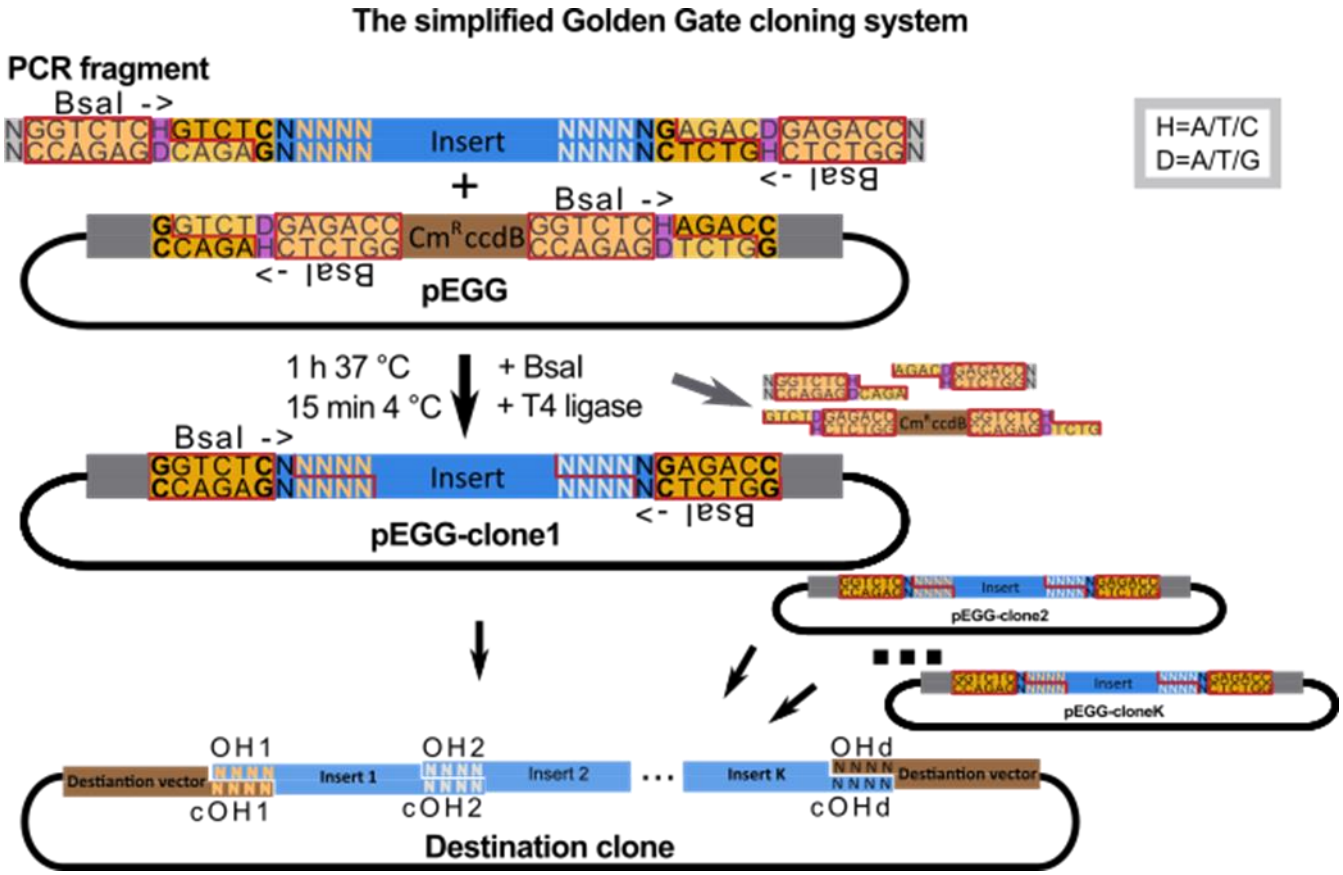
Objective 1: Development and adaptation of synthetic biology tools for rhizobia to manipulate their gene expression and the activity of their proteins:

A simplified GoldenGate cloning system to assemble multiple fragments

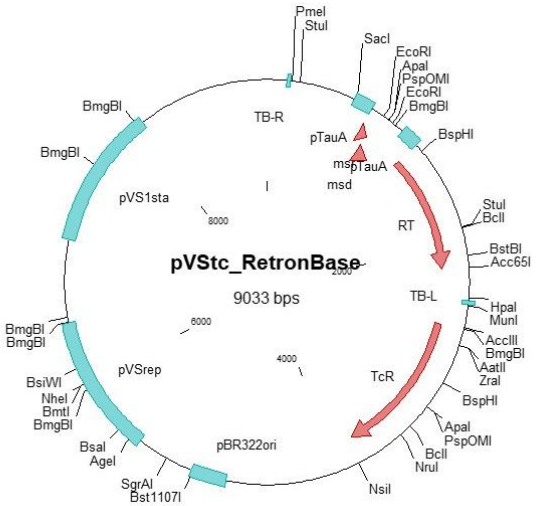
To simplify and to make more efficient our cloning efforts, we developed a simplified GoldenGate cloning system to assemble multiple fragments. The Golden Gate cloning and derived techniques to assemble multiple fragments have drawbacks such as the need for multiple entry vectors or to create the entry clones from PCR fragments employing completely different methods like TA/TOPO cloning or using the classical type IIP restriction endonucleases or type IIS enzymes other than used for the assembly. To simplify and make the procedure cost-effective, an approach using a universal entry vector, a novel primer design and a single type IIS restriction endonuclease producing four nucleotide overhang (Eco31I/BsaI as an example) in conjunction with T4 ligase was developed. The cloning of the PCR fragments amplified with primers of our unique design, the newly ligated insert can be released with any kind of desired sticky ends from the possible 240 (4^4-4^2) non-palindromic four nucleotide long overhang sequences, allowing the seamless cloning of one or more fragments into any compatible destination vector. In classical single-tube digestion-ligation GoldenGate reactions for cloning a fragment into an entry or destination vector, only the desired ligation products, the entry or destination clones lacking the restriction enzyme recognition site are stable, while all other products are re-digested with the enzyme. In the approach we devised, however, two recognition sites for the type IIS enzyme are formed in the resulting entry clones that can be cleaved by the enzyme making the ligation product unstable. For maximizing the ratio of circularized entry clones before transformation, it would be possible to heat inactivate both enzymes and re-start the ligation by adding ligase and ATP to the reaction mix. However, to make the whole process simpler, we showed that the reaction kinetics could be effectively shifted towards ligation only by lowering the reaction temperature to 4 °C after the digestion phase at 37 °C.

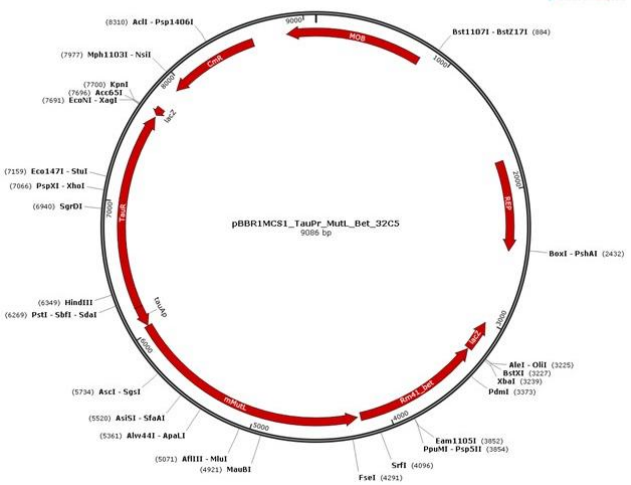


*Adaptation of the pORTMAGE, **port**able **m**ultiplex **a**utomated **g**enome **e**ngineering (MAGE) method for rhizobia.*

To modify the original pORTMAGE set-up, the E. coli and λ phage genes were replaced with *Sinorhizobium* genes that are supposed to be more efficiently expressed and translated in the α-*Proteobacteria* than their *E. coli* counterparts. For this purpose, we cloned the mutant gene of the *Sinorhizobium meliloti* phage 16-3ts coding for a thermosensitive cI repressor and the promoter that is regulated by this repressor. Similarly, a mutant version of *S. meliloti* MutL has been constructed and cloned. We also searched for rhizobial single-stranded DNA-annealing proteins to replace the Red recombinase of phage lambda used in the original pORTMAGE protocol that is not so efficient outside of Enterobacteria. The genes coding for these rhizobial proteins and the mutated MutL protein were cloned behind inducible promoters.

It turned out that transformation efficiency in rhizobia was very low even when oligonucleotides were introduced with the help of electroporation. That is why, a new approach was developed that is based on retron technology: 80-120 bp long double-stranded oligonucleotides are cloned into a reverse transcriptase target gene and both the reverse transcriptase coding gene and its target gene are under the regulation of a chemically inducible promoter. The same promoter regulates the mutant *mutL* gene and the gene coding for the single-stranded DNA binding protein. When this system is induced in a strain deficient in nucleases acting on single-stranded DNA, the reverse transcriptase creates single-stranded oligonucleotides from the target RNA that will facilitate the formation of mutations as the transformed oligonucleotides in the pORTMAGE method. After confirmation of the mutations, they can be moved into the fully wild-type or any mutant strain by phage transduction. For this purpose, we created a binary system: One of the vectors that can be mobilized into rhizobia and can replicate therein contains the codon optimized gene coding for the reverse transcriptase and the gene responsible for the production of its RNA partner and into which the mutant oligonucleotides can be cloned, both under the control of a taurine inducible promoter. The second plasmid, which can also be mobilized into rhizobia and can replicate therein and is compatible with the first vector, harbours a taurine-regulated operon with the gene coding for a thermo-sensitive mutant version of the MutL protein to suppress mismatch repair and another gene coding for the single-stranded DNA binding Bet protein of *S. meliloti*.





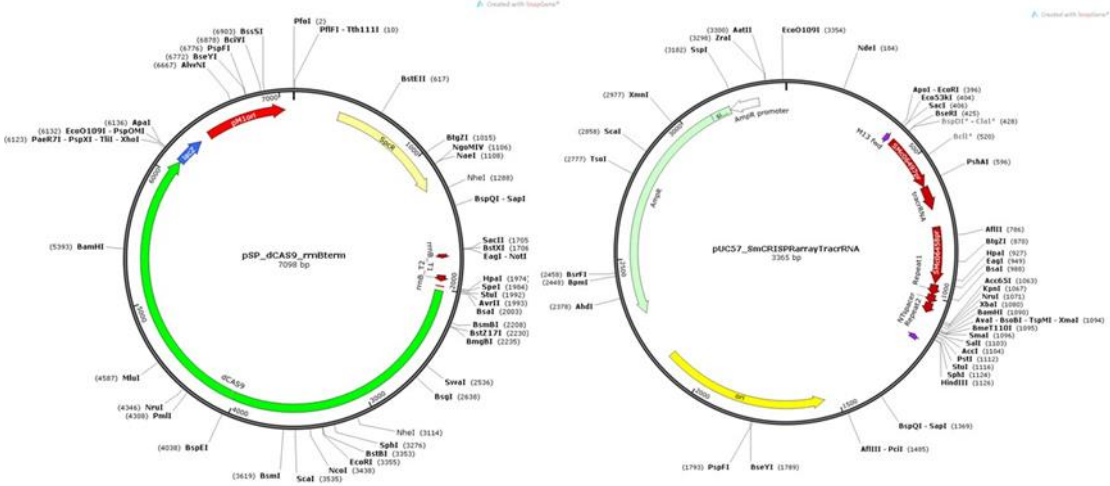
In addition, we created single and double mutant strains with mutations in the genes coding for the RecJ and XseA exonucleases to test whether elimination of either or both activities increases the efficiency of the mutagenesis. As a test system we are targeting the E41* non-sense mutation of the exoB gene in the strain AK631 of compact colony morphology because the reversion of the mutation results in an easily recognizable phenotype, the colonies produces high amount of exopolysaccharides, thus become mucoid and fluorescent on Calcofluor containing media.

Development and adaptation of regulated transcriptional activation and repression methods for rhizobia.

The purpose of this task is to make it possible to investigate whether the role of those *S. meliloti* genes, which are required for Nod Factor production or infection thread initiation/growth (for example, *nodC*, *exoY*, *phaA*, *bacA*, …), is restricted to the early phases or that are also required for the later phases of symbiotic nodule development and function.

First, we constructed a new broad host range Tn7 transposon-based vector with a multiple cloning site and providing the possibility of blue/white selection and gentamycin resistance that allow its selection in different wild-type and mutant Sinorhizobium strains. The promoterless *dCAS9* gene along with the tracrRNA and the CRISPathBrick were cloned into the new vector. To quantify the effectiveness of the exoY gene silencing, a gene coding for the ExoY protein tagged with the hemagglutinin (HA) epitope was constructed and introduced into strain Rm41 carrying a deletion that removed the whole *exoY* gene. As a control, we selected the *SMa0130* gene, which has approximately the same expression level as the *exoY* gene and codes for a protein that has a larger molecular weight than ExoY. We constructed a version that codes for the SMa0130 protein tagged with the HA epitope. This construct was introduced into strain Rm41 and into the strain carrying the gene coding for the ExoY-HA protein.

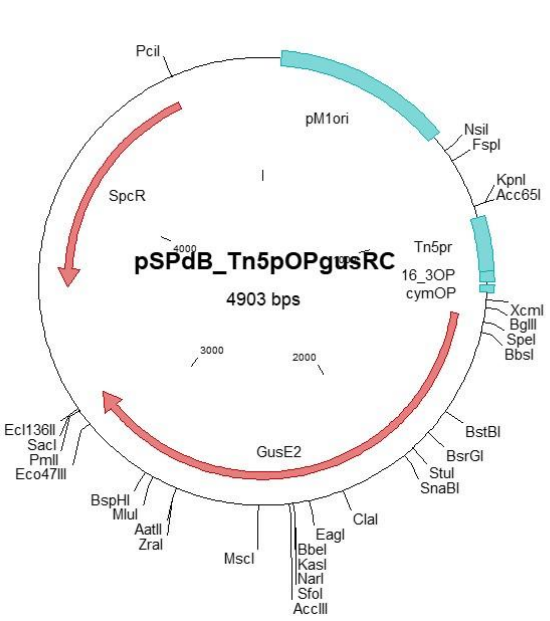
Later, we identified two problems: 1) The transposition of Tn7 in *S. meliloti* happened preferentially at the *nodM* gene instead of the canonical *glmS* gene, thus, disrupting the *nodM-nolGHIJ-nodN* operon and inhibiting nodulation. 2) The cloning of the spacer-repeat blocks into the CRISPathBrick vector proved to be very difficult and the silencing (most probable because of the foreign promoters in front of the RNA coding genes) was not efficient. That is why, we created a split CRISPathBrick system by separating the dCAS9 coding gene from the precursor CRISPR RNA (pre-crRNA) and transactivating crRNA (tracrRNA) genes and by cloning them into compatible broad-host-range vectors.

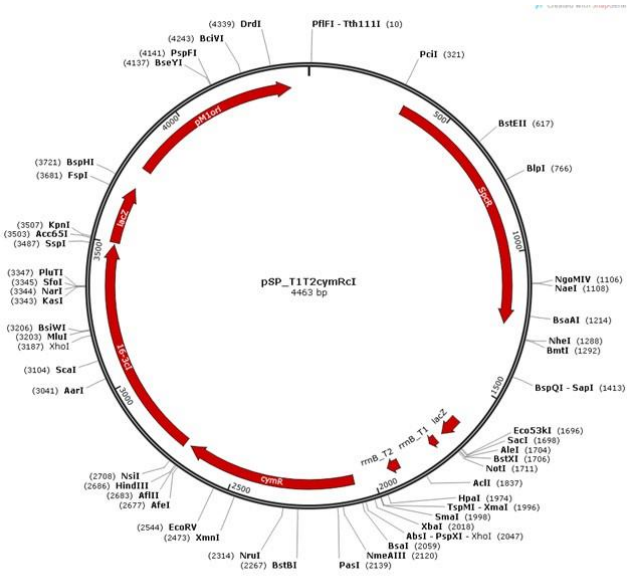


In front of the dCAS9 coding gene, we placed a multiple cloning site where a promoter of choice can be cloned, while the pre-crRNA and tracrRNA genes were placed under the regulation of S. meliloti constitutive promoters driving the expression of non-coding RNAs.

To investigate whether the *nodA* and *exoY* genes have a role in the later phases of symbiotic nodule development and functioning, three spacer-repeat blocks targeting known promoter elements for both genes were cloned into the RNA expressing vector.

Besides the RNA-based silencing system, we created a repressor protein based binary depletion system to inhibit the activity of genes at will. For this purpose, we constructed a new synthetic promoter that provides strong expression but can be repressed by the cI repressor of rhizobium phage 16-3 and/or the CymR repressor protein. The coding sequence of the *uidA* (β-glucuronidase) gene and a number of *S. meliloti* genes were cloned behind this synthetic promoter in a broad-host-range vector and at the same time, the corresponding rhizobial genes were deleted from the *S. meliloti* strain Rm41 genome. To provide the other half of the binary system, the two repressor coding genes were cloned as an operon in a compatible broad-host-range vector with a multiple cloning site preceding the operon to accept different promoters.





A number of promoters to drive the expression of the dCAS gene and the operon coding for the repressor proteins were selected and cloned into the corresponding vectors:

GENE		ENCODED PROTEIN	Total reads LCM (Deseq-normalized)	PERCENTAGE OF THE READS IN THE ZONE					PROMOTER IS ACTIVATED IN
				Zone I	Zone IId	Zone IIp	IZ	Zone III	
SMc02361	cycH	Cytochrome C-type biogenesis	1873,80	10,31	14,20	26,28	31,16	18,05	free-living and symbiotic conditions
SMb20605		ABC transporter	2435,81	1,43	43,60	31,47	9,54	13,96	Zone IId
SMb20425		Putative transcriptional regulator	598,45	0,00	9,47	45,57	31,05	13,91	Zone II(d)p
SMc00819	katA	Catalase	10941,62	2,13	0,89	16,62	61,29	19,07	Zone IIp
SMa0822	fixA	FixA electron transfer flavoprotein beta chain	38185,38	0,80	0,19	0,06	53,36	45,59	IZ
SMc01226		Transcriptional regulator	52751,17	1,04	0,51	0,47	2,43	95,55	Zone III

The *tauA* promoter is regulated by the chemical taurine, the cycH promoter is expressed in both free-living and symbiotic conditions, while the other promoters are induced consecutively in the different nodule zones (*SMb20605*: Zone IId; *SMb20425*: Zone II(d)p; *katA*: Zone IIp; *fixA*: Interzone; *SMc01226*: Zone III). The expression pattern of the selected promoters was confirmed by cloning them in front of the *uidA* gene and testing the reporter gene's activity in root nodules.

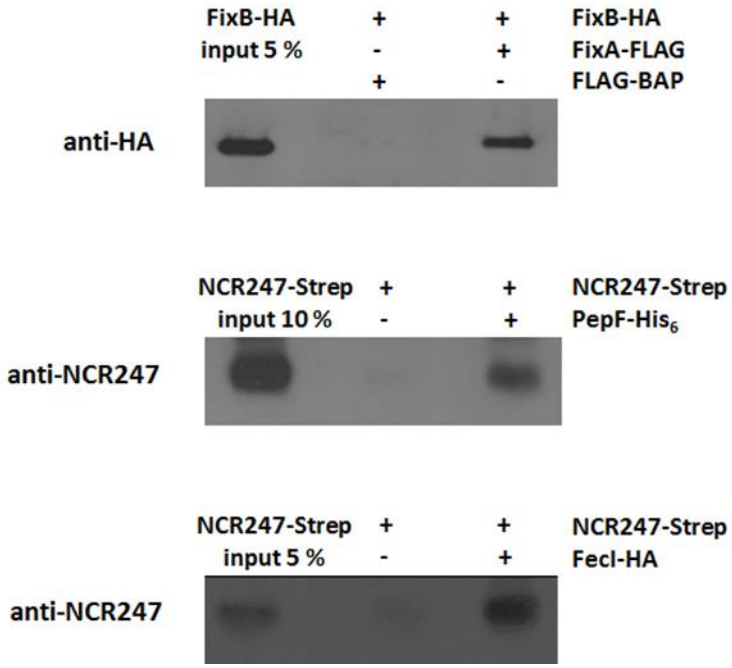
Construction of two-hybrid libraries to investigate interactions of rhizobial proteins.

Individual proteins often do not act alone, but are part of multiprotein complexes. Their interactions can be important in signal transduction, metabolic processes, transport of molecules/ions/electrons, etc., and therefore understanding their interactions with other proteins might also facilitate their functional annotation. Moreover, in nodule cells, symbiosomes and rhizobia therein are targeted by peptides, proteins of plant origin affecting the activity and/or interactions of rhizobial proteins. To allow the study of these interactions a yeast and a bacterial two-hybrid library were constructed from an ORFeome library containing all predicted proteins of *S. meliloti* strain 1021 and cloned into a mobilizable Gateway entry vector. The inserts from this library were transferred *in vitro* into pGADT7GW, a modified, Gateway recombination compatible AD vector. The clones were grouped for the Gateway reactions into 67 sets, of which all but three contained 96 clones. To test the success of the in vitro recombination, five pools of the destination clone library were sequenced on an Illumina MiSeq desktop sequencer. The presence of at least 90 ORFs in each pool could be detected after aligning the sequence reads to the reference genome indicating that the *en masse* recombination and transformation for creation of the library were effective.

Because of the high costs of the in vitro recombination reactions, we decided to use an *in vivo* recombination technique described for the ORFeome library to create the bacterial two-hybrid library. For this purpose, we constructed a mobilizable Gateway compatible destination derivative of the pKT25 library vector for the bacterial two-hybrid screens that can be transferred between E. coli strains by T7 phages. This last feature became necessary because the entry clones are also mobilizable between the strains and have the same antibiotic resistance marker as the original pKT25 plasmid. To create the library, the *in vivo* recombination reactions with the 67 pools were performed accorrding to the original protocol but then, the pKT25-based products were "purified" from the entry clones by transferring them via T7 phage transduction into another E. coli strain. Sequencing of the pools revealed the successful construction of the library. The library then can be transferred into the specific bait vector containing strain used for the screening via biparental mating.

To test the functionality of the libraries, pairwise investigations and mini screens with the sequenced pools were performed with proteins known to form dimers or complexes, then, screens with the NCR247 plant peptide known to be targeted to the bacteroids were performed. In order to evaluate our system we verified some selected protein-protein interactions detected in Y2H by classical biochemical methods (pull-down assay and immunoprecipitations):

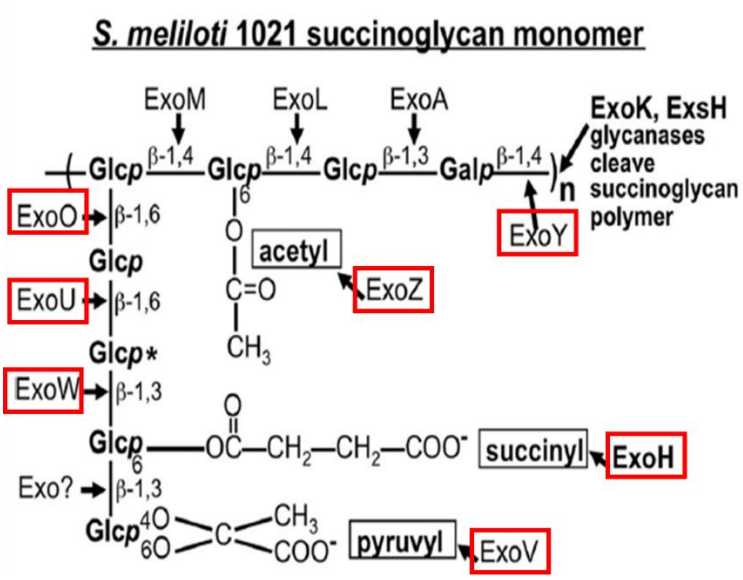
Weak physical interaction between FixB and FixA was verified by immunoprecipitation: Beads covered with FixA-FLAG or FLAG-BAP (negative control) were incubated with HA-tagged FixB, then, bound proteins were eluted by boiling followed by SDS-PAGE and immunoblotting with polyclonal HA antibody that was used to detect the specific in vitro interaction between FixA-FLAG and FixB-HA.



The physical interaction between the NCR247 peptide and oligoendopeptidase F as well as RNA polymerase sigma factor FecI observed in Y2H full screen was verified by pull-down assay. The tagged proteins were immobilized on antigen-specific magnetic beads and were incubated with Strep-tagged synthetic NCR247 peptides. Washed beads were boiled, SDS-PAGE and immunodetection using anti-NCR247 polyclonal antibody was used to demonstrate the specific *in vitro* binding of the peptide to both the strong interactor PepF and the weak interactor FecI.

Objective 2. **Investigation of the effect of surface polysaccharide mutants on the interaction with the host-plant.**

We created a number of marker-free in frame deletion mutants of *S. meliloti* strain Rm41 that produce no EPS (*exoY*) or EPSs with different structural defects such as devoid of branching (*exoO*), succinyl (*exoH*), acetyl (*exoZ*), pyruvyl (*exoV*) modifications or shorter side chains (*exoU*, *exoW*).



All these mutants established effective nitrogen-fixing symbiosis with *Medicago sativa* as well as with *M. truncatula* ecotypes DZA315 and A20 but formed empty, ineffective pseudonodules on the roots of ecotypes R108, F83005, GRC 033-B2, DZA045 and F 11008-C. In addition, the phenotype of the R108 *epr1* mutant was similar with all mutants to the one of the wild-type plant indicating that the function of the EPR1 receptor in *Medicago* is different from its role in *Lotus* (or determinate nodule forming) plants.

As stated above, we could identify a number of ecotypes with opposing phenotypes in association with a rhizobial mutant deficient in normal EPS production. To investigate the generality of these phenotypes, we investigated whether the *exoY* mutants (EPS-deficient derivatives) of other *S. meliloti* strains have the same phenotype as the Rm41 *exo* mutant strains. It turned out that neither the EPS-deficient reference strain 1021 nor the exoY mutant of the highly effective FSM-MA strain can form functional nodules on any of the plants. Earlier. we showed that the presence of a specific K-antigen (KPS) on the surface of strain Rm41 is responsible for the successful invasion of the *M. sativa* nodules. A strain Rm41 derivative deficient in both EPS and KPS production formed empty nodules on all plants tested indicating that in ecotypes DZA315 and A20, there is a genetic determinant that recognizes the presence of the specific KPS and allows the infection by the EPS-deficient strain. To identify this genetic determinant, we crossed ecotypes DZA315 and F 11008-C, then by self-pollinating the hybrid F1 plants an F2 mapping population was established. After phenotyping the individuals of the F2 population, the genetic mapping of the trait is in progress.

The ratio of those root hairs in which infection threads are formed compared to those that respond to Nod Factors is so low (3-10 root hairs per root) that expression changes in these root hairs are masked. That is why, alternative approaches to increase this ratio or to sequence only the RNA populations of the infected root hairs were looked for. We tried to introgress the *sunn* mutation causing a supernodulating phenotype into ecotype F83005 by crossing it to the sunn mutant of Jemalong but the F1 hybrids were almost completely sterile and formed only a very few seeds, thus, we abandoned this approach. As the costs of the new single-cell technologies are high above the budget of an OTKA grant, we decided to employ the Translating Ribosome Affinity Purification Followed by RNA Sequencing Technology (TRAP-SEQ). For this purpose, a gene coding for the FLAG-tagged RPL18 ribosomal protein under the control of the *M. truncatula NS1* promoter (active in the infected cells) were created. The first attempt to use this technology failed because the quantity and quality of the library were not sufficient. Currently, higher number of plants are being transformed to obtain more root hairs from the transformed hairy roots.