Final Report of Grant KH-129547

During the nitrogen-fixing symbiotic interaction between *Medicago truncatula* or related IRLC (Inverted Repeat-Lacking Clade) legumes and rhizobia, the bacterial partner undergoes terminal differentiation, and this differentiated form of rhizobia is able to convert atmospheric nitrogen to ammonia. This terminal bacteroid differentiation is directed by hundreds of nodule specific cysteine-rich peptides (NCRs) produced by the host plant. The genome of *M. truncatula* contains more than 600 *NCR* genes and until recently, it was generally assumed that most of these peptides, if not all, act redundantly. The deletion of two *NCR* genes, *NCR169* and *NCR211* indicated that certain members of the *NCR* gene family can have unique function in the terminal differentiation of bacteroids (Horvath et al.2015, Kim et al. 2015). Some other NCRs are not required for bacterial differentiation but determine the compatibility of the symbiotic interaction between *M. truncatula* and rhizobia (Yang et al. 2017, Wang et al. 2017). Until now, the role of only few NCRs has been defined but the function of the members of the diversified and amplified family of NCR peptides remained to be unclear.

Our former RT-qPCR analysis revealed that some *NCR* genes showed differential expression in two, widely studied wild-type *M. truncatula* lines, *M. truncatula* subsp. *truncatula* (Jemalong) and *M. truncatula* subsp. *tricycla* (R108). We presumed that the analysis of the differentially expressed NCR genes in two wild-type *M. truncatula* lines could identify the function of some additional *NCR* genes. In this project, we aimed to identify the differentially activated NCR genes between the two *M. truncatula* lines using a high-throughput transcriptome analysis of nodules elicited by two effective rhizobial strains. In order to facilitate the study of NCR genes we also planned to apply the targeted mutagenesis of selected NCR genes and develop an efficient way to analyze the gene edited transformed roots of *M. truncatula*.

Identification of *NCR* peptide genes differentially expressed in the nodules of two wild-type *M. truncatula* lines

In order to gain more comprehensive data about the differential expression profile of genes in the *M. truncatula* wild-type Jemalong and R108 nodules, whole transcriptome sequencing (RNA-seq) was carried out on nodule samples. Most of the NCR genes are induced at 7 days post inoculation with rhizobia, therefore total RNA was isolated from three biological replicates at this time point. The RNA samples were sequenced using Illumina sequencing platform at the SEQOMICS Ltd. (Mórahalom, Hungary). Following preliminary data processing (trimming adaptors, quality check), the reads were mapped to the reference genome sequence of *M. truncatula*. The detailed expression analysis of 396 genes of NCR peptides that possess signal peptide and contained four or six cysteine residues in conserved positions was carried out. The „kallisto" és „sleuth" softwares were used with the parameters to identify *NCR* genes showing at least fivefold altered activation between the two wild-type plant lines. The analysis identified 26 upregulated and 36 downregulated NCR genes in the genotype R108 compared to Jemalong line (Figure 1).

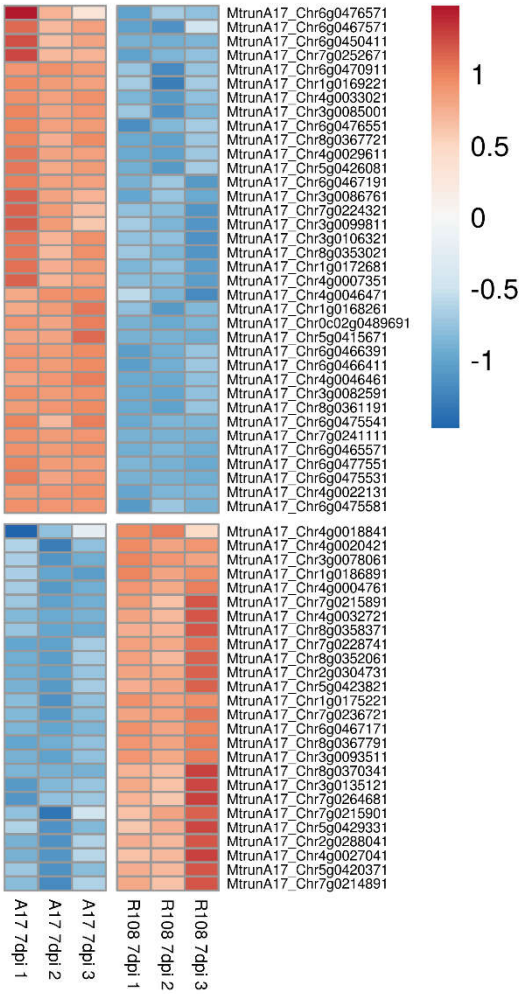

Figure 1. *NCR* genes differentially expressed in *M. truncatula* wild-type Jemalong and R108 nodules at 7 days post inoculation with rhizobia.

The functional analysis *NCR* genes differentially expressed in the nodules of two wild-type *M. truncatula* lines

The promoter regions of the *NCR* genes showing difference between the two wild-types *M. truncatula* plant were analyzed to reveal DNA elements that might be responsible for the altered transcription. No large deletions or other genomic rearrangements were identified in the analyzed promoter regions. The preliminary motif search using the MEME Suite (Motif-based sequence analysis) webtool identified a DNA element that might bind to GATA transcription factors. Some members of the GATA transcription factors (TFs) are nitrate-inducible and thus their involvement in the induction of the nodule-specific cysteine rich-peptide genes is conceivable. We analyzed the expression pattern of GATA TFs of the transcriptome database obtained by laser-capture microdissection of different *M. truncatula* nodule zones (Roux et al. 2014). We identified and selected the genes coding for GATA TFs that are active in the interzone and differentiation zone of *M. truncatula* nodules. These TF genes can potentially regulate *NCR* genes. In order to study the function of these GATA TF genes, mutant lines with *Tnt1* transposon insertions were searched for in the collection of *M. truncatula* mutants. Mutant lines probably carrying insertion in three *GATA TF* genes were found and we ordered seeds to generate homozygous mutant plants for further analysis.

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. We needed to solve few problems to apply the targeted gene editing for the functional study of *NCR* genes: (i) although few experiments using the CAS9/CRISPR system in *M. trucatula* have been published, we needed to establish this system using the *A. rhizogenes*-mediated transformation in our laboratory, and (ii) we had to develop how the gene edited nodules can be genotyped and analyzed for symbiotic phenotype, (iii) and test the possibility to regenerate mutant plants from *A. rhizogenes*-transformed roots. In order to establish the CAS9/CRISPR gene editing method using the M. truncatula hairy root transformation system mediated by A. rhizogenes, 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 (Figure 2).

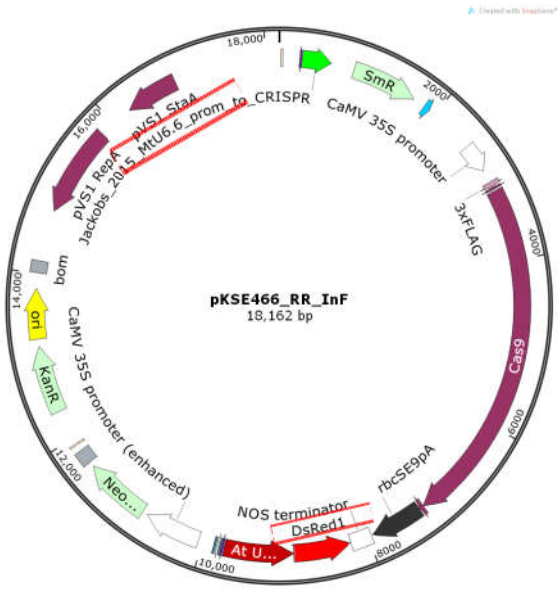

Figure 2. The map of the pKSE466_RR based on pKSE401. The prAtUBQ10::DsRed transformation marker gene was inserted into the vector backbone and the promoter upstream of the guide RNA was exchanged.

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) 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 3).

To test the function of NCR genes showing differential expression activity between M. truncatula Jemalong and R108, 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 plants with Agrobacterium rhizogenes-mediated transformation. The scoring of the transformed roots for the symbiotic phenotype and for the success of targeted gene editing are in progress.

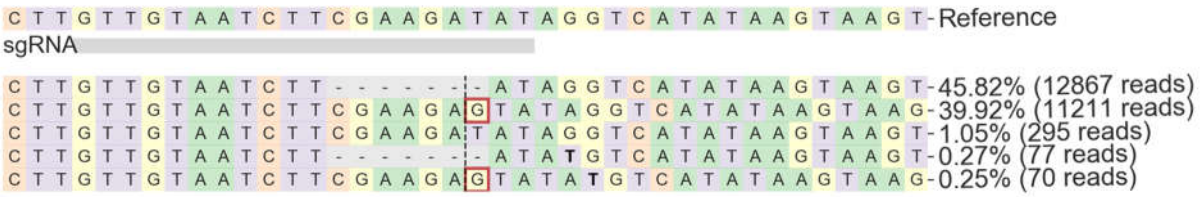

Figure 3. The sequence analysis of nodules on roots of NCR169/21 revealed the high rate (~ 98.6%) of mutant alleles of *NCR169* in mutant nodules.

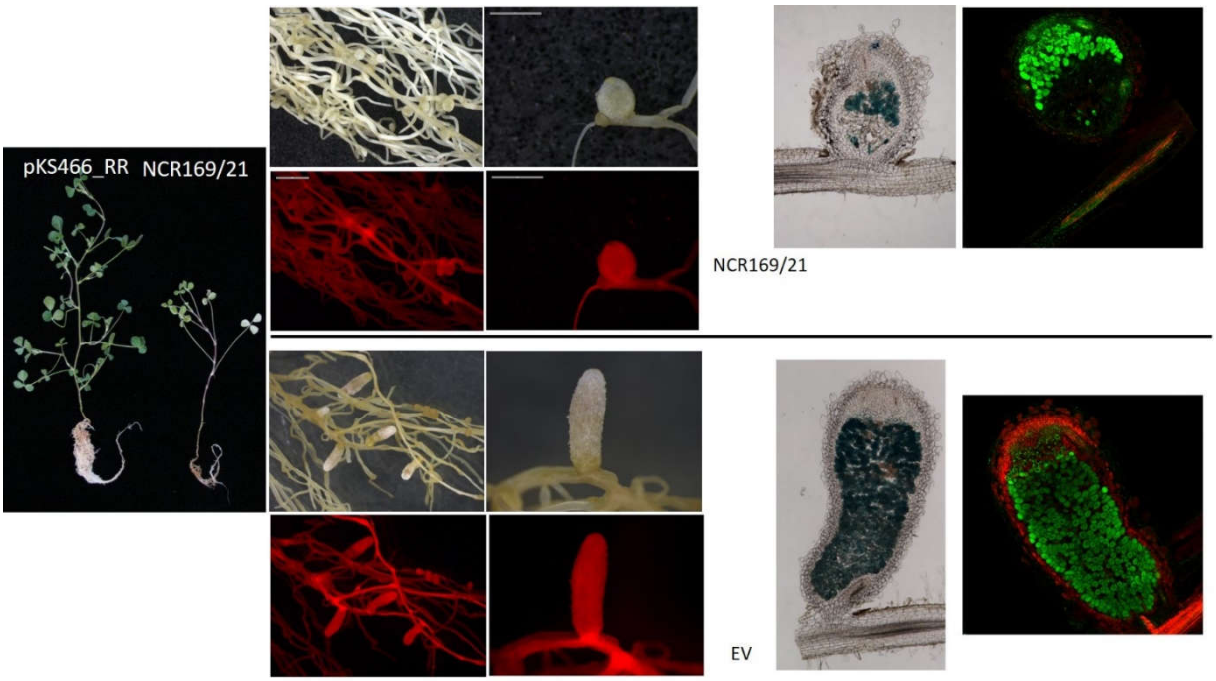

Figure 4. The symbiotic phenotype of wild-type *M. truncatula* Jemalong plants transformed with empty vector (pKS466_RR) or the construct targeting the *NCR169* gene. Functional pink nodules developed on EV-transformed roots and these nodules showed rhizobia colonized nodule zones characteristic for 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).

We identified a mutant FN9285 during a genetic screen for symbiotic mutants in a *M. truncatula* fast neutron-bombarded population. The mutant FN9285 showed the symptoms of nitrogen starvation under symbiotic conditions and mutant nodules were defective in colonization of interzones and nitrogen fixation zones (Figure 5).

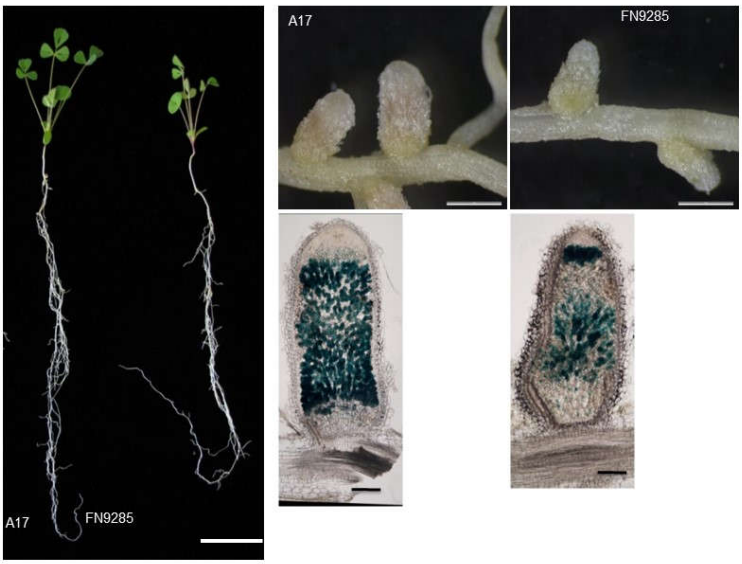

Figure 5. The symbiotic phenotype of wild-type (A17) and FN9285 symbiotic mutant *M. truncatula* plants 3 wpi with *S. medicae* WSM419 carrying the β-galactosidase marker gene (pXLGD4 plasmid). Wild-type plants developed functional pink nodules which showed the characteristic zonation of indeterminate nodules and were colonized with rhizobia. White and slightly elongated nodules showing nodulation zones were formed on FN9285 mutant roots. Nodule cells were colonized only in the infection zones in mutant nodules.

We identified the position of the symbiotic mutant locus of FN9285 on the upper arm of chromosome 3 between genetic markers at a distance about 4 Mbp. Parallel to genetic mapping, we carried out the total transcriptome analysis of mutant nodules to identify the deletion in the mutant locus. The mapping of the reads of mutant nodules indicated a large deletion in the genome of FN9285 corresponding to the defined symbiotic locus. PCR-based markers were used to confirm the deletions of FN9285 mutants spanning 250 kbp. The analysis of the gene content in the deleted region identified thirty genes, predicting nineteen genes as hypothetical protein genes. The analysis of the expression profile of the genes in deleted region of FN9285 in the *M. truncatula* Gene Expression Atlas revealed several nodule-specific genes and among them, we discovered nine *NCR* (*Nodule-specific Cysteine-Rich*) peptide genes (A1, A2, B-H). We cloned these *NCR* peptide genes but gene E which did not show any expression according to our RT-qPCR analysis. We introduced the constructs coding for each *NCR* peptide genes into the roots of FN9285 plants using the *A. rhizogenes*-mediated transformation system. Transformed plants were scored for nodulation phenotype 4 weeks post inoculation with *S. meliloti* strain 2011. Only the *NCRA1* and *NCRA2* were able to restore the effective symbiotic interaction of FN9285 mutant plants indicating that the deletion of these *NCR* peptide genes were responsible for the symbiotic defect of FN9285 (Figure 6.).

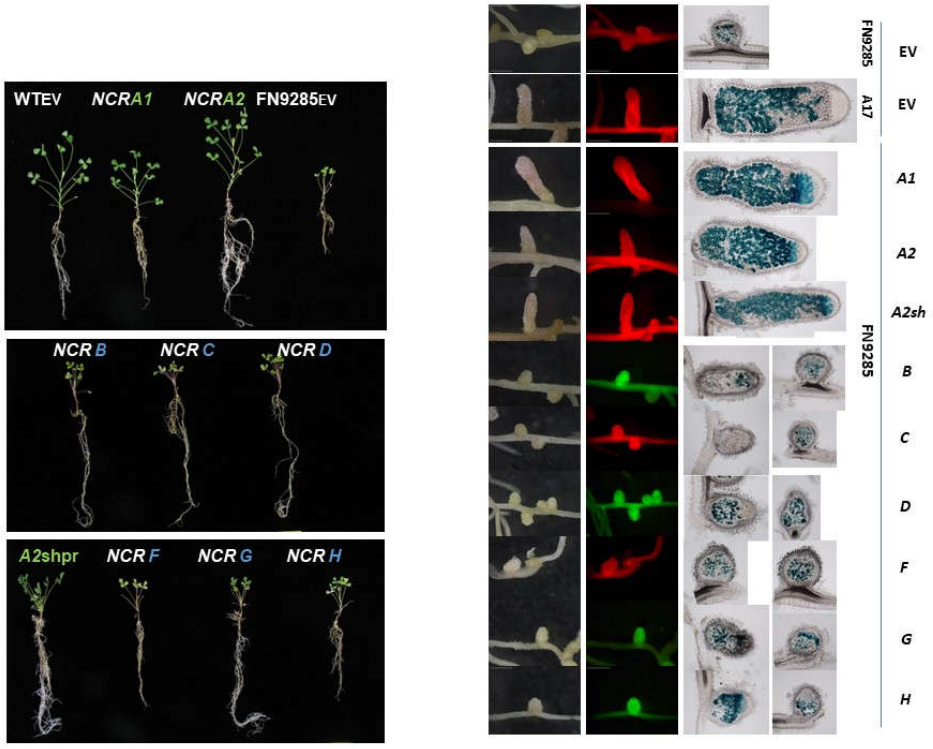

Figure 6. Genes *NCRA1* and *NCRA2* restore the symbiotic phenotype of FN9285.

The sequence analysis of the *NCRA1* and *NCRA2* genes revealed that encoded mature peptides are identical and their signal peptides share high similarity (90% identity). However, although both genes were able to restore the symbiotic phenotype of NF9285, we could detect significant difference in the activity of *NCRA1* and *NCRA2* genes. The promoter of *NCRA2* is expressed in nodules about 200-fold higher level compared with the promoter of *NCRA1*.

We analyzed and compared the cluster of *NCR* genes between *M. truncatula* Jemalong and R108 lines and found that only a single gene $NCRA^{R108}$ corresponding to the Jemalong genes *NCRA1* and *NCRA2* is present in R108 and the *NCRC* is absent in the R108 genome (Figure 7). We have analyzed the promoter sequences of *M. truncatula* Jemalong *NCRA1* and *NCRA2* genes and detected the absence of a potential element of aGATAttt in the promoter of *NCRA1*. This element is a potential binding site of Myb-related putative transcription factors. In order to determine the relevance of this element in the altered expression level of genes *NCRA1* and *NCRA2*, we started to generate mutant derivatives of potential *NCRA2* and analyze the expression pattern of Myb-like transcription factors to identify genes potentially regulating the activity of *NCR* genes.

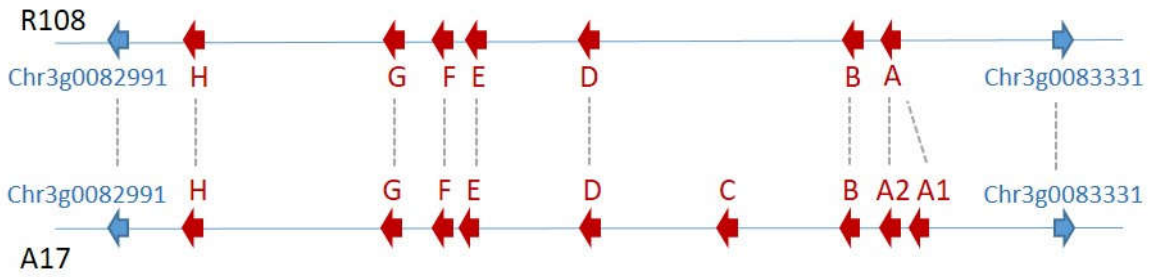

Figure 7. The synteny of the *NCR* peptide gene cluster between *M. truncatula* Jemalong and R108 lines. The Jemalong line contains nine (A1-H) NCR genes including a duplication of *NCRA* genes. In contrast, the genome of R108 contains a single *NCRA* gene and NCRC is absent in R108.

The preparation of the following manuscripts is in progress:

1. The manuscript describing the construction of the vector used for targeted editing of NCR peptide genes and the development of the system to regenerate mutant plants from roots transformed with *A. rhizogenes*. In the case we will be able to detect different requirement of *NCR* peptide genes between R108 and Jemalong, the manuscript will be expanded with these results.

2. The manuscript presenting the requirement of *NCRA1* and *NCRA2* genes and the differential gene content of the *NCR* peptide gene clusters between *M. truncatula* lines Jemalong and R108 in the symbiotic locus of mutant FN9285.

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 both coding for NCR peptides. We identified that the peptides encoded by the Jemalong alleles of the *NFS1* and *NFS2* genes resulted in 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 demonstrated that certain NCR peptides can function as a negative regulator of the symbiotic interaction by controlling the incompatibility of the host plant with certain bacterial strains (Wang et al. 2017).

Our additional projects that are not closely related to NCR peptides, but analyzed different aspects of symbiotic nitrogen fixation, resulted in publication of three 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 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).

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

Kovács, S, Fodor, L, Domonkos, Á, Ayaydin, F, Laczi, K, Rákhely, G and Kaló P (2021) Amino Acid Polymorphisms in the VHIID Conserved Motif of Nodulation Signaling Pathways 2 Distinctly Modulate Symbiotic Signaling and Nodule Morphogenesis in *Medicago truncatula.* Front. Plant Sci. 2021 12:709857. doi: 10.3389/fpls.2021.709857

Roux, B.; Rodde, N.; Jardinaud, M.F.; Timmers, T.; Sauviac, L.; Cottret, L.; Carrere, S.; Sallet, E.; Courcelle, E.; Moreau, S.; et al. (2014) An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. Plant Journal 77: 817-837.

Sós-Hegedűs, A, Domonkos, Á, Tóth, T, Gyula, P, Kaló, P*, Szittya, G* (2020) Suppression of *NB-LRR* Genes by miRNAs Promotes Nitrogen-fixing Nodule Development in *Medicago truncatula***.** Plant Cell and Environment 2020 43(5): 1117-1129

Yang SM, et al. (2017) Microsymbiont discrimination mediated by a host-secreted peptide in Medicago truncatula. Proc Natl Acad Sci USA 114: 6848-6853

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

Walton, JH, Kontra-Kovats, G, Green, RT, Domonkos, A, Horvath, B, Brear, EM, Franceschetti, M, Kaló, P, Balk, J (2020) The *Medicago truncatula* Vacuolar iron Transporter-Like proteins VTL4 and VTL8 deliver iron to symbiotic bacteria at differentstages of the infection process. New Phytologist 228: 651–666.

**corresponding authors*