Charles University

Faculty of Science

Experimental Plant Biology

Cellular and Molecular Biology of Plants

Bc. Karel Raabe

Characterization of subunit A of the Eukaryotic translation initiation factor 3 in *Arabidopsis thaliana*

Charakterizace podjednotky A eukaryotického translačního iniciačního faktoru 3 a její role v *Arabidopsis thaliana*

Diploma thesis

Supervisor: Christos Michailidis, Ph.D.

Consultant: doc. RNDr. David Honys, Ph.D.

Praha. 2020

Declaration

I hereby declare that I have completed this thesis independently, using the listed literature and resources only. Content of the thesis or any part of it has not been used to gain any other academic title.

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

Prague, 06. 08. 2020

…….

Karel Raabe

The presented work was carried out at the Laboratory of Pollen Biology, Institute of Experimental Botany ASCR, v.v.i., Prague 6, Czech Republic.

Acknowledgements:

I wish to express my thanks to my supervisor Christos Michailidis, Ph.D. for all his advice, discussion, patience and professional guidance through the completion of this thesis, where he lead me to deep understanding of the methods and pushed me to schedule, perform, analyse and/or fail the experiments on my own, while still being there to help whenever I truly needed. Many thanks belong to RNDr. Oldřich Navrátil CSc. as well, as he was teaching us how to successfully produce stable transgenic Tobacco. I´m also very grateful to my consultant doc. RNDr. David Honys, Ph.D. for his precious advice and the given opportunity to be a part of the team of the Laboratory of Pollen Biology.

Abstract

In plants, translation regulation plays an important role during progamic phase, fertilization and seed development. The process of translation is mostly regulated in its initiation phase, where Eukaryotic translation initiation factor 3 (eIF3) is the largest and most complex initiation factor, consisting of 12 different subunits. In plants, single eIF3 subunit mutants caused various growth and development defects, depending on the particular subunit that was mutated. However, not all the plant eIF3 subunits were characterized to this date. The objective of this work was to functionally characterize the eIF3 subunit A using *Arabidopsis thaliana* as the main model plant. We described in this work that plant eIF3A proteins share high levels of homology and domain organization with eIF3A subunits from non-plant eukaryotic species but contain regions specific only to plants. Next we described that *Arabidopsis thaliana AteIF3A* gene is transcribed in highly proliferating tissues, its protein product localizes to cytoplasm and around pollen vegetative cell nucleus and observed an increased frequency of defective pollen grains and defects in seed formation in plants with T-DNA insertion localized to the *AteIF3A* gene. We also produced stable transgenic *Nicotiana tabacum* lines expressing heterologous AteIF3A protein from Arabidopsis.

Key words:

plants, reproductive tissues, gametophyte, translation initiation, translation regulation

Abstract in Czech:

U rostlin hraje regulace translace důležitou roli během progamické faze, oplození a vývoje semene. Regulace syntézy proteinů je soustředěna převážně v iniciační fázi translace, kde je největším a nejsložitějším iniciačním faktorem Eukaryotický translační iniciační faktor 3 (eIF3), který se skládá z 12 nestejných podjednotek. Mutace jednotlivých podjednotek v rostlinách způsobily defekty v různých fázích růstu a vývoje rostliny v závislosti na dané mutované podjednotce. Charakterizace některých podjednotek rostlinného eIF3 komplexu však doposud nebyla provedena. Předmětem této práce bude funkční charakterizace podjednotky A převážně v modelové rostlině *Arabidopsis thaliana*. V této práci jsme popsali vysoce shodnou homologii a doménovou organizaci proteinů podjednotky eIF3a mezi rostlinnými a nerostlinnými druhy eukaryotických organismů, která má však svá rostlinná specifika. Dále jsme popsali, že gen *AteIF3A* je transkribován v rychle se dělících pletivech, jeho proteinový produkt je lokalizován v cytoplasmě a v okolí jádra vegetativní buňky pylu a pozorovali jsme zvýšenou frekvenci defektu pylových zrn a defektu v utváření semen v rostlinách s T-DNA inzercí v genu *AteIF3A*. Také jsme vytvořili stabilní transgenní linie rostliny *Nicotiana tabacum* s heterologní expresí AteIF3A proteinu z Arabidopsis.

Key words in Czech:

rostliny, reprodukční orgány, gametofyt, iniciace translace, regulace translace

List of frequently used abbreviations

CaMV	-	Cauliflower mosaic virus
Cryo-EM	-	Cryogenic Electron Microscopy
CTD	-	C-terminal domain
eEF	-	eukaryotic translation elongation factor
eIF	-	eukaryotic translation initiation factor
eIF3a	-	eukaryotic translation initiation factor subunit A
EPP	-	EDTA/puromycin-resistant particle
eRF	-	eukaryotic translation release factor
GTP	-	guanosine triphosphate
GFP	-	green fluorescent protein
IRES	-	internal ribosome entry site
MFC	-	multi-factor complex
MPN	-	Mpr-Pad1-N-terminal domain
MS	-	Murashige-Skoog media
MSA	-	Multiple Sequence Alignment
NTD	-	N-terminal domain
PCI	-	Proteasome, COP9 signalosome, eIF3 domain
PIC	-	pre-initiation complex
SDS-PAGE	-	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
TC	-	ternary complex
UTR	-	untranslated region
YFP	-	yellow fluorescent protein
YLC	-	yeast-like core

Contents

Introduction

In Angiosperms, male gametophyte is reduced to a three-cell organism, the pollen grain. During the male gametophyte development, especially the translation is highly regulated and alters between several translational states, starting with massive production of transcripts and inactive mRNA storage during microsporogenesis and controlled activation and localized translation during pollen germination and pollen tube growth.

Eukaryotic translation initiation factor eIF3 is a 12-subunit complex that, during the translation initiation phase, serves as a scaffold for mRNA and 40S ribosomal subunit assembly. In eukaryotes, eIF3 research has a long history of discoveries on the roles and functions pertained by this complex. Moreover, the literature points on possible eIF3 subcomplexes with non-identical composition within one cell that could lead to different regulatory potential. Our laboratory is currently focusing on characterization of eIF3 subunits in plants, as several subunits of the eIF3 complex were found in *Nicotiana tabacum* pollen proteomic analysis as a component of large mRNA storage particles, described as non-translating monosomes. Recently, we summarized the eIF3 research in plants (Raabe et al., 2019). Now, using the reverse genetic approach, we are trying to characterize eIF3 subunits with the focus on pollen development and its fitness.

The conceptualization of this thesis is the novel characterization of subunit A (eIF3A), the largest subunit of the eIF3 complex, which is encoded by single gene in *Arabidopsis thaliana*. The aims of this work are to compare the plant eIF3A protein *in silico* with its better characterized orthologues from non-plant eukaryotic species, to analyse the *AteIF3A* gene expression pattern, describe subcellular localization of its protein product, AteIF3A and to isolate the epitope tagged AtEIF3A protein. Another important aim of this work is to characterize a mutant of the *AteIF3A* gene. Moreover, as the initial studies done in our lab were using *Nicotiana tabacum* pollen, we aimed to introduce a stable *N. tabacum* transgenic lines expressing *Arabidopsis* eIF3 subunits in pollen.

Aims of the thesis:

- *In silico* analysis of plant eIF3A protein sequence and its comparison with other eukaryotic eIF3A orthologues
- Phenotypical characterization of *Arabidopsis thaliana* T-DNA insertion lines in the *AteIF3A* gene locus At4g11420
- Expression analysis of *AteIF3A* gene in *Arabidopsis thaliana* stable transgenic lines expressing GFP-GUS fusion protein driven by the native eIF3A promoter
- Subcellular localization of AteIF3A protein in *Arabidopsis thaliana* stable transgenic lines expressing the AteIF3A protein fused to GFP, driven by native, constitutive sporophytic or male gametophyte specific promoters
- Stable transformation and regeneration of *Nicotiana tabacum* plants expressing fluorescent markers or heterologous *Arabidopsis* AteIF3A subunit fused to GFP

Literature overview

Plants as model organisms

Plant characteristics and life cycle

Flowering plants are multicellular autotrophic eukaryotic organisms capable of photosynthesis. They are characteristic for their sessility, totipotency, non-determined growth, modularity of their body and position-determined differentiation of cells and tissues/organs. Moreover, during their sexual life cycle, an alteration of generations occurs, where both haploid and diploid cells mitotically divide and produce a complete generation. The generation of haploid cells, the gametophyte, is highly reduced in flowering plants and mostly depends on the sporophyte.

This thesis will focus on two plant model organisms: *Arabidopsis thaliana* and *Nicotiana tabacum*. Arabidopsis is the main model for plant research. It is a small self-pollinating plant that belongs to *Brassicaceae* family. Its genome is relatively small, consisting of 5 chromosomes with total genome size of 125 Mb. Arabidopsis plants are plants of small size with quick lifespan (Figure 1A) that are easily cultivated, producing a reasonably great number of seeds per plant. Under standard conditions (16/8-hour light/dark period and 22°C) Arabidopsis seeds germinate to seedlings with fully opened cotyledons within a week. After a month since sowing, first flowers open on the emerging inflorescence. After fertilization, flowers turn into a silique, a fruit containing around 60 seeds in Col-0 ecotype used in our laboratory. After ripening and senescence, seeds of the next generation are ready to be harvested after approx. 2 months from sowing.

Nicotiana tabacum is a commercially cultivated crop plant belonging to *Solanaceae* family. It is an allotetraploidic hybrid species (2n=4x=48) of *N. sylvestris* and *N. tomentosiformis*. In our cultivation conditions (same as for Arabidopsis), *N. tabacum* seeds germinate within two weeks into seedlings, and first flowers are produced in 2-3 months. Seeds can be harvested from dry pods in about 4-5 months after seed sowing. Although the process of stable transformation is more difficult for Tobacco species than for Arabidopsis, the advantages for pollen biology experiments and their possible use for crop plants cultivation make the *N. tabacum* an interesting model organism to use.

Male gametophyte and its development

In plants, reproduction cycle alters between diploid sporophyte and haploid gametophyte. Sporophyte produces male microspores and female megaspores by meiotic division. Haploid spores then develop to gametophyte that produces gametes, male sperm cells or female egg cells. In Angiosperms, gametophyte is reduced to a small number of cells and most of its life is surrounded by generative tissues of the sporophyte.

The development of the pollen takes place in anthers, where diploid microspore mother cells (microsporocyte) meiotically divide into four haploid microspores forming a tetrad. Microspores are then released from tetrads, forming a large vacuole that pushes the nucleus from central position to the edge. Asymmetric division (pollen mitosis I; PMI) then produces large vegetative cell and small generative cell. Shortly after that, the generative cell is engulfed into vegetative cell cytoplasm and divides once more to form two sperm cells (pollen mitosis II; PMII). PMII occurs either before or after full pollen maturation, varying between species. Tobacco and Arabidopsis are amongst the better known examples of plants producing two- or three-cell mature pollen grain, respectively (Figure 1B). During anthesis, mature pollen grains leave the anther in a dehydrated state. That is the only time when male gametophyte is freely outside of the sporophyte, and can be easily collected and studied, representing a full haploid model organism for cell fate, cell growth, cellular differentiation and cell signalling research (Hafidh et al., 2016).

After reaching the surface of stigma, pollen is rehydrated and activated, followed by a germination event where pollen tube emerges and subsequently penetrate the female tissues by tip growth of the pollen tube. Tip growth is one of the fastest cell growths described in the plant kingdom and requires a tight coordination of cytoskeleton, vesicular trafficking, ion transport and cell signalling. Callose plugs are formed at the opposite site to the growing tip to ensure the constant volume of the cytoplasm of the vegetative cell. The growing of the pollen tube is also dependent on new cell wall formation at the tip, calcium ion and proton exchange with the environment and their gradient within the tube (Honys, 2018).

Moreover, to fuel the pollen tube growth, a massive storage of saccharides, proteins, nucleic acids and fats is formed in the vegetative cell during pollen development. This phenomena unsurprisingly uncovered large translational regulation events, where mRNA, ribosomes and other proteins connected to translational machinery are stored in a large particles of non-translating monosomes, formerly called EPP, which stands for EDTA/puromycin resistant particles (Hafidh et al., 2018). Large amounts of transcripts produced during pollen

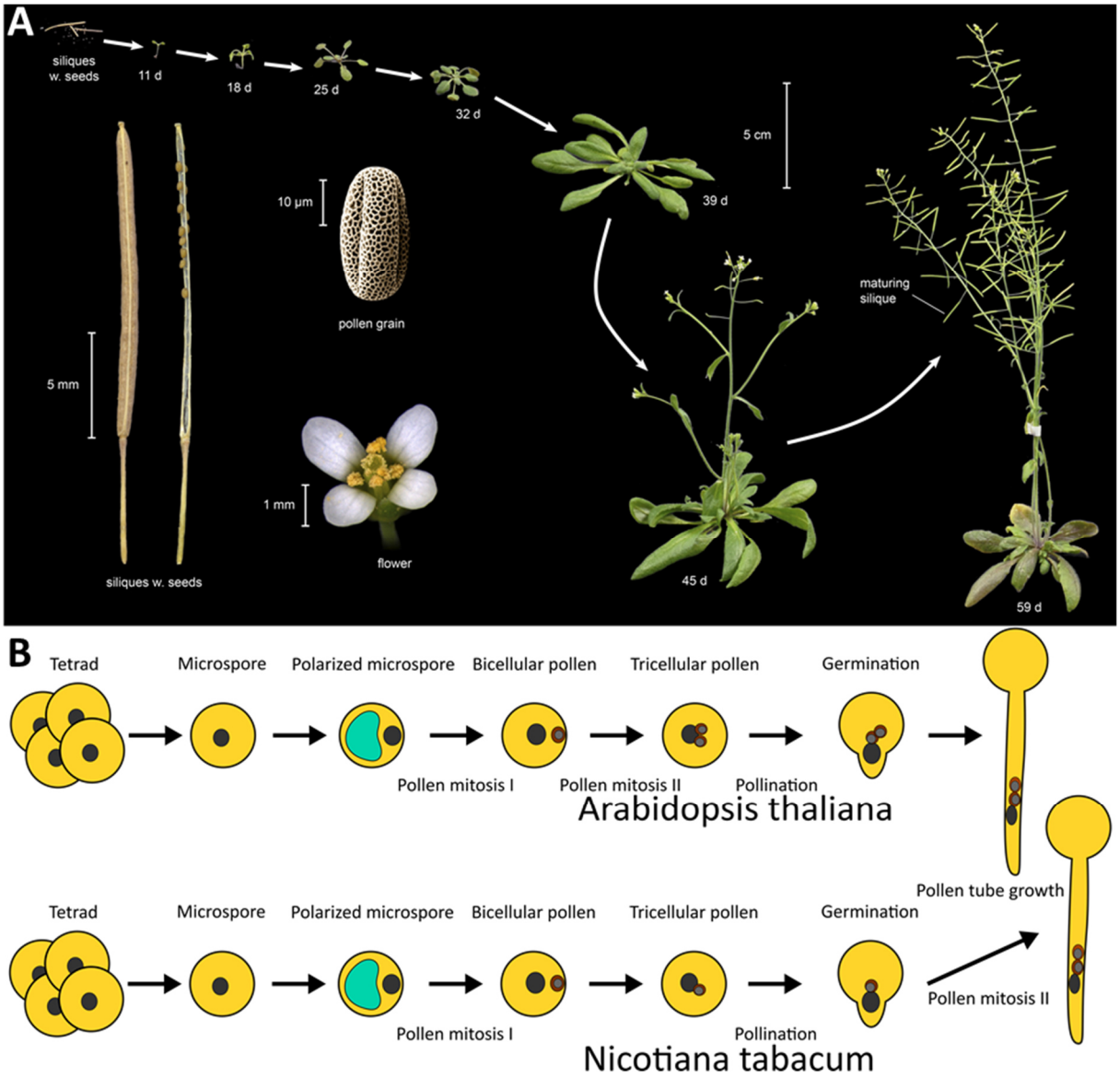

Figure 1: Life cycle of Angiosperms. (A) Life cycle of *Arabidopsis thaliana* sporophyte from seeds to senescent plant with indicated age as number of days. Detailed view on Arabidopsis pollen grain, open flower and silique is shown in left lower section. **(B)** A schematic picture of *Arabidopsis thaliana* (upper section) and *Nicotiana tabacum* (lower section) microgametogenesis and progamic phase. Developmental stages and division events are indicated. Note that the mature pollen of Arabidopsis and tobacco differ in the total number of cells as the timing of Pollen mitosis II is scheduled to different developmental stages. **(A)** is derived from (Krämer, 2015)

development are not translated immediately, but are stored in an inactive form. After pollination, such complexes of non-translating ribosomes are specifically activated during pollen activation, germination and in different phases of pollen tube growth and pollen tube guidance to the ovule. Moreover, these particles are transported across the pollen tube and translated locally at specific parts of the pollen tube (Urquidi Camacho et al., 2020).

Pollen tube is guided to the ovule all the way from the germination event to double fertilization event. Transmitting tract tissues are specifically differentiated tissues that guide the growing pollen tube through the pistil, while ovules themselves secrete a number of

chemoattractants to guide the tube from pistil to the ovary and to the ovule. When the pollen tube enters the embryo sac through the micropylar point, it bursts into one of the synergid cells and releases two sperm cells. One sperm cell nucleus fuses with the egg cell nucleus, the second with the two nuclei of the embryo sac to produce triploid endosperm. This process is called double fertilization and it is a phenomena specific for Angiosperms (Honys, 2018).

Translation and translation initiation

Translation cycle and its components

Translation is the final step of the central dogma, in which the sequence of four nucleotide residuals in nucleic acid are translated into the sequence of 20 amino acid residuals in peptides, following the rules of the genetic code. Among all eukaryotes, translation is a highly conserved, multistep, multifactorial process that requires a great number of proteins with structural, enzymatic and regulation function and consumes a great amount of the cell energy. Moreover, three types of RNA play key informational, structural and functional roles in protein synthesis as well: mRNA, tRNA and rRNA.

Eukaryotic ribosome is a large ribonucleoprotein particle (80S), consisting of rRNAs and as many as 80 Ribosomal Proteins (RPs). One 80S ribosome consists of two ribosomal subunits; small 40S and large 60S. Small 40S subunit, more important for the proper translation initiation, contains one 18S rRNA molecule and around 30 additional ribosomal proteins. Assembled 80S ribosome possess three sites crucially important for its function during all phases of the translation: aminoacyl-tRNA binding site (A site), peptidyl-tRNA binding site (P site) and exit site (E site). Both subunits contribute to all three sites with their intersubunit side. mRNA molecules carry the sequence information from the genomic DNA to cytoplasm. Canonical mRNA is capped with 7-methylguanosine on its 5′ end and tailed with poly-Adenine sequence on its 3′ end. mRNAs are always associated with proteins in ribonucleoprotein particles (mRNP). For proper translation, mRNAs have to be activated by the eIF4 complex and multiple copies of the cytoplasmic Poly(A)-binding protein (PABP) (Figure 2). tRNA are a group of non-coding RNA molecules with well-described secondary structure and various base modifications. tRNAs are specifically loaded with amino acid residues and are crucial in the translation, ensuring the correct decoding of the codon to its respective amino acid. Detailed overview of all components of translation machinery in plants is given here (Browning and Bailey-Serres, 2015).

Translation cycle has four stages; initiation, elongation, termination and recycling (simplified in Figure 2). During initiation, pertained by the so-called scanning mechanism in eukaryotes, the initiating Met-tRNA$_i$ needs to be activated by its binding to the eIF2 complex to form so-called Ternary Complex (TC). TC then joins to the 40S subunit together with eIF1, eIF1A, eIF3 and eIF5. This complex is referred as 43S Pre-Initiation complex (43S PIC). Assembled 43S PIC is loaded with mRNA activated by capping protein eIF4E, helicases eIF4A and eIF4B, scaffolding protein eIF4G and PABP. After mRNA joining, 48S PIC is formed and scanning of the mRNA begins. The mRNA is unwinded from the 5´UTR until start codon in Kozak context is reached. After AUG recognition, eIFs are released and 60S subunit joins the 40S with the help of eIF5B, forming the 80S initiation complex (80S IC).

During elongation, factor eEF1A ensures proper decoding of the next codon by placing a correct aminoacyl-tRNA in the A site. Then, a new peptide bond is formed, followed by a peptydityl-tRNA translocation to the P site with the help of eEF2. Former P-site tRNA, now without amino acid, translocates to the E site, where it leaves the ribosome. The elongation process is repeated with each codon in the mRNA. When stop codon (UAA, UGA, UAG) appears in the A site, eRF1 and eRF3 bind to the A site to release the polypeptide from the 80S ribosome. Beginning of the ribosome recycling is also pertained by the eRF1 together with ABCE1 (ATP-binding cassette 1), and the process involves separation of ribosomal subunits and release of the mRNA. Detailed overview of all stages of translation in plants is given here (Browning and Bailey-Serres, 2015).

Translation initiation: P_{out} and P_{in} PIC conformations

Translation initiation has been described functionally and structurally in detail here (Valášek, 2012; Hinnebusch, 2017; Hashem and Frank, 2018), while known plant specificities are reported here (Browning and Bailey-Serres, 2015). Majority of eukaryotic mRNA translation is initiated by the scanning mechanism, where recruited mRNA is scanned from the 5´cap base-by-base, seeking for codon-anticodon complementarity by the PIC with Met-tRNA$_i$ being pre-loaded in the P-site of 40S. The whole process is accompanied by many structural rearrangements of the PIC (Reviewed in Guca and Hashem, 2018). During translation initiation, 40S subunit together with eIFs occur in two conformational states; open conformation (P_{out}) and closed conformation (P_{in}). Until start codon recognition, the PIC is held by the eIFs in P_{out} state, a conformation prone to scanning of the mRNA, where Met-tRNA$_i$ placement in P site has a little shifted orientation and is sub-optimal as opposed to P_{in}. The shift of conformations

is triggered after the start codon recognition. Then, the PIC shifts to P_{in} state that is restrictive for further scanning of the mRNA. In P_{in}, IFs dissociate from the PIC, leaving the 40S subunit with Met-tRNA$_i$ tightly placed in P site, with its anticodon bound to the start codon of translated mRNA.

Translation initiation: eIFs of the 43S PIC

The key factor in the initiation is the eIF2, a heterotrimeric protein complex composed of three subunits (eIF2α, -β, -γ). eIF2 is a small GTPase that binds the Met-tRNA$_i$ in eIF2-GTP state. eIF2-GTP-(Met-tRNA$_i$) is also named as the Ternary Complex (TC) that binds to the P-site position of free 40S subunit. The GTP hydrolysis is activated by its GAP factor, the eIF5 and occurs after start of the mRNA scanning. However, the hydrolysed phosphate (P_i) is restricted to leave the eIF2-GDP by eIF5-CTD, eIF1 and eIF1A at non-AUG triplets, and remains bound to the PIC until AUG recognition. After AUG recognition, Met-tRNA$_i$ shifts to the P-site, changing the PIC to P_{in} state. The P_i is then released and eIF2-GDP leaves the 40S, keeping the Met-tRNAi tightly in P site and the PIC in P_{in} conformation (Reviewed in Hinnebusch, 2017). Another factor, eIF1, binds the 40S subunit near P site and enhances the TC binding to the 40S and favours the P_{out} state before AUG is in the P site. After start codon recognition, eIF1 dissociates from the PIC, easing the P_i release and the transformation into P_{in} conformation. Factor eIF1A binds to the 40S near A-site and enhances the TC binding, probably by promoting P_{out} state together with eIF1.

Largest initiation factor, eIF3, is a multisubunit complex composed of 12 individual subunits (a, b, c, e, f, g, h, i, k, l, m) and forms multiple interactions with most if not all other IFs and various interaction are mapped along the 40S. The structure, known interactions and other eIF3-specific features are discussed below. In both yeast and mammals, a protein named eIF3j is necessary for eIF3 binding to 40S but is only loosely attached to the eIF3 complex and not present in the purified yeast eIF3 complex (Phan et al., 2001). Despite its name, eIF3j is currently not considered as a proper eIF3 subunit and has been characterized as an eIF3-associated factor. The function is still not well determined, but it is known that eIF3j promotes the assembly of PIC (Valášek et al., 2017).

eIF5 is a GTPase activating protein (GAP) of the eIF2-GTP. After mRNA joining, eIF5 triggers GTP hydrolysis, but sterically restricts P_i to be released from eIF2-GDP-Pi during scanning for the start codon. eIF5 also interacts with eIF4G during mRNA recruitment to the 43S PIC.

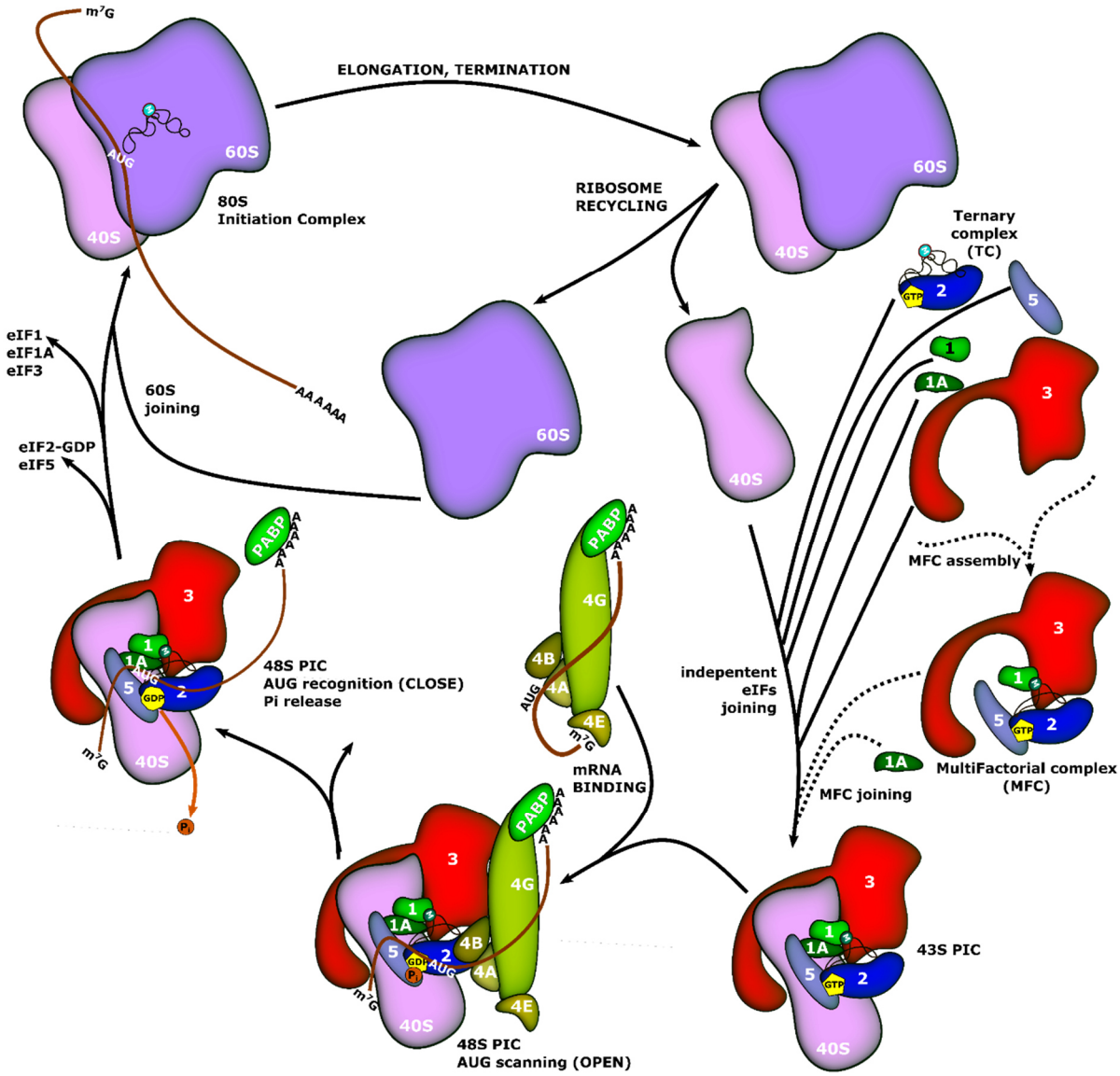

Figure 2: Model of the Scanning Mechanism of Eukaryotic Translation Initiation shown as a series of steps. After ribosomal subunits separation, the 40S (pink) subunit is joined by eIFs to form 43S PIC. eIFs are joining the 40S individually or as a large MFC complex. Activated mRNA (brown) bound to eIF4s is then joined to the 43S PIC to form 48S PIC. GTP hydrolysis on eIF2 starts subsequent scanning of the mRNA by unwinding the mRNA by eIF4A. AUG recognition shifts the complex to its closed conformation, triggering the release of eIF1, P_i, and eIF5-eIF2-GDP. Releasing of eIFs enables joining of the 60S subunit to the PIC. Formed 80S IC is then ready for elongation. After termination, 80S ribosome is disassembled again. Simplified scheme is based on the Figure 1 in (Browning and Bailey-Serres, 2015).

Translation initiation: The Multi-Factorial Complex

The Multi-Factorial Complex (MFC) is a supercomplex of initiation factors that assemble freely in cytoplasm in 40S-independent way and after that is able to bind to 40S. MFC comprised from most IFs; eIF1, eIF2-GTP-Met-tRNAi (TC), eIF3, and eIF5. Its formation is induced by phosphorylation of the factors and the inability to form MFC decreases cell growth rate in yeast. MFC was found in yeast, mammals and plants (Asano et al., 2000; Dennis and Browning, 2009; Sokabe et al., 2012).

Translation initiation: The mRNA-activating eIF4s

eIF4s are a class of eIFs that activate the mRNA and brings it to the 43S PIC (Villa et al., 2013). eIF4A is an ATP-dependent RNA-helicase that helps unwinding the mRNA during the scanning process. eIF4B is an enhancer of the eIF4A, important for structured 5´UTRs. eIF4E is a 5´cap binding protein. By capping the mRNA, it recruits the eIF4G and eIF4A to the 5´ end. PABP binds to the poly-A tail on 3´ of mRNA. eIF4G is a large scaffold protein, able to bind the eIF4E, eIF4A/eIF4B and PABP. Together, eIF4A, eIF4B, eIF4E and eIF4F form a supercomplex recognized as eIF4F in the literature. Interestingly, plants encode two functionally different versions of the eIF4E and eIF4G that operate on different mRNA pools (Browning and Bailey-Serres, 2015).

The eIF3 complex

The eIF3 complex is the largest initiation factor that is composed of multiple different protein chains. The eIF3 complex was isolated from rabbit reticulocytes in the 1970´s and since, research revealed its subunit composition, core functions and structural insights (Reviewed in Cate, 2017; Valášek et al., 2017). With the rising number of characterized subunits coming from different organisms, unification of the subunit nomenclature was proposed instead of using respective molecular weights or abbreviations derived from the mutant phenotypes (Browning et al., 2001; Burks et al., 2001). The canonical function of eIF3 is to promote initiation machinery assembly, providing a scaffold that forms numerous interactions with almost all factors involved. Although the complex was found present in all eukaryotic organisms so far, mammalian eIF3 and budding yeast eIF3 complexes were examined in most of the studies. In plants, several subunits were characterized with interesting mutant phenotypes and specific functions (Reviewed in Raabe et al., 2019).

Subunit composition of the eIF3 complex

In higher eukaryotes, eIF3 is composed out of 12 subunits (a, b, c, d, e, f, g, h, i, k, l, m). Homologues of all 12 subunits were in mammals, plants and also in *Neurospora crassa*, (Valášek et al., 2017). Not all organisms kept the complete set of twelve subunits. In *Excavata* supergroup, generally speaking, computational research of eIF3 composition suggest similarities to the mammalian model with the evidences of missing presence of some eIF3 subunits in several of inspected *Excavata* species proteomes (Rezende et al., 2014; Han et al.,

2015; Meleppattu et al., 2015). The most reduced eIF3 complex identified so far is also the best understood model of the budding yeast eIF3. *Saccharomyces cerevisiae* possess an eIF3 complex composed of only five subunits (homologues of the mammalian eIF3a, -3b, -3c, -3i, and -3g) (Asano et al., 1998; Phan et al., 1998; Khoshnevis et al., 2012). Even though so reduced, the yeast complex is still capable of performing the canonical eIF3 functions, suggesting on the essential subunits that pertain the basal functions. However, some eIF3 subunits that are not presented in the minimalistic budding yeast complex are still essential in other organisms (Valášek et al., 2017). Moreover, experiments suggest that partial eIF3 sub-complexes could be formed, are stable, functioning in the initiation and that they exist *in vivo* (Smith et al., 2016). In plants, for example, eIF3 complex lacking the subunit eIF3h is functional in *Arabidopsis*, but impairs eIF3h-specific function in reinitiation after short upstream ORFs, which is an important post-transcriptional regulation pathway, and leads to pleiotropic growth defects (Kim et al., 2004; Kim et al., 2007). Plant eIF3 is similar to the mammalian 12-subunit complex, ass plant genomes possess encode full set of genes for the 12 eIF3 subunits (Browning and Bailey-Serres, 2015; Raabe et al., 2019).

eIF3 structure and assembly

The 12-subunit model of eIF3 could be divided into two structural modules: 1) rigid Octamer composed of eight subunits (a, c, e, f, h, k, l and m) and 2) flexible part composed of three subunits (b, i and g), named Yeast-Like Core (YLC) (Figure 3). YLC is named after the conserved functional core in the reduced budding yeast complex. Both modules are bridged by the connection of YLC to the C-terminal end of the eIF3a subunit.

The formation of the Octamer is enabled by a scaffold structure of 6 subunits (a, c, e, k, l and m) containing a PCI (stands for Proteasome, CSN, eIF3) domain together with 2 subunits (h and f) containing a MPN (stands for Mpr1-Pad1-N-terminal) domain (Siridechadilok et al., 2005; Querol-Audi et al., 2013). The structure of the Octamer is stabilized by two interaction hubs; first being the six PCI domains forming a structure known as β-sheet arc, and second being a helical bundle in which every subunit of the Octamer joins with one of its C-terminal α-helix (des Georges et al., 2015). The architecture comprised of the 6xPCI/2xMPN domains and the α-helical bundle is shared between three functionally unrelated protein complexes; eIF3, COP9 signalosome (CSN) and the lid of the 26S proteasome (Pick et al., 2009; Enchev et al., 2010; Ellisdon and Stewart, 2012). All these octameric complexes, however, form another structural feature, an alpha helical bundle, in which each of the eight subunits contributes with

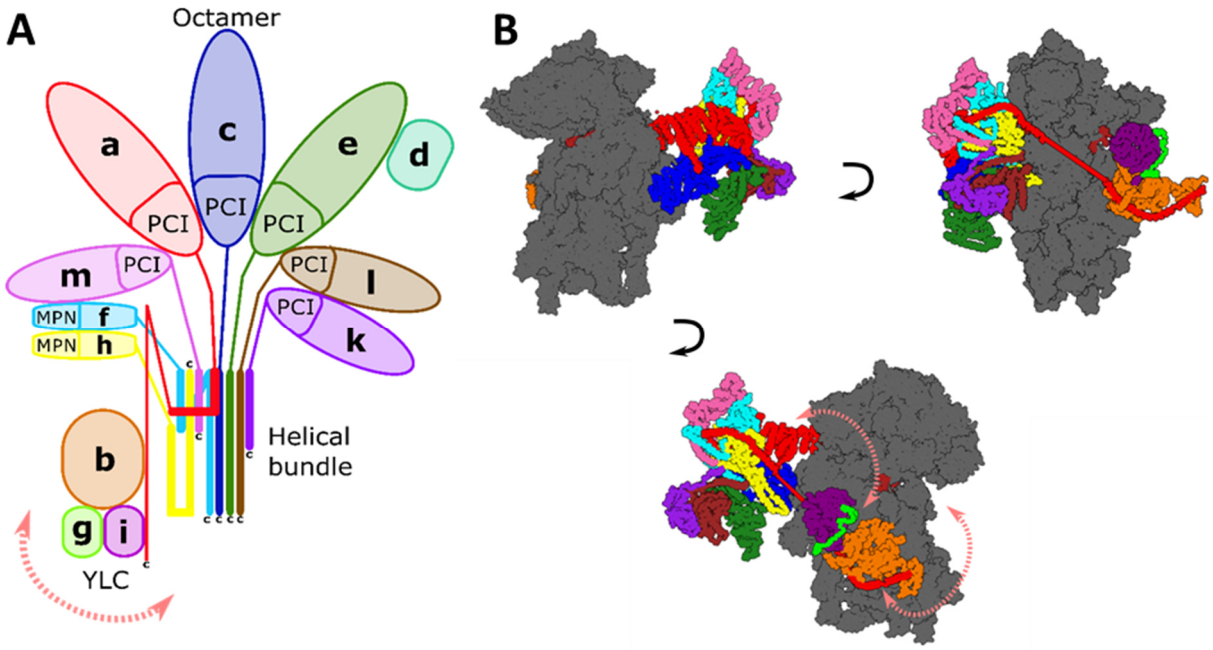

Figure 3: Mammalian eIF3 complex. (A) Simplified cartoon model of the 12-subunit mammalian eIF3 **(B)** Series of figures depicting position of the eIF3 on the 40S subunit. The Octamer module binds to the solvent side of the 40S ribosomal subunit. Via flexible C-terminus, eIF3a binds subunits of the YLC, a module that undergoes multiple structural rearrangements throughout the initiation phase. Pink arrows represent the flexibility of the YLC. eIF3a and eIF3g structures are incomplete, as the flexibility of the YLC makes it difficult to visualize the connection between the Octamer and YLC. eIF3 position on the ribosome was assembled using USCF Chimera from mammalian 48S-PIC (PDB code: 6FEC), combined with incomplete yeast YLC that was aligned to partial mammalian eIF3b (PDB code: 6FEC) from yeast 48S-PIC (PDB code: 6FYY) using USCF Chimera MatchMaker tool. The cartoon model modified from (Raabe et al., 2019).

one alpha helix. In the case of eIF3, this helical bundle was proposed as the first and main contributor to the assembly of the eIF3 itself, whereas the PCI/MPN domain is more necessary for the perfect final mutual orientation of the Octamer subunits. The second module, YLC, consists of subunits eIF3b, i and g that are connected to the flexible C-terminus of the eIF3a, which enabes increased plasticity of the YLC relative position with respect to the Octamer. eIF3b and eIF3i possess another protein-protein interaction domains: 9- and 7-bladed WD40 β-propeller, respectively. Moreover, eIF3b and eIF3g possess an RNA-binding domain (RRM), which in this case do not form RNA-protein interactions, but protein-protein interactions (with eIF3a) (Reviewed in Valášek et al., 2017). The YLC is one of the most difficult structural challenges that remains to be solved within the translation initiation, as it was captured by few Cryo-EM studies of the PIC only as a partial density (For example see Eliseev et al., 2018). Moreover, the position of the YLC within the PIC varies on the stage of the PIC captured. Therefore, the YLC is considered the flexible part of the eIF3 complex.

The assembly of the eIF3 is started by the eIF3a binding to eIF3b as the first nucleation core, then subunits of the YLC module (i, g) join together with eIF3c and form a subcomplex closely resembling the complete yeast complex. Moreover, in mammals, this subcomplex already binds

40S and joins activated mRNA *in vivo* (Wagner et al., 2016, Wagner et al., 2014). Assembly of other PCI/MPN domain containing subunits of the Octamer then proceeds. The last, twelfth subunit eIF3d, was found to be bound to the eIF3e subunit (Wagner et al., 2016). Although the proposed -3a/-3b nucleation dimer is needed for the eIF3 assembly, the whole human PCI/MPN Octamer (e.g. without eIF3b) was succesfully reconstituted *in vitro* (Sun et al., 2011), suggesting that the eIF3 assembly hierarchy might not be that strict. Recent studies in budding yeast showed that eIF3 subunits might assemble co-translationally on the ribosome that translates mRNA of eIF3 subunit (Wagner et al., 2020). In addition, such co-translational assembly might occur also for the whole MFC complex (Wagner et al., 2020).

As partial structural model of the eIF3 has been mostly solved in different contexts of PIC by Cryo-EM, we have very poor structural information about how eIF3 complex behaves freely in cytoplasm, especially in mammals. In recent budding yeast eIF3 study, cross-linking experiments suggested that the reduced 5-subunit eIF3 has a globular structure, where eIF3a-eIF3c PCI domains do not interact with each other and the YLC is tightly packed to the eIF3a-eIF3c dimer (Zeman et al., 2019). Moreover, we still lack the exact atom models of most of the eIF3 subunits. From budding yeast, there are only partial atom models for all five yeast subunits. The 12-subunit mammalian eIF3 structural model is based on Cryo-EM densities of the whole complex, where the models of individual eIF3 subunits were fitted, using different sources as template: 1) the partial yeast atom models of eIF3a, eIF3b, eIF3c, eIF3i, eIF3g; 2) atom model of eIF3k; 3) homology modelled eIF3 subunits based on structurally/architecturally related subunits of the COP9 signalosome protein complex and the 26S proteasome lid protein complex (Erzberger et al., 2014; des Georges et al., 2015).

eIF3 complex within the PIC

In eukaryotic translation initiation, the basal eIF3 function is to scaffold all other factors together with the 40S subunit and the mRNA. After binding to the 40S subunit, eIF3 enhances joining of the TC, eIF1, eIF1A and eIF5 (Valášek, 2012). Moreover, eIF3 also participates in the forming of the MFC (Sokabe et al., 2012). In budding yeast, the tightly packed globular architecture of free eIF3 undergoes large structural rearrangements during the attachment to the 40S, in which PCI domains of eIF3a and eIF3c form the PCI-PCI dimer near mRNA exit channel on the solvent-side of the 40S (Zeman et al., 2019). The YLC unpacks from the eIF3a/eIF3c and, still attached to the eIF3a, is able to move more to the intersubunit side of 40S like a mechanical arm (Figure 3B). The position of YLC probably changes during different

phases of the initiation process (Zeman et al., 2019). Recent structural study done with the reconstituted yeast eIF3 complex suggested that also the N-terminal part of eIF3c interacts with eIF3b on the PIC, which would mean that eIF3 could form a complete circle around the small ribosomal subunit at some stage of the initiation process at least (Llacer et al., 2018). In mammalian PICs, the Octamer module is bound to the 40S on the solvent side with eIF3a-NTD reaching the mRNA exit channel, and YLC module is placed on the intersubunit side near the mRNA entry channel (des Georges et al., 2015), where interactions with the 40S are maintained by a 9-bladed β-propeller of eIF3b (Liu et al., 2014). The recruitment of mRNA to PIC is enabled by the direct binding of the eIF4G to a surface made from eIF3 subunits; eIF3c, eIF3d and eIF3e (Villa et al., 2013). eIF3a interacts directly with the mRNA at the exit channel, stabilizing the mRNA on both entry and exit channels (Aitken et al., 2016). The whole YLC module undergoes structural rearrangements, relocating from the solvent exposed site to the intersubunit surface and stabilizing the interaction of the scanning complex, while upon AUG recognition, YLC is relocated back to the solvent exposed site (Simonetti et al., 2016; Eliseev et al., 2018; Llacer et al., 2018; Zeman et al., 2019). The latter rearrangement seems to be triggered by conformational change of eIF2 (Eliseev et al., 2018).

Apart from translation initiation, additional functions were desribed for eIF3 during translation termination, recycling and during some non-canonical translational events such as reinitiation after upstream open reading frames (uORFs), stop codon read-through, ribosome recycling or in the nonsense-mediated mRNA decay (Reviewed in Valášek et al., 2017).

The subunit A of the eIF3 complex

eIF3a is the largest subunit of the eIF3 complex, being an essential part of the minimal functional core of eIF3 that still can execute the function in translation initiation. eIF3a is considered an essential house-keeping gene. This is in agreement with the findings that, at least in mammals, expression of eIF3a is usually present in all organs and tissues, although on different levels, being most abundantly expressed in proliferating and developing tissues (Pincheira et al., 2001b). Dysregulated expression of eIF3a is commonly found in cancer tissues (Reviewed in Yin et al., 2018)

Protein evolution

A homologue of eIF3a subunit is found in most eukaryotes, so far with the exception of some *Excavata* species, where a *bona fide* homologue was not found *in silico* (Rezende et al., 2014).

Noteworthy is the difference in size of the eIF3a homologues from different species. For example in yeast (Figure 4A), eIF3a subunit is approximately 1000 amino acid residues long, whereas the mammalian eIF3a subunit is longer, reaching to 1350 amino acid residues. In *X. laevis*, the eIF3a reaches even 1420 amino acid residues (Uniprot, search for "eif3a AND reviewed:yes"). It seems that the growth in size is caused by additional C-terminal domain in animals. The additional few hundred amino acid residues at the C terminal end are composed of Glutamine- and Asparagine-rich repeats of 10 amino acids that are repeated 21-25x in the protein sequence chain, forming a so-called DRYG domain. It was proposed that this region interacts with phosphorylated eIF4B as a part of the interferon-inducible signalling (Méthot et al., 1996). However, since interferons are associated only with immune system of vertebrates, it is not surprising that the DRYG domain is not present in non-vertebrate species such as yeast and plants.

Structure and domains

The eIF3a protein is mostly helical polypeptide with two characteristic domains: PCI domain and Spectrin-binding domain. The DRYG domain present in vertebrates is discussed in the previous subchapter. The PCI domain (Ineterpro domain ID: IPR000717; also called PCID or PINT motive) is around 190 amino acid residues long alpha helical structural domain residues that is well recognized by domain search software. The PCI domain is found in multisubunit complexes, such as CSN, 19S proteasome lid and eIF3 complex (Ellisdon and Stewart, 2012). In these large complexes, six subunits that interact with each other contain the PCI domain, forming a partial horseshoe architecture. PCI domain is also associated with the correct assembly of these relatively complicated protein complexes. PCI domain is composed of two main subdomains, more conserved C-terminal WH subdomain, that is necessary for PCI:PCI interactions, and N-terminal helical repeat region. The latter region is generally less conserved part of the PCI domain with hardly interpreted boundary, causing differences in PCI domain length characzterized by different databases such as PROSITE, Pfam and SMART. The WH subdomain has an "αβααββ" arrangement, where the β-strands are very short and form an antiparallel sheet (Scheel and Hofmann, 2005). Normally, PCI domain is localized to the C-terminal end of its protein. However, PCI of the eIF3A protein is situated more to the middle region of the protein. That might be due to the overall longer protein chain in eIF3a than other PCI-containing proteins.

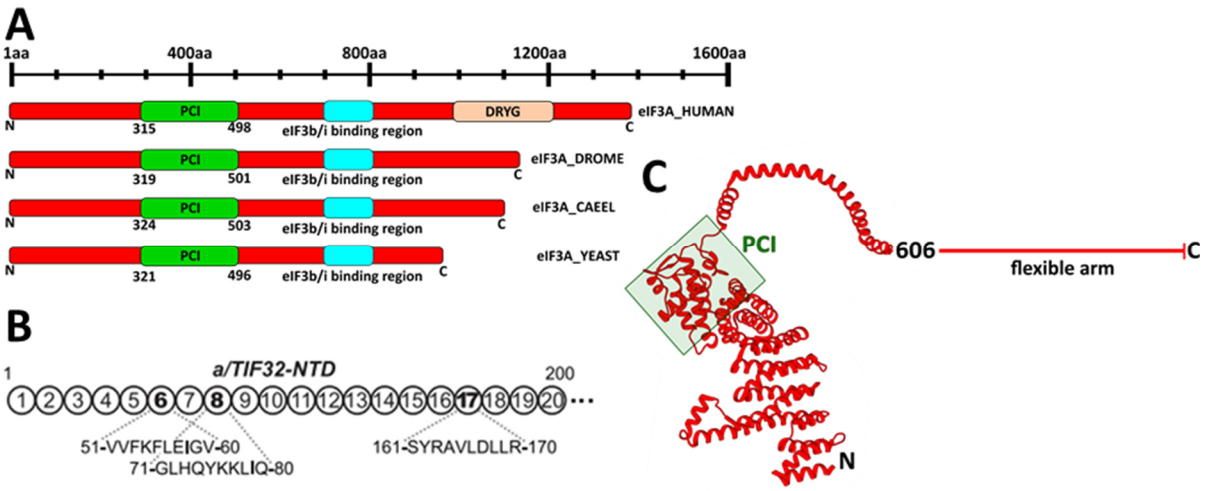

Figure 4: eIF3 protein structure and domain organisation. (A) Graphical visualization of eIF3A protein chain lenght and domain organisation in four different eukaryotic species (For Uniprot asccention codes, please see Table 6). PCI domains and eIF3b/i binding regions are shown at their relative position. DRYG repeat region is depicted in the human sequence. **(B)** 5´UTR interacting region characterized in budding yeast, taken from (Munzarová et al., 2011) **(C)** Partial eIF3A structure from mammalian 43S PIC structure with highlighted PCI domain (PDB code: 5A5T), visualization of the protein structure was done using the USCF Chimera.

More to the C-terminal end of eIF3a, the binding domain for eIF3b and eIF3i is present. Based on sequence homology, it was named Spectrin domain. Older studies suggested, based on the sole presence of Spectrin domain, some association between eIF3a and cytoskeleton that were further backed up by evidences of eIF3a localization on membranes and even binding to actin itself (Pincheira et al., 2001a). However, any of these findings were not further investigated and from the beginning of the 21st century, no evidences of further development in this field are apparent in the literature. As already mentioned, this domain was found responsible for assembly of the eIF3 (a:b:i:g) subcomplex formation. In this subcomplex, eIF3b and eIF3i bind directly to eIF3a, both on a very short sequence region; amino acid residues 756– 765 for eIF3b and 766–775 for eIF3i (Dong et al., 2013), while eIF3g does not interact with eIF3a directly and joins the subcomplex via interaction with eIF3b. The single amino acid residue Phe760 in the Spectrin domain is essential for the interaction between eIF3a and eIF3b. Sequence alignment showed that this residue is completely conserved across species (Dong et al., 2013). Within the binding site of eIF3i there are three conserved residues in mammals (Arg768, Tyr772, and Lys775), that could be essential in contributing to the binding by eIF3i (Dong et al., 2013) However, Dong et al. included in his alignment only species from animal kingdom.

Known interactions and functions

As it comes to interactions of eIF3a in the context of the eIF3 complex, eIF3a interacts with most, if not all other eIF3 subunits. This is in the agreement with the α-helical interaction bundle that serves as an intersubunit interaction hub of the Octamer. Exceptions might be the eIF3d that binds eIF3e peripherally and the eIF3g that is joined to the YLC by interactions with eIF3b, as mentioned earlier. The eIF3a interacts with the ribosome as a part of the eIF3 that contacts the ribosome on multiple sites (Reviewed in Hinnebusch, 2017). Interestingly, it was shown that NTD of yeast eIF3a interacts with RPS0A *in vitro* near the mRNA exit channel and CTD of the yeast eIF3a interacts with 18S rRNA helices 16–18 of domain I in 18S rRNA near the mRNA entry channel (Valásek et al., 2003). Additionally, eIF3a was also found crosslinked to 5′ UTR of -14/-17 nucleotides downstream from AUG, stabilizing the mRNA-PIC complex at the exit channel (Pisarev et al., 2008). This is altogether in the agreement with the structural findings suggesting eIF3a extends the mRNA exit channel for a few interactions between the translational machinery and the mRNA (Aitken et al., 2016). eIF3a is also involved in the reinitiation after short uORF in the 5′UTR in yeast, where it interacts with specific secondary structures of the mRNA, specifically with amino acid regions 51-60, 71-80, 161-170 (Szamecz et al., 2008; Munzarová et al., 2011) (Figure 4B). eIF3a and eIF3g were proven to regulate scanning for AUG and also promoting the shift from open to close conformation (Chiu et al., 2010). eIF3a and eIF3c also interact with RNA in some Internal-Ribosome-entry-sites (IRES) elements (Reviewed in Valášek et al., 2017).

Arabidopsis *AteIF3A* gene organisation, expression and general functions in plants

In Arabidopsis, eIF3a is encoded by a single copy gene. *AteIF3A* gene is located on the reverse strand of the chromosome 4 (AGI code At4g11420). The gene is 5004 bp long and contains 13 introns and 14 exons (Figure 5). The mRNA is 3485 bp long, giving a protein product of 987 amino acid residues and molecular weight 114 kDa. No alternative splicing of *AteIF3A* gene was reported (Gene organisation data are obtained from TAIR, Berardini et al., 2015). Based on available transcriptomic data, strongest expression of the *AteIF3A* gene is located to both plant meristems, imbibed seeds, ovary tissues and developing pollen during microgametogenesis (Figure 5) (Klepikova et al., 2016). Up to date, no evidence of any plant mutant of the eIF3a subunit has been reported. AtEIF3A protein was recognized in several studies as an interacting partner with histone deacetylase GCN5 (Servet et al., 2008), as

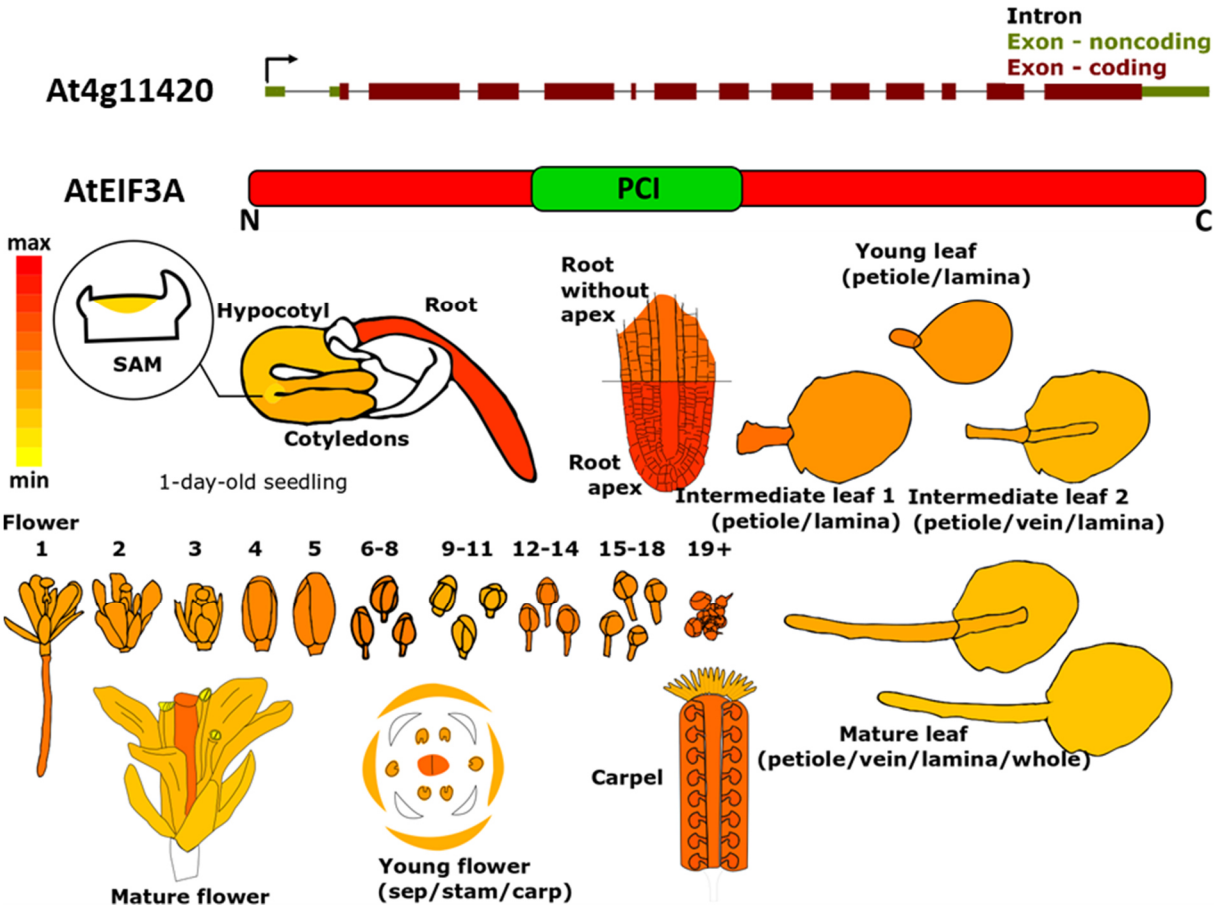

Figure 5: *AtEIF3A* gene structure, protein product and expression. Graphical visualization of *AteIF3A* gene organisation in *Arabidopsis* is given in the upper part of the figure. In the mid-part, an Uniprot-based domain organisation of the AteIF3A protein is shown. Lower part shows relative *AtEIF3A* (At4g11420) expression in major plant organs. The expression data were obtained from http://bar.utoronto.ca/eplant/ and is based on RNA-seq transcriptomics.

substrate of phosphorylation in the shift from light to dark (Boex-Fontvieille et al., 2013), as a ubiquitin-conjugate protein (Kim et al., 2013), and as part of the non-translating storage polysomes in tobacco pollen (Honys et al., 2009; Hafidh et al., 2018). Interestingly, AteIF3A was recognized as a mRNA-binding protein in various plant tissues (Köster et al., 2020). Another interaction was discovered with plant reinitiation supporting protein (RISP), which increases reinitiation of the *Cauliflower mosaic virus* (CaMV) mRNA (Thiébeauld et al., 2009). A Chinese study (available only in authors original language, with the exception of abstract) presented rice *OsEIF3A* as a gene expressed in various *Oryza sativa* tissues, strongly in root tips, leaf or in stigma, while its expression was induced by auxin (Li et al., 2003). Most interestingly, there was an attempt to complement the budding yeast eIF3a mutant (*rpg1-1* alelle) with *Zea mays ZmeIF3A* cDNA clone isolated from corn meristematic tissues (Sabelli et al., 1999). Although the complementation of the yeast mutant failed (similarly the attempt to complement it with human eIF3a also failed (Valášek et al., 1998)), this study analysed

expression of *ZmeIF3A* as well, using the radioactively labelled antisense *in situ* hybridisation and northern hybridisation. Strong signal was detected in the root meristematic region surrounding the central cylinder, young root, the male inflorescence, the developing cob and in the seeds (Sabelli et al., 1999).

Working hypotheses

Based on the summarized literature, we state following hypotheses for this work:

- eIF3a subunit is present in plants and is conserved with its respective non-plant orthologues
- *AteIF3A* gene is an essential house-keeping gene expressed in all cells and tissues that actively translate in Arabidopsis
- AteIF3A is cytoplasmic protein
- Modulation of *AteIF3A* expression would lead to observable alterations in plant viability

Materials and Methods

Plant material and cultivation conditions

Plant material

The main model plant organism used in this project is *Arabidopsis thaliana* ecotype Colombia-0. Arabidopsis was used for most experiments in this work. Occasionally, male sterile mutant *ms1-1* line of *Arabidopsis thaliana* ecotype Landsberg was used for the aniline blue staining experiments, as there was no need to emasculate the flowers to avoid undesired self-pollination in the experimental procedure. Four mutant *Arabidopsis thaliana* lines obtained from NASC were investigated: SALK_123762, SAIL_124_G07, SAIL_227_B02 and SAIL_545_B08. Lines SAIL_124_G07 and SAIL_227_B02 have *quartet (qrt)* mutant background, therefore their pollen grains do not separate from the tetrads. *Nicotiana tabacum* (L. cv. Samsun) was used for stable tobacco transformation experiments.

Plant cultivation conditions in vitro and in soil

With almost no exception, all seeds of both Arabidopsis and Tobacco were first germinated *in vitro*, then transferred to soil. Seeds were first sterilized and germinated on MS media with 0.8% agar (Sigma) of following strengths. For Arabidopsis seeds, half strength MS salts (Sigma) with additional vitamin supplements (2.2 g/L MS basal salts; 100 mg/L myo-inositol; 500 mg/L MES; 0.5 mg/l Nicotinic Acid; 0.5 mg/L Pyridoxine·HCl, 1.0 mg/L Thiamine·HCl and 1M KOH to adjust pH to 5.7) is sufficient. For *Nicotiana tabacum* seeds, 2 strength MS media (Sigma) (8.8g/ MS media with vitamin supplement and 1M KOH to adjust pH to 5.7) was needed. The melted media was poured into plastic 9-cm diameter Petri plates or 12x12cm square plates and left to cool down and solidify. After sowing the seeds, plates were sealed by Micropore^{TM} tape (CM Company). Sown seeds were stratified by keeping them at 8 °C for 3 days. Then, 7-10 days old seedlings with developed cotyledons and first true young leaves were transferred to sterile soil (jiffy tablets) and placed either directly on white plastic tray or in a transformation pot. All plants were grown under long day (16 hours light, 8 hours dark) at 22 °C and 60% RHH either in a growth room or cultivation chamber. The recommended addition of 5 g/L sucrose to the media has little or no effect on the autotrophic growth under light for the short period of time needed for germination. Moreover, addition of sugars causes faster

yeast/fungi contamination of *in vitro* cultures, therefore the sucrose was mostly omitted from the media.

Seeds harvesting, sterilization and sowing

When siliques started to turn brown and yellow, plant were moved to maturation room to dry out. Seeds were harvested from completely dried plants by releasing the seeds with pressing dried siliques between fingers on a paper. The remnants of siliques were removed from the paper using forceps or gentle mouth-blow. Seeds were collected to a 1.5 mL Eppendorf tube and left open on the bench for 3-14 days to prevent fungi contamination by additional drying. Then, harvested seeds were kept in -20 °C for 2-3 days to prevent insect contamination and stored at room temperature. For sowing, an aliquot of seeds to be sown was taken from the stock into a separate 1.5 ml Eppendorf tube. In a laminar flow hood, the aliquot was first washed with 70% ethanol, mixed by shaking and incubated for 1 minute. Ethanol was then removed by pipetting or decanting and seeds were sterilized in 20% SAVO solution (sodium hypochlorite-based commercial disinfectant) for 5-10 minutes. After removing the SAVO solution by decanting, seeds were washed 5 times with sterile freshly autoclaved tap water cooled down to room temperature. Sterile seeds were placed on media plate one by one using pipette with cut yellow 200 µL tip. This sterilization protocol was used successfully for both Arabidopsis and Tobacco seeds without any observable negative effect on the germination.

Plant crossing

In order to cross different lines into combined progeny or to outcross the mutant line, we needed to apply pollen from one parent (father) on a stigma from the second parent (mother). *Arabidopsis thaliana* and *Nicotiana tabacum* flowers normally self-pollinate to produce genetically identical progeny. Therefore, we needed first to remove anthers from unopened flowers of the mother. In order to accomplish the procedure in the case of *Arabidopsis*, working with binolupe was inevitable. Using the forceps, young unopened flower buds were emasculated by gentle excision of one sepal and one petal and subsequent removal of all anthers from the bud. Using this approach, the stigma was still covered and protected by remaining sepals and petals and left to mature for 2 days. Post-anthesis anthers with mature pollen were excised from the male parent line‘s flower and used to pollinate the matured pistil. Plants were then kept until the crossed silique was mature and ready to be collected. For experiments where seeds were not needed for further experiments, for example the aniline blue staining, *ms1-1*

mutants were used as the mother, as no need for anthers removal saves a lot of time. In case of *Nicotiana tabacum*, the flowers are relatively big and use of binolupe is not needed. During the corolla lobes opening, anthers were not yet opened and could have been removed using forceps. Mature pollen was released by rubbing post-anthesis anthers held in forceps against the visibly moist mature stigma. In several cases, we used this procedure to pollinate stigmas with pollen from another flower of the same individual plant, as some regenerated transgenic individuals sometimes encountered problems with self-pollination.

In vitro pollen tube growth

Arabidopsis pollen could be germinated and pollen tubes could be cultivated *in vitro*. Pollen germination media was used with agarose to make small agarose pads on the sample glass, where gently rubbing of open flowers released and stacked pollen grains to the solidified media. The germination media (1.6 mM H_3BO_3; 5 mM $CaCl_2$; 5 mM KCl; 1 mM $MgSO_4$; 10% sucrose; 1.5% low-melting agarose; 1M KOH to adjust pH to 7.5) was always prepared fresh by mixing the chemicals before adjusting the pH, subsequent addition of agarose and melting in the microwave. Cooled but not solidified media was then pipetted on the sample glass to form a 1x3 cm pads and left to solidify. Glasses with agarose pads were kept in a plastic box with wet cotton tissues to avoid drying of the media. At least 3 freshly opened flowers, each from different individual plant, were excised and rubbed against the agarose pad to release pollen on the media. Sample glasses with pollen were kept at 22 °C in plastic boxes with wet tissues, for at least 4 hours before pollen tubes were observed. The sample glass with agarose pad was mounted to the microscope for observation.

De novo regeneration of transformed N. tabacum plants by indirect organogenesis

In order to produce stable Tobacco transgenic lines from Agrobacterium-infiltrated leaf discs where only a few cells were transformed, we needed to regenerate whole plants from the leaf discs. We used *in vitro* indirect organogenesis on media containing growth regulators. The aim was to induce callus formation first, from which shoots emerge to be cut and transferred on media that favours rooting. After Agrobacterium-mediated leaf discs infiltration (protocol is given in its respective chapter below), leaf discs were transferred on Callus inducing media (CIM) (2 strength MS media with vitamins; 2% sucrose; 0.8% agar; 5 mg/L Naphthaleneacetic acid; 0.1 mg/L Benzylaminopurine) with addition of 250 mg/L cefotaxin in order to prevent additional growth of the Agrobacterium plus 50 mg/L expression cassette selection antibiotic

(kanamycin) to let only the transformed cells produce a callus. Plates were then covered by Micropore^{TM} tape (CM) and incubated in sterile growth room at standard conditions. Then, leaf discs were transferred on Shoot inducing media (SIM) (2 strength MS media with vitamins; 2% sucrose; 0.8% agar; 0.2 mg/L NAA; 0.2 mg/L Gibberellic acid ; 2 mg/l zeatin riboside with addition of 250mg/l cefotaxin and 50mg/l plant selection antibiotic). On SIM, explants were passaged every week for maximum of 11 weeks, during which calli grew in size and occasionally a shoot emerged. Shoots were cut when the stem was at least 3 cm long and small leaves were produced. Excised shoots were transferred into magenta boxes in Rooting media (RIM), which was plain MS medium (two strength MS media with vitamins; 2% sucrose; 0.8% agar). For about 4-7 weeks, shoots were incubated to develop a rich root system. Then, plantlets were carefully removed from the RIM, carefully cleaned from the leftover media and transferred into soil (jiffy tablets). In order not to stress the *de novo* regenerated plantlet too much by the direct transfer from *in vitro* into environment of the cultivation room, the plantlet was left for acclimation in a autoclaved glass jar with sterile hydrated jiffy tablet. The jar lid had a small hole filled with common kitchen sponge, therefore little gas-exchange was allowed. After 3-5 weeks, jar with healthy plantlet with around 6 green true leaves were opened for 2-3 days. After that, plantlet was removed from the jar and transferred to a black transformation pot with soil from jiffy tablets and placed in standard cultivation conditions described above. Occasionally, this T_0 generation was fed with flowering-inducing fertilizer in order to induce flower formation and seed production. Simplified scheme of this process is shown in Figure 6.

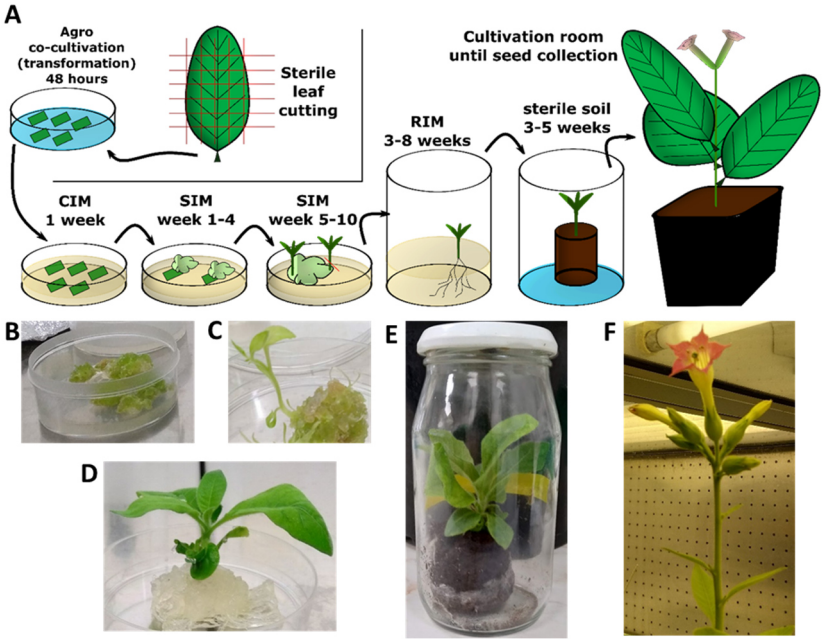

Figure 6: *De novo* regeneration of transformed *N. tabacum* plants by indirect organogenesis. **(A)** Schematic depiction of the *de novo* regeneration process that starts with *Agrobacterium*-mediated transformation of leaf discs and ends with transgenic individual plant in soil. Arrows indicate the order of the sequential process. Callus, Shoot and Root inducing media (CIM, SIM, RIM, respectively) indicate media composition (see text for further info) and an aproximate time scale is indicated as well. **(B-F)** Images of most but not all the steps shown in **(A)**.

Phenotype analysis

Sporophytic phenotype

Sporophytic phenotype was determined by direct visual observation of 7-day seedlings grown vertically in vivo on half MS media, as well as grown flowering individual plants against control grown in exactly same conditions.

Male gametophyte phenotype – pollen morphology

For observation of overall pollen morphology of inspected Arabidopsis mutant lines, 3 mature post-anthesis flowers were collected in 150 µL GUS buffer (0.1 M NaPO4 pH 7.0; 10 mM EDTA; 0.1% Triton) in Corning 96 well microscope plates and mounted in Nikon TE2000E inverted microscope. After 10 minutes, most of the pollen sedimented on the bottom of the well. Observation in bright field allowed to distinguish pollen shape defects from normal pollen. Generally, at least 100 pollen grains (but usually around 250) from at least seven plants of identical genotype background were counted.

Male gametophyte phenotype – DAPI staining – pollen anatomy

Mature pollen grain contains a determined number of cells (one vegetative and two sperm cells in case of Arabidopsis). In order to determine defects in pollen anatomy, DAPI fluorescent dye was used to specifically bind DNA, visualizing the vegetative and sperm cell nucleus. For observation of inspected mutant lines, 3 mature post-anthesis flowers were collected in 150 µL of GUS buffer with addition of DAPI (8 µL of DAPI solution in 10 mL of GUS buffer) in Corning 96 well microscope plates and mounted in Nikon TE2000E inverted microscope. After 30 minutes, DAPI diffuses throughout the pollen grain and stains the DNA. Observation under UV light allows to visualize the DAPI fluorescence and determine the number of nuclei inside the pollen grain. Generally, at least 100 pollen grains (but usually around 250) from at least seven plants of identical genotype background were counted.

Male gametophyte phenotype – Alexander´s staining - pollen viability

For observation of mature pollen viability of inspected *Arabidopsis* mutant lines, Alexander´s staining was used to distinguish viable and non-viable pollen grains. Alexander's staining colours cell walls of all pollen grains in blue-green, and viable and non-aborted pollen grains and spores in magenta-red. We used simplified protocol for Alexander´s staining (Peterson et

al., 2010), as the aim was only observation of pollen outside anthers and there was no point in using highly toxic fixatives and long fixation procedures. For the analysis, 3 post-anthesis flowers were tapped and squeezed in 40 µL of Alexander´s staining (10% v/v ethanol; 0.01% malachite green; 0.05% Acid fuchsin; 0.005% orange G; 4% v/v glacial acetic acid) sample glass, carefully plates and mounted in Nikon TE2000E inverted microscope. After 10 minutes, most of the pollen sedimented on the bottom of the well. Observation in bright field allows to clearly distinguish between viable red and non-viable green pollen grains. Generally, at least 100 pollen grains (but usually around 250) from at least seven plants of identical genotype background were counted.

Aniline blue staining of pollen tube in Arabidopsis pistil

In plants, aniline blue dye can be used to identify callose, a polysaccharide associated with plasmodesmata, pollen tubes, and microsporocytes. The aim of this experiment was to visualize pollen tube growth through the pistil and observe phenotypical differences between inspected lines and the wild type control. To observe a true male-based differences in pollen tube growth, pollen from mutant lines was used to pollinate female pistils of one identical background, therefore self-pollinated pistils were not used. Crossed flowers at stage 15-17 were removed from plants and their pistils were excised and incubated in 120 µL of the fixative solution (Acetic acid:ethanol in 1:3 v/v) for 2 hours at room temperature. Fixative solution was then exchanged with stepwise ethanol series: 120 µL of 70%, 50% and 30% ethanol and finally dH_2O, in each step, samples were incubated for 10 minutes at room temperature. Then, water was exchanged for 120 µL of the Alkaline treatment solution (8M NaOH), in which pistils were incubated overnight at room temperature. Next day, pistils were washed in dH_2O and incubated in the decolorized alkaline blue solution (0.1% aniline blue in 108 mM K_3PO_4 pH = 11, incubated at 4°C overnight and filtered through filter paper with one added teaspoon of active carbon) in dark until observation, but at least for 24 hours. Pistils were then mounted on sample glass for observation under UV. Pistils of the *ms1-1* mutant background were used in these experiments.

Transmission defects – genotyping

Defect in transmission of the mutant allele through male or female gametophyte was analysed by reciprocal back-crossess of the heterozygous mutant line with wild type, subsequent genotyping of the progeny and suitable statistical testing.

Transmission defects – siliques dissection

Dissection of Arabidopsis siliques served to investigate seed to analyse any defects in mutant allele transmission. In one Arabidopsis pistil, from 50 to 60 ovules is usually formed during the carpel development (Col-0 ecotype), which corresponds with the number of seeds in one silique. Generally, defects in seed formation are easily observable under binolupe. However, determination of the causality why seed formation is disrupted is not that easy and requires more detailed analyses, as defects in seed formation could be caused by numerous reasons (e.g. defects in embryo sac formation, pollen tube guidance or post-fertilization embryo lethality). In order to investigate whether particular mutant line has some effect on seed formation, siliques from analysed mutant background plants were excised, stacked to double-sided tape on a sample glass, cut alongside the replum using sharp needle and, using forceps, valves were opened to expose seeds undamaged by the procedure. Seeds from at least three siliques (but usually four) from one plant were scored and at least seven plants of identical genotype background were analysed.

Molecular cloning

Bacterial strains and common cultivation media

For molecular cloning, our lab uses *Escherichia coli* strain TOP10a for plasmid amplification and verification by Colony PCR and RE digest. The method of TOP10a transformation used was the transformation by a heat shock, where the plasmid enters the cell through cell membrane pores generated by a sudden change of temperature in calcium-rich environment. *Agrobacterium tumefasciens* strain GV3101 was used for plant transformation assays. The GV3101 strain has a chromosomal background containing cassette with resistance to rifampicin and Ti helper plasmid cassette with resistance to gentamicin. The method of GV3101 transformation used was the transformation by electroporation, where the plasmid enters the cell trough cell membrane pores generated by electric impulse. Several common media were used to successfully cultivate bacteria in vitro. For *E. coli* TOP10 cells, LB solution media (Duchefa) is used in all procedures. For *Agrobacterium tumefaciens* strain GV101, YEB media (6 g/L yeast extract; 5 g/L peptone from casein; 5 g/L sucrose; 0.5 g/L $MgSO_4$. $7H_2O$, NaOH to adjust pH to 7.2) is used in most of the procedures.

Competent bacteria used for molecular cloning and plant transformation

Chemically competent *E. coli* cells were prepared by growing the inoculum in 250 mL of sterile SOB media (5 g/L yeast extract; 20 g/L tryptone; 0.6 g/L NaCl; 0.2 g/L KCl; 2.03 g/L $MgCl.6H_2O$; 2.47 g/L $MgSO_4.7H_2O$) at 20 °C and 160 rpm to reach OD600 = 0.3. After reaching desired OD600, the culture was centrifuged at 3000 rpm and 4 °C for 10 minutes in 50 mL Falcon flasks, resuspended in 80 mL total of the sterile precooled CCMB80 buffer (10% v/v glycerol; 11.76 g/L $CaCl_2$; 3.96 g/L $MnCl_2.4H_2O$; 2.03 g/L $MgCl.6H_2O$; 0.1 g/L KOAc) and incubated for 20 minutes on ice. The suspension was then centrifuged again at 3000 rpm, resuspended in 10 mL of precooled sterile CCMB80 and incubated on ice for 20 minutes. OD600 was measured and adjusted to 5. Then the suspension was aliquoted in precooled 1.5 ml Eppendorf tubes (50 µL aliquots), and slow-freezed by placing the tubes in -80°C. *Agrobacterium tumefasciens* electro-competent cells were prepared by growing the inoculum first in 5 mL YEB media (6 g/L yeast extract; 5 g/L peptone from casein; 5 g/L sucrose; 0.5 g/L $MgSO_4$. $7H_2O$; NaOH to adjust pH to 7.2) with 50 mg/L gentamycin, 100 mg/L rifampicin overnight to full saturation of the culture. Next day, 1 L Erlenmeyer flask with 200 mL YEB with 50 mg/L gentamycin, 100 mg/L rifampicin was inoculated with the culture and cultivated overnight. Next day, the culture fitting within OD600 between 0.5 and 1 was centrifuged at 3000 rpm and 4°C for 20 minutes in 50 mL Falcon tubes and resuspended in 120 mL total of sterile H_2O. The suspension was centrifuged at 3000 rpm and 4°C for 20 minutes and resuspended in 10 ml of 10% glycerol, centrifuged again at 3000 rpm and 4°C for 20 minutes and resuspended in 2 ml of 10% glycerol. From this suspension, aliquots of 30 µl were prepared in 1.5 ml Eppendorf tubes and slow-freezed by placing the tubes in -80°C.

Polymerase chain reaction (PCR)

DNA amplification by Polymerase Chain Reaction (PCR) is common technique of molecular biology. DNA template was mixed with specific primers, free nucleotides and DNA dependent DNA polymerase. For PCR amplification of domesticated fragment that were meant to be cloned into GBparts and further cloning, Phusion polymerase (Life Technologies) with proofreading activity was used to minimize errors in the sequence. For other PCR amplifications, Merciáza Taq polymerase enzyme (Merci) was used. In the final PCR mixture, concentration of dNTPs was 0.2 mM, final concentration of every primer added was 0.5 mM together with 1.5 U of Taq polymerase enzyme, all in 1x Merciáza buffer (Merci). When the Phusion enzyme was used, all concentrations were identical, only appropriate HF buffer (Life

Technologies) was used for optimal activity of the Phusion enzyme. Oligonucleotides used in PCR reactions are shown in Table 4.

Agarose gel electrophoresis

To visualize the DNA product of PCR reaction or enzymatic DNA cleavage, samples were separated by electrophoresis through agarose gel. To the sample reaction mixture, appropriate volume of 10x OrangeG loading dye (45% sucrose with a tip of small spatula of the OrangeG powder) was added to the mixture and the reaction mix was separated on 1 % agarose gel casted in Thermo Scientific™ Owl™ EasyCast™ (Thermo) with 1-2 drops of Ethidium Bromide (PanReac AppliChem), set up at 100 V for 45-65 minutes until the Orange G line (migrating in the at aprox. 100-500bp) almost reached the bottom of the gel. DNA was visualized under UV light for the Ethidium Bromide excitation. If fragment of <500 bp was expected, the gel was additionally visualized after approx. 30 minutes of separation.

GoldenBraid cloning system

GoldenBraid is a cloning system for plant synthetic biology and Agrobacterium-mediated transformation that enables to build transcription units from multiple blocks in one reaction (Sarrion-Perdigones et al., 2011). It is based on a hierarchy of backbone plasmids and restriction/ligation reactions with the use of Type IIS restriction enzymes and T4 DNA ligase. Type IIS restriction enzymes recognize specific non-palindromic dsDNA sites and cut non-specifically outside of the recognition sequence, while T4 DNA ligase seals nicks in dsDNA by catalysing a formation of a phosphodiester bond between 5' phosphate group and 3' hydroxyl group. GoldenBraid uses Type IIS restriction enzymes that generate a pre-defined four-nucleotide overhang on one building block (e.g. promoter) and complementary overhang on the second building block (e.g. CDS). Complementary overhangs renature in the reaction only to be joined together with T4 DNA ligase. By designing different four-nucleotide overhangs on different building blocks (GBparts) and the destination vectors, the GoldenBraid system enables multi-block assembly in one reaction. The hierarchy of plasmids is divided in three levels: pUPD2 level, Alpha level and Omega level. Each level plasmids possess different selection marker (based on antibiotic resistance) and different multiple cloning site for type IIS restriction enzymes. Alpha and Omega vectors are binary vectors suited for replication in *E.coli* and *Agrobacterium* and contain LB and RB sequences for Agrobacterium-mediated plant transformation. On each level, several versions of the same plasmid (e.g. Alpha1, Alpha2) differ

only in the flanking sequence that will form the four-nucleotide overhangs when digested by appropriate Type IIS restriction enzyme. Originally, each plasmid has a lacZ cistrone cassette in between two multiple cloning sites, that is replaced during the restriction/ligation reaction with desired fragment. The process of building a transformation cassette for plant transgenesis and expression in plants is divided into three phases: domestication of template DNA into pUPD2 GBpart, assembly of the transcription unit in Alpha vector and the transcription unit assembly with plant selection markers in Omega vector. Each assembly reaction produces a vector that is transformed in *E.coli* strain TOP10a, selected by blue/white selection and PCR verification, purified by GeneJET plasmid MiniPrep Kit (Thermo Scientific) and verified by restriction and/or sequencing. The domestication of template DNA is a process of removing undesired recognition sites of type IIS restriction enzymes used in GoldenBraid system from the template sequence and providing appropriate cloning sites to both ends of the fragment. The domestication protocol is designed online by GB domestication tool (available here: https://gbcloning.upv.es/do/domestication/) that generates sequences of domestication oligonucleotides with appropriate linkers. These oligonucleotides are used as primers in one or several PCR reaction(s), dependent on the number of recognition sites for type IIS restriction enzymes used). Domestication PCRs products are then used in the first restriction/ligation reaction, using BsmBI restriction enzyme (Thermo Scientific) to assembly the template DNA into pUPD2 plasmid backbone. The pUPD2 vector with the cloned fragment represents one building block (GB part) of the GoldenBraid system. Each GBpart has a flanking sequence at the beginning and at the end, which will eventually turn into the overhang after digestion by Type IIS restriction enzyme. Then in the next step, GBparts in pUPD2 vectors are mixed in a reaction with the backbone Alpha level plasmid vector and BsaI restriction enzyme (Thermo Scientific). When digested, the overhangs of all GBparts complement in a way that will form correctly ordered transcription unit inside of the Alpha level plasmid. Similarly to the transcription unit assembly, the purified Alpha vector with desired transcription unit cassette and plant selection marker (Alpha2 vector with KanFAST selection or BASTA resistance cassettes) are cloned into another assembled into Omega vector, using again BsmBI restriction enzyme. Omega vector is then transformed into *Agrobacterium tumefasciens* strain GV3101 that is used for plant transformation. Simplified scheme of the GoldenBraid system is shown in Figure 7. Sequences of novel GBparts domesticated in this work are shown in Table 1. GBparts and promoters used in this work are shown in Table 2. Expression cassettes assembled in this work are shown in Table 3.

GoldenBraid restriction/ligation reaction

In one GoldenBraid restriction/ligation reaction, 50 ng of the purified assembly vector and 50 ng of each lower level GoldenBraid plasmids were added together with 5 U of the restriction enzyme and 3 U of the T4 DNA ligase (Thermo Scientific) in 1 x T4 ligase buffer (Thermo Scientific) in total amount of 10 µl. The mixture is incubated in the following programme in the basic bench PCR cycler: initial 10 minutes at 37 °C, 45 cycles of 5 minutes ligation at 16 °C and 2 minutes restriction at 37 °C, followed by 30 minutes of final restriction at 37 °C, 5 minutes of final ligation at 16 °C and 20 minutes of enzyme inactivation at 80 °C. The reaction was then cooled down to 4 °C and stored at -20 °C.

Blue/white selection and construct verification

After GoldenBraid restriction/ligation reaction, selection and verification of the plasmid is needed. Because the T4 ligase buffer contains ATP that could be toxic for the *E.coli* cells in the following transformation, the mixture was diluted with 5 µL ddH_2O. From this diluted mixture, a volume of 5 µL was used for *E.coli* heatshock transformation. On plates containing the proper antibiotics, only cells that were successfully transformed with the destination vector backbone were resistant to the antibiotics and form a colony. However, there is a chance that some destination vectors were not cut at all and were undesirably transformed during the heat shock as well. Second selection called the blue/white selection was therefore needed to distinguish between the original destination vector with lacZ cistrone and the correctly assembled construct. All backbone plasmids originally contain lacZ cistrone gene that is removed in the reaction and replaced by the construct of interest. The presence of the lacZ cistrone could be visualized by addition of 0.5 mM IPTG, the lacZ expression-inducing chemical, and 40 mg/L 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal), the lacZ substrate that produces blue dye after enzymatic cleavage by lacZ enzyme. Blue coloured colonies formed on agar plate were therefore false positives in the antibiotic resistance selection. Plates incubated overnight at appropriate temperature were placed in the fridge for 1 hour to better development of the blue signal. After 1 hour, only white and not blue colonies were picked for quick PCR verification of presence of the correct insert in the individual colony. For that purpose, a 200 µL yellow pipette tip was used to touch a single white colony. The tip was then placed in 500 µL small test tube with 30 µL of ddH_2O. The tip was then connected to the pipette and cells were resuspended in the water by sucking up and down with the pipette. 5 µL of this cell suspension

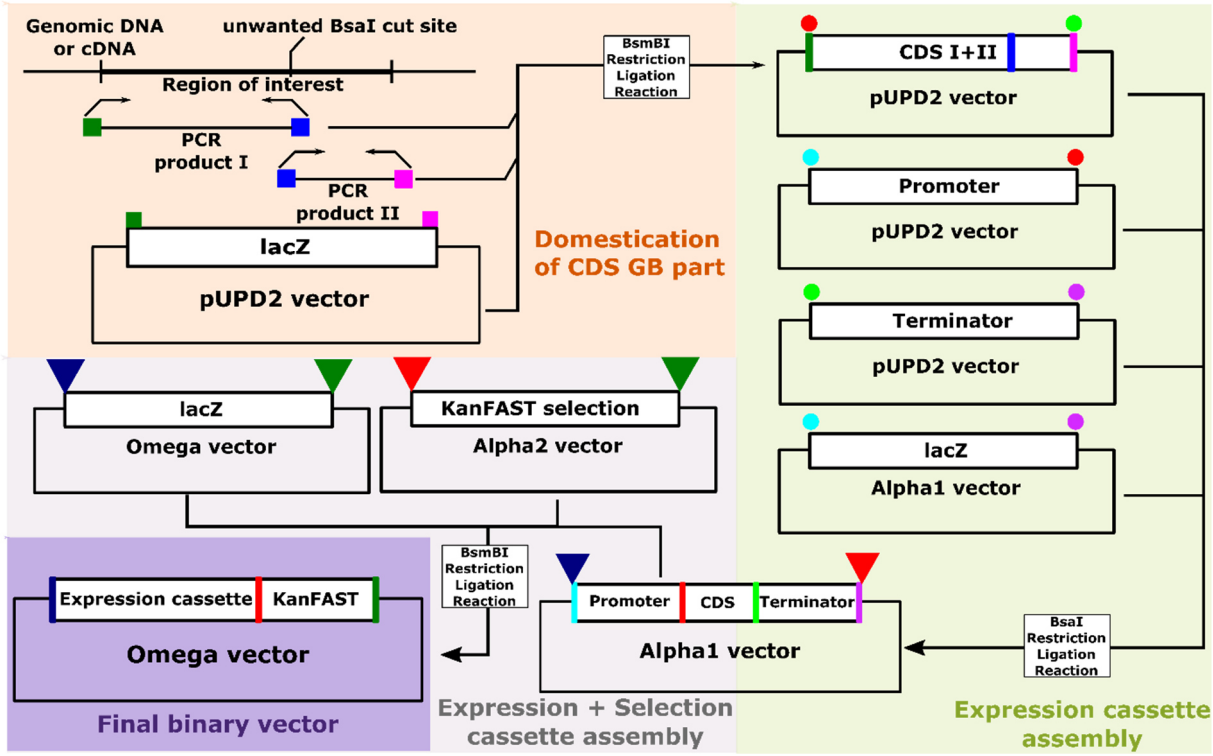

Figure 7: GoldenBraid cloning system from template sequence to expression cassette. Schematic depiction of the domestication process (upper left corner) from template sequence to pUPD2 GBpart, the expression cassette assembly (right section) from pUPD2 GBparts into expression cassette and the assembly of final binary vector from expression cassette and selection marker expression cassette. Rectangles, cirles and triangles of the same color represent complementary overhangs produced by the restriction enzyme. Color borders within the insertion indicate joining sites of the complementary overhangs from previous assembly step.

was then used for modified PCR reaction of total volume 20 µL with primers that would produce a construct-specific band. The PCR programme was slightly modified with prolonged initial 96 °C denaturing temperature to 10 minutes in order to degrade the cell wall and release the plasmid into the reaction mixture.

Plasmid amplification, isolation and storage

With a yellow 200 µL tip or a toothpick, colony verified by PCR was gently touched on the agar plate and inoculated into 5 mL liquid LB broth low salt media with appropriate antibiotic selection. The culture was then incubated overnight at 37 °C and 160 rpm to full saturation. Next day, the well-grown culture was processed using the GeneJET Plasmid Miniprep Kit (Thermo Scientific) to plasmids isolated and purified. 4 mL of saturated culture were centrifuged at 13000 rpm for 2 minutes, and resuspended by vortexing in 250 µL Resuspension solution. Then, 250 µL of Lysis solution was added, the suspension was mixed well by inverting the tube up and down 5 times. Immediately, 350 µL of neutralization solution was added and the sample was mixed the same way as before and placed on ice. After that, the sample was

centrifuged at 13000 rpm for 5 minutes and the supernatant was transferred in GeneJET Spin Column. After 1 minute centrifugation at 13000 rpm, the column was washed twice by adding 500 µL of Wash solution to the column and centrifuging for 1 minute. The plasmid DNA was eluted by addition of 50-100 µL of the Elution buffer, incubated for 6 minutes at room temperature and centrifuged for 2 minutes at 13000 rpm. Concentration of the dsDNA in the eluted solution was then measured on NanoDrop One Microvolume UV-Vis Spectrophotometer. Samples were stored at -20°C and used in further experiments.

Plasmid verification by enzymatic cleavage and/or sequencing

After isolation of the plasmid, subsequent steps of verification proceeded. First, the correctness of inserted sequence was verified by specific enzymatic cleavage of the plasmid DNA. 500-1000 µg of purified plasmid DNA was mixed together with the 1 U of common restriction enzymes (Thermo Scientific) that cut in the insert sequence of the plasmid, plus appropriate buffer in 15 µL of total volume. The mixture was incubated for 4 hours or overnight at temperature optimal for activity of the enzyme. Plasmid fragments produced by enzymatic cleavage were separated by electrophoresis on agarose gel and visualized band pattern was compared to the *in silico* digest of the plasmid performed by Benchling software. Whenever plasmid was verified by the digest, 400-500 µg of the purified plasmid DNA was mixed with 2.5 µL of 10 mM specific primer and sent to sequencing company Eurofins Genomics as LightRun tube sample (Sanger sequencing). Obtained sequencing results were aligned *in silico* to the plasmid sequence using MAFFT algorithm incorporated in Benchling. Whenever new GBpart is domesticated and cloned into pUPD2, the whole sequence needs to be verified, whereas after transcription cassette is assembled form GBparts that are already verified by sequencing, the insert was sequenced only for the correct promoter-CDS connection.

Escherichia coli transformation by heat-shock

E.coli strain TOP10a chemically competent cells were placed from -80 °C on ice to thaw. 100 ng of purified plasmid or 5 µL of 0.66 times dilution of the GoldenBraid reaction were mixed with the competent cells and left 30 minutes on ice. Sample tube was then placed in water bath at 42 °C for exactly 1 minute and resuspended in 975 µL of liquid LB broth high salt medium. The suspension was then mixed and placed in a static incubator set at 37 °C for 1.5 hour. 100-150 µL of the suspension was then plated on solid LB broth high salt medium with 100 mg/L appropriate plasmid selection. Plates were then incubated at 37 °C overnight.

Agrobacterium tumefaciens transformation by electroporation

Agrobacterium tumefaciens strain GV3101 electrocompetent cells was placed from -80 °C on ice and mixed with 200 ng of purified plasmid. After 3 minutes of incubation on ice, the suspension was pipetted in an ice cold Gene Pulser Electroporation Cuvettes (Bio-Rad; 0.1 cm electrode gap) and electroporated in the Eporator® (Eppendorf) (settings: 2 200 kV, 200 Ω, capacity 25 µF). 970 µL of ice cold liquid YEB medium was added immediately to the cuvette to resuspend the cells. Cells suspension was then transferred into ice cold 1.5 ml Eppendorf test tube, mixed through and incubated at 28 °C for 2 hours in a static incubator. 80-150 µL of the cell suspension were spread on solid YEB media with 50 µg/mL gentamycin, 100 µg/mL rifampicin and 100 µg/mL plasmid selection (kanamycin or spectinomycin). Inverted plates were incubated at 28 °C for at least 40 hours before colonies were formed.

Table 1: Table showing novel sequences domesticated and used in this work.

Name	5' to 3' sequence
p3A N-	gattttatgtaaatcaattgtttttttgaaatgaatttttgtagtggcattttatgtaatatttaatatgttgttggctatttacttaaatgtatgtttagt taataatatagaaatatcaaagttaattgctaaaaatattattaaacaaaatatagagagtatattgtaaatgataattcgaaaatgaaatccata aattttgttgatctcaattgaatcacatcccagtattagacaaattgggttttataaacactaagaagataaaacggggtcatttcgaaaccgtaa
p3A C-	gaaaataggatgatttctgcaaatatgggtaatatatatgtagaacgataatattaataccagaaaccctagagttgcttccgcttcgcttctcgtc tttcttatcagtcttctcttttcacggctacaaggttagtcctaagtttcggctctgttcttgcatcagtgattggatctctactgccttttaatcaaatct cattgaattgatctgggttagggtttgggtcaggttatcgaatctttctgtgttattttgattctatgaaatgaatttagggtttgggcaagttttttct ggtctatgtagcttccatccttttctgagatatctgattgggtttttcggattctggtgatctctgtaacaaaatcgcttctctgtgtttttcaggctaat tgagagtgtcccttgcgttagagtttgttactagctttttttcgtctgaaac
3A RNAi sense	tacagctaagtattgtcagtggggagaagggaaaagataggtctgaccgtgagttggtcaccccatggtttaagtttctgtgggagacatacagg acagtccttgagatattgcgaaacaactcaaagttggaagcactttatgcgatgacagcacataaagccttccagttctgtaagcagtacaagcg aacaacagagttccgtaggctttgtgaaattataagaaatcatttggcaaacctgaacaagtacagagacaaaagggaccgacctgacttatca gcccctgagagcttgcagctgtatctggacacaaggtttgatcagttgaaagttgctactgagcttggactttggcaggaagcttttcgctctgttg aagatatctatggattaatgtgcatggtcaaaaaaacgcccaagtcatctttattgatggtttattattccaaattgactgagatcttctggatttctt ctagccatctttatcatgcttatgcatggtttaagctgttta
3A RNAi antisense	taaacagcttaaaccatgcataagcatgataaagatggctagaagaaatccagaagatctcagtcaatttggaataataaaccatcaataaaga tgacttgggcgttttttgaccatgcacattaatccatagatatcttcaacagagcgaaaagcttcctgccaaagtccaagctcagtagcaactttc aactgatcaaaccttgtgtccagatacagctgcaagctctcaggggctgataagtcaggtcggtccctttggtctttgtacttgttcaggtttgccaa atgatttcttataatttcacaaagcctacggaactctgttgttcgcttgtactgcttacagaactggaaggctttatgtgctgtcatcgcataaagtg cttccaactttgagttgtttcgcaatatctcaaggactgtcctgtatgtctcccacagaaacttaaaccatggggtgaccaactcacggtcagacct atcttttcccttctccccactgacaatacttagctgta

Table 2: Table showing used Gbparts and promoters in this work and their basic description.

Name	Description	Overhangs		Origin
		start	end	
CaMV 35S Double	35S double promoter	GGAG	TACT	GB kit
5U-TMV	TMV 5´UTR adaptor	TACT	CCAT	GB kit
pLat52	Lat52 promoter for N-terminal fusions	GGAG	CCAT	Created by Christos Michailidis
p3A N-	AteIF3A promoter for N-terminal fusion	GGAG	CCAT	This work
p3A C-	AteIF3A promoter for C-terminal fusion	GGAG	AATG	This work
pJas	Jason promoter for N-terminal fusions	GGAG	CCAT	Created by Christos Michailidis
pAsy	Lat52 promoter for N-terminal fusions	GGAG	CCAT	Created by Christos Michailidis
eIF3Astop	AteIF3A CDS for N-terminal fusion	AATG	GCTT	Created by Lenka Steinbachová
eIF3Anostop	AtEIF3A CDS for C-terminal fusion	CCAT	TTCG	Created by Lenka Steinbachová
GUSstop	β-Glukoronidase CDS for N-terminal fusion	AATG	GCTT	GB kit
NT-GFP	GFP N-terminal fusion	CCAT	AATG	GB kit
NT-mCherry	mCherry N-terminal fusion	CCAT	AATG	GB kit
NT-6xHA	6x HA epitope tag for N-terminal fusion	CCAT	AATG	GB kit
NT-mOrange	mOrange N-terminal fusion	CCAT	AATG	Created by Christos Michailidis
CT-GFP	GFP C-terminal fusion	TTCG	GCTT	GB kit
CT-3xFLAG	3x FLAG epitope tag for C-terminal fusion	TTCG	GCTT	GB kit
CT-Hellfire	CT-6xHA-3xFLAG epitope tag for C-terminal fusion	TTCG	GCTT	Provided by Daniel Sanchez
3A RNAi sense	AteIF3A region for RNAi - sense	CCAT	AGCC	This work
3A RNAi antisense	AteIF3A region for RNAi - reverse complement	GCAG	GCTT	This work
Kannibal intron	Kannibal intron for RNAi	AGCC	GCAG	Provided by Tomáš Moravec
NOS-t	3´UTR+Nos terminator	GCTT	CGCT	GB kit

Promoter	Gene origin	Expected	Reference
pAsy	ASY2 (AT4G32200)	prophase I	(De Storme et al., 2016)
pJas	JASON (AT1G06660)	male metaphase II	(De Storme et al., 2016)
pLat52	Tomato LAT52	Pollen specific	(Twell et al., 1989)
35Sdouble	duplication of CaMV 35S	Constitutive sporophytic	(Kay et al., 1987)

Table 3: Table showing all assembled expression cassettes in this work and their organization

GB cloning	Omega2 assembly						
Construct label	Alpha1 insert						Alpha2 insert
	GB parts						
	Promoter		N-tag	CDS		Terminator	Selection cassette
pLat52::AtEIF3A-GFP	pLat52		eIF3Anostop	CT-GFP		NOS-t	KanFAST
pLat52::GFP-AtEIF3A	pLat52		NT-GFP	eIF3Astop		NOS-t	KanFAST
p3A::AtEIF3A-GFP	p3A N-		eIF3Anostop	CT-GFP		NOS-t	KanFAST
p3A::GFP-AtEIF3A	p3A N-		NT-GFP	eIF3Astop		NOS-t	KanFAST
p3A::GFP-GUS	p3A N-		NT-GFP	GUSstop		NOS-t	KanFAST
p3A::AteIF3A	p3A C-		eIF3Astop			NOS-t	
pLat52::AtEIF3A-HF1	pLat52		eIF3Anostop	CT-Hellfire		NOS-t	KanFAST
pLat52::AtEIF3A-FLAG	pLat52		eIF3Anostop	CT-3xFLAG		NOS-t	KanFAST
pLat52::HA-AtEIF3A	pLat52		NT-6xHA	eIF3Astop		NOS-t	KanFAST
35S::AtEIF3A-GFP	CaMV 35S Double	5U-TMV	eIF3Anostop	CT-GFP		NOS-t	KanFAST
35S::mOrange-AtEIF3A	CaMV 35S Double	5U-TMV	NT-mOrange	eIF3Astop		NOS-t	KanFAST
Construct label	RNAi constructs						
	Promoter		Sense	Intron	Antisense	Terminator	Selection cassette
pLat52::RNAi 3A	pLat52		3A RNAi sense	Kannibal intron	3A RNAi antisense	NOS-t	BASTA
p3A::RNAi 3A	p3A N-					NOS-t	BASTA
pAsy::RNAi 3A	pAsy					NOS-t	BASTA
pJas::RNAi 3A	pJas					NOS-t	BASTA
35S::RNAi 3A	CaMV 35S Double	5U-TMV				NOS-t	BASTA

Table 4: Table showing oligonucleotides used in this work

Used for	Primer name	5' to 3' sequence
T-DNA genotyping (SAIL 124, SAIL 545)	SAIL_124_G07.V1 LP	GAGGACCTGTGACGATAGCAG
T-DNA genotyping (SAIL 124, SAIL 545)	SAILSEQ_545_B08.4 RP	TTGAGAAGCAAAGGCAGAGAG
T-DNA genotyping (SAIL 227)	SAILSEQ_227_B02.0 LP	GATTCAAGGGAAGGTCCAAAG
T-DNA genotyping (SAIL 227)	SAILSEQ_227_B02.0 RP	CCTTCTTCGGTTGGATCTTTC
T-DNA genotyping (SALK 762)	SALK_123762.31.45.X RP	AAGCCCAAGCTTTACTGGAAG
T-DNA genotyping (SALK 762)	SALKSEQ_123762.0 LP	CTCTAGGACCAGAACCCCAAC
T-DNA genotyping (SAIL 124, 227, 545)	SAIL_LB	TAGCATCTGAATTTCATAACCAATCTCGATACAC
T-DNA genotyping (SALK 762)	SALK_LB	TGGTTCACGTAGTGGGCCATCG
p3A domestication	p3A int1 fw	GCGCCGTCTCGCTCGGGAGGATTTTATGTAAATCAATTGTTTTTTTG AAAT
p3A domestication	p3A int1 rv	GCGCCGTCTCGAAAGACGAGAAGCGAAGCGG
p3A domestication	p3A int2 fw	GCGCCGTCTCGCTTTCTTATCAGTCTTCTCTTTTC
p3A C- domestication	p3A_AATG_int2 rv	GCGCCGTCTCGCTCACATTGTTTCAGACGAAAAAAGCTAGTAA
p3A N- domestication	p3A_CCAT_int2 rv	GCGCCGTCTCGCTCAATGGGTTTCAGACGAAAAAAGCTAGTAA
3A RNAi sense domestication	3A_RNAi3_sense_F1	GCGCCGTCTCGCTCGCCATTACAGCTAAGTATTGTCAGTGG
3A RNAi sense domestication	3A_RNAi3_sense_F2	GCGCCGTCTCGACAAAAGGGACCGACCTGAC
3A RNAi sense domestication	3A_RNAi3_sense_R1	GCGCCGTCTCGTTGTCTCTGTACTTGTTCAGG
3A RNAi sense domestication	3A_RNAi3_sense_R2	GCGCCGTCTCGCTCAGGCTTAAACAGCTTAAACCATGCATAAG
3A RNAi antisense domestication	3A_RNAi3_rev_compl_F1	GCGCCGTCTCGCTCGGCAGTAAACAGCTTAAACCATGCATAAG
3A RNAi antisense domestication	3A_RNAi3_rev_compl_F2	GCGCCGTCTCGCTTTGTACTTGTTCAGGTTTGC
3A RNAi antisense domestication	3A_RNAi3_rev_compl_R1	GCGCCGTCTCGAAAGACCAAAGGGACCGACC
3A RNAi antisense domestication	3A_RNAi3_rev_compl_R2	GCGCCGTCTCGCTCAAAGCTACAGCTAAGTATTGTCAGTGG
eIF3A construct verification	IF3A_sekvF2	AAGCCAAAAGGATTTGCAG
eIF3A construct verification	eIF3A_sekv(Peter)	GCAGATACTGTTGAGAAAGAGC
eIF3A construct verification	eIF3A_RT_F	AACCAGAGAACGCGTTGAAG
eIF3A construct verification	eIF3A_RT_R	CTTGCTGGCAAACAATACGA
3A RT-qPCR, primer set 2	eIF3A_RT_F2	AAGCACTTTATGCGATGACAG
3A RT-qPCR, primer set 2	eIF3A_RT_R2	CTTCCTGCCAAAGTCCAAGC
3A RT-qPCR, primer set 4	eIF3A_RT_F4	GCAAGCAGGAAGAAGCTGAG
3A RT-qPCR, primer set 4	eIF3A_RT_R4	TGGTGCGGTAGTTCCAACAG
ACT8 RT-qPCR	ACT8_f	TTCATCGGCCGTTGCATTTC
ACT8 RT-qPCR	ACT8_rv	AATGTCATCAGCATCGGCCA
pUPD2 vector verification, sequencing	pUPD2_F	CCGATCAACTCGAGTGC
pUPD2 vector verification, sequencing	pUPD2_R	TGTTCTTTCCTGCGTTATCC
alpha/omega vector verification, sequencing	LB_F	TGGCAGGATATATTGTGGTG
alpha/omega vector verification, sequencing	RB_R	GTTTACCCGCCAATATATCC
Lat52 containing vector verification	Lat52_F_Said	TTGAGGAATGATCGATTCTGG
GFP containing vector verification	GFP_F	ATGGTGAGCAAGGGCGAGGA
GFP containing vector verification	YFP/GFP R	GAAGTTCACCTTGATGCCG
KanFAST containing vector verification	Seq_FAST_F3	AGAAACCCGCTAAGAACC
Hellfire containing vector verification	Fwd Hellfire	TCACCATCACCATCACGATTA
Hellfire containing vector verification	RV Hellfire	TTGTCATCGTCATCCTTGTAGT
FLAG containing vector verification	Flag_F	GATTATAAGGACCATGACGG
FLAG containing vector verification	Flag_R	CTTATCGTCATCGTCCTTAT
HA containing vector verification	seq_HA_Fr	CCAGATTACGCCACTAGAG
HA containing vector verification	seq_HA_Rv	CTCTAGTGGCGTAATCTGG
mCherry containing vector verification	mCherry_R1	GATGAACTCGCCGTCCTG

Plant transformation

Stable transformation of Arabidopsis thaliana by the floral dipping technique

Picked and verified *Agrobacterium tumefaciens* colony carrying the plasmid with expression cassette was inoculated in 5 mL YEB media with 50 mg/L gentamycin, 100 mg/L rifampicin and 100 mg/L plasmid selection and grown overnight till saturation. From this starting culture, 200 µL of the cell suspension was inoculated in 1 L Erlenmeyer flask with 200 mL of the medium with identical composition as described for the starting culture and grown for 20-24 hours till saturation. The saturated culture was then poured into 250 mL centrifuge tubes (BeckMan) and centrifuged at 3900 rpm for 30 minutes. After discarding the supernatant, the cell pellet was resuspended in fresh Infiltration media (2.165 g/L MS salts; 5% sucrose; 1 ml/L Gamborg vitamin solution; 0.5 g/L MES; 10 µg/L Benzylaminopurine). Before dipping, 300 µL/L of Silwet L-77 was added to the suspension. Freshly bolting Arabidopsis thaliana Col-0 plants rich in secondary inflorescences (usually 9-12 days after cutting the primary inflorescence to encourage branching) placed in transformation pots (5 plants/pot) were removed of siliques and opened flowers and dipped in the cell suspension for exactly 1 minute. After dipping, plants were covered in a black plastic bag overnight and tied up after 1 week after transformation.

Stable transformation of N. tabacum by Agrobacterium-mediated leaf discs infiltration

The floral dipping technique has not been proven to successfully transform *Nicotiana tabacum*. Therefore, different transformation technique was used in order to produce stable transgenic lines: the *Agrobacterium*-mediated transformation of tobacco leaf discs. This method of transformation uses the *Agrobacterium* to incorporate the insertion cassette in the genomic DNA of mesophyll cells. Subsequently, leaf discs were cultivated *in vitro* on series of media containing antibiotic selection and growth regulators in order to regenerate Tobacco plants from the transformed cell *de novo* using indirect organogenesis (see the corresponding chapter above for the cultivation technique). Sterile conditions were absolutely necessary during this procedure and every manipulation with sterile plants was performed in laminar box. Sterile tobacco seeds were germinated in vitro on full MS media and seedlings with developed cotyledons were transferred to magenta boxes with 5 cm layer of MS media and grown further. For transformation of leaf discs, the appropriate age of the plants was approximately 1 month, Picked and verified *Agrobacterium tumefaciens* colony carrying the plasmid with expression

cassette was inoculated and grown overnight in 5 mL YEB with 50 µg/mL gentamycin, 50 µg/mL rifampicin and 100 µg/mL plasmid selection (kanamycin) in 10 mL vial to full saturation. From this culture a volume of 1/1000 of the final volume was inoculated in 6 mL of YEB with 50 µg/mL gentamycin and 100 µg/mL plasmid selection in 100 mL Erlenmeyer flask and grown overnight to full saturation. 400 µL of the saturated culture was centrifuged at 13000rpm for 2 minutes. Then, the pellet was washed two times by resuspending in 2 mL 2MS media and centrifuging at 13000 rpm for 2 minutes. Finally, the pellet was resuspended in 700 µL of Two strength MS media and placed on ice. This was the Agro stock. Young leaves were excised from the *in vitro* grown *N. tabacum* plant using sterile scalpel and forceps, and placed on an empty plastic 9cm-diameter Petri dish. Excised leaf was cut to 0.5x0.5 cm pieces using the scalpel. Using forceps, leaf discs were placed in another Petri dish with 10 ml of the 2MS media in a way that their adaxial side was contacting the media surface. After whole surface of the media was covered with leaf discs, 25 µL of the Agro stock was s inoculated in the media and gently mixed. At least 5 Petri dishes per construct were prepared and kept in dark at 26°C for 48 hours, during which the Agrobacterium transformed the mesophyll cells. After two days, plates were opened and the cell suspension was removed by pipetting. Using forceps, leaf discs were carefully removed from the Petri dish and placed on dry sterile filter paper to remove remnants of the cell suspension. From filter paper, leaf discs were transferred on fresh CIM media, where the procedure of the *de novo* regeneration of transformed *N. tabacum* plants by indirect organogenesis followed (described in its respective chapter above).

Selection of positive transformants

Selection of successful transgenic individuals is a crucial step after plant transformation. The most common technique is the presence of antibiotic or herbicide resistance cassette in the construct introduced in plant genomic DNA, driven by strong sporophytic promoter. Seeds harvested from transformed T_0 plants are cultivated *in vitro* on media containing proper antibiotic or herbicide. However, the transformation is usually low-efficient and hundreds or even thousands of seeds are needed to be sown in order to get a reasonable number of individual transgenic lines. Such a big number of seeds on a selection plate causes insufficient sterilization and higher chances of contamination. Therefore, an advanced double selection approach was used, the KanFAST selection. KanFAST selection is a combination of two expression cassettes where: 1) kanamycin resistance gene is driven by the strong sporophytic Nos promoter with 5′

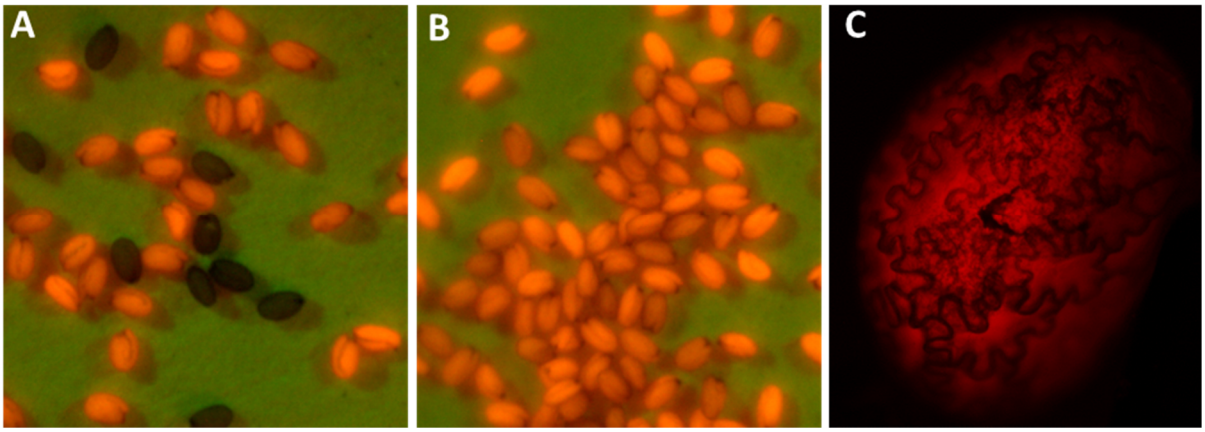

Figure 8: FAST selection. Stereomicroscope with filter adjustors allows observation of the Oleosin-RFP fluorescence in the seeds. For *N.tabacum* seeds expressing FAST cassette, the stereomicroscope setting was insufficient to observe the signal. However, under common inverted fluorescent microscope, the RFP signal could be observed. **(A)** and **(B)** are representative pictures of *Arabidopsis thaliana* seeds harvested from T2 generation plants of Lat52::GFP-AtEIF3A transgenic lines. **(A)** progeny of a heterozygous plant, approx. one third of seeds has no RFP signal. **(B)** progeny of homozygous plants, all seeds are RFP-positive. **(C)** Transgenic *N. tabacum* seed expressing FAST cassette, progeny of T0 plant transformed with Lat52::AtEIF3A-GFP construct.

leader of the TMV and 2) Oleosin-RFP fusion protein is driven by its native Oleosin promoter. Oleosin is a structural protein that accumulates on the surface of oil bodies, which are single layer organelles abundant in oily seeds. The expression of Oleosin-RFP leads to strong RFP signal that enables distinguishing transgenic from non-transgenic seeds under fluorescent microscope or binolupe adjusted with optical filters. In this work, a binolope with optical filters on the light source and the objective was used. The filters on light source allow transmission of RFP-excitation wavelengths, while the objective filter allows transmission of RFP-emission wavelengths only. Therefore, successfully transformed seeds expressing the Oleosin-RFP in seeds are clearly distinguishable from non-transformed seeds. In practice, seeds harvested from T_0 plants were spread on a filter paper that fitted in the lid of plastic 9-cm diameter Petri dish and observed under the binolupe with filters. Around 30 RFP-positive seeds were picked up with forceps and collected in 1.5 mL Eppendorf tube. Selected seeds were sterilized and cultivated on media containing 50 µg/mL Kanamycin or 50 µg/mL BASTA herbicide. After this double selection, at least 10 grown seedlings were cultivated further, representing 10 individual transgenic lines of one particular construct. Seeds harvested from T_1 plants (heterozygous in the insertion site) were again checked for the RFP signal in order to select independent lines with 3:1 segregation (RFP-positive:RFP-negative) to avoid lines with possible double or multiple insertion events. The example of Oleosin-RFP observation in transgenic seeds is shown in Figure 8.

Nucleic acid extractions from plant tissue

DNA extractions from plant tissue

Young rosette leaves from single plant specimen were collected to 1.5 mL Safe-lock Eppendorf tube with 4-8 sterilized sea sand stone pieces. The sample was immediately frozen in liquid nitrogen and grinded for 30-40 seconds in MagNA lyser (Roche) set at 4000 rpm. Immediately after grinding, 250 µL of CTAB extraction buffer (1.4 M NaCl; 20 mM EDTA; 100 mM Tris pH 8; 3% w/v Cetyl trimethylammonium bromide) were added, samples were mixed by shaking or vortexing. After 20 minutes of incubation at room temperature, 250 µL of Chloroform:Isoamyl alcohol (24:1 v/v) were added to the sample. The emulsion was mixed by shaking or vortexing and subsequently centrifuged for 10 minutes at 18000 rpm. After centrifugation, aqueous layer containing solute DNA was transferred to a fresh 1.5ml Eppendorf tube containing isopropanol (0.7 transfer volume), gently mixed and incubated at room temperature for 5 min. Then, sample was gently mixed again and centrifuged at 18000 rpm for 7 minutes. After centrifugation, supernatant was discarded by decanting and the pellet was washed with 500 µL of 70% ethanol, gently mixed and centrifuged at 18000 rpm for 5 minutes. Then, ethanol was removed by vacuum driven pipette and the tube was left open to dry out the pellet. The pellet was then resuspended with 30-50 µL of ddH_2O, the sample was well vortexed and frozen in -20 °C before the extracted DNA was used for ongoing experiments. Samples were then stored in -20°C.

RNA extraction and DNAse treatment

For extraction of RNA from plant tissue, we used RNeasy plant Mini Kit (Qiagen). Maximum amount of 100 mg of plant tissue was collected and frozed in liquid nitrogen. Using mortar and pestle cooled with liquid nitrogen, the sample was grinded into powder. 450 µL of RLT buffer were added and homogenized with the powder and left at room temperature to thaw. Obtained homogenate was transferred to QIAshredder spin column and centrifuged at 13000 rpm for 2 minutes. The supernatant was transferred to a 1.5 ml Eppendorf tube and mixed with 99.8% ethanol (0.5 transfer volume). The solution was transferred in an RNeasy spin column placed in 2 mL collection tube and centrifuged at 8000 rpm for 15 s. 700 µL of RW1 buffer was added and the sample was centrifuged for at 8000 rpm for 15 s to wash the column. 500 µL of RPE buffer was then added and the sample was centrifuged at 8000 rpm for 2 minutes to wash the column. The RNA was eluted with 30 µL of RNAse free water added directly on the column

placed in a fresh 1.5 ml Eppendorf tube, followed by centrifugation at 8 000 rpm for 1 minute. The RNA concentration was measured on NanoDrop One Microvolume UV-Vis Spectrophotometer. Ambion® DNA-freeTM kit was used for removing any undesirable DNA contamination. 10x DNase I buffer, DNase I enzyme, extracted RNA and Nuclease-free water was mixed and incubated at 37°C for 30 minutes. To deactivate the DNase enzyme, 0.1 reaction volume of DNase Inactivation Reagent was added and the solution was then centrifuged at 8000 rpm for 1 minute. Supernatant containing RNA was transferred to a new 1.5 mL Eppendorf tube. RNA sample was stored at -80°C and processed in downstream experiments as soon as possible.

Generation of cDNA by reverse transcription

Obtaining the cDNA from RNA samples was accomplished by reverse transcription. The ImProm-IITM Reverse Transcription System (Promega) was used for this purpose. 1 µg of DNase-treated RNA was mixed with 1 µL of oligo-dT 18mers and nuclease-free water in the volume of 5 µL, denatured for 5 minutes at 70°C and immediately cooled on ice. 5 µL of the mixture was added to 15 µL of the reaction mix (4 µL of the ImProm-IITM 5x reaction buffer; 2 mM $MgCl_2$ solution; 0.67 mM dNTP mix; 20 U of Recombinant RNasin® Ribonuclease Inhibitor; 1 µL of ImProm-IITM Reverse Transcriptase stock in nuclease-free water) in small PCR tube. In a bench thermocycler, the reaction consists of 5 minutes annealing at 25°C, elongation of 60 minutes at 42°C and enzyme inactivation of 70°C for 15 minutes. Alternatively, M-MLV reverse transcriptase (Promega) was used, using the same denaturing procedure for 2 µg of the extracted RNA mixed with 2 µL mRNA specific oligo-dT 18mers in the volume of 10 µL, that are then mixed with 15 µL reaction mix (5 µL of M-MLV 5X Reaction Buffer, 0.67 mM dNTP mix; 20 U of Recombinant RNasin® Ribonuclease Inhibitor, 1 µL ofM-MLV reverse transcriptase stock in nuclease free water). The reaction conditions were the same. Obtained cDNA was stored at -20°C and used as a template in subsequent PCR reactions.

Protein extraction, SDS PAGE and western blotting

Native protein extraction

Sample material was collected, weighted and immediately frozen in liquid nitrogen. Using sterile mortar cooled with liquid nitrogen, the sample was grinded and homogenized in 3:1

volume:sample weight of DEM protein extraction buffer (0.1 M tris-HCl pH 8.8; 100 mM EDTA; 1mM phenylmethansulfonyl fluoride; 5 mM dithiothreitol; 0.1M KCl; 10 mM Na_2SO_3; 0.1M glycine). Collected homogenate was centrifuged in 1.5 mL Eppendorf tube at 20000 rpm and 4 °C for 10 minutes. The supernatant was transferred to a new ice-cold Eppendorf tube, ready to be used in downstream experiments.

SDS-PAGE

MiniProtean 1 mm Kit (Biorad) was used for SDS-PAGE. Gel chamber glass was cleaned with aceton using cotton tissue and fixed to the stand. Resolving gel (10% acrylamide/bisacrylamide solution in 37.5:1 ratio; 0.375 M Tris-Cl pH 8.8; 0.1 M SDS; 0.1% freshly prepared ammonium persulfate; 0.05% Temed) was mixed and loaded with 5 mL pipette to the gel chamber approximately 2 cm under the top edge and topped with isopropanol or ddH_2O. After 45 minutes, water or isopropanol was removed and the stacking gel (5% acrylamide/bisacrylamide solution in 37.5:1 ratio; 0.125 M Tris-Cl pH 6.8; 0.1 M SDS; 0.06% freshly prepared ammonium persulfate solution; 0.2% Temed) was carefully loaded on top of the resolving gel. Appropriate comb was quickly but carefully inserted to the stacking gel. After approximately 30 minutes, the gel chamber glass was removed from the polymerization stand, transferred to the electrophoresis stand and filled with the electrode buffer (25 mM Tris; 0.192 M glycine; 0.1% SDS). Comb was then removed. Excessive remnants of the solid gel were carefully removed using 10 µL pipette tip. 2 volumes of the sample were mixed with one volume of 3x sample buffer (0.15 M Tris-Cl pH 6.8; 30% v/v glycerol; 6% w/v SDS; 0.15 M DTT; 0.03% bromophenol blue), heated for 5 min at 95 °C and loaded on the gel together with appropriate ladder. Constant voltage was used to separate the samples. At first, the voltage was set at 75 V for 30 minutes to concentrate the extracted proteins at the beginning of the resolving gel, then the sample was separated in the resolving gel at 150 V for approximately 50 minutes and switched off just before the bromophenol blue line reached the bottom of the gel. The gel was then removed by careful opening of the glass gel chamber and the stacking gel was excised using a small plastic spatula.

Coomasie brilliant blue staining

The polyacrylamide gel was placed in the Coomasie brilliant blue staining solution (5% aluminium-sulphate-(14-18)-hydrate; 0.02% Coomasie brilliant blue G250 dye; 10% ethanol and 2% ortho-phosphoric acid) and stained overnight. Next day, the staining solution was

removed and the gel was destained with distilled water for 1-3 hours until the background was washed away.

Semi-dry CSH transfer method for Western blotting

The gel was placed on top of 5-10 filter papers soaked in CSH Semi-dry transfer buffer (20% v/v methanol; 5.8 g/L Tris base; 2.9 g/L glycine; 0.37 g/L SDS), covered by blotting membrane soaked in the same buffer. Finally, another 5-10 filter papers soaked in the same buffer were placed on top. The whole sandwich was placed in the Semi-dry blotting machine, excess drops of the buffer were removed from the surrounding and the lid was tighten on the top. Transfer at constant voltage 25 V was ran for 22 minutes. After the transfer, the gel was stained by Coomassie brilliant blue dye staining solution.

Western blotting

The membrane was washed in the TBST buffer (10 mM Tris;150 mM NaCl; 0.05% Tween) 3 times for 5 minutes, blocked in milk solution (5% w/v milk powder in TBST) and washed once with TBST for 5 minutes. Then the membrane was incubated in the appropriate primary antibody solution (500-5000x diluted primary antibody, 0.5% bovine serum albumin in TBST) overnight at 8 °C. Next day, the membrane was washed 3 times in TBST and incubated in Alkaline-Phosphatase (AP) conjugated secondary antibody for 1 hour at 8 °C. Next, the membrane was washed 3-times in TBST, followed by an incubation in the AP buffer (100 mM Tris-Cl pH 9.5; 100 mM NaCl, 5 mM MgCl; 0.33 mg/mL nitro blue tetrazolium chloride, 0.15 mg/mL 5-bromo-4-chloro-indolylphosphate) for development of the AP signal (usually 1-15 minutes). After sufficient signal development, the reaction was stopped by replacing the AP buffer with distilled H_2O. The list of used antibodies is showed in Table 5.

DNA/RNA/Protein All-in-one extraction

In one case, we failed to obtain desired results using the native protein extraction with DEM buffer. We therefore repeated the extraction using commercial kit, the AllPrep®

Table 5 Table showing antibodies used in this work.

Type	Host / Target	From	Product code
primary	Rabbit anti-GFP (polyclonal)	Agrisera	AS15 2987
primary	Rabbit anti-HA (polyclonal)	Thermo Life	715500
primary	Rabbit anti-FLAG (polyclonal)	Sigma/Merck	F7425-
secondary	Goat Anti-Rabbit AP (polyclonal)	Sigma-Aldrich	A3687-.25ML

DNA/RNA/Protein Mini (Qiagen), which allowed to obtain genomic DNA, total RNA and proteins from one sample. Sample material was collected, weighted and immediately frozen in liquid nitrogen. Using sterile mortar filled with liquid nitrogen, the sample was grinded and homogenized in 100 µL of buffer RLT with 1% β-Mercaptoethanol. The homogenate was then transferred in AllPrep DNA spin column placed in 2 mL collection tube and centrifuged at 13000 rpm for 3 minutes. The column was placed in new collection tube, washed first with 500 µL of AW1 buffer and centrifuged at 13000 rpm for 15 seconds. Supernatant was discarded and the column was washed with 500 µL of AW2 buffer and centrifuged at 13000 rpm for 15 seconds. Then the column was placed in a fresh Eppendorf tube, 100 µL of EB buffer were added and the sample was incubated at room temperature for 2 minutes, followed by a centrifugation at 13000 rpm for 1 minute. The flow-through contained the genomic DNA of the sample. 72 µL of 96% ethanol were added to the flow-through from homogenate centrifugation and the solution was transferred to RNeasy spin column placed in a 2 mL collection tube and centrifuged at 13000 rpm for 15 seconds. From this point, the process of RNA isolation with RNeasy spin column placed is identical to the protocol described in RNA extraction section. To the flow-through from RNeasy spin column, 600 µL of buffer APP were added and the sample was incubated at room temperature for 10 minutes. Then, the sample was centrifuged at 18000 rpm for 10 minutes. Supernatant was removed and the pellet was left to dry out for 10 minutes. Then, 500 µL of 70% ethanol was added and the sample was centrifuged at 18000 rpm for 1 minute, the supernatant was removed and the pellet was left to dry out for 10 minutes. Then, 50 µL of buffer ALO were added and the sample was mixed to dissolve the pellet, followed by incubation at 95°C. The sample was then stored at -20 °C and used for further experiments.

Expression analysis

Real Time quantitative PCR

In order to quantify the expression of the eIF3A gene in inspected T-DNA mutant lines, the cDNA obtained from isolated RNA by Oligo-dT reverse transcription was analysed by Real Time quantitative PCR (RT-qPCR). The reactions were performed using GoTaq® qPCR Master Mix (Promega), each reaction consisting from 6 µL of the 2x GoTaq® qPCR Master Mix, 0.83 mM of each primer and 2 µL of sample cDNA, all in nuclease free water. For each sample, three technical replicas were tested on LightCycler 480 (Roche) for each primer pair used. Obtained Ct values were normalized to ACT8 (Actin 8; AT1G49240) endogenous control

to dCt (Ct_{sample} - Ct_{ACT8}) and relative expression in % of ACT8 expression was counted using following formula: 100/(2^dCt).

GUS staining for promoter analysis experiments

Expression analysis using the β-glucuronidase (GUS) reporter enzyme expressed under plant promoter of the inspected gene is a common tool to visualize the tissues that transcribe the studied gene. The activity of the enzyme was visualized by incubating transgenic plant tissue with solution containing the substrate of enzymatic cleavage: 5-bromo-4-chloro-3-indolyl glucuronide (X-Gluc). After cleavage, the product gives characteristic blue colour that does not migrate between tissues. Studied plant tissue (seedling, leaf, and inflorescence) was immersed in the staining solution (0.1 M NaPO4 pH 7.0, 10 mM EDTA, 0.1% Triton X-100, 1 mM X-Gluc, 1 mM FerroCyanide, X-Gluc) and vacuum infiltrated for 10 minutes to penetrate the whole tissue with the buffer. Then the sample container was covered in tinfoil and incubated at 37°C for signal development. Different incubation time was used to optimize the strength of the signal. After desired incubation time passed, the sample was transferred to 70% ethanol and destained either by incubation in 70% for 3 days or by ethanol dilution series: 70%, 90% and 70% ethanol, each for 1 hour. After destaining, series of 70%, 50% and 30% ethanol dilutions each for 1 hours was used, with final rinse with water and observation in water or 15% glycerol. For long term storage, the samples are rinsed with water and stored in 30% ethanol in dark at 8°C to overcome contamination of the sample.

Microscopy, statistical testing and Software used

Microscopy and image manipulation

Microscopy images were acquired on Nikon TE2000E fluorescent inverted microscope, ZEISS Stemi 508 stereomicroscope or Nikon laser spinning disc. For detailed specifications of the Nikon laser spinning disc visit here: http://www.ueb.cas.cz/en/if/equipment. Protein gels and western blot membranes were scanned on common office scanner on high DPI and the images acquired were then saved in TIFF formate. Acquired images were processed in Fiji (ImageJ) software or less usually in GIMP 2 software. Schematic figures were illustrated in Inkscape vector graphics.

Statistical testing

Where appropriate, statistical testing was used to support the observations. Mean and standard deviations of the measurements are written as "mean ± st. deviation" format. For comparison of expected/observed genotype frequencies, goodness-of-fit test based on χ^2 statistics was used. For comparison of defect frequency measurements, one-tailed T-test for 2 independent means was used.

Computational biology

In order to investigate the previously unexplored eIF3 subunit A in plants, a basic computation biology survey of eIF3A protein sequences was performed. The overall process of questions asked and steps taken in order to answering them is depicted in the Figure 9, together with the online software used in all the procedures. For sequence and MSA organization and easy handling, Seaview software was used For the eIF3a computational biology analyses and comparisons, we searched several plant and non-plant species: *Arabidopsis thaliana* (taxid: 3702), *Nicotiana tabacum* (taxid: 4097), *Oryza sativa* (taxid: 4530), *Amborella trichopoda* (taxid: 13333), *Physcomitrella patens* (taxid: 3218), the green algae *Chlamydomonas reinhardtii* (taxid: 3055) and four non-plant species; *Homo sapiens* (taxid: 9606), *Drosophila melanogaster* (taxid: 7227), *Caenorhabditis elegans* (taxid: 6239), *Saccharomyces cerevisiae* (taxid: 4932). The exact software used for every particular *in silico* analysis is stated in the respective subchapter in Results for better understanding of the analysis flow. Generally, we used all the software with its basic settings.

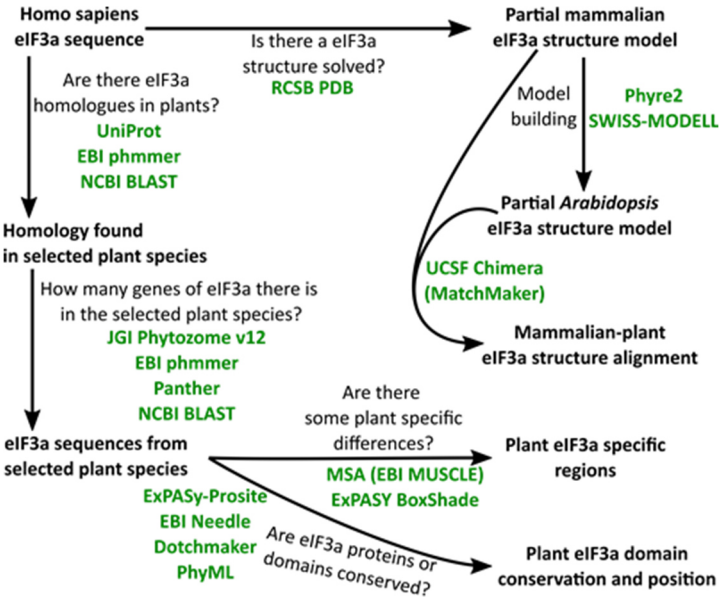

Figure 9: Computational biology mind map. Schematic depiction of the computational biology analysis steps. The aim was to search the databased for eIF3a homologues in plant species and their comparison with non-plant eukaryotic species. The main comparisons were done on Multiple sequence alignment and structural alignment of modelled plant eIF3a subunits. For that purpose, numerous software was used, most of which are available online. In each step, a question motivating the procedure is asked and software used is noted.

Results

Computational biology

As no literature reported on eIF3A protein characterization in plants, we tried to perform basic computational biology analysis, using the known and well characterized eIF3a protein sequences from other eukaryotes as templates to find similar or different domains and regions with described function.

Usually one eIF3A protein (gene) is found in plant and non-plant species

First, we searched the protein sequence databases using (UniProt) for obtaining the reviewed mammalian eIF3A protein sequence (*Homo sapiens* eIF3A, Uniprot code: Q14152) and used it as template for phmmer online search (https://www.ebi.ac.uk/Tools /hmmer/search/phmmer). The search found a hit in all organisms included in the search. With exception of 5 hits in *N. tabacum* and 3 hits in *P. patens*, all other organisms contained one homologue of human eIF3A. To reduce the possibility of hits that are actually fragments of the same gene, which is common between non-reviewed entries in databases, the results of the phmmer search were then compared with Uniprot, Panther and plant Phytosome databases to confirm the number of paralogue sequences found in two plant species mentioned. This reduced the number of *N. tabacum* sequences from 5 to 4, as one hit was a protein fragment of the same locus (hits Uniprot IDs: A0A1S3X135 and Q40554, both locus LOC107760201). The resulting summary of protein sequences found by the search are shown in Table 6. In non-plant species, one homologue was found in each species without exceptions. It was apparent that between the non-plant species, the eIF3A length differs and tends to increase towards more complex organisms. Nevertheless, this analysis worked with limited non-plant species to elucidate such conclusions. In *H. sapiens,* the extra 400 amino acids are explained by the addition of DRYG domain that is a part of Vertebrate-specific interleukin signalization pathway (see subchapter eIF3A - protein evolution). In plant species, the length of eIF3A orthologues was persistently ranging around the same length (from 955 amino acids of the eIF3A.1_TOBAC to 1021 amino acids of the eIF3A.3_PHYPA).

Table 6: Table showing summarized homology search results. The table is a result of search for eIF3A homologues in eukaryotes. We compared results from phmmer search engine using human eIF3a sequence as a template and available protein databases (Phytozome, Uniprot, Panther) in search for plant eIF3A homologues. Four non-plant annotated eIF3a orthologues were added as a reference sequences for further comparisons. The lenght and Uniprot accession codes are included in the table and working names are proposed for simplification in

Species	Homologues	Protein length	Uniprot accession code	Working codename
Homo sapiens	1	1382	Q14152	eIF3A_HUMAN
Drosophila melanogaster	1	1140	Q9VN25	eIF3A_DROME
Caenorhabditis elegans	1	1076	P34339	eIF3A_CAEEL
Saccharomyces cerevisiae	1	964	P38249	eIF3A_YEAST
Arabidopsis thaliana	1	987	Q9LD55	eIF3A_ARATH
Nicotiana tabacum	4	955	A0A1S4ASL9	eIF3A.1_TOBAC
		958	A0A1S3Z3K3	eIF3A.2_TOBAC
		956	A0A1S3X5F8	eIF3A.3_TOBAC
		958	Q40554	eIF3A.4_TOBAC
Oryza sativa	1	986	Q5ZEN1	eIF3A_ORYSJ
Amborella trichopoda	1	956	W1NY10	eIF3A_AMBTC
Physcomitrella patens	3	1008	Pp3c4_26450	eIF3A.1_PHYPA
		1013	Pp3c3_10600	eIF3A.2_PHYPA
		1021	Pp3c26_11290	eIF3A.3_PHYPA
Chlamydomonas reinhardtii	1	1033	A0A2K3CZ51	eIF3A_CHLRE

eIF3A orthologues share high levels of sequence identity and similarity in plants

We next compared the protein sequences of all eIF3A orthologues found in the search. Computing sequence identity and sequence similarity is a useful and reliable computational biology strategy to determine the rate of homology between sequences. We used EMBOSS Needle online tool to compute the sequence similarity from pairwise global alignments between the eIF3A orthologues (default settings). Although Needle computes the sequence identity as well, a better option was available, based on Multiple sequence alignment (MSA) of full protein sequences using EMBOSS MUSCLE online tool (default settings) that computes the sequence identity between multiple inputs (unfortunately MUSCLE does not show similarity). The results of protein sequence identity and similarity are shown in Table 7. The overall results are clearly showed how eIF3A sequences from plant species with multiple eIF3A paralogs (*N. tabacum* and *P. patens*) share very high levels of identity (87-97 % and 84-87 %, respectively) and similarity (92-98 % and 91-93 %, respectively). Also, eIF3A proteins share high levels of identity (60-72 %) and similarity (70-84 %) when plant eIF3A proteins are compared. Both variables differ more when plant eIF3A proteins are compared with green algae *C.reindhartii* (44-47 % identity and 58-62 % similarity). Not surprisingly, comparison of plant and non-plant

Table 7: Table showing sequence Identity (% ID) and sequence similarity (% SIM) between eIF3A orthologues. Sequence Identity values (percents) were obtained from multiple sequence alignment (Emboss MUSCLE online tool, default settings) of full length protein sequences. Sequence Similarity values (percents) were obtained from respective pairwise sequence alignments (EMBOSS Needle online tool, default settings) of full length protein sequences. The plant and non-plant eIF3A comparisons are graphically separated. Heatmap depicts obtained values from lowest value (red) to the highest value (green).

% ID / % SIM	eIF3A_HUMAN	eIF3A_DROME	eIF3A_CAEEL	eIF3A_YEAST	eIF3A_ARATH	eIF3A.1_TOBAC	eIF3A.2_TOBAC	eIF3A.3_TOBAC	eIF3A.4_TOBAC	eIF3A_ORYSJ	eIF3A_AMBTC	eIF3A.1_PHYPA	eIF3A.2_PHYPA	eIF3A.3_PHYPA	eIF3A_CHLRE
eIF3A_HUMAN	100	48	35	27	36	37	37	37	36	36	37	38	38	38	35
eIF3A_DROME	53	100	33	27	33	34	33	34	33	32	34	35	35	34	30
eIF3A_CAEEL	43	50	100	24	30	31	31	32	31	30	32	31	31	31	27
eIF3A_YEAST	32	39	42	100	27	28	26	27	26	27	28	28	29	30	28
eIF3A_ARATH	42	49	47	44	100	72	71	71	70	65	67	61	61	61	45
eIF3A.1_TOBAC	42	48	48	43	84	100	89	97	87	69	69	63	62	61	45
eIF3A.2_TOBAC	42	47	46	46	82	94	100	88	96	68	69	63	62	61	45
eIF3A.3_TOBAC	42	47	48	46	83	98	94	100	87	69	68	63	62	60	45
eIF3A.4_TOBAC	42	47	45	45	82	93	97	92	100	67	68	63	62	61	45
eIF3A_ORYSJ	42	48	47	46	77	80	80	80	78	100	69	62	61	60	44
eIF3A_AMBTC	42	46	45	45	79	81	81	81	80	79	100	70	70	68	46
eIF3A.1_PHYPA	43	51	48	45	73	73	73	72	72	70	76	100	89	87	47
eIF3A.2_PHYPA	44	49	47	42	72	71	71	71	71	71	75	93	100	84	46
eIF3A.3_PHYP	42	48	46	44	73	74	72	73	72	72	78	91	91	100	47
eIF3A_CHLRE	41	48	43	41	62	59	58	60	59	61	61	62	59	62	100

eIF3A sequences gave relatively lower amount of identity and similarity (26-38 % and 41-51 %, respectively). Notably, comparing the sequence identity only, the most distant eIF3A orthologue to plants was the one of *S. cerevisiae* (26-30 % identity). However, at the level of sequence similarity, the levels are very close to other non-plant species (41-46 %).

Evolutionary relationships of eIF3A proteins separates plant and non-plant sequences

To visualize evolutionary relationships between found eIF3A orthologues, we built a phylogenetic tree based on MSA alignment (EMBOSS MUSCLE) of first 900 amino acid residues (Figure 10). We used the Seaview software with integrated PhyML tool and created a phylogenetic tree based on estimating maximum likelihood phylogenies (LG model, 1000 Bootstrap replicates, optimized across site rate variation, BioNJ starting tree with optimized tree topology and 5 random starts). The evolutionary relationships of eIF3A visualized with the cladogram clearly separated plant and non-plant eIF3A proteins, as well as *C. reindhartii* and true plant eIF3A. Within the plant branch, the eIF3A relationships showed no links on diversification in the eIF3A family and clearly follow the current view on the plant evolutionary relationships, moreover confirming that one eIF3A gene is the usual copy number, at least in plant species investigated in this work.

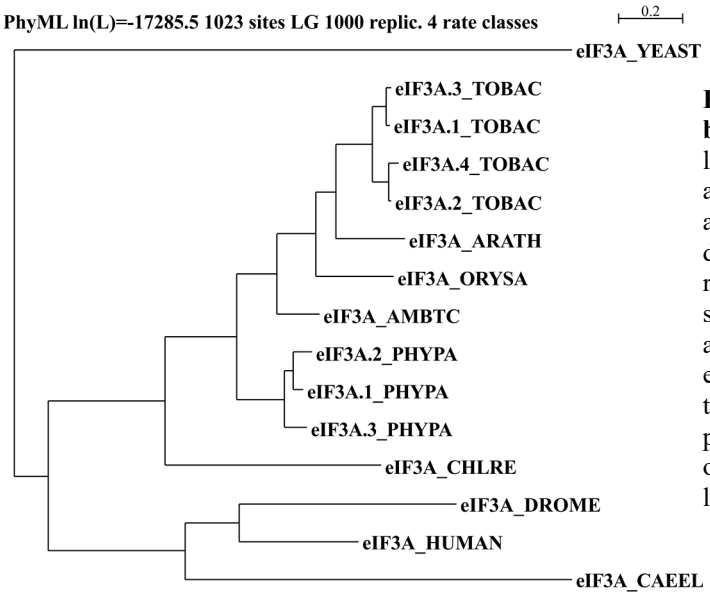

Figure 10: Evolutionary relationship between eIF3a homologues. Maximum likelihood based cladogram build on MSA alignment of the well conserved first 900 amino acids of the eIF3a protein sequences downloaded from UNIPROT. The tree respects the divergence of plant eIF3a sequences from other non-plant species and also separates the yeast eIF3a from other eukaryotes. Although in P. patens and N. tabacum, eIF3a gene multiplicated into 3 or 5 paralogue genes, respectively, the ground state of eIF3a family in plants seems to be at least/usually one gene in the genome.

Plant eIF3A sequence complexity is similar to the non-plant homologues

To easily visualize the overall complexity of protein sequences, we used EMBOSS Dotmatcher online tool (default settings: window size 10, threshold 23, BLOSUM62 matrix) to create self-dotplots of found eIF3A protein sequences. We challenged all found eIF3A protein sequences to this analysis to visualize sequence complexity and possible repetitive or low complexity regions. The representative Dotplots are shown in Figure 11. It was apparent from the analysis that all plots have a highly similar pattern, where approximately first 600 amino acids form a unique region. Then, more to the C-terminus, a region of low complexity follows, spreading from 600-900 amino acids in the sequence. It is notable that the human eIF3A possess this low complexity region at the same position within the protein that is followed by approx. 500 amino acids of another region of low complexity. Within this region, the Vertebrate-specific DRYG repeat domain was clearly visualized. Therefore, the extraordinary length of human eIF3A is truly an enlargement of the C-terminal part that follows the first low complexity region, which is present in most but not all the plots. The only exception was the budding yeast eIF3A, where the low complexity region was not as well apparent as it was in all other plots.

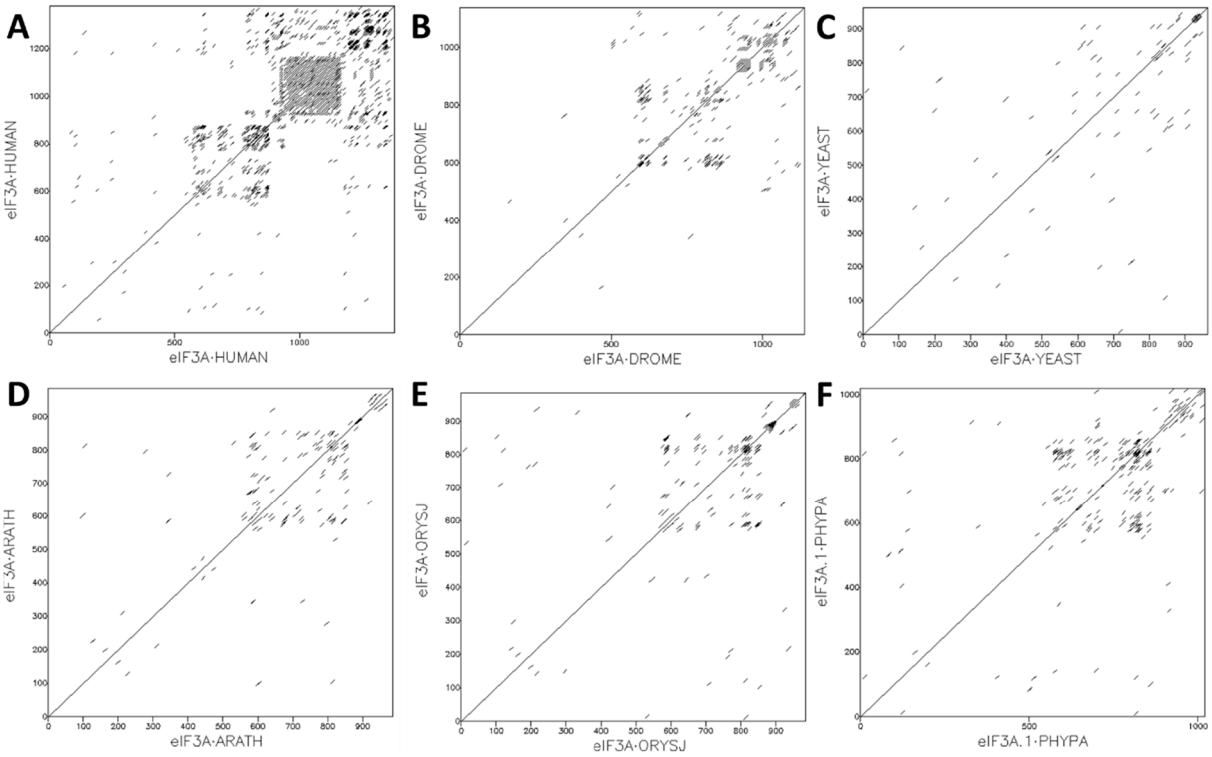

Figure 11: Self-dotplots of 3 non-plant and 3-plant species. Dotplots were generated with EMBOSS dotmatcher online tool. **(A-F)** Self-dotplots of H. sapiens, D. melanogaster, S. cerevisiae, A. thaliana, O. sativa, P.patens protein chains, respectively. Protein codenames are indicated at both axes of each chart.

The length of PCI domain differs in plant and non-plant species

We challenged found sequences to the Expasy-Prosite domain search (default settings) to annotate the plant eIF3 proteins. In all selected plant species, the PCI domain was recognized by the algorithm with a strict position in the protein sequence (from amino acid 316 to 513, spanning the region of 197 amino acid residues). Only in green algae, *C. reindhartii*, PCI domain position was not identically placed. In non-plant eIF3A sequences, the PCI domain was recognized and localized to the same relative position, although exact start-stop position was slightly more variable between chosen species. Strikingly, when the overall length of the PCI domain is calculated ($PCI_{(end.position)}$-$PCI_{(start.position)}$), its length differs between plant and non-plant species. While in non-plant species eIF3A PCI domain length varied around 180 amino acids (175-183), the plant PCI domain length in eIF3A was exactly 197 amino acid residues in all cases, which is approx. 15 amino acids longer when compared with non-plant species. The green algae *C.reindhartii* eIF3A PCI domain length was even longer than the plant ones (205 amino acid residues). Altogether, plant eIF3A proteins contain the structural PCI domain on the same position as non-plant species, but the recognized PCI domain is longer in all plant species, when compared to the non-plant species.

PCI domain of plants possess two additional regions and missing one region

To inspect the PCI domain differences more specifically, we built a Multiple Sequence Alignment (MSA), using MUSCLE online tool with default options on all plant and non-plant protein sequences as input. The MSA highlighted three regions within the PCI domain that are different specifically in plants (Figure 12A). Two additional stretches of 8 (Region 1) and approx. 10 (Region 2) amino acid residues are present in all plant species and highly conserved between them, positioned on 365-372 (Region 1) and 425-434 (Region 2) segments. These two regions appear to be plant-specific, at least in this analysis, as all plant species contain these regions, while non-plant species do not. Moreover, these regions are quite highly conserved between plant species. Even more surprisingly, at the C-terminal PCI border (as it is recognized by the Prosite algorithm), all plant eIF3A proteins lack a stretch of amino acids (Region 3), that has also variable length in non-plant species. Accordingly to the alignment, between amino acid 510 and 511 of all plant eIF3A protein sequences there are additional 16, 16, 22, 10 amino acid residues in *H. sapiens*, *D. melanogaster*, *C. elegans* and *S. cerevisiae*, respectively. It is notable that in non-plant species, the PCI domain recognized by the Prosite ends within the region plants do not contain. If these extra regions would be included in non-plant species, it would compensate a portion of the length differences mentioned in previous subchapter.

mRNA interacting regions conserved in plants

A study in yeast revealed that eIF3a subunit is involved in the reinitiation after short uORF in the 5′UTR, where it interacts with specific secondary structures of the mRNA, specifically with aa regions 51-60, 71-80, 161-170 (Szamecz et al., 2008; Munzarová et al., 2011). The MSA done in previous subchapter showed that in plants, these regions share a high number of identical or similar amino acid residues and are highly conserved among all studied species (Figure 12B). Evidences that eIF3a is a RNA-binding protein have been reported in plants (see eIF3a in plants subchapter in the literature overview for reference). However, there is no direct evidence that in plants, these regions have the mRNA interacting/binding ability.

eIF3b/i binding platform has conserved organization in plants

eIF3a binds eIF3b and eIF3i with its C-terminal end in yeast and mammals. In mammals, it has been shown that regions 756-762 and 766-775 are responsible for direct interaction with eIF3b and eIF3i, respectively, and that this region is conserved in animal kingdom. Key amino

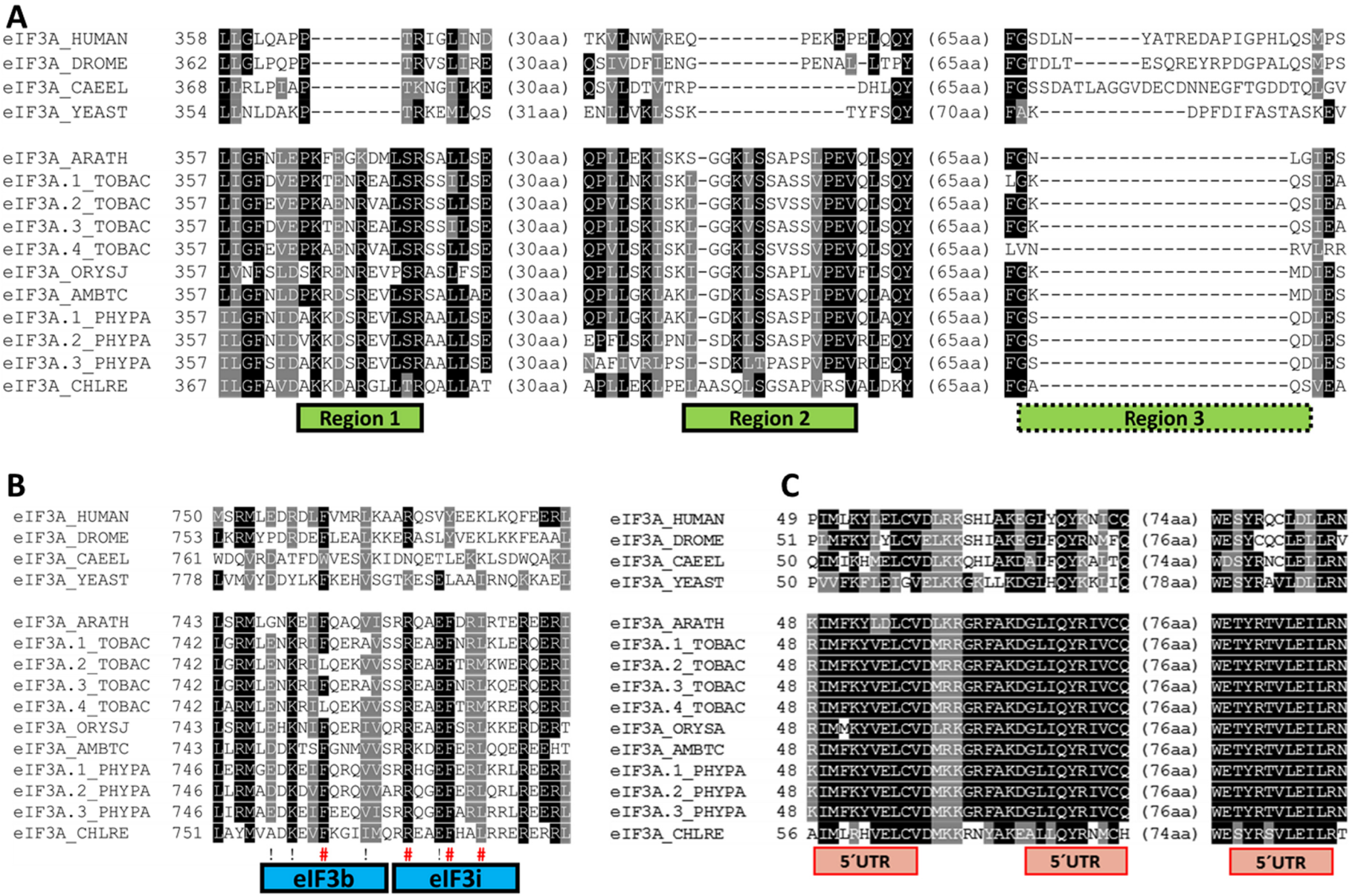
A
eIF3A_HUMAN 358 LLGLQAPP--------TRIGLIND (30aa) TKVLNWVREQ---------PEKEPELQQY (65aa) FGSDLN------YATREDAPIGPHLQSMPS
eIF3A_DROME 362 LLGLPQPP--------TRVSLIRE (30aa) QSIVDFIENG---------PENAL-LTPY (65aa) FGTDLT------ESQREYRPDGPALQSMPS
eIF3A_CAEEL 368 LLRLPIAP--------TKNGILKE (30aa) QSVLDTVTRP--------------DHLQY (65aa) FGSSDATLAGGVDECDNNEGFTGDDTQLGV
eIF3A_YEAST 354 LLNLDAKP--------TRKEMLQS (31aa) ENLLVKLSSK-------------TYFSQY (70aa) FAK-------------DPFDIFASTASKEV
eIF3A_ARATH 357 LIGFNLEPKFEGKDMLSRSALLSE (30aa) QPLLEKISKS-GGKLSSAPSLPEVQLSQY (65aa) FGN----------------------LGIES
eIF3A.1_TOBAC 357 LIGFDVEPKTENREALSRSSILSE (30aa) QPLLNKISKL-GGKVSSASSVPEVQLSQY (65aa) LGK----------------------QSIEA
eIF3A.2_TOBAC 357 LIGFEVEPKAENRVALSRSSLLSE (30aa) QPVLSKISKL-GGKLSSVSSVPEVQLSQY (65aa) FGK----------------------QSIEA
eIF3A.3_TOBAC 357 LIGFDVEPKTENREALSRSSILSE (30aa) QPLLSKISKL-GGKVSSASSVPEVQLSQY (65aa) FGK----------------------QSIEA
eIF3A.4_TOBAC 357 LIGFEVEPKAENRVALSRSSLLSE (30aa) QPVLSKISKL-GGKLSSVSSVPEVQLSQY (65aa) LVN----------------------RVLRR
eIF3A_ORYSJ 357 LVNFSLDSKRENREVPSRASLFSE (30aa) QPLLSKISKI-GGKLSSAPLVPEVFLSQY (65aa) FGK----------------------MDIES
eIF3A_AMBTC 357 LLGFNLDPKRDSREVLSRSALLAE (30aa) QPLLGKLAKL-GDKLSSASPIPEVQLAQY (65aa) FGK----------------------MDIES
eIF3A.1_PHYPA 357 ILGFNIDAKKDSREVLSRAALLSE (30aa) QPLLGKLAKL-GDKLSSASPIPEVQLAQY (65aa) FGS----------------------QDLES
eIF3A.2_PHYPA 357 ILGFNIDVKKDSREVLSRAALLSE (30aa) EPFLSKLPNL-SDKLSSASPVPEVRLEQY (65aa) FGS----------------------QDLES
eIF3A.3_PHYPA 357 ILGFSIDAKKDSREVLSRAALLSE (30aa) NAFIVRLPSL-SDKLTPASPVPEVRLEQY (65aa) FGS----------------------QDLES
eIF3A_CHLRE 367 ILGFAVDAKKDARGLLTRQALLAT (30aa) APLLEKLPELAASQLSGSAPVRSVALDKY (65aa) FGA----------------------QSVEA
Region 1
Region 2
Region 3
B
eIF3A_HUMAN 750 MSRMLEDRDLFVMRLKAARQSVYEEKLKQFEERL
eIF3A_DROME 753 LKRMYPDRDEFLEALKKERASLYVEKLKKFEAAL
eIF3A_CAEEL 761 WDQVRDATFDWVESVKIDNQETLEKKLSDWQAKL
eIF3A_YEAST 778 LVMVYDDYLKFKEHVSGTKESELAAIRNQKKAEL
eIF3A_ARATH 743 LSRMLGNKEIFQAQVISRRQAEFDRIRTEREERI
eIF3A.1_TOBAC 742 LGRMLENKRIFQERAVSSREAEFNRLKLERQERI
eIF3A.2_TOBAC 742 LGRMLENKRILQEKVVSSREAEFTRMKWERQERI
eIF3A.3_TOBAC 742 LGRMLENKRIFQERAVSSREAEFNRLKQERQERI
eIF3A.4_TOBAC 742 LARMLENKRILQEKVVSSREAEFTRMKRERQERI
eIF3A_ORYSJ 743 LSRMLEHKNIFQERIVQRREAEFSRLKKERDERT
eIF3A_AMBTC 743 LLRMLDDKTSFGNMVVSRRKDEFERLQQEREEHT
eIF3A.1_PHYPA 746 LERMGEDKEIFQRQVVSRRHGEFERLKRLREERL
eIF3A.2_PHYPA 746 LLRMADDKDVFQRQVVARRQGEFERLQRLREERL
eIF3A.3_PHYPA 746 LIRMAEDKEIFEEQVISRRQGEFARLRRLREERL
eIF3A_CHLRE 751 LAYMVADKEVEKGIIMQRREAEFHALRREREREL
eIF3b
eIF3i
C
eIF3A_HUMAN 49 PIMLKYLELCVDLRKSHIAKEGIYQYRNICQ (74aa) WESYRQCLDLLRN
eIF3A_DROME 51 PLMFKYIYLCVELKKSHIAKEGIFQYRNMFQ (76aa) WESYCQCLELLRV
eIF3A_CAEEL 50 QIMIKHMELCVDLKKQHIAKDAIFQYKAITQ (74aa) WDSYRNCLELLRN
eIF3A_YEAST 50 PVVFKFLEIGVELKKGKLLKDGIHQYKKIIQ (78aa) WESYRAVLDLLRN
eIF3A_ARATH 48 KIMFKYLDLCVDLKRGRFAKDGLIQYRIVCQ (76aa) WETYRTVLEILRN
eIF3A.1_TOBAC 48 RIMFKYVELCVDMRRGRFAKDGLIQYRIVCQ (76aa) WETYRTVLEILRN
eIF3A.2_TOBAC 48 RIMFKYVELCVDMRRGRFAKDGLIQYRIVCQ (76aa) WETYRTVLEILRN
eIF3A.3_TOBAC 48 RIMFKYVELCVDMRRGRFAKDGLIQYRIVCQ (76aa) WETYRTVLEILRN
eIF3A.4_TOBAC 48 RIMFKYVELCVDMRRGRFAKDGLIQYRIVCQ (76aa) WETYRTVLEILRN
eIF3A_ORYSA 48 RIMMKYVELCVDLRKGRFAKDGLIQYRIVCQ (76aa) WETYRTVLEILRN
eIF3A_AMBTC 48 RIMFKYVELCVDMRRGRFAKDGLIQYRIVCQ (76aa) WETYRTVLEILRN
eIF3A.1_PHYPA 48 KIMFKYVELCVDMKKGRFAKDGLIQYRIVCQ (76aa) WETYRTVLEILRN
eIF3A.2_PHYPA 48 KIMFKYVELCVDMKKGRFAKDGLIQYRIVCQ (76aa) WETYRTVLEILRN
eIF3A.3_PHYPA 48 KIMFKYVELCVDMKKGRFAKDGLIQYRIVCQ (76aa) WETYRTVLEILRN
eIF3A_CHLRE 56 AIMLRHVELCVDMKKRNYAKEALLQYRNMCH (74aa) WESYRSVLEILRT
5´UTR
5´UTR
5´UTR

Figure 12 (previous page): Regions of interest in the MSA of selected eIF3a orthologues. (A) Multiple sequence alignment highlighting three regions in the PCI domain, where plants clearly differ. **(B)** Conservation of the eIF3b/eIF3i binding platform found in mammals (eIF3A_HUMAN in the first row) between plant and non-plant species. Described important amino acids are highlighted with # or ! depending on their overall conservation in the presented MSA **(C)** Conservation of the mRNA interacting segments characterized in budding yeast (eIF3A_YEAST in the fourth row).

acids were Phe760 for eIF3b binding and Arg768, Tyr772 and Lys775 for eIF3i binding (Dong et al., 2013). The MSA done in previous subchapter showed that in plants, a region located on similar position as eIF3b and eIF3i binding platform in *H. sapiens* shares identical or similar amino acid organization in the key positions (Figure 12C). However, there is no direct evidence that in plants, this region truly is an eIF3b/i interaction platform.

AteIF3A structural model locates the additional regions of plant PCI domain into interhelical loops

Homologous proteins are usually more conserved in their tertiary structure than in primary protein sequence. Since the eIF3A identity/similarity between all species investigated in this work were relatively high for such distant species, we tried to build a plant eIF3A structural model. We used the eIF3A partial structure from the eIF3 Octamer model (PDB code 5A5T, Figure 13A) as a template to build a plant eIF3A protein structure model based on the N-terminal 600 amino acid query of the A. thaliana protein sequence. We used 2 different online structure modelling tools; Swiss-Model and Phyre2 to build two different models (Figure 13B-C). Then, using the software UCSF Chimera Matchmaker, we coloured and aligned the .pdb maps of the structures obtained from the PBD (template 5A5T) and AteIF3A from modelling software (AteIF3A(Swiss) and AteIF3A(Phyre2)) (Figure 13D, middle section). Then, since the template partial eIF3A model contains the PCI domain, we were able to visualize the PCI differences specific to plants in the context of eIF3A tertiary structure (Figure 13). All three regions were modelled by both software used as loops between two following helixes, which suggests these regions might not interfere with the overall eIF3A protein structure and might provide some specific interactions. Region 2 is an interhelical loop, according to both modelling software at least, and is located to the protein surface between helix 22 and 23 (Figure 13D, right section). Region 1 is an interhelical loop as well and is located between helix 17 and 18. When compared to the mammalian model, the Region 3 is removing one short helix 25 in plants and brings helix 24 and 26 closer together (Figure 12D, left section). Interestingly, Region 1 and Region 3 are placed in a close proximity to each other in the tertiary structure of the protein, despite being almost 200 amino acid residues apart from each other on the primary protein

structure (Figure 13D, left section). The interhelical loop of Region 2 is occupying a part of the same area in plants that is otherwise filled with Region 3 in non-plant species (Figure 13D, left section).

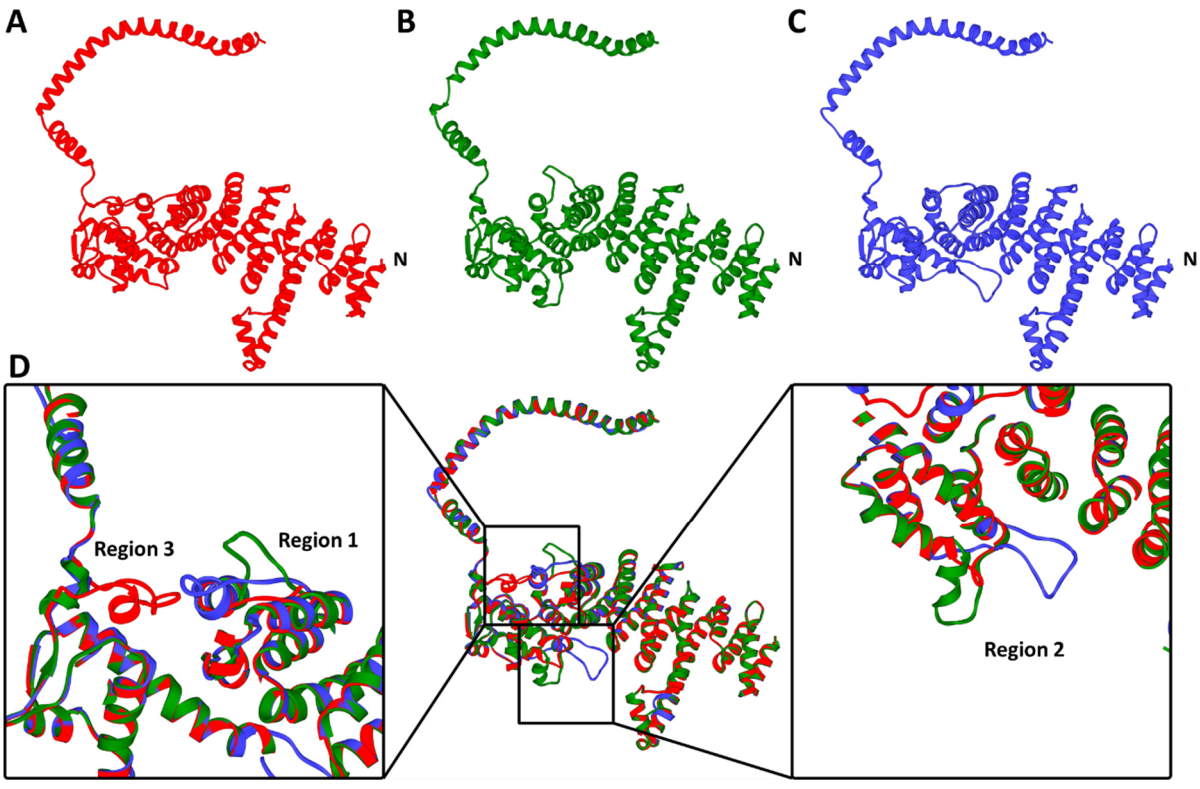

Figure 13: Arabidopsis eIF3a models with highlighted regions in the eIF3a PCI domain. Using available partial mammalian eIF3a structure (PDB code: 5A5T), we made models of Arabidopsis thaliana eIF3a subunit. **(A)** original partial mammalian eIF3a (N-terminal 600 aa; red; PDB code: 5A5T); **(B)** partial model of Arabidopsis thaliana eIF3a modelled by SWISS-MODELL (green) using 5A5T as template **(C)** partial model of Arabidopsis thaliana eIF3a modelled by Phyre2 (blue) using 5A5T as template. **(D)** Two Arabidopsis thaliana models in structural alignment with the template sequence. Regions 1-3 in the PCI domain are showed in greater detail (see alignment in Figure 12A) Files obtained in .pdb format from PDB, SWISS-MODEL and Phyre2 were colored and aligned using the USCF Chimera tool MatchMaker.

Expression pattern of the *AteIF3A* gene

GFP-GUS fusion driven by the 750bp native AteIF3A promoter was used for revealing the gene expression pattern

To investigate the expression of *AteIF3A* gene (At4g11420), Columbia-0 plants were transformed with the expression cassette of β-glucuronidase reporter gene fused with green fluorescent protein driven by the native promoter, finally forming the p3A::GFP-GUS expression cassette. The native *AteIF3A* promoter consists of the 750 bp long region upstream of the At4g11420 transcription start site. Choosing longer region than 750 bp would interfere with the coding sequence of neighbouring gene At4g11430. In the T_1 generation, quick screen for GFP activity in mature pollen was performed, marking GFP-positive lines with no mature pollen defects. Seeds from 11 independent positive lines were harvested and plated on ½ MS containing kanamycin. After 7 days, seedlings of independent lines were analysed by GUS staining, while some were cultivated further. Observation of the GFP signal in the p3A::GFP-GUS was done using Nikon laser spinning disc.

GUS staining revealed AteIF3A expression in 7-Day old seedlings

7-days old kanamycin resistant seedlings of each line were aseptically removed from the media and submerged into GUS-staining solution. After 1 hour, 3 hours and 24 hours, two or three seedlings per line were removed and destained by subsequent ethanol dilutions series and stored in H_20 before being analysed under stereomicroscope for blue precipitate formed in tissues expressing GUS.

Staining of 1 hour revealed the blue histochemical signal in the root system in most of the lines, strongly staining the meristematic zone of main and lateral roots (Figure 14A,C), the central cylinder and root primordia emerging from the pericycle (Figure 14D). Few lines showed signal at the cotyledon tip, cotyledon vascular bundles (Figure 14E) or in the hypocotyl or young leaves. After 3 hours of staining, root system was strongly stained and blue signal became saturated in few lines. Almost all lines had the cotyledon tip and cotyledon vascularity stained, few lines developed a signal in mesenchymal cells cotyledons or within the young leaves (Figure 14E). Staining of the young leaves and SAM was present in several lines as well (Figure 14G). Interestingly, some lines had unevenly stained leaves or small patches of expression in the leaf tissues. Moreover, in these lines, some lateral root tips were stained unevenly or not stained at all (not shown), whereas the main root or lateral roots were stained

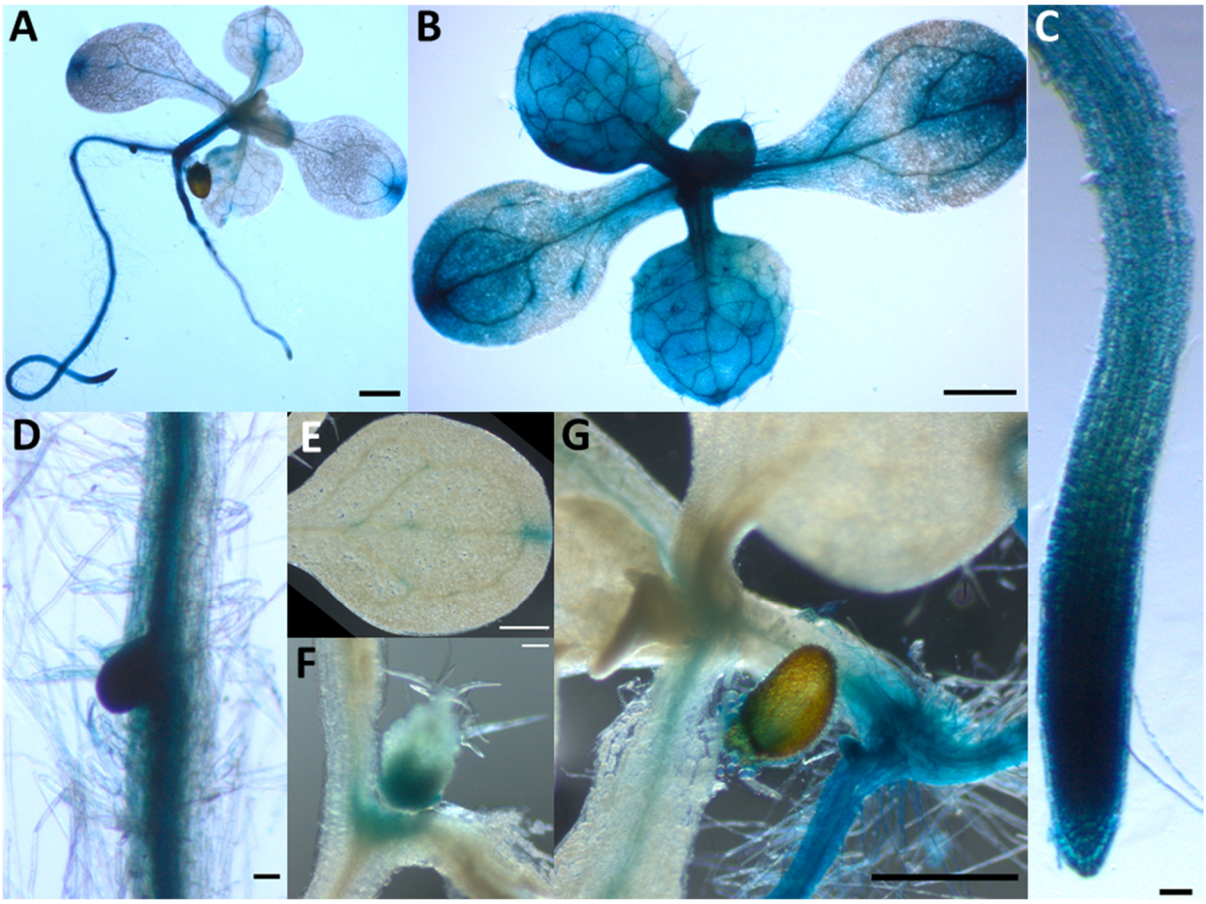

Figure 14: **Native AteIF3A promoter expression in 7-day old seedlings.** GUS staining of transgenic p3A::GFP-GUS 7-day old seedlings. **(A)** GUS expression in the seedling, Scale bar = 1mm. **(B)** Shoot view from above, Scale bar = 1mm. **(C)** Expression in the root meristem Scale bar = 50µm. **(D)** Expression in the central cylinder and root primordia Scale bar = 50µm. **(E)** Detailed look on the cotyledon, Scale bar = 500µm. **(F)** Expression in SAM, Scale bar = 50µm. **(G)** Expression in hypocotyls and vascular bundles of the shoot, Scale bar = 1 mm.

evenly. 24 hours of staining had not revealed any novel tissue with GUS expression, when compared to the 3-hour stain. Root system of all lines was overstained, while staining of cotyledons and young leaves varied greatly between independent lines, but not between individuals from the same line. The variability of the signal in leaves was observed in a range from all tissues strongly stained to small patches of blue in the whole cotyledon or young leaf only. However, majority of lines fitted somewhere between the two extremes observed. In lines, where shoot signal was weaker, the strong signal present in roots had a clear border at the hypocotyl (Figure 15D).

Therefore, general conclusion was that *AteIF3A* gene is strongly expressed in dividing cells of the root system, root and shoot vascular tissues, while expression in cotyledons and young leaves surrounding the SAM has an unclear pattern. To inspect the shoot expression on more developed shoots, a decision was made to repeat this procedure with 14-days old seedlings as well.

GUS staining of 14-days old seedlings confirmed the expression of AteIF3A in Arabidopsis

14-days old kanamycin resistant seedlings of each line were aseptically removed from the media and submerged into GUS-staining solution. After 1 hour, 3 hours and 24 hours, two or three seedlings per line were removed and destained by ethanol dilutions series and the signal was observed.

Staining of 1 hour had similar results as with 7-days old seedlings, staining strongly the root system, mainly in the meristematic zone of root tips, within the central cylinder and root primordia. Signal in the shoot was weaker again, when compared to the root, but corresponding to the 7-Day old 1 hour-stained seedlings, being consistent in each respective independent line. After 3 hours of staining, the root system almost saturated in most lines. All lines showed some signal developed in the shoot, most commonly the cotyledon or true leaf tip being blue, together with leaf vascularity (Figure 15). Again, several lines developed a signal in mesenchymal cells in cotyledons or within the young leaves. Blue-coloured trichomes were present occasionally, but when coloured, it was almost exclusively on young leaves (Figure 15E). 24-hour staining only confirmed the previously observed expression. Root system of all lines was overstained, while staining of cotyledons and young leaves varied between lines (Figure 15A-C), from one extreme staining (shoot fully blue) to partially stained leaves or only patches of cells stained (mostly the vascular bundles). Although the 14-day old seedling staining has not revealed any novel tissue expressing the native promoter expression cassette, it confirmed the previous 7-day old seedlings staining and shown the variability of the expression pattern in the shoot.

GUS staining of inflorescence showed AteIF3A promoter is active during pollen development

Kanamycin resistant p3A::GFP-GUS plants of T2 generation were cultivated in soil under standard conditions until flowering. Young and richly flowering primary inflorescences were cut off, submerged in GUS staining solution and incubated at 37 °C. After 1 hour, 3 hours and 24 hours, one or two inflorescences per line were removed and destained by subsequent ethanol dilutions series and finally in water before being analyzed.

Staining of 1 hours clearly revealed that the native promoter is active in the floral tissues (Figure 16). Anthers of developing pollen buds were stained blue from a very young stage. The source of signal in anthers was mostly from developing pollen and mature pollen was stained strongly as well. Whether the signal was present in sporophytic tissues of the anther was not

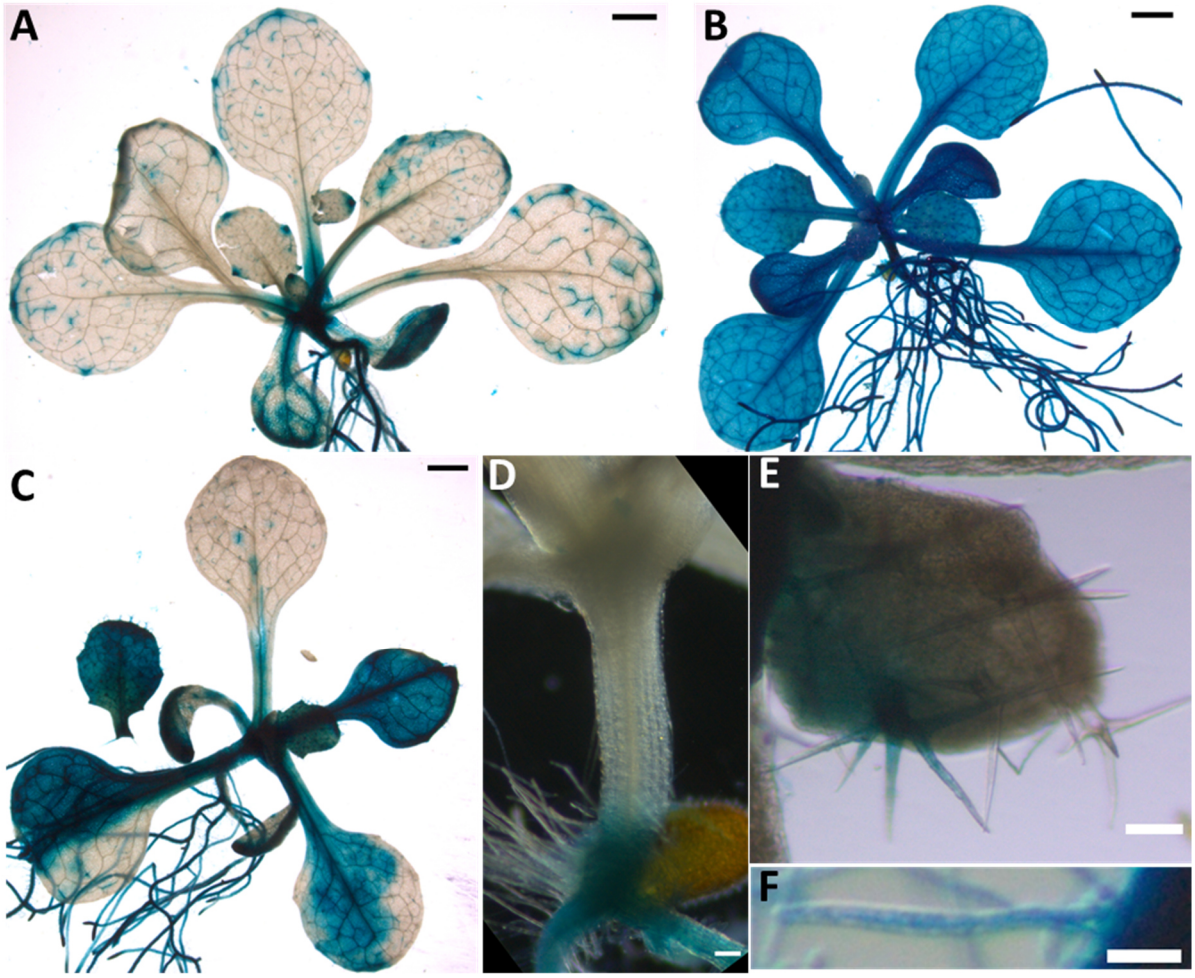

Figure 15: Shoot AteF3A expression variability in 14-day old seedlings. GUS staining of transgenic p3A::GFP- GUS 14-day old seedlings. **(A)** Weak GUS expression shoot pattern, independent line 18. **(B)** Strongest GUS expression shoot pattern, independent line 5. **(C)** Patchy GUS expression shoot pattern, independent line 4. **(A-C)** Scale bar = 1 mm. **(D)** Expression border in the hypocotyl of line with weak native promoter activity. **(E)** Detailed look on uneven patchy expression in young developing leaf and trichomes. **(F)** GUS signal in root hair. **(D-F)** Scale bar = 50 µm.

clear, as the blue precipitate might have non-specifically spread within the anther as well (Figure 16B). The signal was also observed on stigma and young developing siliques. Staining of 3 and 24 hours revealed the expression within the pistil of both developing and open flowers (Figure 16B-C). The signal was present in the whole pistil, more strongly staining the ovules in comparison to the surrounding tissues. Anthers and pollen were already overstained after 3 hours. Weak staining of petals and sepals was also observed. However, in petals, the signal was present mostly in the tissues that were in contact with strongly stained anthers, raising the question of the blue signal spreading from anthers, causing a false positive result. The great variability observed in shoot promoter activity was present as well, manifesting mainly in the variability of stem, sepals and petals GUS signal strength, but not that apparent in the anthers, pollen and pistil tissues. Taken together, the AteIF3A native promoter is strongly expressed

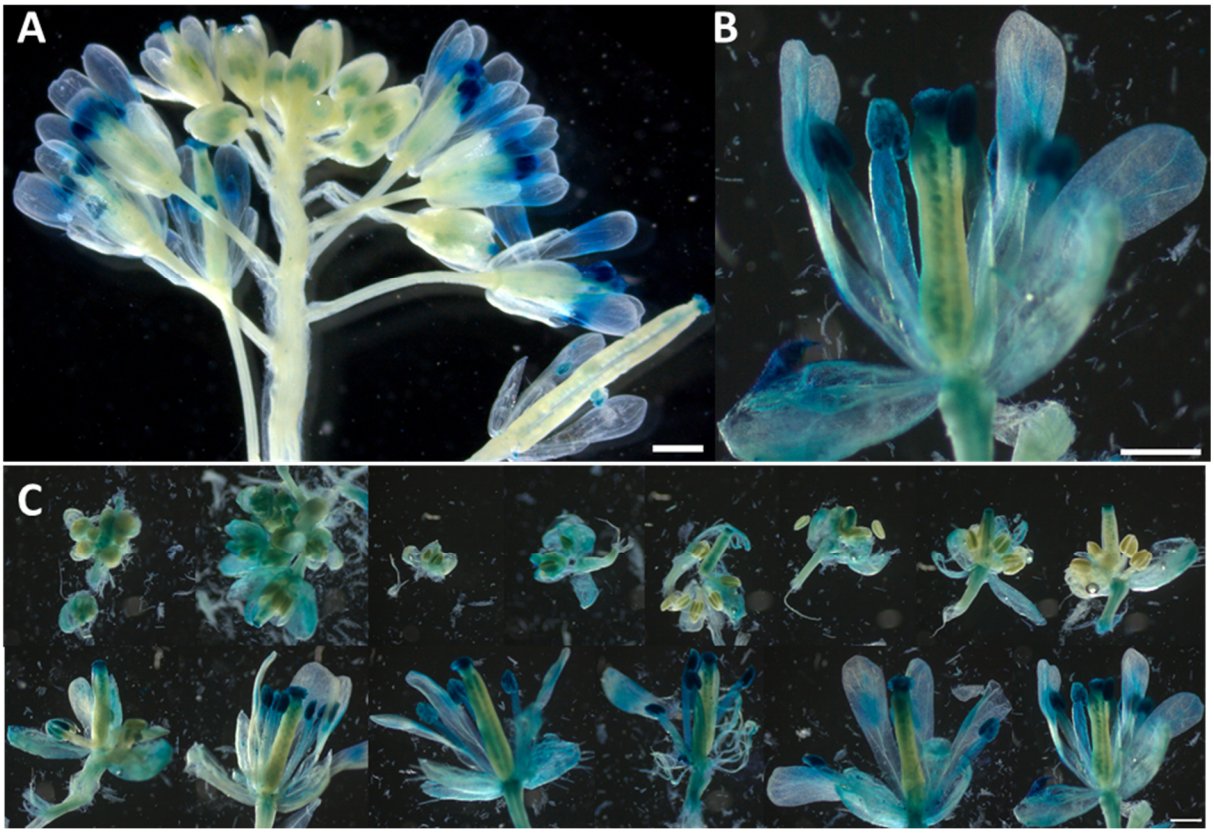

Figure 16: AteIF3A promoter activity in floral organs. GUS staining of transgenic p3A::GFP-GUS Inflorescences. **(A)** GUS expression pattern in the whole inflorescence, Scale bar = 1 mm. **(B)** Detail on expression in open flower, Scale bar = 500 µm. **(C)** Series of flower images after inflorescence dissection, starting in the upper right corner from the youngest bud to the oldest open flower in the lower left corner. Scale bar = 500 µm in the lower left corner is common for all parts of the series.

during flower development mainly in anthers and/or developing pollen, also in mature pollen and female tissues within the flower, especially the ovules.

Fluorescent microscopy of the p3A::GFP-GUS shows patchy pattern in the roots

Seedlings cultivated *in vitro* were examined at the age of 5 to 8 days. Seedlings were stained in ½ MS with 10 µg/mL propidium iodide dye for 5-15 minutes in order to gain better orientation in the root tissue by visualizing the cell walls. The GFP signal was clearly observable and located in the cytoplasm of root cells, being strongest in the meristematic zone, central cylinder and in root primordia (Figure 17), which was consistent with the observation of the GUS staining analysis. However, the expression pattern in the root tip was highly variable not only between independent lines, but also between individual plants and even between different root tips of one seedling (Figure 17A-B). The expression in some root meristems was restricted only to one or a few cells in the division zone and subsequently in their daughter cells, leading to a strip of cells where the promoter is active (Figure 17A-B). In the elongation zone,

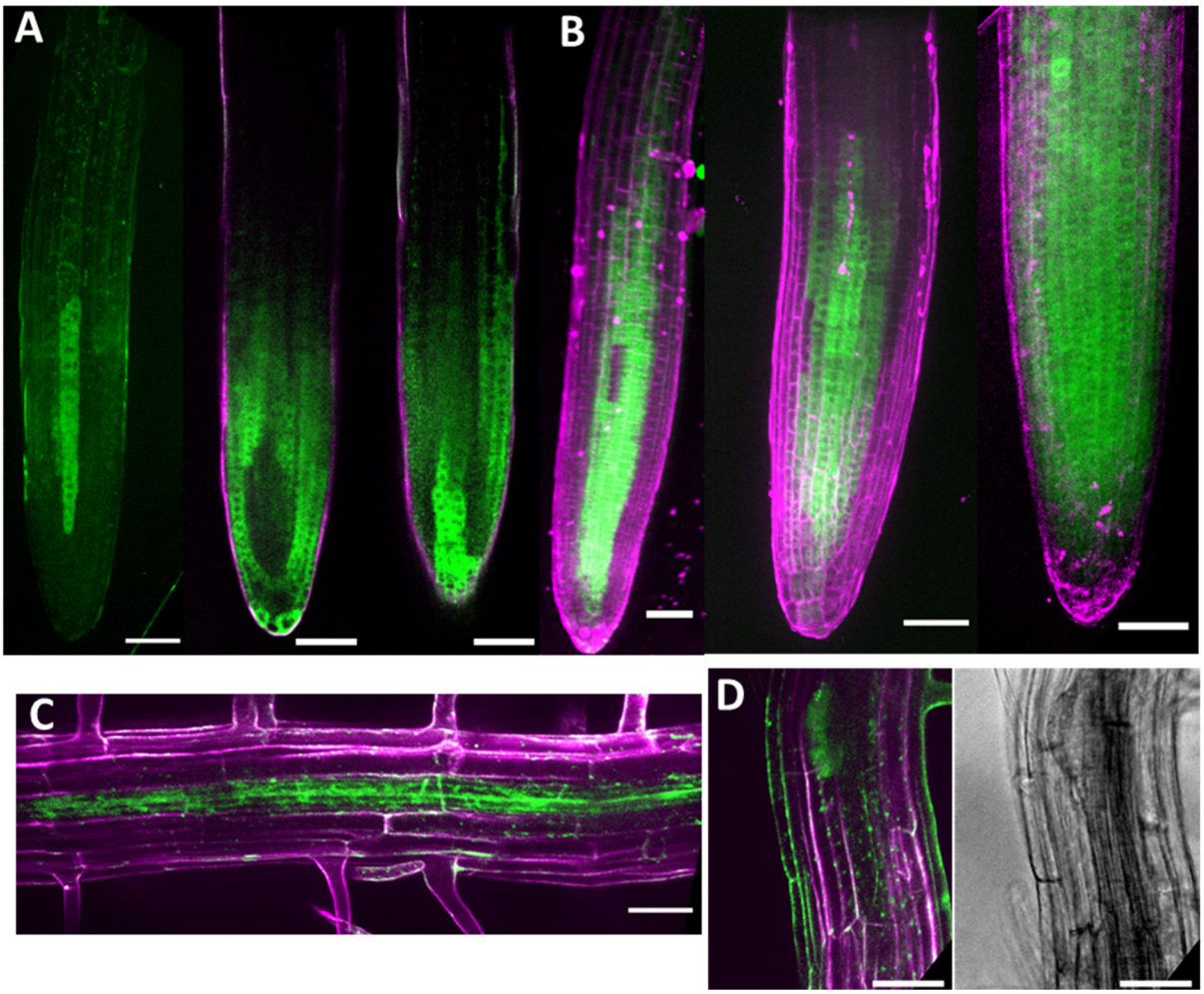

Figure 17: GFP fluorescent signal in p3A::GFP-GUS root system. GFP fluorescence signal (green), cell walls stained by propidium iodide (magenta) **(A)** Series of single confocal section of the GFP expression in the root meristem. Variable expression is shown, as only one file of cells is expressing GFP-GUS (left) to more cell files expressing GFP expression pattern (middle and right). **(B)** Series of Z-projections (max intensity) of the GFP-GUS expression in the root meristem. Variable patchy expression is shown, as more cell files expressing GFP are shown (left and middle) and whole meristem expression of GFP (right). **(C)** Detail on expression in mature zone of the root with GFP-GUS expression in central cylinder. (**D)** GFP-GUS expression in forming young and older root primordia emerging from pericycle. Scale bar = 50 µm is common for all parts of the figure.

as cells became larger, the signal fainted, suggesting that the promoter is active mostly in the division zone, whereas in the daughter cells, the expression stops but the GFP-GUS was still present in cytoplasm. In other cases, the expression was observed in all the cells of the division zone and then fainting again in the elongation zone (Figure 17B). The signal was also observed in the central cylinder of the mature zone (Figure 17C) and in emerging primordia of the root (Figure 17D). Unfortunately, the propidium iodide, in our conditions, insufficiently stained the young cell walls of the primordia and sometimes failed to properly stain even the root tip. Observation of the GFP signal in the shoot was almost non-distinguishable (not shown) on the Nikon spinning disc and will need further investigation using more sensitive confocal setup (preferably the scanning microscope with Airyscan detector). Taken together, the GFP

observation clearly follows the GUS expression pattern in the sporophyte and reveals a variable patchy pattern in the promoter activity, which might also explain the GUS signal variability and will be discussed in its respective chapter in the Discussion.

GFP-GUS is sequestered into small or large foci in mature pollen

Observation of the GFP signal in mature pollen brought another interesting feature for the GFP-GUS fusion protein. The GFP signal was clearly observable in mature pollen, which was in acceptance with the GUS staining analysis of floral tissues (Figure 18A-B). Interestingly, the signal was concentrated to hundreds or thousands small foci in the vegetative cell of the mature pollen (Figure 18A-B). It is most probable that the AtEIF3A native promoter is highly active during the development, so the GFP-GUS is either forming large aggregates on its own and/or is so abundant that the vegetative cell is actively sequestering the GFP-GUS into the foci observed. Similar phenomena was present in the control line Lat52::GFP-GUS, where the GFP-GUS expression lead to similar small foci or, in some cases, into one large aggregate and lower

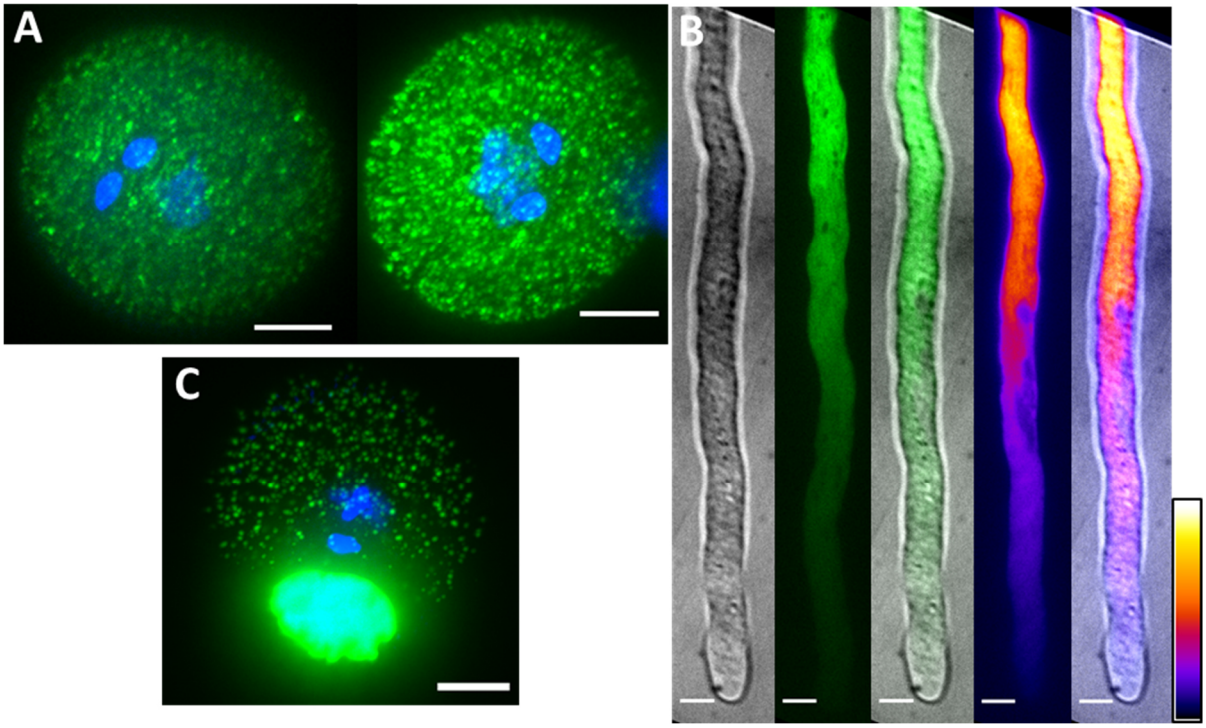

Figure 18: GFP fluorescent signal in p3A::GFP-GUS pollen and pollen tubes. GFP fluorescence signal (green) or coloured as Fire LUT option in ImageJ, DAPI staining of vegetative nucleus and sperm cells (blue) **(A)** Series of Z-projections (max intensity) of lower amount of sections (left) and higher amount of sections (right) of the GFP-GUS expression and foci localization in mature pollen. **(B)** Series of single section images of GFP-GUS expression in a pollen tube. From left: DIC+ image of the pollen tube; green channel image showing GFP signal; merge of first two images; GFP signal intensity coloured in Fire; merge of the first image and the fourth. **(C)** Lat52::GFP-GUS expression in pollen shown as a control for GFP-GUS localization. Scale bar = 5 µm is common for all parts of the figure. Intensity scale bar is shown in the lower left corner of **(B).**

number of small foci (Figure 18C), suggesting this could really be a specific feature of the GFP-GUS fusion protein. Expression was localized predominantly in the vegetative cell. If there is some expression in the sperm cells was not clearly observable, as the foci were so abundant and strong in signal that they were creating high level of noise outside their confocal section. However, the over-expression of the GFP-GUS did not reduced the pollen ability to germinate into growing pollen tube, where the signal was strongest in the region more distant from the pollen tube tip. Towards the tip, the signal is fainting. However, the small foci were not that apparent in the growing pollen tube after 4 hours of incubation on pollen germination media.

GUS staining and fluorescent microscopy of Col-0 control plants showed no detection of signal described in p3A::GFP-GUS

Col-0 wild type plants were included in every GUS staining analysis and during all confocal microscopy sessions as internal control for possible non-specific GUS staining or false positive fluorescence signal (Figure 19). Col-0 seedlings were treated identically to the transgenic p3A::GFP-GUS lines. After 24 hours of GUS staining and subsequent destaining in ethanol dilution series, no blue precipitate formed in the tissues of the Col-0 seedling (Figure 19A, 19E). Identically treated inflorescences of Col-0 plants showed no blue staining of any tissue (Figure 19B-C). Fluorescent microscopy of roots and pollen grains (identical acquisition setup)

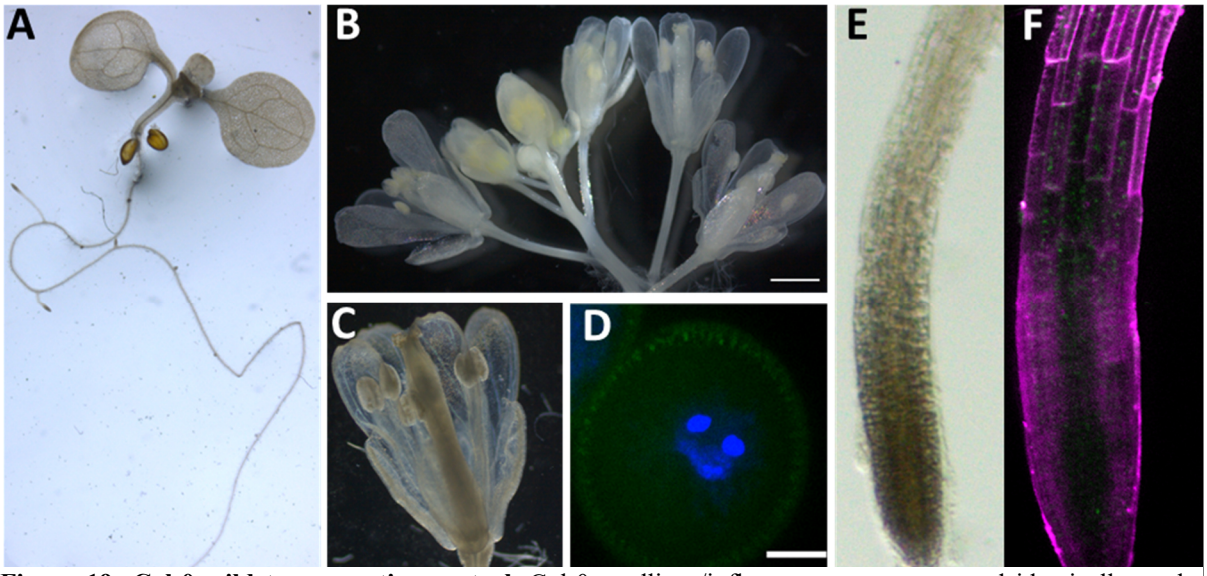

Figure 19: Col-0 wild type negative control. Col-0 seedlings/inflorescences were treated identically to the p3A::GFP-GUS lines, fluorescent images were processed in the same way. **(A)** Col-0 7-day old seedling after 24 hours of GUS staining, scale bar = 1 mm. **(B)** Col-0 inflorescence after 4 hours of GUS staining, Scale bar = 1mm. **(C)** Col-0 Inflorescence after 4 hours of GUS staining, Scale bar = 500 µm. **(D)** Max intensity Z-projection covering the MGU of mature pollen grain, green channel and DAPI staining (blue) Scale bar = 5µm. **(E)** Col-0 root meristem after 24 hours of GUS staining Scale bar = 50 µm. **(F)** Max intensity Z-projection of green channel in Col-0 root meristem stained in propidium iodide (magenta). Scale bar = 50 µm.

showed no or very little background noise in the green channel (Figure 19D, 19F). Altogether, identically treated control plants showed no signal that we previously described in p3A::GFP-GUS transgenic lines, confirming that the signal observed in p3A::GFP-GUS was truly observation of the promoter activity.

AteIF3A protein subcellular localization

AteIF3A fusions with GFP driven by native, 35S sporophytic and Lat52 pollen specific promoters were used for subcellular localization study

For studying the subcellular localization of AteIF3A, we created expression cassettes expressing AteIF3A fused with GFP either on N-terminal or C-terminal end driven by different promoters. The native promoter activity was described in the previous chapter. Despite the troubles with the variability of the p3A promoter expression pattern, we showed that the p3A is expressed strongly in root meristem, root primordia and floral tissues. Expression cassettes p3A::GFP-AteIF3A and p3A::AteIF3A-GFP were investigated in this analysis. We also decided to overexpress the AteIF3A. For that purpose, we used strong promoters for the sporophyte and the male gametophyte. For overexpression analyses in the sporophyte, we created expression cassettes driven by 35Sdouble promoter, 35Sdouble::eIF3A-GFP and 35Sdouble::mOrange-eIF3A. For overexpression in the male gametophyte, we created expression cassettes driven by Lat52 promoter, Lat52::eIF3A-GFP and Lat52::GFP-eIF3A. Seeds from at least 12 independent in the T1 generation were harvested and four lines segregating in approx. ratio of 3:1 (RFPpositive:RFPnegative seeds) were grown and investigated further. Observation of the GFP signal in the p3A::GFP-GUS was done using Nikon laser spinning disc. Observation of the fluorescence in the shoot epidermal cells (not shown) was insufficient at the Nikon spinning disc confocal and will be investigated in future experiments using the more sensitive Zeiss LSM880 scanning confocal with Airyscan detector.

GFP-AteIF3A localized to the cytoplasm, surrounded the vegetative cell of mature pollen and forms foci around the nucleus in the root

Transgenic p3A::GFP-AteIF3A seedlings of 5-8 days were investigated by fluorescent microscopy for the GFP fluorescence. The GFP signal was clearly observable in the root system, giving similar pattern variability as the p3A::GFP-GUS expression analysis (F. Again, the construct was either expressed in only a few cell lines of the cell division zone of the root

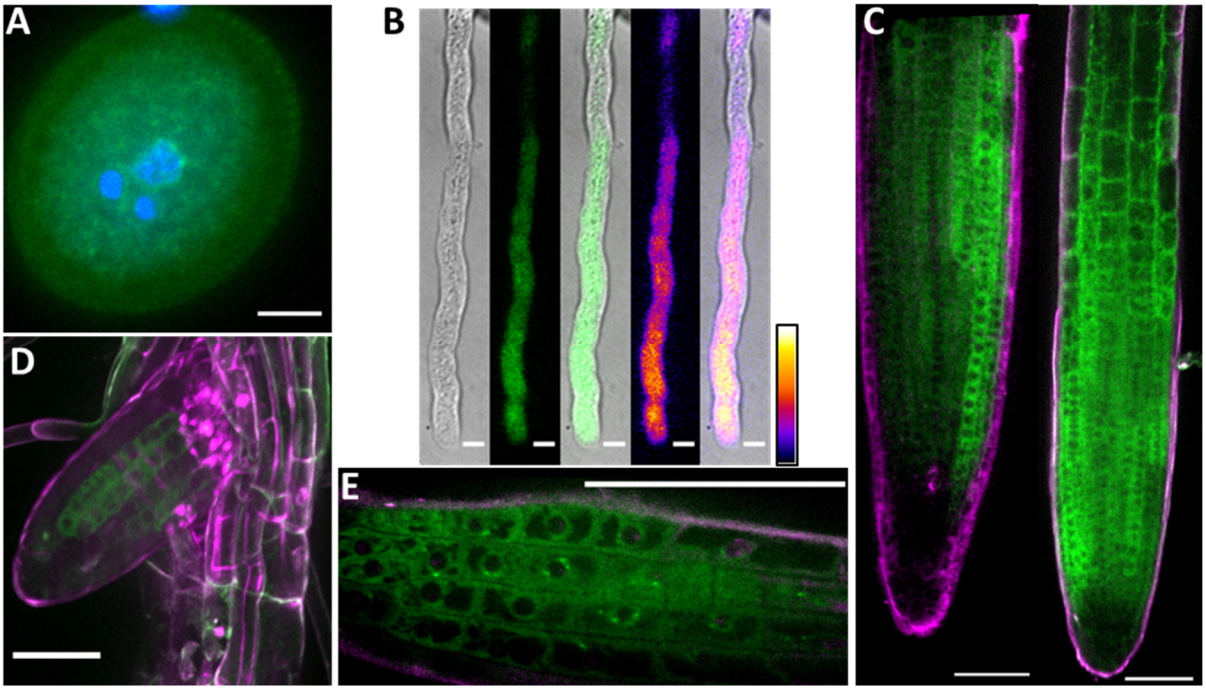

Figure 20: Expression and localization of the GFP-AtEIF3A fusion driven by native promoter. (A) max intensity Z-projection covering the MGU of mature pollen grain series, GFP-EIF3A localization in the vegetative cell cytoplasm (green), stained in DAPI (blue). **(B)** Single sections of GFP-EIF3A expression in the growing pollen tube and its localization in the shank region of the tube. From left: DIC+ image of the pollen tube; green channel image showing GFP signal; merge of first two images; GFP signal intensity colored in Fire LUT; merge of the first image and the fourth. **(C)** Single series of the GFP-EIF3A expression in the root tip (left), max intensity Z-projections of 2 series of the expresion in the root tip (right), stained with propidium iodide (magenta). **(D)** Max intensity Z-projection of the patchy GFP-EIF3A expression in the growing lateral root, stained with propidium iodide (magenta). **(E)** Detail on subcellular localization of GFP-EIF3A in the division zone of the root meristem, stained with propidium iodide (magenta). Scale bar = 5 µm is common for figures **(A)** and **(B)**. Scale bar = 50 µm is common for figures **(C-E)**. Intensity scale bar for the Fire LUT is shown in the lower left corner of **(B)**.

meristem (Figure 20C-left section) or in the whole cell division zone, slowly fading in the elongation zone (Figure 20C-right section). Uneven expression was also observed in developing lateral roots (Figure 20D) Interestingly, when compared to the p3A::GFP-GUS we observed more roots with promoter activity in the whole meristematic region, in the independent lines inspected. The GFP-eIF3A fusion protein was clearly localized in the cytoplasm of the root cells, best visualized in the meristematic zone of the root, where vacuoles were small and cytoplasm filled a large portion of the cell (Figure 20E), whereas it was clearly not present in the nucleus (round black circles) or in the vacuole. In the transition to the elongation zone, the GFP signal was moved more towards the subcortical region of the cell, as the vacuole enlarged and started to fill more space in the cell (Figure 20C-right section). Moreover, one or a few foci of the GFP fluorescence per cell were detected localized around the nucleus but not around the vacuoles (Figure 20E). GFP-AteIF3A localization in the cytoplasm was observed in the mature pollen, where higher GFP fluorescence was detected around the vegetative nucleus (Figure 20A). In cytoplasm, the GFP-AteIF3A was not evenly

distributed and resembled cloudy distribution. No fluorescence was apparent in the sperm cells on single section images, but the GFP-AteIF3A was more concentrated around the sperm cells (Figure 20A). In comparison with the p3A::GFP-GUS, the foci/aggregates were not present in the p3A::GFP-eIF3A and the overall GFP fluorescence was much weaker. In pollen tubes, the GFP-eIF3A protein was localized in the vegetative cell cytoplasm, where the fluorescence was stronger in the zone of the pollen tube tip while not present in the apical region (Figure 20B). However, better imaging would be needed in the effort to focus more on the subcellular localization, as the pollen cultivation on agarose pads was causing insufficient imaging to make clearer observations of the whole pollen tubes.

AtEIF3A C-terminal fusion with GFP localization was similar to the N-terminal fusion but was expressed evenly in the full root meristem of all lines

Transgenic p3A::AteIF3A-GFP seedlings of 5-8 days were investigated by fluorescent microscopy for the GFP fluorescence with similar but not identical results as described previously with the N-terminal GFP fusion (Figure 21). The GFP signal was clearly observable in the root cell cytoplasm, surrounding the cell nucleus and forming vacuoles in the young dividing cells as well as in the cells of elongation zone (Figure 21E). AteIF3A-GFP was also present in young emerging root primordia (Figure 21D) Most strikingly, all investigated p3A::eIF3A-GFP independent lines showed promoter activity in all cell files of the root meristem division zone (Figure 21C). The zone around the quiescent center was more depleted of the GFP signal (Figure 21C-left), which was common for all lines expressing the native promoter but not so apparent with the previously described lines showing variable expression. Also, the small foci surrounding the nucleus were not apparently present in the p3A::AteIF3A-GFP lines investigated in this analysis. In mature pollen, similar localization was observed. The AteIF3A-GFP protein was present in the cytoplasm of the vegetative cell, around the vegetative nucleus and sperm cells, but not in the nucleus of the vegetative cell or inside the sperm cells. Moreover, the cloudy distribution of AteIF3A-GFP signal in cytoplasm was more apparent in the case of C-terminal fusion. In pollen tubes, AteIF3A-GFP localized to the cytoplasm of the pollen tube (Figure 21B). Stronger fluorescence signal was observed and localized in the shank region of the pollen tube tip but not in the apical region.

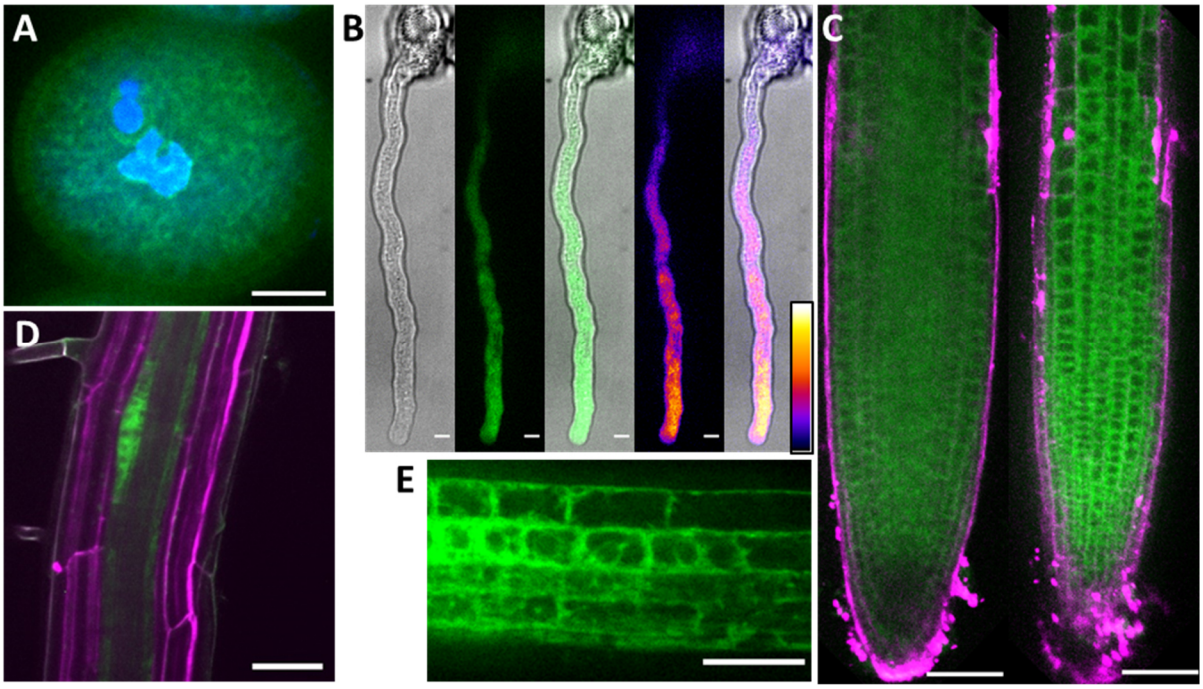

Figure 21: Expression and localization of the AtEIF3A-GFP fusion driven by native promoter. (A) Max intensity Z-projection covering the MGU of mature pollen grain series, EIF3A-GFP localization in the vegetative cell cytoplasm (green), stained in DAPI (blue). **(B)** Single sections of EIF3A-GFP expression in the growing pollen tube. From left: DIC+ image of the pollen tube; green channel image showing GFP signal; merge of first two images; GFP signal intensity colored in Fire LUT; merge of the first image and the fourth. **(C)** Single series of the evenly distributed EIF3A-GFP expression in the root tip (left), max intensity Z-projection in the root tip (right), stained with propidium iodide (magenta). **(D)** max intensity Z-projection of the EIF3A-GFP expression in the emerging lateral root, stained with propidium iodide (magenta). **(E)** Detail on subcellular localization of GFP-eIF3A in the division zone of the root meristem. Scale bar = 5 µm is common for figures **(A)** and **(B)**. Scale bar = 50 µm is common for figures **(C-E)**. Intensity scale bar for the Fire LUT is shown in the lower left corner of **(B)**.

Identical localization and similar protein abundance was observed in mature pollen overexpressing eIF3A fusions with GFP under Lat52 promoter

Transgenic lines carrying Lat52::GFP-eIF3A and Lat52::eIF3A-GFP expression cassettes were investigated by fluorescent microscopy for the GFP fluorescence in mature pollen grains and pollen tubes. The signal of both fusions had similar cloudy distribution in cytoplasm and clearly surrounded the vegetative nucleus and sperm cells (Figure 22A-B). Interestingly, the overall fluorescence signal of both N-terminal and C-terminal fusions was only slightly increased in comparison with the native p3A promoter. In pollen tubes, both eIF3A fusions localized to the cytoplasm of the pollen tube, being stronger in the shank region of the pollen tube tip but not in the apical region (Figure 22C-D). Interestingly, the Fire visualization (ImageJ Fire LUTs) showed partial localization to fibrous network that was more apparent in the Lat52::eIF3A-GFP pollen tube (Figure 22D). Such organisation is typical for F-actin filaments in the pollen tube tip. However, better imaging and co-localization experiments are necessary to confirm this observation.

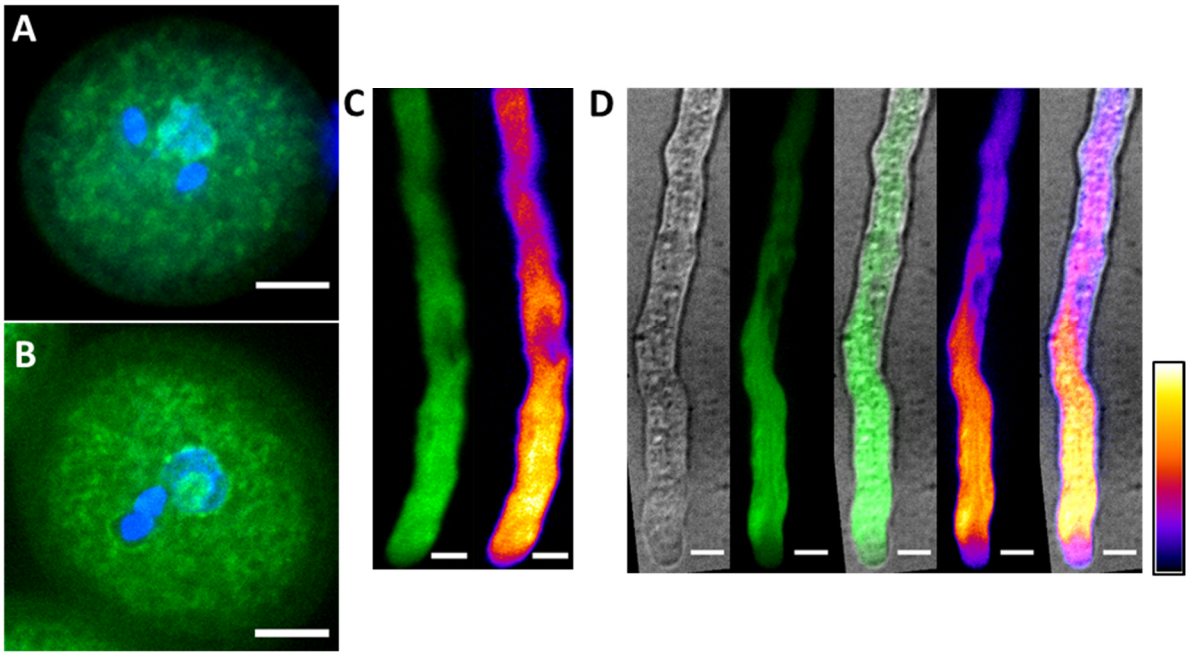

Figure 22: Expression and localization of the AtEIF3A GFP fusions driven by Lat52 pollen specific promoter. (A) Max intensity Z-projection covering the MGU of the Lat52::GFP-eIF3A mature pollen grain series, GFP-EIF3A localization in the vegetative cell cytoplasm (green), stained in DAPI (blue). **(B)** Max intensity Z-projection covering the MGU of the Lat52::eIF3A-GFP mature pollen grain series, eIF3A-GFP localization in the vegetative cell cytoplasm (green), stained in DAPI (blue). **(C)** Single sections of GFP-eIF3A driven by Lat52 in the growing pollen tube. From left: DIC+ image of the pollen tube; green channel image showing GFP signal; merge of first two images; GFP signal intensity colored in Fire LUT; merge of the first image and the fourth. **(D) s**ingle sections of eIF3A-GFP driven by Lat52 in the growing pollen tube. Green channel image showing GFP signal in green (left) and GFP signal intensity colored in Fire LUT (right). Scale bar = 5 µm is common for all figures. Intensity scale bar for the Fire LUT is shown in the lower left corner of **(D)**.

Observation of sporophytic overexpression of AteIF3A protein fusions driven by 35Sdouble promoter was not possible due to germination and seedlings growth arrest

Fluorescent signal in transgenic lines carrying 35Sdouble::AteIF3A-GFP and 35Sdouble::mOrange-AteIF3A was not observed due to impaired possibility to cultivate these lines. We failed to cultivate the KanFAST-positive transgenic seeds of both lines *in vitro* on ½ MS media with or without kanamycin selection and with/without supplement of additional sucrose (data not shown). Cultivation of the seeds placed directly in soil failed as well. Very little seeds of the 35Sdouble::AteIF3A-GFP germinated and produced a seedling. However, in those seedlings that managed to do so, no fluorescent signal was observed in the root and shoot tissues (data not shown). All seeds of the 35Sdouble::mOrange-AteIF3A either failed to germinate or produced a small root and then arrested the growth after approximately 3 days after sowing (data not shown).

Identification of the protein products revealed the presence of full AteIF3A-GFP and GFP-AteIF3A fusion proteins with possible products of degradation

Proteins were extracted from collected inflorescences of p3A::GFP-AteIF3A, p3A::AteIF3A-GFP, Lat52::eIF3A-GFP and Lat52::GFP-eIF3A transgenic lines, using the DEM native extraction protocol. Protein samples were then mixed with appropriate volume of Sample buffer, sonicated and loaded on polyacrylamide gel in two replicas, once on experimental gel and once on loading control gel. After SDS-PAGE separation, experimental gel was used for western blotting and the loading control gel was stained with Coomassie brilliant blue dye right away (Figure 23A). As all samples contained AteIF3A fused to GFP, a protein product of 141 kDa was expected as the full recombinant protein in all samples. The rabbit anti-GFP primary antibody and anti-rabbit secondary antibody fused to alkaline phosphatase were used for signal development. As internal positive control, protein sample of Lat52::YFP transgenic line (provided by Christos Michailidis) was included in the analysis with expected molecular weight of 27 kDa. Although the Lat52::YFP sample tissue was not identical (open flowers), this positive control was still suitable in this experiment, as the main purpose was only to confirm the presence of full protein. Moreover, the YFP and GFP share same molecular weight and are recognized by the same anti-GFP antibody. After signal detection, a band of apparent molecular weight higher than the 130 kDa marker pointed on the presence of

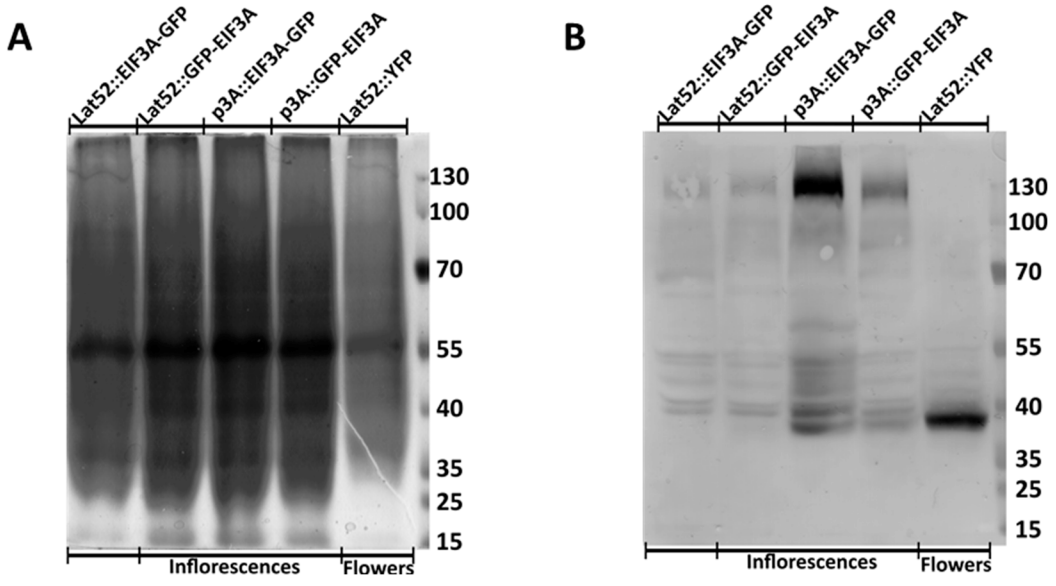

Figure 23: Western blot analysis of various AteIF3A fusions with GFP. (A) SDS-PAGE gel loading control, stained in Coomassie briliant blue after separation. Lat52::YFP positive control is included. **(B)** Western blot membranes after signal development. Molecular weights of the protein marker are indicated.

full 141 kDa protein fusions in all experimental samples and was not present in the control sample. The control sample contained one band with apparent molecular weight between 35 and 40 kDa, which was unexpectedly higher than YFP alone should be. Whether it was due to some biological cause or due to uneven polymerization of the polyacrylamide gel that would change the relative migration was not clear. However, several weaker bands were detected in the experimental samples. Though, it is possible that AteIF3A GFP fusions could have been partially degraded either in the inflorescence tissue or during the isolation procedure, which was supported by the fact that one degradation band had similar apparent molecular weight as the free YFP, suggesting that degraded AteIF3A leaves some fraction of the free GFP in the tissue/sample. Although the same amount of sample tissue was processed for every sample and treated in a similar way so the protein concentration was similar, an apparent difference in signal strength was observed between the samples. That could be explained by different activity of the promoters. Nevertheless, this experiment was not planned to include quantification analysis and altogether confirmed that the transgenic lines contained the correct protein product.

Arabidopsis *AteIF3A* mutant characterization

Two investigated SAIL T-DNA insertion lines were mapped outside the AtEIF3A gene

First line inspected as a possible *AtEIF3A* gene mutant was the SAIL_124_G07.v1 (SAIL 124). TAIR database identified the flanking region of SAIL 124 insertion in the very end of the 3′ UTR of the *AtEIF3A* gene (AGI code: At4g11420), very close to the beginning of neighbouring gene (At4g11410). The T-DNA insertion in SAIL 124 contains a BASTA resistance cassette in the *qrt* mutant background of Columbia-0 ecotype. ½ MS plates containing sucrose and BASTA selection were used to germinate the stock seeds obtained from NASC. However, no seed managed to germinate on this media. Another set of seeds was therefore sown on ½ MS plates without selection. On this plate only one seed managed to germinate and grow into wild type-like seedling (data not shown). After transfer to soil, this one seedling produced a normal looking plant indistinguishable from wild type Col-0 plants that were cultivated alongside at same conditions. This single SAIL 124 plant was not analysed. Progeny of this plant was tested for BASTA resistance. This time, all seeds germinated and produced wild type-like seedlings and plants. This generation was tested in further experiments. First, plants were genotyped using insertion specific SAIL LB primer and genome specific SAILSEQ_545_BO8.4 RP and SAIL_124_GO7.V1 LP primers in a three primer genotyping reaction. All SAIL 124 plants investigated were clearly homozygous in the insertion, as the

PCR product contained only one band that was smaller than genomic band in Col-0 control plant (not shown). The insertion band was sequenced using the SAIL LB and SAIL_545_BO8.4 RP primers in order to map the insertion site at the border closer to the beginning of *AteIF3A* gene. The sequencing result was aligned to the genomic sequence of *AteIF3A*. The alignment showed that the LB site of the insertion is not inside the At4g11420 gene but in the 5′ UTR of the neighbouring gene, the At4g11410 (Figure 24A). Nevertheless, even when the insertion does not directly interfere with the inspected gene, it could have some effect on the gene expression. Therefore, basic phenotypical analysis was performed, focusing mainly on defects in the male gametophyte. SAIL 124 mature pollen morphology, anatomy and viability were tested (Figure 24F). No increase of defective pollen grains was observed as averaged 0.80 ± 0.76 percent out of total pollen counted in SAIL 124, which was not significantly different to 0.56 ± 0.42 percent observed in Col-0 control. No significant increase of non tricellular pollen was observed as averaged 3.56 ± 0.66 percent out of total pollen counted in SAIL 124, which was not different to 3.03 ± 0.72 percent observed in Col-0 control. No significant increase of non-viable pollen was observed as averaged 3.75 ± 1.87 percent out of total pollen counted in SAIL 124, which was not different to 2.86 ± 1.03 percent observed in Col-0 control. The great variability of the observations was caused mainly by the very small number of defective, non tricellular or non-viable pollen grains in the samples from which the statistics was calculated (for example in Col-0, 6 defective pollen grains per 1000 counted). Altogether, no general increase in pollen defects was observed in the SAIL 124. Moreover, pollen tube growth in the pistil was analysed using the aniline blue staining (Figure 24C). The overall pollen tube growth was not impaired in the SAIL 124 line as pollen tubes successfully grew through the tissues of transmitting tract all the way to the bottom of the pistil. Next, the seed formation was analysed (Figure 24D, 24F). Morphology of SAIL 124 siliques was no different from the Col-0 control. Defects in seeds formation averaged 2.56 ± 0.97 abnormal seeds per silique and was not different from Col-0 plants that averaged 1.75 ± 0.92 seeds per silique. Last analysis pertained with SAIL 124 was relative quantification of *AteIF3A* expression. Main inflorescences of seven individual SAIL 124 plants were used as sample material for RT-qPCR analysis (3 biological replicas, each in 3 technical replicas). Two primer sets were used for AteIF3A mRNA amplification, targeting a CDS region in 5′ (primer set 2) and 3′ (primer set 4). A small decrease in *AteIF3A* expression was observed as the levels of AteIF3A mRNA averaged 30.29 ± 3.96 (primer set 2) and 50.21 ± 7.00 (primer set 4) percent of ACT8 expression when compared to averaged 38.31 ± 4.27 (primer set 2) and 70.00 ± 11.78 (primer set 4) percent observed in Col-0 control. However, great variability was observed in these measurements, which

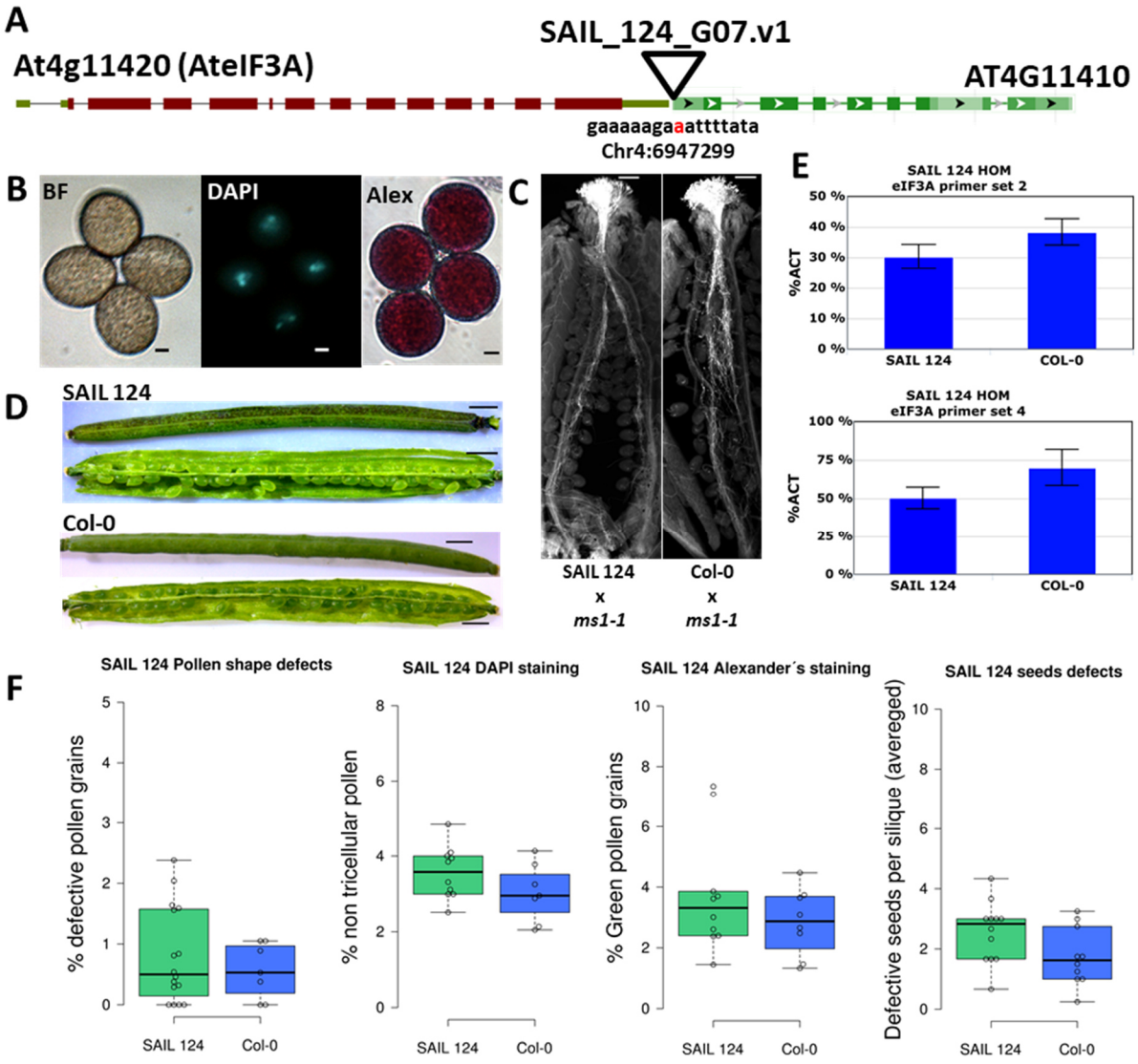

Figure 24: SAIL 124 analysis. (A) Schematic illustration of SAIL_124.G07.v1 T-DNA insertion mapped to the *Arabidopsis* genomic sequence. The detail of first wild type nucleotide affected by the insertion is showed under the insertion site. (B) Images of SAIL 124 pollen morphology (BF), anatomy DAPI, and viability (Alexander´s staining). Scale bar = 5 µm. (C) Aniline blue staining of pollen tubes within the pistil. Col-0 results is given as control. Scale bar = 200 µm. (D) Siliques before and after dissection. Col-0 silique is shown for comparison. Scale bar = 1 mm. (E) RT-qPCR results of the *AteIF3A* relative expression in the SAIL 124 samples. (F) Series of graphs showing the observations in the SAIL 124 and Col-0 control.

is commonly known for this method. Moreover, great inconsistence in *AteIF3A* expression was observed even between individual samples of Col-0 control, therefore this analysis was not statistically tested and needs to be repeated after reproducible measurements will be reached at least between the wild type samples. Altogether, the insertion in the SAIL_124_G07 line was mapped outside the inspected gene and no phenotypical differences in the SAIL 124 homozygous plants were observed.

Next T-DNA insertion line potentially located in the *AteIF3A* gene was the SAIL_227_B02 (SAIL 227). TAIR database identified the flanking region of SAIL 227 insertion in the At4g11420 3′ UTR. The T-DNA insertion in SAIL 227 contains a BASTA resistance cassette in the *qrt* mutant background of Columbia-0 ecotype. ½ MS plates containing sucrose and BASTA selection were used to germinate the stock seeds obtained from NASC. All seeds germinated and produced BASTA resistant seedlings. Plants transferred to soil were genotyped using insertion specific SAIL LB primer and genome specific SAILSEQ_227:B02.0 RP and SAILSEQ_227:B02.0 LP primers in a three primer genotyping reaction. Genotyping revealed that a mix of heterozygous and homozygous plants was present in the seed stock, as all PCR reactions of analysed plants had one band smaller than the genomic band present in Col-0 plants and some had both the genomic and the insertion band. The insertion band was sequenced using the SAIL LB and SAILSEQ_227:B02.0 RP primers in order to map the insertion site at the border closer to the beginning of *AteIF3A* gene. The sequencing result was aligned to the genomic sequence of *AteIF3A*. The alignment showed that the LB site of the insertion is not inside the *AteIF3A* gene but in the 5′ UTR of the neighbouring gene, the At4g11410 (Figure 25A). Identically to the SAIL 124, phenotypical analysis was performed with similar results (Figure 25F). No significant increase of defective pollen grains was observed as averaged 0.45 ± 0.43 percent out of total pollen counted, no significant increase of non tricellular pollen was observed as averaged 2.11 ± 0.71 percent out of total pollen counted and no significant increase of non-viable pollen was observed as averaged 3.37 ± 1.05 percent out of total pollen counted in SAIL 227. Col-0 control statistics are given earlier in this subchapter. Pollen tube growth in the pistil was analysed using the aniline blue staining (Figure 25C). Similarly to the SAIL 124, the aniline blue staining showed no pollen tube growth defects in the SAIL 227 line. Seed formation was analysed as well (Figure 25D, 25F), where the morphology of SAIL 227 siliques was no different from the Col-0 control. Defects in seeds formation averaged 1.48 ± 0.72 abnormal seeds per silique, showing again no significant increase in seed defects. Main inflorescences of seven individual SAIL 227 plants were used as sample material for RT-qPCR analysis (3 biological replicas, each in 3 technical replicas) with identical primer sets

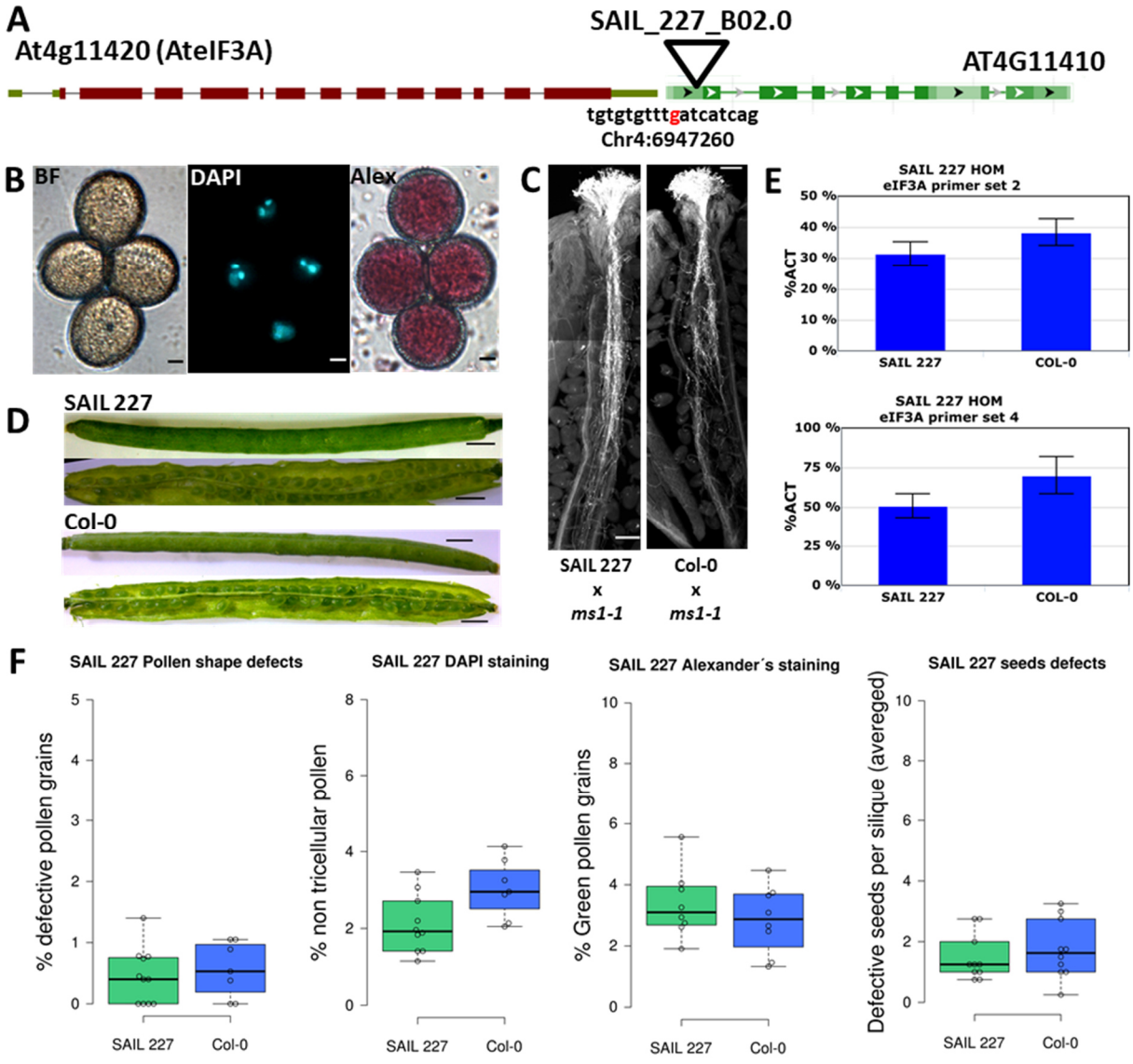

Figure 25: SAIL 227 analysis. (A) Schematic illustration of SAIL_227.B02.0 T-DNA insertion mapped to the *Arabidopsis* genomic sequence. The detail of first wild type nucleotide affected by the insertion is showed under the insertion site. **(B)** Images of SAIL 227 pollen morphology (BF), anatomy DAPI, and viability (Alexander´s staining). Scale bar = 5 µm. **(C)** Aniline blue staining of pollen tubes within the pistil. Col-0 results is given as control. Scale bar = 200 µm. **(D)** Siliques before and after dissection, Col-0 silique is shown for comparison. Scale bar = 1 mm. **(E)** RT-qPCR results of the *AteIF3A* relative expression in the SAIL 227 samples. **(F)** Series of graphs showing the observations in the SAIL 124 and Col-0 control.

as for SAIL 124. The qPCR of SAIL 124 and SAIL 227 was performed in one LightCycler480 run, therefore the Col-0 statistics (described earlier in this subchapter) are identical for SAIL 227 as well. A decrease in *AteIF3A* expression was observed as the levels of AteIF3A mRNA averaged 31.40 ± 3.90 (primer set 2) and 50.75 ± 7.63 (primer set 4) percent of ACT8 expression (Figure 25E). Statistical testing was not done due to the similar reason explained earlier in this chapter. Altogether, the insertion in the SAIL_227_G07 line was mapped outside the inspected gene and no phenotypical differences in the SAIL 227 homozygous plants were observed.

SAIL 545 insertion in AteIF3A gene does not affect plant viability

Last SAIL line inspected as a possible *AtEIF3A* gene mutant was the SAIL_545_B08 (SAIL 545). TAIR database identified the flanking region of SAIL 545 insertion in the last exon 14 of the *AteIF3A* gene. The T-DNA insertion in SAIL 545 contains a BASTA resistance cassette in Columbia-0 background. ½ MS plates containing sucrose and BASTA selection were used to germinate the stock seeds obtained from NASC. All seeds germinated and produced BASTA resistant seedlings. Plants transferred to soil were genotyped using insertion specific SAIL LB primer and genome specific SAILSEQ_545_BO8.4 RP and SAIL_124_GO7.V1 LP primers in a three primer genotyping reaction. Genotyping revealed that a mix of homozygous (SAIL 545 HOM) and wild type-like plants (SAIL 545 WT) were present in the seed stock, as PCR reactions of some plants had one band smaller than the genomic band present in Col-0 plants and some had only the genomic band (data not shown). This points to two or more insertions in the SAIL 545 line, where only some possess the insertion in the inspected gene. The insertion band was sequenced using the SAIL LB and SAIL_545_BO8.4 RP primers in order to map the insertion site at the border closer to the beginning of *AteIF3A* gene. The sequencing result was aligned to the genomic sequence of *AteIF3A.* The alignment showed that the LB site of the insertion is inside the coding sequence for the At4g11420 gene, more specifically in the exon 14 (Figure 26A). If a protein is produced from this allele, it would affect the protein sequence with changing codon encoding Alanine957 to Glycine and following protein sequence to PCQQN before stop codon is reached. Thus, the SAIL 545 HOM insertion would not affect any domain or region characterized within the eIF3A protein, only last 36 amino acids would not be present in the protein product. Nevertheless, it could affect the expression of *AteIF3A*. With SAIL 545 HOM plants, a basic phenotypical analysis was performed. The SAIL 545 WT plants were analysed as well to ensure that any phenotype observed in the SAIL 545 HOM is caused by the insertion in *AteIF3A* gene and not by any other background insertion. The statistics of

the observations were compared between SAIL 545 HOM and SAIL 545 WT plants. No significant increase of defective pollen grains was observed as averaged 0.73 ± 0.56 percent out of total pollen counted in SAIL 545 HOM plants, which was not different to 0.56 ± 0.45 percent observed in the SAIL 545 WT control. No significant increase of non tricellular pollen was observed as averaged 2.99 ± 1.25 percent out of total pollen counted in SAIL 545 HOM plants, which was not different to 3.35 ± 0.72 percent observed in the SAIL 545 WT control. No significant increase of non-viable pollen was observed as averaged 4.66 ± 1.01 percent out of total pollen counted in SAIL 545 HOM plants, which not different to 4.93 ± 1.50 percent observed in the SAIL 545 WT control (Figure 26F). For comparison with Col-0 (non-significant in all observations), please revisit the respective statistics given earlier in the previous subchapter. Altogether, no general increase in pollen defects was observed in the SAIL 545 HOM plants. Then, pollen tube growth in the pistil was analysed using the aniline blue staining (Figure 26D). The overall pollen tube growth was not impaired in the SAIL 545 HOM and SAIL 545 WT plants pollen tubes successfully grew through the tissues of transmitting tract to the bottom of the pistil. The seed formation was analysed as well (Figure 26C, 26F). Morphology of SAIL 545 HOM siliques was no different from either the SAIL 545 WT or the Col-0 control. Defects in seeds formation averaged 2.58 ± 2.29 and 1.83 ± 1.12 abnormal seeds per silique in SAIL 545 HOM and SAIL 545 WT, respectively. Both observations were not different from Col-0 plants that averaged 1.75 ± 0.92 seeds per silique. At last, analysis of relative quantification of *AteIF3A* expression was performed in SAIL 545 HOM and SAIL 545 WT plants. Main inflorescences of 7 individual SAIL 545 HOM or SAIL 545 WT plants were used as sample material for RT-qPCR analysis with identical primer sets as for the other two SAIL lines. No decrease in *AteIF3A* expression was observed between the SAIL 545 HOM and SAIL 545 WT samples as the levels of AteIF3A mRNA averaged 20.38 ± 1.69 (primer set 2) and 40.59 ± 4.14 (primer set 4) percent of ACT8 expression in SAIL 545 HOM when compared to averaged 22.15 ± 5.58 (primer set 2) and 53.18 ± 14.24 (primer set 4) percent observed in SAIL 545 WT control. Similarly to the SAIL 124 and SAIL 227, a decreased expression was observed when compared to averaged 34.58 ± 5.94 (primer set 2) and 77.22 ± 18.26 (primer set 4) percent observed in Col-0 control. In this case, the variability of relative expression between the biological replicas was striking and only increased the doubts about the reliability of this measurements. However, even though the difference between SAIL 545 HOM and Col-0 samples was apparent (again not statistically tested), only small differences were observed between SAIL 545 HOM and SAIL 545 WT.

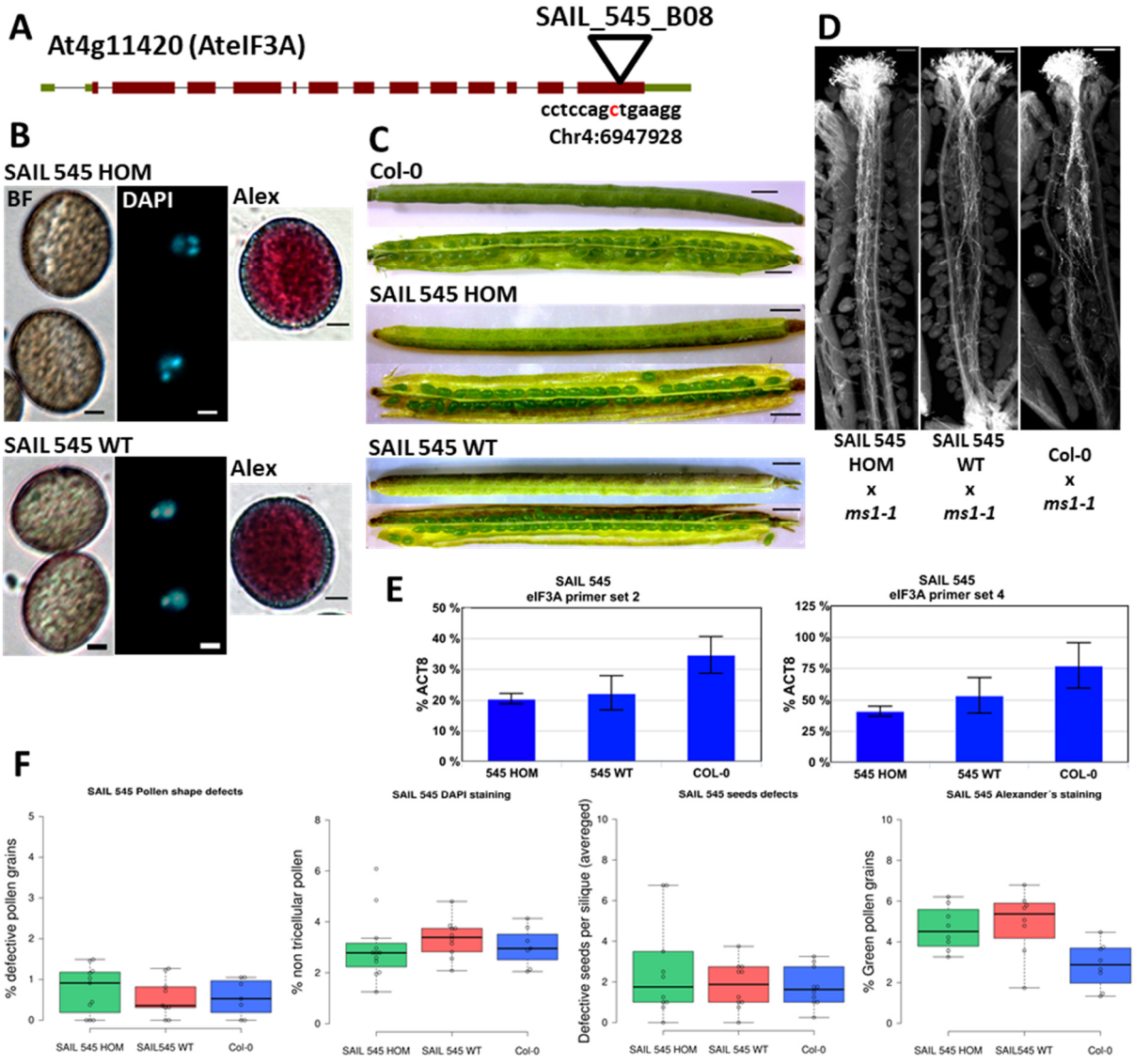

Figure 26: SAIL 545 analysis. (A) Schematic illustration of SAIL_545.B08 T-DNA insertion mapped to the *Arabidopsis* genomic sequence. The detail of first wild type nucleotide affected by the insertion is showed under the insertion site. **(B)** Images of SAIL 545 HOM and WT pollen morphology (BF), anatomy DAPI, and viability (Alexander´s staining). Scale bar = 5 µm. **(C)** Aniline blue staining of pollen tubes within the pistil. Col-0 results is given as control. Scale bar = 200 µm. **(D)** Siliques before and after dissection.. Scale bar = 1 mm. **(E)** RT-qPCR results of the *AteIF3A* relative expression in the SAIL 545 HOM and WT samples. **(F)** Series of graphs showing the observations in the SAIL 545 HOM and WT.

Altogether, the insertion in the SAIL_545_G07 homozygous line was mapped at the very end of the coding sequence of the inspected gene and no phenotypical differences in the SAIL 545 homozygous plants were observed.

SALK 762 mapped to AteIF3A gene and plants show pollen/ovule abortion

Last T-DNA insertion line inspected as a possible *AtEIF3A* gene mutant was the SALK_123762 (SALK 762). TAIR database identified the flanking region of SALK 762 insertion in the short exon 13 of the *AtEIF3A* gene (AGI code: At4g11420). The T-DNA insertion in SALK 762 contains a kanamycin resistance cassette in background of Columbia-0 ecotype. ½ MS plates containing sucrose and kanamycin selection were used to germinate the stock seeds obtained from NASC. However, no seed managed to germinate on this media. Another set of seeds was therefore sown on ½ MS plates without selection. On this plate all seeds managed to germinate and grew into wild type-like seedling (data not shown). Therefore, the kanamycin resistance was probably silenced. After transfer to soil, seedlings managed to grew in a wild type-like plants. First, we genotyped these plants using insertion specific SALK LB primer and genome specific SALK_123762.31.45.X RP and SALKSEQ_123762.0 LP primers in a three primer genotyping reaction. Investigated plants SALK 762 showed a mixture of heterozygous (SALK 762 HET) and wild type plants (SALK 762 WT), as the products of PCR reaction contained either two bands, one similar to the genomic band observed in Col-0 control and one band smaller than the genomic or only the genomic band (not shown). The insertion band was sequenced using the SALK LB and SALK_123762.31.45.X RP primers in order to map the insertion site at the border closer to the beginning of *AteIF3A* gene. The sequencing result was aligned to the genomic sequence of *AteIF3A.* The alignment showed that the LB site of the insertion is inside the coding sequence of the *AT4G11420* gene, specifically in the exon 13 (Figure 27A). When protein is produced from this allele, it would affect the protein sequence with changing codon encoding Isoleucine768 to Valine and the following protein sequence to GLFY before stop codon would be reached. The Isoleucine768 is inside the region that was described in mammals as eIF3i binding platform, which we showed has a similar amino acid organisation in plants. Altogether, the SALK 762 HET insertion allele could affect the ability of its protein product to bind eIF3i, additionally missing 219 amino acids of the C-terminal end. Pollen morphology was analysed in the SALK 762 HET and its internal control, the segregating SALK 762 WT plants (Figure 27D). Surprisingly, although an significant increase in collapsed aborted pollen was observed in most SALK 762 HET plants

(SALK 762 HET+) as averaged 28.25 ± 3.67 percent out of total pollen counted in comparison to 0.93 ± 0.72 observed in SALK 762 WT (one-tailed t test, p-value <0.00001), there were several SALK 762 HET plants with no increase in aborted pollen as it averaged 1.61 ± 0.44 in 6 plants (named SALK 762 HET-) we analysed. This meant that there is either second background insertion that is causing the phenotype or that the mutation allele has decreased penetrance. The latter was supported by the fact that no phenotype segregation was observed in the SALK 762 WT lines. To uncover which of these possibilities is more likely, SALK 762 HET+ plants were reciprocally crossed with wild type Col-0 plants in order to clean the possible background insertions from the one in the inspected gene. In the progeny of these crosses (SALK 762* generation, 40 plants analysed), we observed only heterozygous plants with the aborted pollen phenotype (SALK 762* Het+), averaging 31.09 ± 3.87 percent aborted pollen grains per total pollen counted, and wild type-like plants (SALK 762* WT-), averaging 0.96 ± 0.23 percent aborted pollen grains per total pollen counted. Statistical tests were omitted in this case, as the sample number (n=3) of The SALK 762* WT- was limited. However the visual observation clearly showed direct link of the phenotype to the genotype in this case. Moreover, heterozygous plants were observed in both progeny of the crosses, where SALK 762 HET+ was

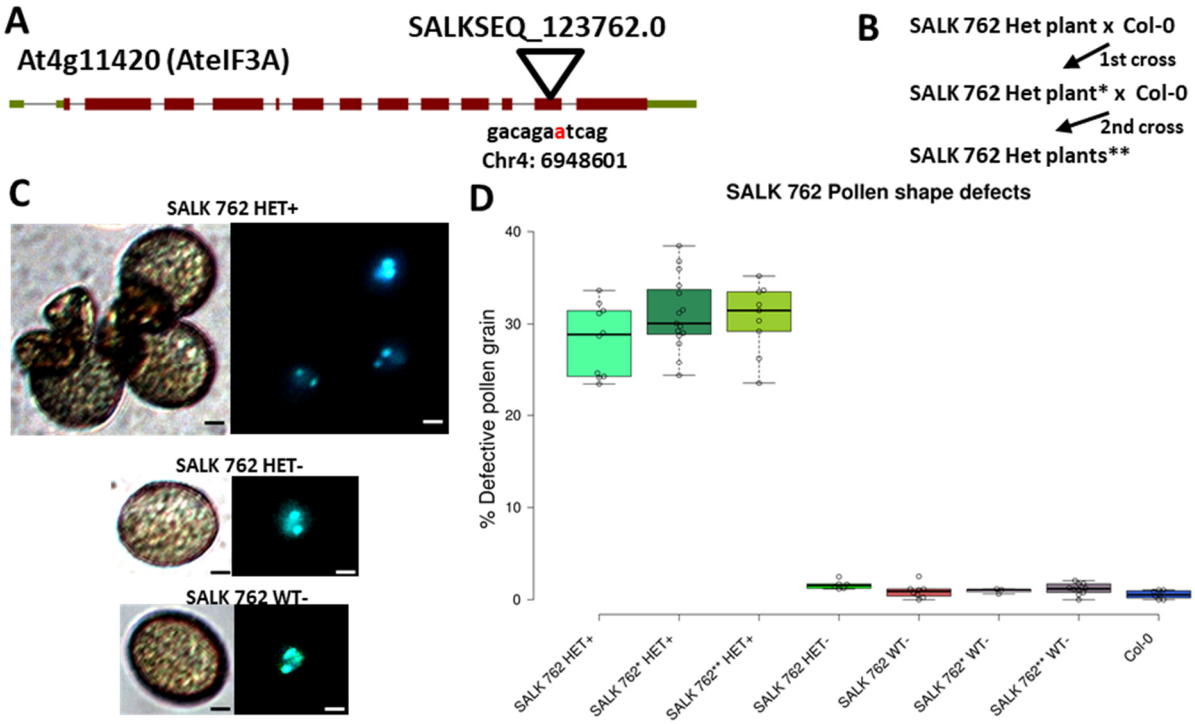

Figure 27: SALK 762 pollen morphology analysis in initial and crossed generations. (A) Schematic illustration of SALKSEQ_123762 T-DNA insertion mapped to the *Arabidopsis* genomic sequence. The detail of first wild type nucleotide affected by the insertion is showed under the insertion site. **(B)** Scheme of the crossing process pertained with SALK 762 line **(C)** Images of SALK 762 pollen morphology (BF), anatomy DAPI, and viability (Alexander´s staining). Scale bar = 5 µm. **(D)** Graphs showing the pollen defects in the SALK 762 initial and crossed generations.

used either as the father or mother plant. This means that the allele is transmittable through both male and female germline, although we observed slightly decreased number of the SALK 762* HET individuals that were a progeny from cross where SALK 762 HET+ was used as father plant. However, insufficient number of plants (n=20) was analysed for statistical testing of transmission efficiency and further analysis of more plants was not done due to the following discovery. The SALK 762* WT- plants were screened for any background SALK T-DNA insertion, using primers specific for the Kanamycin resistance cassette. One PCR product was present in all samples of the SALK 762* WT- plants in this analysis, and not in Col-0 control (data not shown), meaning that background insertions were truly present in the genome of SALK 762 line. Therefore, a second reciprocal crossing with Col-0 followed in order to clean the line once more. In the progeny of these crosses (SALK 762** generation, 40 plants analysed), we again observed only heterozygous plants with the aborted pollen phenotype (SALK 762** Het+), averaging 30.58 ± 3.54 percent aborted pollen grains per total pollen counted, and wild type-like plants (SALK 762** WT-), averaging 1.16 ± 0.58 percent aborted pollen grains per total pollen counted (one-tailed t test, p-value <0.00001). Again, the aborted pollen phenotype was linked to the genotype of SALK 762** Het+ in this generation. Additionally, using the frequencies of observed individuals in this crossing progeny, differences in transmission efficiency were tested with significant reduction of transmission through male germline using goodness-of-fit test based on 2-values χ^2 statistics ($\chi^2 = 6.919$, p=0.009), but not through the female ($\chi^2 = 1.719$, p=0.190). Crossing table is shown in Figure 29. Additionally, the SALK 762** generation was used in further phenotypical analyses, using the SALK 762** WT- plants as the main internal control. Most interestingly, significantly increased number of aborted/unfertilized seeds was found in SALK 762** Het+ plants (Figure 28) as averaged 29.82 ± 3.38 aborted seeds per silique, compared to 0.56 ± 0.87 observed in SALK 762** WT- (one-tailed t test, p-value <0.00001) (Figure 28A). In this observation, the category of aborted/unfertilized ovules was excluded from a small number of white/yellow collapsed seeds observed at similar frequency in both genotypes, as averaged 1.79 ± 1.72 and 2.81 ± 2.25 white/yellow seeds per silique in SALK 762** Het+ and SALK 762** WT- (p-value=0.242), respectively. The increased number of aborted/unfertilized ovules also reciprocally decreased the number of normal seeds per silique (statistics not presented) and lead to observably decreased length of the silique (not measured) (Figure 28B). The pollen anatomy was analysed in both genotypes of SALK 762** generation, where only non-aborted

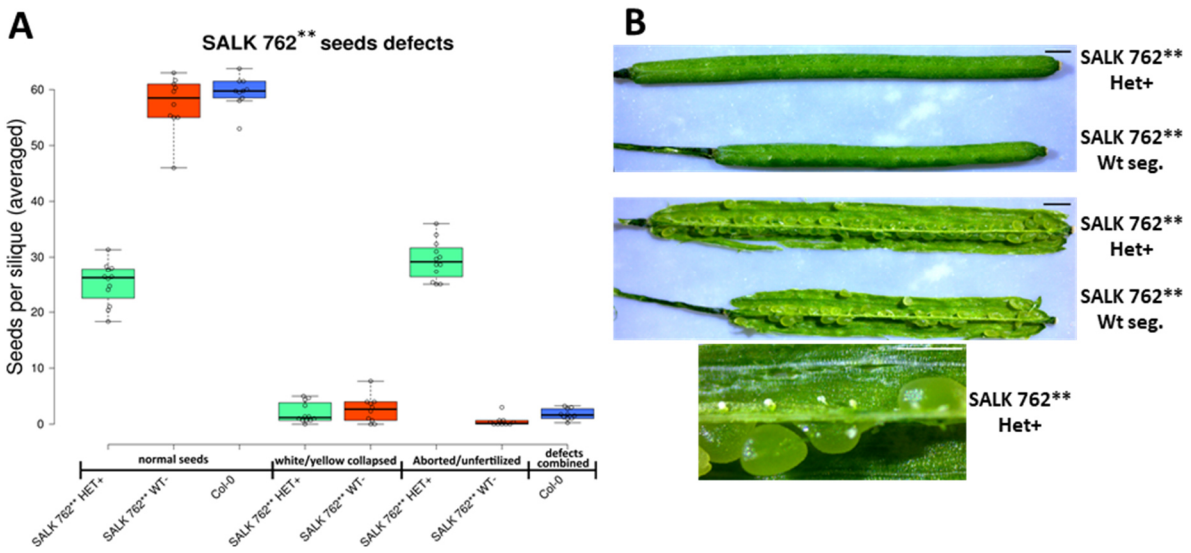

Figure 28: SALK 762 siliques dissection analysis. (A)** Graph showing observed defects in the seed formation in the SALK 762** generation. **(B)** Siliques of the SALK 762** generation before and after dissection. Scale bar = 1 mm.

pollen of the SALK 762** HET+ genotype were scored, as the aborted pollen showed no DAPI signal, while all pollen grains were included in the SALK 762** WT- plants. The results observed point to no increase in the non tricellular pollen frequency in the SALK 762** line, as averaged 1.81 ± 0.86 percent of non-tricellular pollen (from the normally shaped non-aborted pollen) in SALK 762** HET+ and 1.42 ± 0.42 percent of non-tricellular pollen (total pollen counted) in SALK 762** WT- plants. Analysis of the Col-0 control is given in previous chapters and was not different from the SALK 762** observations. Statistical testing is not included due to the inconsistency of methodical approach between the two samples. No segregation of the phenotype between the plants of identical genotype was observed in the progeny of the crosses and lead to conclusion that the SALK 762 HET- plants observed in the initial generation could have been a result of mislabelling done somewhere during the genotyping of the initial generation. Altogether, the SALK 762 insertion was mapped to the *AteIF3A* coding sequence, presumably disrupting the eIF3i binding platform. Mutant plants were observed only in heterozygous genotype, which leads to increased number of aborted pollen, decreased transmission through male germline and decreased seed production due to aborted/unfertilized ovules.

*Complementation and further analyses of SALK 762 failed due to sudden re-occurrence of the segregation in the progeny of SALK 762***

After basic phenotypical characterization of the cleaned SALK 762 line, a more detailed study was planned in order to uncover 1) how the insertion affects the *AteIF3A* expression, 2)

at what stage of the pollen development the pollen is arrested, 3) whether the seed formation defects are caused by decreased ovule viability or by post-fertilization defects 4) what is the segregation ratio of the self-pollinated clean SALK 762** generation and 5) could the phenotype be rescued by complementation with the CDS of eIF3A driven by the native p3A promoter. However, the progress of this analysis was halted as the phenotype-genotype segregation re-occurred in the progeny of SALK 762** generation. However, again only SALK 762 HET- segregants were observed, while no SALK 762 WT+ plants were detected. Nevertheless, this was an unexpected experience and aforementioned more detailed analyses of the segregation were postponed in order to investigate the causes of the segregation and include more plants in the study to obtain more precise data for testing hypothesis based on common genetic rules. Only analysis done was the aniline blue staining, where all genotype-phenotype combinations, SALK 762 HET+, SALK 762 HET- and SALK 762 WT- were tested in pollen tube growth ability (Figure29C). In all cases, pollen tubes grew all the way through the transmitting tissues and no observable defect was detected. The very last analysis done within the final days of this work was the analysis of *AteIF3A* expression on limited number of samples per genotype-phenotype category (2 biological replicas for SALK 762 HET+ and SALK 762 HET-, only one for SALK 762 WT-). The measurements showedgreat variability in relative expression, mostly in the SALK 762 HET+ samples, as averaged 23.58 ± 20.19 (primer set 2) and 35.77 ± 29.98 (primer set 4) percent of ACT8 expression. Moreover, the primer set 2 showed increased relative expression SALK 762 HET-, as averaged 53.09 ± 7.24, 31.31 ± 3.84 and 30.29 ± 4.50 percent of ACT8 expression in SALK 762 HET-, SALK 762 WT- and Col-0,

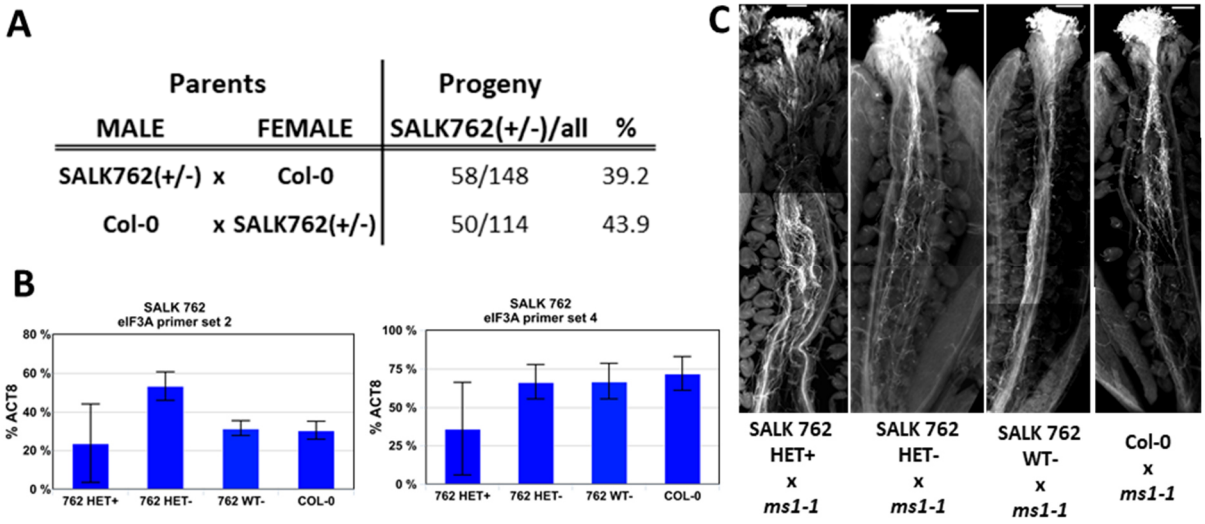

Figure 29: Additional SALK 762 phenotype analysis. (A) Crossing table of analysed genotypical frequencies in the progeny of crossed plants. (upper part) **(B)** RT-qPCR results of the *AteIF3A* relative expression in different SALK 762 genotype-phenotype populations on limited samples (see text for further info.) **(C)** Aniline blue staining of pollen tubes within the pistill. Col-0 results is given as control. Scale bar = 200 µm.

respectively. This difference was not observed using the primer set 4, as averaged 66.26 ± 11.19, 66.26 ± 11.66 and 71.82 ± 10.96 percent of ACT8 expression in SALK 762 HET-, SALK 762 WT- and Col-0, respectively. Nevertheless the great variability observed in all measurements was again striking (therefore not statistically tested) and lead to conclusions that it needs to be repeated after consistent results within the control samples are reached.

Modulated expression of eIF3A using RNAi interference and Crispr/Cas9 approach

As the analysis of T-DNA insertion lines did not produce any sufficiently characterized ateIF3A mutant alelle, we started to think about different approaches how to modulate the expression of *AteIF3A* in plants. We decided that we could downregulate the *AteIF3A* expression by RNAi interference. Therefore we cloned a non-coding expression cassettes of sense and reverse complement region of the AteIF3A CDS linked by Kannibal intron and driven by different sporophytic and pollen specific promoters: 35Sdouble, p3a, pLat52, pJas and pAsy.. pJas and pAsy are male gametophyte-specific promoters expressed during different stages of microgametogenesis. As a selection marker, we used BASTA resistance cassette, because we wanted to introduce the 3A-RNAi cassettes into homozygous p3A::eIF3A-GFP and Lat52::eiF3A-GFP background. In ideal situation, the 3A-RNAi-induced silencing would also affect the GFP fusion constructs and decrease the fluorescent signal, serving as a control of the silencing efficiency. However, time constraints only allowed us to transform Arabidopsis plants with the assembled expression cassettes, and the T_0 plant were still maturing at the time of this thesis completion,. In order to produce a knockout allele, we also plan to edit the of the *AteIF3A* gene using Crispr/cas9 approach in future experiments.

Isolation of epitope tagged AteIF3A

Lat52 pollen specific promoter was used to drive AteIF3 fusions with epitope tags

One of the aims in the eIF3 research in our lab is to isolate, purify and identify the whole complex in the male gametophyte. As the eIF3A is interacting with most if not all the other subunits of the eIF3 complex in yeast and mammals, it makes it an ideal candidate for fusing it with a small epitope tag commonly used in recombinant protein purification. Therefore, we cloned expression cassettes driven by the Lat52 pollen specific promoter that would produce AteIF3A fused to either Hemaglutinin, FLAG or Hellfire (6xHis + 3x FLAG) epitope tag. We transformed plants with Lat52::HA-eIF3A, Lat52::eIF3A-FLAG and Lat52::eIF3A-Hellfire expression cassettes, that should express products of 122, 118 and 121 kDa specifically in the pollen, respectively. We screened T_1 generation plants, using flowers from each independent line as sample material for protein extraction and recombinant protein detection.

Detection of AteIF3A epitope tag fusion protein failed from native protein extraction lysate

Proteins were extracted from collected flowers of independent transgenic lines Lat52::HA-eIF3A, Lat52::eIF3A-FLAG and Lat52::eIF3A-Hellfire using the native protein extraction protocol. To reach the most identical concentration between samples, we used similar volume of DEM extraction buffer per flower. Protein lysate samples were then mixed with appropriate volume of Sample buffer, sonicated and loaded on polyacrylamide gel in two replicas, once on experimental gel and once on loading control gel. All gels included Columbia flowers sample and Sample buffer diluted in DEM buffer as negative control samples. After SDS-PAGE separation, one gel was used for western blotting and subsequently stained with Coomassie brilliant blue dye, while the other gel was stained with Coomassie brilliant blue dye right away. Loading control gels are shown in Figure 30. The proteins were transferred successfully at the lower molecular weights (lower than 60 kDa) and only partially at the higher molecular weights. Nevertheless, all bands of the protein marker were visibly transferred on all blotting membranes (not shown). However, we failed to observe any band of correct size on the membrane after secondary antibody signal development (Figure 31). With rabbit anti-FLAG antibody, both used for Lat52::eIF3A-FLAG and Lat52::eIF3A-Hellfire samples, we observed a great abundance of non-specific signal development at multiple molecular weights. The non-specifity was apparent

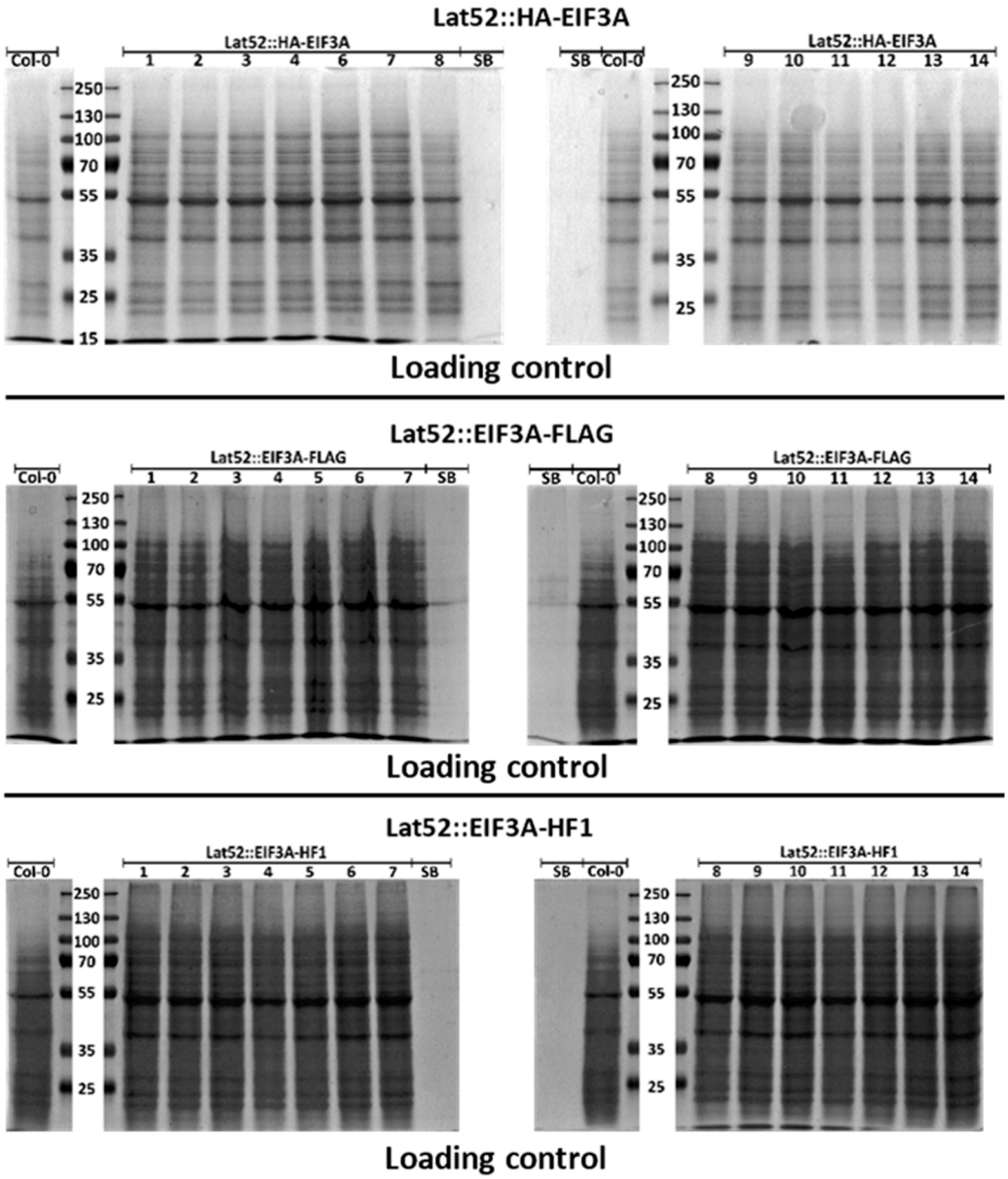

Figure 30: SDS-PAGE loading control gels. Gels running Lat52::HA-eIF3A, Lat52::eIF3A-FLAG or Lat52::eIF3A-Hellfire samples Samples are marked with numbers that represent independent lines. Col-0 negative control is included in every gel, together with sample buffer diluted with DEM extraction buffer (SB samples). Molecular weights of the marker are indicated.

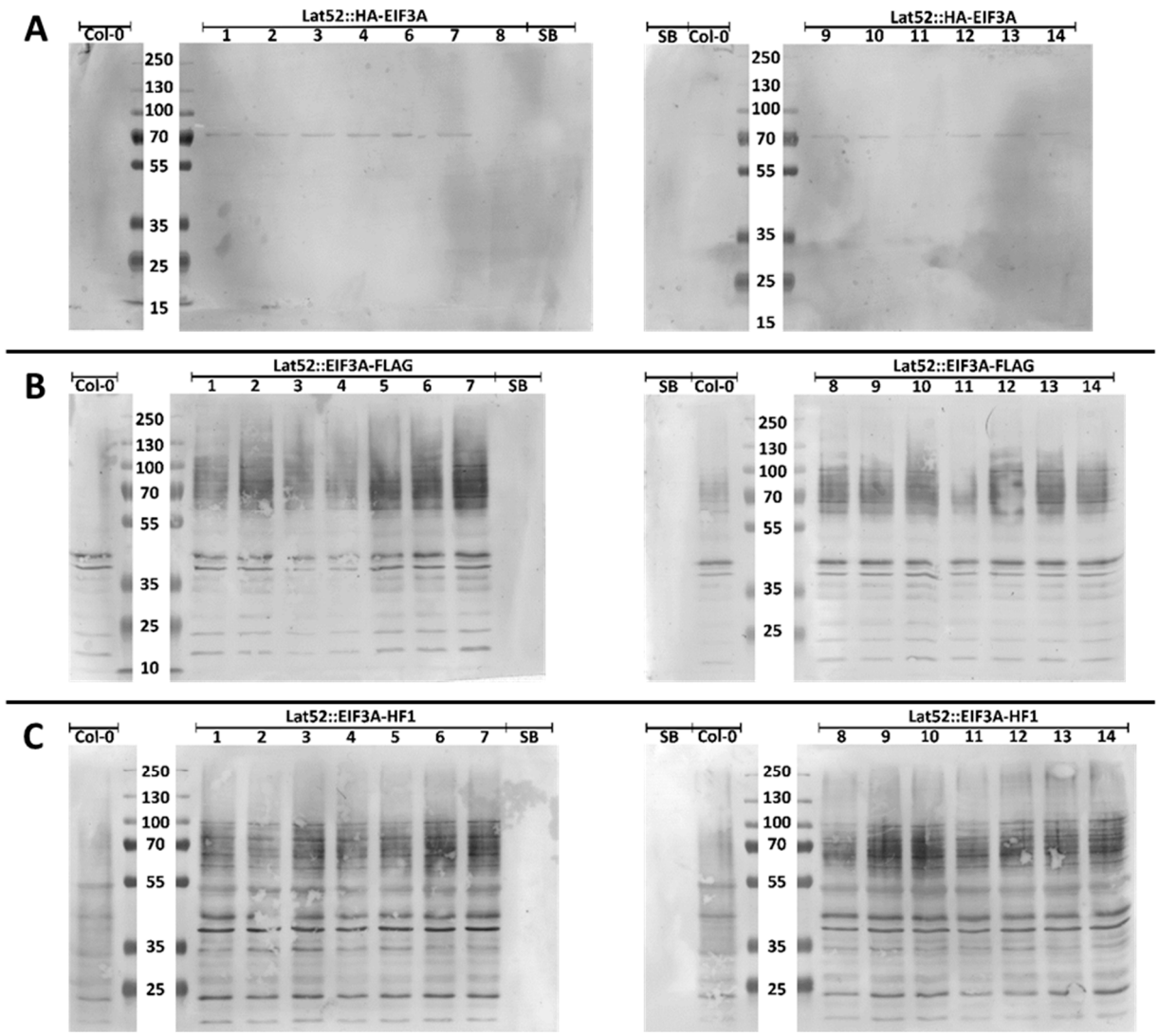

Figure 31: Western blot membranes after signal development. Membranes with Lat52::HA-eIF3A, Lat52::eIF3A-FLAG or Lat52::eIF3A-Hellfire samples are separated by black borderline, Samples are marked with numbers that represent independent lines. Col-0 negative control is included in every gel, together with sample buffer diluted with DEM extraction buffer (SB samples). Molecular weights of the marker are indicated.

as similar non-specific band pattern was present in the negative control sample of Col-0 as well, but not in the DEM buffer sample. At higher molecular weights, a region of non-distinguishable signal development with several distinguishable weak bands was observed, that was much weaker in the Col-0 sample, but still visibly present. Using the rabbit anti-HA antibody for Lat52::HA-eIF3A samples, we observed a weak band at approx. 70 kDa after secondary antibody signal development. This band was not present in the Col-0 and Sample/DEM buffer controls. However, it goes almost 50 kDa under the expected molecular weight of HA-eIF3A. Therefore, no combination of eIF3A epitope tag fusions was present at the expected protein product size in all lines inspected.

No improvements observed after All-in-one extractions of pooled AteIF3A epitope tag fusions

After we failed with the detection of eIF3A epitope tag fusions using the native protein extractions with DEM buffer, we tried different approach. We combined flowers of 5 or 4 lines together in three pooled samples per each construct and performed the All-in-one extraction Kit for simultaneous DNA, RNA and protein extraction. We then repeated the SDS-PAGE from the extracted proteins as described in previous subchapter with freshly prepared primary rabbit anti-HA, rabbit anti-FLAG and secondary anti-rabbit antibodies. However, the result was the same as in the previous experiment. After equally successful transfer to the blotting membrane and signal development (Figure 32A), there was no apparent band at expected molecular weight that would clearly differ from the Col-0 control in either of the samples. Moreover, the non-specific signal development was present even in the freshly prepared anti-HA antibody, giving a pattern of dozens of bands of different molecular weights. When we were not able to detect any protein in the eIF3A epitope tag lines, we used the RNA isolated to check if the expression cassettes are producing any transcripts. After the DNase treatment and Reverse transcription, we used the cDNA as a template for qPCR with primers specific for the eIF3A (primer set 2) and a combination of Epitope tag and eIF3A specific primers that would produce a PCR product detectable by the LightCycler480. ACT8 specific primers were used to measure the levels of ACT8 mRNA for data normalization. For each primer set, three reactions serving as technical replicas were tested. The qPCR showed no expression using the epitope tag specific primers. The averaged results normalized to % of ACT8 using the relative expression formula are shown in Figure 32B. It is clear from the charts, that although AteIF3A is expressed in flowers of inspected lines and in Columbia-0 control as well, the expression cassette itself is not expressed

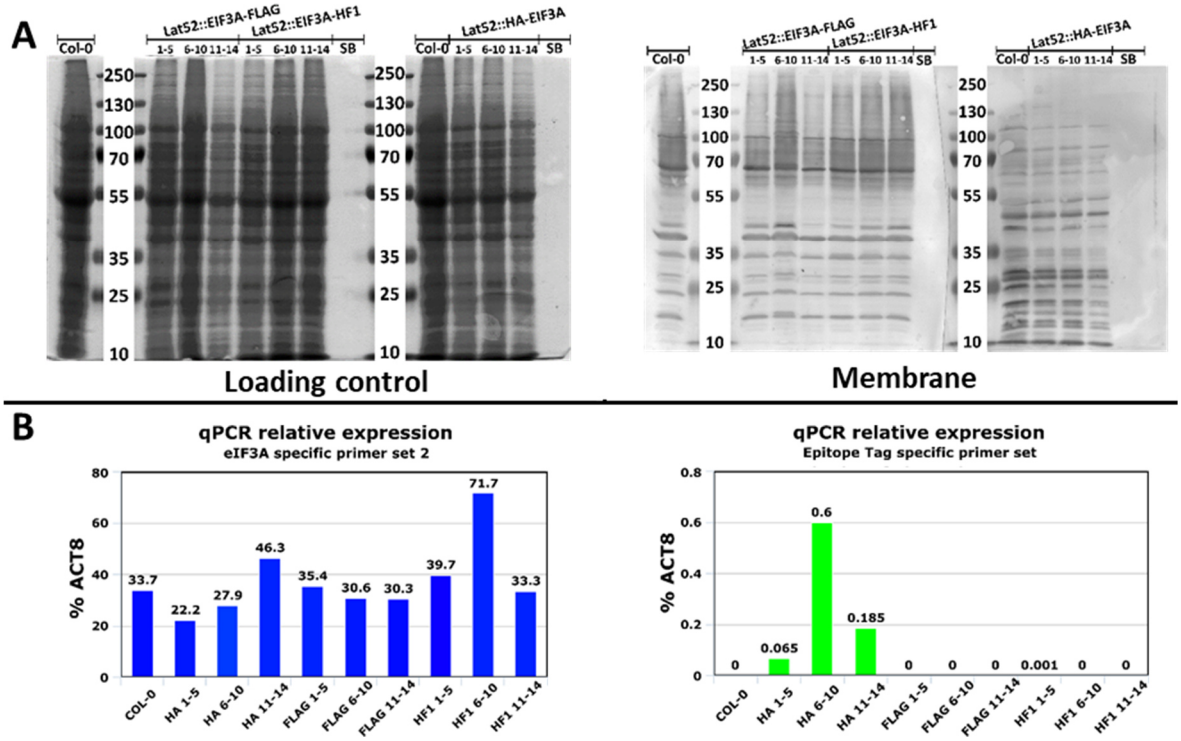

Figure 32: Western blot and qPCR analysis of pooled samples after All-in-one extraction. **(A)** SDS PAGE loading control and the blotting membrane after signal development. Samples are marked with the construct name and a number that represent pooled independent lines. Col-0 negative control is included together with sample buffer diluted with DEM extraction buffer (SB samples). **(B)** RT-qPCR analysis showing relative expression of eIF3A using the *AteIF3A* specific and epitope tag specific primers. Due to great variability, the standard deviation is not shown.

at all. Only detection of the Lat52::HA-EIF3A showed some extremely low level of expression, which is not consistent with the fact that Lat52 driven constructs should reach significantly higher expression, although in pollen only. Interestingly, the endogenous expression of eIF3A is highly variable between the samples itself, assuming that the expression cassette is not at all or minimally additive to the total eIF3A CDS-containing transcripts. Altogether, the detection of AteIF3A protein fused to epitope tag was not successful.

Stable *Nicotiana tabacum* transgenic lines

Transformation experiments lead to multiple independent lines per construct

As Tobacco pollen is one of the most widely used model system for pollen biology research, we decided to produce stable *N. tabacum* transgenic lines expressing eIF3 subunits under pollen-specific promoters. As was shown in previous chapters, *N. tabacum* eIF3A orthologues are closest to the *Arabidopsis* eIF3A that was the main object of study in this work. Therefore we used the expression cassettes with CDS of *Arabidopsis* eIF3 subunits. Subsequently, several marker lines were produced by this method that would serve as internal controls in following experiments. This relatively time demanding procedure included more simultaneous transformation experiments, which were evenly accomplished by several members of our lab under the supervision of RNDr. Oldřich Navrátil CSc. from the Laboratory of Virology at the Institute of Experimental Botany. The results presented in this work only include those on which I participated with major contributions. In our hands, at least 2, but usually more than 5 independent transgenic plants (i.e. not regenerated from the same callus) were obtained by the procedure. These plants usually produced a small number of flowers that were able to self-pollinate and produce seeds. Before this text was written, the time-demanding procedure together with long life-cycle of the *N. tabacum* enabled us only to confirm the expression in the pollen of transformed T_0 generation of pLat52::FLAG-YFP, pLat52-AtCenH3:YFP and pLat52::AtEIF3A-GFP and in T_1 generation seedlings of the 35Sdouble::GFP-GUS.

Stable transgenic N. tabacum expressing fluorescent markers in pollen

Two and seven independent transgenic plants carrying pLat52::FLAG-YFP and pLat52::AtCenH3-YFP were obtained, respectively. When T_0 plants started to produce, pollen was collected from 1 anther and the signal was observed with Nikon TE2000E fluorescent inverted microscope (Figure). In the pLat52::FLAG-YFP lines, clear GFP signal was observed in the cytoplasm of vegetative cell and was slightly enriched in the vegetative nucleus but not in the generative cell. The FLAG-YFP fusion should be localized freely in cytoplasm but is small enough to slip through the nuclear pore complex inside the vegetative nucleus, and the observations are in agreement with that. In the pLat52::AtCenH3-YFP line expressing the *Arabidopsis* centromeric variant of the histone H3 fused to GFP, a clear signal was observed in the vegetative nucleus but not in the generative cell or in the cytoplasm of the vegetative cell. This observation is in agreement with the histone being a strict nuclear protein. However,

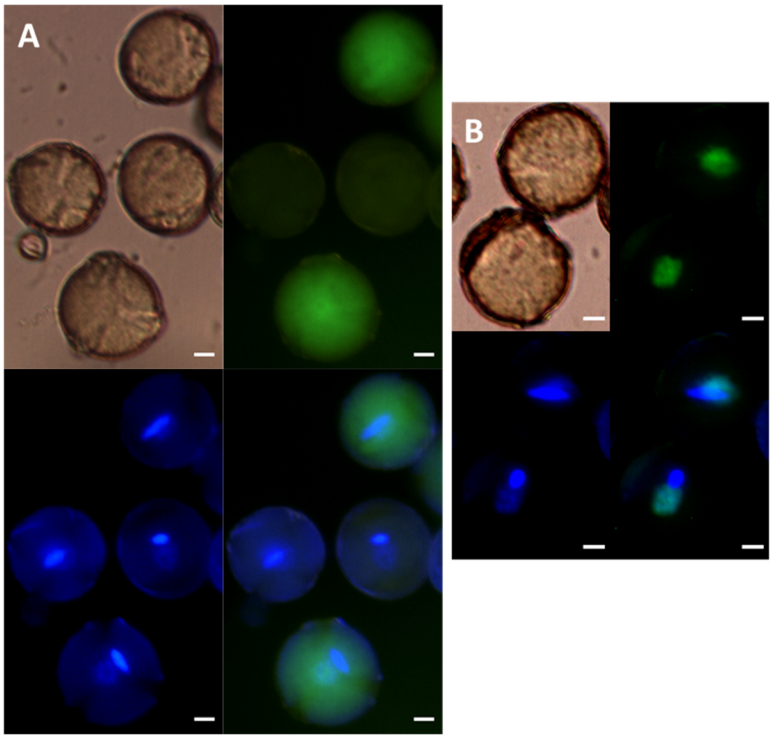

Figure 33: Expression of FLAG-YFP and AtCenH3-YFP in *N. tabacum* mature pollen grains. (A) Series of single section images of the FLAG-YFP driven by pLat52 pollen specific promoter. From left upper corner to right lower corner: bright field, green channel, DAPI (blue) and merge of the latter two. **(B)** Series of single section images of the AtCenH3-YFP driven by pLat52 pollen specific promoter. From left upper corner to right lower corner: bright field, green channel, DAPI (blue) and merge of the latter two. Non-transgenic pollen graines with no YFP expression are included as negative control. Scale bar = 10 µm is common for all images.

whether *Arabidopsis* CenH3 is bound in the histone complex or is just freely in the nucleoplasm was not observable.

Stable transgenic N. tabacum expressing marker protein in the sporophyte

The expression cassette with 35Sdouble::GFP-GUS was used as the sporophytic marker for successful transformation event. Seeds from 8 T_0 plants were harvested and sown on selection media containing 300 µg/mL kanamycin (a strong selection recommended by Oldřich Navrátil). When kanamycin-resistant seedlings had fully developed cotyledons and were clearly distinguishable from kanamycin-sensitive seedlings, they were analysed in GUS staining assay. After 1 hour, 2 to 3 seedlings were taken from the GUS staining solution, destained in ethanol series and observed under ZEISS Stemi 508 stereomicroscope. In this assay, 5 out of 8 independent lines initially obtained showed GUS signal development. The signal was clearly present in the root and shoot tissues, while not present in the wild-type plant (Figure 34). The rest of the seedlings were left in the GUS staining solution for another 23 hours in order to see if the lines with no signal after 1 hour of staining would develop some overnight. However, it was not the case and only the 5 independent lines showing GUS signal development already after 1 hour were overstained while the other 3 lines had no GUS signal.

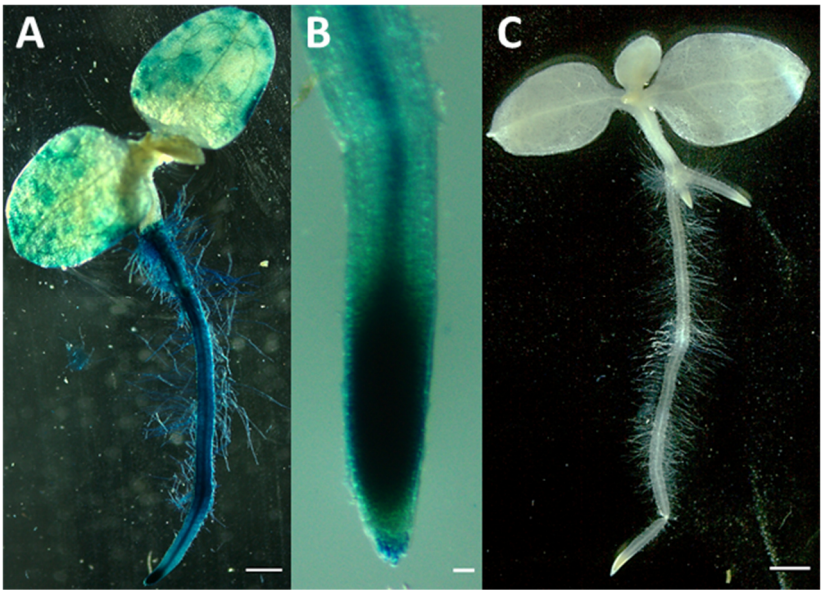

Figure 34: Expression of GFP-GUS in *N. tabacum* sporophyte. **(A)** GUS staining of the transgenic 35Sdouble::GFP-GUS *N.tabacum* seedling (left). Scale bar = 1 mm. **(B)** Detail on expression in the root meristem (right) Scale bar = 50 µm. **(C)** Equally treated seedling of the wild type control. Scale bar = 1 mm.

Stable transgenic N. tabacum expressing Arabidopsis AteIF3A in mature pollen

Seven independent transgenic *N. tabacum* plants expressing *Arabidopsis* AtEIF3A-GFP fusion protein under Lat52 pollen specific promoter were regenerated after leaf disc transformation. When T_0 plants started to produce flowers, pollen was collected from 1 anther and the signal was observed with Nikon TE2000E fluorescent inverted microscope. In this case, however, the GFP signal was hardly observable and the experiment was repeated using the Nikon spinning disc microscope. With the confocal microscope, the difference between transgenic and non-transgenic pollen was clearly distinguishable, although still very weak (Figure 35). The AteIF3A-GFP was clearly localized to the cytoplasm and not in the generative cell. Whether the signal was present in the vegetative cell was not clear, as vegetative nucleus was not properly observable as the DAPI staining was insufficient.

Verification of the protein product expressed in the pollen

During the flowering of the T_0 generation, 2-3 anthers were collected from 1 open flower. Using the native protein extraction protocol, proteins were isolated from this tissue, separated on SDS-PAGE and western blotting was done using rabbit anti-GFP primary and goat anti-rabbit secondary antibody conjugated with alkaline phosphatase (Figure 36). We expected protein products of 141, 30 and 48 kDa in lines pLat52::AtEIF3A-GFP, pLat52::FLAG-YFP and pLat52::AtCenH3-YFP, respectively. In the case of pLat52::AtCenH3-YFP, no clear band was distinguishable on the membrane after signal development. It could be due to insufficient breakage of the vegetative nucleus in the proceduce of protein extraction. In case of

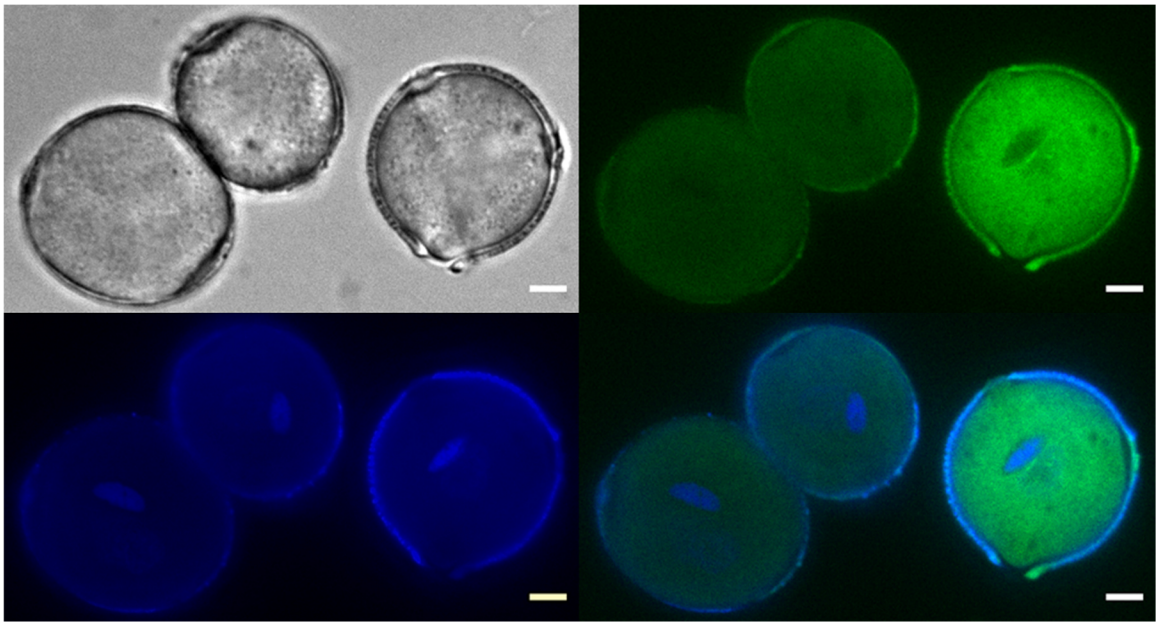

Figure 35: Expression of AtEIF3A-GFP in *N. tabacum* mature pollen grains. (A) Series of single section images of the AtEIF3A-GFP driven by pLat52 pollen specific promoter. Non-transgenic pollen graines with no GFP expression are included as negative control. From left upper corner to right lower corner: bright field, green channel, DAPI (blue) and merge of the latter two. Scale bar = 10 µm is common for all images.

pLat52::FLAG-YFP, one clear band was observed on the membrane. Although it probably is the FLAG-YFP fusion protein, it was above the 40 kDa mark instead of being below it. Therefore, the protein was not of a correct size expected. In case of pLat52::AtEIF3A-GFP, a very weak faint band was observed above the 130 kDa mark, probably being the AtEIF3A-GFP. However, two strong bands were observed between 40 and 50 kDa marks. The smaller band closely resembled the band observed in pLat52::FLAG-YFP in its size, although being slightly larger. Two other very weak fainted bands were observed under the 50 kDa mark and under the 70 kDa mark. It is possible that *Arabidopsis* eIF3A is degraded when expressed in the Tobacco pollen under strong promoter, although this claim will need further experiments to be proven.

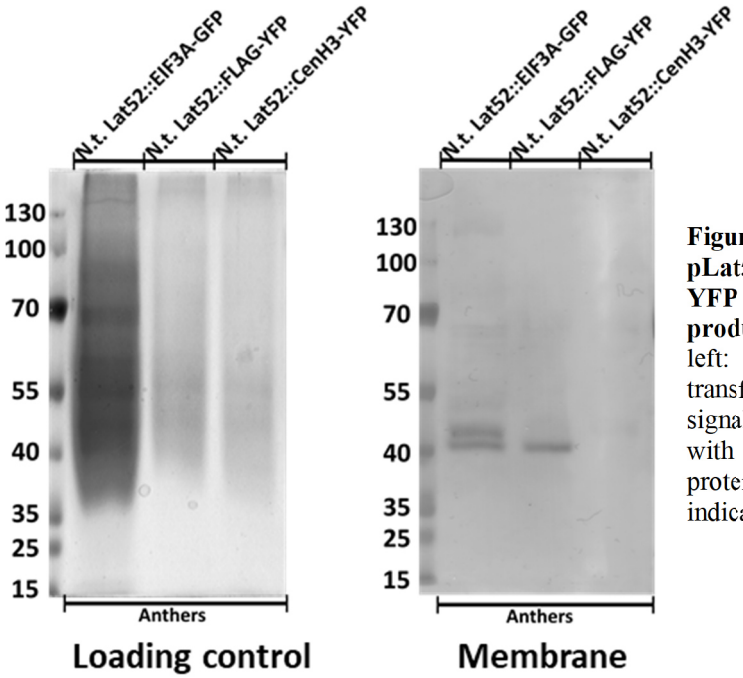

Figure 36: Western blot analysis of pLat52:AtEIF3A-GFP, pLat52::FLAG-YFP and pLat52::AtCenH3-YFP protein products in mature pollen grain. From left: SDS-PAGE gel before and after transfer, Western blot membrane after signal development. Samples are marked with the construct name. Where visible, protein marker molecular weights are indicated.

Discussion

Plant species possess at least one eIF3A homologue in their genome

To understand a never before studied gene in plants, we first analysed the available sequence databases of plant proteins. Commonly known phenomena is that the protein products of any given homologous genes are more conserved than its nucleotide sequences. We therefore omitted the DNA sequence comparisons and analysed the eIF3a homologues on the level of protein. The eIF3a subunit was found present in almost all eukaryotic species, including plants, where such effort was made, using either molecular or *in silico* –based approach. We knew already, that the whole eIF3 complex was found in plants and resembles the mammalian 12-subunit complex (Discussed in Raabe et al., 2019). Nevertheless, we started with formulating simple questions about eIF3a in plants: Is there a homologue of eIF3a in plants? If yes, how similar to the characterized orthologues from other eukaryotes it is? And is there one gene or many paralogue genes that encode eIF3a in plant species? Our initial hypotheses was that at least one eIF3a homologue should be found in all plant species, although it might have diverged in many aspects in the evolution and, as it is quite usual for genes in plant species, might be present in multiple copies per genome. For solving these questions, we performed *in silico* based analysis using simple computational biology methods. We used the mammalian eIF3A protein sequence to search protein databases of limited number of plant species for homology. Moreover, we added 4 non-plant eukaryotic species that served as control for plant and non-plant species comparisons. We found that at least one homologue of eIF3a was found in all plant species included in the search and that, with exceptions, there is usually just one gene encoding the eIF3A subunit. The exceptions of 3 and 4 eIF3A paralogues found in two distant species, *P. patens* and *N. tabacum*, respectively, could be simply caused by recent duplication events. This is supported by several facts: 1) other investigated plant species, including *A. trichopoda*, the most basal Angiosperm, contained one only one eIF3A subunit; 2) eIF3A proteins in both *P. patens* and *N. tabacum* share highest protein sequence similarity and identity only with their respective paralogues and not between orthologues from other plant species; 3) *N. tabacum* is an allotetraploidic hybrid, therefore only one duplication event inside the *Nicotiana* genus would lead to 4 paralogues in the hybrid inspected. The latter is also supported by the fact that inside the *N. tabacum* clade, 4 eIF3A paralogues form two distant clades that most probably originate from either of the original *N. sylvestris* or *N. tomentosiformis*. The only disadvantage of our search was that the template used was the well characterized mammalian

eIF3A protein, which is longer than most of its eIF3A orthologues. However, the human additional 400 amino acids are composed of highly repetitive low complexity regions, while the first approx. 900 amino acids are well conserved between mammals and yeast at least, therefore true homologues were still expected to appear in the search. Although more detailed analysis with more plant species would be needed for making a strong claim that all plants have and eIF3A homologue, we altogether confirmed our initial hypotheses with the findings that every plant species included in this analysis contains at least one eIF3A homologue.

High levels of homology are shared between plant eIF3A subunits

Next we asked how similar the eIF3a orthologues in plants are. Our initial hypotheses were that plant eIF3 subunits share higher levels of homology than plant and non-plant eIF3A subunits and that eIF3a, as a part of highly conserved translational machinery, is not a highly variable protein and the shared homology would be high. We used alignments algorithms to calculate protein sequence identity and similarity, two commonly used parameters of sequence homology. The results of this analysis could be summarized in following points: 1) The highest eIF3A homology is shared between paralogues of *P. patens* and *N. tabacum;* 2) eIF3A homology shared between plant species is overall very high; 3) eIF3A homology shared between plant species is way higher than between plant and non-plant species; 4) eIF3A homology between relatively distant non-plant species is similar to the homology between plant and non-plant species; 5) *S.cerevisie* eIF3A subunit shares the lowest homology with both plant and non-plant species. Overall, this supports both our hypotheses. Although the identity was not remarkably high between plant and non-plant species, one must consider that the numbers were counted from full-length protein sequences of evolutionary very distant species. Moreover these levels of homology are common between plant and non-plant eIF3 subunits (Burks et al., 2001) or between eIF3 subunits of non-plant species (For example Asano et al., 1997; or Crane et al., 2000). Additionally, the yeast eIF3A difference could be explained by the fact that yeast eIF3 complex is different in composition from other species and individual subunits (Reviewed in Valášek et al., 2017). Altogether we confirmed our hypotheses about high levels eIF3A homology in plants.

Plant-specific differences in the PCI domain

After we confirmed that eIF3A subunits are overall well conserved in plants, we aimed at the differences between plant and non-plant eIF3 subunits. Our initial hypotheses were that major

domains and their organisation would be identical between the eIF3A orthologues. We used self-dotplot charts, domain search algorithms, multiple sequence alignments combined with structural modelling and the knowledge about regions with particular function described in literature to characterize the plant eIF3A protein sequences. The results of this analysis are summarized in following points: 1) Self-dotplots of plant and non-plant species have similar pattern, showing a low complexity region in the C-terminal half of the protein sequence; 2) N-terminal mRNA-interacting region described in yeast eIF3A orthologue is highly conserved between all species in the analysis; 3) eIF3b/i binding platform described in mammals shares common amino acid organisation pattern in all species in the analysis; 4) the position of PCI domain is strictly positioned to the N-terminal half of the eIF3A in all species in the analysis; 5) plants differ in the PCI domain length due to variability in the length of three interhelical loops; 6) Plant-specific interhelical loops are well conserved between plant species in the analysis. Two sequences differed in the self-dotplot analyses, the human and yeast. The human eIF3A showed a second region of repetitive low complexity in its C-terminal end. This region was already described as the DRYG repeat domain and is specific to Vertebrata, as it was shown to be a part of interleukin signalling pathway (Saletta et al., 2010). The lack of the otherwise obvious low complexity region in yeast could be explained again by the different composition of the eIF3 complex in yeast, although direct evidences to support this claim are lacking in the literature. Overall, this analysis supports our hypotheses about conserved eIF3A domain organisation, although the PCI domain differences found in the analysis were highly surprising. The PCI domain is a structural domain with high and low conserved regions (described in Literature overview), however, despite its overall good conservation in all species, additional 8 and 10 amino acids were found in all plant species, including *C. reindhartii*, while were not present in non-plant species. Additionally, another region present in non-plant species is not present in plants. Whether these regions alter the overall structure of plant eIF3A proteins are not known. However, the *Arabidopsis* eIF3A model built in this work suggests that all three regions are located between helices, therefore a sever disruption of the protein structure could be not expected. Moreover, the additional Region 1 and the lacking Region 3 are in a very close proximity to each other in the tertiary structure of the protein, which could mean that in plants, the additional Region 1 may compensate the missing Region 3 in its putative function (unknown) and/or was just given the opportunity to enlarge and fill the missing space of Region 3 without disrupting the overall structure. Moreover, the major differences in the regions of the PCI domain between plants and non-plant species could be behind the reason why both mammalian and plant eIF3A failed to complement the yeast eIF3A mutant *rpg1-1* (Valášek et

al., 1998; Sabelli et al., 1999). Altogether, we confirmed our initial hypotheses about conservation of eIF3A domain organization in plants and found novel plant specific-regions inside the eIF3A PCI domain. It would be interesting to place the plant eIF3A structure model to the mammalian eIF3 complex structure and identify where these different regions clash with the other subunits. Also, performing a similar computational biology analysis for other subunits of the eIF3A complex is an interesting challenge to the following doctoral project. Moreover, heterologous expression of yeast or human eIF3A, particularly in *eif3a* mutant, could help us uncover more information about the differences between the eIF3A in plant and non-plant species.

AteIF3A gene is expressed in proliferating tissues of the sporophyte

When we had a clearer idea on how eIF3A subunits looks like in plants and what characteristics common with other eukaryotes it possess, we were curious where eIF3A subunit is expressed in the plant and whether it is expressed in the male gametophyte. Our initial hypotheses were that eIF3A is expressed in most if not all cell tissues as the translation initiation is needed in all cells that translate. Moreover, proliferating and developing tissues need new sets of translation machinery and high levels of translation to cover the needs for new sets of proteins. Such tissues would therefore express the eIF3A more than fully developed tissues. Although expression databases already contain supporting information on this hypotheses, the expression analysis is one of the key steps in any gene characterization analysis. We used the reporter system with GFP-GUS recombinant protein fusion driven by the native promoter. Usually, a region about 1-2 kbp long is considered as standard for native promoter length. However the *AteIF3A* gene is in close sequence proximity to other gene at its 5′ end. Therefore such long region would include the neighbouring gene. Therefore, we used 750 bp region upstream from the *AteIF3A* gene start codon to the neighbouring gene border as the native promoter, p3A. Based on the observed p3A::GFP-GUS expression pattern observed in *Arabidopsis*, we conclude that: 1) the p3A promoter is highly active in the root system, being strongest in the cell division zone of the root meristem and less in the central cylinder, root hair and root primordia; 2) the p3A promoter is relatively less active in the vegetative shoot in the comparison with the root, where it is active in cotyledon tips, cotyledon and leaf vascular tissues; 3) the p3A promoter is active in floral tissues, being strongly expressed in developing anthers, the stigma and ovules and 4) the p3A::GFP-GUS showed a great variability of the expression pattern. This altogether confirmed our hypotheses about AteIF3A being expressed

where new set of the translation machinery is needed, although we approached great variability of the promoter activity between independent lines. Similar observation the expression pattern of eIF3A in plants was observed in maize and rice (Sabelli et al., 1999; Li et al., 2003) also in promoter analyses of other eIF3 subunits in our lab (unpublished). The variability observed in the expression pattern could be caused either by 1) position effect of insertion, 2) incomplete insertion of the expression cassette; 3) random silencing of the expression cassette in dividing cells or 4) regulatory elements not present in the native promoter. While 1) and 2) being the most probable and commonly observed phenomena, some facts are suggesting it could be a combination of 1) and/or 2) and at least one more of the given possibilities. The 3) is supported by the patchy expression observed in GUS staining of leaves and in GFP fluorescence in root meristems. If this is the case, individual cells in the meristem would either have the p3A active or silenced, which would be then the case for all their daughter cells as well, giving the chimeric patchy expression pattern. The 4) is strongly supported by the fact that the p3A::GFP-eIF3A expression was similar to the p3A::GFP-GUS, while evenly distributed p3A::eIF3A-GFP expression was observed in all cells of the root meristem of all inspected plants from multiple independent lines carrying this expression cassette. This would mean that the AteIF3A CDS possess some cis-regulatory elements and the expression is fully activated only when the CDS directly follows the native promoter. Cis-regulatory elements in the downstream of start codon were already described in plants, for example for the floral homeotic gene *AGAMOUS* (Hong et al., 2003). To uncover the details behind this topic, we would have to clone and analyse expression of the p3A with partial fragments of the CDS fused to reporter gene. Moreover, the GUS expression pattern observed in vascular tissues and cotyledon tips slightly resembles expression patterns of auxin synthesis genes in *Arabidopsis*, for example YUCCAs (Müller-Moulé et al., 2016). The possible relation of eIF3A to auxin was already hinted in rice (Li et al., 2003), although the outcomes of this publication are limited to study as it is written in Chinese original language only. We were also surprised by the numerous foci of GFP-GUS signal in the mature pollen, but it is possible that either the abundance of the protein causes self-aggregates of the GFP-GUS, given that GUS holoenzyme is a tetramer, or the vegetative cell is actively sorting the plant unrelated GFP-GUS into foci. Moreover the GBpart of GUS gene contains an intron, which sometimes increases the levels of transcription. Nevertheless, we confirmed our hypotheses about the general expression pattern of AteIF3A.

AteIF3A is localized in cytoplasm and enriches around the nucleus in pollen

Next, we analysed the subcellular localization of the eIF3A protein. We hypothesized that eIF3A, as translation initiation factor, would localize to cytoplasm, where all eukaryotic translation is situated. We used the eIF3A N- and C- terminal fusions with GFP driven by different promoters to analyse the subcellular localization. From our observations, we can conclude that: 1) eIF3A fusions are strictly localized to cytoplasm; 2) eIF3A fusions are localized to cytoplasm of the vegetative cell with cloudy distribution in pollen and are enriched around the vegetative nucleus and sperm cells; 3) GFP-eIF3A driven by native promoter forms a small number of foci around the nucleus in the root meristem cells; 4) full protein product is present in inflorescences. Altogether, our hypothesis was confirmed, as no tissue observed localized either of the eIF3A GFP fusions to nucleus, vacuoles or other cell compartments. This is in agreement with numerous studies that found eIF3 bound to ribosomes (For example Zeman et al., 2019). Moreover, the pollen signal showing localization around vegetative cell nucleus and sperm cells is common for other eIF3 subunits investigated in our lab (unpublished) and was not caused by the addition of the GFP on either terminus of the eIF3A. It is also not caused by the GFP alone, as we observed similar distribution with Lat52::mCherry-eIF3A (data not shown). It is documented in the literature, that eIF3E changes its localization to nucleus when driven by strong promoters (Yahalom et al., 2001), but we have not observed any changes in the localization within the pollen when eIF3A was driven by strong Lat52 pollen specific promoter. What are the GFP signal foci found forming around root cells nuclei we cannot explain, but it is possible that it is an artefact caused by the propidium iodide staining that stresses the seedlings. This phenomena was already observed with eIF3i subunit in fission yeast (Dunand-Sauthier et al., 2002). Interestingly, we observed a filamentous organisation of the eIF3A-GFP signal in the pollen tubes. Such filamentous organisation is typical for F-actin filaments in the pollen tube. This would be in agreement that eIF3A was found in Tobacco pollen cytoskeleton-associated RNP complexes (Honys et al., 2009). If this is the case only for eIF3A-GFP fusion and not GFP-eIF3A was not determined, but it is possible that the microscopy setup we used in this analysis was insufficient to observe such localization in other transgenic lines as well. It would be interesting to clone the whole genomic region of AteIF3A that would also include introns to observe whether the expression would increase as well. There are evidences of eIF3A attachment to the cytoskeleton, however they are described in mammals, where the eIF3A possess some additional functions due to its prolonged C-terminal

end. We will investigate this phenomena using better confocal microscopy setup, the scanning confocal microscope with Airyscan detector.

Analysis of T-DNA insertion lines brought one still unclear candidate line

Mutation analysis is probably the most important part of gene characterization using reverse genetic approach, where well-characterized knockout and knockdown mutants are crucial for uncovering the gene function. We hypothesized that *AteIF3A* is a single copy house-keeping gene that is essential for cell viability. This hypothesis was supported by the literature, where eIF3a subunits is essential for example in *S. cerevisiae, H. sapiens and N. crassa* (Valášek et al., 2017). Therefore, any mutations affecting its expression and/or disrupting its protein product would lead to observable phenotype, at least at the level of haploid pollen grains. For the purpose of eIF3A mutant characterization, we screened four T-DNA insertion lines with following results. T-DNA insertion lines SAIL 124 and SAIL 227 were not mapped inside the At*eIF3A* gene, are viable in homozygous state and have no phenotypical differences from the Col-0 plants. SAIL 545 insertion was located to the very end of the At*eIF3A* gene, disrupting the sequence so that last 36 amino acids of the protein product would be presumably affected, however, homozygous plants did not differ from Col-0 plants in any phenotypical analysis pertained. SALK 762 was mapped to the At*eIF3A* gene exon 13, where its insertion interferes with the presumable eIF3i binding platform. We only observed plants in heterozygous state, which was supported by the genotyping PCR reactions and segregation of the genotype in heterozygous and wild type plants in all progeny tested (data not shown). In most but not all heterozygous plants, we observed significantly increased number of aborted/collapsed pollen grains and significantly increased number of aborted/unfertilized ovules that impaired proper seed formation and development. However, not all heterozygous plants had this phenotype. This suggested on background mutations present in the SALK 762 line that co-segregated with the *AteIF3A* T-DNA insertion. Against this claim were the facts that we haven´t observed any plant that would segregate in the last category of two-traits Mendelian inheritance, having wild type genotype and the pollen/seeds defects. To observe the segregation and/or to clean the SALK 762 line, we reciprocally crossed the heterozygous plants with phenotype (SALK 762 HET+) with Col-0 plants in two consecutive crosses and analysed the progeny, SALK 762* and SALK 762**. In the crossing progeny, no phenotype segregation was detected and all heterozygous plants showed the described phenotype. We observed that the insertion is transmitting through both male and female germline, although frequency of the transmission

through male germline was significantly reduced. We concluded that the phenotype observed is linked to the phenotype and properly analysed the progeny of the second outcross. Nevertheless, in subsequent experiments with limited number of the SALK 762** progeny, we again observed the one-way orientated segregation, again with no plant showing wild type genotype and the observed pollen phenotype (data not shown) and no homozygous plant in the *AteIF3A* insertion. All this leads to suggestions that the mutation of eIF3A has reduced penetrance and the line will be investigated further in this aspect and on larger number of individuals. Also, our results from qPCR of all lines investigated were truly far from satisfying, we encountered great variability between the samples, including Col-0 control. This is probably caused by the methodological approach itself. While the qPCR itself is highly sensitive to any suboptimal sample treatment, we also think that non-optimal sample material was chosen, as the inflorescences were not identical when it comes to their robustness and the number of developing buds. Moreover, inflorescences are a mixture of sporophytic and gametophytic tissues, therefore the aborted pollen phenotype might have increased the variability even more. We tried to reduce this variability from the beginning by combining 21 main inflorescences of one genotypic background in three independent replica samples, but the range of relative expression was still too big to assume we got to biologically relevant measurements. We are not sure if we re-visit the SAIL lines in this aspect after proper optimization, but we will definitely continue to struggle with the SALK 762. As a last comment on the topic of T-DNA insertion lines, it is interesting how limited is the number of lines available in the databases that target the *AteIF3A* gene. The SALK 762 insertion present in the last ¼ of the coding sequence is the furthest CDS insertion available to the 5′ end of relatively large gene. Whether it is due to the lethality of such insertions generated in these collections is of course not known.

Modulation of eIF3A expression requires different approach

As the T-DNA analysis produced insufficient outcomes, we turned to different methodology to modulate eIF3A expression in *Arabidopsis.* We knew mainly from mammals, that modulation of eIF3 subunits expression is commonly found in many tumorous tissues (Gomes-Duarte et al., 2018; Yin et al., 2018), but had no clue what could be the outcome of enhanced expression in plants. We were able to express the eIF3A GFP fusions in the pollen under strong Lat52 promoter, but we observed similar or only slightly increased GFP signal in the pollen, with no difference in eIF3A protein localization in the cytoplasm of vegetative cell and no observable defect on pollen. We also tried overexpression by driving eIF3A fusions under

35Sdouble and TMV Ω 5′ leader, but we couldn´t germinate seeds that we selected as transformants on any media composition we tried (data not shown). The sporophytic overexpression would require modifications and we will probably test some inducible systems that would allow seeds to germinate before we increase the eIF3A abundance. It is also possible that cloning of the AteIF3A CDS without introns might lead to post-transcriptional gene silencing that could affect the endogenous expression as well. It would be interesting to clone the whole genomic region of AteIF3A that would also include introns to observe whether the expression would increase as well. If such induced eIF3A levels would lead to growth arrest in seedlings, we could conclude that the germination arrest observed in 35Sdouble driven expression is also caused directly by the higher eIF3A abundance. Moreover, we started with knocking down the AteIF3A mRNA levels using RNAi interference, but the time constraints not allowed us to complete this analysis before this thesis was finished. Moreover, in following doctoral project, we would like to target the *AteIF3A* gene coding sequence by Crispr/Cas9 in order to edit the genomic sequence more at the 5′ site and in the described domains and regions of interest.

AteIF3A protein isolation encountered insufficient expression

We clearly failed in the effort to isolate the AteIF3A protein fusions with epitope tags driven by Lat52 promoter. Whether it was due to insufficient expression of the recombinant protein is unclear but probable, as we have not detected the transcript with qPCR. Most probably, we will generate transgenic lines expressing the C-terminally-tagged epitope fusions of eIF3A under the native promoter, which would allow us to verify the construct in the sporophyte. Nevertheless, we were able to verify the size and presence of eIF3A fusions with GFP, using anti-GFP antibody. However, we detected several bands that might be degradation products of the GFP fusion protein. Such degradation of eIF3a subunit was described in the literature as well (For example Milburn et al., 1990; Asano et al., 1997). We will progress further in this aspect and test whether these constructs could be integrated into the eIF3 complex (Immunoprecipitation) and whether such complexes are able to bind the 40S ribosomal subunit in plants (Polysome profiling).

Stable *N. tabacum* transgenic lines were produced for further analyses

Alongside the continuous work with *Arabidopsis*, we produced stable transgenic lines of *Nicotiana tabacum* expressing fluorescent markers and Lat52::eIF3A-GFP. We successfully

regenerated transgenic plants and verified the expression of the constructs with fluorescent microscopy. However, the enormous time demands of this procedure allowed us only to harvest the seeds from T_0 plants and start growing the next generation before this thesis was finished. We will continue to work with these lines after we obtain homozygous lines for at least two independent lines. The subsequent work will include the protein product verification and downstream experiments that are similar to the ones that are described in the previous subchapter with eIF3A in *Arabidopsis*.

Future perspectives

- Analysis of other plant subunits of the eIF3 complex using the *in silico* pipeline determined in this work
- Analysis of the possible downstream cis-regulatory elements in the *AteIF3A* gene
- Better characterization of the AteIF3A subcellular localization in pollen tubes, co-localization analysis with cytoskeleton markers and with other eIF3 subunits
- Modulation of *AteIF3A* expression using the RNA interference approach
- Generation of *AteIF3A* knockout mutant lines using the Crispr/Cas9 editing system
- Sporophytic overexpression using the inducible systems
- Precipitation of the whole eIF3 complex using the AteIF3A epitope tag fusions driven by native promoter
- Characterization of the effects of heterologous expression of *Arabidopsis* eIF3A subunit in stable *Nicotiana* transformants
- Analysis of the heterologous expression of non-plant eIF3a in *Arabidopsis*
- Characterization of the ribosome-binding ability of the AteIF3A GFP fusions using polysome profiling
- Characterization of the mRNA-binding specifity of the AteIF3A using CLIP-seq analysis

Conclusions

The aims of this thesis were to characterize the single copy eIF3a subunit in plants, *using Arabidopsis thaliana* as the main model organism. As literature contains only hints about eIF3a in plants, this investigation began with *in silico* analysis of the eIF3a subunit in several plant species that was based on known information from other eukaryotes. Using methods of computational biology, we confirmed that plants contain well-conserved eIF3A subunits that share protein homology, complexity and domain organisation with their non-plant counterparts. To the contrary, we described novel regions in the PCI domain that were, in our *in silico* analysis, exclusively plan-specific.

In *Arabidopsis thaliana*, we described the expression of *AteIF3A* gene and subcellular localization of its protein product in the sporophyte and in the male gametophyte. We confirmed our hypotheses that *AteIF3A* is expressed in proliferating tissues that need high levels of translation and that the protein product AteIF3A is strictly cytoplasmic and enriches around vegetative nucleus in pollen. Moreover, we produced stable transgenic lines of *N. tabacum* expressing heterologous eIF3A from *Arabidopsis.*

During the characterization of the mutant, we encountered biologically interesting phenotype in the heterozygous SALK line plants, while no phenotype was observed in other lines investigated, as the insertion were positioned at the very end of the *AteIF3A* gene or completely outside of it. However, due to probable background mutation or decreased penetrance of this mutation of the only T-DNA insertion line showing phenotype, we have not progressed further due to efforts in cleaning the line. Lastly, we encountered difficulties in efforts to modulate the *AteIF3A* expression and to isolate the protein with desired epitope tag. Altogether, this work started with characterization of a gene from scratch, producing sufficient answers for several initial hypotheses, while leaving many opened questions that we hope to answer during the following doctoral project, for which this work serves as a good basis.

Publications and conferences

Impacted publications

Raabe, K., Honys, D., Michailidis, C., 2019. The role of eukaryotic initiation factor 3 in plant translation regulation. Plant Physiol. Biochem. 145.
https://doi.org/10.1016/j.plaphy.2019.10.015

Conferences

Results of the bioinformatical survey of eIF3A were presented at international conference Plant Biology CS, 2019 in a form of a poster during student days and the general conference, accompanied with 5 min „Flash talk“ oral presentation at lecture session of the general conference.

Awards

The poster presented at international conference Plant Biology CS, 2019 was selected by the audience for a Company of Biologists Young Researcher Award for best poster of the general conference.

References

Aitken CE, Beznosková P, Vlčkova V, Chiu W-L, Zhou F, Valášek LS, Hinnebusch AG, Lorsch JR (2016) Eukaryotic translation initiation factor 3 plays distinct roles at the mRNA entry and exit channels of the ribosomal preinitiation complex. Elife. doi: 10.7554/eLife.20934

Asano K, Clayton J, Shalev A, Hinnebusch AG (2000) A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA(Met) is an important translation initiation intermediate in vivo. Genes Dev **14**: 2534–46

Asano K, Kinzy TG, Merrick WC, Hershey JWB (1997) Conservation and Diversity of Eukaryotic Translation Initiation Factor eIF3. J Biol Chem **272**: 1101–1109

Asano K, Phan L, Anderson J, Hinnebusch AG (1998) Complex Formation by All Five Homologues of Mammalian Translation Initiation Factor 3 Subunits from Yeast Saccharomyces cerevisiae. J Biol Chem **273**: 18573–18585

Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E (2015) The arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. genesis **53**: 474–485

Boex-Fontvieille E, Daventure M, Jossier M, Zivy M, Hodges M, Tcherkez G (2013) Photosynthetic control of Arabidopsis leaf cytoplasmic translation initiation by protein phosphorylation. PLoS One **8**: e70692

Browning KS, Bailey-Serres J (2015) Mechanism of cytoplasmic mRNA translation. Arab B **13**: e0176

Browning KS, Gallie DR, Hershey JWB, Hinnebusch AG, Maitra U, Merrick WC, Norbury C (2001) Unified nomenclature for the subunits of eukaryotic initiation factor 3. Trends Biochem Sci **26**: 284

Burks EA, Bezerra PP, Le H, Gallie DR, Browning KS (2001) Plant initiation factor 3 subunit composition resembles mammalian initiation factor 3 and has a novel subunit. J Biol Chem **276**: 2122–31

Cate JHD (2017) Human eIF3: from "blobology" to biological insight. Philos Trans R Soc Lond B Biol Sci **372**: 20160176

Chiu W-L, Wagner S, Herrmannová A, Burela L, Zhang F, Saini AK, Valášek L, Hinnebusch AG (2010) The C-terminal region of eukaryotic translation initiation factor 3a (eIF3a) promotes mRNA recruitment, scanning, and, together with eIF3j and the eIF3b RNA recognition motif, selection of AUG start codons. Mol Cell Biol **30**: 4415–34

Crane R, Craig R, Murray R, Dunand-Sauthier I, Humphrey T, Norbury C (2000) A fission yeast homolog of Int-6, the mammalian oncoprotein and eIF3 subunit, induces drug resistance when overexpressed. Mol Biol Cell. doi: 10.1091/mbc.11.11.3993

Dennis MD, Browning KS (2009) Differential phosphorylation of plant translation initiation factors by Arabidopsis thaliana CK2 holoenzymes. J Biol Chem **284**: 20602–14

Dong Z, Qi J, Peng H, Liu J, Zhang J-T (2013) Spectrin Domain of Eukaryotic Initiation Factor 3a Is the Docking Site for Formation of the a:b:i:g Subcomplex. J Biol Chem **288**: 27951

Dunand-Sauthier I, Walker C, Wilkinson C, Gordon C, Crane R, Norbury C, Humphrey T (2002) Sum1, a Component of the Fission Yeast eIF3 Translation Initiation Complex, Is Rapidly Relocalized During Environmental Stress and Interacts with Components of the 26S Proteasome. Mol Biol Cell **13**: 1626–1640

Eliseev B, Yeramala L, Leitner A, Karuppasamy M, Raimondeau E, Huard K, Alkalaeva E, Aebersold R, Schaffitzel C (2018) Structure of a human cap-dependent 48S translation pre-initiation complex. Nucleic Acids Res **46**: 2678–2689

Ellisdon AM, Stewart M (2012) Structural biology of the PCI-protein fold. Bioarchitecture **2**: 118–23

Enchev RI, Schreiber A, Beuron F, Morris EP (2010) Structural insights into the COP9 signalosome and its common architecture with the 26S proteasome lid and eIF3.

Structure **18**: 518–27

Erzberger JP, Stengel F, Pellarin R, Zhang S, Schaefer T, Aylett CHS, Cimermančič P, Boehringer D, Sali A, Aebersold R, et al (2014) Molecular architecture of the 40S□eIF1□eIF3 translation initiation complex. Cell **158**: 1123–1135

des Georges A, Dhote V, Kuhn L, Hellen CUTT, Pestova T V., Frank J, Hashem Y (2015) Structure of mammalian eIF3 in the context of the 43S preinitiation complex. Nature **525**: 491–5

Gomes-Duarte A, Lacerda R, Menezes J, Romão L (2018) eIF3: a factor for human health and disease. RNA Biol **15**: 26–34

Guca E, Hashem Y (2018) Major structural rearrangements of the canonical eukaryotic translation initiation complex. Curr Opin Struct Biol **53**: 151–158

Hafidh S, Fíla J, Honys D (2016) Male gametophyte development and function in angiosperms: a general concept. Plant Reprod **29**: 31–51

Hafidh S, Potěšil D, Müller K, Fíla J, Michailidis C, Herrmannová A, Feciková J, Ischebeck T, Valášek LS, Zdráhal Z, et al (2018) Dynamics of the Pollen Sequestrome Defined by Subcellular Coupled Omics. Plant Physiol **178**: 258–282

Han H, Kong C, Dong H, Zhu S, Zhao Q, Zhai Q, Liang S, Li S, Yang S, Huang B (2015) Molecular characterization and functional analysis of subunit 7 of eukaryotic initiation factor 3 from Eimeria tenella. Exp Parasitol **154**: 118–26

Hashem Y, Frank J (2018) The Jigsaw Puzzle of mRNA Translation Initiation in Eukaryotes: A Decade of Structures Unraveling the Mechanics of the Process. Annu Rev Biophys. doi: 10.1146/annurev-biophys-070816-034034

Hinnebusch AG (2017) Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. Trends Biochem Sci. doi: 10.1016/j.tibs.2017.03.004

Hong RL, Hamaguchi L, Busch MA, Weigel D (2003) Regulatory Elements of the Floral Homeotic Gene AGAMOUS Identified by Phylogenetic Footprinting and Shadowing. Plant Cell **15**: 1296–1309

Honys D (2018) Růst pylové láčky: od opylení k oplození. Nová Bot **1**: 9–11

Honys D, Rěnák D, Feciková J, Jedelský PL, Nebesárová J, Dobrev P, Capková V (2009) Cytoskeleton-associated large RNP complexes in tobacco male gametophyte (EPPs) are associated with ribosomes and are involved in protein synthesis, processing, and localization. J Proteome Res **8**: 2015–31

Kay R, Chan A, Daly M, McPherson J (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science (80-). doi: 10.1126/science.236.4806.1299

Khoshnevis S, Hauer F, Milón P, Stark H, Ficner R (2012) Novel insights into the architecture and protein interaction network of yeast eIF3. RNA **18**: 2306–19

Kim B-H, Cai X, Vaughn JN, von Arnim AG (2007) On the functions of the h subunit of eukaryotic initiation factor 3 in late stages of translation initiation. Genome Biol **8**: R60

Kim D-Y, Scalf M, Smith LM, Vierstra RD (2013) Advanced proteomic analyses yield a deep catalog of ubiquitylation targets in Arabidopsis. Plant Cell **25**: 1523–40

Kim T-H, Kim B-H, Yahalom A, Chamovitz D a, von Arnim AG (2004) Translational regulation via 5' mRNA leader sequences revealed by mutational analysis of the Arabidopsis translation initiation factor subunit eIF3h. Plant Cell **16**: 3341–56

Klepikova A V., Kasianov AS, Gerasimov ES, Logacheva MD, Penin AA (2016) A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. Plant J. doi: 10.1111/tpj.13312

Köster T, Reichel M, Staiger D (2020) CLIP and RNA interactome studies to unravel genome-wide RNA-protein interactions in vivo in Arabidopsis thaliana. Methods **178**: 63–71

Krämer U (2015) Planting molecular functions in an ecological context with Arabidopsis thaliana. Elife. doi: 10.7554/eLife.06100

Li H, Shen SS, Xu ZH, Xue HW (2003) [Isolation and expression pattern analysis of rice auxin-induced eIF3a coding gene]. Shi Yan Sheng Wu Xue Bao **36**: 54–60

Liu Y, Neumann P, Kuhle B, Monecke T, Schell S, Chari A, Ficner R (2014) Translation initiation factor eIF3b contains a nine-bladed β-propeller and interacts with the 40S ribosomal subunit. Structure **22**: 923–30

Llacer JL, Hussain T, Gordiyenko Y, Lorsch J, Hinnebusch A, Ramakrishnan V (2018) The beta propellers of eIF3b and eIF3i relocate together to the ribosomal intersubunit interface during translation initiation. bioRxiv 453688

Meleppattu S, Kamus-Elimeleh D, Zinoviev A, Cohen-Mor S, Orr I, Shapira M (2015) The eIF3 complex of Leishmania-subunit composition and mode of recruitment to different cap-binding complexes. Nucleic Acids Res **43**: 6222–35

Méthot N, Song MS, Sonenberg N (1996) A region rich in aspartic acid, arginine, tyrosine, and glycine (DRYG) mediates eukaryotic initiation factor 4B (eIF4B) self-association and interaction with eIF3. Mol Cell Biol. doi: 10.1128/mcb.16.10.5328

Milburn SC, Duncan RF, Hershey JWB (1990) Immunoblot analysis of the structure of protein synthesis initiation factor eIF3 from HeLa cells. Arch Biochem Biophys. doi: 10.1016/0003-9861(90)90002-G

Müller-Moulé P, Nozue K, Pytlak ML, Palmer CM, Covington MF, Wallace AD, Harmer SL, Maloof JN (2016) YUCCA auxin biosynthetic genes are required for Arabidopsis shade avoidance. PeerJ **4**: e2574

Munzarová V, Pánek J, Gunišová S, Dányi I, Szamecz B, Valášek LS (2011) Translation reinitiation relies on the interaction between eIF3a/TIF32 and progressively folded cis-acting mRNA elements preceding short uORFs. PLoS Genet **7**: e1002137

Peterson R, Slovin JP, Chen C (2010) A simplified method for differential staining of aborted and non-aborted pollen grains. Int J Plant Biol **1**: 13

Phan L, Schoenfeld LW, Valášek L, Nielsen KH, Hinnebusch AG (2001) A subcomplex of three eIF3 subunits binds eIF1 and eIF5 and stimulates ribosome binding of mRNA and tRNAiMet. EMBO J **20**: 2954–2965

Phan L, Zhang X, Asano K, Anderson J, Vornlocher HP, Greenberg JR, Qin J, Hinnebusch AG (1998) Identification of a translation initiation factor 3 (eIF3) core complex, conserved in yeast and mammals, that interacts with eIF5. Mol Cell Biol **18**: 4935–46

Pick E, Hofmann K, Glickman MH (2009) PCI complexes: Beyond the proteasome, CSN, and eIF3 Troika. Mol Cell **35**: 260–4

Pincheira R, Chen Q, Huang Z, Zhang JT (2001a) Two subcellular localizations of eIF3 p170 and its interaction with membrane-bound microfilaments: implications for alternative functions of p170. Eur J Cell Biol **80**: 410–8

Pincheira R, Chen Q, Zhang JT (2001b) Identification of a 170-kDa protein over-expressed in lung cancers. Br J Cancer. doi: 10.1054/bjoc.2001.1828

Pisarev A V, Kolupaeva VG, Yusupov MM, Hellen CUT, Pestova T V (2008) Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes. EMBO J **27**: 1609–21

Querol-Audi J, Sun C, Vogan JM, Smith MD, Gu Y, Cate JHD, Nogales E (2013) Architecture of human translation initiation factor 3. Structure **21**: 920–8

Raabe K, Honys D, Michailidis C (2019) The role of eukaryotic initiation factor 3 in plant translation regulation. Plant Physiol Biochem **145**: 75–83

Rezende AM, Assis LA, Nunes EC, da Costa Lima TD, Marchini FK, Freire ER, Reis CRS, de Melo Neto OP (2014) The translation initiation complex eIF3 in trypanosomatids and other pathogenic excavates--identification of conserved and divergent features based on orthologue analysis. BMC Genomics **15**: 1175

Sabelli PA, Burgess SR, Valášek L, Shewry PR (1999) Molecular cloning and characterisation of a maize cDNA for a homologue of the large subunit of the eukaryotic initiation factor 3 (eIF3). Mol Gen Genet **261**: 820–830

Saletta F, Rahmanto YS, Richardson DR (2010) The translational regulator eIF3a: The tricky eIF3 subunit! Biochim Biophys Acta - Rev Cancer **1806**: 275–286

Sarrion-Perdigones A, Falconi EE, Zandalinas SI, Juárez P, Fernández-del-Carmen A, Granell A, Orzaez D (2011) GoldenBraid: An iterative cloning system for standardized assembly of reusable genetic modules. PLoS One. doi: 10.1371/journal.pone.0021622

Scheel H, Hofmann K (2005) Prediction of a common structural scaffold for proteasome lid, COP9-signalosome and eIF3

complexes. BMC Bioinformatics **6**: 71

Servet C, Benhamed M, Latrasse D, Kim W, Delarue M, Zhou D-X (2008) Characterization of a phosphatase 2C protein as an interacting partner of the histone acetyltransferase GCN5 in Arabidopsis. Biochim Biophys Acta **1779**: 376–82

Simonetti A, Brito Querido J, Myasnikov AG, Mancera-Martinez E, Renaud A, Kuhn L, Hashem Y (2016) eIF3 Peripheral Subunits Rearrangement after mRNA Binding and Start-Codon Recognition. Mol Cell **63**: 206–217

Siridechadilok B, Fraser CS, Hall RJ, Doudna JA, Nogales E (2005) Structural roles for human translation factor eIF3 in initiation of protein synthesis. Science **310**: 1513–5

Smith MD, Arake-Tacca L, Nitido A, Montabana E, Park A, Cate JH (2016) Assembly of eIF3 Mediated by Mutually Dependent Subunit Insertion. Structure **24**: 886–96

Sokabe M, Fraser CS, Hershey JWB (2012) The human translation initiation multi-factor complex promotes methionyl-tRNAi binding to the 40S ribosomal subunit. Nucleic Acids Res **40**: 905–13

De Storme N, Keçeli BN, Zamariola L, Angenon G, Geelen D (2016) CENH3-GFP: A visual marker for gametophytic and somatic ploidy determination in Arabidopsis thaliana. BMC Plant Biol. doi: 10.1186/s12870-015-0700-5

Sun C, Todorovic A, Querol-Audí J, Bai Y, Villa N, Snyder M, Ashchyan J, Lewis CS, Hartland A, Gradia S, et al (2011) Functional reconstitution of human eukaryotic translation initiation factor 3 (eIF3). Proc Natl Acad Sci U S A **108**: 20473–8

Szamecz B, Rutkai E, Cuchalova L, Munzarova V, Herrmannova A, Nielsen KH, Burela L, Hinnebusch AG, Valasek L (2008) eIF3a cooperates with sequences 5' of uORF1 to promote resumption of scanning by post-termination ribosomes for reinitiation on GCN4 mRNA. Genes Dev **22**: 2414–2425

Thiébeauld O, Schepetilnikov M, Park H-S, Geldreich A, Kobayashi K, Keller M, Hohn T, Ryabova L a (2009) A new plant protein interacts with eIF3 and 60S to enhance virus-activated translation re-initiation. EMBO J **28**: 3171–84

Twell D, Wing R, Yamaguchi J, McCormick S (1989) Isolation and expression of an anther-specific gene from tomato. MGG Mol Gen Genet. doi: 10.1007/BF02464887

Urquidi Camacho RA, Lokdarshi A, Arnim AG (2020) Translational gene regulation in plants: A green new deal. WIREs RNA. doi: 10.1002/wrna.1597

Valášek L, Trachsel H, Hašek J, Ruis H (1998) Rpg1, the Saccharomyces cerevisiae homologue of the largest subunit of mammalian translation initiation factor 3, is required for translational activity. J Biol Chem. doi: 10.1074/jbc.273.33.21253

Valášek LS (2012) 'Ribozoomin'--translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIFs). Curr Protein Pept Sci **13**: 305–30

Valášek LS, Zeman J, Wagner S, Beznosková P, Pavlíková Z, Mohammad MP, Hronová V, Herrmannová A, Hashem Y, Gunišová S (2017) Survey and summary: Embraced by eIF3: Structural and functional insights into the roles of eIF3 across the translation cycle. Nucleic Acids Res **45**: 10948–10968

Villa N, Do A, Hershey JWB, Fraser CS (2013) Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome. J Biol Chem **288**: 32932–40

Wagner S, Herrmannová A, Hronová V, Gunišová S, Sen ND, Hannan RD, Hinnebusch AG, Shirokikh NE, Preiss T, Valášek LS (2020) Selective Translation Complex Profiling Reveals Staged Initiation and Co-translational Assembly of Initiation Factor Complexes. Mol Cell. doi: 10.1016/j.molcel.2020.06.004

Wagner S, Herrmannová A, Malík R, Peclinovská L, Valášek LS (2014) Functional and biochemical characterization of human eukaryotic translation initiation factor 3 in living cells. Mol Cell Biol **34**: 3041–52

Wagner S, Herrmannová A, Šikrová D, Valášek LS (2016) Human eIF3b and eIF3a serve as the nucleation core for the assembly of eIF3 into two interconnected modules: the yeast-like core and the octamer. Nucleic Acids Res **44**: 10772–10788

Yahalom A, Kim TH, Winter E, Karniol B,

Von Arnim AG, Chamovitz DA (2001) Arabidopsis eIF3e (INT-6) associates with both eIF3c and the COP9 signalosome subunit CSN7. J Biol Chem **276**: 334–340

Yin Y, Long J, Sun Y, Li H, Jiang E, Zeng C, Zhu W (2018) The function and clinical significance of eIF3 in cancer. Gene. doi: 10.1016/j.gene.2018.06.034

Zeman J, Itoh Y, Kukačka Z, Rosůlek M, Kavan D, Kouba T, Jansen ME, Mohammad MP, Novák P, Valášek LS (2019) Binding of eIF3 in complex with eIF5 and eIF1 to the 40S ribosomal subunit is accompanied by dramatic structural changes. Nucleic Acids Res. doi: 10.1093/nar/gkz570