AGRICULTURAL ACADEMY

AGROBIOINSTITUTE – SOFIA

Assoc. Prof. Dr. ANELIA VENEVA IANTCHEVA

The role of model species *Medicago truncatula, Lotus japonicus* and *Arabidopsis thaliana* in plant biotechnology and functional genomics studies of legumes

ABSTRACT

OF THE DISSERTATION WORK TO OBTAIN THE SCIENTIFIC DEGREE

"DOCTOR OF SCIENCE" IN PROFESSIONAL FIELD 4.3 BIOLOGICAL SCIENCES

Sofia

2019

The most commonly used abbreviations

DSE- direct somatic embryogenesis

ISE-indirect somatic embryogenesis

2.4-D - 2.4-Dichlorophenoxyacetic acid

IAA - indolylacetic acid Cb - carbenicillin

NAA - α-naphthylacetic acid

TDZ-1-phenyl-3- (1,2,3-thiadiazol-5-yl) urea, thidiazuron-TDZ thiadiazuron

PAT - polar auxin transport

Nod factors - specific signaling lipochitinoligosaccharide molecules

Gm - gentamicin

Km - kanamycin

Rif - rifampicin

Sp - spectomycin

GFP - a marker gene for green fluorescent protein

GUS is a marker gene for β-glucuronidase activity

OE lines - overexpression lines

RNAi lines - lines with suppressed expression

amiRNA lines - lines with repressed expression

qRT-PCR - Quantitative real-time PCR

HU-hydroxyurea

BLAST- Basic Local Alignment Search Tool - A basic search engine for searching and comparing sequences

NCBI - National Center for Biotechnology Information

TD - Transposon display

IPCR - Inverse PCR /

TAP- Tandem Affinity Purification

BiFC - Bimolecular Fluorescent Complimentation

INTRODUCTION

A major contribution of legumes to sustainable agriculture and the nitrogen cycle is their ability to fix atmospheric nitrogen in agricultural ecosystems. This process is carried out using specialized organs, root noduls in symbiosis with soil bacteria, of the genus *Rhizobium*. The ability of legumes to enter into symbiosis with nitrogen fixing bacteria provides them with a unique advantage over other plant species.

The legume family includes a variety of plants of high economic value - soybean, alfalfa, clover, peas, peanuts and other grain and fodder legumes. Unfortunately, these economically important crops have large polyploid genomes that are difficult to study at the molecular and genetic level. Major discoveries in biology are based on model plant species, and the knowledge accumulated in these species can be transferred to crop plants. Model plants are valuable to plant biologists because they are susceptible to tissue cultivation, have simple genome organization and short development cycles, and have multiple genetic and genomic approaches applicable to them.

Over the past two decades, the model species *M. truncatula* and *L.japonicus* have been established for molecular genetic and biochemical studies. The knowledge gained from model legumes and model plants for the plant biology, *Arabidopsis thaliana* is a source of useful information for crop improvement. The development of modern trends in biology - genomics, proteomics, metabolomics, collectively known as "omics" is a combined scientific approach that is applied worldwide. Genomics analyzes the genetic information encoded in the genomes of living organizms, which gives prospects for deciphering that information and using it for human benefit.

Genomics of legumes is a relatively new scientific field in which two model legumes - *Medicago truncatula* and *Lotus japonicus*, as well as the *Arabidopsis thaliana* model plant - are the main subjects of study. In Bulgaria, studies on *M. truncatula* begin with the participation of the AgroBioInstitute (ABI) in a project under the Fifth Framework Program of the European Union (EU) "Molecular and Cellular Foundations of Somatic Embryogenesis in Alfalfa" (INCO-COPERNICUS - IC15-CT96-0906). Based on the scientific results achieved and the good interaction with one of the partners, the ABI Functional Genetics Group was invited as a partner in a project under the EU's Sixth Grain Legumes Framework Program (FOOD-CT-2004-506223). One of the major achievements of ABI's work in WP 5.2 "*Tnt1* insertion mutagenesis" is the creation of a collection of *Tnt1* insertion mutants of *M. truncatula*. Based on the created mutant collection of *M.truncatula*, the team of Functional Genetics of legumes, became the coordinator of the project "Integrated functional and comparative genomics of model legumes *M. truncatula* and *L.japonicus*", funded by the Bulgarian National Science Fund of the Ministry of Education and Science in the "Ideas" competition, announced

in 2008. Within the framework of this project, research on the functional genomics of legumes in Bulgaria was started by applying the methods of forward and reverse genetics. Key genes responsible for plant development and the process of nitrogen fixation have been cloned and transferred into the genome of the model species *Medicago truncatula*, *Lotus japonicus* and *Arabidopsis thaliana*, to study their function and their interaction with other genes.

Long standing experience in working with the model plant *M. truncatula*, as well as the results obtained, are the basis for building a platform for functional and comparative genomics of legumes and create a prerequisite for integrating Bulgarian genomic research with other European and world programs in this field.

AIM AND TASKS

The purpose of this dissertation is to confirm the role of model legumes *Medicago truncatula*, *Lotus japonicus* and model plant *Arabidopsis thaliana* in modern plant biotechnology research, comparative and functional genomics and their importance in understanding the plant biology of legumes.

To achieve this aim, the following main tasks are set up:

1. Demonstration of asymmetry of the first cell division initiating the process of direct somatic embryogenesis in a liquid medium and tracing the process from single cell to plant.

2. Establishment of genome size in *Medicago* diploid species and the relationship between genome size and competence for inducing direct somatic embryogenesis in a liquid medium.

3. Construction of stable transgenic plants by genetic transformation of cell suspension culture and indirect somatic embryogenesis from *M.truncatula*.

4. Development of regeneration protocols for efficient insertional mutagenesis by *Tnt1* retrotransposon and creation of a collection of mutant lines in *M.truncatula* cv. Jemalong.

5. Development of a system for *Tnt1* insertion mutagenesis in the second model plant *Lotus japonicus.*

6. Identification of the MERE1 retroelement and demonstration of its transposition during *in vitro* regeneration of *M.truncatula*.

7. Phenotyping of selected *Tnt1* insertion mutant lines from the created collection under different growing conditions -*in vitro, in vivo*, aeroponic and hydroponic culture.

8. Construction of vectors for overexpression, knockdown expression (by RNAi and amiRNA methods) and analysis of the promoter activity of the studied genes, as well as constructs for TAP (Tandem Affinity Purification) and BiFC (Bimolecular Fluorescence Complementation) assays.

9. Analysis of the promoter activity of the studied genes in the process of indirect somatic embryogenesis in *M. truncatula*.

10. Comparative genomics analysis to study the function of the gene encoding the F-box protein in model plants *M. truncatula* and *A. thaliana*.

10.1. Determination of gene expression level of *MtF-box* and *AtF-box* by quantitative real-time PCR (qRT-PCR) in progeny of transgenic plants with modified expression and controls and analysis of promoter activity in plant tissues and organs of transgenic plants of *M. truncatula* and *A. thaliana*.

10.2. Morphological analyses and phenotyping of T_1 transgenic plants of *M. truncatula* and T_3 homozygous transgenic plants of *A. thaliana*.

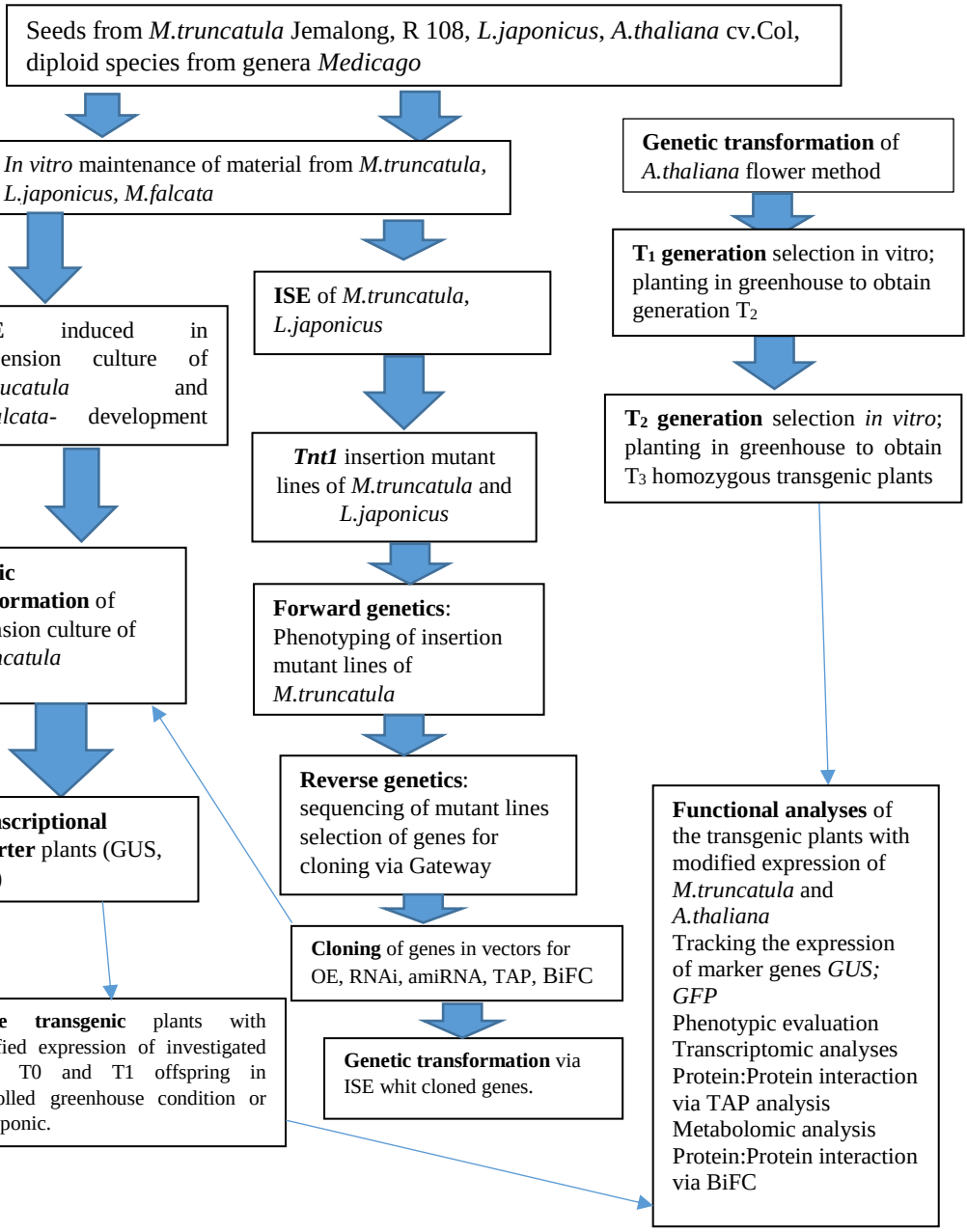
10.3. Determination of maximal activity of the *MtF-box* and *AtF-box* gene during the cell cycle.

10.4. Determination of the gene expression level of the direct interactor of the *MtF-box* and *AtF-box* gene in transgenic plants with modified expression and controls identified after TAP analysis.

10.5. Validation of the results of TAP analysis in *M.truncatula* transgenic plants with modified *MtF-box* expression and control by quantification of free branched amino acids.

10.6. Validation of the results of TAP analysis in *A.thaliana* by conducting a Bimolecular Fluorescence Complementation Experiment.

Schematic presentation of Material amd Methods used in disseratation.



RESULTS AND DISCUSSION

I. Somatic embryogenesis. DSE in a liquid culture medium

The original experimental systems for DSEs and ISEs for model plants presented in this dissertation are summarized in a chapter on a book on somatic embryogenesis, the first edition of the *M.truncatula* model plant handbook, and a review (Iantcheva et al. 2006a; Iantcheva et al. 2008; Iantcheva and Revalska 2018). The accumulated knowledge and experience of working with the model plant is a prerequisite for successful participation in project and project management of the EU's Sixth Framework Program (FOOD-CT-2004-506223 "Grain legumes") and the BNSF project "Integrated Functional and comparative genomics of model legumes *Medicago truncatula* and *Lotus japonicus* "-DO-02-268, 2008.

The results of the somatic embryogenesis studies obtained in the implementation of the above mentioned projects shed light on the processes of regeneration of plants from single-cell to whole plant, contribute to the creation of a collection of insertion mutants from model plants, *Medicago truncatula* and *Lotus japonicus*, and establishment of studies of key genes involved in plant development.

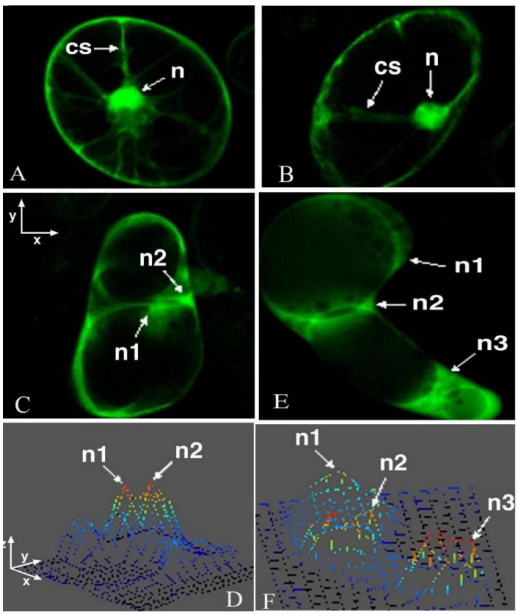
In DSE in liquid medium, somatic embryos develop directly from the epidermal and sub-epidermal layers of the original explant. The protocols developed for *M.truncatula* (Iantcheva et al. 2006a, Iantcheva et al. 2006b) allow the induction of somatic embryos from different plant explants - leaf, petiole, root, single cells. Characteristic of the suspension cultures of *M.truncatula* is the formation of a fine cell fraction. 15 to 20 days after induction of somatic embryogenesis, single cells, cell clusters of 3-5 cells, small (up to 20 cells) and large (more than 20 cells) and cell aggregates are separated from explants in the culture medium. These cells are of various shapes spherical, elliptical or elongated with a size of 100-200 μm and have embryogenic potential. Released in a fresh induction medium, they re-form embryos and can continue to develop throughout the plant. In experimental setings based on cell suspension cultures of transgenic plants of *M. falcata* expressing the *GUS* marker gene under the control of promoters of cell cycle regulatory genes and transgenic plants of *M. truncatula* expressing the green fluorescence marker gene /*GFP*/, under 35S promoter, the asymmetry of the first cell division initiating the process of DSE has been observed and proven (Iantcheva et al. 2006a; Iantcheva et al. 2006b; Iantcheva and Revalska 2018).

1. Asymmetry of the first cell division in model systems *M. falcata* and *M. truncatula* in liquid culture medium

Observations by a confocal microscope on a cell suspension induced by transgenic *M.truncatula* plants expressing the *GFP* marker gene show that the cell fraction is

composed of three cell types - spherical, elliptical and elongated, Fig. 1 A, B (Iantcheva et al. 2006a). These cells have a dense cytoplasm and, when transferred to fresh induction medium containing auxin 2,4-D, are reactivated to division. The green fluorescence signal is observed in the cell nucleus where it slowly accumulates, as well as in the cytoplasmic filaments. In cells competent for division, the nucleus is located in the periphery (Fig. 1B). In mesophilic protoplasts from *M.sativa*, Dijak and Simmonds (1988) also found a nucleus localization in the periphery just before division. The studies done confirm the asymmetric division of the cell into two unequal parts (Fig. 1C). Using the software of the confocal microscope, it is possible to determine the strength of the fluorescence signal in cells and cell structures. The peaks of green fluorescence indicated by the program are two and are connected to the two nuclei of the newly divided cell into two unequal parts. Subsequent observations indicated three peaks of fluorescence and the formation of a three-cell embryo (Fig. 1D, E, F). The development of cellular structures continues with the formation of globular and torpedo structures over a period of 20-30 days.

Figure 1. Single cell DSE of *Medicago truncatula*. **A-** a spherical cell with a marked nucleus and cytoplasmic filaments n-nucleus; cs - cytoplasmic filaments; **B**-oval cell with a nucleus positioned in the periphery; **C**-an asymmetric division with two nuclei n1, n2. **D**-the fluorescence level after the first cell division, a fish-net image of the signal intensity of C. **E**-tricellular proembryo with nuclei n1, n2, n3. **F** - "fish-net display" image of signal intensity of "E".



Genetic transformation in *M. falcata* 47/1/150 line with high regeneration potential is rapid and effective with the production of transgenic plants within two months. 60 independent transgenic plants were regenerated from the two constructs used /pAtcyc3a::GUS, pAtcdc2a::GUS/ and all were verified by histochemical analysis to demonstrate glucuronidase activity, and 10 randomly selected plants were checked for presence of the selective gene neocin phosphotransferase (Iantcheva et al. 2006a). Transgenic plants as well as the starting line are sterile and maintained only in *in vitro* culture.

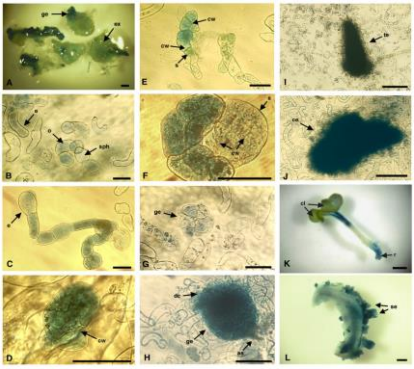
Proven transgenic plants expressing the marker *GUS* gene under the control of promoters of cell cycle regulatory genes (cyclin A /cyc3A/), which is expressed during the S/M phase of dividing cells and cyclin dependent kinase type 2A /cdc 2A/ (a gene expressed during the G1/M phase/ of dividing cells as well as those responsible for division) have been used to induce suspension culture and to produce a cellular fraction. The type of first cell division and the acquisition of embryogenic competence may be associated and visualized with the expression of the marker gene controlled by the above promoters. In *M.falcata*, a fine cell fraction is formed 10-15 days after induction of suspension culture, the small leaf explants form embryoids at their edge and surface, separating the cells competent for division in the culture medium Fig.2 A. Cells, similar to the cell fraction of *M.truncatula*, are similar in shape and size - spherical, elliptical and elongated, possessing embryogenic potential, able to form embryos that develop to plant. Most often, spherical and elliptical cells have dense cytoplasm and are smaller in size up to 100 μ (Fig. 2 B, C). This type of cell is preferably divided asymmetrically into two unequal parts. In most of the dividing cells observed, one cell of asymmetric division forms a suspensor, typically consisting of 2-5 cells /Fig. 2 D, E, F/ and the other cell after asymmetric division continues to divide and form the somatic embryo /Fig. 2 F, G/. The asymmetry of the first cell division and the establishment of polarity are prerequisites for the subsequent development of the somatic embryo. Similar results have been observed in mesophilic protoplasts of *M.varia* genotype A2 from (Dudits et al. 1995). At a later stage in the development well-formed embryos in the globule stage, the suspensor aborts, (Fig. 2H). It is not clear exactly what role the suspensor plays in the liquid nutrient medium, but it is most likely related to the establishment of polarity in developing embryos, or perhaps to serve as a channel for uptake of nutrients and growth regulators from the culture medium until the development of a vascular system in the embryos and subsequently aborted. Following the expression of the marker *GUS* gene under the control of the cyc 3a promoter, it is documented that cells of the aborted suspensor in a well-developed globular embryos do not stain in blue /Fig. 2H/, unlike cells in a pro-globular embryos, which are still competent for division Fig. 2F. Following the DSE procedure, development from day 20 to 40 was continued in liquid medium containing PEG without growth regulators. Dividing cells are observed along

the surface of the formed embryos in the torpedo stage /Fig. 2 I/, the asymmetric divisions decrease by 3-4 times and are equal with the symmetric divisions. The development of the structures up to the cotyledonary stage Fig. 2 J, after day 40, was continued in MS1 liquid medium in the presence of low concentration of cytokinine and casein hydrolyzate. Part of the cotyledonary embryos are converted into plants, and a process of secondary embryo formation is also observed /Fig. 2 K, L/.

Developed cell suspension culture protocols for *M. falcata* and *M.truncatula* are essential in studying early events - first cell division, induction of embryogenic potential in somatic embryos originating from single cells. These systems are the perfect tool to track the process of plant development from one cell to the whole plant. The development of modern genomics and the presence of such systems for somatic embryogenesis allows the identification of key genes and the tracking of their expression in dynamics throughout the plant development process.

Last but not least, the presence of regeneration from cell suspension culture presupposes the development of gene transfer systems and the origin of transgenic plants from single transformed cells. Such a gene transfer system is developed and described in the next section of this dissertation.

Figure 2. Process of somatic embryo formation from single cell to plant in *Medicago falcata*. Expression of the *GUS* marker gene was visualized by blue staining. **A** - leaf explant from the mother suspension with globular embryoids. (pcyc3a gus); ex - explant: ge - globular embryo, bar 1 mm; **B** - single cell types observed in the single-cell fraction (pcyc3a gus suspension culture), o-ellipsoid, sph - vesicle, e - elongated, at least 100 µm; **C** - disorganized elongated cell structure (pcdc2a gus); **E** - elongated cell, at least 100 µm; **D** - asymmetric division into an ellipsoid cell (pcyc3a gus); cw - cell wall, bar 100 µm; e -three-cell pro-embryo (pcyc3a gus A); s - suspensor, bar 100 µm; **F, G** - proglobular and globular embryo (pcyc3a gus), bar 100 µm; **H** - well-formed globular embryo (cdc2a gus A suspension culture): dc - dividing cells along the periphery with embryogenic potential, as-aborted suspension, bar 200 µm; **I** -somatic embryo in the torpedo stage (pcyc3a gus); te – torpedo embryo, at least 1 mm; J - cotyledonary stage embryo (cdc2a gus A): ce - cotyledonary stage embryo, bar 1 mm; **K** - late cotyledon stage (pcyc3a gus A): cl - cotyledonary leaves, r - root, bar 1 mm; **L** - secondary embryo formation (cdc2a gus A): se - secondary embryo, at least 1 mm.



2. Genome size in genera Medicago. Relationship between genome size and competence for inducing DSE

The genome size of diploid species from 17 populations originating from Algeria, including 8 annuals / *M.orbicularis, M.arabica, M.laciniata, M.minima, M.ciliaris, M.truncatula*, annual hybrid, *M. littoralis* x *M.truncatula, M.scutellata/*, and a perennial *M.sativa* was studied in the course of the research (Fyad-Lameche et al. 2016). Genome size was estimated by flow cytometry by a method developed by (Ochatt 2008), in addition to the procedure published in the literature, originally developed by (Blondon et al. 1994). The chromosomal number of this diverse population of diploid species was also analyzed. The genome size ranged from 2C = 0.94 pg in *Medicago orbicularis* (L.) Bartal., the smallest, to 2C = 1.80 pg in *Medicago laciniata* (L.) Miller. Data for the investigated collection originating in Algeria are presented in **Table 1**. The relationship between genome size and the ability to induce DSE and embryogenic response was investigated in four of these species belonging to different populations of *M. orbicularis* DZ19C, (2C = 0.94 pg), *M. truncatula* Gaertn., DZ2B (2C = 1.08 pg) , *M. scutellata* (L.) Miller 274, (2C = 1.11 pg); and *M. arabica* (L.) Huds., INA46, (2C = 1.22 pg). *Medicago orbicularis*, having the smallest genome, has been found to form somatic embryos for the shortest period, with the highest percentage of reacting explants and the number of somatic embryos. It is followed by *M. truncatula, M. scutellata* and *M. arabica*, in fact in order of increasing of genome size **Tab.2**. Representatives of genus *Medicago* in the Algerian collection studied are characterized by two ploidy levels - diploids and tetraploids. Most of the species studied are diploid 2n = 14 or 2n = 16 and the major chromosomal number x = 7 and 8. One of the species studied is tetraploid, *M. scutellata* 2n = 30 (Table 1). The genome size of the annual species of this collection (Table 2) ranges from 2C = 0.94 pg (*M. orbicularis* population DZ19C) to 1.11 pg (characteristic of *M. truncatula*, newly identified for *M. polymorpha* and *M. scutellata*), to 1.80 pg (*M. laciniata*). During the analyzes, reference clones were used, which were tested regularly and were of constant value: *M. truncatula* cv. Jemalong (2C = 1.14 pg), ecotype R108-1 (2C = 0.98 pg); and the tetraploid *M. sativa* "A2" (2C = 3.44 pg).

The competence for the direct formation of somatic embryos in liquid nutrient B5 media has been tested for 4 species of *M. orbicularis, M. truncatula, M. scutellata* and *M. arabica* (Fyad-Lameche et al. 2016). After 10-15 days of induction of suspension culture, globular embryos were formed on the explants or separated in the liquid medium. The percentage of reacted explants (leaves, petioles and roots) varies with the species studied (Table 2). In *M. orbicularis*, leaves and petiols, as well as most of the root explants, form somatic embryos; *M. truncatula* explants are characterized by a weaker embryogenic response, followed by *M. scutellata* explants and the lowest frequency in *M. arabica*, where only one-third of the explants respond. The number of

somatic embryos formed for explant is in the same order (Table 2). The induction period is shortest in *M. orbicularis* (leaves and petioles), 6-8 days; by the tenth day, both leaves and petiols have a significant number of somatic embryos, both associated with explants and separated in the liquid culture medium. Root tissue is characterized by a weaker embryogenic response. Similar results are found for *M. truncatula*.

The determined chromosome number (Table 1) is in agreement with previous data published by Lesins and Lesins (1979). The studied populations of annual diploid species have a haploid number of chromosomes x = 8. In previous studies of (Lesins and Lesins 1979) *M. polymorpha* was determined by a diploid number of chromosomes 2n = 14 (and two satellites), and the chromosomal number for *M. scutellata* 2n = 30 was found in the study of Bauchan and Elgin (Bauchan and Elgin 1984). In our studies, we determined the genome size in 17 populations of Algerian origin in eight annual *Medicago* species, and one hybrid, using 5 *Medicago* species as reference (Table 2). The results obtained are very close to those of Blondon et al. (1994): *M. orbicularis*, originating in Algeria, has 2C = 0.94-0.96 pg (vs. 0.99 pg determined by Blondon et al. 1994); *M. arabica* - 1.22 pg (vs. 1.16 pg). In some cases, significant differences are observed, for example for Algerian *M. minima* 2C = 1.10-1.15 pg (vs. 1.73 pg) calculated from Essad in Blondon et al. (1994) and Algerian *M. ciliaris* by 1.56 pg (vs. 1.62-2.06 pg determined by Blondon et al. 1994). No previous genome size data for *M. laciniata*, *M. polymorpha*, and *M. scutellata* have been found in the literature. The observed significant intraspecific variation in genome size at *M. laciniata*, *M. polymorpha* and *M. scutellata*. The observed significant intraspecific variation in genome size in *M. truncatula*, where Jemalong is 16% larger R108-1, is identical to that reported by Blondon et al. (1994). The Algerian populations of *M. truncatula* are quite homogeneous (identical letters Tab. 2) and have a significantly larger genome by ~ 10% than that of the R108-1 ecotype. In the study, the Algerian annual diploid, *M. orbicularis* (2C = 0.95 pg), has the smallest genome size. The phylogenetic analyzes made by Steele et al. (2010), based on nucleotide sequencing data from a plastid DNA encoding a trnK/matK gene (a chloroplast DNA complex of a mutaraseK gene localized in the trnK intron) and nuclear DNA encoding (giberillin 3-β-hydrolase), *M. orbicularis* is not is part of no characterized group and is defined as an "orphan type". The following is a group of representatives with a genome size of 1.08-1.15 pg including *M. minima*, *M. polymorpha*, *M. scutellata* and *M. truncatula*. The spontaneous hybrid (littoralis × truncatula) also falls into this group, although *M. littoralis*, as a putative parent, has been reported with a much larger genome (2C = 1.34 pg by Blondon et al. 1994). These four types of different sub sections of the Spirocarpos section have almost the same 2C value. Although tetraploid, *M. scutellata* has an unexpectedly low genome size of 2C = 1.11 pg. This result contrasts with those obtained in other genera, where the 2C value increases with the ploidy level (Kharrat-Souissi et al. 2013). The chromosomal number

in *Medicago polymorpha*, 2n = 14, is characterized by diminishing ploidy, which is most likely due to Robertsonian translocation (Singh 2010; Steele et al. 2010). *M. scutellata* is known as tetraploid 2n = 30 (Bauchan and Elgin 1984; Hanson et al. 1988) resulting from hybridization between species with 2n = 14 and 2n = 16 (Lesins and Lesins 1979; Bauchan and Elgin 1984). Based on sequences of the low-repetitive NORK genes, Eriksson and Pfeil (2012) identify a group formed by *M. polymorpha*, *M. lesinsii*, and *M. murex* as the sister group to *M. soleirolii*, *M. doliata*, and *M. scutellata*, and usually cluster with *M. truncatula* in other research. Therefore, this group of annual wild alfalfa with a genome size of about 1.12 pg contains species with a chromosomal number of 2n = 14, 16, as well as atypical *M. scutellata* (2n = 30). *Medicago arabica* is characterized by a relatively large genome, 2C = 1.22 pg, followed by *M. ciliaris* with 1.56 pg and *M. laciniata* with 1.80 pg. *Medicago arabica* and *M. orbicularis* have been identified as sister species in a study based on EST sequencing data (Steele et al. 2010), although genome size values for the two species are quite different. The size of the genome differs significantly in some species, such as a 59% increase between M. *polymorpha* (2n = 14) and *M. laciniata* (2n = 16), although they belong to the same sub-section of Leptospireae. In determining the chromosomal number of *M. laciniata* is a diploid (2n = 16) with a significant genome size twice as large as the genome of *M. orbicularis*. The significant differences found in the genome size of different diploid species have provoked us to investigate the relationship between genome size and the ability to directly induce somatic embryos and embryogenic potential in a liquid medium. Applying the protocol of (Iantcheva et al. 2006b) to the 4 species selected, we observed an inverse relationship between genome size and induction of embryogenic potential. *Medicago orbicularis* (2C = 0.94) responded within the shortest period and with the formation of the highest number of somatic embryos followed by *M. truncatula* (2C = 1.08 pg), *M. scutellata* (2C = 1.11 pg) and *M. arabica* (2C = 1.22 pg). Generally, genome size is a parameter for checking the homogeneity of plants produced by *in vitro* cultivation, and helps to isolate polyploid and aneuploid regenerants. Genome-wide verification also helps to identify and separate plant tissue composed of cells with different levels of ploidy, which is unsuitable for cellular tissue manipulation and predisposes to the production of increased ploidy regenerants (Catrice et al. 2006; Ochatt 2008). Genome size analysis should also be applied to taxonomic studies in different populations. Populations studied from *Cenchrus ciliaris* in Tunisia differ in ploidy levels (Kharrat-Souissi et al. 2013).

Table 1. Genome size and chromosomal number observed in genus *Medicago* originated from Algeria. Spirocarpos subsections are separated by lines, each homogeneous group are marked with the same letters (at $P < 0.05$). The genome size can also be expressed as base pairs using the conversion of 1 pg DNA ≈ 978 Mbp.

Species	population	Chromosom number 2n	Genome size 2C (pg)	SD (pg)	Groups P<0.05 a	Number of repeats
Section *Orbiculares*						
M. orbicularis	DZ14A	16 0.960			-	
“	DZ19C	16	0.943	0.018	a	8
Section *Spirocarpos*						
M. arabica	46	16		1.220	-	
M. laciniata	239	16	1.800	0.007		4
M. minima	216	16	1.099	0.005	bdefghi	6
“	230		1.149	0.011	dghi	6
M. polymorpha	16	14	1.111	0.028	cdefghi	14
“	511	14	1.114	0.010	defghi	12
“	DZ15A	14	1.125	0.011	cdefghi	8
“	DZ20D	14	1.109	0.006	bdefghi	14
“	Serena	14	1.173	0.011		2
M. scutellata	274	30 1.110			-	
M. ciliaris	DZ13C	16	1.562	0.012		8
M. truncatula	28	16	1.120	0.009	cefghi	10
“	448	16	1.115	0.006	bdefghi	8
“	DZ2B	16	1.079	0.009	be	10
“	DZ15B	16	1.099	0.043	bcefgi	23

“	Jemalong(Gif)	16	1.137	0.022	cdefghi	6
“	Jemalong(ITGC)	16	1.136	0.010	cdefghi	12
“	R108-1	16	0.976	0.003	-	3
Section *Medicago*						
M. sativa	A2	32	3.444	0.026	k	12
M. sativa	Sitel	32	3.429	0.028	k	6
Hybrid of annuals						
littoralis×*truncatula*	DZA3175-2	?	1.122	0.002	cdefghi	2

Table 2. Potential for somatic embryoid formation in 4 annual alfalfa arranged in ascending genome size. a, Mean ± sd of three independent experiments, average of 10-20 explants in each experiment

Species		Period for induction		Frequencey of reacted explants (%)		*Number of embryos per explant*	
	(дни)	Лист	л.др.	корен	лист	л.др. корен	
M. orbicularis DZ19C	8	95±1	95±1	70±1	15±1	10±1	4±1
M. truncatula DZ2B	10	90±1	80±1	70±1	10±2	9±1	7.5±1
M. scutellata 274	10-15	70±2	70±2	60±2	6.5±2	5±1	3±2
M. arabica INA46	10-15	40±1	20±1	50±1	5±1	4±1	6.5±1

II. Genetic transformation

The development of an effective plant regeneration protocol is a main requirement in the creation of transgenic plants. The presence of an effective genetic transformation system facilitated by *Agrobacterium tumefaciens* is a challenge for plant biologists doing research in legume biology.

Genetic transformation of cell suspension culture

The developed system for the genetic transformation of cell suspension culture (Iantcheva et al. 2014; Iantcheva and Revalska 2019) allows the creation of stable transgenic plants regenerated by single cells or small cell clusters, is a new strategy for creating transgenic plants with modified gene expression and plants to study promoter activity. The origin of single-cell transgenic plants eliminates the creation of chimeric transgenic plants. Transformed suspension culture offers unlimited possibilities for studying the mechanisms of somatic embryogenesis and functional analyses based on current systems biology methods for studying the function of the genes studied, as well as the interactions of proteins encoded by these genes.

Four milliliters of refined fine cell fraction treated with acetosyringone for 24 h was co-cultured with 400 μl of bacterial suspension (pMtARF-B3: GUS / GFP or 35S: CycF-GFP) on a shaker. After co-cultivation, the cells and clusters of the suspension culture together with the bacterium form a total dense cell mass. This dense mass was transferred to a solid CIM medium for another 48 h of co-cultivation, and then a period of induction of callus structures in the form of small pellets in a selective CIM medium began (Figure 3, A). In the course of the experimental work, two concentrations of acetosyringone were tested to improve the transformation efficiency. A concentration of 25 μM was found to have a positive effect on the transformation frequency observed in the formed callus beads expressing the green fluorescence marker gene. The steps of the regeneration process of the transformed suspension culture are presented in (Figure 3). Callus structures resembling balls are observed for 15-20 days (Fig. 3B). Callus structures of size (1-2 mm) that are growing are transferred to a fresh selective medium (Figure 3 C). During the second passage of selective medium, a second screening of callus beads (1-4 mm) for green fluorescence is performed. Structures with a clear gfp signal are present and transferred to media for embryo induction and development. During the two weeks of embryo induction medium on the surface of the callus balls, the formation of green zones and the subsequent development of embryos were observed (Fig. 3 D, E). The presence of auxin NAA and cytokine BAP in a ratio of 1: 3 in the medium for induction of somatic embryos positively influenced this process. Subsequent development of somatic embryos continues in the presence of low concentrations of cytokinin BAP and casein hydrolyzate, which stimulate embryo elongation and the development of well-formed bipolar structures (Figure 3 F). The transgenic nature of the embryo structures formed in the globular, torpedo and cotyledonary stage is confirmed by histochemical analysis for the presence of GUS activity. The promoter activity of MtpARF-B3 was visualized by histochemical analysis of embryos at different stages (Fig. 4 A, B), and the presence of a green fluorescence signal was also confirmed in the nuclei of epidermal cells of late torpedo embryos (Fig. 4 C). Cotyledonary stage embryos with a well-defined bipolar structure continue to develop into plants on a selective medium. In some cases, clusters of embryos are

formed which are carefully separated to facilitate the root formation process. The period for the conversion of somatic embryos into plantlets is 20-40 days.

Rooted on the selective-medium plants were subjected to PCR analysis to demonstrate the presence of a selective gene (Fig. 4 D). Plants showing a positive signal are transferred to a greenhouse for offspring. Regenerated plants by transformation of cell suspension culture have morphology, flowering time and seed pods formation such as control plants. The efficacy of the developed gene transfer system was evaluated in the projeny in the presence of a selective gene in the genome, as well as a positive signal for β-glucuronidase activity or presence of green fluorescence signal (Figure 4. E, F, G). The harvested T_1 seeds of transgenic plants were grown in a selective medium to trace inheritance in the offspring. The pool of tested seeds is characterized by a Mendelian inheritance 3: 1.The developed system for the genetic transformation of cell suspension cultures of *M. truncatula* has many advantages, but one of the most important is that stable and fertile single-cell transgenic plants are regenerated. Transformed suspension cell cultures support the study of various physiological, biochemical processes that take place at the cellular and subcellular levels. They are a powerful tool for functional analysis and analysis of promoter activity of studied genes and protein-protein interactions. The Suspension Cell Transformation Protocol provides an opportunity to characterize specific gene functions and key stages in the development of legumes.

Figure 3. Steps from the process of transformation of cell suspension culture from *M.truncatula*: **A**-Cell suspension plated after co-cultivation. **B**-first passage of callus structures of selective CIM medium. **C**-callus structures ready for screening for the presence of GFP signal. **D**-formation of embryogenic zones of EIM medium. **E, F** Embryo development and extension. Scale bar 2 mm.

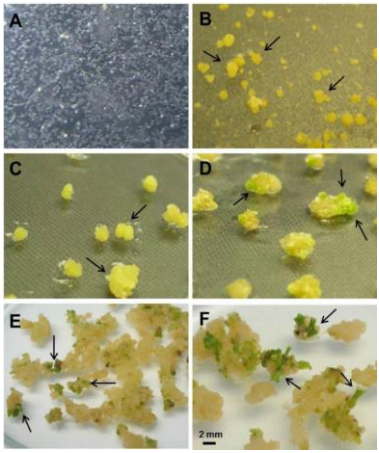
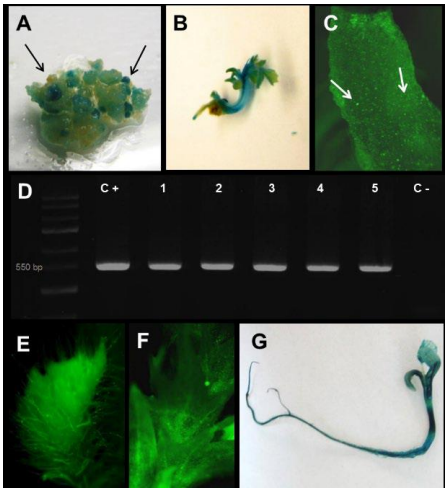


Figure 4. A, B. Expression of the marker *GUS* gene under the control of the MtARF-B3 promoter in globular and cotyledon embryos. **C** - Expression of the *GFP* marker gene in a torpedo embryo. **D** - PCR analysis for the presence of the *nptII* selective gene in T_0 plants after transformation with the Mt35S: Cyc-F-box-GFP construct. **E, F** - Leaf and sepals of T_1 GFP positive plants. **G**-GUS controlled by MtARF-B3 positive T_1 seeds.



Genetic transformation by indirect somatic embryogenesis

Generation of stable transgenic plants through a process of indirect somatic embryogenesis is most common among legume researchers. Despite the long and time-consuming period for producing transgenic plants, this method is most used by plant biologists. The construction of most of the stable transgenic plants with modified expression of the studied genes, as well as those for visualizing promoter activity, through the process of ISE in the studies of the Functional Genetics legume group has been documented in our published research papers (Revalska et al. 2015; Iantcheva et al. 2015; Boycheva et al. 2015; Revalska et al. 2016a; Boycheva et al. 2017; Revalska et al. 2017; Revalska and Iantcheva 2018a; Revalska and Iantcheva 2018b). ISE is characterized by a phase of callus tissue formation and subsequent induction of somatic embryos, which undergo development and conversion phases. The continuous experimental work with representatives of the genera *Medicago* allowed us to determine the optimal parameters for efficient generation of stable transgenic plants. The most commonly used plant explants are leaf and petioles of *in vitro* plant material aged 30-35 days. For better efficiency in the gene transfer process, starter explants are usually pretreated (for 2-4 days) on a callus tissue medium containing auxin 2.4-D at relatively high concentrations of 4 to 5 mg/l, which provokes explant cells to reactivate cell division. The optimal density of the bacterial suspension varies from OD = 0.5-0.7, with

the desired density being attained by dilution using the maintenance-specific *Agrobacterium* medium YEB. This creates favorable conditions for the attachment of the bacterium to the plant tissue and subsequent infection. In the course of the experimental work, we found that inoculating the pre-treated explants with a bacterial suspension for 1 hour on a horizontal shaker at low rpm of 40 rpm allows each explant to be in direct contact with the bacterium. The transformed plant material was co-cultured for 48 hours in the dark of the medium used for pre-treatment.

The application of selection immediately after the co-cultivation period, as well as the interruption for a period of one passage of the medium to induce somatic embryos, positively affects the regeneration of the transformed tissue and at the same time minimizes the number of regeneration of non-transgenic plants. The optimal dose of the selective agent most commonly found to be the antibiotic kanamycin is 50 mg / l Km. The choice of an antibiotic to eliminate bacterial infection is also important in the process of regenerating transgenic plants. In our experiments, the antibiotic used was carbenicillin (Cb) (Carbenicilin disodium, Duchefa), which is more effective than others like Claforan (Cefotaxime), Augmentin. It is used at a concentration of 400 mg/l with subsequent reduction to 200 mg/l during embryo induction and development, which does not adversely affect the regeneration process and successfully eliminates bacterial infection. The somatic embryos induction medium is characterized by the presence of auxin naphthyl acetic acid NAA and cytokinin benzylaminopurine, which positively influences the formation of green embryogenic zones and the subsequent induction of somatic embryos by them. In our previous studies on DSE (Iantcheva et al. 2006b; Iantcheva et al. 2005), the positive effect of cytokinin at a low concentration of 0.05 mg/l BAP on the process of somatic embryo development, supplemented by the action of casein hydrolyzate 250 mg/l. The addition of free amino acids in the form of casein hydrolyzate in the process of embryo development assists in elongation of embryos and root formation and with subsequent development of cotyledonary leaves. At this stage of development, somatic embryos with the morphology described, undergo a process of convertion to plant. The process of forming a well shaped root and rooting of the putative transgenic plants on medium with the presence of a selective agent, is positively affected by halving the concentration of macro elements and sucrose.

The pool of transgenic plants obtained is subjected to massive screening most often to demonstrate β-glucuronidase activity or the presence of a green fluorescence signal. From the plants possess positive signal, few are randomly selected to be subjected to PCR analysis for the presence of *nptII* gene for antibiotic resistance. Proven transgenic plants are acclimated to an adaptation room in soil and raised for offspring. T_0 transgenic plants thus regenerated and their progeny T_1; T_2 is used in subsequent morphological, physiological, molecular and functional analyses.

III. Establishment of a collection of *Tnt1* insertion mutants in *M.truncatula*

The first step for establishment of an extensive collection of insertion mutant lines from *M.truncatula* cv. Jemalong is evidence that *Tnt1* retrotransposon is activated during *in vitro* regeneration. Initiall, transgenic lines bearing a small number of copies of *Tnt1* were created, the so-called "starting lines" - A, B and C, respectively with 8, 4 and 2 *Tnt1* inserts, which are characterized by high regeneration potential. Initially, the starting lines were regenerated using the R1 protocol (Chabaud et al. 1996). The application of this protocol on lines A and B causes transposition that varies widely in regenerated plants with up to 10 new inserts or the absence of new transpositions. These results are completely different from those obtained for the R108 genotype, which observed a frequency of new inclusions of *Tnt1* to 25 in regenerated plants (d'Erfurth et al. 2003, Tadege et al. 2008). A similar heterogeneous transposition in which only 40% of the regenerated plants are characterized by new *Tnt1* inserts has also been observed in lettuce (Mazier et al. 2007). It should be noted that the tested starting lines A and B, when applying protocol R1, do not show a significant difference related to the frequency of transposition (Table 3). None of the 14 plants tested regenerated from the C line with 2 copies of *Tnt1* showed new transpositions.

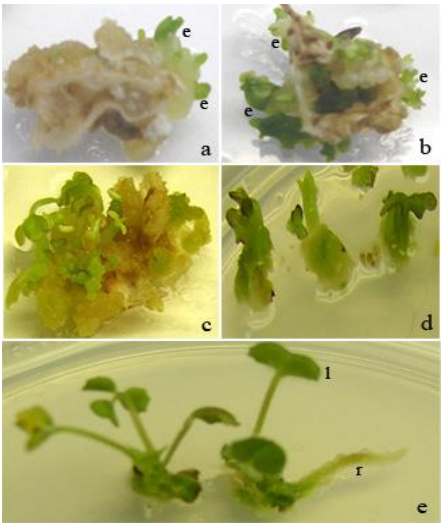
Table 3. Percentage of plants with new *Tnt1* inserts *in vitro* regenerated of A and B start lines and using protocols R1, R2, R3 and R5. Statistical analyses made using χ^2, which shows a significant difference $\chi^2 > 3.84$ (p value = 0.05). There is no significant difference between the two starting lines when applying the same protocol. The comparison of the protocols is done by combining the results of the two starting lines / A, B /. The analysis shows that protocols R2, R3, and R5 do not differ statistically in terms of the efficiency of *Tnt1* transposition (χ^2 not specified), but protocol R1 is statistically less efficient when compared to the other three protocols. (χ^2 bolded values).

Starting line	Number of analysed plats	Plants with new *Tnt1* inserts Number	%	Comperison between starting lines χ^2 стойност	Comparison between R1 and other protocols χ^2 стойност
Протокол R1					
A	75	28	37	1.77	-
B	67	18	27		
Протокол R2					
A	21	13	62	-	**4.04**
Протокол R3					
A	28	22	77	0.52	**24.96**
B	16	11	68		
Протокол R5					
BCX-N	285	193	66	0.02	**50.51**

The unsatisfactory results of the transposition frequency using the R1 protocol for the Jemalong genotype necessitated the development and testing of new R2 and R3 protocols for ISE (Iantcheva et al. 2008) in order to achieve a higher transposition rate

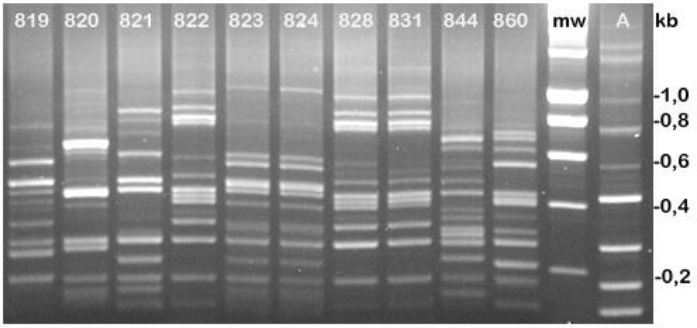
during the *in vitro* regeneration process. Protocols R2 and R3 are characterized by the presence of callus tissue and application of auxin 2.4-D in the induction medium at a higher concentration or difference in the ratio of cytokinins: auxins compared to protocol R1. The newly developed R2 protocol has a relatively short regeneration period with the induction of small callus tissue, such as the nutrient medium for inducing callus is based on MS macro, micro elements and vitamins. In the newly developed R3 protocol, the medium is composed on the basis of B5 salts and vitamins, and the concentration of 2.4-D is increased to 4 mg/l in combination with cytokinine kinetin at a concentration of 0.8 mg/l, which favors the formation of large callus tissue. In protocol R5, the concentration of auxin in the callus tissue induction medium is 5 mg/l, but the basic mineral composition and vitamins are based on the SH culture medium borrowed from the regeneration protocol at genotype R 108 1. The changing concentration of 2.4D in the various protocols, the frequency of *Tnt1* transposition was monitored. Protocol R2 is characterized by the rapid and efficient formation and development of somatic embryos and the main points in the course of regeneration under this protocol are presented in (Fig. 5). About 200 mutant lines were regenerated through this protocol, of which 21 were randomly selected and analyzed by the IPCR method. 13 of the plants regenerated by this protocol (62%) had new inclusions of the *Tnt1* retrotransposon (Table 3) with an average of 6-7 new copies per regenerated line.

Figure 5. ISE regeneration using protocol R2. a- induction of a small amont of callus tissue in the presence of 1 mg/l 2.4D leaf. b-formation of somatic embryos; c, d - development of somatic embryos and formation of cotyledonary leaves; e-conversion to a plant.



About 400 plants were regenerated according to protocol R3, the main steps in regeneration being shown in (Fig.6), using start lines A and B. 44 of these plants were randomly selected and analyzed by TD-PCR, 33 of them were characterized by new insertion events (Table 3) with an average of up to 11 new retrotransposon inclusions in each regenerated plant (Fig. 7). The achieved transposition efficiency is approximately 75%.

Figure 7. Transposon display analysis of mutant lines regenerated by protocol R3. A - start line; mw marker



More than 500 plants were regenerated using the R5 protocol, 339 of which were analyzed. About 67% owned at least one new copy of Tnt1 (Table 3). Most of the regenerated plants are characterized by 3-5 new copies, but only a few have more than 10 new copies, with an average frequency of 3 to 4 new *Tnt1* inclusions per regenerated

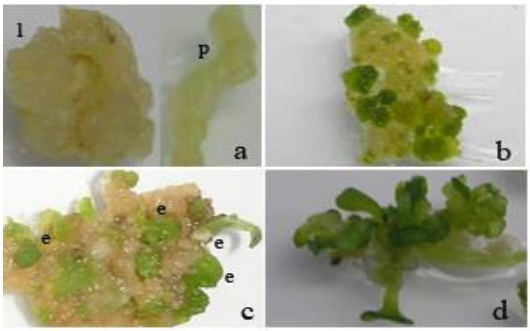


Figure 6. ISE regeneration using protocol R3. a-induction of callus tissue in the presence of 4 mg/ l 2.4D of leaf-l and petiole-p. b-formation of embryogenic zones; c-formation and development of somatic embryoids; d-development of cotyledonary leaves.

plant reported for this protocol. For the protocols tested, there was no difference in the transposition frequency associated with the starting lines (Table 3).

In the Grain legumes project, Work Package 5.2, the Functional Genetics Legume Group was the main producer of mutant insertion lines. On our initiative, a new approach to increasing the transposition frequency has been developed. In the course of experimental work, pretreatment of osmotic shock explants (5M sucrose) before initiating regeneration was tested to increase the transposition rate of the *Tnt1* retroelement (Iantcheva et al. 2009a). Previous studies have shown that transcription of the *Tnt1* retroelement can be induced by various treatments with chemical agents or treatments by applying abiotic stress (Mhiri et al. 1997, Vernhettes et al. 1997). To determine whether such treatments stimulate *Tnt1* transposition in *M.truncatula* cv. Jemalong during *in vitro* regeneration we tested the effect of two types of chemical agents in the step of the regeneration procedure. Salicylic acid pre-treatment and strong osmotic shock were tested (5M sucrose) of the original plant tissue. Salicylic acid treatment during callus tissue induction in regeneration protocol R1 was found to have no effect on the transposition of *Tnt1* in *M.truncatula* cv. Jemalong. In contrast to these experiments, a significant increase in the efficiency of *Tnt1* transposition was observed by pretreatment of the original plant tissue by the application of osmotic shock and subsequent regeneration according to protocols R1, R2 and R3. The effect of osmotic pretreatment on *Tnt1* transposition was evaluated by immersing the initial leaf explants in sucrose solution prior to application of protocols R1, R2 and R3. We found that osmotic pretreatment results in a shortening of the callus tissue induction phase, similar to that described by Iantcheva et al. (2005a), but it also adversely affects the viability of explants. Sucrose treatment and application of the R2 protocol results in necrotizing explants and subsequent formation of only 3-5 explant embryos instead of 5-15 reported without pretreatment. In protocol R3, leaves and petioles are able to form a soft callus, but the number of embryos formed is reduced to 5. Despite the negative effects that osmotic treatment has on the explants and the number of embryos formed, approximately 350 plants have been regenerated and randomly selected regenerants from the two starting lines A and B were analysed by TD-PCR. The results show that, regardless of the regeneration protocol, osmotic pretreatment increases the percentage of regenerants with new *Tnt1* inserts (Table 4). Among the lines tested, we found 6 to 11 new copies per line of the retrotransposon. From the experiments performed in the application of the different protocols, we can conclude that osmotic pretreatment increases the efficiency of transposition under protocol R1 (Table 4) and R2 and significantly increases the efficiency under protocol R3. The osmotic pretreatment duration is (3-5h) for protocols R2 and R3 compared to protocol R1 (1h). With optimal pretreatment duration (3-5h) and application of the R2 and R3 protocol for regeneration, the rate of regeneration of lines with new inserts increased to 95% (Table 4).

Table. 4. Effect of sucrose pretreatment on *Tnt1* transposition in regenerated plants using starter lines A and B and protocols R1, R2 and R3. Statistical analyses were performed using a χ^2 test that showed a significant difference of $\chi^2 > 3.84$ (p value = 0.05). the positive effect of sucrose pretreatment on the number of regenerated plants with new inserts was statistically significant for protocols R1 and R2 (χ^2 values in bold), but not for protocol R3.

Pre-treatment conditions	Number of plants analysed	Plants with new *Tnt1* inserts		Statistical affect of the pretreatments with sucrose on *Tnt1* transposition
		Number plants	%	χ^2 value
Protocol R1/ starting line B				
control	27	3	12	**6.2**
+ sucrose	50	19	42	
Protocol R2/ starting line A				
control	19	11	62	**7.1**
+ sucrose	19	18	95	
Protocol R3/ starting line B				
control	16	11	68	3.5
+ sucrose	17	16	94	

The process of creating a collection of mutant lines in *M. truncatula* Jemalong helped to determine important parameters related to the efficiency of transposition of *Tnt1*. *In vitro* generation of a collection of insertion mutants derived from lines with a small number of retrotransposon copies leads to different transposition efficiency of *Tnt1*. Osmotic shock treatment helps to increase the percentage of *Tnt1* transposition in 95% of the lines recovered. This treatment increases not only the number of lines with new inserts, but also the number of new inserts for each individual plant. Establishing optimal transposition conditions assists in the process of constructing a collection of insertion mutants and saturating the genome with mutations in *M. truncatula* Jemalong. The new plant regeneration protocols developed are of benefit to legume researchers in the creation of transgenic plants required for functional genomics studies.

IV.Creation of a system for *Tnt1* insertion mutagenesis in *L. Japonicus*

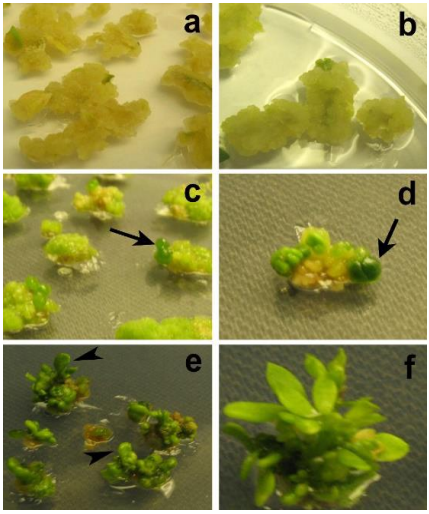
One of the main tasks of the IFKOSMO project is to develop a system for *Tnt1* insertion mutagenesis in the second model legume *L. japonicus*. The availability of an alternative to the endorethrotransposon LORE1 to create mutant lines in this model plant would support comparative genomics research (Iantcheva et al. 2016). A major step in the creation of a collection of mutant lines through the *Tnt1* retrotransposon is the availability of an effective plant regeneration/transformation system. From the experience of creating a collection of *M.truncatula* insertion mutants, we have identified the important role of the step of dedifferentiation of plant tissue cells by forming an unorganized callus mass that facilitates the re-activation of the retroelement and its transfer to new sites in the genome. To this end, the newly created regeneration/transformation procedure was developed in the presence of auxin 2.4-D,

which is not a typical growth regulator used for regeneration in *L. japonicus*. The presence of auxin 2,4-D in the callus tissue induction medium (CIM) causes tissue cell dedifferentiation and soft and loose callus formation (Fig. 8 a, b). The use of light and dark conditions in inducing callus tissue from starter explants and its subsequent increase is equally effective and results in a positive response in 85% of the explants placed. There is a difference in the color of the induced tissue (yellow under dark and light green when light is cultured) but is of equal capacity in the embryo formation process (Fig. 8 a, b). The presence of cytokinin BAP and auxin NAA in the process of embryogenic zone formation (EIM) leads to the appearance of green spots, which subsequently develop to globular embryoids (Fig. 8 c) over a period of 30-40 days. The newly developed embryonic development medium (ED) in the presence of cytokinin BAP alone makes it possible to further development of the formed globular embryos. Subsequent 2 to 3 passages of this medium resulted in the successful formation of cotyledonary embryos in 50% of the induced globular embryos (Fig. 8 d, e, f).

Genetic transformation

Pretreatment of explants of solid CIM for 48 h before inoculation with a bacterial suspension with an optical density of OD600 = 0.3 for 1 h has established the optimal conditions for maintaining the viability of the explants. The selective pressure of 5.5 mg/l hygromycin (Hyg) does not completely eliminate callus formation in control explants, but promotes rapid callus formation in the transformed tissue. The presence of a selective agent in the first two regeneration steps - callus induction and embryo formation helps to separate the transformed tissue, but removal of selective agent during the development of somatic embryos aids the successful growth of the structures. The optimal gene transfer parameters in *L. japonicus* thus established lead to the regeneration of 57 well-developed transgenic plants. The PCR analysis confirmed the inclusion of the *hptII* gene in 86% of the transgenic lines recovered. The reported 14% plants showing a negative result are most likely a consequence of the interrupted selection during the embryo development process.

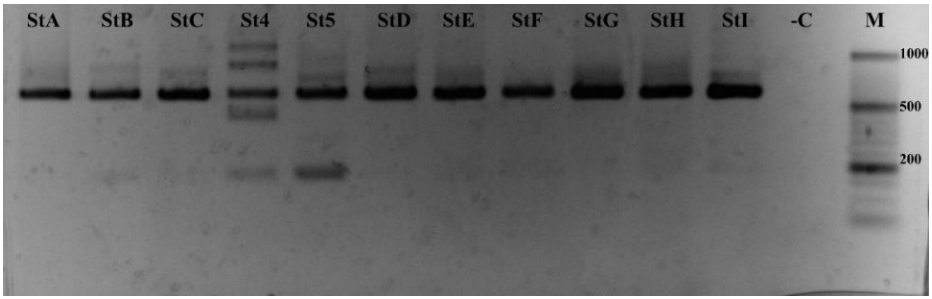
Figure 8. Steps in the regeneration process of L. japonicus using the newly developed callus phase initiation protocol. a, b soft and loose callus induced in light and dark; c, d embryogenic areas of the medium (EIM) that develop to globular embryoids; is the formation of cotyledon embryoids; f-formation of the plant.



Determination of Tnt1 copy number in L. japonicus starter lines

The number of *Tnt1* copies in the 57 transgenic plants obtained, called'starting lines' (St), was analyzed by the IPCR method. This method is effective in proving mutant lines and has been used in our previous studies (Iantcheva et al. 2009a; Ratet et al. 2010; Revalska et al. 2011) and confirms the rapid and accurate demonstration of *Tnt1* copy numbers in characterizing starter lines from *L. japonicus*, their regenerants and their offspring (Fig. 9). In most of the transgenic lines studied, a single band of 600 bp was observed (Fig. 9). This fragment corresponds to the flanking sequences of the Tnt1 retroelement (pCambia1381XC: Tnt1 T-DNA), which is integrated into the genome of transgenic plants and confirms their transgenic nature. Fragments smaller than 200 bp observed in lanes 4 and 5 (start lines St4 and St5, Fig. 9) and St6 and St8 (Fig. 10) correspond to the two terminal repeats of the retroelement and several base pairs of the genome and represent free copies of the element in transgenic plants. These free copies show the expression and reverse transcription of the retroelement in these lines. In the St4 starter line, 3 to 4 additional fragments are identified that correspond to new insertion points of the element and indicate that it is active during the transformation/regeneration process. The profile of the St4 line makes it a good candidate as a starter line in subsequent transposition experiments.

Figure 9. IPCR analysis of *Tnt1* starter lines (StA-I) with the most commonly observed insertion profile. Lanes 1 to 11, Tnt1 positive lines; lane 4 and 5, St lines 4 and 5 with 5 and 3 Tnt1 copies; lane 12 (-C), negative control; lane 13 (M), DNA marker.



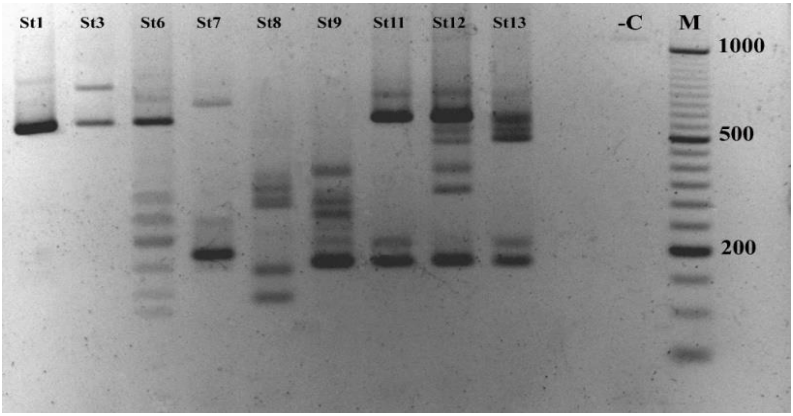
The results shown in Figure 10 show that some of the generated starter lines (St7, St8 and St9) do not include in their genome a 600 bp fragment corresponding to the T-DNA of the construct. However, these lines are PCR positive for the flanking *Tnt1* sequences. These results are confirmation of the results obtained by d'Erfurth et al. (2003) and show that *Tnt1* insertions in the Lotus genome are T-DNA independent during the transformation / regeneration process. Starting line, St9 contains several flanking sequences that represent several insertion sites of the *Tnt1* retroelement in the Lotus genome. Among the lines analyzed, some are characterized only by T-DNA copy (St1) or T-DNA copy and additional stand-alone copies of the starting lines St3, St6, St11, St12 and St13, which is confirmation of transposition into the Lotus genome during plant transformation.

Transposition of Tnt1 is reactivated during in vitro regeneration by ISE

Reactivation of the transposition of the *Tnt1* tobacco element during *in vitro* regeneration has been demonstrated in several heterologous plant species of soybean, potato, lettuce, wild alfalfa (Courtial et al. 2001; Cui et al. 2013; d'Erfurth et al. 2003; Duangpan et al. 2013; Mazier et al. 2007; Iantcheva et al. 2009a). Our studies confirm that during the regeneration process of *L. japonicus* via ISE *Tnt1* is successfully transposed. The starter lines created are characterized by good regeneration potential, which makes it possible to obtain regenerants from the starter lines and to study the possibility of the retroelement being transmitted through ISE. The new regeneration protocol for *L. japonicus* developed was designed to efficiently transpose *Tnt1*, with the St3, St6, and St11 starting lines selected for these reactivation experiments (Fig. 10) and their regenerants and T_1 progeny re-analyzed. The transposition profiles of randomly selected regenerants R12, R14 and R18 recovered from the St11 line (with 4 *Tnt1* inserts) and the regenerants R13, R15 and R16 recovered from the St3 line (with 2

inserts), as well as R6 recovered from the St6 line (with 8 *Tnt1* inserts) were analyzed. The regenerated R12 and R14 lines contain 3 and 4 new *Tnt1* inserts compared to the St11 starting line. In regenerant R13, 2 new inserts were noted, and in R15, 4 new transitions were noted. The transposition rate in the genome of *L. japonicus* ecotype B-129 Gifu due to *Tnt1* reactivation varies from 1 to 4 new copies and is significantly lower than that reported in *M. truncatula* cv. Jemalong, where 8 to 12 new transpositions are observed after *in vitro* regeneration from an 8-copy start line.

Figure 10. IPCR of Tnt1 starter lines. Number of *Tnt1* copies: lane 1 and 2, St1 with 1 copy; St3 with 2 copies; lane 3, St6 with 8 copies; lane 4, St7 with 3 copies; lane 5, St8 with 4 copies; lane 6, St9 with 5 copies; lane 7, St11 with 4 copies; lane 8, 7 copies with St12; lane 9, St13 with 5 copies; lane 11, negative control; lane 12, DNA marker.



Tnt1 inserts are stable and are inherited in T1 generation

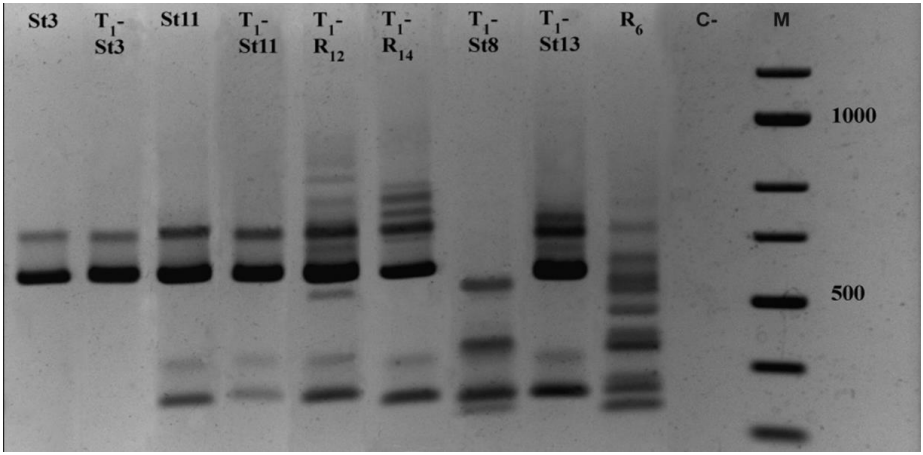
Well-developed *in vitro* starter plants and their regenerants have been adapted under greenhouse conditions to produce offspring. Plants are characterized by a normal flowering period and form pods. Only regenerant R6 derived from the St6 start line is sterile and no progeny have been obtained. T_1 offspring is derived from all starter lines and their regenerants. Seeds of T_1 from several of the starter lines and seeds of their regenerants are germinated in a selective medium in the presence of 5.5 mg/l hygromycin. The lines examined show a 3:1 ratio decay, except for the T_1 offspring of St6 and St13, in which a 1:1 ratio decay was observed. We suppose that this deviation from the expected decay can be explained by the small seed population analyzed (12 seeds only).

The *Tnt1* retroelement would be a suitable insertion mutagen to create a collection of *L. japonicus* mutants only if *Tnt1* remains stable during plant development and in the offspring. To confirm the stability of the retroelement, we compared the transposition profile of some starter lines during their growth in a greenhouse and in the offspring obtained by IPCR analysis. The transposition profile of T_1 progeny of the regenerants R14 and R12 and the starting lines (St3, St11, St8 and St13) is presented in Fig. 11. No new transitional events were noted during growth and in progeny, which confirms that the Tnt_1 retroelement is not active during the life cycle.

Sequencing of (FSTs)

To demonstrate that *Tnt1* is involved in the genome and also has hits in the coding regions, DNA samples from the offspring of two starter lines and the R6 regenerant were sent for lab sequencing on Prof. Pascal Rathe. The fragments (FSTs) obtained from IPCR experiments on lines R6, T_1 St3, and T_1 St12 were cloned and sequenced. A total of 18 fragments were sequenced. As indicated above, the fragments correspond to autonomous copies of the *Tnt1* retroelement with two to several base pairs between the two boundary repeats were obtained at all three lines (10 sequences in total). Bound *Tnt1* borderline repeats with T-DNA were also detected in all three lines. From T_1 St3 line 4 sequences were obtained representing the same FST fragment and corresponding to *Tnt1* insert in a putative gene. These sequences (493 bp) are not represented as part of the Lotus genome in accessible databases but correspond to a putative cytochrome P450 coding gene, and show high homology (90% AK matching) with cytochrome P450 (XP_006596023.1) soybean gene *Glycine max*. The sequence of this IPCR fragment identified 5 bp repeats (TTTTC) characteristic of *Tnt1* transposition. This result indicates that *Tnt1* is active in this lineage and can be involved in DNA coding regions.

Figure 11. IPCR of selected starter lines, regenerated lines and T_1 offspring: lane 1, St3 in greenhouse; lane 2, T_1 of St3; lane 3, St11 in the greenhouse; lane 4, T_1 of St11; lane 5, T_1 of R12; lane 6, T_1 of R14; lane 7, T_1 of St8; lane 8, T_1 of St13; lane 9, R6 regenerant of St6; lane 10, negative control C-; lane 11, DNA marker.



The created small collection of *Tnt1* insertion mutants in *L. japonicus* ecotype B-129 Gifu includes starter lines with a small number of retroelement copies obtained by transformation with *A. tumefaciens*, regenerators of starting lines with new copies of the retroelement. The included *Tnt1* copies are inactive throughout the life cycle of the plant and are inherited in the offspring. Last but not least, the established protocol for efficient transformation/regeneration in the second model legume plant is important. This protocol has been successfully used in our functional and comparative genomics studies in model legumes presented in some of our publications (Revalska et al. 2015; Revalska et al. 2017; Boycheva et al. 2016, Boycheva et al. 2015).

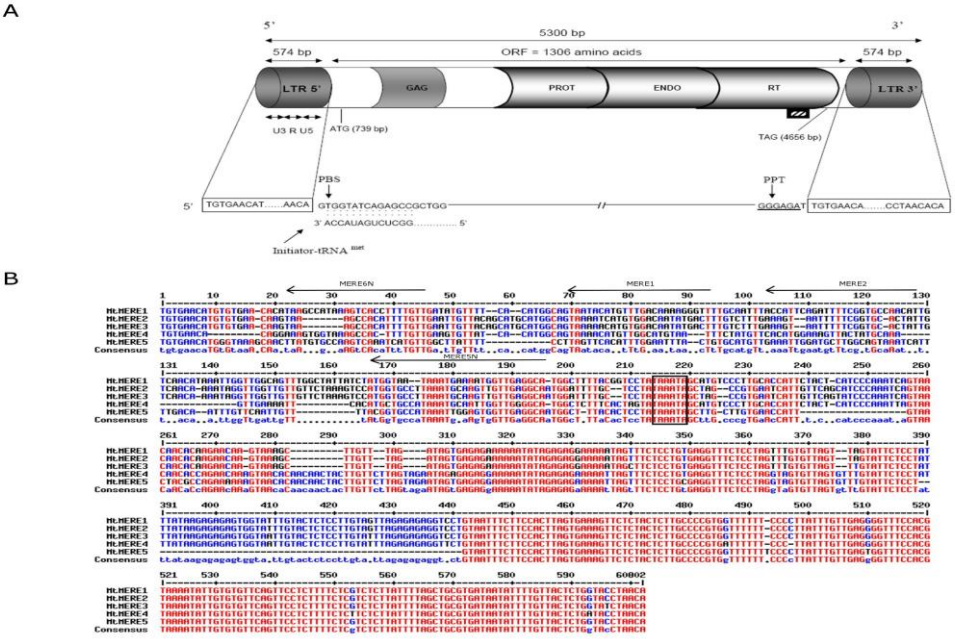
V. Identification of a new MERE1-1, active retroelement in *Medicago truncatula*

The identification of the MERE1-1 retroelement was established in a joint study with the group of Prof. Pascal Rate, and the analyses were performed at the Plant Science Institute at the Center for Scientific Research, Jives-sur-Yvette, France. A *Medicago truncatula* copia-like retroelement called MERE1-1 (Medicago RetroElement1cpoia 1) has been found as an insert into the *NSP2* gene, one of the genes responsible for the symbiotic nitrogen fixation process. When screening for mutant lines in a collection of T-DNA mutants (Scholte et al. 2002), a non-nodulating mutant (nod-) called (ms219) was detected (Brocard et al. 2006). In this mutant we found that the somaclonal mutation did not correspond to the inclusion of T-DNA in this line. There is no deformation of the root hairs or in the formation of root primordia after inoculation with *Sinorhizobium*

meliloti, indicating that the symbiotic nitrogen fixation process is blocked at a very early stage. Expression of marker genes tipical for beginning of root nodule development (ENOD11: Journet et al. 2001; ENOD12: Journet et al. 1994; MtN6 and Rip1: Cook et al. 1995) is absent (ENOD11 and ENOD12) or is reduced (MtN6 and Rip1) in the mutant after exposure to the Nod factor. Comparison of the ms219 mutant line and other mutant lines genes tpipical for activating the signaling cascade upon Nod factor uptake, nfp (Amor et al. 2003), dmi1, dmi2 and dmi3 (Catoira et al. 2000), showed that these genes were not affected in the ms219 mutant line. The mutation is localized, in a 3 cM region on chromosome 3 bordering BACs AC149303 and AC126782. This region of the genome includes the *NSP2* gene (Kalo et al. 2005). The putative mutated gene was considered to be a candidate gene involved in symbiotic nitrogen fixation, regardless of the different phenotype specific of the nsp2 mutant (branched hairs). Amplification of the NSP2 coding sequence in DNA from isolates from independent ms219 plants using DNA polymerase to amplify large fragments showed that the NSP2 sequence contained an insert of approximately 5.3 kb, which is responsible for the mutation. Subsequent sequencing of this insert identifies the retrotransposon included in bp1062 of the *NSP2* gene coding sequence. The retrotransposon thus identified is named MERE1-1, coming from "MEdicago Retro Element 1 -copy number 1".

Figure 12. Structure of the MERE1-1 retroelement representative of the MERE1 family. (A) Schematic representation of MERE1-1. The gray boxes represent the long end repeats (LTR), divided into three parts: U3 (promoter), R (polyadenylation signal), and U5. The dark gray cylindrical boxes indicate the sequence of regions encoding the characteristic proteins of the retrotransposon. From the N- to C terminal region, the MERE1-1 ORF shows homology to AKs with copies of GAG (structural proteins), PROT (protease involved in GAG protein maturation), ENDO (endonuclease included in host DNA integration), and reverse transcriptase (RT) domains. The putative primer binding site (PBS) and poly purine tract (PPT) required for DNA retrotransposon synthesis, as well as the complementary site between PBS and 3 'ends of leguminous methylated transport RNA (tRNA met). (B) Compare sequences of long LTRs endpoints from MERE1-1 to MERE1-5. The U5 region exhibits a high degree of identity between the 5 sequences, while the U3 region shows considerable variability. The supposed TATA box is outlined. The positions of the different nucleotides used in this study are indicated by arrows above the sequence. MERE6N and 5N represent the

position of the specific oligonucleotides MERE61, 62, 63, 64, 51 and 52 (sequences are described in material and methods).



MERE1-1 belongs to a small family of copia retroelements

The MERE1-1 retroelement identified at the NSP2 gene locus has a length of 5300 bp and two long terminal repeats (LTR) regions of 574 bp (Fig. 12a). This element is enclosed by 5 bp repeats in the host DNA corresponding to target site duplication (TSD) repeat target sites. The open reading frame (ORF) includes 1305 AK, and starts at bp 746 and ends at bp 4660, and after in silico analysis shows high similarity to GAG-POL proteins from class I retrovirus elements. Positioning of reverse transcriptase (RT) after endonuclease (endo) classifies MERE1 as Ty1-copia retrotransposon. BLAST analyses using genome sequence data of *M. truncatula* Jemalong allowed the identification of 5 MERE1 elements as follows MERE1-1 to MERE1-5 corresponding to the Mtr16 family described by Wang et al. (2008). MERE1-1 corresponds to the element inserted into the NSP2 gene. Three of the elements (MERE1-1, MERE1-3 and MERE1-4) have an open reading frame of 1305 AK. In the MERE1-2 sequence, an additional T base at nucleotide 2852 creates a truncated reading frame of 707 AK. The MERE1-5 sequence has a stop codon at position 1950, resulting in an ERA of 425 AK. The three MERE1

elements (MERE1-1, -3, -4) have complete ORF and can be active retroelements and the other two are inactive. Comparison of the sequence of long LTR endpoint repeats in the MERE1 family revealed a highly conserved U5 region, but there were many differences in the U3 promoter region (Fig. 12b). Among the five elements described in MERE1-1, 1-2 and 1-3 are the identity of long end repeats of LTRs. There are 3 bp differences in the two LTRs of MERE1-4, while there are 6 bp differences in the MERE1-5 LTRs.

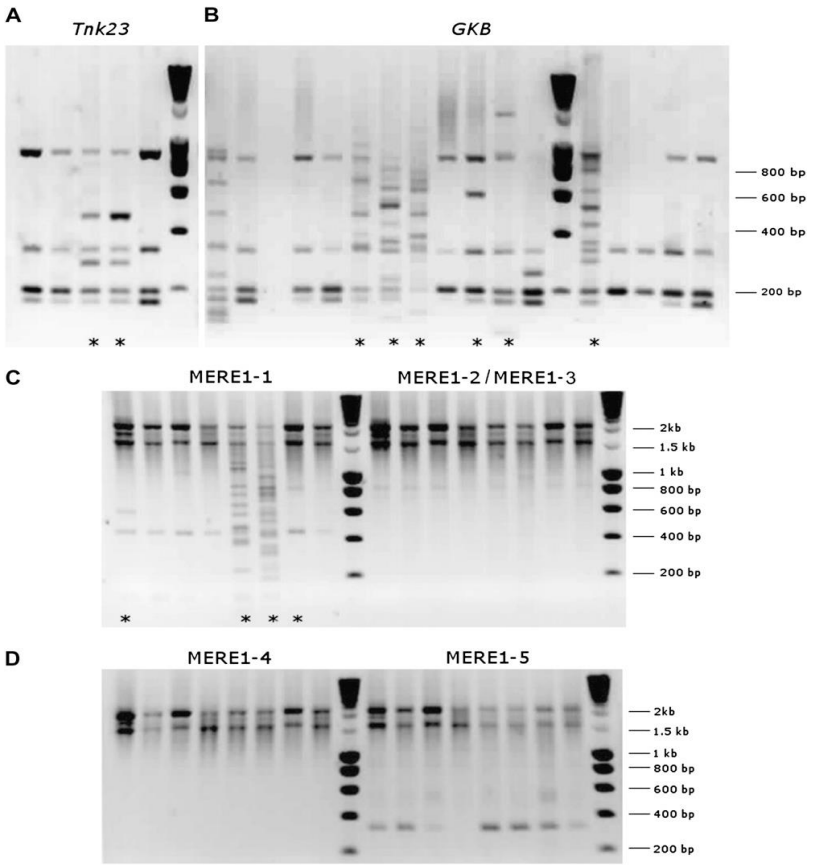
MERE1 is transposed into Medicago truncatula mutant lines regenerated in vitro

The MERE1 retroelement is small in copy number and stable in wild-type *Medicago truncatula* plants, so we hypothesized that its transposition into the NSP2 gene is induced by *in vitro* regeneration of transgenic plants. To exclude that the transposition of the MERE1 element is restricted to the ms219 nod mutant only, we examined mutant lines from different collections generated by *in vitro* cultivation to determine the presence of MERE1 in these lines. The transposition of MERE1 was first examined in two Medicago R108 collections: the *Tnt1* R108 collection (250 Tnk mutant lines, d'Erfurth et al. 2003) and the T-DNA collection (500 GKB mutant lines, Scholte et al. 2002; Brocard et al. 2006). The presence of different MERE1 boundary fragments obtained in this experiment is evidence of transposition into some of the mutant lines examined (Fig. 13 a, b). The summarized results of this experiment confirm the transposition in 73 of the 183 analysed plants or 40% of the plants with new transpositions with 1 to 10 new inserts. MERE1 activity was also monitored in *Tnt1* insertion mutant lines obtained by *in vitro* regeneration of *M.truncatula* cv.Jemalong 2HA. Specific primers for the3 'end of LTR were selected for the various MERE1 representatives, which allowed amplification of the different boundary regions (MERE1-2 and MERE1-3, MERE1-4 could not be distinguished by this technique). In this experiment, the new MERE1-specific fragments obtained from some of the lines examined are an indicator of new transpositions in the genome of the regenerated plants of this collection. New copies were obtained for the MERE1-1 element only, suggesting that only this copy could be reactivated by *in vitro* regeneration. (Fig.13c, d).

The *Medicago truncatula* copia-like retroelement MERE1-1 (Medicago RetroElement1-1) has been identified as an insertion into the symbiotic NSP2 gene. MERE1-1 belongs to a low-copy retrotransposon family in the genome of *Medicago truncatula* represented by MERE1-1; 1-2; 1-3; 1-4; 1-5. These elements are closely related, but only three of them have a complete coding region, and polymorphism is observed between the long end repeats of these copies. This retrotransposon family is represented in all *M. truncatula* ecotypes tested, as well as in the second model plant *Lotus japonicus*. Only MERE1-1 is active during *in vitro* regeneration in both *Medicago*

truncatula genotypes R108 and Jemalong used to construct collections of insertion mutants. In these, MERE1-1 is preferably incorporated into genes.

Figure 13. Transposon display analysis of MERE1 elements in *in vitro* regenerated plants. Total DNA was digested with an AseI restriction enzyme. The additional fragments observed in the lines are marked with asterisks corresponding to new insertion sites of MERE1. Bands larger than 1,500 bp are artifacts of the PCR reaction. A, Tnk23 regenerated plants of the *Tnt1* collection (R108 genotype); 5 of the 23 analyzed plants are shown. PCR amplification was performed with MERE51 + MERE52 / Ase1 and MERE61 / Ase2 primer pairs. B, GKB regenerated plants from the T-DNA collection (R108 genotype); 16 of the 39 analyzed plants are shown. PCR amplification was done by primer pairs MERE51 + MERE52 / Ase1 and MERE61 / Ase2. C and D, Transposition of the various MERE1 elements in regenerated 2HA Jemalong plants. The profile of 8 of the 96 lines analyzed is shown. PCR amplification was done by primer pairs MERE51 + MERE52 / Ase1 for PCR1 followed by MERE61 / Ase2 for MERE1-1, MERE62 / Ase2 for MERE1-2 / MERE1-3, MERE63 / Ase2 for MERE1-4, and MERE64 / Ase2 for MERE1- 5 for PCR2. * mutant lines with new transpositions of MERE1-1.



VI. Forward Genetics. Study of the phenotype of selected *Tnt1* mutant lines

The investigation of part of the *Tnt1* mutant lines from the Functional Genetics legume collection, ABI, is based on the two main "forward and reverse genetics" strategies and begins with the implementation of the Grain legumes and IFOSMO projects. The results of this study have been published in national and international journals with Impact Factor (Iantcheva et al. 2009a; Iantcheva et al. 2009b; Vassileva et al. 2010; Revalska et al. 2011; Iantcheva et al. 2014; Boycheva et al. 2014 ; Revalska et al. 2015; Boycheva et al. 2015; Iantcheva et al. 2015; Iantcheva et al. 2016; Revalska et al. 2016a; Boycheva et al. 2016; Revalska et al. 2016b; Revalska et al. 2017; Iantcheva and Revalska 2018; Revalska and Iantcheva 2018a; Revalska and Iantcheva 2018b, Iantcheva and Revalska 2019) as well as the basis for the initiation of three new projects started in the beginning of 2017 and 2018, funded by the BNSF, MES.

Phenotypic evaluation of selected mutant lines. The collection of insertion mutants in *M. truncatula* is a valuable resource for studying the function of genes in legumes. The establishment of new copies of the *Tnt1* retrotransposon in the recovered lines from the collection was confirmed by PCR-based transposable assay (TD) and reverse two-step (IPCR) methods. The dissertation presents the morphological characterization of 11 lines using the criteria for evaluation of mutant lines (Table 5) from the *M.truncatula* R108 collection, Noble Foundation, presented in the publication of Tadege et al. (2008). Preliminary analyses (TD, IPCR) prove that there are new inclusions of *Tnt1* other than those of starting line A and B. From each tested starting line: e.g. T_0 5106A-1 (where T_0 5106A is the baseline of the collection and 1 is plant 1 planted to produce seeds / 1 to 8 cloned T_0 plants are planted from each seed line /) at random from plant 1 to 8 are selected from 10 seeds of a T_1 or T_2 generation plant (eg T_1 5106A-1 or T_2 5227A-2). The numbering of the mutant lines is identical to that of the ABI collection for the Grain Legumes project, and the numbers are also preserved in the digital collection of sequenced insertion mutants (http://bioinfo4.noble.org/mutant/database) of the Nobel Foundation, with the Bulgarian lines numbering between 5000 and 6000 and including one additional index before the numbers e.g. So5945 / index for Sofia, Bulgaria /. The line numbering is the same in the seed banks in Szeged, Hungary and Dijon, France. The 11 mutant lines selected for phenotypic evaluation are (T_1 5106A-1, T_1 5628A-1, T_1 5644A-1, T_1 5938A-1, T_1 6120A-1, T_1 6123A-2, T_1 6124A-3, T_1 6142A-1, T_1 6142A-3, T_1 6142A-4, T_1 5945A-1, T_1 5945A-2, T_1 5945A-4, T_1 5945A-5, T_2 5048A-1, T_2 5227A-2, T_2 5227A-3). As a control for the characterization of *M. truncatula Tnt1* mutant lines, an 8-copy start line was used, which is one of the baselines for creating the collection (Iantcheva et al. 2009a). Starting line A does not show deviations from the *M. truncatula* 2HA wild-type control line, both morphologically and physiologically.

Table 5. Morphological indicators for evaluation of selected 11 mutant lines

Main phenotypic characteristics in *M. truncatula*
Symbiotic indicators
Not forming nodules - T_1 6123A-2, T_1 6124A-3
Over production of nodules (more than usual)
Not functional nodules (white) - T_1 5106A-1, T_1 5628A-1, T_1 5644A-1, T_1 5938A-1, T_1 6120A-1, T_1 6142A-1, T_2 5048A-1, T_2 5227A-2
Modified tubers (rod-shaped, brown / black, large
Morphology
Highly pronounced dwarf habit- T_1 5628A-1, T_1 6142A-3, T_1 6124A-3
Medium Dwarf Habitus- T_1 5106A-1, T_1 5644A-1, T_1 6120A-1, T_1 6123A-2, T_1 5945A-1, T_1 5945A-2, T_1 5945A-4, T_1 5945A-5 T_1 5945A-1, T_1 5945A-2, T_1 5945A-4, T_1 5945A-5 T_2 5048A-1, T_2 5227A-2
Underdeveloped Dwarf Habitus- T_1 5628A-1p2
Different branching and late flowering- T_1 6142A-1, T_1 6142A-4
Epinastic leaves - T_2 5227A-2, T_2 5227A-3
Multicolored leaves - T_2 5048A-1
Chlorotic leaves - T_1 5644A-1, T_1 5938A-1
Short and compact roots
Deformed roots- T_1 6124A-3, T_2 5048A-1
Long and thin roots- T_1 5106A-1, T_1 5628A-1, T_1 5644A-1, T_1 5938A-1, T_1 6120A-1, T_1 6123A-2, T_1 6142A-1, T_1 6142A-3, T_1 6142A-4, T_1 5945A-1, T_1 5945A-2, T_1 5945A-4, T_1 5945A-5 T_1 5945A-1, T_1 5945A-2, T_1 5945A-4, T_1 5945A-5, T_2 5227A-2
Development and flowering
Single-leafed and truncate peduncles
Multilaflet (4 or 5 leflets) - T_1 6142A-3
Late flowering
Shrubby, late flowering
Shrub-non-flowering (lacking stems)
Early flowering- T_1 5628A-1, T_1 5106A-1, T_2 5048A-1, T_1 6124A-3, T_1 6142A-1, T_1 6142A-3, T_1 6142A-4
Defective peduncles and flowers- T_1 6142A-3, T_1 6123A-2, T_1 5945A-1, T_1 5945A-2, T_1 5945A-4, T_1 5945A-5 T_1 5945A-1, T_1 5945A-2, T_1 5945A-4, T_1 5945A-5
Pigmentation
Purple leaves (more anthocyanin)- T_2 5048A-1
Yellow to light green (slightly anthocyanin) - T_1 6142A-1, T_1 6142A-3, T_1 6142A-4, T_2 5227A-2, T_2 5227A-3

Pale green leaves (light or yellow-green)- T_2 5938A-1-5, T_2 5628A-1-1

Dark green leaves

Albino

The phenotypic characterization of Tnt1 mutant lines provides an opportunity to identify changes in plant morphology and allows a link to be made between the observed phenotype and genes involved in plant development, the nitrogen fixation process or regulatory pathways. This research approach is fundamental to the strategy of 'forward genetics'. In characterizing the phenotype, a great pool of data and photographic material are accumulated, which is a prerequisite for creating a catalog of mutant lines.

The selected *Tnt1* mutant lines were planted in a greenhouse to evaluate plant morphology, growth, flowering and pigmentation for two consecutive life cycles, e.g. of T_1 and subsequent T_2 offspring, the results being illustrated with tables and figures. In the present summary only two figures are presented, illustrating the results of the study of two of the selected lines (Figs. 14, 15). Following the two-stage assessment made under greenhouse conditions, several phenotypic classes are formed related to the morphology of the aboveground part: lines characterized by dwarfism and mid-dwarfism. Almost all lines tested fall into this phenotypic class. Of particular interest is line T_1 6142A-3, T_1 6124A-3 plant 2 and T_1 5628A-1 plant 2, the latter two

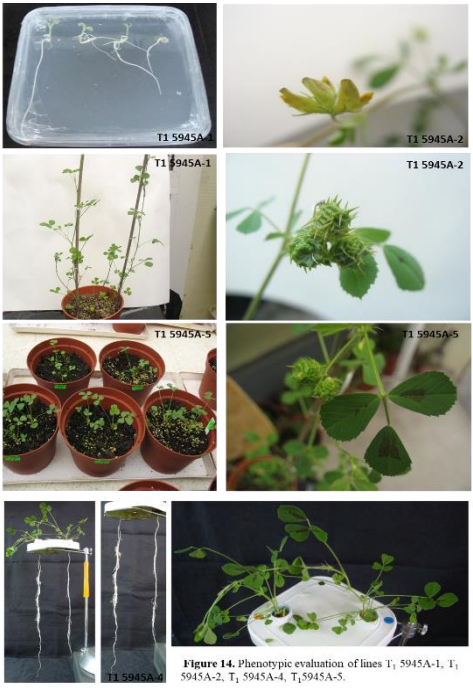


Figure 14. Phenotypic evaluation of lines T_1 5945A-1, T_1 5945A-2, T_1 5945A-4, $T_1$5945A-5.

characterized by a pronounced dwarf and undeveloped dwarf habit. The second phenotypic class consists of lines with altered or defective colors, yellow and yellow-green leaves, multicolored leaves, early flowering, three-leafed and multicolored leaves. Of particular interest is the T_1 6142A-3 line, where deviations are observed in the flowering period, the shape of the flowers, the color and the number of leaves in the tripartite leaf. All plants in this line are sterile. Root growth and morphology as well as the symbiotic profile of the lines studied were evaluated *in vitro*, as well as in aeroponic culture and hydroponic culture. The symbiotic profile of 8 of the studied lines (T_1 5106A-1, T_1 5628A-1, T_1 5644A-1, T_1 5938A-1, T_1 6120A-1, T_1 6142A-1, T_2 5048A-1, T_2 5227A-2) is characterized with white non-functional nodules (Nod +, Fix–), mutants in the ability to fix atmospheric nitrogen. Two of the lines do not form root nodules (Nod-), T_1 6123A-2, T_1 6124A-3. One of these two T_1 6123A-2 lines is susceptible to the bacterium symbiont, rhizobial bacteria enter the host plant, but infectious filaments do not progress and nodule primordia is not observed. Root nodule formation was not observed at line T_1 6124A-3 (Figure 15). Line T_1 5945A-1, T_1 5945A-2, T_1 5945A-4, T_1 5945A-5 form nodules, some of which are functional, that is, capable of fixing atmospheric nitrogen (Figure 14).

Nine of the lines investigated fall into the group of long, thin and unbranched roots. Two of the lines fall into the group with deformed roots. Of particular interest is the T_1 6124A-3 line, which is characterized by a specific root morphology, in which the secondary root branches are located parallel to both sides of the main root resembling fish bone (Fig. 15). Because of its interesting root profile, this mutant line was called "Fish Bone" and DNA isolated from it was sent to Prof. Pascal Rathe's laboratory for sequencing the fragments with *Tnt1* and *MERE1-1* flanking insert sequences.

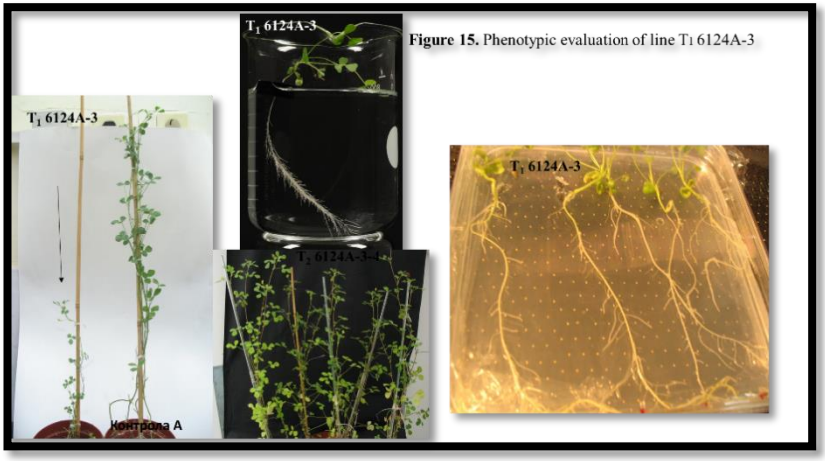


Figure 15. Phenotypic evaluation of line T1 6124A-3

A total 42 fragments of them were sequenced, 32 corresponding to autonomous copies of *Tnt1* and MERE1-1 retroelements with several to several hundred bases of the T_1 6124A-3 genome between the two boundary repeats. From *Tnt1*, 14 sequences representing autonomous copies of the retroelement were obtained, and from MERE1-1, 18 sequences were obtained. One of the A32-18 sequenced fragments is a MERE1-1 insert, the sequence is 494 bp in size and in accessible databases corresponds to the gene encoding WAT1 (Walls are thin 1) related protein (Ranocha et al. 2010), AT5G07050, from *Arabidopsis thaliana*, ortholog of nodulin 21, MtN21 of the genome of *Medicago truncatula*, Medtr8G041390, a membrane-localized transport protein. Plant genes whose expression is induced in legumes by the symbiotic bacterium of the genus Rhizobium during the nodule formation process are referred as nodulins. A study by Pascal Gama et al. since 1996 (Gamas et al. 1996) identified 29 genes whose expression is related to the interaction of the plant with the bacterium, these genes have been designated nodulin (MtN1-MtN29). In the same study, MtN21 was also identified in differential screening of transcripts of root nodules in the presence of the symbiont *Sinorhizobium meliloti*. Nodulin-like proteins have been found in other plant species Arabidopsis, poplar, rice, maize. A large number of 'nodulin' -like genes have been identified in the Arabidopsis genome, 132. For legume 'nodulin' genes are directly related to the early stages of the nitrogen fixation process, such as Enod11 (Svistoonoff et al. 2010), but in other plant species these genes are involved in various aspects of plant development, such as the transport of nutrients, salts, amino acids and hormones. Until now, it is known that an ortholog of *NODULIN 21* from pine is induced by auxin and is involved in the process of secondary root formation (Busov et al. 2004). Available literature studies have suggested that *MtN21* is involved in the nodulation process and is induced by auxin. The established symbiotic profile in aeroponic culture of the T_1 6124A-3 mutant line (Nod-) does not form nodules and has a characteristic phenotype in the number and location of the secondary roots. In addition, an inoculation experiment with *Sinorizobium meliloti* was performed *in vitro*. Observations of root growth 30 days after inoculation confirmed a lack of nodules (Fig. 16 A). These results, sequencing information, and literature data were prerequisites for conducting an experiment to complement the mutant phenotype by inducing hairy roots with *Agrobacterium rhizogenes* carri a vector for overexpression of the gene encoding auxin transmembrane carrier *LAX3*). We hypothesized that overexpression of this auxin carrier would compliment the mutant phenotype and restore nodule formation. For this purpose, "composite plants" were created in which the root part of the mutant line was eliminated, the plant base was infected with *Agrobacterium rhizogenes* carrying an *LAX3* overexpression vector (35S*MtLAX3::GFP*) and "hairy roots" were produced. The appearance of the first non-true roots is observed after about 10 days. The roots obtained are characterized by the presence of many hairs and appear more fluffy. They grow rapidly on a basic nutrient medium, and proving their transgenic nature is easy after

scanning with a fluorescence scanner. Over a 20-day period, these roots grow sufficiently and are infected with the symbiotic bacterium (Figure 16 B). 18-20 days after infection, globular nodule like structures are observed on the roots, which do not form on the mutant line. When viewed with a microscope, a stereoscope shows that these structures begin to form adventitious roots (Figure 16 C, D). The formation of one or two adventitious roots of nodule-like structures is observed. The expression of the auxin transmembrane carrier initiates the formation of nodule-like structures, but these structures have no nodule identity, are not functional and their development continues with the formation of adventitious roots. Similar structures – nodules lost their identity forming roots from their own vascular system are observed in Mtnoot1noot2 double mutants of the *MtNOOT1* and *MtNOOT2* genes, which determine the identity of the formed nodules (Magne et al. 2018).

Genetic approaches are routinely used to study plant-bacterial interaction and root organogenesis. In these studies, mainly non-forming nodules of mutant plants (nod-) or mutant plants forming non-functional nodules (fix-) are generated. Nod mutant lines allow the identification of genes responsible for the first stages of interaction (plant-bacterial recognition) and the onset of nodule organogenesis. In addition to these two main groups of mutant lines, was also identified Medicago mutant named noot (nodule root) that develops roots from the apical part of the root nodules (Couzigou et al. 2012;

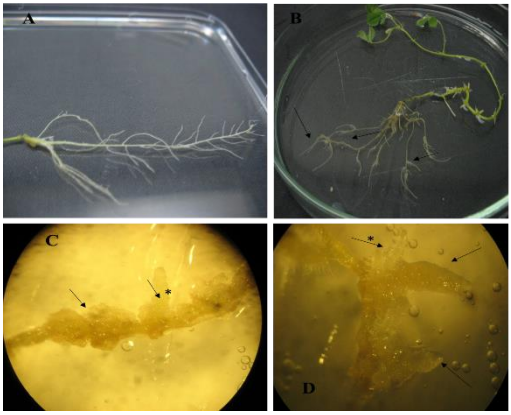


Figure. 16

A- *In vitro* symbiotic profile of line T1 6124A-3 30 after inoculation with *Sinorhizobium meliloti* nodule not observed. **B-**"Composite plant" forming „hairy roots" after transformation with *Agrobacterium rhizogenes* -35S *MtLAX3*::GFP „hairy roots" infected with *Sinorhizobium meliloti* .
C - 18-20 on root a ball like structutres are visible /arrow/ resemble nodules, and adventitious root formation is beginning /arrow with asterix/.
D –Formation of new root from nodule like structure /arrow with asterix and growing of firstly formed roots /arrows/.

Couzigou et al. 2013). This mutant provides an opportunity to study the identity of root nodules in legumes forming nodules of undetermined type at the molecular level. Changing organ identity occurs at the beginning of the nodule initiation process and is an interesting genetic model for understanding the process of symbiotic nodule formation and origin (Hofer and Ellis, 2014; Magne et al. 2018).

VII. Reverse genetics. Cloning and functional characterization of selected genes

Deciphering the function of plant genes and their role in the genome under study is the subject of functional genomics. Functional genomics studies require mutant line collections and are based on two strategies forward and reverse genetics. The previous section of the results reflects studies related to the phenotype of the mutant lines, and the information obtained from their sequencing suggests the identification of certain genes and the ability to study their function in the genome of the studied plant.

As mentioned above, the flanking sequences enclosing a portion of the *Tnt1* (Flanking Sequence Tag (FST) genomic DNA can be identified by PCR-based transposon display methods and reverse two-step PCR (TD and Inverse PCR). These FSTs are stored in the database of the Nobel Foundation's Samuel Roberts Noble Foundation (http://bioinfo4.noble.org/mutant/). A small portion of the *Tnt1* mutant lines created in ABI is also part of this database, and the sequences are available at http://bioinfo4.noble.org/mutant. Within the two projects coordinated by the Agrobioinstitute and funded by the BNSF in the 2008 competitions, DO-268- IFOSMO and DO 105 Center for Sustainable Development of Plant and Animal Genomics under Project 7, part of the scientific program's tasks include cloning and study the functions of genes involved in plant development in legumes. Following analysis of the sequenced mutant lines, the genes of interest and purpose of the planned study were selected (Iantcheva et al. 2009b, Revalska et al. 2011). In three of the mutant lines (So6142A, So5945A, So5828A), *Tnt1* incorporation was found in the exons of genes. The BLAST analysis performed in the specialized database for comparative genomics PLAZA3,0 https://bioinformatics.psb.ugent.be revealed that the sequences correspond to the genes MT5G085330, 1 aminocyclopropane-1 carboxylate oxidasa, (line So6142A which in 2009 according to the NCBI (http://www.ncbi.nlm.nih.gov/) was annotated as 2OG-Fe (II) oxygenase), MT7G105660, Histone acetyltransferase, transcription cofactor, (So5945A line, insert 5), MT2G007220, F-box plant like protein previously annotated as, Cyclin-like F-box protein, (So5945A line, insert 7) and MT7G078300, Disease resistant protein (CC-NBS-LRR class), (So5828A line, insert 1). The other three genes (MT3G072870, Auxin influx carrier protein, LAX3, MT5G040880, Auxin response factor, MT2G026250 GRAS family transcription factor) partially correspond to insert 5 on line So5860A, insert 3 on line So5928A, and insert 6 on line So5945A. In the sequenced line So5823A, insert 9 corresponds to a gene encoding a zinc finger CCHC type protein, NCBI AC140549. These genes are selected for cloning and study of their function because they participate in major processes of plant development and signaling cascades, e.g. secondary rooting, seed maturation, symbiotic nitrogen fixation and somatic embryogenesis, response to environmental stress, pathogen resistance, ethylene synthesis and degradation.

The eight genes listed above were cloned and examined to determine their function in the plant development of model legumes *Medicago truncatula* and *Lotus japonicus*, as well as the model plant *Arabidopsis thaliana*.

The dissertation describes very briefly the studied genes, their established functions in the conducted studies, most of which have been published or are in the process of research in the current projects of the Functional Genetics-legume group. A detailed comparative study of the gene encoding the F-box protein of model plants *M.truncatula* and *A.thaliana* is presented in this dissetation.

Participation of the studied genes in the process of somatic embryogenesis

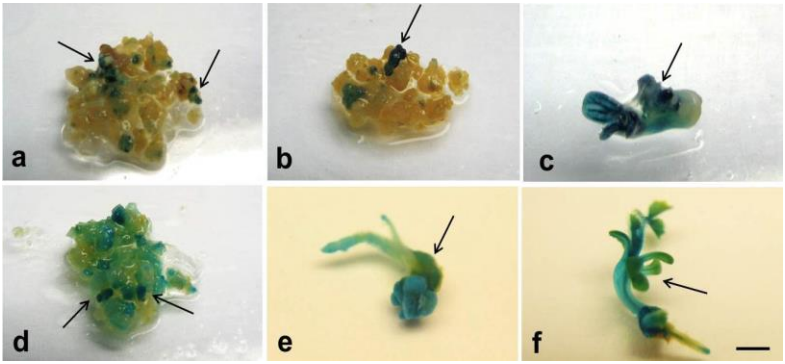
The long experience that the Legume Functional Genetics Group has in the studies of somatic embryogenesis is a prerequisite the investigated genes to be studied in the induction and development of somatic embryos through indirect somatic embryogenesis. A recent transcriptomic study of zygotic embryogenesis in *M.truncatula* sheds light on genes involved in the development of the zygote embryo, as well as seed maturation, and determines the crucial role of some transcription factors in seed maturation (Kurdyukov. 2014). In somatic embryogenesis study on cellular and molecular mechanisms of the process was recently published by Rose 2019. Our investigated genes have not been studied for somatic embryogenesis by other authors, and some of the initial results accumulated in the course of experimental work are presented in our publications (Iantcheva and Revalska 2018; Boycheva et al. 2015; Boycheva et al., 2016). These results are a precondition for future studies to trace the functional linkage of the genes examined in the context of the process, as well as other processes occurring during cellular dedifferentiation and subsequent differentiation. The experiments performed to investigate the role of the studied genes in the process of indirect somatic embryogenesis were performed using transgenic plants for localization of gene expression by fusion of the *GUS* and *GFP* marker genes under the control of the promoters of the studied genes by MtLAX3:: GUS-GFP; pMtARFB3::GUS-GFP, pMtF-box::GUS-GFP; pMtHAC::GUS-GFP; pMtGRAS::GUS-GFP, pMtDRP::GUS-GFP).

As already noted in our studies, we confirmed the extremely important role of auxin 2,4-D in the process of somatic embryogenesis. Most of the auxin enters the cells by polar transport through the cell membrane using transmembrane transporters. Activation of the auxin signaling cascade permits the participation of two of the studied *MtLAX3* and *MtARFB3* genes in the process of somatic embryogenesis. The transgenic plants obtained expressing the *GUS* and *GFP* marker genes under the control of promoters of these two genes, as well as their offspring, were used for histochemical analyzes to track the expression of the *GUS* marker gene in the process of indirect

somatic embryogenesis in *M. truncatula*. The activity of the *GUS* marker gene under the control of the LAX3 promoter was observed as points in the induction of callus tissue by explant, but gradually the signal strength increased in the formed globular embryos and became intense in embryos in the torpedo stage (Fig. 17 a, b). The activity of the marker gene was also reported at a later stage than the development of the somatic embryo - the cotyledon stage, with the signal being observed throughout the bipolar structure (Fig. 17c). The role of auxin transmembrane carriers responsible for the entry of auxin into the cell is well documented, as is their combined action with transporters for the removal of auxin from the cell. These two types of transporters are major players in auxin transport, cell division and auxin distribution during the stages of zygotic embryogenesis (Bassuner et al. 2007; Firml et al. 2002a; Jenic and Barton 2005). The combined action of membrane transporters, as well as the proven important role of auxin in the process of plant development (Saini et al. 2013; Davis 2010; Quint et al. 2006) allows us to conclude that the *MtLAX3* gene is involved in the process of somatic embryogenesis.

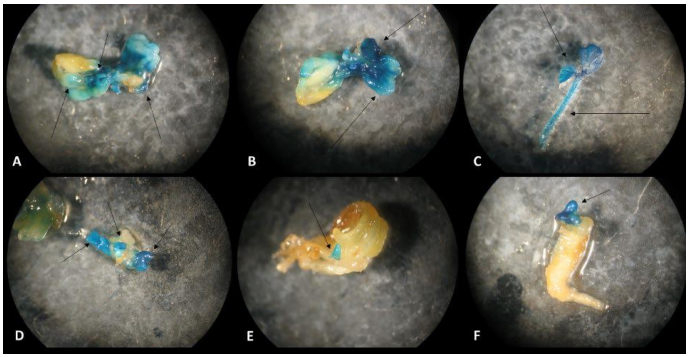
The involvement of auxin responsible factors in the process of zygotic embryogenesis was demonstrated in the study by Kurdyukov et al. (2014). Our analyzes to monitor the activity of the marker *GUS* gene under the control of a promoter of the MtARF-B3 gene tested confirmed its involvement in the process of indirect somatic embryogenesis in *M. truncatula*. A weak signal was initially observed upon the induction of callus tissue by explant, which became more intense in the formed globular embryos (Fig. 17d) and their subsequent development to the torpedo stage and the early and late cotyledon stage (Fig. 17 e). Localized expression of the marker gene is less pronounced in the formed small plants, but is observed throughout the well-shaped bipolar structure (Fig. 17 f).

Figure 17. The ISE process in *M.truncatula* visualized by expression of the *GUS* marker gene under the control of endogenous promoters of *MtLAX* and *MtARF* genes pLAX3::GUS-GFP (a, b, c) and pARF :: GUS-GFP (d, e , f).



Expression of the *GUS* marker gene under the control of promoters of the examined *MtF-box* and *MtHAC1* genes was also monitored in the course of indirect somatic embryogenesis. Under the control of the pMtF-box promoter, signal from the marker gene is observed in the formed embryogenic zones as well as in the formed globular and torpedo embryos (Fig. 18A). The activity of the marker gene is reported in the formed cotyledon leaves, as the signal intensity increases, and is even more pronounced in the leaves and petioles of the formed small plants (Fig. 18 B, C). Expression of the marker gene under the control of the pMtHAC1 promoter is localized exclusively in the formed embryos in the heart, torpedo stage, the signal is not very intense (Fig. 18 D, E) and in cotyledonary embryos with a well-formed bi-polar structure is localized in the cotyledonary leaves (Fig. 18 F).

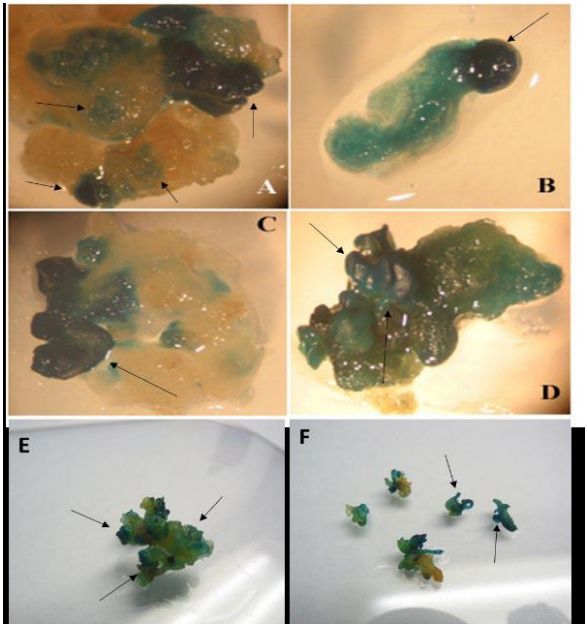
Figure 18. The ISE process in *M.truncatula* visualized by expression of the *GUS* marker gene under the control of endogenous promoters of the *MtF-box* and *MtHAC1* genes pF-box::GUS-GFP (A, B, C) and pHAC1::GUS-GFP (D, E, F).



The GRAS family of transcription factors is numerous and represented by many members in different plant species. In the *M.trucatula* model plant, 68 members belong to this family grouped into 16 groups (Song et al. 2017). These genes are involved in various processes of plant development, and are associated with plant response to abiotic stress. There is no literature data related to their involvement in the somatic embryogenesis process. The involvement of the investigated gene encoding the GRAS transcription factor in the process of indirect embryogenesis was documented in all observed stages by monitoring the expression of the marker *GUS* gene under the control of the gene promoter. The signal is reported in the callus formed, becoming more intense in the induced embryogenic zones (Fig. 19 A, B) and the subsequently formed heart and torpedo embryos (Fig. 19 C, D). Although disease resistance genes are primarily associated with this function, expression of the test gene (encoding Disease resistant protein CC-NBS-LRR), traced through the marker *GUS* gene, has also been observed

in various stages of somatic embryogenesis. The signal is observed in the formed embryogenic zones and becomes more intense in the forming structures, and is observed in well-separated bipolar cotyledonary embryos Fig.19 (E, F). The functions of the genes encoding Disease resistant protein possessing NBS-LRRs domains are associated with participation in other processes of plant development besides responding to pathogenic attacks.

Figure 19. The IEE process in M.truncatula visualized by expression of the GUS marker gene under the control of endogenous promoters of MtGRAS and MtDRP genes pGRAS :: GUS-GFP (A, B, C, D) and pDRP :: GUS-GFP (E , F).



The histochemical analyses performed to track the expression of the marker *GUS* gene under the control of the promoters of six of the studied genes in the process of indirect somatic embryogenesis shed light on their function in the process, but also raised many questions that could be addressed through in-depth studies and multilateral analyses.

Transcriptional profiles of the *MtF-box* and *AtF-box* genes in model transgenic plants with modified expression

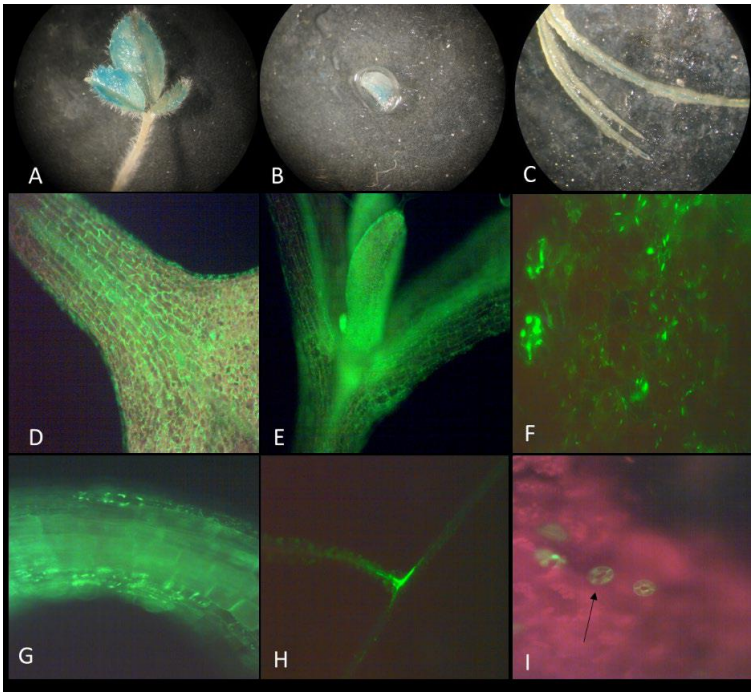
A very important step in elucidating the function of the genes studied is the production of stable transgenic plants with modified expression. The resulting T_0 transgenic plants of *M.truncatula* were screened initially for the presence of the *nptII* selective gene by performing PCR assays. Proven transgenic plants are planted under greenhouse

conditions to produce offspring. Determination of the transcriptional profile is crucial in conducting subsequent functional analyses. Three to five independent transgenic T_1 lines with altered MtF-box transcript levels are selected and used for subsequent phenotypic evaluation, as well as five independent transgenic T_1 lines carrying the MtF-box promoter fused to the *GUS* and *GFP* marker genes (MtF-box::GUS/GFP) were selected to localize gene expression. Transcript levels in transgenic T_1 OE lines were slightly increased, whereas transcript levels in knockout plants showed a significant decrease in expression compared to controls. To determine the expression level of AtF-box-OE, the seeds of eight homozygous T_3 lines were germinated and the seed were grown in a selective medium containing kanamycin. Leaves were separated from the *in vitro* plants for subsequent transcriptional analysis and plants showing a change in expression profile were planted in the greenhouse. Three of the eight lines examined showed a more pronounced level of transcript increase of 57/17, 58/14 and 60/17. These lines were used in subsequent phenotypic and functional analyses. Similarly, inactivated expression of AtF-box-amiRNA lines was made. Of these, three lines with the lowest transcript level 3/4;3/3; 1/1were selected.

Localization of *F-box* gene expression in transgenic plants of the two model species at different stages of plant development

In the offspring of transgenic plants of *M. truncatula*, expression of the *GUS* marker gene under the control of the endogenous F-box promoter (pF-box::GUS:GFP) has been detected, except in the various stages of the somatic embryogenesis process and in young leaves, primary roots and secondary root branches (Fig. 20 A, B, C). Green fluorescence signaling has been observed in trichomes of young *in vitro* leaves (Boycheva et al. 2015). The expression of the marker gene in the nitrogen fixing nodules has not been established. It is possible for the marker gene to be expressed at very early stages in the nodule formation process when primordia of the forming structures are initiated. These early stages of the process have not been monitored, but no signal is localized in mature nodules. In contrast to *M.truncatula*, in the other model legume *L.japonicus*, expression of the marker gene was detected in the determined mature root nodules (Boycheva et al. 2015). In *A. thaliana*, expression of the green fluorescence marker gene (*GFP*) fused with the coding sequence of the tested *At-F-box* gene was monitored. Observations made with a fluorescence microscope visualize the signal in the epidermal cells of the leaf, the apex of the plant, as well as in the trichomes (Fig. 20 D, E). At higher magnification, expression was detected at the subcellular level, with the signal observed in chloroplasts of leaf parenchyma cells (Fig. 20 F), secondary root branches in the endoplasmic reticulum and plastids of root epidermal cells, as well as in the stomata of the leaves (Figure 20 G, H, I).

Figure 20. Expression of the marker *GUS* gene under the control of an endogenous pMtF-box::GUS-GFP promoter from *M.truncatula* (A, B, C); expression of the marker *GFP* gene fused to the At35S-F-box::GFP gene coding sequence from *A.thaliana* (D, E, F, G, H, I).



The results obtained in both model plants relate the expression of marker genes to young organs with intensively dividing cells, which proves the involvement of the gene in plant development, and subcellular localization also implies its likely involvement in metabolic pathways. The results are in support of previously published studies on *A. thaliana* that detail *in silico* expression of Arabidopsis SKP proteins and various F-box (FBX) proteins in almost all plant organs including cotyledons, apical stem meristem (SAM), inflorescence, flower, leaves, roots, root apex, root epidermis, pericycle, apex, pedicel, pollen, anther, hypocotyl, stele, stalk, protoplast, zygous embryo suspension, rosette and young leaves (Kuroda et al. 2012). Studies on the AFB4 gene encoding another F-box protein have shown its major role in plant growth and development. afb4 mutants have a pleiotropic phenotype characterized by severely reduced primary root growth, shorter lateral roots, very short hypocotyl, greater leaf thickness, and a significant delay in flowering time (Hu et al. 2012). The results obtained are in agreement with the published and confirm the expression of the investigated gene in

plant organs and tissues with actively dividing cells and at different stages of somatic embryogenesis. The role of the studied gene /F-box protein/ in plant growth and development has been confirmed in both model species.

Phenotypic evaluation of transgenic lines with modified F-box expression in M.truncatula and A.thaliana

In the overexpressing transgenic lines of *M.truncatula*, apart from the difference in the size of the tripartite leaves of one plant, the specific deviation observed in these lines is the positioning of the leaves in the tripartite leaf. The top leaf or either side is located at almost right angles to the leaf petula, a deviation which was not observed in the control. The most common abnormalities observed in transgenic *M. truncatula* lines with F-box (OE) overexpression are in the shape, size, jaggedness of the leaflets of the tripartite leaf, sometimes leading to the rudimentality of the individual leavlets compared to the control. In many cases, the trichomes are located on the edge of the leaf and are much smaller on the leaf surface. The leaves of transgenic lines with inactivated expression of *in vitro M. truncatula* plants are characterized by an oval-shaped leaf and varying toothed leaves, but almost indistinguishable from the control. Electron microscopic analyses revealed the presence of single cells longer than those of controls and overexpression lines. Both types of modified expression lines have highly open stomata. The test of the transgenic lines in the hydroponic system showed no significant deviations from the control. In over-expressing *M. truncatula* lines, a less developed and loose root system was observed compared to the control. Separate, single nodules along the root length were reported, and their number was comparable to that observed in the control. The aerial part of the transgenic and control lines is comparable in size, with variation in leaf size observed. The root system of the plants with suppressed expression is also less developed and with less branching than the control. The number of nodules scattered along the root is reduced. The aboveground part of the plants is comparable to that of the control and no deviation in leaf shape is observed. The transgenic and control plants of *A. thaliana* were grown *in vitro* for a period of 15 days, after which they were transferred to the soil under greenhouse conditions. At-F-box-OE overexpressing plants are characterized by a variety of leaf-shaped and large leaves in the rosette with few long trichomes. The same long trichomes are also observed on the stem, which is not branched. The phenotypic evaluation of the flower does not show deviations neither in number of flower elements nor in shape. Microscopic observations of the discolored leaves revealed atypical cells in size in lines 5814 and 6017, as well as undeveloped trichomes.

Plants of At-F-box-amiRNA with inactivated expression of the At-F-box gene are characterized by many elongated and almost uniformly shaped leaves in the rosette. The stem is thick covered with short trichomes. No deviations were observed in the number

and shape of the flower parts. The elongated shape of the leaves in the rosette provokes us to perform microscopic observations of discolored leaves. As a result of these observations, giant epidermal cells were found along the length of the leaves. Most likely the elongated shape of the leaves is the result of the presence of this type of cells that are not observed in the control and overexpressing lines. These results are also confirmed by the flow cytometric analyzes presented below.

Morphometric analyzes of transgenic lines with modified expression of At-F-box OE and amiRNA from Arabidopsis thaliana

Morphometric analyses include measurements of leaf area, number and size of epidermal cells, DNA histograms of leaf (flow cytometry) at OE and amiRNA lines from *A. thaliana*. The morphometric measurements of leaf epidermis from F-box OE and amiRNA lines were made with 22 daily *in vitro* leaves, sampled from the first pair of true leaves, one leaf discolored and fixed to capture leaf area parameters, number and size of the epidermal cells, and the opposite leaf of the pair was used as the material for flow cytometric analysis. This method of sampling ensures maximum reliability of the results obtained in both types of analysis. The parameters of the leaf blade were removed from a leaf of the first pair of true leaves and the ploidy level of the cells in the opposite leaf of the same age and floor was determined. The results reported show comparable leaf area on amiRNA lines and overexpression line (6017) and larger than the control. The size of the epidermal cells in one of the tested amiRNA lines (line 1-1) is slightly larger than that of the overexpression line and control and is significantly larger than the size of the other two lines 3-3 and 3-4. The reported average number of epidermal cells was slightly higher for amiRNA lines (line 1-1) compared to the overexpression line and control, and the number of epidermal cells for the other two amiRNA lines was significantly greater than for line 1-1, overexpressing line and control. The results of leaf area measurements, epidermal cell size and the smaller number of epidermal cells found at line / 1-1 / compared to the other two inactivated expression lines correspond to the size of the leaves and giant cells along the same line under greenhouse conditions described above.

In parallel with the morphometric assays made on the leaves of transgenic lines with modified expression of AtF-box OE and RNAi and control, flow cytometric analysis was also performed to determine the ploidy of cells in the levels from the first true leaves of 22-day *in vitro* plants. Compared to the control, leaves from the OE lines showed significantly higher ploidy of the cells of levels (4C), with a significant difference ($p \leq 0.05$). In amiRNAi lines, ploidy levels are characterized by a higher percentage of 8C cells and almost double the number of 16C cells compared to those of overexpressing lines and controls ($p \leq 0.05$), and this is particularly clear in line 1-1 ($p \leq 0.01$).

The results of the morphometric measurements of the leaf epidermis cells and the flow cytometric assay suggest that endoreplication processes are more clearly expressed in the cells on the modified expression lines (OE, amiRNA), which is more pronounced in the amiRNA lines. Similar results having a direct effect on the number and size of epidermal cells as well as the level of ploidy in transgenic plants with modified expression of *A.thaliana* were obtained and described in our published results (Boycheva et al. 2015), where the MtF-box is heterologously expressed in the model reference plant.

The biological meaning of endoreplication is the presence of DNA replication without subsequent mitosis and cytokinesis during the cell cycle, resulting in increased levels of ploidy (Massonnet et al. 2011). Increasing the ploidy cell level without subsequent division is closely related to the size and number of epidermal cells as well as to the leaf area. A key mechanism of the plant development process is the right balance between cell division and cell size (Hemerly et al. 1995). Control of the tissue-specific size of the cells takes place during the cell cycle (Francis 1998). When the cell reaches the optimum size, a cell division process is initiated. Presumably, the presence of high ploidy cells in the transgenic lines studied is caused by the impaired function of the investigated gene encoding the F-box protein involved in the SCF complex related to protein degradation, which affects the cell cycle, preventing mitosis and subsequent cellular differentiation (Ishida et al., 2010). Based on the results obtained, we hypothesize that the gene we study encoding the F-box protein component of the SCF complex may be involved in the degradation of some of the major genes involved in the cell cycle.

Dynamics of root growth observed in transgenic lines with modified expression of *A.thaliana*

The main root length was increased by five-day seedlings, initially recording the primary root length and increasing at 24, 48, 72, 96, and 120 hours. The reported initial root length was the smallest in the lines with inactivated gene expression compared to the control and overexpressing lines. A statistically significant difference ($p \leq 0.0001$) was observed between the tested amiRNA lines / 1-1; 3-3; 3-4; 5-4 / and controls, as well as ($p \leq 0.05$) at OE line 6017 and controls. A statistically significant difference of ($p \leq 0.05$) for the first 24 hours) in root growth of amiRNA lines and the OE line 5814 compared to the control was maintained up to 48 hours ($p \leq 0.001$) and 72 hours at which it ranged from $p \leq 0.05$ to $p \leq 0.001$. The reported difference in root growth at 96 and 120 h is available for amiRNA lines 3-4 and 5-4 and OE lines 5717 and 6017 relative to control and ranges from ($p \leq 0.05$ to $p \leq 0.005$) .

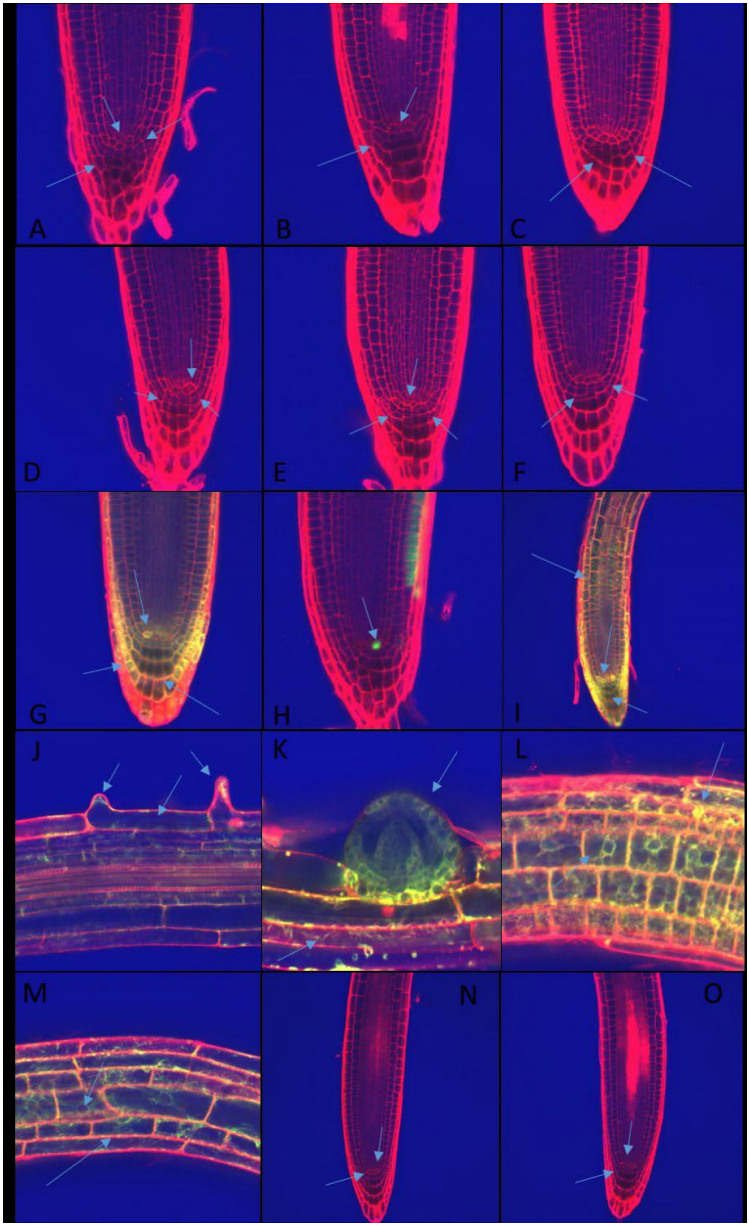
The results of the study provoked us to conduct a detailed microscopic observation with a confocal microscope of 5 day primary roots in *A.thaliana* lines with inactivated expression and those with overexpression compared with conotrol.

Root growth and development in plants is supported by a group of cells called stem cells that form a stem cell niche SCN in the root apical meristem and divide asymmetrically. These stem cells maintain a balance between cell division and differentiation and give rise to different tissues in the root, thereby ensuring its continued growth (Bennett and Scheres, 2010). At the center of the stem cell niche are 4 to 8 cells that form the so-called quiescent center (QC) silent center, which is responsible for the fate of the adjacent stem cells (DSC), (van den Berg et al 1997). The PLETHORA (PLT) and SCARECROW (SCR) transcription factors are included in the QC specification (Aida et al. 2004; Sabatini et al. 2003), and the WOX5 transcription factor (WUSCHEL RELATED HOMEOBOX 5) expressed in QC is responsible for maintaining a non-differentiated state surrounding stem cells. The QC located at the center of the SCN maintains indeterminate root growth, preventing differentiation of surrounding stem cells. Both auxin and a group of transcription factors, including PLTs, SCRs and WOX5, are important for the QC specification and activity of stem cells.

In lines with inactivated expression / 5-4; 1-1; 3-3; 3-4 / the most commonly observed abnormalities are characterized by the variable number and shape of QC cells. These cells are (4-8 in number, most commonly 4) in Arabidopsis are of equal size undifferentiated and at rest. On amiRNA lines, QC cells and distant stem cell (DSC) are irregularly shaped, in some cases chaotically arranged and subjected to multiple disorganized asymmetric divisions (Fig. 21 A, B, C, D, E). Microscopic observations confirmed that in one of the overexpression lines 5814, the GFP signal (fused with the gene coding sequence) was localized in one of the QC cells, and at line 6017 and in the four cells of the QC (G, H, I), similarly of the transcription factor described in Arabidopsis WUSCHEL RELATED HOMEOBOX 5 (WOX5). This transcription factor is specifically expressed in QC cells and is associated with the establishment of a maximal auxin gradient that directs root growth and development by modulating cell division and growth, as well as stem cell differentiation (Tian et al. 2013). Expression of the GFP marker gene in the overexpression lines is also observed in the root hairs forming, secondary roots and their growth, in the endoplasmic reticulum and the plastids of the cells of the transitional zone, located immediately after the meristem and in the region of extension 21. K, L, M). The control observed 4 QC cells that were in the correct shape, as well as the arrangement and distribution of DSC. The slower growth of the primary root in the lines with inactivated expression is likely to be observed, as well as its slower growth observed in the experiment with root growth due to the atypical shape and cellular activity of the QC and DSC cells characterized by chaotic asymmetric

divisions. In the overexpression lines, the cells of QC and DSC are of the correct shape and arrangement, with a few exceptions at line 5717, where cells of different shape at DSC are observed (Fig. 21 F). The expression of the marker gene in QC cells is surprising, at lines 5814 and 6017. The expression marker in QC cells is reported (WOX5) in the literature transcription factor at Arabidopsis (Tian et al. 2013). Recent studies by de Luis Balanguer et al. from 2017, indicate that in addition to WOX5, PERIANTIA (PAN) is another regulator of QC functions. Like WOX5, PAN controls the maintenance of stem cells f columella and divisions of QC. In contrast to WOX5, which limits cell division in QC, PAN induces division in QC. Made by de Luis Balanguer et al. transcriptomic analysis of Arabidopsis root stem cells identified 1625 genes, 2/3 of which are responsible for the ubiquitination and degradation process, as well as 201 transcription factors (de Luis Balanguer et al. 2017). In agreement with these results, the observed expression of the GFP marker gene in QC cells is explicable. In this aspect, the studied AtF-box is of interest, but further comparative studies are needed related to the abovementioned genes fused to different fluorescent marker genes, as well as the presence of double mutant lines.

Figure 21. Microscopic observations with a confocal microscope of 5 day primary roots of modified expression / amiRNA lines - 5-4; 3-6; 3-4 and OE lines - 5717; 5814; 6017 SALK line and control. The roots of the amiRNA, and the OE lines, and the control were labeled with propidium iodite (PI) and the OE lines; 5717;5814; 6017 GFP marker gene is fused with the coding sequence. A- amiRNA line 5-4; B- amiRNA line 1-1; With amiRNA line 3-3a; D- amiRNA line 3-4; E- amiRNA line 3-3-b. Arrows mark unequal numbers, shape and size of QC cells, as well as DSC, with chaotic asymmetric divisions. F-OE lines 5717 arrows indicate asymmetric divisions in DSC; G-line 5814, GFP signal is observed in one of the cells of QC and columella, the signal is colored in yellow / fusion of the fluorescence signal by GFP and PI; H-line 5814, GFP signal is observed in one of the cells of QC, both fluorescent signals are filtered; I-line 6017 GFP signal is observed in four QC cells, cells of the endodermis, cortex and central cylinder (the two fluorescence signals are fused; yellow branch indicates strong fluorescence); J-line 6017 GFP expression in root hairs forming, strong signal at tip of hairs and endoplasmic reticulum; K-line 6017 Secondary root formation GFP expression is strongly expressed at the base of the nasal root primordia and a weaker signal in the endoplasmic reticulum and plastids; L-line 6017 transition zone / GFP expression is more pronounced in the plastids and less in the endoplasmic reticulum; M-line 6017 elongation zone / GFP expression in plastids and endoplasmic reticulum; N-control; O-control. Arrows indicate QC cells.


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O

Observed deviations in the number and shape of cells in the quiescent center (QC) zone, irregular cells and chaotic asymmetric cell divisions from distal stem cells (DSC) in *A.thaliana* transgenic lines with modified expression, as well as variations in size in shape and dentition of leaves in transgenic lines with modified expression in *M.truncatula* and *A.thaliana* provoked us to study a link between the gene, we tested, during the cell cycle in both model plants. Cell cycle research is focused on answering the question of how cell division integrates into the process of plant growth and development. An important aspect of research is knowledge of the mechanisms that drive the transition to mitosis, and especially the pre-phase of the cell cycle preceding the mitosis G2-check point. Much knowledge has been accumulated over the last 15 years on cell cycle transcriptional control, and in particular on the G2 check point, but there is much ambiguity regarding the post-transcriptional level of regulation. In this regard, our studies aim to identify the role of the F-box protein encoding gene we have studied and its involvement in the cell cycle. In this respect, the tested gene of *Medicago truncatula* and its homologue of *Arabidopsis thaliana* have not been studied before. Our pioneering studies already published are related to the localization of gene expression of Mt-F-box by monitoring the expression of the GUS and GFP marker genes in the three model plants of *M.truncatula, L.japonicus A. thaliana*. One of our goals in the study of this gene is to determine the activation time of its expression in the cell cycle in controls and transgenic lines with modified expression of *M. truncatula* and *A. thaliana*. This type of research is related to the presence of a method for synchronizing the root meristem or other type of actively dividing cells when using a replication inhibitor. Such a system has already been developed in our previous studies with the model plant A. *thaliana* (Cools et al. 2010). One of the tasks during the EMBO specialization of Assoc. Prof. Anelia Iantcheva was to establish a system for synchronization of root meristem cells in the model legume plant *M. truncatula*. The results of this study were published (Iantcheva et al. 2015), and the system has been successfully used to detect activation of Mt-F-box gene expression in the cell cycle of the model plant.

The cell cycle has two main phases, the S phase, during which DNA is synthesized and the M phase or mitosis, through which the chromosomes are distributed equally in the daughter cells. Two control points G1 and G2 preceding the S and M phases are also identified. These two transitions G1/S and G2/M to the main phases are characterized as control points/check points, phases for switching to division, differentiation, programmed cell death or rest. The phases and transition points are controlled by serine-threonine kinases, which are activated by regulatory protein cyclins, which is why they are called cell cycle-controlling cyclin-dependent kinases (CDKs). The expression profile of these CDKs is constant throughout the cell cycle. The periodic expression of cyclins provides specific synchronization of the activity of CDKs during the phases.

Functional studies related to the determination of the activity of the studied genes are based on the presence of synchronized cultures. The development of a root tip synchronization system by the application of the toxic toxic agent hydroxyurea provides an opportunity to study genes under endogenously stable conditions and under unified meristem cells. The efficacy of such a system available in the *Arabidopsis thaliana* model plant has been confirmed by the use of modern expression assays (Cools et al. 2011). In most cases, cell culture synchronization is based on the depletion of important nutrients or through the use of inhibitors known as replication inhibitors, such as aphidicolin or hydroxyurea. Application of HU is based on antiproliferative ability leading to the inhibition of small subunits of ribonucleotide reductase, an enzyme responsible for maintaining the required amount of nucleotides (dNTPs) in dividing cells (Saban and Bujak, 2009). Application of this genotoxic agent reduces the concentration of available nucleotides required for DNA replication and causes cell cycle arrest (Wang and Liu, 2006).

The development of a root tip synchronization system for the *M.truncatula* model plant requires preliminary analyzes related to establishing the parameters necessary for the preparation of the methodology. The use of the hydroxyurea (HU) replication inhibitor requires a pre-dose determination non-toxic to plants. For *Medicago truncatula* root meristem cells synchronized system the determined concentration HU is 1mM. For detailed monitoring of changes in cell cycle kinetics, flow cytometric analyzes were performed on root tips of seedlings treated with 1 mM HU for different times (0; 6; 10; 14; 18; 22; 26; 30 hours). A flow cytometric study of *Medicago truncatula* root tips under conditions of treatment with the HU replication inhibitor shows that it is characterized by a prolonged S phase beginning from the 6th to the 14th and second entry into the S phase after the 30th hour. The period from the 14th to the 22nd hour of HU treatment corresponds to the G2-M transition of the cell cycle. Validation of the presented cell cycle model of hydroxyurea synchronized root tips of *Medicago truncatula* was performed by expression analyzes of specific cell cycle marker genes. The expression profiles of these genes were determined at the following time intervals (0; 4; 6; 10; 14; 18; 22; 26 h). Orthologs of Arabidopsis cell cycle marker genes were determined in *Medicago truncatula* using the PLAZA 2.5 comparative genomics database. Six markers gene for S phase and the G2/M transition were used for validation and their profiles were graphically visualized. The expression levels of the S phase marker genes are shown in Figure 22 A. The transcriptional level of the MtCyclinA3-4 gene increases sharply to the 4th hour of treatment, maintaining a constant level between the 6th and 10th hours. Expression reaches its maximum by the 10th hour followed by a drop in transcript level. The second expression peak, which is typical of the cycline markers for the S phase, is formed at the 18th hour, followed by a sharp decrease and again rises with the start of the next S phase at the 26th hour. The other MtCyclin1

marker gene has the same profile, but its expression level is much lower than that of MtCyclinA3-4. Transcript level change is very slight with a pronounced peak at 6 hours and another less pronounced peak at 26 hours at MtCyclin. The results obtained for the expression profile of the selected marker gene designated as MtCyclinA3-4 (ortholog in Arabidopsis AT5G43080 is AtCyclinA3; 1) is an accurate marker for the S phase in the model plant *M.truncatula*. The presented profiles of *M. truncatula* CyclinA3-4 and MtCyclin1 have expression peaks similar to those obtained for the S phase marker genes found in Arabidopsis in the studies of Cools et al., (2010). The results obtained in the study to create a synchronized root tip system in *Arabidopsis* confirm that HU treatment has an effect on the transcription of histone genes. The transcriptional level of two of the histone genes (MtH2A and MtH4) was also studied during the cell cycle. The MtH2A transcript level increased sharply from 0 to 6 h and formed a peak at the 6th hour and slowly decreased, starting to rise again when cells began to enter the S phase again at the 26th hour (Figure 22 B). The expression of the other histone MtH4 gene shows the same expression profile, but its peak appears at the time point in the 4th hour, two hours earlier than MtH2A, which was also observed in the results obtained for *Arabidopsis* (Cools et al. 2010). Based on the flow cytometry assays received, the 14th to the 22nd hour of treatment responds to the G2-M transition of the cell cycle. The expression profiles of the selected G2 / M marker genes are shown in Figure 22 C. The first of the selected G2/M marker genes CyclinA G2/M specific, forming a small expression peak at 4 hours followed by a decrease in expression and a maximum induction of expression. at the 18th hour of HU treatment, followed by a slow decrease to the 22nd hour and a new increase in the transcript as cells enter a new cell cycle. The expression profiles of the other two selected MtCyclinA2 G2/M specific and MtCDK genes show a similar expression profile with a peak at 14 hours followed by a slow decrease. The results of the G2/M expression profiles obtained for the selected marker genes for *M. truncatula* are in agreement with those observed in *Arabidopsis*. The developed system for synchronization of root tips of *Medicago truncatula* was used to position and determine the putative role of the studied Mt-F-box gene during the cell cycle. The expression profile of the gene binds to the long S phase. The transcript level rises to the 4th hour and maintains a near constant level until the 10th hour and slowly decreases until the G2 / M phase is complete. The transcript level increased again as the cell cycle entered the S phase again (Figure 22 D).

The transcriptional profile of the identified marker gene for the S phase of MtCyclinA3-4 was also established at synchronized root tips of *M. truncatula* OE and RNAi lines. The expression level of the MtCyclinA3-4 S phase marker is comparable to that of the Mt-F-box in modified expression transgenic lines (OE, RNAi) and wild-type plants. (Figure 23). Most of the data published in the literature confirm two E3 ubiquitin ligase complexes (SCF and Anaphase-Promoting Complex/Cyclosome (APC/C) complex),

which are responsible for protein degradation during the cell cycle and bind to G1/S and G2/ M transitions (Marrocco et al. 2010). The Mt-F-box gene under study is part of the SCF complex. Based on the analyses performed, we assume that its function in the cell cycle is related to the degradation of target proteins from the cell cycle during the G1/S phase.

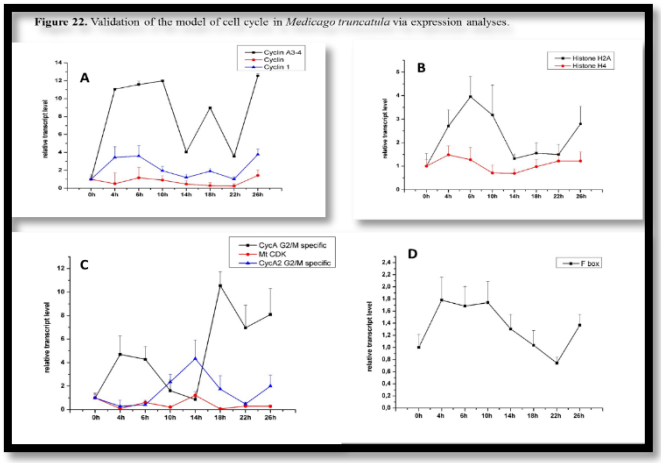
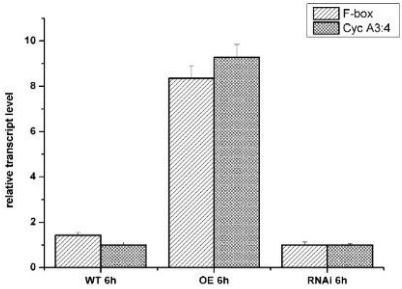


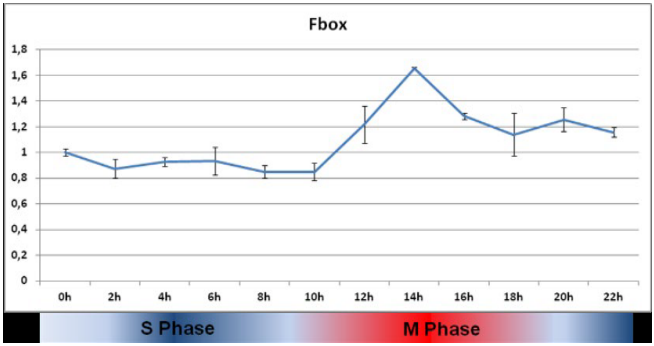
Figure 23. Transcript levels of Mt-F-box and MtCyclinA3-4 in meristem cells from synchronized root tips for 6 hours in transgenic plants with modified expression and control.



Based on the system developed for the *A.thaliana* model plant for root tip synchronization in previous studies (Cools et al. 2010), activation of the *At-F-box* gene during the cell cycle in *A.thaliana* wild-type plants was determined. The investigated At-F-box gene ortholog of the Mt-F-box shows the maximum level of expression in the transition to mitosis or more precisely at the G2/M checkpoint. By the time the S phase

is complete, gene expression is almost constant (10th hour), after which it activates and peaks at 14th hour. Unlike the *Mt-F-box* gene, in which we determined almost constant and highest expression (from the 6th to the 10th hour) during the prolonged S phase, achieving maximum expression of the *A.thaliana* gene differs much more short duration of action Figure 24.

Figure 24. Expression of *At-F-box* during the cell cycle in wild-type *A.thaliana*.

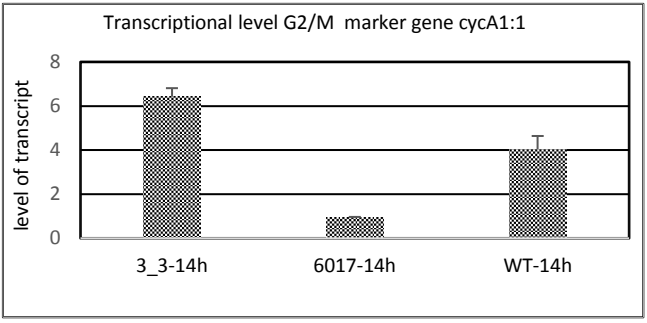


Activation of *At-F-box* gene expression is associated with the G2/M transition from the cell cycle. As an integral part of the SCF complex, the role of the investigated gene is most likely to be related to the degradation of proteins involved in the cell cycle. In *A.thaliana*, the G2/M transition is very well studied in relation to key genes as well as their transcripts. Cyclin dependent kinases CDKB1:1, CDKB2:1 and cyclins CycA1:1, CycA2:1 CycB1:1 have been shown to have maximum expression during the transition to mitosis even as the peak of their transcripts binds to the 14th hour of HU treatment . Cyclin CycA1:1, which is extremely specific for the G2/M transition, was selected for the experiment. Root tips of *A.thaliana* transgenic lines with modified expression 6017OE and amiRNA3-3, as well as control plants WT, were synchronized with 1 mM HU for a period of 14 hours.

The transcriptional levels of the selected marker gene in the assay lines and controls were determined by qRT-PCR. The transcript level of the selected CycA1:1 marker gene is highest in amiRNA3-3line (Fig. 25). At line 6017OE, the reported level of the transcript is much lower than even the control. The results of the experiment allow us to assume that the F-box protein under study is involved in protein degradation, /as part of the structure of SCF complexes/ which are specific for the G2/M transition. According to the database Atted II ver. 9.2, the list of genes co-expressed with studied At1g10780, CycA1; 1 / At1g44110 / directly binds to the F-box protein examined. The level of expression correlation based on Microarray results is very similar for these two

genes /http://atted.jp/. Most likely, the F-box protein binds directly to CycA1;1 and participates in its degradation during the G2/M transition. For this reason, the reported level of CycA1; 1 transcript in the 3-3 repressed expression line is the highest, followed by the control, and the lowest at the 6017 overexpressing At-F-box.

Figure 25. Transcriptional levels of the G2 / M marker gene cycA1: 1 in transgenic lines with modified expression of *Arabidopsis thaliana* and control.



More detailed elucidation of the functions of the investigated F-box protein in *M.truncatula* and *A.thaliana* associated with its activation during the cell cycle, its expression in almost all plant organs and tissues during plant growth and development requires additional protein-level assays. To answer some of these questions, experiments were performed based on the Tandem Affinity Purification (TAP) method developed in both model plants. This method is characterized by affinity purification of proteins or protein complexes and their mass spectral determination (AP-MS) (Wodak et al. 2009; Pflieger et al. 2011; Braun et al. 2013). For the experiment, constructs are prepared that are translational fusions of the protein of interest "bait" and double affinity tag, which are introduced into the plants by genetic transformation, the total proteins are isolated from the resulting transgenic plants, and by a subsequent double affinity procedure. Purification separates the protein complexes and is determined using protein-sensitive high-sensitivity mass spectral detection (Li 2011). This method provides information about protein interactions and gives a more complete picture of the functions of the gene under study.

Isolation of protein complexes by TAP analysis in *M.truncatula* is based on the production of transgenic hairy roots, and in *A.thaliana* transformed suspension cultures. The results of the Tandem Afinity Purification TAP analysis of *M.truncatula* identify direct and indirect functional interactions of the F-box protein encoded by the gene under study. There were 4 direct and 10 putative interactors of the F-box protein examined. According to the results obtained, one of the direct interactors of the F-box

protein is the SKP1 protein. SKP1 as well as the F-box protein are key units that form the SCF complex responsible for ubiquitination of proteins (Kuroda et al. 2002), thus being involved in the process of selective degradation. Our subsequent studies have focused on one of the major interactors of the F-box protein identified as a gene encoding 2-isopropylmalate synthase (2-isopropylmalate synthase, IPMS, Medtr1g116500.1). Available data in the literature define the gene as the coding enzyme isopropylmalate synthase, which is involved in the step in the biosynthesis of the branched amino acid leucine (de Kraker et al. 2007). The amino acid leucine is one of the three branched amino acids-valine (val), leucine (leu), isoleucine (isoleu). Specific to these amino acids is that they are synthesized only by plants. In plant organisms, synthesis and degradation of branched aminoacids are in strict balance, and its disruption causes changes in plant growth and development.

The results of the TAP analysis as well as the available literature suggest that transcriptional analysis of the IPMS gene was performed on T1 transgenic lines of *M. truncatula* with modified expression (MtF-box OE and RNAi) and control. The investigated MtF-box gene encoding the F-box protein is an integral part of the SCF complex and is involved in the selective degradation of the target enzyme isopropylmalate synthase, which decreases its activity and leads to a decrease in the level of IPMS transcripts. With MtF-box OE lines, the level of IPMS transcripts is lower, and slightly increased content of IPMS transcripts was reported in lines with inactivated gene expression. Similar correlation of enzyme activity and transcript was reported in the study of Kuntz et al. 1992 and Yin et al. 2017 as well as found in our studies for the HAC1 gene encoding histone acetyltransfer (Boycheva et al. 2017).

The differences in transcript levels found in the expression lines also suggest differences in the level of synthesized leucine. It is logical to assume that leucine concentrations are expected to be higher in overexpressing lines and higher inactivated expression lines. The amino acid deficiency of leucine in the overexpression lines is probably the reason for the reported leaf morphological changes observed in these transgenic lines. The level of F-box transcripts on MtRNAi lines is lower, leading to higher levels of IPMS transcripts, and possibly an increase in leucine.In order to confirm the validity of the results of TAP analysis and transcript analysis at the MtIPMS gene, a logical next step in our studies is to determine the free amino acid content of *Medicago truncatula* transgenic lines with modified expression and control.The free amino acid content of leu, isoleu and val was determined by GC/MS (gas chromatography/mass spectrometry) analysis in *M. truncatula* plants with modified MtF-box expression and wild-type control. The results obtained confirm the presence of lower levels of the three amino acids in the overexpressing lines and higher in the inactivated lines.(Table 6). The concentration of branched val, leu, and isoleu amino acids in *M. truncatula* lines with

modified MtF-box expression and control was counted for each sample and expressed in μg/g FW. The data obtained clearly demonstrate a low content of all three amino acids - val ($p \leq 0.01$), leu ($p \leq 0.001$) and isoleu $p \leq 0.001$) on the overexpression lines, and significantly higher on the inactivated expression lines ($p \leq 0.001$) as well as compared to controls (Table 6). The results obtained for the determination of the amount of branched amino acids confirm the results of the TAP analysis and the determined transcriptomic profile of the MtIPMS gene and demonstrate the function of the examined gene encoding MtF-box protein in the synthesis of branched amino acids.

Table 6. Valine, leucine and isoleucine (μg/g FW) content of *M. truncatula* transgenic lines with modified expression and control.

	L-valine, μg/g FW		**L-leucine, μg/g FW**		**L-isoleucine, μg/g FW**	
	Average value	±SD	Average value	±SD	Average value	±SD
OE	32.83	±1.89	56.30	±5.29	28.77	±4.27
RNAi	61.36	±6.16	293.73	±19.18	72.80	±5.00
WT	33.25	±1.39	202.21	±15.44	64.66	±4.75

The results of the Tandem Afinity Purification TAP analysis of *A.thaliana* identified 1 direct interactor interacting with the AtF-box protein encoded by the gene under study. One suspected interactor of the investigated F-box protein has been identified.

According to the results obtained, the indirect interactor of the F-box protein is the SKP1 protein. It has been noted above that SKP1 and F-box proteins are the units that form the SCF complex responsible for protein ubiquitination (Kuroda et al. 2002) and thus are involved in the process of selective degradation of target proteins. As with the Mt-F-box protein and in *A.thaliana*, the only direct interactor identified was determined to be 2-isopropylmalate synthase (2-isopropylmalate synthase, IPMS, AT1G74040.1), and subsequent studies focused on this single AtF- box protein.

Similar to the studies on *M. truncatula*, transcriptional analysis of the IPMS gene was performed on transgenic *A.thaliana* transgenic lines with modified expression (AtF-box OE and amiRNA) and control. Transcript levels were examined separately in roots and leaves in transgenic lines and controls due to the very strong expression of the GFP marker gene traced in both plant organs. With AtF-box OE line 6017, the level of IPMS transcripts is lower both in the roots and in the leaf, suggesting that the gene encoding

the AtF-box protein studied by us is directly involved in the degradation of the target enzyme isopropylmalate synthase, leading to reducing the level of IPMS transcripts. As indicated above, there is a correlation between enzyme activity and transcript level. Degradation of the enzyme isopropylmalate synthase also results in a decrease in the level of IPMS transcripts. In lines with inactivated gene expression, a slightly increased content of IPMS transcripts was reported compared to the AtF-box OE line, both in roots and in leaves.

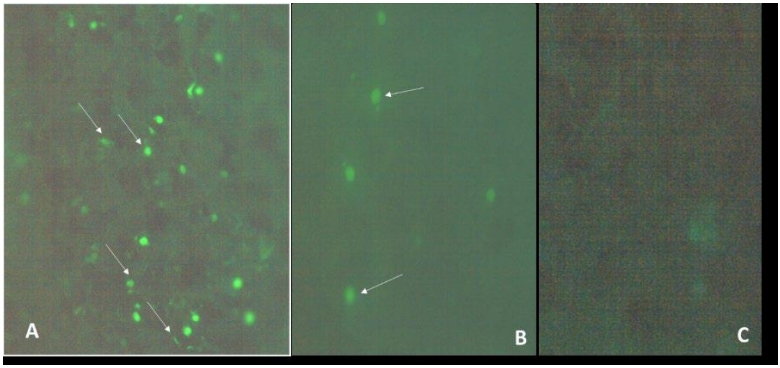
To confirm the validity of the results of the TAP analysis and the transcriptome level analysis performed for the AtIPMS gene, a logical next step in our studies is to perform an analysis showing the interaction of the AtF-box protein encoding gene and the isopropylmalate synthase enzyme-encoding AtIPMS gene for evidence of protein-protein interaction.

Following the complete genomic sequencing of various plant species, as well as advances in decoding whole proteomes, the need for a reliable way of identifying protein-protein interactions becomes an important task in plant research. Bimolecular fluorescence supplementation (BiFC) is an assay based on a non-invasive fluorescence technique that allows the detection of protein interactions in living cells, and can be used to determine the sub-cellular localization of interacting proteins. BiFC analysis is based on the reconstruction of split non-fluorescent variants of GFP, or yellow fluorescent protein (YFP), to form a fluorescent fluorophore (Ghosh et al. 2000; Hu et al. 2002). The BiFC method has also been applied to visualize protein-protein interactions in subcellular structures such as the endoplasmic reticulum (Zamyatnin et al. 2006), chloroplasts (Caplan et al. 2008), the vacuole (Voelker et al. 2006), the nucleus (Stolpe et al. 2005) in a number of plant species such as tobacco (Zamyatnin et al. 2006), parsley (Stolpe et al. 2005) and Arabidopsis (Voelker et al. 2006) in various tissues and cell types - intact leaves, epidermal layers, semen and protoplasts were used in this analysis.

Transformation of infiltration of the leaves of greenhouse plants from *Nicotiana bentamiana* was carried out using different combinations of vectors and to each of these combinations was added an auxiliary vector p19 for genetic transformation and acetosyringone to increase the efficiency of agro-infiltration. A reconstructed signal for green fluorescence in the nuclei of the epidermal cells was observed with the combination of constructs used at 35S: cF-box-stop: heGFP + 35S: cIPMS-stop: teGFP. The established interaction of the two proteins successfully restores the separated non-fluorescent fragments of the GFP protein to an active fluorescent protein, with pronounced nuclear localization. The results of the BIFC analysis performed are shown in Figure 26 and are consistent with the results of the TAP and transcriptomic analysis performed. The experiments, based on the three methods (TAP, transcriptomic profile

and BIFC), prove the participation of the studied gene encoding the AtF-box protein in metabolic processes / branched AK synthesis / in the model plant *A.thaliana*.

Figure 26. Interaction of AtF-box and AtIPMS on Agro-infiltration of *Nicotiana bentamiana* leaves. A-Presence of protein:protein interaction, arrows indicate pronounced nuclear and weaker cytoplasmic localization of the interacting proteins; B-positive control with nuclear localization of the GFP signal; C-negative control no signal. Observations were made with a 10X lens.



FINALE

The research work presented in this dissertation reflects the accumulation and upgrading of knowledge gained in the study of the genus *Medicago*, the model plant for the Legumes family *M. truncatula*, as well as two other model species included in the experimental work - the second model plant of the family Legumes - *Lotus japonicus* and model plant for plant biology - *Arabidopsis thaliana*. Initially obtained key results and accumulated knowledge in the field of plant regeneration through direct and indirect somatic embryoesis in liquid and solid media, are the basis for the development of effective and original genetic transformation protocols for model legumes.

The developed cell suspension protocols for *M. falcata* and *M.truncatula* are essential in the study of early events - first asymmetric cell division, induction of embryogenic potential and somatic embryos originating from single cells. These systems are the perfect tool to track the process of plant development from one cell to the whole plant. The development of modern plant science and the presence of such systems for cell somatic embryogenesis allow the identification of key genes and the tracking of their expression in dynamics throughout the plant development process. Moreover, the presence of cell suspension culture regeneration implies the development of gene transfer systems and the origin of transgenic plants from single transformed cells.

The accumulated original results in the field of plant regeneration and genetic transformation in the model legume plant provide a prerequisite for the initiation of insertional mutagenesis studies by the *Tnt1* retrotransposon and the creation of a collection of mutant lines in *M.truncatula* characterized by efficient transposition of the retroelement and multiple new insertion in the genome of the model plant. In the course of the experimental work on insertional mutagenesis, in partnership with Prof. Pascal Rate's research team, the retro-element MERE1 was identified and investigated, which is also transposed during *in vitro* regeneration and induces additional insertions into the genome of the mutant lines. The insertion mutant lines created are valuable experimental material, both for identifying key plant development genes and for exploring their role in the functioning of the entire plant organism.

The next upgrade step in research is to initiate functional genomics studies in legumes. The created mutant collection of *Medicago truncatula* became the basis for launching functional and comparative genomics studies of model legumes *Medicago truncatula* and *Lotus japonicus*, using the methods of forward and reverse genetics in the framework of a project of which the team of Functional Genetics legumes group, is the coordinator.

The accumulated results and experimental experience in developing a collection of insertion mutant lines in *Medicago truncatula* form the basis for the development of a system for *Tnt1* insertion mutagenesis in the second model legume plant - *Lotus japonicus* and the creation of a small collection of insertion mutant lines. The developed system for insertional mutagenesis in the second model plant would facilitate comparative genomics research, as well as complement the existing strategy for creating mutant lines in *Lotus japonicus* based on the endogenous retrotransposon LORE1.

The ABI insertion mutant lineage collection selected sequenced lines with key genes involved in plant development, nitrogen fixation and secondary rooting, as well as involved in signaling cascades and plant protection. These genes were cloned and transferred into the genome of the model species *Medicago truncatula*, *Lotus japonicus*, and *Arabidopsis thaliana* to study their function. Structured transgenic plants with overexpression, repressed expression and analysis of promoter activity of selected genes in the three model species have been constructed using gene transfer protocols already developed, which support the experiments on functional analysis. For six of the studied genes, their involvement in different stages of the process of indirect somatic embryogenesis in *M. truncatula* was demonstrated for the first time after monitoring the expression of the *GUS* marker gene under the control of their endogenous promoters.

A detailed comparative study on functional genomics for the gene encoding the F-box protein from the genome of the model legume *Medicago truncatula* and the model plant

Arabidopsis thaliana is presented in the dissertation. The results obtained from studies of phenotype, leaf cell morphometry, root growth dynamics, in transgenic plants with modified expression and controls, and localization of expression of the *GUS* and *GFP* marker genes, associate the study gene with actively dividing tissues and organs in *M. truncatula* and *A. thaliana*, and confirm its involvement in growth and plant development.

In the course of the experimental work, the maximum activity of the gene encoding F-box protein during the cell cycle in the model plants *M. truncatula* and *A. thaliana* was determined, which suggests its participation in the degradation of key proteins from the cell cycle. As a result of the TAP analysis, a common direct gene interactor was radiated in both model plants, which revealed its involvement in the metabolic pathway for the synthesis of the branched amino acid leucine. In *M. truncatula*, this function of the gene was confirmed by determining the amounts of free branched AK in transgenic plants with modified expression and control. In the reference plant *A. thaliana*, the involvement of the gene encoding the F-box protein in leucine synthesis was confirmed by carrying out an analysis of bimolecular fluorescence supplementation.

The results of a comparative functional study of the gene encoding an F-box protein from model plants *M. truncatula* and *A. thaliana* determine its specific function in the synthesis of the branched amino acid leucine, as well as its role in the plant cell cycle.

Fundamental discoveries in biology are based on model plant species, and the information accumulated in these species can be transmitted to crops of economic importance. The knowledge gained from the model plant *M. truncatula* can be transferred to closely related legumes-alfalfa and clover, even when it comes to agronomic harvesting characteristics, growth features, which is confirmed by functional analyzes of the *STAY-GREEN* gene, and the results of the model plant has been brought forward to improve the alfalfa forage crop. The knowledge gained from model plants is a source for improving crops.

CONCLUSIONS

Summarizing the results presented in this dissertation obtained in the process of experimental work, the following important conclusions can be drawn:

1. Using liquid suspension cultures of *M.truncatula* and *M.falcata*, the asymmetry of the first cell division was demonstrated, which initiated the process of direct somatic embryogenesis and traced its development from cell to whole plant.

2. The genome size of Medicago diploid species originating from Algeria has been established and the relationship between genome size and competence for inducing direct somatic embryogenesis in a liquid medium has been demonstrated.

3. Effective insertional mutagenesis through *Tnt1* retrotransposon was achieved and a collection of mutant lines was created in *M.truncatula* cv. Jemalong using two new regeneration protocols and an osmotic treatment method.

4. A system for *Tnt1* insertion mutagenesis was developed for the second model legume plant - *Lotus japonicus*, based on the experimental experience of *M.truncatula*.

5. A new retro element *MERE1-1* has been identified and its transposition into the genome of *M.truncatula* mutant lines during in vitro regeneration has been demonstrated.

6. The role of *Tnt*1 as an effective mutagen in the genome of the model plant *M.truncatula* in the phenotyping of selected insertion mutant lines grown under different conditions -*in vitro, in vivo*, airoponic and hydroponic culture has been demonstrated.

7. Vectors for overexpression, knockdown expression and analysis of promoter activity of the studied genes have been established, as well as constructs for TAP and BiFC analyzes in the *M.truncatula* and *A.thaliana* model plants.

8. The involvement of the *MtLAX; MtARFB3; MtF-box; MtHAC; MtGRAS; MtDRP* genes tested at various stages of the indirect somatic embryogenesis process in *M. truncatula* after monitoring the expression of the *GUS* marker gene under the control of endogenous promoters has been demonstrated.

9. Homozygous transgenic plants with overexpression and knockdown expression of thegene coding F-box protein for the model plant *A. thaliana* were obtained.

10. The activity of the *GUS* and *GFP* marker genes was monitored and the expression of the studied F-box gene in plant tissues and organs of transgenic plants of *M. truncatula* and *A. thaliana* was localized.

11. F-box overexpression in *M. truncatula* results in visible leaf-shaped deviations, a shorter and less developed root system with single nodules, and transgenic F-box plants with knockdown expression are characterized by a close-to-control morphology.

12. Homozygous *A. thaliana* plants with F-box overexpression differ in leaf shape and size, as well as dominance of 4C cells in the leaves, and those with suppressed

expression form highly elongated leaves characterized by 8C and 16 cells in leaves compared to overexpressing plants and controls.

13. The maximum activity of the investigated gene encoding the F-box protein during the cell cycle in the model plants *M. truncatula* and *A. thaliana* has been determined, suggesting its participation in the degradation of key cell cycle proteins.

14. As a result of the TAP analysis, a common direct interactor of the gene encoding the F-box protein was determined and its expression profile was defined in transgenic plants with modified expression and controls in the two model plants - *M. truncatula* and *A. thaliana*.

15. The amounts of free branched aminoacids in transgenic plants with modified expression and control in *M. truncatula* validate the results of the TAP analysis and the determined transcriptomic profile of the direct F-box protein interactor.

16. The detection of a green fluorescence signal after Agro-infiltration of a leaf of *Nicotiana benthamiana* validates the results of the TAP analysis and the determined transcriptomic profile of the direct interactor of the *A. thaliana* F-box protein.

CONTRIBUTIONS

I. CONTRIBUTIONS WITH FUNDAMENTAL SCIENTIFIC VALUE

1. Asymmetry of the first cell division, the induction of embryogenic potential and the development of the process from cell to whole plant have been demonstrated in the suspension cell culture of model plants *M.truncatula* and *M.falcata*.

2. The genome size of Medicago diploid species originated from Algeria was examined for the first time and the relationship between genome size and competence for inducing DSE in a liquid medium was established.

3. A collection of *Tnt1* insertion mutant lines in *M.truncatula* cv. Jemalong has been created through efficient transposition protocols.

4. A small collection of *Tnt1* insertion mutant lines was developed for the second model legume plant - *Lotus japonicus*.

5. For the first time, a new retro-element MERE1-1 has been identified in the genome of *M.truncatula* mutant lines that is transposed during *in vitro* regeneration.

6. An original insertion mutant line was identified, named “Fishbone” after phenotypic characterization, and *Tnt1* and *MERE1-1* inserts were sequenced.

7. For the first time, a comparative functional study of a gene encoding an F-box protein in model plants *M.truncatula* and *A.thaliana* was made, including phenotypic, morphometric, transcriptomic, proteomic and metabolic analyses.

II. CONTRIBUTIONS WITH SCIENTIFIC APPLIED VALUE

1. An effective protocol for the genetic transformation of cell suspensions of *M. truncatula* has been created, in which stable transgenic plants are regenerated from single-cell and provide the opportunity for cellular and sub-cellular analyses.

2. A system for synchronization of root-tip meristem cells was developed to support cell cycle studies in the model plant *M.truncatula*.

List of publications included in the dissertation

1. Iantcheva A, Vlahova M, Atanassov A **(2006a)**. Somatic embryogenesis in genera *Medicago*: an overview. In: Mujib A, Samaj J (eds). Somatic embryogenesis. Plant Cell Monographs, vol 2. Springer, Berlin, pp 285–304.
2. Iantcheva A, Vlahova M, Atanassov A **(2008)**. Somatic embryogenesis from leaf explants of *Medicago truncatula* cv. Jemalong genotype 2HA In: Mathesius U, Journet EP, Sumner LW (ed). The *Medicago truncatula* handbook. http://www.noble.org/MedicagoHandbook/.
3. Iantcheva A, Chabaud M, Cosson V, Barascud M, Schutz B, Primard-Brisset C. Durand P, Barker DG, Vlahova M, Ratet P **(2009a).** Osmotic shock improves Tnt1 transposition frequency in Medicago truncatula cv Jemalong during in vitro regeneration, **Plant Cell Reports,** vol 28, pp. 1563-1572.
4. Rakocevic A, Mondy S, Tirichine L, Cosson V, Brocard L, **Iantcheva A,** Cayrel A, Devier B,. Abu El-Heba GA, Ratet P **(2009).** MERE1, a low copy number copia type retrotroelement in Medicago truncatula active during tissue culture**. Plant Physiology**, vol. 151, pp. 1250-1263.
5. Vassileva V, Zehirov G, Ugrinova M, **Iantcheva A (2010).** Variable leaf epidermal morphology in Tnt1 insertional mutants of the model legume *Medicago truncatula*. **Biotechnology &Biotech Eq.** vol. 24/4, 2060-2065.
6. Revalska M, Vassileva V, Goormachtig S, Van Hautegem T, Ratet P, **Iantcheva A (2011).** Recent Progress in Development of Tnt1 Functional Genomics Platform for *Medicago truncatula* and *Lotus* japonicus in Bulgaria. **Current genomics**, 12, 147-152.
7. Iantcheva A, Mysore K, Ratet P **(2013)**. Transformation of leguminous plants to study symbiotic interactions. **International Journal Developmental Biology** 57:577-586.

8. **Iantcheva A**, Revalska M, Zehirov G, Vassileva V **(2014)**. *Agrobacterium*-mediated transformation of *Medicago truncatula* cell suspension culture provides a system for functional analysis. **In vitro Cellular and Developmental Biology – Plant**, DOI 10.1007/s11627-013-9554-4, 50(2), pp. 149-157.
9. Boycheva I, Vassileva V, Revalska M, Zehirov G, **Iantcheva A (2015)**. Cyclin-like F-box protein plays a role in growth and development of the three model species *Medicago truncatula*, *Lotus japonicus*, and *Arabidopsis thaliana***. Research and Report in Biology**. vol 6:117-130, DOI http://dx.doi.org/10.2147/RRB.S84753
10. Iantcheva A, Boycheva I, Revalska M **(2015).** Development of root tips synchronized system for the model legume *Medicago truncatula* upon replication stress. BJAS. 21(6): 1177-1184.
11. Iantcheva A, Revalska M, Zehirov G, Boycheva I, Magne K, Radkova M, Ratet P, Vassileva V **(2016)**. *Tnt1* retrotransposon as a tool for developing an insertional mutant collection of *Lotus japonicus*. **In vitro Cellular and Developmental Biology – Plant, 52(3), 338-347.** DOI: 10.1007/s11627-016-9768-3
12. Fyad-Lameche FZ, **Iantcheva A**, Siljak-Yakovlev S, Brown SC **(2016)**. Chromosome number, genome size, seed storage protein profile and competence for direct somatic embryo formation in Algerian annual Medicago species. **Plant Cell Tiss Organ Cult**. vol 124(3) pp 531-540. DOI 10.1007/s11240-015-0912-2.
13. Iantcheva A, Revalska M **(2018).** Early events during the induction of somatic embryogenesis in genera *Medicago.* **BJAS**, vol. 6:1042-1052.