



Unveiling root nodulation in *Tribulus terrestris* and *Roystonea regia* via metagenomics analysis

Saleh Rahimlou^{1,2} · Mahdieh S. Hosseyni Moghadam¹ · Romina Gazis^{3,4} · Elena Karlsen-Ayala^{4,5} · Mohammad Bahram^{6,7} · Timothy Y. James² · Leho Tedersoo¹

Received: 2 July 2024 / Accepted: 14 December 2024

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

Root nodule symbiosis is traditionally recognized in the Fabales, Fagales, Cucurbitales, and Rosales orders within the Rosid I clade of angiosperms. However, ambiguous root nodule formation has been reported in Zygophyllaceae and Roystonea regia (Arecaceae), although a detailed analysis has yet to be conducted. We aimed to perform morphological analyses of root structures in these plants and utilize metagenomic techniques to identify and characterize the bacterial populations within the nodule-like structures. We collected root samples of *Tribulus terrestris* (Zygophyllaceae) and *Roystonea regia* from West Asia and the Caribbean, respectively. We conducted detailed morphological analyses of nodule-like structures, isolated and genome-sequenced the endophytes, and employed metagenomic techniques to identify the bacterial populations within these formations. We observed nodule-like structures in both plant species. Symbiosomes, which are hallmark structures of nodulating plants, were not detected. Metagenome sequence data analysis revealed potential nodulating and nitrogen-fixing bacteria in the nodule-like structures of both species. Canonical nodulation and nitrogen-fixation genes were identified in microbes associated with the nodules. However, the phylogenomic analysis showed that the bacteria isolated from *T. terrestris* and *R. regia* are within Gammaproteobacteria and Bacilli, which are not typically known as nodulating bacteria. The observed structures differ significantly from traditional nodules found in legumes and actinorhizal plants, suggesting unique characteristics with hosting nitrogen-fixing bacteria. Although bacteria identified through in silico analysis or culture are well-known nitrogen-fixers, their specific role in root nodule formation remains to be investigated.

Keywords Actinobacteria · Actinorhizal plants · Legumes · Nitrogen · Nodule · Rhizobia · Symbiosis

Communicated by Bing Yang.

✉ Saleh Rahimlou
Rahimlou@umich.edu

¹ Institute of Ecology and Earth Sciences, University of Tartu, Liivi 2, 50409 Tartu, Estonia

² Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA

³ Department of Plant Pathology, Tropical Research and Education Center, University of Florida, Homestead, FL 33031, USA

⁴ Department of Plant Pathology, University of Florida, Gainesville, FL 32608, USA

⁵ USDA Forest Service, Northern Research Station, Hamden, CT, USA

⁶ Department of Ecology, Swedish University of Agricultural Sciences, Ulls Väg 16, 756 51 Uppsala, Sweden

⁷ Department of Agroecology, Aarhus University, Slagelse 4200, Denmark

Introduction

Approximately 78% of our atmosphere consists of nitrogen (N), making it the core component of the air that we breathe. The atmospheric nitrogen is present in the form of N₂, which can be biologically utilized by relatively small groups of N-fixing bacteria and archaea. Other organisms such as plants rely on more reactive forms of N such as ammonium and nitrate for growth. Agriculture uses a large amount of synthetic N fertilizers, which are expensive to produce and energy-intensive and may result in severe or irreversible damage to the environment, including nitrate pollution of surface- and groundwater, dead zones on the coast, and increased emissions of nitrous oxides and carbon dioxide (Eickhout et al. 2006). Alternatively, the biological process of N fixation produces between 50 and 70 Tg of bioavailable nitrogen every year for agricultural systems, which is a significant contribution to global food security

(Herridge et al. 2008). The most efficient sources of biologically fixed nitrogen are symbiotic relationships between plants and rhizobia, which cause the root nodules formation.

N-fixing endosymbiosis between bacteria and plants is known to be restricted to angiosperms. One group of gymnosperms (the cycads), several mosses, and the fern *Azolla* all form N-fixing symbioses with cyanobacteria, but in these associations, the bacteria do not infect the plant cells (Peters 1991; Tedersoo et al. 2018). Bacteria inhabiting a specialized compartment of a host cell (symbiosomes/fixation-threads) can be found in legume root nodules—an organ that provides optimal physiological conditions for N fixation (Coba de la Pena et al. 2018). In actinorhizal plants, N fixation occurs within specialized structures known as vesicles. These vesicles provide an anaerobic environment that is essential for the effective functioning of the N-fixing bacteria, such as *Frankia* (Wall 2000). Rhizobia is a polyphyletic diazotrophic group of bacteria that form root nodule symbioses, belonging to the Alpha- and Betaproteobacteria classes. Most legumes form symbiotic and N-fixing associations with species of Alphaproteobacteria. Several legumes, particularly those of the genus *Mimosa*, are nodulated primarily by Betaproteobacteria. There have been a few reports of Gammaproteobacteria inducing nodules on legumes, however, they have not yet been confirmed (Shiraishi et al. 2010; Gyaneshwar et al. 2011; Moulin et al. 2015). *Frankia*, a gram-positive and filamentous genus from Actinobacteria is another bacterial lineage fixing atmospheric N in symbiosis with eight plant families so-called actinorhizal plants. So far, the only known exception to this is *Parasponia* (Cannabaceae, Rosales) which is known to be promiscuously associated with four rhizobial genera: *Bradyrhizobium*, *Rhizobium*, *Mesorhizobium*, and *Sinorhizobium* (Op Den Camp et al. 2012).

While Cyanobacteria form root nodules in Cycadophyta, rhizobial and actinorhizal root nodule symbiosis is thought to be limited to four plant orders (Fabales, Fagales, Cucurbitales, Rosales) that all share a common ancestor in the Rosid I clade of angiosperms (Soltis et al. 1995) and form the N-fixing clade. However, nodulating taxa within these four clades are scattered and interspersed with many non-nodulating lineages. There are two main hypotheses to explain the paraphyletic distribution of nodulating plant taxa. First, nodulation may have evolved independently multiple times, preceded by a single hypothetical predisposition event in the common ancestor of the N-fixing clade within Rosids (Doyle 1998, 2011, 2016; Werner et al. 2014; Li et al. 2015; Martin et al. 2017). In 2018, two studies proposed an alternative hypothesis based on comparative phylogenomics: nodulation originated at the root of the currently known N-fixing clade, followed by massive occasions of independent losses (Griesmann et al. 2018; Van Velzen et al. 2019). This hypothesis is based on the loss of the central symbiotic

regulator gene (*NIN*), the infection-associated gene (*RPG*), and the receptor *NFR5* in most of the non-nodulating species within the N-fixing clade. Additional support was provided to this hypothesis using comparative transcriptomics that revealed gene sets expressed in nodules from all four families and inferred to form the ancestral nodule transcriptomes (Libourel et al. 2023; Zhang et al. 2024).

Putative N-fixing nodules associated with rhizobia have been reported in *Tribulus terrestris* L. and *Roystonea regia* Cook (Issatschenko 1913; Sabet 1946; Mostafa and Mahmoud 1951; Athar and Mahmood 1972; Athar and Mamood 1981; Basu et al. 1997; Mahmood and Athar 1998, 2006). For the first time, Isachenko (1913) revealed that nodular growth in *T. terrestris* roots in the Southern Bug region of Ukraine (formerly part of the Soviet Union) contained a fungus. In the nodules of *T. terrestris*, Mahmood and Athar 1998, 2006 described the presence of abundant rhizobia and cyanobacteria cells using electron microscopy. Allen and Allen (1950), Sahibzada and Khan (1973), and Gresshoff et al. (1984) failed to find any endophyte in the nodular growths on *Tribulus* roots and concluded that nodulation of *Tribulus* was associated with the starch content of the plant. There is a lack of information about hypertrophies observed on *Tribulus* roots, and the specific agents causing these nodule-like structures—whether they are fungi, rhizobia, *Frankia*, or diverse bacterial species—are unclear. Additionally, a study conducted in India revealed that the roots of *R. regia* trees harbor nodules containing *Rhizobium* bacteria (Basu et al. 1997). This was the first time nodules have been reported on the roots of a monocotyledonous tree. Nitrogen fixation was further demonstrated by the presence of nitrogenase (an enzyme used in the N fixation process) and leghaemoglobin, which allows nitrogenase to function by reducing oxygen concentration in the root nodule (Basu et al. 1997).

The plant genus *Tribulus*, commonly known as puncturevine or goat's head, encompasses a group of flowering plants belonging to the family Zygophyllaceae. Phylogenetic analyses have revealed *Tribulus*' placement within the order Zygophyllales, a group characterized by its xeromorphic adaptations and diverse floral structures. These hardy and drought-resistant plants are native to warm temperate and tropical regions worldwide, thriving in arid and semi-arid habitats (Sheahan and Chase 2000). *Tribulus* species are characterized by their prostrate or ascending growth habits and deeply lobed leaves, often adorned with sharp spines or thorns. They produce attractive yellow flowers with five petals and fruit capsules that split into four to six segments, each containing hard, spiny seeds (Wu et al. 2018). Despite their spiny appearance, *Tribulus* species have been utilized in traditional medicine and as a source of food and fodder in various cultures. Additionally, certain *Tribulus* species are recognized for their potential medicinal properties and

have been studied for their pharmacological effects, including anti-inflammatory and aphrodisiac properties (Singh and Gupta 2011; Abbas et al. 2022).

The genus *Roystonea*, belonging to the family Arecaceae, comprises a distinguished assemblage of palm trees recognized for their majestic stature and ornamental value. Endemic to tropical regions of the Caribbean and certain parts of Florida. *Roystonea* species exhibit distinctive morphological characteristics, including towering heights, smooth grey trunks, and gracefully arching fronds (Roncal et al. 2008). Phylogenetic analyses have situated *Roystonea* within the diverse family Arecaceae, shedding light on its evolutionary relationships and taxonomic placement within the plant kingdom (Asmussen et al. 2006). Their preference for warm, humid climates and well-drained soil makes them well-suited for landscaping in tropical and subtropical regions. Beyond their aesthetic appeal, *Roystonea* palms serve as vital components of ecosystems, providing habitat and sustenance for various wildlife species. Their resilience to adverse environmental conditions, including strong winds and coastal environments, further underscores their ecological importance.

In this study, we aimed to investigate the presence of nodule-like structures in the root systems of *T. terrestris* and *R. regia*. To achieve this, we collected root samples of *T. terrestris* from West Asia and *R. regia* from the Caribbean. We conducted detailed morphological analyses of the structures formed in the roots of these plants. Additionally, we employed metagenomic techniques to identify and characterize the bacterial populations residing within the nodule-like structures. This involved sequencing the genomes of bacteria isolated from surface-sterilized nodule-like structures to obtain a comprehensive understanding of the microbial communities associated with these formations. Our integrated approach combining morphological studies and advanced metagenomic sequencing aimed to provide insights into the potential role of these bacteria in the formation and function of nodule-like structures in these plant species.

Materials and methods

Sampling, DNA extraction, and PCR

Root samples were taken from *T. terrestris* growing in semi-arid lands of Iran, Saudi Arabia, South Africa, Benin, and Western Australia and *R. regia* from Florida (USA), Yucatán (Mexico), Dominican Republic, and Cuba between the years 2018–2022. From each sampling site, we selected three plant individuals with at least a 10 m distance. Roots were gently extracted from the soil and washed with tap water. From cleaned roots, 10 root pieces (approx. 1 cm in size) bearing mature nodules were separated. Root pieces were dried

by pressing them into a paper towel and stored in fixation buffer, RNAlater solution (ThermoFisher Scientific), and 1% PBS solution for further analyses. Total DNA from root nodules was extracted using the CTAB protocol (dx.doi.org/<https://doi.org/10.17504/protocols.io.bhx8j7rw>) and DNeasy Plant mini kit (Qiagen) according to the manual. We used several general primer pairs and species-specific custom-designed primers to amplify the structural *nodABC* and *nifHDK* genes (Table 1). Individual PCR reactions were carried out using 5 µl of HOT FIREPol® Blend Master Mix (Solis BioDyne) with 2 mM MgCl₂ final concentration, 1 µl of each primer pair, and 2 µl of DNA sample diluted to a final volume of 25 µl. The PCR conditions were adjusted to initial denaturation at 94 °C for 5 min, followed by 35 cycles of 94 °C for 1 min, suitable annealing temperature for each primer pair for 1 min (Table 1), 72 °C for 1 min, and a final extension at 72 °C for 7 min using Eppendorf 5345 Mastercycler Ep Gradient S PCR Thermal Cycler (Applied Biosystems™). PCR products were purified using PCR Clean-Up kit (Promega) and Sanger sequenced in Eurofins, Germany. Sequences were trimmed, quality filtered using Sequencher v.5.1 and deposited in GenBank (accession numbers provided in Table 1).

Isolation into culture

Nodules were excised from the roots and surface sterilized in 70% ethanol for 1 min and 2.5% Na hypochlorite for 1–3 min and washed with sterile water 3 times. Then nodules were squashed in a 1.5 ml tube using sterile plastic pestles. Nodule juice dispersed across the YMA (Yeast Mannitol Agar) plates (mannitol: 10 g/L, KH₂PO₄: 0.5 g/L, K₂HPO₄: 0.5 g/L, MgSO₄*7H₂O: 0.2 g/L, yeast extract: 0.5 g/L, NaCl: 0.2 g/L, Agar: 15 g/L, pH: 6.8). Plates were incubated upside down overnight at 28 °C. Individual colonies were selected and streaked onto fresh YMA plates. Once re-grown, a single colony from a pure isolate on the YMA plate was taken and used to inoculate a sterile 5 ml TY broth (tryptone 5 g/L, yeast extract 3 g/L, CaCl₂*xH₂O: 0.89 g/L, pH: 6.8). After 24 h (at log phase), 0.9 ml of culture was combined with 0.9 ml of sterile 50% glycerol. The stock was mixed, frozen in liquid nitrogen, and subsequently used for DNA extractions and PCR.

Nodule microscopy

The nodules were removed from the roots, included in 3% agarose gel, and cut into 50 µm sections using a vibratome (Leica VT1000S vibratome). Sections were observed in brightfield with Leica DM RXA2 microscope, Nikon Eclipse Ti inverted widefield microscope or stereo microscope Zeiss Axiozoom V16. Some of the sections were stained with Congo red, WGA-alexa488, and Syto9 at 1/1000 in

Table 1 Sequences of primer sets used to amplify *nod* and *nif* genes from *T. terrestris* and *R. regia* nodules

Gene	Primer sequences	Annealing temperature	Length (bp)	GenBank accession	Reference
<i>nodAB</i>	5'-TGCRGTGGARDCTRYGCTGGGAAA-3' 5'-GNCCGTCRTCRAASGTCARGTA-3'	64	560	MW628881	Elliott et al., 2007
<i>nodB</i>	5'-GCACCTCGATTACATACACGA-3' 5'-CGAATTGCAAAACCACGCTC-3'	64	557	MW628882	This study
<i>nodC</i>	5'-TTATGGGTCAACTCACGGCT-3' 5'-CAAGTGCAGTCAGGTGTCAC-3'	64	690	MW628883	This study
<i>nifH</i>	5'-CTCCAAAATACACTCGCCG-3' 5'-GCCGGAATTTGCATGGATCT-3'	60	696	MW628884	This study
<i>nifD</i>	5'-CAAAACAACAATCGTGCCGG-3' 5'-GCGCTGATAATCGTCGTTGT-3'	60	631	MW628885	This study
<i>nifK</i>	5'-CGTCACTGATCAACGCGAAT-3' 5'-CGCTCTCATGTCCTCGATCT-3'	60	692	MW628886	This study
<i>nodA</i>	5'-TCACARCTCKGGCCCGTTCCG-3' 5'-TGGGCSGGNGCNAGRCCBGA-3'	58	—	—	Moulin et al. 2000
<i>nodC</i>	5'-AGGTGGTYGAYGACGGTTC-3' 5'-CGYGACAGCCANTCKTATTG-3'	58	—	—	Laguerre et al. 2001
<i>nodCI</i>	5'-CCATACGCACCGTGG TGCTCTTGC-3' 5'-TCTTCCATVAWRTGVGTNGTCA-3'	58	—	—	Moulin et al. 2000
<i>nodDC</i>	5'-GTACAGGAGGGCATCGCGAA-3' 5'-CTGCAGCTGCAGCGAATCTG-3'	58	—	—	Zhang et al. 2000
<i>nifH</i>	5'-TACGGNAARGSGGNATCGGCAA-3' 5'-AGCATGTCYTCSAGYTCNTCCA-3'	58	—	—	Laguerre et al. 2001
<i>nifHD</i>	5'-GCCWTCTAYGGNAARGGNGG-3' 5'-ATCAGGCCGATCGGGCATT-3'	58	—	—	Moulin et al. 2000

water and observed with a confocal microscope Leica TCS-SP8-CSU with objective 25x/0.8. Dyes were excited with laser 488 nm, and fluorescence collected between 500 and 550 nm.

Metagenome sequencing and data analysis

To prepare libraries for metagenomics we used Nextera XT DNA Library Preparation Kit (Illumina Inc.) according to the manual. We generated metagenome sequences for ten root nodule samples of *T. terrestris* and four samples of *R. regia* using Illumina NovaSeq 2 × 150 paired-end sequencing machine. Raw reads were trimmed and quality filtered with Trimmomatic v.0.39 (Bolger et al. 2014) using ILLUMINACLIP: Nextera-PE.fa:2:30:10:2:keepBothReads LEADING:3 TRAILING:3 MINLEN:36 parameter values. Spades v.3.15.5 (Prjibelski et al. 2020) with meta flag and default kmer size and MegaHit v.1.1.3 (Li et al. 2016) with kmer sizes ranging from min = 15 to max = 255 with increment = 2 were utilized to assemble the filtered sequences using the Great Lakes HPC system (University of Michigan). Statistics for each metagenome assembly are provided in Table S1. A combination of Meta-bat2 (Kang et al. 2019), MaxBin2 (Wu et al. 2016), and Concoct (Alneberg et al. 2014) was used for clustering

the assembled contigs into taxonomic bins. Based on the results of all binning tools, the optimal and non-redundant set of bins was obtained using DAS Tool v.1.1.6 (Sieber et al. 2018). To enhance assembly contiguity, the obtained bins were reassembled using the MetaWRAP v.1.3.2 pipeline (Uritskiy et al. 2018). CAT/BAT v5.2.3 (Von Meijenfildt et al. 2019) was used for the taxonomic classification of metagenomic bins. To minimize noise from plant sequences in metagenomes, sequencing runs containing fewer than 90% plant sequences were selected. In *R. regia*, only the WF1.FI sample survived this threshold by containing 13% plant sequences. The samples of *T. terrestris* (IR1-10) were dominated by plant sequences between 76 and 94 percent. Since the sequencing depth was low for *T. terrestris* and plant sequences had a higher noise, we combined all 10 runs and reassembled them together. In total, the combined dataset contained 82% of plant sequences. Metagenomic assemblies were annotated using EggNOG-mapper pipeline v.2.1.10 using Diamond or Prodigal searches (Cantalapiedra et al. 2021). Metaxa v.2 (Bengtsson-Palme et al. 2015) and ITSx v.1.1.3 (Bengtsson-Palme et al. 2013) were utilized to extract and calculate the abundance of prokaryotic SSU and fungal ITS sequences from the assemblies respectively.

Genome sequencing and N-fixing genes characterization

Library preparation and sequencing were performed for seven bacterial strains isolated from root nodules of *T. terrestris* and *R. regia* using the PacBio Sequel II and Novaseq instruments. Genome assembling was performed using Canu v.2.2 (Koren et al. 2017) for PacBio and Megahit v.1.1.3 for Illumina reads with default parameter settings. For the Megahit assembler, we used Kmer sizes ranging from min = 15 to max = 255 with increment = 2 after each iteration. Genome completeness was measured with BUSCO v.5.1.12 (Manni et al. 2021) using 'bacteria_odb10' lineage. The assemblies were functionally annotated using EggNOG-mapper pipeline v.2.0.6. Average nucleotide identity (ANIb) analysis was performed with 14 *Kosakonia* genome assemblies retrieved from GenBank as well as the R1.FI isolate genome using Pyani v.0.2 (Pritchard et al. 2016). The GEnView pipeline was used to characterize the structural N-fixing genes in *K. sacchari* (R1.FI) genome and their gene-environment with 50% identity and 60% coverage (Ebmeyer et al. 2022). All genomes were deposited in GenBank under the Bioproject number PRJNA818535.

Phylogenetic analyses

The UBCG pipeline v.3 (Na et al. 2018) was used for species tree construction from 41 bacterial genome assemblies obtained from GenBank as well as seven genomes generated in the current study. Multiple sequence alignments were generated using MAFFT v.7.271 (Katoh 2005) applying the 100% cut-off value for gap-containing positions and the 'codon' method to account for amino acid sequences. Finally, a concatenated alignment of 78,879 and 84,207 nucleotide character states were subjected to tree estimation. The alignment matrices are provided in supplementary materials. IQtree v.2 (Minh et al. 2020) was applied for phylogenetic tree construction with 1000 bootstrap iterations. The phylogenetic trees were visualized using FigTree v.1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree>).

Nucleotide sequences of each *nodABC* and *nifHDK* genes from various rhizobial taxa were retrieved from their corresponding genome sequences in GenBank. Gene datasets were aligned including our sequences from *T. terrestris* nodules using MAFFT through the EMBL-EBI server (www.ebi.ac.uk/Tools/msa/mafft). Following this, the alignments were manually adjusted using Mesquite v.2.75 (Maddison and Maddison 2019). We used IQtree v.2 to run several preliminary phylogenetic analyses using a single gene dataset to check for incongruences in the phylogenetic position of selected taxa. Thereafter the concatenated dataset of *nodABC* and *nifHDK* genes were generated using Mesquite v.2.75. The final alignments for *nod* and *nif* genes included

26 and 22 taxa with 1807 and 2019 character states, respectively. Alignments of the concatenated taxa are provided in supplementary materials. Phylogenetic analyses were performed using IQtree v.2 with the concatenated datasets using 50 replicated likelihood searches followed by 1000 bootstrap iterations. The general time reversible (GTR) model along with gamma distribution was selected as the nucleotide substitution model. The trees were midpoint rooted and visualized using FigTree v.1.4.3.

Results

Nodule morphology

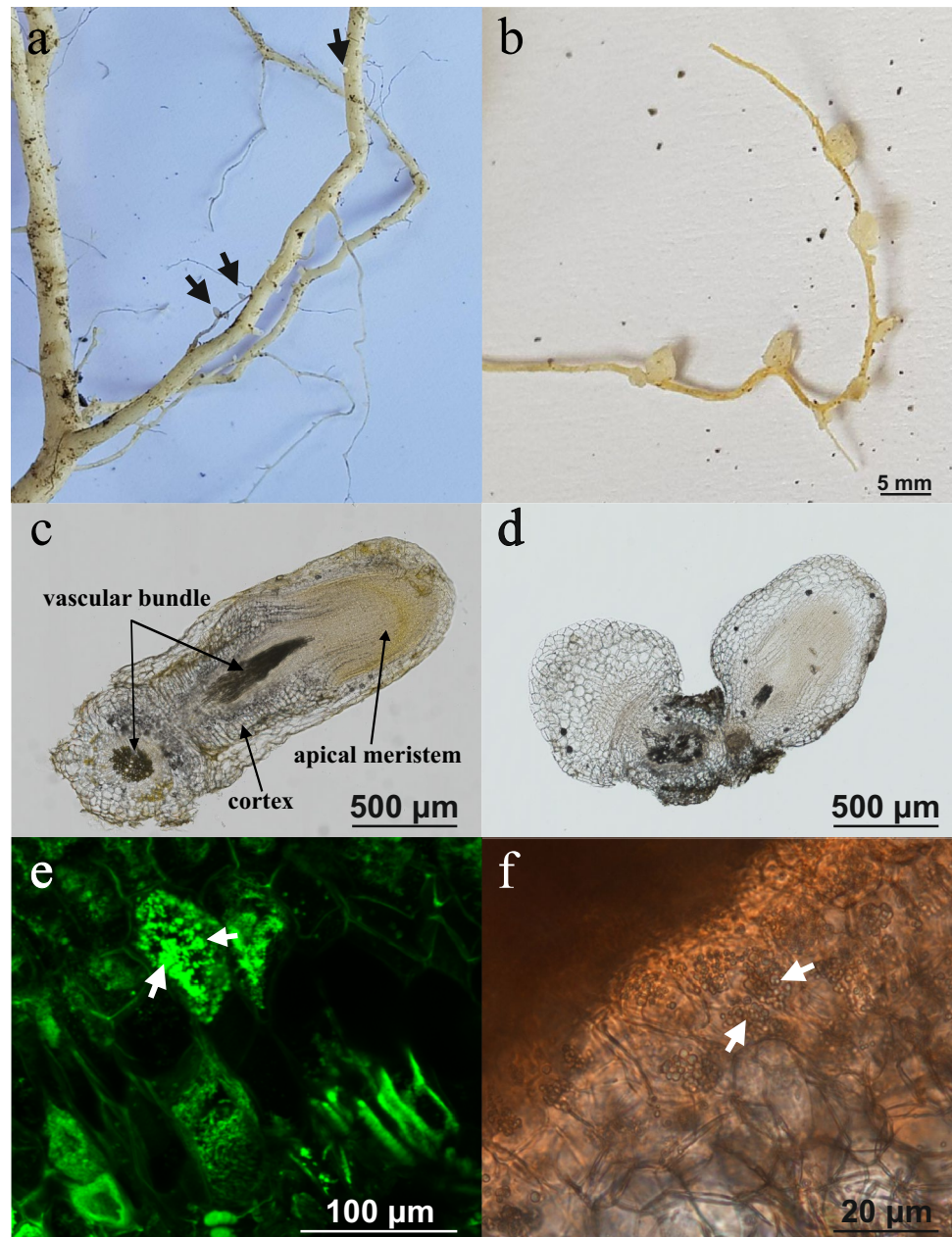
Nodule-like structures were observed on *T. terrestris* main and lateral roots of samples collected from Iran (Figs. 1, S1). The structures averaged 1.5 to 2 mm in diameter and were loosely attached to the root system and are white to yellowish, similar to the color of the roots (Fig. S1). Field observations demonstrated that maximum nodulation occurred in sandy soils and sandy gravels with low moisture content. Root hair deformation was observed in the field which is indicative of early physiological responses initiated by the nod factor perception (Fig. S2). An apical meristem and central vascular bundles are visible in nodules (Fig. 1a). Some bacterial infections were observed in the outer cortex or epidermal cells (Fig. 1c). Cortex cells contained globular structures, which resemble underground storage organs, such as starch granules or amyloplasts (Fig. 1d, e). Root samples collected from Australia, South Africa, and Benin showed no nodules, suggesting that nodule formation may be restricted to specific Middle Eastern varieties, or that the necessary biotic or abiotic environmental conditions required for nodule formation are lacking in the other sampled locations.

Structures resembling nodules were observed on the roots of *R. regia* (Figs. 2, S2) obtained from various locations, including Florida (USA), Yucatán (Mexico), the Dominican Republic, and Cuba. These structures are creamy in color, oval to cylindrical, 2 to 2.5 mm in diameter with an apical meristem and a central vascular system (Fig. S2e, f). These structures are less common in rocky or dry soils under natural conditions. Tubular structures were observed inside the nodule cortical cells that are similar to fungal hyphae in terms of thickness (Fig. 2d, e). Electron microscopy of the nodule samples revealed structures similar to those described by Basu et al. (1997) as fixation threads (Fig. 2f, g).

Metagenomic analysis

The metagenome sequences were primarily comprised of plant sequences, indicating the dominance of the host in the

Fig. 1 *Tribulus terrestris* root nodule-like structure. White to yellowish nodules connected to the root hairs (**a**, **b**; arrows indicate nodules), longitudinal sections (**c**, **d**), central vascular bundle, cortex, and apical meristem (**c**) two young nodules have formed on sides of a rootlet (**d**), structures resembling starch granules indicated by white arrows captured using confocal fluorescence and brightfield microscopy (**e**, **f**)

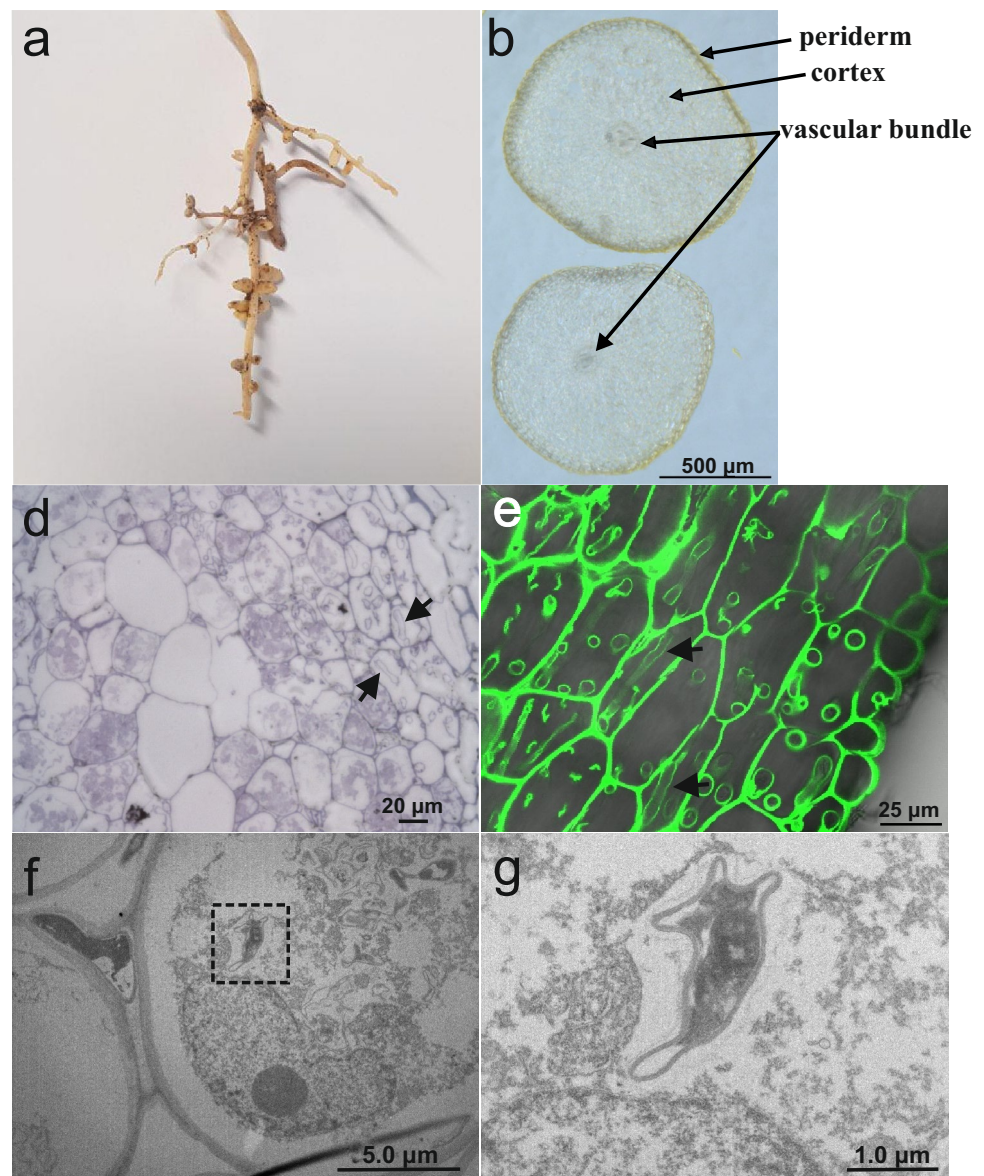


assembly. Following the exclusion of sequences corresponding to Streptophyta, we successfully retrieved a total of 15 and 12 metagenomic bins for the *T. terrestris* and *R. regia* sequencing run respectively. Our taxonomic classification indicates that *R. regia* nodules harbor more diverse bacteria than *T. terrestris* nodules (Figs. 3, 4). Based on the analysis of whole metagenomes sequences and the SSU sequences extracted from assemblies, root nodules of *T. terrestris* were predominantly associated with Actinobacteria, whereas Alphaproteobacteria were the dominant group in *R. regia* nodules (Figs. 3, 4). We recovered genomes of *Streptomyces*, *Glycomyces*, and *Lentzea* from *T. terrestris* assemblies, which are not commonly known as nodule-inducing

genera (Fig. 3) (Rahimlou et al. 2021). *Bradyrhizobium* was the common nodulating bacterial genus whose genome was obtained from *R. regia* root nodules (Fig. 3). According to CheckM v.1.2.2 (Parks et al. 2015), the recovered *Bradyrhizobium* genome is 96.05% complete with 14.54% contamination (Fig. S3). However, we were unable to find the canonical *nodABC* genes associated with *Bradyrhizobium*. Despite many attempts, none of those genera could be isolated in the YMA medium after surface sterilization of nodular tissues.

The taxonomic classification of bacterial SSU sequences extracted from assemblies indicates that Actinobacteria and Alphaproteobacteria are dominant taxa

Fig. 2 *Roystonea regia* root nodule-like structure. Creamy oval nodules attached to the rootlets (a), transverse section indicating central vascular bundle, cortex, and periderm (b), resin-embedded semi-thin section stained with Toluidine Blue, displaying tubular structures marked with arrows, captured using an inverted widefield microscope (d), confocal fluorescent image highlights tubular structures, which are indicated by arrows (e), TEM image shows structures resembling those described by Basu et al. (1997), referred to as fixation threads (f), a magnified view of the structure observed in panel f (g)



colonizing *T. terrestris* and *R. regia* root nodules, respectively (Fig. 4). A number of less abundant SSU sequences were identified matching nodulating/N-fixing bacteria such as *Methylobacterium*, *Paenibacillus*, *Cupriavidus*, *Devosia* and *Rhizobium* in *R. regia* and *Microbacterium*, *Bosea* and *Frankia* in *T. terrestris* (Table S2). Taxonomy assignment of raw reads at the genus level revealed that 93 and 56% of the reads were labeled as unclassified in *T. terrestris* and *R. regia* samples, respectively. In *T. terrestris*, Actinobacteria genera such as *Streptomyces*, *Lentzea*, *Glycomyces*, and *Nocardioides* were more prominent, as we obtained bins corresponding to these taxa (Table S3). In *R. regia*, *Bradyrhizobium* is highly abundant constituted 8% of the reads, and other N-fixing taxa like *Paenibacillus* also noted among the most prevalent taxa (Table S3). To explore the genetic relationship between nodulating

agents in *T. terrestris* and *R. regia* and commonly known nodulating bacteria, we aligned their metagenomic assemblies with 45 genomes of well-established nodulating bacteria. The assemblies of *T. terrestris* and *R. regia* exhibited strong alignment with *F. meridionalis* and *B. japonicum*, respectively (Fig. 5a, b). Supporting this finding, several SSU sequences matching *Frankia* and *Bradyrhizobium* were identified within the metagenomic datasets of *T. terrestris* and *R. regia*, respectively (Table S2). The functional annotation of metagenome assemblies of *T. terrestris* uncovered *nodA* genes from Pseudonocardiales and Micrococcaceae, *nodB* gene from Xanthomonadales, and several *nodC* genes not attributed to any specific taxa (Table S4). Furthermore, within the *T. terrestris* metagenome, we identified *nifH* genes associated with Actinobacteria and Clostridia (Table S4). In our analysis of *R.*

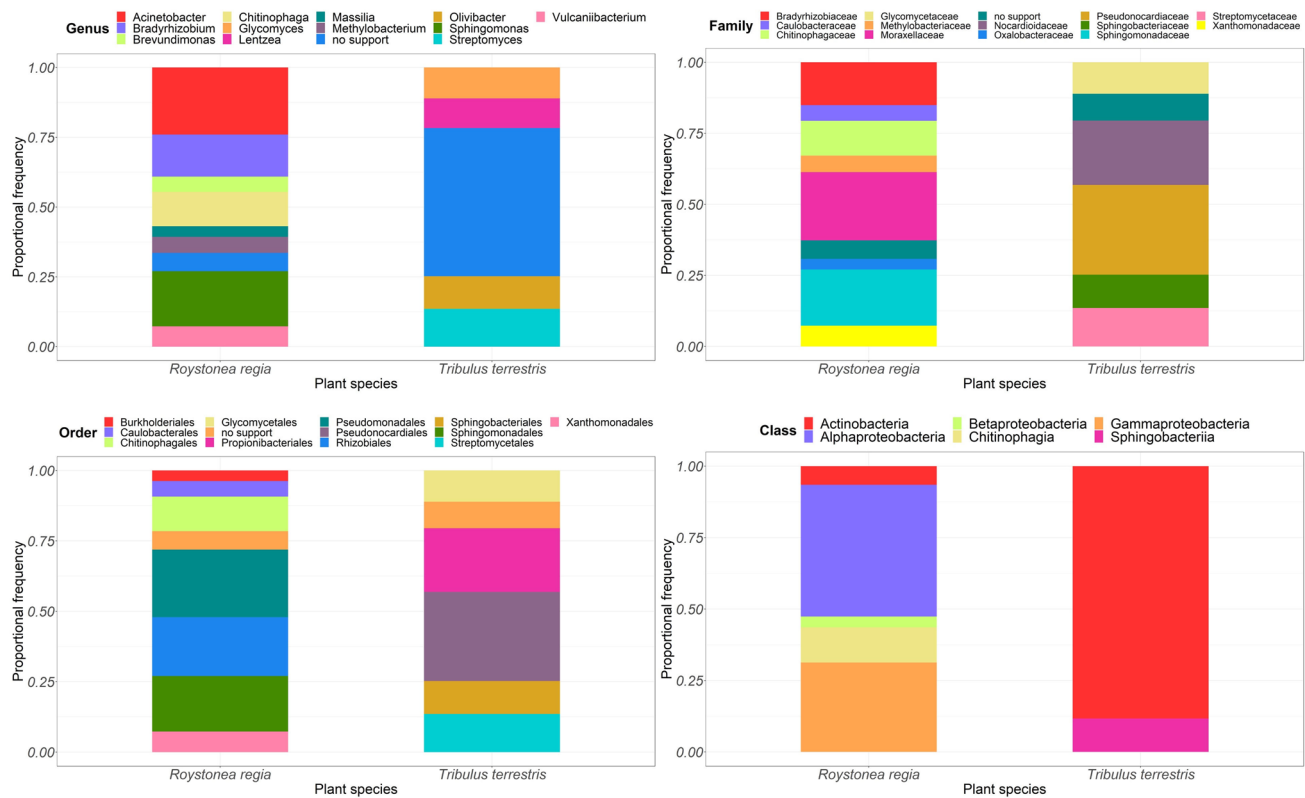


Fig. 3 Taxonomic composition of *R. regia* and *T. terrestris* metagenomic sequencing bins at genus, family, order, and class levels

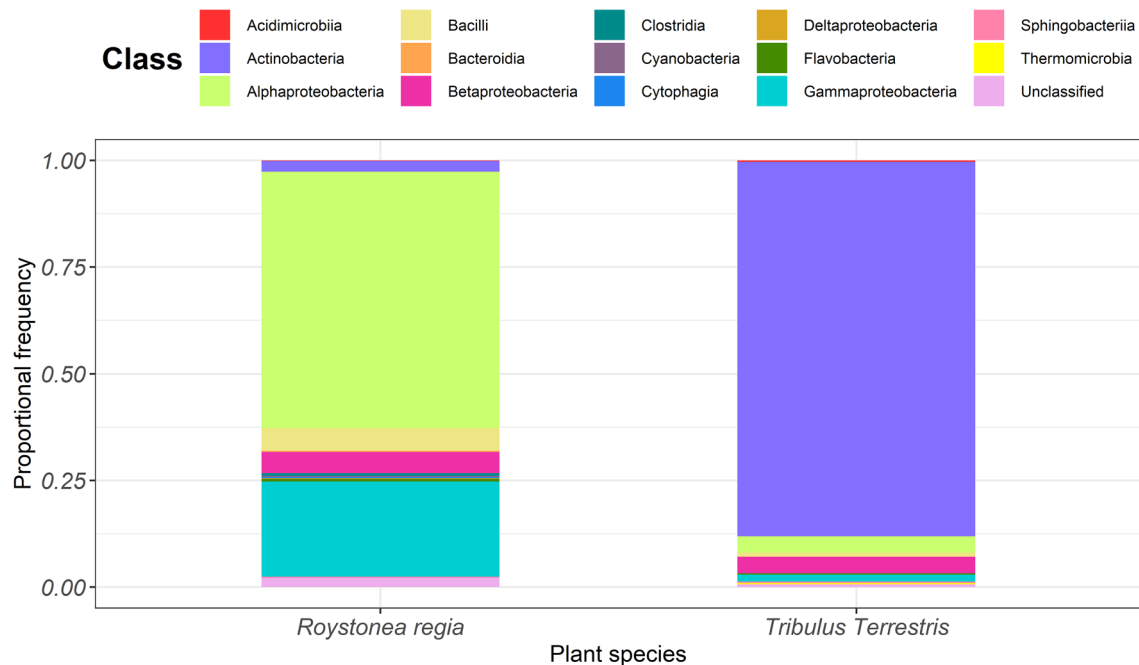


Fig. 4 Taxonomic composition of extracted SSU sequences from *R. regia* and *T. terrestris* metagenomic assemblies at class level

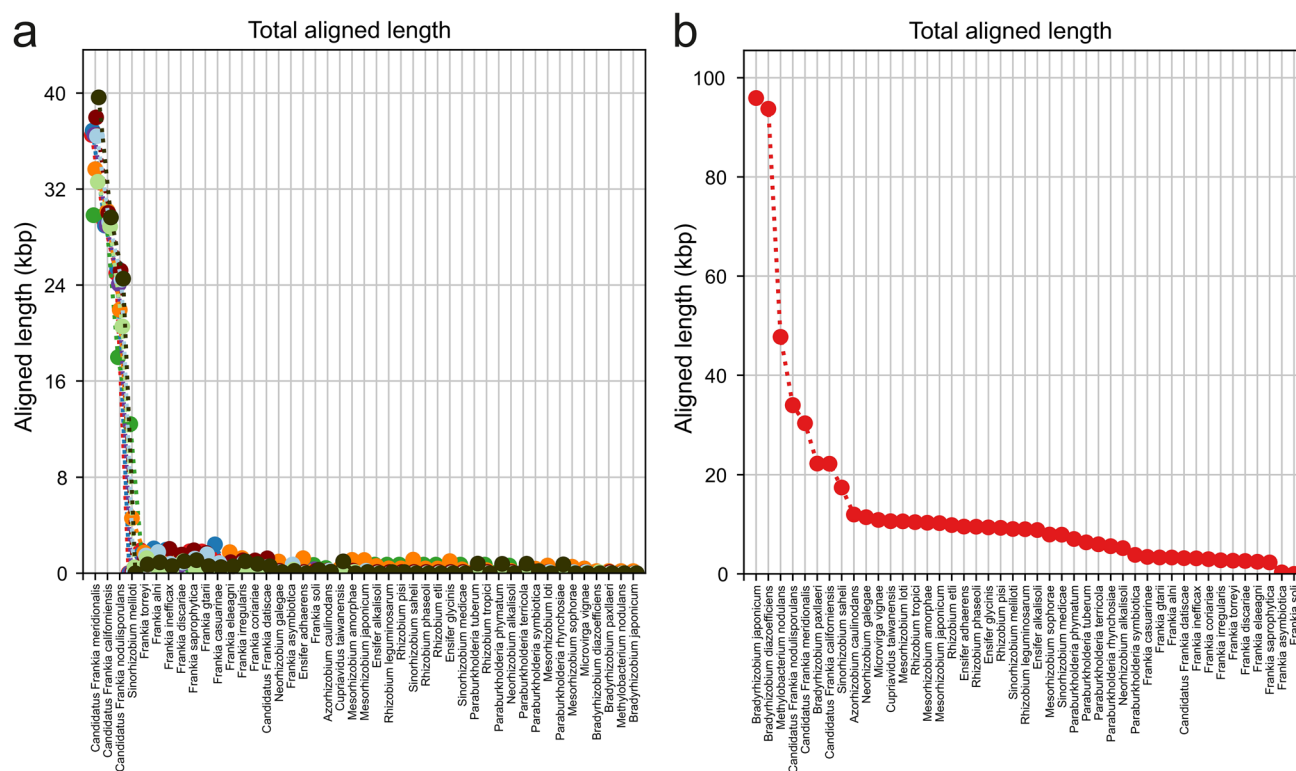


Fig. 5 Alignment of metagenomic assemblies of *Tribulus terrestris* (a) and *Roystonea regia* (b) to 45 well-known nodulating bacteria genomes using QUAST v5.2.0 (Mikheenko et al., 2018)

regia metagenomes, *nodB* genes affiliated with Oxalobacteraceae (Betaproteobacteria) and Bacilli (Firmicutes) were identified (Table S5). However, *nodA* and *nodC* genes were not detected in these metagenomes. We discovered *nifH* genes attributed to Bradyrhizobiaceae and Sphingomonadales. Several *nifD* genes were detected, belonging to Bradyrhizobiaceae, Rhizobiaceae, Nostocales, and Gammaproteobacteria. The detected *nifK* genes were associated with Bradyrhizobiaceae and Methanomicrobia (Table S5).

Amplification of *nod/nif* genes

We've adopted PCR as an alternative approach to identify *nod* and *nif* genes within the DNA extracted from root nodules. Efforts were made to amplify structural nodulating and N-fixing genes using a variety of primer pairs listed in (Table 1). From the nodules of *T. terrestris*, structural *nodABC* and *nifHDK* genes were successfully amplified and sequenced. Phylogenetic analyses showed that the amplified *nod* and *nif* genes were more closely related to *Sinorhizobium meliloti* (see Figs. 4, S5, and S6). Despite employing multiple general primer pairs, *nod* and *nif* genes could not be amplified from the nodules of *R. regia*.

Genome analysis

Phylogenomic analysis indicates that the dominant species isolated from surface-sterilized root nodules of *T. terrestris* (T2.Ur) and *R. regia* (R1.Fl) are classified as *Klebsiella* and *Kosakonia* (Gammaproteobacteria) respectively (Fig. 6). The R1.Fl strain isolated from *R. regia* indicates a high similarity to *Kosakonia sacchari* (GCA_001683395.1). The other less frequent isolates, T1.Ur (*T. terrestris*), R2.Fl (*R. regia*), R3.Fl (*R. regia*), MO2 (*R. regia*), and MO3 (*R. regia*), were classified as *Pseudomonas*, *Pseudoxanthomonas*, *Pseudomonas*, *Priestia*, and *Bacillus* respectively. The R1.Fl strain showed 96 and 95 percent ANI to *Kosakonia* sp. AG348 and *K. sacchari*, respectively (Fig. 7). Genes involved in the nod factor backbone structure synthesis were not detected in genomic annotations of any isolates and there was no evidence that any of the isolated strains were common nodulating bacteria. However, R1.Fl isolated from *R. regia* was identified as *K. sacchari* which is an N-fixing species encoding essential *nifHDK* genes (Fig. S4).

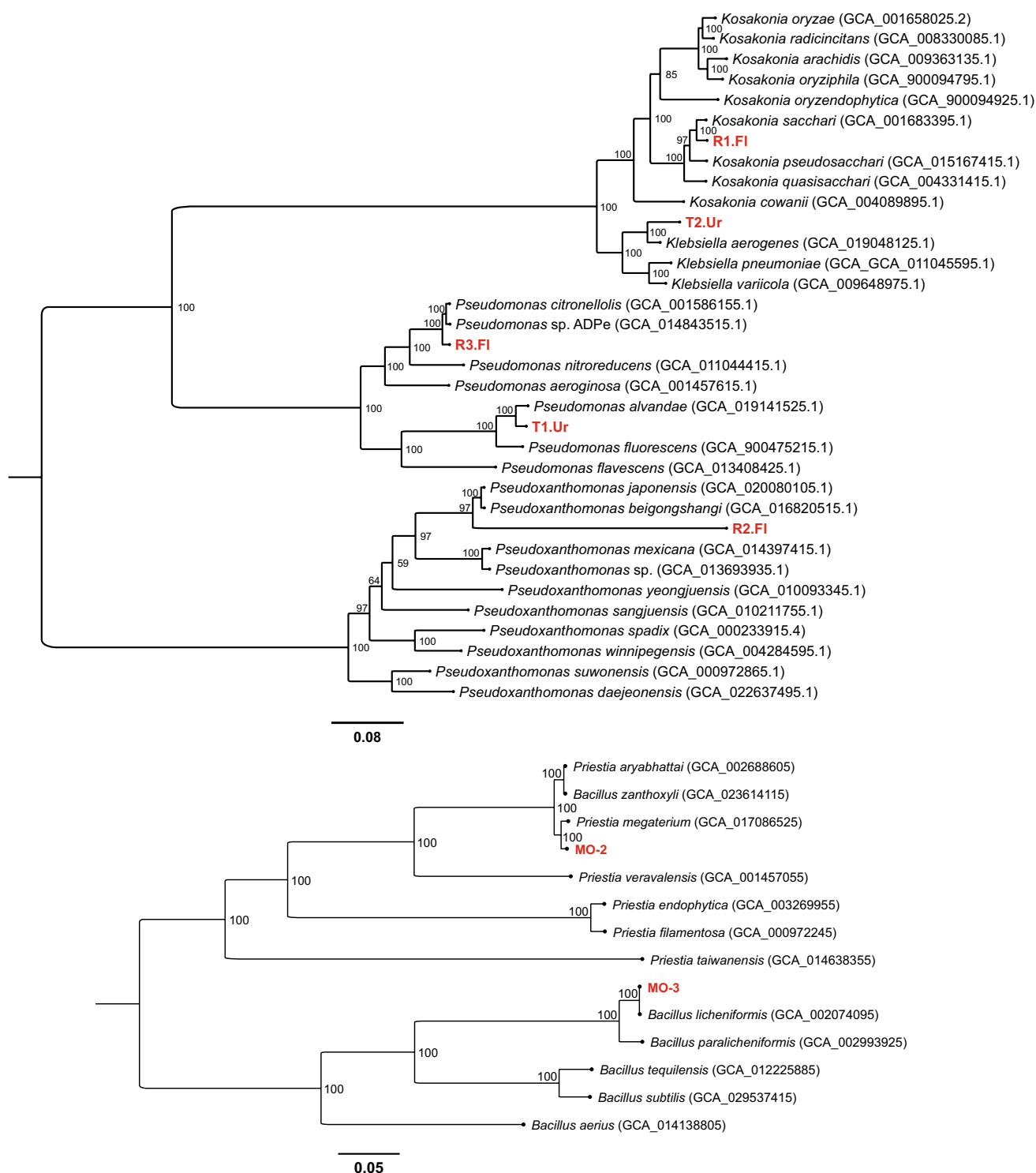


Fig. 6 Phylogenomic analysis of bacterial strains isolated from *T. terrestris* (T1.Ur and T2.Ur) and *R. regia* (R1.FI, R2.FI, R3.FI, MO-2, and MO-3) root nodules

Discussion

It is generally accepted that symbiotic root nodule formation is confined to legumes and actinorhizal plants (Soltis et al.

1995; Sprent 2009). A long time ago it was claimed that some members of the dicotyledonous family Zygophyllaceae are also capable of forming hypertrophies on their root system similar to rhizobial nodules. Nodule-like structures had

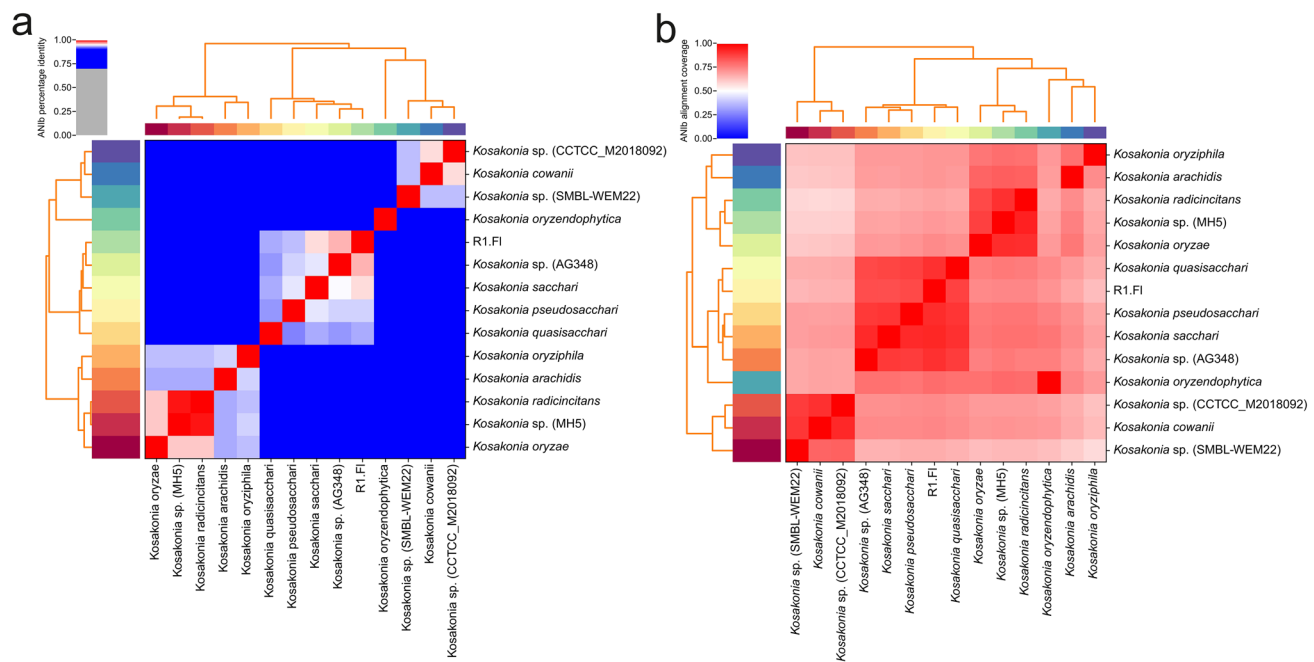


Fig. 7 Average nucleotide identity (ANI) of *Kosakonia* species. Percentage identity (**a**) and alignment coverage (**b**)

been observed on *Tribulus* species and nodulation was attributed to either fungi (Issatschenko 1913), rhizobia (Sabet 1946; Mostafa and Mahmoud 1951; Athar and Mahmood 1972; Basu et al. 1997; Athar and Mamood 1981), or a cyanobacteria '*Newmania karachiensis*' (Mahmood and Athar 1998, 2006). However, the name '*Newmania*' proposed by Mahmood and Athar (1998) has not been officially recognized as a bacterial genus. On the other hand, Allen and Allen (1950), Sahibzada and Khan (1973), and Gresshoff et al. (1984) were unable to detect any endophytes in the nodular growths on *Tribulus* roots and concluded that nodulation of *Tribulus* roots was associated with the plant's starch content. However, cross inoculation experiment performed by Mostafa and Mahmoud (1951) indicated suspension of bacterial isolates from nodules of *Tribulus alatus* could induce nodulation on *Trifolium alexandrinum* and *Arachis hypogaea* roots. Additionally, in 1997, root nodulation was reported in a monocot tree *R. regia* for the first time, and evidence of N fixation was provided by the presence of nitrogenase and leghaemoglobin, a compound that allows nitrogenase to function by reducing the oxygen concentration in the root nodule (Basu et al. 1997). *Rhizobium* was thought to be responsible for nodulation in *R. regia* (Basu et al. 1997). The nodules of *R. regia* were also found to be producing indole acetic acid, which is an important plant hormone (Basu and Ghosh 1998, 2001). Nevertheless, further research has never been undertaken to investigate the bacterial inhabitants of *T. terrestris* and *R. regia* nodules, and the nature of nodulation in Zygothryaceae and Arecaceae has remained unknown.

We observed nodule-like structures in root samples of *T. terrestris* (Zygothryaceae), and *R. regia* (Arecaceae) collected from West Asia and the Caribbean, respectively. The nodule-like structures with an apical meristem and a central vascular system were observed on both *T. terrestris* and *R. regia* root samples (Figs. 1, 2, S1, and S2). We found no evidence of cyanobacterial infection in *T. terrestris* nodules in contrast to Mahmood and Athar (1998, 2006) who reported heavy colonization of nodules by '*Newmania karachiensis*' (Figs. 3, 4). However, we found Actinobacteria are the most abundant taxa colonizing *T. terrestris* nodule-like structures.

Actinobacteria have superior tolerance to desiccation and solute stress among bacteria and they have been isolated from diverse, hostile environments such as arid and hyper-arid deserts (Neilson et al. 2012; Stevenson and Halls-worth 2014) similar to the environments that *T. terrestris* samples were collected. We identified *nodA* genes from Pseudonocardiales and Micrococcaceae, *nodB* gene from Xanthomonadales, and several *nodC* genes not attributed to any specific taxa (Table S2). A comparison of the metagenomic assemblies to the genomes of well-known nodulating bacteria indicates the highest alignment with Candidatus *Frankia meridionalis*, however, the alignment length is short (ca. 40 kb; Fig. 5a). In addition, we observed a small number of *Frankia* SSU sequences within the metagenome assemblies of *T. terrestris* and we were unable to retrieve a genome of *Frankia* from our metagenomic dataset. With the application of universal primer pairs, we successfully amplified the *nodABC* and *nifHDK* genes attributed to *S.*

meliloti from nodular tissues of *T. terrestris* collected in Iran (refer to Figs. S5, S6). However, the abundance of *S. meliloti* in our metagenomic data was found to be very low, less than 0.2%. Furthermore, despite being recognized as a fast-growing species, *S. meliloti* could not be isolated from surface-sterilized nodules.

Our metagenomic analysis indicates that Alphaproteobacteria are the dominant colonizers of nodule-like structures in *R. regia* (Fig. 3). We found that *Bradyrhizobium japonicum* was most closely aligned to the metagenome assembly when compared to genomes of well-known nodulating bacteria (Fig. 5b). Additionally, we identified several SSU sequences from *Bradyrhizobium* species in our metagenome assembly. We recovered a nearly complete *Bradyrhizobium* genome (96.05%) from metagenomic data but could not detect the essential *nod* genes in its genome (Fig. S3). In metagenome sequences we identified *nodB* genes affiliated with Oxalobacteraceae (Betaproteobacteria) and Bacilli (Firmicutes). However, *nodA* and *nodC* genes were not detected. We found *nifH* genes attributed to Bradyrhizobiaceae and Sphingomonadales. Several *nifD* genes were detected, belonging to Bradyrhizobiaceae, Rhizobiaceae, Nostocales, and Gammaproteobacteria (Table S3). However, *Bradyrhizobium* could not be isolated after multiple attempts of cultivating surface sterilized nodules probably due to the long shipping time resulted in bacterial cells losing their viability. Our cultivation attempts resulted in consistent isolation of *K. sacchari* (Gammaproteobacteria) in the YMA medium. *Kosakonia sacchari* is a fast-growing species and is known to be able to colonize roots and fix nitrogen in association with sugarcane plants (Chen et al. 2014). We found the *nifHDK* genes, encoding for the α and β subunits of nitrogenase in the genome of *K. sacchari* isolated from *R. regia* nodules (Fig. S4). It's noteworthy that the morphological traits of the bacteria obtained from *R. regia* nodules in the Basu et al. (1997) initially labeled as *Rhizobium*, align with our *K. sacchari* isolate. While *K. sacchari* may not have induced the development of root nodules in *R. regia*, it does reside within the root nodules as an endophyte. Conducting experiments like the acetylene reduction assay is crucial for assessing N fixation levels in the rhizosphere.

Several hypotheses can potentially elucidate the presence of nodule-like structures in the root systems of these two plant species, with the available information from metagenomics, culturing, and microscopy. First, in the absence of rhizobia, there are plant mutations that can lead to spontaneous root nodule formation (Tirichine et al. 2006, 2007; Gleason et al. 2006). It's plausible that specific varieties of *T. terrestris* underwent a natural mutation, resulting in the formation of nodule-like structures. The likelihood of this hypothesis seems minimal, given that evidence of nodulating taxa is present in the nodules of both plant species and canonical *nodABC* genes are present in *T. terrestris*

metagenomes. However, these taxa do not appear to produce the symbiosome or fixation thread structures. This discrepancy suggests that while nodulation occurs, the formation of symbiosomes or fixation threads is not an inherent trait of these nodulating taxa. Second, small non-N-fixing protuberances often designated as mycorrhizal nodules have been known for more than one century to be formed on the roots of some Gymnosperms (Bond 1963). Members of the Podocarpaceae, Araucariaceae, and Phyllocladaceae families have been reported to form such nodules (McLuckie 1923, Morrison and English 1967). We observed hyphal structures inside the nodular cell in the case of *R. regia*, remaining cell walls of (dead) fungal hyphae remotely resembling mycorrhizal structures. While we only identified the ITS sequences of *Leptosphaerulina*, *Cladosporium*, and *Nigrospora* in the metagenomic assemblies, fungi not typically classified as mycorrhizal which contradicts the hypothesis proposing mycorrhizal involvement in nodulation. Third, pseudonodules devoid of endophytes can be induced by chemical compounds, including metabolites from plants, animals, or microorganisms. There's a potential scenario that nodulation may be triggered by chemicals secreted by bacteria residing on the surfaces of root nodules. For instance, in the case of *T. terrestris*, *S. meliloti* may release such chemicals, even though the bacteria are incapable of penetrating the nodules and forming a symbiotic relationship. This hypothesis implies that the presence of these bacterial chemicals could stimulate the formation of nodules, potentially mimicking the signaling pathways typically initiated by symbiotic bacteria. However, despite their inability to establish direct contact with the plant's root system, these surface-dwelling bacteria may indirectly influence nodulation processes through chemical signaling mechanisms. Last but not least, taking into account the identification of various N-fixing bacteria within nodules and the successful recovery of essential nod factor synthesizing genes, we can deduce that specific diazotrophic bacteria play a crucial role in initiating the formation of root nodules. These bacteria probably colonize the intercellular spaces within the nodules without forming conventional symbiosomes or fixation thread structures and contribute fixed nitrogen to the host plant. This conclusion corroborates earlier studies that proposed the involvement of diazotrophic bacteria in nodulation processes. Notably, previous research indicated that *T. terrestris* plants harboring nodules exhibited robust and vigorous growth, in contrast to the stunted and unhealthy growth observed in plants lacking nodules. In *T. terrestris*, our investigation uncovered canonical nod genes associated with Actinobacteria, implying the possible participation of bacterial genera apart from *Frankia* in nodulation mechanisms. This finding contradicts the belief that *Frankia* is the only genus mediates nodulation in Actinobacteria. The discovery of *nod* genes in diverse Actinobacteria, extending beyond the traditional nodulation

candidates like *Frankia*, sheds light on the potential diversification of nodulation mechanisms within this bacterial phylum. These findings imply that the ability to form nodules may have evolved among a wider array of organisms within the Actinobacteria, suggesting a more intricate and varied landscape of nodulation across this bacterial lineage. Further investigations into the specific mechanisms underlying the interactions between diazotrophic bacteria and host plants are warranted to enhance our understanding of these symbiotic processes and their implications for plant growth and productivity.

Despite significant advancements in research, the evolutionary trajectory of nodulation remains an enigma lacking a unified explanation. While some theories posit independent origins of nodulation within the N-fixing clade, others suggest a single emergence in the common ancestor of Rosids (Doyle 1998, 2011, 2016; Werner et al. 2014; Li et al. 2015; Martin et al. 2017). Kates et al. (2024) conducted an extensive phylogenetic analysis, revealing multiple independent instances of nodulation with numerous gains and losses throughout evolutionary history. Conversely, recent studies propose an alternate hypothesis, suggesting nodulation originated at the root of the N-fixing clade, followed by widespread independent losses, supported by the absence of key symbiotic regulators (Libourel et al. 2023, Zhang et al. 2023). Investigating these non-traditional plant species and their nodulation mechanisms could offer valuable insights into the evolution of root nodule symbiosis. Future research endeavors focusing on genome and transcriptome analyses of these species, as well as exploring nodulation potential in other members of plant families Zygophyllaceae and Arecaceae, are vital for deciphering the underlying mechanisms of nodulation. These plant species are thought to represent early and rudimentary forms of nodulation that predate the emergence of the N-fixing clade. Their nodulation processes appear to be ancient, rudimentary, and not fully developed, suggesting a stage preceding the evolution of the N-fixing clade in evolutionary history.

Conclusion

The formation of nodules on the roots of *T. terrestris* and *R. regia* raises intriguing questions about the underlying nodulation mechanisms. Unlike legumes and actinorhizal plants, the nodules in these species lack infected cells but are nonetheless associated with N-fixing bacteria suggesting a distinct symbiotic relationship despite this absence. Through sequencing efforts, we have uncovered essential nodulation and N-fixation genes within the microorganisms inhabiting these nodules. Our findings point to the pivotal role of specific diazotrophic bacteria in initiating root nodules, which persist within intercellular spaces, thereby providing fixed

nitrogen to the plant. This observation corroborates previous research demonstrating nitrogenase activity within these nodules. Further exploration of these two unconventional plant species holds great promise for shedding light on the intricate evolutionary dynamics governing the symbiotic interactions between plants and N-fixing bacteria.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00438-024-02218-2>.

Author contributions SR and LT designed the study, SR performed the lab experiments and data analyses, MSHM contributed to sampling and lab experiments, RG and EK contributed to sampling, MB contributed to data analysis, and SR, MB, TYJ, and LT wrote the manuscript with input from the other authors.

Funding This research was funded by the Estonian Research Council grant numbers 1399PUT and 635PRG.

Data availability Genomes have been deposited in GenBank under the Bioproject number PRJNA818535. Metagenomic data are available in the GenBank SRA database under the Bioproject number PRJNA1012500. All supplementary materials are available online with this article.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

- Allen EK, Allen ON (1950) The anatomy of the nodular growths on the roots of *Tribulus cistoides* L. *Soil Sci Soc Am J* 14(2):179–183
- Athar M, Mahmood A (1972) Root nodules in some members of Zygophyllaceae growing at Karafhi University Campus. *Pak J Bot* 4:209–210
- Athar M, Mamood A (1981) Extension of Rhizobium host range to Zygophyllaceae. In: Gibson AH, Newton WE (eds) *Current perspectives in nitrogen fixation*. Australian Academy of Science, Canberra, p 481
- Abbas MW, Hussain M, Akhtar S, Ismail T, Qamar M, Shafiq Z, Esatbeyoglu T (2022) Bioactive compounds, antioxidant, anti-inflammatory, anti-cancer, and toxicity assessment of *tribulus terrestris*—In vitro and in vivo studies. *Antioxidants* 11(6):1160
- Alneberg J, Bjarnason BS, De Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C (2014) Binning metagenomic contigs by coverage and composition. *Nat Methods* 11(11):1144–1146
- Asmussen CB, Dransfield J, Deickmann V, Barfod AS, Pintaud JC, Baker WJ (2006) A new subfamily classification of the palm family (Arecaceae): evidence from plastid DNA phylogeny. *Bot J Linn Soc* 151(1):15–38
- Bond G (1963) The root nodules of non-leguminous angiosperms. In: Nutman PS, Mosse B (eds) *Symbiotic Associations*. Cambridge University Press, UK, pp 72–91
- Basu PS, Ghosh AC (1998) News & notes indole acetic acid and its metabolism in root nodules of a monocotyledonous tree *roystonea regia*. *Curr Microbiol* 37:137–140
- Basu PS, Ghosh AC (2001) Production of indole acetic acid in culture by a *Rhizobium* species from the root nodules of a monocotyledonous tree. *Roystonea Regia Acta Biotechnol* 21(1):65–72

- Basu PS, Ghosh AC, Dangar TK (1997) Roystonea regia a monocotyledonous tree, bears rhizobial root nodules. *Folia Microbiol* 42:601–606
- Bengtsson-Palme J, Hartmann M, Eriksson KM, Pal C, Thorell K, Larsson DGJ, Nilsson RH (2015) METAXA2: improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. *Mol Ecol Resour* 15(6):1403–1414
- Bengtsson-Palme J, Ryberg M, Hartmann M, Branco S, Wang Z, Godhe A, De Wit P, Sánchez-García M, Ebersberger I, de Sousa F, Amend A (2013) Improved software detection and extraction of ITS1 and ITS 2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods Ecol Evol* 4(10):914–919
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120
- Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J (2021) eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol* 38(12):5825–5829
- Chen M, Zhu B, Lin L, Yang L, Li Y, An Q (2014) Complete genome sequence of *Kosakonia sacchari* type strain SP1 T. *Stand Genom Sci* 9:1311–1318
- Coba de la Pena T, Fedorova E, Pueyo JJ, Lucas MM (2018) The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organelle? *Front Plant Sci* 8:305846
- Doyle JJ (1998) Phylogenetic perspectives on nodulation: evolving views of plants and symbiotic bacteria. *Trends Plant Sci* 3(12):473–478
- Doyle JJ (2011) Phylogenetic perspectives on the origins of nodulation. *Mol Plant Microbe Interact* 24(11):1289–1295
- Doyle JJ (2016) Chasing unicorns: nodulation origins and the paradox of novelty. *Am J Bot* 103(11):1865–1868
- Ebmeyer S, Coertze RD, Berglund F, Kristiansson E, Larsson DJ (2022) GENView: a gene-centric, phylogeny-based comparative genomics pipeline for bacterial genomes and plasmids. *Bioinformatics* 38(6):1727–1728
- Eickhout B, Bouwman AV, Van Zeijts H (2006) The role of nitrogen in world food production and environmental sustainability. *Agr Ecosyst Environ* 116(1–2):4–14
- Gleason C, Chaudhuri S, Yang T, Munoz A, Poovaiah BW, Oldroyd GE (2006) Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. *Nature* 441(7097):1149–1152
- Gresshoff PM, Newton S, Mohapatra SS, Scott KF, Howitt S, Price GD, Bender GL, Shine J, Rolfe, BG (1984) Symbiotic nitrogen fixation involving *Rhizobium* and the non-legume *Parasponia*. In *Advances in Nitrogen Fixation Research: Proceedings of the 5th International Symposium on Nitrogen Fixation*, Noordwijkerhout, The Netherlands, August 28–September 3, 1983, Springer Netherlands, pp. 483–489
- Griesmann M, Chang Y, Liu X, Song Y, Haberer G, Crook MB, Billault-Penneteau B, Laursseergues D, Keller J, Imanishi L, Roswanjaya YP (2018) Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. *Science* 361(6398):p. eaat1743
- Gyaneshwar P, Hirsch AM, Moulin L, Chen WM, Elliott GN, Bontempes C, Estrada-de Los SP, Gross E, dos Reis FB, Sprent JI, Young JPW (2011) Legume-nodulating betaproteobacteria: diversity, host range, and future prospects. *Mol Plant Microbe Interact* 24(11):1276–1288
- Herridge DF, Peoples MB, Boddey RM (2008) Global inputs of biological nitrogen fixation in agricultural systems. *Plant Soil* 311:1–18
- Issatschenko BL (1913) Über die wurzelknöllchen bei *Tribulus terrestris* L. *bulletin jardine botanique*. St Petersburg 13:23–31
- Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z (2019) MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359
- Kates HR, O'Meara BC, LaFrance R, Stull GW, James EK, Liu SY, Tian Q, Yi TS, Conde D, Kirst M, Ané JM (2024) Shifts in evolutionary lability underlie independent gains and losses of root-nodule symbiosis in a single clade of plants. *Nat Commun* 15(1):4262
- Katoh K, Kuma KI, Toh H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* 33(2):511–518
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM (2017) Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27(5):722–736
- Laguerre G, Nour SM, Macheret V, Sanjuan J, Drouin P, Amarger N (2001) Classification of rhizobia based on nodC and nifH gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. *Microbiology* 147(4):981–993
- Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW (2016) MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102:3–11
- Li HL, Wang W, Mortimer PE, Li RQ, Li DZ, Hyde KD, Xu JC, Soltis DE, Chen ZD (2015) Large-scale phylogenetic analyses reveal multiple gains of actinorhizal nitrogen-fixing symbioses in angiosperms associated with climate change. *Sci Rep* 5(1):14023
- Libourel K, Keller J, Brichet L, Cazalé AC, Carrère S, Vernié T, Couzigou JM, Callot C, Dufau I, Cauet S, Marande W (2023) Comparative phylotranscriptomics reveals ancestral and derived root nodule symbiosis programmes. *Nat Plants* 9(7):1067–1080
- Mahmood A, Athar M (1998) Cyanobacterial root nodules in *Tribulus terrestris* L. (Zygophyllaceae). In *Nitrogen Fixation with Non-Legumes: Proceedings of the 7th International Symposium on Nitrogen Fixation with Non-Legumes*, held 16–21 October 1996 in Faisalabad, Pakistan, Dordrecht: Springer Netherlands, pp. 345–350
- Mahmood A, Athar M (2006) Scanning electron microscopic observations on micro-organisms in the root nodules of *Tribulus terrestris* L. (zygophyllaceae). *Scanning* 28(4):233–235
- Maddison WP, Maddison DR (2019) Mesquite: a modular system for evolutionary analysis, v. 3.61.
- Manni M, Berkeley MR, Seppel M, Simão FA, Zdobnov EM (2021) BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol* 38(10):4647–4654
- Martin FM, Uroz S, Barker DG (2017) Ancestral alliances: plant mutualistic symbioses with fungi and bacteria. *Science* 356(6340):4p. eaad4501
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 37(5):1530–1534
- Morrison TM, English DA (1967) The significance of mycorrhizal nodules of *Agathis australis*. *New Phytol* 66(2):245–250
- Mostafa MA, Mahmoud MZ (1951) Bacterial isolates from root nodules of Zygophyllaceae. *Nature* 167(4246):446–447
- Von Meijenfeldt FB, Arkhipova K, Cambuy DD, Coutinho FH, Dutilh BE (2019) Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. *Genome Biol* 20:1–14
- Moulin L, Debellé F, Giraud E, Mangin B, Dénarié J, Boivin-Masson C (2000) The nodA sequence of rhizobia gives clues on structural features of Nod factors. *Nitrogen Fixation: From Molecules to Crop Productivity*. Springer, Dordrecht, pp 204–204
- Moulin L, James EK, Klonowska A, de Faria SM, Simon MF (2015) Phylogeny, diversity, geographical distribution, and host range of legume-nodulating Betaproteobacteria: what is the role of plant

- taxonomy? In: de Bruijn FJ (ed) Biological Nitrogen Fixation. Wiley, Hoboken
- Na SI, Kim YO, Yoon SH, Ha SM, Baek I, Chun J (2018) UBCG: up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. *J Microbiol* 56:280–285
- Neilson JW, Quade J, Ortiz M, Nelson WM, Legatzki A, Tian F, LaComb M, Betancourt JL, Wing RA, Soderlund CA, Maier RM (2012) Life at the hyperarid margin: novel bacterial diversity in arid soils of the Atacama Desert, Chile. *Extremophiles* 16:553–566
- Op den Camp RH, Polone E, Fedorova E, Roelofs W, Squartini A, Op den Camp HJ, Bisseling T, Geurts R (2012) Nonlegume *Parasponia andersonii* deploys a broad rhizobium host range strategy resulting in largely variable symbiotic effectiveness. *Mol Plant Microbe Interact* 25(7):954–963
- Peters GA (1991) Azolla and other plant-cyanobacteria symbioses: Aspects of form and function. *Plant Soil* 137:25–36
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25(7):1043–1055
- Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK (2016) Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods* 8(1):12–24
- Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A (2020) Using SPAdes de novo assembler. *Curr Protoc Bioinformatics* 70(1):e102
- Rahimlou S, Bahram M, Tedersoo L (2021) Phylogenomics reveals the evolution of root nodulating alpha-and beta-Proteobacteria (rhizobia). *Microbiol Res* 250:126788
- Roncal J, Zona S, Lewis CE (2008) Molecular phylogenetic studies of Caribbean palms (Arecaceae) and their relationships to biogeography and conservation. *Bot Rev* 74:78–102
- Sabet YS (1946) Bacterial root nodules in the Zygophyllaceae. *Nature* 157(3994):656–657
- Sprent JI (2009) Legume Nodulation: A Global Perspective. Wiley, Hoboken
- Sahibzada MA, Khan AG (1973) Starch-filled nodular hypertrophies in *Tribulus terrestris*. *Pak J Bot* 5:87–88
- Sheahan MC, Chase MW (2000) Phylogenetic relationships within Zygophyllaceae based on DNA sequences of three plastid regions, with special emphasis on Zygophylloideae. *Syst Bot* 25(2):371–384
- Singh S, Gupta YK (2011) Aphrodisiac activity of *Tribulus terrestris* Linn. in experimental models in rats. *J Men's Health* 8(S1):S75–S77
- Stevenson A, Hallsworth JE (2014) Water and temperature relations of soil Actinobacteria. *Environ Microbiol Rep* 6(6):744–755
- Shiraishi A, Matsushita N, Hougetsu T (2010) Nodulation in black locust by the Gammaproteobacteria *Pseudomonas* sp. and the Betaproteobacteria *Burkholderia* sp. *Syst Appl Microbiol* 33(5):269–274
- Sieber CM, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF (2018) Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol* 3(7):836–843
- Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, Martin PG (1995) Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proc Natl Acad Sci* 92(7):2647–2651
- Tedersoo L, Laanisto L, Rahimlou S, Toussaint A, Hallikma T, Pärtel M (2018) Global database of plants with root-symbiotic nitrogen fixation: Nod DB. *J Veg Sci* 29(3):560–568
- Tirichine L, James EK, Sandal N, Stougaard J (2006) Spontaneous root-nodule formation in the model legume *Lotus japonicus*: a novel class of mutants nodulates in the absence of rhizobia. *Mol Plant Microbe Interact* 19(4):373–382
- Tirichine L, Sandal N, Madsen LH, Radutoiu S, Albrechtsen AS, Sato S, Asamizu E, Tabata S, Stougaard J (2007) A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. *Science* 315(5808):104–107
- Uritskiy GV, DiRuggiero J, Taylor J (2018) MetaWRAP—A flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 6:1–13
- van Velzen R, Doyle JJ, Geurts R (2019) A resurrected scenario: single gain and massive loss of nitrogen-fixing nodulation. *Trends Plant Sci* 24(1):49–57
- Wall LG (2000) The actinorhizal symbiosis. *J Plant Growth Regul* 19:167–182
- Werner GD, Cornwell WK, Sprent JI, Kattge J, Kiers ET (2014) A single evolutionary innovation drives the deep evolution of symbiotic N₂-fixation in angiosperms. *Nat Commun* 5(1):4087
- Wu YW, Simmons BA, Singer SW (2016) MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32(4):605–607
- Wu SD, Zhang LJ, Lin L, Yu SX, Chen ZD, Wang W (2018) Insights into the historical assembly of global dryland floras: the diversification of Zygophyllaceae. *BMC Evol Biol* 18:1–10
- Zhang XX, Turner SL, Guo XW, Yang HJ, Debellé F, Yang GP, Dénarié J, Young JPW, Li FD (2000) The common nodulation genes of *Astragalus sinicus* rhizobia are conserved despite chromosomal diversity. *Appl Environ Microbiol* 66(7):2988–2995
- Zhang Y, Fu Y, Xian W, Li X, Feng Y, Bu F, Shi Y, Chen S, van Velzen R, Battenberg K, Berry AM (2024) Comparative phylogenomics and phylotranscriptomics provide insights into the genetic complexity of nitrogen-fixing root-nodule symbiosis. *Plant Commun* 5(1):100671

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.