Final Report of Grant K-1196537

Sinorhizobium species undergo an irreversible differentiation in *Medicago truncatula* symbiotic nodules leading to the formation of elongated polyploid nitrogen fixing bacteroids that convert atmospheric dinitrogen into ammonia. This terminal differentiation is directed by the host plant and involves hundreds of nodule specific cysteine-rich (NCR) peptides. It was demonstrated that two members of the *NCR* peptide gene family, *NCR169* and *NCR211*, are essential for effective symbiotic interaction. One of the main objectives of this proposal was the (i) analysis of the regulatory elements in the promoters of genes *NCR169* and *NCR211* and define the residues essential for the function of the encoded peptides. The other main goal of the proposal was the (ii) identification of additional *NCR* peptide genes that are essential for differentiation of rhizobia.

Objective 1: Analyzing of the regulation of genes *NCR*169 and *NCR211* essential for bacteroid differentiation

In order to investigate the regulation of genes *NCR169* and *NCR211*, the function of potential cis-elements (i) and the specificity (ii) of the promoters required for the proper expression were analyzed. (i) In a previous study (Nallu et al. 2013), 41-50 bp long unique DNA motifs were identified in the 1 kb promoter sequences of several *NCR* peptide genes. Five conserved motifs were detected in the promoters of *NCR* peptide genes in a non-conserved position and order. It was hypothesized that these motifs are related to the with nodule- and zone-specific expression pattern of *NCR* peptide genes. In order to identify whether the promoters of *NCR169* and *NCR211* peptide genes contained these conserved motifs, we analyzed their promoter sequence using the by the FUZZNUC nucleic acid pattern search online program (http://www.hpa-bioinfotools.org.uk/pise/fuzznuc.html). We found that the conserved motifs were clustered in the ~400bp long region upstream of the start codons the *NCR169* and *NCR211* genes. To test the function of the identified conserved motifs two truncated promoter versions of previously used *NCR169* (1178 bp) and *NCR211* (~2 kb) promoters were generated. A ~800 bp and a 400 bp region upstream of the start codon long were fused with the *NCR169* and *NCR211* genes and the constructs were tested in a genetic complementation assays in the *dnf7-2* (defective in *NCR169*) and *dnf4* (defective in *NCR211*) mutant plants using the *Agrobacterium rhizogenes*-mediated hairy root transformation system. We found that even the ~400bp long promoter versions were sufficient to restore the effective symbiotic interaction of mutants *dnf7-2* and *dnf4* with rhizobia (Figure 1A and B.) indicating that the regulatory sequences necessary for the proper expression of these two *NCR* peptide genes are located on these minimal promoter regions reside.

The promoter specificity was assayed with exchanging the minimal promoters of *NCR169* and *NCR211* genes. The constructs of the *NCR169* and *NCR211* genes under the control of the minimal (*pNCR169*$_{400bp}$*::NCR211* and *pNCR211*$_{400bp}$*::NCR169*) and the full length promoters (*pNCR169*$_{1178bp}$*::NCR211* and *pNCR211*$_{2kb}$*::NCR169*) were introduced into *dnf7-2* and *dnf4* roots using *A. rhizogenes*-mediated transformation and the capacity of the constructs were assayed for restoring of the symbiotic interaction in mutant plants.

The full length promoter exchanged between the two *NCR* peptide genes did not affect the complementation capacity but the construct of *NCR169* gene driven by ~400bp *NCR211* promoter could not restore the effective symbiotic interaction in *dnf7-2* nodules indicating the lack of cis-elements in the minimal promoter of *NCR211* for the proper expression of *NCR169* (Figure 1A). In contrast, the gene *NCR211* driven by the minimal promoter of *NCR169* was able to complement the symbiotic defect of *dnf4* (Figure 1B).

These complementation experiments also suggested the importance of the conserved motifs in the minimal promoters for the proper expression pattern of *NCR* genes. In order to further analyze the relevance of these cis-regulatory elements, two conserved motifs (CM1 and CM2) were modified with non-specific sequences in the minimal promoter regions of *NCR169* and *NCR211*. The modified version of CM1 was unable to restore the symbiotic phenotype either in *dnf7-2* or *dnf4* indicating the requirement of this motif for the regulation of genes *NCR169* and *NCR211* (Figure 2).

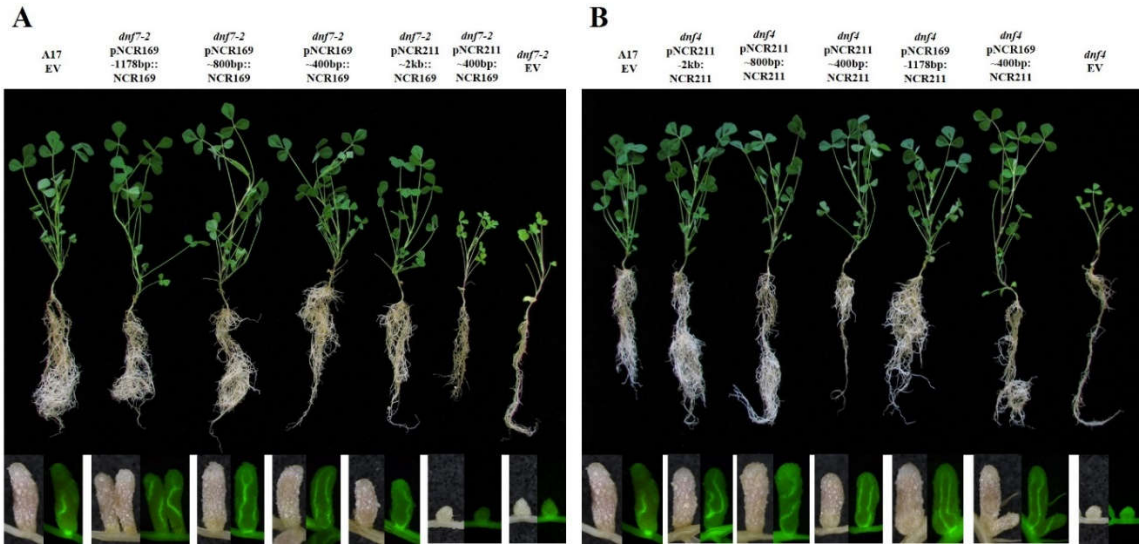


Figure 1: Determination of the minimal promoter region of genes *NCR169* (A) and *NCR211* (B) and the investigation of the specificity of the two NCR promoter. EV: empty vector.

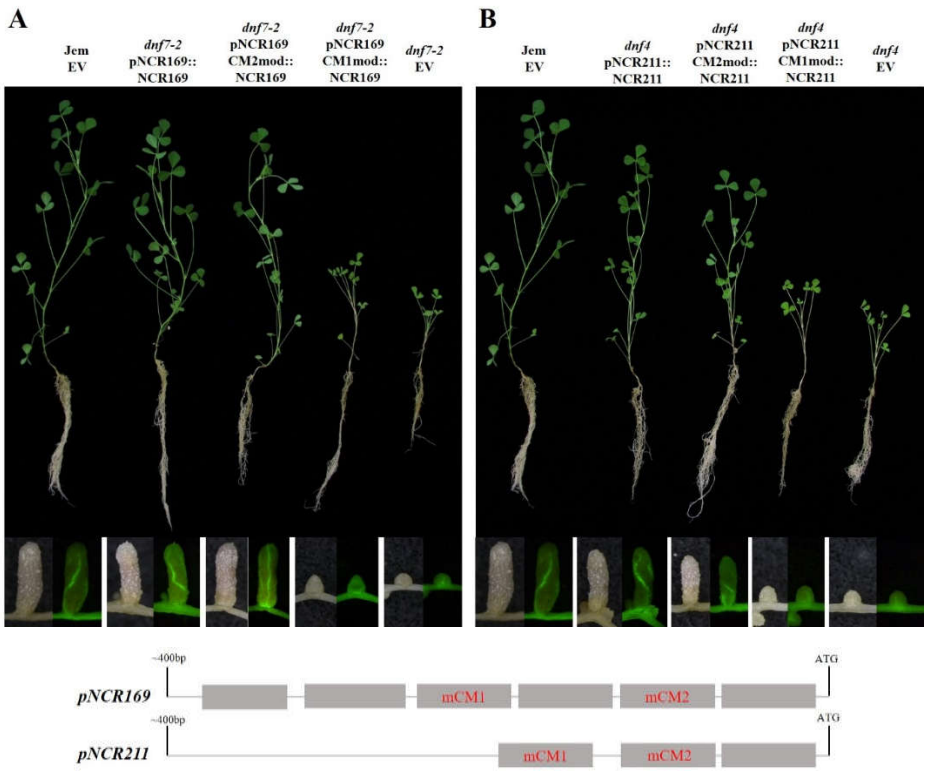


Figure 2: Functional study of CM1 conserved motif for the proper regulation of *NCR169* and *NCR211* genes in *dnf7-2* (A) and *dnf4* (B) mutants. EV: empty vector.

In order to further dissect the function of CM1 for the promoter activity, additional mutant versions of NCR169-CM1 were generated with modification of four larger ~15-20bp long overlapping regions (CM1 A-D) and substitution of 4 and 5 bps for non-specific nucleotides (CM1 E-F). The constructs were introduced into the *dnf7-2* mutant and the two overlapping 31bp long modified CM1 regions - CM1-A and-B - failed to restore the symbiotic phenotype indicating the requirement of these regions for the promoter activity of *NCR169* (Figure 3.).

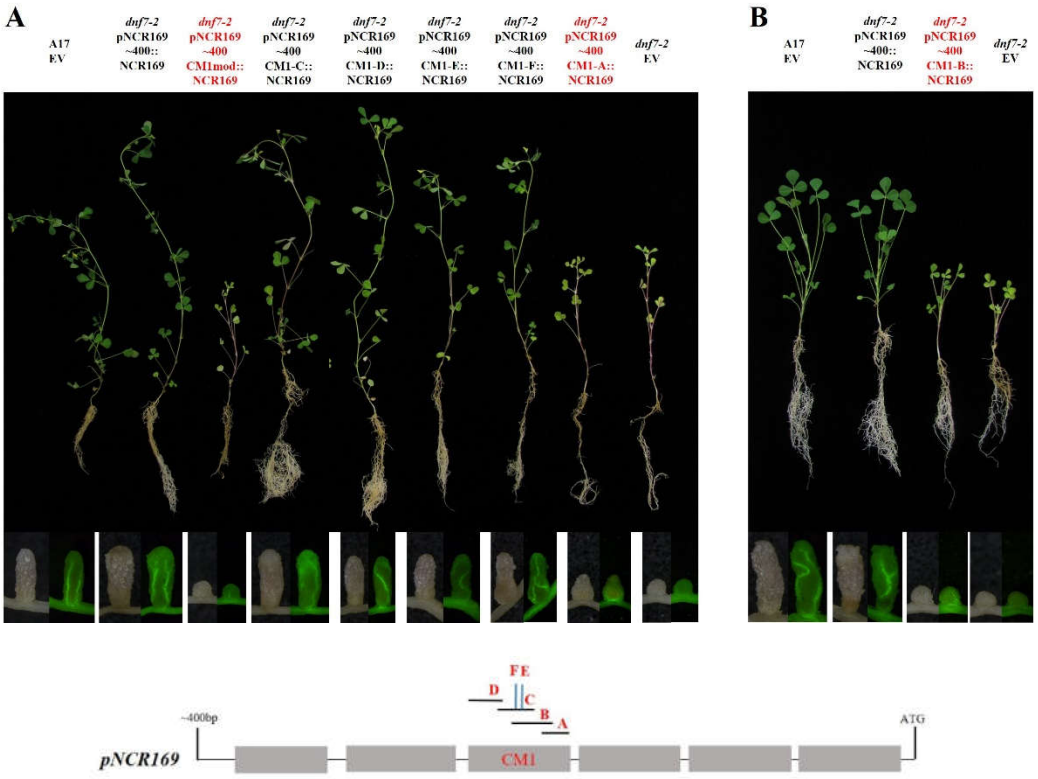


Figure 3: The investigation of complementation capacity of different modified version of NCR169-CM1 region (CM1-A-F) in *dnf7-2* mutant in two experimental series. EV: empty vector

These results indicated the primary role of CM1 motif in the activity of *NCR169*. In order to test the binding capacity of CM1 to transcription factors, we initiated a collaboration with Fernanda De-Carvalho-Niebel (Laboratory of Plant-Microbe Interactions, Toulouse, France) who is an expert in screening for TFs using yeast one-hybrid (Y1H) system. To confirm the transcription activator capacity of CM1 motif, we fused CM1 to the minimal 35S promoter and tested the transcriptional activity of the construct in a GUS reporter system. The constructs were introduced into wild-type *M. truncatula* roots using *A. rhizogenes*–mediated transformation system. Single copy of CM1 motif did not show GUS activity in wild-type nodules, therefore a tetrameric form of the CM1 motif and an extended region of CM1 including CM3 (CM1-3) were synthesized and fused to 35S minimal promoter and the *GUS* gene. The tetramer version of Cm1 and CM1+CM3 were able induce the activity of the *GUS* gene in the infected cells indicating their transcription activator capability in nodules. The tetramer+35Smin promoter constructs were also fused with the gene *NCR169* and tested for genetic complementation. Both constructs were able to restore the symbiotic phenotype of *dnf7-2*. These results indicated that the CM1 and CM1-3 motifs fused with a minimal promoter were competent to activate *NCR169* and appropriate to be used in Y1H screens to identify DNA binding proteins (Figure 4.).

The tetramer constructs were cloned into the yeast compatible vector and the Y1H screens are currently in progress with the collaboration partner. The background expression level of the two tetramer constructs (*CM1+35Smin::GUS* and *CM1+CM3+35Smin::GUS*) were infiltrated into *N. benthamiana* leaves and tested in GUS reporter assay. Both promoter constructs showed relatively low expression background which is an essential criteria for the successful Y1H screen. The constructs were successfully transformed into Yeast YM421 cells and the Y1H screening of a nodule cDNA library is in progress.

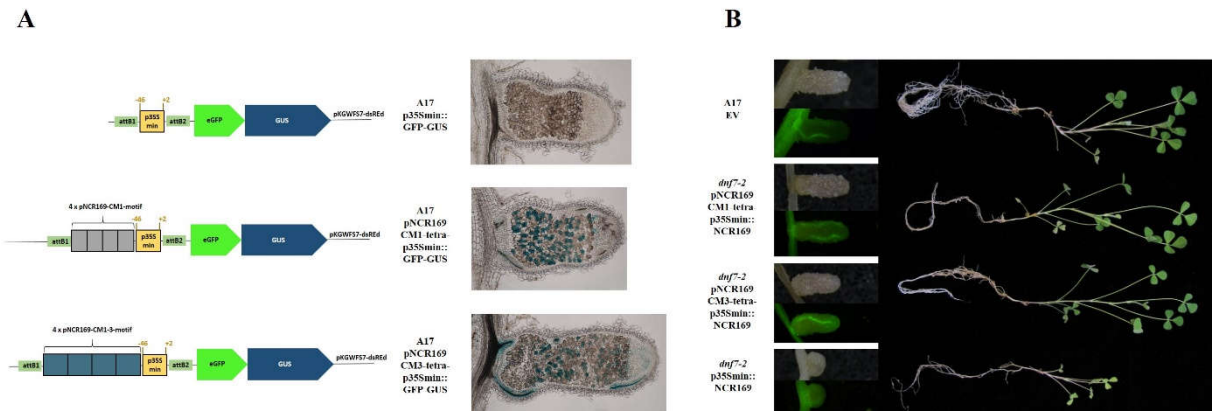


Figure 4: The transcription activator capability of the tetramer form of NCR169-CM1 and -CM3 promoter regions (A). The *NCR169* gene driven by pNCR169-CM1 or CM1+CM3 tetramers fused with 35S minimal promoter were able to restore the symbiotic phenotype of *dnf7-2*. EV: empty vector

The previous results showed that the gene *NCR169* driven by the minimal promoter of *NCR211* could not restore the symbiotic phenotype of *dnf7-2* indicating the lack of *NCR169*-specific gene regulatory element(s) in the minimal promoter of *NCR211*. We carried out a comparative sequence analysis of the minimal promoters of *NCR169* and *NCR211* and found that two conserved motifs, CM4 and CM5 were hardly or even no detectable in the minimal promoter of *NCR211* because of the very low level of similarity to these motifs. We were wondering if these motifs had function in the activity of *NCR169* minimal promoter. Therefore, we generated the mutant versions of the motifs CM4 and CM5 in the minimal promoter of *NCR169* and additionally a shorter 20 bp long region (I-1) which was also undetectable in the minimal promoter of *NCR211*. These mutated promoter versions were able to rescue the symbiotic defect of *dnf7-2* indicating that the sequences of CM4, CM5 and I-1 are not essential for the appropriate expression of *NCR169*. We also inserted the *NCR169*-specific sequences of I-1 and I-2 (in two parts I2-A and I2-B) into the minimal promoter of *NCR211,* fused with *NCR169* and introduced them into the roots of *dnf7-2.* None of these promoter versions were able to rescue the *dnf7-2* Fix- phenotype suggesting regions I-1 and I-2 do not contain *NCR169*-specific sequences for the expression of *NCR169* (Figure 5).

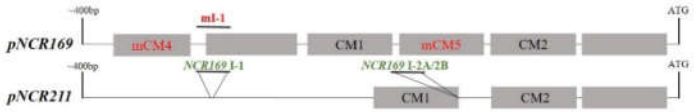


Figure 5: Modification of the CM4, CM5 and I-1 region in minimal promoter of *NCR169* and the insertion of I-1, I2 *NCR169*-specific region into short promoter of *NCR211*.

The *NCR169* gene controlled by the longer versions of *NCR211* promoter (~800bp and 2kb) were able to rescue the Fix- symbiotic phenotype in *dnf7-2* suggesting that the additional regulator elements required for the appropriate expression of *NCR169* are located upstream of the ~400 bp promoter of *NCR211*. To identify these regulatory elements, three overlapping fragments (F1-F3) between 400-800 bp upstream of the start codon of *NCR211* were fused with the 400 bp minimal promoter of *NCR211*, respectively and with the gene *NCR169*. The promoter constructs were tested in complementation assays of mutant *dnf7-2*. None of the promoter constructs could restore the *dnf7-2* symbiotic phenotype which may indicate that either the fragmentation destroyed the required promoter regions or nonsequence-specific regions are needed for the regulation of *NCR169* (Figure 6). Further analysis of promoter region of *NCR211* between 400-800 bp upstream of the start codon is still in progress.

The order of the described motifs is not conserved in the promoters of all *NCR* peptide genes. We noticed that motif CM5 is absent between the motifs CM1 and CM2 in the minimal promoter of *NCR211*. In order to further dissect the differences between the *NCR211* and *NCR169* promoters, we rearranged the order of motifs CM1 and CM5 in the minimal promoter of *NCR169* and inserted CM5 into the minimal promoter of *NCR211* between motifs CM1 and CM2. Both promoter variations driven *NCR169* gene constructs were tested in complementation experiments of mutant *dnf7-2*. The rearranged order of the CM1 and CMI5 in promoter of *NCR169* did not caused any defects in the complementation capability of the construct. In addition, the insertion of CM5 into the promoter of *NCR211* did not result in the rescue of the symbiotic phenotype of *dnf7-2*, therefore, we concluded that the motif CM5 is not essential for the appropriate expression of *NCR169*.

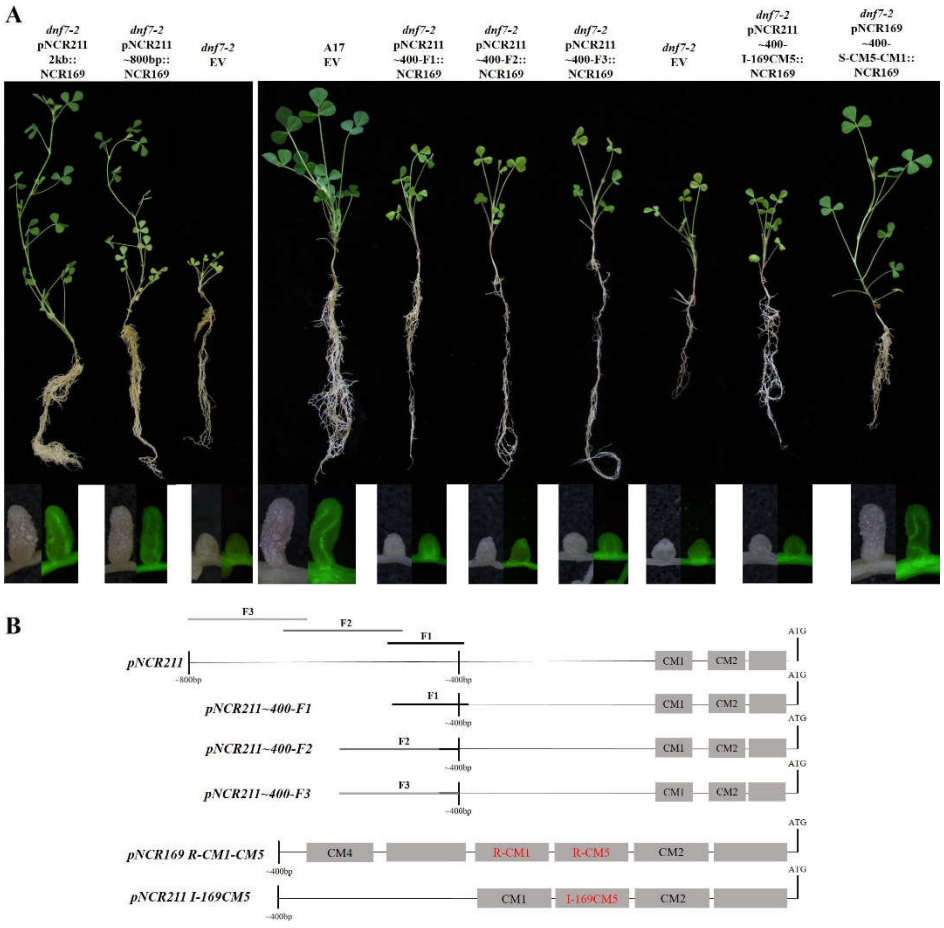


Figure 6: Investigation of the complementation efficiency of *NCR211* promoter variations in *dnf7-2* mutant plant and the role of the CM5 motif in the *NCR169* specific gene regulation. R-CM1 and R-CM5 indicate rearranged order of motifs CM1 and CM5.

The PlantRegMap online program (http://plantregmap.gao-lab.org/index.php) was used to scan the minimal promoters of *NCR169* and *NCR211* to predict transcription factor (TF) binding sites using the *M. truncatula* TF database. We found two predicted binding sites in *NCR169* (C2H2-MYB and HD-ZIP) which have not yet been modified to investigate their role in the regulation of *NCR* genes. We mutated these potential TF binding sites in the promoters *NCR211* and *NCR169*. The promoters with modified TF binding sites were able to rescue the symbiotic phenotypes of the corresponding mutants *dnf4* and *dnf7-2* indicating that the predicted C2H2 and HD-ZIP TF binding sites do not have essential role in the regulation of the two *NCR* peptide genes (Figure 7).

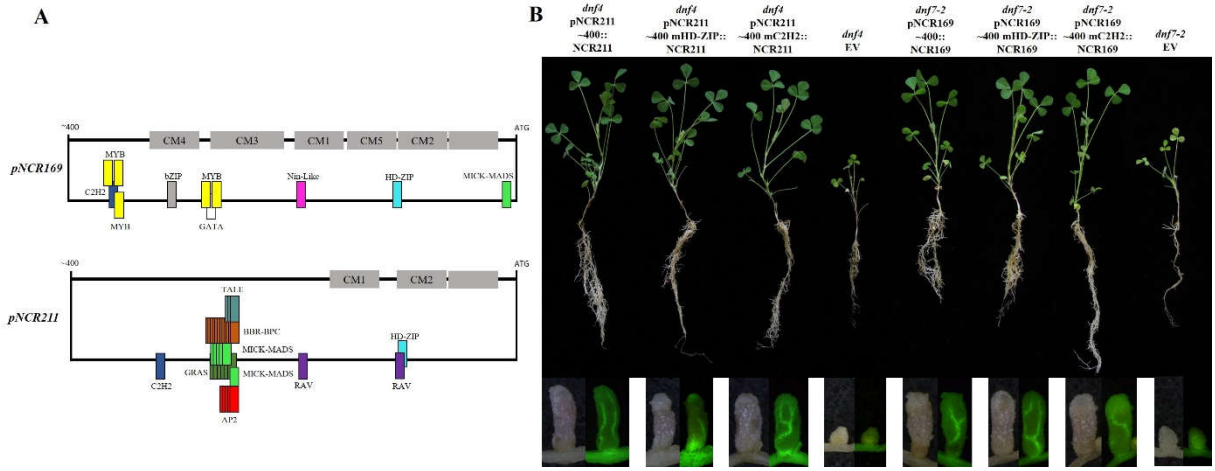


Figure 7: Prediction of transcription factor binding sites in the minimal promoter regions of *NCR169* and *NCR211* (A) and investigation of the role of the predicted C2H2 and HD-ZIP TF binding sites in both promoter regions (B).

Objective 2: Identification of residues required for the function of NCR169 and NCR211 peptides

The NCR169 and NCR211 encode highly divergent mature peptides with four cysteine in conserved positions and with distinct biochemical properties. The NCRs characterized by different charges, NCR169 is slightly cationic (pI = 8.45) while NCR211 is anionic (pI = 5.38). It is known from previous results that the four cysteine residues have an essential role in the function of the peptides. In order to identify additional essential residues for the function of peptides NCR169 and NCR211, we generated constructs coding for the mutant versions of the peptides and tested their complementation efficiency in mutants *dnf7-2* and *dnf4*.

The residues to be modified in mature peptide NCR169 were selected based on alignment of known unique NCR peptides and some representative members of the NCR peptide family of *M. truncatula*. Eight constructs were generated to encode for the substituted residues to amino acids with distinct biochemical features (S1-S8) for the NCR169 mature peptide. Two substitutions between two conserved cysteine residues resulted in defective symbiotic nitrogen fixing nodules on *dnf7-2* transgenic roots indicating the requirement of these two residues for the function of NCR169 (S6). In case of only single amino acid substitution in the peptide NCR169 (S7), the transformed plants were partially deficient in symbiotic nitrogen fixation. Transgenic plants developed pale green shoots indicating an ineffective symbiotic interaction with rhizobia but we could detect a mixture of few wild type smaller pink and several white undeveloped nodules on transgenic roots (S7) indicating that the substitution impaired the function of the NCR169.

Additional constructs were generated which coded for modified versions of peptide NCR169 containing 4-5 NCR211 and NCR343 (a novel unique NCRs identified recently, see Objective III) specific amino acids between the first two and the second two conserved cysteine positions. The complementation capacity of these chimeric gene constructs was tested in *dnf7-2* but none of them ceased the complementation efficiency of the constructs (C1-C8). The results suggest the modified regions can be important for the appropriate peptide function but they are not essential for the unique role of NCR169.

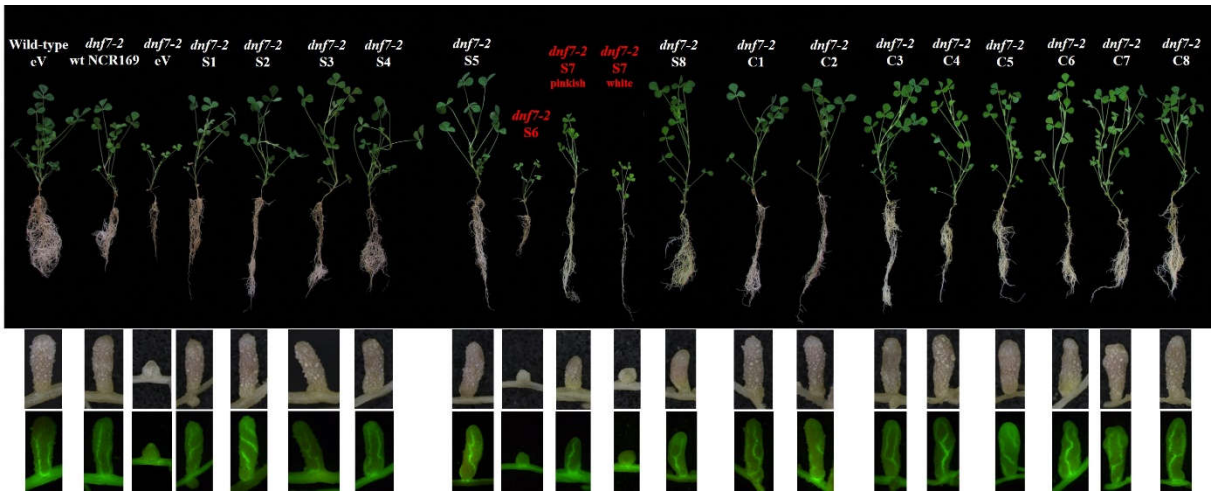


Figure 8: Complementation assays in *dnf7-2* using constructs for modified NCR169 mature peptides. S1-S8: substitution of amino acids with distinct biochemical futures in NCR169 mature peptide and C1-C8: chimeric constructs of NCR169 peptide containing NCR211 and NCR343 specific residues.

It is still unknown how NCR peptides regulate bacteroid differentiation; whether they function alone and simultaneously or they act in complexes with other peptides. In order to identify other interacting peptides or potential other protein partners of NCR169, a differential proteomic analysis was planned following pull-down experiments using wild-type NCR169 and the loss of function version of NCR169 peptide with two modified amino acids. For affinity purification Strep and GFP tags were fused to the C-terminus of the NCR169 wild-type and mutated peptides. We tested the function of the tagged peptides in complementation assays and found that none of the fusion tag affected the function of the wild-type NCR169 (Figure 9). We purified the recombinant proteins from transgenic nodules and attempted to detect them by western blot experiment. Although we could not detect the strep tag-labeled recombinant peptides, the presence of GFP-tagged version of the NCR169 was confirmed by western blot. In addition to the recombinant peptides, the presence of the free GFP was identified which raised concerns about the efficiency of the planned pull-down experiment. To reduce the amount of free GFP, we purified bacteroids from the nodules. In order to perform successful pull-down experiments, we are still working on the improvement of the methodology to enhance bacteroid purification and decrease the amount of free GFP (Figure 9).

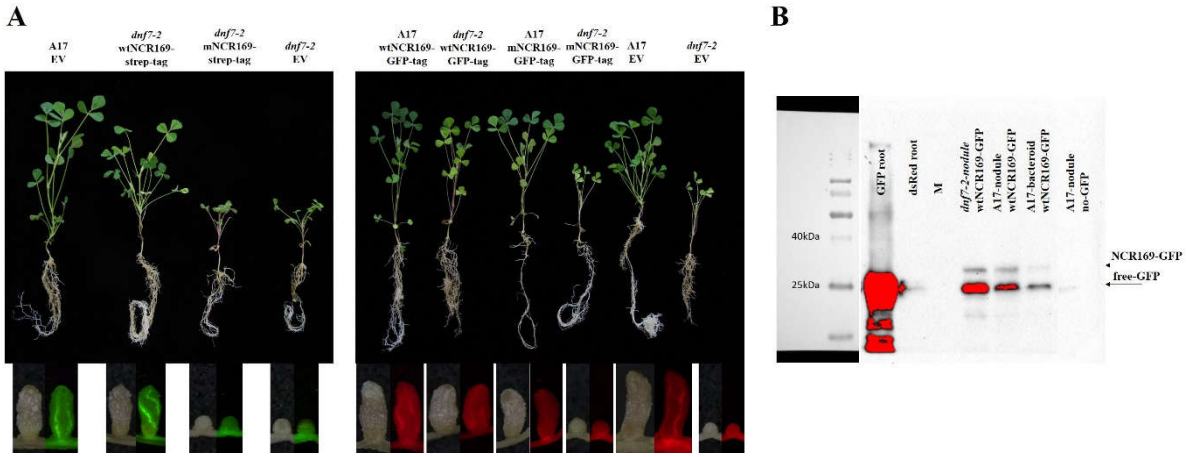


Figure 9: Investigation of the strep-tagged and GFP-tagged NCR169 peptide and its mutated versions (m). NCR169 constructs in a series of complementation assay on *dnf7-2* and wild-type plants (A). Western blot analyses to detect the recombinant GFP-tag labeled wtNCR169 (B).

The deletion in *dnf4* mutant eliminated two *NCR* peptide genes, *NCR211* and *NCR178*. Although peptides show high similarity (62% identity), only *NCR211* was able to rescue the *dnf4* phenotype (Kim et al. 2015). To reveal why NCR178, a highly homologous peptide cannot replace the function of NCR211 when introduced into the *dnf4* mutant lacking both peptide, hybrid genes were constructed and introduced into the *dnf4* mutant via *A. rhizogenes* mediated transformation. The transgenic plants were inoculated by *S. mediacae* strain WSM419 and the ability of the constructs to complement the mutation were evaluated 4-5 weeks post-inoculation (wpi). As the expression level of the two genes is different, first, the promoter sequences were exchanged, but it did not affect the complementation ability of the genes, i.e. the *NCR178* promoter driven *NCR211* gene could complement the mutant, while the *NCR211*-driven *NCR178* gene was incapable of complementation. Complementation experiments with additional hybrid genes created by further sequence swaps (promoter and N-terminal coding sequences from the *NCR178* gene and C-terminal coding sequences from the *NCR211* gene) revealed that the functional difference is originated from the last eight amino acids of the two peptides. The construction of additional genes with single amino acid changes is in progress.

To study the role of NCR peptides in the near future, a new vector system allowing the selection of transgenic roots by naked eye was developed and tested. The vectors – termed as pPurpleRoot plasmids – were developed from the widely used pCAMBIA plasmids and contain the *MtLAP1* gene coding for an R2R3 type MYB transcription factor (Peel et al. 2009). This transcription vector was shown to direct the production and accumulation of the secondary metabolites, anthocyanins with purple color in *Medicago truncatula*, *M. sativa* and *Trifolium repens* as well as with a lower intensity and homogeneity in tobacco after transient (Picard et al. 2013) or ectopic (Peel et al. 2009) expression. As anthocyanins might have antimicrobial activity and its constitutive production might affect the rhizobium - plant interaction, the constitutive *CaMV35S* promoter of pPurpleRootC was replaced by two tissue-specific promoters from *Arabidopsis* (Lee et al. 2006), namely by *pAtS5* (*At5g24800*) and *pAtE47* (*At2g37950*) that in *Lotus japonicus* provided pericycle/phloem and endodermis/pericycle enriched promoter activities, respectively (Gavrilovic et al. 2016). These plasmids were named pPurpleRootP (pericycle/phloem) and pPurpleRootE (endodermis/pericycle). The vectors were successfully tested for the genetic complementation of the nodule developmental defect of the *M. truncatula dnf7* (NCR169-deficient) mutant via *A. rhizogenes* mediated transformation. The complementation experiments showed that species-specific allelic variations (*Melilotus albus* vs. *Medicago truncatula*) and a mutation (K40R) preventing posttranslational acetyl modification of the NCR169 peptide do not affect the symbiotic interaction of *Medicago tuncatula* cv. Jemalong with *Sinorhizobium medicae* strain WSM419.

Objectives 3 and 4: Cloning of genes that govern bacterial differentiation in *M. truncatula* and other legume nodules

Based on the symbiotic phenotype of mutant FN8338 and the non-differentiated thizobia detexcted in mutant nodules, we speculated that line FN8338 is impaired in a NCR peptide gene. The genetic mapping identified the position of the NF8338 mutant locus on chromosome 3 close to the known symbiotic gene, *symCRK*. The sequence analysis of *symCRK* from the NF8338 mutant detected a single-base pair deletion at the position of 1815 bp in the coding sequence which generates a frameshift in the encoded protein from the position of 606 residue. The complementation experiment of the NF8338 mutant with the wild-type *symCRK* gene proved that the deletion in *symCRK* was responsible for the symbiotic phenotype of NF8338.

FN9363, *Mtsym19* (TR183) and *Mtsym20* (TRV43) ineffective symbiotic (Fix-) mutants were identified formerly in genetic screens for symbiotic nitrogen fixation mutants of fast neutron bombarded (NF-FN9363) or γ–ray irradiated (*Mtsym19* and *Mtsym20*) *M. truncatula* A17 (NF-FN9363) or Jemalong (*Mtsym19* and *Mtsym20*) populations. The mutants showed the symptoms of nitrogen starvation under symbiotic conditions and mutant nodules were defective in the persistence of rhizobia in the nitrogen fixation zone of mutant nodules (Figure 10).

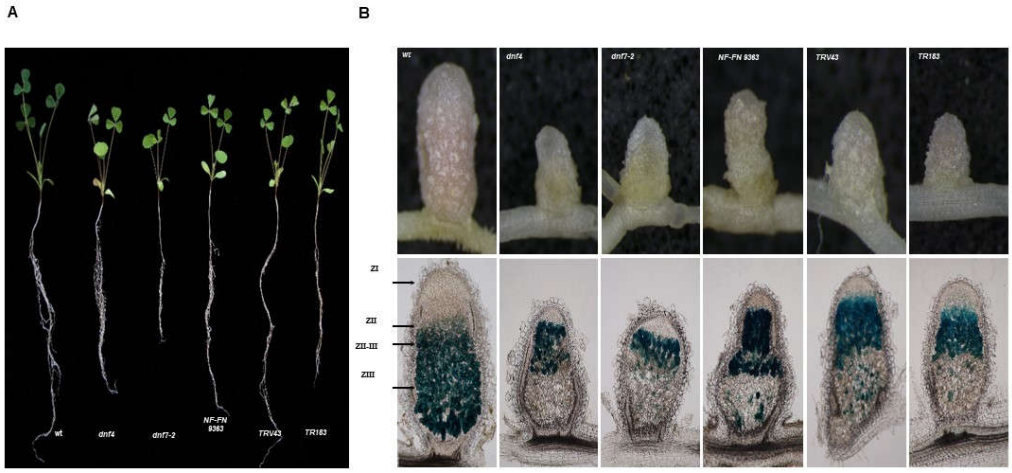


Figure 10. The symbiotic phenotype of wild-type and FN9363, *Mtsym19* and *Mtsym20* symbiotic mutant *M. truncatula* plants 3 wpi with *S. medicae* WSM419 carrying the β-galactosidase marker gene (pXLGD4 plasmid). Wild-type plants developed functional pink nodules which showed the characteristic zonation of indeterminate nodules and were colonized with rhizobia. White and slightly elongated nodules showing nodulation zones were formed on mutant roots. Nodule cells were colonized only in the infection and interzones of mutant nodules.

The mutant locus of FN9363 was positioned on chromosome 6 (LG 6) between genetic markers Crs and MtB178 (Figure 11A). To facilitate the identification of gene defective in NF-FN9363, the genomic position of copy number variations detected in the mutant line using the array-based CGH (https://medicago-mutant.dasnr.okstate.edu/mutant/index.php) was analyzed. One probeset of CHR06FS008747671, corresponding to the genomic position of the symbiotic loci of FN9363, indicated a potential deletion in the mutant genome. The deletion was verified by PCR-based markers that defined an almost 500 kb deletion in the genome of FN9363 (Figure 11B). The deletion in NF-FN9363 removed more than 30 predicted genes or gene models including four *NCR* genes, *NCR341* (6g0467551), *NCR343* (6g0467511), *NCR344* (6g0467581) and *NCR345* (6g0467571) which were the primary candidates responsible for the symbiotic phenotype of FN9363

The mutant loci of *Mtsym19* and *Mtsym20* were positioned into the same genomic region on chromosome 4 (LG 4) between genetic markers 4g0020111, 4g0020421 and 4g0020631 (Figure 11D). The similar position of the mutant loci of *Mtsym19* and *Mtsym20* indicated either the defect of two different genes that are in proximity to each other or the malfunction of the same gene in the two mutants. Former genetic analysis suggested that *Mtsym19* and *Mtsym20* belong to different complementation groups (Morandi et al 2005). In order to verify their allelic relationship, allelism test was carried out using F3 mutant plants selected from the populations of *Mtsym19*xA20 and *Mtsym20*xA20 crosses. The progeny of the three successful crossings were all Fix– (n=8) indicating that *Mtsym19* and *Mtsym20* belong to the same complementation group. We presumed that the mutations in *Mtsym19* and *Mtsym20* affected the level of transcripts of genes impaired in the mutant plants. In order to promote gene identification, the transcript abundance in nodules of *Mtsym19*, *Mtsym20* and *Mtsym18*, belonging to an independent symbiotic complementation group (Morandi et al. 2005), was analyzed at 2 wpi with *S. medicae* WSM419. Filtered sequence reads were mapped against the *M. truncatula* genome assembly (A17r5.0) and the number of reads was analyzed in the genomic region between genetic markers 4g0020111 and 4g0020631. No or reduced number of aligned reads were detected between gene models 4g0020281 and 4g0020321in *Mtsym19*, and 4g0020281 and 4g0020301in *Mtsym20* compared with *Mtsym18* indicating a deletion in this genomic region (Figure 11E). Further analysis by PCR amplifications identified a less than 45-kb and a less than 31 kb

deletion in the *Mtsym19* and *Mtsym20* genomes, respectively. The deletion in *Mtsym20* spanned three genes encoding for two putative proteins and a Nodule-specific Cysteine-Rich (NCR) peptide.

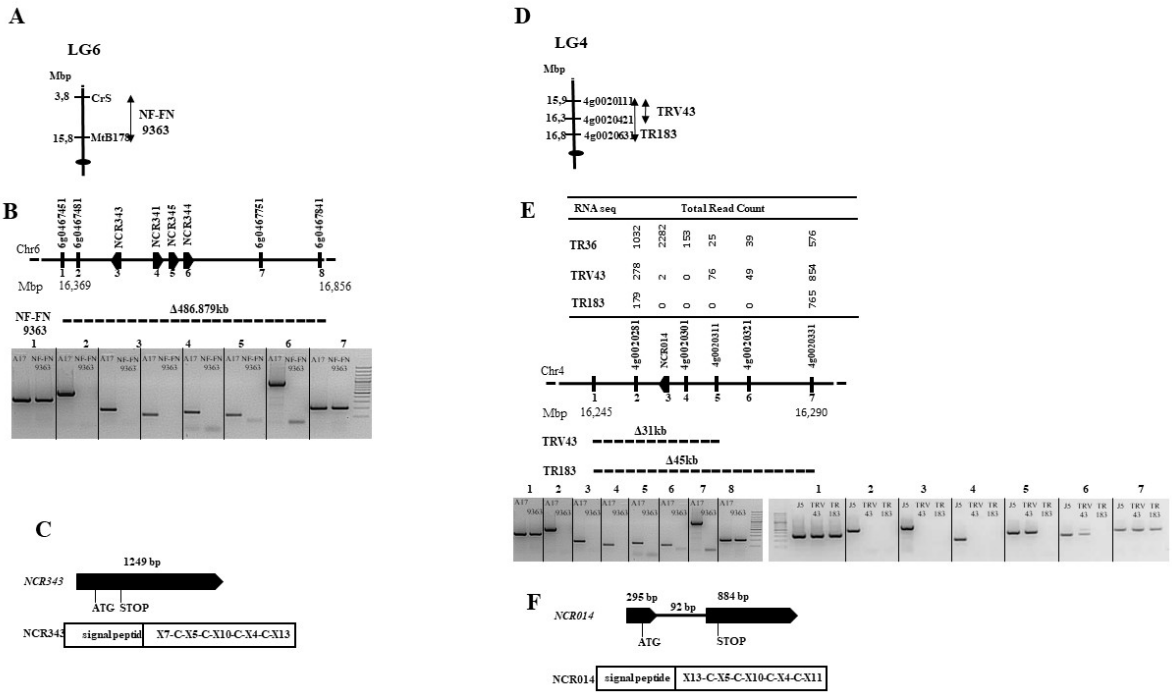


Figure 11. The genetic mapping of mutant loci of FN9363, *Mtsym19* and *Mtsym20* and identification of the deletions in the three symbiotic mutants.

In order to demonstrate which gene or genes were responsible for the symbiotic defect of FN9363, the genomic fragments of the *NCR341*, *NCR343*, *NCR344* and *NCR345* genes containing at least 1.5 kb native promoters and native 3' UTRs were introduced into FN9363 roots using *A. rhizogenes*-mediated hairy root transformation. The roots were inoculated with *S. medicae* WSM419 (pXLGD4) and nodules were stained for β-galactosidase activity at 4 wpi. Nodules cells in nitrogen fixation zone of mutant of FN9363 were devoid of rhizobia (**Fig. 10**), therefore we assessed the complementation of nodules based on the presence of rhizobia in ZIII. White undeveloped nodules with colonized cells by rhizobia in ZII and ZII-III were detected on roots transformed with empty vector or with the genes *NCR341*, *NCR344* and *NCR345* indicating that these *NCR* genes were not able to restore the effective symbiotic interaction in FN9363 (**Fig. 12**). The nodules on the roots of FN9363 transformed with the construct of *NCR343* were elongated and pink, and the bacterial invasion of ZIII in these nodules was similar to wild-type nodules, suggesting that these were functional nodules on FN9363 roots (**Fig. 12**). The rescue of the symbiotic phenotype of FN9363 with gene *NCR343* indicated that the loss of this gene caused the ineffective symbiotic phenotype of mutant FN9363.

The identified deletions in mutants *Mtsym19* and *Mtsym20* overlapped about 30 kb removing three genes including *NCR014*, which was the best candidate for *MtSYM19* and *MtSYM20*. In order to confirm that the deletion of *NCR014* caused the ineffective symbiotic phenotype of *Mtsym19* and *Mtsym20*, roots of the two mutants were transformed with the wild-type *NCR014* gene controlled by an 819 bp native promoter and inoculated with *S. medicae* WSM419 (pXLGD4). Transgenic *Mtsym19* and *Mtsym20* roots expressing the gene *NCR014* developed elongated and pink nodules in contrast to white round-shaped or slightly cylindrical nodules formed on empty vector-transformed roots of mutant plants (**Fig. 12**). In order to visualize bacterial invasion, longitudinal sections of 4-week-old nodules were stained for β-galactosidase activity. Rhizobia-colonized cells were observed in zones II and II-III of nodules formed on empty vector-transformed *Mtsym19* and *Mtsym20* roots (Fig. 12). Nodules developed on mutant roots transformed with the *NCR014* construct showed the typical zonation of indeterminate nodules with colonized cells in the nitrogen fixation zone like nodules on empty vector-transformed wild-type roots (**Fig. 12**). The restoration of nodule colonization in ZIII of mutant nodules with *NCR014* confirmed that this gene corresponds to *MtSYM19* and *MtSYM20*.

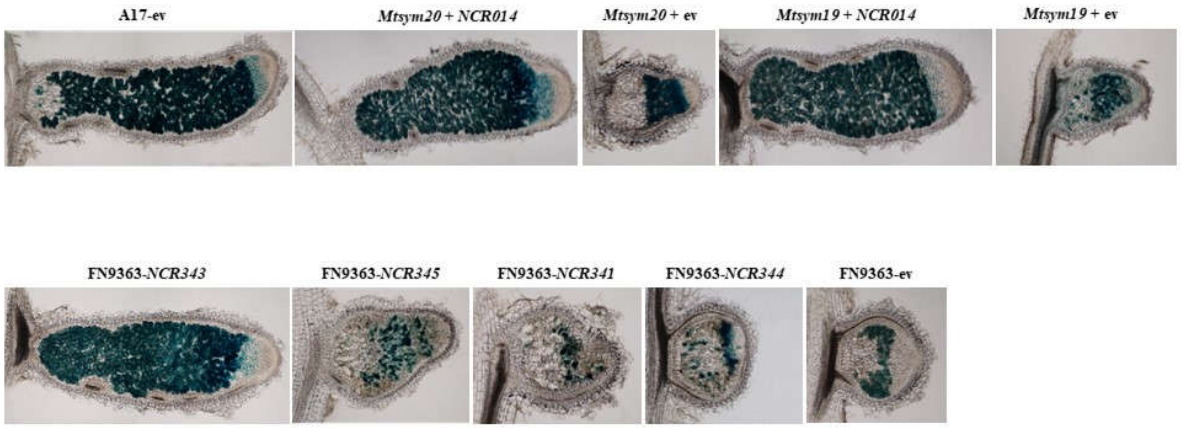


Figure 12. Gene *NCR14* restores the symbiotic phenotype of Mtsym19 and Mtsym20 (upper panel). NCR343 complement the symbiotic phenotype of FN9363. ev: empty vector

In order to analyze the function of *NCR* genes showing distinct expression pattern between *M. truncatula* Jemalong and R108 lines, we planned to apply a targeted genome editing approach of the CAS9/CRISPR system using the *M. truncatula* hairy root transformation system mediated by *A. rhizogenes*. In order to establish the CAS9/CRISPR gene editing method, we obtained the pKSE401 cloning vector which was used successfully for targeted genome editing previously. We cloned the sequence coding for the DsRed fluorescent protein into the pKSE401 vector to allow the detection of the transformed roots easily. We also exchanged the promoter controlling the expression of the guideRNA for *M. truncatula* U6.6 promoter We demonstrated previously that the *NCR169* gene is essential for effective nitrogen fixing symbiosis (Horvath et al. 2015). Therefore, we selected *NCR169* to test the genome editing capacity of the pKSE466_RR vector. Overlapping primer pairs were designed for the gene *NCR169* to generate the guide RNA molecules which have been ligated into the *Bsa*I-digested pKSE401 vector. Following confirmation of the accuracy of the generated construct, we tested this construct to generate mutant version of *NCR169*. We introduced he construct into the roots of wild-type *M. truncatula* plants using the *A. rhizogenes*-mediated transformation system. Transformed roots were detected by the DsRed fluorescent marker and the symbiotic phenotype of nodules developed on transformed roots were analyzed. Nodules displaying defects (small round-shaped and white nodules; **Figure 13**) were detached from the root and cut into halves. The half of the nodules were stained for the presence of rhizobia inside the symbiotic nodules to analyze colonization of the nodules and the other halves were used to amplify the targeted region of *NCR169*. We sequenced the amplified fragments and found that symbiotic nodules with colonization defects carried mutations in the *NCR169* gene indicating that system is suitable to generate gene-edited mutant version of *NCR* genes (**Figure 14**).

To test the requirement of additional *NCR* genes for the nitrogen-fixing symbiotic interaction using the targeted gene editing approach, several *NCR* peptide genes were selected based on the criteria of high expression in invasion zone, having an appropriate signal peptide and cysteine residues in conserved positions. Single-stranded guide RNAs (sgRNA) have been designed to target the selected *NCR* genes using the CRISPR/CAS9 genome editing system. Constructs have been generated, verified and transformed into the roots of wild- type *M. truncatula* 2HA plants with *A. rhizogenes*-mediated transformation.

The analysis of the symbiotic phenotype (colonization of rhizobia) of nodules formed on gene edited roots indicated that genes *NCR068*, *NCR089*, *NCR128*, *NCR139* and *NCR161* are not essential for the symbiotic interaction between *M. truncatula* line 2HA and *S. mediacae* strain WSM419 (**Figure 15** and **16**). The scoring of roots with targeted *NCR* genes using CRISPR/CAS9 genome editing system is in still progress.

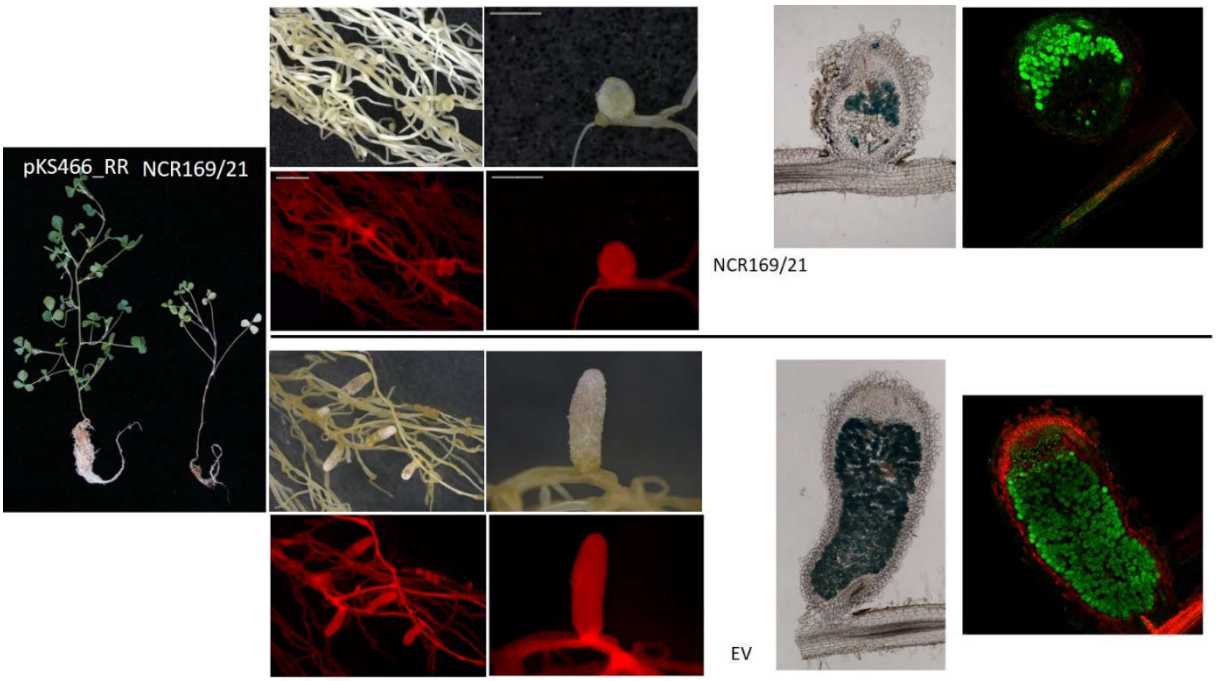


Figure 13. The symbiotic phenotype of wild-type *M. truncatula* Jemalong plants transformed with empty vector (pKS466_RR) or the construct targeting the *NCR169* gene. Functional pink nodules developed on EV-transformed roots and these nodules showed rhizobia colonized nodule zones characteristic for wild-type indeterminate nodules (lower panels). The nodules formed on roots of plant NCR169/21, which was transformed with the *NCR169*-targeting constructs, were similar to the *dnf7* mutant nodules defective in *NCR169* and showed the similar nodule invasion phenotype to the *dnf7* nodules (upper panels).

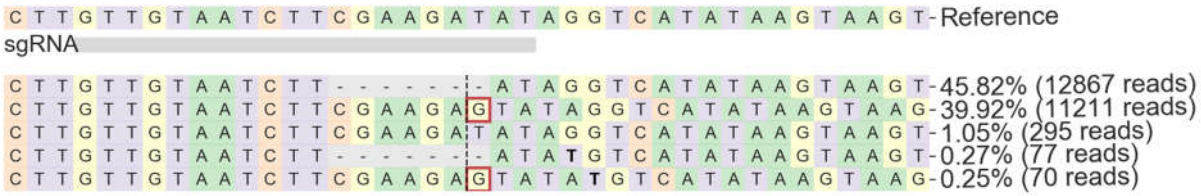
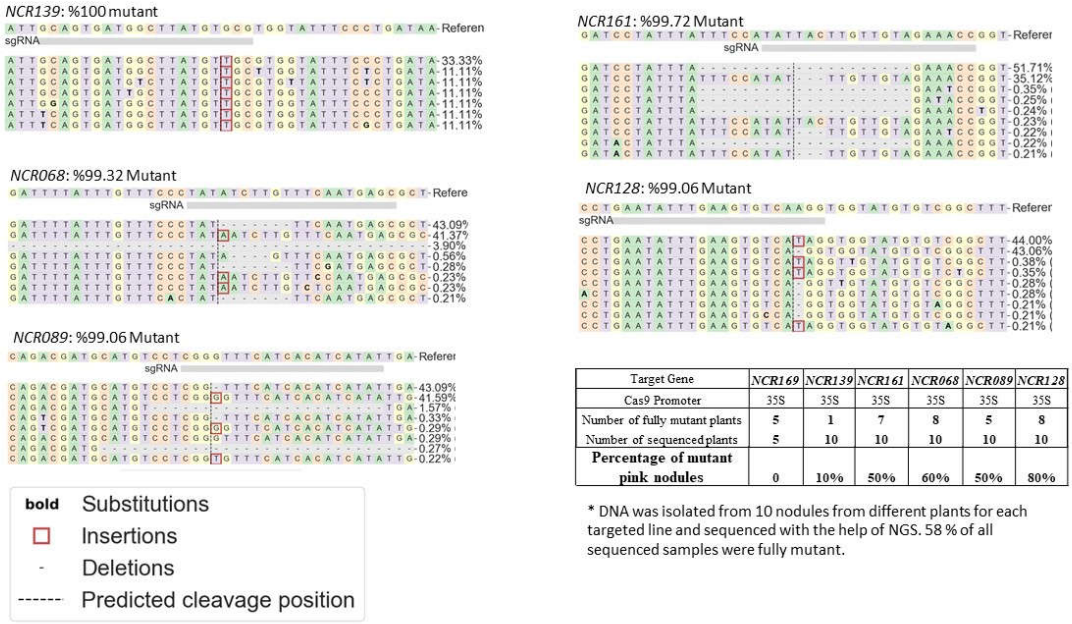


Figure 14. The sequence analysis of nodules on roots of NCR169/21 revealed the high rate (~ 98.6%) of mutant alleles of *NCR169* in mutant nodules.



Target Gene	*NCR169*	*NCR139*	*NCR161*	*NCR068*	*NCR089*	*NCR128*
Cas9 Promoter	35S	35S	35S	35S	35S	35S
Number of fully mutant plants	5	1	7	8	5	8
Number of sequenced plants	5	10	10	10	10	10
Percentage of mutant pink nodules	0	10%	50%	60%	50%	80%

* DNA was isolated from 10 nodules from different plants for each targeted line and sequenced with the help of NGS. 58 % of all sequenced samples were fully mutant.

Figure 15. The sequence analysis of nodules on gene edited roots targeting genes *NCR068*, *NCR089*, *NCR128*, *NCR139* and *NCR161* revealed the high rate of mutant alleles.

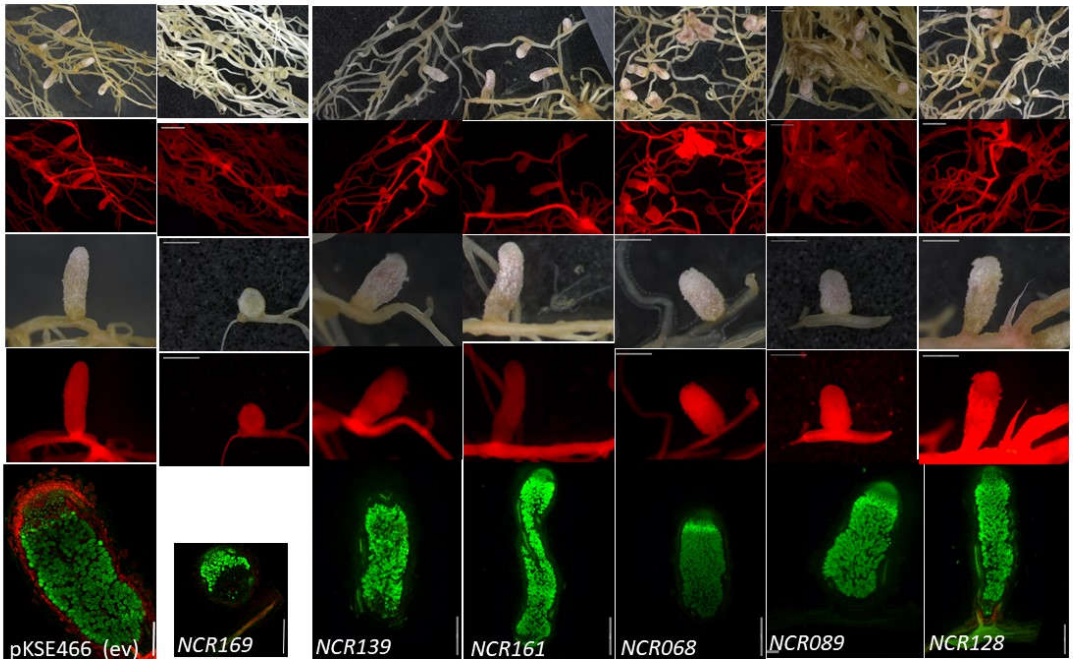


Figure 16. The symbiotic phenotype of wild-type *M. truncatula* Jemalong plants transformed with empty vector (pKS466_RR) or the construct targeting the genes *NCR169* (control), *NCR068*, *NCR089*, *NCR128*, *NCR139* and *NCR161*. Functional pink nodules developed on EV-transformed roots and on roots transformed with the *NCR068*, *NCR089*, *NCR128*, *NCR139* and *NCR161*-targeting constructs. These nodules showed rhizobia colonized nodule zones characteristic for wild-type indeterminate nodules (lower panels) indicating that these NCR peptides are not essential for the symbiotic interaction between *M. truncatula* line 2HA and *S. mediacae* WSM419.

The genetic analysis of the incompatibility of the M. truncatula A17 genotype with the S. meliloti strain 41 (Rm41) was carried out to identify the gene responsible for this non-functional symbiotic interaction. In a collaboration with Hongyan Zhu's group, two genes, *NFS1* and *NFS2* were identified coding for NCR peptides, respectively. We identified that the Jemalong alleles of genes *NFS1* and *NFS2* led to the elimination of the Rm41 bacteria, the alleles of other *M. truncatula* accessions (e.g. A20, DZA315) did not affect the persistence of the Rm41 strain (compatible interactions). This study, showing that NCR peptides can function as a negative regulator of the symbiotic interaction by specifying the incompatibility of the host plant with certain bacterial strains was published in the PNAS (Wang et al. 2017).

In collaboration with the researchers of the Institute for Integrative Biology of the Cell (Gif-sur-Yvette, France), we investigated bacteroid development in the nodules of *Aeschynomene* species that develop determinate nodules, produces NCR-like peptides and impose terminal bacteroid differentiation on their bacterial partners. Terminally differentiated bacteroids in the different *Aeschynomene* species are spherical (S morphotype) or elongated (E morphotype). We used a combination of *Aeschynomene* species inducing E or S-type bacteroids in symbiosis with *Bradyrhizobium* sp. ORS285 to show that S-type bacteroids present a better symbiotic efficiency than E-type bacteroids. We performed a transcriptomic analysis on E- and S-type bacteroids formed by *A. afraspera* and *A. indica* nodules and identified the bacterial functions activated in bacteroids and specific to each bacteroid type. Extending the expression analysis in E- and S-type bacteroids in other *Aeschynomene* species by qRT-PCR on selected genes from the transcriptome analysis narrowed down the set of bacteroid morphotype-specific genes. Functional analysis of a selected subset of 31 bacteroid-induced or morphotype-specific genes revealed no symbiotic phenotypes in the mutants. This highlights the robustness of the symbiotic program but could also indicate that the bacterial response to the plant environment is partially anticipatory or even maladaptive. The antimicrobial peptide transporter BclA is a major determinant of terminal bacteroid differentiation in *Bradyrhizobium* sp. strain ORS285. In the absence of BclA, the bacteria proceed until the intracellular infection of nodule cells, but they cannot differentiate into enlarged polyploid and functional bacteroids. Thus, the *bclA* nodule bacteria constitute an intermediate stage between the free-living soil bacteria and the nitrogen-fixing bacteroids. Metabolomics on whole nodules of *A. afraspera* and *A. indica* infected with the wild type or the *bclA* mutant revealed 47 metabolites that differentially accumulated concomitantly with bacteroid differentiation. Bacterial transcriptome analysis of these nodules demonstrated that the intracellular settling of the rhizobia in the symbiotic nodule cells is accompanied by a first transcriptome switch involving several hundred upregulated and downregulated genes and a second switch accompanying the bacteroid differentiation, involving fewer genes but ones that are expressed to extremely elevated levels. The transcriptomes further suggested a dynamic role for oxygen and redox regulation of gene expression during nodule formation and a nonsymbiotic function of BclA. Together, our data uncover the metabolic and gene expression changes that accompany the transition from intracellular bacteria into differentiated nitrogen-fixing bacteroids. *Bradyrhizobium diazoefficiens* USDA110 was isolated as a soybean symbiont, but it can also establish a functional symbiotic interaction with *A. afraspera*, which – unlike soybean – triggers terminal bacteroid differentiation. A combination of plant metabolomics, bacterial proteomics, and transcriptomics along with cytological analyses were used to study the physiology of USDA110 bacteroids in these two host plants. It was shown that USDA110 establishes a poorly efficient symbiosis with *A. afraspera* despite the full activation of the bacterial symbiotic program. Molecular signatures of high levels of stress in *A. afraspera* bacteroids were found, whereas those of terminal bacteroid differentiation were only partially activated. Finally, we showed that in *A. afraspera*, USDA110 bacteroids undergo atypical terminal differentiation hallmarked by the disconnection of the canonical features of this process.

Our additional projects that are not closely related to NCR peptides, but analyzed different aspects of symbiotic nitrogen fixation, resulted in publication of additional scientific papers. We contributed to the analysis of the suppression of NB-LRR resistance genes by miRNAs that facilitates nodule formation in *Medicago truncatula*. The paper describing this study was published in the Plant Cell and Environment (2020; 43:1117–1129). In another project, we demonstrated that Vacuolar iron Transporter-Like proteins, VTL4 and VTL8 deliver iron to symbiotic bacteria (Walton et al. 2020, New Phytologist 228: 651–666). We also contributed to the functional analysis of the *CYCLOPS* (*IPD3*) gene in the crack entry legume, peanut (Rajlakshmi Das et al. 2018, Plant Science). We also published another study showing that single amino acid alterations in the VHIID motif of the NSP2 transcription factor modulate the symbiotic process

in different ways and the responses of these point mutants is strain-dependent (Kovacs et al. 2021, Frontiers in Plant Science 12: 709857). The principal investigator of the associated project of the consortium (Attila Kereszt) published a paper on how gene regulation studies contributed to better understand the process of nodule organogenesis and function. Plant Cell 32: 42–68, 2020).

References

Horvath, B.; Domonkos, A.; Kereszt, A.; Szucs, A.; Abraham, E.; Ayaydin, F.; Boka, K.; Chen, Y.; Chen, R.; Murray, J.; Udvardi, M, Kondorosi E, Kaló, P (2015) Loss of the nodule-specific cysteine rich peptide, NCR169, abolishes symbiotic nitrogen fixation in the *Medicago truncatula dnf7* mutant. PNAS 2015, 112, 15232-15237

Kim,M, Chen, Y, Xi, J, Waters. C, Chen, R, and Wang, D (2015) An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. PNAS 2015, 112, 15232-15237

Nallu, S, Silverstein, KA, Samac, DA, Bucciarelli, B, Vance, CP, VandenBosch, KA (2013) Regulatory patterns of a large family of defensin-like genes expressed in nodules of Medicago truncatula. PLoS One. 8(4):e60355.

Wang, Q, Yang, SM, Liu, JG, Terecskei,, K, Abraham, E, Gombar, A, Domonkos, A, Szucs, A, Kormoczi, P, Wang, T, Fodor, L, Mao, LY, Fei, ZJ,Kondorosi, E, Kalo, P, Kereszt, A, Zhu, HY (2017) Host-secreted antimicrobial peptide enforces symbiotic selectivity in Medicago truncatula. Proc Natl Acad Sci USA 114: 6854-6859

Morandi, D, Prado, E, Sagan, M, Duc, G (2005) Characterisation of new symbiotic *Medicago truncatula* (Gaertn.) mutants, and phenotypic or genotypic complementary information on previously described mutants. Mycorrhiza 15(4):283-9.