This is a preprint of an article published in Environmental Microbiology. The final authenticated version is available online at: https://doi.org/10.1111/1462-2920.15941

Flemish soils contain rhizobia partners for Northwestern Europe-adapted soybean cultivars

Judith Van Dingenen[1,2#], Sonia Garcia Mendez[1,2,3#], Stien Beirinckx[1,2], Lena Vlaminck[1,2], Annick De Keyser[1,2], Naomi Stuer[1,2], Severine Verschaete[1,2], Alexander Clarysse[1,2], Joke Pannecoucque[4], Stephane Rombauts[1,2], Isabel Roldan-Ruiz[4], Anne Willems[3], and Sofie Goormachtig[1,2*]

ORCID ID: 0000-0001-7887-6000 (J.D.V.), 0000-0002-1140-9237 (S.G.M.), 0000-0002-4830-3225 (S.B.), 0000-0001-7081-2395 (A.D.K.), 0000-0002-0563-6492 (L.V.), 0000-0003-1192-6709 (N.S.), 0000-0003-4916-9416 (A.C.), 0000-0002-5912-1355 (J.P.), 0000-0002-3985-4981 (S.R.), 0000-0001-7340-3386 (I.R.-R.), 0000-0002-8421-2881 (A.W.), 0000-0001-6195-9889 (S.G.)

[1]*Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium.*

[2]*Center for Plant Systems Biology, VIB, 9052 Ghent, Belgium.*

[3]*Department of Biochemistry and Microbiology, Faculty of Sciences, Ghent University, 9000 Ghent, Belgium.*

[4]*Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Plant Sciences Unit, 9820 Merelbeke, Belgium.*

[#]These two authors contributed equally to the manuscript

*For correspondence. E-mail sofie.goormachtig@psb.vib-ugent.be; Tel. +32 9 3313910; Fax +32 9 3313809.

Running title: Soybean-nodulating rhizobia in Flanders

Author contributions

J.V.D., S.G.M., and S.G. designed the experiments, interpreted the data, and wrote the manuscript. S.G.M. performed microbiome analysis, comparative genomics, and ONT analysis. J.V.D., A.D.K., N.S., and S.V. collected soybean and nodule samples. J.V.D., S.G.M., and S.B. performed raw data analysis. S.R. performed ONT sample preparation, sequencing, genome assembly and analysis. J.V.D, S.B., A.D.K., L.V., and A.C. performed nodulation screening assays. J.V.D., S.B., A.D.K., and A.C. performed amplicon sequencing. J.P. provided knowledge on soybean cultivars and the commercial inoculants. I.R.-R., A.W., and S.G. supervised the project. All authors read and revised the manuscript.

Originality-Significance Statement

This study demonstrates the feasibility to detect indigenous rhizobia that nodulate locally adapted early-maturing soybean cultivars in Belgium by nodule trapping. Isolation techniques combined with 16S rRNA PCR and 16S metabarcoding studies, revealed the nodule bacterial diversity, allowing the identification of rhizobial strains most likely responsible for nodule formation. Interestingly, however, by means of ONT sequencing on a single nodule, we generated the genome of a *Bradyrhizobium* species that corresponded to the most abundant ASV detected by metabarcoding, but that differed from the isolated strain belonging to the same ASV, demonstrating an unanticipated rhizobial community inside one nodule. The newly identified rhizobial candidates provide a sound basis to tap from for the development of efficient inoculants to improve local soybean productivity in North-West Europe.

Summary

In Europe, soybean (*Glycine max*) used for food and feed has to be imported, causing negative socioeconomic and environmental impacts. To increase the local production, breeding resulted in varieties that grow in colder climates, but the yield using the commercial inoculants is not satisfactory in Belgium because of variable nodulation efficiencies. To look for indigenous nodulating strains possibly adapted to the local environment, we initiated a nodulation trap by growing early-maturing cultivars under natural and greenhouse conditions in 107 garden soils in Flanders. Nodules occurred in 18 and 21 soils in the garden and greenhouse experiments, respectively. By combining 16S rRNA PCR on single isolates with HiSeq 16S metabarcoding on nodules, we found a large bacterial richness and diversity from different soils. Furthermore, using Oxford Nanopore Technologies sequencing of DNA from one nodule, we retrieved the entire genome of a *Bradyrhizobium* species, not previously isolated, but profusely present in that nodule. These data highlight the need of combining diverse identification techniques to capture the true nodule rhizobial community. Eight selected rhizobial isolates were subdivided by whole-genome analysis in three genera containing six genetically distinct species that, except for two, aligned with known type strains and, were all able to nodulate soybean in the lab.

Introduction

Soybean (*Glycine max*), one of the most important legume crops worldwide, is grown for its protein-rich beans (approximately 40% of the seed) that contain all essential amino acids for human food and animal feed (Hammond *et al.,* 2003). To date, Belgium completely depends on the import of soybean for the production of soy-derived products. On average, 2.5 million tons of soy and soy-derived products, such as soy meal and oil, are imported annually. Almost 65% of the imported soybean originates from high-risk countries, such as Argentina and Brazil, where soybean cultivation results in a land footprint of 2 million hectares with, among others, deforestation and forest degradation as a consequence (Boerema *et al.,* 2016). To reduce this dependency on import and to develop sustainable protein products, Europe, including Belgium, would strongly benefit from an increased local soybean production.

The cultivation of soybean in Northwestern Europe is challenging because the plant is not adapted to cold and wet environments. In the last decade, extensive breeding programs and field trials with hundreds of early-maturing soybean cultivars have been performed to find genotypes suitable for growth in more temperate regions, such as Belgium (Aper *et al.,* 2016; Pannecoucque *et al.,* 2018b) and Germany (Zimmer *et al.,* 2016). As a result, dozens of soybean cultivars, e.g., Primus, Bettina, and Shouna cv., have been bred to be adapted to local environmental conditions, concerning maturity and tolerance to common fungal pathogens in combination with high seed yields and protein content. However, when grown in Belgian reference soils, none of these cultivars meets the bean protein threshold of >42% that has been set by a major Belgian soy processing company to be eligible for human consumption processing (Pannecoucque *et al.,* 2018b).

To guarantee high protein contents in beans, soybean plants must interact symbiotically with nitrogen-fixing bacteria, the so-called rhizobia. This symbiosis results in the formation of

root nodules in which the bacterial strains fix atmospheric nitrogen that is subsequently used for plant protein production. Soybean plants can interact with diverse bacterial genera, such as *Rhizobium*, *Bradyrhizobium, Ensifer (Sinorhizobium)*, *Azorhizobium*, and *(Para)Bulkholderia* as well (Gyaneshwar *et al.,* 2011; Arigas Ramírez *et al.,* 2019; Sharaf *et al.,* 2019). Besides these nodulating bacteria, a variety of other bacterial strains, of which the function is still not entirely clear, are present in soybean nodules (Sharaf *et al.,* 2019; Zhong *et al.,* 2019). Some of these nonrhizobial bacteria have been demonstrated to enhance nitrogen fixation, for instance by increasing the nitrogen availability or improving the presymbiotic steps (Martínez-Hidalgo and Hirsch 2017).

Currently, in North-West Europe, commercially available *Bradyrhizobium* inoculants, such as HiStick® and Biodoz, are used for soybean cultivation, with variable and sometimes unsatisfactory results (Zimmer *et al.,* 2016; Pannecoucque *et al.,* 2018a). These nonindigenous bacterial strains are possibly not adapted to the northern environmental conditions, resulting in insufficient nodulation and subsequently inadequate bean protein contents for the processing industry (Alexandre and Oliveira, 2013; Pannecoucque *et al.,* 2018b). Especially the low temperature of the root zone is a major constraint for the establishment of the legume-rhizobia symbiosis (Zhang *et al.,* 1995; Schmidt *et al.,* 2015). Optimal temperatures for soybean growth and nodule formation are between 25°C and 30°C, which are significantly higher than the Belgian soil temperatures that range from 8°C and 15°C at the sowing time (Miransari *et al.,* 2013). Other environmental parameters strongly influence optimal nodulation as well, i.e. soil type, soil acidity, soil texture, salinity, drought, and the microbial community present in the rhizosphere (Miransari *et al.,* 2013; Liu *et al.,* 2019; Zhong *et al.,* 2019).

In order to find improved soybean nodulators, pre-existing collections of rhizobial strains identified from regions where soybean is endemic could be used, such as the published catalogues from China (Camacho *et al.*, 2002; Thomas-Oates *et al.*, 2003) or Nepal (Adhikari

et al., 2013). As an alternative, and based on the Baas Becking and Beijerinck hypothesis : “everything is everywhere: but the environment selects”, we hypothesized that if we could retrieve indigenous rhizobia adapted to the local Belgian soils and capable of nodulating soybean, we might be able to develop inoculants with less variability in nodulation efficiency. One manner to capture these rhizobia is by so-called ‘trapping’ experiments that have been successfully applied in diverse tropical countries, such as Ethiopia, South-Africa, Afghanistan, and South-America (Naamala *et al.,* 2016; Habibi *et al.,* 2017; Arigas Ramírez *et al.,* 2019; Temesgen *et al.,* 2019; Yuan *et al.,* 2020). Recently the feasibility to identify soybean-nodulating isolates by means of a vermiculite-based system to inoculate soybean plants with local soil samples has been demonstrated in Germany, which has a temperate climate similar to Belgium and also no history of soybean cultivation (Yuan *et al.,* 2020).

Here, we used local climate-adapted soybean varieties to discover native soybean nodulators in Belgian soils. We assigned 107 citizens spread over the northern Belgian region Flanders to perform trapping experiments in their gardens. We did not include agricultural soils as we hypothesized that these soils in Flanders are heavily fertilized compared with garden soils, and the chance to find nodules and thus soybean-nodulating rhizobia would be low or even not existing, as ammonium and nitrate, main fertilizer components, inhibit nodulation formation and biological nitrogen fixation (Yamashita *et al.*, 2019). All volunteers grew three early-maturing soybean varieties (maturity group MG000), i.e. Primus, Shouna, and Bettina cv., in the summer of 2019. In addition, soil samples from these gardens were used to grow the same soybean cultivars under greenhouse growth conditions. The presence of nodules or nodule-like structures was determined in both experiments and both nodules and less developed nodules were collected for further identification and confirmation studies. From a subset of the nodules, bacteria were isolated via classical plating methods and characterized by sequencing the 16S rRNA genes. Additionally, the microbial diversity within the nodules was determined

by means of 16S rRNA amplicon sequencing. Next, the most abundant rhizobia were identified and six candidate rhizobial strains from different experiments, cultivars, and soils were selected, their genomes were compared, and the nodulation capacity was tested on vermiculite-grown soybean plants under greenhouse conditions. Interestingly, using Oxford Nanopore Technologies (ONT), sequencing a crude mixture of plant and imbedded prokaryotic DNA extracted from a single nodule, retrieved the complete genome of the most abundant nodule occupying strain. This strain belonged to another *Bradyrhizobium* species compared to those isolated with the classical techniques.

Results

Selection of 107 gardens with diverse soil types spread over Flanders

To maximize the diversity in soil types, thus increasing the range of rhizobia that could be retrieved (De Meyer *et al.,* 2011), we selected more than 100 gardens of volunteering citizens living in different Flemish regions, Belgium (Fig. 1A). The majority of these gardens (43 out of 107) were located in the province East-Flanders, because participants were recruited via staff of the regional research institutes involved and located in East-Flanders (Fig. 1B). Furthermore, the provinces West-Flanders, Antwerp, and Flemish Brabant were equally represented, whereas the province Limburg and the capital region of Brussels had only very few participating gardens. By consulting the Database Subsurface of Flanders (dov.vlaanderen.be), we could categorize each garden soil to one of the six soil types present in Flanders, namely loam, sand, sandy loam, loamy sand, clay, dune, and anthropogenic (Fig. 1C). However, caution should be taken when analyzing these data, because personal and individual garden soil management and adaptations were not considered. Most garden soils belong to the moderately dry-to-wet sandy

loam or loamy sand soil type (38 out of 107, 36%), followed by sand and loam soils (24 and 18 soils out of 107, 22.4% and 16.8%, respectively). Finally, 19 out of 107 soils (17.8%) are described as anthropogenic or artificial soils in the Database Subsurface of Flanders, meaning that these soils are the result of human activity and do not have a natural origin or with an unknown texture. Finally, four gardens soils were clay and four others were dune soils or polders ("others" in Fig. 1C). Taken together, although the participating gardens were not distributed equally over the different regions in Flanders, six different soil types were represented.

Differences in soybean cultivar-dependent nodulation efficiency between soil types and growth conditions

As both genetic factors and environmental conditions can alter the nodulation efficiency, we performed two trapping experiments in parallel, i.e., in the gardens under environmental conditions and in the greenhouse. Additionally, we included three different soybean varieties as host plants in both experimental growth setups (Supporting Information Fig. S1). These three early-maturing cultivars (Primus, Bettina, and Shouna) exhibit differences in, among others, oil and protein content, thousand kernel weight (weight in grams of 1000 seeds), and disease tolerance (Pannecoucque *et al.,* 2018b) (https://bsl.baes.gv.at/). For the greenhouse experiments, soil samples were transferred from the gardens to the laboratory and three plants of each cultivar were grown in pots. Nodulation on plants of both experiments was assessed after a minimum growth of two months when plants were approximately 60 cm high. At this stage many of the plants were not yet mature or did not yet contain flowers and/or pods.

From the 107 Flemish soils tested, nodules or nodule-like structures on at least one plant (independent of the cultivar) occurred in garden and greenhouse experiments of 18 (or 16.8%)

and 21 (or 19.6%) soils, respectively. For convenience, we refer to all structures as “nodules” although some structures are more nodule-like and might be caused by inefficient bacterial rhizobial infection, underdeveloped cell division or caused by an event that is not linked to rhizobial infection (Hiltenbrand et al., 2016; Rodríguez-Navarro et al., 2011; Fig. 2). Interestingly, only five soils (#19, #43, #67, #76, and #94) had nodules in both experiments (Figure 2, Table 1), but only in soils #19 and #94, nodules occurred for all three soybean cultivars, although the shape and size of the nodules differed strongly between the two experimental setups (Figure 2, Table 1; Supporting Information Fig. S2).

In general, nodulation of the cultivars clearly differed depending on the soil tested. Only in a few soils (7 out of 107, i.e., soils #19, #27, #38, #43, #54, #77, and #94), all three cultivars had nodules per experiment (Figure 2, Table 1). Soybean plants of the Primus cv. clearly outperformed the other two cultivars, with 15 out of 18 (83%) and 17 out of 21 (81%) of the soils, respectively, Primus plants had nodules in both garden as well as in greenhouse experiments. Bettina cv nodulated more frequently under greenhouse conditions, with 14 out of 21 soils (67%), showing nodules, while only in five out of 18 soils (28%), under natural garden conditions, nodules were found on Bettina roots. The opposite was true for Shouna cv., where nodules were detected under greenhouse conditions in only eight out of 21 soils (38%) versus 11 out of the 18 soils (61%) in garden experiments. Most soils that resulted in nodulation in the gardens were sandy (seven out of 18 soils), whereas in the greenhouse experiments they were mostly sandy or sandy/loam (each six out of 21 soils) (Supporting Information Fig. S3A and B).

Nevertheless, a large diversity in nodule size and shape was observed (Table 1; Fig. 2; Supporting Information Fig. S2). The majority of the nodules were small and white/brown, an indication for the absence of nitrogen fixation, irregularly shaped, or located at root branching points (Table 1; Fig. 2). Regular shaped nodules were found in the garden (soils #4, #19, #42,

#43, #50, #64, and #76) and in the greenhouse (soils #12, #19, #43, #60, #67, #70, and #93) experiments (Table 1; Fig. 2; Supporting Information Fig. S2). In three garden (soils #50, #64, and #76) and in four greenhouse (soils #12, #19, #43, and #93) experiments, nodules were seemingly actively fixing nitrogen, as indicated by a light pink to red color (due to the leghaemoglobin). Interestingly, most red nodules were found on the roots of the Primus soybean plants grown in soils #19, #43, #50, #64, and #93, except for soils #12 and #76, where the Shouna cv. carried the fixing nodules. In these soils (#12 and #76) Bettina only had small non-fixing nodules, while Primus showed no nodules. Betina cv. on the other hand, grown in soils #19 and #43, under greenhouse conditions, resulted in large, fixing nodules, whereas the garden-grown nodules were small and non-fixing for all three cultivars.

In conclusion, it is feasible to obtain nodules on soybean plants grown in local Flemish soils both under natural conditions as well as under greenhouse conditions. These results imply the presence of native rhizobia capable of inducing nodules on roots from locally adapted soybean genotypes in temperate soils. Moreover, nodulation efficiency of the three cultivars differs as well as the environment seem to play a role as only five soils, from the 107 tested, resulted in nodules in both garden and greenhouse experiments.

Identification and isolation of indigenous rhizobial bacteria from nodules

To isolate and identify individual rhizobial bacteria strains residing in regular formed nodules, we crushed one nodule (per plant, per cultivar and per experiment) and plated the juice on yeast mannitol agar (YMA) media, allowing the specific growth of rhizobia as well as other bacteria. For each nodule, one to eight bacterial strains were isolated and their taxonomic identify was determined by means of 16S rRNA PCR sequencing.

A total of 217 pure single-colony isolates were obtained from all nodules, with 74 from the nodules harvested in the gardens and 143 from the greenhouse experiments (Fig. 3A). Combination of both experiments provided 88, 42, and 87 isolates from nodules of the cultivars Primus, Bettina, and Shouna, respectively (Fig. 3B). Based on the 16S rRNA gene sequencing, the isolates could be assigned to 25 bacterial genera: *Agrobacterium, Bacillus, Bradyrhizobium, Buttiauxella, Caulobacter, Enhydrobacter, Enterobacter. Erwinia, Ewingella, Frigoribacterium, Kocuria, Lelliottia, Neorhizobium, Nocardioides, Paenibacillus, Pantoea, Pseudarthrobacter, Pseudomonas, Rhizobium, Rhodococcus, Serratia, Sphingomonas, Variovorax, Xylophilus* sp., and one group of uncultured/unknown/unassigned bacteria. The majority of the cultured bacteria were *Bacillus* sp. (n = 92) and *Pseudomonas* sp. (n = 51). Most of these *Bacillus* sp. strains were found in Primus and Bettina nodules, independently of the experimental setup, whereas the Shouna nodules from gardens accounted for most of the *Pseudomonas* sp. isolates in the collection (Fig. 3B).

Out of the 217 isolates, we identified 17 rhizobia, of which eight were identified as *Rhizobium* sp., five as *Neorhizobium* sp., and four as *Bradyrhizobium* sp (Supporting Information Table S1). Seven of the eight bacteria belonging to the *Rhizobium* sp. were obtained from nodules of plants grown under greenhouse conditions: five in soil #12 (Bettina and Shouna), one in soil #19 (Bettina), and one in soil #54 (Shouna), whereas one *Rhizobium* isolate was identified in garden soil #64 (Primus) (Supporting Information Table S1). The five *Neorhizobium* strains were isolated from nodules harvested from greenhouse plants grown in soil #12 (Bettina and Shouna) and from garden plants grown in soil #19 (Primus). The *Bradyrhizobium* strains were found in soils #76 (Shouna) and #42 (Primus) from the garden experiment and from the greenhouse experiment in soil #93 (Primus).

In summary, we were able to retrieve indigenous *Rhizobium, Neorhizobium,* and *Bradyrhizobium* sp. strains from nodules of the three soybean cultivars grown in Flemish soils

under greenhouse and natural conditions. Besides these nodulating bacteria, we isolated a diverse richness of other strains belonging to various bacterial families.

Characterization of the nodule microbiome

To obtain a better overview of the bacterial diversity present in the single nodules from soybean plants grown under different experimental conditions, we carried out a 16S rRNA metabarcoding experiment. After the final filtering steps, the dataset consisted of 63 nodules from which 50 originated from plants grown in gardens and 13 from the greenhouse experiments.

Next, we determined whether different factors, including growth conditions, provenance, soil type, garden soil number (#), plant genotype, nodule color (nitrogen fixing or nonfixing), and shape (regular or irregular) (Table 1), had an influence on the nodule bacterial diversity. To this end, a dissimilarity matrix was generated, and the samples were subjected to a PERMANOVA analysis. Four of the factors seemed to affect somewhat the diversity within the nodules, including the growth conditions (R^2 0.024, *P* value 0.001), the soil type (R^2 0,096, *P* value 0.002), the region where the nodules were collected (R^2 0.087, *P* value 0.005) and the soil number (R^2 0.207, *P* value 0.004). However, when the data were presented in Bray-Curtis dissimilarity-based principal coordinate analysis (PCoA) plots (Supporting Information Fig. S4), the samples apparently did not group separately based on any of the described factors. Nonetheless, the samples of the greenhouse experiments seemed to group together, even though they overlapped with the garden samples, but the latter were more spread within the PCoA plot (Supporting Information Fig. S4). Given the low number of samples from some of the soils, we decided to focus our analysis only on the identification of the general diversity found in the

samples and of the bacteria that may be responsible for the nodule formation and nitrogen fixation, i.e. the rhizobial isolates.

Next, we assessed whether the microbiome analysis could help to identify the nitrogen fixing strain as this one is expected to belong to the highest relative abundant rhizobium related ASV inside one nodule. We first assessed the diversity within the nodules with a red phenotype as for these ones our hypothesis is expected to be true. In total, six out of 63 nodules were red with a regular morphology, from which three were acquired from the gardens and three were obtained in the greenhouse. Two of the red regular nodules from the garden-grown plants contained ASV1 corresponding to *Bradyrhizobium* with the highest relative abundance and accounting for more than 50% of the relative abundances of all samples. One of the nodules was acquired from the Shouna cv. grown in soil #76, whereas the other was obtained from the Primus cv. in soil #50 (Fig. 4). Surprisingly, the other red nodule obtained in the garden did not demonstrate an expected microbial diversity, namely the presence of a highly abundant (*Brady*)*Rhizobium* strain. This nodule from cv. Primus grown in the garden soil #64 was colonized by many different ASVs, with one of the relatively high abundant ASV's belonging to the genus *Escherichia*/*Shigella* (accounting for 17%), whereas the presence of the ASVs belonging to *(Brady)Rhizobium* were equal to or lower than 3% (Fig. 4). In the greenhouse experiment, ASV2 (*Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium*) accounted for more than 50% of the relative abundances in one of the three red regular nodules from cv. Primus grown in soil #93. The other two red nodules did not present the expected diversity. The red regular nodule from cv. Shouna grown in soil #12 in the greenhouse had an ASV accounting for 79.64% of the relative abundances that belonged to the *Burkholderia-Caballeronia-Paraburkholderia* group (ASV6), together with another ASV belonging to the same group (ASV9), which represented 20.35% of the relative abundance. Unexpectedly, the most representative bacteria of the other red nodule present in the greenhouse experiment (cv. Primus

grown in soil #43), were *Escherichia/Shigella* (ASV5), *Pantoea* (ASV13), and *Pseudomonas* (ASV7), accounting for 28%, 29%, and 28% of the relative abundances, respectively (Fig.4).

Next, we studied the bacterial diversity inside the 57 white~~, nonfixing~~ nodules or nodule like structures. We found that ASV1 (*Bradyrhizobium*) appeared in 52 out the 57 nodules from garden and greenhouse experiments, accounting for more than 50% of the relative abundances in 11 of these nodules; hence, together with its presence in the red nodules, it is the most abundant *Bradyrhizobium* sequence in the dataset. These 11 nodules were all collected from gardens: one (cv. Bettina) grown in soil #38, six (cv. Primus) in soils #13, #19, #38, #42, and #50, and four (cv. Shouna) in soils #19, #42 and #55 (Fig. 4). However, the high presence of ASV1 in the nodules could not be associated to a specific soil and cultivar, because it occurred in nodules from different soils and colonized nodules of the three cultivars. Nevertheless, the relative abundances of this ASV was high only in the garden experiments, whereas in the greenhouse experiment, its relative abundance seemed low (ranging from 1%-19%). Another rhizobial ASV commonly found back in white nodules was ASV2 (*Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium*) that was found in 55 out of 57 samples, but represented only more than 50% of the relative abundances in two of the 55 nodules and ranged from 16% to less than 1% in the other 53 nodules. The two nodules in which ASV2 accounted for more than half of the relative abundances were white, irregular, and harvested from Shouna plants grown in garden soil #77 (Fig. 4).

A third rhizobia-associated ASV was assigned to *Bradyrhizobium* ASV12 that also accounted for more than 50% of the relative abundances in one of the white nodules. This nodule had an irregular morphology and was acquired from Bettina plants grown also in the garden, in soil #77 (Fig. 4). This ASV was less abundant than the other two (ASV1 and ASV2), because it was present in only 17 samples across the dataset, all from the garden experiment.

Additionally, other ASVs also belonged to (*Brady*)*Rhizobium* within the dataset, but because of their low abundance in the samples, we decided not to investigate them further.

Subsequently, we analyzed the nodule bacterial diversity beyond the rhizobia strains to examine whether common nonrhizobial bacteria or specific rhizobia-associated bacteria could be found. Within the nonrhizobia-related ASVs, ASV5 (*Escherichia*/*Shigella*) was the most abundant and was present in 62 out of 63 nodules, representing more than 50% of the relative abundances in four out of 13 and one out of 50 nodules from the greenhouse and garden experiments, respectively (Fig. 4). Additionally, other nodule-colonizing bacteria in relatively high abundances were two ASVs belonging to the genus *Pseudomonas* (ASV7 and ASV10), with relative abundances ranging between 28%-56%, *Duganella* (ASV34) with a relative abundance of 70.72% in one nodule, ASVs belonging to diverse groups, such as *Burkholderia-Caballeronia-Paraburkholderia* (ASV6 and ASV9), *Klebsiella* (ASV8), *Tardiphaga* (ASV66), (Fig. 4). Surprisingly, in eight out of the 57 nodules from the garden experiments, an ASV of an unclassified genus belonging to the order Chlamydiales (ASV3) was highly present, accounting for more than 20% of the relative abundances in seven nodules from soil #77 and in one nodule from soil #42 (Fig. 4).

Finally, to investigate whether specific ASVs appeared together in the nodules that were mainly colonized by *Bradyrhizobium* (ASV1 and ASV12), *Rhizobium* (ASV2), or *Burkholderia* (ASV6), because they could hint at nodulation helpers (Martínez-Hidalgo and Hirsch, 2017), we analyzed the diversity of the samples in which these ASVs represented more than 50% of the relative abundances (Supplementary Information Fig. S5). Within the co-occurring bacteria in these nodules, ASV5 (*Escherichia*/*Shigella*) was present in most of them, with only one exception, the nodule, in which ASV6 (*Burkholderia-Caballeronia-Paraburkholderia*) was predominant, had another ASV belonging to *Burkholderia-Caballeronia-Paraburkholderia* group (ASV9) (Supplementary Information Fig. S5). However, other ASVs could be found in

at least two of the nodules in which *(Brady)Rhizobium* accounted for the majority of the relative abundances. Within these ASV, only those were analyzed that represented more than 1% of the relative abundances in at least one of the nodules. These ASVs belonged to different groups, including *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium* (ASV2 and ASV11), *Duganella* (ASV25), *Bradyrhizobium* (ASV1 and ASV12), *Pantoea* (ASV13), *Klebsiella* (ASV8), *Streptomyces* (ASV81), *Burkholderia-Caballeronia-Paraburkholderia* (ASV6), *Rhodopherax* (ASV26), *Bacillus* (ASV17 and ASV23), *Enhydrobacter* (ASV35), and *Pseudomonas* (ASV7, ASV10, ASV14, ASV20, and ASV21) (Supplementary Information Fig. S5).

Phylogenetic analysis of candidate soybean nodulators based on whole genome comparisons

Eight strains belonging to the genera *Bradyrhizobium*, *Neorhizobium,* and *Rhizobium* were selected from the 217 isolates based on nodule phenotypes, i.e. isolates from red fixing or regular shaped nodules were included (Table 1; Supporting Information Table S1). The strains were isolated from soybean nodules of soil samples #12, #64, and #76. Additionally, because the nodule microbiome (16S) indicated that more than one (*Brady*)*rhizobium* sp. was present in the same nodule, impeding the selection of one isolate responsible for the nitrogen fixation, several isolates from the same nodule were included in the final eight (Table 2). Of these eight isolates, five were obtained from nodules of the cultivars Bettina and Shouna grown in the same soil #12, namely, BETTINA12A, BETTINA12B, SHOUNA12A, SHOUNA12B, and SHOUNA12C, whereas the remaining three were isolated from nodules of cv. Primus and Shouna from four different soils: PRIMUS64 from soil #64, PRIMUS42 from soil #42, and SHOUNA76 from soil #76 under garden conditions.

To investigate whether these isolates could relate to the nodule microbiome data, we compared their 16S rRNA sequences with the ASV sequences by means of BLASTn (Altschul *et al.,* 1990). The results indicated that the isolates matched four ASVs. Interestingly, ASV1, the most abundant ASV identified as *Bradyrhizobium*, presented a 100% identity with the isolates *Bradyrhizobium* sp. SHOUNA76 and PRIMUS42. The *Neorhizobium* strains BETTINA12A, SHOUNA12A, and SHOUNA12B were associated with another low-abundant ASV (ASV44), with 100% identity values. Finally, the three strains identified as *Rhizobium* were included in ASV2, the most abundant ASV identified as *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium*. The similarity was 100% for all three *Rhizobium* isolates.

To further evaluate the phylogenetic relationships between these isolates in more detail and to score the differences between these isolated strains, we performed a whole genome sequencing approach with short reads, on a NovaSeq-6000. This data was assembled with the Shovill pipeline, after which their completeness and contamination level was examined with CheckM v1.1.2 (Parks *et al.,* 2015). The different data sets showed different degrees of contamination and the assembled genomes varied in size, ranging from 6.5 Mb to 8.7 Mb (Supporting Information Table S2). After assembly, the genomes were annotated with Prokka v1.14.5 (Seemann, 2014).

Firstly, the strains PRIMUS64, BETTINA12B, and SHOUNA12C, previously identified as *Rhizobium*, were compared with 56 strains from this genus by means of the orthoANIu tool to calculate the Average Nucleotide Identity (ANI; Yoon *et al.,* 2017). Additionally, a phylogenomic tree was generated with the bcgTree v1.1.0 software (Ankenbrand and Keller, 2016) by comparison of 107 core genes. The tree topology together with the ANI values revealed that two distinct groups within the three isolated *Rhizobium* strains (Supporting Information Fig. S6). The *Rhizobium* strains BETTINA12B and SHOUNA12C belonged to the species *R. rhizogenes*, because they shared an ANI value of 99.13 and 99.12, respectively, with

the type strain *R. rhizogenes* NBRC13257. Additionally, the results indicated a high similarity between both strains with ANI values of 99.98 when they were compared to each other, implying that they are probably the same strain. The other *Rhizobium* sp. strain PRIMUS64 seemed to belong to the *R. leguminosarum* species. Its closest relatives were *R. leguminosarum* bv. *trifolii* 23B (ANI value of 98.16) and *R. leguminosarum* bv. *viciae* BIHB 1217 (ANI value of 98.11). Overall, *Rhizobium* sp. PRIMUS64 also shared high ANI values with other strains belonging to this species, including the type strain *R. leguminosarum* bv. *viciae* USDA 2370, sharing an ANI value of 98.11.

Next, we assessed the phylogeny of the *Neorhizobium* strains SHOUNA12A, BETTINA12A, and SHOUNA12B. Here, 22 reference strains were included for the analysis, comprising all the available genomes of the type strains. The analysis revealed a high similarity between the strains SHOUNA12A and SHOUNA12B, with an ANI value of 99.98%. However, the strain BETTINA12A was different, presenting values of 97.08 and 97.07, compared with SHOUNA12A and SHOUNA12B, respectively. Based on the results of orthoANIu (Table 3) and the tree topology (Supporting Information Fig. S7), all three strains probably belong to the *N. galegae* with *N. galegae* bv. *officinalis* HAMBI 1141, followed by *N. galegae* bv. *orientalis* HAMBI 540^{T} and *N. galegae* Ng87, as their closest relatives.

Finally, the phylogeny of the strains PRIMUS42 and SHOUNA76 belonging to the genus *Bradyrhizobium*, was established. In this case, 48 strains, including the available type strains within the genus, were used to calculate the orthoANIu values. *Bradyrhizobium* sp. SHOUNA76 and PRIMUS42 shared ANI values of 99.81 and did not seem to belong to any of the already described species within this genus. Based on the tree topology (Supporting Information Fig. S8), their closest relatives were *B. shewense* ERR11^{T}, *B. ottawaense* OO99^{T}, *B. symbiodeficiens* 85S1MB^{T}, and *B. amphicarpaeae* 39S1MB^{T}, results that were corroborated with orthoANIu (Table 4).

Taken together, whole-genome analysis of the eight selected rhizobial isolates characterized them as six distinct strains in three genus groups (*Rhizobium, Neorhizobium* and *Bradyrhizobium* sp.). The comparative analysis classified both *Neorhizobium* sp. strains SHOUNA12A and SHOUNA12B as well as *Rhizobium* sp. strains BETTINA12B and SHOUNA12C as similar strains.

ONT sequencing of a single root nodule harvested from soil-grown Shouna cv. soybean plant

To identify whether it is possible to recover the whole genome of the most abundant rhizobial strain from a root nodule harvested from soil-grown plants with Oxford Nanopore Technologies (ONT), we used the red nodule from cultivar Shouna grown in soil #76 (nodule S101) as a proof-of-concept. We sequenced bulk DNA from a single nodule, extracted following a powersoil protocol, by means of a MinION Mk 1B sequencer using a SpotON flow cell (type 9.4). In total, 5,595,749 reads (556,5750,767 nucleotides) were sequenced with a read-N50 of 1499 nucleotides (ranging from 67,899 to 100). These reads were assembled using Flye (v2.7.1), obtaining 298 contigs with a mean coverage of 407x. From the 298 contigs, the largest one was 7,673,612 nucleotides long and represented a circular genome belonging to *Bradyrhizobium* sp. The genome had a completeness of 98.05 % and a contamination of 0.82 %, based on the CheckM v1.1.2 results (Parks *et al.,* 2015). This ONT assembled genome of the *Bradyrhizobium* sp. was compared with *Bradyrhizobium* sp. strain SHOUNA76 that was isolated from the same nodule using the orthoANIu tool. Surprisingly, both strains belonged to two different species, sharing ANI values of 89.40, meaning that the most abundant *Bradyrhizobium* sp. strain present in the nodule was not the one that was isolated. The ONT assembled genome shared an ANI value of 97.03 with the type strain *B. symbiodeficiens* 85S1MB indicating it belonged to the *B. symbiodeficiens* species group. This type strain was also isolated from soybean nodules but lacks key nodulation (*nod*) and nitrogen fixation (*nif*)

genes (Bromfield *et al*., 2020). The presence of those genes was investigated in the ONT genome with Prokka v1.14.5 (Seemann, 2014). We located *nifH*, *nifA*, *nifDKN* and *fixX* genes, although, they were spread within the genome instead of clustered together. *nifE*, *nodDYABCSUIJ* and *fixABC* genes could not be found.

Subsequently, to learn more about the bacterial diversity, we compared the taxonomy results from the 16SrRNA amplicon sequencing with the ONT data. By means of amplicon sequencing, only two ASVs were found, ASV1 (*Bradyrhizobium*) and ASV5 (*Escherichia*/*Shigella*), which represented 99.94% and 0.06% of the relative bacterial abundances in this nodule, respectively. As expected, the *Bradyrhizobium* sp. genome assembled with the ONT read also belonged to ASV1. The taxonomical assignment of the other less abundant ONT reads was performed using three different programs: MiniKraken2 v2.0.9 (Wood *et al*., 2019), Centrifuge vs 1.0.4 (Kim *et al*., 2016), and Kaiju (Menzel *et al*., 2016) depending on different databases. In all three programs more microorganisms were identified with a more precise taxonomical assignment compared with the amplicon sequencing, (Fig. 5; Supporting Information Fig. S9). The results obtained with the MiniKraken program assigned 62%, of the total amount of reads, to bacteria, with only small amounts of the reads belonging to archaea (0.02%) or virusses (0.03%). In addition, 37% of the reads had no hits with microorganisms present in the database. Within the bacteria, 97% of the reads belonged to *Bradyrhizobium* sp., and 1% to *Escherichia* sp., whereas 2% of the reads belonged to bacteria from different phyla (Fig. 5). MiniKraken seemed to be better at assigning reads within the bacterial kingdom as 76% of the reads belonged to *B. symbiodeficiens*, whereas Centrifuge did not identified any reads and Kaiju identified only 22 % of the reads as *B. symbiodeficiens* (Supporting Information Fig. S9). Furthermore, reads belonging to other *Bradyrhizobium* sp., including *B. amphicarpaeae*, *B. ottawaense, B. diazoefficiens*, among others, could also be found within the nodule. In the case of the identification obtained with Centrifuge and Kaiju,

58% and 63% of the sequences belonged to bacteria, respectively. Within the bacteria, 79% and 83% of the reads belonged to several species of *Bradyrhizobium*, while the rest belonged to bacteria from other genera (Supporting Information Fig. S9).

To summarize, by means of ONT sequencing, the most abundant *Bradyrhizobium* strain was identified for which the whole genome was recovered, although, belonging to the same ASV, this strain was different than the one isolated. Furthermore, the ONT reads allowed for a more in depth identification of the nodule microbiome compared with the 16S rRNA amplicon sequencing.

Nodulation capacity of six unknown soybean nodulators of Rhizobium, Neorhizobium, and Bradyrhizobium sp.

Whole-genome analysis of the eight selected rhizobial isolates characterized them as six distinct strains in three genus groups. To determine whether these six selected strains were responsible for nodulation on soybean roots, we repeated several inoculation experiments on soybean plants grown in sterile vermiculite under growth room conditions.

The inoculation experiments were carried out on the three cultivars (Primus, Bettina, and Shouna) and with the commercial inoculants Biodoz and HiStick® as positive controls. After four weeks, nodule presence and color were assessed (Fig. 6; Supporting Information Table S3). Inoculation with Biodoz and HiStick resulted in large red nitrogen-fixing nodules on all soybean cultivar roots and the six candidate isolates in nodules on at least one of the cultivar roots (Supporting Information Table S3; Fig. 6). *Neorhizobium* sp. SHOUNA12A repeatedly generated nodules in all experiments that were white and occurred only on the roots of Primus and Bettina cultivars. Inoculation with the other candidate strains, *Rhizobium* sp. strains PRIMUS64 and BETTINA12B and *Bradyrhizobium* sp. strains SHOUNA76 and PRIMUS42,

resulted in nodules on the three cultivars in more than half of the inoculation experiments (Supporting Information Table S3). These nodules were mainly white, but occasionally a red and probably fixing nodule was found for *Bradyrhizobium* sp. SHOUNA76. Inoculation with *Neorhizobium* sp. BETTINA12A also generated white nodules on the roots of all cultivars but only in one of the two performed experimental repeats. In conclusion, nodules were formed after inoculation of soybean plants with pure isolates of *Rhizobium* (BETTINA12B/SHOUNA12C and PRIMUS64), *Neorhizobium* (SHOUNA12A and BETTINA12A), and *Bradyrhizobium* sp. (PRIMUS42 and SHOUNA76).

Discussion

European farmers and industry are interested in local cultivation of soybean plants for human consumption. Therefore, soybeans need to produce protein-rich beans and a high protein content is obtained only upon efficient nodulation of the soybean roots. Previous studies with early-maturing soybean varieties and commercial inoculants in Flanders have revealed that the yearly variable results do not reach the required quota (Pannecoucque *et al.*, 2018a). We hypothesized that indigenous rhizobia strains that are adapted to grow under a temperate climate are essential for an efficient soybean nodulation.

Here, we selected 107 different soils from different regions in North Belgium to trap indigenous rhizobia strains that could nodulate three different early-maturing soybean varieties. The trapping experiment was carried out both in gardens under natural conditions and in pots under greenhouse conditions. Interestingly, both experiments resulted in approximately 20 soils, in which at least one soybean plant had nodules. Given that soybean is not endemic to these regions and that the soybean-adapted rhizobial strains might not be available in the soils, this number is quite high. However, in most cases, the nodules were not efficient, because they

lacked the red nitrogen-fixing central tissue, indicating that although these strains could start the development of a nodule, the interaction was interrupted. For example, some rhizobia might have the specific Nod factors to trigger the nodule organogenesis signaling pathway but do not contain the necessary exopolysaccharides critical for nodule colonization (Clúa et al. 2018).

In only five soils (#19, #43, #67, #76, and #93), nodules were obtained in both setups, but without any clear commonality in soil type, province, or other factor (Table 1). As the overlap between the two setups is small, we speculate that the differences are mainly due to environmental growth conditions between the two setups, such as temperature, watering, and growing space (pot versus open ground). Additionally, in the soils in which nodules occurred in both setups, differences in nodule shape, size, and color were clearly visible (Fig. 2; Table1), further demonstrating that the environmental conditions greatly affect the nodulation process. Indeed, differences in watering and temperature can greatly influence both the host plant and the rhizobial strain involved with an impact on the growth of both organisms, the infection process, the nodule functionality, and the molecular signal exchange between both (Miransari *et al.,* 2013). However, we cannot rule out that the observed differences are caused by the coincidental capture of the rhizobia in the trapping and, hence the formation of nodules, in a particular soil sample used in the greenhouse experiments (0.5 L) and the 1 m^2 used in the gardens (Supporting information Fig. S1). It will be therefore interesting in future experiments, such as the soybean in 1000 garden project (https://sojain1000tuinen.sites.vib.be), to investigate the abundances of the captured rhizobia in the corresponding soils.

The PERMANOVA analysis on the entire dataset similarly revealed a significant effect of growth conditions, soil type, and region on the diversity within the nodules. However, a clear separation of the samples based on these factors could not be observed in the PCoA plots (Supporting Information Fig. S4). The absence of distinctive clusters could relate to the fact that even though these factors had a significant impact on the microbial community, it was not

strong enough to indicate a clear clustering of the samples. A more specific and detailed soil and growth condition analysis as well as an expanded sample size will be necessary to identify common environmental factors that could explain the presence or absence of nodules, the residing rhizobia, and the general bacterial community inside the nodules. Aspects that will be tackled in the soybean in 1000 garden project (https://sojain1000tuinen.sites.vib.be).

We further noticed that in the five common soils, the genotypes that were nodulated differed in the two setups. Additionally, when the entire dataset was considered, nodules occurred more frequently on the cv. Primus roots in both experiments, whereas Bettina and Shouna roots were also nodulated, but differently. Although the PERMANOVA tests did not reveal a significant effect of the plant genotype on the microbial diversity, these observations are in line with the strong genotype influence on nodule formation (Sinclair *et al.,* 1991; Sharaf *et al.,* 2019). From the soybean genetics perspective, it would be interesting to investigate why the Primus cv. can be nodulated when grown in the gardens, because it might help the selection of soybean varieties capable of nodulating with a variety of strains present in the Flemish soils.

Besides an effect on nodule formation, environment and genotype also influence the rhizobia strains residing inside the nodules that determine the nitrogen-fixing capacity and nodulation efficiency (Miransari *et al.,* 2013). Indeed, differences were observed in the nodule bacterial communities in the same soil used in different experimental setups. By using 16S rRNA PCR and metabarcoding, we determined the microbial community and identified the most abundant rhizobia of the nodules. ASV1 (*Bradyrhizobium*) was the most highly represented rhizobial ASV in the dataset and was profusely present in the nodules of garden plants, but, in contrast, was not or very scarcely present in nodules from greenhouse experiments. In these experiments, ASV2 (*Rhizobium*) was commonly found and present in almost all white nodules, as was also the case for soil #76, in which also the nodule phenotypes varied between both experimental setups and between cultivars. In the garden soil #76, large

and regular fixing nodules on Shouna plants were mainly colonized by ASV1 (*Bradyrhizobium*), whereas in the greenhouse, these nodules were white and small, present only on Primus plants and colonized by ASV5 (*Escherichia*/*Shigella*) (Fig. 2; Table 1). Indeed, some *(Brady)Rhizobium* are known to be more sensitive to soil temperatures and water status (Zhang *et al*., 1995; Pannecoucque *et al.,* 2018a; Sharaf *et al.,* 2019). In the same cultivar used under different water conditions, an entirely different community of rhizobia bacteria was found in nodules (Sharaf *et al*., 2019).

By classical plating and 16S rRNA PCR on single isolates, we isolated and identified eight *Rhizobium* sp., five *Neorhizobium* sp, and four *Bradyrhizobium* sp strains, and 200 other non-rhizobial bacteria. The isolation of many non-rhizobial strains can be explained by a faster and better growth of certain strains compared with the slow-growing *Bradyrhizobium* sp. (Jordan, 1982), also on the specialized YMA media which is useful for rhizobial enrichment but is clearly not exclusive. In addition, as can be seen from the microbiome analysis, not all nodule structures contained rhizobia and some nodules were expected to contain rhizobia but for which we did not get a rhizobial strain. Nevertheless, we did isolate many of the nodule-residing strains. To improve the specific isolation of rhizobial strains, future experiments should include additional specialized rhizobial growth media, more nodule juice dilutions, temperature variation and longer incubation times. The two isolated *Bradyrhizobium* sp. strains, SHOUNA76 and PRIMUS42, also belonged to the most abundant ASV of the microbiome dataset, ASV1 (*Bradyrhizobium*). ASV1 was found in almost all white nodules with low relative abundancy and highly abundant in the red nodules of Shouna plants grown in soil #76, from which we isolated *Bradyrhizobium* sp. SHOUNA76. However, by means of ONT sequencing on the same DNA of the same nodule, we revealed the genome of yet another *Bradyrhizobium* species strain related to *B. symbiodefficiens*, belonging to the same ASV1 and sharing 89.39 % homology with *Bradyrhizobium* sp. SHOUNA76 that was isolated. In addition,

the classification of the ONT reads allowed the identification of a higher amount of microorganisms and detected even sequences of viruses and archaea, which is not possible with 16S rRNA amplicon sequencing. These results demonstrate that diverse identification approaches have to be combined to capture the true nodule rhizobial community of a nodule. Additionally, the ONT results show, to our knowledge for the first time, that, although, the genome of soybean is approximately 150x times larger, the entire genome of the most abundant microbial occupant can be revealed using one single ONT run on nodule derived DNA.

The isolated *Rhizobium* sp. strains, including *Rhizobium* sp. BETTINA12B/SHOUNA12C and PRIMUS64, belonged to ASV2 (*Rhizobium)* that was the second most abundant ASV in the dataset and was present in all white nodules at a low relative abundance. The isolated *Neorhizobium* sp. SHOUNA12A/ SHOUNA12B and BETTINA12A, coming from the same cultivars and soil as the *Rhizobium* sp. BETTINA12B/SCHOUNA12C, belonged to a low-abundant ASV, ASV44 (*Neorhizobium*). Due to the small nodule size, it was almost impossible to isolate bacteria from the same nodule as was used for the microbiome analysis. Hence, we cannot conclude or rule out that one plant carries several nodules occupied by different rhizobia.

Interestingly, besides the expected common presence of ASV1 (*Bradyrhizobium*) and ASV2 (*Rhizobium*) in nodules, we also found one red nodule in which two ASVs of the genus *Burkholderia* (ASV6 and ASV9) had a high relative abundance. These ASVs accounted for the entire bacterial community in the red nodule from Shouna plants grown in soil #12. Nevertheless, we could not isolate *Burkholderia* strains belonging to these ASV from this nodule but only strains belonging to ASV44 (*Neorhizobium*), *Neorhizobium* sp. SHOUNA12A/SHOUNA12B strain and ASV2 (*Rhizobium*), *Rhizobium* sp. BETTINA12B/SHOUNA12C strain. This result can be due to the special growth media for *Rhizobium* sp. that do not allow optimal growth of certain other strains. Strains of the (*Para*)*Burkholderia* genera have emerged as new nodule symbionts of various legumes in

Africa (Talbi *et al.,* 2010; Beukes *et al.,* 2019) and of soybean in Venezuela (Arigas Ramírez *et al.,* 2019). To be able to demonstrate that these strains can independently initiate nodule development and nitrogen fixation without the help of other *Rhizobium* sp., we should isolate the strains and carry out inoculation experiments in the laboratory. Interestingly, the *Neorhizobium* sp. SHOUNA12A/SHOUNA12B and *Rhizobium* sp. BETTINA12B/SHOUNA12C strains that had been isolated from the same nodules were capable of nodulating soybean in the laboratory (Figure 5). Future coinoculation experiments of these strains with the still to be isolated *Burkholderia* sp. strains, but also the combination of both (*Neo*)*Rhizobium* sp. strains could have an impact on nodulation efficiency, nitrogen fixation, and plant growth, as demonstrated previously (Hilário de Carvalho *et al.,* 2020).

The six distinct soybean nodulators that we isolated and identified from the Flemish soils all nodulated soybean in the laboratory, albeit only the *Bradyrhizobium* sp. SHOUNA76 strain demonstrated occasionally possible nitrogen-fixing capability. Comparative genomics classified these soybean nodulators in three distinct genera with high similarities to known reference strains, except for the two *Bradyrhizobium* sp. strains SHOUNA76 and PRIMUS42 (ASV1) that seemingly did not belong to any of the already described species within this genus. The closest relative was *B. schewense* ERR11^T, a type strain isolated from a nodule of the leguminous tree *Erythrina brucei* (Aserse *et al.,* 2017). As these two indigenous *Bradyrhizobium* strains can nodulate soybean, although suboptimal causing small white nodules for *Bradyrhizobium* sp. strains PRIMUS42 and occasionally red nodules for *Bradyrhizobium* sp. strains SHOUNA76, they would be perfect candidates for future experimental strategies to improve their nodulation performance for agricultural application. For example, by subjecting these strains to experimental evolution, they could develop an improved compatibility with soybean. Evolutionary adaptation of indigenous rhizobia to soybean could ameliorate not only their nitrogen fixation capacity but also their effect on plant

yield (Zhao *et al.,* 2018). The improved yield, nodulation capacity and nitrogen fixing activity of these evolved rhizobial strains can then be compared with the original strains and/or other commercial inoculants by for instance in-depth phenotypic and microscopic analysis as well as acetylene reductase assays, respectively. Furthermore, testing their nodulation capacity and efficiency in more environmental conditions such as diverse soils under natural conditions, allowing their interaction with rhizosperic microbial community, might improve their symbiotic efficiency (Han *et al*., 2020). On top of that, more effort has to be put in the practical aspects of the rhizobial isolation techniques as *Bradyrhizobium* sp. strain SHOUNA76 was not the most relative abundant *Bradyrhizobium* species of the nodule harvested from Shouna grown in soil #76 under garden conditions. Future combinatorial inoculation experiments of the rhizobial strains within one nodule might optimize both nodulation capacity and nitrogen fixation.

The two *Neorhizobium* sp. strains, SHOUNA12A/SHOUNA12B and BETTINA12A (ASV44), belonged to the *N. galegae* species that resembled most *N. galegae* bv. *officinalis.* The *N. galegae* species consist of two biotypes, *N. galegae* bv. *officinalis* and *N. galegae* bv. *orientales,* that both form an efficient symbiosis with their respective goat's rue (*Galega orientalis* and *G. officinalis*) host plants (Karasev *et al.,* 2019). When these strains interact with another host, they still form nodules, although ineffectively because of the divergence in the nitrogen fixation genes (*nif/fix*) (Karasev *et al.,* 2019). Also here, inoculation of soybean plants in the laboratory with the two *Neorhizobium* sp. SHOUNA12A/SHOUNA12B and BETTINA12A strains reproducibly resulted in nodules in all experiments, but never red and thus probably did not fix nitrogen. Surprisingly, the *Rhizobium* sp. strain BETTINA12B/SHOUNA12C belonged to the pathogenic species *R. rhizogenes* and resulted in nodules when soybean plants were inoculated in the laboratory. *R. rhizogenes* strains are common inhabitants of legume nodules, but surprisingly and unexpectedly, they also can induce nodulation independently. The other strain *Rhizobium* sp. PRIMUS64 belonged to the *R.*

leguminosarum species, known to have a broad host plant range and to easily induce efficient symbiosis. Also in our experiments, inoculation of soybean with *Rhizobium* sp. PRIMUS64 resulted in nodules. Hence, a detailed genomic analysis or genome mining approaches might help shed light on why certain strains nodulate more efficiently compared to others and why other rhizobia strains cannot nodulate at all.

Coinoculation of rhizobia with other bacteria has long been known to promote nodulation and nitrogen availability (Martínez-Hidalgo and Hirsch, 2017). To find these potential and so-called 'helper' bacteria for enhancement of the nodulation of the isolated soybean candidates, we searched the most common nonrhizobia co-occurring ASVs in the microbiome dataset. ASV5 (*Escherichia*/*Shigella*) was the most abundant ASV and co-appeared within almost all nodules. *Escherichia*/*Shigella* spp. are mostly pathogenic bacteria belonging to the family of Enterobacteriaceae. Although detection of ASVs of this family in our nodule microbiome was rather surprising, they have been found in nodules of other legumes as well (Lu *et al.,* 2017). In the past, isolation experiments of rhizobia from legume nodules ignored the identification of nonrhizobial bacteria, resulting in a general lack of knowledge of their true diversity in root nodules. Only now, the recent metagenomics and microbiome approaches, as done here, can give further insights into the presence of these bacteria and their possible roles in nodulation. All nodules were surface-sterilized prior to analysis, decreasing the possibility of contamination, but still ASV5 (*Escherichia*/*Shigella*) was present in all of them. These bacteria could play an unprecedented and surprising role in nodulation. The second most abundant ASVs that colonized the nodules in relatively high abundances belonged to the *Pseudomonas* sp. genus. Besides *Pseudomonas* sp. strains, our bacterial collection also contained many isolates belonging to the Bacillaceae family that were equally represented by abundant ASVs. It is not surprising that many *Pseudomonas* and *Bacillus* sp. isolates were found, because they are commonly known soybean-associated bacteria (Kuklinsky-Sobral *et al.,* 2004; Sharaf *et al.,*

2019). *Bacillus* sp. can act as plant-growth promoting rhizobacteria and the presence of *Bacillus* strains in nodules has been reported for decades (Beijerinck, 1888). Future co-inoculation experiments with certain isolates of the *Bacillus* and *Pseudomonas* sp. and the identified soybean nodulators could result in improved soybean growth (Bai *et al.,* 2003; Sánchez *et al.,* 2014), opening promising perspectives, such as the development of inoculants or seed coatings, for future agricultural processes.

In conclusion, our soybean trapping experiment has revealed the presence of indigenous rhizobia strains, adapted to temperate soils and environment that are compatible with the locally adapted soybean genotypes. We not only identified three promising rhizobial strains that can efficiently and reproducibly nodulate soybean, and can be used in soybean cultivation approaches, but we also found less-effective rhizobia and co-occurring nonrhizobial bacteria that might be perfect candidates for future experimental evolution and genomic comparison analyses. To further help the local soybean cultivation in Europe in the future, more efficient soybean nodulators should be identified by including an expanded sample size, nodule nitrogen fixation assays, more detailed soil analysis together with field trials and optimization of soybean inoculation strategies, such as the soybean in 1000 gardens project (https://sojain1000tuinen.sites.vib.be).

Experimental Procedures

Plant material and growth conditions

For each experiment, three early-maturing soybean cultivars Primus, Bettina, and Shouna were used (https://bsl.baes.gv.at/). The garden experiments were conducted from May until

September 2019. Each volunteer received five seeds per cultivar that were sown on a 30×50 cm square patch with 10 cm between each seed (Supporting Information Fig. S1). Volunteers were instructed not to water the plants to keep them under natural growing conditions. In parallel, the soil samples collected from each garden were directly or kept at 4°C until used in the greenhouse experiments. In these experiments, three seeds per soybean cultivar were germinated and grown in 13-cm round pots containing the different garden soils. Plant were watered twice a week. After two months of growth or when approximately 60 cm high, the plants were harvested and the nodules were assessed directly.

For nodulation confirmation, inoculation experiments in the laboratory, soybean seedlings from the cultivars Primus, Bettina, and Shouna were germinated and grown in sterilized vermiculite soaked in nitrogen-poor Sol-I nutrition media under optimal greenhouse conditions. After 7 days, plants were inoculated with the candidate rhizobial strain at OD = 0.01 and watered twice a week with Sol-I solution. After four weeks, nodule presence and color were assessed.

Nodule harvest and bacterial isolations

Of each plant the nodules were collected and surface sterilized for further experiments. Nodules were surface sterilized with 70% ethanol for 1 min, with 2.5% Na-hypochlorite for 3 min, and washed with sterile water three times. Before isolations or storage at -80°C, a subset of the nodules were tested for surface sterilization by rolling over YMA and testing bacterial growth upon incubation at 28°C for 3 days. One to two nodules per plant were crushed in liquid yeast mannitol agar (YMA) medium and dispersed on solid YMA with Congo Red in *in vitro* plates for the specific cultivation of rhizobia species (http://www.himedialabs.com/TD/M707.pdf). The plated suspensions were incubated at 28°C for a minimum of 3 days until bacterial colonies appeared. Colonies were picked and streaked until pure cultures on YMA for DNA extraction and identification at family level based on the 16S rRNA gene. DNA was extracted with

alkalytic lysis buffer as described (Niemann *et al.,* 1997). The 16S rRNA gene was amplified with the 16S rRNA primers 27F (5'-AGAGTTTGATCMTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3'). PCR amplification was done with 0.2 µl polymerase (IProofTM High Fidelity DNA Polymerase; Bio-Rad, Hercules, CA, USA), 1 µl of 10 µM forward primer, 1 µl of 10 µM reverse primer, 0.4 µl 10 mM dNTP mix, and 4 µl buffer 5× IproofTM buffer (Bio-Rad). PCR conditions were denaturation at 98°C for 3 min, 30 cycles of 98°C for 10 s, 55 °C for 20 s, and 72 °C for 45 s, and a final elongation at 72 °C for 5 min. The resulting PCR products were purified with the GeneJET PCR Purification Kit (Thermo Fischer Scientific, Waltham, MA, USA) and sequenced by Sanger sequencing (Eurofins Genomics, Ebersberg, Germany). Forward and reverse sequences were merged by means of the CLC Main Workbench 7 (QIAGEN Digital Insights, Aarhus, Denmark) and taxonomy was assigned to the resulting sequences with the SILVA database version 128 or the National Center for Biotechnology Information (NCBI).

16S rRNA amplicon sequencing of nodules

Each nodule was ground in liquid nitrogen and DNA was extracted with the DNeasy PowerSoil DNA kit (QIAGEN, Hilden, Germany). The V4 region of the 16S rRNA gene was amplified via PCR as described (Beirinckx *et al.,* 2020). The amplicons of all nodules were pooled and sequenced on the Illumina MiSeq 2×250 platform (VIB, Nucleomics Core, Leuven, Belgium). After sequencing, the samples were de~~multi~~plexed and primers were removed by the company. The reads were then trimmed by means of the DADA2 package in Rstudio, by truncation at 170 and 150 for the forward and reverse reads, respectively, and with the default quality parameters (trunQ = 2; maxEE=c(2,2)). The amplicon sequence variants (ASV) were generated and the taxonomy was assigned with the SILVA vs 138 database. Subsequently, the reads belonging

chloroplastic (Class Chloroplast) and mitochondrial (Order Rickettsiales) DNA were removed, reducing the read count on average by 81.2%. The samples with a read count below 1,000 reads were eliminated.

To observe the effect of the different factors on the bacterial diversity of the nodules, we generated a dissimilarity matrix based on the Bray-Curtis dissimilarity index with the vegdist function in the vegan package, providing the ASV count tables as input. The effect of the metadata (e.g. growth conditions, region, soil type, soil number, plant genotype, and nodule color and shape) on the diversity within the nodule was determined with the function adonis in RStudio, which calculates the Permutational Multivariate Analysis of Variance (PERMANOVA). A Principal Coordinates Analysis (PCoA) plot was generated to represent the data.

Whole-genome analyses

Isolates were grown in liquid YMA and genomic DNA was extracted and purified with the cetyl trimethylammonium bromide/phenol-chloroform method. Genomes were sequenced on the NovaSeq 6000 (Oxford Genomics Center, University of Oxford, Oxford, UK). Libraries were prepared by an in-house adapted protocol of the NEB prep kit (New England Biolabs, Ipswich, MA, USA). The genomes were assembled and annotated as described (Luo *et al.*, 2019). The genome quality was assessed with CheckM (Parks *et al.*, 2014) that allows the determination of the completeness and contamination of the different genomes (Supporting Information Table S2). The software MaxBin (Wu *et al.*, 2016) was used on the assembly for genome separation and the quality was assessed by means of CheckM after the separation of the contigs belonging to each of the strains.

To investigate the phylogeny of the different strains, the type strains from the genus *Bradyrhizobium, Neorhizobium*, and *Rhizobium* were downloaded from the NCBI database and analyzed as described (Luo *et al*., 2019). In the case of *Neorhizobium,* three strains belonging to the genus *Rhizobium* were included in the analysis as an outgroup. For the study of the phylogeny of the strains belonging to the genus *Bradyrhizobium,* one strain of *Neorhizobium*, one of *Agrobacterium,* and one of *Rhizobium* were used as an outgroup. Finally, three strains belonging to the genus *Ensifer* were used as an outgroup in the case of *Rhizobium.*

ONT of a single red nodule from Shouna grown in soil #76 (nodule S101)

The isolated clean DNA was provided as bulk extraction without preliminary filtering or selection. Sequencing was done using Oxford Nanopore Technologies (ONT, MinION). Isolated DNA was purified using AMPure XP beads and resuspended in water prior to library preparation. Sequencing library was prepared using the ligation sequencing kit, SQK-LSK109 according to manufacturer's protocols. No size selection was done as the extraction method would only deliver rather shorter reads. This library was sequenced on a FLO-MIN106 R9.4.1 SpotON flow cell on a MinION Mk 1B sequencer (Oxford Nanopore technologies). Basecalling of the raw data was done using Guppy (version 4.5.2).

The ONT reads were assembled using the Flye software (v2.7.1), with parameters, –meta and –nano-raw, also including 9 rounds of polishing through consensus. The high coverage combined with the relatively long reads was deemed sufficient to allow for proper segregation and assembly of the read mix of plant and bacterial origin. To assign the taxonomy of the ONT reads, three different programs were used, including MiniKraken2 v2.0.9 (Wood *et al*., 2019), Centrifuge vs 1.0.4 (Kim *et al*., 2016) and Kaiju web server (Menzel *et al*., 2016). These programs used the databases minikraken2_v1_8 GB, p_compressed_2018_4_15 (6.3 GB) (for bacteria and archaea), and nr_euk 2021-02-24 (61 GB), respectively. To visualize the

taxonomy, the software Krona (Ondov *et al*., 2011) was used. In addition, the quality of the genome obtained with the ONT reads was checked using CheckM (Parks *et al*., 2014), and was annotated using Prokka v1.14.5 (Seemann, 2014).

Nodulation confirmation experiments

Soybean seeds were surface sterilized first 1 min with 70% ethanol and then 3 min with a bleach solution (29 ml sterile water, 15 ml NaClO, 12–13% (v/v) stock solution, and 1 ml Tween 20), where after they were washed five times for 15 min with sterile water and allowed to germinate for 4 days in the dark at room temperature. Seedlings of each cultivar were sown in sterilized vermiculite in 13-cm round pots and grown under a 16-h light/8-h dark photoperiod at 22°C in the greenhouse. Plants were watered twice a week with nitrogen-poor SOLi solution (Blondon, 1964). After one week, plants were inoculated with 1 ml of the respective bacteria at OD = 0.01. Four weeks after inoculation, the nodule presence was assessed and the nodules were photographed. These screening experiments were repeated independently four to 10 times for each candidate isolate. The commercial *Bradyrhizobium* inoculants, Biodoz and HiStick®, were used as positive controls for nodules. Noninoculated mock plants were used as negative controls.

Data deposition

All raw sequencing data are deposited in Bioproject ID PRJNA776129: indigenous (Belgian) rhizobia and associated bacteria.

Acknowledgments

The authors thank all volunteers, including those who were not incorporated in this study, for their enthusiasm to grow and harvest soybean plants in the summer of 2019 and are grateful to the Oxford Genomics Centre at the Wellcome Centre for Human Genetics (funded by Wellcome Trust grant reference 203141/Z/16/Z) for the generation and initial processing of the sequencing data and are grateful to Yves Van de Peer for his support in this research and Martine De Cock for help in preparing the article. This research was supported by grants from VIB Tech Transfer (T0291 and T0391) and from the Research Foundation-Flanders (project no. G015118N, S.G.M). J.V.D. is a postdoctoral fellow of the Research Foundation-Flanders (grant no. 1213520N).

Conflict of interest

The authors declare no conflict of interest.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web site.

Fig. S1. Representative pictures of garden and greenhouse experimental growth setups.

Fig. S2. Nodule sizes of Primus, Bettina, and Shouna cultivars.

Fig. S3. Nodule-producing soil types organized per experimental setup and cultivar

Fig. S4. Principal coordinate analysis of the bacterial nodule-colonizing communities.

Fig. S5. Diversity of bacteria co-appearing with possible rhizobia.

Fig. S6. Phylogenomic tree of the *Rhizobium* strains.

Fig. S7. Phylogenomic tree of the *Neorhizobium* strains.

Fig. S8. Phylogenomic tree of the *Bradyrhizobium* strains.

Fig. S9. Microbial diversity within the nodule harvested from cultivar Shouna grown in soil #76 (nodule S101).

Table. S1. Bacterial taxonomy classifications of pure single-colony isolates of harvested nodules based on 16S rRNA PCR sequences (xlsx. File)

Table. S2. Genomic description of the different strains.

Table. S3. Summary of nodule confirmation experiments of vermiculite-grown soybean plants in the greenhouse under optimal conditions.

References

Adhikari, D., Kaneto, M., Itoh, K. et al. (2013) Genetic diversity of soybean-nodulating rhizobia in Nepal in relation to climate and soil properties. *Plant Soil* **357**: 131–145.

Alexandre, A., and Oliveira, S. (2013) Response to temperature stress in rhizobia. Crit Rev Microbiol 39: 219-228.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410.

Ankenbrand, M.J., and Keller, A. (2016) bcgTree: automatized phylogenetic tree building from bacterial core genomes. *Genome* **59**: 783-791.

Aper, J., De Clercq, H., and Baert, J. (2016) Agronomic characteristics of early-maturing soybean and implications for breeding in Belgium. *Plant Genet Resour* **14**: 142-148.

Artigas Ramírez, M.D., España, M., Aguirre, C., Kojima, K., Ohkama-Ohtsu, N., Sekimoto, H., and Yokoyama, T. (2019) *Burkholderia* and *Paraburkholderia* are predominant soybean rhizobial genera in Venezuelan soils in different climatic and topographical regions. *Microbes Environ* **34**: 43-58.

Aserse, A.A., Woyke, T., Kyrpides, N.C., Whitman, W.B., and Lindström, K. (2017) Draft genome sequences of *Bradyrhizobium shewense* sp. nov. ERR11^T and *Bradyrhizobium yuanmingense* CCBAU 10071^T. *Stand Genomic Sci* **12**: 74.

Bai, Y., Zhou, X., and Smith, D.L. (2003) Enhanced soybean plant growth resulting from coinoculation of *Bacillus* strains with *Bradyrhizobium japonicum*. *Crop Sci* **43**: 1774-1781.

Beijerinck, M.W. (1888) Die Bacterien der Papilionaceen-knöllchen. *Bot Zeitung* **46**: 725-804.

Beirinckx, S., Viaene, T., Haegeman, A., Debode, J., Amery, F., Vandenabeele, S. et al. (2020) Tapping into the maize root microbiome to identify bacteria that promote growth under chilling conditions. *Microbiome* **8**: 54.

Beukes, C.W., Boshoff, F.S., Phalane, F.L., Hassen, A.I., le Roux, M.M., Stępkowski, T. et al. (2019) Both alpha- and beta-rhizobia occupy the root nodules of *Vachellia karroo* in South Africa. *Front Microbiol* **10**: 1195.

Blondon F. (1964) Contribution à l'étude du développement de graminées fourragères: ray-grass et dactyle. *Rev Gen Bot* **71**: 293–381

Boerema, A., Peeters, A., Swolfs, S., Vandevenne, F., Jacobs, S., Staes, J., and Meire, P. (2016) Soybean trade: balancing environmental and socio-economic impacts of an intercontinental market. *PLoS ONE* **11**: e0155222.

Bromfield, E.S.P., Cloutier, S., and Nguyen, H.D.T. (2020) Description and complete genome sequences of *Bradyrhizobium symbiodeficiens* sp. nov., a non-symbiotic bacterium associated with legumes native to Canada. *Int J Syst Evol Microbiol* **70**: 442-449.

Camacho M., Santamaría C., Temprano F., Rodríguez-Navarro D.N., Daza A., Espuny R., Bellogín R., Ollero F.J., Lyra de M.C., Buendía-Clavería A., Zhou J., Li F.D., Mateos C., Velázquez E., Vinardell J.M., Ruiz-Sainz J.E. (2002) Soils of the Chinese Hubei province show a very high diversity of Sinorhizobium fredii strains. *Syst Appl Microbiol* **25**:592-602.

Clúa J., Roda C., Zanetti M.E., Blanco F.A. (2018) Compatibility between Legumes and Rhizobia for the Establishment of a Successful Nitrogen-Fixing Symbiosis. *Genes* **9** (3):125

De Meyer, S.E., Van Hoorde, K., Vekeman, B., Braeckman, T., and Willems, A. (2011) Genetic diversity of rhizobia associated with indigenous legumes in different regions of Flanders (Belgium). *Soil Biol Biochem* **43**: 2384-2396.

Gyaneshwar, P., Hirsch, A.M., Moulin, L., Chen, W.-M., Elliott, G.N., Bontemps, C. et al. (2011) Legume-nodulating Betaproteobacteria: diversity, host range, and future prospects. *Mol Plant-Microbe Interact* **24**: 1276-1288.

Habibi, S., Ayubi, A.G., Ohkama-Ohtsu, N., Sekimoto, H., and Yokoyama, T. (2017) Genetic characterization of soybean rhizobia isolated from different ecological zones in North-eastern Afghanistan. *Microbes Environ* **32**: 71-79.

Hammond, E.G., Murphy, P.A., and Johnson, L.A. (2003) SOY (SOYA) BEANS | Properties and analysis. In *Encyclopedia of Food Sciences and Nutrition, 2nd Ed*. Caballero, B. (ed). Oxford: Academic Press, pp. 5389-5392.

Han, Q., Ma, Q., Chen, Y., Tian, B., Xu, L., Bai, Y. et al. (2020) Variation in rhizosphere microbial communities and its association with the symbiotic efficiency of rhizobia in soybean. *ISME J* **14**: 1915-1928.

Hilário de Carvalho, R., da Conceição Jesus, E., Favero, V.O., Straliotto, R., and Araújo, A.P. (2020) The co-inoculation of *Rhizobium* and *Bradyrhizobium* Increases the early nodulation and development of common beans. *J Soil Sci Plant Nutr* **20**: 860-864.

Hiltenbrand R., Thomas J., McCarthy H., Dykema K.J., Spurr A., Newhart H., Winn M.E., Mukherjee A. (2016) A Developmental and Molecular View of Formation of Auxin-Induced Nodule-Like Structures in Land Plants. *Frontiers in Plant Science* **7**:1692

Jordan, D. C. (1982) Transfer of Rhizobium japonicum Buchanan 1980 to Bradyrhizobium gen. nov., a Genus of Slow-Growing, Root Nodule Bacteria from Leguminous Plants. *International Journal of Systematic and Evolutionary Microbiology* **32**:136-139.

Karasev, E.S., Andronov, E.E., Aksenova, T.S., Chizhevskaya, E.P., Tupikin, A.E., and Provorov, N.A. (2019) Evolution of goat's rue rhizobia (*Neorhizobium galegae*): analysis of polymorphism of the nitrogen fixation and nodule formation genes. *Russ J Genet* **55**: 263-266.

Kim, D., Song, L., Breitwieser, F.P., and Salzberg, S.L. (2016) Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res* **26**: 1721-1729.

Kuklinsky-Sobral, J., Araújo, W.L., Mendes, R., Geraldi, I.O., Pizzirani-Kleiner, A.A., and Azevedo, J.L. (2004) Isolation and characterization of soybean-associated bacteria and their potential for plant growth promotion. *Environ Microbiol* **6**: 1244-1251.

Liu, F., Hewezi, T., Lebeis, S.L., Pantalone, V., Grewal, P.S., and Staton, M.E. (2019) Soil indigenous microbiome and plant genotypes cooperatively modify soybean rhizosphere microbiome assembly. *BMC Microbiol* **19**: 201.

Lu, J., Yang, F., Wang, S., Ma, H., Liang, J., and Chen, Y. (2017) Co-existence of rhizobia and diverse non-rhizobial bacteria in the rhizosphere and nodules of *Dalbergia odorifera* seedlings inoculated with *Bradyrhizobium elkanii*, *Rhizobium multihospitium*–like and *Burkholderia pyrrocinia*–like strains. *Front Microbiol* **8**: 2255.

Luo, J., Lyu, M., Chen, R., Zhang, X., Luo, H., and Yan, C. (2019) SLR: a scaffolding algorithm based on long reads and contig classification. *BMC Bioinformatics* **20**: 539.

Martínez-Hidalgo, P., and Hirsch, A.M. (2017) The nodule microbiome: N_2-fixing rhizobia do not live alone. *Phytobiomes J* **1**: 70-82.

Menzel, P., Ng, K.L., and Krogh, A. (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* **7**: 11257.

Miransari, M., Riahi, H., Eftekhar, F., Minaie, A., and Smith, D.L. (2013) Improving soybean (*Glycine max* L.) N_2 fixation under stress. *J Plant Growth Regul* **32**: 909-921.

Naamala, J., Jaiswal, S.K., and Dakora, F.D. (2016) Microsymbiont diversity and phylogeny of native bradyrhizobia associated with soybean (*Glycine max* L. Merr.) nodulation in South African soils. *Syst Appl Microbiol* **39**: 336-344.

Niemann, S., Pühler, A., Tichy, H.-V., Simon, R., and Selbitschka, W. (1997) Evaluation of the resolving power of three different DNA fingerprinting methods to discriminate among isolates of a natural *Rhizobium meliloti* population. *J Appl Microbiol* **82**: 477-484.

Ondov, B.D., Bergman, N.H., and Phillippy, A.M. (2011) Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics* **12**: 385.

Pannecoucque, J., Goormachtigh, S., Ceusters, J., Debode, J., Van Waes, C., and Van Waes, J. (2018a) Temperature as a key factor for successful inoculation of soybean with *Bradyrhizobium* spp. under cool growing conditions in Belgium. *J Agric Sci* **156**: 493-503.

Pannecoucque, J., Goormachtigh, S., Heungens, K., Vleugels, T., Ceusters, J., Van Waes, C., and Van Waes, J. (2018b) Screening for soybean varieties suited to Belgian growing conditions based on maturity, yield components and resistance to *Sclerotinia sclerotiorum* and *Rhizoctonia solani* anastomosis group 2-2IIIB. *J Agric Sci* **156**: 342-349.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043-1055.

Rodríguez-Navarro, D.N., Margaret Oliver, I., Albareda Contreras, M. et al. (2011) Soybean interactions with soil microbes, agronomical and molecular aspects. *Agronomy Sust. Developm.* **31**: 173–190

Sánchez, A.C., Gutiérrez, R.T., Santana, R.C., Urrutia, A.R., Fauvart, M., Michiels, J., and Vanderleyden, J. (2014) Effects of co-inoculation of native *Rhizobium* and *Pseudomonas* strains on growth parameters and yield of two contrasting *Phaseolus vulgaris* L. genotypes under Cuban soil conditions. *Eur J Soil Biol* **62**: 105-112.

Schmidt, J., Messmer, M., and Wilbois, K.-P. (2015) Beneficial microorganisms for soybean (*Glycine max* (L.) Merr), with a focus on low root-zone temperatures. *Plant Soil* **397**: 411-445.

Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**: 2068-2069.

Sharaf, H., Rodrigues, R.R., Moon, J., Zhang, B., Mills, K., and Williams, M.A. (2019) Unprecedented bacterial community richness in soybean nodules vary with cultivar and water status. *Microbiome* **7**: 63.

Sinclair, T.R., Soffes, A.R., Hinson, K., Albrecht, S.L., and Pfahler, P.L. (1991) Genotypic variation in soybean nodule number and weight. *Crop Sci* **31**: 301-304.

Talbi, C., Delgado, M.J., Girard, L., Ramírez-Trujillo, A., Caballero-Mellado, J., and Bedmar, E.J. (2010) *Burkholderia phymatum* strains capable of nodulating *Phaseolus vulgaris* are present in Moroccan soils. *Appl Environ Microbiol* **76**: 4587-4591.

Temesgen, D., Maluk, M., James, E.K., Iannetta, P.P.M., and Assefa, F. (2019) The functional characterisation of soybean (*Glycine max* L.) rhizospheric bacteria indigenous to Ethiopian soils. *Afr J Agric Res* **14**: 1659-1673.

Thomas-Oates J., Bereszczak J., Edwards E., Gill A., Noreen S., Zhou J.C., Chen M.Z., Miao L.H., Xie F.L., Yang J.K., Zhou Q., Yang S.S., Li X.H., Wang L., Spaink H.P., Schlaman H.R., Harteveld M., Díaz C.L., van Brussel A.A., Camacho M., Rodríguez-Navarro D.N., Santamaría C., Temprano F., Acebes J.M., Bellogín R.A., Buendía-Clavería A.M., Cubo M.T., Espuny M.R., Gil A.M., Gutiérrez R., Hidalgo A., López-Baena F.J., Madinabeitia N., Medina C., Ollero F.J., Vinardell J.M., Ruiz-Sainz J.E.. (2003) A catalogue of molecular, physiological and symbiotic properties of soybean-nodulating rhizobial strains from different soybean cropping areas of China. *Syst Appl Microbiol* **26**:453-65.

Wood, D.E., Lu, J., and Langmead, B. (2019) Improved metagenomic analysis with Kraken 2. *Genome Biol* **20**: 257.

Wu, Y.-W., Simmons, B.A., and Singer, S.W. (2015) MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**: 605-607.

Yamashita, N., Tanabata, S., Ohtake, N., Sueyoshi, K., Sato, T., Higuchi, K., Saito, A., and Ohyama, T. (2019). Effects of Different Chemical Forms of Nitrogen on the Quick and Reversible Inhibition of Soybean Nodule Growth and Nitrogen Fixation Activity. *Frontiers in Plant Science* **10,** doi:10.3389/fpls.2019.00131

Yoon, S.-H., Ha, S.-m., Lim, J., Kwon, S., and Chun, J. (2017) A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek* **110**: 1281-1286.

Yuan, K., Reckling, M., Artigas Ramirez, M.D., Djedidi, S., Fukuhara, I., Ohyama, T. et al. (2020) Characterization of rhizobia for the Improvement of soybean cultivation at cold conditions in Central Europe. *Microbes Environ* **35**: ME19124.

Zhang, F., Lynch, D.H., and Smith, D.L. (1995) Impact of low root temperatures in soybean [*Glycine max*. (L.) Merr.] on nodulation and nitrogen fixation. *Environ Exp Bot* **35**: 279-285.

Zhao, R., Liu, L.X., Zhang, Y.Z., Jiao, J., Cui, W.J., Zhang, B. et al. (2018) Adaptive evolution of rhizobial symbiotic compatibility mediated by co-evolved insertion sequences. *ISME J* **12**: 101-111.

Zhong, Y., Yang, Y., Liu, P., Xu, R., Rensing, C., Fu, X., and Liao, H. (2019) Genotype and rhizobium inoculation modulate the assembly of soybean rhizobacterial communities. *Plant Cell Environ* **42**: 2028-2044.

Zimmer, S., Messmer, M., Haase, T., Piepho, H.-P., Mindermann, A., Schulz, H. et al. (2016) Effects of soybean variety and *Bradyrhizobium* strains on yield, protein content and

biological nitrogen fixation under cool growing conditions in Germany. *Eur J Agron* **72**: 38-46.

Table 1. Nodules from Flemish soils in garden and greenhouse experiments.
G, garden; GH, greenhouse; P, Primus cv., B; Bettina cv., S, Shouna cv; red/pink nodules are indicated as fixing. The available nodule microbiome dataset indicated with "X". Soils that resulted in nodules in both setups are underlined.

Soil	Exp.	Province	Soil type	Cultivar	Nodule characteristics	Rhizobial isolates in collection	Microbiome
#1	GH	Flemish Brabant	Clay	P, S	Small		
#4	G	E Flanders	Anthropogenic	S, B	Regular		X (S, B)
#12	GH	E Flanders	Sandloam	S, B	Large fixing (S); small (B)	*Rhizobium* and *Neorhizobium* sp. (S, B)	X (S)
#13	G	E Flanders	Sandloam	P, S	Round (P) and irregular shaped (S) nodules		X (P)
#17	GH	Brussels	Anthropogenic	P	Small		X
<u>#19</u>	G, GH	E Flanders	Anthropogenic	P, S, B	G, small irregular; GH, large fixing (P, B) and small white (S)	*Rhizobium* sp. (GH, B) and *Neorhizobium* sp. (G, P)	X (GH, S; G, P, S, B)
#26	GH	E Flanders	Sandloam	P	Small, white		X
#27	GH	E Flanders	Anthropogenic	P, S, B	Small		
#38	G	Antwerp	Sand	P, S, B	Irregular, white		X (P, B)
#41	GH	Antwerp	Sand	P, B	Nonfixing		
#42	G	Antwerp	Anthropogenic	P, S	Large and regular nodules	*Bradyrhizobium* sp. (P)	X (P, S)
<u>#43</u>	G, GH	E Flanders	Loam	P, S, B	G, small, regular, (S, P); GH, small, and large red (P, B), small (S)		X (GH, P, S; G, S)
#48	GH	W Flanders	Other	S	Small, white		
#50	G	Flemish Brabant	Loam	P	Large, fixing and small		X
#51	GH	W Flanders	Sandloam	P, B	Many small		X (P)
#53	GH	Flemish Brabant	Sandloam	P, B	Small		
#54	GH	E Flanders	Sand	P, S, B	Small irregular	*Rhizobium* sp. (S)	
#55	G	Flemish Brabant	Sandloam	S	Many small		X
#60	GH	Antwerp	Sand	B	Many small		
#61	G	Antwerp	Sandloam	P	Many small		
#64	G	W Flanders	Sandloam	P	Fixing and regular shapes (P)	*Rhizobium* sp.	X
#65	GH	E Flanders	Anthropogenic	P, B	White		
#66	G	E Flanders	Sand	P	Large		

<u>#67</u>	G, GH	E Flanders	Sand	P, B	G, small, white (P); GH, small (P and B)		X (GH, P; G, P)
#69	GH	W Flanders	Sandloam	P, B	Small		
#70	GH	E Flanders	Sand	B	One large irregular nodule		
#74	GH	Flemish Brabant	Loam	P	Small		X
<u>#76</u>	G, GH	E Flanders	Sand	P, S	G, few, regular, round fixing (S); GH, small, white (P)	*Bradyrhizobium* sp. (G, S)	X (GH, P; G, S)
#77	G	E Flanders	Sand	P, S, B	Many irregular		X (G, P, S, B)
#80	G	W Flanders	Sand	P	Small		
#86	G	Antwerp	Sand	P, S	Few		
#93	GH	E Flanders	Sandloam	P	One large fixing	*Bradyrhizobium* sp.	X
<u>#94</u>	G, GH	W Flanders	Clay	P, S, B	G, many, small, irregular; GH, small, white		X (GH, P)
#107	G	E Flanders	Anthropogenic	P	Many small		X

Table 2. Candidate *Bradyrhizobium, Neorhizobium* and *Rhizobium* sp. strains.

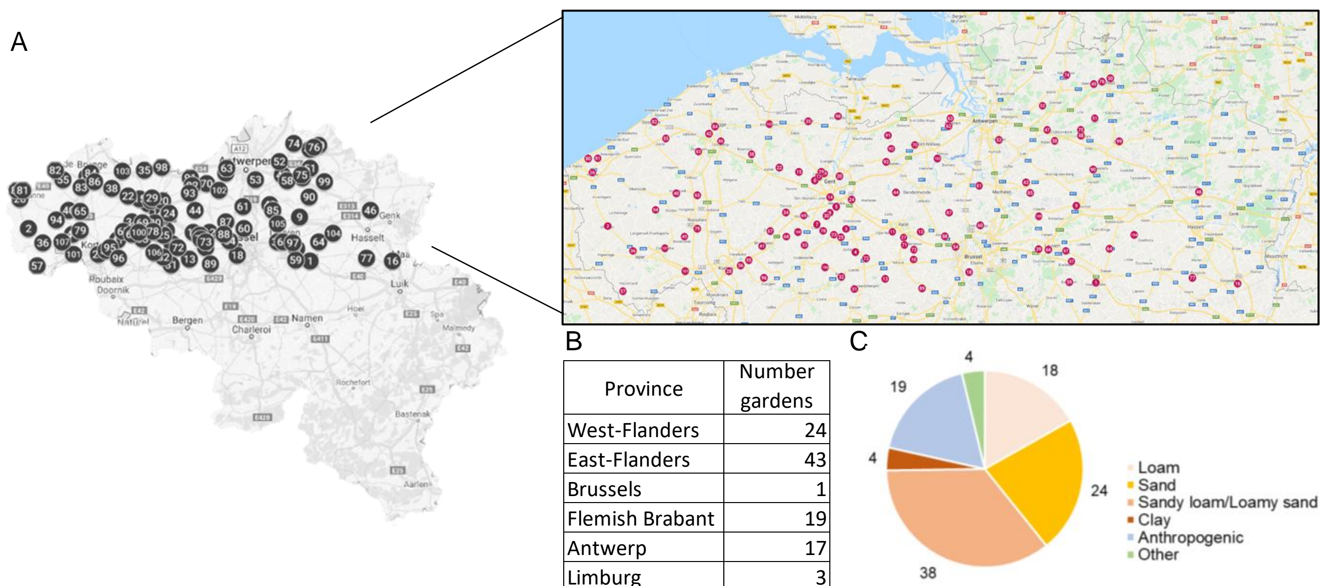
Nick Name	**Assigned Name**	**Experiment**	**Cultivar**	**Soil**	**16S rRNA identity**	**ASV identity**
PRIMUS64	*Rhizobium* sp. PRIMUS64	Garden	Primus	#64	*Rhizobium* sp.	ASV2
SHOUNA76	*Bradyrhizobium* sp. SHOUNA76	Garden	Shouna	#76	*Bradyrhizobium* sp.	ASV1
PRIMUS42	*Bradyrhizobium* sp. PRIMUS42	Garden	Primus	#42	*Bradyhizobium* sp.	ASV1
BETTINA12A	*Neorhizobium* sp. BETTINA12A	Greenhouse	Bettina	#12	*Neorhizobium* sp.	ASV44
BETTINA12B	*Rhizobium* sp. BETTINA12B	Greenhouse	Bettina	#12	*Rhizobium* sp.	ASV2
SHOUNA12A	*Neorhizobium* sp. SHOUNA12A	Greenhouse	Shouna	#12	*Neorhizobium* sp.	ASV44
SHOUNA12B	*Neorhizobium* sp. SHOUNA12B	Greenhouse	Shouna	#12	*Neorhizobium* sp.	ASV44
SHOUNA12C	*Rhizobium* sp. SHOUNA12C	Greenhouse	Shouna	#12	*Rhizobium* sp.	ASV2

Table 3. Average Nucleotide Identity (ANI) values of the *Neorhizobium* strains.

Closest relative	***Neorhizobium* sp. SHOUNA12B**	***Neorhizobium* sp. SHOUNA12A**	***Neorhizobium* sp. BETTINA12A**
N. galegae bv. *officinalis* HAMBI 1141	97.23	97.23	97.28
N. galegae bv. *orientalis* HAMBI 540^T	94.99	95.06	95.00
N. galegae Ng87	94.8	94.79	94.96

Table 4. Average Nucleotide Identity (ANI) values of the *Bradyrhizobium* strains.

Closest relative	*Bradyrhizobium* sp. SHOUNA76	*Bradyrhizobium* sp. PRIMUS42
B. shewense ERR11	89.96	89.92
B. ottawaense OO99	89.96	89.90
B. symbiodeficiens 85S1MB	89.56	89.46
B. amphicarpaeae 39S1MB	89.36	89.25



Province	Number gardens
West-Flanders	24
East-Flanders	43
Brussels	1
Flemish Brabant	19
Antwerp	17
Limburg	3

Fig 1. Geographic distribution of the participating gardens and the different soil types in Flanders (Belgium). (A) Distribution of the 107 locations used to grow soybeans and collect soil samples. Left, sample sites indicated with numbers right, close-up of Flanders with locations of participating gardens highlighted with purple circles. (B) Number of participating gardens per Flemish provinces. (C) Number of garden soils predicted to belong to a specific soil type: loam, sand, sandy loam or loamy sand, clay, anthropogenic (built-up areas), and other soil types, such as polders and dunes.

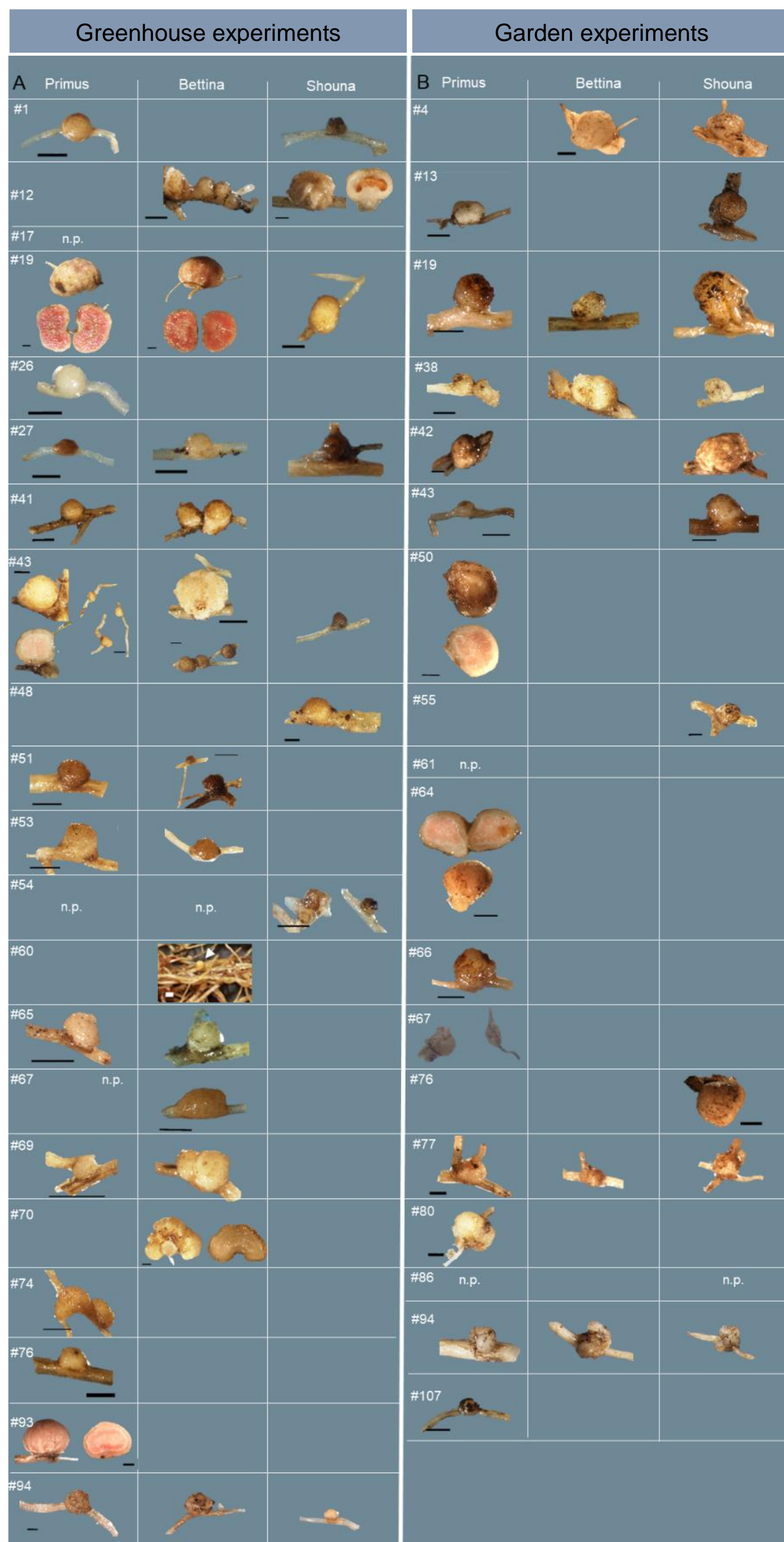


Fig 2. Nodules obtained in garden and greenhouse experiments. Plants were grown in 107 Flemish soils under (A) greenhouse conditions and (B) in gardens under natural conditions during the summer of 2019 (May-July/August). In each soil, three soybean cultivars Primus, Bettina, and Shouna were grown and after a minimum of two months, nodulation was assessed.

Of the 107 soils, 18 and 21 resulted in nodules on at least one plant in the gardens and under greenhouse conditions, respectively. In five soils, nodules occurred in both experimental setups. n.p., no representative picture available. Scale bars, 1 mm.

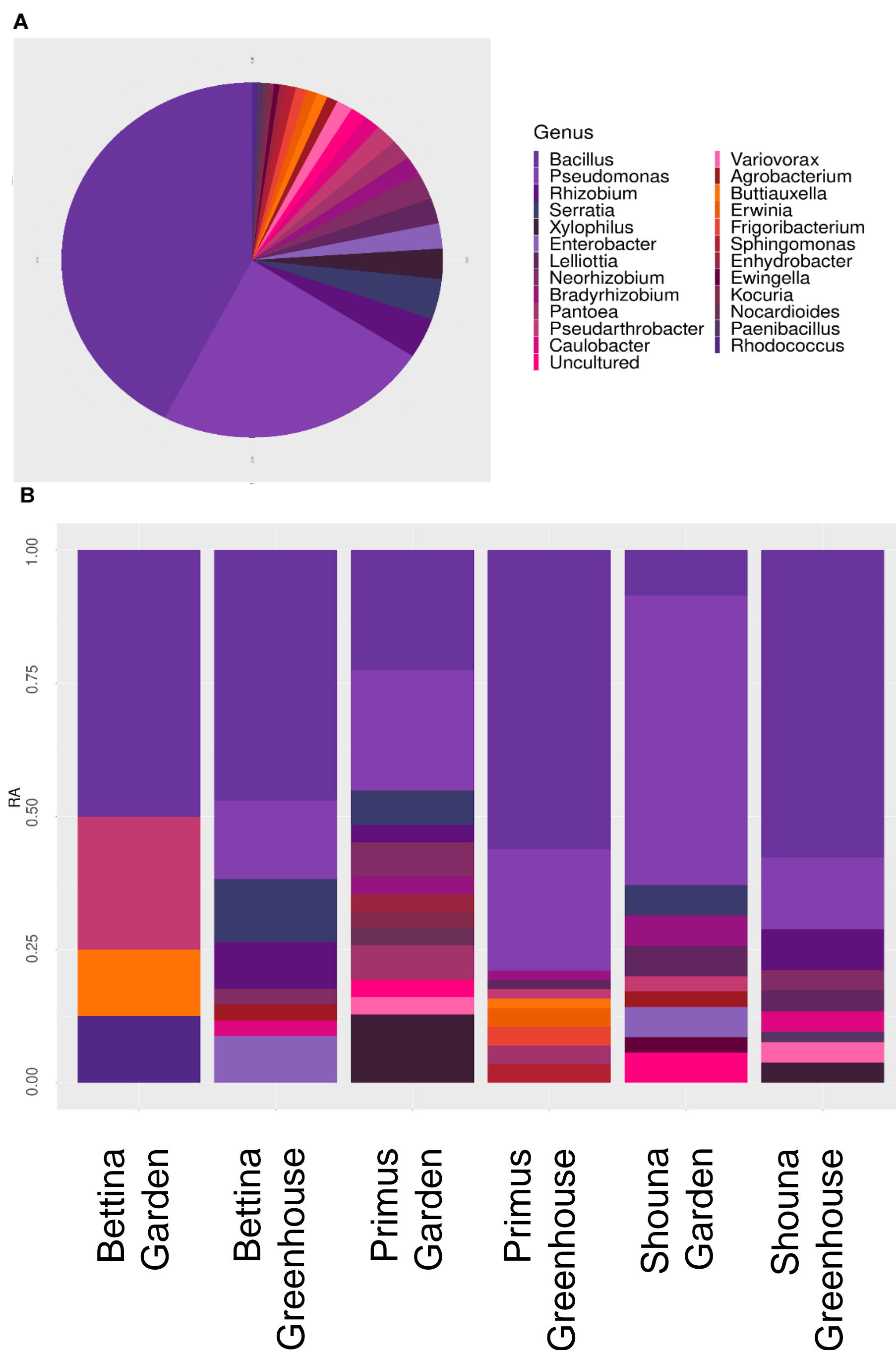


Fig 3. Bacterial isolates from nodules of plants grown in gardens and in the greenhouse. One nodule per plant, cultivar, and experimental setup was crushed and plated on selective media to allow bacterial growth. (A)

Pie diagram of 218 pure single-colony bacterial isolates identified at the taxonomical genus level based on 16S rRNA PCR sequences. (B) Relative abundance of bacterial genera in nodules harvested from each soybean cultivar in garden or greenhouse experimental setup.

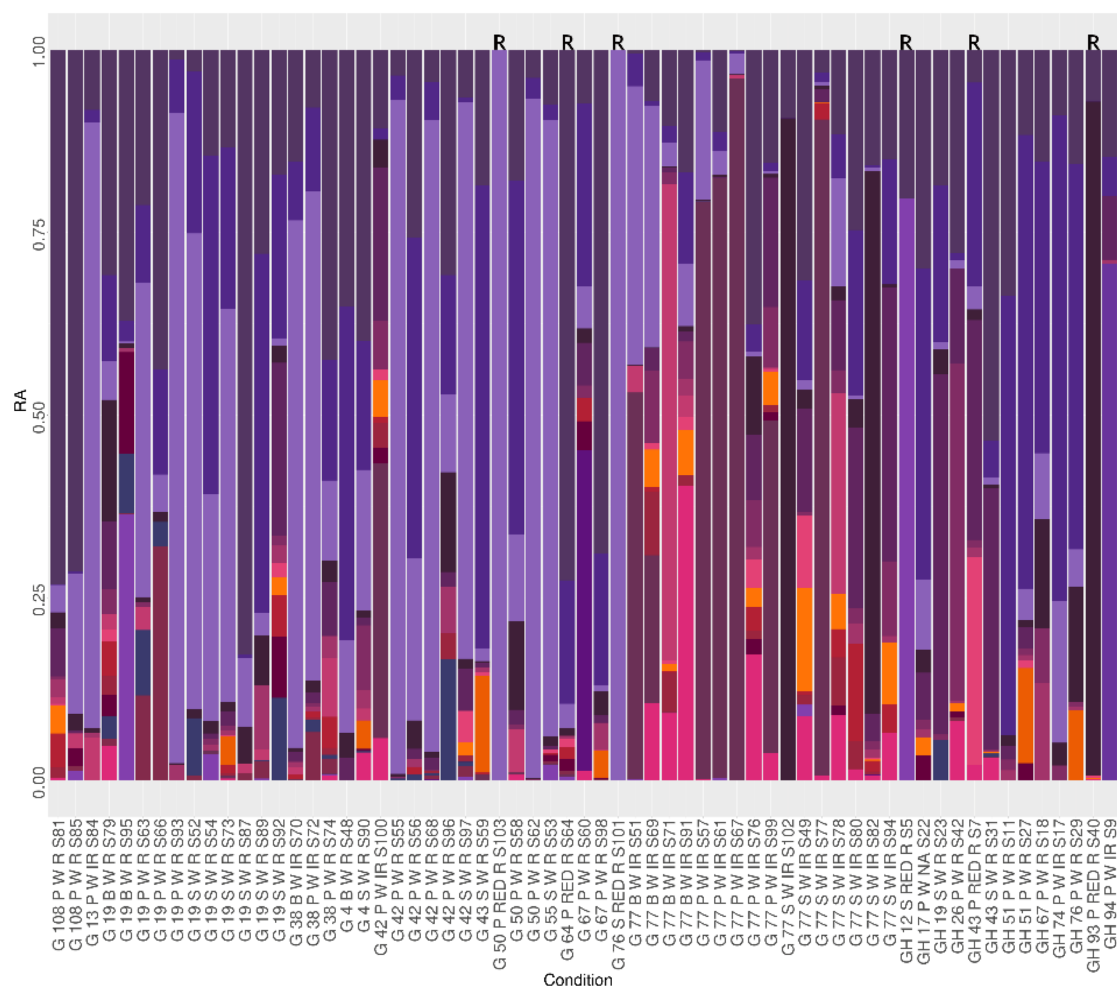


Fig 4. Microbial diversity of the soybean nodules. Representation of the 20 most abundant ASVs across the nodule dataset. Each bar constitutes a sample, with indication of G, garden; GH, greenhouse; X, soil number; P, cv. Primus; B, cv. Bettina; S, cv. Shouna (S); W, white

nodule; RED, red nodule, R, regular; and IR, irregular; and S, nodule number. Red nodules are marked with an R at the top of the chart.

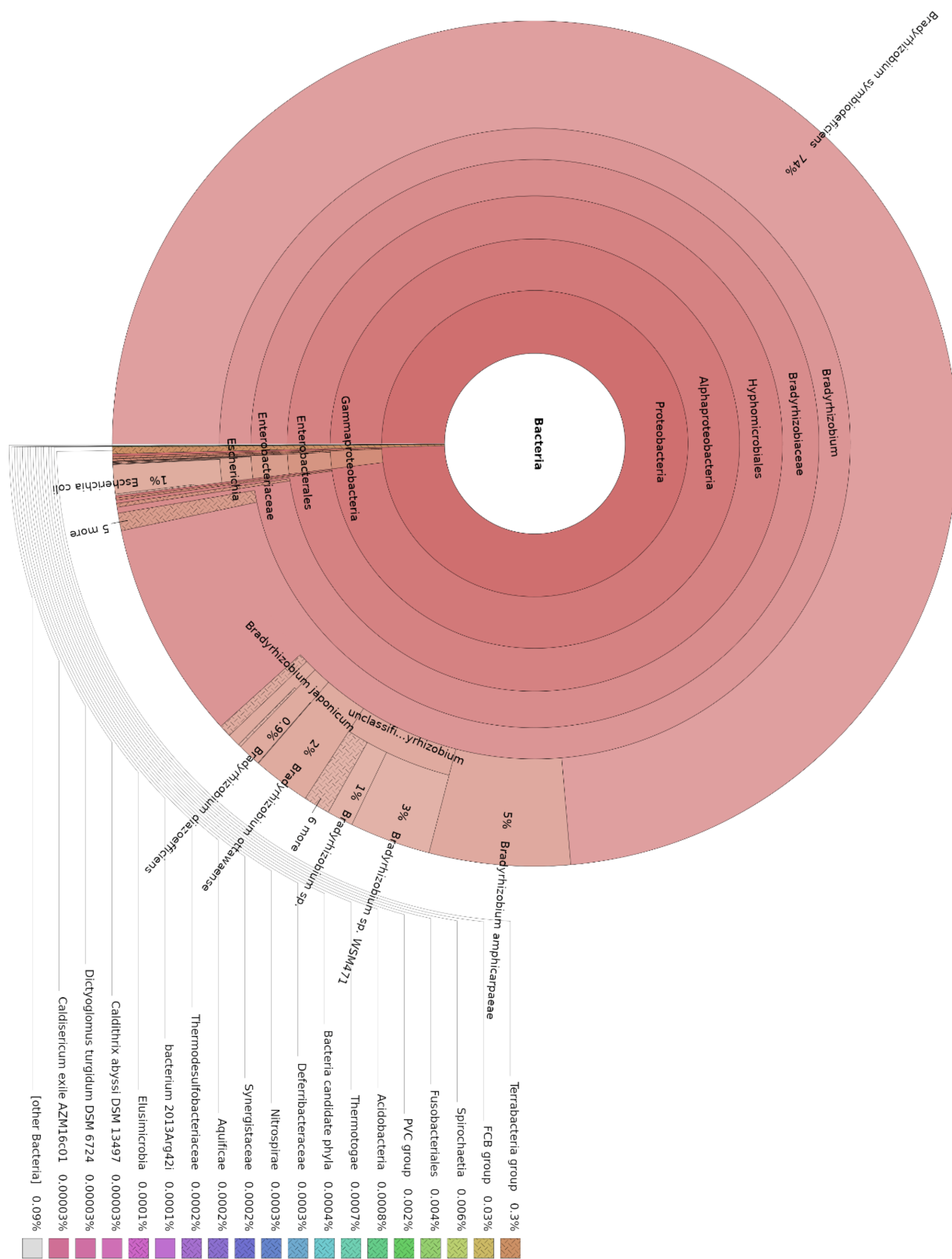


Fig 5. Microbial diversity within the nodule harvested from cultivar Shouna grown in soil #76 (nodule S101). Krona chart representing the taxonomical assignment of the ONT reads

belonging to the bacterial kingdom with the percentage of each identified group, by means of the MiniKraken software.

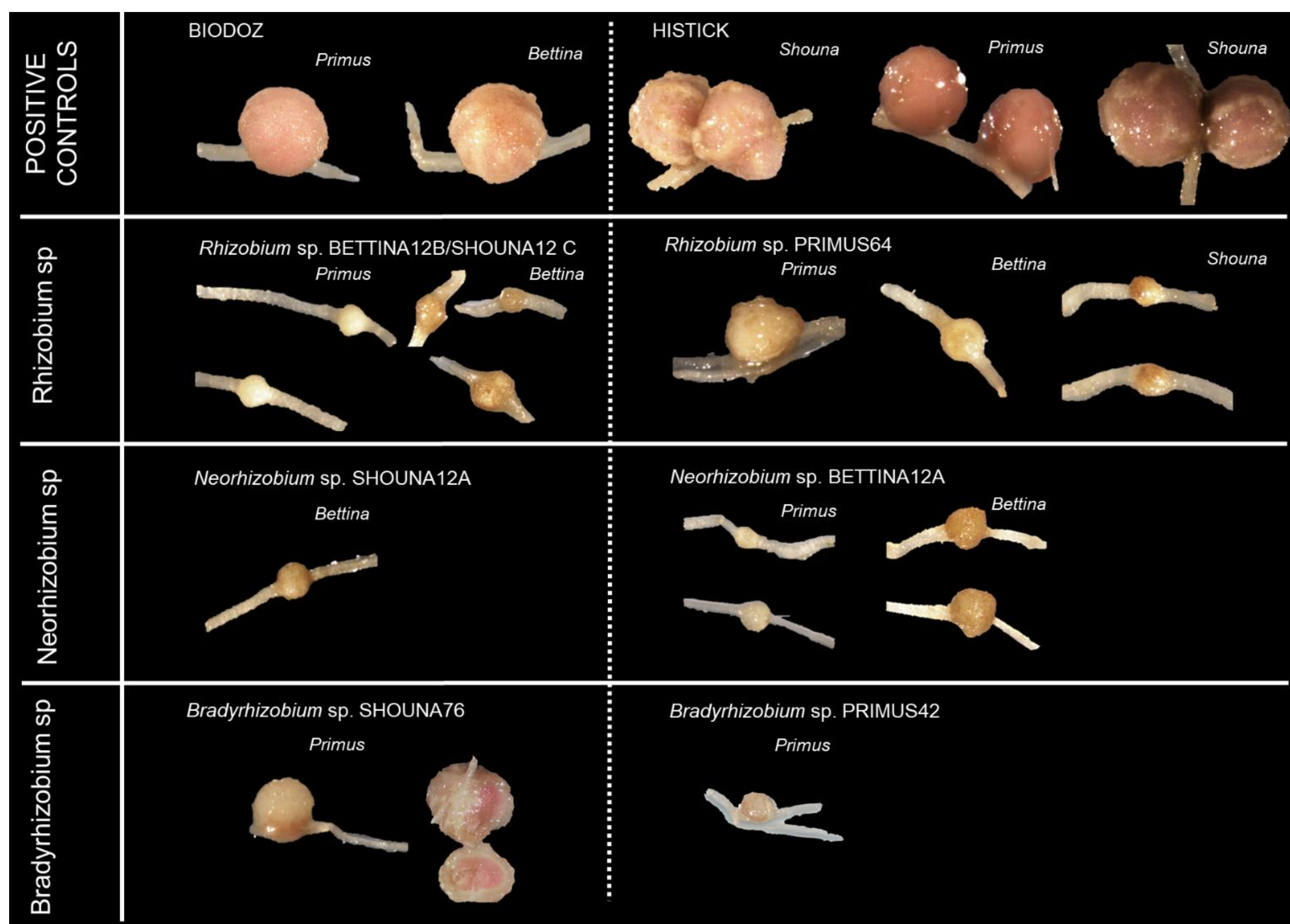


Fig 6. Nodules of vermiculite-grown soybean plants inoculated with candidate rhizobial strains. Soybean seedlings from the cultivars Primus, Bettina, and Shouna were inoculated with the candidate rhizobial strain at OD = 0.01 and after four weeks, nodules were assessed and photographed. As positive controls for nodulation, plants were inoculated with the commercial strains Biodoz and HiStick. Representative pictures of the resulting nodules are depicted.