MASTERARBEIT | MASTER'S THESIS

Titel | Title

Characterizing Microbial Communities in Individual Root Nodules of Clover

verfasst von | submitted by

Carina Gratz BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 830
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Molecular Microbiology, Microbial Ecology and Immunobiology
Betreut von | Supervisor:	Univ.-Prof. Dr. Jillian Petersen

Table of contents

I. List of Figures

II. List of Tables

III. List of Abbreviations

16S rRNA	16S ribosomal ribonucleic acid
ASV	Amplicon sequence variant
DNA	Desoxyribonucleic acid
dPCR	Digital polymerase chain reaction
EtOH	Ethanol
NAB	Nodule associated bacteria
OTU	Operational taxonomic unit
PCR	Polymerase chain reaction
PERMANOVA	Permutational multivariate analysis of variance
pPNA clamp	Plastid peptide nucleic acid clamp

Acknowledgements

First, I would like to express my gratitude to Prof. Jillian M. Petersen, who welcomed me into her group and has always been supportive. Further, I want to thank Dr. Christina Straub, who was always here for me and assisted me throughout the whole master's thesis. She was always available for answers and assisted me whenever I needed help.

Further, I want to thank everybody in Prof. Petersen´s group for their kindness and support.

I want to thank my fellow master students of the symbiosis group, Carmen Christine Fontanes Eguiguren, Tobias Leberfinger, Philipp Schmelz and Stefan Eckensperger, as well as all other master students at DOME for always being there for me. We have become a great community, helping each other and giving support to whoever needs it.

Further, I want to thank the gardeners Andreas Schröfl and Thomas Joch for caring for my plants.

I also want to thank the DOME TA team, who were always willing to help with lab questions. Further, I want to thank the JMF team, who sequenced my DNA.

I want to express my gratitude to my family, who have always supported and cared for me. I especially want to thank my mom, who has devoted her life to making mine the best it can be. Lastly, I want to thank Tobias Schwarzinger, who always had a listening ear for me and is always incredibly supportive.

1. Abstract

Legumes form symbiotic relationships with nitrogen-fixing rhizobia in root nodules. This symbiosis positively influences plant growth and development. The nodules also host nodule associated bacteria (NAB), whose diversity and roles remain poorly understood. Further, the composition of individual nodules within single clover (*Trifolium* sp.) plants has not been investigated. I examined the composition of individual clover root nodules from plants collected in Vienna and from a potting experiment. Three different varieties of clover: red clover (*Trifolium pratense* var. *Global)*, red clover (*Trifolium pratense* var. *Reichersberger*), and crimson clover (*Trifolium incarnatum*) were sown in organic farmer's soil in May 2024. Plants were harvested at two timepoints (growth phase and flowering phase). Nodules were collected from each plant, surface sterilized, and the DNA was extracted. Amplicon sequencing of 16S rRNA and *nifH* genes was used to assess bacterial diversity and dPCR to quantify bacterial cells.

This thesis revealed *Rhizobium leguminosarum* as the main symbiont in *Trifolium* sp.. *Bacillales* were the most common NABs, but nodules were also colonized by *Pseudomonadales, Corynebacterales* and *Enterobacterales*. I could show a plant-specific preference for NAB infection, where 95% of the 19 nodules from two individual plants were occupied by *Enterobacterales* that were absent from all other samples. No significant differences in community composition were found between the clover varieties or sampling timepoints. In comparison, the city samples showed a different bacterial community, with higher NAB concentrations and *Bacillales* as the most common NAB. These findings highlight the importance of studying the diversity of individual nodule microbiomes.

2. Zusammenfassung

Leguminosen gehen eine symbiotische Beziehung mit stickstofffixierenden Rhizobien in speziellen Wurzelorganen, den Knöllchen, ein. Diese Symbiose ist für beide Partner vorteilhaft. Die Wurzelknöllchen werden auch von knöllchenassoziierten Bakterien (NAB) bewohnt, deren Zusammensetzung und Rolle noch unzureichend bekannt ist. Die bakterielle Zusammensetzung einzelner Knöllchen von einer Kleepflanze (*Trifolium* sp.) ist bisher nicht erforscht worden. In dieser Arbeit wurde diese Zusammensetzung in einzelnen Wurzelknöllchen von Kleepflanzen in Abhängigkeit von Sorte und Wachstumsphase untersucht.

Dazu wurden Kleepflanzen in der Stadt Wien gesammelt und drei verschiedene Kleesorten: Rotklee (*Trifolium pratense* var. *Global)*, Rotklee (*Trifolium pratense* var. *Reichersberger*) und Blutklee (*Trifolium incarnatum*) im Mai 2024 in Töpfen angebaut. Die Pflanzen wurden zu zwei verschiedenen Zeitpunkten geerntet (Wachstumsphase und Blütezeit). Von jeder Pflanze wurden einzelne Knöllchen geerntet, oberflächensterilisiert und nach DNA-Extraktion mittels Amplikon-Sequenzierung des 16S rRNA- und *nifH*-Gens die bakterielle Diversität bestimmt. Die Quantifizierung der Bakterien in den Knöllchen erfolgte mittels digitaler PCR.

Ich habe *Rhizobium leguminosarum* als primären Symbionten in Klee identifiziert. Neben *Bacillales* als häufigste NABs, wurden auch andere nicht-rhizobielle Mikroorganismen wie*, Pseudomonadales* und *Enterobacterales*, nachgewiesen. Ich konnte eine pflanzenselektive Bevorzugung der NAB-Infektionen zeigen. In 95 % aller Knöllchen von zwei Pflanzen wurden *Enterobacterales* nachgewiesen, in den übrigen Proben waren diese nicht nachweisbar. Ansonsten wurde kein signifikanter Unterschied in der bakteriellen Diversität zwischen Pflanzensorten und Zeitpunkten festgestellt. Die städtischen Proben wiesen eine andere Bakteriengemeinschaft auf als die Universitätsproben, mit höheren Konzentrationen von knöllchenassoziierten Bakterien und *Bacillales* als häufigste NAB. Diese Ergebnisse zeigen die Wichtigkeit, die Vielfalt des Mikrobioms in den einzelnen Knöllchen zu untersuchen.

3. Introduction

Climate change greatly impacts the planet's food security, as extreme weather events happen more often and more severely. With increasing changes, crops encounter biotic and abiotic stressors in the field. Mainly due to changes in temperature, precipitation patterns, droughts or floods, and pest and disease outbreaks, crop quantity and quality are being negatively affected worldwide (Munaweera et al., 2022). With a growing population and climate change, ensuring a sustainable food supply is a major challenge. Attempts to increase the adaptability, resilience, and productivity of major crops are taken to solve the problem (Munaweera et al., 2022). A practical solution is needed to provide sustainable, high-quality crops without expanding agricultural land.

One useful approach is to implement legume-inclusive agricultural systems, which can utilize intercrops to improve nitrogen availability in the soil. Additionally, legumes are important because they are nitrogen-fixing plants that serve as sustainable food and feed sources (Stagnari et al., 2017).

Leguminous plants belong to the order *Fabales* and consist of more than 19500 species, with a wide range of valuable crops, including high-protein sources like soy (*Glycine max*). (Kazmierczak et al., 2020). Other important legumes include beans, peanuts, lentils, lupins, lucerne, and clover.

Clover (*Trifolium* sp.) is an important food crop for livestock farming and is among the top three most widely cultivated legume crops in Austria with 42571 ha planted in 2024 (STATISTIK AUSTRIA, 2024). Clover is also used as a cost-saving way to improve or support the biodiversity of flower-visiting insects by providing an abundant source of nectar and pollen. Thus, the decline in pollinator activity, which affects crops, can be mitigated (Harris & Ratnieks, 2022).

Farmers use inorganic fertilizers to boost plant growth and crop yield by improving soil nitrogen. However, these fertilizers come with a significant energy cost and a long-term loss of macro- and micronutrients in the environment (Jones et al., 2013). Excess mineralized nitrogen harms the environment, as nitrate leaching and microbial processes release nitrous oxide, a potent greenhouse gas. Including legumes in crop systems sustains the agricultural productivity of subsequent crops planted in the same field by

adding nitrogen to the soil and reducing environmental impact by requiring less fertilizer. In addition, the use of *Trifolium* sp. in the field is a beneficial economic decision for farmers, as the cost of inorganic fertilizers increases (Harris & Ratnieks, 2022).

Legumes have a unique ability to host nitrogen-fixing bacteria in specialized root compartments, called nodules, by forming a symbiotic relationship (Larrainzar et al., 2020). Rhizobia, soil-dwelling bacteria, infect root cells to form nodules, where they fix nitrogen in exchange for plant-provided carbon (Etesami, 2022)

3.1.Nodulation process

Efficient nitrogen fixation in legumes is an intricate developmental process, which must be coordinated between the plant and the microorganisms (Oldroyd et al., 2011).

The plant and microorganisms interact via extracellular factors. The complex development of the nodule requires signal exchange during the different steps to allow for coordinated regulation of the process (Crespi & Gálvez, 2000). This interaction is coordinated via many different genes of the host plant and the microbial symbiont (Figure 1).

For infection, rhizobia attach to susceptible parts of the plant root hairs. Now, the plant recognizes the Nod factors secreted by the bacteria and induces root hair curling (Kazmierczak et al., 2020). Root invasion can occur via two mechanisms, (1) the "crack entry", which involves rhizobia infecting plants through natural wounds or (2) "intercellular invasion"**,** which occurs when rhizobia enter intact epidermal cells, typically at root-hair junctions or the interface between a root hair and a neighboring epidermal cell (Zhang et al., 2024). The "crack entry" process is the most prevalent; consequently, the following discussion will concentrate on this mechanism.

Legumes secrete flavonoids into the rhizosphere, which the rhizobia will recognize and respond to by secreting Nod factors, initiating nodulation (Mergaert et al., 1997). The flavonoid signal molecules are only recognized by compatible rhizobia, allowing for specificity in recognition between rhizobia and legume species. The Nod factors can induce various cellular responses

related to the nodule formation in the plant, such as deformation of the root hairs or division of the cortical cells (Crespi & Gálvez, 2000). Rhizobia are trapped within the curled root hair, where they multiply and establish microcolonies. This leads to a high concentration of Nod factors, allowing for host plasma membrane invagination within root-hair cells (Zhang et al., 2024). The microorganisms can now degrade the plant cell walls, causing the development of an infection thread around the bacteria (Crespi & Gálvez, 2000). During the nodule development, *nodulin* genes are activated in the plant. The *nodulin* genes are upregulated in the nodules in a spatiotemporal expression pattern (Dupont et al., 2012). The expression of *nodulin* genes allows for the proper development and function of the nodules. Many of the *nodulin* genes code for proteins associated with the cell walls. Therefore, they are important in wall modifications during infection and organogenesis (Crespi & Gálvez, 2000). Different specific interactions between plant and bacterial cells are possible through this stepwise regulation (Crespi & Gálvez, 2000).

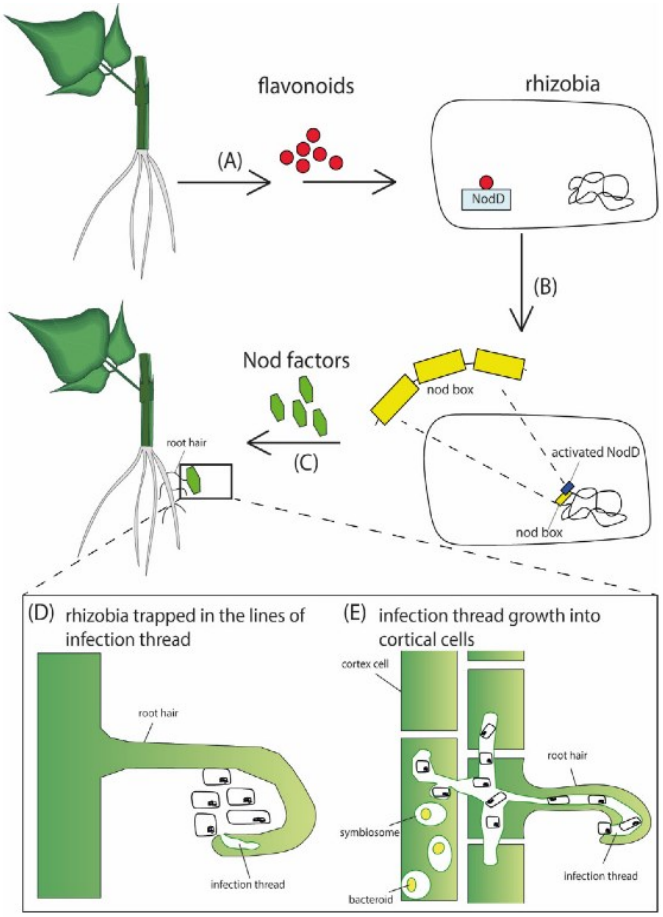

Figure 1: Mechanisms of root nodule development. The plant secretes flavonoids(A), which are detected by rhizobia (B). They, then release NOD factors (C), which allows the formations of an infection thread (D) and the infection of the plant cells by the bacteria (E) (Wekesa et al., 2022).

If the infection thread reaches the nodule primordium, bacteria can enter the cortical cells in an endocytosis-like process. Each bacterium is surrounded by a cell membrane derived from the host. Together with the encapsulated bacteria, the membrane forms an organelle-like structure, the symbiosomes (Ivanov et al., 2010). Each symbiosome consists of a symbiosome membrane, a symbiosome space, as well as the bacteria. Thousands of symbiosomes exist, within the infected cells (Zhang et al., 2024).

In the symbiosome the bacteria differentiate into bacteroids, which can fix nitrogen (Crespi & Gálvez, 2000). The mature nodules now house the bacteroids, which fix nitrogen, using the enzyme nitrogenase, until senescence.

3.2.Nitrogen fixation in the nodules

Rhizobia express the nitrogenase enzyme, which is encoded by the *nif* gene. The enzyme catalyzes the reduction and, thus, fixation of gaseous nitrogen (N_2) to ammonia (NH_3) (Goswami et al., 2021; Markl et al., 2019) (Figure 2). Nitrogenase is a complex metalloenzyme that can break the strong triple bond of N_2. It contains complex iron-sulfur cofactors, which can be inhibited by the presence of oxygen (Mus et al., 2019).

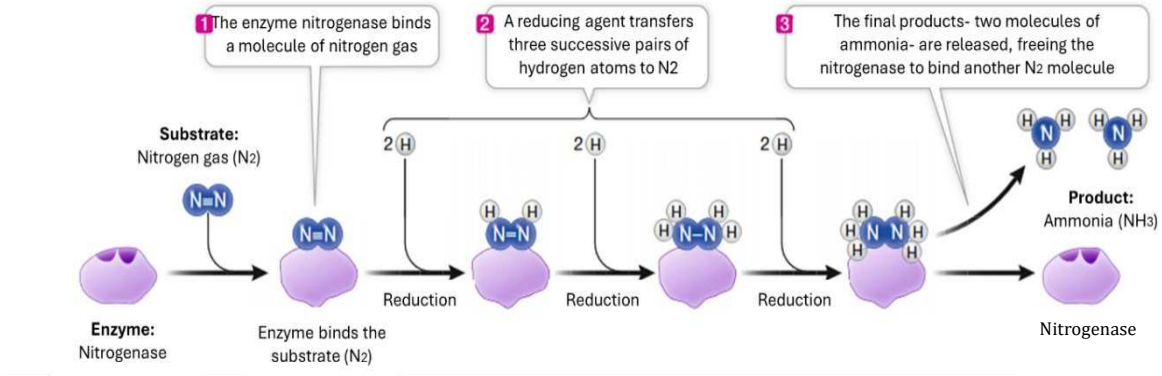

Figure 2: Reaction of Nitrogenase. The nitrogenase binds nitrous gas and reduces it to ammonia. Translated from Markl et al. (2019).

The inhibition of nitrogenase is a key driver in the establishment of symbiotic relationships between nitrogen-fixing bacteria and various plant species. The bacteria colonize plant roots to escape oxygen exposure. In legume root nodules, the plant produces an oxygen-free environment, via expression of leghemoglobin, which binds and removes the free oxygen (Larrainzar et al., 2020). Nodules have a limited life span, and with increasing age, the senescence

zone grows, turning the nodule from red to green as leghemoglobin is lost, consequently hindering nitrogen fixation (Kazmierczak et al., 2020). After about 10-12 weeks the nodules become fully senescent, and the plant cells and some of the bacteria die (Navascués et al., 2012). Although the expression of nitrogenase is a defining characteristic of the rhizobia and explains their presence in the nodules of legumes, they are not the only bacterial inhabitants of the symbiosomes.

3.3.Nodule occupants

3.3.1. Symbiotic partners

The root nodule symbiosis involves different rhizobial partners, depending on the plant type and location. Rhizobia are gram-negative rod-shaped bacteria that belong to the alphaproteobacteria (Zakhia & de Lajudie, 2001). These bacteria are naturally found in the soil, where they persist as free-living organisms until they encounter a suitable legume host. Depending on the legume species, different rhizobial species are involved in this symbiosis. For example *Rhizobium leguminosarum* primarily forms a symbiosis with clover (*Trifolium* sp.*)*, lentils (*Lens)* or the common bean *(Phaseolus vulgaris)*, whereas *Sinorhizobium meliloti* undergoes symbiosis with burclover (*Medicago* sp.*)* or sweet clover *(Melilotus* sp.) (Zakhia & de Lajudie, 2001)*. Bradyrhizobium japonicum*, on the other hand, is a known symbiont of soybean (*Glycine max*) (Mayhood & Mirza, 2021).

The ability of rhizobia to colonize different host plants is due to their distinct Nod factors, which determine bacterial acceptance and host-symbiont specificity (Zgadzaj et al., 2015). The selection of rhizobial endophytes by the plant is well-studied and has been reported to be influenced by many abiotic factors, such as nutrient availability, pH and water (Herridge et al., 1984; Mason et al., 2018; Sharaf et al., 2019). Different rhizobial species might perform better depending on soil conditions, such as pH, soil type, moisture or nutrient content. Thus, different species might be present in different field conditions (Mendoza-Suárez et al., 2020).

Symbiotic associations are not restricted to a single rhizobial species for each plant. In fact, individual nodules can harbor two or more distinct rhizobial species or strains of rhizobia. This variation can be observed between different nodules of a single plant (Figure 3) (De Meyer et al., 2011).

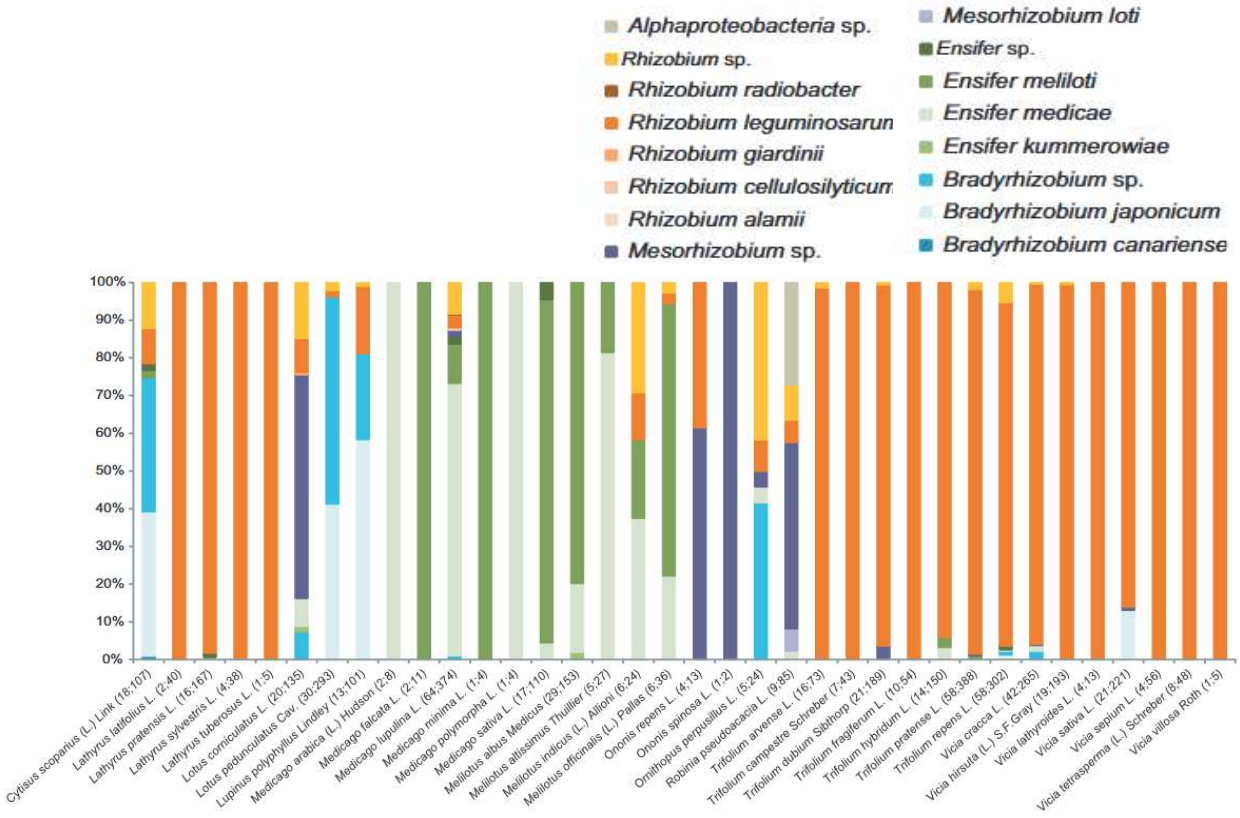

Figure 3: "Stacked column chart showing the host plant species with their rhizobial endosymbionts. Plant names are listed followed by in brackets the number of plants sampled and the number of isolates" (De Meyer et al., 2011). Each legume species forms a symbiotic relationship with a view specific rhizobial species.

For example, in pea, nodules with mixed rhizobial populations represented about 20.6% of the population (Mendoza-Suárez et al., 2020).

Additionally, ineffective rhizobia with no or reduced nitrogen-fixing abilities can be found in some root nodules. Less effective strains benefit from "hitchhiking" with more effective strains or by forming a cooperative partnership (Mendoza-Suárez et al., 2020). These rhizobia cannot form nodules on legumes, as they lack the key loci for nodulation like *nifD,* the nodD-A spacer or *nolL* (Kelsey et al., 2016; Sachs et al., 2010). Instead, they can coinfect the nodules, when identical but nodulating strains are present. Interstrain competition between different nodulating and nonnodulating rhizobial species can have a negative fitness effect on the plant, which is not completely understood yet (Kelsey et al., 2016).

3.3.2. Nodule associated bacteria (NAB)

While rhizobia play the most important role in nodule development and nitrogen fixation, they are not the only inhabitants in the nodules. A range of studies from as early as 1888, described the presence of bacteria other than rhizobia in legume root nodules (Beijerinck, 1888; Martínez-Hidalgo & Hirsch, 2017) These nodule associated bacteria (NAB), which can be nitrogen-fixing or non-nitrogen-fixing, are highly diverse and plentiful in the root nodules. Some of these bacteria also possess *nif* and *nod* genes and trigger nodule formation (Martínez-Hidalgo & Hirsch, 2017). Previous studies found 1-40% of the bacterial community within root nodules to be NABs (Mayhood & Mirza, 2021; Sharaf et al., 2019). The presence of different rhizobial and non-rhizobial species can be partly dependent on cultivar and water status (Sharaf et al., 2019).

The presence of NAB in root nodules has been shown to have positive effects on the plant, such as increased plant growth, resistance to environmental stressors, or protection against plant pathogens (Etesami, 2022). For example, in soybean, non-nitrogen-fixing bacteria, like *Pseudomonas* and *Enterobacteria*, were found (Sharaf et al., 2019). Nodule symbiosomes can be occupied by a variety of bacterial species, for example, *Bradyrhizobiaceae*, *Pseudomonadaceae* and *Enterobacteriaceae* were found to be the dominant species in soy root nodules (Sharaf et al., 2019).

Pseudomonas, *Enterobacter* and *Burkholderia* were found to improve the shoot nitrogen in plants, as well as biomass accumulation when co-inoculated (Subramanian et al., 2015). The most abundant NAB genus previously identified in root nodules was *Bacillus* (Xu et al., 2014). *Bacillus* was shown to have a positive effect on plant growth by suppressing plant pathogens and expressing antifungal compounds (Dhole & Shelat, 2022).

There is still a lack of understanding of the differences in bacterial diversity of individual root nodules within plants. Overlooking differences in nodule communities of the same plant means missing the possibility that each nodule may harbor a distinct microbial composition, with variations in the types and abundance of both rhizobia and NABs. However, these differences might influence a plant's survival, stress response and growth.

A previous study by Mayhood and Mirza (2021) investigated bacterial nodule communities from nine soybean plants. Through this investigation, they identified *Bradyrhizobia* within most nodules. Alongside the *Bradyrhizobia*, they identified rhizobial symbionts belonging to the genera *Rhizobum*, *Neorhizobium*, and *Mesorhizobium*. Additionally, the study identified the non-rhizobial genera *Nitrobacter* and *Tardiphaga* within almost all root nodules. Furthermore, they identified various NABs like, *Bacillus*, *Pseudomonas*, *Flavobacterium*, and *Variovorax* sp. at lower abundances. The researchers thought that the presence of these organisms was not influenced by the location of the nodule. Rather, they postulated that it was due to a random process (Mayhood & Mirza, 2021).

Both rhizobia and non-rhizobia bacteria are crucial for growth and survival of plants. Clover is a frequently used crop in agriculture worldwide but remains relatively understudied. Previous analyses of the rhizobial community in clover have primarily relied on pooled nodules or bacterial isolates cultured on agar plates (Schulz et al., 2013; Stefan et al., 2018; Zhang et al., 2016). Similarly, NAB community in clover has also been previously studied primarily using isolated bacterial cultures. These isolation-based approaches revealed the presence of various NABs in *Trifolium pratense*, including *Bacillus*, members of the order *Pseudomonadales*, *Curtobacterium*, and *Phyllobacterium* (Sturz et al., 1997).

4. Aim of this study

In this master's thesis, I wanted to establish bacterial symbionts of the root nodules within individual clover plants. I investigated individual nodules of three clover types at two different time points during their developmental stage: red clover (*Trifolium pratense* var. *Global*), red clover (*Trifolium pratense* var. *Reichersberger*) and crimson clover (*Trifolium incarnatum*).

I aimed to determine whether all root nodules were occupied by the same bacterial symbiont, and if there were variations in the symbionts among the different clover varieties. Then, I wanted to understand whether the relative abundance of NAB varied across nodules within plants, between plants and between different clover varieties. Further, I wanted to understand the impact

of nodule age on the bacterial community composition by comparing two different timepoints.

Studying the variations between individual nodules within a plant provides insight into the preferential selection of bacterial endophytes within different root nodules of a single plant.

5. Materials and Methods

5.1.Cultivating *Trifolium* sp.

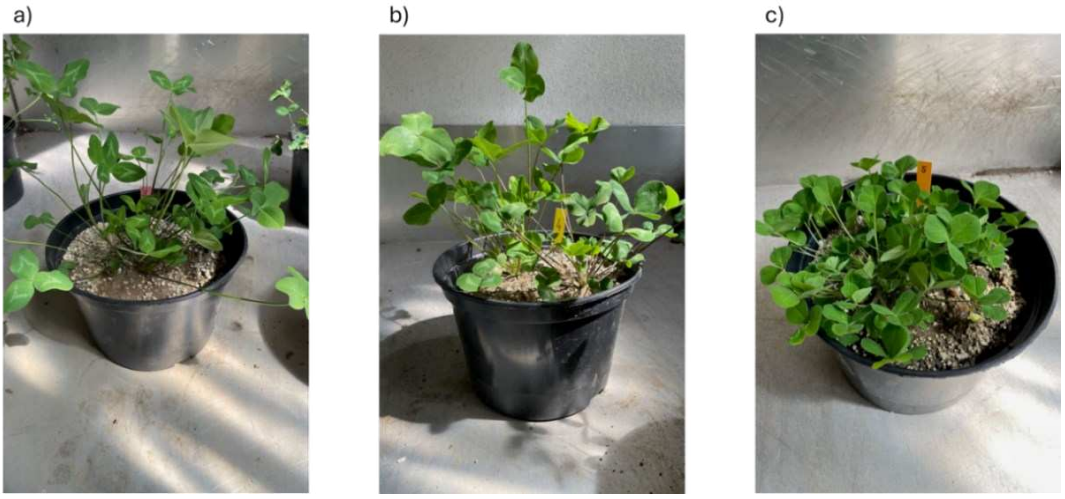

Figure 4: Potted Trifolium plants during growth phase. a) *Trifolium pratense* var. *Reichersberger* b) *Trifolium pratense* var. *Global* c) *Trifolium incarnatum*

Three clover varieties were grown in pots in the garden of the University of Vienna: *T. pratense* var. *Global*, *T. pratense* var. *Reichersberger* (red clover) and *T. incarnatum* (crimson clover). All varieties were seeded in soil collected from an organic agricultural field in Mank (Lower Austria) in early May 2024. The soil was sieved with a 5 mm sieve and mixed with perlite (2:1) to improve drainage and aerate the soil. For every *Trifolium* variety, seeds were surface sterilized for 1 min in 70% ethanol and sown in 15 pots on May 29th, 2024. After germination, three seedlings were kept, and once plants were established, the number of plants per pot was reduced to one.

5.2.Plant sampling

Growth of the plants and nodule size was monitored by manual inspection of surplus plants. This determined the first sampling point at 11 weeks after seeding (growth phase, August 12th, 2024) and the second sampling point at 15

weeks (flowering phase, September 9th, 2024), when nodules from five plants of each variety were harvested

Additionally, on July 9th, 2024, six flowering *T. pratense* plants were collected in six locations in Vienna: The collection began in front of outside Schönbrunn Palace, Auer von Welsbach Park, the garden of a residential complex- in the Linzer Straße (14th district) and two plants growing beside the Wien River (Table S 1).

5.3.Nodule harvest

Plants were shaken vigorously to remove excess soil. Nodules were cut off the root systems using a scalpel, ensuring 1-2 mm of root were left attached on either side of the nodule. Ten nodules were collected from across the root system per plant.

The nodules were then surface sterilized using the following protocol: 1 min in 70% ethanol, 3 min in 4% bleach, followed by three washes in autoclaved Milli-Q water. Then the nodules were weighed and crushed in an Eppendorf with 100 µL sterile Milli-Q water using a pestle.

5.4.Soil sampling

At the first sampling point, rhizosphere soil was collected by shaking off the soil attached to the root systems. For the second sampling point, the soil was too muddy, hence the root systems were washed in sterile Milli-Q water to remove excess soil. Then the rhizosphere was harvested by manually scraping the remaining soil off the roots using a spatula. Additionally, I sampled the rhizoplane by further washing the root systems in 50 ml of 0.9% NaCl for two minutes.

Additionally, for each clover variety, one soil sample per timepoint was sent to a collaborator at the University of Natural Resources and Life Sciences (BOKU, Vienna) to analyze basic soil parameters like pH, potassium (K), phosphate (P), copper (Cu) and iron (Fe).

5.5.DNA extraction

The DNA was extracted using the Qiagen DNeasy PowerSoil Pro Kit according to the manufacturer's instructions. Extraction was performed on single nodules for all samples, except for *T. incarnatum*, where five nodules were pooled together for each plant due to the small size. For soil samples, 0.2 g of rhizosphere soil were used for DNA extraction. The DNA was eluted in 70 µL and stored at -80°C until further use. One of the DNA extraction negative controls was contaminated. Sequencing results revealed rhizobia as the main contaminant in that control, hence this was attributed to human error during DNA extraction and no further decontamination of the dataset was performed.

5.6.Amplicon sequencing

Amplicon sequencing was performed using the marker genes 16S rRNA and *nifH*. The dinitrogenase reductase gene (*nifH*) was used as a marker for the diazotrophs, to be able to differentiate between different *Rhizobium* species present in the nodules and the soil.

Amplicon sequencing was performed on 269 samples, including 11 extraction controls, on the V4 region of 16S rRNA gene and the *nifH* gene (primers in Table 1). A pPNA clamp (Lundberg et al., 2013) was used to prevent sequencing the plant chloroplast. Sequencing services were provided in-house by the Joint Microbiome Facility at the Centre for Microbial and Environmental Science.

Sequencing of the nodules yielded an average sequencing depth of 5188 and 9380 for 16S rRNA and *nifH* respectively. Every amplicon sequence variant (ASV) that was present less than 20 times in all the samples was removed from both datasets. Rarefaction using the function, rarefy_even_depth in R, was performed to 2000 and 5000 reads for 16S rRNA and *nifH* respectively. This resulted in the removal of 10 samples and 774 taxa for the 16S rRNA dataset (Figure S 3) and 13 samples and 23 taxa for the *nifH* dataset (Figure S 5). Further quality checks included the removal of 32% of mitochondrial and 1.32% chloroplast reads for the 16S rRNA dataset.

Soil sequencing yielded an average sequencing depth of 7685 and 9635 for 16S rRNA and *nifH* respectively. Every ASV present less than 50 times, for 16S rRNA,

and less than 40 times, for *nifH* were filtered out. Rarefaction using the function, rarefy_even_depth in R, was performed to 4500 and 5000 reads for 16S rRNA and *nifH* respectively. This resulted in the removal of 2551 taxa for the 16S rRNA dataset (Figure S 7) and 4 samples and 3596 taxa for the *nifH* dataset (Figure S 9).

Table 1: Primers for 16S rRNA and *nifH* sequencing

PCR primers	**Target gene, size**	**Sequence (5'-3')**
515F Parada (Parada et al., 2016)	16S rRNA, ~253 bp	GTGYCAGCMGCCGCGGTAA
806R Apprill (Apprill et al., 2015)		GGACTACNVGGGTWTCTAAT
nifH-Ueda19F (Ueda et al., 1995)	*nifH*, ~420 bp	GCIWTYTAYGGIAARGGIGG
nifH-R6 (Ueda et al., 1995)		GCCATCATYTCICCIGA

5.7.Digital PCR (dPCR)

Digital PCR was performed to determine the total number of bacterial cells in the nodules. The same primers and chloroplast clamp were used, as for the 16S rRNA amplicon sequencing. All samples were diluted to a concentration ranging from 0.1 to 0.9 ng/µL for further use. Primer and reaction mixes were prepared (Table 2, Table 3) and the reactions were pipetted into 96-well nanoplates for use in the Qiagen QIAcuity dPCR machine with the thermal cycler settings listed in Table 4.

The dPCR results were then used to calculate the number of copies of 16S rRNA present in the samples. The absolute quantification calculations (Table 5) were performed on the previously weighed samples, excluding the samples from the city samples and T. incarnatum, from quantification.

Table 2: dPCR primers for 16S rRNA dPCR

PCR primers	**Target gene, size**	**Sequence (5'-3')**
515F Parada (Parada et al., 2016)	16S rRNA, ~253 bp	GTGYCAGCMGCCGCGGTAA
806R Apprill (Apprill et al., 2015)		GGACTACNVGGGTWTCTAAT

Table 3: dPCR master mix

Component	**Volume (µL) for 1 reaction**
3x EvaGreen PCR Master Mix	4
10x primer mix	1,2
Rnase-free water	4,8
Template (DNA /cDNA)	2
Total reaction volume	**12**

Table 4: Thermal cycler setting

Target gene	**16S rRNA**		
	Temp (°C)	**Time**	
Initial denaturing	95	3min	
Clamp annealing	78	10sec	40 cycles
Denaturing	95	30sec	
Annealing	52	40sec	
Elongation	72	30sec	
Final Elongation	40	5min	

Table 5: Absolute quantification calculations for dPCR

Value	Calculation
Expected value **λ**	$\lambda = -ln(\frac{nr\ of\ valid\ partitions\ -\ nr\ of\ positive\ partitions}{nr\ of\ valid\ partitions}$
Copies per µL in the PCR	$cp_{Vol} = \frac{\lambda}{Vol\ of\ partition\ [nl]} * 1000$
Copies per µL in the sample tube	$cp_{Vol} = \lambda * \frac{PCR\ reaction\ volume\ (= 12\ \mu L)}{template\ volume\ (= 2\ \mu L)}$
Copies in undiluted sample	$cp_{undiluted} = cp_{tube} * dilution\ factor$
Copies per g	$cp_{total} = cp_{undiluted} * Vol_{elution} * \frac{entity}{sample\ mass}$

5.8.Bioinformatics

The JMF processes raw reads using an in-house R pipeline which does the

following: demultiplexing, trimming (primer, barcode and linker removal), merging read pairs and removal of bimera. Taxonomic classification was performed using the SILVA database (version 138.2) for 16S rRNA and the moyn database (version 1.1.0) for *nifH* (Moynihan, 2020). The JMF identified known buffer contaminants in their internal controls, for the 16S rRNA data, which were removed. I used the contaminant-free ASV count table and the taxonomic classification for subsequent analyses. All data analysis was performed with R (version 4.3.3), using the packages phyloseq (version 1.46.0), vegan (version 2.6.4), ggplot2 (version 3.5.1) and microViz (version 0.12.5).

6. Results

6.1. Nodule and soil sampling

At the first timepoint, due to the small plant size, not all five plants had 10 nodules, so I collected 35 and 47 nodules for *T. pratense* var. *Reichersberger* and *T. pratense* var. *Global* respectively. Nodule weight ranged from 0.1-2.5 mg for *T. pratense* var. *Reichersberger* and, 0.2-3.99 mg for *T. pratense* var. *Global* (Table 6). For *T. incarnatum*, five nodules were pooled for each plant due to the small size, resulting in a total of 5 samples for each sampling point.

For the first timepoint, 15 soil samples were collected, whereas for the second timepoint both rhizosphere and rhizoplane samples were collected, yielding 30 soil samples with a total of 45 soil samples.

Table 6: Number of nodules and nodule weight for each timepoint.

	Growth Phase		**Flowering**	
	Nodules	Weight [mg]	Nodules	Weight [mg]
T. pratense var. *Reichersberger*	35	0.2-1.7	49	0.2-3.47
T. pratense var *Global*	47	0.1-2.5	50	0.35-3.99
City samples			22	

For each variety, one soil sample per timepoint was analyzed by a collaborator at the BOKU for phosphorus and potassium content in the soil (Table 7), which ranged from 21.5 to 46.6 and 44.2 to 52.3 for phosphorus and potassium. Analysis of the other soil parameters is currently in progress.

Table 7: Potassium and phosphorus concentration in the soil for all varieties and timepoints.

Soil sample	**Phosphorus [mg/kg soil]**	**Potassium [mg/kg soil]**
T. incarnatum, sampling 1	24.63	49.84
T. incarnatum, sampling 2	36.41	45.36
T. pratense var. *Reichersberger, sampling 1*	24.65	48.97
T. pratense var. *Reichersberger, sampling 2*	21.58	44.29
T. pratense var. *Global, sampling 1*	46.61	47.69
T. pratense var. *Global, sampling 2*	24.70	52.38

6.2. DNA extraction

DNA extraction yielded concentrations ranging from 0.5 to 6.7 ng/µL for individual nodules and between 1.4 and 14.5 ng/µL for the pooled nodules of *T. incarnatum*. Soil samples yielded between 74.9 and 281 ng/µL (Table S 6, Table S 7).

6.3. Amplicon sequencing

Rhizobiales accounted for an average of 96.1% of the bacterial symbionts across all *Trifolium* varieties and sampling points (Figure 17). Overall, the samples from the city showed a higher colonization of NABs, with 7.1% of the 16S rRNA amplicon reads accounting for NABs, compared to 0.6% of NAB reads for the university experiment. When *Rhizobiales* were excluded (Figure 18), *Bacillales* emerged as the dominant NAB, present in 45.4% of all NAB containing samples, followed by *Burkholderiales*, *Pseudomonadales*, and *Corynebacteriales*.

NifH sequencing allowed identification of *Rhizobiales* down to the species level. *Six* dominant *nifH* ASVs were found across all three varieties, and all were classified as *Rhizobium leguminosarum*, (Table S 4). In contrast, in the city samples, three different rhizobial species were observed: *Rhizobium leguminosarum, Rhizobium anhuiense,* and *Sinorhizobium meliloti*. *R. leguminosarum* was present as the dominant symbiont (Table S 5).

Comparing the three *Trifolium* varieties showed that 15 out of 41 of the 16S rRNA ASVs were shared, including the rhizobia ASVs. The number of ASVs and, thus, the bacterial diversity was higher in both *T. pratense* species than in *T. incarnatum* (Figure 5). *T. pratense* var. *Reichersberger* had 2 unique and the city samples ten unique ASVs.

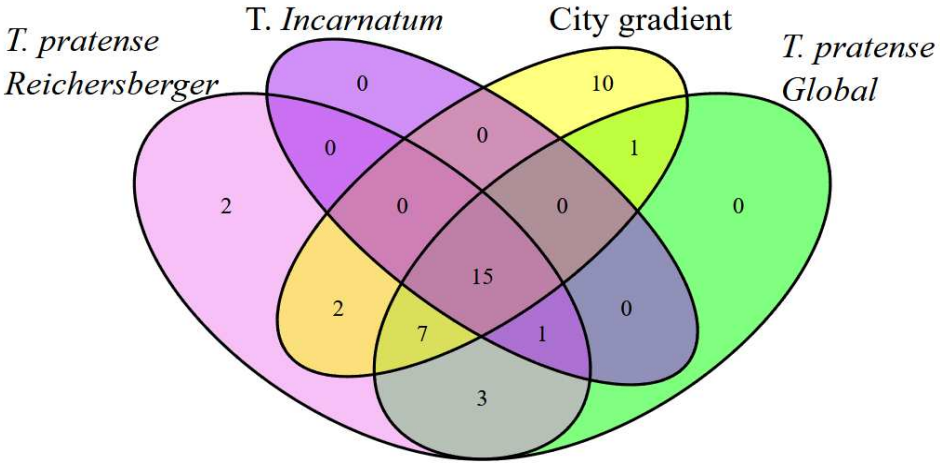

Figure 5: 16S rRNA shared ASVs among the three *Trifolium* varieties and the city samples.

Each sample exhibited a low Chao1 index, indicating low alpha diversity across all varieties, ranging from values of 10 to 20 (Figure 6, Figure 9). This reduced diversity was primarily driven by the dominance of *Rhizobiales* in each sample. A small difference in the Chao1 Index was observed between the two sampling timepoints for *T. pratense* var. *Global* and *Reichersberger.* At the first sampling point, *T. pratense* var. *Global* exhibited a higher mean Chao1 index, with 13.25, compared to the mean of 12.7 in the second timepoint. In contrast, *T. pratense Reichersberger* displayed a mean Chao1 diversity of 14.36 at the first timepoint and a higher Chao 1 index with 14.55 at the second timepoint (Figure 6). The Shannon diversity of root nodules of the different varieties ranged from 0.13 to 0.29 for both sampling timepoints (Figure 8).

Beta diversity analysis revealed low diversity within the nodule associated communities, as indicated by a Bray-Curtis NMDS clustering around 0. No significant changes in community composition were observed between plant varieties or across a temporal scale (Figure 7). The bacterial community composition did not exhibit significant clustering based on clover variety. Permutational multivariate analysis of variance (PERMANOVA) indicated no statistically significant differences ($p = 0.09$). Similarly, the sampling time points did not contribute to any clustering (PERMANOVA, $p = 0.14$).

Samples collected from the city showed low Chao 1 alpha diversity, ranging from 11 to 36, indicating small bacterial diversity. However, samples taken from

the second location at the Auer von Welsbach Park (AWPB 1 and AWPB 2) showed higher alpha diversity compared to the other samples, and thus expressed a higher community diversity (Figure 9). The Shannon index ranged from 0.1 to 0.85, with the highest values for the three samples, HO1, AWPB1 and AWPB2 (Figure 10).The Bray-Curtis NMDS analysis revealed low beta diversity, indicating no significant differences in the bacterial community composition of the nodules between the sampling locations.

Beta-diversity analysis using PERMANOVA indicated a significant difference in community composition when comparing the plants from the university and the city sampling ($p = 0.041$). However, this only accounted for 7.73% of the total variation, leaving 92.27% of the variation to other factors ($R^2 = 0.0773$).

A correlation was observed between nodule size and ASV abundance (Figure 11). Additionally, a correlation was observed between nodule weight and quantity of DNA, supporting the fact that larger nodules yielded larger quantity of extracted DNA (Figure S 1).

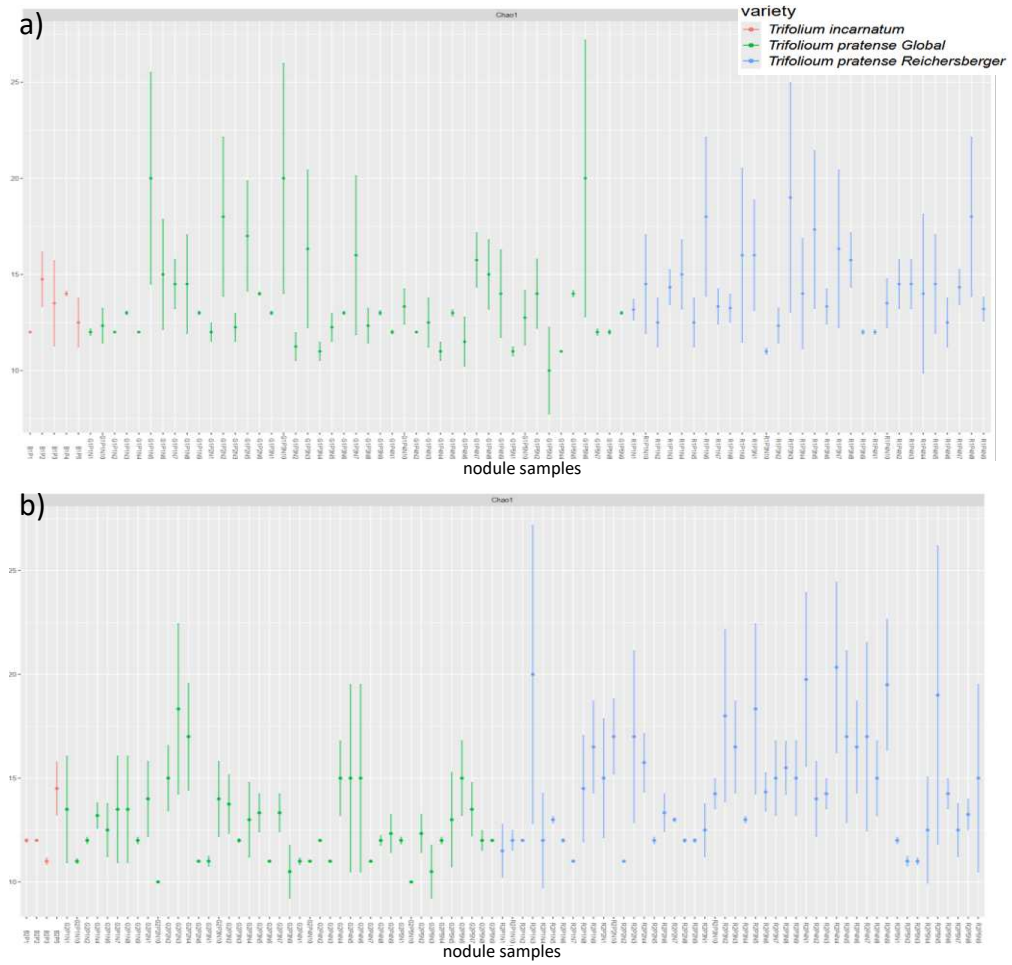

Figure 6: Chao 1 alpha diversity of 16S rRNA amplicons of the root nodules. a) nodules of each variety for the first sampling timepoint. b) nodules of each variety for the second timepoint

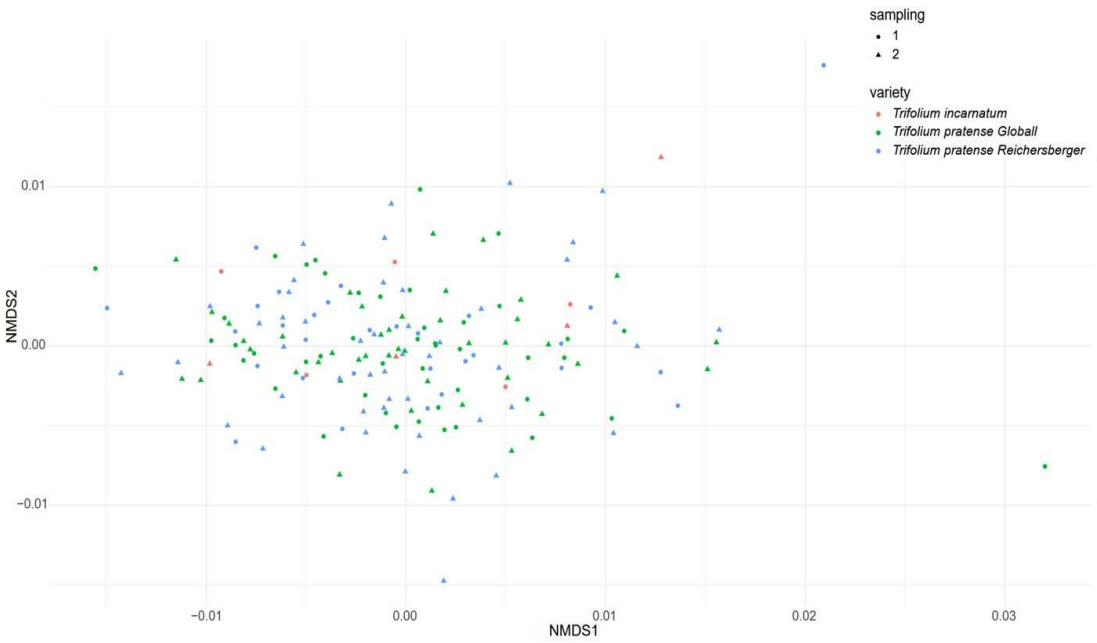

Figure 7: Bray- Curtis NMDS of 16S rRNA amplicons of the root nodules, both sampling points.

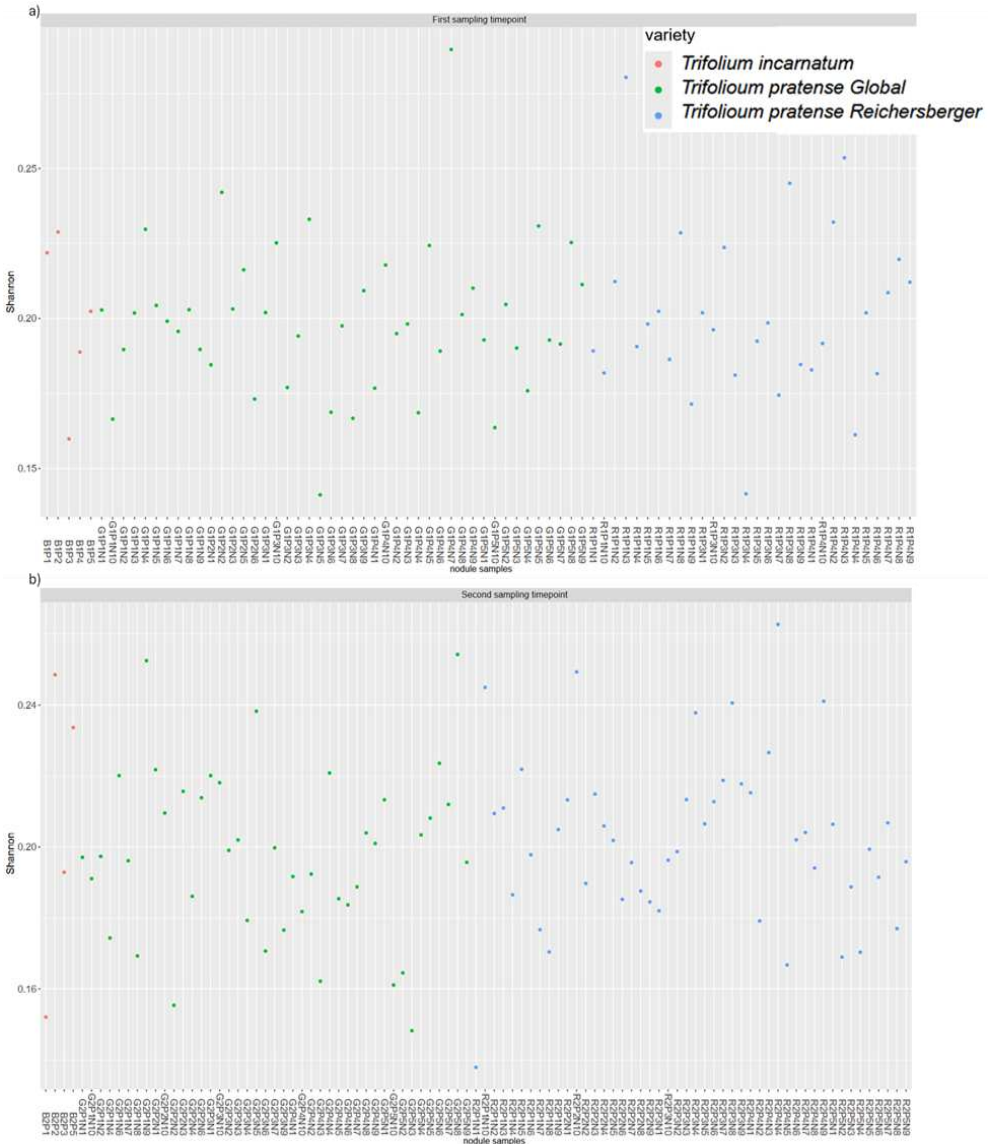

Figure 8: Shannon alpha diversity of 16S rRNA amplicons of the root nodules. a) nodules of each variety for the first sampling timepoint. b) nodules of each variety for the second timepoint

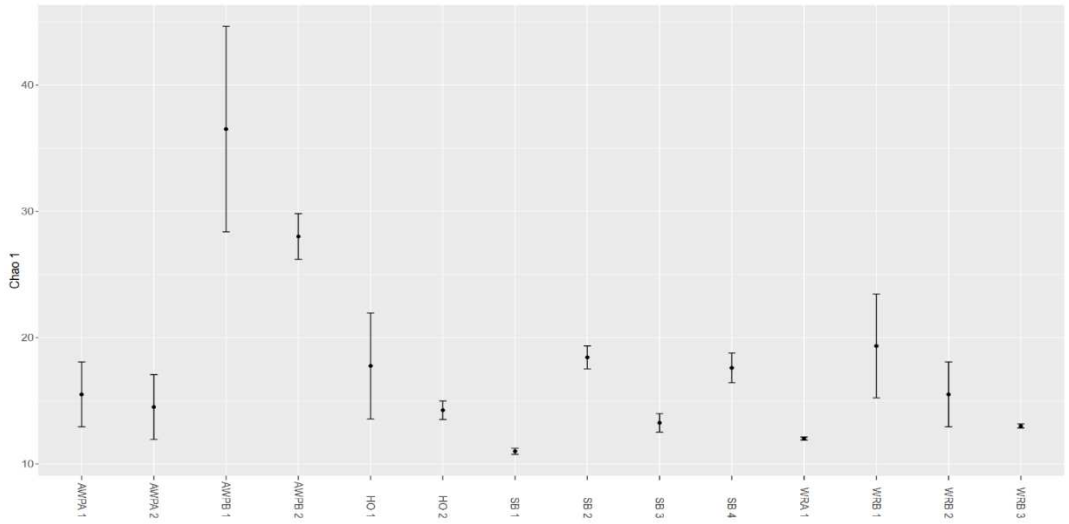

Figure 9: Chao 1 alpha diversity of 16S rRNA amplicons of the root nodules, sampled in the city

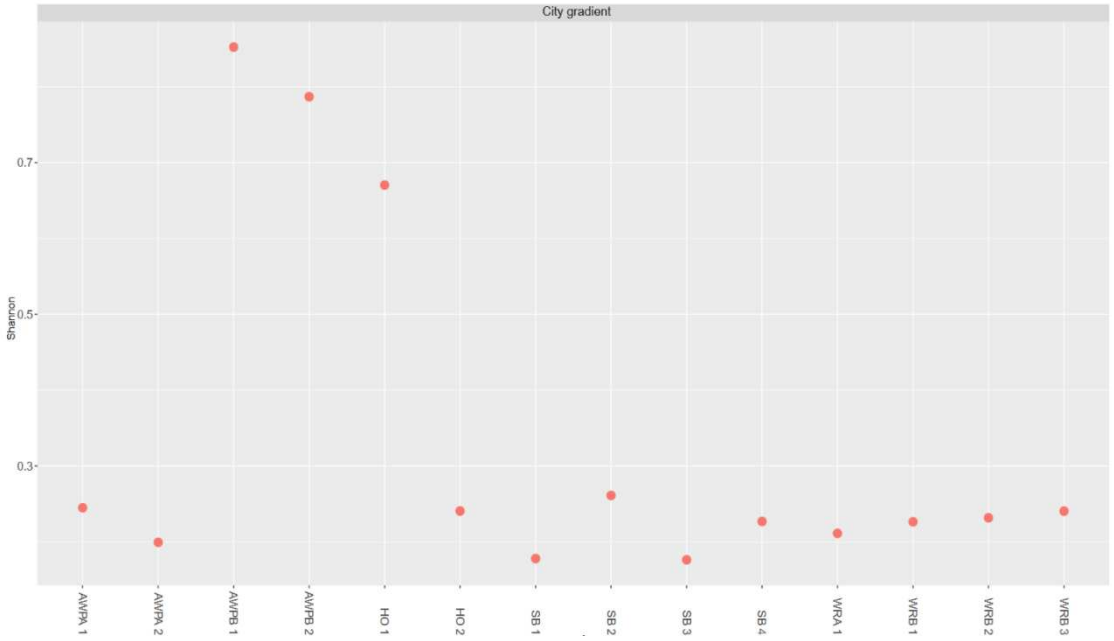

Figure 10: Shannon alpha diversity of 16S rRNA amplicons of the root nodules, sampled in the city.

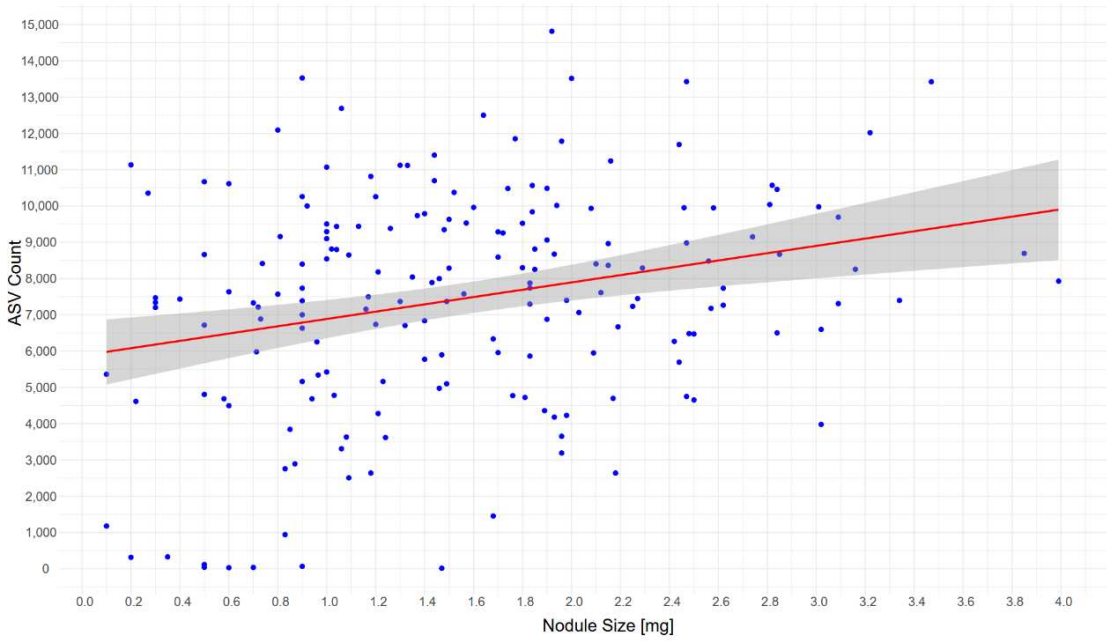

Figure 11: Correlation of nodule weight and ASV count, sequenced. Red: regression line.

Soil samples showed a high Chao1 index, ranging from 1320 to 1523, indicating a rich and diverse bacterial community (Figure 12). Chao1 alpha diversity slightly decreased at the second sampling time point, from a mean value of 1426 to 1393, suggesting a reduction in bacterial diversity over time or due to the different sampling methods. In comparison, rhizoplane samples displayed lower bacterial diversity than the rhizosphere samples (Figure 12). The Shannon diversity in the soil ranged from 5.9 to 6.4 (Figure 14).

A comparison of soil beta diversity using Bray-Curtis NMDS revealed low beta diversity. Bacterial community composition remained relatively stable across all soil samples. However, 11 samples showed an NMDS1 value higher than 0. These 11 samples were all derived from the rhizoplane, suggesting a shift in bacterial composition between the rhizosphere and the rhizoplane (Figure 13). Analysis of rhizosphere soil samples using PERMANOVA found that both the sampling timepoint and clover species significantly influenced bacterial community composition ($p = 0.002$, $p=0.003$). However, the R^2 values were 8.1% and 4.5% respectively. Therefore, the plant species and sampling timepoint only accounted for a small proportion of the observed diversity within the soil.

The soil samples from the city samples did express a high Chao 1 index, ranging from 807 to 1111, indicating a high alpha diversity in the sampled soil (Figure 9). The Shannon diversity of soil samples from the city samples ranged from 5.5 to 6.1 (Figure 10) Beta diversity indices could not be calculated due to the small number of city samples.

Four of the predominant *R. leguminosarum* ASVs identified in the plant samples were also detected in the corresponding soil samples. The two most abundant ASVs, ASV_kew_w1i and ASV_kvg_1lq, were consistently present in the soil across all three plant varieties. Not all ASVs detected in the plant samples were identified in the soil (Table S 4).

The rarefaction curve analysis reveals that the number of bacterial species has not yet reached a plateau for the soil samples. Additional sequencing might capture further bacterial diversity (Figure S 7).

The soil samples were analyzed for potassium and phosphorus concentration.

Phosphorus concentration of the soil decreased over time from an average of 31.9 mg/kg soil to 27.5 mg/kg soil over all samples. Potassium decreased from an average of 49.4 mg/kg soil to 45 mg/kg soil during the growth period of *T. pratense* var. *Reichersberger* and *T. incarnatum*. Only *T. pratense* var. *Global* showed an increase in potassium concentration from 47.7 mg/kg soil to 52.4 mg/kg soil during plant growth (Table *7*).

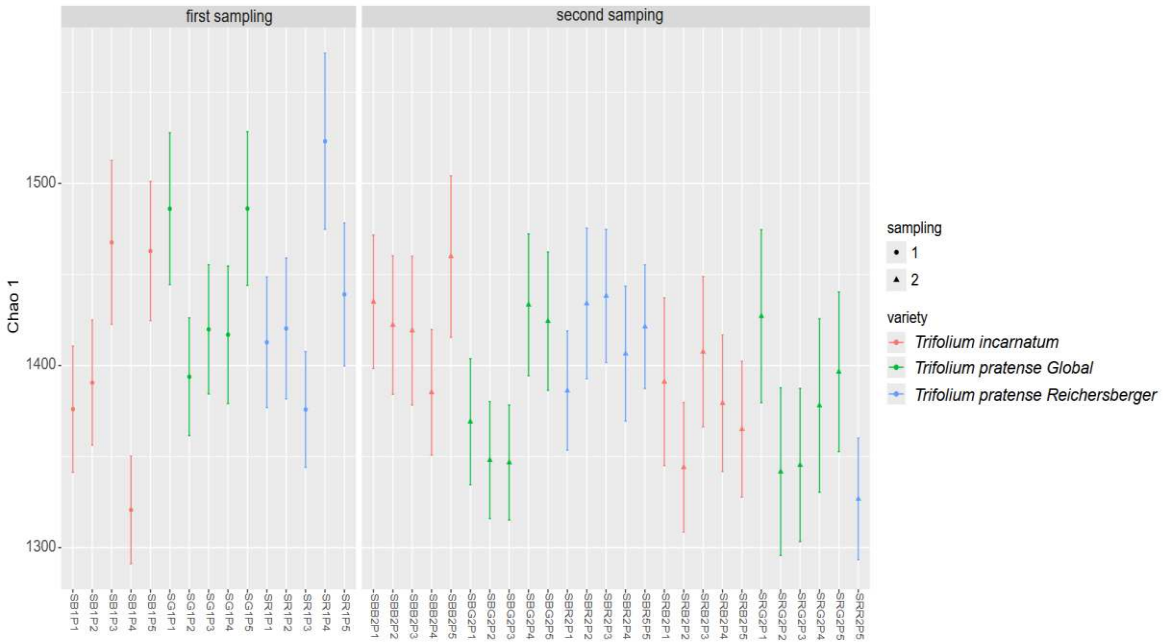

Figure 12: Chao 1 alpha diversity of 16S rRNA amplicons of the soil samples. The first and the second sampling points can be seen.

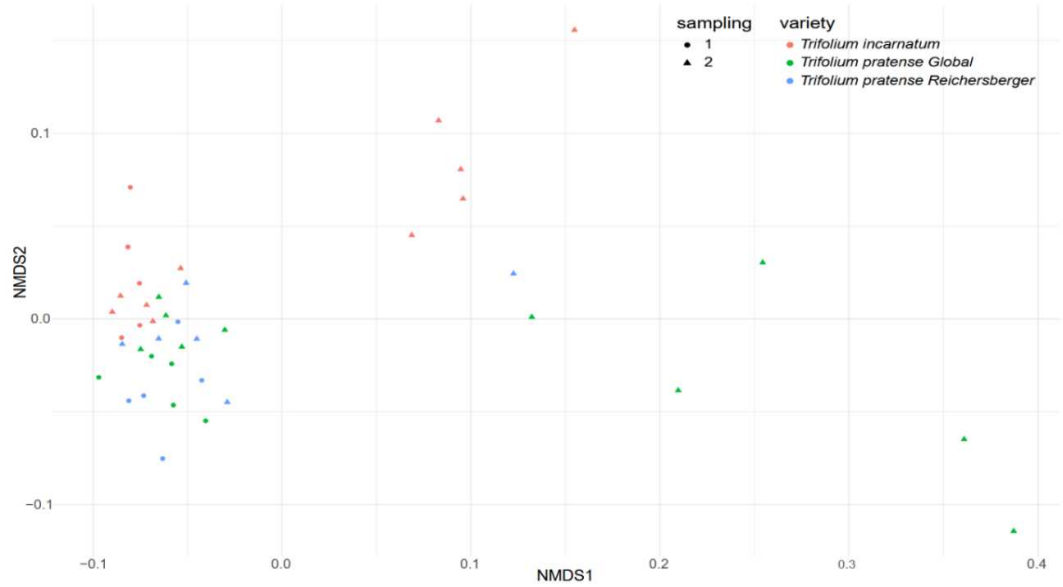

Figure 13: Bray- Curtis NMDS of 16S rRNA amplicons of the soil samples, both sampling points.

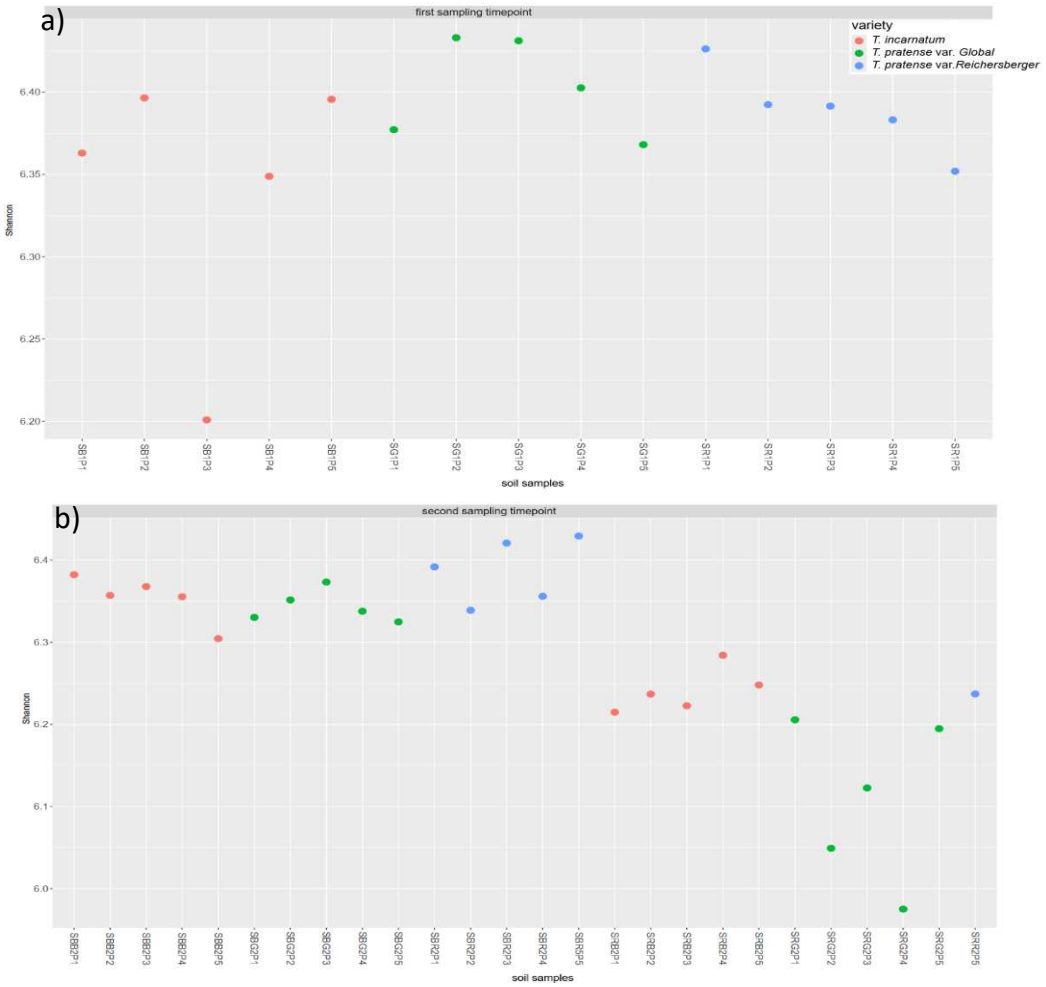

Figure 14: Shannon alpha diversity of 16S rRNA amplicons of the soil. a) soil of each variety for the first sampling timepoint. b) soil of each variety for the second timepoint

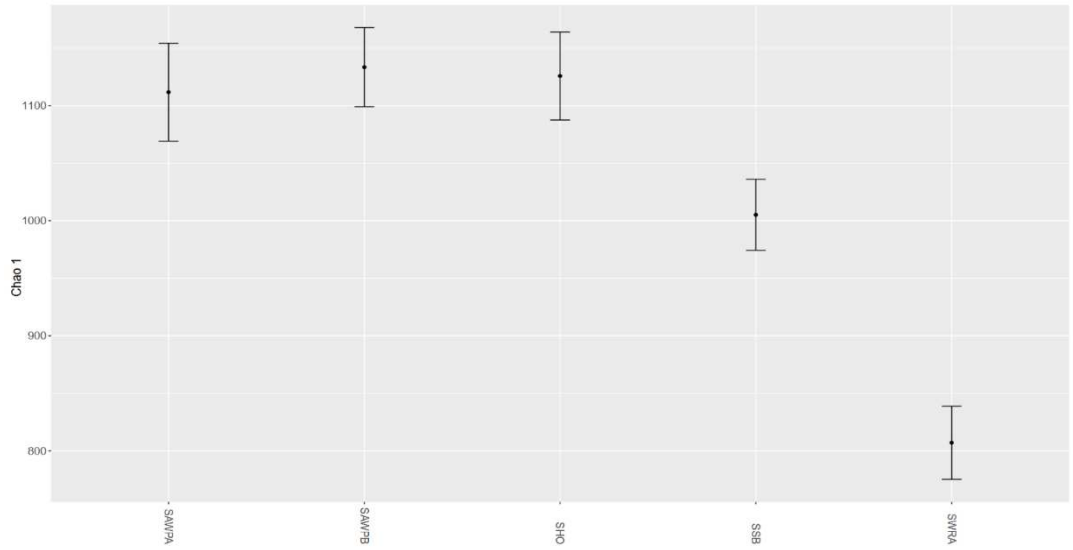

Figure 15: Chao 1 alpha diversity of 16S rRNA amplicons of the soil, sampled in the city

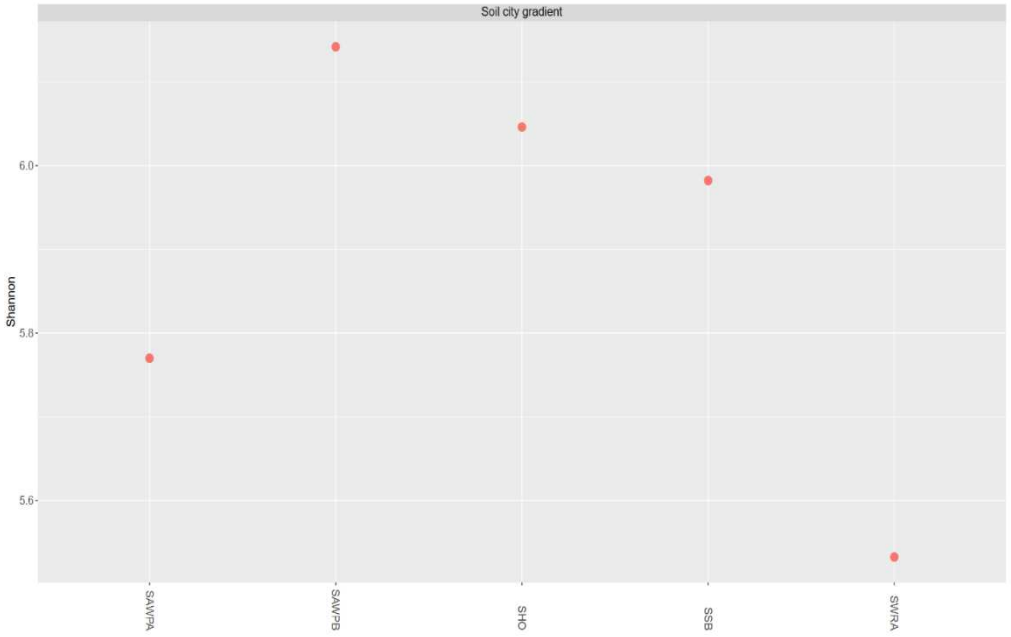

Figure 16: Shannon alpha diversity of 16S rRNA amplicons of the soil, sampled in the city

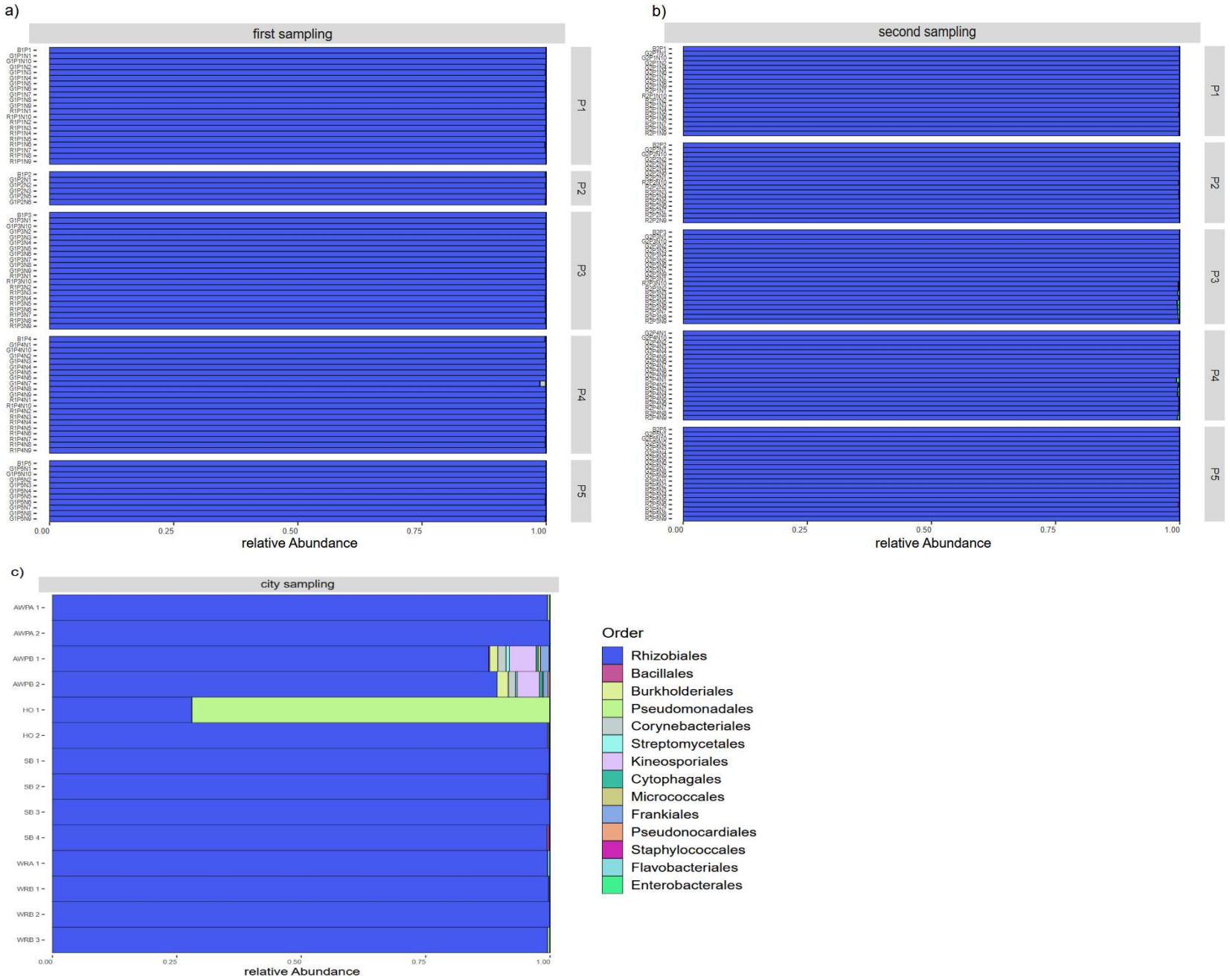

Figure 17: Relative abundance of bacterial community in nodules obtained from 16S rRNA amplicon sequencing. a) First sampling timepoint, b) second timepoint and c) city samples. Each bar represents one individual nodule, P1 to P5 = Plant 1 to 5.

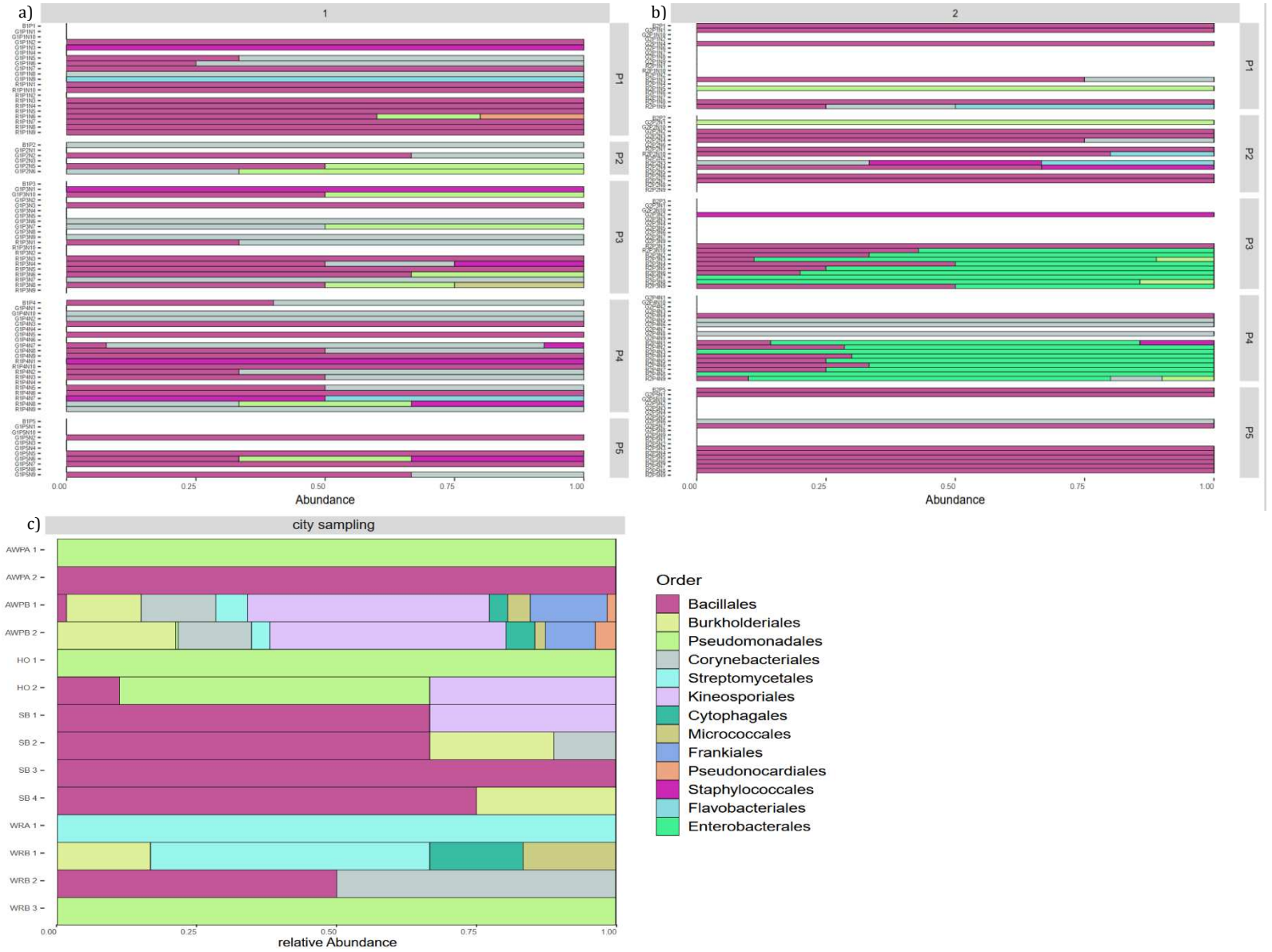

Figure 18: Relative abundance of bacterial community in nodules obtained from 16S rRNA amplicon sequencing with *Rhizobiales* removed. a) First sampling timepoint, b) second timepoint and c) city samples. Each bar represents one individual nodule. Empty rows are nodules inhabited by 100% *Rhizobiales*.

The following sections present the results for each variety individually, showing key differences and trends.

6.3.1. T. pratense var. *Reichersberger*

6.3.1.1. Nodules

Rhizobiales were identified as the dominant symbiont across all plants and sampling points with an average relative abundance of 97.67% (Figure 19). In the first sampling timepoint, five nodules and in the second timepoint, 10 nodules contained only *Rhizobiales*. In all other nodules *Bacillales* were the most abundant NAB, which comprised 31.12% of all NABs present in the nodules. Six main ASVs for *Bacillales* were identified (Table S *2*). Co-infection by multiple NABs alongside *Rhizobiales* was observed, at both timepoints (Figure 20).

However, for two nodules collected during the first timepoint (R1P3N7 and R1P4N9) *Corneybaceriales* were the sole NAB present. In four nodules (R2P3N7, R2P3N8, R2P4N3 and R2P4N8) *Enterobacterales* were the only present NAB.

While *Bacillus* was present as NAB in most of the nodules during both timepoints, *Enterobacterales* emerged as dominant NAB in plants three and four (18 out of 19 nodules) during the flowering phase, However, it could not be detected in the other three plants sampled at the flowering phase, showing a plant-specific trend.

Looking at the *nifH* level, the nodules were dominated by four ASVs belonging to *R. leguminosarum* with no changes in ASV abundance over both sampling points (Table S 4).

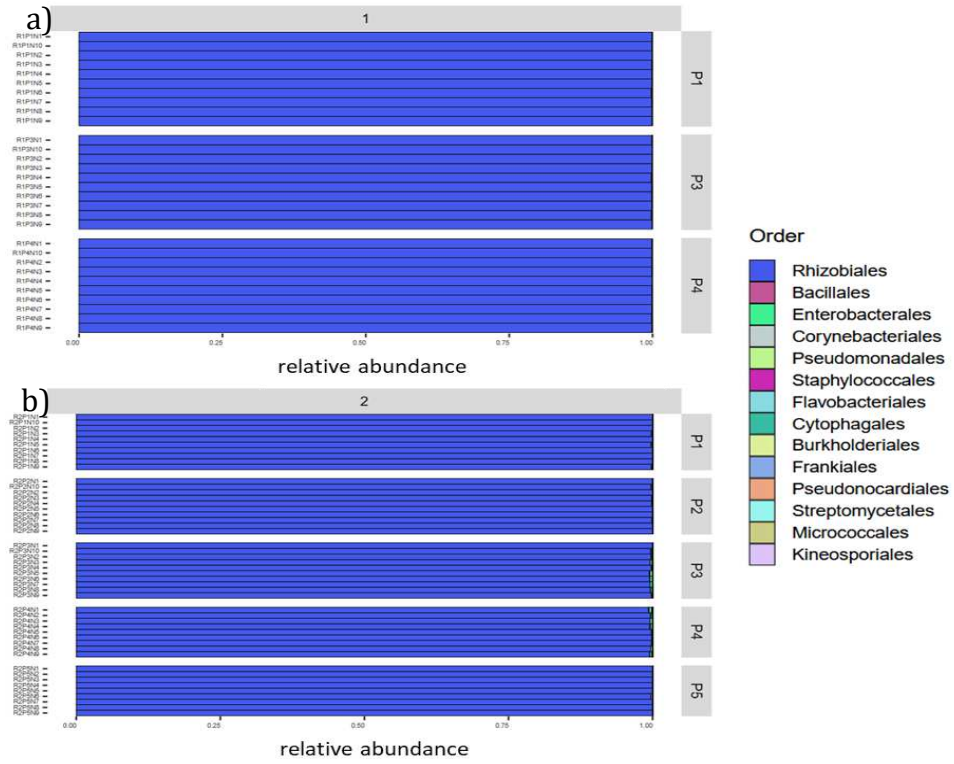

Figure 19: Relative abundance of nodule bacterial community in *T. pratense* var. *Reichersberger* based on 16S-rRNA amplicon sequencing. Each bar represents one nodule. a) First sampling timepoint. b) second sampling timepoint.

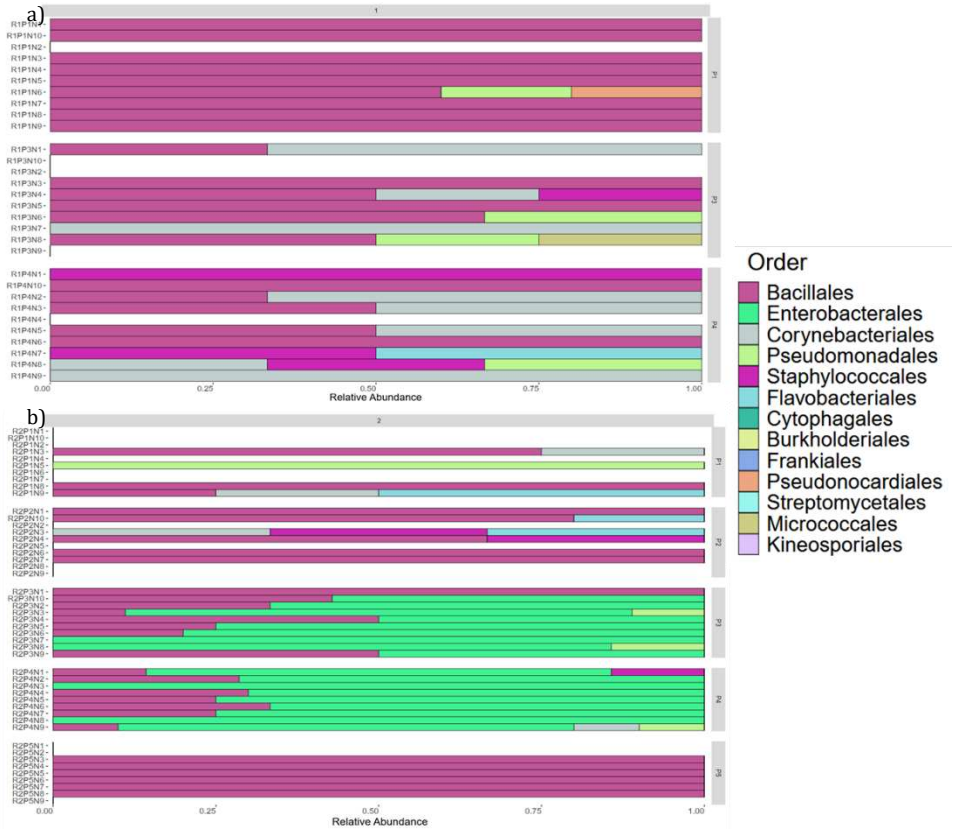

Figure 20: Relative abundance of nodule bacterial community in *T. pratense* var. *Reichersberger* based on 16S-rRNA amplicon sequencing with *Rhizobiales* removed. Each bar represents one nodule. a) First sampling timepoint. b) second sampling timepoint.

6.3.1.2. *Soil*

No soil community shifts were observed during the growth period (Figure 21). Four out of five rhizoplane samples were filtered out during prevalence filtering of the 16S rRNA amplicons. Comparing the rhizosphere sample with the rhizoplane sample from plant 5 shows a slight shift in the community composition (Figure 21). A comparison of the ASVs belonging to *Bacillales* in the nodules and soil, showed that the ASVs present in the nodules were also detected in the soil. Those ASVs were among the most abundant *Bacillales* ASVs in the soil. Additionally, one high abundant *Bacillales* ASV was detected, ASV_4z5_fws_llw, that could not be found in the nodules (Table S 2).

A slight variation in bacterial composition was observed among rhizosphere samples from the growth phase. Sample SR1P3 had a higher abundance of *Sinorhizobium,* which was present as a small fraction in the other four rhizosphere samples.

NifH sequencing showed an overall increase in *Rhizobiales* abundance during the flowering stage of the plant. Rhizosphere samples SBR2P1 and SBR2P2 showed a slight increase in *Rhizobiales* presence compared to the first timepoint. They contained 8.3% and 13.7% *Rhizobiales*, respectively. In contrast, the rhizosphere samples SBR2P4 and SBR2P5 exhibited a higher abundance of *Rhizobiales*, with 26.6% and 27.7% respectively.

In contrast all rhizoplane samples were dominated by *Rhizobiales*, with an average of 47.8% *Rhizobiales* in these five samples. Indicating a distinct shift in bacterial community composition close to the plant (Figure 22).

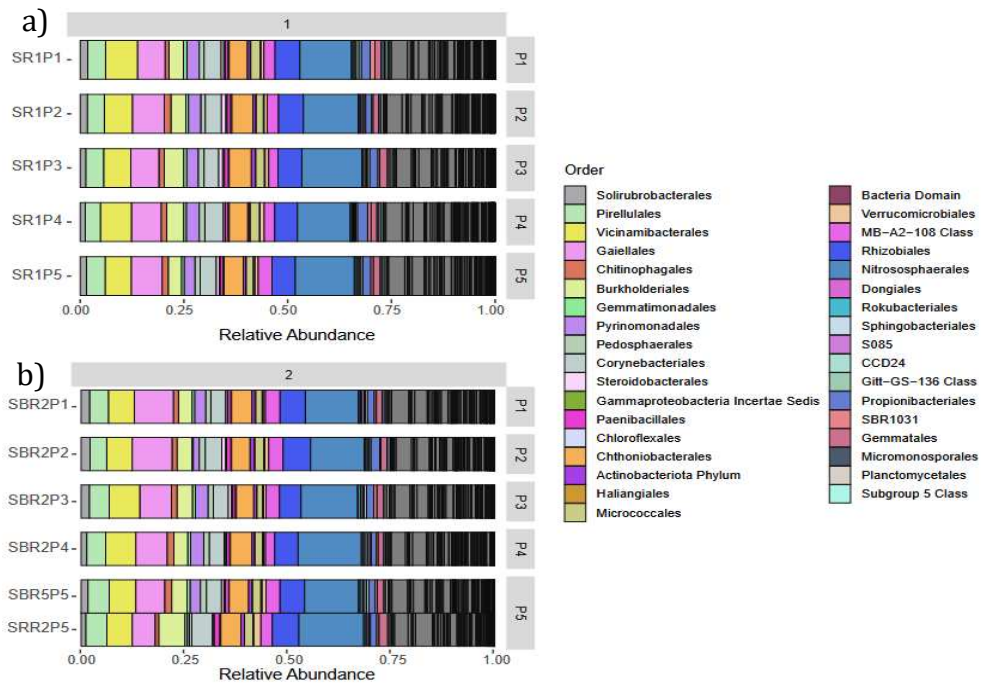

Figure 21: Relative abundance of soil bacterial community in *T. pratense* var. *Reichersberger,* based on 16S rRNA V4 amplicons a) first sampling point, rhizosphere was sampled b) second sampling point, rhizosphere and rhizoplane was sampled. Rhizoplane: SRR2P5.

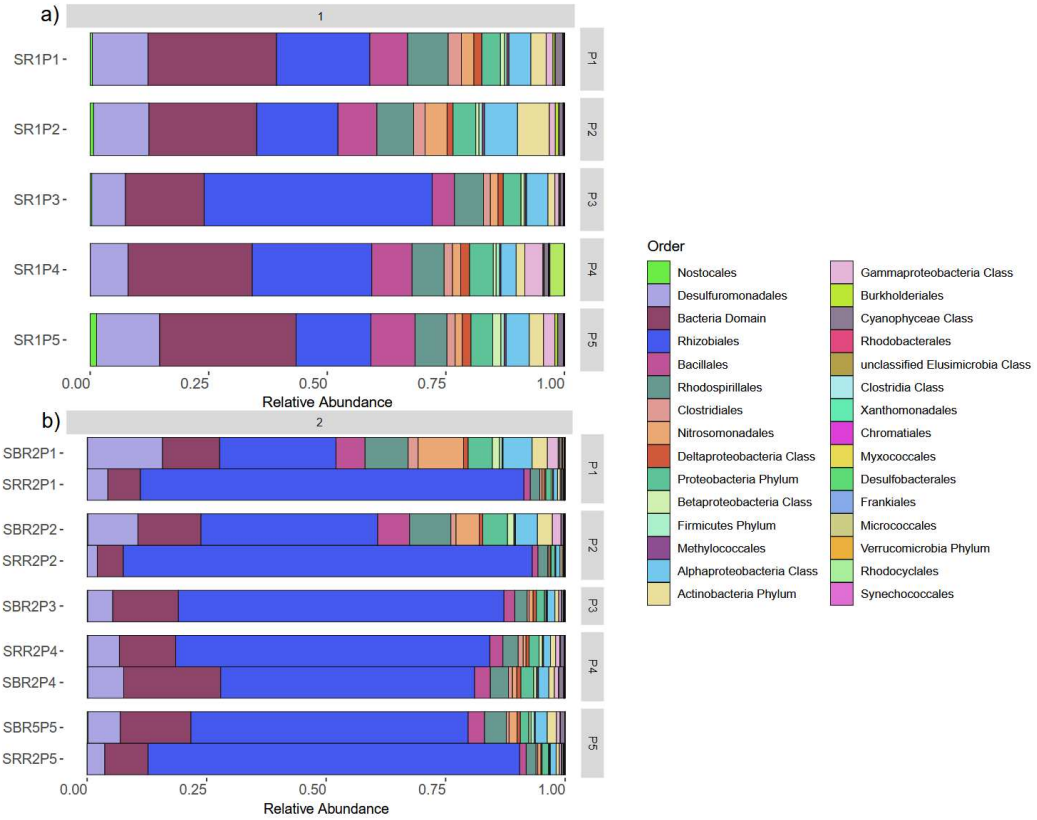

Figure 22: Relative abundance of soil bacterial community in *T. pratense* var. *Reichersberger*, based on nifH amplicons. a) first sampling point b) second sampling point. Rhizoplane: SRR2P1, SRR2P2, SRR2P3, SRR2P4, SRR2P5

6.3.2. T. pratense var. *Global*

6.3.2.1. Nodules

Rhizobiales were identified as the dominant symbiont across all plants and sampling points with an average relative abundance of 99.94% (Figure 23). In the first and second sampling timepoint, 17 and 28 nodules were inhabited solely by *Rhizobiales,* respectively. The main NAB was *Bacillales*, making up 49% of all NABs in *T. pratense* var. *Global*. Overall, nodules were dominated by *Bacillales, Corynebacteriales* and *Staphylococcales* as main NABs. Co-infection of nodules by multiple NABs occurred mainly in the first sampling timepoint. The same six main ASVs for *Bacillales* as in *T. pratense* var. *Reichersberger* could be identified (Table S 2).

The primary NABs in the second timepoint were *Bacillales* and *Corynebacteriales* (Figure 24). Except for one nodule (plant 2), all nodules contained only a single NAB. Notably, in plant 3, a single nodule (G2P3N2) was colonized by *Staphylococcales*, while the remaining nodules of this plant showed no presence of NABs.

Based on *nifH* amplicons *R. leguminosarum* was identified as the main symbiont. At the first sampling timepoint five predominant *R. leguminosarum* ASVs were found, whereas for the second sampling timepoint it was four ASVs that remained dominant (Table S 4).

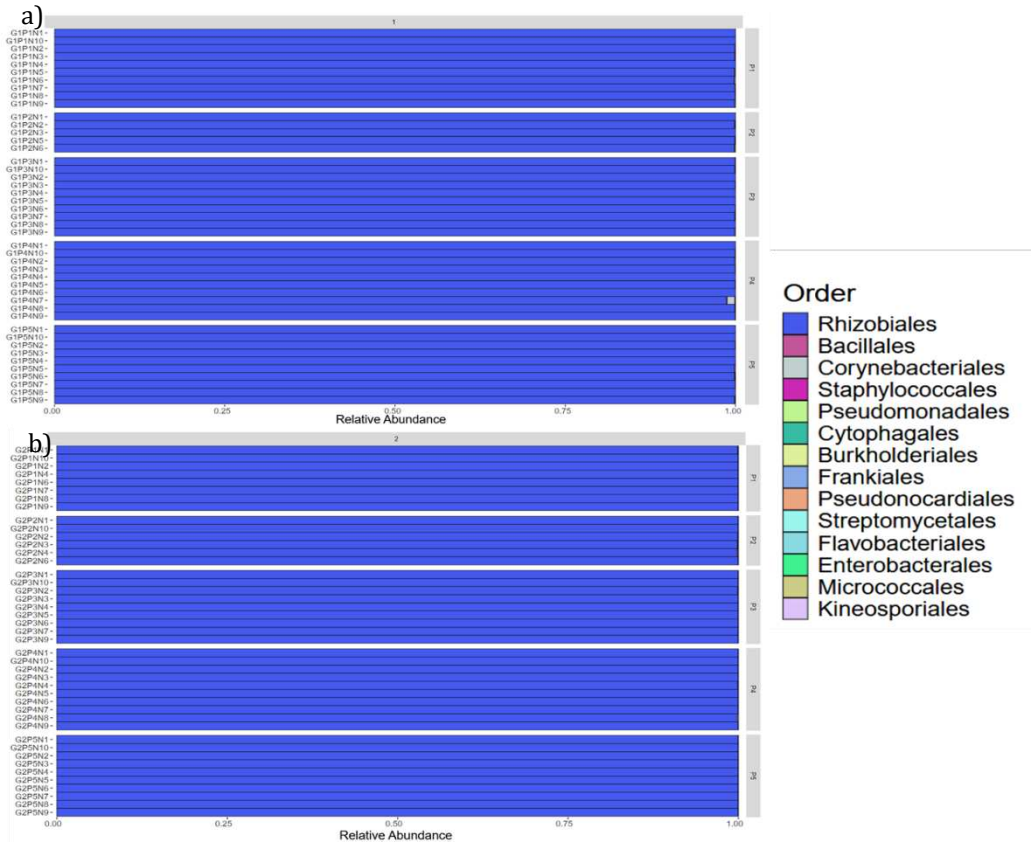

Figure 23: Relative abundance of nodule bacterial community in *T. pratense* var. *Reichersberger* based on 16S-rRNA amplicon sequencing. Each bar represents one nodule. a) First sampling timepoint. b) second sampling timepoint.

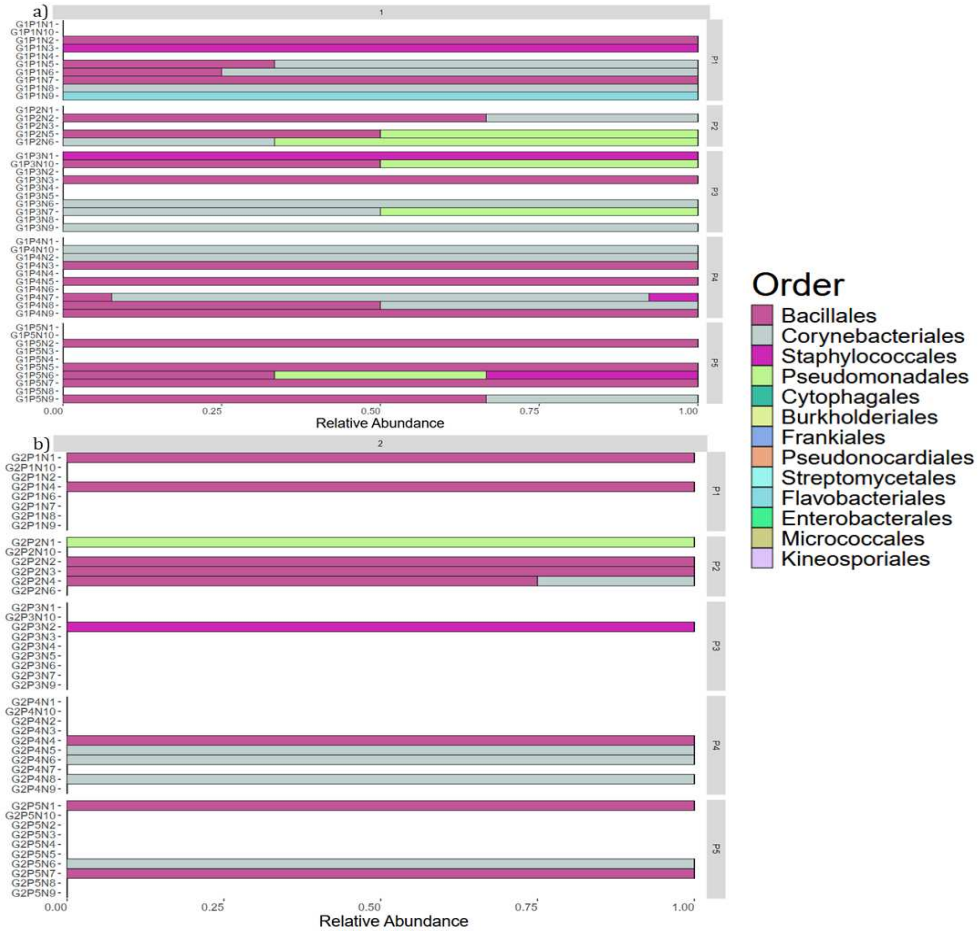

Figure 24: Relative abundance of nodule bacterial community in *T. pratense* var. *Reichersberger* based on 16S-rRNA amplicon sequencing with *Rhizobiales* removed. Each bar represents one nodule. a) First sampling timepoint. b) second sampling timepoint.

6.3.2.2. Soil

The soil microbial community in the rhizosphere remained stable throughout plant growth (Figure 25), including no changes in the presence of *Rhizobiales*. During the growth and the flowering phase, an average of 5.8% of the amplicon reads were detected from *Rhizobiales* (Figure 25).

In contrast, the bacterial community in the rhizoplane exhibited a slight shift in the bacterial community, and *Rhizobiales* abundance, compared to the rhizosphere to an average of 7.8% (Figure 25). A comparison of the ASVs belonging to *Bacillales* in the nodules and soil showed that the most common *Bacillales* ASVs in the nodules were also the most prevalent *Bacillales* ASVs in the soil. Additionally, one high abundant *Bacillales* ASV was detected in the soil, ASV_4z5_fws_llw, that could not be detected in the nodules (Table S 2)

Among the *nifH* gene carrying bacteria, the relative abundance of *Rhizobiales* varied across different soil samples from 16.7% to 56%, in the rhizosphere. The rhizoplane samples, from the second timepoint exhibited a distinct bacterial composition compared to the rhizosphere, showing a higher representation of *Rhizobium* species near the plant roots, with 58% to 86.6% of the ASVs being *Rhizobiales* (Figure 26).

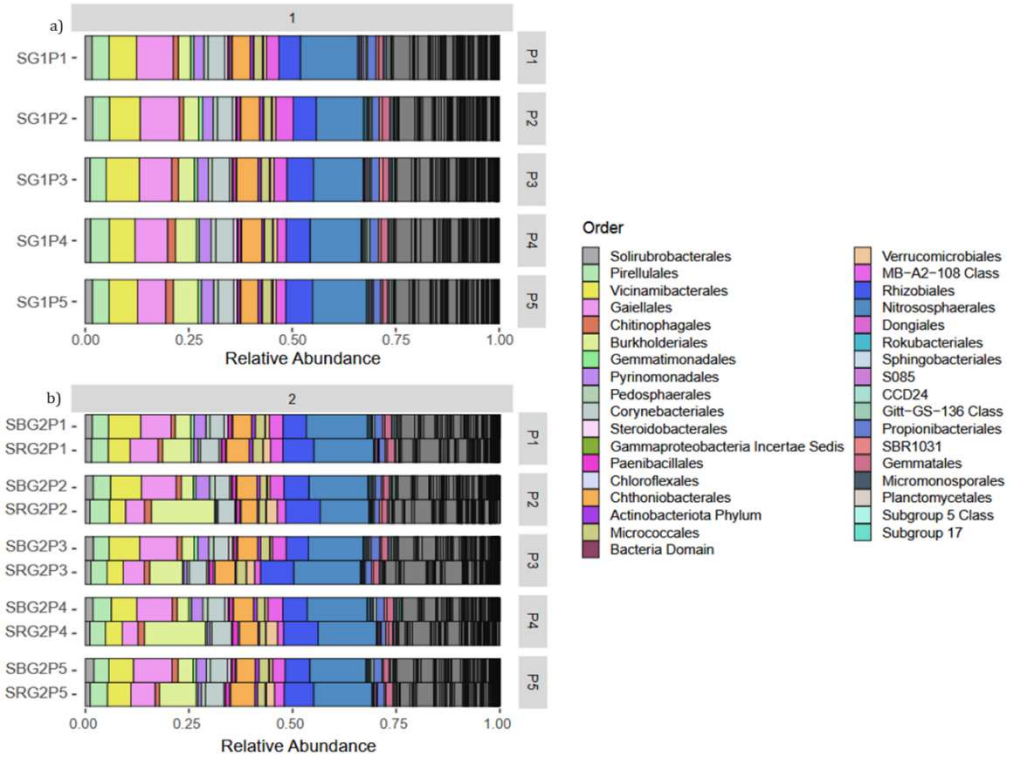

Figure 25: Relative abundance of soil bacterial community in *T. pratense* var. *Global,* based on 16S rRNA V4 amplicons a) first sampling point, rhizosphere was sampled b) second sampling point, rhizosphere and rhizoplane was sampled. Rhizoplane samples: SRG2P1, SRG2P2, SRG2P3, SRG2P4, SRG2P5.

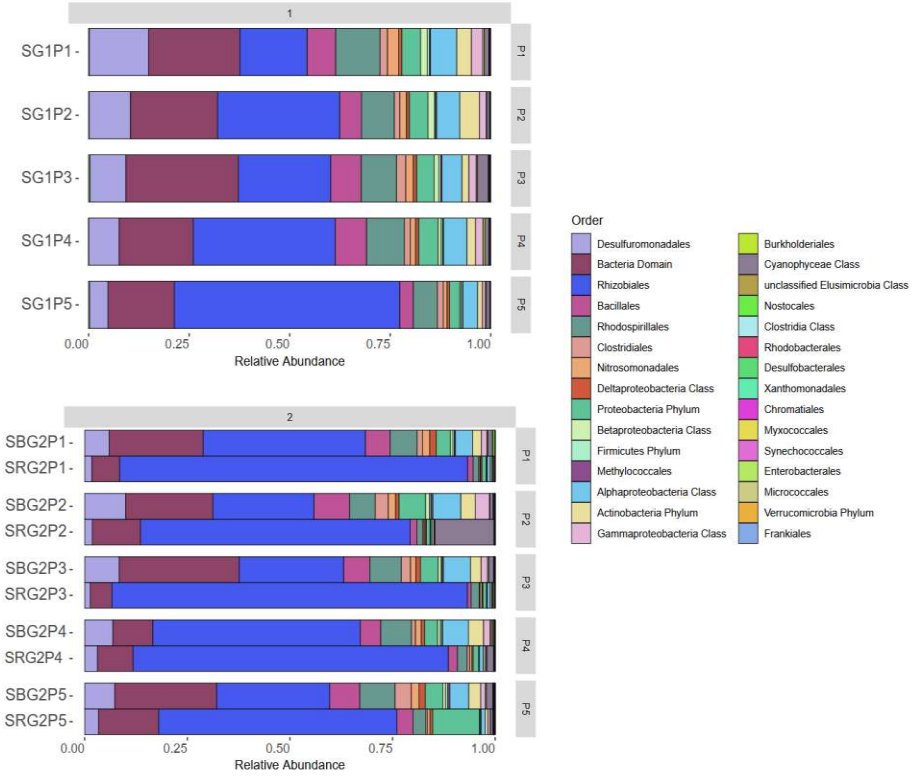

Figure 26: Relative abundance of soil bacterial community in *T. pratense* var. *Global,* based on *nifH* amplicons. a) first sampling point b) second sampling point. Rhizoplane: SRG2P1, SRG2P2, SRG2P3, SRG2P4, SRG2P5

6.3.3. T. incarnatum

For *T. incarnatum* five nodules were pooled for each plant, due to the small size.

6.3.3.1. Nodules

All plants were dominated by bacteria from the order *Rhizobiales* with an average abundance of 99.93% (Figure 27).

At the first sampling timepoint, nodules of two plants were also occupied by NABs, with *Corynebacteriales* present as the dominant NAB; the remaining three plants were exclusively colonized by *Rhizobiales* (Figure 28). By the second sampling timepoint, *Bacillales* emerged as the primary NAB in two plants, though two of four plants remained solely inhabited by *Rhizobiales*. Two ASVs for *Bacillales* could be detected in the samples, each of which were only present in a small abundance (Table S 2).

NifH sequencing of *T. incarnatum* samples showed *R. leguminosarum* as the sole nitrogenase carrying symbiont, with five dominant ASVs. By the second sampling timepoint, three of these ASVs remained dominant, while two ASVs were lost during the transition to the flowering phase (Table S 4).

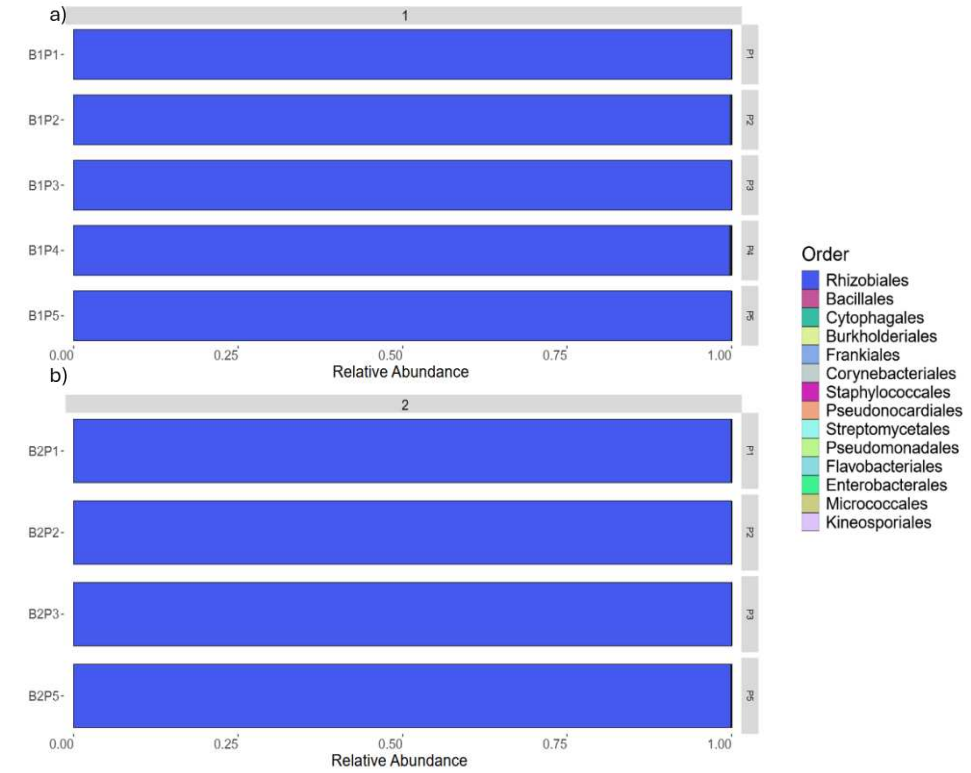

Figure 27: Relative abundance of nodule bacterial community in *T. incarnatum* based on 16S-rRNA amplicon sequencing. Each bar represents five pooled nodules of one plant. a) First sampling timepoint. b) second sampling timepoint.

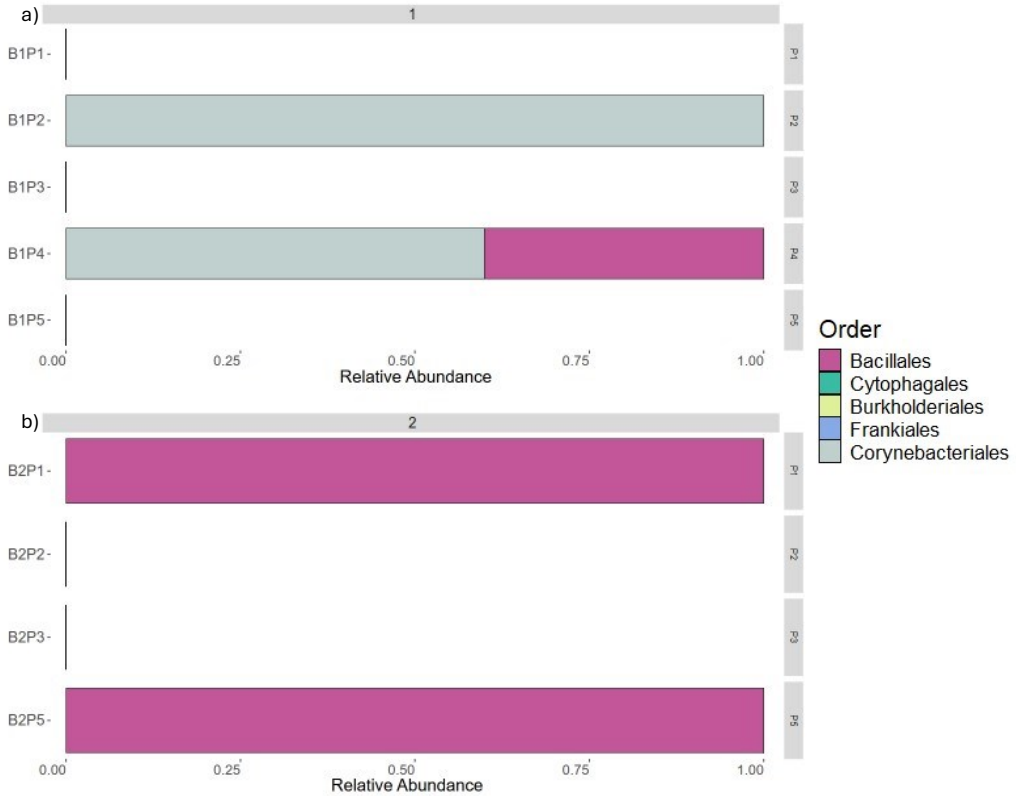

Figure 28: Relative abundance of nodule bacterial community in *T. incarnatum* based on 16S-rRNA amplicon sequencing with *Rhizobiales* removed. Each bar represents five pooled nodules of one plant. a) First sampling timepoint. b) second sampling timepoint

6.3.3.2. Soil

The microbial community in the rhizosphere was very similar for all plants across timepoints. No changes were observed in the presence of *Rhizobiales* when comparing rhizoplane sampling timepoints. During the growth phase, an average of 5.7% of the ASVs were from *Rhizobiales*, and during the flowering phase, the average was 5.6% (Figure 29). In the rhizoplane the average read of *Rhizobiales* was slightly higher with an average of 6.5%. Overall, the community of the rhizoplane showed minor differences to the rhizosphere samples.

The same six main *Bacillales* ASVs could be detected, as in the other clover varieties, including the two ASVs present in the plant nodules of *T. incarnatum.* Additionally, a highly abundant *Bacillales* ASV, ASV_4z5_fws_llw, was detected in the soil but not in the nodules (Table S 2).

Amplicon sequencing of *nifH* showed consistent results for the first timepoint, with only minor variations in ASV counts. The relative abundance of *Rhizobiales* varied across different soil samples from 9.2% to 36.6%, in the rhizosphere. In the second sampling timepoint, the rhizoplane samples exhibited a distinct bacterial composition compared to the rhizosphere, showing a higher representation of *Rhizobium* species near the plant roots, with 34.6% to 83% of the *nifH* ASVs being *Rhizobiales* (Figure 30).

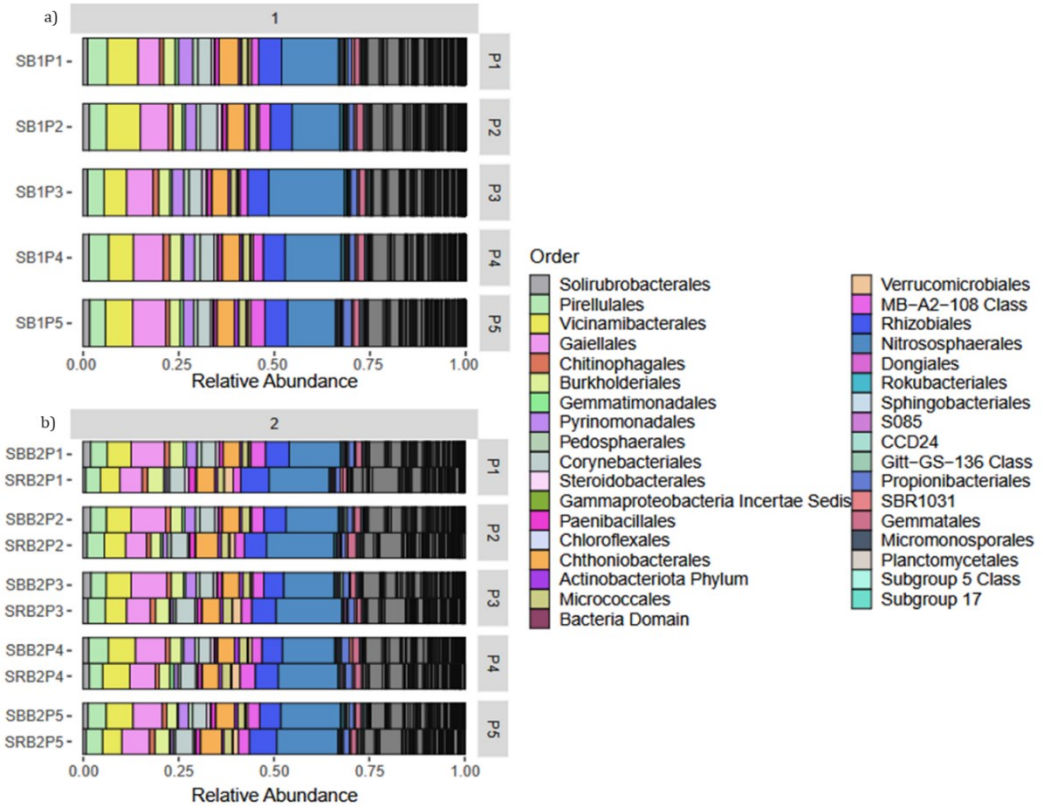

Figure 29: Relative abundance of soil bacterial community in *T. incarnatum,* based on 16S rRNA V4 amplicons a) first sampling point, rhizosphere was sampled b) second sampling point, rhizosphere and rhizoplane was sampled. Rhizoplane samples: SRB2P1, SRB2P2, SRB2P3, SRB2P4, SRB2P5.

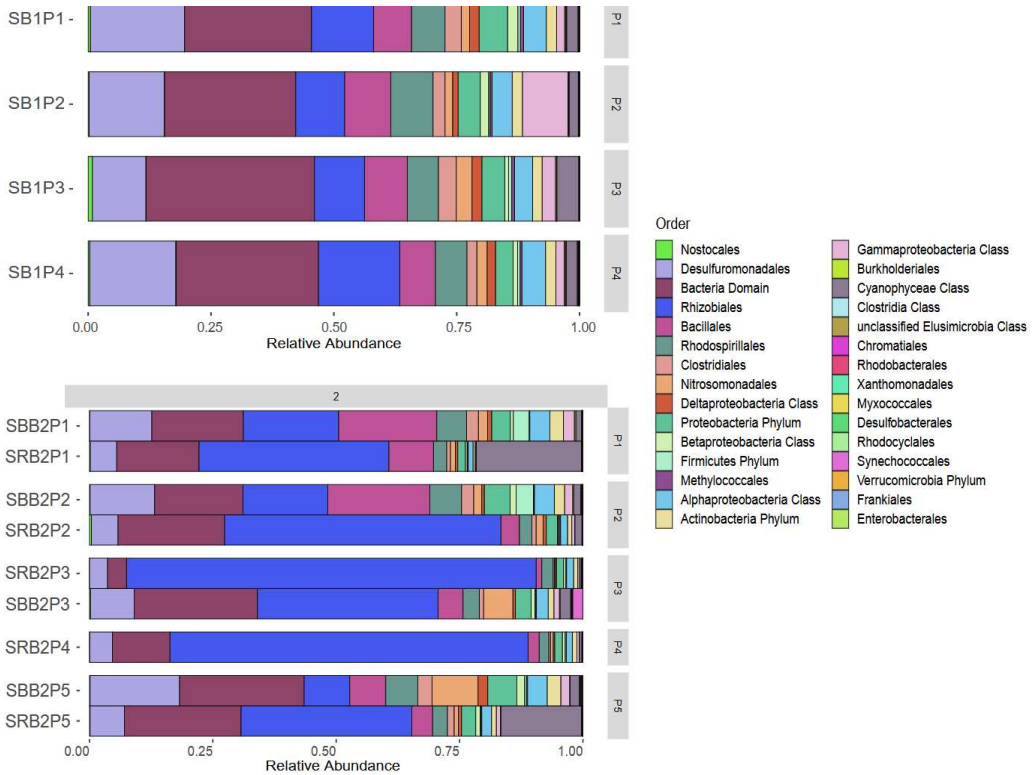

Figure 30: Relative abundance of soil bacterial community in T. incarnatum, based on nifH amplicons. a) first sampling point b) second sampling point. Rhizoplane: SRB2P1, SRB2P2, SRB2P3, SRB2P4, SRB2P5

6.3.4. City *samples*

6.3.4.1. Nodules

Rhizobiales were the predominant symbionts across all samples, comprising 92.95% of the total bacterial sequence reads. Three nodules had a higher NAB concentration than the average: HO 1 (71.9% NAB) collected in the residential garden, AWPB 1 (12.3%) and AWPB 2 (10.65%), collected from a frequently populated area of the park (Figure 31). NAB composition varied by location, e.g. the Schönbrunn samples were dominated by *Bacillales*. *Pseudomonadales* were the sole NAB in three samples (HO 1, WRB 3, and AWPA 1). Most nodules showed multi-infection with different bacterial orders (Figure 32). At the ASV level, the most abundant symbionts were *Bacillales, Burkholderiales, Cytophagales, Frankiales, Corynebacteriales, Streptomycetales, Kineosporiales* and *Pseudomonadales* (Table S 3).

NifH amplicon sequencing identified *R. leguminosarum* as the dominant symbiont in all nodules. At the ASV level, five major *nifH* ASVs were identified, with one ASV from *R. leguminosarum* present, in all six sampling locations. Overall, ASVs belonging to various rhizobial species were identified in the different samples. These included *R. leguminosarum*, *R. anhuiense* and *Sinorhizobium meliloti*. The highest *Sinorhizobium* sp. levels were found in Auer von Welsbach Park samples AWPB 1 and AWPB 2, containing 3.32% and 4.16%, respectively (Table S 5).

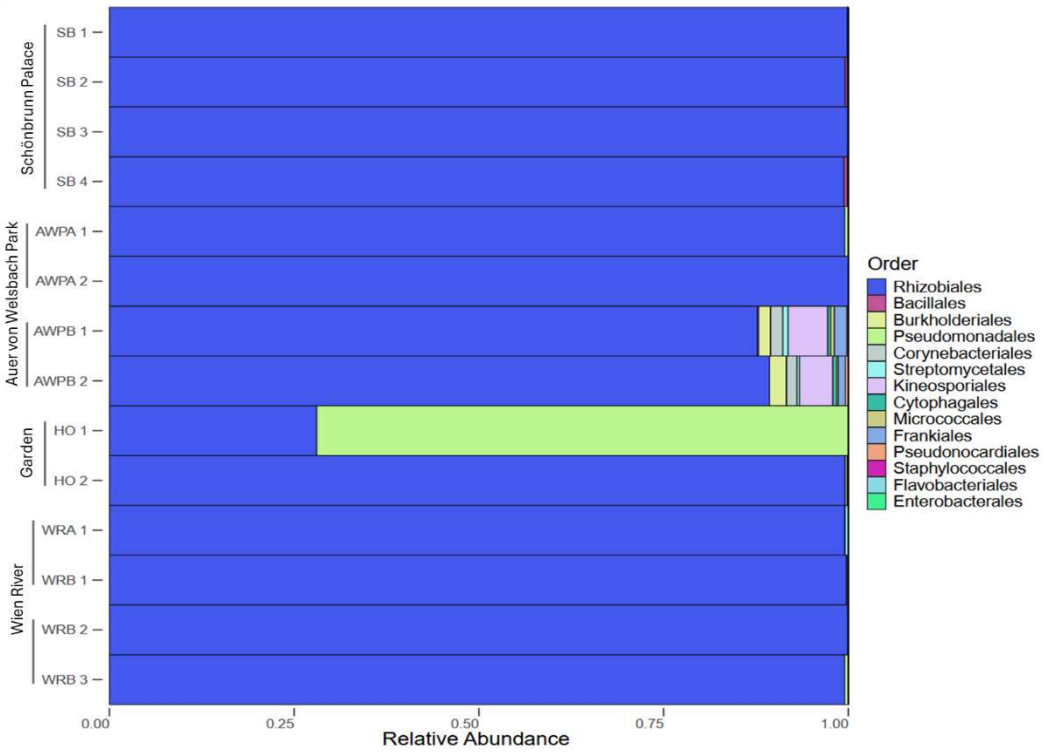

Figure 31: Relative abundance of nodule bacterial community of the city samples based on 16S-rRNA amplicon sequencing. Each bar represents one nodule.

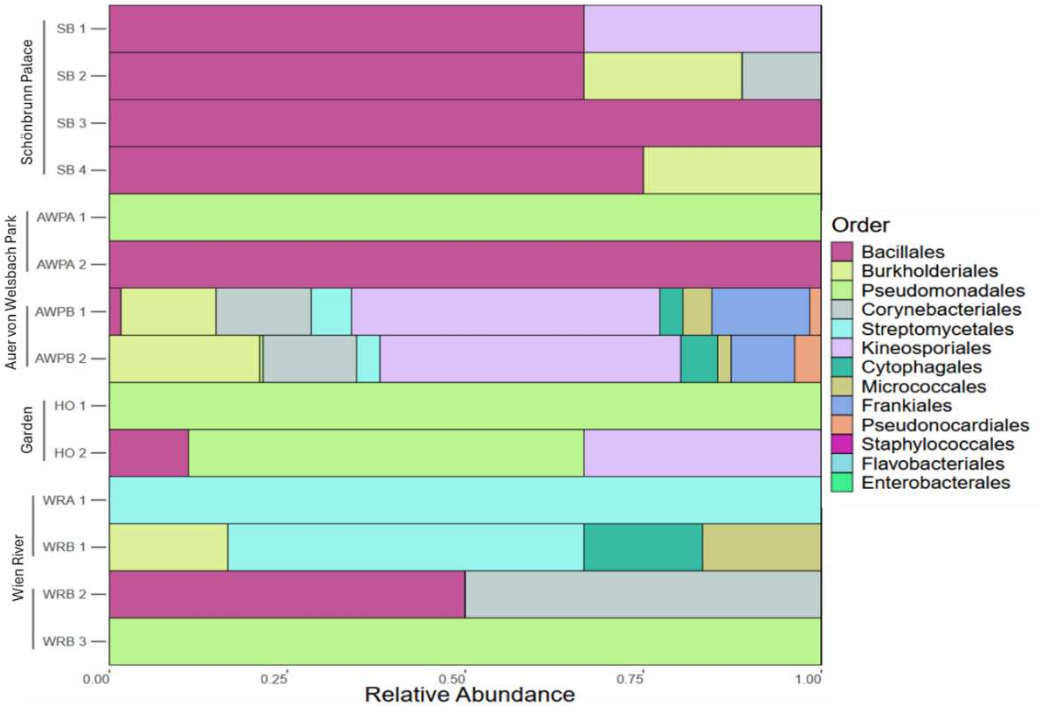

Figure 32: Relative abundance of nodule bacterial community of the city samples based on 16S-rRNA amplicon sequencing with *Rhizobiales* removed. Each bar represents one nodule.

6.3.4.2. Soil

The bacterial community did show high diversity and differences in abundance according to the sampling locations, with an average presence of 6.2% *Rhizobiales.* The soil sample taken at the Wien River (SWRA) showed a higher bacterial diversity compared to the other five locations (Figure 33).

Among the bacteria, carrying the *nifH* gene, the relative abundance of *Rhizobiales* varied across different soil samples from 8.9% to 47.1%. The samples collected from the Wien River, which had the lowest abundance of *Rhizobiales,* had a higher prevalence of *Nostocales* (70.5%) (Figure 34). Six *Rhizobiales* ASVs were found in both soil and plant nodules. ASV_kvg_1lq was the most prevalent, present in both. Certain ASVs appeared only in nodules or soil. Overall ASV richness was lower in soil compared to the nodules (Table S 5).

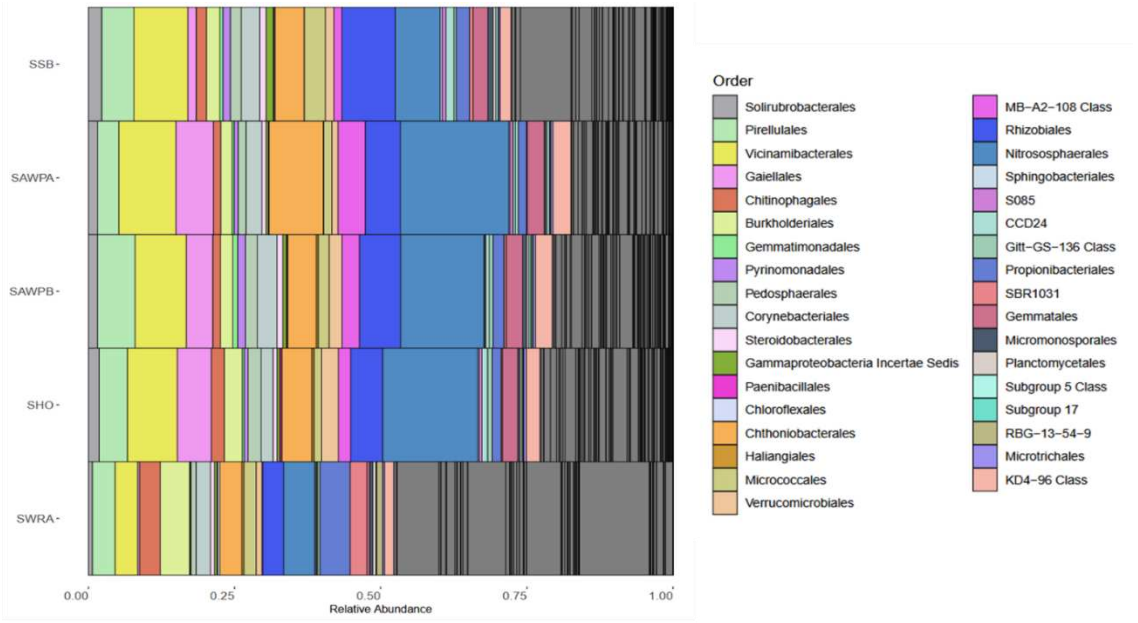

Figure 33: Relative abundance of soil bacterial community in the city experiment, based on 16S rRNA V4 amplicons

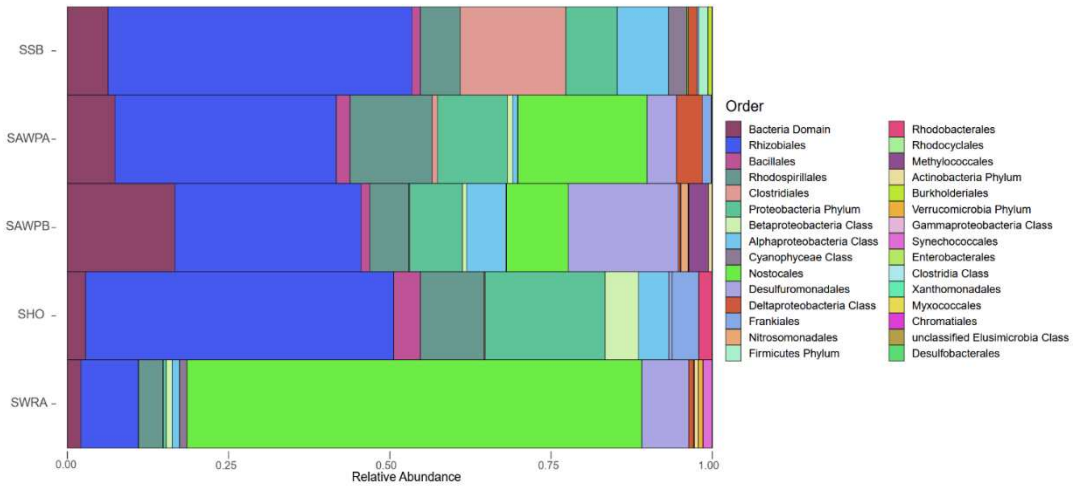

Figure 34: Relative abundance of soil bacterial community in the city experiment*, based on nifH* amplicons.

6.4.dPCR

The total number of bacterial cells in nodules could only be calculated for the two varieties *T. pratense* var *Global and T. pratense* var *Reichersberger.* The count of 16S rRNA ranged from $2.7x10^{7}$ to $1.7x10^{12}$ per gram nodule tissue with an average of $3.28x10^{10}$ copies of 16S-rRNA per gram nodule tissue.

The analysis showed minimal variation in 16S rRNA counts between the two varieties and across the timepoints. For *T. pratense* var *Global,* the average was $8.82x10^{10}$ and $7.8x10^{10}$ copies per gram nodule tissue, respectively, for the first and second sampling timepoint. For *T. pratense* var. *Reichersberger* the average was slightly lower with $5.23x10^{10}$ and $3.72x10^{10}$ copies per gram nodule tissue, respectively, for the first and second sampling timepoint

Based on the assumption that >99% of nodule inhabitants were *R. leguminosarum* and carry three copies of the 16S rRNA gene, there was an average of $1.09x10^{10}$ bacterial cells per gram of nodule tissue.

For soil the total number of the 16S rRNA gene ranged from $3.28x10^{8}$ to $1.29x10^{11}$ with an average of 5.65×10^{9} copies. The number of copies was not affected by plant variety and showed no significant difference between rhizoplane and rhizosphere soil.

7. Discussion

Rhizobia and NABs are important for plant growth and development. While their presence is understood, little is known about the variation in nodule microbial communities within individual plants or whether nodule size influences the presence of NABs and the number of bacterial cells present in root nodules.

This study provides a better understanding of bacterial communities in individual legume nodules, and how soil bacterial communities serve as source for bacteria infecting legume roots. By sampling nodules from three different clover varieties grown in the same soil, I investigated whether the clover variety had an impact on community composition. I further investigated whether nodule communities changed over time during plant development by sampling during the growth phase, as well as at the flowering stage. A small number of clover plants were also collected in the city of Vienna and investigated to provide information about nodule communities in urban natural settings.

7.1. Bacterial diversity

I found a high abundance of *Rhizobiales* in all nodules, with an average relative abundance of 96.1% in all 203 sampled nodules. Previous studies found 60-99% of the bacterial community within legume root nodules to be *Rhizobiales* (Mayhood & Mirza, 2021; Sharaf et al., 2019). Consequently, analyzing nodule associated bacteria (NABs) presented a challenge by focusing on a relatively small subset of the microbial community.

Bacterial diversity showed small alpha diversity values for both nodule datasets, with a Chao1 index ranging from 10 to 20 in the university experiment and from 11 to 36 in the city experiment. The Shannon index exhibited a similar trend, ranging from 0.13 to 0.29 in the university experiment and from 0.1 to 0.85 in the city experiment. These findings suggest a limited bacterial composition within clover nodules, which is not unexpected given the dominance of rhizobia. Beta diversity analysis of the nodule samples showed no clustering of samples based on plant varieties or sampling timepoints for the university experiment (Figure 7), suggesting the plants

harbored similar bacterial species (PERMANOVA, p = 0.09, p = 0.14). However, PERMANOVA indicated a significant difference in community composition when comparing the plants from the university and the city sampling (p = 0.041) suggesting that location-specific factors influenced the microbial community.

Previous studies by Mayhood and Mirza (2021) reported low beta diversity and a Shannon diversity index ranging from 2.03 to 2.83 in soybean root nodules (Mayhood & Mirza, 2021), indicating higher bacterial diversity in soybean nodules compared to clover. Further trends of higher bacterial diversity were found during the comparison of three legumes; *Glycine max*, *Glycine soja* and *Sesbania cannabina,* with a mean Shannon diversity ranging from 5.28 to 5.78, and a mean Chao1 richness ranging from 112 to 284 (Zheng et al., 2020).

In comparison, alpha diversity in the soil samples was high, with a Chao1 index ranging from 1320 to 1523 in the university experiment and 807 to 1111 in the city experiment. The Shannon index exhibited a similar trend, ranging from 5.9 to 6.4 in the university experiment and from 5.5 to 6.1 in the city experiment. There was little variation between soil from the three *Trifolium* varieties since they were potted in soil from the same location. Bray-Curtis NMDS analysis, with clustering around 0, indicated low beta diversity. This finding indicates that the microbial composition among the diverse soil samples is the same. Furthermore, analysis of rhizoplane soil samples using PERMANOVA found that both the sampling timepoint and clover species significantly influenced bacterial community composition (p = 0.002, p=0.003).

Greater diversity of NAB ASVs was observed for the city samples compared to the university experiment for 16S rRNA. However, these ASVs were not detected in the respective soil samples, due to the effects of prevalence filtering. In the unfiltered data, all the present ASVs could be detected, but in small abundance, making them more susceptible to being filtered out during data normalization.

With dPCR, I calculated the amount of the bacterial cell count, yielding an average of $3.28x10^{10}$ copies of 16S rRNA per gram of nodule tissue. A study performed by Zheng et al. (2020) showed lower values in other legumes, with quantitative PCR of 16S rRNA. The average values obtained for *Glycine max*, *Glycine soja* and *Sesbania cannabina,* were $5.6x10^5$, $4.2x10^5$ and $2.4x10^5$ cells per gram sample, respectively (Zheng et al., 2020).

The potassium and phosphorus contents of the soil samples decreased during the growth period as the plant took up the nutrients. This did not influence the bacterial community in the soil and the nodules.

7.2. *Rhizobium leguminosarum*

The main bacterial symbiont found for the potting experiment was *R. leguminosarum*. Their consistent high abundance in all three *Trifolium* variety root nodules suggests that the clover variety did not influence the selection of rhizobia under consistent experimental conditions, even though other possible symbionts like *Mesorhizobium, Sinorhizobium* sp. or *Bradyrhizobium* sp. were present in the soil. The most abundant *R. leguminosarum* ASV, ASV_kvg_1lq, was present in all plants, and the top four *nifH* ASVs were consistently abundant in the soil, suggesting that the most prevalent soil rhizobia infected the nodules.

For the city samples, *R. leguminosarum* was detected as the dominant rhizobial symbiont in all samples as well. However, 14 out of 16 of the city sample nodules analyzed with *nifH* were co-infected with both *Rhizobium* and *Sinorhizobium*. Three different rhizobial species were observed in the city samples: *Rhizobium leguminosarum, Rhizobium anhuiense,* and *Sinorhizobium meliloti*, highlighting the complexity of the nodulation process and plant-bacteria crosstalk, allowing for the simultaneous presence of different species. This co-infection might portray a more natural setting, as the plants were collected from different sampling sites. Mixed rhizobial infections could be influenced by environmental factors, such as soil nutrients, microbial competition, or by plant-related factors, such as plant age. Previous studies have shown that co-infection by different rhizobial species within a single plant and nodule is possible. For example, a study of pea nodules revealed the

presence of up to six distinct rhizobial strains within a single nodule (Mendoza-Suárez et al., 2020).

Overall, 36.3% of the sampled nodules from both experiments were inhabited exclusively by *Rhizobiales*. In the city experiment, all nodules also had NABs, while in the university experiment, nodules containing only *Rhizobiales* increased from 31.25% to 43.69% at the second timepoint. This suggests plants shape their microbiome during growth to meet nutrient needs. The nodules inhabited only by *Rhizobiales* may be inhabited by more competitive strains that may have outcompeted the other bacteria in the given environmental conditions. It has been shown that competitiveness is influenced by different abiotic factors such as soil pH or nutrient availability, affecting the symbiosis and its involved microorganisms (Ledermann et al., 2021). Whether or not the nodules were colonized by NABs did not correlate with nodule weight.

R. leguminosarum was present in all samples, regardless of soil location, including both the Mank field and the six city sites, suggesting that the location did not significantly affect the dominant symbiont. In other legumes, such as soybean, it was also shown that the geographical location of the sampled plants did not influence the selection of the preferred rhizobial symbiont (Mayhood & Mirza, 2021). However, Fields et al. (2023) found significant changes in the bacterial composition, depending on the white clover (*Trifolium repens*) variety. They also found that local growth conditions, such as soil type and nutrient availability, played a crucial role in the selection of rhizobial endophytes (Fields et al., 2023). Interestingly, we observed the presence of other rhizobial species, like *Bradyrhizobium* sp.*, Sinorhizobium meliloti* and *Mesorhizobium* sp., in the soil that have not been taken up into the nodules.

Bacterial communities differed between rhizoplane and rhizosphere samples, with *Rhizobiales* averaging 24% in the rhizosphere compared to an average of 61.9% in the rhizoplane. Three of the five present *R. leguminosarum* ASVs from the nodules were also present in high abundance in the rhizosphere samples. In the rhizoplane, *R. leguminosarum* was enriched, while other *Rhizobiales*, like *Bradyrhizobium* sp.*, Sinorhizobium meliloti,* and *Mesorhizobium* sp., were

present in lower abundances. This change suggests that *R. leguminosarum* was not only the preferred endosymbiont in the root nodules but may have been selectively favored by the plant. Thus, increasing its abundance in the rhizoplane. These findings are consistent with previous studies by Mayhood and Mirza (2021) in soybean. They showed that the soy plant favors *Bradyrhizobium japonicum* as an endosymbiont and enriches it in the rhizoplane (Mayhood & Mirza, 2021). It is important to note that we did collect the rhizoplane only for the second sampling timepoint. No comparison between the different timepoints can be made, and thus, it cannot be said if the community composition of the rhizoplane shifted between the initial nodulation process (first timepoint) and the well-established nodules at the second timepoint.

7.3. Nodule associated bacteria

The most common NAB in the nodules was shown to be *Bacillales*, which were present in 45.4% of all the NAB containing nodules across all varieties and replicate plants. The most abundant *Bacillales* ASVs found were also present in the soil samples suggesting that the most abundant *Bacillales* likely invaded the nodules from the soil alongside the nodule-inducing *Rhizobiales*.

Bacillus sp. was previously identified in root nodules of different legumes, and some studies suggest it is the most abundant NAB genus (Hnini & Aurag, 2024; Pang et al., 2021; Xu et al., 2014). It was also shown that *Bacillus* might have a positive effect on plant growth by suppressing plant pathogens and expressing anti-fungal compounds (Dhole & Shelat, 2022; Hnini & Aurag, 2024). However, possible negative impacts of the presence of *Bacillales* and *Pseudomonadales* on the growth of *Medicago truncatula,* as well as on the number of nodules were shown as well (Kosmopoulos et al., 2024). Thus, *Bacillales* may enhance plant growth and stress resistance but can also potentially hinder plant growth.

Corynebacteriales were detected as NAB in 19% of NAB containing nodules across all clover varieties. *Corynebacteriales* were previously found in other legumes, such as in buckthorn nodules (Liu et al., 2024). *Pseudomonas* sp. were frequently detected as NAB in soybean (Mayhood & Mirza, 2021) and they

were also found during cultivation studies of clover root nodules (Sturz et al., 1997).

Some NABs like *Streptomycetales, Staphyloccocales, Kinesiosporales, Cytophagales* could be detected in low abundance in a small number of nodules. The low abundance and inconsistent presence of these genera in root nodules within the same plant suggest that the host plant does not actively choose these endophytes. *Streptomyces* were found to have antifungal and plant growth enhancement abilities in the root nodules (Martínez-Hidalgo & Hirsch, 2017). This indicates that the NABs present in small abundances, might still have an impact on plant growth and survival rate.

Interestingly, *Enterobacterales* were identified in 18 out of 19 sampled nodules from two plants taken at the second sampling timepoint of *T. pratense* var. *Reichersberger*, but in none of the other samples, showing a plant-selective preference of NAB infection. Interestingly, the corresponding *Enterobacterales* ASVs were not detected in the respective soil. *Enterobacterales* were found to be present in the root nodules of chickpea, where they are thought to enhance nodulation (Martínez-Hidalgo & Hirsch, 2017).

7.4. Limitations

This study captures the diversity of individual nodules across different developmental stages of three clover varieties grown in farmers' soil. This, however, represents only a small portion of the natural diversity, as clover plants in different agricultural settings and soils might host different symbionts.

The sample size of the city samples was insufficient to capture the natural diversity of clover nodules in an urban environment. Most of the clover plants sampled in the city had no nodules. Consequently, only one plant was collected from each sampling site, limiting the number of nodules available for analysis. I suggest a follow-up experiment with a larger number of plants and nodules collected from different sites across an urban gradient.

The results of the rhizoplane and rhizosphere samples should be treated with caution, as sampling method differed between the two timepoints due to the wetness of the soil, and the rhizoplane was only sampled in the second

timepoint.

The sequencing depth for the 16S rRNA data from the nodules is low, with 2000 reads. This may impact the accuracy of the observed community composition, as the normalization could have potentially filtered out ASVs, which might be present at low abundance in the nodules. To achieve a more in-depth analysis of the nodule associated microbiome, the use of a peptide nucleic acid (PNA) clamp designed to target rhizobia (Straub, January 2025, personal communication) should be considered, to increase the sequencing depth of non-rhizobia.

Despite introducing a potential bias, I decided to do normalization, as rarefaction is still thought to be the most effective method for data comparison. It is important to note that rarefaction comes with certain limitations, as it might remove bacterial species, present in a low abundance. Nevertheless, it remains the best approach to ensure comparability of the data (Schloss, 2024). With testing different parameters for the normalization process, I tried to ensure that NABs were accurately represented by minimizing the risk of unintentionally removing present bacterial taxa.

For the analysis of dPCR data, I calculated the results based on three copies of the 16S rRNA gene for the *R. leguminosarum* genome. However, it is likely that the *R. leguminosarum* bacteroids within the symbiosomes replicate their DNA without dividing. This phenomenon is thought to occur in *Sinorhizobium meliloti*, which can increase their DNA content more than 24-fold compared to their free-living counterparts (Jones et al., 2007; Kondorosi et al., 2013). Consequently, an increase in the expression of the 16S rRNA gene for *R. leguminosarum* in nodules is possible. The cells may carry more than the usual three copies, therefore, depending on the 16S rRNA expression, the bacterial number could be overestimated.

8. Conclusion and Perspectives

In summary, I performed the first within-plant analysis for individual root nodules of different *Trifolium* sp. varieties. With amplicon sequencing of 16S rRNA and *nifH*, detecting the bacterial community composition and abundance in individual nodules was possible. The microbiome was primarily composed of *Rhizobium leguminosarum*, averaging 96.1% of all amplicon reads. The primary NAB detected, amongst a diverse range of bacteria, were *Bacillales*. The bacterial community composition within the nodules was compared to that of the soil community, showing that the most abundant ASVs in the soil were also the ones found in the nodules for both rhizobia and NABs. I have also shown plant-selective colonization by NAB, where *Enterobacterales* were found 95% of nodules from two plants at a single timepoint, but in none of the other sampled plants.

Further experiments must be conducted to understand the NAB community in *Trifolium* and its impact on plant growth. Since clover is a perennial plant, the timeframe of this experiment could be extended to conclude whether nodule bacterial communities markedly differ between different nodule generations, possibly due to environmental conditions, or whether the plant selects the same rhizobia and NAB. This would help understand how the NAB community changes during the natural aging process of the plant and between nodule generations, with a typical nodule life span being 10 to 12 weeks. Additionally, this study could be expanded to include soils with different nitrogen, pH and water concentrations, as well as bacterial compositions, enabling the assessment of NAB influence across different environmental conditions.

Metagenomic, proteomic and transcriptomic analyses could be conducted to investigate the function, role and impact of NABs further. These studies would offer insights into NABs' metabolic activity, allowing the analysis of their impact on plant health and development.

9. Supplementary Data

9.1. Nodule sampling

Table S 1: Sampling location and GPS coordinates for the six city samples

Sample	Sampling location	
AWPA	Auer von Welsbach Park A	48°11'23.1"N 16°18'58.6"E
AWPB	Auer von Welsbach Park B	48°11'20.5"N 16°19'02.4"E
SB	Schönbrunn	48°11'14.5"N 16°18'29.3"E
HO	Garten	48°11'38.4"N 16°17'29.3"E
WRA	Wien Fluss A	48°11'21.5"N 16°17'50.1"E
WRB	Wien Fluss B	48°12'04.7"N 16°15'01.4"E

9.2. DNA extraction

Correlation between nodule weight and extracted DNA amount:

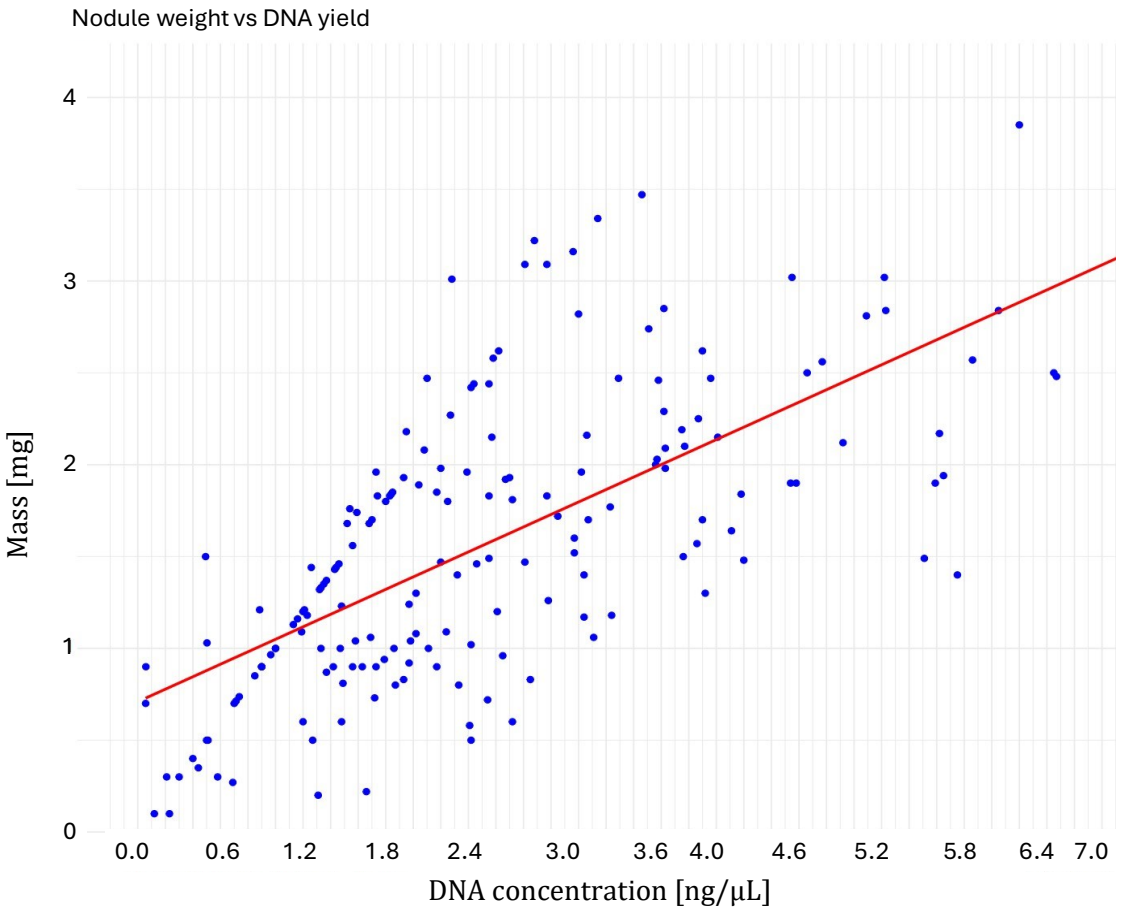

Figure S 1: correlation of nodule weight and extracted DNA. Red: trendline.

9.3. Rarefaction of amplicon sequencing data

9.3.1. 16S rRNA Nodules

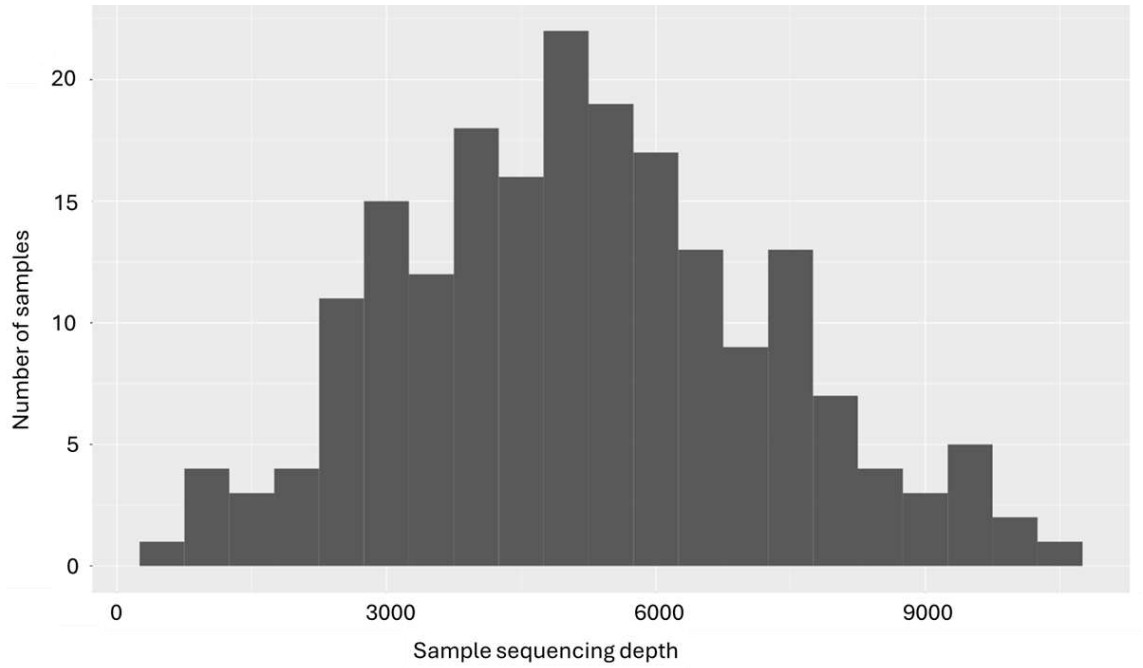

Figure S 2: Sequencing depth of 16S rRNA amplicon sequencing of the V4 region of the nodule DNA.

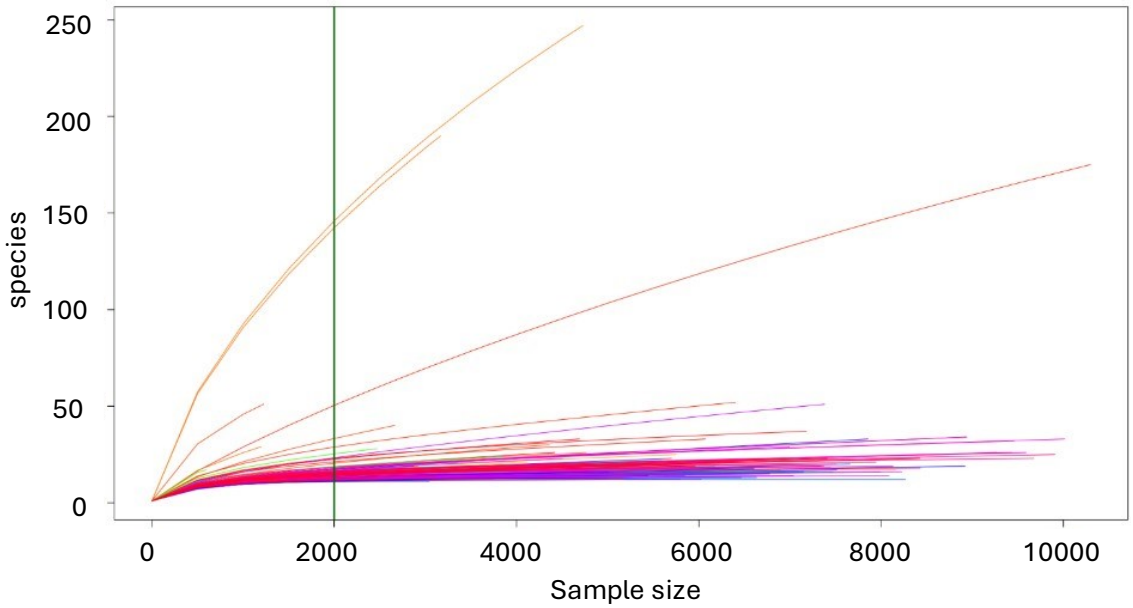

Figure S 3: Rarefaction curve of 16S rRNA amplicon sequencing of the V4 region of the nodule DNA. Rarefaction was performed to 2000, indicated by the vertical green line.

9.3.2. *nifH* Nodules

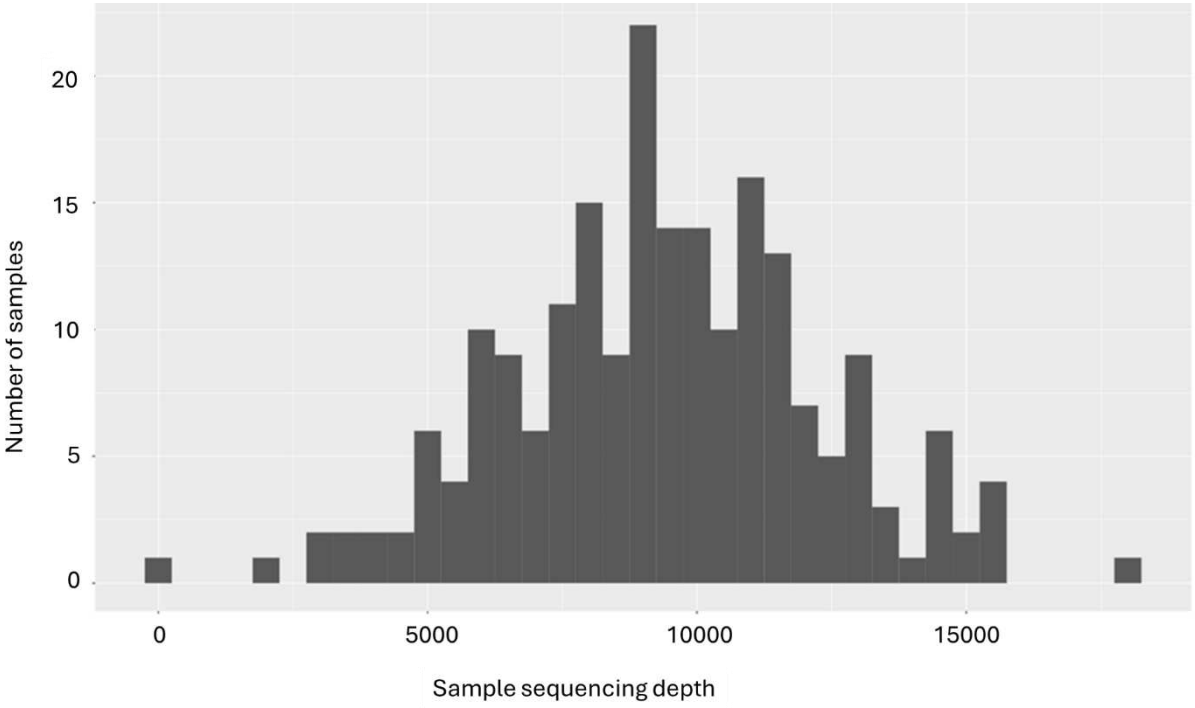

Figure S 4: Sequencing depth of *nifH* amplicon sequencing of the nodule DNA

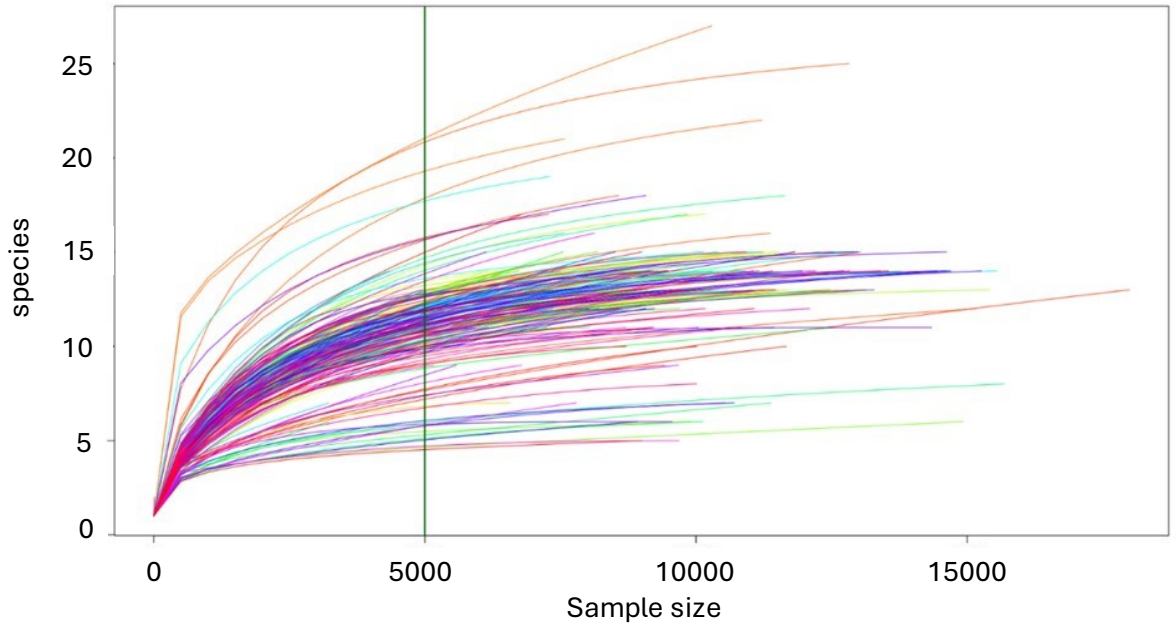

Figure S 5: Rarefaction curve of *nifH* amplicon sequencing of the nodule DNA. Rarefaction was performed to 5000, indicated by the vertical green line.

9.3.3. 16S rRNA Soil

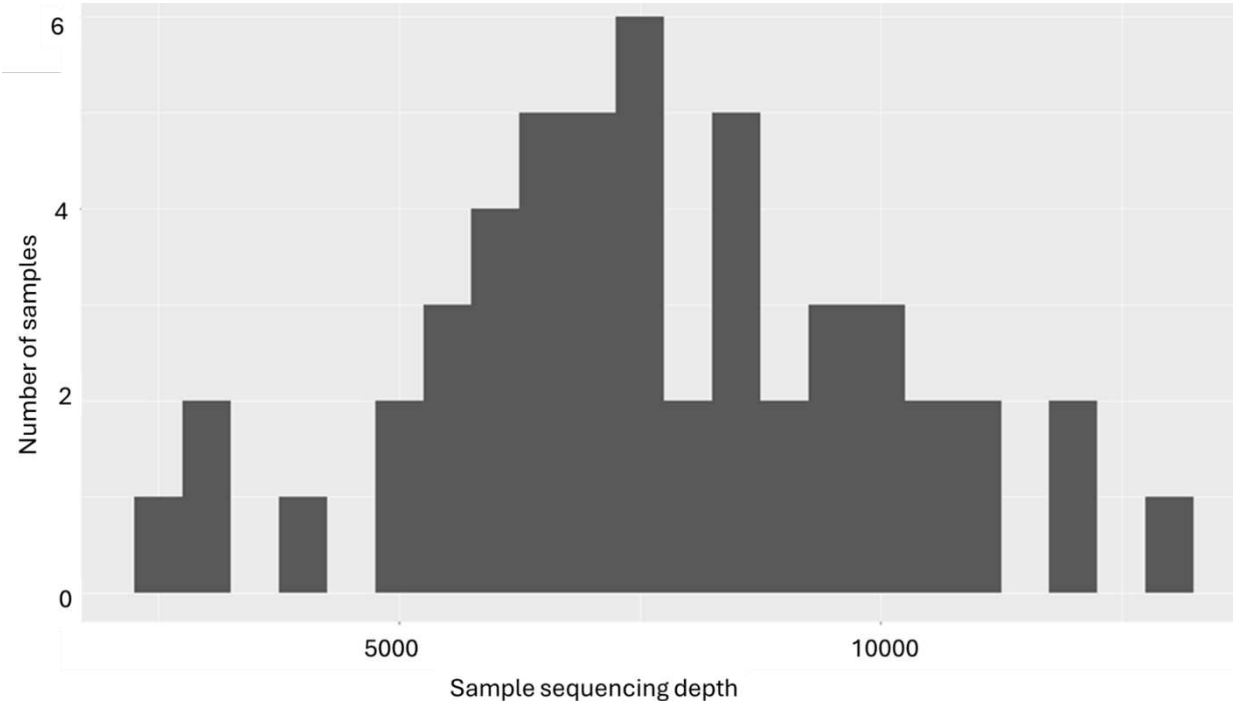

Figure S 7: Sequencing depth of 16SrRNA amplicon sequencing of the V4 region of the soil DNA

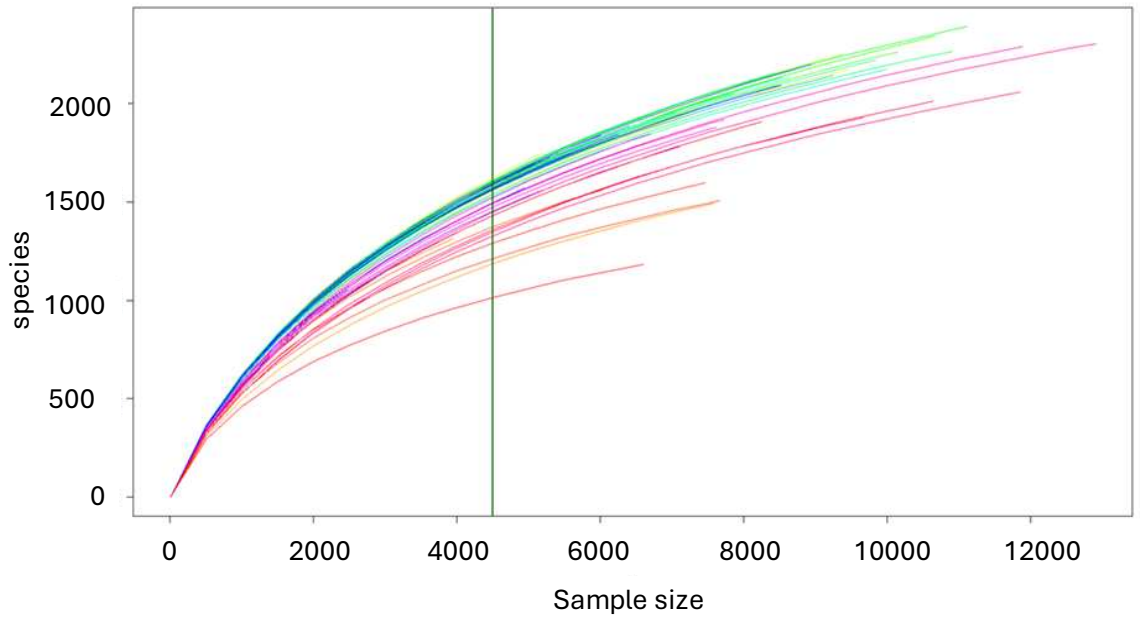

Figure S 6: Rarefaction curve of 16S rRNA amplicon sequencing of the V4 region of the soil DNA.
Rarefaction was performed to 4500, indicated by the vertical green line.

9.3.4. *nifH* Soil

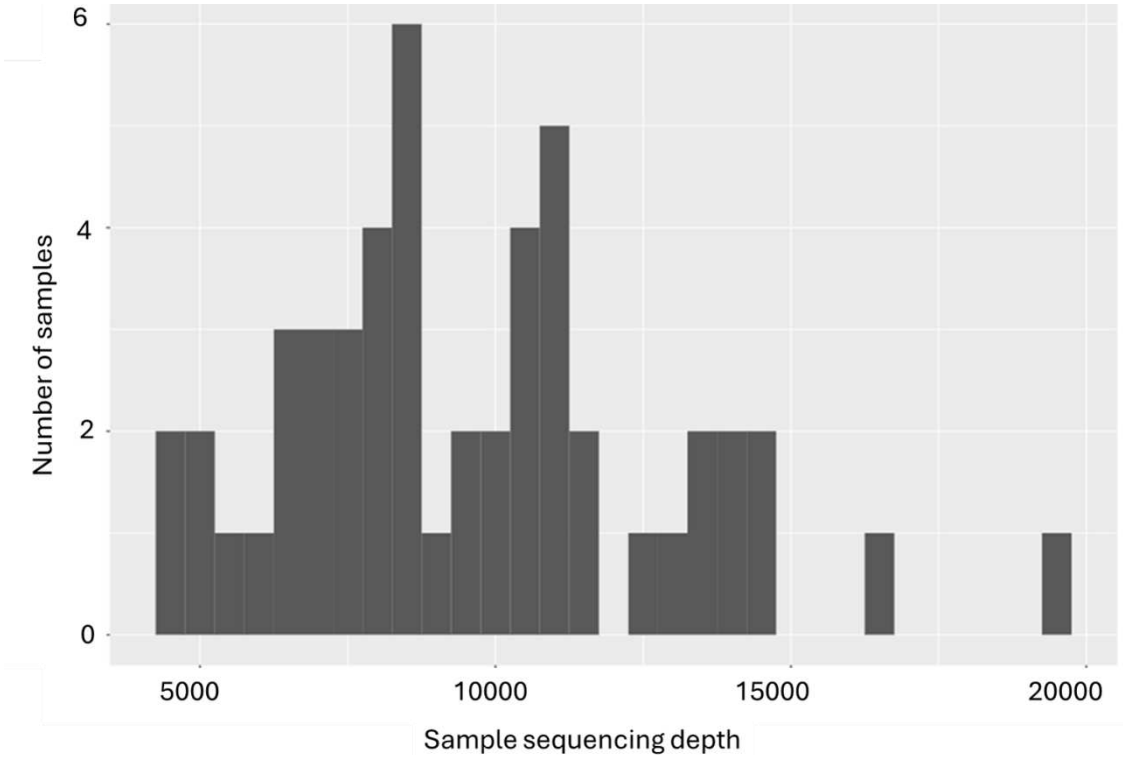

Figure S 8: Sequencing depth of *nifH* amplicon sequencing of the soil DNA

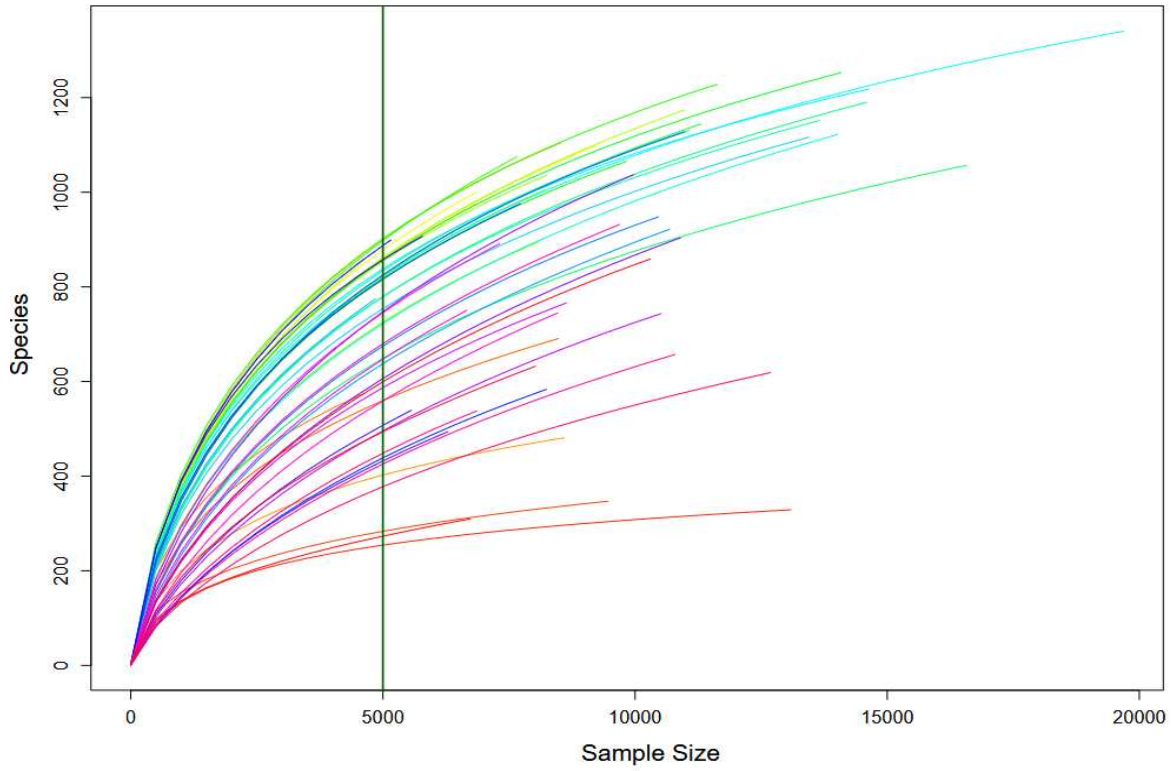

Figure S 9: Rarefaction curve of *nifH* amplicon sequencing of the soil DNA. Rarefaction was performed to 5000, indicated by the vertical green line.

9.4.ASV counts

9.4.1.16S rRNA amplicon sequencing

Table S 2: ASV counts for the top 7 *Bacillales* ASVs in nodule and soil samples.

		ASV count/ clover variety		
		nodules		
ASV	Bacterial order	*T. pratense* var. *Global*	*T. incarnatum*	*T. pratense* var. *Reichersberger*
ASV_792_kv9_a7d	*Bacillales*	7	1	4
ASV_cmz_ztb_ud6	*Bacillales*	3		39
ASV_fhc_0jv_2xd	*Bacillales*	12		9
ASV_kjo_ml7_s23	*Bacillales*	12	4	14
ASV_mtb_0m6_wuz	*Bacillales*	1		1
ASV_rjw_hvp_67t	*Bacillales*	5		21
ASV_4z5_fws_llw	*Bacillales*			
		soil		
ASV	Bacterial order	*T. pratense* var. *Global*	*T. incarnatum*	*T. pratense var. Reichersberger*
ASV_792_kv9_a7d	*Bacillales*	59	62	31
ASV_cmz_ztb_ud6	*Bacillales*	200	167	123
ASV_fhc_0jv_2xd	*Bacillales*	409	392	237
ASV_kjo_ml7_s23	*Bacillales*	298	370	200
ASV_mtb_0m6_wuz	*Bacillales*	224	291	154
ASV_rjw_hvp_67t	*Bacillales*	216	234	130
ASV_4z5_fws_llw	*Bacillales*	181	191	96

Table S 3: ASV counts for the top NAB 16S rRNA ASVs from the nodules from the city experiment

		ASV count/sampling location				
ASV	Bacterial species	SB	AWPA	AWPB	HO	WRB
ASV_1k9_aqq_7z6	*Burkholderiales*	4		43		
ASV_792_kv9_a7d	*Bacillales*	4				
ASV_fhc_0jv_2xd	*Bacillales*	12				
ASV_rjw_hvp_67t	*Bacillales*	2				
ASV_17k_yb8_exf	*Cytophagales*			19		
ASV_7g4_0zz_tzn	*Frankiales*			53		
ASV_7yo_o9d_hjn	*Corynebacteriales*			19		
ASV_amg_agi_h37	*Streptomycetales*			20	3	9
ASV_qyw_8gg_4z6	*Kineosporiales*			197		
ASV_rde_k4b_z48	*Corynebacteriales*			31		
ASV_cme_yc0_c7z	*Pseudomonadales*		10		1439	10
ASV_kjo_ml7_s23	*Bacillales*		1			

9.4.2. *nifH* amplicon sequencing

Table S 4: ASV counts for the top 6 rhizobia *nifH* ASVs from the three clover varieties in nodules and soil.

		ASV count/ clover variety		
		nodules		
ASV	Bacterial order	*T. pratense var. Global*	*T. incarnatum*	*T. pratense var. Reichersberger*
ASV_1x8_n4v	*Rhizobium leguminosarum*	406	973	19340
ASV_kew_w1i	*Rhizobium leguminosarum*	34482	172	14845
ASV_kvg_1lq	*Rhizobium leguminosarum*	397511	46695	344169
ASV_oig_afn	*Rhizobium leguminosarum*	9738	860	
ASV_pm5_ku4	*Rhizobium leguminosarum*	5636	659	4915
ASV_uy4_m28	*Rhizobium leguminosarum*	4907		
		soil		
ASV	Bacterial order	*T. pratense var. Global*	*T. incarnatum*	*T. pratense var. Reichersberger*
ASV_1x8_n4v	*Rhizobium leguminosarum*	180		162
ASV_kew_w1i	*Rhizobium leguminosarum*	579	391	527
ASV_kvg_1lq	*Rhizobium leguminosarum*	23294	14151	24673
ASV_oig_afn	*Rhizobium leguminosarum*	1926	1075	1926
ASV_pm5_ku4	*Rhizobium leguminosarum*			
ASV_uy4_m28	*Rhizobium leguminosarum*			

Table S 5: ASV counts for the top 7 rhizobial *nifH* ASVs from the city experiment in nodules and soil.

		ASV count/sampling location					
		Nodules					
ASV	Bacterial species	SB	AWPA	AWPB	HO	WRA	WRB
ASV_6qz_kz2	*Rhizobium anhuiense*	4888					
ASV_kvg_1lq	*Rhizobium leguminosarum*	4901	9768	2314	9800	3	4875
ASV_ljo_5f7	*Rhizobium leguminosarum*	9827					
ASV_tli_pvo	*Rhizobium leguminosarum*			6433			4924
ASV_oig_afn	*Rhizobium leguminosarum*					9828	
ASV_eho_sfq	*Sinorhizobium meliloti*			306			
ASV_74f_2k7	*Sinorhizobium meliloti*			61			
		Soil					
ASV	Bacterial species	SB	AWPA	AWPB	HO	WRA	WRB
ASV_6qz_kz2	*Rhizobium anhuiense*	236					
ASV_kvg_1lq	*Rhizobium leguminosarum*	187	1195	237	1258	279	
ASV_ljo_5f7	*Rhizobium leguminosarum*	224					
ASV_tli_pvo	*Rhizobium leguminosarum*		134				
ASV_oig_afn	*Rhizobium leguminosarum*			115			
ASV_eho_sfq	*Sinorhizobium meliloti*	573					
ASV_74f_2k7	*Sinorhizobium meliloti*						

9.5.Protocol for DNeasy Powersoil Pro Kit

1. Add up to 200 mg of soil or crushed nodules and 800 µl of Solution CD1
2. Vortex briefly to mix
3. Vortex at maximum speed for 10 min
4. the PowerBead Pro Tube at 15,000 x g for 1 min
5. Transfer the supernatant to a clean 2 ml Microcentrifuge Tube
6. Add 200 µl of Solution CD2 and vortex for 5 s
7. Centrifuge at 15,000 x g for 1 min. Transfer 700 µl of supernatant to a clean 2 ml Microcentrifuge Tube
8. Add 600 µl of Solution CD3 and vortex for 5 s
9. Load 650 µl of lysate to an MB Spin Column. Centrifuge at 15,000 x g for 1 min
10. Discard the flow-through and repeat step 8 to ensure that all of the lysate has passed through the MB Spin Column
11. Carefully place the MB Spin Column into a clean 2 ml Collection Tube.
12. Add 500 µl of Solution EA to the MB Spin Column. Centrifuge at 15,000 x g for 1 min
13. Discard the flow-through and place the MB Spin Column back into the same 2 ml Collection Tube
14. Add 500 µl of Solution C5 to the MB Spin Column. Centrifuge at 15,000 x g for 1 min
15. Discard the flow-through and place the MB Spin Column into a new 2 ml Collection Tube
16. Centrifuge at up to 16,000 x g for 2 min. Carefully place the MB Spin Column into a new 1.5 ml Elution Tube
17. Add 50–100 µl of Solution C6 to the center of the white filter membrane
18. Centrifuge at 15,000 x g for 1 min
19. Store the extracted DNA at -80°C

9.6.Sample data

Table S 6: Sample data for the nodule samples

sample ID	Plant variety	Experiment	collection date 2024	DNA concentration [ng/µL]	mass [mg]	Copies of 16S rRNA/gram nodule tissue
WRA 1	*T. pratense*	Wien Fluss	9.07.	4.68		
WRA 2	*T. pratense*	Wien Fluss	9.07.	1.65		
AWPB 1	*T. pratense*	Auer von Welsbach-Park	9.07.	1.41		
AWPB 2	*T. pratense*	Auer von Welsbach-Park	9.07.	1.17		
HO 1	*T. pratense*	Garten	9.07.	1.67		
HO 2	*T. pratense*	Garten	9.07.	6.12		
WRB 1	*T. pratense*	Wien Fluss	9.07.	1.66		
WRB 2	*T. pratense*	Wien Fluss	9.07.	2.88		
WRB 3	*T. pratense*	Wien Fluss	9.07.	3.27		
AWPA 1	*T. pratense*	Auer von Welsbach-Park	9.07.	1.39		
AWPA 2	*T. pratense*	Auer von Welsbach-Park	9.07.	1.53		
AWPA 3	*T. pratense*	Auer von Welsbach-Park	9.07.	1.04		
SB 1	*T. pratense*	Schönbrunn	9.07.	5.62		
SB 2	*T. pratense*	Schönbrunn	9.07.	3.22		
SB 3	*T. pratense*	Schönbrunn	9.07.	3.71		

SB 4	*T. pratense*	Schönbrunn	9.07.	3.49		
R1P1N1	*T. pratense* var. *Reichersberger*	University	12.08.	4.1	1.7	7,45E+10
R1P1N2	*T. pratense* var. *Reichersberger*	University	12.08.	2.33	0.8	4,85E+10
R1P1N9	*T. pratense* var. *Reichersberger*	University	12.08.	1.87	0.8	6,28E+10
R1P1N3	*T. pratense* var. *Reichersberger*	University	12.08.	1.47	1	2,29E+10
R1P1N4	*T. pratense* var. *Reichersberger*	University	12.08.	2.42	0.5	1,09E+11
R1P1N5	*T. pratense* var. *Reichersberger*	University	12.08.	1.48	0.6	9,18E+10
R1P1N6	*T. pratense* var. *Reichersberger*	University	12.08.	1.31	0.2	1,57E+11
R1P1N7	*T. pratense* var. *Reichersberger*	University	12.08.	1.27	0.5	6,46E+10
R1P1N8	*T. pratense* var. *Reichersberger*	University	12.08.	2.17	0.9	3,20E+10
R1P1N10	*T. pratense* var. *Reichersberger*	University	12.08.	1.2	0.6	3,88E+10
R1P2N1	*T. pratense* var. *Reichersberger*	University	12.08.	0.059	0.9	
R1P2N2	*T. pratense* var. *Reichersberger*	University	12.08.	/	0.5	
R1P2N3	*T. pratense* var. *Reichersberger*	University	12.08.	0.057	0.7	
R1P2N4	*T. pratense* var. *Reichersberger*	University	12.08.	/	0.5	
R1P2N5	*T. pratense* var. *Reichersberger*	University	12.08.	/	0.6	
R1P3N1	*T. pratense* var. *Reichersberger*	University	12.08.	3.96	1.5	3,73E+10
R1P3N2	*T. pratense* var. *Reichersberger*	University	12.08.	2.61	1.2	6,71E+10
R1P3N3	*T. pratense* var. *Reichersberger*	University	12.08.	2.02	1.3	2,65E+10

R1P3N4	*T. pratense* var. *Reichersberger*	University	12.08.	1.33	1	4,31E+10
R1P3N5	*T. pratense* var. *Reichersberger*	University	12.08.	1.63	0.9	5,44E+10
R1P3N6	*T. pratense* var. *Reichersberger*	University	12.08.	1.73	0.9	4,38E+10
R1P3N7	*T. pratense* var. *Reichersberger*	University	12.08.	1.86	1	3,17E+10
R1P3N8	*T. pratense* var. *Reichersberger*	University	12.08.	1.56	0.9	7,99E+09
R1P3N9	*T. pratense* var. *Reichersberger*	University	12.08.	2.11	1	5,51E+10
R1P3N10	*T. pratense* var. *Reichersberger*	University	12.08.	1.42	0.9	
R1P4N1	*T. pratense* var. *Reichersberger*	University	12.08.	1.85	0.9	6,46E+10
R1P4N2	*T. pratense* var. *Reichersberger*	University	12.08.	0.718	1	2,50E+10
R1P4N3	*T. pratense* var. *Reichersberger*	University	12.08.	1.04	0.9	4,65E+10
R1P4N4	*T. pratense* var. *Reichersberger*	University	12.08.	1.17	1	2,43E+10
R1P4N5	*T. pratense* var. *Reichersberger*	University	12.08.	0.678	0.9	3,63E+10
R1P4N6	*T. pratense* var. *Reichersberger*	University	12.08.	1.26	1	8,08E+10
R1P4N7	*T. pratense* var. *Reichersberger*	University	12.08.	1.22	0.3	6,46E+10
R1P4N8	*T. pratense* var. *Reichersberger*	University	12.08.	1.06	0.7	3,06E+10
R1P4N9	*T. pratense* var. *Reichersberger*	University	12.08.	1.14	0.5	2,17E+10
R1P4N10	*T. pratense* var. *Reichersberger*	University	12.08.	2.51	0.4	5,26E+10
G1P2N1	*T. pratense* var. *Global*	University	12.08.	0.69	0.27	6,88E+10

G1P2N2	*T. pratense* var. *Global*	University	12.08.	0.58	0.3	6,32E+09
G1P2N3	*T. pratense* var. *Global*	University	12.08.	0.51	0.5	1,04E+10
G1P2N4	*T. pratense* var. *Global*	University	12.08.	0.12	0.1	1,47E+10
G1P2N5	*T. pratense* var. *Global*	University	12.08.	0.23	0.1	4,23E+10
G1P2N6	*T. pratense* var. *Global*	University	12.08.	0.21	0.3	1,07E+10
G1P2N7	*T. pratense* var. *Global*	University	12.08.	/	0.2	8,86E+08
G1P5N1	*T. pratense* var. *Global*	University	12.08.	4.78	1.9	6,61E+10
G1P5N2	*T. pratense* var. *Global*	University	12.08.	3.27	1.7	5,76E+10
G1P5N3	*T. pratense* var. *Global*	University	12.08.	6.65	2.5	3,66E+10
G1P5N4	*T. pratense* var. *Global*	University	12.08.	2.32	1.4	3,34E+10
G1P5N5	*T. pratense* var. *Global*	University	12.08.	0.493	1.5	7,56E+11
G1P5N6	*T. pratense* var. *Global*	University	12.08.	4.12	1.3	5,63E+10
G1P5N7	*T. pratense* var. *Global*	University	12.08.	4.86	2.5	4,03E+10
G1P5N8	*T. pratense* var. *Global*	University	12.08.	4.74	1.9	3,40E+10
G1P5N9	*T. pratense* var. *Global*	University	12.08.	3.17	1.6	1,64E+10
G1P5N10	*T. pratense* var. *Global*	University	12.08.	2.25	1.8	2,03E+10
G1P3N1	*T. pratense* var. *Global*	University	12.08.	0.737	0.2	
G1P3N2	*T. pratense* var. *Global*	University	12.08.	1.35	0.4	9,62E+09

G1P3N3	*T. pratense* var. *Global*	University	12.08.	2.27	1	1,17E+11
G1P3N4	*T. pratense* var. *Global*	University	12.08.	1.84	0.7	6,80E+10
G1P3N5	*T. pratense* var. *Global*	University	12.08.	1.43	0.7	2,53E+10
G1P3N6	*T. pratense* var. *Global*	University	12.08.	0.85	0.6	1,25E+10
G1P3N7	*T. pratense* var. *Global*	University	12.08.	1.32	0.2	2,95E+10
G1P3N8	*T. pratense* var. *Global*	University	12.08.	1.93	0.5	5,14E+10
G1P3N9	*T. pratense* var. *Global*	University	12.08.	0.713	0.3	
G1P3N10	*T. pratense* var. *Global*	University	12.08.	1.8	0.6	1,28E+11
B1P2	*T. incarnatum*	University	12.08.	1.36		
B1P3	*T. incarnatum*	University	12.08.	1.27		
B1P4	*T. incarnatum*	University	12.08.	1.4		
B1P5	*T. incarnatum*	University	12.08.	14.5		
B1P1	*T. incarnatum*	University	12.08.	2.37		
G1P4N1	*T. pratense* var. *Global*	University	12.08.	2.42	1.1	1,02E+11
G1P4N2	*T. pratense* var. *Global*	University	12.08.	1.7	0.4	1,66E+11
G1P4N3	*T. pratense* var. *Global*	University	12.08.	3.16	1.2	9,14E+10
G1P4N4	*T. pratense* var. *Global*	University	12.08.	2.08	1.1	6,54E+10
G1P4N5	*T. pratense* var. *Global*	University	12.08.	2.62	1	9,76E+10
G1P4N6	*T. pratense* var. *Global*	University	12.08.	3.34	0.8	2,56E+11

G1P4N7	*T. pratense* var. *Global*	University	12.08.	2.44	0.6	6,83E+10
G1P4N8	*T. pratense* var. *Global*	University	12.08.	1.44	0.1	2,35E+11
G1P4N9	*T. pratense* var. *Global*	University	12.08.	1.16	0.1	1,51E+11
G1P4N10	*T. pratense* var. *Global*	University	12.08.	1.68	0.6	5,38E+10
G1P1N1	*T. pratense* var. *Global*	University	12.08.	1.13	0.32	7,59E+10
G1P1N2	*T. pratense* var. *Global*	University	12.08.	1.21	0.13	1,63E+11
G1P1N3	*T. pratense* var. *Global*	University	12.08.	2.58	0.25	1,31E+11
G1P1N4	*T. pratense* var. *Global*	University	12.08.	0.965	0.38	
G1P1N5	*T. pratense* var. *Global*	University	12.08.	1.46	0.11	2,34E+11
G1P1N6	*T. pratense* var. *Global*	University	12.08.	1.85	0.43	6,02E+10
G1P1N7	*T. pratense* var. *Global*	University	12.08.	1.33	0.26	7,44E+10
G1P1N8	*T. pratense* var. *Global*	University	12.08.	1.56	0.94	1,82E+10
G1P1N9	*T. pratense* var. *Global*	University	12.08.	1.83	0.37	4,03E+10
G1P1N10	*T. pratense* var. *Global*	University	12.08.	1.37	0.35	8,68E+10
R2P1N1	*T. pratense* var. *Global*	University	09.09.	1.79	0.94	2,70E+10
R2P1N2	*T. pratense* var. *Reichersberger*	University	09.09.	2.02	1.08	3,04E+10
R2P1N3	*T. pratense* var. *Reichersberger*	University	09.09.	2.46	1.46	1,11E+10
R2P1N4	*T. pratense* var. *Reichersberger*	University	09.09.	2.98	1.26	2,74E+10

R2P1N5	*T. pratense* var. *Reichersberger*	University	09.09.	2.41	0.58	6,42E+10
R2P1N6	*T. pratense* var. *Reichersberger*	University	09.09.	2.1	2.47	7,68E+09
R2P1N7	*T. pratense* var. *Reichersberger*	University	09.09.	1.66	0.22	9,46E+10
R2P1N8	*T. pratense* var. *Reichersberger*	University	09.09.	1.98	1.04	2,06E+10
R2P1N9	*T. pratense* var. *Reichersberger*	University	09.09.	1.49	0.81	3,33E+10
R2P1N10	*T. pratense* var. *Reichersberger*	University	09.09.	1.26	1.44	1,91E+10
R2P2N1	*T. pratense* var. *Reichersberger*	University	09.09.	2.97	3.09	7,74E+09
R2P2N2	*T. pratense* var. *Reichersberger*	University	09.09.	3.44	1.18	6,38E+10
R2P2N3	*T. pratense* var. *Reichersberger*	University	09.09.	2.88	3.22	1,37E+10
R2P2N4	*T. pratense* var. *Reichersberger*	University	09.09.	2.28	3.01	6,38E+09
R2P2N5	*T. pratense* var. *Reichersberger*	University	09.09.	6.25	2.84	5,41E+10
R2P2N6	*T. pratense* var. *Reichersberger*	University	09.09.	3.31	1.06	8,99E+10
R2P2N7	*T. pratense* var. *Reichersberger*	University	09.09.	4.97	2.56	1,90E+10
R2P2N8	*T. pratense* var. *Reichersberger*	University	09.09.	3.83	1.98	4,35E+10
R2P2N9	*T. pratense* var. *Reichersberger*	University	09.09.	1.59	1.74	8,65E+09
R2P2N10	*T. pratense* var. *Reichersberger*	University	09.09.	3.83	2.09	4,41E+10
R2P3N1	*T. pratense* var. *Reichersberger*	University	09.09.	3.97	2.1	3,48E+10
R2P3N2	*T. pratense* var. *Reichersberger*	University	09.09.	5.43	2.84	2,90E+10

R2P3N3	*T. pratense* var. *Reichersberger*	University	09.09.	5.71	1.49	1,00E+11
R2P3N4	*T. pratense* var. *Reichersberger*	University	09.09.	4.21	2.15	4,48E+10
R2P3N5	*T. pratense* var. *Reichersberger*	University	09.09.	3.22	1.96	5,59E+10
R2P3N6	*T. pratense* var. *Reichersberger*	University	09.09.	3.78	2.46	1,13E+10
R2P3N7	*T. pratense* var. *Reichersberger*	University	09.09.	2.55	2.44	3,51E+10
R2P3N8	*T. pratense* var. *Reichersberger*	University	09.09.	4.4	1.48	5,84E+10
R2P3N9	*T. pratense* var. *Reichersberger*	University	09.09.	4.31	1.64	2,14E+10
R2P3N10	*T. pratense* var. *Reichersberger*	University	09.09.	5.85	1.94	6,26E+10
R2P4N1	*T. pratense* var. *Reichersberger*	University	09.09.	3.66	3.47	2,89E+10
R2P4N2	*T. pratense* var. *Reichersberger*	University	09.09.	1.97	0.92	5,20E+10
R2P4N3	*T. pratense* var. *Reichersberger*	University	09.09.	2.67	1.92	4,71E+10
R2P4N4	*T. pratense* var. *Reichersberger*	University	09.09.	3.76	2	6,31E+10
R2P4N5	*T. pratense* var. *Reichersberger*	University	09.09.	4.38	1.84	6,22E+10
R2P4N6	*T. pratense* var. *Reichersberger*	University	09.09.	2.55	1.83	1,20E+10
R2P4N7	*T. pratense* var. *Reichersberger*	University	09.09.	3.05	1.72	1,41E+10
R2P4N8	*T. pratense* var. *Reichersberger*	University	09.09.	2.42	1.02	6,12E+10
R2P4N9	*T. pratense* var. *Reichersberger*	University	09.09.	3.17	1.52	4,52E+10
R2P4N10	*T. pratense* var. *Reichersberger*	University	09.09.	2.2	1.47	3,23E+10

R2P5N1	*T. pratense* var. *Reichersberger*	University	09.09.	2.72	1.81	2,61E+10
R2P5N2	*T. pratense* var. *Reichersberger*	University	09.09.	2.72	0.6	1,15E+09
R2P5N3	*T. pratense* var. *Reichersberger*	University	09.09.	1.48	1.23	5,40E+10
R2P5N4	*T. pratense* var. *Reichersberger*	University	09.09.	3.24	1.17	5,55E+10
R2P5N5	*T. pratense* var. *Reichersberger*	University	09.09.	2.85	0.83	1,67E+10
R2P5N6	*T. pratense* var. *Reichersberger*	University	09.09.	0.502	1.03	2,70E+07
R2P5N7	*T. pratense* var. *Reichersberger*	University	09.09.	2.17	1.85	4,90E+10
R2P5N8	*T. pratense* var. *Reichersberger*	University	09.09.	1.58	1.04	3,48E+10
R2P5N9	*T. pratense* var. *Reichersberger*	University	09.09.	1.72	0.73	2,90E+10
B2P1	T. incarnatum	University	09.09.	9.29		
B2P2	T. incarnatum	University	09.09.	6.39		
B2P3	T. incarnatum	University	09.09.	11.4		
B2P4	T. incarnatum	University	09.09.	8.28		
B2P5	T. incarnatum	University	09.09.	7.11		
G2P1N1	*T. pratense* var. *Global*	University	09.09.	5.29	2.81	4,43E+10
G2P1N2	*T. pratense* var. *Global*	University	09.09.	4.1	2.62	4,04E+10
G2P1N3	*T. pratense* var. *Global*	University	09.09.	1.23	1.18	1,11E+10
G2P1N4	*T. pratense* var. *Global*	University	09.09.	5.79	1.9	6,99E+10
G2P1N5	*T. pratense* var. *Global*	University	09.09.	1.37	0.87	1,49E+10
G2P1N6	*T. pratense* var. *Global*	University	09.09.	2.54	0.72	7,95E+10

G2P1N7	*T. pratense* var. *Global*	University	09.09.	4.75	3.02	3,77E+10
G2P1N8	*T. pratense* var. *Global*	University	09.09.	2.04	1.89	8,67E+09
G2P1N9	*T. pratense* var. *Global*	University	09.09.	2.65	0.96	2,05E+10
G2P1N10	*T. pratense* var. *Global*	University	09.09.	2.97	1.83	2,24E+10
G2P2N1	*T. pratense* var. *Global*	University	09.09.	2.24	1.09	2,49E+10
G2P2N2	*T. pratense* var. *Global*	University	09.09.	1.54	1.76	1,52E+10
G2P2N3	*T. pratense* var. *Global*	University	09.09.	0.885	1.21	1,77E+10
G2P2N4	*T. pratense* var. *Global*	University	09.09.	2.81	1.47	2,51E+10
G2P2N5	*T. pratense* var. *Global*	University	09.09.	0.44	0.35	4,99E+09
G2P2N6	*T. pratense* var. *Global*	University	09.09.	2.55	1.49	5,62E+10
G2P2N7	*T. pratense* var. *Global*	University	09.09.	1.95	2.18	1,68E+09
G2P2N8	*T. pratense* var. *Global*	University	09.09.	1.69	1.06	2,62E+10
G2P2N9	*T. pratense* var. *Global*	University	09.09.	1.52	1.68	
G2P2N10	*T. pratense* var. *Global*	University	09.09.	1.97	1.24	1,65E+10
G2P3N1	*T. pratense* var. *Global*	University	09.09.	3.95	2.19	
G2P3N2	*T. pratense* var. *Global*	University	09.09.	2.7	1.93	2,78E+10
G2P3N3	*T. pratense* var. *Global*	University	09.09.	3.24	1.4	4,22E+10
G2P3N4	*T. pratense* var. *Global*	University	09.09.	3.82	2.85	2,85E+10

G2P3N5	*T. pratense* var. *Global*	University	09.09.	1.74	1.83	2,87E+10
G2P3N6	*T. pratense* var. *Global*	University	09.09.	5.12	2.12	2,62E+10
G2P3N7	*T. pratense* var. *Global*	University	09.09.	4.16	2.47	3,14E+10
G2P3N8	*T. pratense* var. *Global*	University	09.09.	1.93	0.83	3,57E+10
G2P3N9	*T. pratense* var. *Global*	University	09.09.	2.2	1.98	2,22E+10
G2P3N10	*T. pratense* var. *Global*	University	09.09.	1.19	1.09	
G2P4N1	*T. pratense* var. *Global*	University	09.09.	3.43	1.77	5,23E+10
G2P4N2	*T. pratense* var. *Global*	University	09.09.	3.71	2.74	1,64E+10
G2P4N3	*T. pratense* var. *Global*	University	09.09.	2.81	3.09	2,11E+10
G2P4N4	*T. pratense* var. *Global*	University	09.09.	3.26	2.16	
G2P4N5	*T. pratense* var. *Global*	University	09.09.	2.57	2.15	
G2P4N6	*T. pratense* var. *Global*	University	09.09.	4.06	1.57	8,88E+09
G2P4N7	*T. pratense* var. *Global*	University	09.09.	3.2	2.82	1,85E+10
G2P4N8	*T. pratense* var. *Global*	University	09.09.	2.39	1.96	1,13E+10
G2P4N9	*T. pratense* var. *Global*	University	09.09.	3.49	2.47	3,84E+10
G2P4N10	*T. pratense* var. *Global*	University	09.09.	3.82	2.29	2,77E+10
G2P5N1	*T. pratense* var. *Global*	University	09.09.	4.07	2.25	6,35E+10
G2P5N2	*T. pratense* var. *Global*	University	09.09.	6.4	3.85	2,40E+10

G2P5N3	*T. pratense* var. *Global*	University	09.09.	1.73	1.96	1,45E+10
G2P5N4	*T. pratense* var. *Global*	University	09.09.	9.96	3.99	4,28E+10
G2P5N5	*T. pratense* var. *Global*	University	09.09.	5.95	1.4	1,01E+11
G2P5N6	*T. pratense* var. *Global*	University	09.09.	3.77	2.03	1,92E+10
G2P5N7	*T. pratense* var. *Global*	University	09.09.	5.42	3.02	1,70E+12
G2P5N8	*T. pratense* var. *Global*	University	09.09.	6.67	2.48	1,06E+11
G2P5N9	*T. pratense* var. *Global*	University	09.09.	5.82	2.17	4,55E+11
G2P5N10	*T. pratense* var. *Global*	University	09.09.	6.06	2.57	1,06E+10

Table S 7: Sample data for the soil samples

sample ID	Plant variety	Experiment	collection date 2024	DNA concentration [ng/µL]	mass [mg]	Copies of 16S rRNA/gram soil
SWRA	*T. pratense*	Wien Fluss	9.07.	268		9,16E+09
SWRB	*T. pratense*	Wien Fluss	9.07.	253		4,27E+09
SAWPA	*T. pratense*	Auer von Welsbach-Park	9.07.	184		8,52E+09
SAWPB	*T. pratense*	Auer von Welsbach-Park	9.07.	131		4,20E+09
SHO	*T. pratense*	Garten	9.07.	366		2,90E+09
SSB	*T. pratense*	Schönbrunn	9.07.	267		2,99E+09
SB1P1	*T. incarnatum*	University	12.08.	110	199	8,29E+09
SB1P2	*T. incarnatum*	University	12.08.	115	200	1,22E+10
SB1P3	*T. incarnatum*	University	12.08.	74.9	200	
SB1P4	*T. incarnatum*	University	12.08.	88.1	206	
SB1P5	*T. incarnatum*	University	12.08.	114	204	
SR1P1	*T. pratense var. Reichersberger*	University	12.08.	135	200	6,09E+09
SR1P2	*T. pratense var. Reichersberger*	University	12.08.	113	203	
SR1P3	*T. pratense var. Reichersberger*	University	12.08.	132	204	1,77E+09
SR1P4	*T. pratense var. Reichersberger*	University	12.08.	138	202	6,58E+09
SR1P5	*T. pratense var. Reichersberger*	University	12.08.2024	111	201	6,80E+08
SG1P1	*T. pratense var. Global*	University	12.08.	107	208	5,26E+09
SG1P2	*T. pratense var. Global*	University	12.08.	126	201	5,16E+09
SG1P3	*T. pratense var. Global*	University	12.08.	137	214	5,85E+09

SG1P4	*T. pratense var. Global*	University	12.08.	124	207	8,76E+09
SG1P5	*T. pratense var. Global*	University	12.08.	114	200	7,63E+09
SBG2P1	*T. pratense var. Global*	University	09.09.	150	215	6,69E+09
SBG2P2	*T. pratense var. Global*	University	09.09.	123	213	3,63E+09
SBG2P3	*T. pratense var. Global*	University	09.09.	138	208	3,40E+09
SBG2P4	*T. pratense var. Global*	University	09.09.	119	214	5,91E+09
SBG2P5	*T. pratense var. Global*	University	09.09.	112	208	6,27E+09
SBR2P1	*T. pratense var. Reichersberger*	University	09.09.	109	199	4,84E+09
SBR2P2	*T. pratense var. Reichersberger*	University	09.09.	119	207	4,74E+09
SBR2P3	*T. pratense var. Reichersberger*	University	09.09.	134	213	6,52E+09
SBR2P4	*T. pratense var. Reichersberger*	University	09.09.	118	203	6,22E+09
SBR5P5	*T. pratense var. Reichersberger*	University	09.09.	126	203	5,59E+09
SBB2P1	*T. incarnatum*	University	09.09.	113	204	4,29E+09
SBB2P2	*T. incarnatum*	University	09.09.	110	205	7,13E+09
SBB2P3	*T. incarnatum*	University	09.09.	105	201	6,09E+09
SBB2P4	*T. incarnatum*	University	09.09.	146	200	5,53E+09
SBB2P5	*T. incarnatum*	University	09.09.	125	208	3,89E+09
SRR2P1	*T. pratense var. Reichersberger*	University	09.09.	120		4,37E+09
SRR2P2	*T. pratense var. Reichersberger*	University	09.09.	257		3,28E+08
SRR2P3	*T. pratense var. Reichersberger*	University	09.09.	281		1,46E+10
SRR2P4	*T. pratense var. Reichersberger*	University	09.09.	244		4,82E+09

SRR2P5	*T. pratense var. Reichersberger*	University	09.09.	225		5,21E+09
SRB2P1	*T. incarnatum*	University	09.09.	145		3,99E+09
SRB2P2	*T. incarnatum*	University	09.09.	224		2,62E+09
SRB2P3	*T. incarnatum*	University	09.09.	181		5,43E+09
SRB2P4	*T. incarnatum*	University	09.09.	286		5,23E+09
SRB2P5	*T. incarnatum*	University	09.09.	128		9,55E+09
SRG2P1	*T. pratense var. Global*	University	09.09.	226		6,24E+09
SRG2P2	*T. pratense var. Global*	University	09.09.	123		6,41E+09
SRG2P3	*T. pratense var. Global*	University	09.09.	135		1,29E+11
SRG2P4	*T. pratense var. Global*	University	09.09.	119		6,77E+09
SRG2P5	*T. pratense var. Global*	University	09.09.	235		3,29E+09

9.7.R code used for analysis

9.7.1. Create phyloseq object and plot it

```
#loading the required packages
library(microViz)
library("plyr")
library(dplyr)
library(phyloseq)
library(patchwork)
library(ggplot2)
library(tidyr)
library(microbiome)
library(readxl)
library(vegan)
library(tidyverse)
library(readxl)
library(ggtext)
library(gridExtra)
```

```
library(pals)
library(Polychrome)
library(grid)
library(futile.logger)
library(VennDiagram)

#create a phyloseq object
metadata <- read_excel("~/Uni/Masterarbeit/R/metadata.xlsx")
Taxa <- read.delim("~/Uni/Masterarbeit/JMF-2409-10__all__nifH_A__combined/DADA2_ASVs.nifH.moyn413_reference.DADA2_classified.tsv")
ASV_nifH <- read.delim("~/Uni/Masterarbeit/JMF-2409-10__all__nifH_A__combined/DADA2_counts_as_matrix.tsv")
real_IDs <- metadata$my_ID
otumat = as.matrix(ASV_nifH[, -1])
rownames(otumat) <- c(ASV_nifH[, 1])
taxmat = as.matrix(Taxa[, -1])
rownames(taxmat) <- c(Taxa[,1])
taxmat[is.na(taxmat) | taxmat == ""] <- "Unknown"
SAM=sample_data(metadata[,-1])
rownames(SAM) <- c(metadata[, 1])
OTU = otu_table(otumat, taxa_are_rows = TRUE)
TAX = tax_table(taxmat)
physeq = phyloseq(OTU,TAX,SAM)
physeq <- tax_fix(physeq, verbose = FALSE)
physeq
sample_names(physeq)<- real_IDs
#remove controls
ps_filtered <- ps_filter(physeq, type != "control", plant!= "control")

#remove Rhizobia, Chloroplasts and Mitochondria
physeq_f <- subset_taxa( ps_filtered,
 !(Order %in% c("Chloroplast") |
```

```
    Class %in% c("Chloroplast") |
    Genus %in% c("Chloroplast") |
    Phylum %in% c("Chloroplast") |
    Family %in% c("Mitochondria") |
    Order %in% c("Rhizobiales"))
)
# filter for nodules or soil
physeq_n <- physeq_f %>% ps_filter( type == "nodule")
otu_table_matrix <- as(otu_table(physeq_n), "matrix")
sample_sums<- colSums (otu_table_matrix, na.rm = FALSE, dims = 1)

# Compute the average sequencing depth
average_sequencing_depth <- mean(sample_sums)
cat("The average sequencing depth is:", average_sequencing_depth, "\n")

# Plot the sequencing depth distribution
qplot(sample_sums, geom = "histogram", binwidth = 500) +
 xlab("Sample sequencing depth") +
 ylab("Number of samples") +
 theme(
   axis.text.y = element_text(size = 15, hjust = 1),
   axis.text.x = element_text(size = 15, hjust = 1)
 )+ ggtitle("Sequencing depth distribution nifH amplicon sequencing")

# Plot rarefaction curves
if (taxa_are_rows(physeq_n)) {
 otu_table_matrix <- t(otu_table_matrix)
}
rarecurve(otu_table_matrix, step = 500, col =rainbow(nrow(otu_table_matrix)), cex
= 0.2, label = FALSE)
par(cex.axis = 1.2)  # Change the axis text size globally

# Filter taxa based on row sums > 20 takes out 23 taxa
```

```
physeq_filtered_n <- filter_taxa(physeq_n, function(x) sum(x) >= 20, TRUE)

#rarefaction
otu_table_matrix2 <- as(otu_table(physeq_filtered_n), "matrix")
sample_sums<- colSums (otu_table_matrix2, na.rm = FALSE, dims = 1)
set.seed(123)
physeq_rarefied_n<-rarefy_even_depth(physeq_filtered_n, sample.size = 5000,
                        rngseed = 123, replace = FALSE, trimOTUs = TRUE, verbose =
TRUE)
otu_table(physeq_rarefied_n) <- otu_table(as.matrix(otu_table(physeq_rarefied_n)),
taxa_are_rows = TRUE)
physeq_rarefied_n

tax_fix(physeq_rarefied_n, verbose=FALSE)

#plot the relative abundance
nod_ids<-sort(sample_names(physeq_rarefied_n), decreasing=TRUE)

all1 <- physeq_rarefied_n %>%
 subset_samples(sampling == "1")  # Filter by plant_sampling type 1
ida1<-sort(sample_names(all1), decreasing=TRUE)
all1<- all1 %>% comp_barplot(
 tax_level = "Genus",
 n_taxa = 30,
 sample_order = ida1,
 tax_order = prev,
 merge_other = FALSE
) +facet_grid(plant_sampling~sampling ,scales = "free", space = "free")+
 theme(
  axis.text.y = element_text(size = 4, hjust = 1),
  axis.text.x = element_text(size = 4, hjust = 1)
 ) +ggtitle("First Sampling")+
 coord_flip()
```

```
all2 <- physeq_rarefied_n %>%
  subset_samples(sampling == "2")  # Filter by plant_sampling type 2
ida2<-sort(sample_names(all2), decreasing=TRUE)
all2<- all2 %>% comp_barplot(
  tax_level = "Genus",
  n_taxa = 30,
  sample_order = ida2,
  merge_other = FALSE) + facet_grid(plant_sampling~sampling ,scales = "free",
space = "free")+
  theme(
    axis.text.y = element_text(size = 4, hjust = 1),
    axis.text.x = element_text(size = 4, hjust = 1),
    legend.text = element_text(size = 6),
    legend.title = element_text(size = 7),
    legend.key.size = unit(0.3, "cm")
  ) + ggtitle("Secound Sampling")+
  coord_flip()
grid.arrange(all1+ theme(legend.position = "none"), all2, ncol = 2,
          widths = c(2.5,3),
          heights = c(1,1))
```

9.7.2. Beta diversity calculation

```
#select for nodules or soil and sampling timepoint
physeq_rarefied_n <- ps_filter(physeq_nr, , sampling !="city")
physeq_rarefied_n <- physeq_rarefied_s %>%
  subset_samples(sampling == "1")

# Calculate Bray-Curtis distance
dist_bray <- distance(physeq_rarefied_n, method = "bray")

# Perform PCoA
ordination <- ordinate(physeq_rarefied_n, method = "NMDS", distance = dist_bray)
# Plot the ordination
```

```
plot_ordination(physeq_rarefied_n, ordination) +
 geom_point(size = 2) +
 theme_minimal() +
 labs(title = "Beta Diversity (Bray-Curtis NMDS)",
    x = "NMDS1",
    y = "NMDS2")+
 facet_wrap(~type) +
 theme(strip.text = element_text(size = 14, face = "bold"))

#perform permanova analysis
metadata <- as(sample_data(physeq_rarefied_n), "data.frame")
adonis2(dist_bray ~ species, data = metadata, permutations = 999)
```

9.7.3. Alpha diversity calculation

```
#select for nodules or soil and sampling
physeq_rarefied_n <- ps_filter(physeq_rarefied_n, , sampling !="city")
physeq_rarefied_1<- physeq_rarefied_n %>% ps_filter ( sampling=="1")

# Compute Shannon diversity
diversity_data <- estimate_richness(physeq_rarefied_1, measures = "Shannon")
print(diversity_data)

#plot alpha diversity
plot_richness(physeq_rarefied_1)

#plot Shannon diversity
plot_richness(physeq_rarefied_1, x="my_ID", measures=c("Shannon"),
color="species") +
 theme(axis.text.x = element_text(hjust = 1,size=12),
     axis.text.y = element_text(hjust = 1,size=12))

#plot Chao1 diversity
plot_richness(physeq_rarefied_1, x="my_ID", measures=c("Chao1"),
```

```
color="species") +
  theme(axis.text.x = element_text(hjust = 1,size=12),
        axis.text.y = element_text(hjust = 1,size=12))
```

9.7.4. Venn diagram of ASVs

```
# select for plant and sampling timepoint
group1 <- subset_samples(physeq_rarefied_n %>% ps_filter( species == "Reichersberger"), sampling == "1")
group2 <- subset_samples(physeq_rarefied_n %>% ps_filter( species == "Reichersberger"), sampling == "2")

# Extract taxa present in each condition
taxa_group1 <- taxa_names(prune_taxa(taxa_sums(group1) > 0, group1))
taxa_group2 <- taxa_names(prune_taxa(taxa_sums(group2) > 0, group2))

# Create a list of taxa
taxa_list <- list(  R1 = taxa_group1,
          R2 = taxa_group2)

# Generate the Venn diagram
venn.plot <- venn.diagram(
  x = list(taxa_list[[1]], taxa_list[[2]]),
  filename = NULL,  # Set to NULL to display directly in R
  fill = c( "light green", "purple"),
  alpha = 0.6,
  cat.cex = 1.5,
  main = "Taxa Overlap Between Sampling Points"
)
# Plot the Venn diagram
grid.draw(venn.plot)
```

9.7.5. Analyze the presence of specific ASVs

```
#filter for the plants wanted
```

```
ps <- ps_filter(physeq_rarefied_n, , sampling !="city")
taxonomy_table <- tax_table(ps)

taxonomy_table_df <- as.data.frame(taxonomy_table)

#Filter the rows for a specific species
target_species <- "Bacillales"  # Replace with the species name you want
filtered_ASVs <- rownames(taxonomy_table_df[taxonomy_table_df$Order == target_species, ])

view(filtered_ASVs)

#check all ASVs for the sample
p1 <- ps%>%
 subset_samples(sampling == "1")
reich <- p1 %>%
 subset_samples(species == "Reichersberger")

otu_table_data <- otu_table(reich)
otu_table_df <- as.data.frame(otu_table_data)
otu_table_data <- otu_table(reich)
if (taxa_are_rows(reich)) {
 asv_counts <- rowSums(otu_table_data)
} else {
 asv_counts <- colSums(otu_table_data)
}
print(asv_counts)
```

List of Figures:

List of Tables:

10. Disclaimer

During the writing of this thesis, I have used Grammarly, DeepL and OpenAI ChatGPT-4 for grammar correction, translation, and harmonization of certain text passages. Additionally, I used AI as support during data analysis with R Studio.
After using AI, I carefully reviewed and edited the content as necessary. I therefore take full responsibility for the content of this publication.

11. References

Apprill, A., McNally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, *75*(2), 129–137. https://doi.org/10.3354/ame01753

Beijerinck, M. W. (1888). *The root-nodule bacteria.*

Crespi, M., & Gálvez, S. (2000). Molecular Mechanisms in Root Nodule Development. *Journal of Plant Growth Regulation*, *19*(2), 155–166. https://doi.org/10.1007/s003440000023

De Meyer, S. E., Van Hoorde, K., Vekeman, B., Braeckman, T., & Willems, A. (2011). Genetic diversity of rhizobia associated with indigenous legumes in different regions of Flanders (Belgium). *Soil Biology and Biochemistry*, *43*(12), 2384–2396. https://doi.org/10.1016/j.soilbio.2011.08.005

Dhole, A., & Shelat, H. (2022). Non-Rhizobial Endophytes Associated with Nodules of Vigna radiata L. and Their Combined Activity with Rhizobium sp. *Current Microbiology*, *79*(4), 103. https://doi.org/10.1007/s00284-022-02792-x

Dupont, L., Alloing, G., Pierre, O., El, S., Hopkins, J., Hrouart, D., & Frendo, P. (2012). The Legume Root Nodule: From Symbiotic Nitrogen Fixation to Senescence. In T. Nagata (Ed.), *Senescence*. InTech. https://doi.org/10.5772/34438

Etesami, H. (2022). Root nodules of legumes: A suitable ecological niche for isolating non-rhizobial bacteria with biotechnological potential in agriculture. *Current Research in Biotechnology*, *4*, 78–86. https://doi.org/10.1016/j.crbiot.2022.01.003

Fields, B., Moeskjær, S., Deakin, W. J., Moffat, E. K., Roulund, N., Andersen, S. U., Young, J. P. W., & Friman, V.-P. (2023). Rhizobium nodule diversity and composition are influenced by clover host selection and local growth conditions. *Molecular Ecology*, *32*(15), 4259–4277. https://doi.org/10.1111/mec.17028

Goswami, R. K., Mehariya, S., Obulisamy, P. K., & Verma, P. (2021). Advanced microalgae-based renewable biohydrogen production systems: A review. *Bioresource Technology*, *320*, 124301. https://doi.org/10.1016/j.biortech.2020.124301

Harris, C., & Ratnieks, F. L. W. (2022). Clover in agriculture: Combined benefits for bees, environment, and farmer. *Journal of Insect Conservation*, *26*(3), 339–

357. https://doi.org/10.1007/s10841-021-00358-z

Herridge, D. F., Roughley, R. J., & Brockwell, J. (1984). Effect of rhizobia and soil nitrate on the establishment and functioning of the soybean symbiosis in the field. *Australian Journal of Agricultural Research*, *35*(2), 149–161. https://doi.org/10.1071/ar9840149

Hnini, M., & Aurag, J. (2024). Prevalence, diversity and applications potential of nodules endophytic bacteria: A systematic review. *Frontiers in Microbiology*, *15*. https://doi.org/10.3389/fmicb.2024.1386742

Ivanov, S., Fedorova, E., & Bisseling, T. (2010). Intracellular plant microbe associations: Secretory pathways and the formation of perimicrobial compartments. *Current Opinion in Plant Biology*, *13*(4), 372–377. https://doi.org/10.1016/j.pbi.2010.04.005

Jones, D. L., Cross, P., Withers, P. J. A., DeLuca, T. H., Robinson, D. A., Quilliam, R. S., Harris, I. M., Chadwick, D. R., & Edwards-Jones, G. (2013). REVIEW: Nutrient stripping: the global disparity between food security and soil nutrient stocks. *Journal of Applied Ecology*, *50*(4), 851–862. https://doi.org/10.1111/1365-2664.12089

Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E., & Walker, G. C. (2007). How rhizobial symbionts invade plants: The Sinorhizobium-Medicago model. *Nature Reviews. Microbiology*, *5*(8), 619–633. https://doi.org/10.1038/nrmicro1705

Kazmierczak, T., Yang, L., Boncompagni, E., Meilhoc, E., Frugier, F., Frendo, P., Bruand, C., Gruber, V., & Brouquisse, R. (2020). Chapter Seven - Legume nodule senescence: A coordinated death mechanism between bacteria and plant cells. In P. Frendo, F. Frugier, & C. Masson-Boivin (Eds.), *Advances in Botanical Research* (Vol. 94, pp. 181–212). Academic Press. https://doi.org/10.1016/bs.abr.2019.09.013

Kelsey, A., Gano-Cohen, Peter J. Stokes, Mia A. Blanton, Camille E. Wendlandt, Amanda C. Hollowell, John U. Regus, Deborah Kim, Seema Patel, Victor J. Pahua, & Joel L. Sachs. (2016, September). *Nonnodulating Bradyrhizobium spp. Modulate the Benefits of Legume-Rhizobium Mutualism*. *82*(17). https://doi.org/10.1128/AEM.01116-16

Kondorosi, E., Mergaert, P., & Kereszt, A. (2013). A Paradigm for Endosymbiotic Life:

Cell Differentiation of Rhizobium Bacteria Provoked by Host Plant Factors. *Annual Review of Microbiology*, *67*(Volume 67, 2013), 611–628. https://doi.org/10.1146/annurev-micro-092412-155630

Kosmopoulos, J. C., Batstone-Doyle, R. T., & Heath, K. D. (2024). Co-inoculation with novel nodule-inhabiting bacteria reduces the benefits of legume–rhizobium symbiosis. *Canadian Journal of Microbiology*, *70*(7), 275–288. https://doi.org/10.1139/cjm-2023-0209

Larrainzar, E., Villar, I., Rubio, M. C., Pérez-Rontomé, C., Huertas, R., Sato, S., Mun, J.-H., & Becana, M. (2020). Hemoglobins in the legume–Rhizobium symbiosis. *New Phytologist*, *228*(2), 472–484. https://doi.org/10.1111/nph.16673

Ledermann, R., Schulte, C. C. M., & Poole, P. S. (2021). How Rhizobia Adapt to the Nodule Environment. *Journal of Bacteriology*, *203*(12), 10.1128/jb.00539-20. https://doi.org/10.1128/jb.00539-20

Liu, H., Ni, B., Duan, A., He, C., & Zhang, J. (2024). High Frankia abundance and low diversity of microbial community are associated with nodulation specificity and stability of sea buckthorn root nodule. *Frontiers in Plant Science*, *15*. https://doi.org/10.3389/fpls.2024.1301447

Lundberg, D. S., Yourstone, S., Mieczkowski, P., Jones, C. D., & Dangl, J. L. (2013). Practical innovations for high-throughput amplicon sequencing. *Nature Methods*, *10*(10), 999–1002. https://doi.org/10.1038/nmeth.2634

Markl, J., Sadava, D., Hillis, D. M., Heller, H. C., & Hacker, S. D. (2019). Mineralstoffversorgung der Pflanzen. In D. Sadava, D. M. Hillis, H. C. Heller, S. D. Hacker, & J. Markl (Eds.), *Purves Biologie* (pp. 1077–1097). Springer. https://doi.org/10.1007/978-3-662-58172-8_35

Martínez-Hidalgo, P., & Hirsch, A. M. (2017). The Nodule Microbiome: N2-Fixing Rhizobia Do Not Live Alone. *Phytobiomes Journal*, *1*(2), 70–82. https://doi.org/10.1094/PBIOMES-12-16-0019-RVW

Mason, M. L. T., Tabing, B. L. C., Yamamoto, A., & Saeki, Y. (2018). Influence of flooding and soil properties on the genetic diversity and distribution of indigenous soybean-nodulating bradyrhizobia in the Philippines. *Heliyon*, *4*(11). https://doi.org/10.1016/j.heliyon.2018.e00921

Mayhood, P., & Mirza, B. S. (2021). Soybean Root Nodule and Rhizosphere Microbiome: Distribution of Rhizobial and Nonrhizobial Endophytes. *Applied*

and Environmental Microbiology, *87*(10), e02884-20. https://doi.org/10.1128/AEM.02884-20

Mendoza-Suárez, M. A., Geddes, B. A., Sánchez-Cañizares, C., Ramírez-González, R. H., Kirchhelle, C., Jorrin, B., & Poole, P. S. (2020). Optimizing *Rhizobium*-legume symbioses by simultaneous measurement of rhizobial competitiveness and N_2 fixation in nodules. *Proceedings of the National Academy of Sciences*, *117*(18), 9822–9831. https://doi.org/10.1073/pnas.1921225117

Mergaert, P., Van Montagu, M., & Holsters, M. (1997). Molecular mechanisms of Nod factor diversity. *Molecular Microbiology*, *25*(5), 811–817. https://doi.org/10.1111/j.1365-2958.1997.mmi526.x

Moynihan, M. A. (2020). *moyn413/nifHdada2: V1.1.0* (Version v1.1.0) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.4283278

Munaweera, T. I. K., Jayawardana, N. U., Rajaratnam, R., & Dissanayake, N. (2022). Modern plant biotechnology as a strategy in addressing climate change and attaining food security. *Agriculture & Food Security*, *11*(1), 26. https://doi.org/10.1186/s40066-022-00369-2

Mus, F., Colman, D. R., Peters, J. W., & Boyd, E. S. (2019). Geobiological feedbacks, oxygen, and the evolution of nitrogenase. *Free Radical Biology and Medicine*, *140*, 250–259. https://doi.org/10.1016/j.freeradbiomed.2019.01.050

Navascués, J., Pérez-Rontomé, C., Gay, M., Marcos, M., Yang, F., Walker, F. A., Desbois, A., Abián, J., & Becana, M. (2012). Leghemoglobin green derivatives with nitrated hemes evidence production of highly reactive nitrogen species during aging of legume nodules. *Proceedings of the National Academy of Sciences*, *109*(7), 2660–2665. https://doi.org/10.1073/pnas.1116559109

Oldroyd, G. E. D., Murray, J. D., Poole, P. S., & Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annual Review of Genetics*, *45*, 119–144. https://doi.org/10.1146/annurev-genet-110410-132549

Pang, J., Palmer, M., Sun, H. J., Seymour, C. O., Zhang, L., Hedlund, B. P., & Zeng, F. (2021). Diversity of Root Nodule-Associated Bacteria of Diverse Legumes Along an Elevation Gradient in the Kunlun Mountains, China. *Frontiers in Microbiology*, *12*, 633141. https://doi.org/10.3389/fmicb.2021.633141

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock

communities, time series and global field samples. *Environmental Microbiology*, *18*(5), 1403–1414. https://doi.org/10.1111/1462-2920.13023

Sachs, J. L., Ehinger, M. O., & Simms, E. L. (2010). Origins of cheating and loss of symbiosis in wild Bradyrhizobium. *Journal of Evolutionary Biology*, *23*(5), 1075–1089. https://doi.org/10.1111/j.1420-9101.2010.01980.x

Schloss, P. D. (2024). Rarefaction is currently the best approach to control for uneven sequencing effort in amplicon sequence analyses. *mSphere*, *9*(2), e00354-23. https://doi.org/10.1128/msphere.00354-23

Schulz, S., Engel, M., Fischer, D., Buegger, F., Elmer, M., Welzl, G., & Schloter, M. (2013). Diversity pattern of nitrogen fixing microbes in nodules of *Trifolium arvense* (L.) at different initial stages of ecosystem development. *Biogeosciences*, *10*(2), 1183–1192. https://doi.org/10.5194/bg-10-1183-2013

Sharaf, H., Rodrigues, R. R., Moon, J., Zhang, B., Mills, K., & Williams, M. A. (2019). Unprecedented bacterial community richness in soybean nodules vary with cultivar and water status. *Microbiome*, *7*(1), 63. https://doi.org/10.1186/s40168-019-0676-8

Stagnari, F., Maggio, A., Galieni, A., & Pisante, M. (2017). Multiple benefits of legumes for agriculture sustainability: An overview. *Chemical and Biological Technologies in Agriculture*, *4*(1), 2. https://doi.org/10.1186/s40538-016-0085-1

STATISTIK AUSTRIA. (2024, April). *Anbau auf dem Ackerland 2023*. https://www.statistik.at/services/tools/services/publikationen

Stefan, A., Van Cauwenberghe, J., Rosu, C. M., Stedel, C., Labrou, N. E., Flemetakis, E., & Efrose, R. C. (2018). Genetic diversity and structure of *Rhizobium leguminosarum* populations associated with clover plants are influenced by local environmental variables. *Systematic and Applied Microbiology*, *41*(3), 251–259. https://doi.org/10.1016/j.syapm.2018.01.007

Sturz, A. V., Christie, B. R., Matheson, B. G., & Nowak, J. (1997). Biodiversity of endophytic bacteria which colonize red clover nodules, roots, stems and foliage and their influence on host growth. *Biology and Fertility of Soils*, *25*(1), 13–19. https://doi.org/10.1007/s003740050273

Subramanian, P., Kim, K., Krishnamoorthy, R., Sundaram, S., & Sa, T. (2015). Endophytic bacteria improve nodule function and plant nitrogen in soybean on co-inoculation with Bradyrhizobium japonicum MN110. *Plant Growth Regulation*, *76*(3), 327–332. https://doi.org/10.1007/s10725-014-9993-x

Ueda, T., Suga, Y., Yahiro, N., & Matsuguchi, T. (1995). Remarkable N2-fixing bacterial diversity detected in rice roots by molecular evolutionary analysis of nifH gene sequences. *Journal of Bacteriology*, *177*(5), 1414–1417. https://doi.org/10.1128/jb.177.5.1414-1417.1995

Xu, L., Zhang, Y., Wang, L., Chen, W., & Wei, G. (2014). Diversity of endophytic bacteria associated with nodules of two indigenous legumes at different altitudes of the Qilian Mountains in China. *Systematic and Applied Microbiology*, *37*(6), 457–465. https://doi.org/10.1016/j.syapm.2014.05.009

Zakhia, F., & de Lajudie, P. (2001). Taxonomy of rhizobia. *Agronomie*, *21*(6–7), 569–576. https://doi.org/10.1051/agro:2001146

Zgadzaj, R., James, E. K., Kelly, S., Kawaharada, Y., Jonge, N. de, Jensen, D. B., Madsen, L. H., & Radutoiu, S. (2015). A Legume Genetic Framework Controls Infection of Nodules by Symbiotic and Endophytic Bacteria. *PLOS Genetics*, *11*(6), e1005280. https://doi.org/10.1371/journal.pgen.1005280

Zhang, J. J., Jing, X. Y., de Lajudie, P., Ma, C., He, P. X., Singh, R. P., Chen, W. F., & Wang, E. T. (2016). Association of white clover (Trifolium repens L.) with rhizobia of sv. Trifolii belonging to three genomic species in alkaline soils in North and East China. *Plant and Soil*, *407*(1), 417–427. https://doi.org/10.1007/s11104-016-2899-9

Zhang, X., Wu, J., & Kong, Z. (2024). Cellular basis of legume-rhizobium symbiosis. *Plant Communications*, *5*(11), 101045. https://doi.org/10.1016/j.xplc.2024.101045

Zheng, Y., Liang, J., Zhao, D.-L., Meng, C., Xu, Z.-C., Xie, Z.-H., & Zhang, C.-S. (2020). The Root Nodule Microbiome of Cultivated and Wild Halophytic Legumes Showed Similar Diversity but Distinct Community Structure in Yellow River Delta Saline Soils. *Microorganisms*, *8*(2), Article 2. https://doi.org/10.3390/microorganisms8020207