in root nodules, indicating the inadequacy of bioassays for the detection of specific nodulating frankiae in soil and the role of plants in the selection of frankiae from soil for root nodule formation. The data also highlight the necessity of using a combination of different assessment tools to adequately address methodological constraints that could provide contradictory data sets, or to quantify apparently small, threshold levels of infectious soil frankiae for specific hosts. Future studies could therefore use the availability of pure-culture representatives of different *Frankia* clusters to assess the reliability of individual assessment tools, including qPCR analyses and Illumina sequencing approaches, for both qualitative and quantitative analyses of *Frankia* populations in soils and nodules.

## MATERIALS AND METHODS

**Soil microcosms.** Soil microcosms were each established in 50-ml Falcon tubes using 70 g (dry weight) of soil collected from a dense stand of tall (4- to 5-m) *Elaeagnus angustifolia* plants established on a Drummer silty clay loam (taxonomic classification, fine-silty, mixed, superactive, mesic Typic Endoaquoll) in Illinois, USA. The climate is continental, with mean annual precipitation of 940 mm and a mean annual air temperature of 11°C. The fertile soil, formed under tallgrass prairie vegetation at an elevation of 214 m, is slightly acidic (pH 6.9) and rich in organic matter (ca. 3%). The site is adjacent to a temperate deciduous forest and has a history of row cropping, but not in the past 25 years (40.207797, -88.366524). Seeds of all plant species were surface sterilized in 3% hydrogen peroxide for 10 min, washed twice in sterile water for 10 min each time, and then germinated on a water agar. Three microcosms were set up for each plant species, with three seedlings per microcosm, and three microcosms without plants were used for comparison. All microcosms were kept at 23°C with a photoperiod of 16/8 h (day/night, respectively), and a matric potential of -0.005 MPa using round-bottom ceramic suction tubes (8 cm long, with an outer diameter of 6 mm; product no. 0652X02-B01M1; Soilmoisture Equipment Corp., Santa Barbara, CA) and 50-cm water columns (18).

Microcosms were harvested after 7 months, and root nodules as well as rhizosphere soil (defined as the soil adhering to plant roots when the plants were removed from the remaining bulk soil) were collected. Root nodules were kept in 100% ethanol at  $-20^{\circ}\text{C}$ , while soil was stored at  $-20^{\circ}\text{C}$  until further analysis.

**Qualitative analyses of *Frankia* populations in root nodules by sequencing.** Root nodules generally consisted of single lobes, which were homogenized in 1 ml of sterile water with a mortar and pestle. The homogenate was centrifuged at  $14,000 \times g$  for 3 min, and the pellet was washed once in 1 ml of 1% sodium pyrophosphate. The pellet was then dissolved in 95  $\mu\text{l}$  of sterile water, to which 5  $\mu\text{l}$  of proteinase K solution ( $10 \text{ mg ml}^{-1}$ ) was added (15, 24). This solution was transferred to a 0.5-ml PCR tube and was incubated in a thermocycler at  $37^{\circ}\text{C}$  for 20 min, after which 0.5  $\mu\text{l}$  of 10% SDS solution was added. After incubation at  $37^{\circ}\text{C}$  for 30 min, the proteinase K was inactivated by incubation at  $80^{\circ}\text{C}$  for 20 min.

Primers nifHf1 (5' GGC AAG TCC ACC ACC CAG C) and nifHr (5' CTC GAT GAC CGT CAT CCG GC) (38) were used to amplify 606-bp *nifH* gene fragments in a reaction volume of 50  $\mu\text{l}$ , containing 1  $\mu\text{l}$  of a 10 mM deoxynucleoside triphosphate (dNTP) mixture, 1  $\mu\text{l}$  each primer ( $0.2 \mu\text{M}$ ), 8.2  $\mu\text{l}$  bovine serum albumin (BSA) ( $30 \mu\text{g ml}^{-1}$ ), 5  $\mu\text{l}$  of  $10\times$  PCR buffer with 15 mM  $\text{MgCl}_2$ , 2  $\mu\text{l}$  root nodule lysate, and 0.2  $\mu\text{l}$  *Taq* DNA polymerase ( $5 \text{ U } \mu\text{l}^{-1}$ ; GenScript, Piscataway, NJ). Amplification conditions included an initial incubation at  $96^{\circ}\text{C}$  for 10 min, followed by 35 rounds of temperature cycling ( $96^{\circ}\text{C}$  for 30 s,  $60^{\circ}\text{C}$  for 30 s, and  $72^{\circ}\text{C}$  for 45 s) and a final 7-min incubation at  $72^{\circ}\text{C}$ . Amplified *nifH* gene fragments were cleaned using shrimp alkaline phosphatase and exonuclease I (Affymetrix, Santa Clara, CA) according to the manufacturer's protocols and were then sequenced bidirectionally using a BigDye Terminator cycle sequencing kit, v3.1 (Applied Biosystems, Foster City, CA), with primers nifH1 and nifHr. Sequences were analyzed on a 3500 genetic analyzer (Life Technologies, Carlsbad, CA).

Sequences were assembled in Geneious, v9.1.4 (Biomatters Ltd., Auckland, New Zealand), and were checked in the GenBank/EMBL databases using the BLAST algorithm (39). Representative sequences from confirmed *Frankia* strains of all clusters representing nitrogen-fixing frankiae were added from the GenBank/EMBL databases, aligned by the Geneious alignment tool, and trimmed to 516 bp to match those in the database (15). Geneious was also used to build a neighbor-joining topology based on a distance matrix generated using HKY85 correction from the resulting alignment. Clade confidence values were estimated using 10,000 bootstrap iterations. Maximum likelihood (ML) analyses (40) to evaluate the relationships among the sequences were carried out through the Cyber Infrastructure for Phylogenetic Research project portal ([www.phylo.org](http://www.phylo.org)) (41). Model parameters were estimated via GTRCAT with 25 rate categories (42) and were used as input in an ML heuristic search using RAXML, v8.2.10 (43). Rapid bootstrapping was conducted (44) for 1,000 ML iterations, searching for the best-scoring ML tree.

**Quantitative analyses of *Frankia* populations in soil by qPCR.** DNA was extracted from rhizosphere samples in triplicate 250- $\mu\text{g}$  soil samples using the SurePrep Soil DNA isolation kit (Fisher Scientific, Houston, TX) (22). Soils had been amended with *Salmonella enterica* serovar Typhimurium (ATCC 14028) cells to account for extraction efficiencies as described previously (22). DNA was used in Sybr green-based qPCRs to quantify *Frankia* clusters or subclusters with either *nifH* gene fragments (clusters 1 and 3 together) or 23S rRNA gene fragments (clusters 1, 2, 3 and 4, representing the entire genus, and clusters 1a/d, 1b, 1c, 2, 3, and 4 separately) as targets (24). Amplicons of 23S rRNA gene fragments (about 240 bp generated using primers 23Fra1533f [5' GTT GAT ATT CCC GTA CCG] and 23Fra1769r [5' GGC TCG GCA TCA GGT CTC AG]) (24) were also used to reanalyze endophytes in nodules that had been assigned to a phylogenetically distinct group of yet uncultured frankiae without cultured relatives (referred to as cluster 1e) in our nodule analyses (Fig. 1). PCR and sequence analyses were carried out as described above, and selected sequences were deposited in GenBank. Sequences were aligned in Geneious, v9.1.4, and the alignment amended with sequences of target and nontarget organisms from GenBank databases. This alignment was used to design a forward primer specific for 23S rRNA gene sequences of cluster 1e, i.e., primer 113f (5' GGA TGT GTG TGT GAG GTC GGG A), which could be used with reverse primer 23Fra1769r in Sybr green-based qPCR (23, 24). Primer 113f had 3 mismatches to *Frankia* strain ARgP5 (previously assigned to cluster 1d) and 5 or more mismatches to sequences of all other previously tested strains and uncultured endophytes in root nodules (23, 24). The potential of self-dimer and heterodimer formation, target specificity, and qPCR conditions were checked as described before (24). Briefly, low potential of self- and heterodimer formation was checked in OligoAnalyzer, v3.1 (Integrated DNA Technologies), while target specificity was checked using TestPrime 1.0 and TestProbe 3.0 (45) from the SILVA rRNA database project ([www.arb-silva.de](http://www.arb-silva.de); accessed 29 July 2017) (46). TestProbe revealed a high specificity of primer 113f, with at least 4 or 5 mismatches to nontarget organisms (a *Sulfolobus* sp. and *Haemophilus influenzae*, respectively), while TestPrime on the combination of primers 113f and 23Fra1769r generally retrieved organisms showing at least 9 mismatches (5 and 4) to primers 113f and 23Fra1769r, respectively. Sequences of target organisms for cluster 1e were not available in the database, and neither TestPrimer nor TestProbe retrieved *Frankia* sequences based on the search limit of 5 mismatches.

Sybr green-based analyses were carried out in triplicate in a total volume of 10  $\mu\text{l}$  containing 5  $\mu\text{l}$  of SsoAdvanced Sybr green mixture (Bio-Rad, Hercules, CA), 0.125  $\mu\text{l}$  of forward and reverse primers (100 nM each), and 1  $\mu\text{l}$  of DNA template using an initial denaturation at  $95^{\circ}\text{C}$  for 5 min, 40 cycles of denaturation at  $95^{\circ}\text{C}$ , annealing at 62, 64, or  $66^{\circ}\text{C}$ , depending on the primer combination, and extension at  $72^{\circ}\text{C}$ , each for 30 s (21, 23). The annealing temperature for primer 113f was  $64^{\circ}\text{C}$ . The amplifications were followed by melting curve analyses. The results of all analyses were corrected for extraction efficiencies, determined as the ratio of inoculated *Salmonella* cells detected by qPCR-based quantification of a 268-bp *invA* gene fragment before and after extraction (22).

**Qualitative analyses of *Frankia* in soil by Illumina sequencing.** A second set of triplicate rhizosphere samples was used to extract DNA without the addition of *Salmonella* cells. DNA was used in a nested PCR using the *nifH* gene of *Frankia* as the target for amplification as described before (25). Briefly,

606-bp fragments were retrieved using primers nifH1f and nifHr as described above for root nodule endophytes. These were cleaned using the UltraClean 15 DNA purification kit (Mo Bio Laboratories, Inc., Carlsbad, CA) and were quantified on a 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA) using the Agilent DNA 7500 kit. Dilutions of these amplicons were used as templates in nested PCRs using modified primers nifHf1 and nifH269 to create a 263-bp product (17). In the nested reactions, primer nifH269 was barcoded, and both primers were modified with Illumina adapters, primer pads, and linkers (25, 47). Samples were analyzed at the Genomic Sequencing and Analysis Facility (GSAF) at the University of Texas, Austin, on an Illumina MiSeq system with paired-end 250-bp reads using the respective sequencing and index sequence primers (25).

Illumina sequence data were analyzed as described previously (25). Briefly, paired-end reads were aligned in PEAR, v0.9.6, with default parameters and were then parsed and filtered in QIIME, v1.8.0 (48), using *split\_library.py* accepting reads with a Phred quality score of 26 or better (-q 25). The *nifH* gene sequence from *Frankia* strain EulK1 (U53362) was used as a BLAST reference, in addition to the seq.fna file from *split\_library.py* output, to run *exclude\_seqs\_by\_blast.py* so as to exclude any chimera or truncated reads due to PCR errors (25). The output file (matching.fna) served to get unique reads by running *pick\_denovo\_otus.py* with a similarity threshold set at 1.0. Sequences representing <1% of the total numbers of reads per sample were removed using the resulting BIOM table and a customized R script (25). The collected unique sequences were aligned in Geneious and were compared to *nifH* gene sequences of confirmed *Frankia* strains, as well as uncultured endophytes from nodules or soils obtained from GenBank.

**Accession number(s).** Representative sequences obtained in this study have been deposited at GenBank under accession numbers LT840167 to LT840183. Illumina sequence data were deposited in the NCBI Sequence Read Archive (SRP119405).

## ACKNOWLEDGMENTS

We are indebted to the Graduate College (Doctoral Research Support Fellowship to S. Ben Tekaya) and the Department of Biology at Texas State University for financial support.

## ARTICLE INFORMATION

Appl Environ Microbiol. 2018 Mar 1; 84(5): e02248-17.

Published online 2018 Feb 14. Prepublished online 2017 Dec 15. doi: [10.1128/AEM.02248-17](https://doi.org/10.1128/AEM.02248-17)

PMCID: PMC5812924

PMID: [29247058](https://pubmed.ncbi.nlm.nih.gov/29247058/)

Seifeddine Ben Tekaya,<sup>a</sup> Trina Guerra,<sup>a</sup> David Rodriguez,<sup>a</sup> Jeffrey O. Dawson,<sup>b</sup> and Dittmar Hahn<sup>a</sup>✉

Shuang-Jiang Liu, Editor

Shuang-Jiang Liu, Chinese Academy of Sciences;

<sup>a</sup>Texas State University, Department of Biology, San Marcos, Texas, USA

<sup>b</sup>University of Illinois at Urbana-Champaign, Department of Natural Resources and Environmental Sciences, Urbana, Illinois, USA

✉Corresponding author.

Address correspondence to Dittmar Hahn, [dh49@txstate.edu](mailto:dh49@txstate.edu).

**Citation** Ben Tekaya S, Guerra T, Rodriguez D, Dawson JO, Hahn D. 2018. *Frankia* diversity in host plant root nodules is independent of abundance or relative diversity of *Frankia* populations in corresponding rhizosphere soils. Appl Environ Microbiol 84:e02248-17. <https://doi.org/10.1128/AEM.02248-17>.

Received 2017 Oct 10; Accepted 2017 Dec 9.

Copyright © 2018 American Society for Microbiology.

All Rights Reserved.

This article has been cited by other articles in PMC.

Articles from Applied and Environmental Microbiology are provided here courtesy of American Society for Microbiology (ASM)

## REFERENCES

1. Chaia EE, Wall LG, Huss-Danell K. 2010. Life in soil by the actinorhizal root nodule endophyte *Frankia*. A review. Symbiosis 51:201–226. doi:10.1007/s13199-010-0086-y. [CrossRef] [Google Scholar]
2. Benson DR, Dawson J. 2007. Recent advances in the biogeography and genecology of symbiotic *Frankia* and its host plants. Physiol Plant 130:318–330. doi:10.1111/j.1399-3054.2007.00934.x. [CrossRef] [Google Scholar]

3. Dawson JO. 1986. Actinorhizal plants: their use in forestry and agriculture. *Outlook Agric* 15:202–208. doi:10.1177/003072708601500406. [[CrossRef](#)] [[Google Scholar](#)]
4. Dobritsa SV. 1998. Grouping of *Frankia* strains on the basis of susceptibility to antibiotics, pigment production and host specificity. *Intl J Syst Bacteriol* 48:1265–1275. doi:10.1099/00207713-48-4-1265. [[CrossRef](#)] [[Google Scholar](#)]
5. Hahn D, Starrenburg MJC, Akkermans ADL. 1988. Variable compatibility of cloned *Alnus glutinosa* ecotypes against ineffective *Frankia* strains. *Plant Soil* 107:233–243. doi:10.1007/BF02370552. [[CrossRef](#)] [[Google Scholar](#)]
6. Mirza MS, Hahn D, Akkermans ADL. 1992. Isolation and characterization of *Frankia* strains from *Coriaria nepalensis*. *Syst Appl Microbiol* 15:289–295. doi:10.1016/S0723-2020(11)80103-7. [[CrossRef](#)] [[Google Scholar](#)]
7. Mirza MS, Hahn D, Dobritsa SV, Akkermans AD. 1994. Phylogenetic studies on uncultured *Frankia* populations in nodules of *Datisca cannabina*. *Can J Microbiol* 40:313–318. doi:10.1139/m94-051. [[PubMed](#)] [[CrossRef](#)] [[Google Scholar](#)]
8. Baker D, Newcomb W, Torrey JG. 1980. Characterization of an ineffective actinorhizal microsymbiont, *Frankia* sp. Eu11 (*Actinomycetales*). *Can J Microbiol* 26:1072–1089. doi:10.1139/m80-180. [[PubMed](#)] [[CrossRef](#)] [[Google Scholar](#)]
9. Baker DD. 1987. Relationships among pure cultured strains of *Frankia* based on host specificity. *Physiol Plant* 70:245–248. doi:10.1111/j.1399-3054.1987.tb06139.x. [[CrossRef](#)] [[Google Scholar](#)]
10. Dobritsa SV, Novik SN, Stupar OS. 1990. Infectivity and host specificity of strains of *Frankia*. *Microbiology (New York)* 59:210–214. [[Google Scholar](#)]
11. Huang JB, Zhao ZY, Chen GX, Liu HC. 1985. Host range of *Frankia* endophytes. *Plant Soil* 87:61–65. doi:10.1007/BF02277648. [[CrossRef](#)] [[Google Scholar](#)]
12. Huss-Danell K, Myrold DD. 1994. Intragenic variation in nodulation of *Alnus*: consequences for quantifying *Frankia* nodulation units in soil. *Soil Biol Biochem* 26:525–531. doi:10.1016/0038-0717(94)90238-0. [[CrossRef](#)] [[Google Scholar](#)]
13. Maunuksela L, Hahn D, Hahtela K. 2000. Effect of freezing of soils on nodulation capacities of total and specific *Frankia* populations. *Symbiosis* 29:107–119. [[Google Scholar](#)]
14. Pokharel A, Mirza BS, Dawson JO, Hahn D. 2011. *Frankia* populations in soil and root nodules of sympatrically grown *Alnus* taxa. *Microb Ecol* 61:92–100. doi:10.1007/s00248-010-9726-2. [[PubMed](#)] [[CrossRef](#)] [[Google Scholar](#)]
15. Mirza BS, Welsh A, Rasul G, Rieder JP, Paschke MW, Hahn D. 2009. Variation in *Frankia* populations of the *Elaeagnus* host infection group in nodules of six host plant species after inoculation with soil. *Microb Ecol* 58:384–393. doi:10.1007/s00248-009-9513-0. [[PubMed](#)] [[CrossRef](#)] [[Google Scholar](#)]
16. Welsh A, Mirza BS, Rieder JP, Paschke MW, Hahn D. 2009. Diversity of frankiae in root nodules of *Morella pensylvanica* grown in soils from five continents. *Syst Appl Microbiol* 32:201–210. doi:10.1016/j.syapm.2009.01.002. [[PubMed](#)] [[CrossRef](#)] [[Google Scholar](#)]
17. Mirza BS, Welsh AK, Rieder JP, Paschke MW, Hahn D. 2009. Diversity of frankiae in soils from five continents. *Syst Appl Microbiol* 32:558–570. doi:10.1016/j.syapm.2009.07.008. [[PubMed](#)] [[CrossRef](#)] [[Google Scholar](#)]
18. Nickel A, Hahn D, Zepp K, Zeyer J. 1999. *In situ* analysis of introduced *Frankia* populations in root nodules of *Alnus glutinosa* grown under different water availability. *Can J Bot* 77:1231–1238. doi:10.1139/b99-066. [[CrossRef](#)] [[Google Scholar](#)]
19. Zepp K, Hahn D, Zeyer J. 1997. *In-situ* analysis of introduced and indigenous *Frankia* populations in soil and root nodules obtained on *Alnus glutinosa*. *Soil Biol Biochem* 29:1595–1600. doi:10.1016/S0038-0717(97)00041-2. [[CrossRef](#)] [[Google Scholar](#)]
20. Maunuksela L, Zepp K, Koivula T, Zeyer J, Hahtela K, Hahn D. 1999. Analysis of *Frankia* populations in three soils devoid of actinorhizal plants. *FEMS Microbiol Ecol* 28:11–21. doi:10.1111/j.1574-6941.1999.tb00556.x. [[CrossRef](#)] [[Google Scholar](#)]
21. Samant S, Huo T, Dawson JO, Hahn D. 2016. Abundance and relative distribution of *Frankia* host infection groups under actinorhizal *Alnus glutinosa* and non-actinorhizal *Betula nigra* trees. *Microb Ecol* 71:473–481. doi:10.1007/s00248-015-0643-2. [[PubMed](#)] [[CrossRef](#)] [[Google Scholar](#)]
22. Samant S, Sha Q, Iyer A, Dhabekar P, Hahn D. 2012. Quantification of *Frankia* in soils using SYBR green based qPCR. *Syst Appl Microbiol* 35:191–197. doi:10.1016/j.syapm.2011.12.004. [[PubMed](#)] [[CrossRef](#)] [[Google Scholar](#)]
23. Samant S, Amann RI, Hahn D. 2014. Evaluation of the 23S rRNA gene as target for qPCR based quantification of *Frankia* in soils. *System Appl Microbiol* 37:229–234. doi:10.1016/j.syapm.2013.11.001. [[PubMed](#)] [[CrossRef](#)] [[Google Scholar](#)]

24. Ben Tekaya S, Ganesan AS, Guerra T, Dawson JO, Forstner MRJ, Hahn D. 2017. Sybr green- and TaqMan-based quantitative PCR approaches allow assessment of the abundance and relative distribution of *Frankia* clusters in soils. *Appl Environ Microbiol* 83:e02833-. doi:10.1128/AEM.02833-16. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
25. Rodriguez D, Guerra TM, Forstner MRJ, Hahn D. 2016. Diversity of *Frankia* in soil assessed by Illumina sequencing of *nifH* gene fragments. *Syst Appl Microbiol* 39:391–397. doi:10.1016/j.syapm.2016.06.007. [PubMed] [CrossRef] [Google Scholar]
26. Cotin-Galvan L, Pozzi AC, Schwob G, Fournier P, Fernandez MP, Herrera-Belaroussi A. 2016. In-plant sporulation capacity enhances infectivity and rhizospheric competitiveness of *Frankia* strains. *Microb Environ* 31:11–18. doi:10.1264/jsme2.ME15090. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
27. Markham JH. 2009. Variability of nitrogen-fixing *Frankia* on *Alnus* species. *Botany* 87:1011. (Erratum.) doi:10.1139/B09-086. [CrossRef] [Google Scholar]
28. Zimpfer JF, Smyth CA, Dawson JO. 1997. The capacity of Jamaican mine spoils, agricultural and forest soils to nodulate *Myrica cerifera*, *Leucaena leucocephala* and *Casuarina cunninghamiana*. *Physiol Plant* 99:664–672. doi:10.1111/j.1399-3054.1997.tb05370.x. [CrossRef] [Google Scholar]
29. Diem HG, Dommergues YR. 1990. Current and potential uses and management of Casuarinaceae in the tropics and subtropics, p 365–385. In Schwintzer CR, Tjepkema JD. (ed), *The biology of Frankia and actinorhizal plants*. Academic Press, San Diego, CA. [Google Scholar]
30. Samant SS, Dawson JO, Hahn D. 2015. Growth responses of indigenous *Frankia* populations to edaphic factors in actinorhizal rhizospheres. *Syst Appl Microbiol* 38:501–505. doi:10.1016/j.syapm.2015.07.005. [PubMed] [CrossRef] [Google Scholar]
31. Wilcox DA, Cowan DA. 2016. Diversity of *Frankia* in root nodules of six *Morella* sp. from the Cape flora of South Africa. *Plant Soil* 406:375–388. doi:10.1007/s11104-016-2881-6. [CrossRef] [Google Scholar]
32. Samant S, Dawson JO, Hahn D. 2016. Growth responses of introduced *Frankia* strains to edaphic factors. *Plant Soil* 400:123–132. doi:10.1007/s11104-015-2720-1. [CrossRef] [Google Scholar]
33. Mirza BS, Welsh AK, Hahn D. 2009. Growth of *Frankia* strains in leaf litter-amended soil and the rhizosphere of a non-actinorhizal plant. *FEMS Microbiol Ecol* 70:132–141. doi:10.1111/j.1574-6941.2009.00746.x. [PubMed] [CrossRef] [Google Scholar]
34. Rönkkö R, Pennanen T, Smolander A, Kitunen V, Kortemaa H, Haahtela K. 1994. Quantification of *Frankia* strains and other root-associated bacteria in pure cultures and in the rhizosphere of axenic seedlings by high-performance liquid chromatography-based muramic acid assay. *Appl Environ Microbiol* 60:3672–3678. [PMC free article] [PubMed] [Google Scholar]
35. Rönkkö R, Smolander A, Nurmiäho-Lassila EL, Haahtela K. 1993. *Frankia* in the rhizosphere of nonhost plants: a comparison with root-associated nitrogen-fixing *Enterobacter*, *Klebsiella* and *Pseudomonas*. *Plant Soil* 153:85–95. doi:10.1007/BF00010547. [CrossRef] [Google Scholar]
36. Mirza BS, Welsh A, Hahn D. 2007. Saprophytic growth of inoculated *Frankia* sp. in soil microcosms. *FEMS Microbiol Ecol* 62:280–289. doi:10.1111/j.1574-6941.2007.00382.x. [PubMed] [CrossRef] [Google Scholar]
37. Battenberg K, Wren JA, Hillman J, Edwards J, Huang LJ, Berry AM. 2017. The influence of the host plant is the major ecological determinant of the presence of nitrogen-fixing root nodule symbiont cluster II *Frankia* species in soil. *Appl Environ Microbiol* 83:e02661-. doi:10.1128/AEM.02661-16. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
38. Normand P, Simonet P, Bardin R. 1988. Conservation of *nif* sequences in *Frankia*. *Mol Gen Genet* 213:238–246. doi:10.1007/BF00339587. [PubMed] [CrossRef] [Google Scholar]
39. Pearson WR, Lipman DJ. 1988. Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A* 85:2444–2448. doi:10.1073/pnas.85.8.2444. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
40. Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690. doi:10.1093/bioinformatics/btl446. [PubMed] [CrossRef] [Google Scholar]
41. Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees, p 1–8. In *Proceedings of the Gateway Computing Environments Workshop (GCE)*, New Orleans, LA. [Google Scholar]
42. Tavaré S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences, p 57–86. In *Lectures on Mathematics in the Life Sciences*, vol 17 American Mathematical Society, Providence, RI. [Google Scholar]
43. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. doi:10.1093/bioinformatics/btu033. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

44. Felsenstein J. 1985. Confidence limits of phylogenies: an approach using the bootstrap. *Evolution* 39:783–791. doi:10.1111/j.1558-5646.1985.tb00420.x.

[PubMed] [CrossRef] [Google Scholar]

45. Klindworth A, Priesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner F-O. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41(1):e1. doi:10.1093/nar/gks808. [PMC free article] [PubMed] [CrossRef]

[Google Scholar]

46. Quast C, Priesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner F-O. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41(Database issue):D590–D596. doi:10.1093/nar/gks1219. [PMC free article] [PubMed] [CrossRef]

[Google Scholar]

47. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6:1621–1624.

doi:10.1038/ismej.2012.8. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

48. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Tumbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336.

doi:10.1038/nmeth.f.303. [PMC free article] [PubMed] [CrossRef] [Google Scholar]