Chromosomal integration of Golden Gate compatible plasmids in non-symbiotic Bradyrhizobium
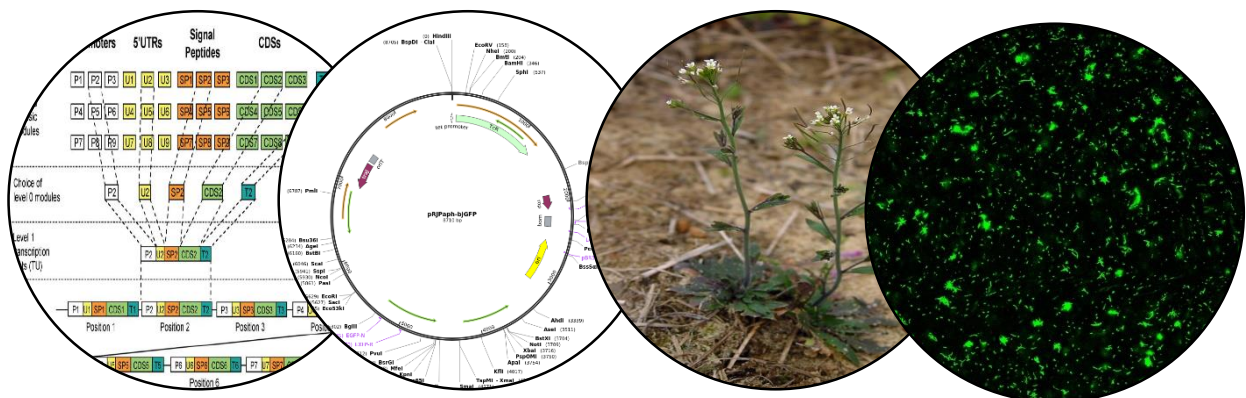
Name   : Paul Nijhuis
Institution   : Wageningen University
E-mail   : Paul.nijhuis@wur.nl
Study program   : Bsc. Biotechnology (BBT)
Registration number : 960509613050
Chair group   : Molecular biology
Supervisor   : Martinus Schneijderberg

Abstract

Bradyrhizobium spp. are well known for their capacity to initiate a beneficial relationship with a large variety of legume species and the non-legume *Parasponia.* Through initiation of nodulogenesis in plants, *Bradyrhizobium* spp. are able to utilize organic molecules derived from photosynthesis in exchange for fixed nitrogen. The genes essential for the nitrogen-fixation and nodulation trait are clustered chromosomally on so called symbiotic islands, which are wildly accepted to be of foreign origin. However, previously published *Bradyrhizobium* ecotypes have been shown to be non-symbiotic. These non-symbionts were found to be highly abundant in soils lacking legumes and do not harbour the nitrogen fixation and nodulation gene clusters. Recently, a study showed that a multitude of *Bradyrhizobium* strains are highly abundant in the root endophytic compartment of eight different plant species originating from the same ecological test field. From a single *Arabidopsis thaliana* (*Arabidopsis*) root, four genetically and functionally highly diverse *Bradyrhizobium* strains were isolated. These isolates, designated *Bradyrhizobium* sp. MOS001 to MOS004, were shown to be non-symbiotic. However, the nature of the interaction between these non-symbiotic *Bradyrhizobium* isolates and the root endophytic compartment remained elusive. In this study, we show the first step to identify the nature of this host-microbe interaction through localization of these non-symbionts on roots of model species *Arabidopsis*. Moreover, we show the construction of strain specific plasmids which become stably integrated at a uniform locus in the chromosome of MOS001, MOS004 and type strain *Bradyrhizobium diazoefficiens* USDA110 (USDA110). These constructs enable expression of a single copy gene for a *Bradyrhizobium* optimized green fluorescent protein. Here, we report successful construction of green fluorescent protein-expressing USDA110, MOS001 and MOS004 strains. In our experimental setup, we were unable to detect these labelled strains in the root endophytic compartment of *Arabidopsis.*

Contents

1 Background

A large variety of legume species can be triggered to initiate nodule formation. In nodules, symbiotic bacteria and archaea are able to exchange fixed atmospheric nitrogen for biomolecules derived from photosynthesis (1). Genes essential for the nitrogen fixation and nodulogenesis trait are either chromosomally clustered on symbiotic island or on symbiotic plasmids. (2). It is widely accepted that symbiotic islands have been distributed over symbionts by horizontal gene transfer (3). Transposability of symbiotic traits presumably enables bacteria with a free-living life style to gain a nitrogen-fixing symbiotic lifestyle by living in close association with symbionts (4) .

Bradyrhizobium spp. are well studied for their ability to initiate a host-microbe symbiosis with a variety of legume species and the non-legume *Parasponia*. However, recent studies show the existence of *Bradyrhizobium* ecotypes incapable of nitrogen fixation and nodulation of legume root hairs (5). A common trait of these free-living *Bradyrhizobium* spp. is the lack of nodulation and nitrogen fixation gene clusters and the presence of genes associated with complex carbon metabolism (6). These findings suggest that non-symbiotic *Bradyrhizobium* strains occupy distinct ecological niches in the soil, other than nitrogen fixation.

In addition, a recent study revealed *Bradyrhizobium* spp. to be highly abundant in the root endophytic compartment of eight different plant species originating from the same ecological test field (Figure 1) (7). This group of highly abundant *Bradyrhizobium* spp. was found to encompass at least four genetically and functionally highly diverse strains, which were isolated from a single *Arabidopsis thaliana* (*Arabidopsis*) root. These strains were designated *Bradyrhizobium* sp. MOS001 to MOS004. Further genomic analysis of the MOS strains showed absence of nitrogen fixation and nodulation gene clusters (Figure 2), implying the ability of non-symbiotic *Bradyrhizobium* spp. to colonize the endophytic compartment of *Arabidopsis*. However, the nature of this interaction remains elusive.
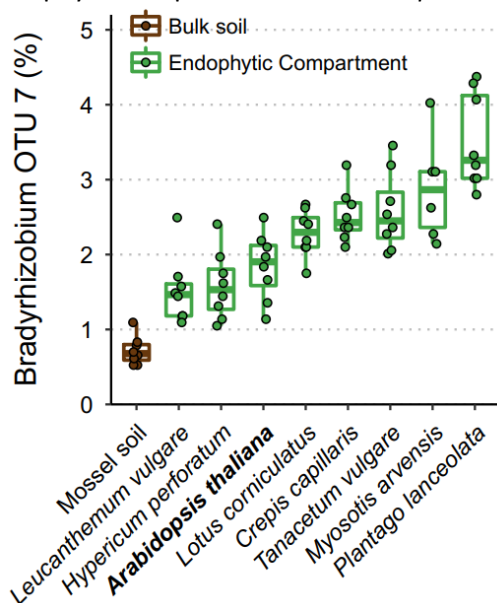

Figure 1 16S V4 rDNA based culture independent approach showing relative abundance of OTU 7 (%) in the root endophytic compartment of various plant species and in bulk soil. This OTU was later shown to encompass at least four genetically and functionally highly diverse *Bradyrhizobium* strains. Further genomic analysis showed these strains to be non-symbiotic
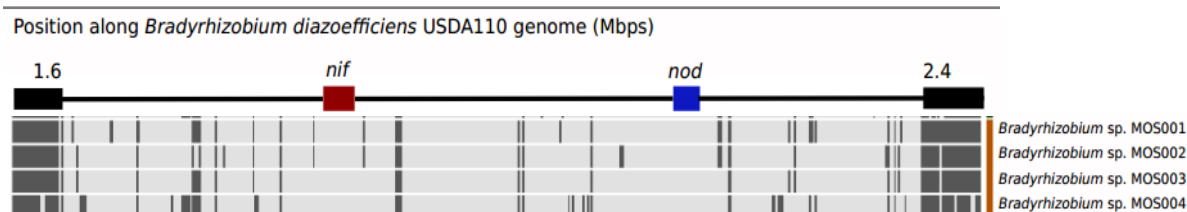
Figure 2 Reciprocal best BLAST hit results of MOS001 to MOS004 and type strain USDA110. These results show the absence of nitrogen fixation (nif) and nodulation (nod) gene clusters in all *Bradyrhizobium* spp. MOS strains.

A good starting point to determine the interactive behaviour between the *Bradyrhizobium* MOS. isolates and *Arabidopsis* is to localize where exactly the bacteria are located. In this study, we aimed at localization of the four *Bradyrhizobium* sp. MOS isolates, and type strain *Bradyrhizobium diazoefficiens* USDA110 (USDA110) on roots of *Arabidopsis.* Although, we were doubting whether transformation of the *Bradyrhizobium* sp. MOS strains would succeed with plasmid DNA because *Bradyrhizobium* spp. lacking endogenous plasmids are known to rapidly dispose plasmids in absence of selection pressure (e.g. in soil or root nodules) (8). To bypass this obstacle we intended to transform MOS001-MOS004 and USDA110 with a marker gene that stably integrates in the chromosome at a uniform locus for all strains. For integration to occur, the constructs must contain a strain specific region of homology. In a recent study, successful transformation of USDA110 was shown to work with a plasmid designated pRJPaph-bjGFP (9). This plasmid contains a *Bradyrhizobium* optimized GFP gene, a homologous sequence specific for USDA110 and a mobility region, making the plasmid compatible for bi-parental mating. However, because the region of homology is specific for USDA110, this plasmid cannot be used for transformation of the MOS isolates. Therefore, we intended to use this plasmid as template for the construction of MOS001 – MOS004 strain specific plasmids. Finally, we decided to limit ourselves to transformation of MOS001 and MOS004 because these strains have the largest genetic diversity of the four MOS strains. This enabled us to investigate whether our transformation protocol for *Bradyrhizobium* was expected to be useful for phylogenetically distinct *Bradyrhizobium* spp.

2   Results and discussion

2.1   Construction and assembly of Golden Gate compatible plasmids for transformation of USDA110, MOS001 and MOS004.

We aimed to investigate the nature of the interaction between the non-symbiotic *Bradyrhizobium* sp. MOS isolates and the endophytic compartment of *Arabidopsis*. To enable localization of these non-symbiotic strains we planned to transform these *Bradyrhizobia* with plasmids containing a single copy of a green fluorescent protein gene. Here, we show the construction and assembly of such plasmids.

2.1.1   Design and assembly of a *Bradyrhizobium* specific GFP module using Golden Gate modular cloning.

The construct used for successful transformation of USDA110 contains a *Bradyrhizobium* optimized green fluorescent protein gene (bjGFP). Because the bjGFP gene from pRJPaph-bjGFP (supplementary figures S1) contains a recognition sequence of a digestive enzyme used in Golden Gate (GG) modular cloning, we aimed to introduce a synonymous mutation in bjGFP to remove the cut site. We intended to make this polymorphism by extracting the N and C terminus of bjGFP separately form pRJPaph-bjGFP, using special primers containing the synonymous mutation. Back ligation of both terminuses

would result in a functional bjGFP gene lacking any recognition sites that may interfere with further cloning procedures.

We show the results of the extraction of bjGFP-N, bjGFP-C and Paph from pRJPaph-bjGFP in Figure 3 A. From these amplicons we created individual bjGFP and promoter (Paph) GG modules. Next, we sequenced both modules and aligned them with *in silico* versions of the constructs (data not shown). Both constructs showed a near perfect nucleotide alignment. However, we found a single adenosine insertion that was possibly present in a fraction of the reverse primer used to amplify Paph from pRJPaph-bjGFP (Figure *3* D). No deleterious effect of this adenosine insertion was expected because it was eleven nucleotides downstream of the annotated promoter region and not inside the open reading frame of bjGFP. When observing the nucleotide alignment of bjGFP from our GG module with bjGFP from pRJPaph-bjGFP, one is able to identify the induced synonymous mutation. We therefore conclude successful removal of the restriction site from bjGFP. From both gene and promoter modules a promoter-gene fusion was successfully assembled (Figure *3* C).
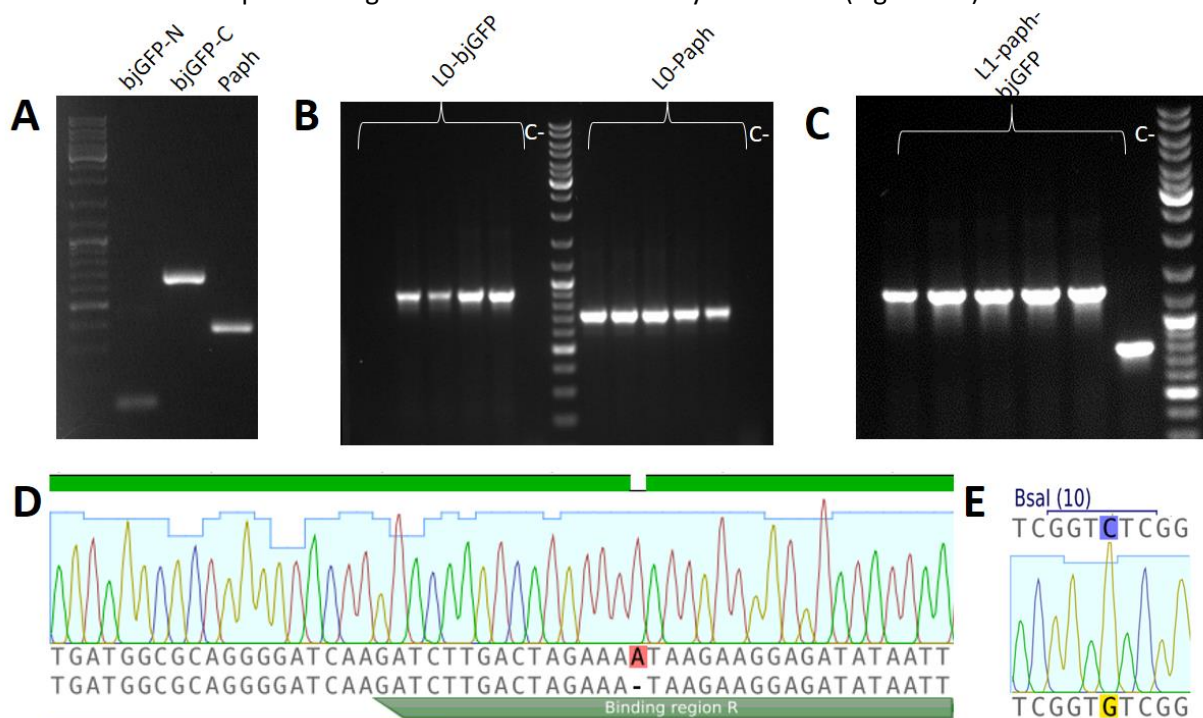
Figure 3 **A** bjGFP-N, bjGFP-C and Paph amplicons amplified from pRJPaph-bjGFP. Generuler DNA ladder mix was used as size marker. **B** Colony PCR indicating successful assembly of bjGFP and promoter GG modules. For both bjGFP and promoter modules, we used a forward primer that bound to the plasmid backbone and a reverse primer that bound to the insert. As negative control the empty acceptor vector was used. **C** Colony PCR experiment on *E.coli* transformants carrying the promoter-gene fusion modules. All transformants showed the fragment of expected size (1240 bp). Both forward and reverse primer bound to the acceptor vector backbone. As negative control we used the empty acceptor vector, resulting in amplicons of 766 bp. **D** Fraction of the nucleotide alignment from the GG promoter module with the *in silico* version of module. An insertion of one adenosine can be seen at the primer binding site. We expected this insertion to be present in a proportion of the reverse primer used to amplify Paph from pRJPaph-bjGFP. **E** Small section of the nucleotide alignment between bjGFP from the GG plasmid and bjGFP from pRJPaph-bjGFP. We deduce that the BsaI recognition sequence is successfully removed by inducting a synonymous mutation (C->G).

2.1.2 Amplification of strain specific homologous regions and assembly of the Golden Gate mobility region module.

For stable inheritance, we aimed to integrate the marker gene in the chromosomes of USD110, MOS001 and MOS004. However, for integration to occur the constructs required the presence of a homologous sequence for a uniform locus in the chromosome of each strain. Furthermore, a mobility region was necessary to make the constructs compatible for bi-parental mating.

The homologous regions for all four *Bradyrhizobium* sp. MOS isolates were successfully amplified from purified genomic DNA. (Figure 4A). We further extracted the region of homology, specific for USDA110, from the original plasmid (Figure 4 B). We additionally made a GG module for the mobility region (Figure 4 C, D). Sequencing results showed a perfect alignment between the mobility region module and the *in silico* version of the plasmid (data not shown).
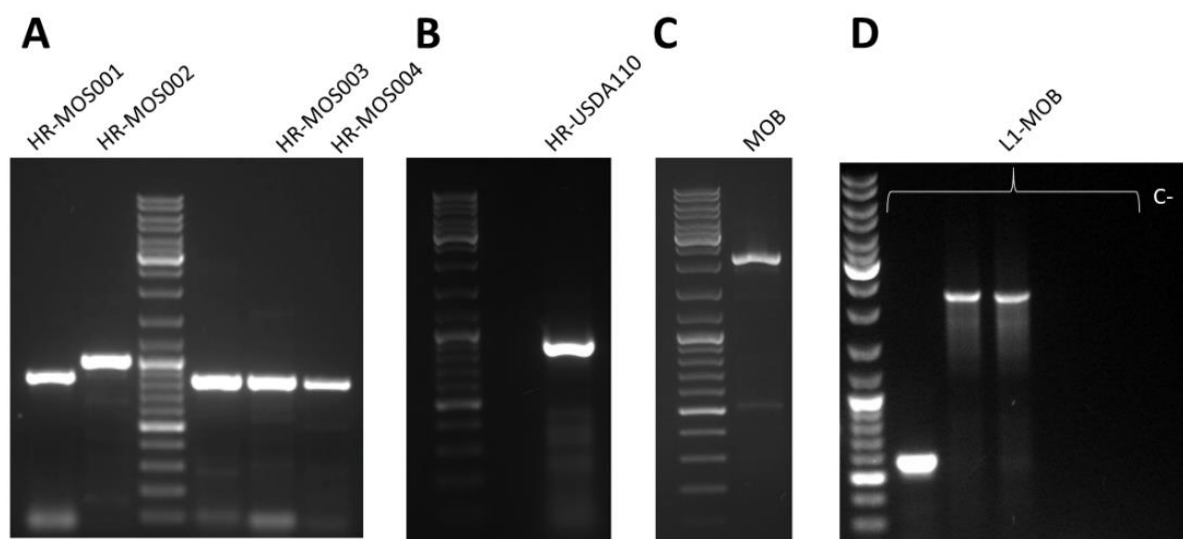

Figure 4 : **A** Amplification of the homologous region from genomic DNA of MOS001 – MOS004 (Toolbox65). Amplicons are of expected size (~800 bp). Generuler DNA ladder mix was used as size marker. **B** Amplification of USDA110 specific region of homology from pRJPaph-bjGFP. **C** Mobility region amplicon from pRJpaph-bjGFP. A note must be made of the feint fragment of approximately 500 bp. We expect a proportion of transformants to contain the insert of ~500 bp. **D** Colony PCR result from colonies harboring the mobility region GG module. The forward primer bound the plasmid backbone whilst the revese primer bound to the mobility region insert. One *E.coli* colony most likely contained the PCR by-product of ~500 bp that was created upon amplification of the mobility region from pRJpaph-bjGFP (see B).

2.1.3 Assembly of final *Bradyrhizobium* strain specific plasmids.

We aimed to combine the synthesized Golden Gate (GG) modules for the construction of USDA110, MOS001 and MOS004 specific plasmids, schematically depicted in figure 5 A. Furthermore, to determine whether we successfully assembled the final constructs we performed a PCR experiment to validate the presence of bjGFP. Amplicons of expected size were established for all final plasmids (Figure 5). Next, pairwise sequence alignments of all final plasmids show successful assembly of the final plasmids (Figure 5 C, D and E). The mutation visible between the annotated promoter and start of the bjGFP open reading frame is the aforementioned adenosine insertion that was possibly present in a fraction of the reverse primer used to amplify the promoter from the original plasmid (Figure 3 D). Moreover, we argued that mutations in the homologous region of USDA110 are either sequencing errors or physical polymorphisms (Figure 5 C). Either way, we expected no detrimental effects from a few mutations in the homologous region. Sufficient homologous nucleotides are still present to ensure successful integration of the construct in the chromosome of USDA110.
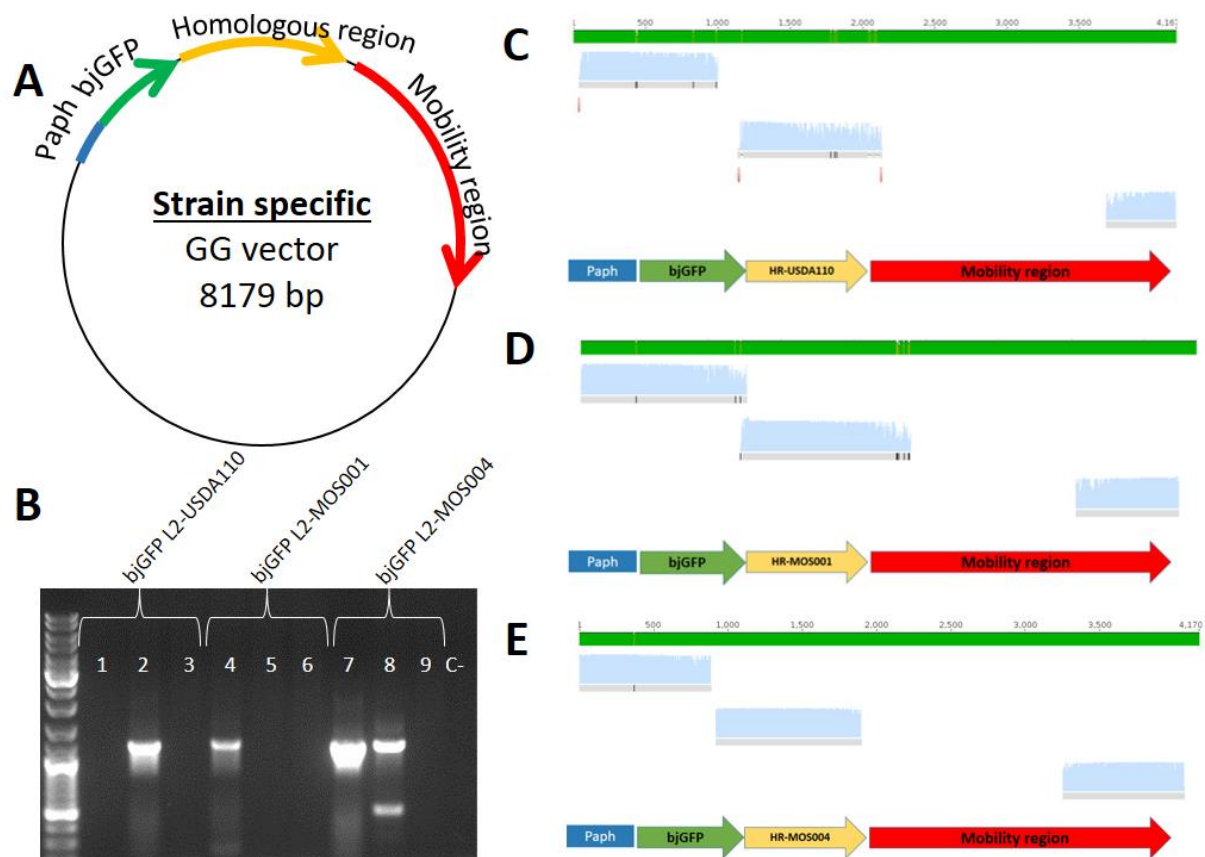
Figure 5 **A** Graphical representation of the final, strain specific golden gate module. Annotated, from left to right, is the codon optimized bjGFP, strain specific region of homology and the mobility region. **B** Colony PCR results of the three final strain specific GG modules. Plasmids from colonies showing a positive result (2 ,4, 7 and 8) for the bjGFP insert were send out for sequencing. Fragments of ~1.2 Kb result from the forward primer annealing to the plasmid backbone and the reverse primer binding the C-terminus of bjGFP. As negative control the empty acceptor vector was used. Generuler DNA ladder mix was used as size marker. **C** Nucleotide alignment of the USD110 specific GG plasmid with the *in silico* version of the construct. No detrimental effects are expected from the mutations in the homologous region. **D** Pairwise nucleotide alignment of the MOS001 specific GG plasmid with the *in silico* version of the plasmid. Mismatches in the alignment of HR and MOB are likely due to a decrease in per base quality at the end of the sequence reads. **E** nucleotide alignment of the MOS004 specific GG plasmid with the *in silico* version of the construct.

2.2 Transformation of USDA110, MOS001 and MOS004 with Golden Gate assembled plasmids.

Our first objective was to successfully transform USDA110 with the original plasmid to test the experimental setup in our laboratory. We expected the same results for USDA110 transformants carrying either the original plasmid or the final strain specific GG plasmid, considering both constructs consist of the same functional elements. Next, we intended to transform MOS001 and MOS004 with the GG plasmids. Further, we questioned whether the plasmid would become integrated chromosomally and in which way it is oriented.

2.2.1 Exponential growth of USDA110, MOS001 and MOS004 starts four days post inoculation.

We aimed to transform of USDA110, MOS001 and MOS004 with the final strain specific plasmids. For efficient transformation we used bacterial conjugation. Plasmid transfer with this procedure is most efficient when both donor and recipient strain are growing exponentially. To determine the duration of the lag phase, we constructed a growth curve using wild type USDA110, MOS001 and MOS004 (Figure 6). It appears that exponential growth starts four days (96 hours) post inoculation. This time point in growth can be used to organize the moment of bi-parental mating to increase transformation efficiency. In addition, we are unable to explain the decrease in absorption of USDA110 after reaching a value of one, considering higher absorptions have been measured in other experiments. We propose repetition of the experiment with both biological and technical triplicates to reduce variation caused by technical errors.

Next, we found the donor strain for bi-parental mating to be able to grow in growth conditions used for *Bradyrhizobium*. Further, we confirmed the donor strain to be unable to grow on counter selective medium. These results show that the donor strain is suitable for bi-parental mating with *Bradyrhizobium*.
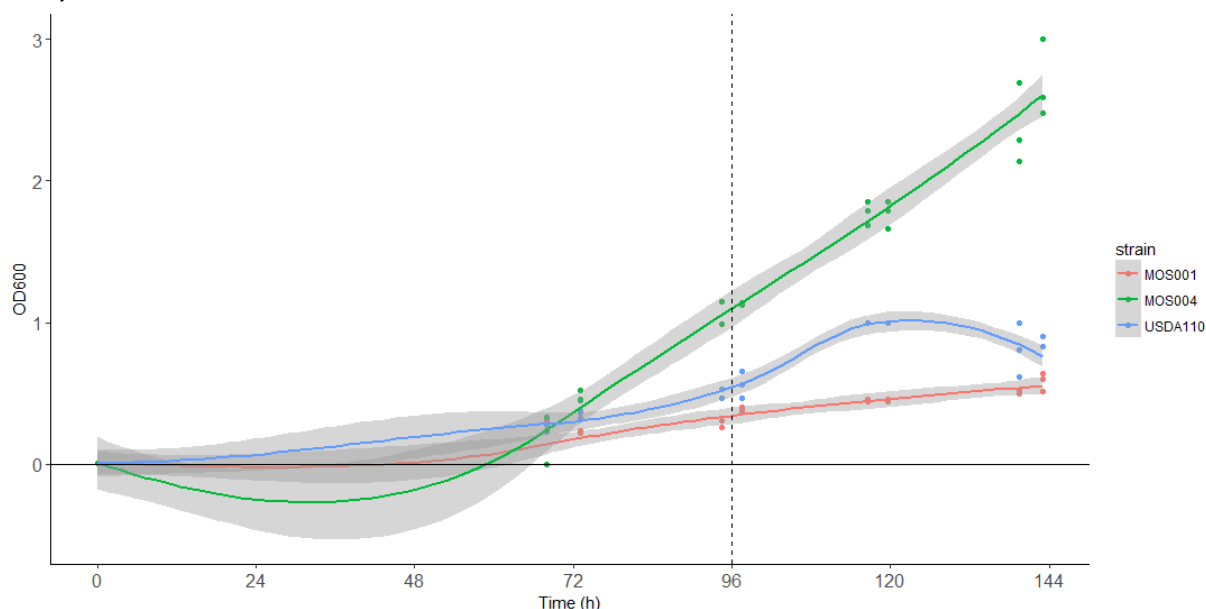
Figure 6 Growth curve of USDA110, MOS001 and MOS004. The lag phase of all three strains appears to end four days past inoculation (96 hours). This time point can be used to organize the moment of bi-parental mating to increase transformation efficiencies of USDA110, MOS001 and MOS004.

2.2.2 Implementing and optimizing the transformation protocol of USDA110 with pRJPaph-bjGFP and the final Golden Gate construct.

We aimed to transform USDA110, MOS001 and MOS004 with the strain specific plasmids. Firstly, because transformation of *Bradyrhizobium* spp. had never been done in our laboratory, we aimed at successful transformation of USDA110 with pRJPaph-bjGFP. Secondly, we wanted to use the implemented and optimized protocol for transformation of USDA110 with the GG plasmid. We expected the same result for both constructs, considering both plasmids contain the same functional elements.

After transformation of USDA110 with the original plasmid and the GG construct, we produced images of the transformant cells using a confocal microscope (Figure *8*). As expected, wild type

USDA110 does not show fluorescence emission. USDA110 transformants carrying the original plasmid were used as positive control for the presence of GFP. Images showing fluorescence emission from USDA110 transformants, carrying either the original plasmid or the GG construct, imply that we successfully implemented and optimized the transformation protocol.

To ascertain whether we successfully integrated the plasmid in the chromosome of USDA110 we conducted a PCR experiment (figure 7 A). The presence of a fragment that resulted from primers binding to the GFP marker gene and the genome of USDA110 gives the first indication of successful chromosomal integration of the construct. However, because of the presence of fragments that were expected when the construct was not integrated, we argued that not all transformants have the plasmid integrated chromosomally. We encourage making serial inoculums of transformants in non-selective medium over an extended period of time. In this way we expect transformants with cytoplasmic plasmids to dispose plasmids due to absence of selection pressure (8). Additionally, we conducted a BLASTn with all primers to determine the odds of off-target primer annealing to chromosomal DNA (data not shown). Beside the fact that one off-target effect is probable for the bjGFP forward primer (Figure 7 C), is it quite unlikely that used primer combinations used would result in amplicons of expected size. Finally, we amplified bjGFP from isolated genomic DNA (data not shown). However, this result only shows the presence of bjGFP, but gives no insight about the success of plasmid integration.

2.2.3 Successful transformation of MOS001 and MOS004 with the final strain specific GG modules.

After succeeding in optimizing the transformation protocol for USDA110, we aimed to transform MOS001 and MOS004 with the final GG plasmids. Confocal microscopy of MOS001 and MOS004 transformant cell cultures showed that both emit fluorescence emission, suggesting successful transformation of both MOS strains with the GG plasmid (Figure 8). As expected, no fluorescence emission could be detected for wild type cells of MOS001 and MOS004.

Nonetheless, the presence of fluorescence emission from MOS001 and MOS004 transformants provides no elucidation of successful chromosomal integration of the construct. To ascertain whether the plasmids integrated chromosomally, and in what orientation, we conducted a PCR experiment. Only successful PCR results were obtained for MOS004 transformants (Figure 7 B). Amplicons resulting from the marker gene and genome specific primers indicate successful plasmid integration in the chromosome of MOS004. Absence of amplicons, expected from wild type MOS004 genomic DNA, make it impossible to determine whether we have successful plasmid integration in all transformants or primer malfunctioning. We propose repetition of the experiment with new primers specific for the chromosomes of MOS001 and MOS004 (Figure 7 C). We further advise designing the new primers in such a way that at least two fragments are expected per plasmid integration possibility. Having more fragments per orientation will provide more conclusive results about the integration efficiency. In addition, we conducted a nucleotide BLAST with the used PCR primers against the genome of MOS001 and MOS004 (data not shown). This resulted in the same findings as the nucleotide BLAST against the genome of USDA110.
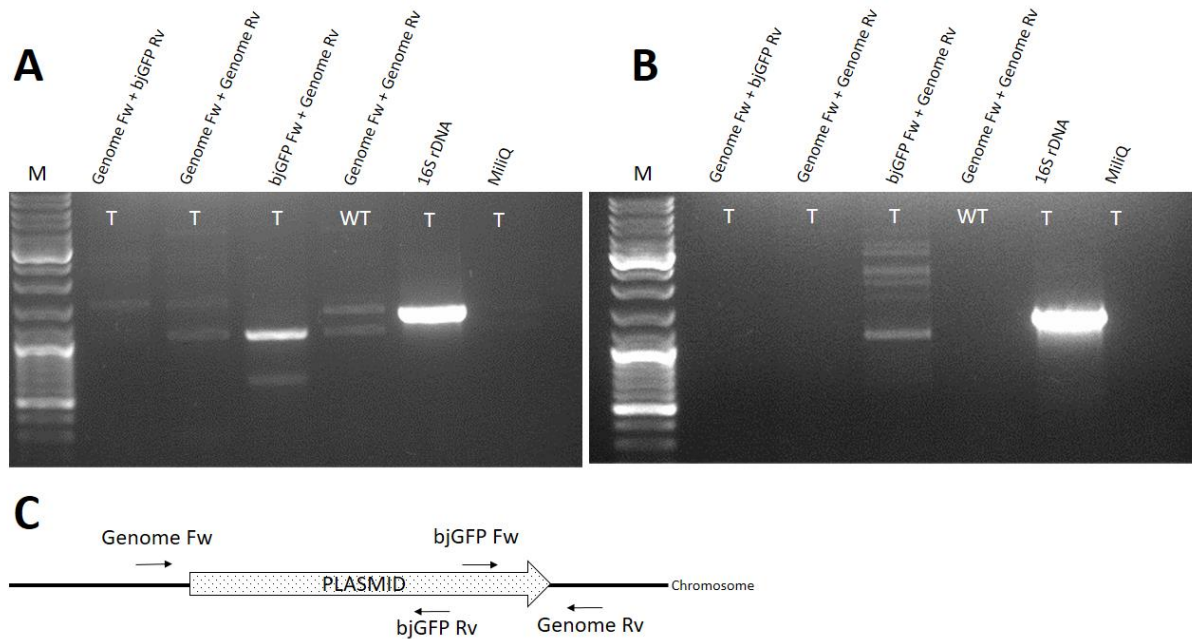
Figure 7 **A** PCR results of different primer combinations on isolated genomic DNA from USDA110 transformants with pRJpaph-bjGFP. T and WT stand for transformant and wild type isolated genomic DNA respectively. The bright amplicon in lane 3 is of expected size and suggest successful l integration of the plasmid in the chromosome. Fragments in lane 2 and 4 suggest that a proportion of the transformants have not yet successfully integrated the plasmid chromosomally. **B** PCR result of multiple different primer combinations on isolated DNA from wild type (WT) and transformant (T) MOS004. The band of ~1 Kb in lane 3 implies successful plasmid integration and its orientation in the genome. Note the absence of fragments in the PCR using WT genomic DNA and the genome forward and reverse primers. Absence of this fragment makes us unable to determine whether absence of amplicons from the transformants with the same primer combination (lane 2) is due to successful plasmid integration or primer malfunctioning. **C** Graphical representation of expected plasmid integration and used primer combinations in this PCR reaction.
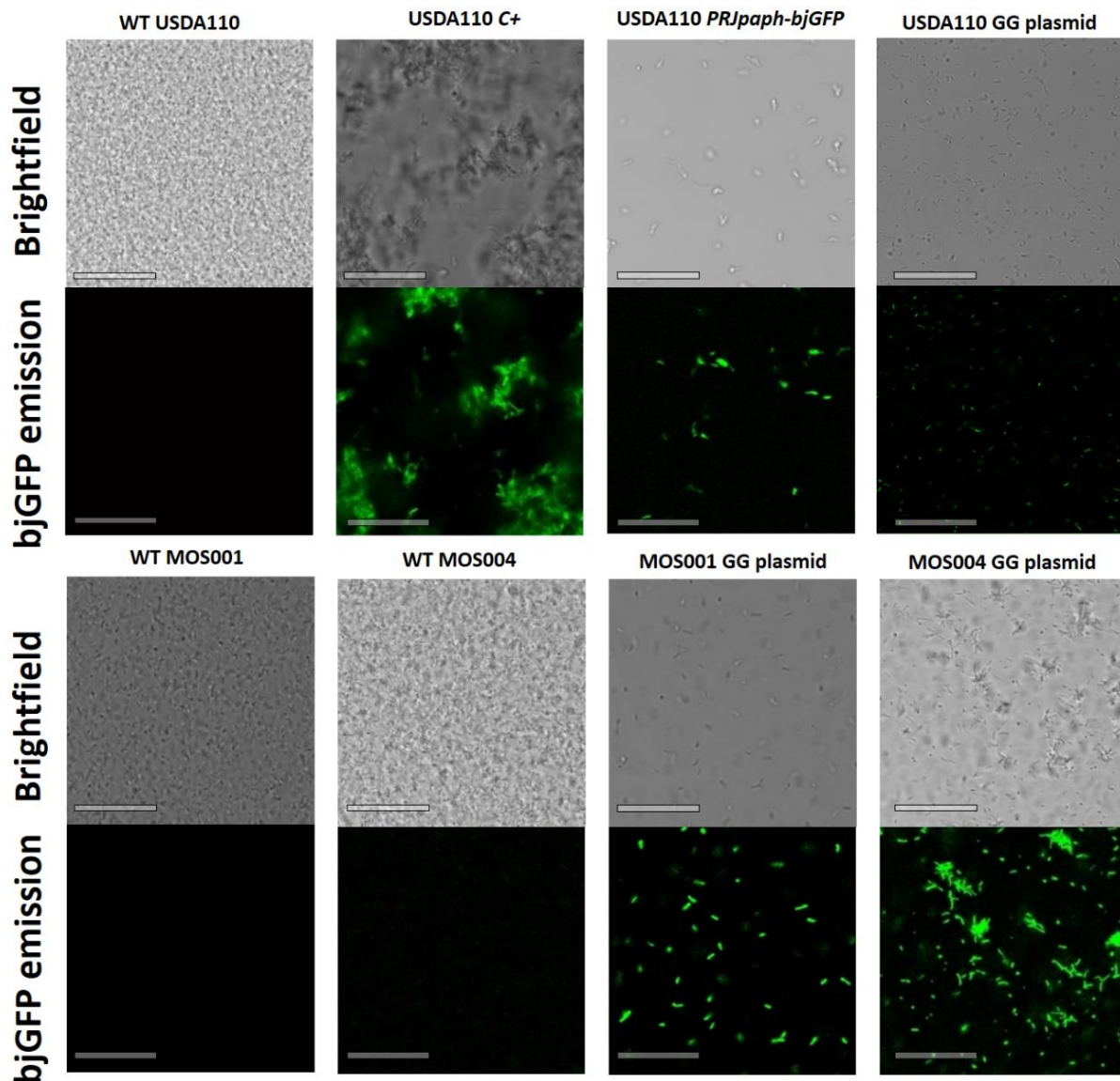
Figure 8 Top images show brightfield and bjGFP emission from wild type USDA110, USDA110 C+, USDA110 PRJpaph-bjGFP and USDA110 transformed with the final GG module. No fluorescence emission is detectable from wild type USDA110. Fluorescent emission is detectable for USDA110 transformants carrying either the original plasmid (pRJPaph-bjGFP) or the final USDA110 specific plasmid. The bottom images show brightfield and bjGFP emission of wild type MOS001, MOS004 and transformant strains thereof. Wild type MOS001 and MOS004 emit no fluorescence emission. Both MOS001 and MOS004 transformants show presence of bjGFP. However, the presence of fluorescence gives no certainty whether the constructs are integrated chromosomally. Scale bar size corresponds to 25 μm.

2.3 Both USDA110 and MOS004 appear incapable of colonizing the endophytic compartment of *Arabidopsis.*

After successful implementing the transformation protocols for USDA110, MOS001 and MOS004, we aimed at localizing these GFP-tagged strains on *Arabidopsis* roots. We spot inoculated one week old *Arabidopsis* root tips. After seven days we examined the area that was previously spot inoculated using confocal microscopy. We used propidium iodide (PI) to visualize the root apoplast. Images were taken using phase contrast, propidium iodide and fluorescence emission channels.

Confocal images show both USDA110 and MOS004 to colonize the rhizoplane of *Arabidopsis* roots (Figure 9, 10). Unfortunately, we were unable to trace back MOS001 transformants on root

sections. From these results, USDA110 and MOS004 appear incapable of entering the endophytic compartment of *Arabidopsis* under tested experimental conditions. These results do not confirm previous findings (7). However, it remains possible that tested conditions evoked an immune response in the roots which prevented endophytic colonization of USDA110 and MOS004 (10).

Moreover, it has been shown that nitrogen fixation and nodulation gene clusters are transferred horizontally between bacteria sharing the same ecological niche. Therefore, it remains paradoxical that symbiotic lifestyles are restricted to a relatively small amount of bacterial genera, given the taxonomic diversity of rhizospheric populations and co-occurrence of distinct non-rhizobial bacteria in root nodules (3) (11). Consequently, we argue that symbiotic *Bradyrhizobium* strains established a symbiotic lifestyle through exclusive acquisition of the nitrogen fixation and nodulation gene clusters.

Former studies showed that microbial communities in the root endophytic compartment of *Arabidopsis* were strongly influenced by soil types. Likewise, host developmental stages appeared to be of minor importance on the community composition of the endophytic compartment (12). In addition, recent inoculation experiments on *Arabidopsis* showed inability of MOS001 and MOS004 to establish growth promoting effects under tested experimental conditions (7). Nevertheless, it remains possible that in the setting of a more complex microbiome both plant and microbes have greater benefit from closer associations.
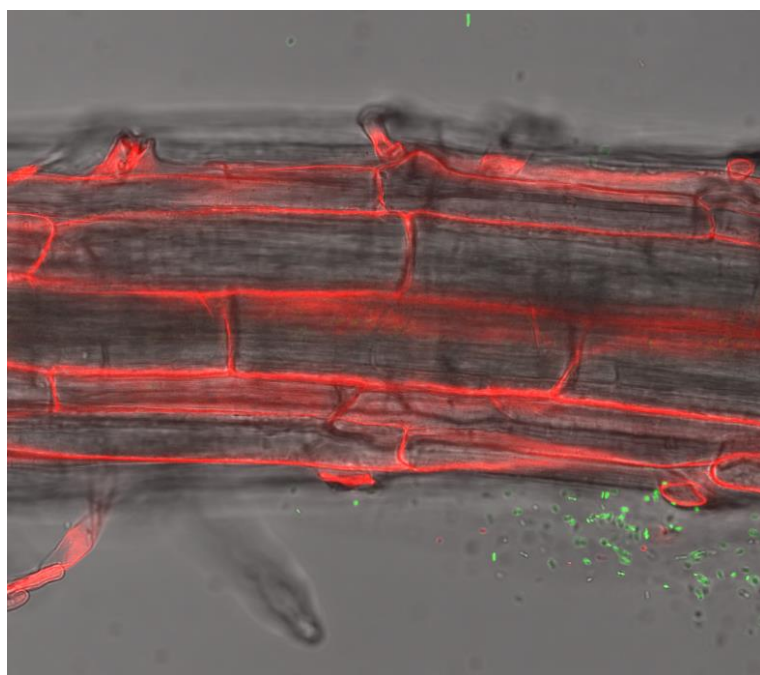
Figure 9 Confocal overlay image of the root section from *Arabidopsis* previously inoculated with USDA110 transformants harbouring pRJPaph-bjGFP. Propidium iodide was used to stain the root apoplast. USDA110 appears to colonize at the rhizoplane of the root, but unable to enter the endophytic compartment under tested experimental conditions.
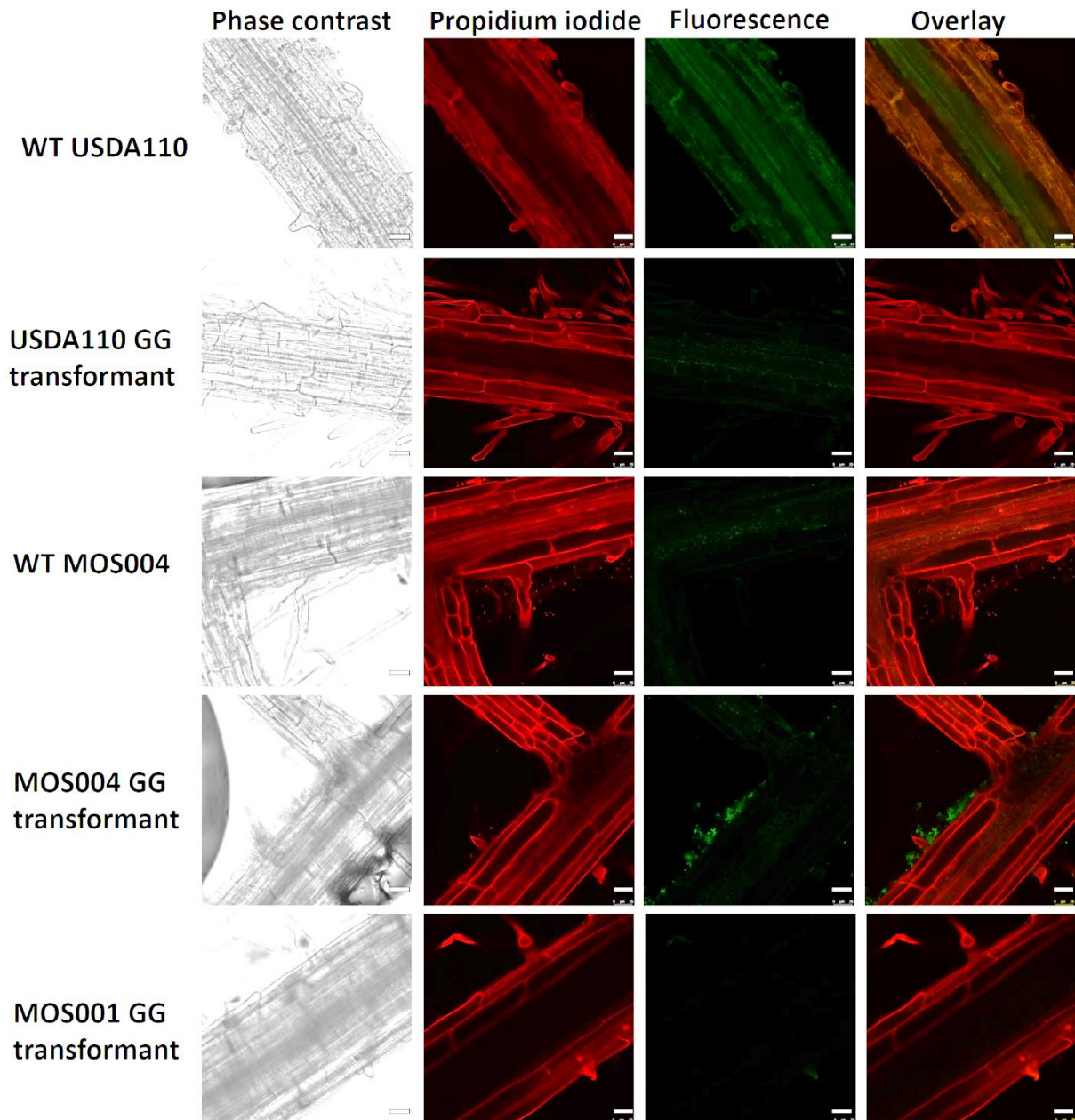
Figure 10 Confocal images of Arabidopsis roots previously spot inoculated with WT USDA110, USDA110 GG transformant, WT MOS004, MOS004 GG transformants and MOS001 GG transformants. Note that no images were made of roots inoculated with wild type MOS001. Propidium iodide was used to stain the root apoplast. Of each root, photographs were taken using a phase contrast, propidium iodide channel and a fluorescent emission channel. *Bradyrhizobium* MOS004 appears to colonize at the rhizoplane of Arabidopsis roots, but seems incapable of entering the endophytic compartment under tested experimental conditions.

2.3.1 Efficiency of bjGFP qPCR primers is sufficient for quantification of root colonization.

To quantify root colonization of USDA110, MOS001 and MOS004 transformants, we determined the relative abundance of the MOS isolates in the root endophytic compartment of *Arabidopsis* (supplementary figures S7), with the use of qPCR. We planned to quantify root colonization with *Bradyrhizobium* specific primers, *Arabidopsis* specific primers, and primers specific for *Bradyrhizobium* strains harbouring the marker gene. These primers would show whether the relative abundance of transformants corresponds with the relative abundance of *Bradyrhizobium* in the endophytic compartment. However, because of limited time we chose to not quantify root colonization of our strains on *Arabidopsis* (Figure 10 , 10).

Nonetheless, we constructed a calibration curve to determine the efficiency of the transformant specific qPCR primers. These primers bind to the bjGFP gene. We constructed a plot containing mean cq values against the [10] log of the DNA copy numbers (Figure *11* A, B). Primer efficiencies of 170.55% and 102.39% for USDA110 pRJPaph-bjGFP DNA and L0-bjGFP respectively were calculated. Primer efficiencies <90% and >110% are considered unacceptable for qPCR (13). We argue that the calculated primer efficiency of 170.55% is so high due to presence of PCR inhibitors. Additionally, we consider the calibration curves to be useful because R-squared values are of sufficiently large value. The complete datasets can be found in supplementary figures S5.
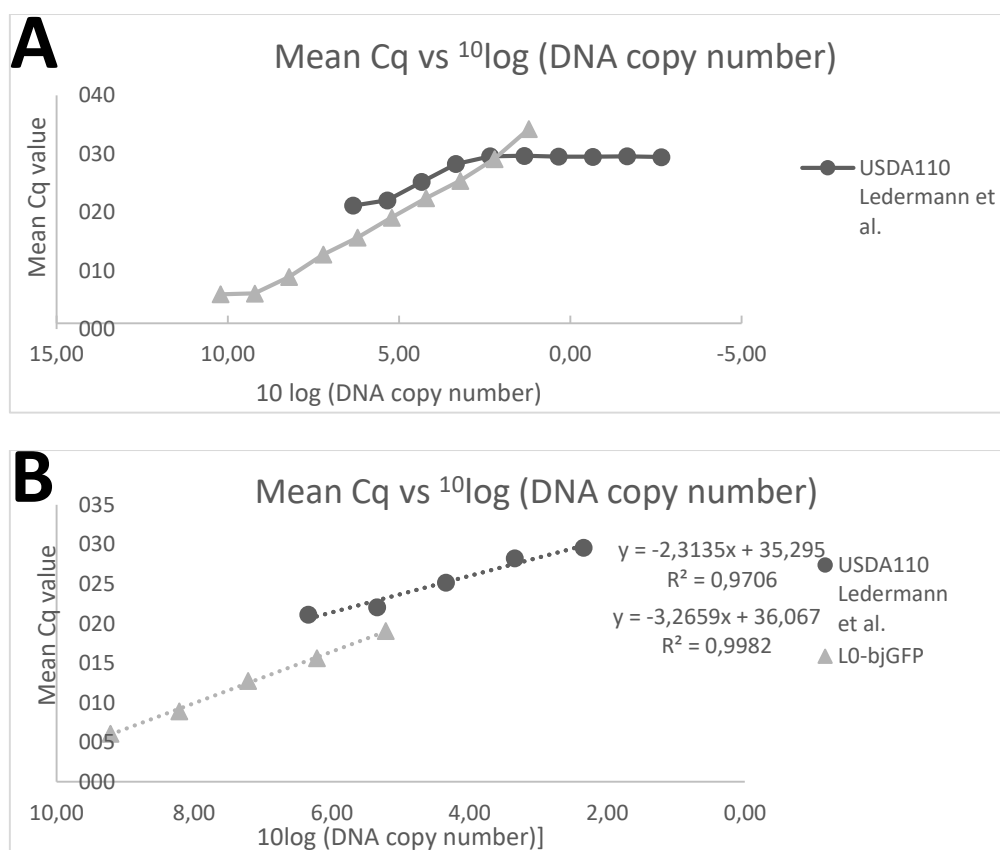
Figure 11 **A** plot of measured Cq values against the 10 log of DNA template copy numbers. **B.** Linear regression curve of the first five data points from figure 9A. The slope of this curve was used to calculate the primer efficiency of bjGFP forward and reverse primers. Total dataset can be found in supplementary figure S5.

3 Conclusion

This study shows the successful construction of Golden Gate compatible plasmids that become stably integrated at a uniform locus in the chromosome of USDA110, MOS001 and MOS004. Because Golden Gate cloning is module-based, are these plasmids usable for further research endeavors. No evident phenotypic side effects were detected after transformation with these constructs, apart from the GFP tag and the antibiotic resistance affiliated with it. Transformation of the genetically and functionally highly diverse *Bradyrhizobium* sp. MOS isolates and type strain USDA110, demonstrate that the optimized transformation protocols are useful for phylogenetically diverse *Bradyrhizobium* spp.

Moreover, the visualization of these genetically tagged *Bradyrhizobium* strains provides the first evidence that both the symbiotic USDA110 and the non-symbiotic MOS004 colonize at the rhizoplane of *Arabidopsis* roots, but are unable the enter the root endophytic compartment. Taken together, tools developed in this study have the potential to greatly expand knowledge about the associative behaviour of microbes and plants, allowing a deeper understanding of what it takes to develop a nitrogen-fixing symbiotic lifestyle.

4   Methods

Purification pRJPaph-GFP form E.coli carriers.
The template plasmid (pRJPaph-bjGFP) used for amplification of Paph, GFP-N, GFP-C and HR-USDA110 was obtained from *E.coli* stock solutions carrying the plasmid. 2 µL of the -80 °C stock was spread on an 1,5% LB-agar plate containing (10 µg/mL) tetracycline and incubated 16 hours at 37 °C. The next day a single colony was picked from the overnight plate using a sterile 200 µL pipetting tip. This tip was immediately put in 10 mL liquid LB medium in a sterile 50 mL tube containing tetracycline (10 µg/mL) and grown overnight at 37 °C in a 200 rpm shaker. The next day 2 mL of the overnight culture was miniprepped using the E.Z.N.A. Plasmid Mini Kit (Omega biotek) using the manufacturer's protocol.

Plasmid isolation from plasmid carriers
For plasmid purification 2 mL of an overnight culture was used. Purification of plasmids was done with the E.Z.N.A plasmid mini Kit (omega Biotek). Use was made of the manufacturer's protocol. After plasmid purification was the DNA concentration measured using SimpliNano™. Elution buffer provided in the plasmid mini kit was used as blanco.

Purification of PCR products
Purification of PCR amplicons directly from PCR mixture was done with the Themo scientific GeneJet Gel Extraction kit. Purification amplicons either from gel or directly from the PCR mixture was done according to the manufacturer's protocol.

Sequencing of L0-univeral acceptor (pAGM9121) and Pro+5U L0 acceptor (pAGM41295)
250 ng of each plasmid together with 20 pmol pBR322ori-GFPtest-F plasmid backbone primers was send to Macrogen Inc for high quality Sanger sequencing. Obtained sequences were aligned with the *in silico* vectors (Geneious 8.1.8) as to validate whether the level 0 acceptors were correct.

Amplification PCR of Paph, GFP-N, GFP-C, GFP-CT, HR-USDA110 and MOB
The amplification of the Paph (348 bp) , GFP-N (84 bp) , GFP-C (643 bp) , HR-USDA110 (837 bp), and MOB (2106 bp) region was performed with Thermo Scientific Phusion High Fidelity DNA Polymerase (2U/µL). The protocol used was provided by Thermofisher scientific. Optimal annealing temperatures for PCR primers for were calculated using the Tm calculater of Thermofisher using only the binding regions of the primers. For amplification of Paph, GFP-N, GFP-C, HR-USDA110 and MOB pRJPaph-bjGFP was used as template. The names of the primers used and specific primer sequence can be seen in Table 1. PCR reaction settings used can be found in Table 4

Golden Gate primer design
For the construction of L0-bjGFP, L0-Paph and L1-MOB from level -1 fragments amplified from pRJPaph-bjGFP, special golden gate compatible primers have been designed by toolbox65. All primers used for the amplification of bjGFP-N, bjGFP-C, HR (strain specific) and MOB from pRJPaph-bjGFP can be seen in table 1. Type IIS endonuclease recognition sites are added at the 5' end of the annealing site of the primers. A synonymous SNP (C->G) is induced in the overhang created after BpiI digestion of TB65_bj_L0_GFP-N_ C_R2 and TB65_bj_L0_GFP-C_F as to remove an internal BsaI recognition sequence. We have designed TB65_bj_L0_GFP-N_F and TB65_bj_L0_GFP-C_R2 in such a way that the first three nucleotides of the overhang created after BpiI digestion are part of the BsaI recognition site of pAGM9121. This enables seamless ligation of the bjGFP coding sequence to Paph in L1-Paph-bjGFP. An overview of all primers used can be found in Table 1. An overview of pRJPaph-bjGFP and a graphical

representation of the restriction site removal can be seen in supplementary figures S1 and S2 respectively.

Golden gate cloning of level 0 modules (L0-GFP & L0-Paph)

GFP-N, GFP-C and Paph purified PCR products each were added in a 3:1 molecular ratio with the appropriate pAGM9121 (universal level 0) and pAGM41295 (L0-acceptor-Pro5U) respectively. 10 units of BpiI were used for digestion. 10U of T4 DNA ligase were used for ligation. A detailed scheme of reaction mixtures can be found in Table 6**.** An overview of the reaction cycles used for the construction of the level 0 modules can be found in Table 7**.** Subsequent digestion/ligation reactions were carried out by means of T100TM Thermal Cycler for PCR.

General transformation of competent E.coli (DH5α)

E.coli cells present in 1.5 mL Eppendorf's 40 µL aliquots were taken from the -80 °C freezer and directly put on ice as to keep the temperature as close to 0 °C as possible. On these cells 2 µL of exoconjugants were added. After 15 min defrosting at 0 °C, the cell and plasmid mixture was added to pre-cooled electrophoreses cuvettes and placed in the electrophoreses apparatus. After electrophoresing the cells with 2500 V, 1 mL of SOC medium was added to the cells. After a minimal of incubation time of 60 min in the 37 °C 200 rpm shaker, 50 µL of the cells suspension was added to a 1.5% LB plate containing the proper screening antibiotic. Thereafter were the cells spun for 5 min, 4000 rpm as to collect all cells in a pellet. This pallet is thereafter spread over a 1.5% agar plate containing the proper antibiotic for selection of recombinants. These inoculated plates are thereafter placed in a 37 °C oven for minimal 16 hours.

Colony PCR to check L0-GFP and L0-Paph

To verify whether the white colonies on the spectinomycin plates truly were recombinants, a colony PCR has been performed. Colonies on which a PCR was performed were dissolved in 10 µL sterile MiliQ as to be able to inoculate fresh liquid medium with the same colonies on which the PCR was performed. A colony PCR mixture was used using 5 µL 10X Taq buffer, 20 µL 0.2 mM dNTPs, 1 µL of both the forward and reverse primer (10 µM stock), 3 µL of the dissolved colony, 0.25 µL DreamTaq polymerase and 19.75 µL MiliQ as to make a total volume of 50 µL. As forward primer pBR322ori-GFPtest-F was used and anneals to the vector backbone. The primer sequence (5'→3') of pBR322ori-GFPtest-F is GGGAAACGCCTGGTATCTTT. The reverse primers used are TB65_bj_L0_GFP-C_R2 and TB65_bj_L0_pAph-5U_R for L0-GFP and L0-Paph respectively (table 1). As negative control pRJPaphbjGFP was used in combination with both primer pairs. No fragment was expected because pBR322ori-GFPtest-F ought not to bind to the vector backbone of pRJPaphbjGFP. The complete PCR program used for amplification of fragments can be seen in table 9 . 3 µL of each dissolved colony was then directly added to 10 mL liquid LB medium in a 50 mL redcap tube containing spectinomycin (100 µg/mL). The liquid cultures were placed in a 37.0 °C 200 rpm oven for at least 16 hours. From these overnight cultures only the cultures with positive PCR results were miniprepped using E.Z.N.A. Plasmid Mini Kit (Omega biotek) using the manufacturer's protocol. The concentration of these plasmids was measured using SimpliNanoTM.. Elution buffer provided by the plasmid purification kit was used as a blanco in this measurement.

LB medium

Lysogeny broth (LB) medium was made by addition of Bactotryptone (10 g/L), NaCl (5 g/L), Yeast extract (5 g/L) to MiliQ water. In case of solid LB medium was agar added in a concentration of 15 g/L (1,5 %).

Plasmid sequencing L0-GFP and L0-Paph

Positive colony PCR Level 0 GG plasmids containing GFP and the Paph promoter were purified from DH5α transformants giving positive results in the colony PCR experiment. These plasmids were purified using E.Z.N.A. Plasmid Mini Kit (Omega biotek) according to the manufacturer's protocol. Approximately 250 ng of L0-GFP and L0-Paph were send with together with 20 pmol of TB65-bj-L0-GFP-N_F and TB65_bj_L0_pAph-5_F respectively. Sequencing results were aligned with the respective in silico level 0 plasmid using Genious version 8.1.8

Golden gate cloning of level 1 modules (L1-PaphbjGFP & L1-MOB)

To produce golden gate level 1 vectors. Level 0 GFP, Level 0 Paph and pICH47732 (L1-1 acceptor) were added to the golden gate reaction mixture in a 3:1 ratio (insert:acceptor). A total of 0.15 pmol insert vector (L0-GFP and L0-Paph) and 0.05 pmol pICH47732 were used in this reaction. A total of 10 U of BsaI and T4 DNA ligase were used for digestion and ligation respectively. The total reaction mixture can be found in Table 111**.** For the construction of L1-MOB (mobility region) use is made of pICH47751 as level 1 position 3 acceptor vector. Again, use was made of 3:1 insert: acceptor ratio. A total of 10 units of BsaI and T4 DNA ligase were used for digestion and ligation respectively. A detailed scheme of the reaction cycles used can be seen in Table 7**.** Subsequent digestion and ligation reactions were carried out by means of T100^{TM} Thermal Cycler

Making competent E.coli cells (DH5α and S17-1 λ- pir)

Of a 10 mL overnight culture in NaCl deficient liquid LB medium, 5 mL is added to 500 mL of the same medium in a 2 L flask. These cells are grown in a 37 °C, 200 rpm shaker to an OD_{600} of approximately 0.6. After reaching the correct OD_{600}, flasks were put on ice to bring the temperature to 0 °C. These cells were than centrifuged in a 4 °C precooled rotor centrifuged and resuspended in ice cold sterile MiliQ water. This washing step is repeated multiple times as to remove any LB medium present. Thereafter have the cells been resupended in ice cold 20% glycerol. After again centrifugation and resuspension of cells in 10% glycerol, the suspension was divided over 40 µL aliquots in sterile 1,5 mL Eppendorf tubes. These tubes were then directly cooled in liquid nitrogen (N_2). Thereafter are the competent cells stored in the -80 °C freezer.

Transformation of DH5α with L1-PaphbjGFP and L1-MOB

Competent DH5α cells, present in 1.5 mL Eppendorf's 40 µL aliquots, were taken from the -80 °C freezer and directly put on ice as to keep the temperature as close to 0 °C as possible. On these cells 2 µL of exoconjugants were added. After defrosting during a minimum of 15 min at 0 °C, the cell and plasmid mixture was added to pre-cooled electrophoreses cuvettes and placed in the electrophoreses apparatus. After electrophoresing the cells with 2500 V, 1 mL of SOC medium was added to the cells. After a minimal of incubation time of 60 min in the 37 °C 200 rpm shaker, 50 µL of the cells suspension was added to an 1.5% LB plate containing ampicillin (100 µg/mL). Thereafter were the cells spun for 5 min at 4000 rpm as to collect all cells in a pellet. This concentrated pallet was then again plated on a 1.5% LB plate containing ampicillin (100 µg/mL). Transformants were grown for a minimum of 16 hour in a 37 °C oven.

Gel electrophoresis

To verify the presence of DNA (i.e. PCR products, (digested) plasmids, RNA, cDNA etc.) use is made of gel electrophoresis. This methods relies on the sieve like properties of agarose gel. The higher the % (i.e. w/v (g/mL) ratio) of agarose, the more suitable the gel for separating smaller molecules due to decreasing pore seize. After adding a suitable amount of agarose to 1x Tris-Acetate-EDTA (TAE) buffer, the agarose is dissolved by heating in the microwave. After the agarose has dissolved ethidium bromide is added in a concentration of 0.1 µg/mL (i.e 5 µL of a 1 mg/mL stock in 50 mL liquid agarose solution). Ethidium bromide has intercatalitic activity with DNA/RNA which enables it to become fluorescent only when interspaced between double stranded DNA/RNA. After successfully dissolving agarose in 1x TAE and addition of ethidium bromide is the solution added to a mall. After polymerization of the agarose (± 30 min) is the gel added to a tray containing an anode and cathode. Due to the negative charge of DNA shall DNA migrate towards the cathode. After enough time is the gel transferred to the Biorad imager as to take a digital picture of the migration pattern of DNA/RNA in the agarose gel.

Colony PCR to check L1-PaphbjGFP, L1-MOB and inoculation

In order to verify whether the level 1 GG reaction was successful and whether the white colonies really carried the correct plasmid, a colony PCR was performed. Of each transformant, being either L1-PaphbjGFP and L1-MOB, 5 colonies were picked from each level 1 transformant with a sterile 200 µL pipetting tip. Each of these colonies was then dissolved individually in a sterile 1.5 mL eppendorf tube containing 10 µL sterile MiliQ water. 3 µL of each dissolved colony was added as template to the PCR reaction mixture containing 5 µL 10x Taq buffer, 20 µL 0.5 mM dNTP's, 1 µL of both forward and reverse primer (10 µM stock), 0.25 DreamTaq DNA polymerase (5 U/ µL) and 19.75 µL sterile MiliQ to make a total of 50 µL. 3 µL of each dissolved colony was then directly added to 10 mL liquid LB medium in a 50 mL redcap tube containing ampicillin (100 µg/mL). These inoculated tubes were than placed in a 200 rpm 37.0 °C shaker for growth overnight. This enables purification of L1 vectors from positively screened colonies. As forward primer for both L1-PaphbjGFP and L1-MOB use is made of MAP69_CRISPR-LB-FW which anneals to the level 1 acceptor vector backbone. As reverse primer for L1-PaphbjGFP and L1-MOB use is made of MAP47_CRISPR_RB_Rv and TB65_bj_L1_MOB_R respectively. Of these reverse primers MAP47_CRISPR_RB_Rv also binds to the vector backbone, thereby, in combination with MAP47_CRISPR_RB_Rv creating amplicons that completely span the Paph-bjGFP insert. As negative control for the L1-Paph-bjGFP PCR use is made of pICH47732 (L1-1 acceptor). Amplification of pICH47732 using MAP69_CRISPR-LB-FW and MAP47_CRISPR_RB_Rv should result in a fragment of 766 bp, containing LacZ. As negative control for L1-MOB, use is made of pICH47751. Because the TB65_bj_L1_MOB_R binds the MOB insert, no fragment is expected upon amplification of pICH47751 using MAP69_CRISPR-LB-FW and TB65_bj_L1_MOB_R. No positive control was taken into account because the expected fragment size was already known. The expected PCR size for L1-PaphbjGFP is 1240 bp and L1-MOB is 2186 bp. The expected fragment sizes were determined with an in silico PCR reaction in Geneious version 8.1.8 A clear overview of the PCR reaction mixture for one 50 µL PCR reaction and the PCR cycle settings used can be seen in Table 3 and Table 9 respectively.

Plasmid sequencing L1-PaphbjGFP and L1-Paph

PCR Level 1 GG plasmids containing GFP and the Paph promoter have been purified from DH5α transformants giving positive results in the colony PCR experiment. These plasmids have been purified using E.Z.N.A. Plasmid Mini Kit (Omega biotek) using the manufacturer's protocol. Approximately 250 ng of L1-PaphGFP and L1-MOB were send with together with 20 pmol of MAP69_CRISPR-LB-FW. Sequencing results were aligned with the respective *in silico* constructed L1-Paph-bjGFP and L1-MOB using Genious version 8.1.8.

Golden Gate cloning of final level 2 modules (L2-USDA110, L2-MOS001 and L2-MOS004)

To construct the strain specific level 2 modules L1-PaphbjGFP, L1-MOB, HR (strain specific) and L3E (end linker position 4) and PMC-01270 (level 2 acceptor vector) were added in a 3:1 (insert:vector) molecular ratio. Of L1-PaphbjGFP, L1-MOB, HR and L3E (end linker position 4) 0.15 pmol was added whilst of PMC-01270 0.05 pmol was added to the golden gate reaction mixture. A total of 10 U of BpiI and T4 DNA ligase were used for digestion and ligation respectively. The complete reaction mixtures can be seen in table 12 . A detailed scheme of the reaction cycles used can be found in table 7. Subsequent digestion and ligation reactions were carried out by means of T100™ Thermal Cycler

Colony PCR on Level 2 vectors and inoculation

Because PMC-01270 (level 2 acceptor vector) does not contain the LacZ gene between BpiI restriction sites, is blue white screening of level 2 transformants not possible. Therefore have three colonies of each level 2 transformants, being DH5α transformants with either L2-USDA110, L2-MOS001 or L2-MOS004, be selected randomly. Each colony was dissolved in 10 µL sterile MiliQ water. Because the complete level 2 insert is approximately 3000 bp, has the choice been made to amplify the level 2 vector in two overlapping parts, as to verify the presence of all functional elements. The first part/amplicon is made a primer binding the level 2 acceptor backbone and one binding the c-terminus of bjGFP. These primers are designated MAP378_L2_bb_Fw and TB65-bj-L0-GFP-C_R2. The fragment size expected upon amplification with this primer pair is 1166 bp. As negative control an empty level 2 acceptor vector (PMC-01270) was used. Because this empty vector does not contain bjGFP, no fragment was expected upon amplification with MAP378_L2_bb_Fw and TB65-bj-L0-GFP-C_R2. The second amplicon was created using MAP364_bjGFP_out1 as forward primer and MAP379_L2_bb_Rv as reverse primer. MAP364_bjGFP_out1 binds the c-terminal part of bjGFP. MAP379_L2_bb_Rv binds PMC-01270 backbone. The expected amplicon size is 3110 bp and spans a small part of bjGFP, the strain specific HR and the mobility region (MOB). As negative control use is made of an empty level 2 acceptor vector (PMC-01270). Because the empty acceptor vector does not contain bjGFP, no amplicon is expected upon amplification with MAP364_bjGFP_out1 and MAP379_L2_bb_Rv. All primer sequences can be seen in table 8. Each PCR reaction contained 5 µL 10x Taq buffer, 20 µL 0.5 mM dNTP's, 1 µL of both forward and reverse primer (10 µM stock), 0.25 µL DreamTaq DNA polymerase (5 U/ µL) and 19.75 µL sterile MiliQ to make a total of 50 µL. 3 µL of each colony was used as template for the PCR reaction. 3 µL of each dissolved colony was then directly added to 10 mL liquid LB medium in a 50 mL redcap tube containing ampicillin (100 µg/mL) and spectinomycin (100 µg/mL). These inoculated tubes were then place in a 200 rpm shaker at 37.0 °C for overnight growth. This enabled plasmid purification of positive colonies >16 hours after colony PCR screening. Because per colony two amplicons have to be synthesized is 3 µL of each dissolved colony added to either of the different PCR. The complete colony PCR reaction settings can be seen in table 9.

Sequencing purified Level 2 vectors

From the level 2 colony PCR experiment, all colonies have been miniprepped with the use of E.Z.N.A. Plasmid Mini Kit (Omega biotek) according to the manufacturer's protocol. Of all three miniprepped L2-USDA110, L2-MOS001 and L2-MOS004 colonies, thus nine colonies in total, only one plasmid solution of each vector type has been send for sequencing. Because the level 2 insert, encompassing Paph-bjGFP, HR and MOB, is too large to sequence in one reaction, we have chosen to sequence each plasmid with one different primer at a time. The first forward primer used is MAP378_L2_bb_Fw which binds the level 2 vector backbone. The second forward primer used for L2-USDA110 and L2-MOS001 is MAP364_bjGFP_out_1, sequencing L2-2 HR. The second primer used for L2-MOS004 is MAP365_bjGFP_out_2 because of unknown sequencing failures using MAP364_bjGFP_out_1. The third primer used for L2-USDA110, L2-MOS001 and L2-MOS004 is reverse primer MAP379_L2_bb_Rv which again binds the level 2 acceptor vector backbone. This primer enables inward sequencing of, at

least a part, of MOB. For sequencing 250 ng of each level 2 vector, together with 20 pmol of primer is send to Macrogen Inc for high quality Sanger sequencing. All three sequenced fractions of the level 2 vector were aligned with the in silico level 2 vector in Genious version8.1.8 Sequence validation of the strain specific HR region was used to validate whether the designated L2-MOS001 and L2-MOS004 vectors were indeed strain specific for MOS001 and MOS004.

Inoculation of E. coli S17-1 λ- pir

A minor volume of -80 °C 40% glycerol stock *E. coli S17-1 λ- pir* stock was taken using a sterile disposable inoculation loop. Inoculation was done on two 1.5% LB agar plates. One plate contained streptomycin (50 µg/mL) as to select for *E. coli S17-1 λ- pir.* A second plate containing 1.5% agar LB lacking antibiotics was used a control. Inoculation is done at 37 °C overnight.

E. coli S17-1 λ- pir on solid PSY medium / LB medium.

As to verify whether *E.coli* strain *S17-1 λ- pir* is able to grow successfully on PSY medium at 28.0 °C an experiment has been carried out. 4x LB 1.5% agar plates and 4x PSY 1.5% agar plates were inoculated with equal amounts of a suspension of *S17-1 λ- pir.* Of these eight agar plates, two of each medium type were put in the 37.0 °C oven, and two were put in the 28.0 °C oven. After at least 16 hours of growth were all eight plates analyzed for colonies.

Inoculation of Bradyrhizobium strains in liquid PSY medium

Of MOS001, MOS004 and USDA110 one single colony is added to 10 mL liquid PSY medium in a sterile 50 mL tube. Colonies on solid PSY plates (1.5% agar) were picked using a sterile 200 µL pipetting tip. Colonies were grown at 28 °C in a 200 rpm shaker for at least four days until the desired OD_{600} was reached.

Growth curve USDA110, MOS001 and MOS004

To determine the growth rate of USDA110, MOS001 and MOS004 a growth curve is constructed. Three 50 mL redcap tubes per *Bradyrhizobium* strain have been filled with 10 mL liquid PSY medium containing chloramphenicol (25 µg/mL). Of each strain, a high density suspension was made in a 1.5 mL Eppendorf tube containing 500 µL sterile MiliQ. This suspension was made by scraping cultures from PSY plates that stored in the 4.0 °C fridge. The OD_{600} of this suspension was measured with an Eppendorf BioSpectrometer® basic. As blanco for measuring the OD_{600} PSY containing chloramphenicol was used. 10x dilutions of each suspension were made in a 1 mL cuvette (100 µL culture + 900 µL PSY). From this suspension was calculated how much of each suspension had to be added to 10 mL liquid PSY cultures to make a start OD_{600} of 0.01. This time point is designated t=0. Twice a day, for a total of six days, the OD_{600} was measured. Averages of the absorption of the triplets was plotted against the accumulative time in hours. This data was then used to construct a growth curve using R version 1.1.383

Genomic DNA isolation USDA110, MOS001 and MOS004

Genomic DNA from preciously inoculated wild type and transformed USDA110, MOS001 and MOS004 was isolated using QIAGEN DNeasy Blood & Tissue kit. Isolation was done according to the manufacturer's protocol. Some deviations however were made to increase the yield of DNA extraction. According to the protocol no more than $2x10^9$ cells should be used. In this experiment, all cells present in the 10 mL inoculum were collected by rotating the 50 mL redcap tubes 10 min at 4500 rpm. This bacterial culture was then incubated 60 min at 60 °C together with protein kinase K. Genomic DNA concentrations were measured using SimpliNano™. Elution buffer provided in the kit was used as blanco prior to measuring.

Biparental mating of Bradyrhizobium

Transformation of *Bradyrhizobium* with exoconjugants is done with the *Escherichia coli* strain S17-1-λpir. First, one colony was picked from S17-1-λpir transformants harboring the strain specific plasmid and added to a 50 mL tube containing 10 mL liquid LB with the right antibiotic. This culture was then added to a 200 rpm 37 °C for overnight growth. The next day fresh LB medium containing the right antibiotic(s) was inoculated with 300 µL of overnight culture. This inoculum was grown till the OD_{600} was between 1 and 2. At least 5 days ahead an inoculum of *Bradyrhizobium* was made in liquid PSY containing chloramphenicol (25 µg/mL) as to ensure and OD_{600} 0.5 and 1.0 at the day of the biparental mating experiment. The cells were then washed two times by centrifugation 10 min at 4000 rpm and resuspension in 0.9% (154 mM) NaCl. After washing were the cells resuspended to an OD_{600} of 4.0. Of this suspension 750 µL recipients (USDA110/MOS001/MOS004) and 250 µL donor (S17-1-λpir) were mixed in a 1.5 mL Eppendorf tube, centrifuged (10 min 4500 rpm) and resuspended in 100 µL 0.9% NaCl. This mixture was then plated on a 1,5% PSY agar plate lacking antibiotics. After incubation for at least 72 hours at 28 °C was the bacterial cake scratched from the PSY plate and resuspended in 750 µL 0.9% NaCL. The cells were washed two times with 0.9% NaCl. To make a final suspension was 100 µL 0.9% NaCL added to the cell pellet after washing. This bacterial suspension was then spread in aliquots of (10, 20, 40, 80, 100, 200 and the rest) on PSY plates containing chloramphenicol (25 µg/mL) for counter selection of S17-1-λpir and the right antibiotics for selection of *Bradyrhizobium* transformants.

PCR to verify whether the strain specific Golden Gate plasmids were integrated chromosomally, and in what orientation (USDA110, MOS001 and MOS004)

To verify whether pRJpaph-bjGFP or L2-USDA110, L2-MOS001, L2-MOS004 were successfully integrated in the chromosome did we carry out a PCR experiment using primers from table 13. Both wild type and transformant strains of USDA110, MOS001 and MOS004 were grown for seven days. Thereafter genomic DNA was purified using QIAGEN DNeasy Blood & Tissue kit. Isolation was done according to the manufacturer's protocol. This genomic DNA was then used in the PCR experiment. PCR reaction settings can be found in table 14

Sterilization of *Arabidopsis* seeds

Arabidopsis seeds were sterilized by first washing the seeds with 70% (v/v) ethanol. Thereafter were the seeds incubated max. 10 min at room temperature in 4x diluted bleach. After incubation was the 4x diluted bleach solution removed and were the seeds washed at least 5 times with sterile MiliQ. Thereafter were the seeds put on a petri dish containing a sterile filter paper. These seeds were put in a 4 °C fridge for at least two days to give the seeds a cold shock. Thereafter were the seeds added to 1/2 Murashine and skoog 10 (MS) medium containing both 0.8% (g/mL) Diashin agar and sucrose (10 g/L).

Inoculation of wild type Mossel *Arabidopsis*

Sterilized *Arabidopsis* seeds were put in the 4 °C fridge for two days prior to plating on ½ Murashine and skoog 10 (MS) medium containing both 0.8% (g/mL) Diashin agar and sucrose (10 g/L). Plants were grown at least seven days till the roots were > 3 cm in length. For each strain we used three biological replicates (i.e. three plates each containing five seeds). Liquid inoculums of each strain were centrifuged for 10 min at 4000 rpm. After washing the cells at least two times with sterile MiliQ were the cells suspended to an OD_{600} of 0.1. This OD_{600} roughly corresponds to a cell/mL ratio of 1.0E5. Root tips were inoculated with 2 µL of each cell suspension using a sterile 200 µL pipetting tip. This corresponds to roughly 2.0E5 cells / root tip.

Localization study of *Bradyrhizobium* transformants

Propidium Iodide (PI) was used to stain the apoplast of *Arabidopsis* roots. Roots were cut with a sterile scalpel 3 mm above and below the site of inoculation because here we had the largest chance of finding back cells. The 6 mm root segments were added to 50 μL drops of PI on thin microscope slides. Phase contrast, propidium iodide and fluorescence emission images were made using a Leica confocal microscope.

Calibration curve qPCR bjGFP primers

To determine the efficiency of MAP365_bjGFP_out2_Fw and MAP390_GFPins_ins_Rv a calibration curve is made. Primers can be found in table 15 The qPCR reaction mixture contained 5 μL iQ^{TM}SYBR® Green Mix (Biorad), 1.5 μL forward and reverse primer (1 μM) and 2 μL DNA template to make a final volume of 10 μL. For making the calibration curve of the aforementioned bjGFP primers use is made of genomic DNA isolated from USDA110 pRJPaph-bjGFP transformants grown in liquid PSY for at least five days. As control is the Golden Gate level 0 bjGFP vector used. A 10 fold dilution series of genomic DNA and L0-bjGFP was made starting at 11 ng/ μL and 25.15 ng/μL respectively. The protocol for qPCR was as followed: 95 °C (5 min); 39x [95 °C (30 s), 60 °C (30 s)]. The reaction was stopped by heating 7 minutes at 95 °C. To determine the primer efficiency we have plotted the mean cq value (of the technical replicates) against the 10 log of the DNA template copy numbers.

Table 1 Primers used for amplification of bjGFP-N, bjGFP-C (terminal), Paph, HR-USDA110 and MOB from pRJPaphbjGFP. Also are the strain specific HR primers mentioned which were used for amplification of HR from genomic DNA of MOS001 and MOS004. The template binding sequence of the primers are designated in Italic. All sequences given in 5'→3' orientation.

Primer	Type IIS endonuclease recognition sequence (5' →3')	BpiI overhang (1st)	BsaI overhang (2nd)	Primer sequence (5' →3')
TB65_bj_L0_GFP-N_F	GAAGAC	CTCA	AATG	TGTGAAGACATCTCAAATG*TCGAAGGGCGAGGAGC*
TB65_bj_L0_GFP-N_R	GAAGAC	ACAC	—	TGTGAAGACATACA*CCGAGAACTTGTGGCCGTTGAC*
TB65_bj_L0_GFP-C_F	GAAGAC	GTGT	—	TGTGAAGACATGTGT*CGGGCGAGGGCG*
TB65_bj_L0_pAph-5U_F	GAAGAC	GGAG	—	TGTGAAGACATGGAG*CACGCTGCCGCAAGCAC*
TB65_bj_L0_pAph-5U_R	GAAGAC	CATT	—	TGTGAAGACATCATT*AATTATATCTCCTTCTTATTTCTAGTCAAGATC*
TB65_bj_L0_GFP-C_R2	GAAGAC	CTCG	AGCG	TGTGAAGACATCTCGAGCG*CCAATTGTCATCACTTGTACAGCTC*
TB65-HR-mos1-bpi-F1	GAAGAC	GCAA	—	TGTGAAGACATGCAACAGTCCTCTCCACATCGTCATG
TB65-HR-mos1-bpi-R1	GAAGAC	TAGT	—	TGTGAAGACAT*GCAACAGTCCTCTCCACATCGTCATG*
TB65-HR-mos4-bpi-F1	GAAGAC	GCAA	—	TGTGAAGACAT*GCAACGGATGGTCTATAAGCCGTTCAATTC*
TB65-HR-mos4-bpi-R1	GAAGAC	TAGT	—	TGTGAAGACATT*AGTCGCAGCGCATCGTTATCCG*

Table 2 Overview template plasmid and acceptor vectors with the antibiotic resistance gene used for Golden Gate cloning

Plasmid	Plasmid functionality	Antibiotic selection marker
pRJPaph-bjGFP	Paph, bjGFP, MOB, HR-USDA110template	Tetracycline
pAGM9121	universal L0 acceptor	Spectinomycin
pICH41295	Pro+5U L0 acceptor	Spectinomycin
pICH47732	L1-1 acceptor	Ampicillin
pICH47751	L1-3 acceptor	Ampicillin
pICH41766	L3E end linker	Spectinomycin
PMC-01270	Level 2 acceptor	Ampicillin + Spectinomycin

Table 3 General composition of one 50 µL PCR reaction

Component	Volume [µL]
10x DreamTaq buffer	5
dNTP's 0.5 M	20
Forward primer	1
Reverse primer	1
Template DNA	*Dependent on template concentration*
DreamTaq (5 U/ µL)	0.25
MiliQ water	*Dependent on template volume*

Table 4 PCR reaction settings used to amplify Paph, bjGFP-N, bjGFP-CT (terminal), MOB and HR-USDA110 from pRJPaph-bjGFP colony PCR.

Phase	Temperature [°C]	Time	Cycles
Initial denaturation	98	1 min	—
Denaturation	98	15 s	
Annealing	59 (HR-USDA110) 60 (MOB, bjGFP-N, bjGFP-CT)	15 s	35 cycles
Extension	72	15 s	
Final Extension	72	7 min	—
Hold	4	∞	—

Table 5 Overview of Level 0 Golden gate insert fragment and acceptor vector characteristics.

DNA fragment	Fragment type [insert or vector]	Length DNA fragment [bp]	Mw [kg/mol]	Concentration [ng/µL]	Acceptor vector
Paph	Insert	348	229.68	35	pICH41295
bjGFP-N	Insert	84	55.4	44.5	pAGM912
bjGFP-CT	Insert	643	424.38	31	
pICH41295	Universal L0 acceptor	2659	1754.94	117	—
pAGM9121	Pro+5U L0 acceptor	2845	1877.94	72.5	—

Table 6 Composition of Golden Gate mixtures belonging to L0-bjGFP L0-Paph. Insert and vector are added in a 3:1 molecular ratio.

L0-bjGFP	Volume [µL]
PCR purified GFN-N amplicons, 4x diluted stock (44.5 ng/µL)	0.7
Purified bjGFP-CT amplicons, 4x dilute stock (31 ng/µL)	8.2
pAGM9121 (Universal L0 acceptor) 10x diluted stock (117 ng/µL)	7.5
Restriction enzyme BpiI (10U/ µL)	1
T4 DNA ligase (10U from high concentration ligase 30 U/µL)	0.5
10X T4 DNA ligase buffer	2
BSA (0.1%)	2
MiliQ	0

L0-Paph	Volume [µL]
PCR purified Paph amplicons, 4x diluted stock (35 ng/µL)	7.5
pICH41295 (Pro+5U L0 acceptor) 10x diluted stock (72.5 ng/µL)	1.3
Restriction enzyme BpiI (10U/ µL)	1
T4 DNA ligase (10U from high concentration ligase 30 U/µL)	0.5
10X T4 DNA ligase buffer	2
BSA (0.1%)	2
MiliQ	9.3

Table 7 General golden gate assembly steps used for digestion – ligation reactions for L0-GFP, L0-Paph, L1-PaphGFP, L1-MOB, L2-USDA110/MOS001/MOS004

Golden gate step	°C	Time [min]
1	37	5
2	37	3
3	16	4
4	GOTO ->	step 2, 25x
5	50	10
6	80	10
7	4	∞

Table 8 All primers used for L0 bjGFP, L0 Paph, L1-Paph-bjGFP, L1-MOB and L2-USDA110/MOS001/MOS004 colony PCR experiments

Primers name	Primer sequence (5' →3')
pBR322ori-GFPtest-F	*GGGAAACGCCTGGTATCTTT*
TB65_bj_L0_GFP-C_R2	*TGTGAAGACATCTCGAGCGCCAATTGTCATCACTTGTACAGCTC*
TB65_bj_L0_pAph-5U_R	*TGTGAAGACATCATTAATTATATCTCCTTCTTATTTCTAGTCAAGATC*
MAP69_CRISPR-LB-FW	CATTGCGGACGTTTTTAATGTACTG
TB65_bj_L1_MOB_R	*TGTGGTCTCAAGCGCTGTGGCGCCGGTGATG*
MAP47_CRISPR_RB_Rv	CTAATAAACGCTCTTTTCTCTTAGGT
MAP378_L2_bb_Fw	CGCAATAGTTGGCGAAGTAATC
MAP364_bjGFP_out1	CTGCTGGAGTTCGTCACC

Table 9 PCR reaction settings for colony PCR on E.coli colonies harboring L0-GFP, L0-Paph, L1-MOB, L1-PaphbjGFP and L2-USDA110/MOS001/MOS004

Phase	Temperature [°C]	Time	Cycles
Initial Denaturation	98	8 min	—
Denaturation	95	30 s	
Annealing	58	30 s	
Extension	72	1 min (L0-GFP & L0-Paph) 1,5 min (L1-MOB & L1-PaphbjGFP) 2.15 min (L2-USDA110/MOS001/MOS004)	35 cycles
Final Extension	72	7 min	—
Hold	4	hold	—

Table 10 Overview of Level 1 Golden gate insert fragment and acceptor vector characteristics. These inserts and vectors were used for construction of L1-PaphGFP and L1-MOB

DNA fragment	Fragment type [insert / vector]	Length DNA fragment [bp]	Mw [kg/mol]	Concentration [ng/µL]	Acceptor vector
L0-bjGFP	Insert	2787	1839.42	251.5	pICH47732
L0-Paph	Insert	2599	1715.34	148	
MOB (fragment)	Insert		1389.96	33.5	pICH47751
pICH47732	L1-1 acceptor vector	4968	3278.88	149	—
pICH47751	L1-3 acceptor vector	4968	3278.88	69.5	—

Table 11 Composition of Golden Gate mixtures belonging to L1-PaphbjGFP and L1-MOB. Insert and vector are added in a 3:1 molecular ratio.

L1-PaphbjGFP	Volume [µL]
Purified L0-Paph vector, 1x diluted stock (148 ng/µL)	1.7
Purified L0-bjGFP vector, 1x diluted stock (251.5 ng/µL)	1.1
pICH47732 (L1-1 acceptor vector) 1x diluted stock (149 ng/µL)	1.1
Restriction enzyme BsaI (10 U/ µL)	1.0
T4 DNA ligase (30 U/µL)	0.5
10X T4 DNA ligase buffer	2.0
BSA (0.1%)	2.0
MiliQ	10.6

L1-MOB	Volume [µL]
PCR purified MOB amplicons, 1x diluted stock (33.5 ng/µL)	6.2
pICH47751(L1-3 acceptor vector) 1x diluted stock (69.5 ng/µL)	2.4
Restriction enzyme BsaI (10U/ µL)	1.0
T4 DNA ligase (30 U/µL)	0.5
10X T4 DNA ligase buffer	2.0
BSA (0.1%)	2.0
MiliQ	5.9

Table 12 Golden Gate reaction mixtures used for the construction of final level 2 Golden Gate modules specific for USDA110, MOS001 and MOS004. Insert and acceptor vector were added in a 3:1 molecular ratio.

L2-USDA110	Volume [µL]
L1-MOB (102 ng/µL)	1
L1-PaphGFP (543 ng/µL)	6.3
HR (USDA110) (13 ng/µL)	6.4
L3E (224 ng/uL)	1.5
L2 acceptor vector (45.5 ng/uL)	3
Restriction enzyme (10U) BpiI	1
T4 DNA ligase (10U from high concentration ligase 30 U/µL)	0.5
10X T4 DNA ligase buffer	2
BSA (0.1%)	2
MiliQ	0

L1-MOS001	Volume [µL]
L1-MOB (102 ng/µL)	6.3
L1-PaphGFP (543 ng/µL)	1
HR (MOS001) (22.5 ng/µL)	3.7
L3E (224 ng/uL)	1.5
L2 acceptor vector (45.5 ng/uL)	3
Restriction enzyme (10U) BpiI	1
T4 DNA ligase (10U from high concentration ligase 30 U/µL)	0.5
10X T4 DNA ligase buffer	2
BSA (0.1%)	2
MiliQ	0

L1-MOS004	Volume [µL]
L1-MOB (102 ng/µL)	6.3
L1-PaphGFP (543 ng/µL)	1
HR (MOS004) (21 ng/µL)	3.9
L3E (224 ng/uL)	1.5
L2 acceptor vector (45.5 ng/uL)	3
Restriction enzyme (10U) BpiI	1
T4 DNA ligase (10U from high concentration ligase 30 U/µL)	0.5
10X T4 DNA ligase buffer	2
BSA (0.1%)	2
MiliQ	0

Table 13 Primers used to validate the orientation in which the final level 2 construct becomes integrated into the genome.

Primer name	Primer sequence (5' →3')	Tm [°C]	Primer length (nt)
MAP390_GFPins_ins_Rv	CGGGGTGTTCTGCTGGTAG	60.1	19
MAP365_bjGFP_out_2_Fw	ATCACCGCCGACAAGCAG	59.0	18
MAP162_16S_63F	CAGGCCTAACACATGCAAGTC	56.4	21
MAP163_16S_1389R	ACGGGCGGTGTGTACAAG	54.9	18
MAP384_USDA110_ins_R	AGATATTCCGGGCGGATTTC	60.0	20
MAP385_USDA110_ins_Fw	CCAAGGACGGCGACTACAC	60.4	19
MAP386_MOS001_ins_Fw	GAAGCGGTATCTCTGAGGGC	60.0	20
MAP387_MOS001_ins_Rv	GATAGGGGTGCAGCTCGG	59.6	18
MAP388_MOS004_ins_Fw	GGAAAGGGCGGTCAGAGAT	59.1	19
MAP389_MOS004_ins_Rv	ATCTTCGGCGATGCGATCT	59.3	19

Cycle step	Temperature [°C]	Time	Cycles
Initial Denaturation	95	5 min	-
Denaturation	95	1 min	
Annealing	58	1 min	25 cycles
Extension	72	1 min	
Final Extension	72	7 min	-
Hold	4	hold	-

Table 14 Primers used for the calibration of the qPCR to quantify Arabidopsis root colonization. There primers are used to determine the relative abundance of Bradyrhizobium transformants carrying the bjGFP marker gene.

Primer name	Primer sequence (5' → 3')
MAP365_bjGFP_out_2	ATCACCGCCGACAAGCAG
MAP390_GFPins_ins_Rv	CGGGGTGTTCTGCTGGTAG

Table 15: qPCR reaction settings used during calibration of qPCR bjGFP primers.

	Temp [°C]	Time
Initial denaturation	95	5 min
39 cycles		
Step 1	95	30 s
Step 2	60	30 s
Melting curve		
	95	7 min
	65	5 s
	95	50 s

5 Bibliography

1. Ott T, Van Dongen JT, Günther C, Krusell L, Desbrosses G, Vigeolas H, et al. Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. Curr Biol. 2005;15(6):531–5.

2. Batut J, Andersson SGE, O'Callaghan D. The evolution of chronic infection strategies in the α - proteobacteria. Nat Rev Microbiol. 2004;2(12):933–45.

3. Remigi P, Zhu J, Young JPW, Masson-Boivin C. Symbiosis within Symbiosis: Evolving Nitrogen-Fixing Legume Symbionts. Trends Microbiol [Internet]. 2016;24(1):63–75. Available from: http://dx.doi.org/10.1016/j.tim.2015.10.007

4. Sachs JL, Skophammer RG, Regus JU. Evolutionary transitions in bacterial symbiosis. Proc Natl Acad Sci [Internet]. 2011;108(Supplement_2):10800–7. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.1100304108

5. VanInsberghe D, Maas KR, Cardenas E, Strachan CR, Hallam SJ, Mohn WW. Non-symbiotic Bradyrhizobium ecotypes dominate North American forest soils. ISME J [Internet]. 2015;9(11):2435–41. Available from: http://dx.doi.org/10.1038/ismej.2015.54

6. Jones FP, Clark IM, King R, Shaw LJ, Woodward MJ, Hirsch PR. Novel European free-living, non-diazotrophic Bradyrhizobium isolates from contrasting soils that lack nodulation and nitrogen fixation genes - A genome comparison. Sci Rep [Internet]. 2016;6(May):1–10. Available from: http://dx.doi.org/10.1038/srep25858

7. Schneijderberg M, Schmitz L, Cheng X, Polman S, Franken C, Geurts R, et al. A genetically and functionally diverse group of non-diazotrophic Bradyrhizobium spp. colonizes the root endophytic compartment of Arabidopsis thaliana. BMC Plant Biol. 2018;18(1):1–9.

8. Cytryn EJ, Jitacksorn S, Giraud E, Sadowsky MJ. Insights learned from pBTAi1, a 229-kb accessory plasmid from Bradyrhizobium sp. strain BTAi1 and prevalence of accessory plasmids in other Bradyrhizobium sp. strains. ISME J. 2008;2(2):158–70.

9. Ledermann R, Bartsch I, Remus-Emsermann MN, Vorholt JA, Fischer H-M. Stable Fluorescent and Enzymatic Tagging of *Bradyrhizobium diazoefficiens* to Analyze Host-Plant Infection and Colonization. Mol Plant-Microbe Interact [Internet]. 2015;28(9):959–67. Available from: http://apsjournals.apsnet.org/doi/10.1094/MPMI-03-15-0054-TA

10. Reinhold-Hurek B, Hurek T. Living inside plants: Bacterial endophytes. Curr Opin Plant Biol [Internet]. 2011;14(4):435–43. Available from: http://dx.doi.org/10.1016/j.pbi.2011.04.004

11. Smillie CS, Smith MB, Friedman J, Cordero OX, David LA, Alm EJ. Ecology drives a global network of gene exchange connecting the human microbiome. Nature [Internet]. 2011;480(7376):241–4. Available from: http://dx.doi.org/10.1038/nature10571

12. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the core Arabidopsis thaliana root microbiome. Nature. 2012 Aug;488(7409):86–90.

13. Rogers-Broadway KR, Karteris E. Amplification efficiency and thermal stability of qPCR instrumentation: Current landscape and future perspectives. Exp Ther Med. 2015;10(4):1261–4.

14. Okazaki S, Noisangiam R, Okubo T, Kaneko T, Oshima K, Hattori M, et al. Genome analysis of a novel bradyrhizobium sp. doa9 carrying a symbiotic plasmid. PLoS One. 2015;10(2):1–18.
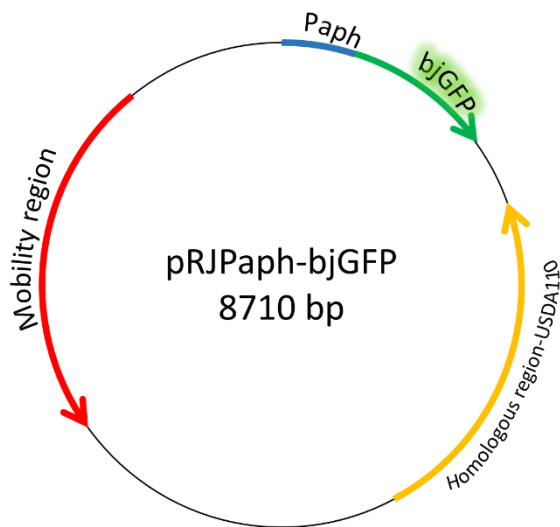
Figure S1 representation of pRJpaph-bjGFP, this plasmid is created by Ledermann et al. and is GG incompatible. This plasmid is used for chromosomal integration downstream of USDA110 scoI. We are going to make this plasmid compatible and specific of USDA110, MOS001 and MOS004. The construction of these plasmids will be done with Golden Gate modular cloning. Having strain specific Golden Gate plasmids extends the use for further research endeavours.
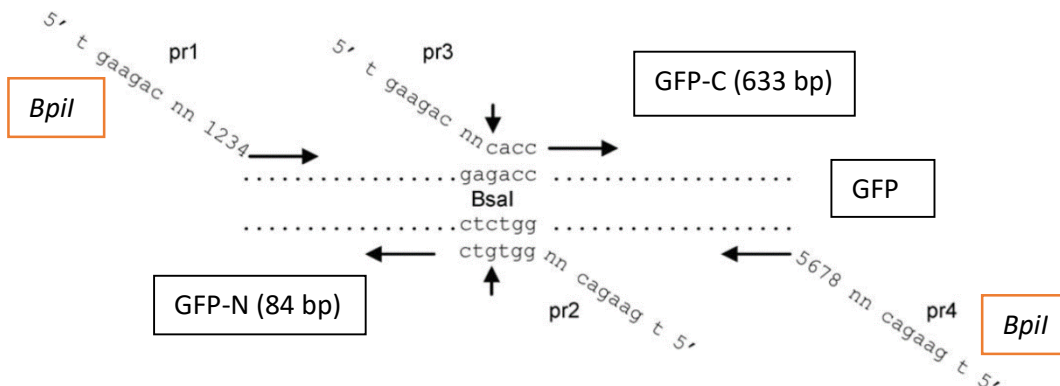

Figure S2: The GFP gene present in the pRJPaph-bjGFP plasmid contains an internal BsaI restriction site. We will remove this site by ampliying GFP in two parts, GFP-N and GFP-C. This will be done with primers containing BpiI recognition sequences (gaagac). After successful digestion of GFP-N and GFP-C and ligation into a universal level 0 acceptor can the GFP gene be ligated to the Paph promoter and inserted into a L1-1 acceptor plasmid.
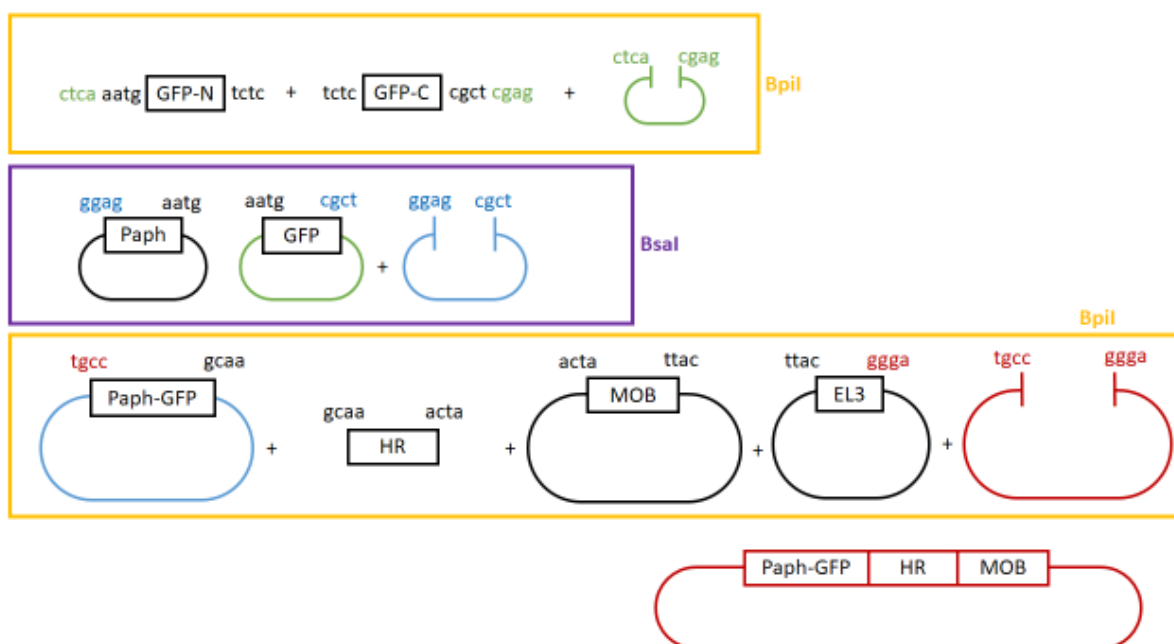
Figure S4: overview of the restriction enzymes used for each level, the overhangs created upon digestion and the final position of each construct in the level 2 acceptor. Note that HR will not be put in a level 1 acceptor plasmid because this ''module'' is specific for each strain and ensures homologous recombination of the plasmid with the chromosome.

Content	Cq (rep1)	Cq(rep 2)	average Cq	Template ng	template pg	# DNA template	10 log (# DNA template)
USDA110Ledermann	21.36	20.90	21.13	2.20E+01	2.20E+04	2.20E+06	6.34
USDA110Ledermann	21.21	22.89	22.05	2.20E+00	2.20E+03	2.20E+05	5.34
USDA110Ledermann	23.31	27.09	25.20	2.20E-01	2.20E+02	2.20E+04	4.34
USDA110Ledermann	27.04	29.50	28.27	2.20E-02	2.20E+01	2.20E+03	3.34
USDA110Ledermann	29.02	30.16	29.59	2.20E-03	2.20E+00	2.20E+02	2.34
USDA110Ledermann	29.99	29.32	29.65	2.20E-04	2.20E-01	2.20E+01	1.34
USDA110Ledermann	29.43	29.63	29.53	2.20E-05	2.20E-02	2.20E+00	0.34
USDA110Ledermann	29.28	29.72	29.50	2.20E-06	2.20E-03	2.20E-01	-0.66
USDA110Ledermann	29.79	29.35	29.57	2.20E-07	2.20E-04	2.20E-02	-1.66
USDA110Ledermann	29.28	29.62	29.45	2.20E-08	2.20E-05	2.20E-03	-2.66
water	29.88	30.09	29.99	0.00E+00	0.00E+00		
water	30.91		30.91	0.00E+00			
L0-GFP	5.21	6.72	5.97	5.03E+01	5.03E+04	1.65E+10	10.2
L0-GFP	6.13	6.08	6.10	5.03E+00	5.03E+03	1.65E+09	9.2
L0-GFP	8.53	9.31	8.92	5.03E-01	5.03E+02	1.65E+08	8.2
L0-GFP	12.50	12.99	12.75	5.03E-02	5.03E+01	1.65E+07	7.2
L0-GFP	15.65	15.63	15.64	5.03E-03	5.03E+00	1.65E+06	6.2
L0-GFP	19.48	18.67	19.07	5.03E-04	5.03E-01	1.65E+05	5.2
L0-GFP	22.55	22.27	22.41	5.03E-05	5.03E-02	1.65E+04	4.2
L0-GFP	25.37	25.41	25.39	5.03E-06	5.03E-03	1.65E+03	3.2
L0-GFP	28.32	29.91	29.11	5.03E-07	5.03E-04	1.65E+02	2.2
L0-GFP	28.93	39.55	34.24	5.03E-08	5.03E-05	1.65E+01	1.2
water	30.11		30.11	0.00E+00	0.00E+00		
water				0.00E+00	0.00E+00		

Figure S5: Total dataset used to calibrate bjGFP specific qPCR primers.

Endogeneous *Bradyrhozium* plasmids

Endogenous plasmids are by definition present in addition to the main chromosome. Further, they are circular, vary in size between approximately 1 Kbp to 1 Mbp, do not contain genes essential for the main survivability of the bacterium under normal conditions and replicate independent of the main chromosome. The genes located on plasmids often enhance the survivability of the organism carrying it.

Several *Bradyrhizobium* strains are reported to harbor either symbiotic or non-symbiotic plasmids (14). Symbiotic plasmids accommodate the nodulation and nitrogen fixation gene clusters, essential for the establishment of a nitrogen-fixing symbiotic lifestyle (8). Possibly due to methodological limitations have few reports been written about the presence of endogenous plasmids in *Bradyrhizobium* spp. strains. To ascertain whether the presence of accessory plasmids in *Bradyrhizobia* were more common than previously reported, assays were carried out by Cytryn et al. They used a PFGE-CHEF assay to screen 46 *Bradyrhizobium* spp. strains. All strains originated from geographically diverse locations (i.e. United States, China and Thailand). Choosing geographically diverse locations excluded the presence of plasmids due to geographic origin. Results showed that one quarter of the strains examined in the assay harbored high-molecular weight plasmids (±300 Kb). Moreover, 16S rDNA analysis revealed these strains to be either *Bradyrhizobium elkani* or *Bradyrhizobium japonicum*. The majority only contained one single plasmid. A few of the strains harbored either two or three plasmids. These results are in contrast to plasmid copy numbers found in earlier experiments carried out by (Masterson et all., 1982).

Finally, the fact that some *Bradyrhizobium* spp. contain endogenous plasmids, whilst others do not, raises questions. One reason for this observation could be that *Bradyrhizobium* spp. containing endogenous plasmids gain evolutionary advantage due to environmental significant genes located on the plasmid. One example of this is *Bradyrhizobium* sp. BTAi1, which contains a 300 Kb plasmid harboring genes for heavy metal resistance (Cytryn et al., 2008). A second theory for strain specific plasmid compatibility is the presence of a *Bradyrhizobium* compatible origin of replication.

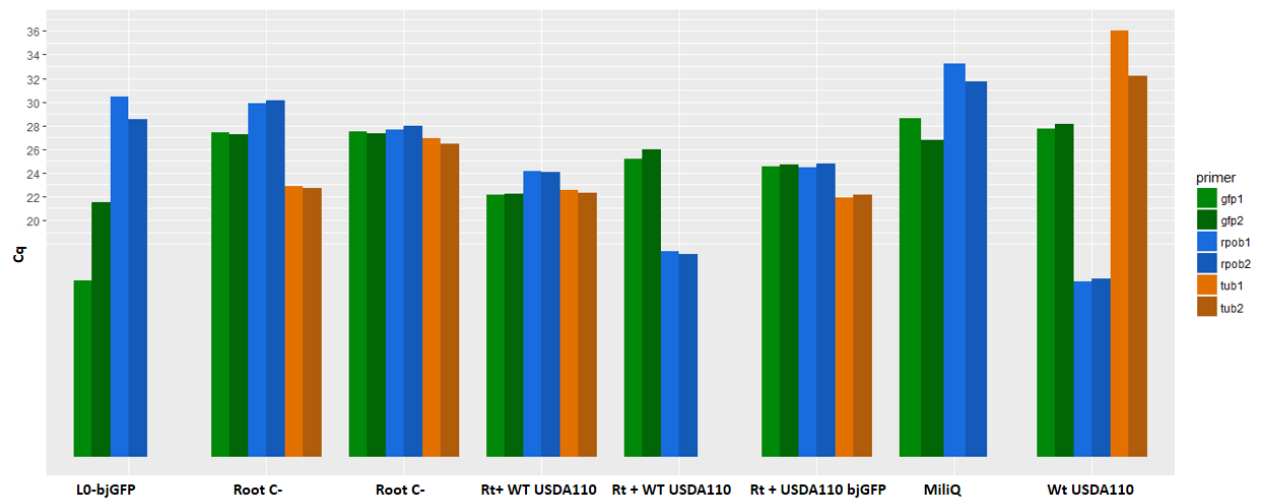**Figure S6** additional information about endogenous Bradyrhizobium spp. plasmids.

Figure S7: qPCR results on genomic DNA isolated from *Arabidopsis* roots inoculated with WT USDA110 and USDA110 transformants carrying pRJPaph-bjGFP. This data was generated to test whether the primer pairs were able to show the expected results.