Molecular and functional characterization of the THO/TREX complex in *Arabidopsis thaliana*

DISSERTATION ZUR ERLANGUNG DES
DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT)
DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG

Vorgelegt von

Brian Bitsch Sørensen

Aus Herning, Dänemark

im April 2016

Molecular and functional characterization of the THO/TREX complex in *Arabidopsis thaliana*

DISSERTATION ZUR ERLANGUNG DES
DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT)
DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG

Vorgelegt von

Brian Bitsch Sørensen

Aus Herning, Dänemark

im April 2016

Das Promotionsgesuch wurde eingereicht am: 29.04.2016

Die Arbeit wurde angeleitet von: Prof. Dr. Klaus D. Grasser

Unterschrift:

Brian Bitsch Sørensen

Table of Contents

List of Figures V
List of Tables VII
Abbrevations IX
1 Introduction 1
1.1 The nuclear pore complex as a permeability barrier. 1
1.1.1 The general principle of NPC translocation by NTRs 2
1.1.2 An overview of the different types of RNA transport 3
1.2 mRNA undergoes extensive processing before being exported 5
1.2.1 5' Capping of mRNA helps recruiting mRNA biogenesis factors 6
1.2.2 Splicing is interconnected with the other processing steps 6
1.2.3 Polyadenylation is the last processing step 8
1.2.4 Only transcripts passing mRNA surveillance are exported 8
1.2.5 Packaging of mRNPs affects the choice of mRNA export pathway 9
1.2.6 NXF1 and the THO/TREX complex are required for mRNA export 10
1.2.7 Export adaptors enhance the mRNA binding activity of NXF1 11
1.2.8 Recruitment of the export machinery 13
1.2.9 Transport to the nuclear pore and beyond 14
1.3 Export of mRNAs in plants 15
1.3.1 Identified proteins affecting mRNA export 15
1.3.2 The *Arbidopsis* THO/TREX complex is conserved 16
1.4 Aim of the project 18
2 Material and Methods 19
2.1 Materials 19
2.1.1 Instruments 19
2.1.2 Chemicals and enzymes 19
2.1.3 Oligonucleotides 20
2.1.4 Vectors 21
2.1.5 Antibodies 22
2.1.6 Bacterial and yeast strains 23
2.1.7 T-DNA lines and plant cell culture 23
2.1.8 Antibiotics 23
2.1.9 Software and internet sources 24
2.2 Methods 24
2.2.1 Microbial work 24
2.2.1.1 Cultivation of bacteria 24
2.2.1.2 Production of heat shock competent *E.coli* and *A. tumefaciens* 24

2.2.1.3 Transformation of *E.coli* 25
2.2.1.4 Transformation of *A. tumefaciens* 25
2.2.1.5 Cultivation of yeast cells 25
2.2.1.6 Production of heat shock competent yeast 26
2.2.1.7 Transformation of yeast. 26
2.2.2 Plant work and Cell biological methods. 26
2.2.2.1 Cultivation of *A. thaliana* on soil 26
2.2.2.2 Cultivation of *A. thaliana* on MS plates and filter paper 27
2.2.2.3 Stable transformation of *A. thaliana.* 27
2.2.2.4 Cultivation of *Arabidopsis thaliana landsberg erecta* PSB-D cell culture. 27
2.2.2.5 Transformation of *Arabidopsis thaliana landsberg erecta* PSB-D cell culture. 28
2.2.2.6 Crossing of plants. 29
2.2.2.7 Soil-based Phenotyping 29
2.2.2.8 Phenotypical analysis of roots. 30
2.2.2.9 Isolation of protoplasts from *Arabidopsis* seedlings 30
2.2.2.10 Extracting root nuclei and immunolocalization analysis. 31
2.2.2.11 Whole mount in situ hybridization of *Arabidopsis* seedlings. 31
2.2.2.12 *In vivo* Localization of GFP and RFP fused proteins in *Arabidopsis*. 32
2.2.2.13 Electron microscopy 33
2.2.3 Molecular Biology methods 33
2.2.3.1 Polymerase chain reaction 33
2.2.3.2 Agarose gel electrophoresis 34
2.2.3.3 Extraction of Genomic DNA from Arabidopsis seedlings. 34
2.2.3.4 Extraction of total RNA from Arabidopsis seedlings 34
2.2.3.5 Northern blot analysis of Ta-siRNA and miRNA. 35
2.2.3.6 Construction of cDNA by reverse transcriptase 36
2.2.3.7 RT-PCR 36
2.2.3.8 RT-qPCR 37
2.2.3.9 Construction of plasmids 37
2.2.3.10 Protein extraction 38
2.2.3.11 Protein separation by SDS-page. 39
2.2.3.12 Immunoblotting 39
2.2.3.13 Pulse labelling 40
2.2.3.14 Affinity purification. 40
2.2.3.15 Trypsin digest 41
2.2.3.16 Identification of protein interaction-partners by tandem mass spectrometry. 42
3 Results 43
3.1 Characterization of ALY1, ALY2, ALY3, ALY4 and UAP56. 43
3.1.1 UAP56 interacts with ALY1, ALY2, ALY3 and ALY4 *in Arabidopsis*. 43

3.1.2 GFP fused ALY1, ALY2, ALY3, ALY4 and UAP56 exhibits differential sub nuclear localization **46**
3.1.3 Analysis of T-DNA lines with deficient transcript levels of *ALY1*, *ALY2*, *ALY3* and *ALY4*. **55**
3.1.4 Loss of *ALY1* and *ALY2* transcript in single and double mutants does not lead to developmental phenotypes. **58**
3.1.5 Loss of *ALY1* and *ALY2* transcripts does not lead to mRNA accumulation. **60**
3.1.6 ALY3 and ALY4 are involved in the transition into reproductive phase **61**
3.1.7 The aly3-1 aly4-1 double mutant is accumulating mRNA in the nucleus. **63**
3.2 Characterization of TEX1 ..... **65**
3.2.1 Molecular characterization of the t*ex1-4* T-DNA line. **65**
3.2.2 Loss of the *TEX1* transcript leads to morphological and pleiotropic phenotypes. **65**
3.2.3* GFP fused TEX1 complements the earlier flowering phenotype in *tex1-4 ..... **69**
3.2.4 TEX1 fused to GFP exhibits specific subnuclear localization in root cells. **73**
3.2.5 Loss of the *TEX1* transcript does not lead to mRNA accumulation. **76**
3.2.6 The THO complex and MOS11 can be affinity purified together with UAP56. ..... **76**
3.3 Characterization of TEX1 and MOS11 ..... **83**
3.3.1 TEX1 and MOS11 colocalize, but TEX1 exhibits unique accumulation in speckles. **83**
3.3.2 Analysis of single and double mutants deficient in *TEX1* and *MOS11* transcript levels. **89**
3.3.3 TEX1 and MOS11 exhibit genetic interaction in multiple phenotypes. **92**
3.3.4 TEX1 affects tasiRNA biogenesis independently of MOS11. ..... **96**
3.3.5 Simultaneous loss of the *TEX1* and *MOS11* transcripts leads to severe phenotypes of trichomes and reproductive organs. **98**
3.3.6 A simultaneous loss of *TEX1* and *MOS11* transcripts leads to deficient mRNA export and reduced protein synthesis rates. **105**
4 Discussion **109**
4.1 The core composition of the TREX complex **109**
4.2* Subcellular localization of TREX components in *Arabidopsis thaliana **112**
4.3 The four *Arabidopsis* ALY proteins and potential redundancy **114**
4.4 The TREX interactome **115**
4.5 Function of the TEX1 and the TREX complex **120**
4.6 Function of MOS11 and the interplay with THO/TREX **124**
4.7 TEX1 and MOS11 in plant development **127**
4.8 Outlook **129**
5 Summary **131**
6 References **133**
7 Appendix **155**
7.1 Statistical analysis **155**

7.2 Anova tables. **156**
7.3 RT-qPCR **165**
7.4 Mass spectrometry **167**
8 List of Publications **173**
8.1 Manuscript: *TEX1* of the *Arabidopsis* THO/TREX complex plays a role at different steps of the biogenesis of mRNAs and small RNAs **174**
Acknowledgements **203**

List of Figures

Figure 1. Structure of the nuclear pore complex. 2
Figure 2. Overview of export of different types of RNA. 3
Figure 3. Co-transcriptional coupling of capping, polyadenylation and export. 5
Figure 4. Overview of pre-mRNA splicing. 7
Figure 5. TREX mediated activation of NXF1. 12
Figure 6. Overview of bulk mRNA export in metazoans. 15
Figure 7. Multiple sequence alignment of ALYREF orthologs. 44
Figure 8. All four ALY proteins interact diretly with UAP56. 45
Figure 9. Publicly available expression data based on microarray experiments. 47
Figure 10. UAP56 exhibit specific sub-nuclear localization in the nucleoplasm 47
Figure 11. GFP-NLS exhibit no specific subnuclear in the root. 48
Figure 12. UAP56 is mainly localized in nucleoplasm 49
Figure 13. ALY1-GFP is mainly localized in the nucleoplasm 51
Figure 14. Aly2 exhibit same localization pattern as Aly1. 52
Figure 15. ALY3-GFP is found in both nucleoplasm and nucleolus 53
Figure 16. ALY4-GFP exhibits subnuclear localization comparable to that of ALY3 54
Figure 17. Molecular characterization of *aly1-1*, *aly2-1*, *aly3-1* and *aly4-1* T-DNA lines. 57
Figure 18. Single and double mutants of ALY1 and ALY2 shows no obvious phenotypes. 59
Figure 19. Single and double mutants of *ALY1* and *ALY2* are not deficient in mRNA export. 61
Figure 20. The two mutants *aly3-1* and *aly3-1 aly4-1* exhibit minor phenotypes. 62
Figure 21. The double mutant aly3-1 aly4-1 exhibit a potential mRNA export block. 64
Figure 22. Molecular characterization of *tex1-4*. 65
Figure 23. The *tex1-4* mutant exhibits pleiotropic phenotypes. 68
Figure 24. p*TEX1*::*TEX1-GFP* can restore expression of the TEX1 transcript in*tex1-4.* 70
Figure 25. *pTEX1::TEX1-GFP* can rescue the *tex1-4* phenotype. 72
Figure 26. The TEX1-GFP fusion protein is expressed in the transformed lines. 73
Figure 27. TEX1 fused to GFP exhibit a broad expression profile. 74
Figure 28. TEX1-GFP localizes mainly in the nucleoplasm, with a faint cytoplasmic signal. 75
Figure 29. Bulk mRNA export is not affected in *tex1-4*. 76
Figure 30. Protein interaction partners of TEX1, UAP56 and MOS11. 79
Figure 31. An enrichment of splicing factors is seen in the TEX1 affinity purification. 80
Figure 32. MOS11 fused to RFP is mainly found in the nucleoplasm. 84
Figure 33. TEX1 and MOS11 exhibit a high degree of colocalization in root nuclei. 85
Figure 34. TEX1 and MOS11 exhibit a high degree of colocalization in leaf nuclei. 86
Figure 35. TEX1-GFP localizes in Speckle-like structures. 88
Figure 36. Molecular characterization of *tex1-4 mos11-2* double mutant. 89

Figure 37. *TEX1* and *MOS11* transcript levels analysed by RT-qPCR 92
Figure 38. The double mutant tex1-4 mos11-2 shows strong phenotypes. 95
Figure 39. Expression of *tasiRNA* is only affected in *tex1-4*,. 97
Figure 40. Single and double mutants of TEX1 and MOS11 show an increased number of branchpoints per trichome. 100
Figure 41. The double mutant *tex1-4 mos11-2* exhibit altered inflorescence morphology 101
Figure 42. The double mutant *tex1-4 mos11-2* exhibits flower meristem phenotypes. 103
Figure 43. Morphology of reproductive organs are affected in *tex1-4 mos11-2*. 105
Figure 44. MOS11 is required for bulk mRNA export, whereas TEX1 is required for efficient protein synthesis. 107
Figure 45 Three-dimensional structure of THO 121
Figure 46. The work flow followed for the two-way ANOVA 156
Figure 47. Linear regression analysis of the primers used for RT-qPCR in this analysis. 166

List of Tables

Table 1. *Arabidopsis* protein orthologs of the human THO/TREX complex and cofactors. 17
Table 2. Instruments used in this study 19
Table 3. Oligonucleotides used in this study.....20
Table 4. Modified oligonucleotides used in this study 21
Table 5. List of vectors used in this study 22
Table 6. Overview of antibodies used in this study. 22
Table 7. Different microbial strains used in this work. 23
Table 8. Overview of T-DNA lines used in this study. 23
Table 9. Antibiotics used in this study. 23
Table 10. List of Software and internet sources used in the creation and analysis of the data presented in this study. 24
Table 11. PCR cycle programs for Taq and KAPA HiFi polymerases used in this thesis.....33
Table 12. Reagents used for Taq and KAPA HiFi PCR in this study. 33
Table 13. Aly1/Aly2 and ALY3/ALY4 show high protein sequence similarity. 43
Table 14 An enrichment of splicing factors is seen in the TEX1 affinity purification. 79
Table 15. Protein interaction partners of TEX1, UAP56 and MOS11 involved in RNA biogenesis identified by LC-MS-MS. 80
Table 16. Colocalization analysis of root and leaf nuclei. 87
Table 17. Trichome branches of wild type, *tex1-4*, *mos11-2* and *tex1-4 mos11-2* plants 100
Table 18. Subcellular localization overview of putative TREX components.....112
Table 19. TEX1 and MOS11 exhibit different kinds of genetic interaction patterns in regards to different phenotypes and gene expression levels. 126
Table 20. Anova: Bolting time analysis of *aly1-1 aly2-1* 156
Table 21. Anova: mRNA export deficiency assay.....156
Table 22. Anova: Bolting time.....157
Table 23. Anova: mRNA export deficiency assay.....157
Table 24. Anova: Bolting time- long day conditions 157
Table 25. Anova: Bolting time- short day conditions.....158
Table 26. Anova: Relative expression of *TEX1* (P29+P30).....158
Table 27. Anova: Relative expression of *TEX1* (P31+P32). 158
Table 28. Anova: Relative expression of *TEX1* (P33+P34).....159
Table 29. Anova: Relative expression of *MOS11* (P35+P36) 159
Table 30. Anova: Relative expression of *MOS11* (P37+P38) 159
Table 31. Anova: Relative expression of *MOS11* (P39+P40). 159
Table 32. Anova: Bolting time- long day conditions 160
Table 33. Anova: Relative expression of *FLC* 160
Table 34. Anova: Number of primary inflorescences. 160

Table 35. Anova: Lateral root density in the Branching zone **160**
Table 36. Anova: Root length **161**
Table 37. Anova: Relative expression of *TAS1* transcript **161**
Table 38. Anova: Relative expression of *TAS3* transcript **161**
Table 39. Anova: Relative expression of *miR173* transcript **161**
Table 40. Anova: Relative expression of *miR390* transcript **162**
Table 41. Anova: Relative expression of *ARF3* transcript **162**
Table 42. Anova: Relative expression of *ARF4* transcript **162**
Table 43. Anova: Trichomes with one branch point **162**
Table 44. Anova: Trichomes with two branch points **163**
Table 45. Anova: Trichomes with three branch points **163**
Table 46. Anova: Trichomes with four branch points **163**
Table 47. Anova: Trichomes with five branch points **163**
Table 48. Anova: Trichomes with seven branch point **164**
Table 49. Anova: Percentage of defective embryos **164**
Table 50. Anova: Total number of normal and non-developed ovules **164**
Table 51. Anova: mRNA export deficiency assay **164**
Table 52. fragment length and efficiency of the primers used in RT-qPCR **166**
Table 53. RT-qPCR: The calibrated normalized relative quantities (CQNR) **166**
Table 54. TEX1 affinity purification. **167**
Table 55. UAP56 affinity purification **171**
Table 56. MOS11 affinity purification. **172**

Abbrevations

Ade	Adenine
ALYREF	Ally of AML-1 and LEF1/ RNA Export Factor
ANOVA	Analysis of variance
bp	base pair
CDS	Coding sequence
C_t	Treshold cycle value
CBC	Cap binding complex
CBP20	Cap binding protein 20
CBP80	Cap binding protein 80
CFIm	Clevage factor I
CFIIm	Clevage factor II
CHTOP	Chromatin Target Of PRMT1
CIP29	Cytokine-induced protein of 29 kDa
CLSM	Confacal laser scanning microscopy
CPSF	Clevage and polyadenylation specificity factor
CSTF	Clevage stimulation factor
CRM1	Chromosomal maintenance 1/Exportin1
CTD	C-terminal domain
CTR	C-terminal repeat region of SPT5
DAPI	4’,6-Diamidin-2-phenylindol
DAS	Days after stratification
DDO	Double drop out medium
DDX39	DEAD box polypeptide 39
DNA	Deoxyribonucleic acid
EF1-α	Translation elongation factor 1 alpha
FG-NUPs	Phenylalanine–glycine repeat containing nucleoporins
FLC	Flowering locus C
GFP	Green fluorescent protein
His	Histidine
hnRNP	Heterogenous nuclear ribonucleoproteins
HPR1	Hyper recombinant 1
Leu	Leucine

LRR	Leucine rich Repeat
LUZP4	Leucine zipper protein 4
M7G	7meGpppA cap on mRNA
MEX67	Messenger RNA export factor of 67 kDA molecular weight
miRNA	MicroRNA
MOS11	Modifier of SNC1, 11
mRNA	Messenger
mRNP	Messenger ribonucleoprotein
MS	Murashige-Skoog
MSA	Multiple Sequence Alignment
MTR2	mRNA transport defective 2
NAB2	Nuclear polyadenylated RNA-binding 2
NASC	Nottingham *Arabidopsis* Stock Center
N/C ratio	Ratio of nucleoplasmic to cytoplasmic mRNA populations
NES	Nuclear export signal
NMD	Nonsense mediated decay
NLS	Nuclear localization signal
NPC	Nuclear pore complex
NPL3	Nuclear protein localization 3
NTR	Nuclear transport receptor
NTR2L	Nuclear export factor 2- like
NXF1	Nuclear RNA export factor 1
NXT1	NTF2-related export protein 1
NUPs	Nucleoporin proteins
PAS	Poly-A site
PHAX	Phosphorylated adaptor of RNA export
QDO	Quadruple drop out medium
RBD	RNA binding domain
RFP	Red fluorescent protein
RNA	Ribonucleic acid
RNAPII	RNA polymerase II
RRM	RNA recognition motif
SAC3	Suppressor of actin 3

Ser2	Serine 2 phosporylated CTD
Ser5	Serine 2 phosporylated CTD
snRNA	Small nuclear RNA
snRNP	Small ribonucleo proteins
SR	Serine rich
SUB2	Suppressor of Brr1-1, 2
tasiRNA	Trans-acting siRNA
ter	Terminator
TEX1	Trex component 1
THO	A suppressor of the transcriptional defects of HPR1 mutants by overexpression
THP1	THO2/HPR1 phenotype 1
TREX	Transcription-export complex
TREX2	Transcription-export complex 2
tRNA	Transfer RNA
Trp	Tryptophan
UAP56	U2AF65 associated protein
UIF	UAP56-interacting factor
UBA	Ubiquitin associated domain
Y2H	Yeast-two-Hybrid
YRA1	Yeast RNA annealing protein 1

1 Introduction

1.1 The nuclear pore complex as a permeability barrier.

In eukaryotes, the nucleoplasm and cytoplasm are divided by a double-membrane nuclear envelope. The presence of the membrane leads to a physical barrier separating processes like nuclear transcription and cytoplasmic translation of proteins (Schmidt & Görlich, 2015). To facilitate the exchange of macromolecules, the cytoplasm and nucleoplasm are connected by the nuclear pore complexes (NPC) (Watson, 1954; Stevens & Swift, 1966). The NPC is a multiprotein complex embedded in the nuclear envelope. The NPC is estimated to be approximately 50 and 125 MDa in *S. cerevisiae* (yeast) and vertebrates, respectively (Reichelt et al., 1990; Rout et al., 2000). The NPC consists of a core scaffold, a central channel, the transmembrane regions, the nuclear basket and protruding filaments on the cytoplasmic side (Figure 1) (Grünwald et al., 2011; Alber et al., 2007; Terry & Wente, 2009). Metabolites, ions and molecules smaller than approximately 40 kDa can passively diffuse through the NPC. The bigger the molecule the slower the passive diffusion will be (Ribbeck & Görlich, 2001; Mohr et al., 2009). Larger macromolecules like larger proteins, mRNPs and ribosome subunits typically cannot diffuse through the NPC. Macromolecules are helped through the nuclear pore. This translocation is facilitated by nuclear transport receptors (NTRs) (reviewed in Görlich & Kutay, 1999; Weis, 2003; Kabachinski & Schwartz, 2015; Sloan et al., 2015). The NPC is assembled from multiple copies of approximately 30 different nucleoporin proteins (NUPs) in yeast, vertebrates and plants (Tamura et al., 2010; Cronshaw et al., 2002; Rout et al., 2000). The NUPs can be divided into anchoring integral membrane NUPs, structural NUPs and phenylalanine-glycine repeat containing NUPs (FG-NUPs). The ability to allow diffusion of low weight molecules through the NPC, and retain bigger molecules is believed to be regulated by the FG-NUPs (Terry & Wente, 2009). A number of models (see Schmidt & Görlich, 2015 for review) have been proposed to explain how the FG-NUPs form the transport barrier. One model is the selective-phase model (Ribbeck & Görlich, 2001). This model assumes that the central channel is lined with FG-NUPs that extend their FG-repeats into the middle of the channel. The high local concentration of the FG-repeats generates a hydrogel, in which FG repeats binds cohesively to form a sieve with a mesh size of approximately 5nm. Macromolecules bigger than 5nm (~40 kDa) are translocated by an NTR, that binds

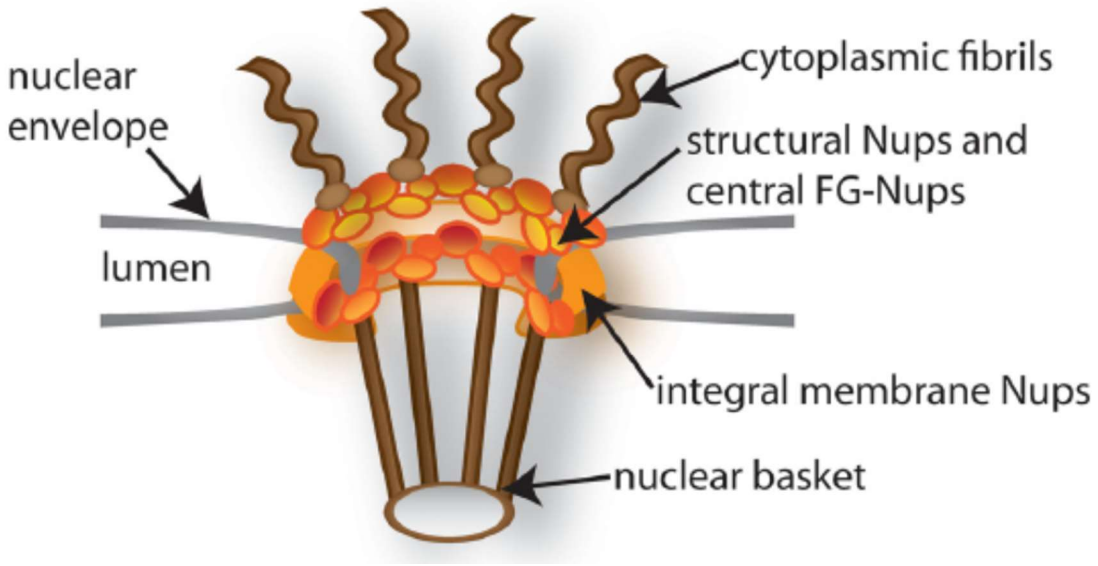


Figure 1. Structure of the nuclear pore complex. The structural NUPs and the membrane NUPs make up the NPC scaffold. The central channel is coated with FG-NUPs, which are suspected to give the NPC its selectivity. (Terry & Wente, 2009).

to the FG repeats thereby locally interrupting the inter-FG repeat interactions. The FG repeat interaction are re-established behind the NTR-cargo complex to maintain the transport barrier (reviewed in Kabachinski & Schwartz, 2015; Schmidt & Görlich, 2015). Even though the selective-phase model does not explain everything observed, it has been supported by stronger evidence than the other models (Hülsmann et al., 2012; Labokha et al., 2013).

1.1.1 The general principle of NPC translocation by NTRs

There are different nuclear transport pathways, each transporting specific macromolecules in or out of the nucleoplasm (Stewart, 2007). Most NTRs belong to the importin β family (Karyopherins). They can be divided in importins and exportins depending on the direction of the transport relative to the nucleus. Karyopherins bind to the FG NUPs and to the GTPase Ran (Wente & Rout, 2010). NTRs are generally recruited to their cargo molecules by interacting with a specific nuclear export signal (NES) or nuclear localization signals (NLS). The signals can be short peptides, nucleotide sequences or structural features of the substrate (Sloan et al., 2015). The directionality of the transport for most macromolecules is determined by a gradient of RanGTP (Wente & Rout, 2010). In the nucleoplasm, Ran is kept in the GTP-bound state by the Ran guanine nucleotide exchange factor (RanGEF; RCC1/PRP20 in *Homo sapiens* (Humans) and yeast respectively). In the cytoplasm Ran is kept in the GDP bound state by the Ran GTPase activating protein (RanGAP; RANGAP1/Rna1

in humans and yeast respectively) that activate the intrinsic GTPase activity of Ran (Sloan et al., 2015; Stewart, 2007). Karyopherins of the importin β family bind RanGTP on the nucleoplasmic side and are translocated to the cytoplasmic side, due to the lower levels of RanGTP there. On the cytoplasmic side, RanGTP is converted to RanGDP that leads to disassociation of Ran and the Karyopherin. The NTF2 protein binds to RanGDP and translocates it to the nucleoplasmic side, where the RanGDP levels are lower (Ribbeck, 1998). Exportins bind their Cargo in the nucleus when associated with RanGTP. The cargo is released upon hydrolysis of RanGTP to RanGDP in the cytoplasm. For importins association with their cargo and RanGDP, leads to nuclear translocation. The cargo is released upon binding of RanGTP (Sloan et al., 2015). This basic principle of Karyopherins coupled with asymmetrical RanGTP distribution between nucleus and cytoplasm are utilized by most protein and Ribonucleic acid (RNA) populations for translocation through the NPC barrier. (Terry & Wente, 2009).

1.1.2 An overview of the different types of RNA transport

Different types of RNA are utilizing different NTRs to translocate through the NPC (Figure 2). The transfer RNA (tRNA) and microRNA (miRNA) molecules are bound

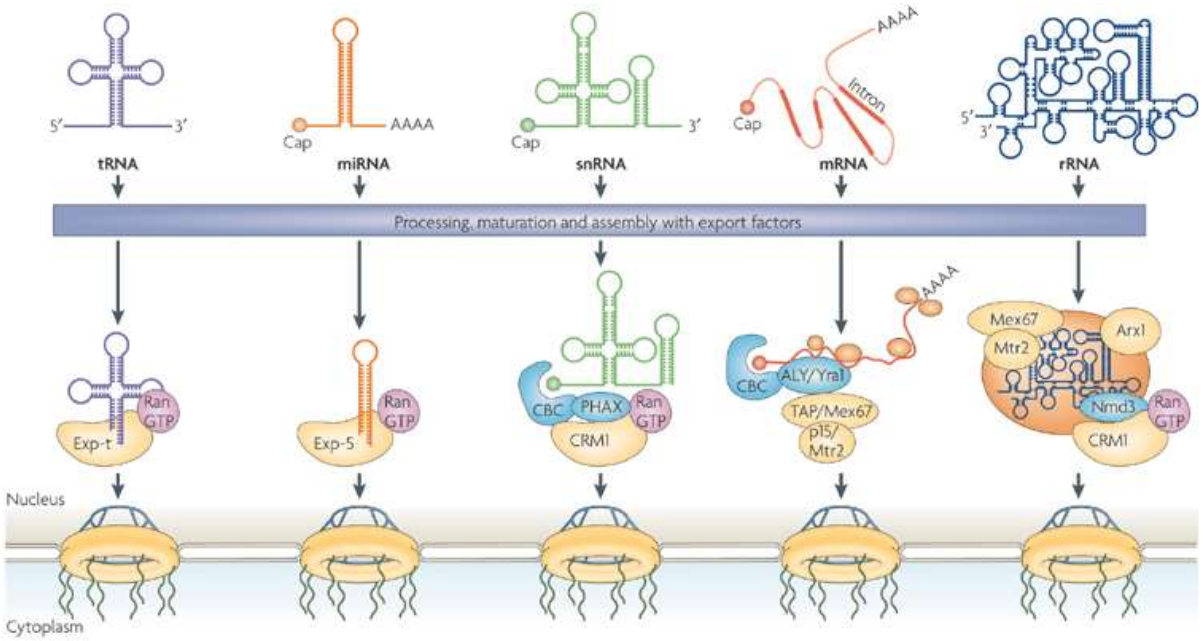


Figure 2. Overview of export of different types of RNA. RNAs like tRNA, miRNA and snRNA are translocated by RanGTP dependent karyopherin NTRS (Exportin-t, Exportin-5 and CRM1). The mRNA is translocated by a non-karyopherin NTR, TAP/MEX67. For rRNA a combination of nuclear transport receptors are used. (Köhler & Hurt, 2007).

directly by their Karyopherins Exportin-t/LOS1 and Exportin-5, respectively, and translocated (Arts et al., 1998; Hellmuth et al., 1998; Bohnsack et al., 2004). The two exportins utilize the RanGTP pathway for export. Small nuclear RNA (snRNAs) are exported to the cytoplasm to undergo maturation. The export is mediated by association of the cap-binding complex (CBC) and the Phosphorylated adaptor of RNA export (PHAX) with the snRNA. These two proteins then mediate the interaction between the pre-snRNA and the RanGTP bound form of the general protein export NTR, Chromosomal maintenance (CRM1) / Exportin1. Binding to CRM1 leads to export (Sloan et al., 2015). Export of bulk messenger RNA (mRNA) is mechanistically different from the export of the other RNAs, as a non-karyopherin NTR is utilized that does not directly depend on the RanGTP-RanGDP gradient. The mRNA is also not exported alone but is a part of a large ribonucleoprotein complex (mRNP) (Köhler & Hurt, 2007). The mRNP proteome in HEK and HeLa cells consists of ~ 800 proteins that have been found to associate with mRNA, underlining the potential complexity of the mRNP molecules (Castello et al., 2012; Baltz et al., 2012). The general NTR for general bulk mRNA export in yeast is a dimeric complex consisting of Messenger RNA Export factor of 67 kDa molecular weight and mRNA transport defective 2 (MEX67-MTR2). In metazoans the main mRNA bulk NTR is the dimeric complex of Nuclear RNA export factor 1 and NFT2-related export protein 1 (NXF1-NXT1) (Segref et al., 1997; Grüter et al., 1998). A subset of mRNAs are exported by eIF4E and CRM1 in mammalian cells, in an NXF1-independent way (Topisirovic et al., 2009; Culjkovic et al., 2006; Assouline et al., 2009). In yeast MEX67-MTR2 also work together with CRM1 to the export both of the pre-40S and pre-60S ribosomal subunits (reviewed in Sloan et al., 2015).

1.2 mRNA undergoes extensive processing before being exported

The different steps of biogenesis and export of mRNA are interconnected and involve a lot of different proteins and enzymatic activities. The process of mRNA biogenesis starts with the Deoxyribonucleic acids (DNA) being transcribed into Pre-mRNA by the RNA polymerase 2 (RNAPII) in the nucleus (reviewed in Moore & Proudfoot, 2009). The pre-mRNA is extensively processed before it is being exported as mature mRNA. Co-transcriptional 5'capping, splicing and 3'end processing are strongly interwoven with export (Hocine et al., 2010). This interconnectivity is governed by the C terminal (CTD) domain of RNAPII that recruits many of the factors involved in these processes. The CTD consist of multiple highly conserved hepta repeat, YSPTSPS, peptides. The phosphorylation state of Serine 5 (Ser5) and Serine 2 (Ser2) changes over the transcription cycle, and determines the proteins that associate with the CTD (Figure 3) (Katahira, 2015; Moore & Proudfoot, 2009).

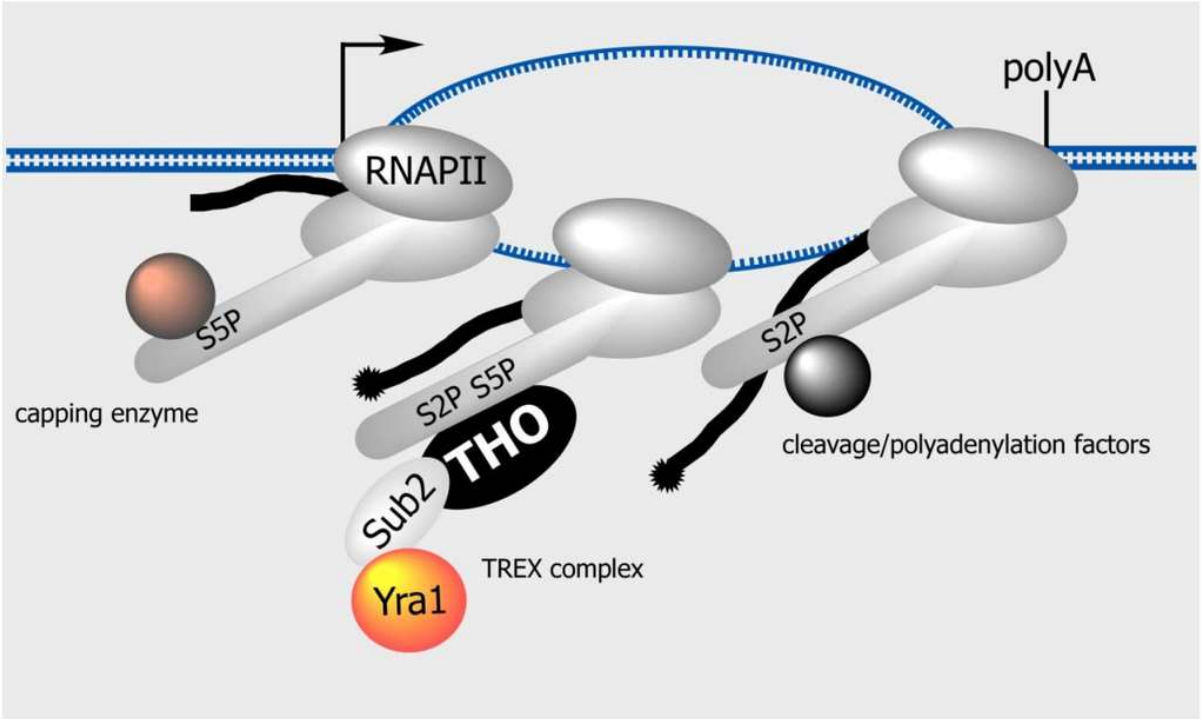


Figure 3. Co-transcriptional coupling of capping, polyadenylation and export. The phosphorylation state of the RNAPII CTD determines which mRNA processing factors that are recruited to the mRNA. The phosphorylation of the CTD by kinases on the Ser5 position is associated with transcription initiation and promoter clearance and recruits the capping enzymes to the transcript co-transcriptionally. As the RNAPII, progress through the template the overall level of Ser5 phosphorylation decreases, but the levels of phosphorylated Ser2 increases towards the end of the transcript. The polyadenylation factors are recruited at the end of transcription, where the Ser2 levels are highest. In yeast the THO/TREX complex involved in mRNA export specifically interacts with the ser2/ser5 phosphorylated CTD. Reviewed in Katahira, 2015 and Hajheidari et al., 2013. Picture was taken from Katahira, 2015.

1.2.1 5’ Capping of mRNA helps recruiting mRNA biogenesis factors

5´capping is the conversion of the pppA 5´ terminus of the primary RNA transcript into 7meGpppA (m7G) by a triphosphatase, a guanyl transferase and methyltransferase activity(Moore & Proudfoot, 2009). In yeast, it has been shown that the C-terminal repeat region (CTR) of the transcript elongation factor SPT5 and the Ser5 CTD of the RNAPII are required for recruitment of the capping enzymes (Schroeder, 2000; Lidschreiber et al., 2013). The (m7G) cap protects the transcript from degradation (Schwer et al., 1998). Many proteins have been reported to bind the methylated cap (Gonatopoulos-Pournatzis & Cowling, 2014). One of these important for mRNP biogenesis is the cap-binding complex (CBC), which consist of the two subunits 80kDa Nuclear cap-binding protein (CBP80) and the 20kDa Nuclear cap-binding protein (CBP20) in yeast and *Arabidopsis*. In humans, CBP80 and CBP20 are also called NCBP1 and NCBP2. Recently a third protein NCBP3 was identified, and the existence of two cap binding complexes NCBP1/NCBP2 and NCBP1/NCBP3 was shown (Gebhardt et al., 2015). These two cap-binding complexes exhibit some different specificity for targets. CBC is a central factor known to orchestrate most RNA biogenesis processes like splicing, 3´ processing, nonsense mediated decay, export of mRNA/snRNAs and recruitment of translation factors in the cytoplasm (Gonatopoulos-Pournatzis & Cowling, 2014; Gebhardt et al., 2015). This is for example illustrated by the CBC stimulating mRNA export by directly interacting with the mRNA export adaptor Ally of AML-1 and LEF1 / RNA Export Factor (ALYREF) both on spliced and intronless mRNA *in vitro* (Hong Cheng et al., 2006; Nojima et al., 2007). In yeast CBC couples transcription and pre-mRNA splicing, as depletion of the CBC inhibits co-transcriptional spliceosome assembly (Görnemann et al., 2005).

1.2.2 Splicing is interconnected with the other processing steps

Besides capping, the removal of intervening RNA sequences (introns) from the pre-mRNA transcript is required for proper maturation of the pre-mRNA. This is done by the spliceosome complex consisting of the U1, U2, U4, U5 and U6 small nuclear snRNPs (Hocine et al., 2010).The spliceosome is functionally conserved in yeast, *Drosophila*, *Arabidopsis* and humans with some species specific variation in the protein composition (Koncz et al., 2012). The overall process (Figure 4) that leads to the excision of the intron and re-ligation of the exons is rather complicated (For some extensive reviews see Lee & Rio, 2015 and Papasaikas & Valcárcel, 2015. Splice

site sequences (5' splice site, branch point and 3' splice site) of canonical (U2) introns in yeast are much conserved, whereas these sequences in metazoans are more loosely defined (Will & Lührmann, 2011; Hocine et al., 2010). In metazoans the 5' splice site selection by U1 can be affected by many factors (Serine rich (SR) proteins, Heterogenous nuclear ribonucleoproteins (hnRNPs), exonic or intronic enhancer/silencing elements, RNA-RNA interactions, histone modifications and RNAPII velocity), leading to alternative splicing of approximately 92-94% human transcripts (Lee & Rio, 2015; Hocine et al., 2010). During splicing, a specific set of proteins that constitute the Exon junction complex (EJC) is deposited onto the exon junction of spliced mRNA. The EJC is composed of the four core proteins Y14, Magoh, eIF4AIII and MLN51/Barentz (BTZ). The EJC is involved in splicing, mRNA export, nonsense-mediated mRNA decay (NMD) and translational control (Reviewed in Bono & Gehring, 2011). All core components of the EJC in humans are involved in NMD, a mechanism that is part of the mRNA surveillance and leads to rapid degradation of

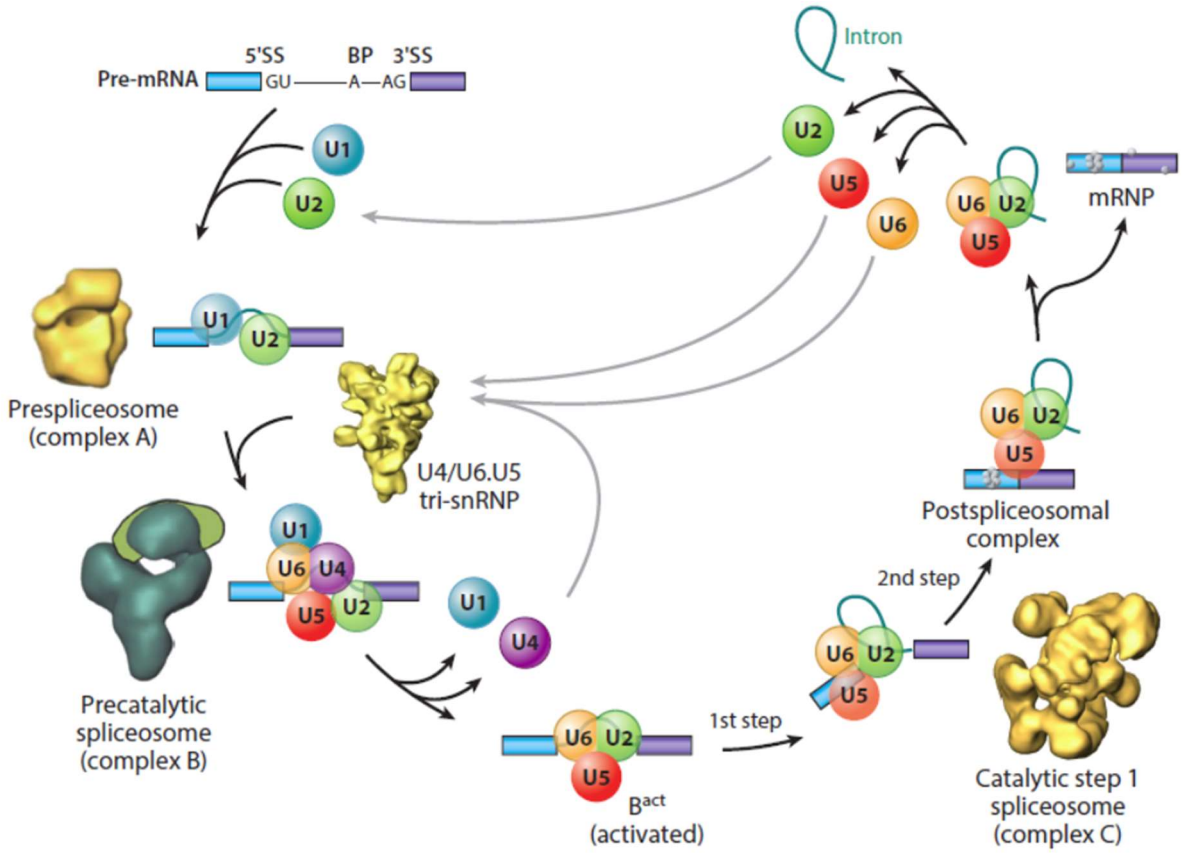


Figure 4. Overview of pre-mRNA splicing. The composition of the spliceosome changes multiple times during the splicing process. The U1, U2, U4, U5 and U6 snRNPs are involved in different stages of splicing. The splicing cycle starts with recognition of the 5' splicing site (5'SS) by U1 and ends with the excision of the intron. (Lee & Rio, 2015)

transcriptswith premature stop codons. Deposition of the human EJC core upstream of the spliced junction is mediated by eIF4AIII and CWC22, a non-core component of the spliceosome-activating PRP19 (also known as NTC) complex (Chuang et al., 2015). The yeast NTC complex plays a role in regulating spliceosome conformation and fidelity (Reviewed in Hogg et al., 2010). In humans, subunits of the CBC, EJC, NTC and the THO/TREX (Described below) complex co-purify with spliceosomes showing the interconnectivity between these complexes and processes (Jurica et al., 2002; Hartmuth et al., 2002; Zhou et al., 2002; Rappsilber et al., 2002).

1.2.3 Polyadenylation is the last processing step

The last of the co-transcriptional maturation processes is the poly-adenylation, where recognition of a Poly-A site (PAS) in the 3´end of the mRNA leads to cleavage of the transcript followed by a ~ 200-300 nucleotide poly(A) tail addition (Keller et al., 2000). These processes are done by the Poly(A) polymerase(PAP), Cleavage and polyadenylation specificity factor (CPSF), Cleavage stimulation factor (CstF) and Cleavage factors I and II (CFIm and CFIIm) complexes. The various steps of polyadenylation are reviewed in Proudfoot, 2004 and Shi & Manley, 2015. Polyadenylation is strongly linked to transcription, splicing and mRNA export (reviewed in Ruepp et al., 2011). An example of this is that the CFIm68 subunit binds to NXF1 and contributes to mRNA export (Ruepp et al., 2009). When the mature RNA has been correctly processed, like described above, it is released and exported to the cytoplasm.

1.2.4 Only transcripts passing mRNA surveillance are exported

Sometimes errors occur during processing leading to aberrant transcripts that could lead to poisonous proteins. There are multiple RNA surveillance mechanisms in place to avoid this (Katahira, 2015). Cytoplasmic mRNA surveillance is reviewed in Schoenberg & Maquat, 2012. Cytoplasmic surveillence is often coupled to translation, like in the case of NMD, where transcripts with premature stop codons are degraded. The degradation of aberrant mRNA is mainly done by 5’ to 3’ and 3’ to 5’ exonucleases. The degradation process is not completely understood, but it also involves deadenylation and decapping enzymes (Schoenberg & Maquat, 2012).

In addition to the cytoplasmic surveillance, there are also various mechanisms in play that controls mRNA in the nucleus (Schmid & Jensen, 2008; Tutucci & Stutz, 2011). In the yeast nucleus, the mRNA degradation is mediated by Rat1p and the

Exozyme/TRAMP complex showing 5' to 3' and 3' to 5' exonuclease activity respectively. Transcripts that fail the surveillance can also be stopped from export by retention at the gene locus or at the NPC (Schmid & Jensen, 2008).

1.2.5 Packaging of mRNPs affects the choice of mRNA export pathway.

There is evidence supporting that mRNPs also undergo compaction during processing. In yeast electron microscopy pictures of mRNPs copurified with the Nuclear polyadenylated RNA-binding (NAB2) protein, show that the mRNPs are much shorter than anticipated from mRNA length (Batisse et al., 2009). *In vivo* studies of fluorescently labelled mRNPs showed that increasing the length of an mRNA does not lead to the same increase in diffusion coefficients, indicating that bigger mRNPs are more tightly packed (Mor et al., 2010). Electron microscopy pictures of the Balbini ring in *Chironomus tentans* also supports the notion the mRNP being packed (Stevens & Swift, 1966). There is data supporting that hnRNPs are involved in packaging the pre-mRNP (reviewed in Müller-McNicoll & Neugebauer, 2013). The RNA binding protein hnRNP C form a heterotetramer of hnRNP1 and hnRNP2 showing 3:1 stoichiometry. The tetramer binds uridine tracts and has 150-250 nucleotides of RNA around it (McAfee, Soltaninassab, et al., 1996; McAfee, Shahied-Milam, et al., 1996; Huang et al., 1994). Recent *in vivo* data from an iCLIP experiment showed the transcriptome wide binding sites of hnRNP C. The hnRNP C was mostly found binding introns. There was also found a bias towards binding uridine tracts with 160-300 bp in between, which is consistent with the length of RNA previously observed to be wrapped around a tetramer of hnRNP C. This indicates that RNA is wrapped around hnRNP C adjacent to one another, making a particle (König et al., 2010). In the same study, it was shown that the position of the hnRNP C also affects alternative splicing. The packaging is speculated to happen co-transcriptionally (Müller-McNicoll & Neugebauer, 2013). An advantage of co-transcriptional packaging could be avoiding that the RNA binds to the DNA, and makes harmful R-loops (Reviewed in Aguilera & García-Muse, 2012).

Recently it was shown that hnRNP C affects which export pathway RNAPII transcript utilize (McCloskey et al., 2012). Both snRNA and mRNA are transcribed by RNAPII, but they are exported via different pathways (PHAX/CRM1 or NXF1/NFT1, respectively). Both types of transcript are capped, and bound by CBC. If the transcript becomes longer than 200-300 bp hnRNP C binds and the transcript is exported by

NXF1/NFT1. The hnRNP C protein inhibits binding of PHAX/CRM1 to transcripts. If the transcript is shorter hnRNP C does not bind and the transcript is bound by PHAX/CRM1. This illustrates how there is an connection between mRNP structure and mRNA export (McCloskey et al., 2012). Splicing has also been suspected to play a role in compaction. In Singh et al., 2012 they utilized a tandem affinity purification strategy combined with high throughput sequencing to show that the EJC together with associated SR proteins make high molecular weight multimers.

1.2.6 NXF1 and the THO/TREX complex are required for mRNA export

The canonical mRNA NTR in metazoans is the heterodimeric NXF1/NXT1 (MEX67-MTR2 in yeast) complex (Segref et al., 1997; Grüter et al., 1998; Köhler & Hurt, 2007). Metazoans have several paralog genes of NXF1. These genes exhibit tissue specific expression and are involved in other pathways. Thus generally mRNA is exported by NXF1 (Katahira, 2015). Knockdown of the metazoan NXF1 leads to accumulation of bulk poly(A) mRNA in the nucleus, and is essential for cell viability. (Katahira et al., 2015; Herold et al., 2001; Katahira et al., 1999). In yeast a thermo sensitive mutant of MEX67 accumulates poly(A) RNA in the nucleus under non-permissive temperatures (Segref et al., 1997). NXF1 consists of multiple domains: RNA binding domain (RBD), RNA recognition motif (RRM), Leucine rich repeat (LRR), Nuclear export factor 2- like (NTF2L) and Ubiquitin-associated (UBA) domain (Viphakone et al., 2012; Katahira et al., 2015). The NTF2L and UBA domain contains FG-repeat-binding sites and are required for translocation through the NPC in both human and yeast (Strässer et al., 2000; Katahira et al., 2002). The RRM, LRR and NTR2L (together with NXT1) domains are involved in NXF1 binding of the retroviral Conserved transport element (Katahira et al., 2015). In the same study it was shown that mutations in both the NTR2L and RRM, lead to less association of NXF1 with bulk mRNA. NXF1/NXT1 and MEX67/MTR2 bind RNA non-specifically and cannot distinguish RNAs on their own (Segref et al., 1997; Katahira et al., 1999; Santos-Rosa et al., 1998). Key factors like the Transcription–export (TREX) complex associates with NXF1 in order to bind mRNAs with high affinity (Katahira et al., 2015; Viphakone et al., 2012; Hung et al., 2010). TREX consists of the core THO (A suppressor of the transcriptional defects of HPR1 mutants by overexpression) complex (HPR1,THOC2, hTEX1, THOC5, THOC6 and THOC7 in mammals and HPR1,THO2, MFT1, THP2 and TEX1 in yeast) and the proteins UAP56 (SUB2), ALYREF (YRA1)

and CIP29 (THO1) (Dufu et al., 2010; Sträßer et al., 2002; Chi et al., 2013). The THO complex assembles with the U2AF associated (UAP56), ALYREF and Cytokine-induced factor of 29 kDa (CIP29) in an ATP dependent manner (Dufu et al., 2010). UAP56 (And its yeast ortholog SUB2 (Suppressor of Brr1-1,2)) is a DEAD box RNA helicase that is involved in mRNA export (Jensen et al., 2001; Strässer & Hurt, 2001; Luo et al., 2001; MacMorris et al., 2003) and involved in pre-spliceosome assembly (Fleckner et al., 1997; Libri et al., 2001; Zhang & Green, 2001; Kistler & Guthrie, 2001). In humans UAP56 has a paralog, DEAD box polypeptide 39 (DDX39/URH49) that is 90% identical on the sequence level (Pryor et al., 2004). Only simultaneous knockdown of both DDX39 and UAP56 in humans leads to a severe accumulation of poly(A) RNA, indicating some level of functional redundancy (Kapadia et al., 2006). Both UAP56 and DDX39 interacts directly with CIP29 and ALYREF (Dufu et al., 2010) Another study found that UAP56 preferentially associates with THO and ALYREF whereas DDX39 prefers to bind CIP29 in the so-called the AREX complex exporting a subset of mRNAs involved in mitosis (Yamazaki et al., 2010). In the same study, it was observed that single knockdown of UAP56 and DDX39 lead to different mitotic defects suggesting that UAP56 and DDX39 might not be completely redundant.

1.2.7 Export adaptors enhance the mRNA binding activity of NXF1

Proteins that mediate RNA binding specificity to the export receptors by direct interaction are called export adaptors (Katahira, 2015). ALYREF is one of these important adaptors which interact directly with NXF1/NXT1 (YRA1 also interacts with MEX67/Mtr2) (Strässer & Hurt, 2000; Rodrigues et al., 2001). ALYREF has mRNA binding activity through its arginine rich motifs in the variable N and C regions, in addition to weak RNA binding of the RRM (Golovanov et al., 2006; Hautbergue et al., 2008). Recent data indicates that a small WQHD motif in the N-terminal variable region is required for the binding of mRNA (Gromadzka et al., 2016). Metazoan ALYREF enhances the RNA binding activity of NXF1, and simultaneously hands over the RNA to NXF1 in a process that is regulated by asymmetric dimethylation of the arginines in the variable N and C regions (Hautbergue et al., 2008; Hung et al., 2010). The mechanism behind how adaptors enhance the RNA binding activity of NXF1 was recently elucidated. Examination of the interaction between TREX complex and NXF1 showed that the two export adaptors ALYREF and

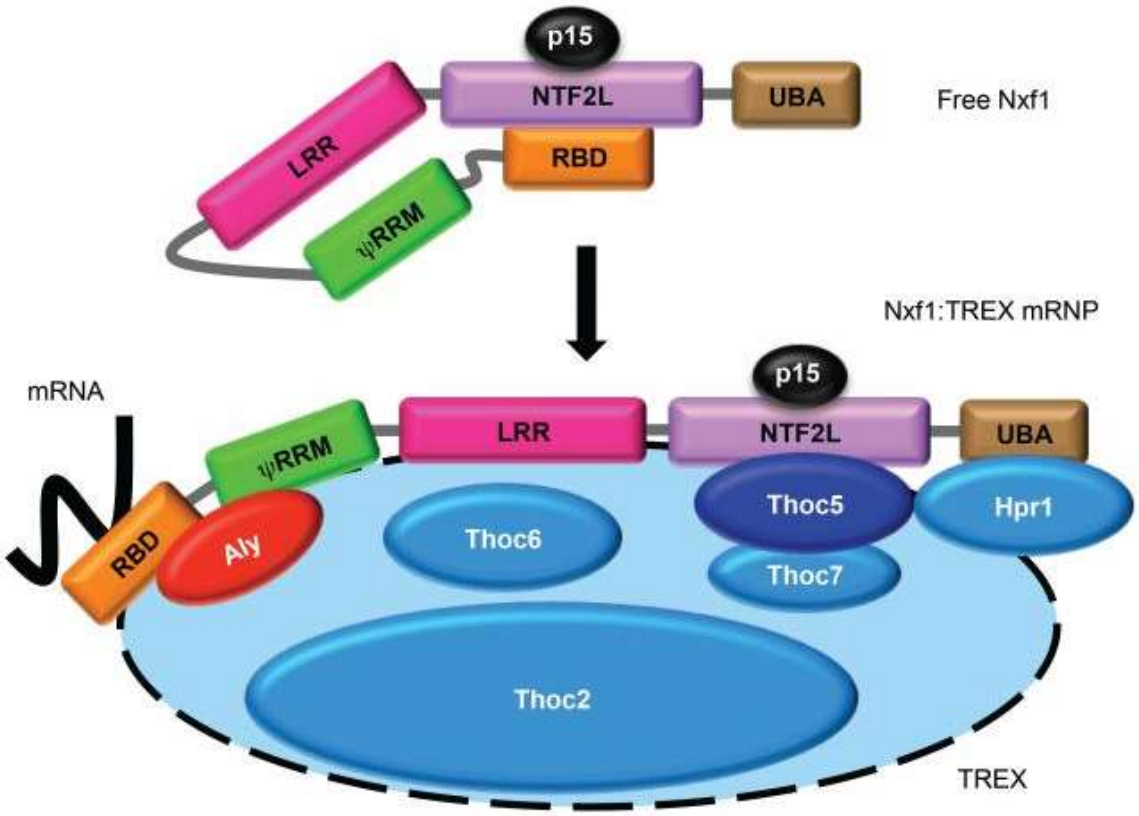


Figure 5. TREX mediated activation of NXF1. Binding of ALYREF and THOC5 to the N-terminal and NTF2L domain of Nxft1/P15(Nxt1), leads to a conformational change that enhance the RNA binding activity of NXF1. Viphakone et al., 2012

THOC5 enhance RNA binding activity of NXF1 by inducing a conformational change of NXF1 upon binding (Viphakone et al., 2012). The conformational change leads to the interruption of an intramolecular interaction between the NTF2L and RBD domain of NXF1, exposing the RNA binding activity of NXF1 to the RNA (Figure 5). Other export adaptors than ALYREF have also been discovered in the recent years that bind NXF1. Known TREX associated adaptors in mammals are Chromatin Target Of PRMT1 (CHTOP), Leucine zipper protein 4 (LUZP4) and UAP56-interacting factor (UIF). They all bind to UAP56 and NXF1 (Chang et al., 2013; Hautbergue et al., 2009; Viphakone et al., 2015). These adaptors show some functional redundancy with ALYREF. Knockdown of ALYREF leads to a weak accumulation of poly(A) RNA compared to knockdown of NXF1, whereas simultaneous knockdown of ALYREF/THOC5, ALYREF/UIF, ALYREF/CHTOP leads to a severe accumulation (Viphakone et al., 2012; Hautbergue et al., 2009; Chang et al., 2013). Like ALYREF, the two proterins CHTOP and CIP29 stimulate the ATPase and helicase activity of UAP56, showing some functional similarity. More adapters are found among the serine-arginine (SR) rich proteins. The SR proteins 9G8 and SRp20 are TREX inde-

pendent adaptors, that enhance the RNA binding efficiency of NXF1 (Hautbergue et al., 2008). In yeast Nuclear protein localization 3 (NPL3) and NAB2 were found to be adaptors that also mediate the recruitment of MEX67/MTR2 (Iglesias et al., 2010; Gilbert & Guthrie, 2004; Batisse et al., 2009).

1.2.8 Recruitment of the export machinery

The TREX complex plays an important role in mediating NXF1 bulk mRNA export. How TREX is recruited to the genes is still not completely clear (Katahira, 2015). Earlier studies have indicated that yeast TREX recruitment is coupled to transcription elongation (Sträßer et al., 2002; Luna et al., 2005). THO is thought to be recruited to all protein coding genes in yeast (Gómez-González et al., 2011) by direct interaction with the ser2/ser5 phosphorylated CTD of RNAPII and RNA (Meinel et al., 2013). YRA1 is recruited independently of SUB2 and requires the PCF11 subunit of the CF1A Cleavage factor involved in 3' mRNA processing (Johnson et al., 2009). In the same study, it was shown that PCF11 and SUB2 binds the same domain of YRA1. After recruitment of YRA1 and SUB2 to the TREX complex, it is thought that YRA1 and SUB2 are transferred to the mRNA where they can recruit MEX67-MTR2 (Abruzzi et al., 2004; Sträßer et al., 2002; Strässer & Hurt, 2000). The Prp19 (NTC) complex is also required for the occupancy of yeast TREX on transcribed genes, except at the 5´end (Chanarat et al., 2011). In humans there are *in vitro* data indicating that recruitment of the TREX complex is splicing dependent (Masuda et al., 2005; Hong Cheng et al., 2006). Splicing has been shown to increase export efficiency and kinetics of spliced mRNAs compared to their cDNA counterparts *in vivo* (Valencia et al., 2008). ALYREF and UAP56 are also recruited to intronless RNAs in a CBC dependent manner and require the ATP-bound form of UAP56 (Taniguchi & Ohno, 2008; Nojima et al., 2007). A Recent *in vitro* study showed that ALYREF, UAP56 and DDX39 preferentially (compared to intronless RNA) are recruited in a cooperative manner to spliced mRNA containing both the EJC and the cap (Gromadzka et al., 2016). The binding of ALYREF and RNA to the ATP bound form of UAP56, stimulates the intrinsic ATPase activity of UAP56, leading to dissociation of UAP56 from the complex (Taniguchi & Ohno, 2008). UAP56/SUB2 and NXF1/MEX67 binds ALYREF/YRA1 at the same domain in an exclusive manner (Strässer & Hurt, 2001; Golovanov et al., 2006; Taniguchi & Ohno, 2008). This all lead to the model that after

recruitment of TREX to the RNA, UAP56 binding of ALYREF is replaced by NXF1, leading to export (Katahira, 2015; Hautbergue et al., 2008).

1.2.9 Transport to the nuclear pore and beyond

In yeast it has been shown that genes like GAL7, GAL1 and GAL10 upon Galactose activation are positioned to the NPC. This relocation, also called gene gating, requires components of the SAGA and the Transcription-export complex 2 (TREX2) complex (Cabal et al., 2006; Chekanova et al., 2008; Jani et al., 2014; Drubin et al., 2006). SAGA is a transcriptional co-activator complex and TREX2 is a complex anchored to the nuclear pore basket on the nucleoplasmic side of the NPC (Fischer, 2002; Rodríguez-Navarro et al., 2004). Deleting any of the TREX2 subunits (Supressor of actin 3 (SAC3), THO2/HPR1 phenotype 1 (THP1), CDC31, SUS1 or SEM1) leads to defective mRNA export in yeast (Okamura et al., 2015). SAC3 of the TREX2 complex interacts directly with MEX67/MTR2(Fischer, 2002).

In mammalian cells TREX2 associates at the nuclear basket independent of SAGA, is required for bulk mRNA export, co-purifies with NXF1 and interacts with the nuclear pore basket (Umlauf et al., 2013; Wickramasinghe, McMurtrie, et al., 2010; Wickramasinghe et al., 2014; Jani et al., 2012). Knocking down GANP (SAS3) and NXF1 in humans leads to poly(A) accumulation in nuclear speckles, and not at the nuclear periphery (Wickramasinghe, McMurtrie, et al., 2010). It has been speculated that TREX2 helps to chaperone the RNA from the place of transcription (Nuclear transcription factories) to the NPC.(Wickramasinghe, Stewart, et al., 2010). TREX2 in humans has been indicated to only be required for a subset of NXF1 transported genes (Wickramasinghe et al., 2014). It is not clear if SAGA and TREX-2 in higher eukaryotes function like the proteins in yeast (Okamura et al., 2015).

When a mRNP/export receptor complex passes though the NPC, remodelling occurs on the cytoplasmic side that is essential for mRNA dissociation from the export receptor (Okamura et al., 2015). The ATP-dependent RNA helicase DBP5 (DBP5/DDX19 in vertebrates) is the main player in remodelling (Alcázar-Román et al., 2006; Tran et al., 2007; Weirich et al., 2006). A study in yeast has indicated that MEX67 is being displaced from the mRNP by DBP5 (Lund & Guthrie, 2005). DBP5 localizes with NPC filaments on the cytoplasmic side and is required for efficient mRNA export (Snay-Hodge et al., 1998; Schmitt et al., 1999; Tseng et al., 1998). The RNA helicase activity is stimulated by GLE1 and inositol

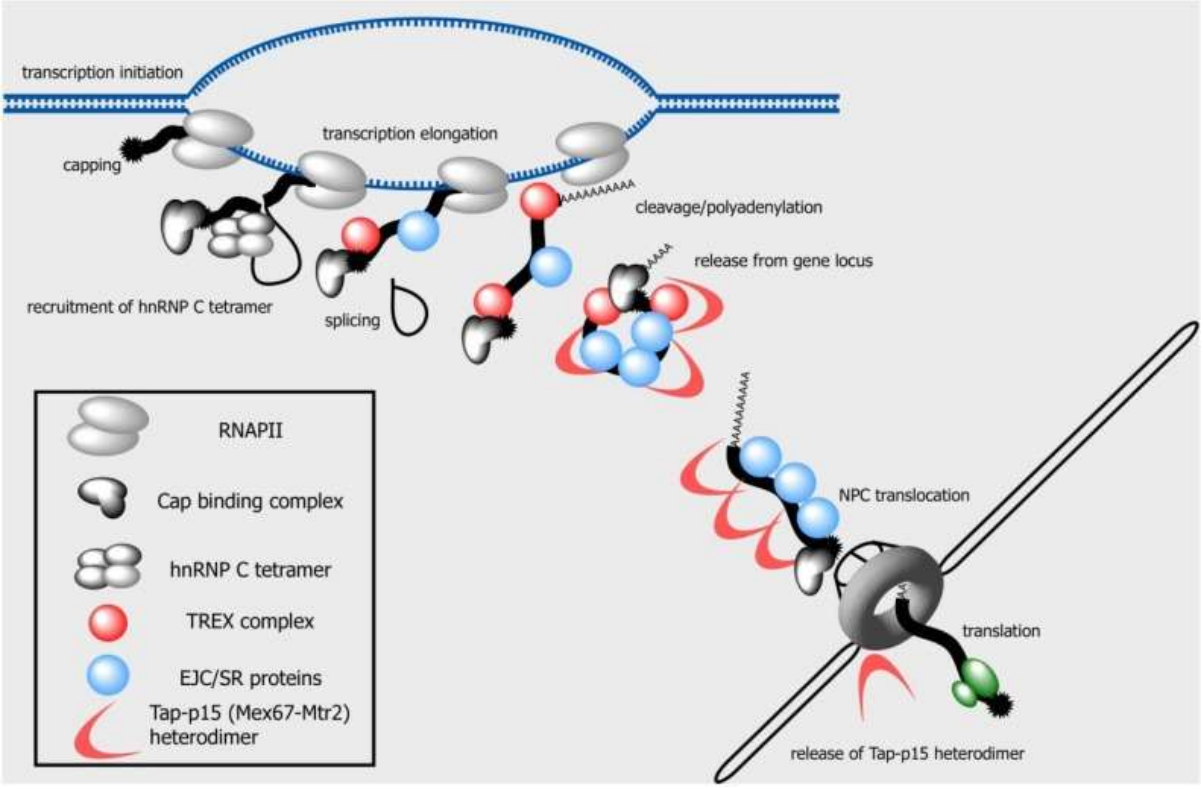


Figure 6. Overview of bulk mRNA export in metazoans. Pre-mRNA transcribed by RNAPII is co-transcriptionally packed, capped, spliced and polyadenylated. The TREX complex together with different export adaptors recruits the export receptor NXF1. NXF1 mediates translocation through the NPC. On the cytoplasmic side of the NPC the mRNP is remodelled, and NXF1 is released. The mRNA can then be translated. Picture is from Katahira, 2015.

hexakiphosphate (IP_6) (Alcázar-Román et al., 2006; Tran et al., 2007). The activity of DBP5 mediated mRNP remodelling depends on stimulation by GLE1/IP_6 and NUP159 triggered ADP release from DBP5 (Hodge et al., 2011; Noble et al., 2011). The DBP5 mediated remodelling on the cytoplasmic side is thought to drive mRNA export from nucleoplasm to cytoplasm (Köhler & Hurt, 2007). After the export, the mRNA can be translated. An overview of all the processes involved in NXF1/NXT1 mediated mRNP export is shown in Figure 6.

1.3 Export of mRNAs in plants

1.3.1 Identified proteins affecting mRNA export

Little is known about mRNA export in plants. Most of the information is from the flowering plant *Arabidopsis thaliana*. The components of the *Arabidopsis* NPC has been identified and a homology analysis showed that the protein composition of *Arabidopsis* NPCs is closer to metazoans than to yeast (Tamura et al., 2010). The atNUP160, atNUP96 and SEH1 subunits of the NUP107-160 subcomplex and the NPC associated proteins atTPR/NUA and HOS1 are required for bulk mRNA export

and exhibit pleiotropic phenotypes upon downregulation (Dong et al., 2006; Parry et al., 2006; Jacob et al., 2007; Xu et al., 2007; Wiermer et al., 2012; MacGregor et al., 2013). The *Arabidopsis* ortholog of the ATP dependent RNA helicase DBP5 (LOS4) together with its co-factor GLE1 are also required for mRNA export (Lee et al., 2015; Gong et al., 2005). RNA, GLE1 and IP_6 also stimulate ATPase activity of LOS4 in *Arabidopsis* (Lee et al., 2015). Another NPC associated complex involved in mRNA export that has been described in *Arabidopsis* is TREX2. A Yeast-two Hybrid screen indicated that the *Arabidopsis* TREX2 complex consists of THP1, SAC3A, SAC3B, CEN1, CEN2 and DSS1. THP1 and SAC3b interacts with NUP1 at the nuclear periphery. THP1 is required for bulk mRNA export (Lu et al., 2010).

1.3.2 The *Arbidopsis* THO/TREX complex is conserved

Interestingly, the bulk mRNA receptor allowing mRNPs to transverse the NPC is unknown in *Arabidopsis*. No orthologs of NXF1 and MEX67 have been identified in *Arabidopsis* (Pendle et al., 2005). The THO/TREX complex is conserved though, as orthologs of human TREX and TREX associated adapters have been identified by sequence similarity (Table 1). The core *Arabidopsis* THO complex was found to consist of Hyper recombinant 1 (HPR1), THO2,TEX1, THO5-1, THO5-2, THO6, THO7-1 and THO7-2 by affinity purification of tagged TEX1 (Yelina et al., 2010). The non-THO subunits of TREX were not co-purified and identified by mass spectrometry in that study. Interrupting the THO2 gene by inserting T-DNA into exon 16 or exon 18 of both alleles leads to lethality during embryo development indicating that the *THO2* gene is essential in *Arabidopsis* (Furumizu et al., 2010). Three other T-DNA and mutation lines positioned in the 5´UTR or the end of the *THO2* gene, have been analysed in another study (Francisco-Mangilet et al., 2015). The three mutant lines showed a wide range of pleiotropic phenotypes. *HPR1* and *TEX1* were identified in two simultaneous screens for mutants defective in spreading of RNA silencing (Jauvion et al., 2010; Yelina et al., 2010). In *tex1*, *hpr1* and *tho6* mutants, the levels of trans-acting si-RNA (tasiRNA) are reduced (Jauvion et al., 2010; Yelina et al., 2010). The *Arabidopsis* THO complex has also been implicated in mRNA export. Loss of the *HPR1* transcript leads to accumulation of bulk poly(A) RNA in the nucleus (Pan et al., 2012; Xu et al., 2015; Furumizu et al., 2010). The same studies showed that HPR1 plays a role in disease resistance, Ethylene response, senescence and miRNA processing. HPR1 is also required for expression of the *ERECTA*

Table 1. *Arabidopsis* protein orthologs of the human THO/TREX complex and cofactors. N/A means no ortholog was found by a BLAST homology search (Arabidopsis.org).

Complex	Human protein	Arabidopsis ortholog (AGI)	Reference (*Arabidopsis*)
TREX	CIP29	MOS11 (AT5G02770	(Germain et al., 2010)
	UAP56/DDX39	UAP56-1 (AT5G11170)	(Kammel et al., 2013)
		UAP56-2 (AT5G11200)	(Kammel et al., 2013)
	ALYREF	ALY1 (AT5G59950)	(Pendle et al., 2005; Uhrig et al., 2004)
		ALY2 (AT5G02530)	(Pendle et al., 2005; Uhrig et al., 2004)
		ALY3 (AT1G66260)	(Pendle et al., 2005; Uhrig et al., 2004)
		ALY4 (At5g37720)	(Pendle et al., 2005; Uhrig et al., 2004)
	UIF	UIF (AT4G10970)	(Hautbergue et al., 2009)
		UIF (AT4G23910)	(Hautbergue et al., 2009)
THO	HPR1	HPR1 (AT5G09860)	(Yelina et al., 2010)
	THO2	THO2 (AT1G24706)	(Yelina et al., 2010)
	TEX1	TEX1 (AT5G56130)	(Yelina et al., 2010)
	THOC5	THO5(AT5G42920)	(Yelina et al., 2010)
		THO5(AT1G45233)	(Yelina et al., 2010)
	THOC6	THOC6 (AT2G19430)	(Yelina et al., 2010)
	THOC7	THO7 (AT5G16790)	(Yelina et al., 2010)
		THO7 (AT3G02950)	(Yelina et al., 2010)

and *REVERSION TO ETHYLENE SENSITIVITY 1* (*RTE1*) genes (Pan et al., 2012; Xu et al., 2015; Furumizu et al., 2010). Some of the non-THO subunit orthologs of human TREX have been identified in *Arabidopsis*, but has not been described in depth. UAP56 in *Arabidopsis* is coded by two genes coding for the same protein. It localizes to the euchromatic areas of the nucleus, but not the heterochromatic chromocenters (Kammel et al., 2013). UAP56 was shown in the same study to associate with actively transcribed genes but not repressed transposons. The *Arabidopsis* UAP56 has retained its helicase and ATPase activity. UAP56 also interacts *in vitro* with ALY2 and the CIP29 ortholog Modifier of SNC1,11 (MOS11). MOS11 was identified in a screen finding modifiers of *snc1* mediated pathogen resistance. MOS11 is required for bulk mRNA export, but the *mos11* mutant shows no obvious phenotypes (Germain et al., 2010). The four ALY orthologs (Mentioned as ALY1, ALY2, ALY3 and ALY4 in this study) have all been identified in a proteomics analysis of the nucleoli proteome in *Arabidopsis* cell culture (Pendle et al., 2005). The four ALY proteins have been reported to exhibit differential sub-nuclear localization (Uhrig et al., 2004; Pendle et al., 2005). The *Tomato bushy stunt virus* P19 is a repressor of RNA silencing, that interact directly with *Arabidopsis* and *Nicotiana benthamiana* ALY proteins (Uhrig et al., 2004).

1.4 Aim of the project

In *Arabidopsis thaliana* no orthologues protein of the mRNA export receptor NXF1/MEX67 has been found, but the orthologs of the human THO/TREX complex, which recruit NXF1/MEX67 have been identified. The *hpr1* mutant of THO accumulate poly(A) mRNA in the nucleus, making THO a good candidate for gaining more information about mRNA export in *Arabidopsis*. The aim of this project is to generate more knowledge about the TREX complex, regarding both composition and function. In order to achieve this aim a reverse genetics approach will be applied for ALY1, ALY2, ALY3, ALY4, MOS11 and TEX1 utilizing T-DNA insertion mutants. To investigate if the four ALY proteins are redundant or exhibit specific functions, phenotyping of single and double mutants, combined with subcellular localization of GFP fused proteins, will be utilized. Furthermore, it will explored if all the four ALY proteins show conserved interaction with the *Arabidopsis* UAP56 protein.

The role of the THO component TEX1 in other organisms is largely unknown. As it has been shown that *Arabidopsis tex1* shows a defect in biogenesis of tasiRNA, it will be examined if *tex1* exhibits other defects by phenotypical analysis. It has so far not been described if *TEX1* is required for bulk mRNA export, so this will be investigated. As only the core THO complex so far has been found to associate with TEX1, it will be attempted to expand the TREX proteome in *Arabidopsis* by affinity purifications. CIP29 (MOS11) in humans has been reported to be closely associated with the TREX complex. It will be investigated if MOS11 interacts physically and genetically with THO/TREX in *Arabidopsis*. Both through the aforementioned affinity purifications and through the creation of double mutants with TEX1. Thorough phenotypical analysis, combined with subcellular localization of GFP fused proteins, will be used to gain information about the relationship between TEX1 and MOS11. Lastly, the effect of TEX1 and MOS11 on mRNA export will be examined. All this together will reveal new insight into the role of the TREX complex in *Arabidopsis thaliana*.

2 Material and Methods

2.1 Materials

2.1.1 Instruments

Table 2. Instruments used in this study

Instrument	Manufacturer /model
Centrifuges	-Sorvall™ Evolution RC with SLA1500/SR34 rotor (Thermo Fisher Scientific - USA) -Sorvall Lynx ™ 4000 with TH13 rotor (Thermo Fisher Scientific – USA) - 5417R (Eppendorf - Germany)
Gel documentation	-Biometra TI5 system (Biometra - Germany)
Hemocytometer	- Fuchs-Rosenthal 0.2mm (Marienfeld-superior - Germany)
Homogenizer	-Silamat S5 homogenizer (Ivoclar vivadent - USA)
Hybridization Oven	-Hybridisierungsöfen (Uniequip - Germany)
Imager immunoblotting	-FluorChem Multiimage II FC2 imager (Alpha Innotech - Germany)
Immunoblotting system	-Semy-Dry-Blotting System (Carl Roth - Germany)
Microscope (Confocal Laser Scanning Microscopy)	-SP8 (Leica)
Microscope (Counting protoplasts)	-Primo star (Zeiss)
Microscope (Crossing)	-Nikon SMZ645 stereo microscopes attached to a Schott KL 1500 LCD
Microscope (Documenting phenotypes)	-Zeiss Discovery V8 stereo with an Axiocam MRc5 digital camera and two Zeiss KL1500 LCD
Microscope (Epifluorescence)	-Zeiss Axio scope with an Axiocam MRc5 digital camera
Microscope (Structure illumination)	-Zeiss Imager M2 ApoTome.2 with Axiocam 503 mono and 105 colour
Nanodrop	-Nanodrop ND1000 spectrophotometer (Thermo scientific - USA)
PCR cycler	-Tgradient or a T3000 PCR thermocycler (Biometra - Germany)
Phosphor imager	-Cyclone™ Storage phosphor imager (Canberra Packard - Austria)
Plant incubator	-Plant incubator (Percival Scientific - USA)
RT-qPCR Cycler	-Mastercycler® ep RealPlex (Eppendorf-Germany)
Sonicator	-UW2070 MS73 (Bandelin electronic - Germany)

2.1.2 Chemicals and enzymes

Laboratory grade chemicals and reagents were purchased from Carl Roth (Germany), Fluka (Switzerland), USBiologcal (USA), Clonetech, AppliChem (Germany), Merck (Germany), Duchefa (Netherlands), Sigma-Aldrich (Germany) and VWR (USA). Enzymes were purchased from Thermoscientific (USA), Peqlab/VWR (USA), Duchefa (Netherlands) or New England Biolabs (USA). L-[35S]-Methionine and [γ-32P] ATP was obtained from Hartmann analytic (Germany).

2.1.3 Oligonucleotides

Oligonucleotides used in this study (Table 3 and Table 4) was obtained from MWG (Germany).

Table 3. Oligonucleotides used in this study. CDS stands for coding sequence, and primer sequences underlined signify restriction cutsites.

Primer	description	sequence	Lab. ID.
P1	Genotyping T-DNA insertion SAIL_LBI	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC	802
p2	Genotyping T-DNA insertion Wiscdxlox LB DS3 (p745)	AACGTCCGCAATGTGTTATTAAGTTGTC	2167
P3	Genotyping T-DNA insertion SALK_LBb	GTTGCCCGTCTCACTGGTGA	812
P4	Genotyping T-DNA insertion GABI-KAT_LB 8409	ATATTGACCATCATACTCATTGC	1595
P5	Genotyping and Expression , TDNA insertion *aly1-1* fw	ACCCTAATCGGAAATCAACCCGATC	2963
P6	Genotyping and Expression , TDNA insertion *aly1-1* rv	TTTGAATTCTTAGTTTGTCTCCATATCTCCAGAATGG	2850
P7	Genotyping , TDNA insertion *aly2-1* fw	GTGGCTCGTCTATCGAGACTG	2153
P8	Genotyping , TDNA insertion *aly2-1* rv	TCCATTGCCTCTTTGTGGTAC	2154
P9	Genotyping , TDNA insertion *aly3-1* fw	AAAGCCTGAAAATGTGTGTGG	2351
P10	Genotyping , TDNA insertion *aly3-1* rv	GTCTCCCAGAAGGACCTGAAC	2350
P11	Genotyping , TDNA insertion *aly4-1* fw	CTCTTGCTGTGAATGCTCGTCCAT	2816
P12	Genotyping , TDNA insertion *aly4-1* rv	CATAACGCTCTACCTCCCCAATCTC	2817
P13	Genotyping , TDNA insertion *tex1-4* fw1	GAGGAGACGACGATTCCTTTC	2159
P14	Genotyping , TDNA insertion *tex1-4* rv1	AACAGCTACATGTGTCCCGTC	2160
P15	Genotyping , TDNA insertion *tex1-4* fw2	AAACTCGAGATGGAGGAGACGACGATTCCTTTC	2914
P16	Genotyping , TDNA insertion *tex1-4* rv2	TTCCCGGTCTCAAAGTTCATCCTCT	3215
P17	Genotyping , TEX1-GFP fw	TTCCCGGTCTCAAAGTTCATCCTCT	2919
P18	Genotyping , TEX1-GFP rv	GGACAACTCCAGTGAAAAGT	860
P19	Genotyping , TDNA insertion *mos11-2* fw	CGGGATCCATGGCGACCAACGGAGAG	2808
P20	Genotyping , TDNA insertion *mos11-2* rv	ACGCGTCGACTTAGGCAGCGCTTCCTGA	2809
P21	Expression of *ALY2* fw, RT-PCR	CTCTGAGGTTGGTGATCTTAAGCGAT	2735
P22	Expression of *ALY2* rv, RT-PCR	TTGTAACGCTTGACAGCTGCCAA	2736
P23	Expression of *ALY3* fw, RT-PCR	GCCCGCAGGAAGAGGAGCTT	2737
P24	Expression of *ALY3* rv, RT-PCR	CATTTGTCACTCCCTGATCAAGGTTG	2738
P25	Expression of *ALY4* fw, RT-PCR	TAGGGTCAGGAGCTTGCCATGG	2958
P26	Expression of *ALY4* rv, RT-PCR	CTTCATTTGTCACACCCTGGTCCA	2959
P27	Expression of *TEX1* fw1, RT PCR	AAACGGCTCGCATTTGGAATATCG	2739
P28	Expression of *TEX1* rv1, RT PCR	CAGAGGCAGTAGCAACAAGATCTGA	2740
P29	Expression of *TEX1* fw2, QPCR	GGAGACGACGATTCCTTTCAAGAGC	3936
P30	Expression of *TEX1* rv2, QPCR	CCGTGTGGTTCGATATTCCAAATGCG	3937
P31	Expression of *TEX1* fw3, QPCR	TAAACCACTTGACACTCTCACAGCCC	3940
P32	Expression of *TEX1* rv3, QPCR	AGCTTATTGTTCTGACTGGCCACTCA	3941
P33	Expression of *TEX1* fw4, QPCR	GCTCGCATTTGGAATATCGAACCACA	3950
P34	Expression of *TEX1* rv4, QPCR	TCTCTCCACTAAGTTCTACCTGCTGTGT	3951
P35	Expression of *MOS11* fw1, QPCR	TACGGTGGTGAATGGCGGAGG	3952
P36	Expression of *MOS11* rv1, QPCR	CTCAGCACGTCGAATCTTCTTCTGGA	3953
P37	Expression of *MOS11* fw2, QPCR	ATAGATTTGGGGTTCCTTCTGCGACC	3954
P38	Expression of *MOS11* rv2, QPCR	GCTGCCTCCTTGCCAATAATTTGCTT	3955
P39	Expression of *MOS11* fw3, QPCR	TATTGGATCGACGGCTGGTG	3998
P40	Expression of *MOS11* rv3, QPCR	TCGCAGAAGGAACCCCAAAT	3999
P41	Expression of *ACTIN8* fw, RT-PCR	TGCTGGTCGTGACCTTACTGATTACC	1473
P42	Expression of *ACTIN8* rv, RT-PCR	TCTCCATCTCTTGCTCGTAGTCGACA	1474
P43	Expression of *ACTIN2* fw, QPCR	CGTACAACCGGTATTGTGC	3384
P44	Expression of *ACTIN2* rv, QPCR	GTCCAGCAAGGTCAAGACG	3385
P45	Expression of *EF1 alpha* fw, QPCR	GTTGTAACAAGATGGATGCCA	3382

P46	Expression of *EF1 alpha* rv, QPCR	GGACAAATGGGATTTTGTCAG	3383
P47	Expression of *FLC* fw1, RT-PCR	CAAACGTCGCAACGGTCTCA	1320
P48	Expression of *FLC* rv1, RT-PCR	TATCGCCGGAGGAGAAGCTG	1321
P49	Expression of *FLC* fw2, QPCR	AGCCAAGAAGACCGAACTCA	3376
P50	Expression of *FLC* rv2, QPCR	TTTGTCCAGCAGGTGACATC	3377
P51	Expression of *ARF3* fw, QPCR	TGTCTGTGGTGTCTTCGAATTCTGCT	3946
P52	Expression of *ARF3* rv, QPCR	ACTTCATCTGTAGTCGTCTCTGCGTG	3947
P53	Expression of *ARF4* fw, QPCR	TTGTGCTGGTCCTCTCACTTGTCTTC	3942
P54	Expression of *ARF4* rv, QPCR	CCCCTCCCCATTTAGCATCGAAAACT	3943
P55	Amplification of ALY1 CDS for pGADT7	GGAATTCCATATGATGTCGACTGGATTAGATATGTCTC	2438
P56	Amplification of ALY1 CDS for pGADT7	CCGCTCGAGTTAGTTTGTCTCCATATCTCCAGAATG	2439
P57	Amplification of UAP56 CDS for pGBKT7	GGAATTCCATATGATGGGAGACGCTAG	2398
P58	Amplification of UAP56 CDS for pGBKT7	TTCTGCAGTTAAGAAGGCATGTA	2399
P59	Amplification of TEX1p::TEX1 for PgreenII0229-3´GFP	AAAGTCGACCTCGAAGACAACAGAATCAGAG-TAAATCTG	2891
P60	Amplification of TEX1p::TEX1 for PgreenII0229-3´GFP	AAACCCGGGGCCGAGCTCTCAAAACCAAATATCCGG	2892
P61	Amplification of MOS11p::MOS11 for PgreenII0179-3´RFP	CATCTAGATGGTTCATGCTTACCCAATGGTGACTA	3582
P62	Amplification of MOS11p::MOS11 for PgreenII0179-3´RFP	CAGGAGCTCGGCAGCGCTTCCTGATACAGC	3583
P63	Amplification of UAP56 CDS for pCam-bis2300-35S::GS	GCTCTAGAATGGGAGACGCTAGA	2776
P64	Amplification of UAP56 CDS for pCam-bis2300-35S::GS	TACCCGGGCCAGAAGGCATGTAGGTTGAAG	2819
P65	Amplification of TEX1 CDS for pCam-bis2300-35S::GS	AAATCTAGAATGGAGGAGACGACGATTCCTTTCAA	3180
P66	Amplification of TEX1 CDS for pCam-bis2300-35S::GS	AAACCCGGGTCCGAGCTCTCAAAACCAAATATCCGGA	3181
P67	Amplification of MOS11 CDS for pCam-bis2300-35S::GS	CGGGATCCCCATGGCGACCAACGGA	2948
P68	Amplification of MOS11 CDS for pCam-bis2300-35S::GS	CAGGAGCTCGGCAGCGCTTCCTGATACAGC	3583
P69	Probe for Northern, TAS1 (sir255)	TACGCTATGTTGGACTTAGAA	4104
P70	Probe for Northern, TAS3 (5´D8)	AAAGGCCTTACAAGGTCAAGA	4105
P71	Probe for Northern, Mir173a	GTGATTTCTCTCTGCAAGCGAA	4106
P72	Probe for Northern, mir390	GGCGCTATCCCTCCTGAGCTT	4107
P73	Probe for Northern, U6	TCATCCTTGCGCAGGGGCCA	4119

Table 4. Modified oligonucleotides used in this study

Oligo	sequence	modification
poly (dT) probe	48 X T	5´ Alexa 488

2.1.4 Vectors

Vector used in this study are listed in Table 5. Plasmids used from the Lab collection were mainly produced by the Students Anna Geiger and Dominic Fiegle as a part of their bachelor projects in connection with this work. The rest were made by Christine Kammel and Hans Ernsberger. The insert and vector for each construct is listed. For the constructs made in this study, the primers used for amplification are also listed.

Table 5. List of vectors used in this study

vector	Description	Insert, Vector, primer	Source
Vec1	ALY1 -GAL4 AD for Y2H assay	ALY1 CDS, pGADT7, P55+P56	This study
Vec2	ALY2 -GAL4 AD for Y2H assay	ALY2 CDS, pGADT7	Lab collection
Vec3	ALY3 -GAL4 AD for Y2H assay	ALY3 CDS, pGADT7	Lab collection
Vec4	ALY4 -GAL4 AD for Y2H assay	ALY4 CDS, pGADT7	Lab collection
Vec5	UAP56 f-GAL4 AD for Y2H assay	UAP56 CDS, pGADT7	Lab collection
Vec6	ALY1-GAL4 BD for Y2H Assay	ALY1 CDS, pGBKT7	Lab collection
Vec7	ALY2-GAL4 BD for Y2H assay	ALY2 CDS, pGBKT7	Lab collection
Vec8	ALY3-GAL4 BD for Y2H assay	ALY3 CDS, pGBKT7	Lab collection
Vec9	ALY4-GAL4 BD for Y2H assay	ALY4 CDS, pGBKT7	Lab collection
Vec10	UAP56-1-GAL4 BD for Y2H assay	UAP56-1 CDS, pGBKT7, P57+P58	This study
Vec11	ALY1-GFP for transformation into Col-0	ALY1p::ALY1, pGreenII0229-3´GFP	Lab collection
Vec12	ALY2-GFP for transformation into Col-0	ALY1p::ALY1, pGreenII0229-3´GFP	Lab collection
Vec13	ALY3-GFP for transformation into Col-0	ALY1p::ALY1, pGreenII0229-3´GFP	Lab collection
Vec14	ALY4-GFP for transformation into Col-0	ALY1p::ALY1, pGreenII0229-3´GFP	Lab collection
Vec15	UAP56-GFP for transformation into Col-0	UAP56-1p::UAP56-1, pGreenII0229-3´GFP	Lab collection
Vec16	TEX1-GFP for transformation into *tex1-4*	TEX1p:TEX1, pGreenII0229-3´GFP, P59+P60	This study
Vec17	MOS11-RFP for transformation into tex1-4/TEX1-GFP	MOS11p::MOS11, pGreenII0179-3´RFP, 61+P62	This study
Vec18	SR30-RFP transformation into PSB-D	SR30 CDS, pGreenII0179-3´RFP	Lab collection
Vec19	UAP56 -SG tag for affinity purification	UAP56 CDS, pCambia2300-35S-3´SG, P63+P64	This study
Vec20	TEX1-SG tag for affinity purification	TEX1 CDS, pCambia2300-35S-3´SG, P65+P66	This study
Vec21	MOS11-SG tag for affinity purification	MOS11 CDS, pCambia2300-35S-3´SG, 67+P68	This study

2.1.5 Antibodies

The antibodies used for western blot and immunofluorescence analysis in this study, are listed in Table 6. The secondary antibodies were conjugated to horseradish peroxidase (HRP), fluorescein isothiocyanate (FITC) or Indodicarbocyanine (CY3).

Table 6. Overview of antibodies used in this study.

Epitope	Host	Conjugate	Application	Dilution	Company
GFP	mouse	none	Immunoblotting	1:2000	Roche
Mouse IgG	rabbit	HRP	Immunoblotting	1:5000	Sigma
UAP56	rabbit	none	Immuno-localization	1:300	Kammel et al.,2013
Fibrillarin	mouse	none	Immuno-localization	1:300	Abcam
Mouse IgG	goat	FITC	Immuno-localization	1:300	Dianova
Rabbit IgG	goat	CY3	Immuno-localization	1:300	Dianova

2.1.6 Bacterial and yeast strains

Table 7. Different microbial strains used in this work.

Name	Resistance	purpose	company
-*E. coli* XL1-blue	Tetracycline	Plasmid amplification	Stratagene - USA
-*A. tumefaciens* GV3101 with pSoup	Gentamycin / Tetracycline	Plant transformation	DSMZ - Germany
-*S.cerevisiae AH109*	-trp -leu -ade -his	Yeast-two hybrid protein interaction experiment	Clontech - Japan

2.1.7 T-DNA lines and plant cell culture

The T-DNA lines used for this thesis are listed in Table 8. PSB-D cell culture of *Arabidopsis landsberg erecta* used in this study was acquired from Geert De Jaeger (VIB, Belgium).

Table 8. Overview of T-DNA lines used in this study. Most of the seeds were bought at Nottingham Arabidopsis Stock Centre (NASC)

Name	T-DNA	AGI	Source
aly1-1	SAIL_381_E08	AT5G59950	NASC
aly2-1	wiscDsLox461-464N10	AT5G02530	NASC
aly3-1	SALK_063320.55.50.x	AT1G66260	NASC
aly4-1	GK-497B06	AT5G37720	NASC
mos11-2	SAIL_266_E03	AT5G02770	NASC
tex1-4	Salk 100012	AT5G56130	NASC
pid-14	SALK_049736	AT2G34650	U. Hammes, University of Regensburg

2.1.8 Antibiotics

Antibiotics used for selection of transgene bacteria and plants are listed in Table 9, together with information about working concentrations, solvent and provider.

Table 9. Antibiotics used in this study.

Name	Concentration stock solution	Final concentration	Solvent	Provider
Ampicillin	100 mg/ml	100 µg/ml	H_2O	Roth
Carbenicillin	50 mg/ml	500 µg/ml	H_2O	Duchefa
Gentamycin	50 mg/ml	50 µg/ml	H_2O	Duchefa
Kanamycin	50 mg/ml	50 µg/ml	H_2O	Roth
Hygromycin B	500 mg/ml	30 µg/ml	H_2O	Duchefa
Tetracycline	6 mg/ml	12 µg/ml	2-propanol	Sigma-Aldrich
Vancomycin	50 mg/ml	500 µg/ml	H_2O	Duchefa

2.1.9 Software and internet sources

Table 10. List of Software and internet sources used in the creation and analysis of the data presented in this study.

Software and internet sources	Description	Reference
ImageJ 1.49d	Image processing and quantification	(Schneider et al., 2012)
R 2.15.2	Statistical Package	(R Core Team, 2015)
Rcmdr (plugin R)	Interface plugin for R	(Fox, 2005)
Microsoft office 2013	Office suite	
Seaview	Multiple sequence alignments	(Gouy et al., 2010)
Alpha view® Software Version 3.0.3.0	Software for Immunoblot detection	Alpha Innotech - Germany
RealPlex Software Version 1.5	Software for RT-qPCR	Eppendorf – Germany
qBase+ (Biogazelle)	Analysis of RT-qPCR data	(Hellemans et al., 2007)
Primer3	Design of primers for RT-qPCR	(Untergasser et al., 2012)
Zen 2 blue edition	Software for ApoTome Image processing	Zeiss - Germany
The Arabidopsis Information Resource (TAIR)	sequence and gene information	www.arabidopsis.org
AgriGO	GO analysis	(Du et al., 2010)
Protein Identifier Cross-reference (EMBL-EBI)	Mapping GI identifier to AGI identifiers	(Wein et al., 2012)
Needle (EMBOSS)	Pairwise sequence alignment	(McWilliam et al., 2013)
Vector NTI ver. 8	*In silico* construction of Plasmids	IThermo Fisher - USA
Conserved domains database	Domain search	(Marchler-Bauer et al., 2014)
Smart database	Domain Search	(Letunic et al., 2015)
SignalP 4.1	Signal peptide detection	(Petersen et al., 2011)

2.2 Methods

2.2.1 Microbial work

2.2.1.1 Cultivation of bacteria

For all bacterial strains (Table 7) used in this work, sterile LB-medium (1% w/v NaCl, 1% w/v Tryptone and 0.5% w/v yeast extract) was used as growth medium. When making plates with LB-medium 1.5% w/v agar was added prior to autoclaving. For selection, sterile filtered antibiotics (Table 7 and Table 9) was added to the medium after autoclaving. The *E.coli* strain XL1B was grown at 37°C and the A. tumefaciens strain GV3101 was grown at 28°C. Growth in liquid LB-medium was under agitation at 200rpm.

2.2.1.2 Production of heat shock competent *E.coli* and *A. tumefaciens*.

5ml of sterile LB medium was inoculated O/N with XL1B or GV3101 with the proper selection (Tetracycline and Tetracycline/ Gentamicin respectively). The inoculum was then added to 100ml sterile fresh LB medium with selection to a final OD_{600} of

0.1. After growing in an incubator at 30°C or 37°C at 200rpm and reaching an OD_{600} of 0.4, the cells were spun down by 10min centrifugation at 3000xg. The pellet was suspended in 30ml sterile filtered TBF1 (100mM RbCl, 10mM $CaCl_2$, 50mM $MnCl_2$, 30mM NaOAc and adjust pH to 5.8 with acetic acid) buffer and incubated on ice for 90min. The cells were spun down by 10min centrifugation at 3000xg and suspended in 4ml of sterile filtered TBF2 (10mM MOPS, 10mM RbCl, 75mM $CaCl_2$ and 15% v/v glycerol) buffer and stored at -80°C in aliquots of 200µl.

2.2.1.3 Transformation of *E.coli*

Heat shock competent XL1B cells were thawed on ice. 20ng of the plasmid to be transformed was added, mixed and incubated together on ice for 20min. The heat shock was applied at 42°C for 2min. The cells were then cooled on ice and 1ml of sterile LB medium was added. The cells were incubated in a shaking incubator at 37°C for 1hr. The cells were spun down and suspending in a bit LB medium, before being transferred to solid LB medium with selection. The plates were incubated over night at 37°C.

2.2.1.4 Transformation of *A. tumefaciens*

Heat shock competent GV3101 cells were thawed on ice. 1µg of the plasmid to be transformed was added, mixed and incubated together on ice for 5min. The cells were then incubated in liquid nitrogen for 5min. The heat shock was applied at 37°C for 5min. The cells were then cooled on ice for 5min and 1ml of sterile LB medium was added. The cells were incubated in a shaking incubator at 28°C for 2-4 hr. The cells were spun down and suspending in a bit LB medium, before being transferred to solid LB medium with selection. The plates were incubated over night at 28°C.

2.2.1.5 Cultivation of yeast cells

The yeast strain AH109 was streaked out and grown on plates of solid YPAD medium (2% w/v agar, 2% w/v Tryptone, 1% w/v yeast extract, 2% w/v glucose and 0.004% w/v adenine hemisulfate) at 30°C for 3-4 days. Selection was done on synthetic dropout medium (SD). Two types of SD medium was used i.e. Double dropout (DDO) (2% w/v Glucose, 0.67% w/v Yeast nitrogen base w/o amino acids, 2.2% w/v agar, 0.064% w/v-Leu/-Trp Do supplement pH adjusted to 5.8 and autoclaved for 15min.) medium and Quadruple drop out (QDO) (2% w/v Glucose, 0.67% w/v Yeast

nitrogen base w/o amino acids, 2.2% w/v agar,0.064% w/v Ade/-His/-Leu/-Trp/ Do supplement pH adjusted to 5.8 and autoclaved for 15min.) medium.

2.2.1.6 Production of heat shock competent yeast

A single colony of AH109 was grown O/N in 3 ml liquid sterile YPAD medium. The inoculum was diluted to an OD_{600} of 0.1 in 50ml YPAD medium, and grown in a shaking incubator at 30°C and 200rpm until an OD of 0.5-1 was achieved. The cells were pelleted by centrifugation at 500xg for 5min. The cells were washed in 25ml sterile MQ water and pelleted by centrifugation at 500xg for 5min. The cells were then washed by 5ml sterile filtered SORB buffer (100mM LiOAc, 10mM Tris, 1mM EDTA, 1M sorbitol, adjust to pH8). The cells were spun down, supernatant removed and washed in 500µl SORB buffer. The cells were spun down again and re-suspended in 360µl SORB buffer. 40µl of denatured single stranded 10mg/ml salmon sperm DNA was added. The salmon sperm was denatured at 100°C and quickly cooled on ice, prior to addition. Cells were aliquoted in 50µl portions and storage at -80°C.

2.2.1.7 Transformation of yeast.

Competent AH109 cells were thawed on ice. 500ng of vector was added to the cells, and the suspension was mixed. 300µl of sterile filtered PEG (100mM LiOAc, 10mM Tris, 1mM EDTA, 40% PEG3350 and the pH was adjusted to 8) buffer was added to the 50µl cells. The solution was mixed and incubated 30 min at room temperature, before adding 40µl pure DMSO. Heat shock was applied at 42°C for 15 min. Cells were pelleted 2min at 500xg, The pellet suspended in selective medium (DDO or QDO) and plated out on auxotrophic SD plates at 30°C for 3-4 days.

2.2.2 Plant work and Cell biological methods.

2.2.2.1 Cultivation of *A. thaliana* on soil

Seeds were sown out on soil (10% perlite, 10% sand, 80% Profisubstrat (Einheitserde), 30g osmocote start (Everris)) in pots that were stratified for 48-72hrs at 4°C. After stratification, the pots were moved to a phyto chamber to grow under either Long day (LD) conditions (16hrs light and 8hrs darkness at 23°C) or Short Day (SD) conditions (8hrs light and 16 hrs darkness 18°C). The Light intensity was measured in PPFD and adjusted to $100 \mu mol \cdot m^{-2} \cdot sec^{-1}$. For selection of plants harbouring a construct with the nos-bar cassette (pGreenII0229; Vector 11-16) the young seedlings

were sprayed two to three times with glufosinate solution (100mg/l Basta®, 200µl/l Silwet® in H_2O) with 4 days in between.

2.2.2.2 Cultivation of *A. thaliana* on MS plates and filter paper

Seeds were surface sterilized by 15min incubation in 70% ethanol on a shaking plate. They were then further sterilized by adding a chloride solution for 2 min (37.5% v/v household bleach, 0.1% tween 20) and subsequently rinsed twice with sterile MQ. They were then sown out on solid ½ MS-medium (2.15 g/l Murashige and Skoog media including vitamins, 1% sucrose, 0.7% phyto agar (w/v), and pH 5.8, sterilized by autoclaving). When examining roots of seedlings, MS plates with 1% phyto agar (w/v) were grown vertically to allow unhindered growth of the root. All MS plates were grown in a Plant incubator (Percival Scientific - USA) under LD conditions after stratification. For selection of Plants transformed with vectors conferring Hygromycin resistance (pGreenII0179; Vec 17) transformed seeds were sown out on plates with 30µg/ml Hygromycin B added. For the Immuno-localization studies, seeds were sown out in a Petri dish on wet filter paper, stratified and grown vertically under LD conditions in the Plant incubator.

2.2.2.3 Stable transformation of *A. thaliana.*

Wildtype or mutant plants were transformed by the floral dipping method (Clough & Bent, 1998). Plasmids for transformation into *A. thaliana* were introduced into *A. tumefaciens* by heat stress induced transformation. Positive colonies were tested by colony PCR and grown (O/N) in 5ml LB medium with triple antibiotic selection (Gentamycin, Tetracycline and Kanamycin). The 5ml culture was used to inoculate 500ml LB medium with selection. The *Agrobacterium* were sedimented by 15 min centrifugation at 5000xg. The bacteria pellet was re-dissolved in 500ml infiltration medium (5% sucrose, 10mM $Mgcl_2$, 0.02% Silwet L-77 and 10µM Acetosyringone. The aerial part of the plant with flowers were dipped in the infiltration medium for 1min, left O/N on plastic and moved to plant room.

2.2.2.4 Cultivation of *Arabidopsis thaliana landsberg erecta* PSB-D cell culture.

The cell cultures were weekly diluted in a sterile bench, by adding 7ml old culture to a sterile 100ml Erlenmeyer flask. 43ml of fresh sterile MSMO (0.443% Murashige and Skoog Salt mixture (USBilologcal), 3% sucrose medium, 0.5mg/l NAA (dissolved in 100mM NaOH), 100mg/l myo-inositol, 0.05mg/l Kinetin, 0.4 mg/l thiamine, adjust-

ed pH to 5.7 with 1M KOH and autoclaved) medium. Plates were added 0.8% plant agar before autoclaving. The cells were kept under dark conditions in a shaking incubator with 130rpm at 23°C, in between dilutions.For upscaling, the two freshly diluted 50ml cultures were prepared. On the second week, 40ml of the old cultures were diluted with 180ml fresh MSMO medium in two sterile 1l Erlenmeyer flasks. After this step, selection was omitted form the medium. In the third week, the two old cultures were used to transfer 40ml cell culture into ten 1l Erlenmeyer flasks with 180ml MSMO medium in each. In the fourth week each of the ten cultures were decanted into a 2l Erlenmeyer flask with 800ml sterile MSMO medium in it. After three days of growth, the cells were harvested by passing the medium through Mira cloth suspended over a 5l beaker glass. The cells were weighed off in batches of 15g, packed in aluminium foil and kept at -80°C.

2.2.2.5 Transformation of *Arabidopsis thaliana landsberg erecta* PSB-D cell culture.

The method is based on (Van Leene et al., 2011). Plasmids were transformed into *Agrobacterium* and a single colony was grown overnight in 2ml LB medium with Kanamycin, Gentamycin and Tetracycline selection. The next day the culture was transferred to 20ml fresh medium with selection and grown O/N. The third day the cells were transferred to a 50ml Greiner tube and spun down for 15min at 3000xg. The pellet was suspended in 40ml sterile MSMO and vortexed. The cells were pelleted again and suspended in 40ml fresh sterile MSMO. The OD_{600} was measured. The cells were pelleted again and suspended with MSMO to an OD_{600} of 1.0. Meanwhile 3ml of Arabidopsis cell culture (OD_{600} of 1.2-1.3) that had been diluted three days prior were filled in four wells of a sterile six well multiwell plate. To each well 200µl of the *Agrobacterium* solution was added together with 6µl of a 100mM Acetosyringone stock. The plate was incubated in a shaking incubator (130rpm) at 23°C for three days. The content from two wells (6ml) were pooled in a 25ml Erlenmeyer flask and added 8 ml MSMO medium with Vancomycin and Carbenicillin (Table 9) to kill of the *Agrobacterium*. For selecting the positive transformants 50µg/ml kanamycin (Vector 19-21) or 20µg/ml, Hygromycin B (Vec18) was added. At the same time the culture from the two other wells were washed with 40ml sterile MSMO medium and the cells were transferred to solid MSMO plates with proper selection. After growing a week, the liquid culture was transferred to a 100ml Erlenmeyer flask and

added 25ml MSMO with selection. After another week the cell culture was let to sediment, and the cells were transferred to 35ml fresh MSMO with selection. After this step, only Kanamycin or Hygromycin was used as selection. The cells were diluted like described in Chapter 2.2.2.4 from this point on. If the culture stopped growing, the backup cells from the plate was transferred to liquid MSMO medium, and grown as described above.

2.2.2.6 Crossing of plants.

Plants with different genotypes were used for crossing to obtain double mutants. Unopened flowers of one genotype were emasculated by removing all anthers, sepals and petals with a crossing tweezer. The pollen from the other genotype were then placed on the stigma of the first genotype leading to fertilization, and double heterozygous offspring. The T2 generation was sown out and screened for double-mutants by PCR based genotyping.

2.2.2.7 Soil-based Phenotyping

The plants were cultivated as described above with the exception that the pots were shuffled every day to avoid that position of the pots on the shelves would affect the phenotypes. The time of flowering was measured as the amount of leaves present at the emergence of the flower bud. The amount of auxillary meristem branches was counted by removing all leaves from plants 60 days after stratification (DAS) with a tweezer. Each inflorescence was then separated at the bottom of the plant to avoid confusing primary and secondary inflorescences. For analysis of the flower and siliques, the whole and disassembled flower organs were positioned on a 0.8% phyto agar gel. Pictures were taken with a digital camera or with a Zeiss Discovery V8 stereo microscope. Siliques were bleached by incubating them O/N in a 3:1 ethanol acetic acid solution. The siliques were washed in 70% ethanol two times and observed under the microscope. For counting the amount of non-developed ovules, flowers upon opening were marked for each phenotype. Twelve days later the siliques were collected, and incubated for an hour in water. The siliques were then opened with two 27G syringes under a dissecting microscope, and the non-developed ovules and normal developing embryos were counted.

2.2.2.8 Phenotypical analysis of roots.

Roots were grown lateral on ½ MS plates (1% phyto agar). The amount of lateral roots in the branching zone was counted by looking at the root under a Nikon SMZ645 stereo microscope. All elongated lateral roots protruding the epidermis were counted. The last protruded lateral root was marked, and a picture was taken. For each root, the length of the branching zone was measures from the beginning of the root to the last protruded Lateral roots. The length of the branching zone was measured from the pictures by using the measurement tool in ImageJ. The branching density was calculated by dividing the amount of Lateral roots with the length of the Branching zone for each root of each genotype, like recommended in Dubrovsky & Forde, 2012. The total length of the roots was also measured on these pictures from beginning to the tip of the root. All phenotyping analyses were done at least twice.

2.2.2.9 Isolation of protoplasts from *Arabidopsis* seedlings

Protoplast isolation was done according to Zhai et al., 2009. 2g of seedlings grown on ½ MS medium were harvested 14DAS in a sterile bench. 15ml of sterile filtered TVL (0.3M sorbitol, 50mM $CaCl_2$) buffer was added, and the seedlings were chopped to fine pieces with a razor blade. The seedlings were transferred to a 200ml beaker glass, and 20ml sterile-filtered enzyme solution (0.5M sucrose, 10mM MES-KOH (pH5.7), 20mM $CaCl_2$, 40mM KCl, 1% w/v Cellulase (Onozuka R-10) and 1% w/v Macerozyme (R10)) was added. The seedlings with enzymes were allowed to incubate with gentle shaking (30rpm) for 18hrs. The released protoplasts were separated from the plant debris, by filtering the solution through eight layers of cheesecloth into a 50ml centrifuge tube. The cheesecloth had been pre-soaked with 5ml sterile filtered W5 (0.1% w/v glucose, 0.08% w/v KCl, 0.9% NaCl, 1.84% w/v $CaCl_2$, 2mM MES-KOH (pH5.7)) solution, and was rinsed afterwards with 10ml W5 solution. A layer of W5 buffer was carefully put on top of the filtered solution, giving rise to two separated phases. The solution was centrifuged in a swing-bucket rotor at 100xg for 7min with brakes turned off and slowest acceleration and de-acceleration settings. The protoplasts were then collected by removing 10ml solution from the interphase and transferring it to another 50ml centrifuge tube. 15ml W5 buffer was added to the protoplasts, and they were centrifuged 5min at 60xg. Supernatant was removed and the protoplasts were washed with 15ml W5 solution and spun down 5min at 60xg. The protoplasts were dissolved in 1ml W5 solution. The number of cells were then counted in a Fuchs Rosenthal counting chamber under a microscope.

2.2.2.10 Extracting root nuclei and immunolocalization analysis.

Roots were extracted at three DAS from seedlings grown on filter paper placed vertically. The roots were incubated in fixation solution(4% Paraformaldehyde(dissolved in MQ at 70°C),1x PBS (0.8% w/v NaCl, 0.02% w/v, 0.268 Na_2HPO_4-7 H_2O and 0.024% w/v KH_2PO_4) and pH was adjusted to 7.5 with H_2SO_4) under vacuum at RT for 20min. The roots were washed three times for 5min with 1x PBS, before adding an enzyme solution (0.7% w/v Cellulase R-10, 0.7% w/v Cellulose, 1% w/v Pectolyase and 1% w/v Cytohelicase) for 30min. The enzyme solution was removed and 1x PBS was added to the roots. After stirring root tips were released from the root, and were transferred by pipette to an object slide. A cover slip was added and the root tips were squashed by applying pressure on the cover slide with a toothpick. The object slide was dipped in liquid nitrogen for 10 seconds, and the cover slip was removed by striking it on the edge with razorblade. The object slide was put into a 1x PBS solution. Blocking (4% BSA w/v, 0.1% w/v Tween20, 0.1% w/v Triton X-100 and 1x PBS) solution was added to the object slide, and incubated for 1hr covered in Parafilm. Slides were washed three time with 1x PBS, and primary antibodies (Table 6) diluted in 1xPBS with 1% w/v BSA was added. After incubation, covered in Parafilm O/N at 4°C, the slides were washed three times with 1xPBS. The fluorescently labelled secondary antibodies (Table 6) were added and incubated for 1hr, before being washed three times with 1x PBS. The object slides now containing the isolated and prepared nuclei were mounted in a 4',6-Diamidin-2-phenylindol (DAPI) solution (VECTASIELD® mounting media) and covered with a cover slip. The sub-nuclear distribution of the fluorescent signals were analysed using a Zeiss Axio scope epifluorescence microscope. An Acroplan 100X/1.25 objective with oil was used. Different filter sets were used 36-513 DAPI HC (DAPI), F46-002 EGFP ET (FITC) and F36-504 TxRed HC (CY3) from AHF Analysentechnik.

2.2.2.11 Whole mount in situ hybridization of *Arabidopsis* seedlings.

The method was done according to (Gong et al., 2005). Seedlings grown laterally for six days were collected in a small sieve and submerged in a 1:1 solution of Heptane and fixation (120mM NaCl, 7mM Na_2HPO_4, 3mM NaH_2PO_4, 2.7mM KCl, 0.1% v/v Tween20, 80mM EGTA, 10% v/v DMSO and 5% formaldehyde) buffer for 30min at RT. The seedlings were washed twice with 99.8% ethanol and trice with 100% methanol for five minutes each. The seedlings were then incubated for 30min. in a

1:1 ethanol/Xylene solution. The seedlings were subsequently washed for 5min. twice with ethanol, then twice with methanol and lastly in a 1:1 methanol/fixation buffer without formaldehyde. Post fixation was done by adding the fixation buffer with formaldehyde for 30min. The seedlings were washed twice for 5min in fixation buffer without formaldehyde, and incubated for 5min in PerfectHyb^{tm} plus (Sigma) Hybridization buffer. The hybridization buffer was exchanged for fresh buffer and incubated an hour at 50°C in a hybridization oven. 25pmol of a 48mer of poly(dT) probe labelled at the 5'end with Alexa Fluor® 488 dye was added, and incubated O/N. The next morning the seedlings were washed once for an hour with 2xSCC (0.15M NaCl, 0.015M sodium citrate) solution added 0.1% v/v SDS, and then washed 20 min with 0.2xSCC solution added 0.1% v/v SDS. The seedlings were mounted on an object slide in DAPI solution (VECTASIELD® mounting media), and the roots were examined by CLSM using a Leica SP8 microscope. The Alexa Fluor 488 molecule was excited with a laser at 488nm, and the signal was detected with hybrid detectors at 500-550nm. To localize the nuclei, DAPI was used as a nuclear marker. DAPI was excited at 405nm, and detected at 410-450 nm. A HC PL APO 40X/1.3 objective with oil was used. A z-stack of twelve optical slices with 0.8 µm in between were acquired in the plane around the xylem vessel elements. Quantification of the GFP signals in cytoplasm and nucleoplasm was done in ImageJ. A maximum projection of the z-stack was done. The median intensity value of the nucleus and an area of the same size in the cytoplasm adjacent to the nucleus were quantified. The experiments were repeated twice unless mentioned otherwise.

2.2.2.12 *In vivo* Localization of GFP and RFP fused proteins in *Arabidopsis*.

Seedlings were grown for five days on ½ MS plates placed vertically and mounted in water on an object slide. First three independent lines of each construct, were analysed with a Zeiss Imager M2 ApoTome microscope to assure that the expression pattern was alike for all three lines. A 20X/0.8 plan apochrome dry objective and a 40X/1.4 plan apochrome objective with oil were used. After this, nuclei were analysed more in detail for one of the independent lines using the SP8 confocal microscope with a HC PL APO 63X/1.4 CS2 objective with oil. RFP was exited at 561nm and detected at 570-630nm.

2.2.2.13 Electron microscopy

Tissue to be analysed with electron microscopy, was submerged in fixation (50mM phosphate buffer pH 7.3, 2% v/v glutaraldehyde and 2% v/v formaldehyde) buffer and kept there. The samples were sealed and sent in the fixation buffer to Dr. M. Melzer (IPK Gatersleben) where the electron microscopy was done according to (Lolas et al., 2010)

2.2.3 Molecular Biology methods

2.2.3.1 Polymerase chain reaction

The polymerase chain reaction (PCR) was used to amplify DNA fragments for vector construction, genotyping, colony PCR and reverse transcriptase PCR (RT-PCR). The reaction was performed using a Tgradient or T3000 PCR thermocycler (Biometra). For genotyping, colony and reverse transcriptase PCR the Taq DNA polymerase (Peqlab) was used, whereas the proof reading KAPA HiFi Polymerase was used to amplify inserts for constructing new vectors. The PCR programs and reaction setup used for amplification with these two polymerases can be seen in Table 11 and Table 12. Amplified fragments were analysed on 1-2% agarose gels depending on the size of the fragments. Amplified DNA was stored at -20°C

Table 11. PCR cycle programs for Taq and KAPA HiFi polymerases used in this thesis

PCR steps	Taq DNA polymerase (Temperature, duration)	KAPA HiFi Polymerase (Temperature, duration)	Cycles
Initial denaturation	95°C, 300sec	95°C, 300 sec	1
Denaturation	95°C, 30sec	98°C, 20 sec	Taq (25-36) Kappa (34)
Annealing	58°C, 30.sec	58°C, 15 sec	
Extension	72°C, 1 min/1000bp	72, 1min/1000bp	
Final elongation	72°C, 300sec	72°C, 300 sec	1

Table 12. Reagents used for Taq and KAPA HiFi PCR in this study.

Reagents	Taq DNA polymerase	KAPA HiFi Polymerase
Buffer	1x Reaction buffer Y (Peqlab)	1x KAPA HiFi Fidelity Buffer (Peqlab)
dNTP mix	0.2mM of each dNTP	0.3mM of each dNTP
Forward primer	0.5µM	0.3µM
Reverse primer	0.5µM	0.3µM
Polymerase	0.75U Taq DNA polymerase (Peqlab)	1U KAPA HiFi (Peqlab)
Template	4% v/v extracted DNA or bacterial culture. 8% v/v 1:10 cDNA library for RT-PCR	4% v/v 1:20 cDNA library or 2-20 ng vector
MQ water	Up to 25µl	Up to 50µl

2.2.3.2 Agarose gel electrophoresis

1-2% w/v agarose was added to 1x TAE (40mM Tris pH 8.0, 1mM EDTA) buffer, boiled in a microwave oven. 0.005% ethidium bromide was added and the solution was poured in a gel cassette with combs for wells, and left to solidify. Samples were analysed in a running chamber at 160V, where the agarose gel was submerged in 1x TAE buffer. Samples to be analysed was mixed with 6x loading dye (250mM Tris pH 7.5, 10% w/v SDS, 30% v/v glycerol, 0.5M DTT, 0.1% w/v bromophenol blue) to a final concentration of 1x and loaded in the wells of the agarose gel. DNA and RNA fragments were visualized using a Biometra TI5 system (Biometra).

2.2.3.3 Extraction of Genomic DNA from Arabidopsis seedlings.

A small leaf from a seedling grown on soil or plate was cut off and transferred to an Eppendorf tube with two glass beads, before being quick-frozen in liquid nitrogen. The leaf was ground for 10sec in a Silamat S5 homogenizer. 400µl Edward (200mM Tris pH 7.5, 250mM NaCl, 25mM EDTA, 0.5% SDS) buffer was added and the solution was vortexed for 5sec. The samples were then centrifuged at 20000xg for 5min at RT. The supernatant was transferred to a new tube, and added an equal amount of 100% 2-propanol. The samples were mixed thoroughly and incubated 2min at RT. Afterwards samples were centrifuged 20000xg at 4°C for 5min. The pellet was washed once with 70% ethanol and left to dry shortly before being dissolved in sterile MQ water. These DNA extractions were used for genotyping and amplification of inserts with promoter, 5'UTR, exon and introns for insertion into vectors.

2.2.3.4 Extraction of total RNA from Arabidopsis seedlings

0.1 g of frozen and ground seedlings harvested 10DAS, were added 750µl TRIzol® LS reagent and was vortexed thoroughly. After 10mins centrifugation at 12000xg (4°C), the supernatant was transferred to a fresh tube and incubated 5min at room temperature. Next 0.2ml chloroform was added, and the tubes were vigorously shaken for 15s, before being incubated for 15min at RT. The samples were centrifuged for 15min at 12000xg (4°C), and the upper aqueous phase was transferred to a new tube. To the samples 0.5ml of 100% 2-propanol was added, and was incubated at RT for 10min. The samples were then centrifuged for 10min at 12000xg (4°C). The pellet was washed with 75% ethanol and subsequently centrifuged for 5min at 7500xg (4°C). The supernatant was discarded, the RNA pellet was left to air dry for

5-10min. 70µl RNase free water, was added, and the tube was incubated at 60°C for 10min. The RNA was suspended by pipetting. The concentration and purity of the RNA was measured on a Nanodrop ND1000 spectrophotometer. Only extracted RNA with A_{260}/A_{280} and A_{260}/A_{230} ratios around 2.0 were used. The quality of the total RNA was controlled by loading 200ng on a 1% agarose gel. If no noticeable degradation of the 28S and 18S rRNA bands were observed the RNA was used or stored at -80°C. The RNA was used for northern blot analysis or making cDNA.

2.2.3.5 Northern blot analysis of Ta-siRNA and miRNA.

The method was based on Pall & Hamilton, 2008. A 12% denaturing polyacrylamide Urea SequaGel (National diagnostics) was poured and polymerized O/N. 15µg of total RNA in 20µl RNase free water was mixed 1:1 with RNA sample buffer (deionized Formamide, Xylene cyanol and Bromophenol blue), and heated at 95°C for 5 min. The gel was running for half an hour at 300V to preheat in 1xTBE buffer. The samples were loaded and separated for one hour at 400V. A marker was also loaded consisting of 19, 21 and 24bp long random dNTP oligos, labelled with [γ-32P] ATP as described below. The gel was stained with 1µg/ml Ethidium bromide in 1x TBE buffer for 10min and visualized on a Biometra TI5 gel documentation system. The RNAs were transferred to a nylon membrane, with a semi-dry blotting setup similar to that used for immunoblotting (Chapter 2.2.3.12) with water as buffer. After 30min of blotting at 20V, the RNA was crosslinked to the membrane by incubating the membrane in an EDC (31mg/ml 1-ethyl-3-(3-dimethyl-aminopropyl)-carbodiimide, 0.127M 1-methylimidazol and 12.5mM HCl) solution packed in plastic wrap for one hour at 50°C. The membrane was washed with water and incubated for one hour in 30ml QuickHyb hybridization solution (Agilent). In the meantime probes where labelled with [γ-32P] ATP. A reaction mix (1U T4 Polynucleotide kinase, 1x Polynucleotide kinase buffer A, 1µM dNTP probe and 20µCi [γ-32P] ATP) was incubated for 30min at 37°C. The labelled probe was purified from the unlabelled [γ-32P] ATP using an illustra MicroSpin G-25 Column according to manufacturer's description. The probe was added to the membrane in the hybridization solution and incubated one hour at 50°C. After the hybridization the membrane was washed for 5min twice with 50°C 5XSCC added 1% SDS and once with 1xSCC added 1% SDS. The membrane was wrapped in plastic foil and left 1-3 days on a phosphor imager screen. The autoradiography signal was detected with a Cyclone Storage Phosphorimager (Cyclone™

Storage Phosphorimager). Probes were stripped by incubating once in boiling water added SDS to a final concentration of 0.1% v/v and twice with pure boiling water in a plastic container placed in a water bath at 70°C. The membrane was packed in plastic foil and put 1-3 days on a phosphor imager screen, to ascertain that all probe had been stripped away, which evaluated by the absence of signal. More probes where hybridized and stripped in this manner. The intensity of the signals were quantified in ImageJ using a line scan and calculating the area under the curve. All signals were normalized to the area under the curve of the marker crosslinked to the gel and the U6 transcript as a reference gene. The experiments were done with four biological replicates

2.2.3.6 Construction of cDNA by reverse transcriptase.

Contaminating DNA was removed from the extracted RNA by incubating 3.5 µg RNA with 1x reaction buffer and 2 U DNase I in a reaction volume of 20µl for 105 min at 37°C. The DNase I was inactivated by adding 2µl of 25mM EDTA, and incubating 10min at 65°C. The RNA was kept on ice. The concentration was measured on the Nanodrop. The quality of the RNA was again controlled, by loading 200ng on a 1% agarose gel. If no degradation was observed 1.5µg of the RNA was used for cDNA synthesis. For cDNA libraries used in constructing plasmids the RevertAid™ H Minus M-MuLV reverse transcriptase and random hexamer primers were used according to manufacturer's description of how to generate cDNA (Thermo scientific). For cDNA used in RT-qPCR the superscript® IV reverse transcriptase and Oligo $d(T)_{18}$ primers were used according to manufacturer's description (Thermo scientific). All incubation steps were done in a thermocycler. DNA contamination was controlled by doing an RT- samples for each cDNA without the reverse transcriptase, and subsequent PCR to see that there was no amplicon due to DNA contamination.

2.2.3.7 RT-PCR

The cDNA libraries was diluted 1:10, and PCR was done with primers binding to the exon regions. For each primer pair, the PCR reaction was done more times with different cycle numbers. This was done to ensure the reactions being compared, all were in the exponential phase of the PCR amplification. To normalize the reference genes, ImageJ was used to quantify the intensity of the reference genes in different genotypes analysed on a 1% agarose gel. This was done by measuring the area un-

der the curve of a line scan of each lane with the reference gene. It was calculated how much should be loaded for each genotype on the agarose gel, to have equal amounts of the reference gene.

2.2.3.8 RT-qPCR

To design primers for RT-qPCR the web application primer3 (Untergasser et al., 2012) was used. Melting temperature was set to 65°C and length of the amplicons where in between 150-250 bp. The analyses were done in a Mastercycler® ep RealPlex thermocycler (Eppendorf). 20µl reaction were set up with 1x KAPA SYBR® FAST qPCR Master Mix, 0.2µM of each primer and 2-4µl template. Gradient PCR (50.5 -62.5°C) was used to establish the optimal annealing temperature of each primer pair. The efficiency of each primer was calculated as shown in Chapter 7.3. Primers with an efficiency between 80-110% were used. The RT-qPCR program was set up with first 3min at 95°C followed by 40 repeated cycles of 3sec at 95°C, 20sec at 56 or 62.5°C (depending on primer) and 8sec at 72°C. To ensure that no primer dimers were formed or unspecific PCR products a melting curve analysis was done after the run. The program was 15sec at 95°C, 15sec at 60°C, then a gradient from 60°C to 95°C in 20min and lastly 15sec at 95°C. The data was collected with the Eppendorf RealPlex Software ver.2.2. The quality control of the data and normalization with two reference genes where done in the software package qBase+ according to Hellemans et al., 2007. In qBase+ the Calibrated Normalized Relative Quantities (CNRQ) were calculated. The values were normalized to two reference genes. The CNRQ are the values showed in the figures. Three biological and three technical replicates were done. For the technical replicates a difference in C(t) value lower than 0.5 was accepted. In the rare case one replicate of three was outside the 0.5 range it was removed. RT- minus samples were included in each RT-qPCR analysis.

2.2.3.9 Construction of plasmids

Inserts were amplified either from a vector template or from a cDNAwith primers containing restriction sites. The amplicon was purified using Nucleospin® Extract II DNA purification kit (Macherey Nagel). The kit was used according to the manufacturer's protocol. The destination vector and the amplified insert was digested according to the manufacturer's protocol with the respective restriction enzymes (thermos scientific). The cut DNA was purified by cutting the bands out off a 0.8% agarose gel, un-

der soft UV light (354nm), and extracted using the Nucleospin® kit again. The vector was dephosphorylated using 5U of Antarctic phosphatase (NEB) in a 1x Antarctic phosphatase buffer incubated at 37°C for an hour. After purification with the Nucleospin® kit, the insert and vector was added to a ligation mix (5U T4 DNA ligase and 1x T4 Ligase buffer in total volume of 30µl) in 4:1 molar ratio and incubated at 4°C in a thermos canister O/N. The next day heat shock competent XL1B cells were transformed with the ligation mix like described in Chapter 2.2.1.3 , and plated out on solid LB medium with selection. Twelve of the positive colonies from the LB plate were each used to inoculate 3ml LB medium for a miniprep of plasmid. After growing O/N 1.5ml of each culture was added to an Eppendorf tube, and spun down at 1000xg for 3min. The cells were suspended in 200µl P1 (50mM Tris (pH 8.0), 10mM EDTA and 100µg/l RNaseA) buffer and given 300µl P2 (0.2M NaoH, 1% SDS) buffer. After 5min of incubation at RT, 300µl P3 (3M potassium acetate, pH 4.8) buffer was added, and the solution was incubated on ice for 10min. The solution was spun down 10min at 20000xg (4°C), the supernatant was transferred to a new tube and added an equal volume of 100% 2-propanol. The mixture was spun down for 10min at 20000xg (4°C), and the pellet was washed in 70% ethanol. After air drying, the pellet was dissolved in 50µl MQ water. 2µl of each miniprep was used for O/N restriction digest using the restriction enzymes from the primer, with which the inserts were amplified. The restriction digests were analysed on a 1% agarose gel. Clones were the insert was present according to the restriction analysis, was used to inoculate liquid 100ml LB medium. After growing O/N, the cells were pelleted in 50 ml Greiner centrifuge tubes at 1000xg for 10min. A maxi plasmid prep was then prepared using the Nucleobond® Xtra Midi plasmid DNA purification kit (Macherey Nagel) according to manufacturer's description. The plasmid preparation was then digested with different enzyme combinations to confirm that the insert was present in the right vector. Plasmids showing the right digestion patterns were sent for sequencing (MWG-Germany), with primers leading to sequencing of the borders and the whole insert. Plasmids were stored at -20°C.

2.2.3.10 Protein extraction

Proteins for immunoblotting were extracted from seedlings harvested at 10DAS. It was done according to Tsugama et al., 2011. Two complete seedlings harvested at 10DAS were ground with glass beads and boiled in extraction buffer (0.1M EDTA

(pH 8.0), 4% w/v SDS, 10% v/v β-mercaptoethanol , 5% v/v glycerol, 0.005w/v bromophenol blue) for 10min. This solution was loaded directly on to a 12% polyacrylamide gel and analysed by SDS-PAGE (Chapter 2.2.3.11).

2.2.3.11 Protein separation by SDS-page.

Polyacrylamide gels were cast by pouring the resolving gel (9 or 12% v/v Rotiphorese Gel 30 (37.5:1) added 0.15% w/v bisacrylamide, 0.75M Tris (pH 8.8), 0.2 % v/v SDS, 0.1% w/v ammonium persulfate and 0.02% v/v N,N,N",N"-tetramethylethylenediamine (TEMED)) solution into a Bio-Rad Mini-Protean® 3 Multicaster system (Bio-Rad). After resting and solidifying half an hour, a stacking gel (10% v/v acrylamide:bisacrylamide (30:0.15), 0.14M Tris (pH 6.8), 0.23 % v/v SDS, 0.11% w/v ammonium persulfate and 0.06% v/v TEMED) was added on top of the resolving gel. Combs were positioned in the gel, and after solidification, the gels were stored at 4°C for 3 weeks maximum. Proteins were separated in the polyacrylamide gels placed in a Bio-Rad Mini-Protean® 3 running chamber at 200 V using Laemmli running buffer (192mM glycine, 0.1% v/v SDS, 25mM Tris-HCl, pH 8.3). Prior to loading in the polyacrylamide gel samples not extracted like in Chapter 2.2.3.10 were added 6x SDS loading (250mM Tris (pH 7.5), 10% w/v SDS, 30% v/v glycerol, 0.5M DTT, 0.1% w/v bromophenol blue) buffer to a final concentration of 1x, and heated at 90°C for 10min. SDS gels were used for immunoblotting analysis (Chapter 2.2.3.12) or stained with Coomassie (30% v/v Ethanol, 10% v/v acetic acid, 0.2 w/v Coomassie Brilliant Blue R-250) solution. The gels were destained with (7.5% v/v ethanol and 5% v/v acetic acid), and documented with a digital camera.

2.2.3.12 Immunoblotting

Protein samples prepared like in chapter 2.2.3.10 were separated by SDS-PAGE. The proteins were blotted onto an Immobilon™ polyvinylidene fluoride membrane (Millipore) that had been activated in pure methanol for 30secs and then equilibrated in blotting buffer (20% v/v methanol, 0.2M glycine, 20mM Tris and 0.01 % v/v SDS) for at least 10min. The polyacrylamide gel was placed on the membrane, which was put on top of three pieces of Whatman paper equilibrated with blotting buffer. Three pieces of equilibrated Whatman paper was put on top of the gel, and the blotter was closed. The transfer of protein to the membrane was done with a Semi-dry Blotter (Roth) at 50mA per gel for 3hrs. The membrane was blocked in blocking buffer

(20mM Tris (pH 7.5), 150mM NaCl, 0.05% v/v tween20 and 5% w/v milk powder) buffer for 1hr. The primary mouse anti-GFP antibody (Roche) was added at a 1:2000 dilution and the membrane was incubated O/N at 4°C on a rotating incubator. The membrane was washed three times for 5min with wash buffer ((20mM Tris (pH 7.5), 150mM NaCl, 0.05% v/v Tween20 and 1% v/v Triton X-100). The secondary anti-mouse antibody coupled with horseradish peroxidase, was diluted 1:5000 in blocking buffer and incubated with the membrane for two hours. The membrane was washed 3 times for 5min with washing buffer. Then the membrane was incubated with SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) for 9mins. The chemiluminescence signal was detected on a FluorChem FC2 imager 2 MultiImager (Alpha Innotech) using the software package AlphaView

2.2.3.13 Pulse labelling

Isolated protoplasts (Chapter 2.2.2.9) was diluted in W5 buffer to a concentration $1 \cdot 10^6$ cells/ ml. 100µCi of L-[35S] methionine (Haartmann) was added and the cells were incubated for 1hr at RT. The cells were pelleted at 100xg and dissolved in 2x SDS loading buffer before being heated for 10min at 90°C. The proteins were separated on polyacrylamide gels using SDS-PAGE. The gels were stained with Coomassie staining. The amount of total protein was measured in ImageJ by doing a line scan and measuring the area under the curve. All samples were normalized so the same amount of total protein could be loaded on a new polyacrylamide gel. After staining with Coomassie, the gel was dried on a piece of Whatman paper under vacuum. The gel was put on a phosphor imager screen for 1-4 hours, before the signal was detected on a Cyclone™ Storage Phosphorimager (Cyclone™ Storage Phosphorimager). The total incorporated amount of L-[35S] methionine was measured in ImageJ by making a line scan and measuring the area under the curve. A ratio was calculated between the total L-[35S] methionine signal and the total protein signal.

2.2.3.14 Affinity purification.

The protocol was modified from Van Leene et al., 2011. 15g of ground PSB-D culture expressing a bait protein fused to the SG-tag was dissolved in 20ml extraction buffer (25mM Hepes (pH7.4), 0.05% IGEPAL CA-630, 1mM DTT, 2 mM $MgCl_2$, 5mM EGTA, 10% glycerol, 1 tablet/ 50 ml buffer cOmplete™, EDTA free proteinase inhibitor (Sigma-Aldrich), 1ml/100ml buffer 0.1 M PMSF dissolved in 2-propanol) in a 50ml

Greiner tube. The slurry was thawed and subsequently homogenized by sonicating 5 x 30sec at 27-30% intensity on a UW2070 MS73 (Bandelin) Sonicator. The extracts were added 5mM $MgCl_2$ and 2000u Benzonase endonuclease (Produced in the workgroup), and incubated for half an hour at 4°C to remove nucleic acids. The samples were centrifuged at 40000xg for 50 min at 4°C. The extract was then filtered through a 0.45µm filter. The protein concentration was measured with a Bradford assay, and 100mg of total protein from each extraction was adjusted to a volume of 30 ml. BcMag Epoxy-Activated magnetic beads (Bioclone, USA) were coupled to rabbit IgG (I5006, Sigma-Aldrich) according to manufacturer's description. 100µl beads washed three times with 0.5ml extraction buffer was added to the protein extracts and incubated for an hour at 4°C under rotation. The solutions were centrifuged at 100xg at 4°C for 10min. The supernatant was decanted and the beads were transferred to a 2ml Eppendorf tube. The beads were washed three times in extraction buffer, using a magnet rack to retain the beads. Proteins were eluted at RT by incubating 5min at RT in 300µl 0.1M Glycine (pH 2.7). The eluate was precipitated by addition of 1.2ml ice-cold acetone and incubation O/N at -20°C. Next day the eluate was spun down three times for 10min at 20000xg (4°C) and washed in fresh ice-cold acetone. The precipitate was dissolved in 30µl 2xSDS loading buffer and heated at 90°C for 10min. Residual acetone was removed by 2 min in a SpeedVac. The proteins were separated on a 9% polyacrylamide gel that was stopped when the bromophenol blue band had emigrated 1/3 into the gel. The gel was stained with Coomassie staining. The lanes were cut into five gel slices that digested by trypsin and sent for MS.

2.2.3.15 Trypsin digest

Gel slices were washed four times for 30min with 50mM NH_4HCO_3, 50mM NH_4HCO_3 /Acetonitrile (3/1), 10mM NH_4HCO_3/Acetonitrile (3/1) and 10mM NH_4HCO_3 /Acetonitrile (1/1) respectively. The slices were completely lyophilized and 25µl trypsin mix (0.133µg/µl Promega gold trypsin dissolved in1mM HCl) was given to the gel pieces. 40µl of 50mM NH_4HCO_3 was added and the digest was incubated at 37°C O/N. The next day all free liquid was removed from the gel pieces to a new 0.5ml Eppendorf tube. The peptides were extracted twice by addition of 40µl 100mM NH_4HCO_3 and incubated at 39°C for one hour. The supernatants were transferred to the same tube for each gel slice. In the third extraction step, 40µl 100mM

NH_4HCO_3/Acetonitrile (1:1) was added at 30°C and removed after an hour. The pooled supernatants were lyophilized O/N and stored until analysed by tandem mass spectrometry.

2.2.3.16 Identification of protein interaction-partners by tandem mass spectrometry.

The trypsin-digested peptides were analysed by Liquid chromatography-mass spectrometry (LC-MS/MS) in the lab of Prof. Dr. Deutzmann (Regensburg University). Separation of peptides was performed on an Ultimate3000 RSLC Nano-HPLC system (Thermo Fisher) with a reversed phase chromatography analytical column (ReproSil Pur 120 C18-AQ, 75µm x 25µm). The mobile phase consisted of a linear gradient containing 0.1% v/v formic acid (Eluent A) and 80% acetonitrile v/v in 0.1% v/v formic acid (eluent B). The HPLC-system was coupled to a Q-TOF mass spectrometer (MaXis plus – Bruker Daltonik) via a nano electron spray source. The mass spectrometer was operated in the data dependent acquisition (DDA) mode. Up to five of the most abundant precursor ions were selected for fragmentation by collision induced dissociation (CID). The NCBI database was searched with Mascot (v2.3.02) using the ProteinScape software (Bruker Daltonics) to identify raw MS data. The identified proteins were mapped to Arabidopsis gene identifiers (AGI) using the Protein Identifier Cross-reference (Wein et al., 2012) and the National Center for Biotechnology Information (NCBI) homepage (http://www.ncbi.nlm.nih.gov/). The experimental background was estimated from 4 affinity purifications with the empty SG-tag, and removed from the results. The remaining proteins with a mean score higher than 85 with at least two peptides were retained as possible interaction partners. Of these only those present in 3 out of 4 independent affinity purifications, were further considered. Well known sticky proteins from similar experiments published in Van Leene et al., 2015 were also omitted for final analysis.

3 Results

3.1 Characterization of ALY1, ALY2, ALY3, ALY4 and UAP56.

3.1.1 UAP56 interacts with ALY1, ALY2, ALY3 and ALY4 *in Arabidopsis*.

ALYREF is one of the main mRNA export adaptor in humans (Viphakone et al., 2012; Stubbs & Conrad, 2015). In *Arabidopsis* four paralogs of ALYREF are found. The four *Arabidopsis* ALY proteins are very similar on the sequence level (Table 13). The ALY3 and ALY4 proteins exhibit the highest degree of similarity (70.0% shared identity). The second highest degree of similarity are found between ALY1 and ALY2 with 54.4% shared identity. All the other combinations of the four ALY proteins exhibit a lower shared sequence identity (38.3-41.6%).

Table 13. Aly1/Aly2 and ALY3/ALY4 show high protein sequence similarity. The protein sequences were compared to each other in a pairwise manner using the EMBOSS Water pairwise sequence alignment tool (McWilliam et al., 2013). X represents redundant protein comparisons.

	ALY1	ALY2	ALY3	ALY4
ALY1	X	54.4%	39.4%	41.4%
ALY2	X	X	38.3%	41.6%
ALY3	X	X	X	70.0%
ALY4	X	X	X	X

In humans it is known that ALYREF interacts through the N and C terminal domains with UAP56 (Hautbergue et al., 2009). In order to see if these domains are conserved in *Arabidopsis thaliana* a multiple sequence alignment (MSA) was done (Figure 7). ALY1 (At5g59950) was used as query protein in a BLASTP search of the NCBI database (http://www.ncbi.nlm.nih.gov) to find orthologs in other species. The N and C terminal domains, together with the RNA recognition motif (RRM), seems to be conserved among the four *Arabidopsis* orthologs of ALYREF. The WQHD motif is conserved in ALY1, whereas the motif has insertions and mutations in ALY2 (WGHD), ALY3 (WQNQND) and ALY4 (WQSG). The ALYREF orthologs from the other species used in the MSA analysis have the conserved motif. The conserved N- and C terminal domains indicate that all four ALY proteins could all potentially interact directly with the *Arabidopsis* UAP56 protein.

In order to test this hypothesis the coding sequences (CDS) of *Arabidopsis* ALY1, ALY2, ALY3, ALY4 and UAP56 were introduced into the Matchmaker™ GAL4 yeast two-hybrid vector (Y2H) system (Clonetech). The CDS of the five proteins were introduced into both the pGBKT7 and pGADT7 vectors. Different combinations of

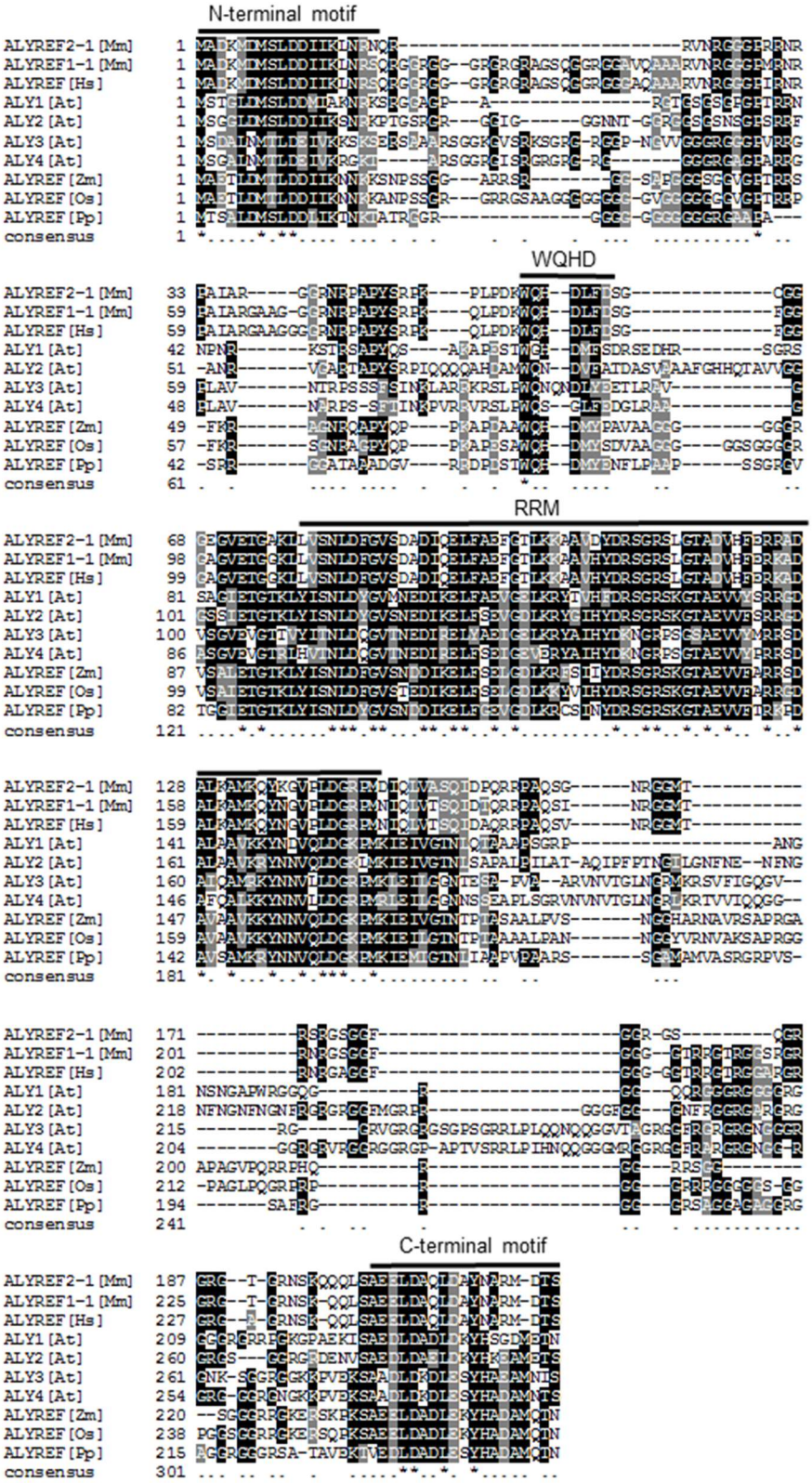


Figure 7. Multiple sequence alignment of ALYREF orthologs. The alignment was generated using the Clustal Omega tool of the Seaview package (Gouy et al., 2010) and the ALYREF sequences of *Mus musculus* [Mm], *Homo sapiens* [Hs], *Arabidopsis thaliana* [At], *Zea mays* [Zm], *Oryza sativa* [Os] and *Physcomitrella patens* [Pp]. Asterisks indicate invariant residues and (.) indicate highly conserved residues. Black residues indicate that the residue is identical to the consensus residue and grey residues indicate highly conserved residues in the consensus sequence. The conserved N-terminal, RRM, WQHD and C-terminal are marked in the alignment.

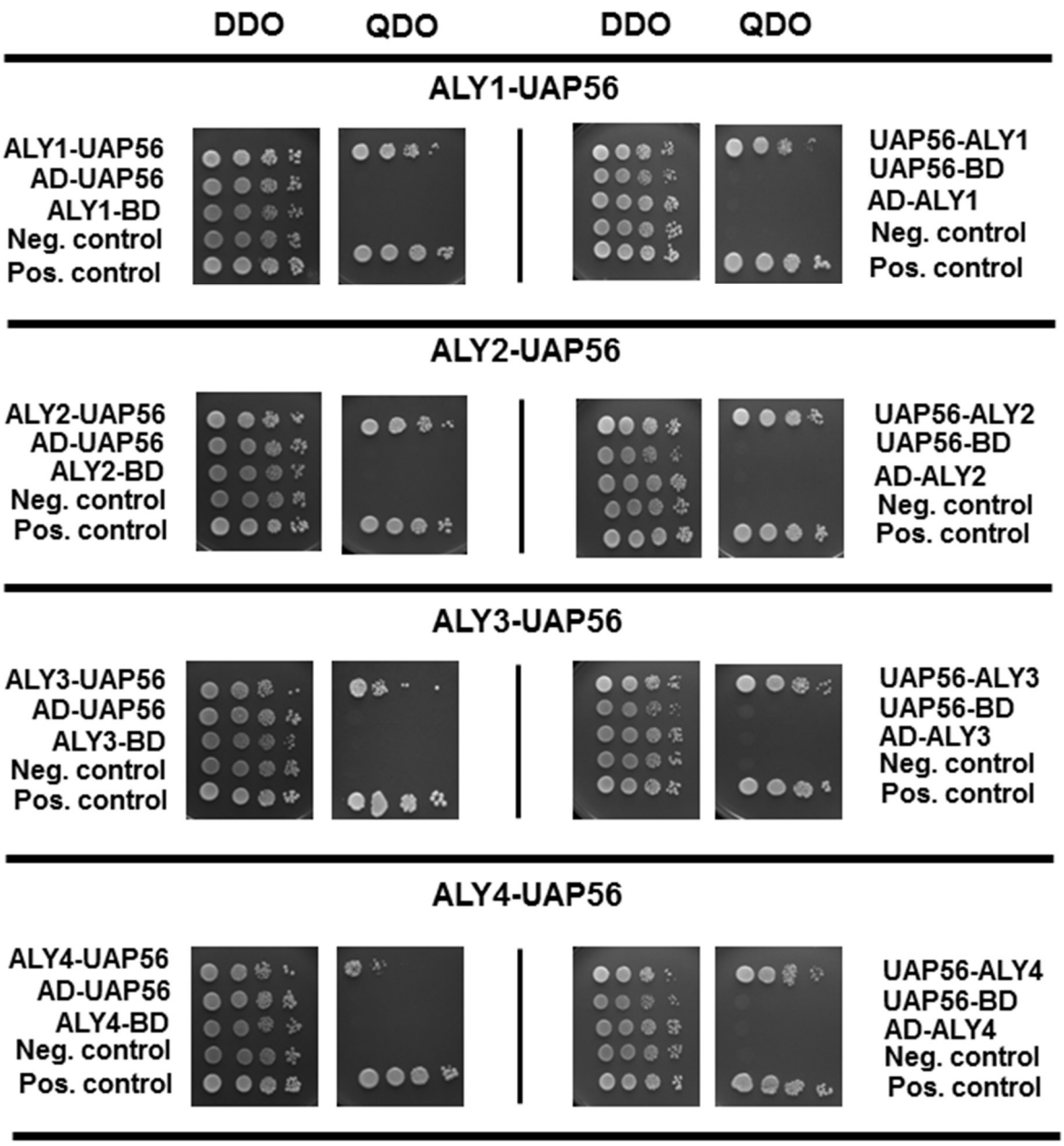


Figure 8. All four ALY proteins interact diretly with UAP56. Yeast cells cotransformed with plasmids leading to expression of the indicated proteins were grown on SD/–Leu/-Trp medium (DDO) and SD/-Ade/-Leu/-Trp/-His (QDO) minimal medium. In the two left side panels UAP56 has been fused to the DNA binding domain of GAL4, and the four ALY have been fused to the activator domain. In the two right side panels, UAP56 is fused to the activator domain and the four ALY proteins are fused to the DNA binding domain. Auto-transcriptional activation of the different fusion proteins, were tested by cotransfecting them with the empty pGADT7 (AD) or pGBKT7 (BD). The positive control was cells transformed with p53 and the SV40 large T-antigen, whereas cells transformed with Lamin and the SV40 large T-antigen was the negative control. The experiments were done twice in both directions.

these vectors were cotransfected into the auxotrophic yeast strain AH109. The Interaction between two proteins can be evaluated as growth on the Quadruple Dropout Minimal Medium (QDO) lacking ADE, LEU, HIS and TRP.

First it was tested if UAP56 fused to the GAL4 DNA binding domain (BD) interacts

with any of the ALYs fused to the GAL4 activator domain (AD) (Figure 8, left panels). Growing cells were seen on the DDO for all construct combinations and controls, indicating the transformations were successful. When examining the QDO plates, growing cells were only observed for the positive control and the yeast cells transfected with UAP56 and each one of the ALYs. For the cells transfected with UAP56 and the AD, no growth were seen, showing that UAP56 does not exhibit auto activation activity of the GAL4 induced promoter. Growing cells were also not seen for the transfections with the four ALYs and the BD. The reciprocal experiment was done with UAP56 fused to the AD, and the four ALYs fused to the BD (Figure 8, right panels). For the reciprocal experiment, growth was observed when the cells were co-transfected with UAP56 and one of the four ALYs, and not in the auto activation controld. Taken together this indicates that *Arabidopsis* ALY1, ALY2, ALY3 and ALY4 interacts directly with UAP56 in the yeast two-hybrid assay.

3.1.2 GFP fused ALY1, ALY2, ALY3, ALY4 and UAP56 exhibits differential sub nuclear localization

When examining publically available microarray expression data, it is seen that ALY1, ALY2 and ALY3 together with some of the other TREX orthologs are equally expressed in different tissues during different development stages, see Figure 9. ALY3 might be expressed more in the reproductive organs, relatively to the other tissues. There are no data available for ALY4. As all proteins are expressed in the roots, it was decided to look at the ALY expression pattern there. The subcellular localization of ALY1, ALY2, ALY3 and ALY4 have already been studied in *Arabidopsis* Col-0 cell suspension culture, Nico*tiana benthamiana* leaves and Nico*tiana benthamiana* BY-2 cells (Pendle et al., 2005; Storozhenko et al., 2001; Uhrig et al., 2004). These constructs were under the control of strong viral promoters I.e. the cauliflower mosaic virus promoter (35S) and tobacco rattle virus promoter. ALY1, ALY3 and ALY4 were in these studies found to be localizing in the nucleoplasm and nucleolus, whereas ALY2 was mostly found in the nucleoplasm. The subcellular localization of UAP56 fused to Green fluorescent protein (GFP) has been examined in *Nicotiana tabacum* BY-2 cell and Nico*tiana benthamiana* leaves (Kammel et al., 2013; Pan et al., 2014). These constructs were also under the control of the 35S promoter. In one study GFP-fused UAP56 was found mainly in the nucleus

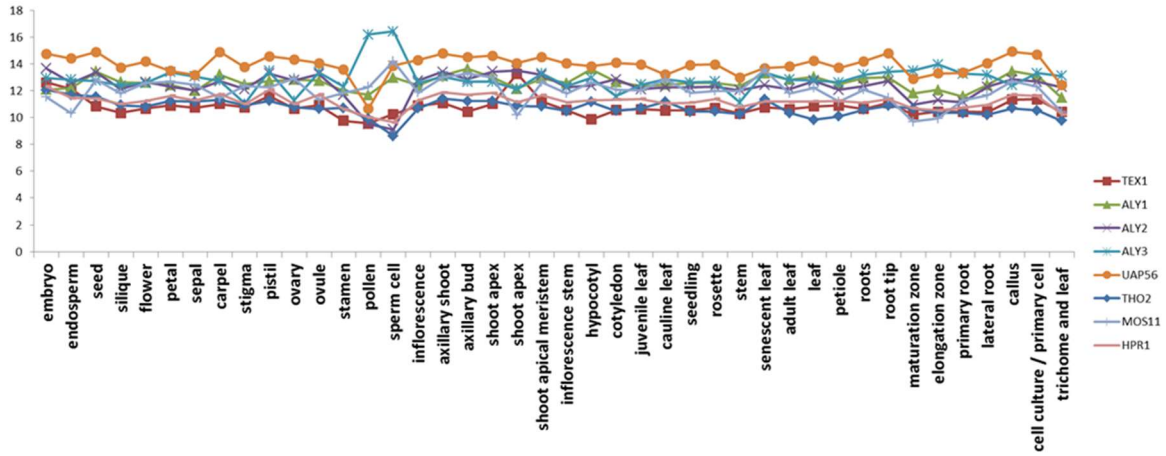


Figure 9. Publicly available expression data based on microarray experiments . The data was accessed through Genevestigator (Hruz et al., 2008), and processed in Excel. The y-axis depicts Level of expression (Log2(signal intensity)) in different tissues. The expression levels in different tissues were tested for the Arabidopsis *TEX*1, *ALY1*, *ALY*2, *ALY3, UAP56*, *THO2*, *MOS11* and *HPR1* genes. According to the Genevestigator interphase, genes are considered to have a medium expression level for log2 intensity values between 8-11.5. Below is considered low, and above is considered high expression

(Kammel et al., 2013). In the other UAP56 was present in both nucleus and nucleolus (Pan et al., 2014). As all these studies were done with strong constitutive promoters, it was decided to look at the expression pattern and sub-nuclear localization of GFP fusions in *Arabidopsis thaliana* roots under the control of the native promoters. In order to define the sub-nuclear compartments, and see the subcellular localization of UAP56, immunolocalization experiments were done. The α-UAP56 antibodies produced in Kammel et al., 2013 and commercial anti-Fibrillarin antibodies were used. Roots from *Arabidopsi*s seedlings harvested three days after stratification (DAS) were enzymatically digested to release root nuclei.

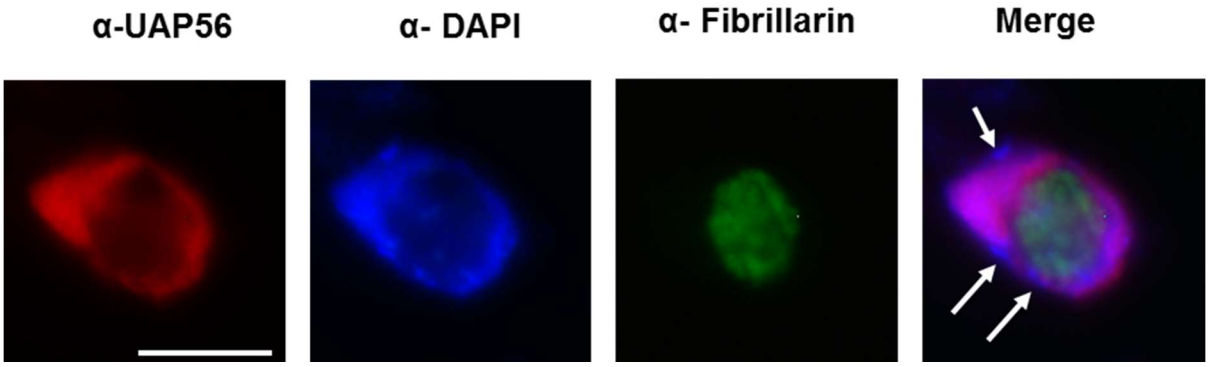


Figure 10. UAP56 exhibit specific sub-nuclear localization in the nucleoplasm The immunostaining was done on isolated root nuclei. Specific antibodies were used to show the nuclear and nucleolar localization of UAP56 and Fibrillarin respectively. DAPI was used as counter stain. A 100X /1.25 objective was used. Scale bar indicates 10µm.

DAPI was used as a marker of DNA defining the borders of the nucleus (Figure 10). The fluorescence signal of the secondary antibody binding to the α-Fibrillarin antibody showed the subnuclear localization of Fibrillarin, which was only found in the nucleolus. UAP56 was in contrast to Fibrillarin mostly present in the nucleoplasm and not prominently localized in the nucleolus. UAP56 was also found.

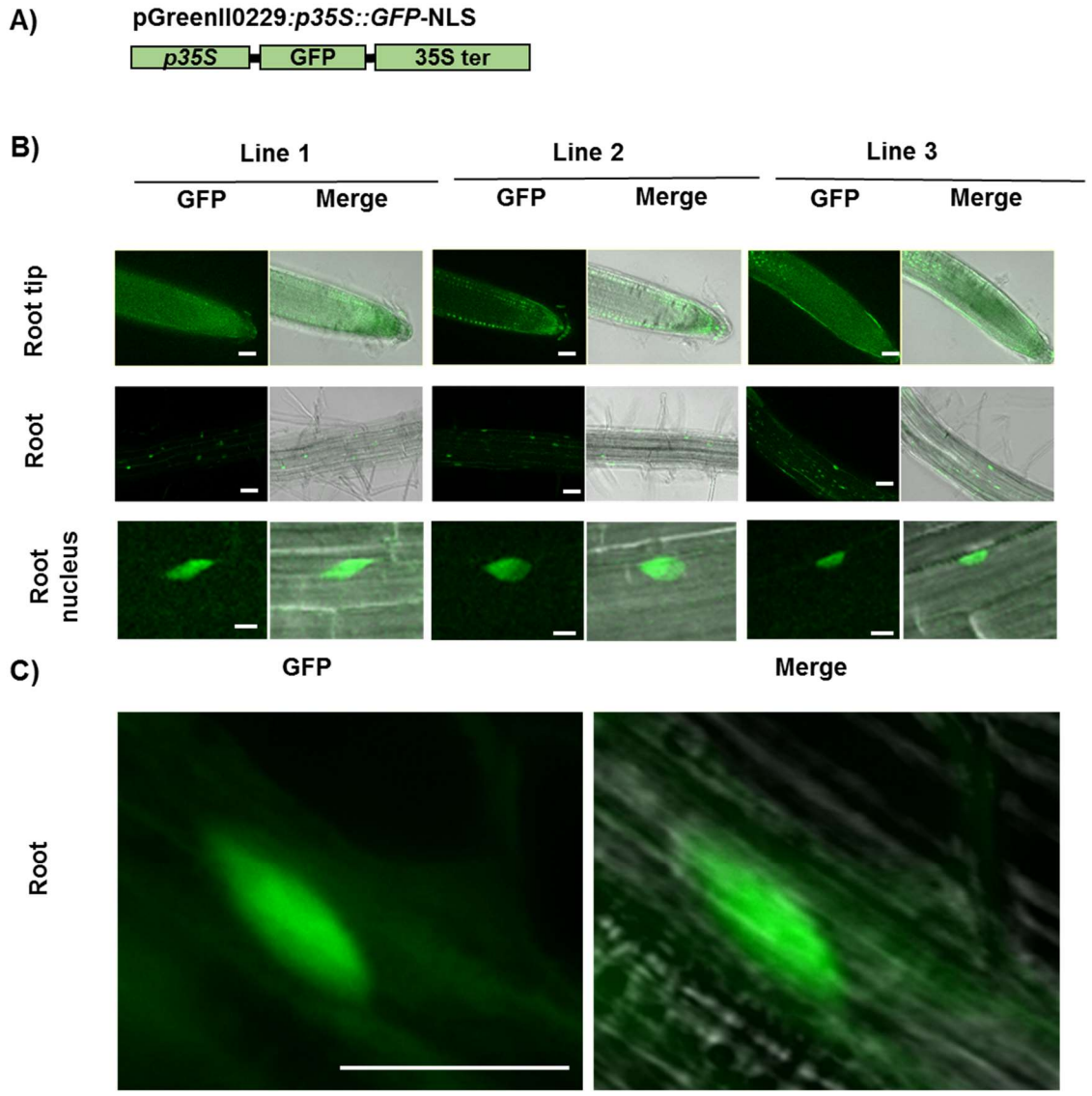


Figure 11. GFP-NLS exhibit no specific subnuclear in the root. A) Schematic representation of the *pGreenII0229:p35S::GFP-NLS* construct under control of the constitutively active 35S promoter. B) ApoTome optical sections made by structured illumination microscopy. The nuclear localization of the GFP protein was analysed in three independent lines transfected with the expression vector. In the upper and middle panel, optical sections of the root tip and root are shown. A 20X/0.8 objective was used. Scale bar indicates 50 µm. The lower panel shows a root nucleus. A 40X/1.4 objective was used. Scale bar indicates 10µm. C) Confocal laser scanning microscopy optical section of root nuclei of Line 1 with a 63X/1.4 objective. Scale bar indicates 10µm. The merge picture, shows an overlay between the GFP and Bright field channel.

to be absent from the dense heterochromatic regions, defined by speckles bound by DAPI. This indicates that UAP56 prefers euchromatic regions in the nucleus. To display unspecific sub nuclear localization in the root, a GFP-NLS construct was used (Figure 11A). The construct was introduced into the wildtype by *Agrobacterium tumefaciens* mediated transformation. Roots of three independent transfected lines were analysed by structured illumination microscopy on an ApoTome microscope

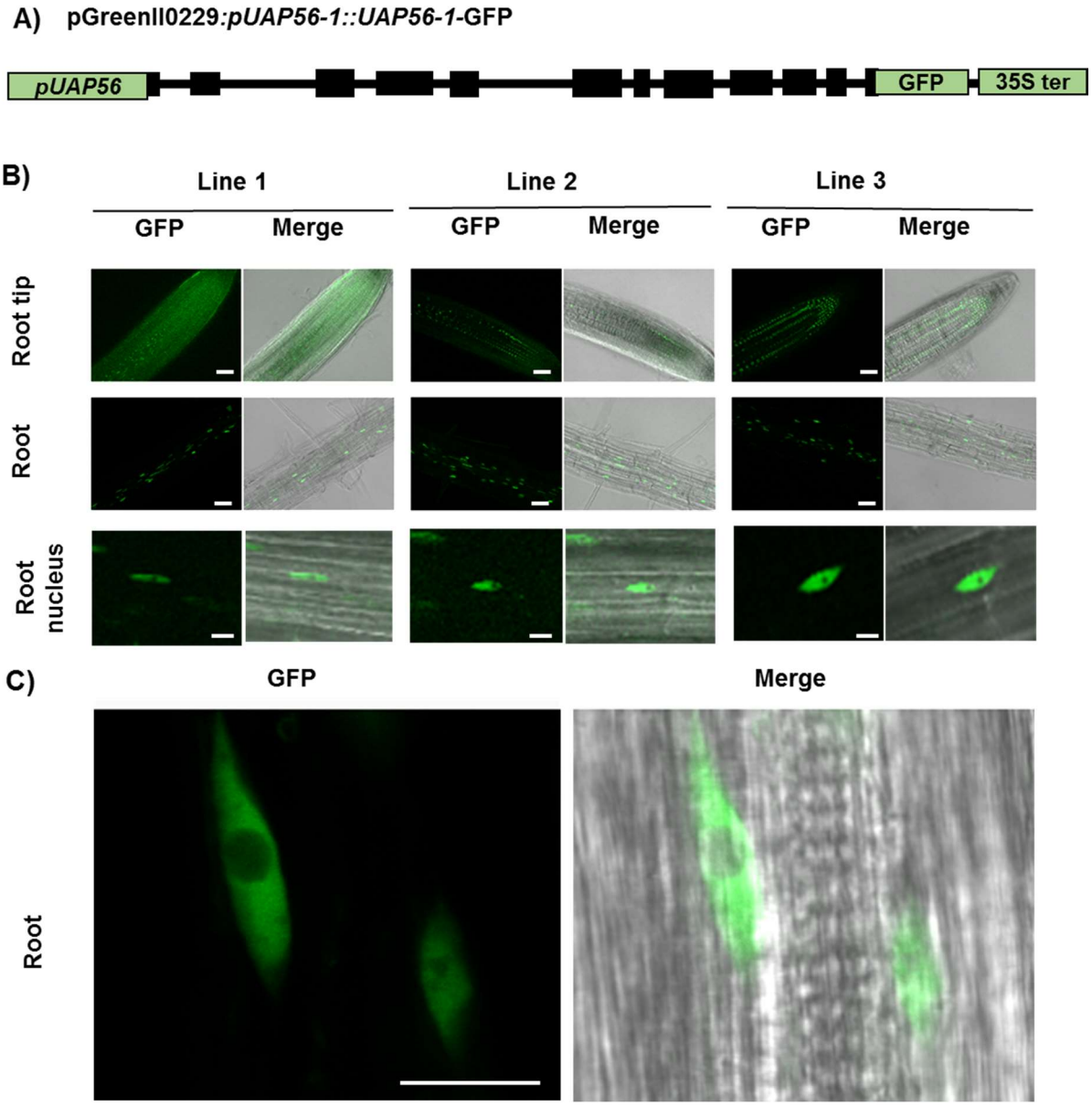


Figure 12. UAP56 is mainly localized in nucleoplasm A) Schematic representation of the *pGreenII0229:pUAP56-1::UAP56-1-GFP* construct under control of the native promoter. B) ApoTome optical sections made by structured illumination microscopy. The nuclear localization of the UAP56 protein was analysed in three independent lines transfected with the expression vector. In the upper and middle panel, optical sections of the root tip and root are shown. A 20X/.08 objective was used. Scale bar indicates 50 µm. The lower panel shows a root nucleus. A 40X/1.4 objective was used. Scale bar indicates 10µm. The optical sections are also shown as merges with the Bright field channel. C) Confocal laser scanning microscopy optical section of root nuclei of Line 1 with a 63X/1.4 objective. Scale bar indicates 10µm.

(Figure 11B). GFP-NLS exhibited similar localization pattern in the three independent lines. It was seen that the GFP-NLS protein is evenly distributed in different parts throughout the root. The distribution in the root tip appears unequal, but that is due to the positioning of the root tip and the thin optical section achieved by the structured illumination microscopy. At greater magnification, it was evident that the GFP-NLS protein showed no specific sub-nuclear localization. This was supported by Confocal Laser Scanning Microscopy (CLSM) done at higher magnification (Figure 11C) of the independent line 1. The GFP-NLS did not exclusively localize to the nucleus, but was also found in the cytoplasm, as detected by the fluorescence signal seen outside of the nucleus surrounding the vacuoles.

In order to see how UAP56 is distributed in the root, a construct of *UAP56* with native promoter fused to C terminal GFP was introduced into wildtype (Figure 12A). Roots of three independent transgenic lines were analysed at 6DAS. On the optical slides acquired by structured illumination microscopy it is seen that UAP56 fused to GFP were found in nuclei evenly distributed all over the root tip and root (Figure 12B). The expression pattern of UAP56 in the three independent lines looks comparable. At greater magnification using CLSM, it can be clearly seen that UAP56 exhibits a similar sub-nuclear localization pattern to what was shown in the immunolocalization analysis (Figure 10 and Figure 12C). UAP56 was equally distributed in the nucleoplasm, exhibiting minor localization in speckles. With the results of the GFP fusion localization study and the immunolocalization showing the same results it can be concluded that UAP56 is mainly found in the nucleoplasm *in vivo* in the roots of *Arabidopsis thaliana*. In order to compare the distribution of UAP56 fused to GFP in roots with the localization of the four ALY proteins, four constructs with GFP fusions were made. All constructs were under control of the native promoter (Figure 13A, Figure 14A, Figure 15A andFigure 16). ALY1 was found in nuclei all over the root tip and root in three independent transgenic lines (Figure 13B). ALY1 fused to GFP also exhibited a mainly nucleoplasmic localization. Unlike UAP56, ALY1 seemed to gather in more pronounced speckle-like structures as seen on the CLSM optical sections in Figure 13 C. The ALY2 protein fused to GFP exhibited a localization pattern with expression in nuclei evenly distributed throughout the root, as also seen for ALY1 and UAP56 (Figure 14B). At the sub-nuclear level ALY2 was shown to localize like ALY1 mainly in the nucleoplasm (Figure 14C).

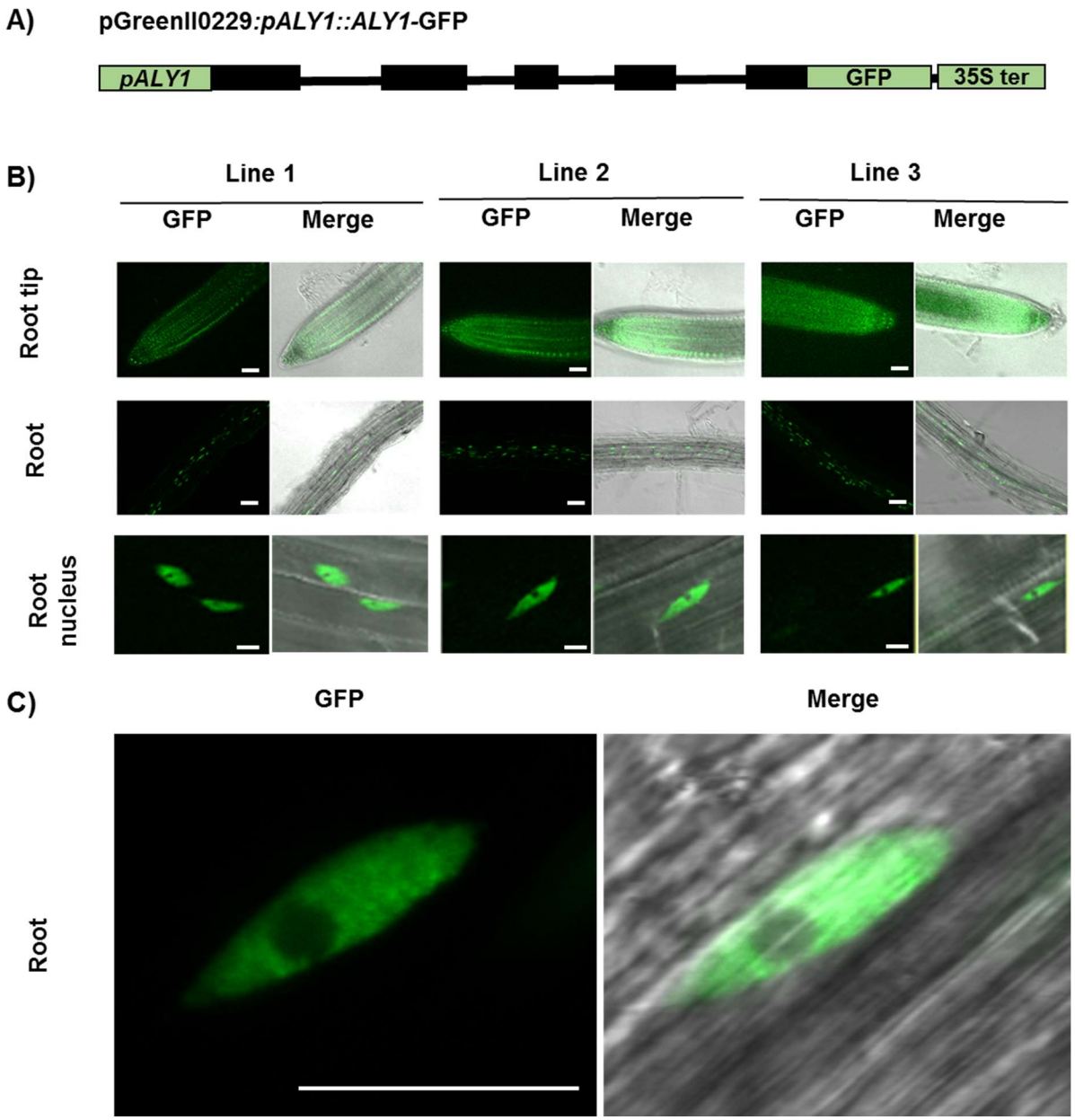


Figure 13. ALY1-GFP is mainly localized in the nucleoplasm A) Schematic representation of the *pGreenII0229:pALY1::ALY1-GFP* construct under control of the native promoter. B) ApoTome optical sections made by structured illumination microscopy. The nuclear localization of the ALY1 protein was analysed in three independent lines transfected with the expression vector. In the upper and middle panel, optical sections of the root tip and root are shown. A 20X/0.8 objective was used. Scale bar indicates 50 µm. The lower panel shows root nuclei. A 40X/1.4 objective was used. Scale bar indicates 10µm. The optical sections are also shown as merges with the Bright field channel. C) Confocal laser scanning microscopy optical section of root nuclei of Line 1 with a 63X/1.4 objective. Scale bar indicates 10µm.

ALY2 is also organized in distinct speckle-like structures, more alike to ALY1 than UAP56. ALY3 fused to GFP localized in the nucleus of cells distributed equally all over the root tip and root. (Figure 15B). The three independent lines exhibited a similar localization pattern of ALY3 fused to GFP. The general localization pattern of ALY3 in the root seems to be like the ones observed for the ALY1, ALY2 and UAP56

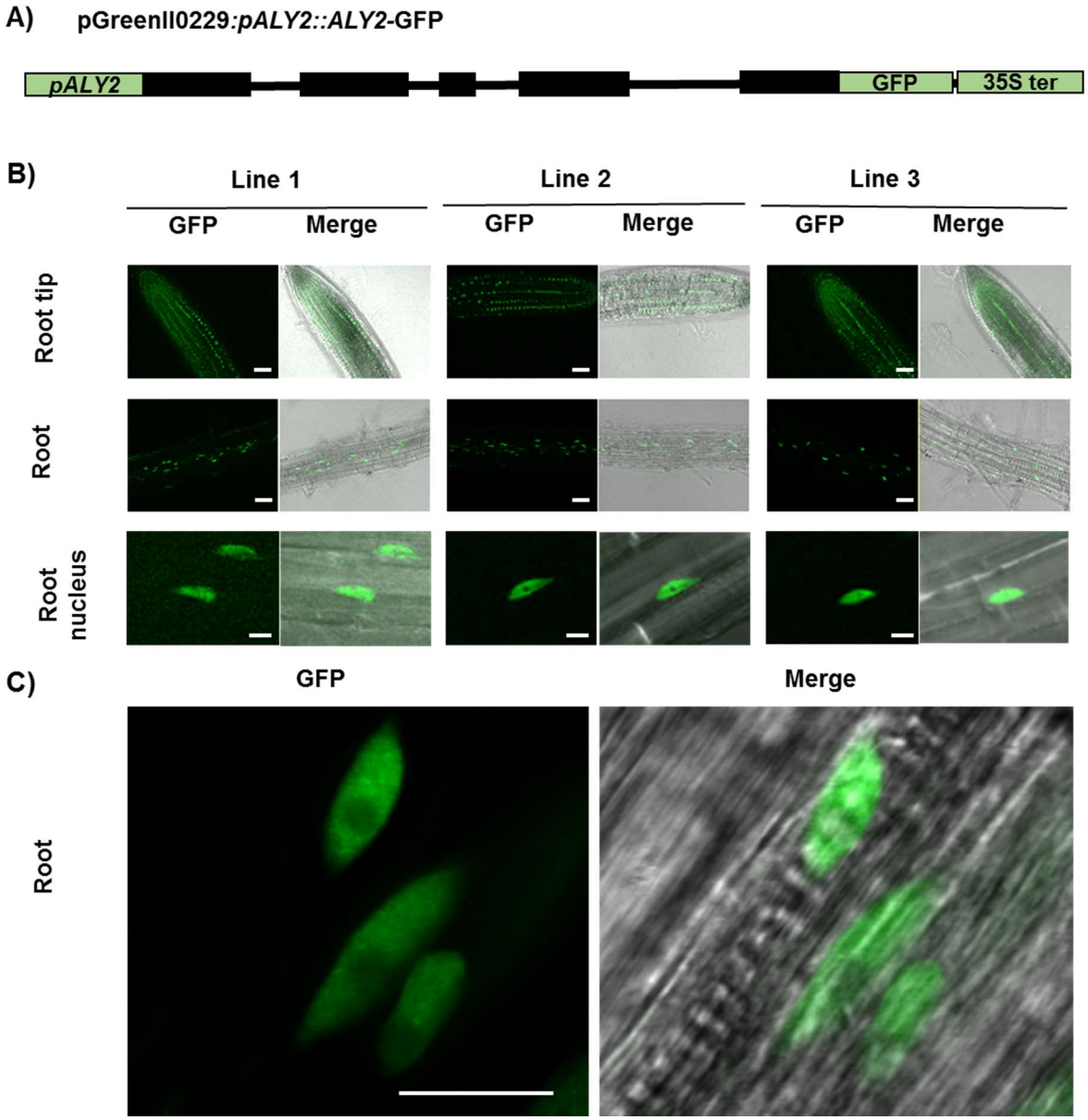


Figure 14. Aly2 exhibit same localization pattern as Aly1. A) Schematic representation of the *pGreenII0229:pALY2::ALY2-GFP* construct under control of the native promoter. B) ApoTome optical sections made by structured illumination microscopy. The nuclear localization of the ALY2 protein was analysed in three independent lines transfected with the expression vector. In the upper and middle panel, optical sections of the root tip and root are shown. A 20X/0.8 objective was used. Scale bar indicates 50 µm. The lower panel shows root nuclei. A 40X/1.4 objective was used. Scale bar indicates 10µm. The optical sections are also shown as merges with the Bright field channel. C) Confocal laser scanning microscopy optical section of root nuclei of Line 1 with a 63X/1.4 objective. Scale bar indicates 10µm.

GFP fusions. When examining the sub-nuclear localization it was seen that ALY3 localized to the nucleoplasm and showed a strong signal in the nucleolus (Figure 15C). This is different from what has been seen for the ALY1, ALY2 and UAP56 GFP fusions. ALY3 was found to localize subtly in speckle-like structures. ALY4 fused to GFP was observed in nuclei distributed equally all over the root tip and root in three

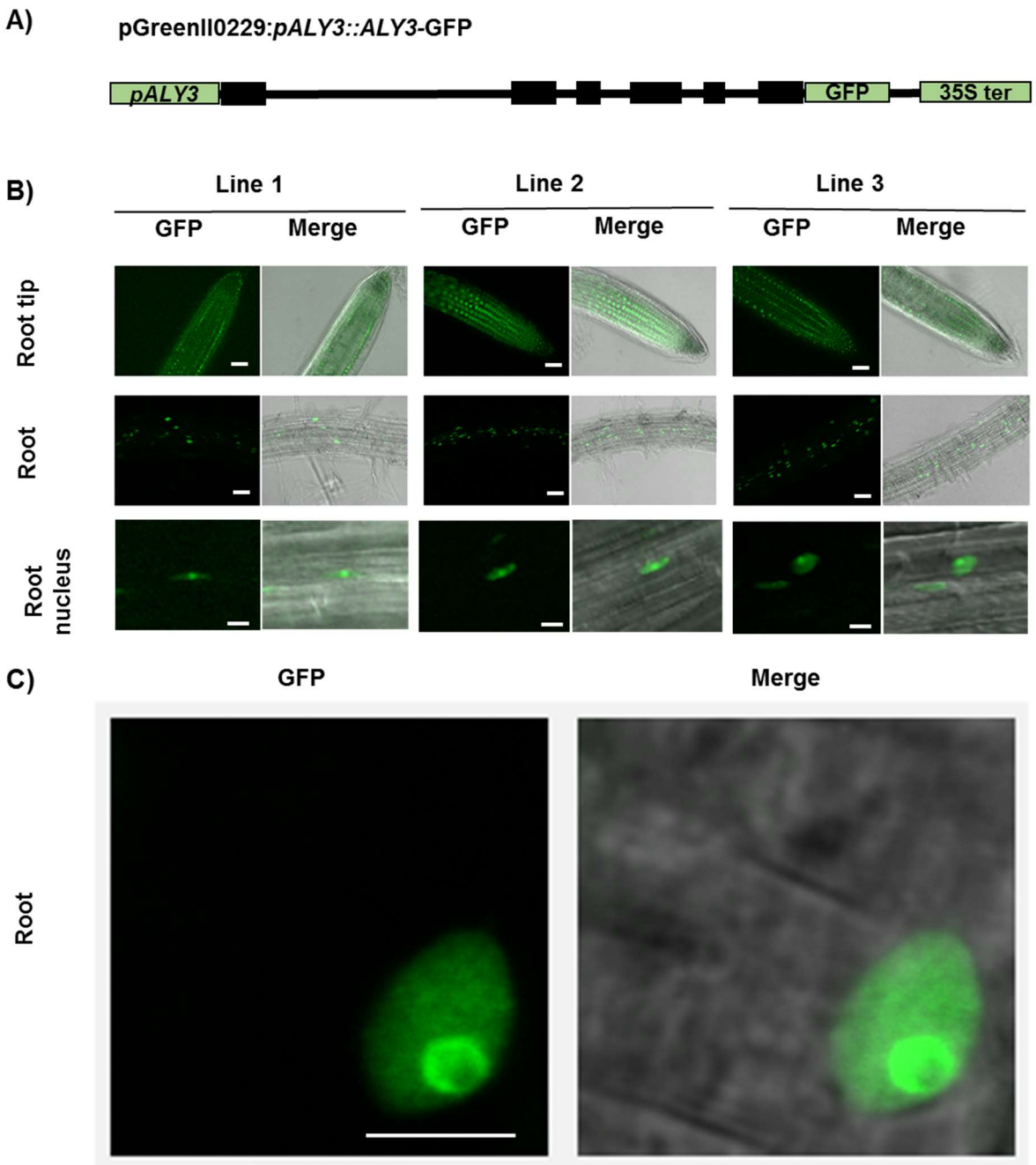


Figure 15. ALY3-GFP is found in both nucleoplasm and nucleolus A) Schematic representation of the *pGreenII0229:pALY3::ALY3-GFP* construct under control of the native promoter. B) ApoTome optical sections made by structured illumination microscopy. The localization of the ALY3 protein was analysed in three independent lines transfected with the expression vector. In the upper and middle panel, optical sections of the root tip and root are shown. A 20X/0.8 objective was used. Scale bar indicates 50 µm. The lower panel shows a root nucleus. A 40X/1.4 objective was used. Scale bar indicates 10µm. The optical sections are also shown as merges with the Bright field channel. C) Confocal laser scanning microscopy optical section of root nuclei of Line 1 with a 63X/1.4 objective. Scale bar indicates 5µm.

independent transfected lines (Figure 16B). At the sub-nuclear level ALY4 seems to localize like ALY3, being present in both nucleoplasm and nucleolus. ALY4 fused to GFP localized in speckle like structures. The speckle-like structures were more prominent in the lines with ALY4 than in the ones with ALY3

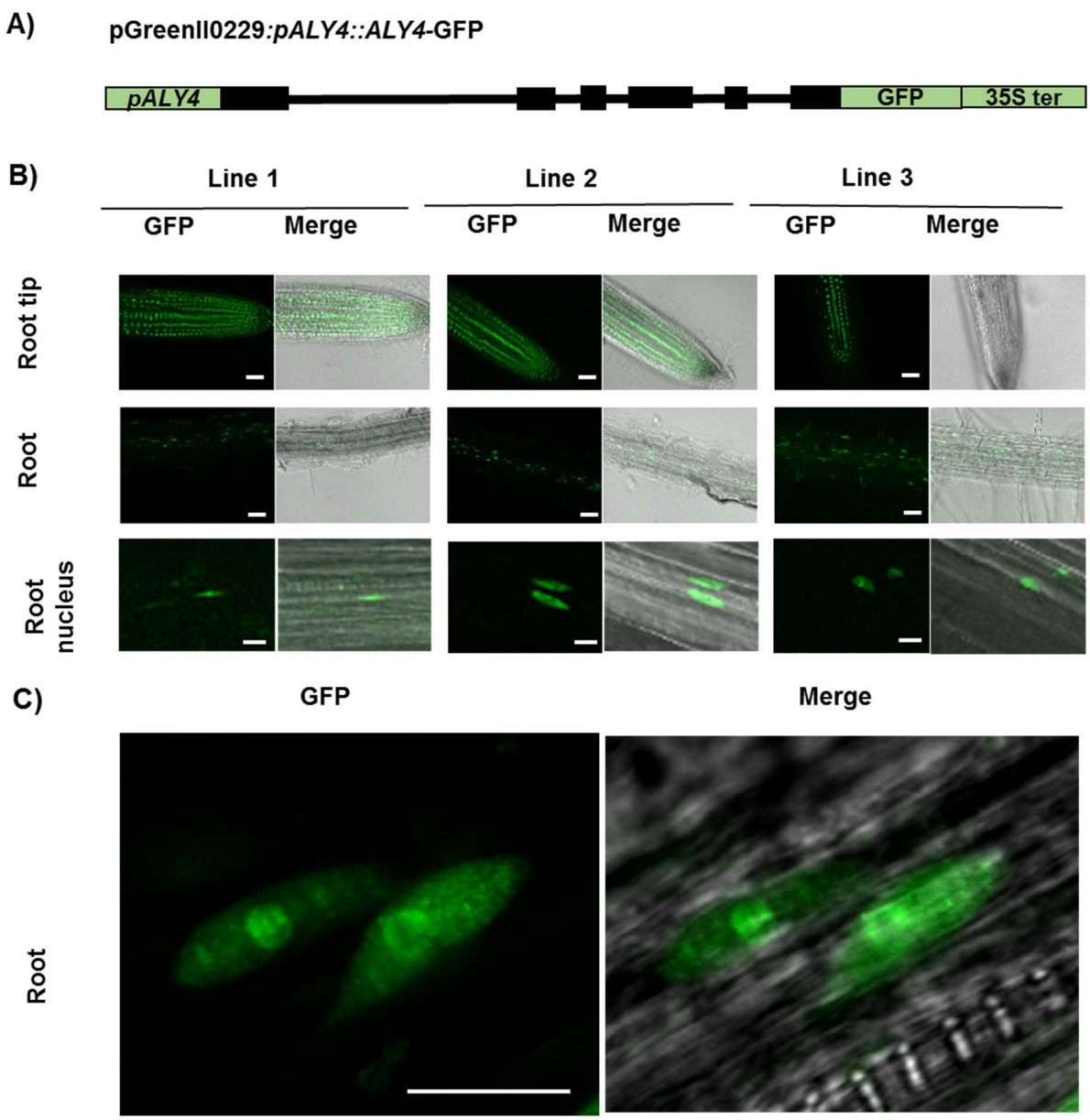


Figure 16. ALY4-GFP exhibits subnuclear localization comparable to that of ALY3 A) Schematic representation of the *pGreenII0229pALY4::ALY4-GFP* construct under control of the native promoter. B) ApoTome optical sections made by structured illumination microscopy. The nuclear localization of the ALY4 protein was analysed in three independent lines transfected with the expression vector. In the upper and middle panel, optical sections of the root tip and root are shown. A 20X/0.8 objective was used. Scale bar indicates 50 µm. The lower panel shows a root nucleus. A 40X/1.4 objective was used. Scale bar indicates 10µm. The optical sections are also shown as merges with the Bright field channel. C) Confocal laser scanning microscopy optical section of root nuclei of Line 1 with a 63x/1.4 objective. Scale bar indicates 10µm.

fused to GFP. It was seen that the GFP fusion proteins of UAP56, ALY1, ALY2, ALY3 and ALY4 were showing an even expression pattern in nuclei all over the root, just like the NLS-GFP control under the 35S promoter. The 35S promoter is constitutive and active all over the root (Holtorf et al., 1995), indicating the same is true for the five GFP fused proteins. At the sub-nuclear level UAP56, ALY1, ALY2, ALY3 and

ALY4 all localize differently than NLS-GFP. NLS-GFP was distributed equally and smoothly in the nucleus, with similar level of GFP signal in the nucleoplasm and nucleolus. UAP56 differed from the GFP-NLS by being excluded from the nucleolus. ALY1 and ALY2 differed in localization from the GFP-NLS, by strongly accumulating in speckle-like structures and mainly localizing in the nucleoplasm. ALY3 and ALY4 differed from GFP-NLS by a stronger accumulation of GFP fused protein in the nucleolus compared to the nucleoplasm. ALY3 and even more so ALY4 accumulated in speckle-like structures. This indicates that ALY1 and ALY2 show similar sub-nuclear localization. The same is seen for ALY3 and ALY4 fused to GFP, although ALY4 seemed to accumulate more in speckles than ALY3. The four ALY proteins could be divided into two groups based on their sub-nuclear localization. One group was not present in the nucleolus (ALY1 and ALY2), the other was (ALY3 and ALY4). UAP56 was mainly excluded from the nucleolus and exhibited a much weaker speckle-like localization pattern than the four ALYs. This suggest that UAP56 might localize differently than ALY1, ALY2, ALY3 and ALY4. The conclusion is that the three groups UAP56, ALY1/ALY2 and ALY3/ALY4 show divergent subnuclear localization patterns when fused to GFP.

3.1.3 Analysis of T-DNA lines with deficient transcript levels of *ALY1*, *ALY2*, *ALY3* and *ALY4*.

In order to learn more about the genetic relationships between the four ALY proteins it was decided to undertake a reverse genetics approach. Four T-DNA lines were obtained from the Nottingham Arabidopsis Stock Centre (NASC). A schematic overview of the four T-DNA lines are found in Figure 17 A,C,E and G. The exact genomic position of the T-DNA insertions was determined by PCR and sequencing. PCR was done with one primer in the genomic sequences and one in the T-DNA insertions, subsequently the PCR amplicon was sent for sequencing (Data not shown). The T-DNA insertion positions are indicated in the schematic representations of the genes (Figure 17A, C, E and G). All four T-DNA lines were then made homozygous for the T-DNA insertions. This was tested by PCR based genotyping. Primer 5 (P5) and primer 6 (P6) spans the site of the *aly1-1* T-DNA insertion. A PCR fragment of the right size was amplified for the extracted wildtype DNA, but not for the *aly1-1* DNA, due to the inserted T-DNA insertion (Figure 17B). P1 binds in the SAIL T-DNA inser-

tion. The PCR fragment of P1 and P5 is only amplified in the *aly1-1* DNA. Combined this indicates that the T-DNA is inserted into both alleles of *ALY1*, so that *aly1-1* is homozygous for the T-DNA insertion. The PCR based genotyping for *aly2-1*, *aly3-1* and *aly4-1* shows that homozygous lines were also obtained for these T-DNA lines (Figure 17D, F and H).

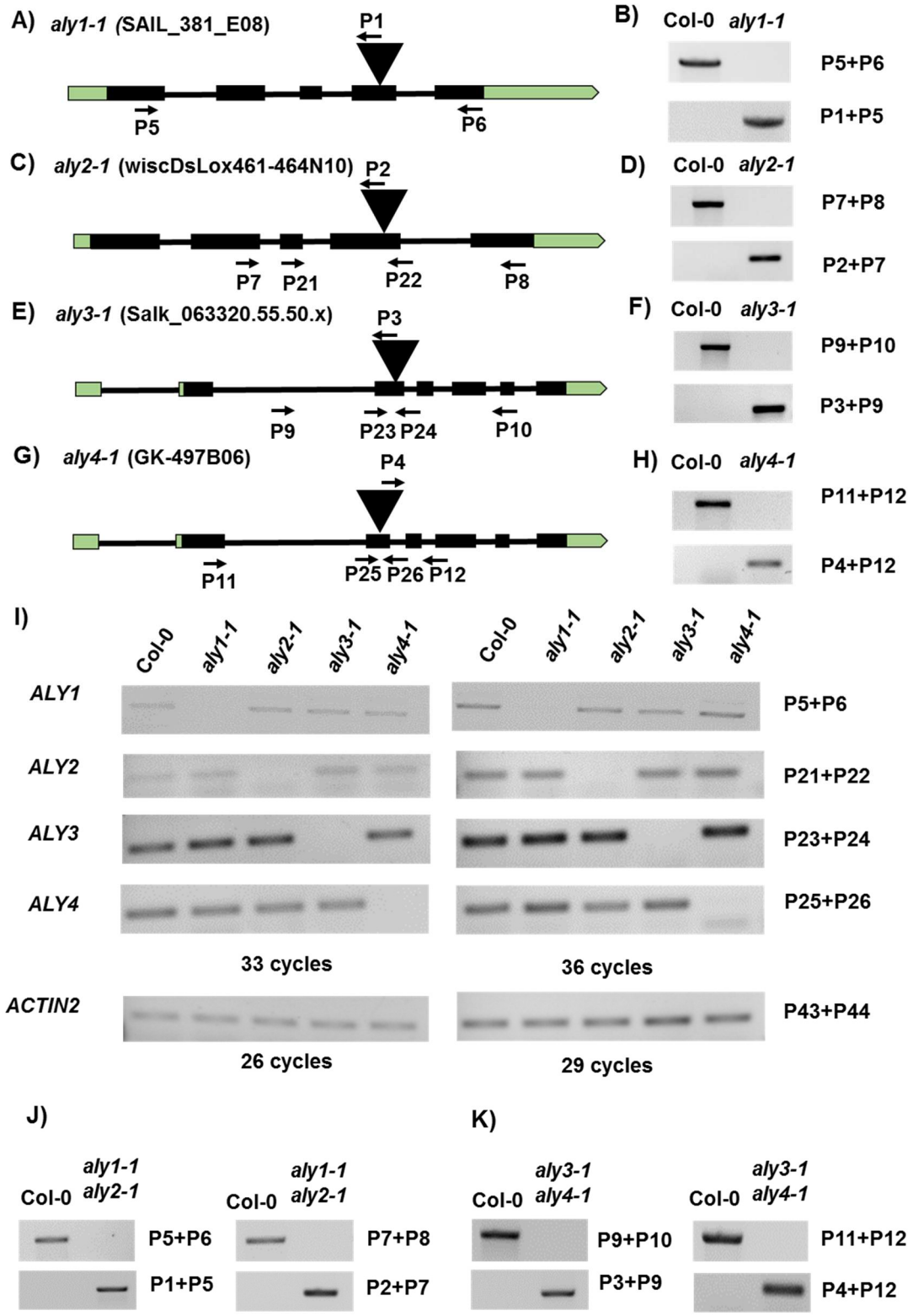


Figure 17. Molecular characterization of *aly1-1*, *aly2-1*, *aly3-1* and *aly4-1* T-DNA lines. A) Schematic representation of *ALY1* (At5g59950) with the T-DNA insertion and primers shown. B) Genotyping of wild type and *aly1-1* using indicated primers. C) Schematic representation of *ALY2* (At5g02530) with the T-DNA insertion and primers shown. D) Genotyping of wild type and *aly2-1* using indicated primers. E) Schematic representation of *ALY3* (At1g66260) with the T-DNA insertion and primers shown. F) Genotyping of wild type and *aly3-1* using indicated primers. G) Schematic representation of *ALY4* (At5g37720) with the T-DNA insertion and primers shown. H) Genotyping of wild type and *aly4-1* using indicated primers. I) Expression analysis. Semi quantitative RT-PCR was performed on RNA extracted from wild type and the four single mutants. For each primer pair the RT-PCRs were performed with two different numbers of amplification cycles to show that the PCR reaction was not saturated. Each primer combination with different amplification cycles, were documented on the same gel, with the same settings. A*CTIN2* was used as reference gene. DNA and RNA was extracted from 10 DAS. Seedlings grown on MS plates. J) Genotyping of wild type and *aly1-1 aly2-1* double mutant using indicated primers. K) Genotyping of wild type and *aly3-1 aly4-1* double mutant using indicated primers.

In order to determine if the T-DNA insertions in the four ALY genomic sequences reduce the transcription levels of the respective mRNA transcripts, semi-quantitative RT-PCR was done. RNA was extracted at 10DAS and converted to cDNA. For the *ALY1* transcript in wildtype and the four T-DNA lines, P5 and P6 was used to amplify the transcript (Figure 17I). It is seen that when comparing the relative expression between the *ALY1* transcripts at 33 cycles and 36 cycles an increase in band intensity is seen for the higher cycle number. This indicates that the PCR was not saturated. Comparing the relative levels of the *ALY1* transcript, between the different genotypes, shows that *ALY1* is altered in *aly1-1*, but not in *aly2-1*, *aly3-1* and *aly4-1* mutants. This insinuates that the *ALY1* transcript is not increased when reducing the expression levels of *ALY2*, *ALY3* and *ALY4*. An increase in transcript levels would be indicative of that more *ALY1* is being transcribed to compensate for the loss of either the *ALY2*, *ALY3* and *ALY4* transcripts. This is not the case according to Figure 17I. In Figure 17I it is also seen that the *ALY2* transcript is downregulated in *aly2-1* and not up or down regulated in any of the other genotypes. Likewise the *ALY3* transcript is downregulated in *aly3-1* and not up or down regulated in any of the other genotypes. *ALY4* is only downregulated in *aly4-1* and not differentially regulated in the other genotypes. Overall, this means that a homozygous T-DNA line interrupting the genomic sequence of each of the four ALYs has been obtained. Each ALY T-DNA line showed no detectable transcript of the respective ALY transcript. None of the *ALY1*, *ALY2*, *ALY3* and *ALY4* transcripts are being upregulated at the RNA level to compensate for the loss of one of the other ALY transcripts.

The localization studies of GFP fused proteins in the root showed that the ALY proteins could be divided in two groups based on sub-nuclear localization (3.1.2). Aly3

and ALY4 both localized in the nucleolus and nucleoplasm, whereas ALY1 and ALY2 mainly localized in the nucleoplasm. Interestingly as earlier shown, ALY3 and ALY4 also had the highest protein sequence similarity of any ALY protein combination (Table 13). The second highest protein similarity was between ALY1 and ALY2. It was decided to cross *aly1-1* with *aly2-1* and *aly3-1* with *aly4-1*, to look for functional redundancy among the proteins that showed similar sub-nuclear localization. Pollen from *aly1-1* was mechanically transferred to the stamens of *aly2-1* emasculated flowers. The T2 generations of the crossing were sown out on soil and genotyped to obtain offspring homozygous of both the *aly1-1* and *aly2-1* T-DNA insertions. The PCR based genotyping lead to the identification of a double homozygous plant, *aly1-1 aly2-1* (Figure 17J). The same were done for *aly3-1* and *aly4*-1, which led to identification of an *aly3-1 aly4-1* double homozygous plant (Figure 17K).

3.1.4 Loss of *ALY1* and *ALY2* transcript in single and double mutants does not lead to developmental phenotypes.

In order to look for functional redundancy in *aly1-1* and *aly2-1*, the double mutant was phenotypically analysed together with the two single mutants and the wild type. The plants were sown out on soil, and grown under long day (LD) light conditions. No difference in plant morphology or growth was observed for *aly1-1*, *aly2-1* and *aly1-1 aly2-1* compared to the wildtype at 15, 28 and 60DAS (Figure 18A). The transition from the vegetative to the reproductive phase in *Arabidopsis* is a crucial step for plant development and can be measured as the number of total leaves present at the emergence of the flower bud , also called bolting (Möller-Steinbach et al., 2010). Leaf number at bolting was counted and quantified for wildtype, *aly1-1*, *aly2-1* and *aly1-1 aly2-1* (Figure 18B). The data was statistically analysed by applying a two way Analysis of Variance (ANOVA). All statistical ANOVA analyses in this thesis are found in Chapter 7.2, as well as the general workflow of the statistical analysis in Chapter 7.1. Three null hypothesis are tested with the two-way ANOVA analysis. 1) The sample means of the first factor are equal 2) The sample means of the second factor are equal 3) there is no interaction between the two factors. The level of significance in this thesis was set at $p < 0.05$. The mean values of the data used for the two-way ANOVA with the corresponding standard deviations used for the bar diagrams are also found in Chapter 7.2

.

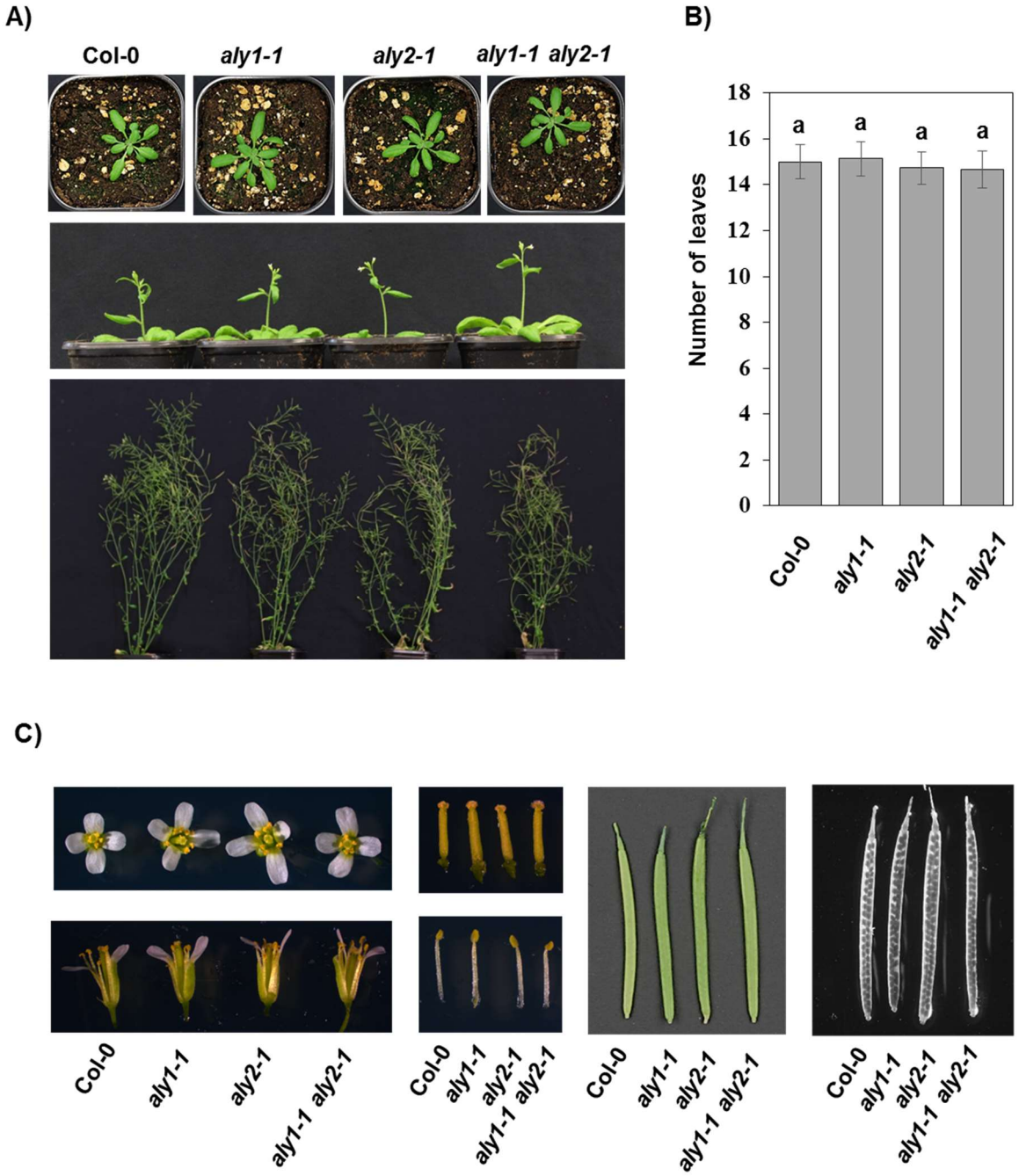


Figure 18. Single and double mutants of ALY1 and ALY2 shows no obvious phenotypes. A) Flowering time. The upper panel shows all the genotypes at 15DAS. The middle panel shows all the genotypes at 28DAS, and the lower panel at 60DAS. B) Quantification of the time of bolting. The average numbers of leaves at bolting for the different genotypes are depicted on the bar diagram (n=15). The error bars depicts the standard deviation. The numbers above the bars indicate if the different genotypes are significantly different from each other. If two genotypes displays the same letter, they are not significantly different. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Appendix Table 20. C) Flower morphology at 44DAS. The two left panels shows the flower seen from above, and from the side with one petal and two sepals removed. In the two middle-left panels, isolated pistils and stamens are shown. In the middle-right panel, fully elongated siliques are shown. The panel to the right shows siliques that has been cleared to make the seeds visible.

For the number of leaves at bolting in the wildtype, *aly1-1*, *aly2-1* and *aly1-1 aly2-1* genotypes all three null hypotheses are confirmed (Figure 18B, Chapter 7.2 (Table 20)). This means neither *ALY1* nor *ALY2* affects the number of leaves at bolting, and the two factors do not show interaction. Flower organs were also phenotypically analysed (Figure 18C). A view from the top of the flowers shows that the general flower morphology was not different in wildtype, *aly1-1*, *aly2-1* and *aly1-1 aly2-1* plants. When dissecting the flowers no obvious developmental defects were observed. Pistil and stamens look alike across all the genotypes. The fertilized pistils later develop into siliques. The overall morphologies of the siliques were also not different among the genotypes. When the siliques were bleached, to make them see-through, it was also seen that the general seed development was unaffected. Single and double mutants of *ALY1* and *ALY2* shows no noticeable effect on growth and development of *Arabidopsis.*

3.1.5 Loss of *ALY1* and *ALY2* transcripts does not lead to mRNA accumulation.

Reducing expression levels of *ALY1* and/or *ALY2* did not led to phenotypes in the *aly1-1*, *aly2-1* and *aly1-1 aly2-1* mutants. In order to investigate if the mRNA export was affected in the four genotypes, whole mount *in situ* hybridization was done on A*rabidopsis* seedlings grown on plates with MS-medium at 6DAS. The experimental setup was done like in MacGregor et al., 2013, using the protocol from Gong et al., 2005 on the root tissue instead of leaf material. The probe binds poly(A) mRNA. Z-stacks of optical slides in the plane running along the xylem vessel elements were made, showing mRNA distribution in nucleus and cytoplasm (Figure 19A). The cytoplasmic and nucleoplasmic levels of the poly(dT) probe, were measured like shown in Figure 19B. The median intensity values of the cytoplasmic and nucleoplasmic areas were quantified. The ratio between the nucleoplasmic and cytoplasmic (N/C ratio) signal for the four genotypes are seen in Figure 19C. The statistical analysis can be found in Chapter 7.2 (Table 21). The two null hypotheses that there were no interaction and that the means of the ALY2 factor were equal, were rejected with p-values at 0.02 and 0.04 respectively. The null hypothesis concerning ALY1 is confirmed. This means there is a significant interaction between ALY1 and ALY2. As significant differences were observed in the Two-way ANOVA, the data were analysed using the *post hoc* Tukey's test. With the *post hoc* test it was seen that the only

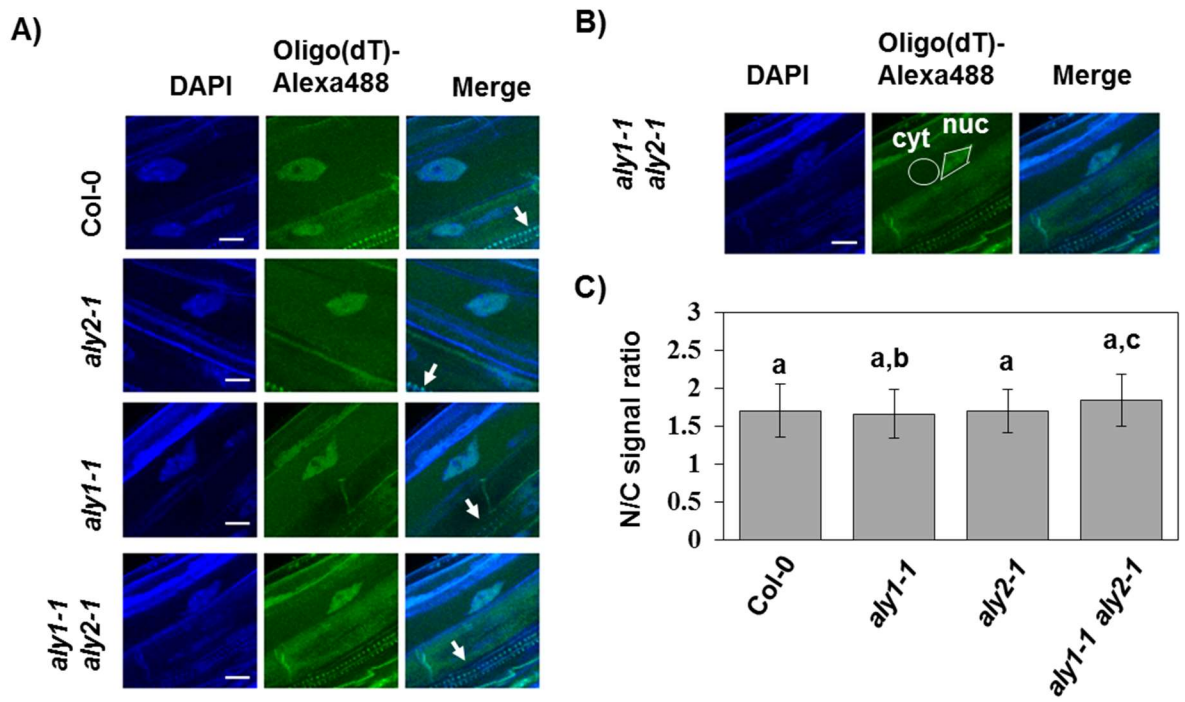


Figure 19. Single and double mutants of *ALY1* and *ALY2* are not deficient in mRNA export. Plate grown seedlings were harvested at 6DAS. A) Whole mount *in situ* hybridization of a 48-mer Oligo (dT) probe fluorescently labelled with Alexa488 to fixated root tissue of the four genotypes. The optical sections were made around the xylem plane using CSLM. A 40X/1.3 objective was used. The xylem vessel elements is indicated with arrows. Twelve z-stacks of 0.8µm were transformed with a maximum projection. DAPI was used as a counterstain. Scale bar indicates 10µm B) Selection for the quantification of Alexa488 signal in nucleus and cytoplasm in root tissue of *aly1-1 aly2-1*. C) Quantification of bulk mRNA export. The ratio of fluorescence signal in nucleus and cytoplasm (N/C ratio) was calculated for 60 nuclei of each genotype. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 21).

significant difference is between the N/C ratios of *aly1-1* and *aly1-1 aly2-1* (Figure 19C). As none of the single and double mutant N/C ratios were significantly different from the wildtype mean it was concluded that reducing *ALY1* and *ALY2* transcript levels does not affect the bulk mRNA export compared to wildtype.

3.1.6 ALY3 and ALY4 are involved in the transition into reproductive phase .

The double mutant of *aly1-1* and *aly2-1* did not seem to affect mRNA export or plant development in any noticeable way. To test if the two other ALYs, ALY3 and ALY4, were redundant and had an effect on plant development the single and double mutants were analysed for abnormal phenotypes. Wildtype, *aly3-1*, *aly4-1* and *aly3-1 aly4-1* were sown out on soil and observed under LD conditions. At 15 and 60DAS, there were no observable difference between the four genotypes (Figure 20A). At 24DAS, it was seen that *aly3-1* and *aly3-1 aly4-1* had switched from vegetative

growth to reproductive growth, whereas wildtype and *aly4-1* plants had not. To quantify the switch to flowering, the number of leaves at bolting were counted and quantified for the single and double mutants (Figure 20B). The statistical analysis for the quantification of leaf number can be found in Chapter 7.2 (Table 22). There was

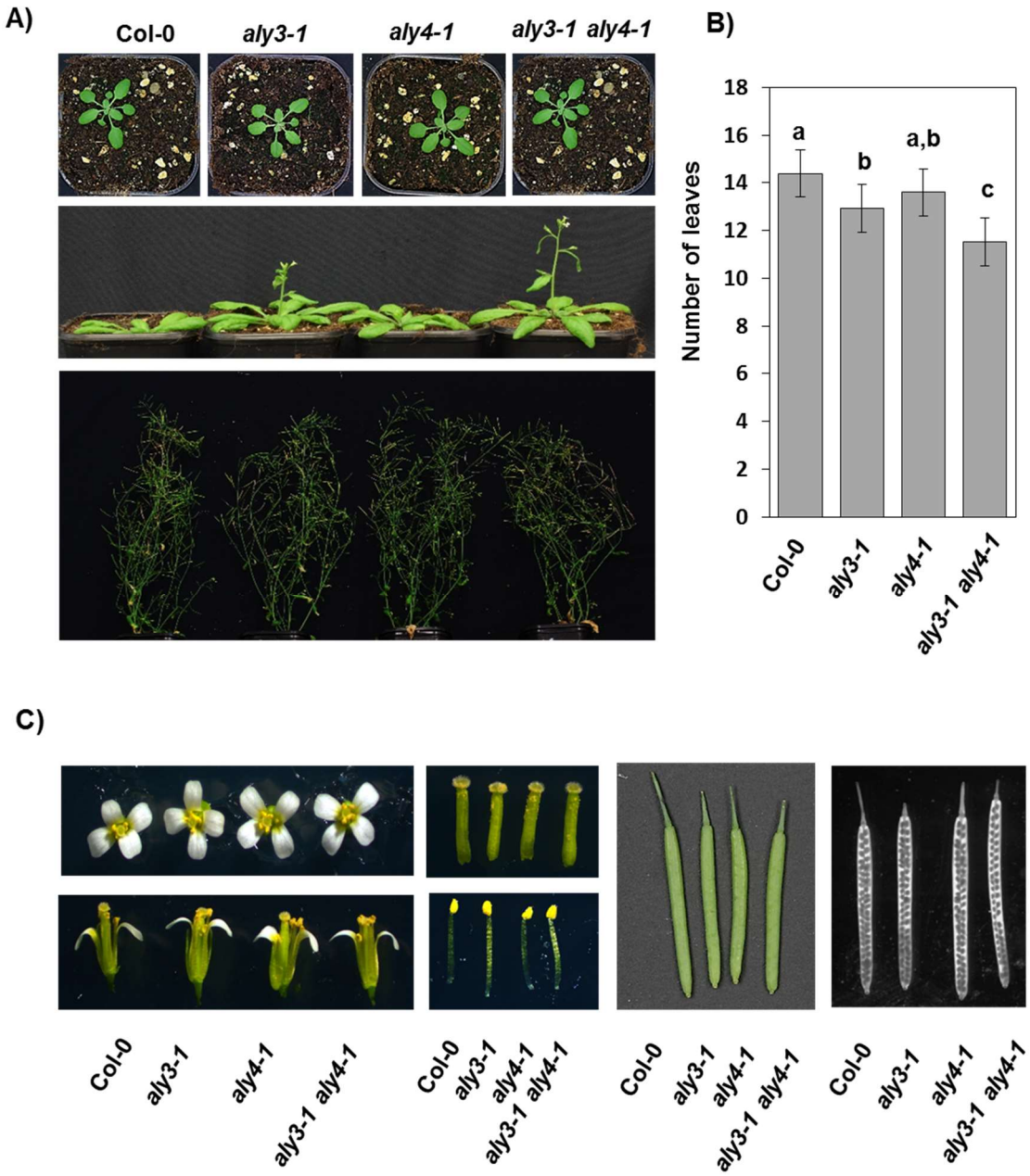


Figure 20. The two mutants *aly3-1* and *aly3-1 aly4-1* exhibit minor phenotypes. A) Flowering time. The upper panel shows all the genotypes at 15DAS. The middle panel shows all the genotypes at 24DAS, and the lower panel at 60DAS. B) Quantification of the time of bolting. The average numbers of leaves at bolting for the different genotypes are depicted on the bar diagram (n=15). The error bars are showing the standard deviations. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 22). C) Flower morphology at 44DAS. The two left panels shows the flower seen from above, and from the side with one petal and two sepals removed. In the two middle-left panels, isolated pistils and stamens are shown. In the middle-right panel, fully elongated siliques are shown. The panel to the right shows siliques that has been cleared to make the seeds visible.

no significant interaction between ALY3 and ALY4 concerning the number of leaves at bolting ($p = 0.17$). The two null hypotheses that there were no differences between the sample means for the two factors ALY3 ($p = 5.11 \cdot 10^{-11}$) and ALY4 ($p = 5.02 \cdot 10^{-6}$) were rejected. This means that ALY3 and ALY4 affects the number of leaves at bolting in an additive and independent way. The Tukey's test was done to find significant differences between the sample means. The results of the Tukey's test has been visualized with the letter codes in Figure 20B. The number of leaves at bolting of the T-DNA lines *aly3-1* and *aly3-1 aly4-1* were significantly different from the wild type. The mean number of leaves were not significantly different between *aly3-1* and *aly4-1*, whereas the double mutant *aly3-1 aly4-1* mean is different from all the other means. This indicates that the reduction of the *ALY3* transcript on its own affects the flowering switching. Reducing the levels of *ALY4* transcript leads to a mild non-significant reduction in number of leaves at bolting. When both the *ALY3* and *ALY4* transcript levels were reduced the leaf number at bolting were reduced in an additive manner. The flower organs were also inspected for noticeable phenotypes (Figure 20C). When inspecting the flowers, seen from above and dissected, no obvious differences between the four genotypes were detected. The isolated stamens, pistils and siliques also looked the same. No differences in the seed sets were observed when inspecting the bleached siliques. The two T-DNA lines *aly3-1* and *aly4-1* did not affect general growth or the proper development of reproductive organs, but earlier flowering is observed in *aly3-1* and *aly3-1 aly4-1*. This indicates that ALY3 and ALY4 are somehow involved in flowering time regulation.

3.1.7 The aly3-1 aly4-1 double mutant is accumulating mRNA in the nucleus.

To observe if the ALY3 and ALY4 were involved in the mRNA export, seedlings from the Wildtype, *aly3-1*, *aly4-1* and *aly3-1 aly4-1* were analysed by whole mount in situ hybridization. The hybridization experiment was done and analysed like described in Chapter 3.1.5. Projected optical sections, showing the distribution of polyadenylated mRNA in *Arabidopsis* roots, are seen in Figure 21A. The quantification of the N/C ratio of mRNA amounts in the four genotypes was done on 60 nuclei (Figure 21B).These data are based on the results from one experiment. The statistical analysis is shown in Chapter 7.2 (Table 23). The null hypothesis that there is no interaction is rejected ($p = 2.5 \cdot 10^{-3}$). With the interaction term significant, the Tukey's test

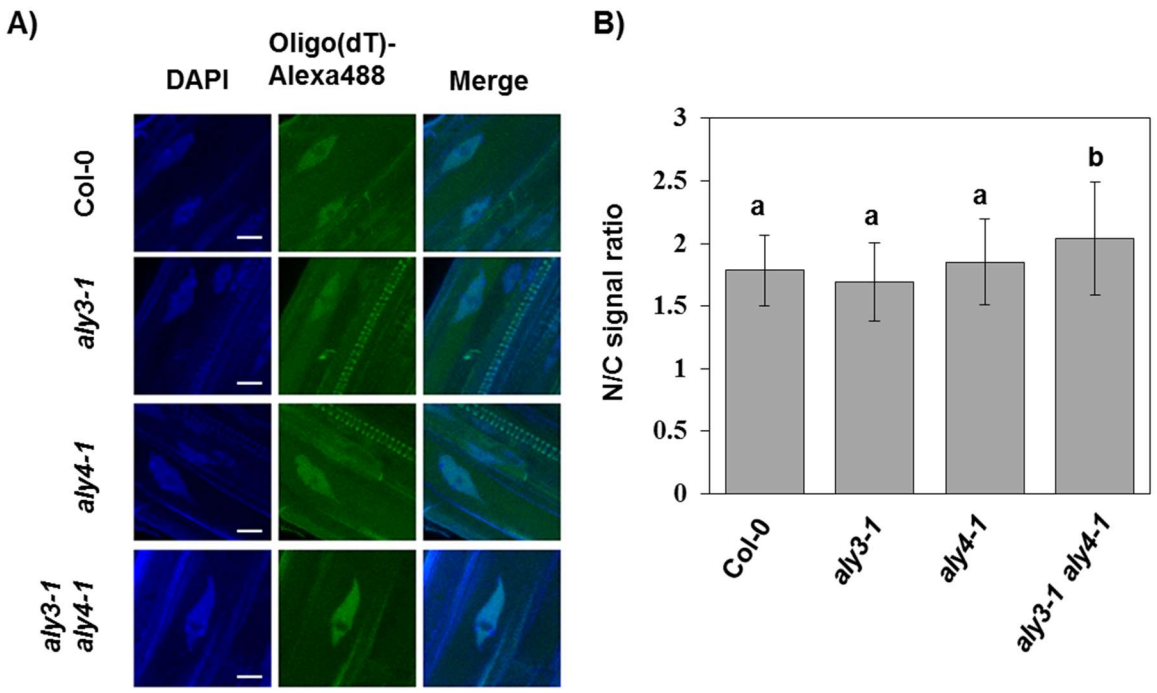


Figure 21. The double mutant aly3-1 aly4-1 exhibit a potential mRNA export block. Plate grown seedlings at 6DAS were used. A) Whole mount *in situ* hybridization of 48-mer Oligo (dT) probe fluorescently labelled with Alexa488 to fixated root tissue of the four genotypes. The optical sections were made in the plane of the xylem vessel elements on a CSLM microscope using a 40X/1.3 objective. Twelve z-stacks of 0.8µm were transformed utilising a maximum projection algorithm. DAPI was used as a counterstain. Scale bar indicates 10µm. B) Quantification of bulk mRNA export. The ratio of fluorescence signal in cytoplasm and nucleus was calculated for 60 nuclei of each genotype. The data was based on one experiment. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 23).

was applied to the data. The sample means of the two single mutants *aly3-1* and *aly4-1* were not significant different from the wild type mean. The sample mean of the double mutant, *aly3-1 aly4-1*, was significantly different from the sample means of all the other genotypes. This indicates that reducing the transcript levels of both *ALY3* and *ALY4* simultaneously leads to a synergistic mRNA export block, increasing the accumulation of polyadenylated mRNA in the nucleus. It is possible that ALY3 and ALY4 share some redundant functions in regards to mRNA export.

3.2 Characterization of TEX1

Not that much is known about the THO complex in plants compared to the human and yeast THO complex. In order to gain more knowledge about its function, one of the THO subunits TEX1 was selected for in detail study.

3.2.1 Molecular characterization of the t*ex1-4* T-DNA line.

The T-DNA line SALK_100012 (*tex1-4*) inserted in *TEX1* has earlier been described to affect tasiRNA biogenesis in a negative manner (Jauvion et al., 2010; Yelina et al., 2010). The T-DNA is inserted in the second intron (Figure 22A). This was confirmed by PCR amplification of DNA extracted from *tex1-4* using p13 and P3. The amplicon was sequenced to identify the exact location of the insertion (Data not shown).
In order to find a *tex1-4* line homozygous for the T-DNA insertion, seedlings from the T2 generation were genotyped by PCR (Figure 22B). It is seen that only *tex1-4* DNA has the T-DNA insertion, and that it is inserted in both *TEX1* alleles. The relative expression levels of TEX1 in tex1-4 compared to wild type were estimated by Quantitative RT-qPCR, as described later on Figure 37B-D). This shows that a line homozygous for the T-DNA inserted in *TEX1* was found, and it reduces the levels of the *TEX1* transcript to near background levels.

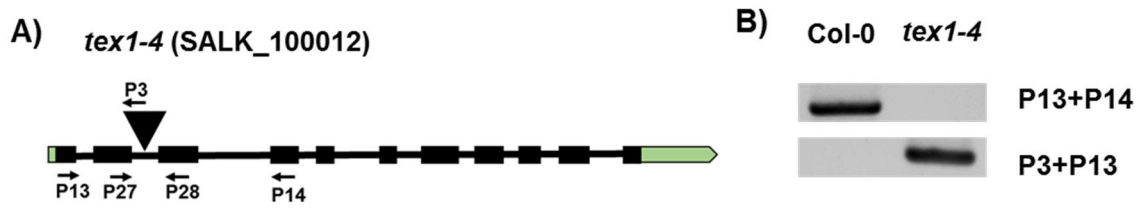


Figure 22. Molecular characterization of *tex1-4*. A) Schematic representation of *TEX1* (At5g56130) with the T-DNA insertion and primers shown. B) Genotyping of wild type and *aly1-1* using indicated primers.

3.2.2 Loss of the *TEX1* transcript leads to morphological and pleiotropic phenotypes.

In the other studies, where TEX1 has been investigated in *Arabidopsis thaliana*, the focus has entirely been on the effect on tasiRNA biogenesis (Jauvion et al., 2010; Yelina et al., 2010). Nothing has been published on if *tex1-4* exhibits any morphological phenotypes. In order to investigate this, seedlings were grown on soil under LD conditions alongside wild type. When visually inspecting the plants at 20DAS it

looked like there was a difference in leaf morphology (Figure 23A). This was confirmed by inspecting excised and flattened leaves of the two genotypes at 35DAS. The leaves of the T-DNA insertion line *tex1-4* looked shorter and wider compared to wild type, giving them a more round appearance. At 24DAS *tex1-4* showed an earlier transition into the reproductive phase than wildtype. At 60DAS *tex1-4* and the wildtype overall looked the same, albeit *tex1-4* appeared more bushy. In order to quantify if there was an earlier flowering phenotype, the number of leaves at bolting was counted (Figure 23B).There were 22% less leaves in *tex1-4* at bolting than what were observed in wildtype. This difference in the means is significant ($p= 6.3 \cdot 10^{-11}$).
The number of leaves at bolting was also counted under non-inductive short day conditions (SD) having 8 hours of light and 16 hours of darkness. Under these conditions wildtype plants stayed in the vegetative phase longer before switching to flowering. Under SD conditions, a visual inspection of *tex1-4* compared to wildtype at flowering showed a clear early bolting phenotype (Figure 23C and D). When quantifying the leaf number at bolting *tex1-4* had 32% less leaves at bolting than wildtype (Figure 23E). The sample mean of *tex1-4* was significantly different from the wild type mean ($p=1.04 \cdot 10^{-6}$) when analysed with a two-tailed students t-test. This means that *tex1-4* also exhibited earlier flowering under non-inductive SD conditions. The same pattern of earlier flowering under both long day and short day conditions has also been observed for mutants of the transcription elongation complex, FACT (Lolas et al., 2010). Down regulation of the floral repressor, Flowering Locus C (FLC), correlated with the earlier flowering in that study. To investigate if downregulation of the *FLC* transcript was observed in *tex1-4*, RT-qPCR was done. *FLC* was significantly downregulated in *tex1-4* to almost background levels (Figure 38C). This indicates that *TEX1* is required for establishing proper *FLC* transcript levels. As mentioned before the *tex1-4* mutant looked bushier than the wild type when growing. At 40DAS, it was seen that *tex1-4* is further developed than wild type, with more activated axillary meristems developed into rosetta branches (Figure 23F). The more branches could be due to *tex1-4* is just further developed due to the earlier flowering or be due to a misregulation in the number of branches. Therefore, the number of rosetta branches were quantified for fully developed plants. The primary inflorescences were ripped of the plants at 60DAS and counted Figure 23G). When the *TEX1* transcript

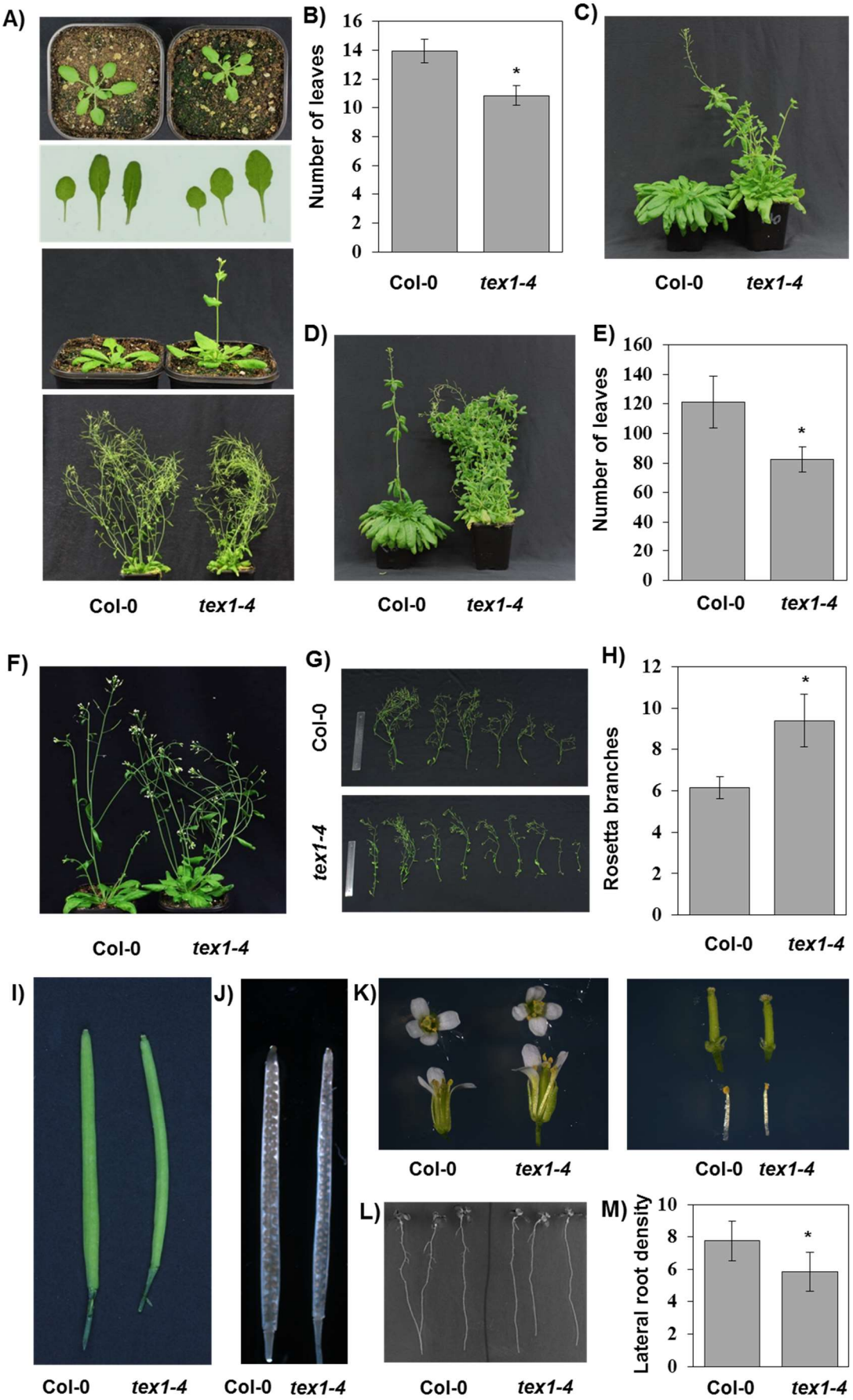

A)
B)
Number of leaves
16
14
12
10
8
6
4
2
0
*
Col-0
tex1-4
C)
Col-0
tex1-4
D)
Col-0
tex1-4
E)
Number of leaves
160
140
120
100
80
60
40
20
0
*
Col-0
tex1-4
Col-0
tex1-4
F)
Col-0
tex1-4
G)
Col-0
tex1-4
H)
Rosetta branches
12
10
8
6
4
2
0
*
Col-0
tex1-4
I)
Col-0
tex1-4
J)
Col-0
tex1-4
K)
Col-0
tex1-4
Col-0
tex1-4
L)
Col-0
tex1-4
M)
Lateral root density
10
8
6
4
2
0
*
Col-0
tex1-4

Figure 23. The *tex1-4* mutant exhibits pleiotropic phenotypes. A) Plant development. The upper panel shows the two genotypes at 20 DAS. Under it a picture of leaf morphology at 35 DAS. The leaves selected are true leaf number 4, 5 and 6 of the two genotypes. The lower-middle panel shows the two genotypes at 24DAS, and the lower panel the plants at 60DAS. B) Quantification of the time of bolting under LD conditions. The average numbers of leaves at bolting for Col-0 (13.93±0.83 leaf) and *tex1-4* (10.86±0.66 leaf) are depicted in the bar diagram (n=14). C) The two genotypes at 100DAS under SD conditions. D) The two genotypes at 120DAS under SD conditions. E) Quantification of the bolting time under SD conditions. The average numbers of leaves at bolting (SD conditions) for Col-0 (121.31±17.45 leaf) and *tex1-4* (82.33±8.35 leaf) are depicted on the bar diagram (n=13 and 12, respectively). F) The two genotypes shown at 40 DAS. G) Showing all elongated rosetta branches removed from a plant of both genotypes grown under LD conditions at 60 DAS. The ruler is 20 cm. long. H) Quantifications of number of rosetta branches in Col-0 (6.14±0.53 branch) and *tex1-4* (9.38±1.26 branch) at LD conditions at 60 DAS. The average number of inflorescences are depicted on the bar diagram (n=14 and 13, respectively). I) Fully elongated siliques at 44 DAS. J) Siliques at 44DAS that has been cleared to make the seed set visible. K) Flower morphology at 44 DAS. In the left panel, flowers are shown from above, and from the side with one petal and two sepals removed. In the right panel, isolated pistils and stamens are shown from the two genotypes. L) Roots of Col-0 and *tex1-1* at 8DAS. M) Quantifications of Lateral root branching density in Col-0 and *tex1-4* at 10DAS. The means of branching density (Lateral roots/cm) for Col-0 (7.74±1.21) and *tex1-4* (5.85±1.21) are depicted on the bar diagram (n=20 and 18, respectively). All the quantifications in these figures were tested for significant differences in the sample means by students T-test. Significance level was set to p<0.05. The error bars depicts the standard deviation.

was downregulated the plants produced 34.5% more rosetta branches than wild type (Figure 23H). The two-tailed student's t-test analysis showed that this difference in the sample means was statistically significant ($P=2.2\cdot10^{-7}$). *TEX1* seems to be involved in regulating the activation of rosetta branches.

The flower organs in *tex1-4* did not deviate in morphology from wild type (Figure 23I, J and K). This indicates that *TEX1* on its own is not required for proper flower and seed set development. TEX1 is associated with tasiRNA biogenesis. One of the tasiRNA genes, *TAS3*, that *tex1-4* has been shown to regulate is involved in releasing repression of lateral root growth (Marin et al., 2010; Jauvion et al., 2010; Yelina et al., 2010). *TAS3* transcript together with *miR390* inhibits expression of the three repressors ARF2, ARF3 and ARF4 (Marin et al., 2010). In order to see if the roots of *tex1-4* were affected, the number of elongated lateral roots per unit length of the branching zone was calculated. The branching zone was defined as the stretch of root beginning at the first elongated lateral root to the last elongated root (Dubrovsky & Forde, 2012). The seedlings were sown out on MS-plates placed vertically and grown 10DAS (Figure 23L). The ratio between the number of elongated lateral roots and the length of the branching zone, was calculated for *tex1-4* and wild type (Figure 23M). Compared to wild type, *tex1-4* has 24.5% reduced lateral root density. The difference was statistically significant. This indicates that TEX1 is involved in the regu-

lation of lateral root elongation in *Arabidopsis*.

When the *TEX1* transcript is downregulated it leads to some pleiotropic phenotypes like earlier flowering, more activated axillary meristems and less elongated lateral roots per unit length, compared to wild type. This indicates that TEX1 plays a general role in plant development.

3.2.3 GFP fused TEX1 complements the earlier flowering phenotype in *tex1-4*

To confirm that the observed pleiotropic phenotypes in *tex1-4* was really due to the lack of *TEX1* transcript it was decided to see if a genomic copy of *TEX1* could rescue the flowering phenotype. A construct with genomic allele of TEX1 fused to GFP was done (Figure 24A). The TEX1-GFP construct was transformed into the *tex1-4* background. Utilizing PCR-based genotyping, three independent lines were found that were homozygous for the *tex1-4* insertion and harboured the pGreenII0229-p*TEX1*::*TEX1*-GFP construct (Figure 24B). The three independent lines were also made homozygous for the GFP construct, by BASTA selection. The parent plants from which no progeny died after selection, were homozygous for pGreenII0229-p*TEX1*::*TEX1*-GFP, which was also confirmed by PCR against the inserted construct (Data not shown). P15 and P16 bound in a way to distinguish between the native *TEX1* allele and the transgene *TEX1* allele. It was seen that a transcript is only amplified for the wildtype. This shows that all three p*TEX1*::*TEX1*-GFP lines have no native *TEX1* allele without the T-DNA insertion. P17 and P18 amplifies only the transformed construct. It is seen that the three- transgene lines all have the p*TEX1*::*TEX1*-GFP insert. To test if the transgene lines had restored the expression levels of *TEX1,* RNA was extracted and cDNA was prepared. The relative expression levels were analysed by semi-quantitative RT-PCR for wildtype, *tex1-4* and the three transgenic lines p*TEX1*::*TEX1*-GFP-1, p*TEX1*::*TEX1*-GFP-2 and p*TEX1*::*TEX1*-GFP-3 (Figure 24C). The two *TEX1* primers, P33 and P34, spans the T-DNA insertion. The expression analysis indicates that TEX1 expression is increased in the three p*TEX1*::*TEX1*-GFP lines, when compared to *tex1-4,* where no endogenous transcript was detected. For p*TEX1*::*TEX1*-GFP-2 and p*TEX1*::*TEX1*-GFP-3 the levels seemed to be around wildtype levels. For p*TEX1*::*TEX1*-GFP-1 it looks like the amount of *TEX1* was increased compared to wildtype. Wildtype, *tex1-4* and the three p*TEX1*::*TEX1*-GFP lines in *tex1-4* background were sown out on

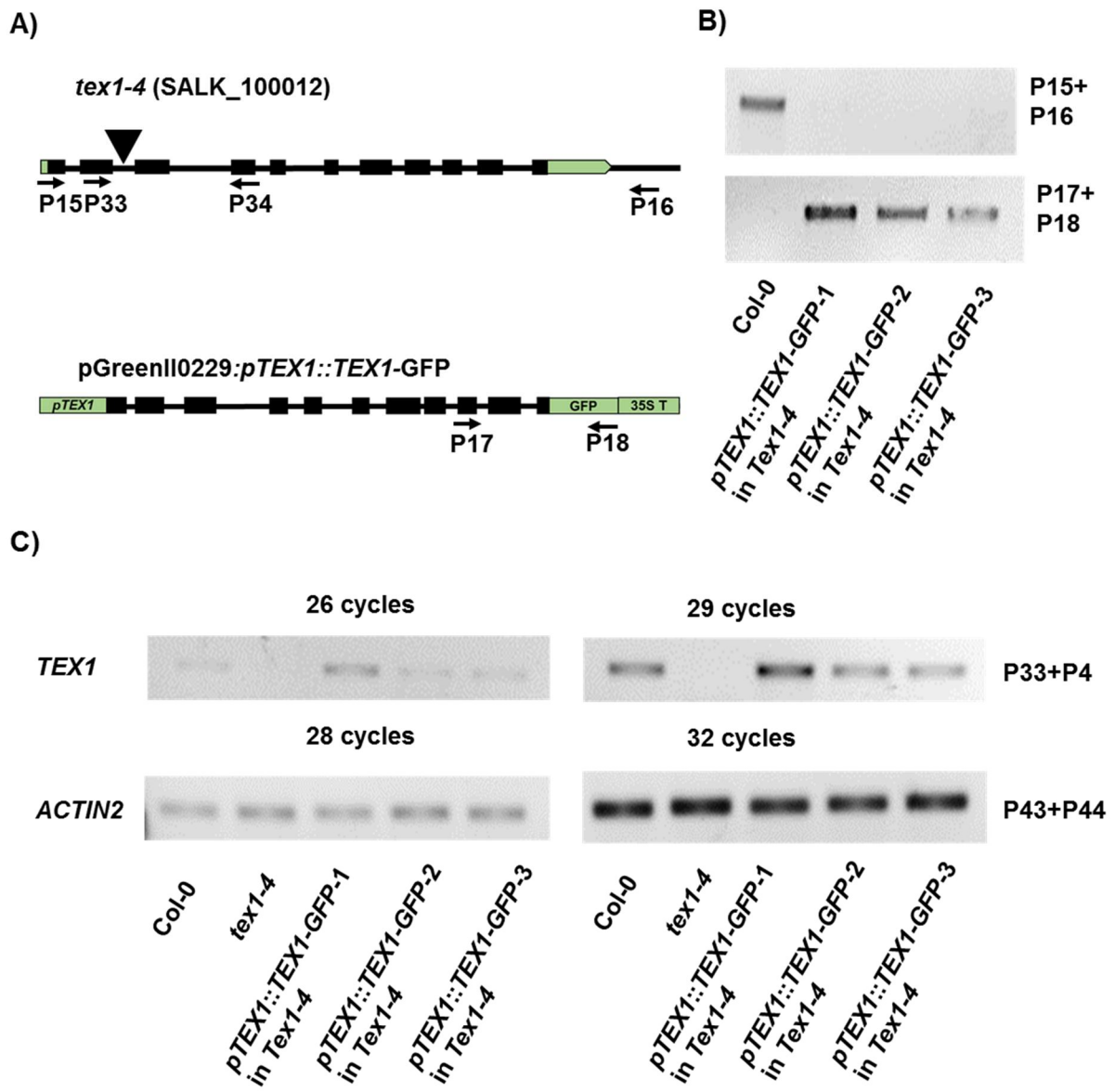


Figure 24. p*TEX1*::*TEX1-GFP* can restore expression of the TEX1 transcript in*tex1-4.* A) Schematic representation of *TEX1* (At5g56130) with the T-DNA insertion and the *pTEX1::TEX1-GFP* construct shown below. The used primers are also shown. B) Genotyping of transfected lines. Upper panel shows primers that amplify the wildtype *TEX1* allele. Lower panel shows primers that amplify a fragment specific for the introduced construct. C) Semi quantitative RT-PCR was performed on RNA extracted from wild type, *tex1* and three independent lines transfected with *pTEX1::TEX1-GFP.* For each primer pair the RT-PCRs were performed with two different numbers of amplification cycles to show that the PCR reaction was not saturated. *ACTIN2* was used as the reference gene. Seedlings were grown on MS plates, and harvested at 10DAS.

soil and grown under both LD and SD conditions. Under LD conditions at 24 DAS it was seen that the three p*TEX1*::*TEX1*-GFP lines looked more like wild type than *tex1-4* (Figure 25A). The T-DNA line *tex1-4* exhibited both deviant leaf morphology and earlier flowering, which was not observed in the other genotypes. To quantify the shift from vegetative to the reproductive state, the number of leaves at bolting were counted (Figure 25B). The results were evaluated by a one-way ANOVA analysis

(Chapter 7.1 (Table 24)). The null hypothesis that all means were equal was rejected ($P=3.39 \cdot 10^{-9}$). This means that at least one sample mean was different from the others. Tukey's test was used to find out which means are different. The mean number of leaves for *tex1-4* was significantly different from wildtype and the three p*TEX1*::*TEX1*-GFP lines in *tex1-4* background. The mean leaf numbers of

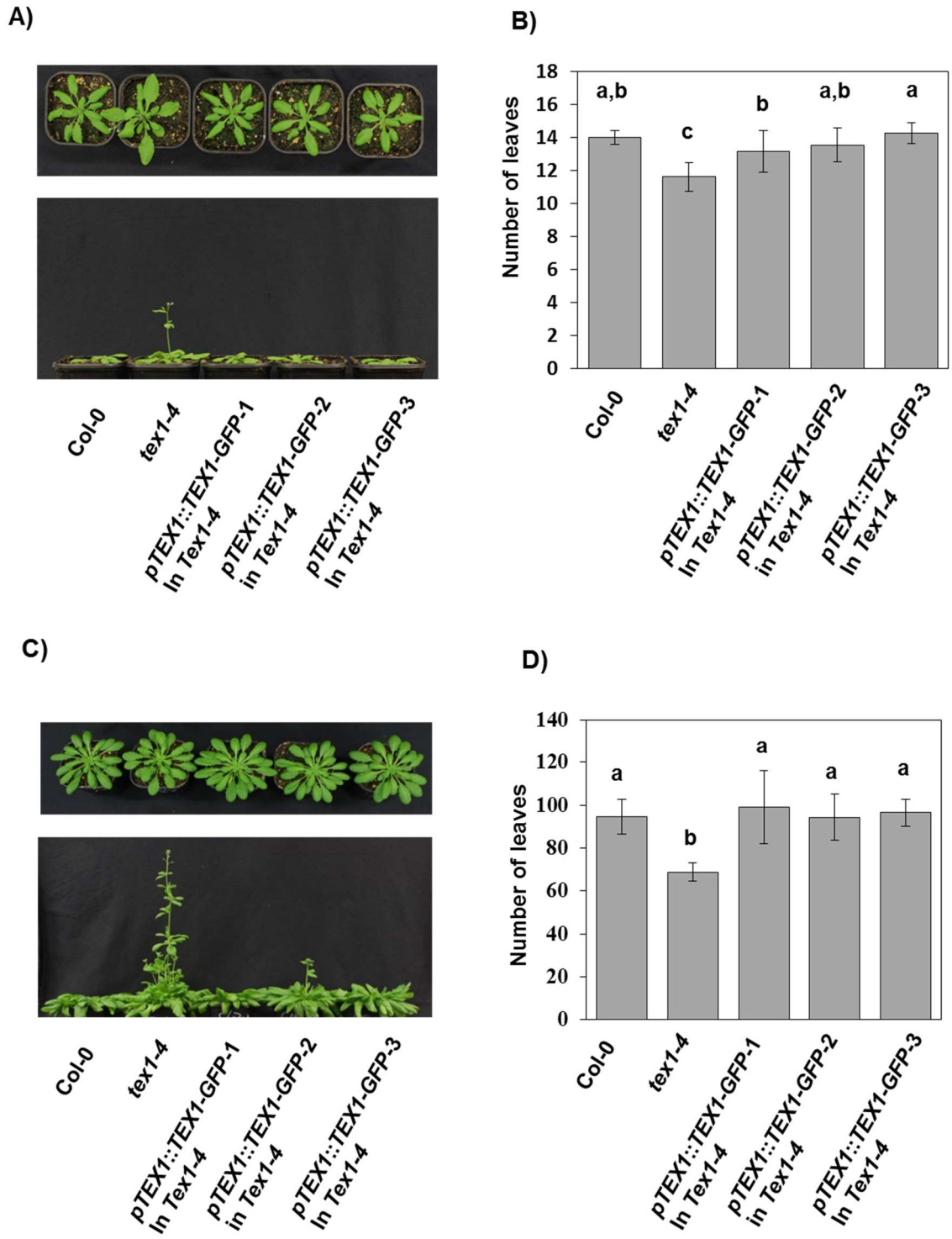


Figure 25. *pTEX1::TEX1-GFP* can rescue the *tex1-4* phenotype. A) Morphology of wild type, *tex1-4* and three independent lines transfected with *pTEX1::TEX1-GFP* at 24 DAS. The plants were grown under LD conditions. B) Quantification of the time of bolting under LD conditions. The average numbers of leaves at bolting for the different genotypes are depicted on the bar diagram (Col-0 (n=13), *tex1-4* (n=13), line1 (n=12), line2 (n=11) and line3 (n=11)). The error bars depicts the standard deviations. The statistical one-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 24). C) Morphology of wild type, *tex1-4* and three independent lines transfected with *pTEX1::TEX1-GFP* grown under SD conditions. Upper panel shows the plants at 70DAS. Lower panel shows the plants at 100DAS. D) Quantification of the time of bolting under SD conditions. The average numbers of leaves at bolting for the different genotypes are depicted on the bar diagram (Col-0 (n=12), *tex1-4* (n=8), line1 (n=10), line2 (n=11) and line3 (n=8)). The error bars depicts the standard deviation. The statistical one-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 25).

p*TEX1*::*TEX1*-GFP-1 and p*TEX1*::*TEX1*-GFP-3 were also significantly different from each other. None of the p*TEX1*::*TEX1*-GFP lines had a significantly different number of leaves compared to wildtype. The experiment was repeated for the five lines under SD conditions. The difference in leaf morphology was not as visible under SD conditions as under LD conditions (Figure 25C). The early flowering phenotype of *tex1-4* also seemed to be rescued under SD conditions. To quantify this the leaf number at bolting was quantified (Figure 25D). According to the one-way ANOVA analysis (Chapter 7.2 (Table 25)) the null hypothesis could also in this case be rejected ($P=1.92 \cdot 10^{-6}$). Tukey's test showed that the *tex1-4* leaf number was significantly different from all the other genotypes. Taken together this indicates that introducing the p*TEX1*::*TEX1*-GFP construct into *tex1-4* can rescue the early flowering phenotype. This suggest that *TEX1* is responsible for the observed phenotypes.

3.2.4 TEX1 fused to GFP exhibits specific subnuclear localization in root cells.

To confirm that the p*TEX1*::*TEX1*-GFP transcripts were translated into a protein fused to GFP, a western blot analysis was done. Proteins were extracted from the three transgenic p*TEX1*::*TEX1*-GFP lines and for a transgenic line of p35S::GFP-NLS. The protein extracts were separated by SDS-PAGE, and analysed by immunoblotting using an α-GFP antibody (Figure 26). It was seen that a protein with the right size of the TEX1-GFP fusion protein was detected for the three p*TEX1*::*TEX1*-GFP lines and not in the NLS-GFP control. Furthermore, a band at the same size as free GFP was observed in the three p*TEX1*::*TEX1*-GFP lines. The expression pattern and subcellular localization of TEX1 protein fused to GFP was analysed in the root (Figure 27 and Figure 28, respectively). When comparing the three independent lines of TEX1 fused to GFP, the protein was found to be expressed in nuclei distributed evenly all over the root and root tip. The subnuclear localization was investigated by CLSM. In contradiction to what was seen for the structure illumination microscopy (Figure 27) both a nucleoplasmic and cytoplasmic signal is observed with the CLSM (Figure 28) in the three independent lines. As the same seeds were used, the observed difference is most likely due to differences in detector

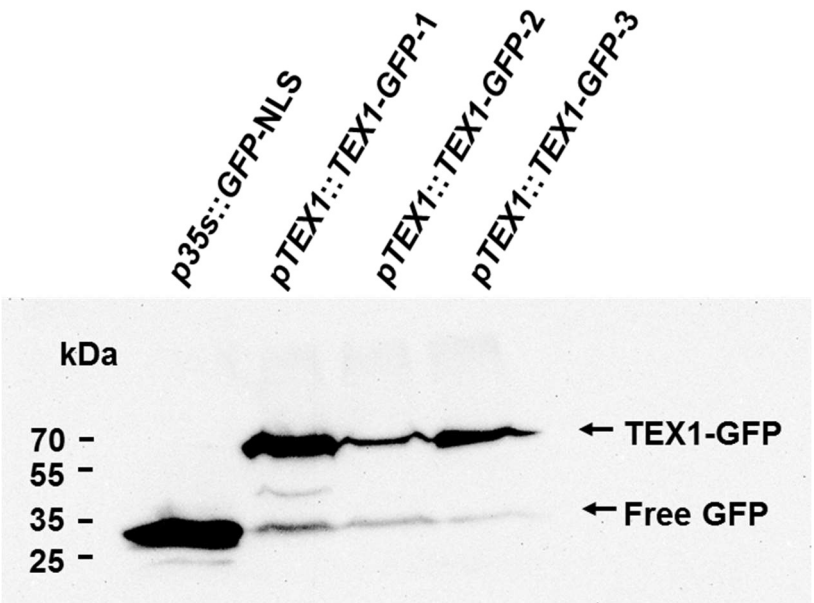


Figure 26. The TEX1-GFP fusion protein is expressed in the transformed lines. Western blot analysis. Protein extracts of independent lines transfected respectively with *pGreenII0229:pTEX1::TEX1-GFP* or *pGreenII0229: p35S::GFP-NLS* were separated on a 9% polyacrylamide gel and analysed by western blot analysis. Primary antibody was α-GFP (G1546, sigma). The TEX1-GFP fusion protein is 62 kDa, and GFP NLS is around 27 kDa. The loaded protein amounts were not normalized, so a semi-quantitative comparison is not possible.

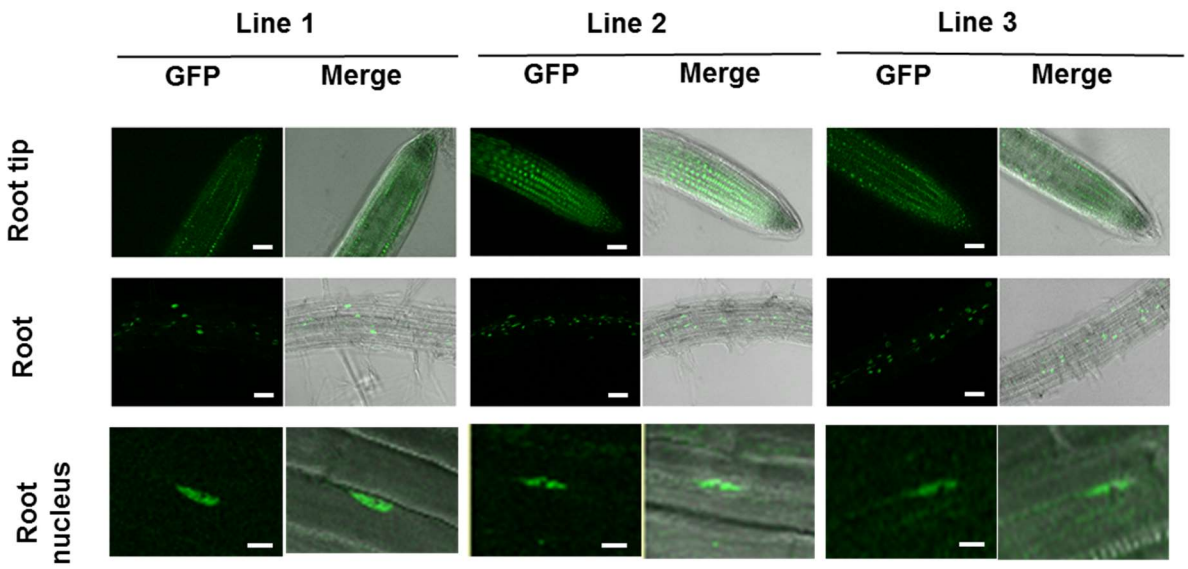


Figure 27. TEX1 fused to GFP exhibit a broad expression profile. ApoTome optical sections made by structured illumination microscopy. The nuclear localization of the TEX1 protein was analysed in three independent lines transfected with the expression vector. In the upper and middle panel, optical sections of the root tip and root are shown. A 20X/0.8 objective was used. Scale bar indicates 50 µm. The lower panel shows a root nucleus. A 40X/1.4 objective was used. Scale bar indicates 10µm. The optical sections are also shown as merges with the Bright field channel. Plants were grown on MS plates and analysed at 5DAS.

sensitivity. The detectors of the CLSM are more sensitive than the ones for the Apo-Tome system. For the p*TEX1*::*TEX1*-GFP-2 line the fluorescence signal was weaker compared to the signal of the two other lines, so the CLSM optical slide was taken in the epidermis cells, instead of the plane around the xylem vessel elements. For p*TEX1*::*TEX1*-GFP-1 and p*TEX1*::*TEX1*-GFP-3 it is seen that the cytoplasmic signal is not equally distributed in the root, but it seems to be stronger around the xylem vessel elements. Due to the potential presence of free GFP (Figure 26) in the lines, it is hard to conclude if TEX1 is really in the cytoplasm or not. No signal peptides were detected in the TEX1 sequence using the SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP/). TEX1 fused to GFP did exhibit a specific sub-nuclear distribution pattern though when comparing to the localization of NLS-GFP in the nucleus (Figure 11). Tex1 fused to GFP accumulated in speckle-like structures mainly throughout the nucleoplasm, like what was seen for ALY1 and ALY2 (Figure 13 and Figure 14). These data indicate that TEX1 is found throughout the whole root and show specific sub-nuclear localization. TEX1 may possibly be found in the cytoplasm with the strongest cytoplasmic signal around the xylem vessel elements.

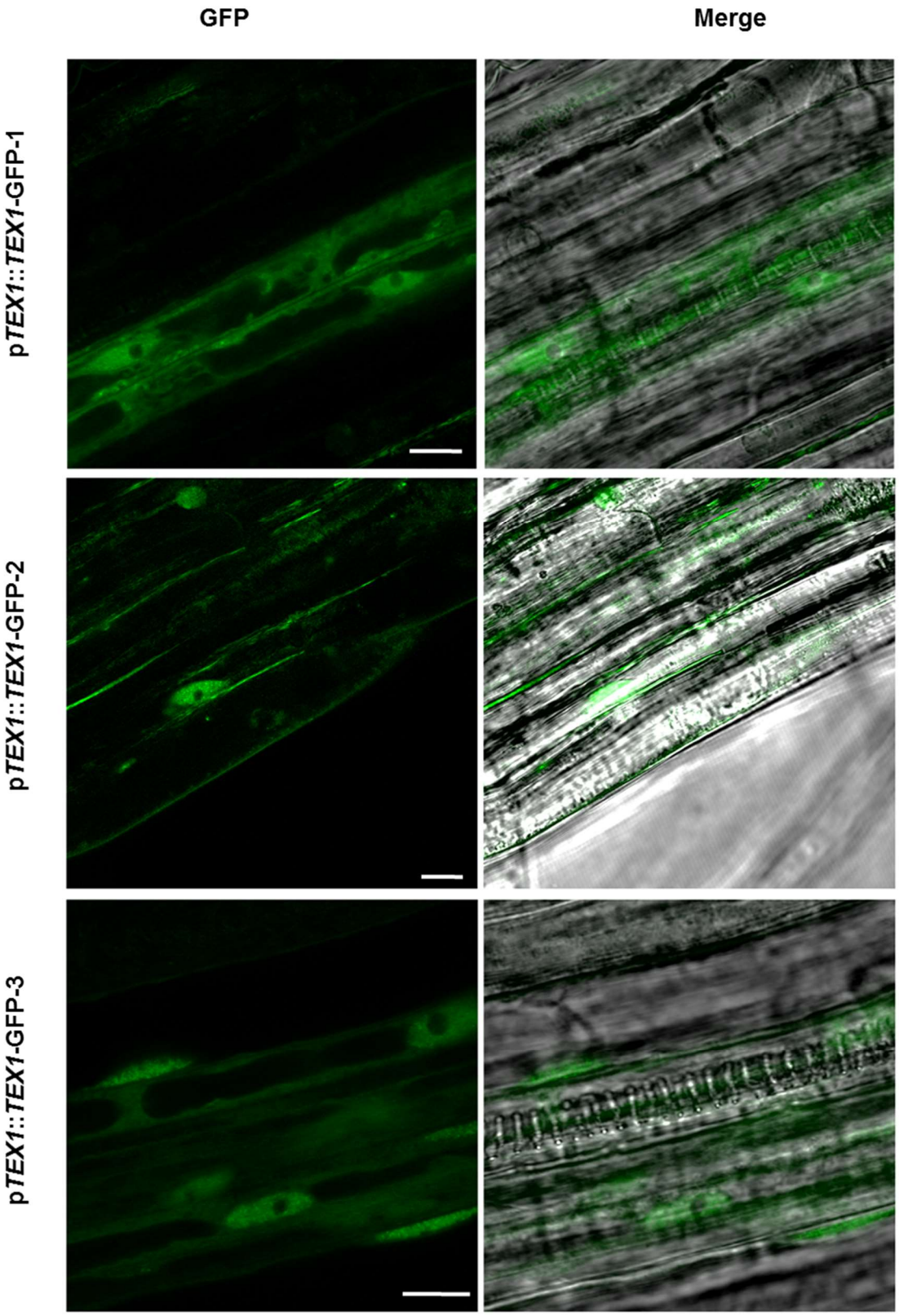


Figure 28. TEX1-GFP localizes mainly in the nucleoplasm, with a faint cytoplasmic signal. Confocal laser scanning microscopy optical section of root nuclei of the three independent lines with a 63X/1.4 objective. Optical slides from all three independent lines are shown. Scale bar indicates 10µm. Plants were grown on MS plates and analysed at 5DAS.

3.2.5 Loss of the *TEX1* transcript does not lead to mRNA accumulation.

To investigate if the phenotypes observed in *tex1-4* were caused by deficient mRNA export, Whole mount *in situ* hybridization was done on seedlings collected 6DAS. The optical slides were processed as described in chapter 3.1.5 and the ratios of

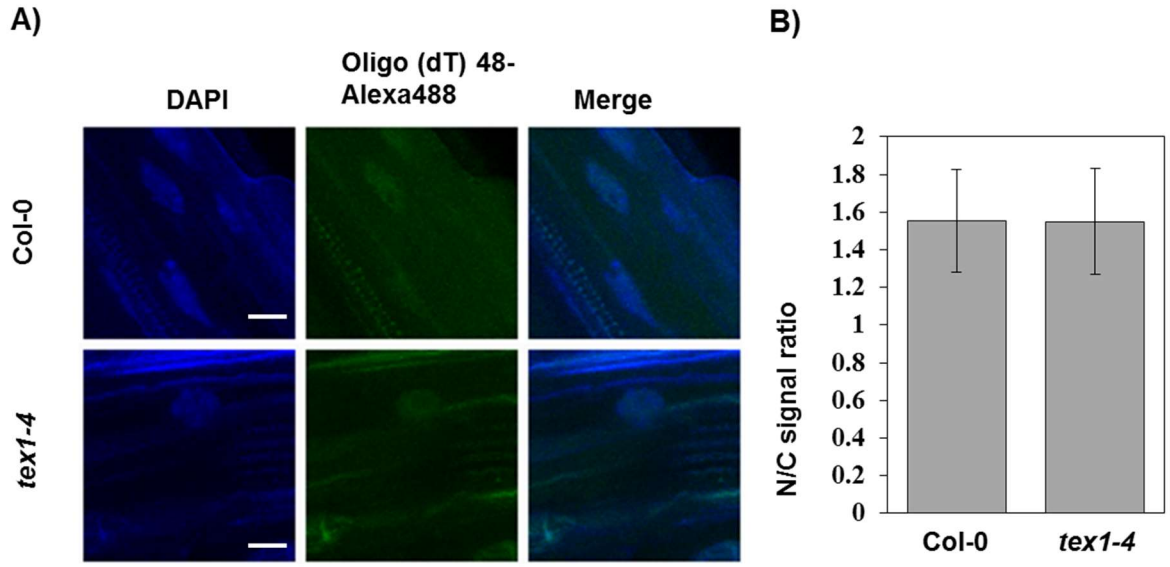


Figure 29. Bulk mRNA export is not affected in *tex1-4*. Plate grown seedlings at 6DAS were used. A) Whole mount *in situ* hybridization analysis. A 48-mer Oligo (dT) probe fluorescently labelled with Alexa488 was hybridized to polya(A) mRNA in fixated root tissue of the four genotypes. The optical sections were made around the plane of the xylem vessels using CSLM. A 40X/1.3 objective was used. Twelve z-stacks of 0.8µm were transformed with a maximum projection algorithm. DAPI was used as a counterstain. Scale bar indicates 10µm. B) Quantification of bulk mRNA export. The N/C ratio was calculated for 60 nuclei of wildtype (1.55±0.27) and *tex1-4* (1.55±0.28). The data was statistically evaluated with students T-test.

nucleoplasmic to cytoplasmic levels of polyadenylated RNA were calculated (Figure 29A and B). No obvious differences were observed by visual inspection of the pictures. When evaluating the quantifications statistically with a two-tailed students t-test no significant difference was observed between the sample means (P=0.92). This indicates that reducing the *TEX1* transcript levels does not lead to a bulk accumulation of mRNA in the nucleus.

3.2.6 The THO complex and MOS11 can be affinity purified together with UAP56.

In *Arabidopsis thaliana,* so far only the eight proteins of the core THO complex have been shown to interact (Yelina et al., 2010). They purified Flag-tagged TEX1 from seedlings, and they did not identify any of the non-THO components of TREX. In order to identify the putative subunit composition of *Arabidopsis* TREX, an alternative

approach was developed by modifying the approach described in Van Leene et al., 2011. The baitproteins were fused to a modified TAP tag GS (The C-terminal version is called SG tag) (Van Leene et al., 2008), under the control of the 35S promoter (Figure 30A). The bait constructs were introduced into PSB-D cell culture of *Arabidopsis landsberg erecta* by *Agrobacterium tumefaciens* mediated transformation. Crude protein extracts were added Benzonase nuclease, to interrupt unspecific interactions through DNA or RNA. The affinity purification was done as a one-step purification only utilizing the Protein G part of the SG tag. Eluates were separated by SDS-PAGE, lanes were cut out, protein were digested in-gel, peptides were extracted and then analysed by Liquid chromatography-mass spectrometry (LC-MS/MS). Four affinity purifications were done with each Bait protein and the empty SG tag as a control (Figure 30B). The SG fused Proteins were seen as the most prominent bands after elution. The bands of the Putative interaction partners were much weaker than the bait proteins, indicating that interaction partners were purified in substoichiometric amounts. After doing the MS analysis, contaminants found in the empty SG tag affinity purification were removed from the data set. These data are found in Chapter 7.4. Only proteins with a protein score higher than 80, more than two unique peptides pr. protein and that was found in at least three of four purifications were selected for further analysis. Only one protein was found in all three affinity purifications, namely UAP56 (Figure 30C). It was seen that more potential protein interaction partners were found in the TEX1 purifications than in the ones of MOS11 and UAP56 (348, 16 and 68 proteins respectively). When analysing the TEX1 data it was observed that a lot of the potential interaction partners were also found on a list of promiscuous proteins published in Van Leene et al., 2015. The list from Van Leene et al., 2015 was created by comparing the MS results from more than 700 *Arabidopsis* bait proteins from 543 affinity purifications with 115 bait proteins divided in 62 divergent groups based on molecular process. As these proteins have affinity for many divergent bait proteins it was decided that this list is also useful for the experiment in this thesis, even though another matrix was used in Van Leene et al., 2015. Proteins that bind many targets with low specificity were not of interest in this study. The cut-off for when a protein is promiscuous, was set as suggested in Van Leene et al., 2015, where proteins found in at least three divergent groups were removed. After removing the promiscuous proteins (Figure 30C) there were still a lot more pro-

teins identified in the TEX1 affinity purification than in the ones for MOS11 and UAP56 (198, 2 and 30 proteins respectively). UAP56 was still found in all affinity

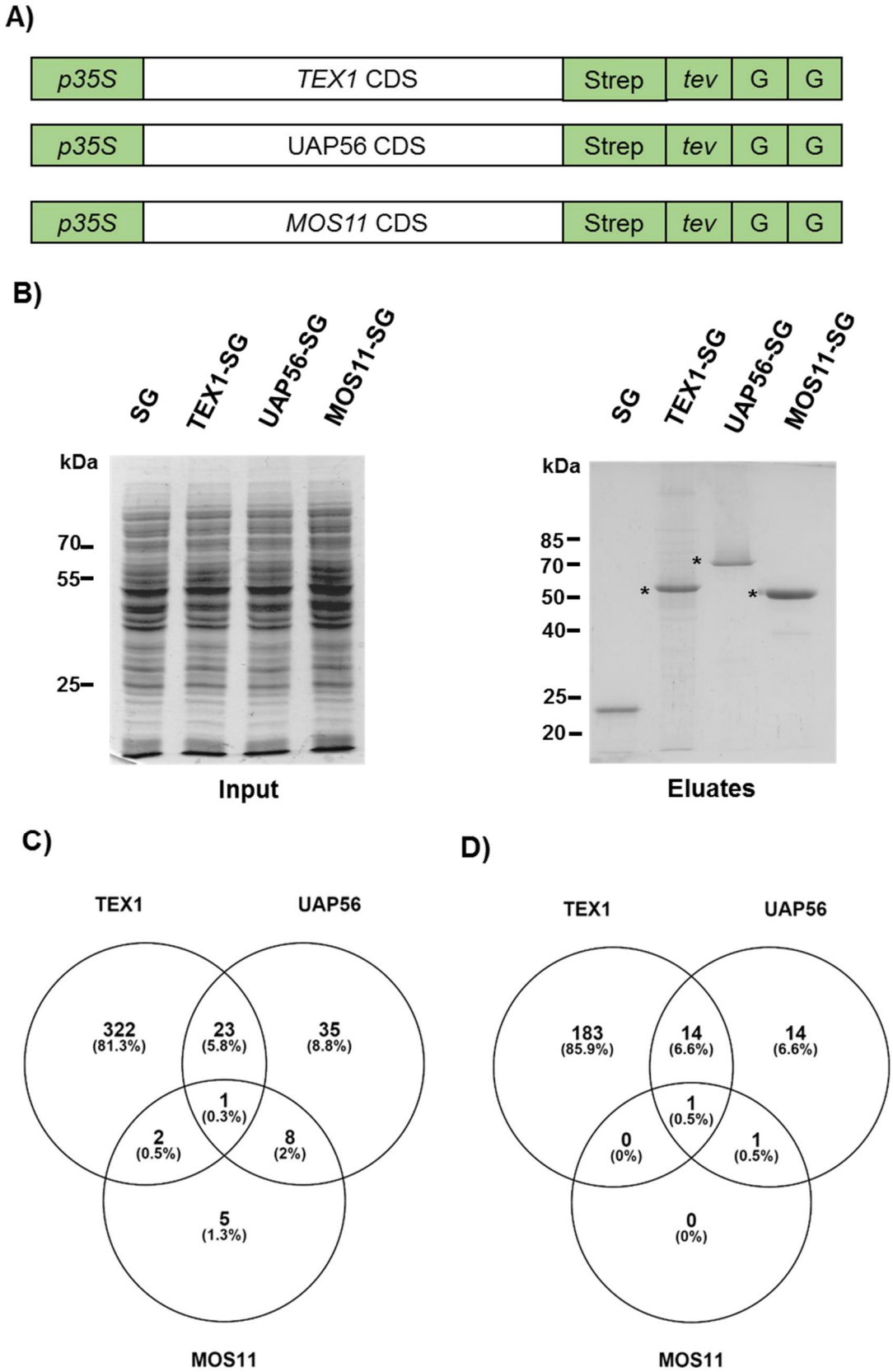


Figure 30. Protein interaction partners of TEX1, UAP56 and MOS11. A) Schematic overview of the TEX1, UAP56 and MOS11 coding sequences introduced into the pCAMBRIA-2300-SG transfection vector. The C-terminal SG tag consist of one streptavidin binding domain, a TEV cleavage site and two Protein G entities. B) Affinity purification of the three constructs, introduced into PSB-D cell culture of *Arabidopsis landsberg erecta* by *Agrobacterium tumefaciens*. Metal beads conjugated to rabbit antibodies were used to fish out the bait proteins with interaction partners. Left panel shows the inputs separated on a 9% polyacrylamide gel by SDS-PAGE. In the right panel, eluates of the three bait proteins together with the empty SG tag as control are seen separated on a 9% polyacrylamide gel by SDS-PAGE. Stars indicate the tagged bait proteins TEX1-SG (56.9 kDa.), UAP56-SG (69.9 kDa.) and MOS11 (44 kDa.). Lanes were excised and protein interaction partners were identified by tandem MS-MS (For all MS data, see Chapter 7.4) Venn diagram comparing all identified proteins with a protein score higher than 80, at least two unique peptides and found in at least 3 of 4 affinity purifications. Proteins found in the empty SG control has been removed. D) Venn diagram of the data from corrected for known sticky proteins found in Van Leene et al., 2015.

purifications. The affinity purifications of UAP56 and TEX1, both shared the entire core THO complex, which made out 57% of the proteins that were found in both purifications (Figure 30D and Table 15). For both UAP56 and TEX1 the interactors with the highest scores were THO2 and THO5, with the other THO subunits also being among the proteins with higher scores (Chapter 7.4(Table 54, Table 55 and Table 56)). In the MOS11 affinity purification only the two proteins MOS11 and UAP56 were found after removing the promiscuous proteins. To get an overview of all the proteins in the TEX1 purification a gene ontology analysis was done to find out if certain groups of proteins were overrepresented. The gene ontology analysis divided the proteins into groups based on which biological process they had been annotated to partake in.. A statistical analysis was done to estimate if any group of proteins are indicated as overrepresented by random. Significance level was set at $p<0.05$. The

Table 14 An enrichment of splicing factors is seen in the TEX1 affinity purification. The TEX1 interaction partners identified by LC MS-MS were analysed by the online gene ontology tool AgriGO (Du et al., 2010). The GO accession number, term, p value and false discovery rate (FDR) are shown for the ten most overrepresented terms. The significance level was set to for p<0.05. The GO analysis were done for Biological process.

GO_accession	Term	pvalue	FDR
GO:0006396	RNA processing	$4.30 \cdot 10^{-13}$	$2.10 \cdot 10^{-10}$
GO:0008380	RNA splicing	$8.40 \cdot 10^{-10}$	$2.00 \cdot 10^{-7}$
GO:0016070	RNA metabolic process	$6.70 \cdot 10^{-9}$	$1.10 \cdot 10^{-6}$
GO:0009791	post-embryonic development	$3.00 \cdot 10^{-8}$	$3.60 \cdot 10^{-6}$
GO:0016071	mRNA metabolic process	$6.60 \cdot 10^{-8}$	$6.30 \cdot 10^{-6}$
GO:0006397	mRNA processing	$2.60 \cdot 10^{-7}$	$2.10 \cdot 10^{-5}$
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	$6.20 \cdot 10^{-7}$	$4.30 \cdot 10^{-5}$
GO:0000375	RNA splicing, via transesterification reactions	$2.20 \cdot 10^{-6}$	$1.30 \cdot 10^{-4}$
GO:0003006	reproductive developmental process	$1.20 \cdot 10^{-5}$	$6.50 \cdot 10^{-4}$
GO:0006807	nitrogen compound metabolic process	$1.60 \cdot 10^{-5}$	$7.20 \cdot 10^{-4}$

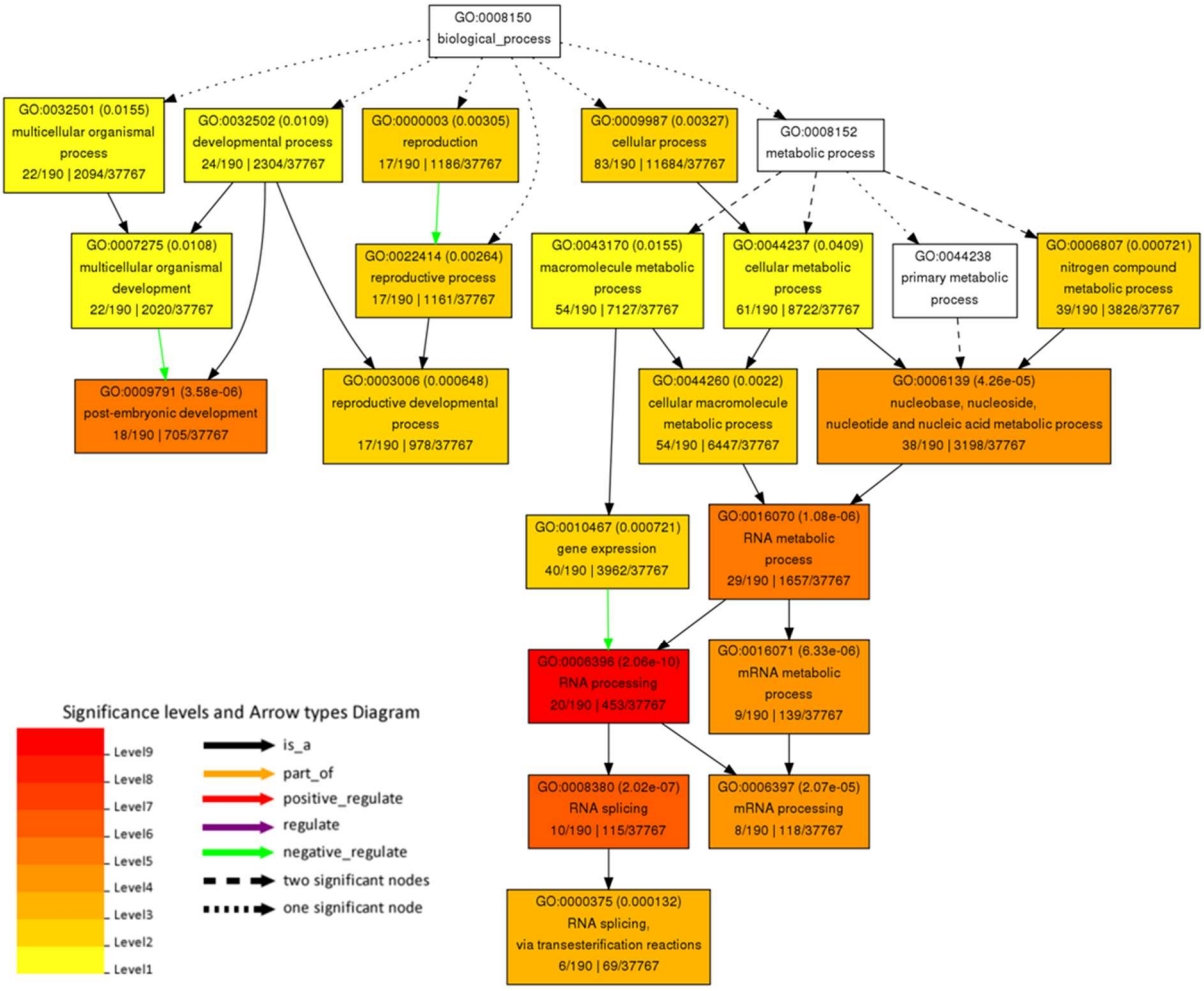


Figure 31. An enrichment of splicing factors is seen in the TEX1 affinity purification. The TEX1 interaction partners identified by LC MS-MS were analysed by the online gene ontology tool AgriGO (Du et al., 2010). This analysis investigates if any groups of proteins, belonging to a certain GO term, are overrepresented in the identified interaction partners. The GO analysis were done for Biological process. Levels of significance are shown in the lower left corner, level 9 is the most significant. The statistical analysis of overrepresented GO terms are found in Table 14. For simplicity, only the ten highest-ranking terms are shown.

summary of the GO Analysis for the proteins identified in the TEX1 affinity purification can be seen in Table 14. To simplify the outputs of the GO analysis only the ten most overrepresented GO terms are shown. Seven out of the ten most statistically significantly over-represented terms were associated with RNA splicing and processing. This is also shown as a visual representation in Figure 31. With the implication of a bias towards proteins involved in RNA processing and splicing, the dataset

Table 15. Protein interaction partners of TEX1, UAP56 and MOS11 involved in RNA biogenesis identified by LC-MS-MS. The different columns contain the *Arabidopsis* gene identifier, name and which complex the protein is found in and in which affinity purification it was detected. The average protein score and number of times found out four IPs are shown for each protein. Orthologs of the spliceosome components were collected from (Koncz et al., 2012), the rest were found by based on sequence similarity.

AGI	protein	complex	TEX1	UAP56	MOS11
AT1G24706	THO2	TREX complex	3355/4	5853/4	
AT5G42920	THO5	TREX complex	1922/4	3421/4	
AT5G56130	TEX1	TREX complex	1815/4	1013/4	
AT5G09860	THO1, HPR1	TREX complex	1534/4	2379/4	
AT5G16790	THO7	TREX complex	452/4	1685/4	
AT3G02950	THO7	TREX complex	430/4	1598/4	
AT2G19430	THO6/DWA1	TREX complex	341/4	1013/4	
AT1G45233	THO5	TREX complex	339/4	1634/4	
AT1G66260	ALY3	TREX complex	492/3		
AT5G02530	ALY2	TREX complex		405/4	
AT5G37720	ALY4	TREX complex		274/4	
AT5G11170	atUAP56b	TREX complex	131/3	4491/4	771/4
AT5G02770	atMOS11	TREX complex/ AMEX		2023/4	3123/4
AT2G19560	AtTHP1, ESSP1	TREX-2 complex		685/4	
AT3G54380	SAC3C	TREX-2 complex		619/4	
AT1G15200	atPinin	EJC/mRNP	171/4		
AT2G45640	HDAC subunit SAP18	EJC/mRNP	206/3		
AT1G51510	Y14	EJC/mRNP	109/3		
AT3G58570	RNA helicase 52	EJC/mRNP		158/3	
AT1G09770	CDC5	Core NTC	532/4		
AT5G28740	SYF1	Core NTC	441/4		
AT2G33340	MAC3B, PRP19B	Core NTC	297/4		
AT1G77180	SKIP	Core NTC	256/4		
AT5G41770	atCRN1c	Core NTC	290/3		
AT2G38770	atAquarius EMB2765	NTC-associated	811/4	159/43	
AT1G07360	RBM22/SLT11	NTC-associated	312/4	145/43	
AT1G80930	CWC22	NTC-associated	333/3		
AT1G32490	DHX16	NTC-associated	156/3		
AT2G46610	atRSp32	SR proteins	303/4		
AT3G49430	AtSR34a	SR proteins	292/4		
AT3G55460	atSCL30	SR proteins	192/4		
AT1G23860	atRSzp21/atSRZ21	SR proteins	301/3		
AT4G35785	RNA-binding protein	SR proteins	160/3		
AT1G55310	atSR33/atSCL33	SR proteins	133/3		
AT1G27650	atU2AF35a	Splice site selection	186/4		
AT1G60900	atU2AF65b	Splice site selection	142/3		
AT4G36690	atU2AF65a	Splice site selection	137/3		
AT5G17440	atLuc7b	U1 snRNP	166/3		
AT2G47580	atU1A	U1 snRNP	116/3		
AT5G25060	atSR140-1	17S U2 associated	301/3		
AT5G64270	atSAP155	17S U2 snRNP	500/4		
AT3G55200	atSAP130a	17S U2 snRNP	337/4		
AT4G21660	atSF3b150	17S U2 snRNP	142/4		
AT1G14650	atSAP114-1a s	17S U2 snRNP	140/3		
AT4G24270	PRP24	U4/U6 snRNP	117/3		
AT4G03430	atU5-102KD, STABILIZED 1	U5 snRNP	222/4		
AT1G20960	atU5-200-2a	U5 snRNP	934/3		
AT2G43770	atU5-40 Prp8 binding protein	U5 snRNP	196/3		
AT3G54230	Suppressor of abi3-5 (SUA)	A complex associated	466/3		
AT3G63400	peptidyl-prolyl cis-trans isomerase	Abundant first in C complex	193/3		
AT4G09980	methyltransferase-like protein 1	B complex associated	132/3		
AT3G26560	DHX8/PRP22	RES	192/4		
AT5G13010	atPrp16	RES	296/3		
AT3G62310	DHX15	RES	243/3		
AT1G17070	NTR1	RES	106/3		
AT3G26420	ATRZ-1A	hnRNP family	413/4		
AT1G60650	ATRZ-1B	hnRNP family	213/4		
AT5G04280	ATRZ-1C	hnRNP family	204/4		
AT2G44710	hnRNP-R3	hnRNP family	230/3		
AT4G21710	NRPB2	Polymerase II	259/3		
AT5G50320	ELP3; ELO3	Elongator	444/4		
AT2G06210	ELF8	PAF-C	122/3		
AT4G25550	CFIm-25	CFIm	158/3	115/4	
AT1G75660	XRN3 5'-3' exoribonuclease 3	Exosome	459/4		

was searched through for known proteins involved in transcription, splicing, mRNA biogenesis and mRNA export. A list of proteins that are part of the spliceosome was obtained from Koncz et al., 2012. Sorting the proteins based on involvement in RNA processing, were also done for the proteins identified in affinity purifications of UAP56 and MOS11. A table was composed that shows what factor interacts with which of the three bait proteins (Table 15). In the TEX1 affinity purification, components of the TREX complex, the EJC complex, the NTC and associated protein, the spliceosome U snRNP, RES splicing complex, Splice site selection factors, splicing related SR proteins, hnRNP proteins, transcript elongation factors, cleveage and polyadenylation factor CFIm25 and the XRN3 exoribonuclease were identified. Of the known non-THO subunits only UAP56 and ALY3 were identified. ALY2 and ALY4 were not detected in any of the four TEX1 affinity purifications. In the UAP56 affinity purification all the core THO complex subunits, MOS11, ALY2, ALY4, SAC3C (TREX2), THP1 (TREX2), RNA helicase 52 (EJC) and two NTC associated proteins were identified. ALY1 was found in the empty SG control and was removed from the data. ALY3 was detected in one of four UAP56 affinity purification with a low score, which is below the set threshold of this thesis (Chapter 7.4).

Regarding potential export receptors in *Arabidopsis* the TREX interactome was searched for proteins with NFT2L or UBA domains, as these has described to be crucial for the translocation through the NPC (Strässer et al., 2000; Katahira et al., 2002). No proteins with these domains were found among the raw affinity purification data using the NCBI conserved domain search (www.ncbi.nlm.nih.gov), meaning no obvious candidate for the mRNA export receptor was found (Data not shown). The interaction between THO and UAP56 was shown in both directions with TEX1 and UAP56 as bait proteins (reciprocal tagging approach). The same was seen for the MOS11 and UAP56 bait proteins in both directions. That the TEX1-UAP56 and UAP56-MOS11 interactions were confirmed from both sides increases the proability that these proteins really interact either directly or indirectly in *Arabidopsis.* Interestingly MOS11 was not found in the TEX1 affinity purification and the other way around, with the selection criteria used in this study. By using a new purification approach in cell culture more interactors of THO/TREX in *Arabidopsis* were identified, than previously shown in Yelina et al., 2010.

3.3 Characterization of TEX1 and MOS11

3.3.1 TEX1 and MOS11 colocalize, but TEX1 exhibits unique accumulation in speckles.

Both TEX1 and MOS11 interacts with UAP56 in cell culture (Table 15). It was decided to test if TEX1 and MOS11 colocalize, as it could reveal more about if MOS11 partake in the TREX complex in *Arabidopsis*. A construct of *MOS11* fused to Red fluorescent protein (RFP) was introduced into wildtype (Figure 32A). MOS11 fused to RFP was found in nuclei throughout the whole root (Figure 32B). Inspecting the sub-nuclear localization of MOS11 fused to RFP, showed that MOS11 was evenly

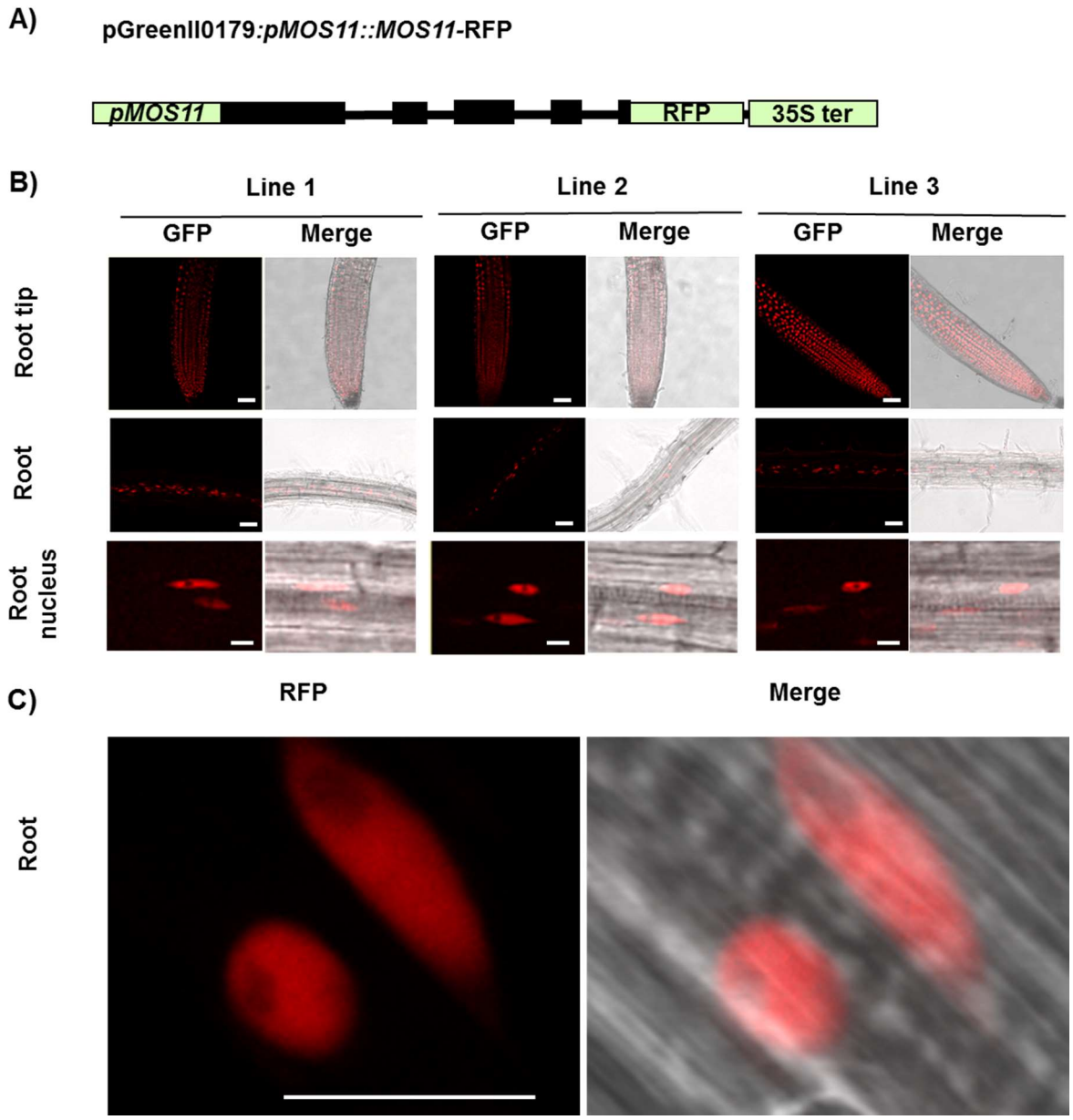


Figure 32. MOS11 fused to RFP is mainly found in the nucleoplasm. A) Schematic representation of the *pGreenII 0179:pMOS11::MOS11-RFP* construct under control of the native promoter. B) ApoTome optical sections made by structured illumination microscopy and confocal microscopy. The nuclear localization of the MOS11-RFP protein was analysed in three independent lines transfected with the expression vector. In the upper panel, confocal laser scanning optical sections of the root tip is shown. A 20X/0.8 objective was used. Scale bar indicates 50 µm. In the middle panel, Apotome optical sections of the root are shown. A 20X/0.8 objective was used. Scale bar indicates 50 µm. The lower panel shows root nuclei. A 40X/1.4 objective was used. Scale bar indicates 10µm. The optical sections are also shown as merges with the Bright field channel. C) Confocal laser scanning microscopy optical section of root nuclei of Line 1 with a 63X/1.4 objective. Scale bar indicates 10µm.

distributed in the nucleoplasm. MOS11 was mostly excluded from the nucleolus, and did not accumulate in speckle-like structures. MOS11 seems to share the overall and subcellular localization pattern of UAP56. To look for colocalization with TEX1 the line p*TEX1*::*TEX1*-GFP-3 in *tex1-4* was additionally transformed with p*MOS11*::*MOS11*-RFP-1. Selected plants from the T1 generation were analysed by CLSM. Ten nuclei from the root (Figure 33A) and ten nuclei from the leaf (Figure 34A) were analysed. The JACoP plugin for ImageJ was utilized for the colocalization analysis. The JACoP plugin and the theory behind the different methods used for the colocalization analysis is described in Bolte & Cordelières, 2006. Scatter plots were made. In the scatterplot, the green channel intensity of each pixel was plotted against the red channel intensity of the same pixel. If two fluorophores colocalized the points on the scatterplot show a linear correlation. For all root nuclei (Figure 33B) and leaf nuclei (Figure 34A), the same linear tendency in the scatterplot was seen. In order to estimate how well the data fit to the linear regression, the Pearson's Correlation Coefficient (PCC) was calculated.-1 and 1 signifies complete negative or positive correlation, whereas zero indicate no correlation. The PCC was calculated for all ten root nuclei and leaf nuclei (Table 16). For the root nuclei the PCCs were between 0.91 and 0.98. For the leaf nuclei, the PCCs were between 0.93 and 0.98. This was a first indication that for both roots and leaves the data fits very well to the linear regression, suggesting a correlation in localization.

To estimate the amount of colocalization, the Mander's' coefficients M1 and M1 (MANDERS et al., 1993) were calculated. A pixel in channel A was considered colocalized with the same pixel in channel B, if the pixel B had an intensity higher than a set threshold. The threshold was calculated automatically by the method proposed by Costes et al., 2004, to avoid user bias. M1 is an indicator of the proportion of channel 1 signal colocalizing with the signal of channel 2 (Bolte & Cordelières,

2006). For the roots TEX1-GFP colocalized with 92-99% of the MOS11-RFP signal (Table 16). MOS11-RFP colocalized with the TEX1-GFP signal in 98-100% of pixels with the calculated threshold. In the leaves, similar results were seen. TEX1-GFP colocalized with 90-99% of MOS11-RFP, MOS11 colocalized with 97-100% of the TEX1 signal. TEX1 appeared to localize a little bit less with MOS11 than the other way around. This could be due to weak cytoplasmic signals observed in the lines

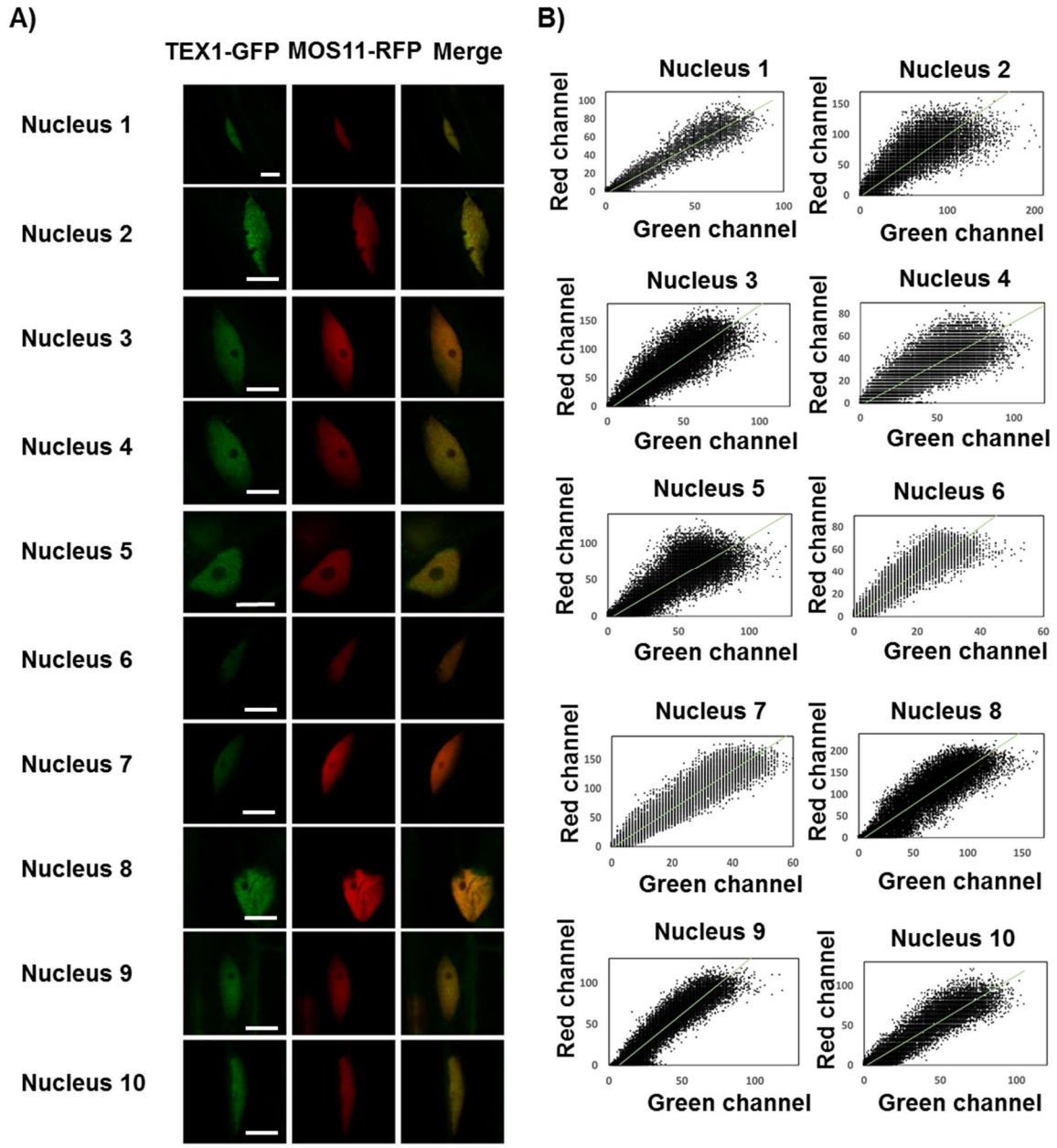


Figure 33. TEX1 and MOS11 exhibit a high degree of colocalization in root nuclei. p*MOS11*::*MOS11*-GFP-1 was transfected into p*TEX1*::*TEX1*-GFP-3 in *tex1-4* background. A) CLSM optical sections of ten root nuclei from 5 independent lines at 6DAS grown on MS plates. The signals from the red and green channels are shown together with a Merge of both. B) Scatter plots of the red and green channel intensities for the ten root nuclei. The line shown in the plots signifies the linear regression of the plotted points. Scale bar indicates 10µm. A 63X/1.4 objective was used.

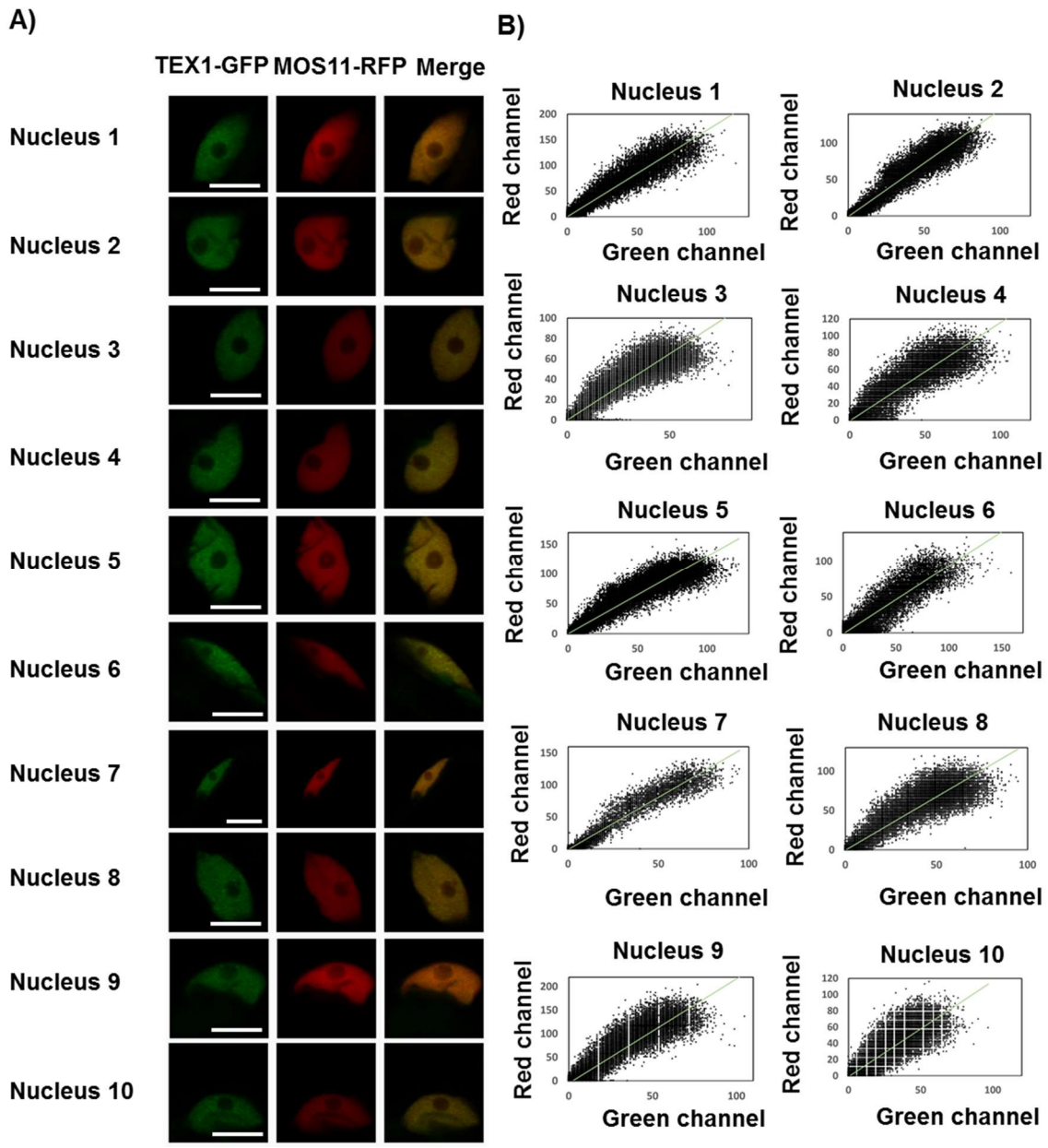


Figure 34. TEX1 and MOS11 exhibit a high degree of colocalization in leaf nuclei. p*MOS11*::*MOS11*-GFP-1 was transfected into p*TEX1*::*TEX1*-GFP-3 in *tex1-4* background A) CLSM optical sections of leaf nuclei from three independent lines at 6DAS grown on MS plates. The signals from the red and green channels are shown together with a Merge of both. B) Scatter plots of the red and green channel intensities for the ten root nuclei. The line in the plots shown signifies the linear regression of the plotted points. Scale bar indicates 10µm. A 63X/1.4 objective was used.

with TEX1 fused to GFP. To estimate if the colocalization is statistically significant, Costes randomization method was used (Costes et al., 2004). For each optical section, the green channel signal was randomized 200 times and compared to the original red channel signal. The PCC of each of the 200 randomizations were compared to the PCC of the original optical slides. A P-value was calculated. 95-100% signified

Table 16. Colocalization analysis of root and leaf nuclei. For each nucleus the Pearson's Correlation Coefficient, Mander's Coefficients M1 and M2 and P-values calculated by Costes Randomization is shown. P>95% signifies that the colocalization is not happening by random.

Nucleus (Root)	PCC	M1	M2	P-value
1	0.97	0.92	0.98	100.00%
2	0.93	0.95	0.99	100.00%
3	0.95	0.94	0.99	100.00%
4	0.94	0.92	0.99	100.00%
5	0.91	0.99	1.00	100.00%
6	0.96	0.95	0.99	100.00%
7	0.98	0.93	0.98	100.00%
8	0.95	0.97	0.99	100.00%
9	0.96	0.92	0.98	100.00%
10	0.96	0.95	0.98	100.00%

Nucleus (Leaf)	PCC	M1	M2	P-value
1	0.97	0.98	0.99	100.00%
2	0.98	0.98	1.00	100.00%
3	0.97	0.99	1.00	100.00%
4	0.96	0.95	0.99	100.00%
5	0.97	0.99	1.00	100.00%
6	0.94	0.92	0.97	100.00%
7	0.98	0.90	0.99	100.00%
8	0.97	0.97	0.99	100.00%
9	0.95	0.94	0.99	100.00%
10	0.93	0.93	0.99	100.00%

a statistically significant difference in the original picture compared to the 200 random ones. All root and leaf nuclei were significantly different from the randomized samples (Table 16). In spite of the high degree of colocalization according to the various tests, it had been observed that TEX1 accumulated in speckles (Figure 28). This was also observed in the lines used for colocalization (Figure 35A). MOS11 on the other hand did not accumulate in speckles as large and well defined as seen for TEX1.

Proteins that are involved in splicing often accumulate in splicing speckles (Reddy et al., 2012). TEX1 was shown to interact with many proteins of the spliceosome (Table 15). In order to see how the splicing speckles look in *Arabidopsis*, the SR protein atSRp30 was fused to RFP. atSRp30 is a paralog of atSRp34 involved in regulating alternative splicing, and localizes in speckle-like structures (Lopato et al., 1999). This work was done together with B.Sc. student D. Fiegle (Fiegle, 2014). The pGreenII0179:p35S::SR30-RFP construct was introduced into PSB-D cell culture of *Arabidopsis landsberg erecta* by *Agrobacterium tumefaciens* mediated transformation. After selection the cell culture, was analysed by CLSM (Figure 35B). SR30 localized in the nucleoplasm and not in the nucleolus. SR30 fused to RFP accumulated in speckle-like structure very similar to what was observed when TEX1 was fused to GFP. Taken together with the results of the affinity purification of TEX1 fused to SG, this indicates that TEX1 might be associated with the splicing machinery in speckles. MOS11 colocalize with TEX1, but does not accumulate in speckles.

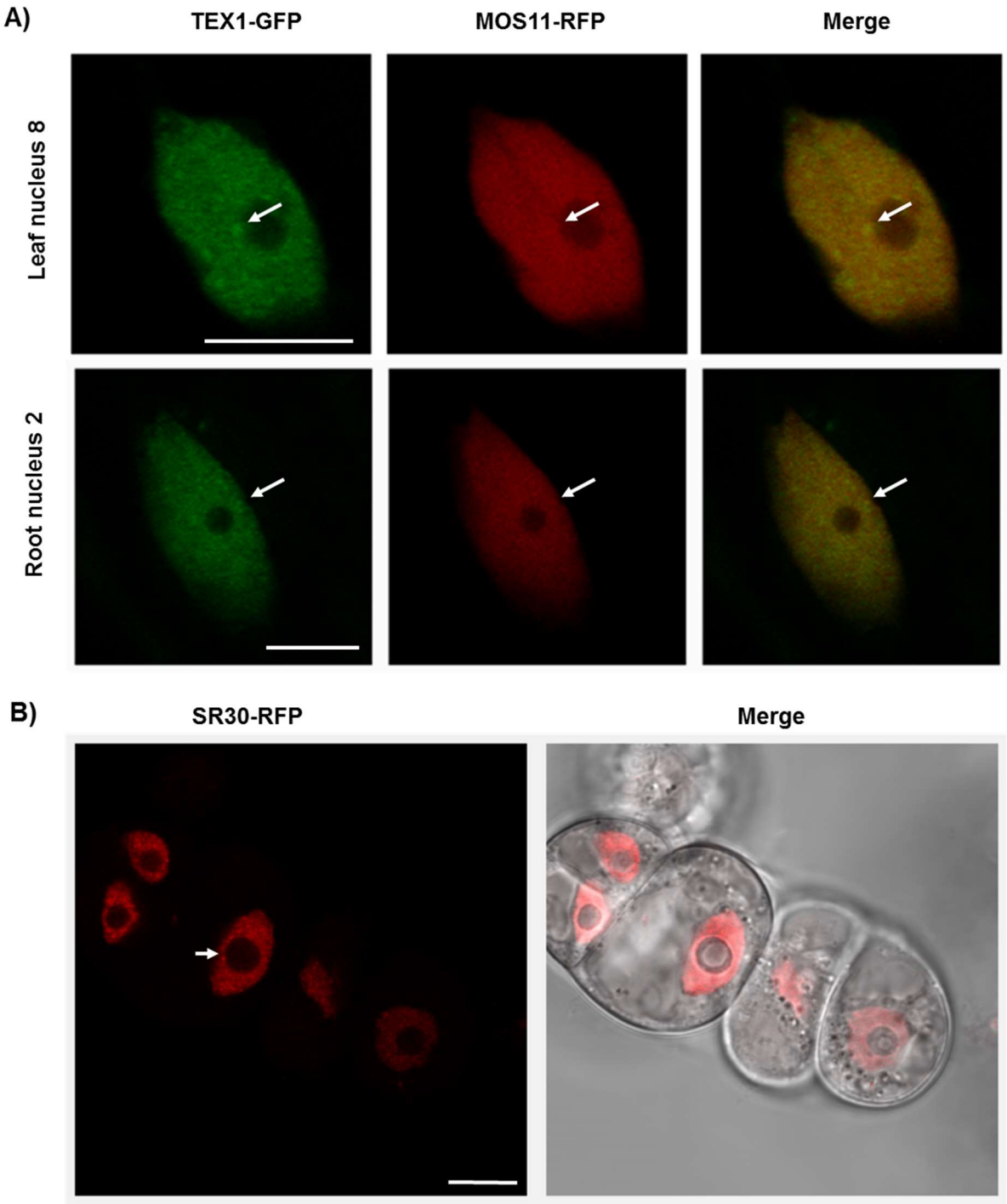


Figure 35. TEX1-GFP localizes in Speckle-like structures. A) Magnification of root nucleus 2 and leaf nucleus 8 from Figure 33A and Figure 34A. The signals from the red and green channels are shown together with a Merge of both. Scale bar indicates 10µm. A 63X/1.4 objective was used. B) Splicing marker. Confocal laser scanning microscopy optical sections of PSB-D cell culture of *Arabidopsis landsberg erecta* transfected with pGreenII 0179:p35S::SR30-RFP. A 20X/0.8 objective was used. The arrows points to an example of a speckle-like structure. The optical sections are also shown as merges with the Bright field channel. The SR30-RFP optical section were done together with B.Sc. student Dominik Fiegle (Fiegle, 2014)

3.3.2 Analysis of single and double mutants deficient in *TEX1* and *MOS11* transcript levels.

To investigate if MOS11 and TEX1 are affecting the same the functional pathways, it was decided to create a double mutant. Function and relationship can be inferred from analysing how the phenotypes are affected in a double mutant. The T-DNA insertion line SAIL-266_E03 (*mos11-2*) was described in Germain et al., 2010. The position of the T-DNA insertion was identified by PCR and sequencing. The sequencing results indicated that the T-DNA insertion was inserted into the second intron, as opposed to the first exon, which was reported in Germain et al., 2010. The two T-DNA lines *mos11-2* and *tex1*-4 were crossed to create a Double mutant. Schematic overviews of the two genes with the primers used for genotyping are shown in Figure 36A. DNA was extracted at 10DAS from plate-grown seedlings of wildtype, *tex1-4*, *mos11-2* and *tex1-4 mos11-2*. PCR with P13 and P14 showed that only Wildtype and *mos11-2* had the uninterrupted *TEX1* wildtype allele (Figure 36B). With primers P3 and P13, it was seen that there was only a T-DNA inserted in *TEX1* in the two *tex1-4* and *tex1-4 mos11-2* genotypes. The uninterrupted wildtype *MOS11* allele, was only detected in wildtype and *tex1-4* with P19 and P20 (Figure 36B). A *MOS11* T-DNA insertion was only present in *mos11-2* and *tex1-4 mos11-2* when using P1 and P19. This meant that single and double mutants of *TEX1* and *MOS11* were now available. To test if the different genotypes have deficient expression levels of *TEX1* and *MOS11* an expression analysis using Quantitative real time PCR

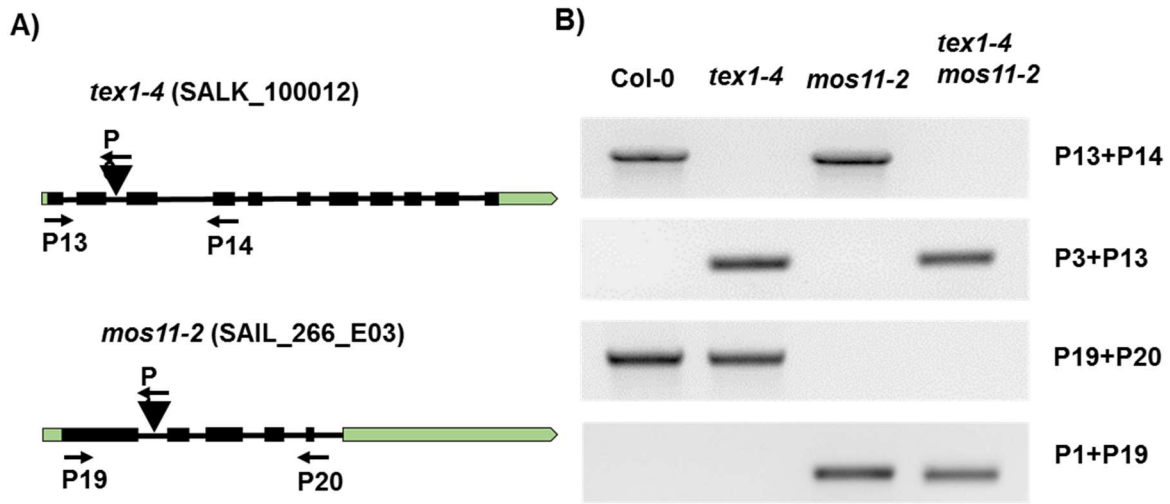


Figure 36. Molecular characterization of *tex1-4 mos11-2* double mutant. A) Schematic representation of *TEX1* (At5g56130) and *MOS11 (*At5g02770) with the T-DNA insertion and primers shown. B) Genotyping of wild type, *tex1-4*, *mos11-2* and *tex1-4 mos11-2* using indicated primers. DNA was extracted from 10DAS. Seedlings grown on MS plates.

(RT-qPCR) was done. The different primers used are seen on the schematic representations of the *TEX1* and *MOS11* alleles (Figure 37A). The Data analysis and quality control were done with the qBase+ (Biogazelle) program based on Hellemans et al., 2007. Additional information about the RT-qPCR data can be found in Chapter 7.3. For all RT-qPCR analyses the reference genes *ACTIN2* (AT3G18780) and Translation elongation factor 1 alpha/*EF1-α* (AT5G60390) were used. They have both been found to be stable reference genes in *Arabidopsis thaliana* (Czechowski et al., 2005). To monitor the expression levels of the *TEX1* transcript three sets of primers were used (Figure 37A). P29 and P30 amplified a transcript upstream of the T-DNA insertion. P31 and P32 amplified a transcript downstream of the T-DNA insertion, whereas P33 and P34 spanned the T-DNA insertion. The statistical analyses are found in Chapter 7.2 (Table 26-Table 28) For the transcript upstream of the T-DNA insertion only the factor TEX1 was significant (Figure 37B). For the transcript downstream of the T-DNA insertion the interaction term was significant ($P=2.90 \cdot 10^{-4}$). To find differences in the sample means the Tukey's test was applied (Figure 37C). Interestingly *tex1-4*, *mos11-2* and *tex1-4 mos11-2* were all significantly different from wildtype. The two mutants *tex1-4* and *tex1-4 mos11-2* were reduced to the same low levels. In *mos11-2* the *TEX1* transcript was expressed significantly at higher levels than in wildtype. The same was seen for the primer pair that spans the T-DNA insertion (Figure 37D). There was an interaction between the two factors TEX1 and MOS11 (P=0.03). In *tex1-4* and *tex1-4 mos11-2* the transcript was equally downregulated according to Tukey's test. In *mos11-2* the *TEX1* transcript was also expressed at higher levels than what was observed in the wildtype. Taken together this indicates that the T-DNA insertion in *TEX1* effectively reduce the transcription levels of *TEX1* to almost background levels. As there was observed no transcript upstream or downstream of the T-DNA insertion, it would indicate that there is also no truncated protein present. This indicates that tex1-4 is a functional knockout mutant of *TEX1*. For the expression of the *MOS11* transcript among the different genotypes, three sets of primers were used (Figure 37A). P35 and P36 amplified a transcript upstream of the T-DNA insertion. P37 and P38 amplified a transcript downstream of the T-DNA insertion, whereas P39, and P40 spanned the T-DNA insertion. The statistical analyses are found in Chapter 7.2 (Table 29-Table 31). The results of the RT-qPCR analysis with the amplicon of P35 and P36, can be seen in Figure 37E. The

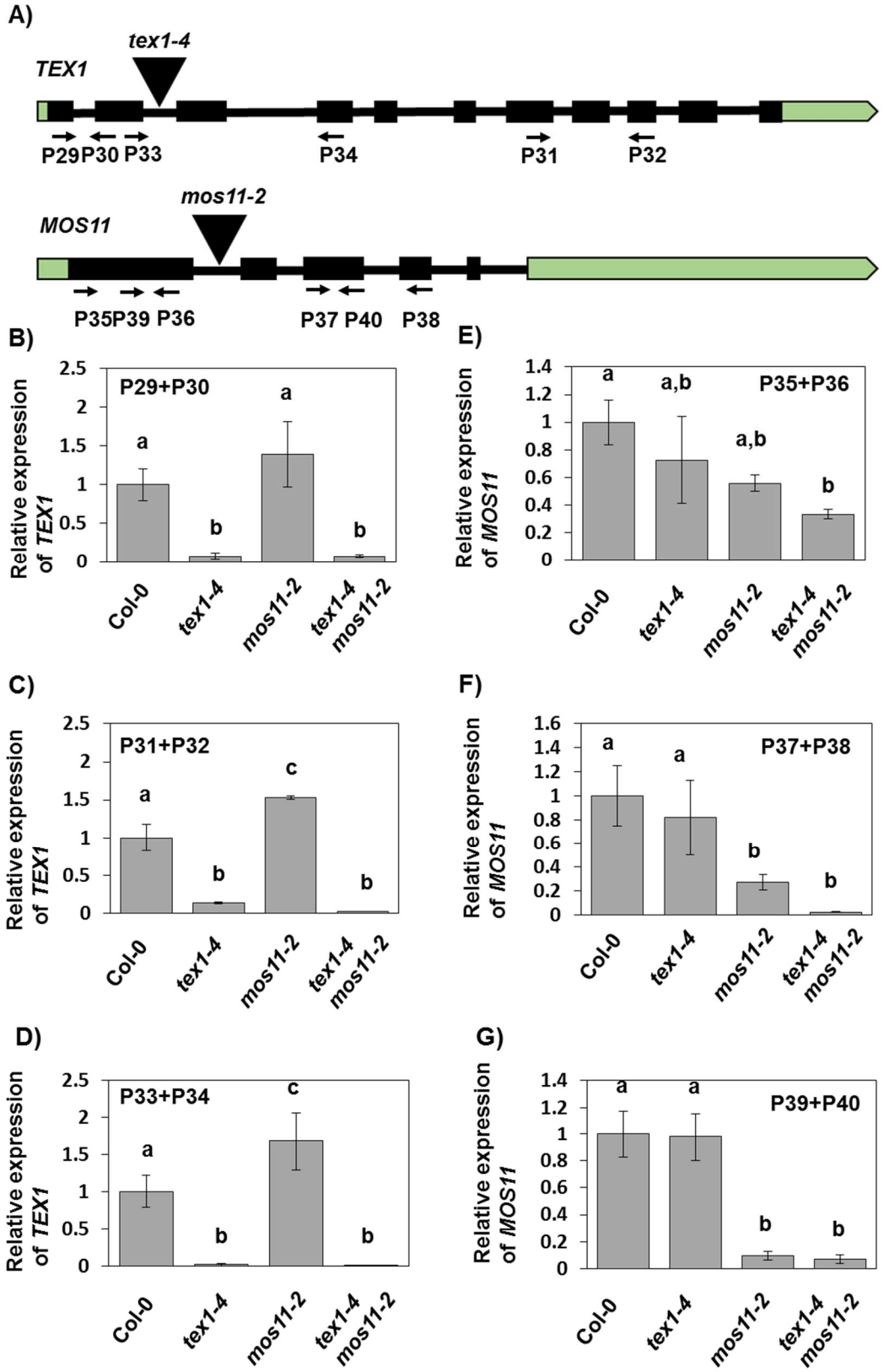

A)
tex1-4
TEX1
P29P30 P33
P34
P31
P32
mos11-2
MOS11
P35P39 P36
P37 P40
P38
B)
P29+P30
Relative expression of TEX1
Col-0
tex1-4
mos11-2
tex1-4 mos11-2
C)
P31+P32
D)
P33+P34
E)
P35+P36
Relative expression of MOS11
F)
P37+P38
G)
P39+P40

Figure 37. *TEX1* and *MOS11* transcript levels analysed by RT-qPCR in the single and double mutants of *tex1-4* and *mos11-2*. A) Schematic representation of *TEX1* (At5g56130) and *MOS11 (*At5g02770) with the T-DNA insertion and primers shown. B-D) Quantification of *TEX1* RNA levels upstream (B), downstream (C) and spanning the T-DNA insertion (D). E-G) Quantification of *TEX1* RNA levels upstream (E), downstream (F) and spanning the T-DNA insertion (G). The average relative expression of the respective genes are depicted in the bar diagram (n=3). The error bars depicts the standard deviation. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 26-Table 31). RNA was extracted from Seedlings 10DAS, grown on MS plates. ACTIN2 (AT3G18780) and EF1α (AT5G60390) were used for normalization.

null hypothesis that there are no interaction, was accepted (P=0.81). The other two null hypotheses, that there are no differences in the means of relative MOS11 mRNA expression levels among the four genotypes in regards to the two factors TEX1 and MOS11, were rejected. When looking for differences in the sample means utilizing the Tukey's test, there was only seen a significant difference between the amount of *MOS11* transcript in wild type and double mutant (Figure 37E). This means the 5' transcript levels of *MOS11* in *tex1-4* and *mos11-2* are not significantly different from wildtype, whereas it is in *tex1-4 mos11-2.* For P37 and P38, the primers binding downstream of the T-DNA, the result are shown in Figure 37F. In the two-way ANOVA analysis, only the factor MOS11 affected the mRNA levels significantly ($P=1.83 \cdot 10^{-4}$). This means that the transcript level of *MOS11* was downregulated to the same levels in *mos11-2* and *tex1-4 mos11-2*. The *MOS11* level in *tex1-4* was the same as in wildtype. The same result was seen for the amplicon spanning the T-DNA insertion (Figure 37G). The *MOS11* transcript was only significantly reduced in *mos11-2* and *tex1-4 mos11-2*. Taken together this points towards that the T-DNA line *mos11-2* has a 5' transcript present at wildtype levels that could lead to a truncated protein containing the first exon. Interestingly this 5' transcript seems to be significantly reduced in the double mutant. As the transcript levels across and downstream of the T-DNA insertion were almost at background levels, it is concluded that the insertion of the T-DNA in *mos11-2* leads to a reduction of full-length *MOS11* transcript.

3.3.3 TEX1 and MOS11 exhibit genetic interaction in multiple phenotypes.

Previously *mos11-2* has been described as having no visible phenotypes (Germain et al., 2010). In order to evaluate if there is genetic interaction between MOS11 and TEX1, phenotypical analyses were done. The phenotypes observed already in *tex1-4* were first examined in order to see the effect of losing both the *TEX1* and *MOS11*

transcripts. Seeds were sown out on soil, and observed at different stages under LD conditions (Figure 38A). Abnormal phenotypes were observed at 24 and 60 DAS for *tex1-4* and *tex1-4 mos11-2*, but not for *mos11-2*. At 24DAS the flowering time transition appeared to be affected in all mutants. The number of leaves at bolting was quantified (Figure 38B). Based on the statistical analysis (Chapter 7.2 (Table 32)) the interaction term is not significant (P=0.26). Both factors TEX1 and MOS11 have been found to contribute significantly to the earlier bolting ($P=2\cdot10^{-16}$ and $P=4.47\cdot10^{-6}$, respectively). This indicates that *TEX1* and *MOS11* interacts genetically in an additive manner on this phenotype. The Tukey's test shows that wild type, *tex1-4*, *mos11-2* and *tex1-4 mos11-2* were all significantly different from one another. This point towards that *tex1-4* and *mos11-2* affect flowering time independently. In order to see how the floral repressor *FLC* is affected in the double mutant, RNA was extracted from seedlings 10DAS, and transcribed into cDNA. *FLC* levels was detected by RT-qPCR (Figure 38C). The null hypothesis that there was no interaction was rejected ($P=4.81\cdot10^{-5}$) (Chapter7.2 (Table 33)). The Tukey's test showed that the mean of relative *FLC* expression levels in *tex1-4*, *mos11-2* and *tex1-4 mos11-2* were significantly different from wild type. In *tex1-4* and *tex1-4 mos11-2,* the *FLC* levels were almost reduced to background levels. In *mos11-2,* there was significantly more *FLC* transcript present compared to the wildtype. Interestingly, *mos11-2* flowers early, even though there was more transcript present of the flower repressor *FLC*. This indicates that TEX1 and MOS11 affect transcription of *FLC* in different ways. TEX1 seems to be dominant over MOS11 in regulating *FLC* expression, as the double mutant has comparable *FLC* levels to *tex1-4*.

As *tex1-4* showed an increased number of activated axillary meristems, leading to more rosetta branches, it was checked how this phenotype was affected in the double mutant. At 60 DAS, *tex1-4* exhibited a bushier appearance, whereas *tex1-4 mos11-2* looked less bushy (Figure 38A). By visual inspection it looked like that, the wildtype and *mos11-2* had the same amount of inflorescences, whereas *tex1-4* and *tex1-4 mos11-2* had more. To quantify this all primary inflorescences were removed from the plants and counted at 60DAS (Figure 38D, Figure 38E). The two-way ANOVA results showed that there was interaction ($P=6.3\cdot10^{-4}$) (Chapter 7.2 (Table 34)). The Tukey's test showed that the number of rosetta branches in wild type and *mos11-2* were not significantly different (Figure 38E). The sample means of *tex1-4*

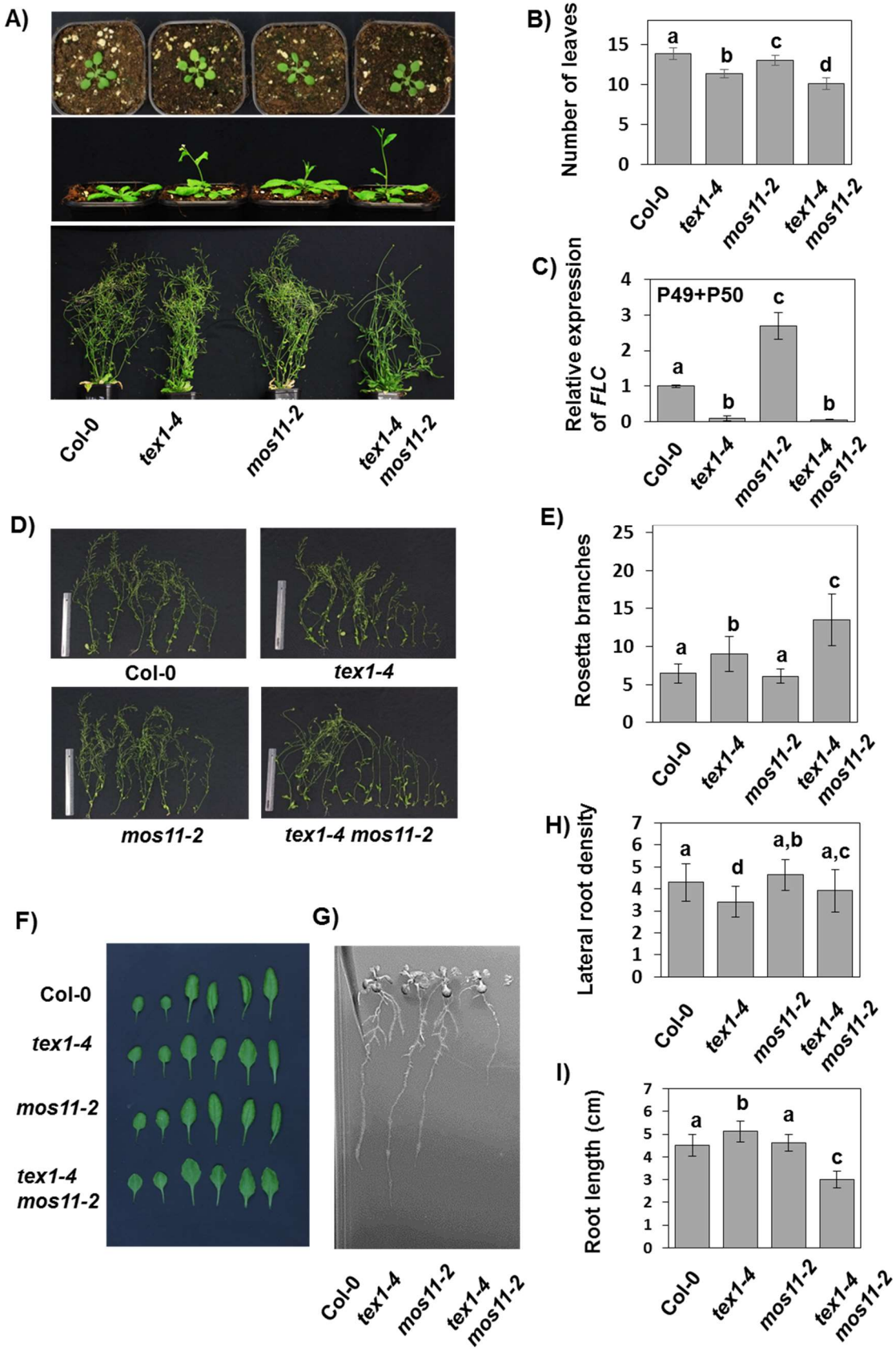

A)
Col-0
tex1-4
mos11-2
tex1-4 mos11-2
B)
Number of leaves
15
10
5
0
a
b
c
d
Col-0
tex1-4
mos11-2
tex1-4 mos11-2
C)
Relative expression of FLC
P49+P50
4
3
2
1
0
a
b
c
b
Col-0
tex1-4
mos11-2
tex1-4 mos11-2
D)
Col-0
tex1-4
mos11-2
tex1-4 mos11-2
E)
Rosetta branches
25
20
15
10
5
0
a
b
a
c
Col-0
tex1-4
mos11-2
tex1-4 mos11-2
F)
Col-0
tex1-4
mos11-2
tex1-4 mos11-2
G)
Col-0
tex1-4
mos11-2
tex1-4 mos11-2
H)
Lateral root density
7
6
5
4
3
2
1
0
a
d
a,b
a,c
Col-0
tex1-4
mos11-2
tex1-4 mos11-2
I)
Root length (cm)
7
6
5
4
3
2
1
0
a
b
a
c
Col-0
tex1-4
mos11-2
tex1-4 mos11-2

Figure 38. The double mutant tex1-4 mos11-2 shows strong phenotypes. Flowering time. The upper panel shows wildtype, *tex1-4*, *mos11-2* and *tex1-4 mos11-2* at 15 DAS. The middle panel shows all the genotypes at 24 DAS, and the lower panel at 60 DAS. B) Quantification of the time of bolting. The average numbers of leaves at bolting for the different genotypes are depicted on the bar diagram (n=11). The error bars depicts the standard deviation. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 32). C) RT-qPCR analysis measuring the FLC levels in wildtype, *tex1-4*, *mos11-2* and *tex1-4 mos11-2*. The average relative expression of the respective genes are depicted in the bar diagram (n=3). The error bars depicts the standard deviation. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 33). RNA was extracted from seedlings 10DAS, grown on MS plates. *ACTIN2* (AT3G18780) and *EF1α* (AT5G60390) were used for normalization. D) Showing all rosetta branches removed from a plant of all four genotypes grown under LD conditions at 60 DAS. The ruler is 20 cm long. E) Quantification of number of branches. The average numbers of rosetta branches per plant for the different genotypes at 60DAS are depicted on the bar diagram (n=11). The error bars depicts the standard deviation. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 34). F) Leaf morphology at 35 DAS. The leaves selected are true leaf number 3, 4, 5, 6, 7 and 8 of the four genotypes. G) Root morphology of the four genotypes at 12 DAS. H) Quantification of Lateral root branching density. The means (Lateral roots/cm) at 10DAS are depicted in the bar diagram (n=37). The error bars depicts the standard deviation. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 35). I) Quantification of root length in the four genotypes. The average root length is depicted in the bar diagram (n=37). The error bars depicts the standard deviation. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 36).

and *tex1-4 mos11-2* were significantly different from all other means and each other. This means that reducing the transcript of *TEX1* leads to misregulation of the number of elongated rosetta branches. Reducing both *TEX1* and *MOS11* transcripts at once, increases the phenotype in a synergistic manner. The leaf morphology was also affected in a synergistic manner in *tex1-4 mos11-2* (Figure 38F). The leaves of *mos11-2* looked like wildtype leaves. For *tex1-4 mos11-2* the leaves were flat like seen for *tex1-4*, but the shape was different and less round.

Another phenotype found in *tex1-4* was a decreased lateral root density (Figure 23L and M). The experiment was repeated with the double mutant (Figure 38G). The quantification of Lateral roots in the branch zone is shown in Figure 38H. The statistical analysis shows that there was no significant interaction term (P=0.19). Both factors contribute significantly to the phenotype in an additive manner (Chapter 7.2 (Table 35)). The Tukey's test showed that only the mean of *tex1-4* is significantly different from the wildtype (Figure 38H). The two mutants *mos11-2* and *tex1-4 mos11-2* were also significantly different from one another. This indicates that TEX1 and MOS11 affect the number of elongated lateral roots in an independent and additive

manner. In addition, the effect of the two mutant seemed to be of opposite direction. As mos11-2 is not significantly different from wildtype it indicates that the main effect comes from TEX1. In the double mutant the length of the main root seemed affected (Figure 38G). The total root length was measured and quantified (Figure 38I). According to the statistical analysis the interaction term was significant (Chapter 7.2 (Table 36)). The Tukey's test shows that the sample means of wildtype and *mos11-2* were not significantly different (Figure 38I). For *tex1-4* the roots were significantly longer than what was observed for the three other genotypes. The double mutant had roots significantly shorter than what is observed for all the other genotypes. TEX1 and MOS11 seems to interact in the regulation of total root length.
When all the above data is taken together it is seen that for all the described phenotypes affected in *tex1-4* there is an increased effect when the *MOS11* transcript is simultaneously downregulated.

3.3.4 TEX1 affects tasiRNA biogenesis independently of MOS11.

As mentioned earlier TEX1 has been described to be involved in tasiRNA biogenesis. When TEX1 is absent *TAS1*, *TAS2* and *TAS3* transcripts are being downregulated (Jauvion et al., 2010; Yelina et al., 2010). As TAS3 is regulating the formation of lateral root primordias, it was speculated that TAS3 might be involved in the lateral root phenotypes observed in *tex1-4*. A northern blot analysis was done to test this. Signals were detected with a phosphor imager screen. The results of the northern blot analysis can be seen in Figure 39A. Probes against TAS1, TAS3, miR173, miR390 and U6 were subsequently hybridized to the blots. The two miR173 and miRNA390, mediate cleavage of respectively the TAS1 and TAS3 precursors, and are required for tasiRNA processing (Felippes & Weigel, 2009). U6 was used as loading control. The signals were quantified by making a line scan of each lane and then calculate the area under the curve of each signal. The data are presented as ratios of the specific probe signal normalized to U6. To account for inter-gel variability a known amount of radioactive probe had been spiked to each gel and each probe was normalized to this signal additionally. The signals were quantified for the *TAS1* specific probe (Figure 39B). According to the statistical analysis (chapter 7.2 (Table 37)) only the factor TEX1 affected the transcript level of TAS1 significantly ($2.13 \cdot 10^{-6}$). The same was seen for TAS3 (Figure 39C). For the *TAS3* specific probe

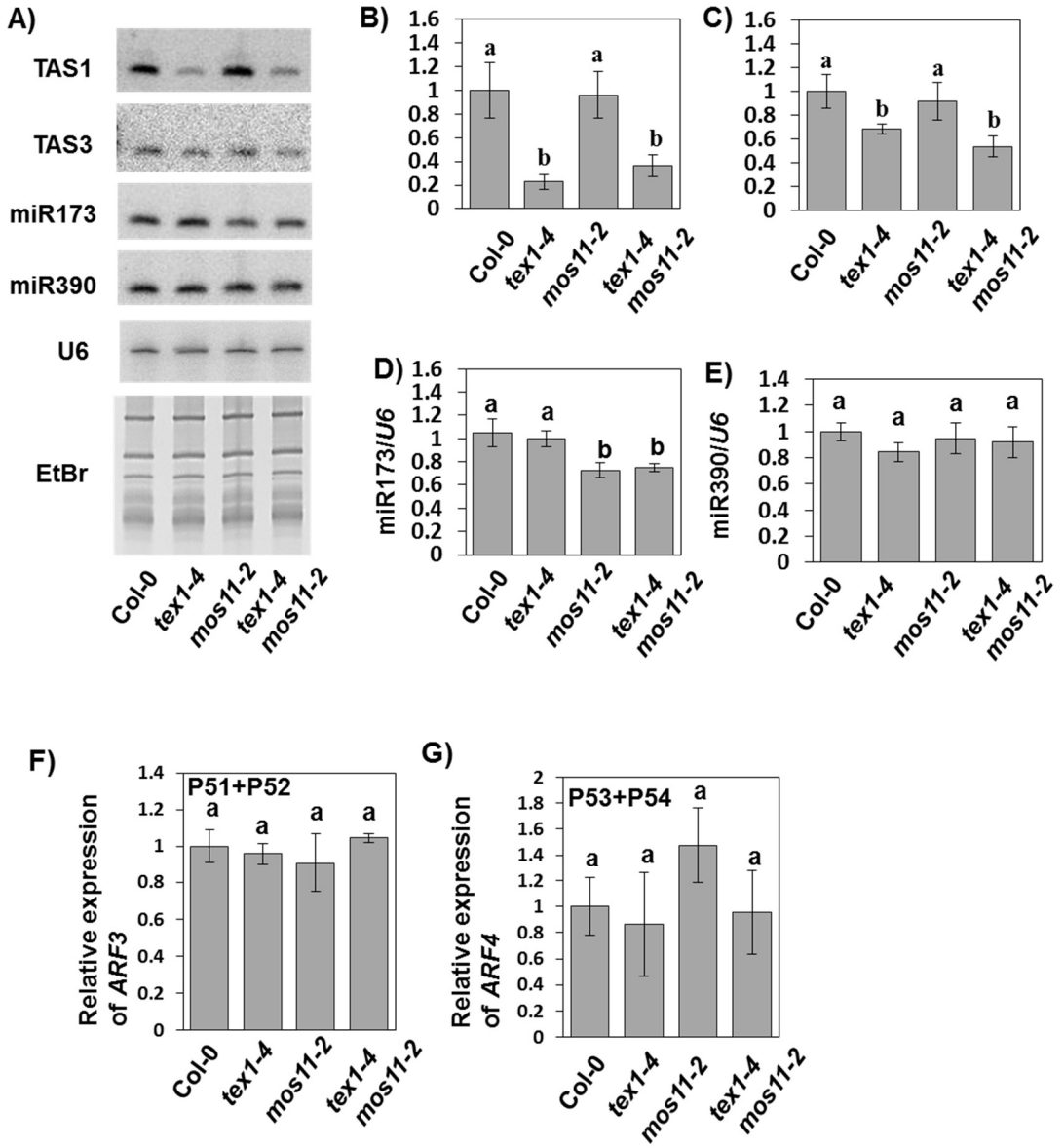


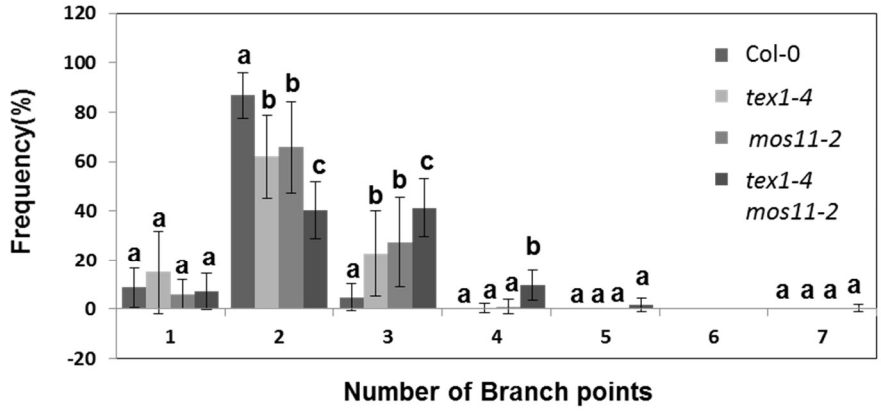
Figure 39. Expression of *tasiRNA* is only affected in *tex1-4*,. A) Northern blot analysis on target RNAs in wildtype, *tex1-4*, *mos11-2* and *tex1-4 mos11-2*. Probes against transcripts of *TAS1*, *TAS2*, *mir173* and *mir390* were radioactively end-labelled with [32P] γ-ATP. The transcript of *U6* was used as a loading control. The EtBr stained polyacrylamide gel with the total RNA extracts are also shown as a loading control. B-E) Quantifications of the expression levels of *TAS1*, *TAS3*, *mir173* and *mir390* transcripts in the four genotypes. Each target gene signal was quantified by calculating the ratio of the area under the curve of a line scan normalized to the area under the curve of the *U6* transcript. To account for inter-gel variations the ratios were normalized to the area under the curve of a known amount of loaded radioactively labelled marker. The means of the normalized target gene/U6 ratios are depicted on the bar diagram (n=4). The error bars depicts the standard deviation. The statistical two-way ANOVA analysis, together with means and standard deviations, can be found in Chapter 7.2 (Table 37-Table 40). F + G) qRT-pcr analysis of *ARF3* and *ARF4* expression levels in *tex1-4*, *mos11-2* and *tex1-4 mos11-2* mutants. The average relative expression of the *ARF3* and *ARF4* transcripts are depicted in the bar diagram (n=3). The error bars describes the standard deviation. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 41 and Table 42).

only TEX1 (chapter 7.2 (Table 38)) affects the expression levels significantly ($P=6.78\cdot10^{-5}$). This indicates that MOS11 is not required for expression of *TAS1* and *TAS3* at all. The miR173, which is required for proper processing of TAS1, was downregulated in *mos11-2* and *tex1-4 mos11-2* (Figure 39D). According to the two-way ANOVA analysis only the factor MOS11 affected the *miR173* (Chapter **7.2** (Table 39)) expression levels significantly ($P=6.81\cdot10^{-6}$). Interestingly the decrease of *miR173* in *mos11-2*, did not lead to a decrease in processed TAS1 (Figure 39B). The miR390 that is involved in processing of TAS3 was not significantly affected in any genotype (Figure 39D and chapter 7.2 (Table 40)). This all indicates that TEX1 and MOS11 affect selected tasiRNA and miRNA transcripts in a specific and independent way. Only TEX1 affects the tasiRNAs transcripts, whereas only MOS11 has an effect on the miR173 expression. In *tex1-4,* TAS3 is downregulated and the mutant has less elongated lateral root primordia pr. unit length. ARF2, ARF3 and ARF4 are repressors of Lateral elongation, being themselves repressed by *TAS3*. It was tested by RT-qPCR if the *ARF3* and *ARF4* transcripts were upregulated, due to less TAS3 transcript in *tex1-4* (Figure 39F and G). For both *ARF3* and *ARF4* transcripts there were no significant difference in relative transcription among the genotypes compared to the wild type (Chapter 7.2 (Table 41 and Table 42)). For the *ARF4* transcript, the non-significant increase in *mos11-2* is surprising. In *mos11-2* the lateral root density is not decreased compared to wildtype, which one would expect with more ARF4 (Figure 39H). These data indicate that the observed effect of *tex1-4* on lateral root density is ARF3/ARF4 independent.

3.3.5 Simultaneous loss of the *TEX1* and *MOS11* transcripts leads to severe phenotypes of trichomes and reproductive organs.

Double mutant plants were also investigated for phenotypes not already observed in *tex1-4*. When investigating the leaves of the double mutant, trichomes with an abnormal amount of branch points were observed in *mos11-2 tex1*-4. In order to investigate if this was a general trend, the branch point distribution among the trichome population was quantified (Figure 40A and Table 17). This was done in collaboration with the lab of Dr. Michael Melzer (IPK Gatersleben). One leaf from 15 independent plants of each genotype were analysed. For each leaf the number of branch points for each trichome was counted. Trichomes were observed to have between two and

A)


Frequency(%)
120
100
80
60
40
20
0
-20
Col-0
tex1-4
mos11-2
tex1-4 mos11-2
1
2
3
4
5
6
7
Number of Branch points

B)

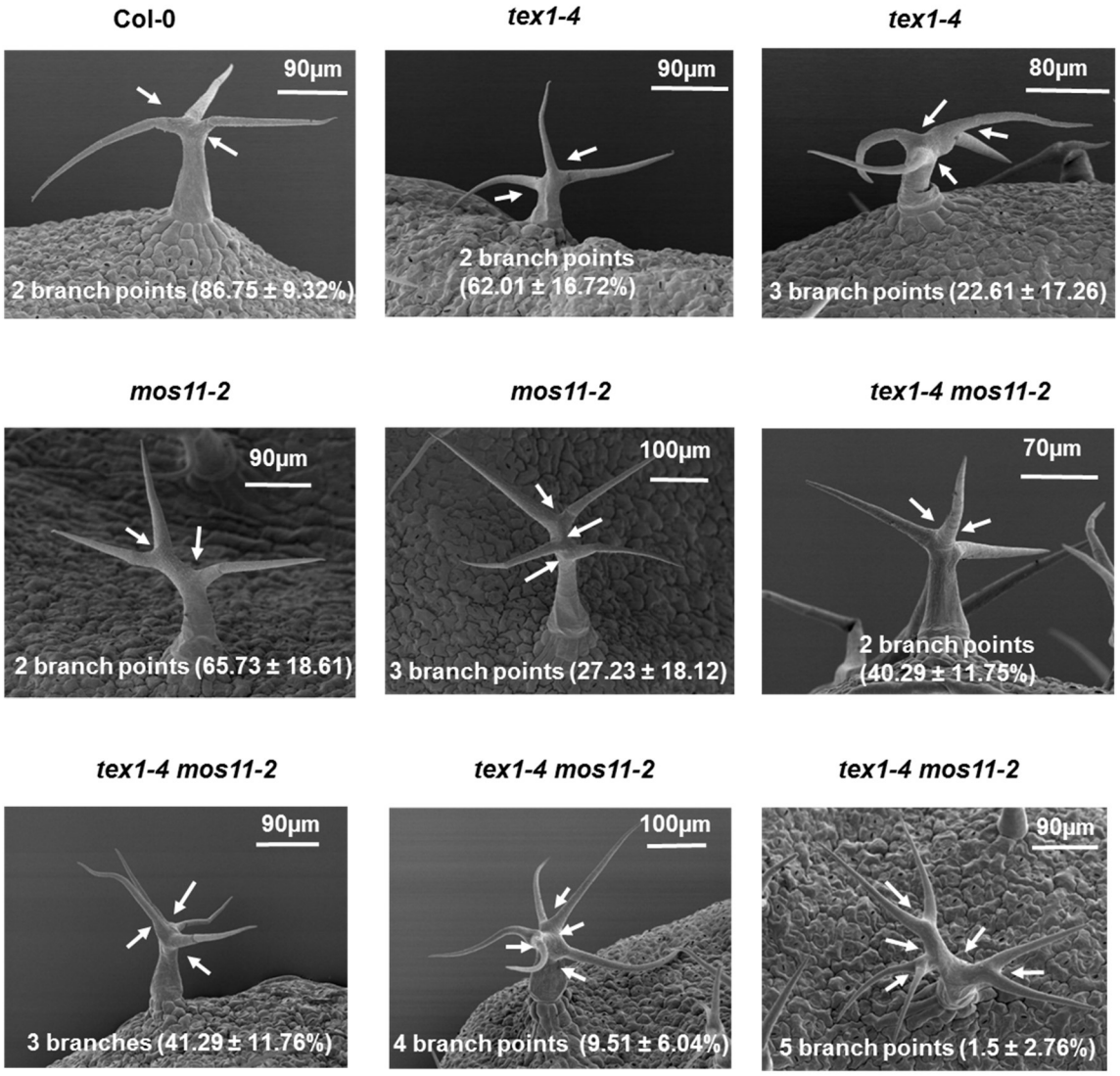

Col-0
90µm
2 branch points (86.75 ± 9.32%)
tex1-4
90µm
2 branch points
(62.01 ± 16.72%)
tex1-4
80µm
3 branch points (22.61 ± 17.26)
mos11-2
90µm
2 branch points (65.73 ± 18.61)
mos11-2
100µm
3 branch points (27.23 ± 18.12)
tex1-4 mos11-2
70µm
2 branch points
(40.29 ± 11.75%)
tex1-4 mos11-2
90µm
3 branches (41.29 ± 11.76%)
tex1-4 mos11-2
100µm
4 branch points (9.51 ± 6.04%)
tex1-4 mos11-2
90µm
5 branch points (1.5 ± 2.76%)

Figure 40. Single and double mutants of TEX1 and MOS11 show an increased number of branchpoints per trichome. A) Quantification of the number of branch points per trichome at 13 DAS. For each genotype one leaf from 15 independent plants were investigated. For each leaf, it was counted how many occurrences there was of trichomes with different amounts of branch points. The mean frequency (percentage) with which trichomes with a certain amount of branch points (2-7) are occurring in each genotype (n=15) are shown in the bar diagram. The error bars depicts the standard deviation. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 43-Table 48). B) EM-pictures of trichomes with different number of branch points in the four genotypes. The percentages presented on the picture are the frequencies from Figure 40A and Error! Reference source not found.**. The arrows shows the position of the branch points. EM pictures and the counting of branch points were done in collaboration with the lab of Dr. Michael Melzer (IPK Gatersleben).**

seven branch points in this study (Table 17), with two branch points being the most prevalent state in wildtype (~87%). For each number of branch points it was analysed statistically if there is a shift in the population between the four genotypes (Chapter 7.2 (Table 43-Table 48)). There were only significant differences in the distribution of trichomes with two, three and four branch points. In the case of trichomes with two branch points, the two-way ANOVA analysis showed that TEX1 and MOS11 affected each other in an additive and independent manner. The Tukey's test showed that *tex1-4* and *mos11-2* were not significantly different from each other, but different from wildtype and the double mutant (Figure 40A). So in *tex1-4* and *mos11-2* there were less trichomes with two branch points than compared to wildtype. In the double mutant there was even less trichomes with two branch points. For the trichomes with three branch points, the distributions had been reversed. There were more trichomes with three branch points in *tex1-4*, *mos11-2* and *tex1-4 mos11-2* than in the wildtype (Figure 40A). The statistical analysis showed that the increase of branch points in *tex1-4 mos11-2* was significantly different from the single mutants and happens in an additive manner. For trichomes with four branch points only the double mutant had significantly more trichomes than what was observed in the wildtype. The null hypothesis that there is no interaction was rejected ($P=3.79 \cdot 10^{-5}$). This means that reducing *TEX1* and *MOS11* transcripts leads to a synergistic increase in trichomes with four branch points. Taken together this indicates that for

Table 17. Trichome branches of wild type, *tex1-4*, *mos11-2* and *tex1-4 mos11-2* plants. The numbers are given as percentages of total branch points, with standard deviation. * signifies a statistic difference in means compared to wildtype (Chapter 7.2 (Table 43-Table 48)).

Branchpoints:	1	2	3	4	5	6	7
wt	8.60 ± 8.07	86.75 ± 9.32	4.65 ± 5.49	0	0	0	0
tex1-4	14.90 ± 16.71	62.01 ± 16.72*	22.61 ± 17.26*	0.48 ± 1.84	0	0	0
mos11-2	6.06 ± 5.84	65.73 ± 18.61*	27.23 ± 18.12*	0.98 ± 2.95	0	0	0
tex1-4 mos11-2	7.00 ± 7.33	40.29 ± 11.75*	41.29 ± 11.76*	9.51 ± 6.04*	1.5 ± 2.76	0	0.39±1.52

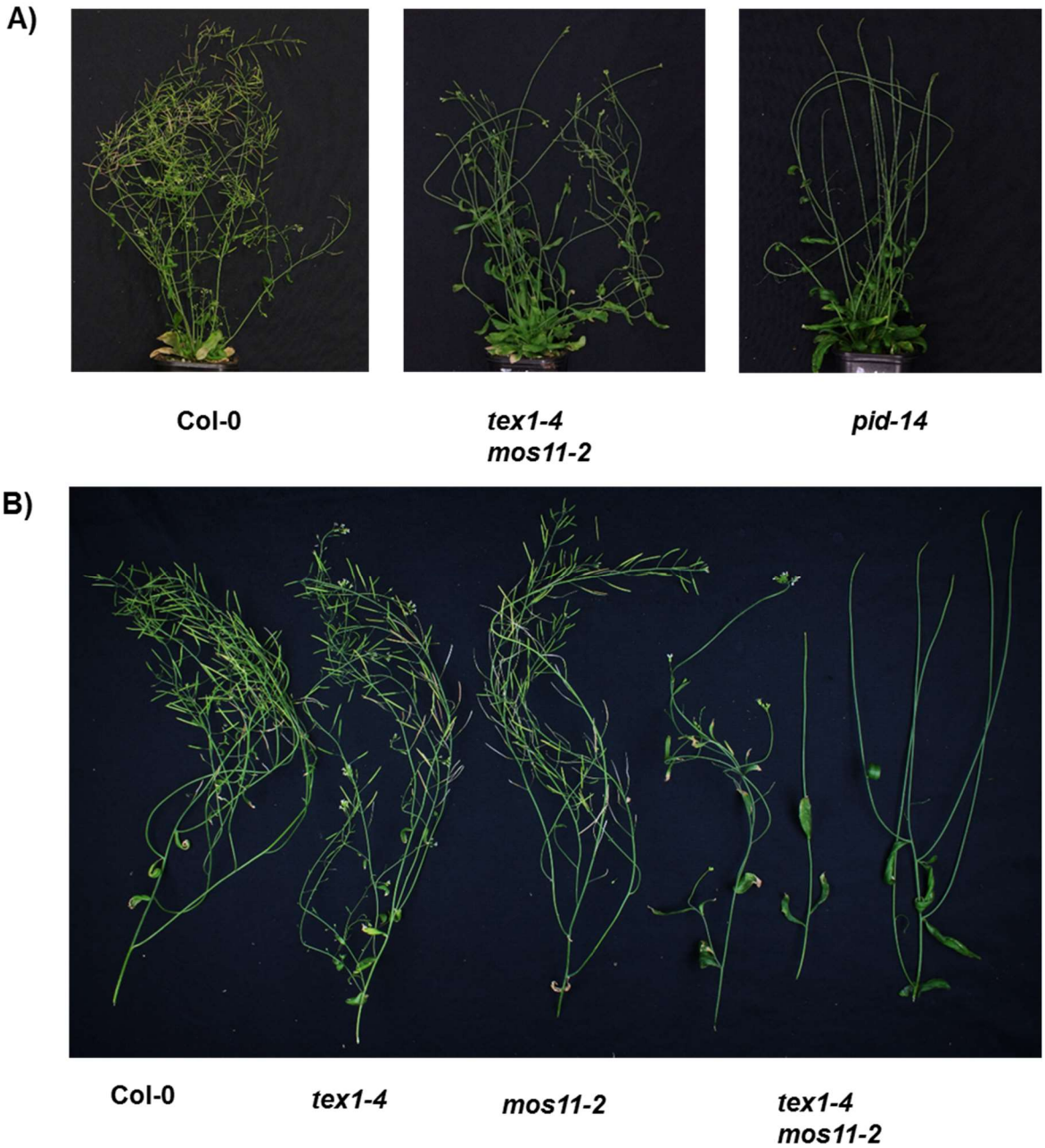


Figure 41. The double mutant *tex1-4 mos11-2* exhibit altered inflorescence morphology . A) Wild type and tex1-4 mos11-2 plants at 60DAS. B) Rosetta branches at 60 DAS. One primary inflorescence is shown for wild type, *tex1-4* and *mos11-2*, and three are shown for *tex1-4 mos11-2*.

tex1-4 and *mos11-2* the regulation of branch density is affected. For both single mutants the percentile decrease in trichomes with two branch points correlates with the percentile increase of branch points with three branch points. For the double mutant a further significant shift is seen towards trichomes with four branch points that was not detected in the single mutants. This indicates that the division of branches is even more misregulated in the double mutant than in the single mutants. Representative examples of the trichomes with a divergent number of branch points of

each genotype can be seen in Figure 40B.

When comparing the general morphology of the four genotypes, differences were observed in the morphology of the primary inflorescences (Figure 38A). Comparing the wild type and the double mutant more in detail revealed that the inflorescences of *tex1-4 mos11-2* looked more barren, with less siliques (Figure 41A). This is seen more clearly when removing the primary inflorescences and rosetta branches from the plants and comparing the four genotypes (Figure 41B). Wildtype, *tex1-4* and *mos11-2* inflorescences looked very alike. The double mutant not only had less siliques but the flower buds also looked affected. Instead of having flowers the primary inflorescence had a pointed tip, reminiscent of mutants deficient in PINOID and PIN (Okada et al., 1991; Bennett et al., 1995). In order to investigate this the flower bud of the first primary inflorescence was analysed in further detail under a dissecting microscope at 30DAS (Figure 42A). Some of the flower primordia from the flower meristem did not develop into flowers in *tex1-4 mos11-2.* It looked like the flower primordia are initiated though. To see this in more detail EM microscopy was done on the inflorescence meristem at 44 DAS in collaboration with Dr. Michael Melzer, IPK Gatersleben(Figure 42B). The Wildtype, *tex1-4* and *mos11-2* inflorescence meristems looks alike. The mutant *pid-14* was used as a control that does not initiate flower primordia. It was seen that the *pid-14* inflorescence tip is completely smooth. For the double mutant the ridged structures at the inflorescence tip indicates that the primordia might be initiated, but then stop in development. In some cases, the flower primordia stayed undeveloped, and in other cases the growth of the primordia was reinitiated (Figure 42B and C).The re-initiation of the primordia on the meristems seemed uncontrolled and lead to a wide variety of misregulated flower bud morphologies at later stages (Figure 42C). This indicates that the loss of both *MOS11* and *TEX1* transcripts negatively affects the onset of flower development from the primordia. As the general flower bud morphology in the double mutant was affected, it was decided also to check the development of the flower organs. For *tex1-4 mos11-2* flowers were selected that were not halted somewhere during the flower development process. The overall flower morphology was alike in the four genotypes (Figure 43A). When isolating the stamens and pistils it was noted thatthere were morphological differences between the double mutant and the other genotypes. The pistil of the double mutant was deform in shape. The stamen filament of *tex1-4 mos11-2*

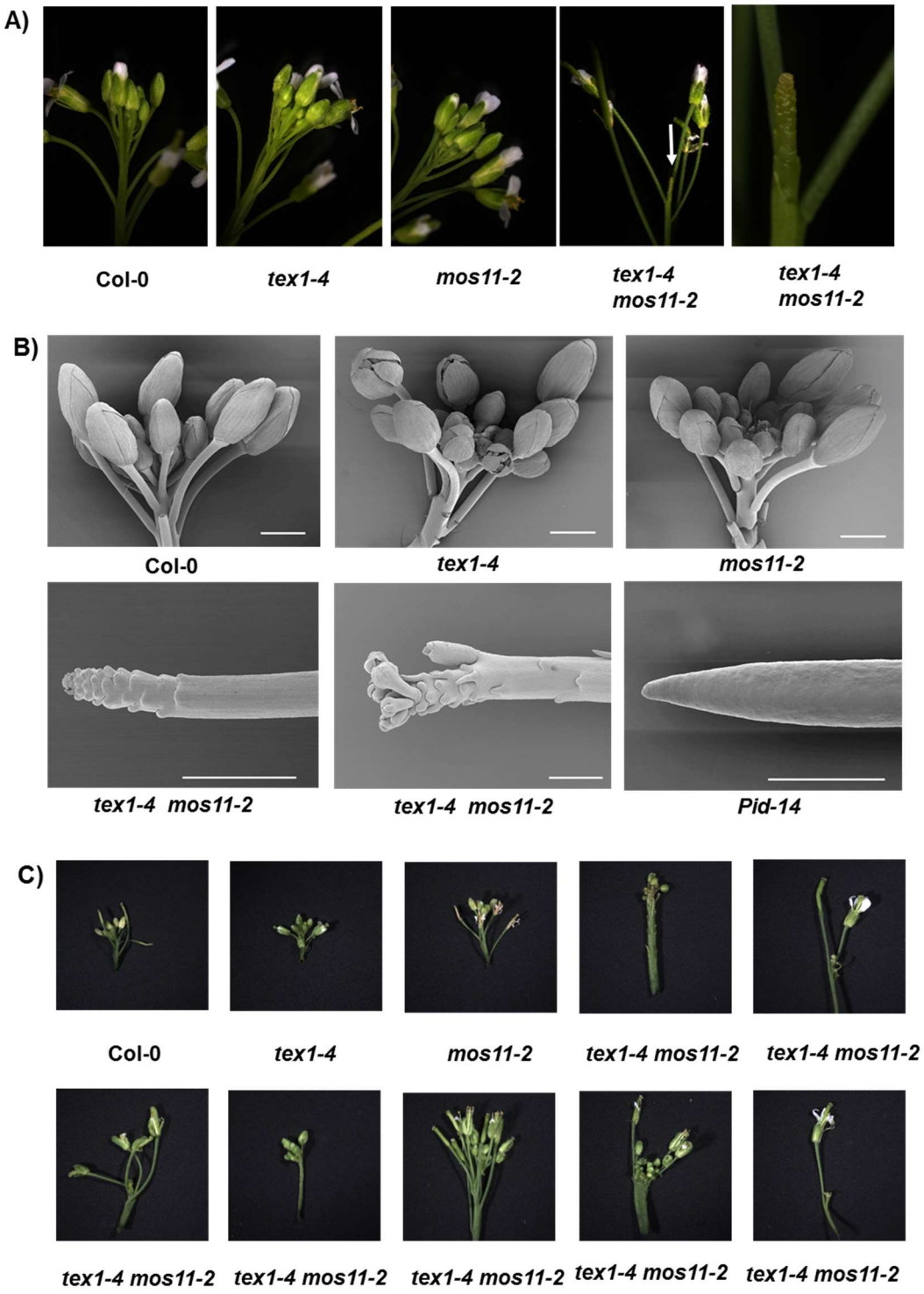


Figure 42. The double mutant *tex1-4 mos11-2* exhibits flower meristem phenotypes. A) The first primary inflorescence meristem of the wildtype*, tex1-4, mos11-2* and *tex1-4 mos11-2*. The right most panel is a magnification of the meristem shown in the second right most panel. Pictures are taken at 30 DAS. B) Electron microscopy pictures of primary inflorescence meristems at 44 DAS. The T-DNA line pid-14 (SALK_049736) was included as a control, for a mutant showing lack of initiation of flower primordia. The scale bar depicts 1mm. The Electron microscopy pictures were done in collaboration with the lab of Dr. Michael Melzer (IPK Gatersleben). C) Pictures of primary inflorescences in the four genotypes at 60 DAS. The double mutant *tex1-4 mos11-2* exhibit a variety of different inflorescence morphologies of which some are shown.

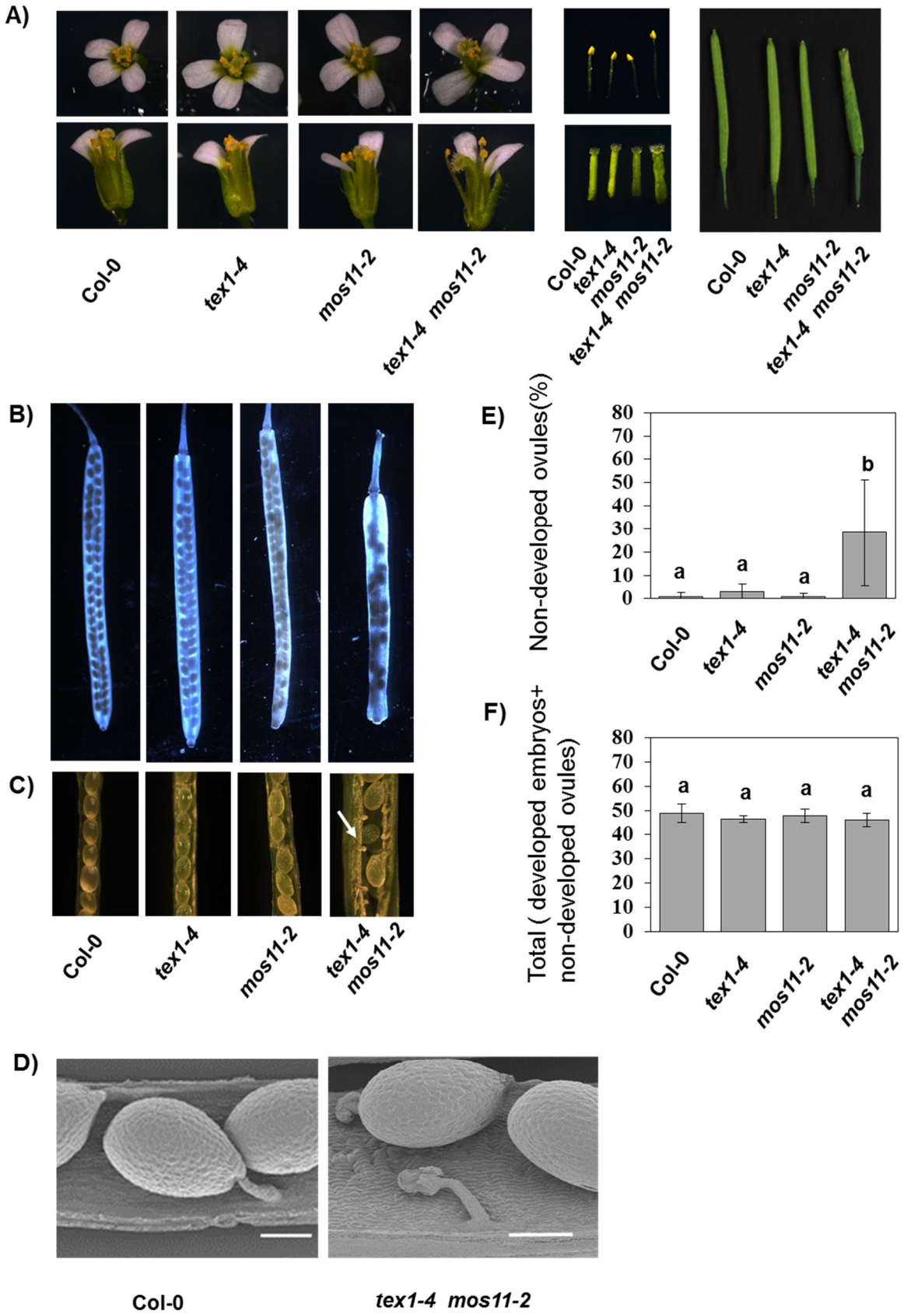

A)
Col-0
tex1-4
mos11-2
tex1-4 mos11-2
B)
C)
D)
Col-0
tex1-4 mos11-2
E)
Non-developed ovules(%)
80
70
60
50
40
30
20
10
0
a
a
a
b
F)
Total (developed embryos+ non-developed ovules)
a
a
a
a

Figure 43. Morphology of reproductive organs are affected in *tex1-4 mos11-2*. Flower morphology at 44DAS. A) In the left side panels flowers of the four genotypes seen from above, and from the side with one petal and two sepals removed. In the middle-right panels, isolated pistils and stamens are shown. In the right-hand panel, siliques are shown. B) Siliques that has been cleared to make the seed set visible. C) Siliques that has been opened to show the seeds. Non-developed ovules are pointed out with an arrow D) Electron microscopy picture of one of the non-developed ovules in *tex1-4 mos11-2*. The scale bar depicts 200 µm. The Electron microscopy pictures were done in collaboration with the lab of Dr. Michael Melzer (IPK Gatersleben). E) Quantification of non-developed ovules in siliques of the four genotypes. Flowers were marked upon opening, and the amount of non-developed ovules were counted 12 days later. Ovules not properly developed was counted as defective. The average number of non-developed ovules are depicted as a percentage of all seed events in the bar diagram (n=250). The error bars describes the standard deviation. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 49). F) The average number of total non-developed ovules and developed embryos pr. silique are depicted in the bar diagram (n=5). The error bars describes the standard deviation. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 50).

seemed to be longer than for the other genotypes. The siliques from the double mutant looked shorter, broader and deform compared to the other genotypes. When looking at the bleached siliques empty spaces were observed where there should be seeds in the double mutant (Figure 43B). When opening the siliques, non-developed ovules were observed in *tex1-4 mos11-2* (Figure 43C). The morphologies of the non-developed ovules were observed with EM microscopy (Figure 43D). The number of non-developed ovules were quantified in the four genotypes (Figure 43E). The two-way ANOVA analysis showed that there was a significant interaction (P=0.03) between TEX1 and MOS11 on the number of non-developed ovules (Chapter 7.2 (Table 49)). The Tukey's test showed that only the double mutant was significant different from the wildtype. Indicating that significantly more ovules are not developing in the double mutant. It was also quantified how many embryos were initiated in total for each silique (Figure 43F). According to the statistical analysis (Chapter 7.2 (Table 50)), there were no effects of TEX1 and MOS11 on the total number of embryos in any of the genotypes. This indicates that the initiation of the ovules is not affected in the double mutant, and that the development is stopped in one of the subsequent processes.

3.3.6 A simultaneous loss of *TEX1* and *MOS11* transcripts leads to deficient mRNA export and reduced protein synthesis rates.

Looking at the observed phenotypes it is clear that TEX1 and MOS11 are important for proper development. In order to narrow down the function of these two proteins

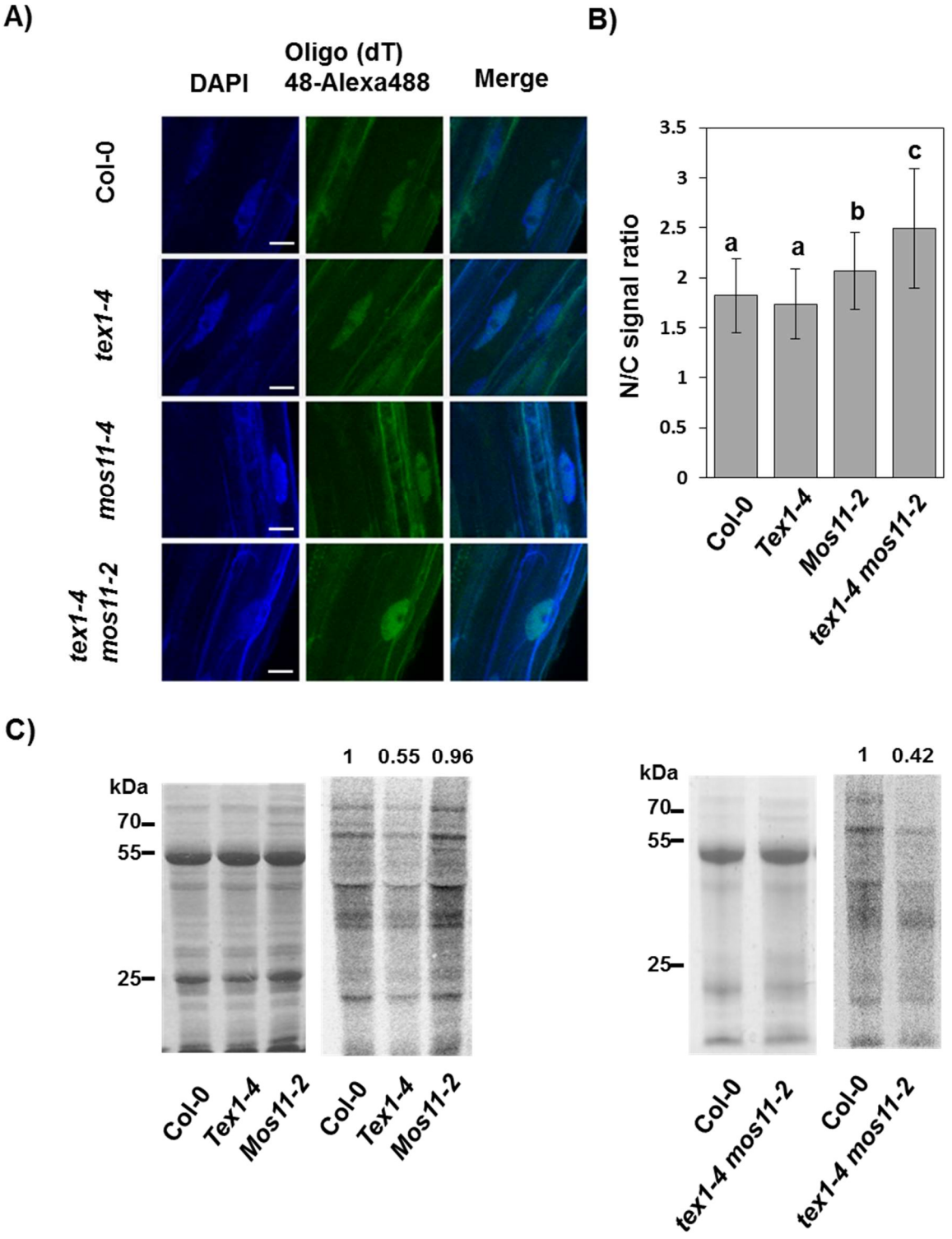

A)
DAPI
Oligo (dT)
48-Alexa488
Merge
Col-0
tex1-4
mos11-4
tex1-4
mos11-2
B)
N/C signal ratio
3.5
3
2.5
2
1.5
1
0.5
0
a
a
b
c
Col-0
Tex1-4
Mos11-2
tex1-4 mos11-2
C)
kDa
70
55
25
1
0.55
0.96
Col-0
Tex1-4
Mos11-2
Col-0
Tex1-4
Mos11-2
kDa
70
55
25
1
0.42
Col-0
tex1-4 mos11-2
Col-0
tex1-4 mos11-2

Figure 44. MOS11 is required for bulk mRNA export, whereas TEX1 is required for efficient protein synthesis. A) Whole mount *in situ* hybridization of 48-mer Oligo (dT) probe fluorescently labelled with Alexa488 to fixated root tissue of the four genotypes. The optical sections were made around the xylem plane using a CSLM microscope with a 40X/1.3 objective. Twelve z-stacks of 0.8µm were transformed with an maximum projection algorithm. DAPI was used as a counterstain. Scale bar indicates 10µm. B) Quantification of bulk mRNA export. The ratio of fluorescence signal in cytoplasm and nucleus was calculated for 60 nuclei of each genotype. The statistical two-way ANOVA analysis, together with means and standard deviations as values, can be found in Chapter 7.2 (Table 51). Seedlings were plate grown for six days under LD conditions. C) Pulse labelling experiments for Wild type, *tex1-4* and *mos11-2*. Protoplasts were enzymatically released and isolated from seedlings of the three genotypes harvested at 14DAS. After 1 hour of incubation with [35s]-Methionine, proteins were extracted and separated on a 9% polyacrylamide gel by SDS-Page (left panel). In the right-side panel, the signal from the Phosphorimager screen is seen. The number above the hosphorimager screen signals are semi quantitative estimations of the amount of synthesized protein. For each genotype a ratio was calculated based on the ratio between the [35s]-methionine signal and the amount of total protein. The [35s]-methionine signal was calculated by taking the area under the curve of each lane of the autoradiography using the ImageJ line scan tool. The amount of total protein was calculated by taking the area under the curve of each lane of the Coomassie stained gel using the ImageJ line scan tool. The ratios were normalized to the wildtype.

first the effect of the double mutant on mRNA export was assessed. The whole mount *in situ* hybridization with fluorescently labelled poly(dT) was done on fixated seedlings at 6DAS (Figure 44A). The ratio of nucleoplasmic to cytoplasmic RNA was quantified (Figure 44B). The two-way ANOVA analysis (Chapter 7.2 (Table 51)) showed that the interaction null hypothesis was rejected ($P=1.41 \cdot 10^{-5}$). Tukey's test showed that the sample means for wildtype and *tex1-4* were not significant different (Figure 44B). The sample mean of *tex1-4 mos11-2* and *mos11-2* were significantly different from *tex1-4*, wildtype and each other. Taken together with the trend seen in Figure 44B, this indicates that the double mutant and *mos11-2* shows increased bulk mRNA export deficiency. The mRNA export deficiency is enhanced in the double mutant in a synergistic manner.

Reducing the *TEX1* transcript on its own, does not lead to blocked mRNA export, but combining it with *mos11-2* leads to more mRNA accumulated in the nucleus. To find out more about the function of TEX1, it was decided to try another approach. In Gatfield et al., 2001 they showed that depletion of UAP56 and NXF1 in *Drosophila* leads to mRNA blockage and decreased protein synthesis rates. In order to test if the protein synthesis rate was affected in *tex1-4* and the other phenotypes, it was decided to pulse label protoplasts with [35S] methionine. The protoplasts were released from two-week-old seedlings of each genotype by enzymatic digest. The isolated protoplasts were incubated 60 min with 100µCi [35S] methionine. The total protein extract from the protoplasts was separated by SDS-PAGE. Incorporation was

detected by exposing a Phosphorimager screen. The analysis was done semi-quantitatively (Figure 44C). Line scans were used to get the area under the curve of the Coomassie stained and autoradiographic signals. The protein synthesis rate was measured as the amount of [35S] methionine signal per amount of total protein in one hour. The synthesis rate was not affected in *mos11-2* compared to wild type. For *tex1-4* and the double mutant there seemed to be less protein synthesized than compared to wild type. The trend was reproducible in three independent experiments. With the semi-quantitative nature of this assay, it is difficult to conclude if the protein synthesis rate is more affected in the double mutant than compared to *tex1-4*. This indicates that MOS11 affects the mRNA transport, whereas TEX1 affects the protein synthesis rates in an undefined manner. When both *TEX1* and *MOS11* are reduced the mRNA export block is enhanced synergistically which might cause abnormal organogenesis.

4 Discussion

The aim of this study has been to gain more information about the existence, composition and function of the *Arabidopsis* TREX complex by analysing TEX1, MOS11, UAP56, ALY1, ALY2, ALY3 and ALY4 with different molecular biological methods.

4.1 The core composition of the TREX complex

In this study the composition of the *Arabidopsis* TREX complex has been expanded from only consisting of the core THO complex (Yelina et al., 2010), to also contain orthologs of TREX interaction partners known from other organisms. TEX1 was shown to co-purify with all THO subunits, UAP56 and ALY3 *in vivo*. UAP56 associates with the THO complex, MOS11, ALY2 and ALY4 *in vivo* and directly binds all four ALY proteins in a yeast-two hybrid analysis. ALY1 was found to be a contaminant under the conditions used for the affinity purification in this study, but showed no auto-activation in the yeast-two hybrid analysis. The *in vivo* affinity purifications have been done with overexpression constructs, in order to compensate for the general higher expression levels due to ploidy level of the *Arabidopsis* cell culture (9n) compared to the diploid wild type plant (Van Leene et al., 2015). Caution must be taken as over expression has been known to disrupt stoichiometry and cause false interactors (Gibson et al., 2013). In this study all interactors of TEX1, MOS11 and UAP56 are found at substoichiometric amounts, which is most likely due to the overexpression. The TEX1 protein is generally associated at stoichiometric amounts with the THO complex in yeast and humans (Chi et al., 2013; Jimeno et al., 2002; Sträßer et al., 2002), though there also have been reports of TEX1 purifying at substoichiometric levels in *Drosophila* and humans (Rehwinkel et al., 2004; Masuda et al., 2005). It is a possibility that there are some false interactors found in this study due to overexpression, but there are arguments talking for that the data are trustworthy. All interactions in between the three bait proteins are confirmed by the “reciprocal tagging” approach, which is recommended as confirmation when working with affinity purification in *Arabidopsis* cell culture (Dedecker et al., 2015). TEX1 copurified all the other THO subunits, like it has been described in Yelina et al., 2010. In that study a HA tagged TEX1 construct under control of native promoter was utilized in seedlings. *Arabidopsis* UAP56 has previously been shown to interact directly with ALY2 and MOS11 *in vitro*, supporting the data found in this study (Kammel et al., 2013). Fur-

thermore other proteins described to interact with THO/TREX in yeast and humans, have been identified (Discussed in Chapter 4.4). All this points towards that the used approach has merit in identifying more putative TREX interactors as a preliminary screen.

In other organisms there have been some contradicting findings in regards to if TEX1 is part of the core THO complex, or an associated member of TREX. TEX1 has been found to not associate with THO2 after immunodepletion at 350mM NaCl in HeLa cells, and to copurify in substoichiometric amount with HPR1 in *Drosophila* (Masuda et al., 2005; Rehwinkel et al., 2004). Contradictory TEX1 has also been found to associate with THO2 after immunodepletion at 350mM NaCl in HeLa cells, to associate with THO at 500mM NaCl in gel filtration fractions (Chi et al., 2013) and to associate stoichiometrically with THO (Peña et al., 2012; Chi et al., 2013; Jimeno et al., 2002; Sträßer et al., 2002). There are more studies supporting that TEX1 is part of the THO core complex. On the contrary ALYREF and UAP56 do not associate stably with THO under high salt conditions (350-500mM) in HeLa cells (Masuda et al., 2005; Chi et al., 2013). Considering the salt sensitivity of the THO and TREX association, the less stringent buffers used in this study are most likely the reason why UAP56, MOS11, ALY2, ALY3 and ALY4 were successfully copurified with TEX1 in this study and not in Yelina et al., 2010. The observed interactions between THO, YRA1/ALYREF and SUB2/UAP56 are well described in both yeast and mammalians (Sträßer et al., 2002; Masuda et al., 2005; Dufu et al., 2010; Viphakone et al., 2012). ALYREF interacts directly with THO2 and THO5 *in vitro*, but it is not known if UAP56 associates directly with THO, or associates through ALYREF in humans (Chi et al., 2013). The association of CIP29 with TREX is not fully understood in humans. Human UAP56 and CIP29, have been described to interact directly in yeast-two hybrid screen (Lehner et al., 2004) and in pull downs using recombinant proteins (Chi et al., 2013). The co-purification of UAP56 and CIP29 has also shown to happen in an ATP-dependent manner in HeLa cells (Dufu et al., 2010). This was confirmed in the same study by in vitro pulldowns with recombinant proteins of UAP56, ALYREF and CIP29. Another study showed that CIP29 almost exclusively binds DDX39 in HeLa cells, whereas UAP56 preferentially interacts with ALYREF and THO (Yamazaki et al., 2010). In Dufu et al., 2010 the addition of ATP has been shown to be crucial for the association of recombinant UAP56, ALY and CIP29. In Yamazaki et al., the

dependency of ATP was not tested. This indicates that CIP29 in humans can be found with TREX in the presence of ATP, but otherwise associate mostly with DDX39 in the AMEX complex. In *Arabidopsis* there are two UAP56 genes, coding for the same protein. The *Arabidopsis* UAP56 protein is the only ortholog of human UAP56/DDX39 found by homology (Gaouar & Germain, 2013), making the data less complicated to interpret in *Arabidopsis* than in humans. UAP56 and MOS11 (CIP29) have been shown to interact directly in *Arabidopsis* in a yeast-two hybrid interaction assay and in *in vitro* pulldown experiments with recombinant proteins (Kammel et al., 2013). The association of CIP29 with THO in HeLa cells is mediated through UAP56 (Dufu et al., 2010). In this study TEX1 and MOS11 both copurify with UAP56 but not with each other, suggesting that MOS11 might also interact through UAP56. No orthologss of CIP29 have been described in the literature for *Drosophila* and *C. elegans*. In yeast it has not been shown that THO1 and SUB2 interact physically. It is known that the THO complex and the nascent RNA is required for recruitment of SUB2 and THO1 (Jimeno et al., 2006). In the same study it was shown that THO1 does not recruit SUB2, and it cannot complement the phenotype of *sub2*. Multicopy THO1 and SUB2 can both repress the phenotypes of THO mutants (Jimeno et al., 2006; Fan et al., 2001), suggesting some overlap in function. This all indicates that the TREX complex seems to be generally evolved around a core setup of the THO complex, ALYREF and UAP56 in eukaryots. CIP29/MOS11 associates with the TREX component UAP56 in mammalians and *Arabidopsis*. Additionally to the core setup of TREX, there has been described a couple of non conserved TREX associated proteins in both humans (LUZP4, PDIP3, ZC11A and CHTOP) and yeast (NAB2) (Viphakone et al., 2015; Chang et al., 2013; Iglesias et al., 2010; Batisse et al., 2009; Folco et al., 2012). None of these have orthologs in *Arabidopsis*, indicating that TREX also can have organism specific subunits.

4.2 Subcellular localization of TREX components in *Arabidopsis thaliana*

In this study the subcellular localization of ALY1, ALY2, ALY3, ALY4, UAP56, MOS11 and TEX1 fused to GFP under native promoters has been described in the root system. According to publically available microarray data (Figure 9) all the subunits are generally distributed throughout all tissues, which corresponds with the localization analysis in this study where all GFP fused proteins were found in nuclei

Table 18. Subcellular localization overview of putative TREX components described in the literature, compared to this study. The abbrevations after the name of the study, describes the settings of the experiment. *N. benthamiana* (Nb), *A. thaliana* (At), Overexpression (Ox), Native promoter (Np), Transient expression (T), Stable transformation (S) and Cell culture (CC). The subcellular localization found in the study is described by other abbreviations. Nucleoplasmic localized (Nu), present in the nucleolus (No), distinct speckle-like localization (DS) and weaker speckle like localization (Ws).

Protein	Uhrig et al., 2004 (Nb, Ox, T)	Pendle et al., 2005 (At, CC, Ox,S)	Pan et al., 2014 Nb, Ox, T)	Germain et al., 2010 (At, Np, S)	This study (At, Np, S)
ALY1	Nu, No, Ds	Nu, No, Ds			Nu, Ds
ALY2	Nu, Ds	Nu, Ds			Nu, Ds
ALY3	Nu, No, Ds	Nu, No, Ds			Nu, No, Ws
ALY4	Nu, No, Ds	Nu, No, Ds			Nu, No, Ds
UAP56-1			Nu, No, Ds		Nu, Ws
UAP56-2		Nu, Ws			
MOS11				Nu	Nu, Ws
TEX1					Nu, Ds

equally distributed all over the root. The GFP fused proteins have been found to exhibit different subnuclear localization. The findings are summarized in Table 18, together with what has been described in the literature and under which conditions the experiments were conducted. ALY2 and ALY4 show a similar subnuclear localization pattern in three independent studies examining GFP fused ALYs (This study, Uhrig et al., 2004; Pendle et al., 2005). For ALY3 there is a difference in how distinct the observed speckles are, otherwise the subcellular localization is similar. In this study the association in speckles is weak, whereas it is more prominent for the other studies (Pendle et al., 2005; Uhrig et al., 2004). In the same two studies ALY1 has been found to localize in the nucleoplasm and nucleolus, whereas ALY1 was mainly nucleoplasmic in this study. UAP56 has been shown to be mainly nucleoplasmic in two studies (Pendle et al., 2005 and this study), and present in the nucleolus in a third (Pan et al., 2014). MOS11 is found mainly in the nucleoplasm in two studies (Germain et al., 2010 and this study). The subcellular localization of TEX1 has only been investigated in this study, where it is found mainly in the nucleoplasm. The THO2 under the native promoter has been found in distinct speckles distributed in

both nucleoplasm and nucleolus, when expressed transiently in *N. benthamiana* (Francisco-Mangilet et al., 2015). HPR1 under control of the native promoter has been found mainly in the nucleoplasm when expressed transiently in *A. thaliana* (Xu et al., 2015). There seems to be an agreement between the literature and this study in how ALY2, ALY3 ALY4 and MOS11 localize in the nucleus and/or nucleolus. Divergent results have been found for ALY1 and UAP56. Furthermore, THO2 has been found in the nucleolus, whereas HPR1 and TEX1 were not. It is surprising that the THO subunits show different subnuclear localization patterns. Based on the findings in humans that the THO complex is very stably associating even under high salt conditions (Chi et al., 2013). One possible explanation for the discrepancies is that the different studies are using different experimental setups. The studies finding that ALY1, UAP56 and THO2 localize in the nucleolus are all utilizing transient expression in *N. benthamiana* or *Arabidopsis* cell culture. For ALY1 and UAP56 the expression is under control of strong constitutive promoters. It is known that transient or overexpression of GFP constructs can lead to GFP self-aggregation into complexes and ectopic protein localization (Gibson et al., 2013). This indicates that the studies using stable transformations of native promoter constructs are more accurate. The implication of this is that ALY1, ALY2, UAP56, HPR1, TEX1 and MOS11 in *Arabidopsis* are most likely mainly nucleoplasmic proteins, whereas ALY3 and ALY4 are also localizing in the nucleolus. For THO2 only transient studies have been done so it is not clear if the localization in the nucleolus is an artefact or not. In mammals ALYREF is found in big distinctive speckles, mainly in the nucleoplasm (Rodrigues et al., 2001; Luo et al., 2001; Gromadzka et al., 2016). Recent FRET studies showed that CHTOP, ALYREF and NXF1 are rather associating at the periphery of these speckles than inside (Teng & Wilson, 2013). Interestingly ALYREF in mammals has been shown to mislocate to the nucleolus, when the WQHD domain is mutated into DQAK (Gromadzka et al., 2016). *Arabidopsis* ALY3 and ALY4, which are found in the nucleolus are also the two orthologs with the least conserved WQHD domains of the four proteins. This could explain why ALY3 and ALY4 are found in the nucleolus also when expressed stably under native promoter. In HeLa cells CIP29 localizes mainly in the nucleoplasm and co-localizes with the splicing factor SC35 and ALYREF (Dufu et al., 2010). ALYREF has also been shown to colocalizes with SC35 (Zhou et al., 2000). In *Drosophila* UAP56 colocalizes with SC35 in speckles but the

two proteins are not colocalizing outside these (Gatfield et al., 2001). UAP56 has been shown to associate independently of ALY and THO, with the non-spliced pre-mRNA in humans (Chi et al., 2013). It has also been shown that upon binding of NXF1 to ALYREF, UAP56 dissociates from ALYREF. This implies that UAP56 is also found in a non-TREX associated state. Microarray analyses in *Drosophila* have shown that less than 20% of genes rely on the THO complex for expression (Rehwinkel et al., 2004). For UAP56 and NXF1 in *Drosophila* ~75% of the mRNA on the microarray are relying on these two factors (Herold et al., 2001), indicating that UAP56 can also regulate genes independent of the THO complex. In this study UAP56 and MOS11 do not accumulate in speckles to the same degree as observed for TEX1 and the four ALY proteins. This supports the indirect evidence present that UAP56 might not only associate with ALY and the THO complex, leading to different localization patterns in the nucleus.

4.3 The four *Arabidopsis* ALY proteins and potential redundancy

In *Arabidopsis* there are four orthologs of ALYREF. In this study data are presented that indicate the four ALYs are largely redundant, but are also involved in different processes. ALY3 and ALY4 localize in the nucleolus and affect flowering time in an additive and independent manner. When both the transcripts of *ALY3* and *ALY4* are reduced an accumulation of poly(A) mRNA is observed. ALY1 and ALY2 are mainly localized in the nucleoplasm, and down regulation of both transcripts does not lead to observable phenotypes or a measurable accumulation of poly(A) in the nucleus. Multiple orthologs of ALYREF are seen in other species. In yeast knockout of the export adaptor YRA1 leads to lethality. Overexpression of the paralog YRA2 can complement the lethality in a dose dependent manner (Zenklusen et al., 2002). In *Mus musculus* there are two paralogs with at least two splice variants each (REF1-1, REF1-2, REF2-1 and REF2-2) (Stutz et al., 2000). Both paralogs bind RNA and help to mediate export (Rodrigues et al., 2001). In *Caenorhabditis* simultaneous knockdown of REF1, REF2 and REF3 leads to a reduced mobility phenotype but no accumulation of poly(A) mRNA (Longman et al., 2003). In *Drosophila* there are two paralogs REF1 and REF2 (Gatfield & Izaurralde, 2002). Only REF1 is expressed in SN2 cells, and knockdown leads to decreased cell growth rate but no accumulation of poly(A) mRNA. Knockdown of human ALYREF leads to a relatively weak accumu-

lation of poly(A) mRNA compared to NXF1 (Hautbergue et al., 2009; Katahira et al., 2009; Viphakone et al., 2012). Only when ALYREF is simultaneously knocked down with other mRNA receptors like UIF, CHTOP and THOC5, respectively, a strong accumulation is seen (Viphakone et al., 2012; Chang et al., 2013; Hautbergue et al., 2009). This indicates that the other export adaptors are partially redundant in function with ALYREF. It has been speculated that the absence of mRNA accumulation when knocking down all the ALYREF orthologs in *C. elegans* and *Drosophila*, might be due to a similar mechanism where other export adaptors can compensate for the loss of ALYREF (Gatfield & Izaurralde, 2002; Hautbergue et al., 2009). The partial reduncancy observed in other organisms, upon ALYREF knockdown, seems be conserved in Arabidopsis. For *Arabidopsis* the early flowering phenotype has been only observed for the two mutants *aly3-1* and *aly3-1 aly4-1* (This study). This signifies that ALY3 and Aly4 proteins might mediate some specific export of a subset of genes affecting flowering. The observed phenotypes in *aly3-1* and *aly3-1 aly4-1* in *Arabidopsis* are relatively weak, indicating that there are other proteins capable of partially compensating for the loss. Partial redundancy is normal in *Arabidopsis* between paralogs (Pérez-Pérez et al., 2009), so it is possible that *ALY1* and *ALY2* can partially compensate for the loss of *ALY3* and *ALY4*. In order to determine if the redundant proteins are ALY1 and ALY2 or some other adaptors, higher order mutants should be produced and analysed. Of the proteins that are known to partially compensate for the loss of ALYREF in humans, only ortholog of UIF (Hautbergue et al., 2009) and THO5 (Yelina et al., 2010), respectively, are conserved in *Arabidopsis*. Investigating how mutants of these factors and the four ALY proteins affect each other, might reveal new insights into the roles of these proteins in *Arabidopsis*.

4.4 The TREX interactome

To gain more information about the *Arabidopsis* THO/TREX complex, the TREX interactome has been investigated by affinity purifications. The TREX interactome in *Arabidopsis* is very similar to what has been observed in yeast and humans. MOS11 was found to interact with UAP56. UAP56 interacts with the core THO complex, MOS11, SAC3C / THP1 (TREX2), RNA helicase 52 (EJC) and two NTC associated proteins. THO purified with tagged TEX1 associates primarily with components involved in splicing (EJC, NTC, U1-U6 snRNPs, SR proteins, hnRNP proteins, splice

site selection and components associated with the A/B/C spliceosome) and transcription (RNAPII, PAF and Elongator). It is surprising that so few interaction partners of MOS11 have been identified in this study when compared to what was identified for the other bait proteins. MOS11 has been described to interact with DDX39, UAP56, THO, ALYREF and FUS/TLS (TAP15) in metazoans (Dufu et al., 2010; Sugiura et al., 2007; Yamazaki et al., 2010). TAP15 has not been identified as an interaction partner of MOS11 in this study. In another study done in *Arabidopsis* MOS11 has also not been found to copurify with TAP15 in *Arabidopsis* (Dong et al., 2016). As very few THO1/CIP29/MOS11 interaction partners have been identified across different organisms it is possible that the interaction partners associate very transiently, like what has been speculated for CIP29 in humans (Sugiura et al., 2007). The strong enrichment of splicing factors in the TEX1 purification in this study is interesting. TEX1 copurifies with subunits of the A, B and C spliceosome complexes, which indicates that TREX is associated with the spliceosome throughout the whole splicing process. Studies in other organisms support this observation. All subunits of the TREX complex have been identified in an *in vitro* study purifying a mixture of all stages of the reassembled human spliceosome (Zhou et al., 2002). In another study with mixed stages of human spliceosomes HPR1 and ALYREF were identified (Rappsilber et al., 2002). ALYREF has also been shown to copurify with the human spliceosome A complex (Hartmuth et al., 2002) and C complex (Jurica et al., 2002). UAP56 and ALYREF associate with the B complex in both humans and *Drosophila* (Herold et al., 2009). In *Arabidopsis* association with the splicing machinery has already been indicated, as HPR1 of the THO complex co-localize with SR33 which is involved in pre-mRNA splicing (Xu et al., 2015). This fits with the findings in this study where SR33 has also been found to affinity purify with tagged TEX1. In yeast THO association with genes requires the splicing associated NTC complex (Chanarat et al., 2011). A genetic and RNA independent biochemical interaction is observed between HPR1 of the THO complex and SYF1 of the NTC complex (Chanarat et al., 2011). SYF1 is found to co-purify with TEX1 in *Arabidopsis* (this study) and associate with the reassembled spliceosomes in humans (Zhou et al., 2002). UAP56 has been described to be more directly involved in splicing. UAP56 is known to interact directly with U2AF65 in HeLa cells and be required for early spliceosome assembly (Fleckner et al., 1997). In yeast SUB2 is also reported to be

required for early spliceosome assembly, and it has been speculated that SUB2 is required to remove MUD2 (U2AF65) and allow branch point recognition by the U2 snRNP (Libri et al., 2001; Zhang & Green, 2001; Kistler & Guthrie, 2001). Other connections to splicing through the association of TREX components with the EJC have been reported. Early studies have indicated that ALYREF, UAP56 and NXF1 are part of the EJC (Gatfield et al., 2001; Le Hir et al., 2001), based on immunoprecipitation of small pieces of RNA with an EJC on it. Recently it has been shown that ALYREF and THO bind to RNA upstream of the EJC (Hong Cheng et al., 2006). ALYREF has been shown to copurify tightly with the EJC, whereas UAP56 and NXF1 did not copurify (Tange et al., 2005). ALYREF is not thought to be a core EJC component though, but rather an outer shell component of EJC (Hong Cheng et al., 2006; Tange et al., 2005). In this study four EJC subunits have been identified in the TEX1 affinity purification and one in the UAP56 purification. The association is most likely through the ALY proteins like described in mammalian cells. In yeast YRA1 has been found to associate with the RNA around the exon exon junction (Le Hir et al., 2000).

Interestingly in this study when UAP56 is used as bait, the spliceosome subunits are not enriched, but some TREX2 subunits are identified. When TEX1 is used as the bait protein UAP56, U2AF65, ALY3, EJC and subunits from the whole spliceosome co-purifies, but TREX2 is absent. Based on the role of UAP56 and ALYREF in splicing assembly and association with splicing factors in other organisms it is very likely the observed interaction of TEX1 in *Arabidopsis* with the splicing factors is through UAP56 or ALY3. Interestingly, ALY3 was found in the TEX1 purification where the spliceosome was observed, but not in the UAP56 purification where the spliceosome was absent. In the TEX1 affinity purification the interaction partners interacting through UAP56 and ALY3, are then those exclusively associated with the THO associated forms of UAP56 and ALY3. Interaction partners that require the THO associated UAP56 protein, and those that do not will both purify in the UAP56 affinity purification, but not in the TEX1 purification. This is one possible explanation between the observed differences between the TEX1 and UAP56 pulldowns. Another explanation is that the UAP56 overexpression perhaps is more disruptive than TEX1 overexpression. Overexpression of UAP56 is lethal in Caenorhabditis (MacMorris et al., 2003). In yeast overexpression of SUB2 leads to an mRNA accumulation, presumably by titrating away other mRNA export factors in a dominant negative manner

(Strässer & Hurt, 2001). This is a known problem of overexpression (Gibson et al., 2013). Overexpression can lead to a wide mixture of incomplete complexes (Dilution effect) and hence less identified interaction partners that require complete complexes to associate. Both explanations are likely contributing to the discrepancy seen between both the composition and amount of TREX interaction partners purified with TEX1 or UAP56 in this study. The THO/TREX complex does not only associate with the splicing machinery, but also with the transcription apparatus. In yeast Yra1 associates directly with the Ser2/Ser5 phosphorylated CTD of RNAPII (MacKellar & Greenleaf, 2011). The core THO complex interacts directly with the Ser2/Ser5 phosphorylated CTD of RNAPII independently of YRA1 (Meinel et al., 2013). In humans it was shown that Ser2 phosphorylated RNAPII affinity purify with HPR1 (Li et al., 2005). This supports the findings of association between TEX1 and the RNAPII in this study. The second largest subunit of RNAPII was identified in the TEX1 affinity purification. The largest subunit, which is the one THO interacts with in yeast and human, has been identified in this study but was only found in one of four IPs (Chapter 7.4). As the protein is very big (>200kDa) it is possible that it migrated above the gel slices that were excised from the gel and sent for LC-MS/MS analysis. Nevertheless this does not change that the data supports a conserved association of TREX and RNAPII in *Arabidopsis*, like seen in yeast and human.

ALYREF, THO and UAP56 recruitment in mammals is depending on both capping and the EJC complex to associate with the RNA (Gromadzka et al., 2016; Hong Cheng et al., 2006). In humans, CBP80 immunoprecipitates with ALYREF, CIP29, THOC2, THOC5 and UAP56 in an RNA independent manner (Hong Cheng et al., 2006; Dufu et al., 2010). In yeast the interaction between YRA1 and CBP20 is bridged by RNA (Iglesias et al., 2010). In *Arabidopsis* CBP20 and CBP80 were not identified in the MOS11 or UAP56 affinity purifications, and only below the threshold in the TEX1 purification (this study, Chapter 7.4). Which could indicate that the interaction between TREX and CBP in *Arabidopsis* is either not conserved, very transient or nuclease sensitive like seen in yeast.

TREX is also known to interact with the polyadenylation machinery. In yeast an interaction has been shown between YRA1 and PCF11 of the CFII complex (Johnson et al., 2009). In humans THOC5 is required for recruitment of CFIm68 to genes and regulating selection of distal polyadenylation sites (Katahira et al., 2013). In this

study TEX1 has been shown to copurify with the smaller CFIm25 subunit. In humans CFIm25 and CFIm68 are found in the CFIm complex. CFIm25 is mediating the sequence specificity whereas CFI68 increases affinity of the CFIm complex and facilitates looping of the RNA (Yang et al., 2011). This is indicating that the interaction between TREX and the 3´processing machinery is conserved in *Arabidopsis*. TREX2 components co-purifying with UAP56 in this study, indicate another interaction that is conserved between yeast, metazoans and *Arabidopsis*. In yeast SAC3 is found to copurify with SUB2 in affinity purifications (Fischer, 2002; Sträßer et al., 2002). It is not known if SUB2 and SAC3 interact directly or through other proteins like YRA1 or MEX67. In yeast MEX67 interacts directly with SAC3 (Fischer, 2002). The same has been seen in humans (Wickramasinghe, McMurtrie, et al., 2010). An allele of *Yra1ΔRRM* and *sac3Δ* interacts genetically in a synthetic lethal manner.

One of the main roles of TREX is to recruit the export receptor NXF1 to the mRNPs through direct interactions with ALYREF and THO5 (Hautbergue et al., 2008; Golovanov et al., 2006; Viphakone et al., 2012; Katahira et al., 2009). In yeast MEX67 is recruited by YRA1 (Zenklusen et al., 2001). YRA1 requires HPR1 to associate with genes (Zenklusen et al., 2002). As no obvious ortholog of NXF1 is found in *Arabidopsis* (Pendle et al., 2005), the interactors of TEX1 and UAP56 were searched for proteins containing a NFT2L or UBA domain in this study, none were found. It is very interesting that the whole machinery around NXF1 and MEX67 seems to be conserved in *Arabidopsi*s but the receptor itself is not. *Arabidopsis* TREX and TREX2 mutants of *hpr1, mos11, thp1, aly3 aly4* and *tex1 mos11* show accumulation of mRNA in the nucleus, indicating that these mRNA pathways are functional in *Arabidopsis* (Pan et al., 2012; Xu et al., 2015; Furumizu et al., 2010; Germain et al., 2010 and this study). Furthermore null mutants of THO2 are embryonic lethal, showing that THO/TREX is essential for development in *Arabidopsis*, like seen for yeast at permissive temperature (Peña et al., 2012), mice (Mancini et al., 2010) and *Drosophila* (Rehwinkel et al., 2004). This all implies that TREX is involved in mRNA export in *Arabidopsis*, and it seems unlikely that *Arabidopsis* utilizes another bulk mRNA pathway. It is speculated that *Arabidopsis* rather use a non-conserved and unidentified mRNA export receptor, but further studies would be needed to support this. The overall TREX interactome seems to be conserved among yeast, humans and *Arabidopsis*, excluding the mRNA export receptor itself. TREX seems to be linked to all

mRNA processing steps in different organisms, pointing out the central role of this complex.

4.5 Function of the TEX1 and the TREX complex

One goal of this study has been to elucidate more about the function of TEX1. It has been shown here that TEX1 is required for efficient protein biosynthesis. Downregulation of the *TEX1* transcript on its own does not lead to accumulated mRNA in the nucleus, but a severe reduction in the amount of [L-35S] methionine incorporation has been observed. The *FLC* gene is downregulated to almost background levels in *tex1-4*, whereas other genes like ARF3 and ARF4 are not differentially regulated. The two tasiRNAs *TAS1* and *TAS3* transcripts are downregulated, whereas the two miRNAs miR173 and miR390 are not. Not much is known about the function of TEX1 generally. In *Arabidopsis* TEX1 has prior to this study only been shown to be required for tasiRNA biogenesis (Yelina et al., 2010; Jauvion et al., 2010). The exact mechanism behind the effect on tasiRNA production is not known. Pri-tasiRNA shares some similarity to mRNAs, as they are transcribed by RNAPII, are capped and polyadenylated before being cleaved into the respective tasiRNAs (reviewed in Yoshikawa, 2013). It has been speculated that THO is involved in the export of these pri-mRNAs (Yoshikawa, 2013). These speculations are based on experiments showing that in *tex1-4* and *hpr1-5* the pri-tasiRNAs are accumulating leading to an decrease in *TAS1*, *TAS2* and *TAS3* transcripts (Yelina et al., 2010; Jauvion et al., 2010). In yeast *Δtex1* is not impaired in growth at 37°C whereas other THO components like THO2 are (Peña et al., 2012). When testing reporter gene expression of exogenous β-galactosidase, the β-galactosidase activity in *Δtex1* has been shown to be twofold lower than compared to wildtype. In comparison the β-galactosidase activity in *Δtho2* is reduced 500 fold (Peña et al., 2012). In HeLa cells it has been shown that downregulation of TEX1 leads to reduced protein levels of the other THO subunits in extracts (Chi et al., 2013). The same has been shown for downregulation of THO2, HPR1 and THOC5 in metazoans (Rehwinkel et al., 2004; Mancini et al., 2010; Yamazaki et al., 2010). Knockdown of TEX1 leads to a milder reduction in the protein levels of the other THO subunits than for example THO2 does (Chi et al., 2013). The indications of this are that all the THO subunits are needed for THO to be stable, but TEX1 is less important for stability than for example THO2. This is sup-

ported in *Arabidopsis* where *tho2* mutants are either embryonic lethal or severely affected in development (Francisco-Mangilet et al., 2015), whereas *tex1* mutants in this study have much weaker phenotypes. It is possible that a reduction of the *TEX1* transcript leads to a less stable and functional THO complex in *Arabidopsis*, like seen in humans. TEX1 has been described to contain multiple WD40 repeats which often give rise to a β-propeller structure (Peña et al., 2012; Rehwinkel et al., 2004). *Arabidopsis* and human TEX1, also contain six to seven predicted WD40 repeats (www.uniprot.org and http://smart.embl.de). WD40 propeller domains are common, and are often scaffolds in complexes and mediate protein-protein in teraction interphases (Stirnimann et al., 2010). Using Electron microscopy on purified yeast THO complex (Figure 45), it was shown that TEX1 protrudes out from the general THO body in a way to allow other proteins to bind it (Peña et al., 2012). Although of very speculative nature this all points towards that TEX1 might be required for THO stability and mediates protein interactions to other complexes. The function of TEX1 is then most likely linked to the function of the THO complex. The THO complex is required for association of NXF1 with bulk mRNA through export

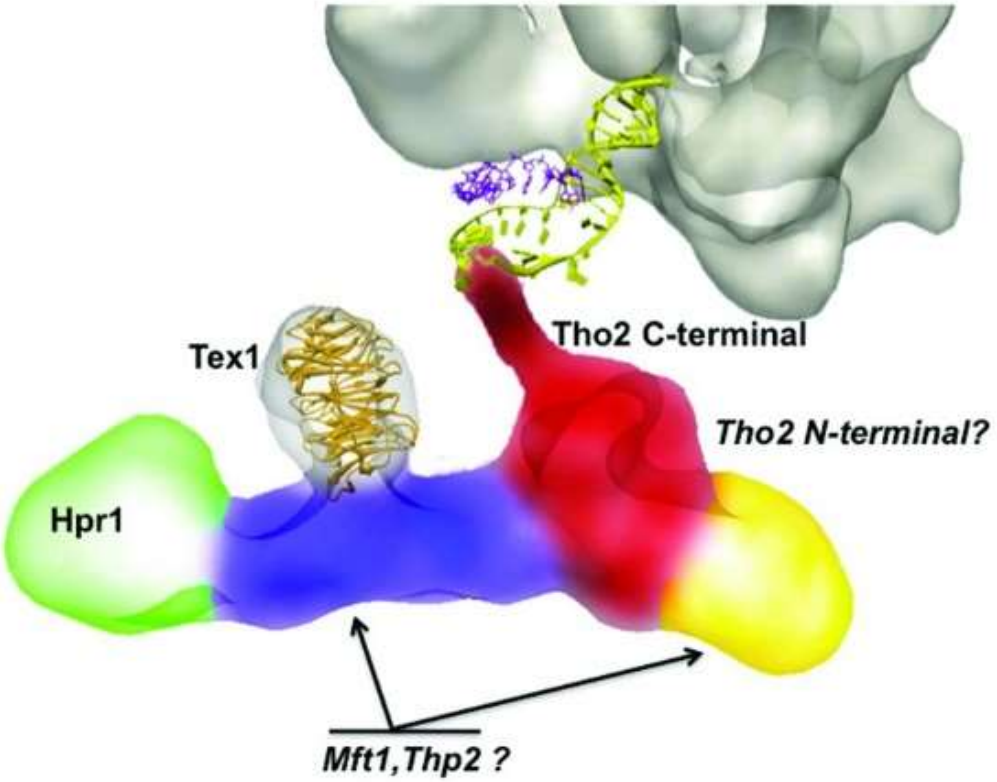


Figure 45. Three-dimensional structure of THO, generated by electron microscopy and image processing. The THO complex comprises THO2, HPR1, MFT1, THP2 and TEX1 proteins. Different electron microscopy experiments have helped to localized HPR1, TEX1 and the C-terminal region of THO2. (Peña et al., 2012)

adaptors like ALYREF in mammals (Viphakone et al., 2012). This correlates with the observed mRNA accumulation seen when knocking down THO subunits in yeast, metazoans and *Arabidopsis* (Viphakone et al., 2012; Sträßer et al., 2002; Furumizu et al., 2010; Pan et al., 2012; Xu et al., 2015). Yeast PAR-CLIP experiments have shown that THO is a general factor associated with RNAPII transcripts (Baejen et al., 2014), whereas the three yeast export adaptors YRA1, NPL3 and NAB2 show distinct crosslinking patterns. This indicates that THO in yeast might have a more generalized role, whereas export adaptors mediate the specificity. In *Drosophila* THO is only required for expression of 20% of the genes (Rehwinkel et al., 2004), indicating that THO might have a more specialized function in higher eukaryotes. There are indications that TREX might affect the splicing process itself. TREX subunits prefer to bind the spliced version of intron containing mRNAs in both human and yeast (Baejen et al., 2014; Gromadzka et al., 2016). SUB2/UAP56 plays a direct role in spliceosome assembly (Fleckner et al., 1997; Libri et al., 2001; Zhang & Green, 2001; Kistler & Guthrie, 2001). Recently it has also been shown in *Arabidopsis* that *hpr1* and *tho2* mutants are affected in alternative splicing (Francisco-Mangilet et al., 2015; Furumizu et al., 2010). This indicates a possible functional role of TREX in splicing. How the THO/TREX complex affects alternative splicing in *Arabidopsis* is unknown and should be investigated further. THO/TREX has also been indicated to influence transcription elongation and genome stability. In yeast THO travels with RNAPII (Sträßer et al., 2002; Zenklusen et al., 2002) by direct interaction with RNAPII on active genes genome-wide (Meinel et al., 2013; Gómez-González et al., 2011). Null mutants of the THO complex and Δ*sub2* are impaired in transcription elongation especially on long and CG rich genes (Rondón et al., 2003; Chávez et al., 2000; Chávez & Aguilera, 1997; Chávez et al., 2001). It is indicated that the processivity of RNAPII is affected and not the elongation rate in THO mutants (Mason & Struhl, 2005). THO null mutants also exhibit strong transcription dependent recombination phenotypes leading to genomic instability (Chávez et al., 2000; Chávez & Aguilera, 1997; Piruat & Aguilera, 1998). The genomic instability has been shown to be due to the presence of RNA loops (DNA:RNA hybrids), as overexpression of RNase H represses the *hpr1*Δ elongation defects (Huertas & Aguilera, 2003). In the same study it was shown that self-splicing ribozymes does not have an elongation defect when being transcribed in *hpr1*Δ, supporting that the RNA causes problems in

*hpr1*Δ. Recently it has been shown in yeast that THO/TREX on a genome-wide scale prevents R-loop dependent replication obstacles. Rrm3 is a helicase that has been shown to accumulate when the replication fork pauses or stalls. A ChIP chip analysis shows that Rrm3 accumulates more at highly transcribed regions in *hpr1*Δ compared to wildtype yeast (Gómez-González et al., 2011). This accumulation of Rrm3 disappears when RNase H is overexpressed in *hpr1*Δ, indicating that THO represses R-loop formation and replication impairment on a genomewide scale in yeast. The yeast THO complex associates with the RNA transcript, indicating that THO might prevent R-loops by binding both RNAPII and the RNA (Meinel et al., 2013; Baejen et al., 2014), separating it from the DNA. When transfecting HeLa cells directly with siRNAs against HPR1/THOC it leads to transcriptional defects, mRNA export defects and increased recombination/DNA breaks (Domínguez-Sánchez et al., 2011; Li et al., 2005). When using THO depleted HeLa extracts it does not affect transcription of plasmid templates *in vitro* (Masuda et al., 2005). It has been shown in HeLa cells that overexpression of RNase H represses the transcription defect, indicating that THO also has an role in repressing R-loops in humans (Domínguez-Sánchez et al., 2011). Consistent with a role of THO in elongation, another study has shown that the TREX subunit ALYREF also affects transcription. When ALYREF is knocked down in 293A TOA cells two thirds of the transcripts are upregulated in the nucleus whereas 98% of the differentially expressed transcripts are down regulated in the cytoplasm, signifying the main effect of ALYREF is in mRNA export (Stubbs & Conrad, 2015). Interestingly 646 transcripts are down regulated in both nucleus and cytoplasm, indicating affected transcription. The changes in RNAs levels are not due to altered RNA stability, and the decrease in transcription correlates with reduced RNAPII density on the genes (Stubbs & Conrad, 2015). In *Arabidopsis* it has been reported that the RTE1 transcript, involved in ethylene signalling, is downregulated in the nucleus of *hpr1-5* independently of mRNA degradation (Xu et al., 2015). This does not correlate with a mRNA export block, where an accumulation in the nucleus would have been expected. Down regulation of HPR1 leads to general accumulation of mRNA in the nucleus (Xu et al., 2015; Furumizu et al., 2010; Pan et al., 2012), indicating that *Arabidopsis* THO/TREX might be involved in both mRNA export and transcription like seen in humans. In this study the *FLC* transcript is found to be downregulated to almost background levels, whereas *ARF3* and *ARF4* are not affected. This would indi-

cate that if TEX1 is involved in transcription it is not on all genes, like seen for ALYREF in humans (Stubbs & Conrad, 2015). Down regulation of *TEX1* transcript in this study does not lead to an observable accumulation of bulk mRNA export. It is interesting that *hpr1-5* shows accumulation of poly(A) mRNA whereas the single mutant of *tex1-4* does not. Only when combined with a simultaneous reduction in *MOS11* transcript, an effect of TEX1 on mRNA export is seen. One possibility could be that TEX1 is not as essential for the mRNA export functionality of THO as the other subunits, but is more important for the function behind efficient protein synthesis rates in *Arabidopsis*. As many defects could lead to decreased protein synthesis further examinations are needed to determine what is the cause of the observed defect in *tex1-4*.

4.6 Function of MOS11 and the interplay with THO/TREX

Another goal of this study has been to gain more information about the role of MOS11 in regards to the TREX complex in *Arabidopsis*. MOS11 has been found to interact genetically and physically with THO/TREX subunits, but acquired data also supports a role independent of THO. Reduction of *MOS11* transcript leads to accumulation of poly(A) mRNA, an earlier flowering phenotype and increased total levels of *FLC, ARF4* and *TEX1* transcript. MOS11 is not required for tasiRNA biogenesis, but the transcript levels of miR173 are decreased. This indicates that the observed reduction of miR173 transcript is not enough to affect the processing of TAS1 in *mos11-2.* In humans CIP29 has been shown to bind RNA and to stimulate the helicase activity of DDX39 in the presence of ATP (Sugiura et al., 2007). In one study overexpression but not knockdown of CIP29 leads to accumulation of poly(A) mRNA (Dufu et al., 2010). In the same study a MS2 tethering assay shows that CIP29 export of a reporter construct is dependent on NXF1. In another study an accumulation of poly(A) mRNA was seen upon knockdown of CIP29 (Yamazaki et al., 2010). This strongly indicates a role for CIP29 in mRNA export. THO1 in yeast is binding to RNA, but no direct connection to mRNA export has been made (Jimeno et al., 2006). The main RNA binding activity is found outside of the SAP domain that both THO1 and CIP29 have. Interestingly *Arabidopsis* MOS11, unlike yeast THO1 and human CIP29, does not contain a conserved Sap domain (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). The SAP domain is suspected

to mediate the DNA binding activity of THO1 (Jimeno et al., 2006). If this leads to a different function of MOS11 compared to THO1 and CIP29 is unknown. The involvement of MOS11 in mRNA export seems to be conserved from CIP29, as the *mos11-2* mutant accumulates poly(A) mRNA (Germain et al., 2010, this study). The mRNA export block in *mos11-2* might be the reason why there is observed higher transcript levels of *FLC*, *ARF4* (non-significant) and *TEX1* in *mos11-2* than wildtype in this study. Along that line it has been shown in humans that upon knockdown of ALYREF ~80% of genes upregulated in total RNA extracts are also differentially upregulated in isolated nuclear RNA extracts (Stubbs & Conrad, 2015). This indicates that knockdown of ALYREF leads to nuclear accumulation of most of the RNAs upregulated in the total extracts. If this is also the case when *MOS11* transcript is reduced in *Arabidopsis*, would need further studies on cytoplasmic and nucleoplasmic RNA populations to clarify. The exact functional relationship between CIP29 and TREX is not clear. In humans CIP29 has been shown to both associate directly with TREX (UAP56) (Dufu et al., 2010) and AMEX (DDX39) (Yamazaki et al., 2010). In the latter study knockdown of UAP56, HPR1, THO2 and ALYREF leads to premature sister chromatid separation (PMSCS) causing chromosome misalignment, whereas knockdown of CIP29 and DDX39 leads to sister chromatids not separating at the chromosome arms. In the UAP56 knockdown genes are downregulated that are known to cause PMSCS upon downregulation, whereas genes are downregulated in DDX39 that are required for chromosome arm resolution (Yamazaki et al., 2010). This indicates that TREX and CIP29 have specific mRNA targets in humans. Interestingly the single mutant *mos11-2* mutant does not show a comparable strong molecular phenotype to what has been seen in humans upon CIP29 knockdown (Germain et al., 2010, This study). Likewise, the null mutant *tho1Δ* shows no growth defects or recombination phenotypes, even in *hpr1Δ* background. This could be an indication of that CIP29 plays a more important role in development, than THO1 do in yeast. In yeast multi copy THO1 can partially rescue THO mutants (Piruat & Aguilera, 1998; Jimeno et al., 2006). Not much are connecting THO1 and TREX in yeast, apart from the observation that THO1 and TREX both are recruited to transcriptionally active genes in a similar manner (Jimeno et al., 2006; Meinel et al., 2013). The ATP dependent protein-protein interaction between CIP29 and UAP56 is known (Dufu et al., 2010), but it is not known if TREX and CIP29 actually work

Table 19. TEX1 and MOS11 exhibit different kinds of genetic interaction patterns in regards to different phenotypes and gene expression levels. Some of the different traits quantified in this paper has been ordered in accordance to the type of genetic interaction seen between MOS11 and TEX1. The type of genetic interaction seen for each trait were evaluated based on the results from statistical two-way ANOVA analysis and the definitions in Pérez-Pérez et al., 2009.

No effect	TEX1 or MOS11 specific	Additivity	Epistasis	Synergy
-*ARF3* transcript	-*TAS1* transcript (*tex1-4*)	-Bolting times	-*FLC* transcript	-Root length
-*miR390* transcript	-*TAS3* transcript (*tex1-4*)			-Leaf morphology
	-*miR173* transcript (*mos11-2*)			-Number of inflorescence
	Lateral root density			-Silique morphology
				-Nondeveloped seeds
				-Accumulation of poly(A) mRNA
				-flower meristem morphology

together in mRNA export. The knockdown of UAP56 and DDX39 leads to 63 shared gene targets (Yamazaki et al., 2010). It is unknown if CIP29 plays any role in the regulation of these genes. In this study it has been shown that MOS11 and TEX1 exhibit a complex pattern of genetic interactions (summarized in Table 19). Additivity is widely accepted as a sign of absence of a functional relationship, whereas epistasis is usually thought of as the phenotypic outcome of genetic interaction (Pérez-Pérez et al., 2009). Synergy between two factors is more complicated to interpret. Synergy can arise from functional redundancy of paralogues, disruption of two pathways that converge at a common node or when one mutation enhances the sensitivity to the effects of another mutation (Pérez-Pérez et al., 2009). In this study it has been found that MOS11 and TEX1 have no functional relationship in regards to the regulation of tasiRNA biogenesis, miR173 expression, the transition from vegetative to reproductive state and lateral root density. *TEX1* and *MOS11* have been found to interact genetically in regards to the number of rosetta branches, total root length, leaf morphology, flower bud development, seed set, silique morphology and accumulation of poly(A) mRNA. As poly(A) mRNA in this study is accumulated in a synergistic manner when both *MOS11* and *TEX1* are downregulated it indicates that the functional redundancy might have to do with mRNA export. That the genetic interactions of MOS11 and TEX1 are varied between different phenotypes suggests that TREX and MOS11 might affect different transcripts, with some overlap. A synergistic effect on mRNA export of interrupting *TEX1* and *MOS11* simultaneously has

not been shown in other organisms so far, which makes the findings interesting. More research should be done to find out if this is conserved among other organisms and to clarify the exact mode of association between TREX and MOS11.

4.7 TEX1 and MOS11 in plant development

In this study it has been shown that TEX1 and MOS11 are important for multiple developmental steps in *Arabidopsis*. Trichome branching is one of the developmental processes that are affected. The single and double mutants of *tex1-4* and *mos11-2* have more trichomes with three branch points than wildtype. The double mutant shows a synergistic increase in trichomes with four branch points. Two classes of mutants have been shown to lead to an increase in branch points. The first class are mutants affected in endoreplication and the other are those that are not (Schwab et al., 2000). Mutants with more DNA after endoreplication have more branches (Perazza et al., 1999). Of the mutants not affecting endoreplication, both positive and negative regulators have been found (An et al., 2012; Schwab et al., 2000; Folkers et al., 1997). The gene NOECK is an example of a negative repressor of trichome branching, *nok* mutants have up to seven branches pr. trichome (Folkers et al., 1997). Further studies would be needed to elucidate by which mechanism TEX1 and MOS11 affect trichome branching. Another development process affected in the double mutant *tex1-4 mos11-2* is seed set development. The double mutant shows more non-developed ovules than the other genotypes. An increase in non-developed ovules has also been observed in the *tho2-6* and *tho2-7* mutants (Francisco-Mangilet et al., 2015). In the *tho2* mutants the anthers are short and deform, whereas the stamen filaments in *tex1-4 mos11-2* seem elongated compared to wildtype. This indicates that *tex1-4 mos11-2* affects stamen filament elongation differently than *tho2-6* and *tho2-7.* Filament elongation has been shown to be dependent on auxin and gibberellin, where mutants deficient in these hormone pathways exhibit shorter filaments (Cheng et al., 2004; Cardarelli & Cecchetti, 2014). Very few cases with elongated filaments have been described in literature. In transgene multicopy plants of the auxin-response genes *SAUR63*, an increase in filament length is seen (Chae et al., 2012). A study using a dominant mutant of another auxin-response gene, IAA19, has shown that stamen elongation and pistil growth needs to be properly timed for efficient pollination (Tashiro et al., 2009). It is possible that the increase in non-

developed ovules in *tex1-4 mos11-2*, is the consequence of misregulated stamen filament growth.

The mutants of *tex1*-4 and *tex1-4 mos11-2* in this study are affected in the number of activated rosetta axillary buds. Most of the axillary buds are normally arrested in a dormant state (Apical dorminance) by the primary inflorescence (Tantikanjana et al., 2001; Pautler et al., 2013). A couple of mutants have been described that show reduced apical dominance and abnormal activation of the dormant axillary meristems like seen in *tex1-4 mos11-2*. T-DNA mutants of genes involved in the MORE AXILLARY GROWTH (MAX) pathway have been shown to cause activation of the dormant meristems (Bennett et al., 2006). Some of the *MAX* genes are involved in biosynthesis of the hormone strigolactone, which in collaboration with auxin inhibits axillary meristem outgrowth (Domagalska & Leyser, 2011; Bennett et al., 2006). The effect is thought to be indirect through altered auxin transport (reviewed in Domagalska & Leyser, 2011). Another gene, *BRC1,* is necessary for auxin induced control of apical dormancy (Aguilar-Martínez et al., 2007). Loss of *BRC1* leads to axillary meristem activation. This suggests that deficient axillary meristem dormancy is either indicative of alterations in auxin transport or a reduction in dormancy factors.

In this study the *tex1-4 mos11-2* double mutant exhibits earlier bolting. The early bolting phenotype of *tex1-4* correlates with deficient expression levels of *FLC*, whereas the additive reduction seen in bolting time of the double mutant is independent of *FLC.* The transition from the shoot apical meristem to the inflorescence meristem can be mediated by multiple pathways I.e. gibberellic acid, autonomous, vernalisation, thermosensory and photoperiod (Blázquez et al., 2003; Mouradov et al., 2002; Simpson & Dean, 2002). It seems unlikely that MOS11 is regulating factors through the autonomous or vernalization pathways as these affect flowering through *FLC* (Amasino & Michaels, 2010). More studies would be needed to find out if MOS11 is involved with the gibberellin, thermosensory or photoperiod pathways.

In the *tex1-4 mos11-2* double mutant the flower meristems do not develop into flowers until late in development. After the switch from shoot apical meristem to inflorescence meristem, the genes LFY and AP1 specify flower meristem identity on the flanks of the inflorescence meristems (Liu et al., 2009). The *pin1*, *yuc*, and *pid* mutants are defective in auxin transport and show defective flower meristem initiation (Vernoux et al., 2000; Youfa Cheng et al., 2006; Zourelidou et al., 2014). The *tex1-4*

mos11-2 double mutant in this study has some morphological resemblance with the mutant of *pid-14* at later developmental stages. For example, the long empty pin like branches, and reduction in siliques and secondary inflorescences on the rosetta branches. The mutant *pid-14* meristems are not initiated whereas they are in *tex1-4 mos11-2* evidenced by the flowers observed at later stages. As the *tex1-4 mos11-2* double mutant also shows defects in pistil morphology, it is speculated that the floral organ identity factors might be affected. These identity factors in *Arabidopsis* are APETALA1, APETALA2, APETAL3, PISTILLATA and AGAMOUS. They are required for proper flower organogenesis from the floral organ primordia on the flower meristem (Wellmer et al., 2014). Further studies would be needed to find out at which stage in flower organogenesis the double mutant *tex1-4 mos11-2* is affected. The *tex1-4 mos11-2* mutant is affected in bulk mRNA export and has a reduced bulk protein synthesis rate. With these defects it is not surprising to see pleiotropic phenotypes. Most of the described phenotypes in the double mutant are the result of a defect in developmental switches like bolting, axillary meristem dormancy and initiation of flower organogenesis. It is speculated that these switches are especially sensitive to the defects of *tex1-4 mos11-2* as proteins might be expressed more slowly, affecting the switch positively or negatively. As most of these phenotypes are also affected by an auxin increase or reduction, it is also possible that factors required for auxin homeostasis are among the targets of TEX1 and MOS11. Further studies would be needed to confirm this.

4.8 Outlook

In this thesis different subunits of THO/TREX have been analysed. The data has in the case of TEX1 and MOS11 indicated that these two factors regulate some transcripts independent of one another. The same was seen for ALY3 and ALY4 on the morphological level, where *aly3-1 aly4-1* showed phenotypes not observed in *aly1-1 aly2-1*. It would be interesting to analyse mutants of subunits in the THO/TREX complex not examined in this study. By creating higher order mutants, this could help elucidate how much the different factors overlap in function and from that indirectly if they affect specific transcripts. Thanks to advances in the last years in the field of high throughput sequencing, this can also be done more directly. To find out with which genes and transcripts the THO/TREX subunits associate, Chromatin Immuno

precipitation (ChIP) and RNA immunoprecipitation (RIP) combined with high throughput sequencing could be utilized. For the THO/TREX subunits not addressed in this thesis either specific antibodies or transgenic plant lines with the bait proteins fused to a Tag, would be needed. It would also be interesting to quantify the nucleoplasmic and cytoplasmic RNA populations in different mutants by RNA sequencing, like it has been done for ALYREF in humans (Stubbs & Conrad, 2015). This would help estimating how many of the transcripts are affected by accumulation in the nucleus, and how many are affected by reduced transcript levels in the nucleus. In *Arabidopsis* due to the presence of a rigid cell wall, cytoplasmic extractions are done with protoplasts (Göhring et al., 2014). Protoplasts are somewhat artificial in this context and this method would have to be optimized for seedlings/plant tissue in order to better reflect the mRNA populations *in vivo*.

5 Summary

In eukaryotic cells, transcription and processing of mRNAs in the nucleus are separated from translation in the cytoplasm by the nuclear membrane. Transport of mRNA from the nucleus to the cytoplasm is facilitated through the nuclear pore complex. The heteromeric THO/TREX complex plays an important role in the export of mRNA, by recruiting the mRNA export receptor that translocates the mRNA through the nuclear pore complex. The composition and function of the TREX complex in *Arabidopsis thaliana* is not well described.In this work the four *Arabidopsis* orthologs ALY1, ALY2, ALY3 and ALY4 of the human TREX subunit ALYREF, were characterized. All four proteins were found to interact directly with another TREX subunit UAP56, as seen in humans. ALY3 and ALY4 were found to localize in both nucleoplasm and nucleolus of root cells, whereas ALY1 and ALY2 only localized in the nucleoplasm. A double mutant of *aly1-1* and *aly2-1* showed no phenotypes. The double mutant of *aly3-1* and *aly4-1* showed an early flowering phenotype and a bulk accumulation of poly(A) mRNA in the nucleus, indicating partial redundancy.

A member of the THO subcomplex of TREX, TEX1 was characterized. A knockdown mutant *tex1-4* exhibited multiple pleiotropic phenotypes, like early flowering, additionally activated axillary meristems and less elongated lateral roots. The *tex1-4* mutant did not accumulate poly(A) mRNA but showed decreased protein synthesis rates. TEX1 was shown to copurify with the rest of the THO complex, UAP56, ALY3 and multiple proteins involved in splicing and transcription. MOS11, an *Arabidopsis* ortholog of the TREX associated human CIP29 protein, was also characterized. MOS11 was shown to copurify UAP56, and to be copurified by UAP56. The *mos11-2* mutant on its own showed a mild flowering phenotype and accumulation of poly(A) mRNA. When *mos11-2* was crossed with *tex1-4* additional pleiotropic phenotypes were observed. Trichomes, auxillary meristems, inflorescences and flower organs were especially affected. TEX1 was found to regulate tasiRNA independent of MOS11, whereas MOS11 was found to regulate miR173 independently of TEX1. In the double mutant a synergistic accumulation of poly(A) mRNA was seen, correlating with the observed increase in phenotype severity. Therefore, this study provides evidence that the general composition of TREX is conserved and that ALY3, ALY4, TEX1 and MOS11 are involved in mRNA export.

6 References

Abruzzi, K.C., Lacadie, S. & Rosbash, M., 2004. Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes. *The EMBO journal*, 23(13), pp.2620–31.

Aguilar-Martínez, J.A., Poza-Carrión, C. & Cubas, P., 2007. Arabidopsis BRANCHED1 acts as an integrator of branching signals within axillary buds. *The Plant cell*, 19(2), pp.458–72.

Aguilera, A. & García-Muse, T., 2012. R loops: from transcription byproducts to threats to genome stability. *Molecular cell*, 46(2), pp.115–24.

Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., Sali, A. & Rout, M.P., 2007. The molecular architecture of the nuclear pore complex. *Nature*, 450(7170), pp.695–701.

Alcázar-Román, A.R., Tran, E.J., Guo, S. & Wente, S.R., 2006. Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. *Nature cell biology*, 8(7), pp.711–6.

Amasino, R.M. & Michaels, S.D., 2010. The timing of flowering. *Plant physiology*, 154(2), pp.516–20.

An, L., Zhou, Z., Su, S., Yan, A. & Gan, Y., 2012. GLABROUS INFLORESCENCE STEMS (GIS) is required for trichome branching through gibberellic acid signaling in Arabidopsis. *Plant & cell physiology*, 53(2), pp.457–69.

Arts, G.J., Kuersten, S., Romby, P., Ehresmann, B. & Mattaj, I.W., 1998. The role of exportin-t in selective nuclear export of mature tRNAs. *The EMBO journal*, 17(24), pp.7430–41.

Assouline, S., Culjkovic, B., Cocolakis, E., Rousseau, C., Beslu, N., Amri, A., Caplan, S., Leber, B., Roy, D.-C., Miller, W.H. & Borden, K.L.B., 2009. Molecular targeting of the oncogene eIF4E in acute myeloid leukemia (AML): a proof-of-principle clinical trial with ribavirin. *Blood*, 114(2), pp.257–60.

Baejen, C., Torkler, P., Gressel, S., Essig, K., Söding, J. & Cramer, P., 2014. Transcriptome maps of mRNP biogenesis factors define pre-mRNA recognition. *Molecular cell*, 55(5), pp.745–57.

Baltz, A.G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., Wyler, E., Bonneau, R., Selbach, M., Dieterich, C. & Landthaler, M., 2012. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Molecular cell*, 46(5), pp.674–90.

Batisse, J., Batisse, C., Budd, A., Böttcher, B. & Hurt, E., 2009. Purification of nuclear poly(A)-binding protein Nab2 reveals association with the yeast transcriptome and a messenger ribonucleoprotein core structure. *The Journal of biological chemistry*, 284(50), pp.34911–7.

Bennett, S.R.M., Alvarez, J., Bossinger, G. & Smyth, D.R., 1995. Morphogenesis in pinoid mutants of Arabidopsis thaliana. *The Plant Journal*, 8(4), pp.505–520.

Bennett, T., Sieberer, T., Willett, B., Booker, J., Luschnig, C. & Leyser, O., 2006. The Arabidopsis MAX pathway controls shoot branching by regulating auxin transport. *Current biology : CB*, 16(6), pp.553–63.

Blázquez, M.A., Ahn, J.H. & Weigel, D., 2003. A thermosensory pathway controlling flowering time in Arabidopsis thaliana. *Nature genetics*, 33(2), pp.168–71.

Bohnsack, M.T., Czaplinski, K. & Gorlich, D., 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA (New York, N.Y.)*, 10(2), pp.185–91.

Bolte, S. & Cordelières, F.P., 2006. A guided tour into subcellular colocalization analysis in light microscopy. *Journal of microscopy*, 224(Pt 3), pp.213–32.

Bono, F. & Gehring, N.H., 2011. Assembly, disassembly and recycling: the dynamics of exon junction complexes. *RNA Biology*, 8(1), pp.24–29.

Cabal, G.G., Genovesio, A., Rodriguez-Navarro, S., Zimmer, C., Gadal, O., Lesne, A., Buc, H., Feuerbach-Fournier, F., Olivo-Marin, J.-C., Hurt, E.C. & Nehrbass, U., 2006. SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. *Nature*, 441(7094), pp.770–3.

Cardarelli, M. & Cecchetti, V., 2014. Auxin polar transport in stamen formation and development: how many actors? *Frontiers in plant science*, 5, p.333.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., Krijgsveld, J. & Hentze, M.W., 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*, 149(6), pp.1393–406.

Chae, K., Isaacs, C.G., Reeves, P.H., Maloney, G.S., Muday, G.K., Nagpal, P. & Reed, J.W., 2012. Arabidopsis SMALL AUXIN UP RNA63 promotes hypocotyl and stamen filament elongation. *The Plant journal : for cell and molecular biology*, 71(4), pp.684–97.

Chanarat, S., Seizl, M. & Sträßer, K., 2011. The prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes. *Genes and Development*, 25(11), pp.1147–1158.

Chang, C.-T., Hautbergue, G.M., Walsh, M.J., Viphakone, N., van Dijk, T.B., Philipsen, S. & Wilson, S. a, 2013. Chtop is a component of the dynamic TREX mRNA export complex. *The EMBO journal*, 32(3), pp.473–86.

Chávez, S. & Aguilera, A., 1997. The yeast HPR1 gene has a functional role in transcriptional elongation that uncovers a novel source of genome instability. *Genes & development*, 11(24), pp.3459–70.

Chávez, S., Beilharz, T., Rondón, A.G., Erdjument-Bromage, H., Tempst, P., Svejstrup, J.Q., Lithgow, T. & Aguilera, A., 2000. A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. *The EMBO journal*, 19(21), pp.5824–34.

Chávez, S., García-Rubio, M., Prado, F. & Aguilera, A., 2001. Hpr1 is preferentially required for transcription of either long or G+C-rich DNA sequences in Saccharomyces cerevisiae. *Molecular and cellular biology*, 21(20), pp.7054–64.

Chekanova, J.A., Abruzzi, K.C., Rosbash, M. & Belostotsky, D.A., 2008. Sus1, Sac3, and Thp1 mediate post-transcriptional tethering of active genes to the nuclear rim as well as to non-nascent mRNP. *RNA (New York, N.Y.)*, 14(1), pp.66–77.

Cheng, H., Dufu, K., Lee, C.-S., Hsu, J.L., Dias, A. & Reed, R., 2006. Human mRNA export machinery recruited to the 5' end of mRNA. *Cell*, 127(7), pp.1389–400.

Cheng, H., Qin, L., Lee, S., Fu, X., Richards, D.E., Cao, D., Luo, D., Harberd, N.P. & Peng, J., 2004. Gibberellin regulates Arabidopsis floral development via suppression of DELLA protein function. *Development (Cambridge, England)*, 131(5), pp.1055–64.

Cheng, Y., Dai, X. & Zhao, Y., 2006. Auxin biosynthesis by the YUCCA flavin monooxygenases controls the formation of floral organs and vascular tissues in Arabidopsis. *Genes & development*, 20(13), pp.1790–9.

Chi, B., Wang, Q., Wu, G., Tan, M., Wang, L., Shi, M., Chang, X. & Cheng, H., 2013. Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Research*, 41(2), pp.1294–1306.

Chuang, T.-W., Lee, K.-M. & Tarn, W.-Y., 2015. Function and pathological implications of exon junction complex factor Y14. *Biomolecules*, 5(2), pp.343–55.

Clough, S.J. & Bent, A.F., 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *The Plant journal : for cell and molecular biology*, 16(6), pp.735–43.

Costes, S. V, Daelemans, D., Cho, E.H., Dobbin, Z., Pavlakis, G. & Lockett, S., 2004. Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophysical journal*, 86(6), pp.3993–4003.

Cronshaw, J.M., Krutchinsky, A.M., Zhang, W., Chait, B.T. & Matunis, M.J., 2002. Proteomic analysis of the mammalian nuclear pore complex. *The Journal of Cell Biology*, 158(5), pp.915–927.

Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M. & Borden, K.L.B., 2006. eIF4E is a central node of an RNA regulon that governs cellular proliferation. *The Journal of cell biology*, 175(3), pp.415–26.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K. & Scheible, W.-R., 2005. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant physiology*, 139(1), pp.5–17.

Dedecker, M., Van Leene, J. & De Jaeger, G., 2015. Unravelling plant molecular machineries through affinity purification coupled to mass spectrometry. *Current opinion in plant biology*, 24, pp.1–9.

Domagalska, M.A. & Leyser, O., 2011. Signal integration in the control of shoot branching. *Nature reviews. Molecular cell biology*, 12(4), pp.211–21.

Domínguez-Sánchez, M.S., Barroso, S., Gómez-González, B., Luna, R. & Aguilera, A., 2011. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS genetics*, 7(12), p.e1002386.

Dong, C.-H., Hu, X., Tang, W., Zheng, X., Kim, Y.S., Lee, B. & Zhu, J.-K., 2006. A putative Arabidopsis nucleoporin, AtNUP160, is critical for RNA export and required for plant tolerance to cold stress. *Molecular and cellular biology*, 26(24), pp.9533–43.

Dong, O.X., Meteignier, L.-V., Plourde, M.B., Ahmed, B., Wang, M., Jensen, C., Jin, H., Moffett, P., Li, X. & Germain, H., 2016. Arabidopsis TAF15b Localizes to RNA Processing Bodies and Contributes to snc1-Mediated Autoimmunity. *Molecular plant-microbe interactions : MPMI*, p.MPMI11150246R.

Drubin, D.A., Garakani, A.M. & Silver, P.A., 2006. Motion as a phenotype: the use of live-cell imaging and machine visual screening to characterize transcription-dependent chromosome dynamics. *BMC cell biology*, 7(1), p.19.

Du, Z., Zhou, X., Ling, Y., Zhang, Z. & Su, Z., 2010. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic acids research*, 38(Web Server issue), pp.W64–70.

Dubrovsky, J.G. & Forde, B.G., 2012. Quantitative analysis of lateral root development: pitfalls and how to avoid them. *The Plant cell*, 24(1), pp.4–14.

Dufu, K., Livingstone, M.J., Seebacher, J., Gygi, S.P., Wilson, S. a. & Reed, R., 2010. ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. *Genes and Development*, 24(18), pp.2043–2053.

Fan, H.Y., Merker, R.J. & Klein, H.L., 2001. High-copy-number expression of Sub2p, a member of the RNA helicase superfamily, suppresses hpr1-mediated genomic instability. *Molecular and cellular biology*, 21(16), pp.5459–70.

Felippes, F.F. & Weigel, D., 2009. Triggering the formation of tasiRNAs in Arabidopsis thaliana: the role of microRNA miR173. *EMBO reports*, 10(3), pp.264–70.

Fiegle, D., 2014. *Characterization of Arabidopsis thaliana mRNA export adaptors*. University of Regensburg.

Fischer, T., 2002. The mRNA export machinery requires the novel Sac3p-Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores. *The EMBO Journal*, 21(21), pp.5843–5852.

Fleckner, J., Zhang, M., Valcárcel, J. & Green, M.R., 1997. U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes & development*, 11(14), pp.1864–72.

Folco, E.G., Lee, C.S., Dufu, K., Yamazaki, T. & Reed, R., 2012. The proteins PDIP3 and ZC11A associate with the human TREX complex in an ATP-dependent manner and function in mRNA export. *PLoS ONE*, 7(8), pp.1–7.

Folkers, U., Berger, J. & Hülskamp, M., 1997. Cell morphogenesis of trichomes in Arabidopsis: differential control of primary and secondary branching by branch initiation regulators and cell growth. *Development (Cambridge, England)*, 124(19), pp.3779–86.

Fox, J., 2005. The R Commander: A Basic-Statistics Graphical User Interface to R. *Journal of Statistical Software*, 14(9), pp.1–42.

Francisco-Mangilet, A.G., Karlsson, P., Kim, M.-H., Eo, H.J., Oh, S.A., Kim, J.H., Kulcheski, F.R., Park, S.K. & Manavella, P.A., 2015. THO2, a core member of the THO/TREX complex, is required for microRNA production in Arabidopsis. *The Plant Journal*, 82(6), pp.1018–1029.

Furumizu, C., Tsukaya, H. & Komeda, Y., 2010. Characterization of EMU, the Arabidopsis homolog of the yeast THO complex member HPR1. *RNA (New York, N.Y.)*, 16(9), pp.1809–17.

Gaouar, O. & Germain, H., 2013. mRNA export: threading the needle. *Frontiers in plant science*, 4, p.59.

Gatfield, D., Hir, L.H., Schmitt, C., Braun, I.C., Ko, T., Wilm, M. & Izaurralde, E., 2001. The DExH / D box protein HEL / UAP56 is essential for mRNA nuclear export in Drosophila. *Current biology*, 11, pp.1716–1721.

Gatfield, D. & Izaurralde, E., 2002. REF1/Aly and the additional exon junction complex proteins are dispensable for nuclear mRNA export. *The Journal of cell biology*, 159(4), pp.579–88.

Gebhardt, A., Habjan, M., Benda, C., Meiler, A., Haas, D.A., Hein, M.Y., Mann, A., Mann, M., Habermann, B. & Pichlmair, A., 2015. mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. *Nature communications*, 6, p.8192.

Germain, H., Na, Q., Cheng, Y.T., Lee, E., Huang, Y., Dong, O.X., Gannon, P., Huang, S., Ding, P., Li, Y., Sack, F., Zhang, Y. & Li, X., 2010. MOS11: A new component in the mRNA export pathway. *PLoS Genetics*, 6(12), pp.1–9.

Gibson, T.J., Seiler, M. & Veitia, R.A., 2013. The transience of transient overexpression. *Nature Methods*, 10(8), pp.715–721.

Gilbert, W. & Guthrie, C., 2004. The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA. *Molecular cell*, 13(2), pp.201–12.

Göhring, J., Jacak, J. & Barta, A., 2014. Imaging of endogenous messenger RNA splice variants in living cells reveals nuclear retention of transcripts inaccessible to nonsense-mediated decay in Arabidopsis. *The Plant cell*, 26(2), pp.754–64.

Golovanov, A.P., Hautbergue, G.M., Tintaru, A.M., Lian, L.-Y. & Wilson, S.A., 2006. The solution structure of REF2-I reveals interdomain interactions and regions involved in binding mRNA export factors and RNA. *RNA (New York, N.Y.)*, 12(11), pp.1933–48.

Gómez-González, B., García-Rubio, M., Bermejo, R., Gaillard, H., Shirahige, K., Marín, A., Foiani, M. & Aguilera, A., 2011. Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. *The EMBO journal*, 30(15), pp.3106–19.

Gonatopoulos-Pournatzis, T. & Cowling, V.H., 2014. Cap-binding complex (CBC). *The Biochemical journal*, 457(2), pp.231–42.

Gong, Z., Dong, C.-H., Lee, H., Zhu, J., Xiong, L., Gong, D., Stevenson, B. & Zhu, J.-K., 2005. A DEAD box RNA helicase is essential for mRNA export and important for development and stress responses in Arabidopsis. *The Plant cell*, 17(1), pp.256–67.

Görlich, D. & Kutay, U., 1999. Transport between the cell nucleus and the cytoplasm. *Annual review of cell and developmental biology*, 15, pp.607–60.

Görnemann, J., Kotovic, K.M., Hujer, K. & Neugebauer, K.M., 2005. Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Molecular cell*, 19(1), pp.53–63.

Gouy, M., Guindon, S. & Gascuel, O., 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular biology and evolution*, 27(2), pp.221–4.

Gromadzka, A.M., Steckelberg, A.-L., Singh, K.K., Hofmann, K. & Gehring, N.H., 2016. A short conserved motif in ALYREF directs cap- and EJC-dependent assembly of export complexes on spliced mRNAs. *Nucleic acids research*.

Grünwald, D., Singer, R.H. & Rout, M., 2011. Nuclear export dynamics of RNA-protein complexes. *Nature*, 475(7356), pp.333–41.

Grüter, P., Tabernero, C., von Kobbe, C., Schmitt, C., Saavedra, C., Bachi, A., Wilm, M., Felber, B.K. & Izaurralde, E., 1998. TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Molecular cell*, 1(5), pp.649–59.

Hajheidari, M., Koncz, C. & Eick, D., 2013. Emerging roles for RNA polymerase II CTD in Arabidopsis. *Trends in plant science*, 18(11), pp.633–43.

Hartmuth, K., Urlaub, H., Vornlocher, H.-P., Will, C.L., Gentzel, M., Wilm, M. & Lührmann, R., 2002. Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *Proceedings of the National Academy of Sciences of the United States of America*, 99(26), pp.16719–24.

Hautbergue, G.M., Hung, M.-L., Golovanov, A.P., Lian, L.-Y. & Wilson, S. a, 2008. Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proceedings of the National Academy of Sciences of the United States of America*, 105(13), pp.5154–5159.

Hautbergue, G.M., Hung, M.L., Walsh, M.J., Snijders, A.P.L., Chang, C. Te, Jones, R., Ponting, C.P., Dickman, M.J. & Wilson, S. a., 2009. UIF, a New mRNA Export Adaptor that Works Together with REF/ALY, Requires FACT for Recruitment to mRNA. *Current Biology*, 19(22), pp.1918–1924.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J., 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome biology*, 8(2), p.R19.

Hellmuth, K., Lau, D.M., Bischoff, F.R., Künzler, M., Hurt, E. & Simos, G., 1998. Yeast Los1p Has Properties of an Exportin-Like Nucleocytoplasmic Transport Factor for tRNA. *Molecular and Cellular Biology*, 18(11), pp.6374–6386.

Herold, A., Klymenko, T. & Izaurralde, E., 2001. NXF1/p15 heterodimers are essential for mRNA nuclear export in Drosophila. *RNA (New York, N.Y.)*, 7(12), pp.1768–80.

Herold, N., Will, C.L., Wolf, E., Kastner, B., Urlaub, H. & Lührmann, R., 2009. Conservation of the protein composition and electron microscopy structure of Drosophila melanogaster and human spliceosomal complexes. *Molecular and cellular biology*, 29(1), pp.281–301.

Le Hir, H., Gatfield, D., Izaurralde, E. & Moore, M.J., 2001. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *The EMBO journal*, 20(17), pp.4987–97.

Le Hir, H., Izaurralde, E., Maquat, L.E. & Moore, M.J., 2000. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *The EMBO journal*, 19(24), pp.6860–9.

Hocine, S., Singer, R.H. & Grünwald, D., 2010. RNA processing and export. *Cold Spring Harbor perspectives in biology*, 2(12), p.a000752.

Hodge, C.A., Tran, E.J., Noble, K.N., Alcazar-Roman, A.R., Ben-Yishay, R., Scarcelli, J.J., Folkmann, A.W., Shav-Tal, Y., Wente, S.R. & Cole, C.N., 2011. The Dbp5 cycle at the nuclear pore complex during mRNA export I: dbp5 mutants with defects in RNA binding and ATP hydrolysis define key steps for Nup159 and Gle1. *Genes & development*, 25(10), pp.1052–64.

Hogg, R., McGrail, J.C. & O'Keefe, R.T., 2010. The function of the NineTeen Complex (NTC) in regulating spliceosome conformations and fidelity during pre-mRNA splicing. *Biochemical Society transactions*, 38(4), pp.1110–5.

Holtorf, S., Apel, K. & Bohlmann, H., 1995. Comparison of different constitutive and inducible promoters for the overexpression of transgenes in Arabidopsis thaliana. *Plant Molecular Biology*, 29(4), pp.637–646.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W. & Zimmermann, P., 2008. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Advances in bioinformatics*, 2008, p.420747.

Huang, M., Rech, J.E., Northington, S.J., Flicker, P.F., Mayeda, A., Krainer, A.R. & LeStourgeon, W.M., 1994. The C-protein tetramer binds 230 to 240 nucleotides of pre-mRNA and nucleates the assembly of 40S heterogeneous nuclear ribonucleoprotein particles. *Molecular and cellular biology*, 14(1), pp.518–33.

Huertas, P. & Aguilera, A., 2003. Cotranscriptionally Formed DNA:RNA Hybrids Mediate Transcription Elongation Impairment and Transcription-Associated Recombination. *Molecular Cell*, 12(3), pp.711–721.

Hülsmann, B.B., Labokha, A.A. & Görlich, D., 2012. The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. *Cell*, 150(4), pp.738–51.

Hung, M.L., Hautbergue, G.M., Snijders, a. P.L., Dickman, M.J. & Wilson, S. a., 2010. Arginine methylation of REF/ALY promotes efficient handover of mRNA to TAP/NXF1. *Nucleic Acids Research*, 38(10), pp.3351–3361.

Iglesias, N., Tutucci, E., Gwizdek, C., Vinciguerra, P., Von Dach, E., Corbett, A.H., Dargemont, C. & Stutz, F., 2010. Ubiquitin-mediated mRNP dynamics and surveillance prior to budding yeast mRNA export. *Genes & development*, 24(17), pp.1927–38.

Jacob, Y., Mongkolsiriwatana, C., Veley, K.M., Kim, S.Y. & Michaels, S.D., 2007. The nuclear pore protein AtTPR is required for RNA homeostasis, flowering time, and auxin signaling. *Plant physiology*, 144(3), pp.1383–90.

Jani, D., Lutz, S., Hurt, E., Laskey, R.A., Stewart, M. & Wickramasinghe, V.O., 2012. Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export. *Nucleic acids research*, 40(10), pp.4562–73.

Jani, D., Valkov, E. & Stewart, M., 2014. Structural basis for binding the TREX2 complex to nuclear pores, GAL1 localisation and mRNA export. *Nucleic acids research*, 42(10), pp.6686–97.

Jauvion, V., Elmayan, T. & Vaucheret, H., 2010. The conserved RNA trafficking proteins HPR1 and TEX1 are involved in the production of endogenous and exogenous small interfering RNA in Arabidopsis. *The Plant cell*, 22(8), pp.2697–2709.

Jensen, T.H., Boulay, J., Rosbash, M. & Libri, D., 2001. The DECD box putative ATPase Sub2p is an early mRNA export factor. *Current Biology*, 11(21), pp.1711–1715.

Jimeno, S., Luna, R., García-Rubio, M. & Aguilera, A., 2006. Tho1, a novel hnRNP, and Sub2 provide alternative pathways for mRNP biogenesis in yeast THO mutants. *Molecular and cellular biology*, 26(12), pp.4387–98.

Jimeno, S., Rondón, A.G., Luna, R. & Aguilera, A., 2002. The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *The EMBO journal*, 21(13), pp.3526–35.

Johnson, S.A., Cubberley, G. & Bentley, D.L., 2009. Cotranscriptional Recruitment of the mRNA Export Factor Yra1 by Direct Interaction with the 3??? End Processing Factor Pcf11. *Molecular Cell*, 33(2), pp.215–226.

Jurica, M.S., Licklider, L.J., Gygi, S.R., Grigorieff, N. & Moore, M.J., 2002. Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA*, 8(4), pp.426–439.

Kabachinski, G. & Schwartz, T.U., 2015. The nuclear pore complex - structure and function at a glance. *Journal of Cell Science*, 128(3), pp.423–429.

Kammel, C., Thomaier, M., Sørensen, B.B., Schubert, T., Längst, G., Grasser, M. & Grasser, K.D., 2013. Arabidopsis DEAD-box RNA helicase UAP56 interacts with both RNA and DNA as well as with mRNA export factors. *PloS one*, 8(3), p.e60644.

Kapadia, F., Pryor, A., Chang, T.-H. & Johnson, L.F., 2006. Nuclear localization of poly(A)+ mRNA following siRNA reduction of expression of the mammalian RNA helicases UAP56 and URH49. *Gene*, 384, pp.37–44.

Katahira, J., 2015. Nuclear export of messenger RNA. *Genes*, 6(2), pp.163–84.

Katahira, J., Dimitrova, L., Imai, Y. & Hurt, E., 2015. NTF2-like domain of Tap plays a critical role in cargo mRNA recognition and export. *Nucleic acids research*, 43(3), pp.1894–904.

Katahira, J., Inoue, H., Hurt, E. & Yoneda, Y., 2009. Adaptor Aly and co-adaptor Thoc5 function in the Tap-p15-mediated nuclear export of HSP70 mRNA. *The EMBO journal*, 28(5), pp.556–567.

Katahira, J., Okuzaki, D., Inoue, H., Yoneda, Y., Maehara, K. & Ohkawa, Y., 2013. Human TREX component Thoc5 affects alternative polyadenylation site choice by recruiting mammalian cleavage factor I. *Nucleic acids research*, 41(14), pp.7060–72.

Katahira, J., Straesser, K., Saiwaki, T., Yoneda, Y. & Hurt, E., 2002. Complex formation between Tap and p15 affects binding to FG-repeat nucleoporins and nucleocytoplasmic shuttling. *The Journal of biological chemistry*, 277(11), pp.9242–6.

Katahira, J., Strässer, K., Podtelejnikov, A., Mann, M., Jung, J.U. & Hurt, E., 1999. The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *The EMBO journal*, 18(9), pp.2593–609.

Keller, R.W., Kühn, U., Aragón, M., Bornikova, L., Wahle, E. & Bear, D.G., 2000. The nuclear poly(A) binding protein, PABP2, forms an oligomeric particle covering the length of the poly(A) tail. *Journal of molecular biology*, 297(3), pp.569–83.

Kistler, A.L. & Guthrie, C., 2001. Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase. *Genes & development*, 15(1), pp.42–9.

Köhler, A. & Hurt, E., 2007. Exporting RNA from the nucleus to the cytoplasm. *Nature reviews. Molecular cell biology*, 8(10), pp.761–73.

Koncz, C., deJong, F., Villacorta, N., Szakonyi, D. & Koncz, Z., 2012. The Spliceosome-Activating Complex: Molecular Mechanisms Underlying the Function of a Pleiotropic Regulator. *Frontiers in Plant Science*, 3(January), pp.1–12.

König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M. & Ule, J., 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature structural & molecular biology*, 17(7), pp.909–15.

Labokha, A.A., Gradmann, S., Frey, S., Hülsmann, B.B., Urlaub, H., Baldus, M. & Görlich, D., 2013. Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes. *The EMBO journal*, 32(2), pp.204–18.

Lee, H.-S., Lee, D.-H., Cho, H.K., Kim, S.H., Auh, J.H. & Pai, H.-S., 2015. InsP6-sensitive variants of the Gle1 mRNA export factor rescue growth and fertility defects of the ipk1 low-phytic-acid mutation in Arabidopsis. *The Plant cell*, 27(2), pp.417–31.

Lee, Y. & Rio, D.C., 2015. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annual review of biochemistry*, 84, pp.291–323.

Van Leene, J., Eeckhout, D., Cannoot, B., De Winne, N., Persiau, G., Van De Slijke, E., Vercruysse, L., Dedecker, M., Verkest, A., Vandepoele, K., Martens, L., Witters, E., Gevaert, K. & De Jaeger, G., 2015. An improved toolbox to unravel the plant cellular machinery by tandem affinity purification of Arabidopsis protein complexes. *Nat Protoc*, 10(1), pp.169–187.

Van Leene, J., Eeckhout, D., Persiau, G., Van De Slijke, E., Geerinck, J., Van Isterdael, G., Witters, E. & De Jaeger, G., 2011. Isolation of transcription factor complexes from Arabidopsis cell suspension cultures by tandem affinity purification. *Methods in molecular biology (Clifton, N.J.)*, 754, pp.195–218.

Van Leene, J., Witters, E., Inzé, D. & De Jaeger, G., 2008. Boosting tandem affinity purification of plant protein complexes. *Trends in plant science*, 13(10), pp.517–20.

Lehner, B., Semple, J.I., Brown, S.E., Counsell, D., Campbell, R.D. & Sanderson, C.M., 2004. Analysis of a high-throughput yeast two-hybrid system and its use to predict the function of intracellular proteins encoded within the human MHC class III region. *Genomics*, 83(1), pp.153–167.

Letunic, I., Doerks, T. & Bork, P., 2015. SMART: recent updates, new developments and status in 2015. *Nucleic acids research*, 43(Database issue), pp.D257–60.

Li, Y., Wang, X., Zhang, X. & Goodrich, D.W., 2005. Human hHpr1/p84/Thoc1 regulates transcriptional elongation and physically links RNA polymerase II and RNA processing factors. *Molecular and cellular biology*, 25(10), pp.4023–33.

Libri, D., Graziani, N., Saguez, C. & Boulay, J., 2001. Multiple roles for the yeast SUB2/yUAP56 gene in splicing. *Genes & development*, 15(1), pp.36–41.

Lidschreiber, M., Leike, K. & Cramer, P., 2013. Cap completion and C-terminal repeat domain kinase recruitment underlie the initiation-elongation transition of RNA polymerase II. *Molecular and cellular biology*, 33(19), pp.3805–16.

Liu, C., Thong, Z. & Yu, H., 2009. Coming into bloom: the specification of floral meristems. *Development (Cambridge, England)*, 136(20), pp.3379–91.

Lolas, I.B., Himanen, K., Grønlund, J.T., Lynggaard, C., Houben, A., Melzer, M., Van Lijsebettens, M. & Grasser, K.D., 2010. The transcript elongation factor FACT affects Arabidopsis vegetative and reproductive development and genetically interacts with HUB1/2. *The Plant journal : for cell and molecular biology*, 61(4), pp.686–97.

Longman, D., Johnstone, I.L. & Cáceres, J.F., 2003. The Ref/Aly proteins are dispensable for mRNA export and development in Caenorhabditis elegans. *RNA (New York, N.Y.)*, 9(7), pp.881–91.

Lopato, S., Kalyna, M., Dorner, S., Kobayashi, R., Krainer, A.R. & Barta, A., 1999. atSRp30, one of two SF2/ASF-like proteins from Arabidopsis thaliana, regulates splicing of specific plant genes. *Genes & Development*, 13(8), pp.987–1001.

Lu, Q., Tang, X., Tian, G., Wang, F., Liu, K., Nguyen, V., Kohalmi, S.E., Keller, W.A., Tsang, E.W.T., Harada, J.J., Rothstein, S.J. & Cui, Y., 2010. Arabidopsis homolog of the yeast TREX-2 mRNA export complex: components and anchoring nucleoporin. *The Plant journal : for cell and molecular biology*, 61(2), pp.259–70.

Luna, R., Jimeno, S., Marín, M., Huertas, P., García-Rubio, M. & Aguilera, A., 2005. Interdependence between transcription and mRNP processing and export, and its impact on genetic stability. *Molecular cell*, 18(6), pp.711–22.

Lund, M.K. & Guthrie, C., 2005. The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim. *Molecular cell*, 20(4), pp.645–51.

Luo, M.L., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M. & Reed, R., 2001. Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature*, 413(6856), pp.644–647.

MacGregor, D.R., Gould, P., Foreman, J., Griffiths, J., Bird, S., Page, R., Stewart, K., Steel, G., Young, J., Paszkiewicz, K., Millar, a. J., Halliday, K.J., Hall, a. J. & Penfield, S., 2013. HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENES1 Is Required for Circadian Periodicity through the Promotion of Nucleo-Cytoplasmic mRNA Export in Arabidopsis. *The Plant Cell*, 25(11), pp.4391–4404.

MacKellar, A.L. & Greenleaf, A.L., 2011. Cotranscriptional association of mRNA export factor Yra1 with C-terminal domain of RNA polymerase II. *Journal of Biological Chemistry*, 286(42), pp.36385–36395.

MacMorris, M., Brocker, C. & Blumenthal, T., 2003. UAP56 levels affect viability and mRNA export in Caenorhabditis elegans. *RNA (New York, N.Y.)*, 9(7), pp.847–57.

Mancini, A., Niemann-Seyde, S.C., Pankow, R., El Bounkari, O., Klebba-Färber, S., Koch, A., Jaworska, E., Spooncer, E., Gruber, A.D., Whetton, A.D. & Tamura, T., 2010. THOC5/FMIP, an mRNA export TREX complex protein, is essential for hematopoietic primitive cell survival in vivo. *BMC biology*, 8(1), p.1.

MANDERS, E.M.M., VERBEEK, F.J. & ATEN, J.A., 1993. Measurement of co-localization of objects in dual-colour confocal images. *Journal of Microscopy*, 169(3), pp.375–382.

Marchler-Bauer, A. et al., 2014. CDD: NCBI's conserved domain database. *Nucleic acids research*, 43(Database issue), pp.D222–6.

Marin, E., Jouannet, V., Herz, A., Lokerse, A.S., Weijers, D., Vaucheret, H., Nussaume, L., Crespi, M.D. & Maizel, A., 2010. miR390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *The Plant cell*, 22(4), pp.1104–17.

Mason, P.B. & Struhl, K., 2005. Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo. *Molecular cell*, 17(6), pp.831–40.

Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N. & Reed, R., 2005. Recruitment of the human TREX complex to mRNA during splicing. *Genes and Development*, 19(13), pp.1512–1517.

McAfee, J.G., Shahied-Milam, L., Soltaninassab, S.R. & LeStourgeon, W.M., 1996. A major determinant of hnRNP C protein binding to RNA is a novel bZIP-like RNA binding domain. *RNA (New York, N.Y.)*, 2(11), pp.1139–52.

McAfee, J.G., Soltaninassab, S.R., Lindsay, M.E. & LeStourgeon, W.M., 1996. Proteins C1 and C2 of heterogeneous nuclear ribonucleoprotein complexes bind RNA in a highly cooperative fashion: support for their contiguous deposition on pre-mRNA during transcription. *Biochemistry*, 35(4), pp.1212–22.

McCloskey, A., Taniguchi, I., Shinmyozu, K. & Ohno, M., 2012. hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science (New York, N.Y.)*, 335(6076), pp.1643–6.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P. & Lopez, R., 2013. Analysis Tool Web Services from the EMBL-EBI. *Nucleic acids research*, 41(Web Server issue), pp.W597–600.

Meinel, D.M., Burkert-Kautzsch, C., Kieser, A., O'Duibhir, E., Siebert, M., Mayer, A., Cramer, P., Söding, J., Holstege, F.C.P. & Sträßer, K., 2013. Recruitment of TREX to the Transcription Machinery by Its Direct Binding to the Phospho-CTD of RNA Polymerase II. *PLoS Genetics*, 9(11).

Mohr, D., Frey, S., Fischer, T., Güttler, T. & Görlich, D., 2009. Characterisation of the passive permeability barrier of nuclear pore complexes. *The EMBO journal*, 28(17), pp.2541–53.

Möller-Steinbach, Y., Alexandre, C. & Hennig, L., 2010. Flowering time control. *Methods in molecular biology (Clifton, N.J.)*, 655, pp.229–37.

Moore, M.J. & Proudfoot, N.J., 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell*, 136(4), pp.688–700.

Mor, A., Suliman, S., Ben-Yishay, R., Yunger, S., Brody, Y. & Shav-Tal, Y., 2010. Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. *Nature cell biology*, 12(6), pp.543–52.

Mouradov, A., Cremer, F. & Coupland, G., 2002. Control of flowering time: interacting pathways as a basis for diversity. *The Plant cell*, 14 Suppl, pp.S111–30.

Müller-McNicoll, M. & Neugebauer, K.M., 2013. How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nature reviews. Genetics*, 14(4), pp.275–87.

Noble, K.N., Tran, E.J., Alcazar-Roman, A.R., Hodge, C.A., Cole, C.N. & Wente, S.R., 2011. The Dbp5 cycle at the nuclear pore complex during mRNA export II: nucleotide cycling and mRNP remodeling by Dbp5 are controlled by Nup159 and Gle1. *Genes & Development*, 25(10), pp.1065–1077.

Nojima, T., Hirose, T., Kimura, H. & Hagiwara, M., 2007. The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. *The Journal of biological chemistry*, 282(21), pp.15645–51.

Okada, K., Ueda, J., Komaki, M.K., Bell, C.J. & Shimura, Y., 1991. Requirement of the Auxin Polar Transport System in Early Stages of Arabidopsis Floral Bud Formation. *The Plant cell*, 3(7), pp.677–684.

Okamura, M., Inose, H. & Masuda, S., 2015. RNA Export through the NPC in Eukaryotes. *Genes*, 6(1), pp.124–49.

Pall, G.S. & Hamilton, A.J., 2008. Improved northern blot method for enhanced detection of small RNA. *Nature protocols*, 3(6), pp.1077–84.

Pan, H., Liu, S. & Tang, D., 2012. HPR1, a component of the THO/TREX complex, plays an important role in disease resistance and senescence in Arabidopsis. *Plant Journal*, 69(5), pp.831–843.

Pan, H., Liu, S. & Tang, D., 2014. The THO/TREX complex functions in disease resistance in Arabidopsis. *Plant Signaling & Behavior*, 7(3), pp.422–424.

Papasaikas, P. & Valcárcel, J., 2015. The Spliceosome: The Ultimate RNA Chaperone and Sculptor. *Trends in biochemical sciences*, 41(1), pp.33–45.

Parry, G., Ward, S., Cernac, A., Dharmasiri, S. & Estelle, M., 2006. The Arabidopsis SUPPRESSOR OF AUXIN RESISTANCE proteins are nucleoporins with an important role in hormone signaling and development. *The Plant cell*, 18(7), pp.1590–603.

Pautler, M., Tanaka, W., Hirano, H.-Y. & Jackson, D., 2013. Grass meristems I: shoot apical meristem maintenance, axillary meristem determinacy and the floral transition. *Plant & cell physiology*, 54(3), pp.302–12.

Peña, Á., Gewartowski, K., Mroczek, S., Cuéllar, J., Szykowska, A., Prokop, A., Czarnocki-Cieciura, M., Piwowarski, J., Tous, C., Aguilera, A., Carrascosa, J.L., Valpuesta, J.M. & Dziembowski, A., 2012. Architecture and nucleic acids recognition mechanism of the THO complex, an mRNP assembly factor. *The EMBO Journal*, 31(6), pp.1605–1616.

Pendle, A.F., Clark, G.P., Boon, R., Lewandowska, D., Lam, Y.W., Andersen, J., Mann, M., Lamond, A.I., Brown, J.W.S. & Shaw, P.J., 2005. Proteomic analysis of the Arabidopsis nucleolus suggests novel nucleolar functions. *Molecular biology of the cell*, 16(1), pp.260–9.

Perazza, D., Herzog, M., Hülskamp, M., Brown, S., Dorne, A.M. & Bonneville, J.M., 1999. Trichome cell growth in Arabidopsis thaliana can be derepressed by mutations in at least five genes. *Genetics*, 152(1), pp.461–76.

Pérez-Pérez, J.M., Candela, H. & Micol, J.L., 2009. Understanding synergy in genetic interactions. *Trends in genetics : TIG*, 25(8), pp.368–76.

Petersen, T.N., Brunak, S., von Heijne, G. & Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods*, 8(10), pp.785–6.

Piruat, J.I. & Aguilera, A., 1998. A novel yeast gene, THO2, is involved in RNA pol II transcription and provides new evidence for transcriptional elongation-associated recombination. *The EMBO journal*, 17(16), pp.4859–72.

Proudfoot, N., 2004. New perspectives on connecting messenger RNA 3' end formation to transcription. *Current opinion in cell biology*, 16(3), pp.272–8.

Pryor, A., Tung, L., Yang, Z., Kapadia, F., Chang, T.-H. & Johnson, L.F., 2004. Growth-regulated expression and G0-specific turnover of the mRNA that encodes URH49, a mammalian DExH/D box protein that is highly related to the mRNA export protein UAP56. *Nucleic acids research*, 32(6), pp.1857–65.

R Core Team, 2015. *R: A language and environment for statistical computing.*, Vienna, Austria.

Rappsilber, J., Ryder, U., Lamond, A.I. & Mann, M., 2002. Large-scale proteomic analysis of the human spliceosome. *Genome research*, 12(8), pp.1231–45.

Reddy, A.S.N., Day, I.S., Göhring, J. & Barta, A., 2012. Localization and dynamics of nuclear speckles in plants. *Plant physiology*, 158(1), pp.67–77.

Rehwinkel, J., Herold, A., Gari, K., Köcher, T., Rode, M., Ciccarelli, F.L., Wilm, M. & Izaurralde, E., 2004. Genome-wide analysis of mRNAs regulated by the THO complex in Drosophila melanogaster. *Nature structural & molecular biology*, 11(6), pp.558–66.

Reichelt, R., Holzenburg, A., Buhle, E.L., Jarnik, M., Engel, A. & Aebi, U., 1990. Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *The Journal of cell biology*, 110(4), pp.883–94.

Ribbeck, K., 1998. NTF2 mediates nuclear import of Ran. *The EMBO Journal*, 17(22), pp.6587–6598.

Ribbeck, K. & Görlich, D., 2001. Kinetic analysis of translocation through nuclear pore complexes. *The EMBO journal*, 20(6), pp.1320–30.

Rodrigues, J.P., Rode, M., Gatfield, D., Blencowe, B.J., Carmo-Fonseca, M. & Izaurralde, E., 2001. REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, 98(3), pp.1030–5.

Rodríguez-Navarro, S., Fischer, T., Luo, M.-J., Antúnez, O., Brettschneider, S., Lechner, J., Pérez-Ortín, J.E., Reed, R. & Hurt, E., 2004. Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell*, 116(1), pp.75–86.

Rondón, A.G., Jimeno, S., García-Rubio, M. & Aguilera, A., 2003. Molecular evidence that the eukaryotic THO/TREX complex is required for efficient transcription elongation. *The Journal of biological chemistry*, 278(40), pp.39037–43.

Rout, M.P., Aitchison, J.D., Suprapto, A., Hjertaas, K., Zhao, Y. & Chait, B.T., 2000. The yeast nuclear pore complex: composition, architecture, and transport mechanism. *The Journal of cell biology*, 148(4), pp.635–51.

Ruepp, M.-D., Aringhieri, C., Vivarelli, S., Cardinale, S., Paro, S., Schümperli, D. & Barabino, S.M.L., 2009. Mammalian pre-mRNA 3' end processing factor CF I m 68 functions in mRNA export. *Molecular biology of the cell*, 20(24), pp.5211–23.

Ruepp, M.D., Schümperli, D. & Barabino, S.M.L., 2011. mRNA 3′ end processing and more-multiple functions of mammalian cleavage factor I-68. *Wiley Interdisciplinary Reviews: RNA*, 2(1), pp.79–91.

Santos-Rosa, H., Moreno, H., Simos, G., Segref, A., Fahrenkrog, B., Panté, N. & Hurt, E., 1998. Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores. *Molecular and cellular biology*, 18(11), pp.6826–38.

Schmid, M. & Jensen, T.H., 2008. Quality control of mRNP in the nucleus. *Chromosoma*, 117(5), pp.419–29.

Schmidt, H.B. & Görlich, D., 2015. Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. *Trends in biochemical sciences*, 41(1), pp.46–61.

Schmitt, C., von Kobbe, C., Bachi, A., Panté, N., Rodrigues, J.P., Boscheron, C., Rigaut, G., Wilm, M., Séraphin, B., Carmo-Fonseca, M. & Izaurralde, E., 1999. Dbp5, a DEAD-box protein required for mRNA export, is recruited to the cytoplasmic fibrils of nuclear pore complex via a conserved interaction with CAN/Nup159p. *The EMBO journal*, 18(15), pp.4332–47.

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), pp.671–675.

Schoenberg, D.R. & Maquat, L.E., 2012. Regulation of cytoplasmic mRNA decay. *Nature reviews. Genetics*, 13(4), pp.246–59.

Schroeder, S.C., 2000. Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes & Development*, 14(19), pp.2435–2440.

Schwab, B., Folkers, U., Ilgenfritz, H. & Hülskamp, M., 2000. Trichome morphogenesis in Arabidopsis. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 355(1399), pp.879–83.

Schwer, B., Mao, X. & Shuman, S., 1998. Accelerated mRNA decay in conditional mutants of yeast mRNA capping enzyme. *Nucleic acids research*, 26(9), pp.2050–7.

Segref, A., Sharma, K., Doye, V., Hellwig, A., Huber, J., Lührmann, R. & Hurt, E., 1997. Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. *The EMBO journal*, 16(11), pp.3256–71.

Shi, Y. & Manley, J.L., 2015. The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. *Genes & development*, 29(9), pp.889–97.

Simpson, G.G. & Dean, C., 2002. Arabidopsis, the Rosetta stone of flowering time? *Science (New York, N.Y.)*, 296(5566), pp.285–9.

Singh, G., Kucukural, A., Cenik, C., Leszyk, J.D., Shaffer, S.A., Weng, Z. & Moore, M.J., 2012. The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell*, 151(4), pp.750–64.

Sloan, K.E., Gleizes, P.-E. & Bohnsack, M.T., 2015. Nucleocytoplasmic Transport of RNAs and RNA-Protein Complexes. *Journal of molecular biology*.

Snay-Hodge, C.A., Colot, H. V, Goldstein, A.L. & Cole, C.N., 1998. Dbp5p/Rat8p is a yeast nuclear pore-associated DEAD-box protein essential for RNA export. *The EMBO journal*, 17(9), pp.2663–76.

Sørensen, B.B., Ehrnsberger, H., Esposito, S., Pfab, A., Meister, G., Merkl, R., Schubert, T. & Längst, G., 2016. TEX1 of the Arabidopsis THO / TREX complex plays a role at different steps of the biogenesis of mRNAs and small RNAs. *In preparation*, pp.1–24.

Stevens, B.J. & Swift, H., 1966. RNA transport from nucleus to cytoplasm in Chironomus salivary glands. *The Journal of cell biology*, 31(1), pp.55–77.

Stewart, M., 2007. Molecular mechanism of the nuclear protein import cycle. *Nature reviews. Molecular cell biology*, 8(3), pp.195–208.

Stirnimann, C.U., Petsalaki, E., Russell, R.B. & Müller, C.W., 2010. WD40 proteins propel cellular networks. *Trends in biochemical sciences*, 35(10), pp.565–74.

Storozhenko, S., Inzé, D., Van Montagu, M. & Kushnir, S., 2001. Arabidopsis coactivator ALY-like proteins, DIP1 and DIP2, interact physically with the DNA-binding domain of the Zn-finger poly(ADP-ribose) polymerase. *Journal of experimental botany*, 52(359), pp.1375–1380.

Strässer, K., Bassler, J. & Hurt, E., 2000. Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export. *The Journal of cell biology*, 150(4), pp.695–706.

Strässer, K. & Hurt, E., 2001. Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature*, 413(6856), pp.648–52.

Strässer, K. & Hurt, E., 2000. Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *The EMBO journal*, 19(3), pp.410–420.

Sträßer, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S., Rondón, A.G., Aguilera, A., Struhl, K., Reed, R. & Hurt, E., 2002. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature*, 417(6886), pp.304–308.

Stubbs, S.H. & Conrad, N.K., 2015. Depletion of REF/Aly alters gene expression and reduces RNA polymerase II occupancy. *Nucleic acids research*, 43(1), pp.504–19.

Stutz, F., Bachi, a, Doerks, T., Braun, I.C., Séraphin, B., Wilm, M., Bork, P. & Izaurralde, E., 2000. REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA (New York, N.Y.)*, 6(4), pp.638–650.

Sugiura, T., Sakurai, K. & Nagano, Y., 2007. Intracellular characterization of DDX39, a novel growth-associated RNA helicase. *Experimental cell research*, 313(4), pp.782–90.

Tamura, K., Fukao, Y., Iwamoto, M., Haraguchi, T. & Hara-Nishimura, I., 2010. Identification and characterization of nuclear pore complex components in Arabidopsis thaliana. *The Plant cell*, 22(12), pp.4084–97.

Tange, T.Ø., Shibuya, T., Jurica, M.S. & Moore, M.J., 2005. Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. *RNA (New York, N.Y.)*, 11(12), pp.1869–83.

Taniguchi, I. & Ohno, M., 2008. ATP-dependent recruitment of export factor Aly/REF onto intronless mRNAs by RNA helicase UAP56. *Molecular and cellular biology*, 28(2), pp.601–8.

Tantikanjana, T., Yong, J.W., Letham, D.S., Griffith, M., Hussain, M., Ljung, K., Sandberg, G. & Sundaresan, V., 2001. Control of axillary bud initiation and shoot architecture in Arabidopsis through the SUPERSHOOT gene. *Genes & development*, 15(12), pp.1577–88.

Tashiro, S., Tian, C., Watahiki, M.K. & Yamamoto, K.T., 2009. Changes in growth kinetics of stamen filaments cause inefficient pollination in massugu2, an auxin insensitive, dominant mutant of Arabidopsis thaliana. *Physiologia plantarum*, 137(2), pp.175–87.

Teng, I.-F. & Wilson, S.A., 2013. Mapping interactions between mRNA export factors in living cells. *PloS one*, 8(6), p.e67676.

Terry, L.J. & Wente, S.R., 2009. Flexible gates: dynamic topologies and functions for FG nucleoporins in nucleocytoplasmic transport. *Eukaryotic cell*, 8(12), pp.1814–27.

Topisirovic, I., Siddiqui, N., Lapointe, V.L., Trost, M., Thibault, P., Bangeranye, C., Piñol-Roma, S. & Borden, K.L.B., 2009. Molecular dissection of the eukaryotic initiation factor 4E (eIF4E) export-competent RNP. *The EMBO Journal*, 28(8), pp.1087–1098.

Tran, E.J., Zhou, Y., Corbett, A.H. & Wente, S.R., 2007. The DEAD-box protein Dbp5 controls mRNA export by triggering specific RNA:protein remodeling events. *Molecular cell*, 28(5), pp.850–9.

Tseng, S.S., Weaver, P.L., Liu, Y., Hitomi, M., Tartakoff, A.M. & Chang, T.H., 1998. Dbp5p, a cytosolic RNA helicase, is required for poly(A)+ RNA export. *The EMBO journal*, 17(9), pp.2651–62.

Tsugama, D., Liu, S. & Takano, T., 2011. A rapid chemical method for lysing Arabidopsis cells for protein analysis. *Plant methods*, 7, p.22.

Tutucci, E. & Stutz, F., 2011. Keeping mRNPs in check during assembly and nuclear export. *Nature reviews. Molecular cell biology*, 12(6), pp.377–84.

Uhrig, J.F., Canto, T., Marshall, D. & MacFarlane, S. a, 2004. Relocalization of nuclear ALY proteins to the cytoplasm by the tomato bushy stunt virus P19 pathogenicity protein. *Plant physiology*, 135(4), pp.2411–2423.

Umlauf, D., Bonnet, J., Waharte, F., Fournier, M., Stierle, M., Fischer, B., Brino, L., Devys, D. & Tora, L., 2013. The human TREX-2 complex is stably associated with the nuclear pore basket. *Journal of cell science*, 126(Pt 12), pp.2656–67.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. & Rozen, S.G., 2012. Primer3--new capabilities and interfaces. *Nucleic acids research*, 40(15), p.e115.

Valencia, P., Dias, A.P. & Reed, R., 2008. Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, 105(9), pp.3386–91.

Vernoux, T., Kronenberger, J., Grandjean, O., Laufs, P. & Traas, J., 2000. PIN-FORMED 1 regulates cell fate at the periphery of the shoot apical meristem. *Development*, 127(23), pp.5157–5165.

Viphakone, N., Cumberbatch, M.G., Livingstone, M.J., Heath, P.R., Dickman, M.J., Catto, J.W. & Wilson, S.A., 2015. Luzp4 defines a new mRNA export pathway in cancer cells. *Nucleic acids research*, 43(4), pp.2353–66.

Viphakone, N., Hautbergue, G.M., Walsh, M., Chang, C.-T., Holland, A., Folco, E.G., Reed, R. & Wilson, S. a., 2012. TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nature Communications*, 3, p.1006.

Watson, M.L., 1954. Pores in the mammalian nuclear membrane. *Biochimica et Biophysica Acta*, 15(4), pp.475–479.

Wein, S.P., Côté, R.G., Dumousseau, M., Reisinger, F., Hermjakob, H. & Vizcaíno, J.A., 2012. Improvements in the Protein Identifier Cross-Reference service. *Nucleic acids research*, 40(Web Server issue), pp.W276–80.

Weirich, C.S., Erzberger, J.P., Flick, J.S., Berger, J.M., Thorner, J. & Weis, K., 2006. Activation of the DExD/H-box protein Dbp5 by the nuclear-pore protein Gle1 and its coactivator InsP6 is required for mRNA export. *Nature cell biology*, 8(7), pp.668–76.

Weis, K., 2003. Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. *Cell*, 112(4), pp.441–51.

Wellmer, F., Graciet, E. & Riechmann, J.L., 2014. Specification of floral organs in Arabidopsis. *Journal of experimental botany*, 65(1), pp.1–9.

Wente, S.R. & Rout, M.P., 2010. The nuclear pore complex and nuclear transport. *Cold Spring Harbor perspectives in biology*, 2(10), p.a000562.

Wickramasinghe, V.O., Andrews, R., Ellis, P., Langford, C., Gurdon, J.B., Stewart, M., Venkitaraman, A.R. & Laskey, R.A., 2014. Selective nuclear export of specific classes of mRNA from mammalian nuclei is promoted by GANP. *Nucleic acids research*, 42(8), pp.5059–71.

Wickramasinghe, V.O., McMurtrie, P.I.A., Mills, A.D., Takei, Y., Penrhyn-Lowe, S., Amagase, Y., Main, S., Marr, J., Stewart, M. & Laskey, R.A., 2010. mRNA export from mammalian cell nuclei is dependent on GANP. *Current biology : CB*, 20(1), pp.25–31.

Wickramasinghe, V.O., Stewart, M. & Laskey, R.A., 2010. GANP enhances the efficiency of mRNA nuclear export in mammalian cells. *Nucleus (Austin, Tex.)*, 1(5), pp.393–6.

Wiermer, M., Cheng, Y.T., Imkampe, J., Li, M., Wang, D., Lipka, V. & Li, X., 2012. Putative members of the Arabidopsis Nup107-160 nuclear pore sub-complex contribute to pathogen defense. *The Plant journal : for cell and molecular biology*, 70(5), pp.796–808.

Will, C.L. & Lührmann, R., 2011. Spliceosome structure and function. *Cold Spring Harbor perspectives in biology*, 3(7).

Xu, C., Zhou, X. & Wen, C.-K., 2015. HYPER RECOMBINATION1 of the THO/TREX Complex Plays a Role in Controlling Transcription of the REVERSION-TO-ETHYLENE SENSITIVITY1 Gene in Arabidopsis. *PLoS Genetics*, 11(2), pp.1–26.

Xu, X.M., Rose, A., Muthuswamy, S., Jeong, S.Y., Venkatakrishnan, S., Zhao, Q. & Meier, I., 2007. NUCLEAR PORE ANCHOR, the Arabidopsis homolog of Tpr/Mlp1/Mlp2/megator, is involved in mRNA export and SUMO homeostasis and affects diverse aspects of plant development. *The Plant cell*, 19(5), pp.1537–48.

Yamazaki, T., Fujiwara, N., Yukinaga, H., Ebisuya, M., Shiki, T., Kurihara, T., Kioka, N., Kambe, T., Nagao, M., Nishida, E. & Masuda, S., 2010. The closely related RNA helicases, UAP56 and URH49, preferentially form distinct mRNA export machineries and coordinately regulate mitotic progression. *Molecular biology of the cell*, 21(16), pp.2953–65.

Yang, Q., Coseno, M., Gilmartin, G.M. & Doublié, S., 2011. Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. *Structure (London, England : 1993)*, 19(3), pp.368–77.

Yelina, N.E., Smith, L.M., Jones, A.M.E., Patel, K., Kelly, K. a & Baulcombe, D.C., 2010. Putative Arabidopsis THO/TREX mRNA export complex is involved in transgene and endogenous siRNA biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 107(31), pp.13948–13953.

Yoshikawa, M., 2013. Biogenesis of trans-acting siRNAs, endogenous secondary siRNAs in plants. *Genes & genetic systems*, 88(2), pp.77–84.

Zenklusen, D., Vinciguerra, P., Strahm, Y. & Stutz, F., 2001. The yeast hnRNP-Like proteins Yra1p and Yra2p participate in mRNA export through interaction with Mex67p. *Molecular and cellular biology*, 21(13), pp.4219–32.

Zenklusen, D., Vinciguerra, P., Wyss, J.-C. & Stutz, F., 2002. Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. *Molecular and cellular biology*, 22(23), pp.8241–53.

Zhai, Z., Jung, H.-I. & Vatamaniuk, O.K., 2009. Isolation of protoplasts from tissues of 14-day-old seedlings of Arabidopsis thaliana. *Journal of visualized experiments : JoVE*, (30).

Zhang, M. & Green, M.R., 2001. Identification and characterization of yUAP/Sub2p, a yeast homolog of the essential human pre-mRNA splicing factor hUAP56. *Genes & development*, 15(1), pp.30–5.

Zhou, Z., Licklider, L.J., Gygi, S.P. & Reed, R., 2002. Comprehensive proteomic analysis of the human spliceosome. *Nature*, 419(6903), pp.182–5.

Zhou, Z., Luo, M.J., Straesser, K., Katahira, J., Hurt, E. & Reed, R., 2000. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature*, 407(6802), pp.401–5.

Zourelidou, M., Absmanner, B., Weller, B., Barbosa, I.C.R., Willige, B.C., Fastner, A., Streit, V., Port, S.A., Colcombet, J., de la Fuente van Bentem, S., Hirt, H., Kuster, B., Schulze, W.X., Hammes, U.Z. & Schwechheimer, C., 2014. Auxin efflux by PIN-FORMED proteins is activated by two different protein kinases, D6 PROTEIN KINASE and PINOID. *eLife*, 3.

7 Appendix

7.1 Statistical analysis

This section contains all one and two-way ANOVA analyses done in this work. The analyses are represented as the ANOVA tables as they are obtained from the R commander package (Fox, 2005) run in the statistical environment R (R Core Team, 2015). The general workflow followed for the two-way ANOVA can be seen inFigure 46. For the two-way ANOVA the following three null hypothesis are tested 1) The population means of the first factor are equal 2) The population means of the second factor are equal and 3) There is no interaction between the two factors.

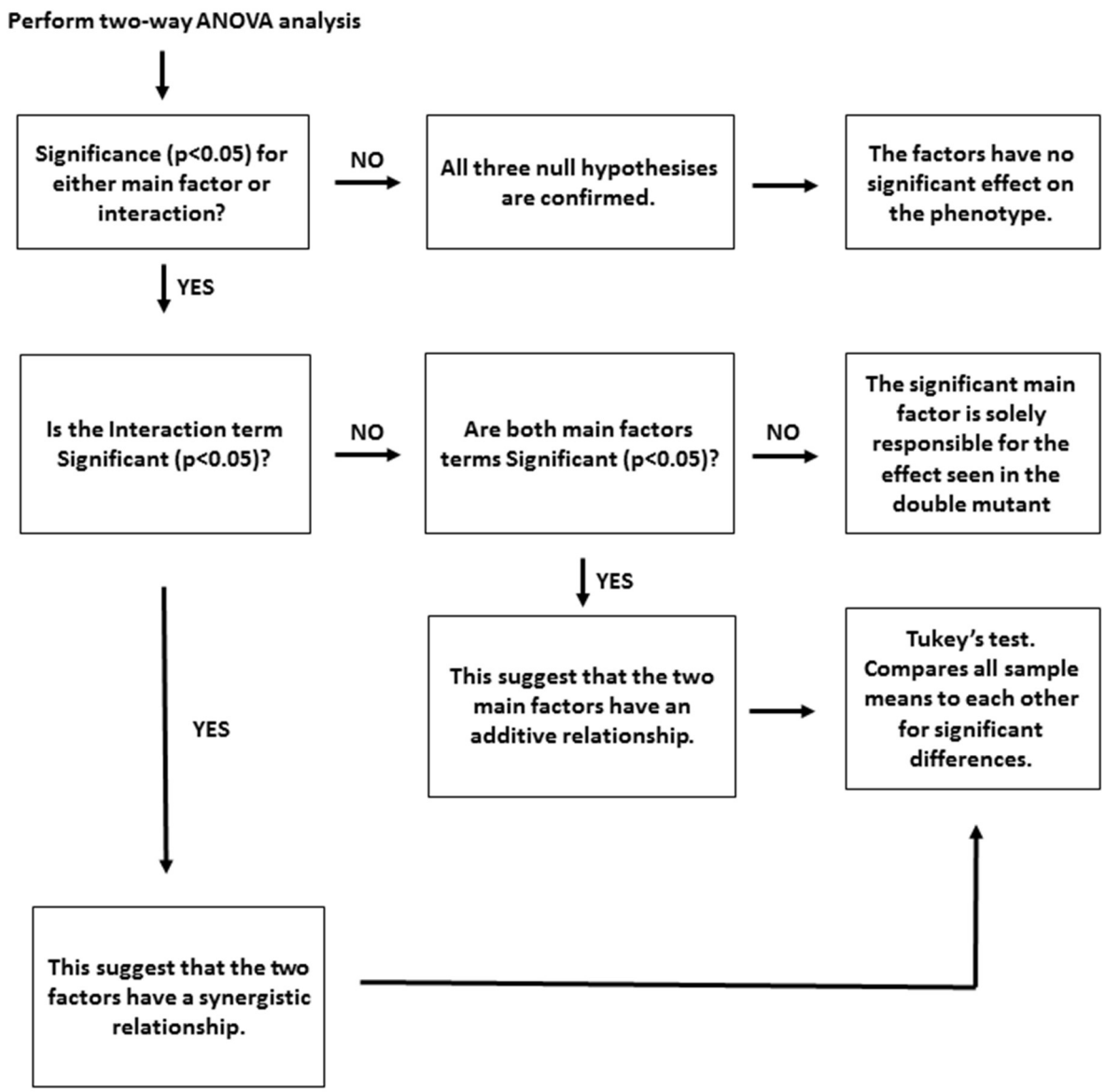


Figure 46. The work flow followed for the two-way ANOVA with two factors with two levels used in this thesis.

For each analysis, more tables are shown. The first table shows the data represented as the mean with the corresponding standard deviation. Next table shows the output of the one- or two-way ANOVA analysis. The three hypothesis tested are shown in the first column (factor one, factor two and Interaction) and the within group variation, also called the error term (Residuals). The following values are shown for each hypothesis in the table a) Degrees of freedom (DF), b) Sum of squares (Sum Sq), c) Mean squares (Mean Sq), d) F-value and e) the significance probability associated with the test F-test (Pr(>F)). In the case of significance, a table showing the outcome of the Tukey's Post HOC test is also shown next to the ANOVA table. Tukey's test is a multiple comparisons test, which compares the means of all groups to each other. For the post hoc test, an adjusted p-value is shown. The factors being tested against each other are separated by (:). In this work significance is evaluated as a significance value (p adj) or (Pr(>F)) below 0.05 and is presented by (*).

7.2 Anova tables.See chapter 7.1 for more details.

Table 20. Anova: Bolting time analysis of *aly1-1 aly2-1* (number of leaves). See Figure 18B.

	Col-0	*aly1-1*	*aly2-1*	*aly1-1 aly2-1*
Mean	15.00	15.13	14.73	14.67
Standard deviation	0.76	0.74	0.70	0.82

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Aly1	1	0.02	0.02	0.03	0.87	
Aly2	1	2.02	2.02	3.53	0.07	
Aly1:Aly2	1	0.15	0.15	0.26	0.61	
Residuals	56	32	0.57			

Table 21. Anova: mRNA export deficiency assay (N/C ratio). See Figure 19C.

	Col-0	*aly1-1*	*aly2-1*	*aly1-1 aly2-1*
Mean	1.70	1.66	1.69	1.84
Standard deviation	0.35	0.32	0.28	0.34

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
ALY1	1	0.17	0.17	1.63	0.20	
ALY2	1	0.45	0.45	4.22	0.04	*
ALY1:ALY2	1	0.58	0.57	5.4	0.02	*
Residuals	236	25.12	0.11			

Tukey's test	p adj	
aly2 : aly1 aly2	0.06	
aly1 : aly1 aly2	0.01	*
Col-0 : aly1 aly2	0.09	
aly1 : aly2	0.95	
Col-0 : aly2	1.00	
Col-0 : aly1	0.88	

Table 22. Anova: Bolting time (number of leaves) of *aly3-1 aly4-1*. See Figure 20B

	Col-0	*aly3-1*	*aly4-1*	*aly3-1 aly4-1*
Mean	14.40	12.93	13.60	11.53
Standard deviation	0.74	0.80	0.99	0.83

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
ALY3	1	46.82	46.82	65.76	$5.11 \cdot 10^{-11}$	*
ALY4	1	18.15	18.15	25.5	$5.02 \cdot 10^{-6}$	*
ALY3:ALY4	1	1.35	1.35	1.9	0.17	
Residuals	56	39.87	0.71			

Tukey's test	p adj	
aly4 : aly3 aly4	$1 \cdot 10^{-7}$	*
aly3: aly3 aly4	$1.72 \cdot 10^{-7}$	*
Col-0 : aly3 aly4	0	*
aly3 : aly4	0.15	
Col-0 : aly4	0.06	
Col-0 : aly3	8.12·10-5	*

Table 23. Anova: mRNA export deficiency assay (N/C ratio). See Figure 21B.

	wt	*aly3*	*aly4*	*aly3-1 aly4-1*
Mean	1.79	1.69	1.85	2.04
Standard deviation	0.28	0.31	0.34	0.45

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
ALY3	1	0.14	0.14	1.1	0.3	
ALY4	1	2.57	2.57	20.68	$8.68 \cdot 10^{-6}$	*
ALY3:ALY4	1	1.16	1.16	9.34	$2.5 \cdot 10^{-3}$	*
Residuals	236	29.34	0.12			

Tukey's test	p adj	
aly4 : aly3 aly4	0.02	*
aly3: aly3 aly4	$1.1 \cdot 10^{-6}$	*
Col-0 : aly3 aly4	$5.78 \cdot 10^{-4}$	*
aly3 : aly4	0.07	
Col-0 : aly4	0.72	
Col-0 : aly3	0.49	

Table 24. Anova: Bolting time- long day conditions (Number of leaves). See Figure 25B

	Col-0	*tex1*	TEX1-gfp Line1	TEX1-gfp Line2	TEX1-gfp Line3
Mean	14.00	11.62	13.17	13.55	14.27
Standard deviation	0.41	0.87	1.27	1.04	0.65

ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Complementation	4	54.53	13.63	17.18	$3.39 \cdot 10^{-9}$	*
Residuals	55	43.65	0.79			

Tukey	Pr(>|t|)	
Col-0 : TEX1-gfp Line1	0.15	
Col-0 : TEX1-gfp Line2	0.72	
Col-0 : TEX1-gfp Line3	0.94	
Col-0 : *tex1*	1·10-3	*
TEX1-gfp Line2 : TEX1-gfp Line1	0.85	
TEX1-gfp Line3 : TEX1-gfp Line1	0.03	*
tex1 : TEX1-gfp Line1	1·10-3	*
TEX1-gfp Line3 : TEX1-gfp Line2	0.32	
tex1 : TEX1-gfp Line2	1·10-3	*
tex1 : TEX1-gfp Line3	1·10-3	*

Table 25. Anova: Bolting time- short day conditions (Number of leaves). See Figure 25D

	Col-0	*tex1*	TEX1-gfp Line1	TEX1-gfp Line2	TEX1-gfp Line3
Mean	94.50	68.75	99.00	94.36	96.50
Standard deviation	8.06	4.33	17.18	10.68	6.21

ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Complementation	4	5101	1275.3	11.42	$1.92 \cdot 10^{-6}$	*
Residuals	44	4913	111.7			

Tukey's test	Pr(>|t|)	
Col-0 : TEX1-gfp Line1	0.86	
Col-0 : TEX1-gfp Line2	1.00	
Col-0 : TEX1-gfp Line3	0.99	
Col-0 : *tex1*	2.5·10-5	*
tex1 : TEX1-gfp Line1	1·10-5	*
tex1 : TEX1-gfp Line2	4.22·10-5	*
tex1 : TEX1-gfp Line3	4.2·10-5	*
TEX1-gfp Line2 : TEX1-gfp Line1	0.85	
TEX1-gfp Line3 : TEX1-gfp Line1	0.99	
TEX1-gfp Line3 : TEX1-gfp Line2	0.99	

Table 26. Anova: Relative expression of *TEX1* (P29+P30). See Figure 37B.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.08	1.39	0.08
Standard deviation	0.20	0.04	0.43	0.02

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	3.77	3.77	67.01	$3.70 \cdot 10^{-5}$	*
MOS11	1	0.12	0.12	2.07	0.19	
TEX1: MOS11	1	0.12	0.12	2.05	0.19	
Residuals	8	0.45	0.06			

Tukey's test	p adj	
mos11 : tex1 mos11	$6.28 \cdot 10^{-4}$	*
tex1 : tex1 mos11	1	
Col-0 : tex1 mos11	$6.13 \cdot 10^{-5}$	*
tex1 : mos11	$6.24 \cdot 10^{-4}$	*
Col-0 : mos11	0.25	
Col-0 : tex1	$6.1 \cdot 10^{-3}$	*

Table 27. Anova: Relative expression of *TEX1* (P31+P32). See Figure 37C

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.13	1.53	0.02
Standard deviation	0.18	0.01	0.03	0.00

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	4.23	4.23	516.85	$1.49 \cdot 10^{-8}$	*
MOS11	1	0.13	0.13	16.26	$3.80 \cdot 10^{-3}$	*
TEX1: MOS11	1	0.31	0.31	37.27	$2.90 \cdot 10^{-4}$	*
Residuals	8	0.07	0.01			

Tukey's test	p adj	
mos11 : tex1 mos11	$1 \cdot 10^{-7}$	*
tex1 : tex1 mos11	0.45	
Col-0 : tex1 mos11	$4.8 \cdot 10^{-6}$	*
tex1 : mos11	$3 \cdot 10^{-7}$	*
Col-0 : mos11	$4.36 \cdot 10^{-4}$	*
Col-0 : tex1	$1.17 \cdot 10^{-5}$	*

Table 28. Anova: Relative expression of *TEX1* (P33+P34) See Figure 37D.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.03	1.68	0.01
Standard deviation	0.21	0.02	0.39	0.00

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	5.24	5.24	107.44	$6.49 \cdot 10^{-6}$	*
MOS11	1	0.33	0.33	6.76	0.03	*
TEX1: MOS11	1	0.36	0.36	7.46	0.03	*
Residuals	8	0.39	0.05			

Tukey's test	p adj	
mos11 : tex1 mos11	$6.98 \cdot 10^{-5}$	*
tex1 : tex1 mos11	0.99	
Col-0 : tex1 mos11	$2.58 \cdot 10^{-3}$	*
tex1 : mos11	$7.51 \cdot 10^{-5}$	*
Col-0 : mos11	0.02	*
Col-0 : tex1	$2.87 \cdot 10^{-3}$	*

Table 29. Anova: Relative expression of *MOS11* (P35+P36) See Figure 37E.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.72	0.56	0.33
Standard deviation	0.16	0.31	0.06	0.03

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	0.19	0.19	5.79	0.04	*
MOS11	1	0.52	0.52	16.11	$3.90 \cdot 10^{-3}$	*
TEX1 : MOS11	1	0.00	0.00	0.06	0.81	
Residuals	8	0.26	0.03			

Tukey's test	p adj	
mos11 : tex1 mos11	0.46	
tex1 : tex1 mos11	0.11	
Col-0 : tex1 mos11	0.01	*
tex1 : mos11	0.68	
Col-0 : mos11	0.07	
Col-0 : tex1	0.31	

Table 30. Anova: Relative expression of *MOS11* (P37+P38) See Figure 37F.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.82	0.27	0.02
Standard deviation	0.25	0.31	0.07	0.01

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	0.14	0.14	3.42	0.1	
MOS11	1	1.74	1.74	42.57	$1.83 \cdot 10^{-4}$	*
TEX1: MOS11	1	0.00	0.00	0.07	0.8	
Residuals	8	0.33	0.04			

Tukey's test	p adj	
mos11 : tex1 mos11	0.48	
tex1 : tex1 mos11	$5.89 \cdot 10^{-3}$	*
Col-0 : tex1 mos11	$1.59 \cdot 10^{-3}$	*
tex1 : mos11	0.04	*
Col-0 : mos11	$9.53 \cdot 10^{-3}$	*
Col-0 : tex1	0.69	

Table 31. Anova: Relative expression of *MOS11* (P39+P40). See Figure 37G.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.98	0.10	0.07
Standard deviation	0.17	0.18	0.03	0.03

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	0.0017	0.0017	0.11	0.75	
MOS11	1	2.46	2.46	157.31	$1.53 \cdot 10^{-6}$	*
TEX1: MOS11	1	0	0	0.001	0.98	
Residuals	8	0.12	0.016			

Tukey's test	p adj	
mos11 : tex1 mos11	0.99	
tex1 : tex1 mos11	$9.44 \cdot 10^{-5}$	*
Col-0 : tex1 mos11	$7.91 \cdot 10^{-5}$	*
tex1 : mos11	$1.16 \cdot 10^{-4}$	*
Col-0 : mos11	$9.7 \cdot 10^{-5}$	*
Col-0 : tex1	0.99	

Table 32. Anova: Bolting time- long day conditions (Number of leaves). See Figure 38B.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	13.81818182	11.36363636	13	10.09090909
Standard deviation	0.750757194	0.504524979	0.632455532	0.70064905

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	79.11	79.11	185.16	$2 \cdot 10^{-16}$	*
MOS11	1	12.02	12.02	28.14	$4.47 \cdot 10^{-6}$	*
TEX1: MOS11	1	0.57	0.57	1.33	0.26	
Residuals	40	17.09	0.43			

Tukey's test	p adj	
mos11 : tex1 mos11	0	*
tex1 : tex1 mos11	$2.64 \cdot 10^{-4}$	*
Col-0 : tex1 mos11	0	*
tex1 : mos11	$4.2 \cdot 10^{-6}$	*
Col-0 : mos11	0.03	*
Col-0 : tex1	0	*

Table 33. Anova: Relative expression of *FLC* (P49+P50) See Figure 38C.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.10	2.69	0.06
Standard deviation	0.04	0.06	0.37	0.01

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	9.39	9.39	260.21	$2.19 \cdot 10^{-7}$	*
MOS11	1	2.04	2.04	56.55	$6.80 \cdot 10^{-5}$	*
TEX1: MOS11	1	2.25	2.25	62.28	$4.81 \cdot 10^{-5}$	*
Residuals	8	0.29	0.04			

Tukey's test	p adj	
mos11 : tex1 mos11	$7 \cdot 10^{-7}$	*
tex1 : tex1 mos11	0.99	
Col-0 : tex1 mos11	$1.32 \cdot 10^{-3}$	*
tex1 : mos11	$8 \cdot 10^{-7}$	*
Col-0 : mos11	$2.08 \cdot 10^{-5}$	*
Col-0 : tex1	$1.76 \cdot 10^{-3}$	*

Table 34. Anova: Number of primary inflorescences. Figure 38E

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	6.45	9.00	6.09	13.55
Standard deviation	1.29	2.28	0.94	3.39

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	275	275	57.18	$3.13 \cdot 10^{-9}$	*
MOS11	1	48.09	48.09	10	$3 \cdot 10^{-3}$	*
TEX1: MOS11	1	66.27	66.27	13.78	$6.3 \cdot 10^{-4}$	*
Residuals	40	192.36	4.81			

Tukey's test	p adj	
mos11 : tex1 mos11	0	*
tex1 : tex1 mos11	$1.05 \cdot 10^{-4}$	*
Col-0 : tex1 mos11	0	*
tex1 : mos11	0.02	*
Col-0 : mos11	0.98	
Col-0 : tex1	$4.5 \cdot 10^{-2}$	*

Table 35. Anova: Lateral root density in the Branching zone (Lateral roots/cm). Figure 38H.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	4.87	3.82	5.29	4.63
Standard deviation	0.68	0.94	0.68	1.16

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	27.13	27.13	34.47	$2.86 \cdot 10^{-8}$	*
MOS11	1	13.95	13.95	17.72	$4.48 \cdot 10^{-5}$	*
TEX1: MOS11	1	1.38	1.38	1.76	0.19	
Residuals	144	113.32	0.79			

Tukey's test	p adj	
mos11 : tex1 mos11	$8.66 \cdot 10^{-3}$	*
tex1 : tex1 mos11	$7.99 \cdot 10^{-5}$	*
Col-0 : tex1 mos11	0.64	
tex1 : mos11	0	*
Col-0 : mos11	0.18	
Col-0 : tex1	$6.6 \cdot 10^{-6}$	*

Table 36. Anova: Root length (cm). See Figure 38I.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	4.51	5.13	4.63	2.99
Standard deviation	0.47	0.46	0.36	0.37

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	10.71	10.71	61.24	$6.49 \cdot 10^{-13}$	*
MOS11	1	41.49	41.49	237.28	$2. \cdot 10^{-16}$	*
TEX1: MOS11	1	51.99	51.99	297.3	$2. \cdot 10^{-16}$	*
Residuals	160	27.98	0.17			

Tukey's test	p adj	
mos11 : tex1 mos11	0	*
tex1 : tex1 mos11	0	*
Col-0 : tex1 mos11	0	*
tex1 : mos11	$1.7 \cdot 10^{-7}$	*
Col-0 : mos11	0.56	
Col-0 : tex1	0	*

Table 37. Anova: Relative expression of *TAS1* transcript (TAS1/U6). Figure 39B.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.23	0.96	0.36
Standard deviation	0.24	0.06	0.19	0.09

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	55.61	55.61	71.46	$2.13 \cdot 10^{-6}$	*
MOS11	1	0.27	0.27	0.35	0.57	
TEX1: MOS11	1	0.90	0.90	1.16	0.30	
Residuals	12	9.34	0.78			

Tukey's test	p adj	
mos11 : tex1 mos11	$1.07 \cdot 10^{-3}$	*
tex1 : tex1 mos11	0.65	
Col-0 : tex1 mos11	$6.17 \cdot 10^{-4}$	*
tex1 : mos11	$1.74 \cdot 10^{-4}$	*
Col-0 : mos11	0.99	
Col-0 : tex1	$1.06 \cdot 10^{-4}$	*

Table 38. Anova: Relative expression of *TAS3* transcript (TAS3/U6) See Figure 39C.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.69	0.92	0.54
Standard deviation	0.14	0.04	0.16	0.09

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	0.40	0.40	35.33	$6.78 \cdot 10^{-5}$	*
MOS11	1	0.04	0.04	3.93	0.07	
TEX1: MOS11	1	$4.20 \cdot 10^{-3}$	$4.20 \cdot 10^{-3}$	0.37	0.56	
Residuals	12	0.14	0.01			

Tukey's test	p adj	
mos11 : tex1 mos11	$2.80 \cdot 10^{-3}$	*
tex1 : tex1 mos11	0.31	
Col-0 : tex1 mos11	$5.768 \cdot 10^{-4}$	*
tex1 : mos11	0.07	
Col-0 : mos11	0.77	
Col-0 : tex1	0.01	*

Table 39. Anova: Relative expression of *miR173* transcript (miR173/U6). See Figure 39D.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.05	1.00	0.73	0.75
Standard deviation	0.12	0.07	0.06	0.04

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	$9.00 \cdot 10^{-4}$	$9.00 \cdot 10^{-4}$	0.14	0.71	
MOS11	1	0.36	0.36	56.93	$6.81 \cdot 10^{-6}$	*
TEX1: MOS11	1	$5.50 \cdot 10^{-3}$	$5.50 \cdot 10^{-3}$	0.88	0.37	
Residuals	12	0.08	$6.30 \cdot 10^{-3}$			

Tukey's test	p adj	
mos11 : tex1 mos11	0.98	
tex1 : tex1 mos11	$2.61 \cdot 10^{-3}$	*
Col-0 : tex1 mos11	$5.8 \cdot 10^{-4}$	*
tex1 : mos11	$1.36 \cdot 10^{-3}$	*
Col-0 : mos11	$3.15 \cdot 10^{-4}$	*
Col-0 : tex1	0.79	

Table 40. Anova: Relative expression of *miR390* transcript (miR390/U6). See Figure 39E.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.84	0.95	0.92
Standard deviation	0.07	0.07	0.12	0.12

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)
TEX1	1	0.04	0.04	3.53	0.08
MOS11	1	$5.40 \cdot 10^{-4}$	$5.40 \cdot 10^{-4}$	0.05	0.83
TEX1: MOS11	1	0.02	0.02	1.75	0.21
Residuals	12	0.14	0.01		

Table 41. Anova: Relative expression of *ARF3* transcript (P51+P52). See Figure 39F.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.96	0.91	1.05
Standard deviation	0.09	0.04	0.13	0.03

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)
TEX1	1	0.0067	0.0067	0.96	0.36
MOS11	1	$1 \cdot 10^{-5}$	0.00001	0.001	0.97
TEX1: MOS11	1	0.024	0.024	3.4	0.1
Residuals	8	0.056	0.007		

Table 42. Anova: Relative expression of *ARF4* transcript (P53+P54). Figure 39G.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.00	0.86	1.47	0.95
Standard deviation	0.22	0.40	0.29	0.32

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)
TEX1	1	0.32	0.32	3.33	0.11
MOS11	1	0.24	0.24	2.47	0.16
TEX1 :MOS11	1	0.11	0.11	1.12	0.32
Residuals	8	0.78	0.10		

Table 43. Anova: Trichomes with one branch point (%). See Figure 40A.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	8.60	14.90	6.06	7.00
Standard deviation	8.07	16.71	5.84	7.33

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)
TEX1	1	196	196.30	1.82	0.18
MOS11	1	409	408.80	3.78	0.06
TEX1: MOS11	1	108	107.60	1.00	0.32
Residuals	56	6053	108.10		

Table 44. Anova: Trichomes with two branch points (%). See Figure 40A.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	86.75	62.01	65.73	40.29
Standard deviation	9.32	16.72	18.61	11.75

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	9442	9442	44.40	$1.24 \cdot 10^{-8}$	*
MOS11	1	6852	6852	32.22	$5.08 \cdot 10^{-7}$	*
TEX1: MOS11	1	2	2	$9.00 \cdot 10^{-3}$	0.93	
Residuals	56	11908	213			

Tukey's test	p adj	
mos11 : tex1 mos11	$7.63 \cdot 10^{-5}$	*
tex1 : tex1 mos11	$8.10 \cdot 10^{-4}$	*
Col-0 : tex1 mos11	0	*
tex1 : mos11	0.9	
Col-0 : mos11	$1.244 \cdot 10^{-3}$	*
Col-0 : tex1	$1.21 \cdot 10^{-4}$	*

Table 45. Anova: Trichomes with three branch points (%).Figure 40A

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	4.65	22.61	27.23	41.29
Standard deviation	5.49	17.26	18.12	11.76

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	3845	3845	19.36	$4.91 \cdot 10^{-5}$	*
MOS11	1	6384	6384	32.14	$5.22 \cdot 10^{-7}$	*
TEX1: MOS11	1	57	57	0.29	0.59	
Residuals	56	11125	199			

Tukey's test	p adj	
mos11 : tex1 mos11	0.04	*
tex1 : tex1 mos11	$3.36 \cdot 10^{-3}$	*
Col-0 : tex1 mos11	0	*
tex1 : mos11	0.81	
Col-0 : mos11	$2.92 \cdot 10^{-4}$	*
Col-0 : tex1	$5.10 \cdot 10^{-3}$	*

Table 46. Anova: Trichomes with four branch points (%). See Figure 40A

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	0.00	0.48	0.98	9.51
Standard deviation	0.00	1.84	2.95	6.04

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	304.2	304.2	25.05	$5.89 \cdot 10^{-6}$	*
MOS11	1	375.9	375.9	30.95	$7.72 \cdot 10^{-7}$	*
TEX1: MOS11	1	243.3	243.3	20.03	$3.79 \cdot 10^{-5}$	*
Residuals	56	680.2	12.1			

Tukey's test	p adj	
mos11 : tex1 mos11	$1 \cdot 10^{-7}$	*
tex1 : tex1 mos11	0	*
Col-0 : tex1 mos11	0	*
tex1 : mos11	0.98	
Col-0 : mos11	0.87	
Col-0 : tex1	0.98	

Table 47. Anova: Trichomes with five branch points (%). See Figure 40A.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	0	0	0	1.52
Standard deviation	0	0	0	2.76

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	8.68	8.68	4.55	0.04	*
MOS11	1	8.68	8.68	4.55	0.04	*
TEX1: MOS11	1	8.68	8.68	4.55	0.04	*
Residuals	56	106.79	1.907			

Tukey's test	p adj	
mos11 : tex1 mos11	0.02	*
tex1 : tex1 mos11	0.02	*
Col-0 : tex1 mos11	0.02	*
tex1 : mos11	1	
Col-0 : mos11	1	
Col-0 : tex1	1	

Table 48. Anova: Trichomes with seven branch point (%). See Figure 40A

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	0.00	0.00	0.00	0.39
Standard deviation	0.00	0.00	0.00	1.52

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	0	0			
MOS11	1	0	0			
TEX1: MOS11	1	0	0			
Residuals	56	0	0			

Table 49. Anova: Percentage of defective embryos (%). See Figure 43E

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	0.85	3.00	0.89	28.34
Standard deviation	1.90	3.22	1.22	22.75

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F-value	Pr(>F)	
Tex1	1	1094.8	1094.8	8.215	0.01	*
Mos11	1	804.9	804.9	6.04	0.03	*
Tex1:Mos11	1	800	800	6.00	0.03	*
Residuals	16	2132.1	133.3			

Tukey's test	p adj	
mos11 : tex1 mos11	$8.36 \cdot 10^{-3}$	*
tex1 : tex1 mos11	0.02	*
Col-0 : tex1 mos11	$8.27 \cdot 10^{-3}$	*
tex1 : mos11	0.99	
Col-0 : mos11	0.99	
Col-0 : tex1	0.99	

Table 50. Anova: Total number of normal and non-developed ovules. See Figure 43F.

	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	48.80	46.40	47.80	46.20
Standard deviation	3.70	1.52	2.86	2.77

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F-value	Pr(>F)	
Tex1	1	20	20	2.51	0.13	
Mos11	1	1.8	1.8	0.23	0.64	
Tex1:Mos11	1	0.8	0.8	0.1	0.76	
Residuals	16	127.6	7.98			

Table 51. Anova: mRNA export deficiency assay (N/C ratio). See Figure 44B.

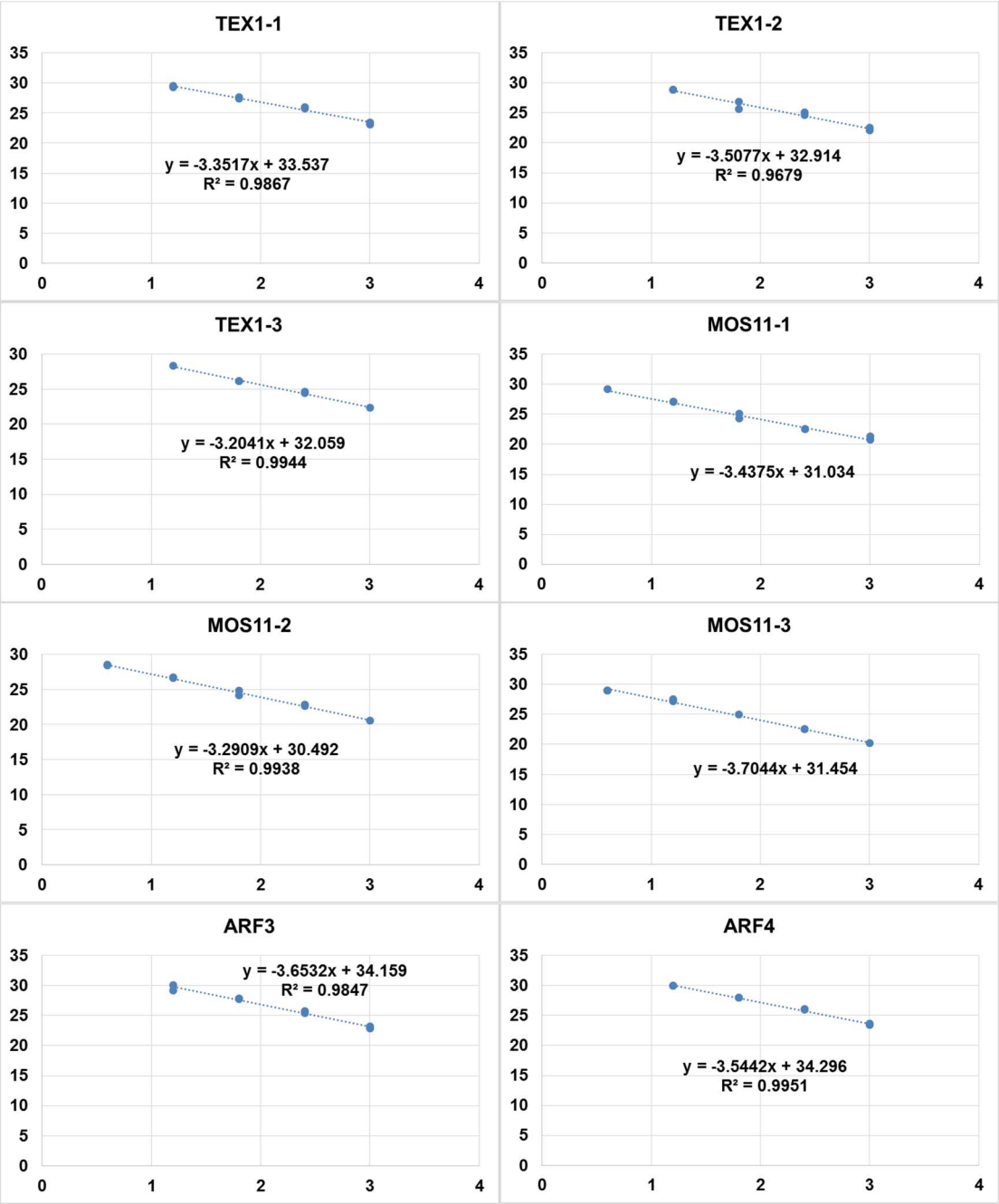
	Col-0	*tex1-4*	*mos11-2*	*tex1-4 mos11-2*
Mean	1.82	1.73	2.07	2.49
Standard deviation	0.37	0.35	0.38	0.60

Two-way ANOVA table	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
TEX1	1	1.65	1.65	8.27	$4.4 \cdot 10^{-3}$	*
MOS11	1	15.12	15.12	75.68	$5.79 \cdot 10^{-16}$	*
TEX1: MOS11	1	3.93	3.93	19.68	$1.41 \cdot 10^{-5}$	*
Residuals	236	47.16	0.2			

Tukey's test	p adj	
mos11 : tex1 mos11	$3 \cdot 10^{-6}$	*
tex1 : tex1 mos11	0	*
Col-0 : tex1 mos11	0	*
tex1 : mos11	$3.07 \cdot 10^{-4}$	*
Col-0 : mos11	0.02	*
Col-0 : tex1	0.69	

7.3 RT-qPCR

To calculate the primer efficiency for the primers used in RT-qPCR a dilution row was made. The cDNA was diluted 1:1, 1:4, 1:16, 1:64, 1:256 and 1:1024. the C_t values (Y-axis) found by RT-qPCR were plottet against the logarithm of the concentration (X-axis), represented by the dilution number. Only the dilutions following linearity were included in the linear regression. The slope of the linear regression was used to calculate the efficiency, using the formula $E = 10^{(-1/S)}$.



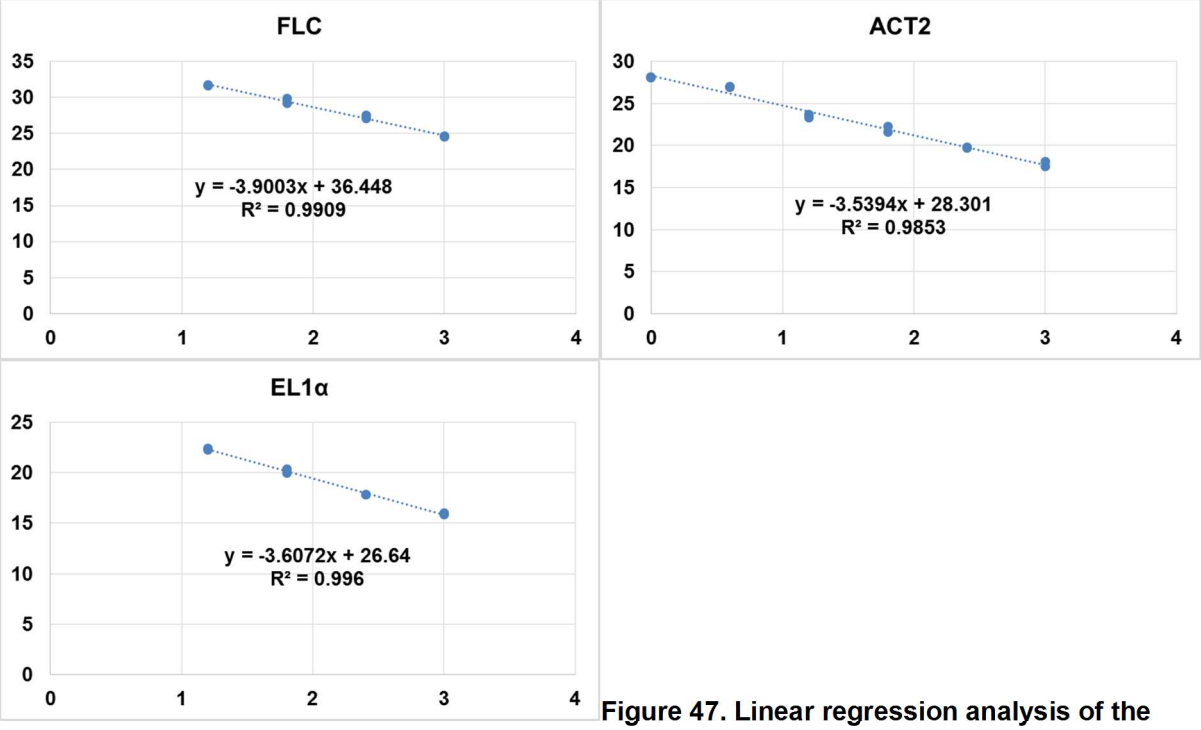


Figure 47. Linear regression analysis of the primers used for RT-qPCR in this analysis. The slope is used to calculate the primer efficiency. X-axis is the log of the concentration. Y-axis is the Ct value.

Table 52. fragment length and efficiency of the primers used in RT-qPCR analysis in this study. All primer pairs had optimal annealing temperature at 56 °C, except TEX1-3 (62.5°C).

	Primers	fragment length (bp)	efficiency (%)
TEX1-1	P29 + P30	150	98.8
TEX1-2	P31 + P32	180	91.4
TEX1-3	P33 + P34	250	105.1
MOS11-1	P35 + P36	203	95.4
MOS11-2	P37 + P38	214	100.0
MOS11-3	P39 + P40	245	86.3
FLC	P49 + P50	155	80.5
ARF3	P51 + P52	233	87.9
ARF4	P53 + P54	236	91.5
ACT2	P43 + P44	108	91.7
EF1α	P45 + P46	121	89.6

Table 53. RT-qPCR: The calibrated normalized relative quantities (CQNR) for all the analyses done in this study. The CQNR was calculated in qBase+. The values are shown for each biological replicate with standard errors. The genes were normalized to ACT2 and EL1α.

	TEX1-1	TEX1-3	TEX1-4	MOS11-1	MOS11-2	MOS11-4	FLC	ARF3	ARF4
wt1	4.18 ± 0.18	4.61 ± 0.34	7.33 ± 0.61	1.96 ± 0.16	4.59 ± 0.64	4.32 ± 0.23	3.02 ± 0.19	1.04 ± 0.07	0.75 ± 0.08
wt2	3.42 ± 0.28	3.64 ± 0.36	5.65 ± 0.65	1.41 ± 0.10	2.75 ± 0.30	3.39 ± 0.25	2.88 ± 0.12	1.11 ± 0.12	1.04 ± 0.09
wt3	2.79 ± 0.12	3.28 ± 0.26	8.74 ± 1.04	1.69 ± 0.14	4.15 ± 0.26	3.15 ± 0.13	3.10 ± 0.20	0.93 ± 0.03	1.18 ± 0.16
tex1	0.17 ± 0.01	0.51 ± 0.04	0.12 ± 0.01	1.82 ± 0.13	4.47 ± 0.56	4.01 ± 0.25	0.22 ± 0.01	1.03 ± 0.06	1.18 ± 0.11
tex2	0.42 ± 0.06	0.56 ± 0.07	0.34 ± 0.02	0.82 ± 0.07	2.64 ± 0.17	2.81 ± 0.20	0.51 ± 0.05	0.97 ± 0.05	0.42 ± 0.02
tex3	0.20 ± 0.01	0.45 ± 0.04	0.13 ± 0.01	1.02 ± 0.04	2.26 ± 0.20	3.81 ± 0.26	0.15 ± 0.01	0.95 ± 0.08	0.96 ± 0.05
mos1	6.49 ± 0.48	5.80 ± 0.54	14.43 ± 1.76	1.05 ± 0.07	1.22 ± 0.16	0.48 ± 0.04	6.92 ± 0.70	0.81 ± 0.11	1.77 ± 0.20
mos2	4.31 ± 0.23	5.99 ± 0.52	13.02 0.61	0.86 ± 0.08	0.75 ± 0.09	0.28 ± 0.05	8.14 ± 0.59	1.08 ± 0.09	1.20 ± 0.11
mos3	3.67 ± 0.28	5.86 ± 0.44	9.03 ± 0.56	0.90 ± 0.07	1.14 ± 0.09	0.27 ± 0.02	9.15 ± 0.74	0.90 ± 0.09	1.41 ± 0.11
texmos1	0.33 ± 0.03	0.09 ± 0.01	0.06 ± 0.01	0.60 ± 0.05	0.07 ± 0.01	0.22 ± 0.03	0.14 ± 0.01	1.08 ± 0.03	1.31 ± 0.12
texmos2	0.21 ± 0.01	0.09 ± 0.01	0.11 ± 0.01	0.49 ± 0.05	0.10 ± 0.01	0.16 ± 0.01	0.16 ± 0.01	1.04 ± 0.09	0.73 ± 0.06
texmos3	0.26 ± 0.00	0.09 ± 0.01	0.05 ± 0.00	0.58 ± 0.05	0.11 ± 0.01	0.38 ± 0.05	0.21 ± 0.01	1.10 ± 0.02	0.80 ± 0.01

7.4 Mass spectrometry

All the Mass spectrometry data are available on the enclosed CD on the Last page. In the following tables the data from three of four affinity purifications are shown. Promiscuous (Van Leene et al., 2015) and contaminating proteins have been removed.

Table 54. TEX1 affinity purification. [1]Times detected out of four affinity purifications.

AGI	NCBI name	Average protein score	Times detected[1]	complex[3]
AT1G24706	THO complex subunit 2	3355.1	4	TREX complex
AT5G42920	THO complex, subunit 5	1922.1	4	TREX complex
AT5G56130	THO complex subunit 3	1814.7	4	TREX complex
AT5G09860	THO complex subunit 1	1533.8	4	TREX complex
AT4G39680	SAP domain-containing protein	1002.2	4	
AT2G38770	unknown protein	810.7	4	NTC-associated
AT2G14120	dynamin-like protein	719.0	4	
AT4G33650	dynamin like protein 2a	663.0	4	
AT1G09770	Myb-like DNA binding protein	531.6	4	Core NTC
AT5G64270	putative splicing factor	499.9	4	17S U2 snRNP
AT5G50850	pyruvate dehydrogenase E1 component subunit beta	494.6	4	
AT1G75660	unknown protein	459.0	4	Exosome
AT5G16790	putative protein	451.6	4	TREX complex
AT5G50320	histone acetyltransferase	444.3	4	Elongator
AT5G28740	pre-mRNA-splicing factor SYF1	440.8	4	Core NTC
AT3G02950	Tho complex subunit 7/Mft1p	429.9	4	TREX complex
AT3G26420	AT3g26420/F20C19_14	413.4	4	hnRNP family
AT4G39050	putative kinesin	394.9	4	
AT2G07690	(cell division control protein 46)	350.2	4	
AT2G19430	THO complex subunit 6	341.3	4	TREX complex
AT3G01540	putative RNA helicase, DRH1	339.9	4	
AT1G45233	THO complex, subunit 5	338.8	4	TREX complex
AT3G01280	mitochondrial outer membrane protein porin 1	338.1	4	
AT3G55200	splicing factor 3B subunit 3	336.8	4	17S U2 snRNP
AT3G08943	armadillo/beta-catenin-like repeat-containing protein	328.5	4	
AT1G07360	pre-mRNA-splicing factor RBM22/SLT11	311.6	4	NTC-associated
AT2G46610	RNA recognition motif-containing protein	303.1	4	SR proteins
AT2G33340	Pre-mRNA-processing factor 19-2	297.1	4	Core NTC
AT1G31817	30S ribosomal protein S11, putative	295.7	4	
AT3G49430	putative splicing factor; 53460-55514	291.5	4	SR proteins
AT2G42710	ribosomal protein .1/L10 family protein	286.0	4	
AT3G54110	uncoupling mitochondrial protein 1	275.5	4	
AT1G02930	glutathione S-transferase 1	272.5	4	

AT3G06480	DEAD-box ATP-dependent RNA helicase 40	270.5	4	
AT2G02090	SNF2 and helicase domain-containing protein	268.8	4	
AT3G55620	Translation initiation factor IF6	265.3	4	
AT4G01100	adenine nucleotide transporter 1	263.5	4	
AT3G13290	varicose-related protein	260.7	4	
AT3G02200	Proteasome component (PCI) domain protein	260.5	4	
AT1G77180	unknown protein	255.6	4	Core NTC
AT5G64260	protein EXORDIUM like 2	239.6	4	
AT2G03510	At2g03510	230.3	4	
AT5G49830	exocyst complex component 84B	223.1	4	
AT4G03430	pre-mRNA-processing factor 6	221.8	4	U5 snRNP
AT1G60650	RNA recognition motif-containing protein	212.6	4	hnRNP family
AT5G04280	RNA recognition motif-containing protein	204.2	4	hnRNP family
AT5G46750	AGD9	202.6	4	
AT4G15020	putative protein	201.1	4	
AT1G64880	Ribosomal protein S5 family protein	197.0	4	
AT3G55460	SC35-like splicing factor 30	192.2	4	SR proteins
AT3G26560	RNA helicase	191.5	4	RES complex
AT2G43030	50S ribosomal protein L3	191.2	4	
AT3G11400	initiation factor 3g	188.9	4	
AT1G27650	splicing factor U2af small subunit A	186.1	4	Splice site selection
AT3G63250	homocysteine S-methyltransferase 2	180.7	4	
AT5G13850	Nascent polypeptide-associated complex protein 3	177.1	4	
AT3G12790	unknown protein	175.7	4	
AT3G45190	SIT4 phosphatase-associated family protein	175.5	4	
AT2G41040	uncharacterized methyltransferase	171.5	4	
AT1G15200	protein-protein interaction regulator family protein	170.5	4	EJC/mRNP
AT4G23620	ribosomal L25/TL5/CTC N-terminal 5S rRNA binding	160.4	4	
AT5G55670	RNA recognition motif-containing protein	160.0	4	
AT1G48650	hypothetical protein	155.4	4	
AT4G17950	AT hook motif DNA-binding family protein	153.4	4	
AT2G22300	calmodulin-binding transcription activator 3	144.3	4	
AT4G21660	splicing factor 3B subunit 2	142.1	4	17S U2 snRNP
AT1G48570	zinc finger (Ran-binding) domain-containing protein	124.1	4	
AT3G07030	Alba DNA/RNA-binding protein	105.3	4	
AT5G26210	nucleic acid binding protein - like	104.9	4	
AT1G20960	putative U5 small nuclear ribonucleoprotein helicase	934.4	3	U5 snRNP
AT3G13930	dihydrolipoyllysine acetyltransferase component 2	755.5	3	
AT5G13530	E3 ubiquitin-protein ligase KEG	556.4	3	
AT4G24680	MODIFIER OF snc1	507.1	3	
AT1G66260	DIP1 protein	491.6	3	TREX complex
AT5G58040	unnamed protein product	489.5	3	
AT3G54230	suppressor of ABI3-5	465.9	3	A complex
AT5G26760	uncharacterized protein	462.6	3	

AT2G01690	uncharacterized protein	441.0	3	
AT2G03150	ATP/GTP-binding protein-like protein	423.4	3	
AT5G53440	uncharacterized protein	400.1	3	
AT4G02570	cullin 1	396.3	3	
AT4G22690	putative cytochrome P450 protein	373.0	3	
AT4G02350	exocyst complex component sec15B	366.9	3	
AT5G55210	uncharacterized protein	353.6	3	
AT2G37230	pentatricopeptide repeat-containing protein	341.6	3	
AT1G80930	pre-mRNA-splicing factor CWC22	332.8	3	NTC-associated
AT5G22040	uncharacterized protein	316.4	3	
AT3G59020	armadillo/beta-catenin-like repeat-containing protein	313.4	3	
AT1G30680	T5I8.13	308.1	3	
AT2G22660	uncharacterized protein	302.4	3	
AT5G25060	U2-associated protein SR140	301.1	3	U2 associated
AT1G23860	RS-containing zinc finger protein 21	301.0	3	SR proteins
AT5G13010	ATP-dependent RNA helicase-like protein	295.7	3	RES complex
AT5G41770	cell cycle control crn (crooked neck) protein-like	289.9	3	Core NTC
AT4G24840	brefeldin A-sensitive Golgi protein-like	287.1	3	
AT1G69830	putative alpha-amylase; 60344-64829	283.3	3	
AT1G24280	glucose-6-phosphate dehydrogenase 3	279.5	3	
AT1G19520	pentatricopeptide repeat-containing protein	270.5	3	
AT5G49720	AT5g49720/K2I5_8	265.8	3	
AT4G21710	DNA-directed RNA polymerase II subunit RPB2	259.1	3	Polymerase II
AT2G38040	carboxyltransferase alpha subunit	252.8	3	
AT1G26750	uncharacterized protein	245.8	3	
AT3G62310	DHX15/PRP43	243.0	3	RES complex
AT1G14900	linker histone protein, putative	240.8	3	
AT1G71270	Vps52 / Sac2-like protein	238.1	3	
AT5G08450	putative protein	238.0	3	
AT5G22770	AT5G22770	237.9	3	
AT5G48680	unnamed protein product	235.3	3	
AT1G44900	minichromosome maintenance protein 2	234.3	3	
AT1G52370	chloroplast 50S ribosomal protein L22,	234.2	3	
AT2G44710	RNA recognition motif-containing protein	229.7	3	hnRNP family
AT5G44780	uncharacterized protein	229.3	3	
AT4G01400	uncharacterized protein	229.2	3	
AT1G47550	exocyst complex component sec3A	229.2	3	
AT2G45140	VAP-like protein 12	226.4	3	
AT5G66420	uncharacterized protein	225.2	3	
AT1G49040	stomatal cytokinesis defective / SCD1 protein (SCD1)	213.8	3	
AT3G55005	hypothetical protein	211.3	3	
AT1G67930	Golgi transport complex-related protein	210.4	3	
AT2G45640	histone deacetylase complex subunit SAP18	206.5	3	EJC/mRNP
AT4G16120	COBRA-like protein 7	205.4	3	
AT1G23180	armadillo/beta-catenin-like repeat-containing protein	200.3	3	

AT5G62270	uncharacterized protein	199.5	3	
ATCG00380	ribosomal protein S4	199.4	3	
AT2G43770	Prp8 binding protein	195.7	3	U5 snRNP
AT1G48900	signal recognition particle subunit SRP54	194.9	3	
AT2G27200	nucleoside triphosphate hydrolase-like protein	194.7	3	
AT3G02650	pentatricopeptide repeat-containing protein	193.8	3	
AT3G63400	peptidyl-prolyl cis-trans isomerase family protein	192.8	3	C complex
AT5G15020	paired amphipathic helix protein Sin3-like 2	192.2	3	
AT1G22730	putative topoisomerase	189.9	3	
AT1G16520	uncharacterized protein	189.5	3	
AT1G79490	pentatricopeptide repeat-containing protein	185.6	3	
AT1G73430	unknown protein	185.0	3	
AT4G01990	hypothetical protein	184.7	3	
AT3G61690	nucleotidyltransferase	184.5	3	
AT3G46220	E3 UFM1-protein ligase 1-like protein	183.5	3	
AT5G50380	unnamed protein product	183.3	3	
AT1G79090	DNA topoisomerase 2-associated protein PAT1	182.8	3	
AT5G35590	multicatalytic endopeptidase complex	178.7	3	
AT2G33620	AT hook motif DNA-binding family protein	172.6	3	
ATMG00290	ribosomal protein S4	168.9	3	
AT4G02150	Importin subunit alpha-2	167.9	3	
AT5G12100	pentatricopeptide repeat-containing protein	166.3	3	
AT1G72560	PAUSED	165.7	3	
AT5G17440	putative protein	165.5	3	U1 snRNP
AT5G18420	uncharacterized protein	164.9	3	
AT4G02060	protein PROLIFERA	164.6	3	
AT2G06990	HUA enhancer 2	164.2	3	
AT5G14580	polyribonucleotide nucleotidyltransferase	161.0	3	
AT5G16840	AT5G16840	160.7	3	
AT4G35785	At4g35785	160.4	3	SR proteins
AT1G23280	F26F24.12	159.2	3	
AT4G25550	cleavage/polyadenylation specificity factor	158.2	3	Polyadenylation
AT1G63160	replication factor C 2	157.1	3	
AT5G08550	GC-rich sequence DNA-binding factor	156.3	3	
AT1G32490	DHX16	156.2	3	NTC-associated
AT1G07990	SIT4 phosphatase-associated family protein	153.0	3	
AT5G16750	WD40-repeat protein	152.3	3	
AT1G55890	pentatricopeptide repeat-containing protein	150.9	3	
AT4G17530	AT4g17530/dl4800c	150.6	3	
AT4G27500	AT4G27500	149.3	3	
AT3G45850	P-loop containing nucleoside triphosphate protein	147.6	3	
AT2G36200	kinesin family member 11	146.7	3	
AT3G05680	unknown protein	145.9	3	
AT3G06670	f IIS longevity pathway SMK-1 domain protein	145.6	3	
AT1G27750	T22C5.20	143.3	3	

AT1G60900	putative U2 snRNP auxiliary factor; 19096-22891	142.2	3	Splice site selection
AT3G21540	transducin/WD40 domain-containing protein	140.3	3	
AT1G14650	splicing factor, putative	139.9	3	U2 snRNP
AT5G40770	prohibitin 3	138.7	3	
AT2G05840	proteasome subunit alpha type-6-B	138.7	3	
AT5G15610	Proteasome component (PCI) domain protein	138.0	3	
AT4G36690	Splicing factor U2af large subunit A	137.2	3	Splice site selection
AT5G37475	Translation initiation factor eIF3 subunit	135.6	3	
AT5G19760	Mitochondrial substrate carrier family protein	133.4	3	
AT1G55310	unknown protein; 47745-45927	133.4	3	SR proteins
AT4G09980	methyltransferase-like protein 1	131.6	3	B complex
AT5G11170	DEAD-box ATP-dependent RNA helicase 56	131.4	3	EJC/mRNP
AT5G60960	unknown	129.8	3	
AT3G48860	uncharacterized protein	129.5	3	
AT1G12470	zinc ion binding protein	126.1	3	
AT1G10170	NF-X1-type zinc finger protein NFXL1	123.3	3	
AT2G06210	putative TPR repeat nuclear phosphoprotein	121.7	3	PAF-C
AT1G70620	unknown protein; 76547-79900	120.6	3	
AT4G24270	protein embryo defective 140	117.3	3	U4/U6 snRNP
AT3G20250	pumilio 5	116.5	3	
AT1G52980	AT1G52980	116.5	3	
AT2G47580	spliceosomal protein U1A	116.4	3	U1 snRNP
AT5G03070	importin alpha-like protein	115.7	3	
AT5G51590	AT hook motif DNA-binding family protein	113.5	3	
AT3G49090	hypothetical protein	109.4	3	
AT1G51510	RNA-binding protein, putative; 35994-37391	109.1	3	EJC/mRNP
AT5G11980	unknown	108.5	3	
AT1G17070	tuftelin-interacting-like protein	105.8	3	RES complex
AT3G17465	50S ribosomal protein L3-2	104.3	3	
AT4G15570	protein MAGATAMA 3	95.7	3	

Table 55. UAP56 affinity purification . [1]Times detected out of four affinity purifications.

AGI	NCBI name	Average protein score	Times detected[1]	complex
AT1G24706	THO complex subunit 2	5853.175	4	TREX complex
AT5G11170	UAP56	4491.325	4	TREX complex
AT5G42920	THO complex, subunit 5	3420.675	4	TREX complex
AT5G09860	THO complex subunit 1	2379.4	4	TREX complex
AT5G02770	MOS11	2023.175	4	TREX complex
AT5G16790	Tho complex subunit 7/Mft1p	1684.9	4	TREX complex
AT1G45233	THO complex, subunit 5	1633.925	4	TREX complex
AT3G02950	Tho complex subunit 7/Mft1p	1597.525	4	TREX complex
AT5G56130	THO complex subunit 3	1013.325	4	TREX complex
AT2G19430	THO complex subunit 6	1013.075	4	TREX complex

AT1G16520	uncharacterized protein	868.575	4	
AT3G04490	uncharacterized protein	693.525	4	
AT2G19560	proteasome-like protein	684.725	4	TREX-2 complex
AT4G15545	uncharacterized protein	639.775	4	
AT3G54380	SAC3C	618.875	4	TREX-2 complex
AT5G02530	THO complex subunit 4	404.65	4	TREX complex
AT3G12790	unknown protein	359	4	
AT5G37720	unnamed protein product	274.25	4	TREX complex
AT5G50850	pyruvate dehydrogenase E1 component subunit beta	264.7	4	
AT2G22230	3R-hydroxymyristoyl ACP dehydrase	180	4	
AT4G25550	cleavage/polyadenylation specificity factor	115.375	4	Polyadenylation-cleavage complex
AT1G56080	uncharacterized protein	714.8333333	3	
AT3G55410	2-oxoglutarate dehydrogenase, E1 subunit-like protein	327.6666667	3	
AT5G65750	2-oxoglutarate dehydrogenase, E1 subunit	245.4333333	3	
AT5G10160	3R-hydroxymyristoyl ACP dehydrase	183.2666667	3	
AT2G38770	intron-binding protein aquarius	158.8333333	3	NTC-associated
AT3G58570	DEAD-box ATP-dependent RNA helicase 52	157.6666667	3	EJC/mRNP
AT1G07360	pre-mRNA-splicing factor RBM22/SLT11	144.7	3	NTC-associated
AT3G07630	putative P-protein: chorismate mutase,	120.2	3	
AT3G06050	unknown protein; 13384-11892	119.8	3	

Table 56. MOS11 affinity purification. [1]Times detected out of four affinity purifications.

AGI	NCBI name	Average protein score	Times detected[1]	complex
AT5G02770	MOS11	3122.5	4	TREX complex
AT5G11170	UAP56	771.1	4	TREX complex

8 List of Publications

Publications related to this work:

- Sørensen, B.B., Ehrnsberger, H., Esposito, S., Pfab, A., Bruckmann, A., Hauptmann, J., Meister, G., Merkl, R., Schubert, T., Längst, G., Melzer, M., Grasser, M. & Grasser, K.D., 2016. TEX1 of the Arabidopsis THO / TREX complex plays a role at different steps of the biogenesis of mRNAs and small RNAs. *In preparation*, pp.1–24.
  -The preliminary manuscript is available in Chapter 8.1

- Kammel, C., Thomaier, M., Sørensen, B.B., Schubert, T., Längst, G., Grasser, M. & Grasser, K.D., 2013. Arabidopsis DEAD-box RNA helicase UAP56 interacts with both RNA and DNA as well as with mRNA export factors. *PloS one*, 8(3).
  – available as PDF on the enclosed CD

Other publications

- Dürr, J., Lolas, I.B., Sørensen, B.B., Schubert, V., Houben, A., Melzer, M., Deutzmann, R., Grasser, M. & Grasser, K.D., 2014. The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in Arabidopsis. *Nucleic acids research*, 42(7), pp.4332–47.
  – available as PDF on the enclosed CD

8.1 Manuscript: *TEX1* of the *Arabidopsis* THO/TREX complex plays a role at different steps of the biogenesis of mRNAs and small RNAs

***Supplemental figures and tables are available on the enclosed CD.**

TEX1 of the *Arabidopsis* THO/TREX complex plays a role at different steps of the biogenesis of mRNAs and small RNAs

Brian B. Sørensen[1], Hans F. Ehrnsberger[1], Silvia Esposito[1], Alexander Pfab[1], Astrid Bruckmann[2], Judith Hauptmann[2], Gunter Meister[2], Rainer Merkl[3], Thomas Schubert[4], Gernot Längst[4], Michael Melzer[5], Marion Grasser[1,*] and Klaus D. Grasser[1,*]

[1]Department of Cell Biology & Plant Biochemistry

[2] Department for Biochemistry I

[3] Department for Biochemistry II

[4] Department for Biochemistry III

Biochemistry Center, University of Regensburg,

Universitätsstr. 31, D-93053 Regensburg, Germany

[5]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstr. 3, D-06466 Stadt Seeland, Germany

To whom correspondence should be addressed:

E-mail Marion.Grasser@ur.de or Klaus.Grasser@ur.de

short title: *Arabidopsis* THO/TREX complex

Abstract

TREX (transcription-export) is a conserved multiprotein complex that plays a central role in the coordination of synthesis, processing and nuclear export of mRNAs. Using a reciprocal tagging approach in combination with affinity purification and mass spectrometry, we characterised the composition of the *Arabidopsis* TREX complex. The previously defined THO core complex consisting of HPR1, THO2, THO5A/B, THO6, THO7A/B and TEX1 was found to associate with the RNA helicase UAP56 and the mRNA export factors ALY2-4 and MOS11. In addition, interactions with the mRNA export complex TREX-2 and multiple spliceosomal components were observed. The phenotype of single-mutant plants defective in the THO component TEX1 or in the RNA-binding mRNA export factor MOS11 (orthologue of human CIP29) is only mildly altered when compared with wild type. However, the *tex1 mos11* double-mutants are more severely affected, displaying various synergistic defects including number and architecture of inflorescences, trichome development and seed set. Northern blot analyses revealed that relative to wild type the levels of tasiRNAs are reduced in *tex1* plants, while miR173 levels are decreased in *mos11* mutants. Examination of subcellular mRNA distribution by *in situ* hybridisation demonstrated increased mRNA accumulation in the nuclei of *mos11* cells, while no mRNA export defect was detected with *tex1* cells. Nevertheless, in *tex1 mos11* double-mutants the mRNA export defect was clearly enhanced relative to *mos11*. Co-localisation experiments show that the subnuclear distribution of TEX1 substantially overlaps with that of splicing-related SR proteins. Moreover, in *tex1*the ratio of certain alternative splicing events is altered. In conclusion, our results demonstrate that *Arabidopsis* TEX1 and MOS11 are involved in distinct steps of the biogenesis of mRNAs and small RNAs, and that they interact regarding some aspects, but act independently in others.

Introduction

Nuclear pre-mRNAs are extensively processed before they are exported to the cytosol for translation. A multitude of mRNA-binding proteins associates co-transcriptionally with the nascent mRNA that brings about various processing steps including 5´end capping, splicing and 3´end polyadenylation. Differential post-translational modifications of the RNA polymerase II (RNAPII) carboxy-terminal domain (CTD) are involved in coordinating transcription and processing of the nascent mRNA [1]. Coupling of transcription and the various processing events is essential for accurate formation of the mature mRNA [2]. Another important step in the formation of export-competent mRNPs is the co-transcriptional recruitment of export factors to the mRNA. The factors mediating mRNA export from the nucleus include so-called export adaptors such as ALY (Yra1 in yeast) and export receptors such as TAP/p15 (Mex67/Mtr2 in yeast), which finally interact with the nuclear pore complex (NPC) upon translocation to the cytosol [3]. Thus, only when mRNA processing is completed and the mRNP is correctly packaged, it is exported from the nucleus for translation [4,5].

The TREX (TRanscription-EXport) multiprotein complex represents a central component in the functional coupling of transcription and other steps of mRNA biogenesis [6,7]. It was first identified in yeast and consists of the THO core complex (Hpr1, Tho2, Mft1 and Thp2) that associates with Tex1, the RNA helicase Sub2 and the mRNA export factor Yra1 forming TREX. Through interaction with the RNAPII-CTD yeast TREX is co-transcriptionally recruited to RNAPII-transcribed genes and operates in transcript elongation as well as the export of mRNAs [8,9]. Moreover, the association of TREX along transcribed genes is increased towards the 3´end [10,11] and mutants affected in TREX components display inefficient mRNA polyadenylation [12]. In metazoa, there are no orthologues of Mft1 and Thp2, but instead three additional subunits of the THO complex were identified that are not found in yeast. Accordingly, the THO complex comprises HPR1, THOC2, THOC5, THOC6, THOC7 and TEX1, which associates with the RNA helicase UAP56 and the mRNA export factor ALY to form metazoan TREX. Additional proteins including the export factor CIP29 (Tho1 in yeast) are interactors of TREX [13,14]. Metazoan TREX is recruited to mRNAs in a capping and splicing dependent manner and it is involved in mRNA export [15–17].

In plants, considerably less is known about the mechanism and the factors involved in co-transcriptional mRNA biogenesis and export of mRNAs. Affinity purification of epitope-tagged TEX1 from *Arabidopsis* plants in combination with mass spectrometric analyses revealed that the plant THO core complex resembles that of metazoa, consisting of HPR1, THO2, THO5A/B, THO6, THO7A/B and TEX1 [18]. The *Arabidopsis* RNA-helicase UAP56 interacts both *in vitro* and in yeast cells with the mRNA export factors ALY2 and MOS11 (the orthologue of mammalian CIP29) [19], but so far it is unclear whether these proteins associate with the THO complex. The phenotypes of *Arabidopsis* mutants defective in THO components reach from no obvious alteration (*tho6* [20]) and relatively mild phenotypes (*tex1, hpr1* [21–23]) to severe/lethal phenotypes (*tho2* [18,20,21,24]). Several lines of evidence demonstrated that *Arabidopsis* plants deficient in the THO components HPR1, THO2 and

TEX1 accumulate reduced levels of small RNAs (siRNAs, miRNAs), but the underlying mechanism is unclear [18,20,21,24]. Moreover, it was reported that *hpr1* mutants display mRNA export defects [22,23] and that *hpr1* and *tho2* mutant plants are affected in alternative splicing [20,24].

In this study, using targeted proteomics we have identified the composition of the *Arabidopsis* TREX complex and its interactions with other proteins. Plants lacking TEX1 display various developmental defects, in particular when combined with a deficiency in the mRNA export factor MOS11. In contrast to *mos11* mutants, *tex1* plants do not show mRNA export defects, but still the *tex1* mutation can enhance the mRNA export defect of *mos11* plants. In addition, *tex1* mutants alter the ratio of some alternative mRNA splicing events and the biogenesis of small RNAs. Therefore, in plants the TEX1 component of the THO/TREX complex appears to be involved in different aspects of the biogenesis of mRNAs and small RNAs.

Results

Composition of the Arabidopsis TREX complex

To learn more about the composition of the plant TREX complex and its interactions with other nuclear proteins, we expressed bait proteins fused to a SG tag (streptavidin-binding peptide and 2x protein G domains) in *Arabidopsis* PSB-D suspension cultured cells [25], an approach that was used to identify other nuclear protein complexes [26–28]. As bait proteins we used the THO component TEX1, the MOS11 protein that was found to be involved in *Arabidopsis* mRNA export [29] and UAP56 that is a central interactor associated with THO in other organisms [6,7]. The Coomassie-stained banding pattern after SDS-PAGE of total protein extracts of cells expressing the SG-fusion proteins or unfused SG was similar and the bands corresponding to the recombinant proteins did not stand out in the extracts of the transformed cell lines (Fig. 1A). To identify proteins interacting with the bait proteins, the SG-fusion proteins and the unfused SG control were comparatively isolated from cell extracts by IgG affinity purification. The eluates were analysed by SDS-PAGE (Fig. 1B) and proteins contained in gel slices were identified after tryptic digestion by mass spectrometry. In line with previous findings [18], all the subunits of the THO core complex were identified in the TEX1-SG eluate (Table 1, TableS1). Consistently, the transcript levels of the genes encoding the THO subunits (except for THO7B, which is expressed at lower levels) display a comparable expression pattern in different *Arabidopsis* tissues (Fig. S2). Moreover, the TREX components UAP56 and ALY3 were found as well as proteins of the exon-junction complex co-purify with TEX1. Since the protein extracts were treated with the endonuclease Benzonase, the observed association of proteins may be due to protein interactions rather than mediated by nucleic acids. Interestingly, many spliceosomal proteins (including proteins of the U1, U2 and U5 snRNPs) and some transcription-related proteins were also identified in the TEX1-SG eluates (Table 1). In agreement with our observation, proteomic analyses of mammalian spliceosomes revealed interactions with THO/TREX proteins [30,31]. All subunits of the THO complex were also found to reproducibly co-purify with UAP56-SG and were detected by mass spectrometry with

high scores. In addition, the mRNA export factors MOS11, ALY2, ALY3 and ALY4 were identified in the UAP56-SG eluate (Table 1, TableS2). ALY1 was also detected, but subtracted from our data, as it was also found in the control experiment with the unfused SG-tag. Moreover, the subunits THP1 and SAC3C of the TREX-2 mRNA export complex were identified in the UAP56-SG affinity purifications (Table 1). The only protein found to co-purify with MOS11-SG was UAP56 (Table 1, TableS3). In conclusion, our proteomics data show that the composition of the plant TREX complex resembles that of metazoa rather than that of yeast (Table S4), and that it interacts with the splicing machinery (Fig. 2). Since our data indicate an in interaction between UAP56 (TREX) and the nuclear pore-associated TREX-2 complex, we did the reverse experiment by analysing proteins that co-purify with THP1-SG (Fig. 1C). We selected THP1 for this experiment, as in *Arabidopsis* THP1 is encoded by a single-copy gene and it shares 35% amino acid sequence identity with its human orthologue PCID2, while there are three versions of SAC3 that are less conserved (10.4%, 19.6%, 5.8% amino acid sequence identity of SAC3A-C, respectively, relative to the human orthologue GANP). In addition to the THP1 bait protein, reproducibly all three variants of SAC3 (SAC3A-C) were identified by mass spectrometry in the eluates of the THP1-SG affinity purification (Table 2, Table S5), but no additional subunits of the putative *Arabidopsis* TREX-2 were detected. SAC3 (GANP) represents the core component of human and yeast TREX-2 around which the other subunits assemble [32,33]. Components of the TREX complex including UAP56 were also found to co-purify with THP1-SG (Table 2), in agreement with the co-purification of SAC3 and THP1 with UAP56-SG (Fig. 2). Moreover, various splicing factors and two nucleoporins were isolated along with the THP1 bait protein (Table S5).

TEX1 and its interplay with MOS11

TEX1 is a conserved 35.4 kDa WD-40 repeat family protein and *Arabidopsis* TEX1 shares 71% and 43% amino acid sequence identity with its rice and human orthologues, respectively (Fig. S2). Since little is known about the function of TEX1, we used various assays to learn more about the role of this conserved TREX component. *Arabidopsis tex1* mutants were reported as viable and phenotypically mildly affected [18,21]. We have now examined more thoroughly the mutant line harbouring a T-DNA insertion in the second intron of the *TEX1* gene (Fig. S3A) and extending previous reports we find that the *tex1* plants are early bolting both under long-day and short-day conditions (Fig. 3A,B; Fig. S4A,B). Moreover, the *tex1* mutants produce an increased number of primary inflorescences relative to Col-0 plants (Fig. 3C; Fig. S4C) and have a decreased density of lateral roots (Fig. 3D; Fig. S4D). Expression of a TEX1-GFP fusion protein in the *tex1* mutants (Fig. S3) fully rescued the early bolting phenotype of the plants under long-day and short-day conditions (Fig. 3E,F; Fig. S4E). Since the TREX complex and MOS11interact (Table 1; [19]) the expression of TEX1-GFP and MOS11-RFP was comparatively analysed by fluorescence microscopy. Both fusion proteins were detected in the nuclei of all cells in the root tips of the transgenic plants (Fig. 4A,B). Closer inspection of leaf and root cell nuclei of plants expressing both fusion proteins revealed that they co-localise to the nucleoplasm and are essentially excluded from the nucleolus (Fig. 4C,D). However, MOS11-RFP seems to be more

uniformly distributed in the nucleoplasm, while TEX1-GFP displays a somewhat speckled pattern. Since the yeast and mammalian orthologues of MOS11 (THO1 and CIP29, respectively) were reported to be RNA-binding proteins [14,34], we examined the RNA-binding capacity of MOS11. 6xHis-tagged MOS11 was efficiently produced in *E.coli* and purified by metal chelate chromatography (Fig. 5A). Using microscale thermophoresis (MST, [35]) we analysed the interaction of recombinant MOS11 with fluorescently labelled 25-nt long ssRNA and dsRNA oligonucleotides in solution. MOS11 interacted with both types of RNA in a concentration dependent manner (Fig. 5B) and quantification of the data revealed that the protein displays a higher affinity for dsRNA (Kd = 93 ± 15 nM) relative to ssRNA (Kd = 1.77 ± 0.4 µM), suggesting that MOS11 binds preferentially to regions of dsRNA such as hairpin loops.

In view of the physical interaction between TREX and MOS11 (this work and [19]) and the co-localisation of TEX1 and MOS11, we generated *tex1 mos11* double mutants to analyse them for possible genetic interactions. The *mos11* single-mutant [29] is like *tex1* phenotypically only mildly affected and fully fertile (Fig. 6A,B; Fig. S5A,B). By contrast, the double-mutant is clearly more severely affected and produces fewer seeds per silique relative to the single-mutants and Col-0. Moreover, the double-mutant plants bolt earlier than the single-mutants and Col-0 (Fig. S5C) and they exhibit an elevated number of primary inflorescences (Fig. S5D). The most striking phenotype of the double-mutant concerns the growth of the inflorescence stem. Whereas Col-0, *tex1* and *mos11* continuously initiate the production of flowers as the inflorescence stem grows, with *tex1 mos11* plants after production of a few flowers the inflorescence stem continues to elongate, but no more flowers are produced, resulting in pin-like structures (Fig. 6B-G) that remind of *pinoid* (*pid*) mutants [36]. Using scanning electron microscopy we compared in more detail the inflorescence tips of Col-0, *tex1 mos11*and *pid* (Fig. 6H-K). Col-0 represents a wild type inflorescence and the inflorescence architecture of *pid* characteristically is pin-shaped, smooth and the flanks are not ridged. The flanks of *tex1 mos11* inflorescences are ridged (cf. Fig. 6I,J and K), suggesting that the floral meristems were initiated, but still they did not produce flowers. Sometimes, after a longer period the elongated pin-shaped inflorescence stems of *tex1 mos11* reinitiate the production of some flowers (Fig. 6B,J; Fig. S5E). In addition, the mutant plants are affected in the branching of leaf hairs called trichomes [37]. Trichomes in Col-0 typically have two branchpoints (87% of trichomes), while the proportion of trichomes with two branchpoints is reduced in *tex1* and *mos11* plants (62% and 66%, respectively) (Fig. 6L-O; Fig. S5F,G). Instead these plants have tendency for trichomes that have three branchpoints (23% and 27% in *tex1* and *mos11*, respectively) when compared with Col-0 (5%). This effect of increased trichome branching is synergistically enhanced in *tex1 mos11* plants having two (40%), three (41%) or four branchpoints (10%) and we even observed one trichome with seven branch points (Fig. S5F,G).

TEX1 and MOS11 in RNA biogenesis

In view of the defects in the accumulation of small RNAs (including tasiRNAs) in *Arabidopsis* plants deficient in the THO components [18,20,21,24], we examined this aspect by Northern blot analyses

comparatively in *tex1 mos11* plants and the corresponding single-mutants as well as Col-0. Using probes specific for the tasiRNAs TAS1 and TAS3 RNA levels similar to Col-0 were detected in *mos11* (Fig. 7A-C). By contrast, comparably reduced levels of both tasiRNAs were measured in *tex1* and the *tex1 mos11* double-mutant, suggesting that in line with previous observations [18,21] TEX1 is required for the efficient biogenesis of these RNAs, but MOS11 seems to be not involved. When the levels of miRNA miR173 were analysed, it was unchanged in *tex1*, but reduced amounts were detected similarly in *mos11* and *tex1 mos11*(Fig. 7D), indicating that MOS11 plays a role in the biogenesis of some miRNAs. Other than with miR173, comparable levels of miR390 were observed in all tested genotypes (Fig. 7E).

Plants defective in the TREX component HPR1 are affected in mRNA export [22,23], which was likewise reported for *mos11* plants [29]. Therefore, we examined bulk mRNA export in *tex1* and *mos11* plants as well as in the corresponding double-mutant. Using *in situ* hybridisation with a fluorescently labelled oligo(dT) probe we analysed the distribution of polyadenylated mRNA in root cells (Fig. 8A). Quantification of the relative nuclear and cytosolic fluorescence revealed that there is increased nuclear retention of mRNAs in *mos11* cells, while in *tex1* the distribution is as in Col-0 (Fig. 8B). Relative to the *mos11* single-mutant, the nuclear accumulation of mRNAs is increased in *tex1 mos11* plants. This suggests that though the *tex1* single-mutant apparently does not affect mRNA export on its own, it enhances the effect when combined with *mos11*. In the literature it is reported that blocks of mRNA export can result in reduced synthesis of new proteins [38]. Therefore, we metabolically pulse labelled protoplasts of the different genotypes with [^{35}S]methionine to visualise newly synthesised proteins. Total protein extracts of comparable number of protoplasts were analysed by SDS-PAGE and Coomassie staining as well as by imaging the radiolabel (Fig. 8C,D). Despite the observed nuclear mRNA accumulation in *mos11*, the incorporation of [^{35}S]methionine into newly synthesised proteins was similar to that seen in Col-0. Perhaps the relatively mild mRNA export deficiency does not result in a detectable reduction in protein synthesis. However, the [^{35}S]methionine incorporation was clearly reduced with *tex1* and *tex1 mos11* protoplasts, indicating that mRNA biogenesis at a step other than export from the nucleus is affected in these plants and that it is caused by the mutation of *TEX1*.

Since many splicing factors were found to interact with TEX1 (Table 1), we tested whether TEX1 might be linked to mRNA splicing. First, we analysed a possible co-localisation of TEX1 and the SR family splicing factors RSZ22 and RSZ33. Leaf cell protoplasts of plants expressing TEX1-GFP (described above) were transiently transformed with plasmids driving the expression of RSZ22-RFP or RSZ33-RFP [39] and analysed by fluorescence microscopy (Fig. 9A,B). While RSZ33 was detected exclusively in the nucleoplasm, RSZ22 in some cells was also located weakly in the nucleolus, which may be due to its dynamic distribution [39,40]. As reported before [39], both SR proteins display a speckled pattern within the nucleoplasm that is similar to that of TEX1. Superimposing the RSZ22/33-RFP and the TEX1-GFP signals demonstrated that there is substantial co-localisation of TEX1 and

these SR proteins (Fig. 9A,B). Second, we examined whether TEX1 (and MOS11) influence alternative splicing events. We analysed the transcripts of three genes encoding SR proteins (At1g09140, At3g61860 and At1g55310 coding for SR30, RS31 and SCL33, respectively), as they are extensively alternatively spliced [41] and are often used to test for alternative splicing. The transcripts were amplified by RT-PCR from RNA isolated from the different mutant lines and Col-0, and assayed using a Bioanalyser (Fig. S6). No changes relative to Col-0 were observed in the ratio of splice variants with *mos11* (Fig. 9C-E). In case of At1g09140 relative to Col-0 a reduced amount of the variant with an alternative 3´ splice site was detected similarly with *tex1* and *tex1 mos11* (Fig. 9C), while with At3g61860 an increased quantity of the variant with an alternative 3´ splice site was measured (Fig. 9D). No alteration in the ratio of splice variants was detected with an exon skipping event in At1g55310 (Fig. 9E). The co-localisation of TEX1 with splicing factors and the observed changes in the ratio of alternative splicing events in *tex1* (and *tex1 mos11*) indicate that TEX1 can influence mRNA splicing in *Arabidopsis*.

Discussion

The *Arabidopsis* THO core complex was isolated and characterised from seedlings revealing its composition and that two subunits (THO5A/B and THO7A/B) apparently occur in alternative variants [18]. Our proteomic analysis of the complex isolated from cultured cells is fully consistent with this finding. Beyond that the reciprocal tagging approach with TEX1, UAP56 and MOS11 shed light upon the composition of the *Arabidopsis* TREX complex and also allowed identifying other interacting proteins. Hence the THO core complex associates with UAP56, MOS11 and ALY2-4, resembling the situation in yeast and metazoa [9,13,16,17]. Additional TREX interactors found in yeast (NAB2, NLP3) or mammals (CHTOP, LUZP4) apparently do not have orthologues in *Arabidopsis* and accordingly were not identified in our experiments. UAP56 was the only interactor recognised in the MOS11-SG affinity purifications, which is in line with the direct interaction of the two proteins seen *in vitro* and in yeast two-hybrid analyses [19]. In human cells there is in addition to UAP56 a closely related paralog termed URH49/DDX39, which appears to be the preferred binding partner of CIP29 [42]. Since in *Arabidopsis* there is only a single type of UAP56 protein (encoded by two distinct genes, [19]) it likely is the most prominent interactor of MOS11. In addition to TREX proteins, we robustly identified in our UAP56-SG purification the THP1and SAC3subunits of the TREX-2 mRNA export complex. This interaction was confirmed by the reciprocal purification of THP1-SG, and it is consistent with the co-purification of yeast Sub2 and SAC3 [43]. However, we did not detect other subunits of the putative *Arabidopsis* TREX-2, although physical interactions with CEN1/2 and DSS1 were reported based on yeast two-hybrid and bi-molecular fluorescence complementation assays [44]. In agreement with this finding, *in silico* analyses of the *Arabidopsis* orthologues of SAC3/GANP (which is the central scaffold subunit of TREX-2) uncovered that there is only limited and local sequence conservation (Fig. S7). Thus, the *Arabidopsis* SAC3 proteins contain a SAC3-GANP domain that in mammals/yeast interacts

with PCID2/THP1 and DSS1/Sem1, but the CID region interacting with CENs/Cdc31 and ENY2/Sus1 [33] does not occur in any land plant protein sequence (Fig. S7). Therefore, the composition/assembly of the plant TREX-2 complex may differ from its yeast/mammalian counterparts and/or it could be unstable under our experimental conditions.

Strikingly, we co-purified with TEX1-SG multiple splicing-related proteins. These include subunits of various spliceosomal complexes such as U1, U2, U5, RES and NTC, suggesting that TREX associates with the spliceosome throughout the mRNA splicing process. Consistent with our findings, TREX components have been identified in various proteomics analyses of metazoan spliceosomes of different catalytic stages [30,31,45,46]. In yeast cells, the splicing complex Prp19/NTC functions also in transcript elongation and is required for TREX occupancy at transcribed genes [47]. The SYF1 subunit of Prp19/NTC biochemically interacts with TREX [47] and SYF1 as well as other subunits of the NTC complex were reproducibly identified in our TEX1-SG purifications. In the literature, the RNA helicase UAP56 (Sub2) has been implicated in various functions related to mRNA splicing [48] such as promoting the U2 snRNP-branchpoint interaction [49] and facilitating spliceosome assembly [50]. However, in our UAP56-SG affinity purifications we did not observe an enrichment of splicing factors. In agreement with the interaction between TREX and spliceosome, TREX components were also found to localise to nuclear speckles in mammalian cells [13,51]. *Arabidopsis* HPR1 was reported to co-localise with SR33 [23] and SR33 (along with other SR proteins) co-purified with TEX1-SG in our proteomics analyses. Moreover, TEX1 co-localised with the splicing factors RSZ22 and RSZ33. Evidence that plant TREX is involved in mRNA splicing came from the observation that in *Arabidopsis* mutants defective in HPR1 and THO2 splicing patterns different from wild type plants were detected for some alternatively spliced genes [20,24]. Likewise our analyses revealed a quantitative change in alternatively spliced transcript variants in *tex1* and *tex1 mos11* (but not in *mos11*) relative to Col-0. Taken together these data indicate that in *Arabidopsis* the THO component TEX1 can modulate certain mRNA splicing events. It will be attractive to examine in future experiments these effects more globally to learn how prevalent *Arabidopsis* TREX affects mRNA splicing genome-wide.

The *Arabidopsis* TREX components HPR1, THO2 and TEX1 are required for efficient biogenesis of small RNAs [18,20,21,24]. In our experiments, reduced levels of the endogenous TAS1 and TAS3 siRNAs were detected in *tex1* plants when compared with Col-0, which is consistent with previous reports [18,21]. Similar levels of both RNAs were measured in *tex1* and the *tex1 mos11* double-mutant, while *mos11* was not affected. This indicates that the reduced TAS RNA levels are caused by the *tex1* mutation and that MOS11 seems to be not required for the production of these RNAs. This was different with the tested miRNAs levels. Here *tex1* acted like Col-0, whereas in *mos11* and *tex1 mos11* reduced amounts of miR173 were detected, suggesting that MOS11 is required for efficient biogenesis of this miRNA. It is still needs to be clarified at which step of the small RNA biogenesis is influenced by the TREX components. Analysis of the subcellular distribution of bulk mRNA revealed relative to Col-0 a nuclear accumulation of mRNAs in *mos11* but not in *tex1* plants, suggesting that

MOS11 is required for efficient mRNA export as previously reported [29]. We have shown here that MOS11 has RNA-binding activity, but it remains to be investigated, whether binding of MOS11 to mRNAs is required to promote their export from the nucleus. Despite the fact that the *tex1* single-mutant did not show an export defect, it resulted in a clearly increased mRNA accumulation (relative to *mos11*) in the *tex1 mos11* double-mutant, suggesting that *TEX1* and *MOS11* synergistically influence mRNA export. The finding that the incorporation of [^{35}S]methionine into newly synthesised proteins is similarly reduced in *tex1* and *tex1 mos11*, but not in *mos11*, indicates that this is not due to decreased mRNA export. Possibly the above-mentioned mRNA splicing defects (provided they occur prevalently) perhaps in combination with other effects on mRNA levels/localisation may cause the reduced protein synthesis rates.

When compared to the Col-0 wild type the phenotype of *tex1* and *mos11* single-mutants is rather mildly affected, while the vegetative and reproductive development of the *tex1 mos11* double-mutant is clearly more severely altered. Thus, *tex1 mos11* plants display various morphological features that are synergistically affected in the double-mutant including number and architecture of inflorescences, trichome development and seed set.

In the functional assays that we performed, a synergistic effect of *tex1* and *mos11* was only observed for mRNA export, but not for biogenesis of small RNAs, mRNA splicing and protein synthesis. Therefore, defects in mRNA export may be critical for the more severe phenotype of the *tex1 mos11* double-mutant, although deficiencies in other aspects of RNA biogenesis likely contribute as well. In yeast cells lacking the THO component HPR1, multicopy *THO1* (yeast orthologue of *MOS11*) could suppress all tested phenotypes including the mRNA export defect [14]. The fact that various phenotypes are affected to different extents in the *Arabidopsis tex1 mos11* double-mutant relative to the corresponding single-mutants could be due to tissue- and/or developmental stage-specific effects or that the two proteins influence the abundance/localisation of a distinct but overlapping set of RNAs. In conclusion, our experiments provide evidence that *Arabidopsis* TEX1 and MOS11 are involved in various steps of the biogenesis of mRNA and small RNAs, and that they interact regarding some aspect, but act independently in others.

Materials and Methods

Plasmid constructions

The required gene or cDNA sequences were amplified by PCR with KAPA DNA polymerase (PeqLab) using *Arabidopsis thaliana* cDNA or genomic DNA as template and the primers (providing also the required restriction enzyme cleavage sites) listed in Table S6. The PCR fragments were inserted into suitable plasmids using standard methods. All plasmid constructions were checked by DNA sequencing, and details of the plasmids generated in this work are summarized in Table S6.

Plant material

Arabidopsis thaliana (Col-0) was grown in soil at 22°C in a growth chamber under long-day (LD) or short-day (SD) conditions (16h or 8h photoperiod per day, respectively). Seeds of the T-DNA insertion lines *tex1-4* [18,21], *mos11-2* [29] were obtained from the European Arabidopsis stock centre (http://www.arabidopsis.info/), and those of *pid-14* (SALK_049736) [52] were a gift from Claus Schwechheimer (München, Germany). After sowing, seeds were stratified in darkness for 48h at 4°C prior to incubation in the plant growth chamber. *tex1 mos11* double-mutants were generated by crossing the parental lines as previously described [53]. Using the plasmid pGreenII-pTEX1::TEX1-GFP (Table S6) and *Agrobacterium*-mediated transformation as previously described [54,55], *tex1* plants were generated that express TEX1-GFP. For root analyses plants were grown on vertically oriented MS plates [56] in a plant incubator (Perzival Scientific) under LD conditions as described before [26].

PCR-based genotyping and reverse-transcribed PCR (RT-PCR)

To distinguish between plants being wild type, heterozygous, or homozygous for the T-DNA insertions, genomic DNA was isolated from leaves. The genomic DNA was used for PCR analysis with *Taq* DNA polymerase (PeqLab) and primers specific for DNA insertions and the target genes (for primer sequences, see Table S6). For RT-PCR total RNA was extracted from ~100 mg of frozen plant tissue using the TRIzol method (Invitrogen), before the RNA samples were treated with DNAse [53]. Reverse transcription was performed using 2 µg of RNA and Revert Aid H minus M-MuLV reverse transcriptase (Thermo Scientific). The obtained cDNA was used for PCR analysis using *Taq* DNA polymerase (PeqLab) and gene-specific primers (Table S6). Amplified DNA fragments were analysed by agarose gel electrophoresis, except for splicing analyses when a Bioanalyser 2100 (Agilent) was used.

Light microscopy

To document plant phenotypic details, a Zeiss Discovery V8 stereo microscope was used.

For co-localisation analyses of TEX1 and RSZ proteins, leaf protoplasts of *tex1* seedlings harbouring pGreenII-pTEX1::TEX1-GFP were prepared and transformed with plasmids pBI35S::atRSZp22-mRFP1 and pBI35S::atRSZp33-mRFP1([39] a gift from Patrick Motte, Liège, Belgium) as previously described [57]. To visualise GFP and RFP fusion proteins in protoplasts and intact cells by confocal laser scanning microscopy (CLSM) a Leica SP8 microscope was used as previously described [58], equipped with a 63 x / 1.4 objective. GFP was excited at 488 nm and emission was detected at 500-530 nm, while RFP was exited at 561 nm and detected at 570 -630 nm.

Scanning electron microscopy (SEM)

Shoot apex and leaf samples were prepared for SEM analysis as previously described [53].

Affinity purification and characterisation of SG-tagged proteins from *Arabidopsis* cells

Arabidopsis suspension cultured PSB-D cells were maintained and transformed as previously described [25]. Protein isolation, purification and mass spectrometric analyses were essentially performed as previously described [26]. In brief, proteins of 15 g cells were extracted (samples were treated with Benzonase (50u/ml extract) as indicated) and SG-tagged proteins were affinity purified using IgG-coupled magnetic beads [59]. Eluted proteins were analysed by SDS-PAGE and digested with trypsin. Peptides were separated on an Ultimate 3000 RSLC nano HPLC System (Dionex) by reversed-phase chromatography using an AcclaimPepmap 100 C18 nano column (ThermoScientific). The LC System was coupled to a MaXis 4G UHR-Q TOF-system (Bruker Daltonics) via a nano electrospray source (Bruker Daltonics). Up to five most abundant precursor ions were selected for fragmentation by collisional dissociation. The data were launched to Mascot using the ProteinScape software (Bruker Daltonics). Mascot (v2.3.02) was used to search the NCBI nr protein data base. The criterion for reliable protein identification were Mascot scores >85. The experimental background of contaminating proteins that were isolated with the unfused SG-tag or that co-purify non-specifically independent of the used bait protein was subtracted. The list of 760 known non-specific *Arabidopsis* proteins is based on 543 affinity purifications with 115 bait proteins [25]. The remaining proteins with a mean score of ≥100 with at least two detected peptides with a score >25 were retained as possible interactors.

Protein/RNA interactions analysed by MicroScale Thermophoresis (MST)

Using plasmid pQE9-MOS11 [19] 6xHis-tagged MOS11 was expressed in *E. coli* M15 cells and purified on Ni-NTA-agarose as previously described [19,60]. The recombinant protein was verified by SDS-PAGE and mass spectrometry. MST binding experiments were carried out with 200 nM Cy3-labeled ssRNA or dsRNA oligonnucleotides in binding buffer (10 mM sodium phosphate, PH 7.0; 1 mM DTT; 1 mM EDTA; 0.5 mM PMSF) with a range of concentrations of recombinant MOS11 (0,916 – 30000 nM) at 20% MST power, 20% LED power in standard capillaries on a Monolith NT.115 device at 25°C (NanoTemper Technologies). A ligand-dependent fluorescence enhancement effect was detected via a protein denaturation test. The raw fluorescence was thus used for data analysis. The recorded fluorescence was normalised to fraction bound (0 = unbound, 1 = bound), and processed using the KaleidaGraph 4.5 software and fitted using the K_D fit formula derived from the law of mass action.

Detection of small RNAs by Northern blot analysis

Small RNAs were analysed by Northern blotting following a previously described method [61]. 15 µg of total RNA were separated on a 12% denaturing polyacrylamide gel and following transfer onto the membrane and crosslinking, hybridisation in QuikHyb solution (Agilent) was performed with [^{32}P]-labelled probes (Table S6). Blots were scanned using a Typhoon FLA 9500 imager (GE Healthcare) and signals were quantified using ImageJ.

Whole mount *in situ* hybridisation

To determine the relative distribution of bulk mRNA in nuclei and cytosol a previously described protocol was adopted [62] using 6 d old seedlings. Hybridisation was performed in PerfectHyb plus solution (Sigma) with an Alexa Fluor 488-labelled 48-nt oligo(dT) probe. Seedling roots were analysed using CLSM with a Leica SP8 microscope. Quantification of the fluorescence signals in cytosol and nucleoplasm was done in ImageJ, analysing areas of same size in the nucleus and in the cytosol adjacent to the nucleus.

Metabolic pulse-labelling with [^{35}S]methionine

Seedling protoplasts were prepared as previously described [63] and $1 \cdot 10^6$ protoplasts per ml were incubated for 1h at 22°C with 100 µCi L-[35S]-methionine. Protein extracts were separated by SDS-PAGE and gels were stained with Coomassie, dried and scanned using a Typhoon FLA 9500 imager (GE Healthcare) and signals were quantified using ImageJ.

Acknowledgements

We thank Jelle Van Leene and Geert De Jaeger for advice regarding cell culture transformation and analysis of proteomics data as well as providing *Arabidopsis* PSB-D cells, Lianng-zi Zhou for help with the protoplast transformation experiment, Patrick Motte for plasmids pBI35S::atRSZp22-mRFP1 and pBI35S::atRSZp33-mRFP1, Claus Schwechheimer for *pid-14* seeds, Eduard Hochmuth for recording mass spectra, the Nottingham Arabidopsis Stock Centre (NASC) for providing *Arabidopsis* T-DNA insertion lines. This work was supported by the German Research Foundation (DFG) through Grant SFB960 to M.G. and K.D.G.

Author Contributions

Conceived and designed the experiments: BBS, MG, KDG. Performed the experiments: BBS, HFE, SE, AP, TS, MM, MG. Analysed the data: BBS, HFE, SE, AP, AB, GM, RM, TS, GL, MM, MG, KDG. Contributed reagents/materials/analysis tools: JH, GM, RM, TS, GL, MM, MG, KDG. Wrote the paper: MG, KDG.

References

1. Hsin JP, Manley JL (2012) The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev 26: 2119-2137.

2. Moore MJ, Proudfoot NJ (2009) Pre-mRNA processing reaches back to transcription and ahead to translation. Cell 136: 688-700.

3. Walsh MJ, Hautbergue GM, Wilson SA (2010) Structure and function of mRNA export adaptors. Biochem Soc Trans 38: 232-236.

4. Köhler A, Hurt E (2007) Exporting RNA from the nucleus to the cytoplasm. Nat Rev Mol Cell Biol 8: 761-773.

5. Wickramasinghe VO, Laskey RA (2015) Control of mammalian gene expression by selective mRNA export. Nat Rev Mol Cell Biol 16: 431-442.

6. Katahira J (2012) mRNA export and the TREX complex. Biochim Biophys Acta 1819: 507-513.

7. Luna R, Rondón AG, Aguilera A (2012) New clues to understand the role of THO and other functionally related factors in mRNP biogenesis. Biochim Biophys Acta 1819: 514-520.

8. Meinel DM, Burkert-Kautzsch C, Kieser A, O´Duibhir E, Siebert M, Mayer A, Cramer P, Söding J, Holstege FC, Sträßer K (2013) Recruitment of TREX to the transcription machinery by its direct binding to the phospho-CTD of RNA polymerase II. PLoS Genet 9: e1003914.

9. Sträßer K, Masuda S, Mason P, Pfannstiel J, Oppizzi M, Rodriguez-Navarro S, Rondón AG, Aguilera A, Struhl K, Reed R, Hurt E (2002) TREX is a conserved complex coupling transcription with messenger RNA export. Nature 417: 304-308.

10. Gómez-González B, García-Rubio M, Bermejo R, Gaillard H, Shirahige K, Marín A, Foiani M, Aguilera A (2011) Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. EMBO J 30: 3106-3119.

11. Kim M, Ahn SH, Krogan NJ, Greenblatt JF, Buratowski S (2004) Transitions in RNA polymerase II elongation complexes at the 3´ends of genes. EMBO J 23: 354-364.

12. Saguez C, Schmid M, Olesen JR, Ghazy MA, Qu X, Poulsen MB, Nasser T, Moore C, Jensen TH (2008) Nuclear mRNA surveillance in THO/sub2 mutants is triggered by inefficient polyadenylation. Mol Cell 31: 91-103.

13. Dufu K, Livingstone MJ, Seebacher J, Gygi SP, Wilson SA, Reed R (2010) ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. Genes Dev 24: 2043-2053.

14. Jimeno S, Luna R, García-Rubio M, Aguilera A (2006) Tho1, a novel hnRNP, and Sub2 provide alternative pathways for mRNP biogenesis in yeast THO mutants. Mol Cell Biol 26: 4387-4398.

15. Cheng H, Dufu K, Lee CS, Hsu JL, Dias A, Reed R (2006) Human mRNA export machinery recruited to the 5' end of mRNA. Cell 127: 1389-1400.

16. Masuda S, Das R, Cheng H, Hurt E, Dorman N, Reed R (2005) Recruitment of the human TREX complex to mRNA during splicing. Genes Dev 19: 1512-1517.

17. Rehwinkel J, Herold A, Gari K, Köcher T, Rode M, Ciccarelli FL, Wilm M, Izaurralde E (2004) Genome-wide analysis of mRNAs regulated by the THO complex in *Drosophila melanogaster*. Nat Struct Mol Biol 11: 558-566.

18. Yelina NE, Smith LM, Jones AME, Patel K, Kelly KA, Baulcombe DC (2010) Putative *Arabidopsis* THO/TREX mRNA export complex is involved in transgene and endogenous siRNA biosynthesis. Proc Natl Acad Sci USA 107: 13948-13953.

19. Kammel C, Thomaier M, Sørensen BB, Schubert T, Längst G, Grasser M, Grasser KD (2013) *Arabidopsis* DEAD-box RNA helicase UAP56 interacts with both RNA and DNA as well as with mRNA export factors. PLoS One 8: e60644.

20. Furumizu C, Tsukaya H, Komeda Y (2010) Characterization of EMU, the *Arabidopsis* homolog of the yeast THO complex member HPR1. RNA 16: 1809-1817.

21. Jauvion V, Elmayan T, Vaucheret H (2010) The conserved RNA trafficking proteins HPR1 and TEX1 are involved in the production of endogenous and exogenous small interfering RNA in *Arabidopsis*. Plant Cell 22: 2697-2709.

22. Pan H, Liu S, Tang D (2012) HPR1, a component of the THO/TREX complex, plays an important role in disease resistance and senescence in *Arabidopsis*. Plant J 69: 831-843.

23. Xu C, Zhou X, Wen CK (2015) HYPER RECOMBINATION1 of the THO/TREX complex plays a role in controlling transcription of the *REVERSION-TO-ETHYLENE SENSITIVITY1* gene in *Arabidopsis*. PLoS Genet 11: e1004956.

24. Francisco-Mangilet AG, Karlsson P, Kim MH, Eo HJ, Oh SA, Kim JH, Kulcheski FR, Park SK, Manavella PA (2015) THO2, a core member of the THO/TREX complex, is required for microRNA production in *Arabidopsis*. Plant J 82: 1018-1029.

25. van Leene J, Eeckhout D, Cannoot B, De Winne N, Persiau G, van de Slijke E, Vercruysse L, Dedecker M, Vandepoele K, Martens L, Witters E, Gevaert K, de Jaeger G (2015) An improved toolbox to unravel the plant cellular machinery by tandem affinity purification of *Arabidopsis* protein complexes. Nat Protoc 10: 169-187.

26. Dürr J, Lolas IB, Sørensen BB, Schubert V, Houben A, Melzer M, Deutzmann R, Grasser M, Grasser KD (2014) The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in *Arabidopsis*. Nucleic Acids Res 42: 4332-4347.

27. Nelissen H, DeGroeve S, Fleury D, Neyt P, Bruno L, Bitonti MB, Vandenbussche F, Van Der Straeten D, Yamaguchi T, Tsukaya H, Witters E, de Jaeger G, Houben A, an Lijsebettens M (2010) Plant Elongator regulates auxin-related genes during RNA polymerase II-mediated transcription elongation. Proc Natl Acad Sci USA 107: 1678-1683.

28. Pauwels L, Barbero GF, Geerinck J, Tilleman S, Grunewald W, Pérez AC, Chico JM, Bossche RV, Sewell J, Gil E, García-Casado G, Witters E, Inzé D, Long JA, de Jaeger G, Solano R, Goossens A (2010) NINJA connects the co-repressor TOPLESS to jasmonate signalling. Nature 464: 788-791.

29. Germain H, Qu N, Cheng YT, Lee E, Huang Y, Dong OX, Gannon P, Huang S, Ding P, Li Y, Sack F, Zhang Y, Li X (2010) MOS11: a new component in the mRNA export pathway. PLoS Genet 6: e1001250.

30. Chen YI, Moore RE, Ge HY, Young MK, Lee TD, Stevens WW (2007) Proteomic analysis of in vivo-assembled pre-mRNA splicing complexes expands the catalog of participating factors. Nucleic Acids Res 35: 3928-3944.

31. Zhou Z, Licklider LJ, Gygi SP, Reed R (2002) Comprehensive proteomic analysis of the human spliceosome. Nature 419: 182-185.

32. Ellisdon AM, Dimitrova L, Hurt E, Stewart M (2012) Structural basis for the assembly and nucleic acid binding of the TREX-2 transcription-export complex. Nat Struct Mol Biol 19: 328-336.

33. Jani D, Lutz S, Hurt E, Laskey RA, Stewart M, Wickramasinghe VO (2012) Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export. Nucleic Acids Res 40: 4562-4573.

34. Sugiura T, Sakurai K, Nagano Y (2007) Intracellular characterization of DDX39, a novel growth-associated RNA helicase. Exp Cell Res 313: 782-790.

35. Zillner K, Jerabek-Willemsen M, Duhr S, Braun D, Längst G, Baaske P (2012) Microscale thermophoresis as a sensitive method to quantify protein:nucleic acid interactions in solution. Methods Mol Biol 815: 241-252.

36. Bennett SRM, Alvarez J, Bossinger G, Smyth DR (1995) Morphogenesis in *pinoid* mutants of *Arabidopsis thaliana*. Plant J 8: 505-520.

37. Schnittger A, Hülskamp M (2002) Trichome morphogenesis: a cell-cycle perspective. Phil Trans R Soc Land Biol Sci 357: 826.

38. Gatfield D, Le Hir H, Schmitt C, Braun IC, Köcher T, Wilm M, Izaurralde E (2001) The DExH/D box protein HEL/UAP56 is essential for mRNA nuclear export in *Drosophila*. Curr Biol 11: 1716-1721.

39. Tillemans V, Dispa L, Remacle C, Motte P (2005) Functional distribution and dynamics of *Arabidopsis* SR splicing factors in living plant cells. Plant J 41: 567-582.

40. Rausin G, Tillemans V, Stankovic N, Hanikenne M, Motte P (2016) Dynamic nucleocytoplasmic shuttling of an *Arabidopsis* SR splicing factor: role of the RNA-binding domains. Plant Physiol 153: 273-284.

41. Palusa SG, Ali GS, Reddy AS (2007) Alternative splicing of pre-mRNAs of *Arabidopsis* serine/arginine-rich proteins: regulation by hormones and stresses. Plant J 49: 1091-1107.

42. Yamazaki T, Fujiwara N, Yukinaga H, Ebisuya M, Shiki T, Kurihara T, Kioka N, Kambe T, Nagao N, Nishida E, Masuda S (2010) The closely related RNA helicases, UAP56 and URH49, preferentially form distinct mRNA export machineries and coordinately regulate mitotic progression. Mol Biol Cell 21: 2953-2965.

43. Fischer T, Sträßer K, Rácz A, Rodriguez-Navarro S, Oppizzi M, Ihrig P, Lechner J, Hurt E (2002) The mRNA export machinery requires the novel Sac3p-Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores. EMBO J 21: 5843-5852.

44. Lu Q, Tang X, Tian G, Wang F, Liu K, Nguyen V, Kohalmi SE, Keller WA, Tsang EW, Harada JJ, Rothstein SJ, Cui Y (2010) *Arabidopsis* homolog of the yeast TREX-2 mRNA export complex: components and anchoring nucleoporin. Plant J 61: 259-270.

45. Herold N, Will CL, Wolf E, Kastner B, Urlaub H, Lührmann R (2009) Conservation of the protein composition and electron microscopy structure of *Drosophila melanogaster* and human spliceosomal complexes. Mol Cell Biol 29: 281-301.

46. Rappsilber J, Ryder U, Lamond AI, Mann M (2002) Large-scale proteomic analysis of the human spliceosome. Genome Res 12: 1231-1245.

47. Chanarat S, Seizl M, Sträßer K (2011) The Prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes. Genes Dev 25: 1147-1158.

48. Shen H (2009) UAP56- a key player with surprisingly diverse roles in pre-mRNA splicing and nuclear export. BMB Rep 42: 185-188.

49. Fleckner J, Zhang M, Valcarcel J, Green MR (1997) U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. Genes Dev 11: 1864-1872.

50. Shen H, Zheng X, Shen J, Zhang L, Zhao R, Green MR (2008) Distinct activities of the DExD/H-box splicing factor hUAP56 facilitate stepwise assembly of the spliceosome. Genes Dev 22: 1796-1803.

51. Dias AP, Dufu K, Lei H, Reed R (2010) A role for TREX components in the release of spliced mRNA from nuclear speckle domains. Nat Commun 1: 97.

52. Cheng Y, Qin G, Dai X, Zhao Y (2007) NPY1, a BTB-NPH3-like protein, plays a critical role in auxin-regulated organogenesis in *Arabidopsis*. Proc Natl Acad Sci USA 104: 18825-18829.

53. Lolas IB, Himanen K, Grønlund JT, Lynggaard C, Houben A, Melzer M, Van Lijsebettens M, Grasser KD (2010) The transcript elongation factor FACT affects *Arabidopsis* vegetative and reproductive development and genetically interacts with *HUB1/2*. Plant J 61: 686-697.

54. Launholt D, Merkle T, Houben A, Schulz A, Grasser KD (2006) *Arabidopsis* chromatin-associated HMGA and HMGB use different nuclear targeting signals and display highly dynamic localization within the nucleus. Plant Cell 18: 2904-2918.

55. Lildballe DL, Pedersen DS, Kalamajka R, Emmersen J, Houben A, Grasser KD (2008) The expression level of the chromatin-associated HMGB1 protein influences growth, stress tolerance and transcriptome in *Arabidopsis*. J Mol Biol 384: 9-21.

56. Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassay with tobacco tissue cultures. Physiol Plant 15: 473-497.

57. Wu FH, Shen SC, Lee LY, Lee SH, Chan MT, Lin CS (2009) Tape-*Arabidopsis* Sandwich - a simpler *Arabidopsis* protoplast isolation method. Plant Meth 5: 16.

58. Antosch M, Schubert V, Holzinger P, Houben A, Grasser KD (2015) Mitotic lifecycle of chromosomal 3xHMG-box proteins and the role of their N-terminal domain in the association with rDNA loci and proteolysis. New Phytol 208: 1067-1077.

59. Hamperl S, Brown CR, Perez-Fernandez J, Huber K, Wittner M, Babl V, Stöckl U, Boeger H, Tschochner H, Milkereit P, Griesenbeck J (2014) Purification of specific chromatin domains from single-copy gene loci in *Saccharomyces cerevisiae*. Methods Mol Biol 1094: 329-341.

60. Grasser KD, Grimm R, Ritt C (1996) Maize chromosomal HMGc: two closely related structure-specific DNA-binding proteins specify a second type of plant HMG-box protein. J Biol Chem 271: 32900-32906.

61. Pall GS, Hamilton AJ (2008) Improved northern blot method for enhanced detection of small RNA. Nat Protoc 3: 1077-1084.

62. Gong Z, Dong CH, Lee H, Zhu J, Xiong L, Gong D, Stevenson B, Zhu JK (2005) A DEAD box RNA helicase is essential for mRNA export and important for development and stress responses in *Arabidopsis*. Plant Cell 17: 256-267.

63. Zhai Z, Jung HI, Vatamaniuk OK (2009) Isolation of protoplasts from tissues of 14-day-old seedlings of *Arabidopsis thaliana*. J Vis Exp 17.

64. Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann JU (2005) A gene expression map of *Arabidopsis* thaliana development. Nat Genet 37: 501-506.

65. Söding J (2005) Protein homology detection by HMM-HMM comparison. Bioinformatics 21: 951-960.

66. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Bateman A (2016) The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44: D279-285.

Figure Legends

Figure 1. Isolation of TREX and TREX-2 complexes. (A) Protein extracts (treated with benzonase) of cells expressing unfused SG, TEX1-SG, UAP56-SG and MOS11-SG after SDS-PAGE and Coomassie-staining of the gel. **(B)** Eluates of the affinity purifications after SDS-PAGE and Coomassie-staining of the gel. The unfused SG-tag and SG-fusion proteins are indicated by asterisks. **(C)** Eluates of the affinity purifications of unfused SG and THP1-SG, which were done in a separate set of experiments, after SDS-PAGE and Coomassie-staining of the gel. The unfused SG-tag and THP1-SG protein are indicated by asterisks.

Figure 2. Scheme depicting the composition of the *Arabidopsis* TREX complex and its interaction with TREX-2 and splicing machinery. Components of the THO core complex (yellow symbols) associate with UAP56, ALYs and MOS11 (green symbols) to form *Arabidopsis* TREX (grey sphere). Most likely UAP56 mediates the association of ALYs and MOS11with THO, as ALYs and MOS11 associate directly with UAP56 (red arrows), but it is unclear how UAP56 exactly interacts with THO. Black arrows depict protein interactions indicated by co-purification with reciprocally tagged proteins (red characters). These experiments suggest that TREX interacts with TREX-2 (orange symbols) and the splicing machinery (blue symbols).

Figure 3. *tex1* plants display various phenotypic alterations. When compared with Col-0, *tex1* plants are early bolting under LD (**A**, 24 days after stratification (DAS)) and SD conditions (**B**, 100 DAS). (**C**) 40 DAS *tex1* plants grown under LD conditions show an increased number of primary inflo-

rescences. (**D**) Roots of 8 DAS plants with *tex1* displaying a reduced lateral root density. The bolting phenotype of Col-0, *tex1* and *tex1* harbouring the construct *pTEX::TEX1-GFP* (plants of three independent lines are shown) under LD (**E**, 24 DAS) and SD conditions (**F**, 100 DAS).

Figure 4. Co-localisation of TEX1 and MOS11 in *Arabidopsis* nuclei. *tex1* plants harbouring constructs that drive the expression of TEX1-GFP and MOS11-RFP under control of the respective native promoters were analysed by fluorescence microscopy. (**A, B**) Overview of root tips showing fluorescent images (left) and the overlay with the bright field image (right). Co-localisation analysis of TEX1-GFP and MOS11-RFP in nuclei of root (**C**) and leaf cells (**D**). A merge of both individual fluorescent signals is shown on the right. Scale bars represent 100 µm (**A,B**) and 10 µm (**C,D**).

Figure 5. MOS11 binds ssRNA and dsRNA. (**A**) 6xHis-tagged MOS11 was expressed in *E. coli* and purified by metal-chelate chromatography. SDS-PAGE and Coomassie-staining revealed in addition to the full-length protein (arrow head) two faster migrating bands. All three bands were identified as MOS11 polypeptides by mass spectrometry, suggesting that the protein is partially degraded, a feature that we observed in several independent protein preparations. (**B**) Increasing concentrations of recombinant MOS11 were analysed for interaction of ssRNA and dsRNA by MST.

Figure 6. The *tex1 mos11* double-mutant is affected in various phenotypic features. (**A**) Phenotype of 60 DAS plants of the different genotypes grown under LD conditions. (**B**) Individual primary inflorescences (60 DAS) and on the right three different *tex1 mos11* inflorescences are shown. (**C-F**) Inflorescence apices (30 DAS) of the first appearing inflorescence and in (**G**) a magnification of the boxed region in F is shown. SEM images of the inflorescence apices of Col-0 (**H**), of two different tex1 mos11 examples (**I,J**) and of pod (**K**). SEM images of trichomes of the different genotypes (**L-O**). Scale bars represent 1 mm (**H-K**) and 90 µm (**L-O**).

Figure 7. *TEX1* and *MOS11* influence the biogenesis of small RNAs. (**A**) Determination of the levels of small RNAs by Northern blot analysis. The hybridisation with an U6 probe and a stain of the gel prior to blotting are shown as loading controls. (**B-E**) Quantification of the levels of the four small RNAs from four independent experiments. Blots were scanned and the signals were normalised relative to the U6 signal, and data were analysed by two-way ANOVA. Error bars indicate standard deviation.

Figure 8. Analysis of bulk mRNA export and protein synthesis. (**A**) mRNA export assay of root cells using whole mount *in situ* hybridisation with a fluorescently-labelled 48mer oligo (dT) probe (green). DAPI-stain (blue) and the merge of the two signals is shown (right panels). Scale bars represent 10µm. (**B**) The fluorescent hybridisation signal of nuclei relative to cytosol was quantified for 60 cells of each genotype. Data were analysed by two-way ANOVA and error bars indicate standard deviation. (**C,D**) Semi-quantitative estimate of the incorporation of [^{35}S]methionine into newly synthesised proteins in the different genotypes. After 1h incubation of seedling protoplasts with [^{35}S]methionine proteins were extracted and separated by SDS-PAGE. Gels were stained with Coo-

massie and the radiolabel was analysed using an imager. Following line scans of both images, the ratios of [^{35}S] to Coomassie signal were calculated (numbers below the [^{35}S] scans, normalised to Col-0). Representative gels from three repetitions are shown.

Figure 9. TEX1 co-localises with splicing factors and influences alternative splicing events. (**A,B**) Co-localisation of TEX1-GFP and RSZ22/33-RFP. Constructs driving the expression of RSZ22-RFP or RSZ33-RFP were transiently expressed in protoplasts of leaf cells derived from plants expressing TEX1-GFP. GFP and RFP fluorescence was analysed by CLSM and a merge of both signals is shown in the right panels. Scale bars represent 5 µm. (**C-E**) Quantification of the ratio of two different splicing variants each with three candidate genes. The splicing variants of the three genes were amplified by RT-PCR from RNA of the different genotypes and quantified using a Bioanalyser. The ratio of the alternatively spliced variant relative to the constitutively spliced variant was calculated and normalised to Col-0. The data of three independent experiments were analysed by two-way ANOVA and error bars indicate standard deviation.

Figure 1

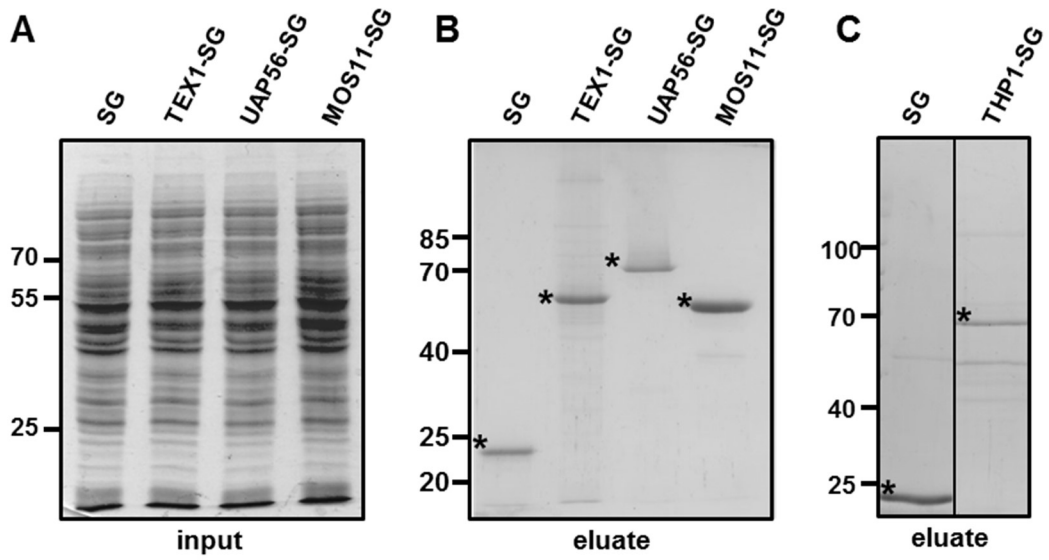


Figure 2

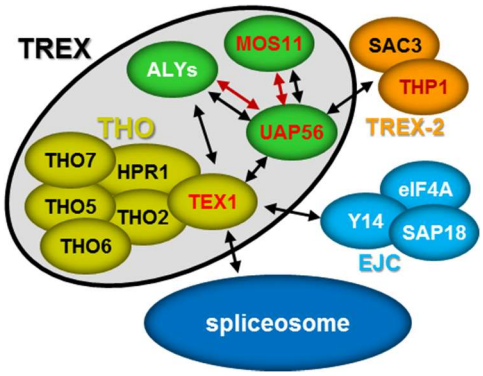


Figure 3

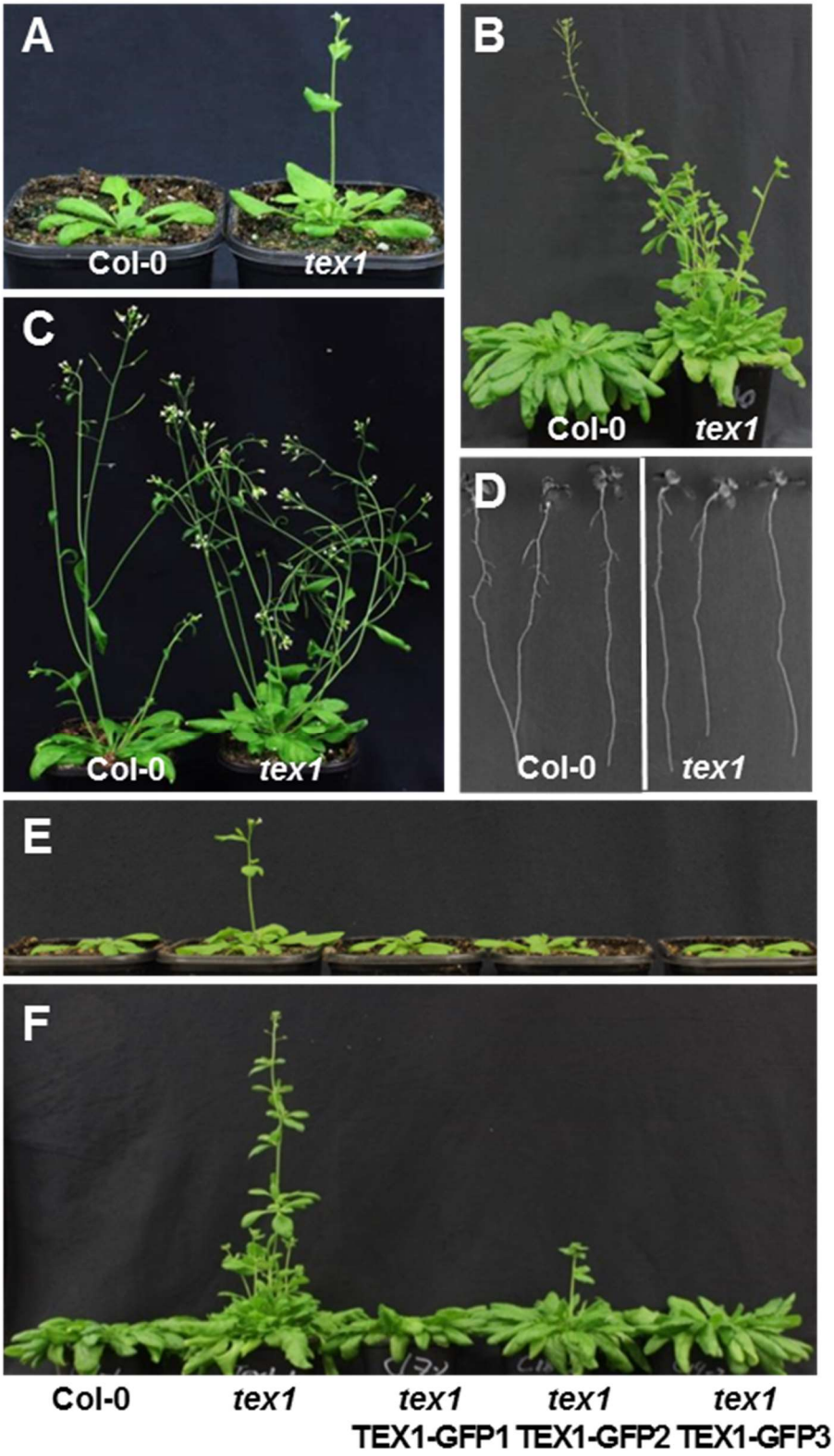

A
Col-0
tex1
B
Col-0
tex1
C
Col-0
tex1
D
Col-0
tex1
E
F
Col-0
tex1
tex1
TEX1-GFP1
tex1
TEX1-GFP2
tex1
TEX1-GFP3

Figure 4

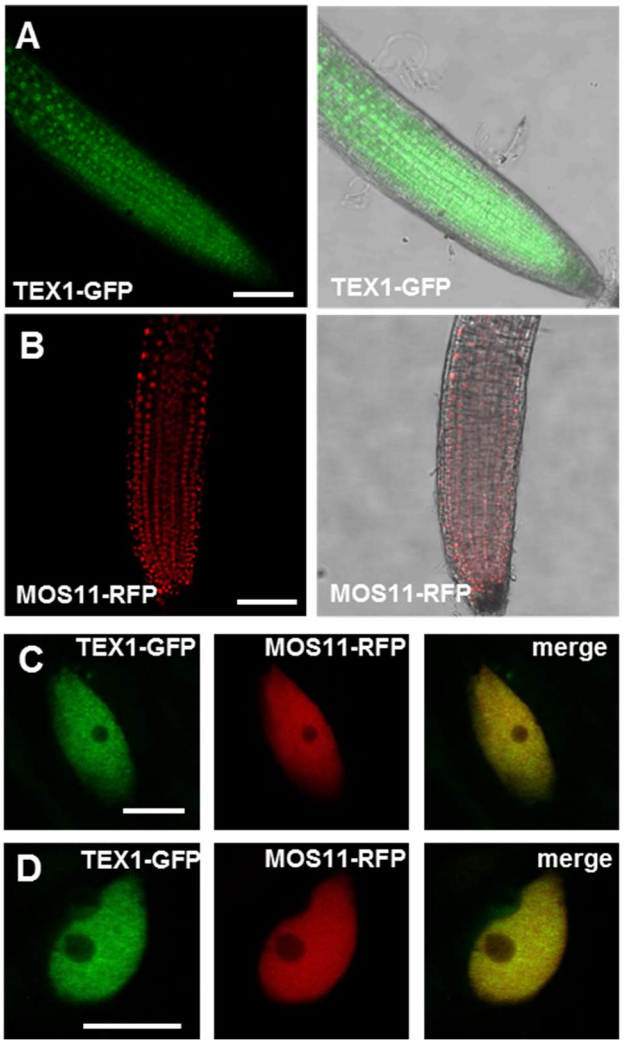


Figure 5

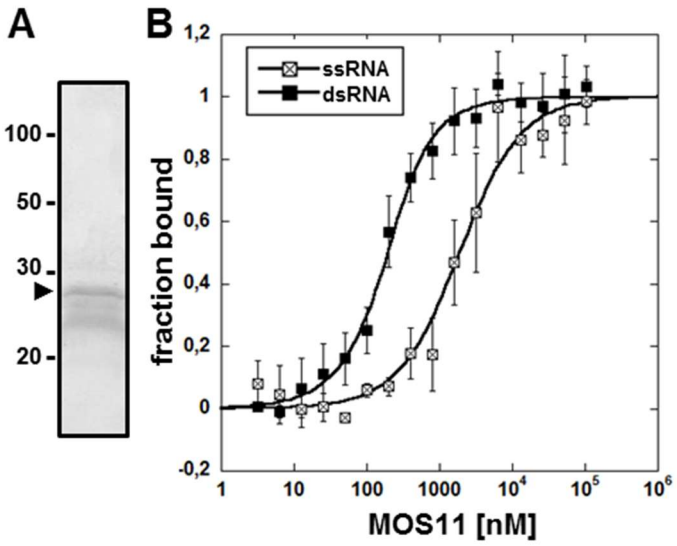


Figure 6

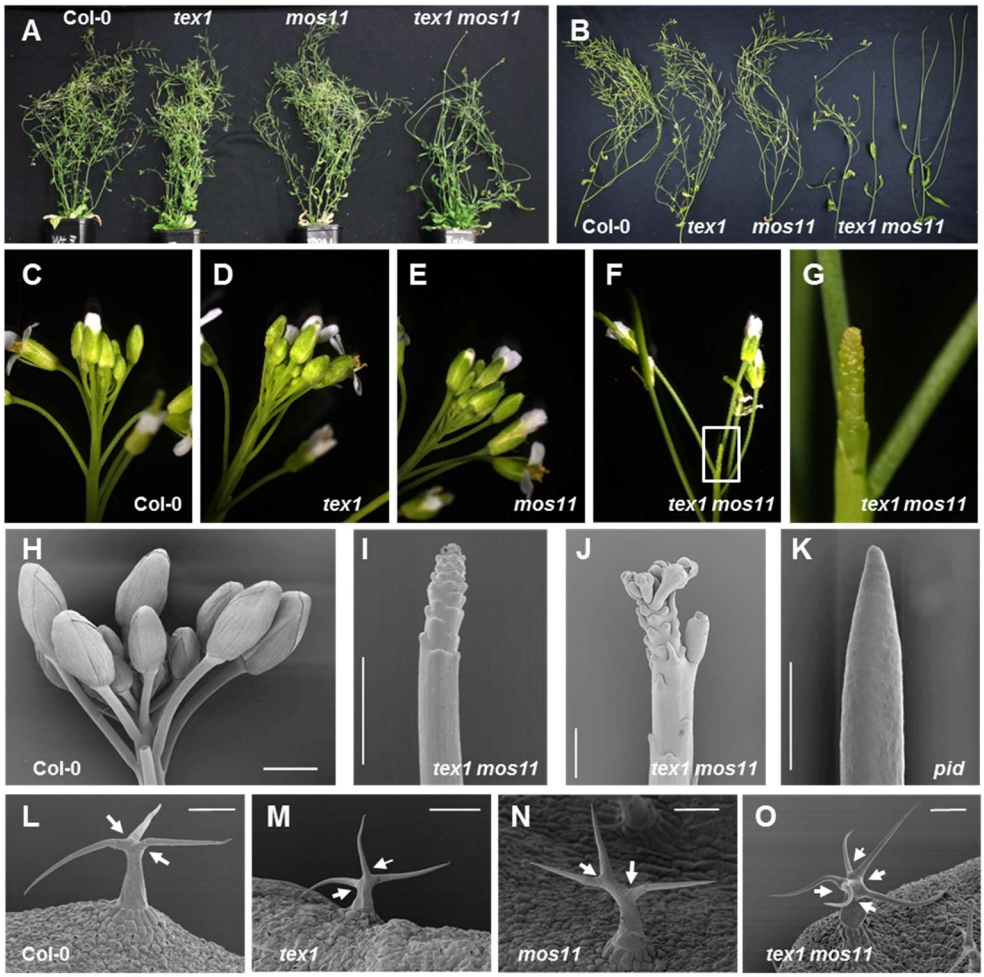


Figure 7

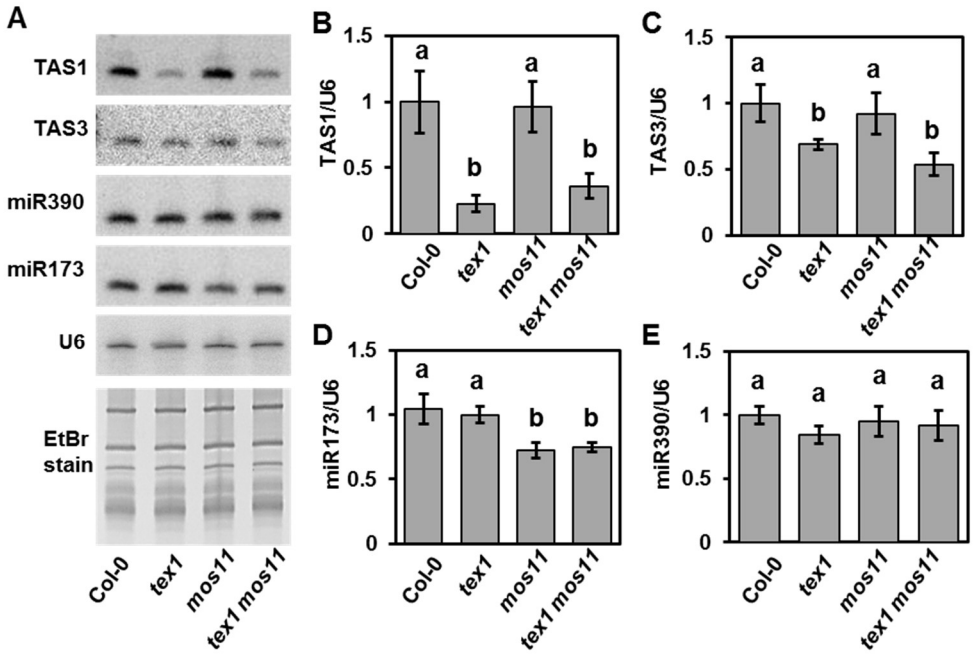


Figure 8

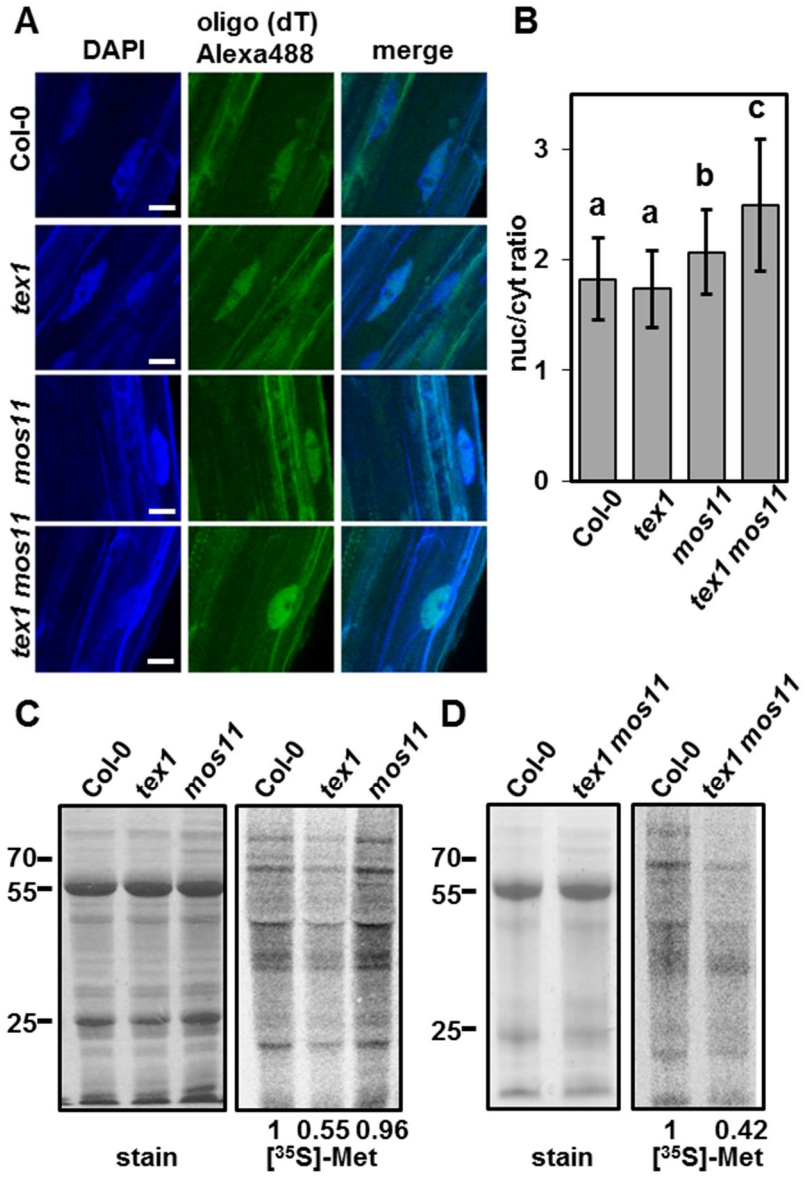

A
DAPI
oligo (dT)
Alexa488
merge
Col-0
tex1
mos11
tex1 mos11
B
nuc/cyt ratio
0
1
2
3
a
a
b
c
Col-0
tex1
mos11
tex1 mos11
C
Col-0
tex1
mos11
Col-0
tex1
mos11
70
55
25
stain
1 0.55 0.96
[35S]-Met
D
Col-0
tex1 mos11
Col-0
tex1 mos11
70
55
25
stain
1 0.42
[35S]-Met

Figure 9

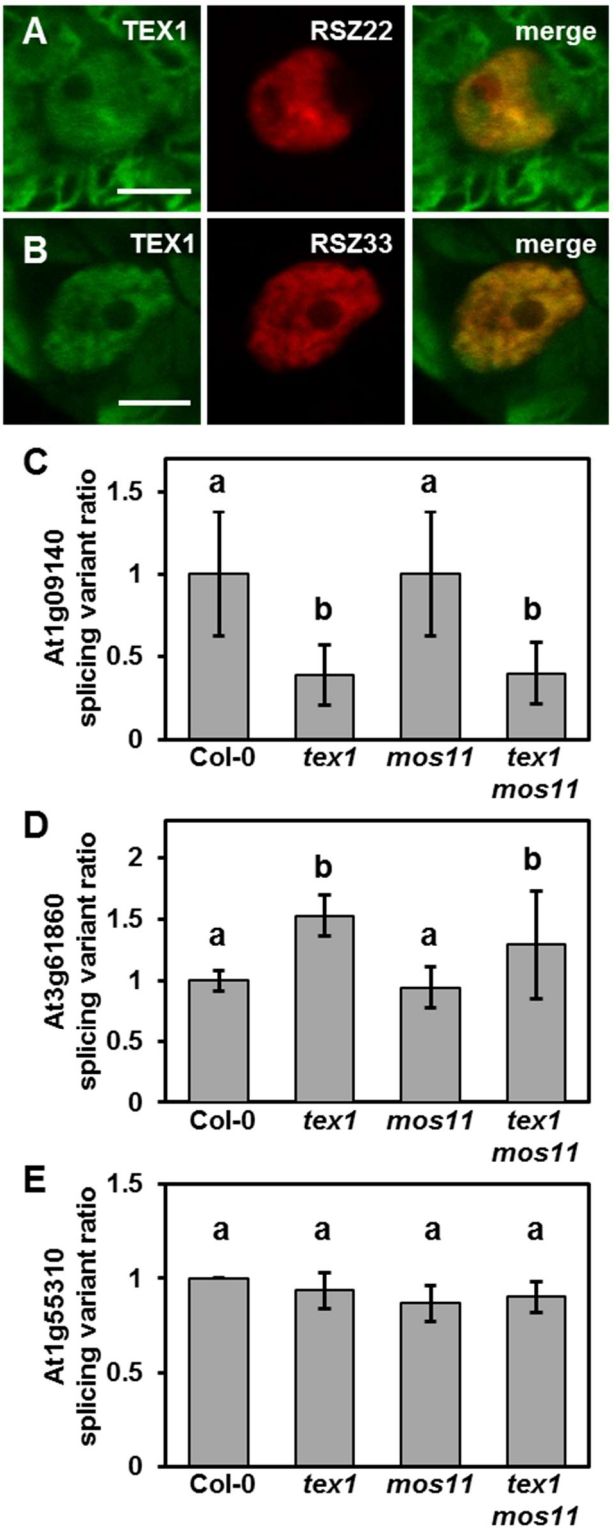

A
TEX1
RSZ22
merge
B
TEX1
RSZ33
merge
C
At1g09140
splicing variant ratio
1.5
1
0.5
0
a
b
a
b
Col-0
tex1
mos11
tex1
mos11
D
At3g61860
splicing variant ratio
2
1.5
1
0.5
0
a
b
a
b
Col-0
tex1
mos11
tex1
mos11
E
At1g55310
splicing variant ratio
1.5
1
0.5
0
a
a
a
a
Col-0
tex1
mos11
tex1
mos11

Table 1 Proteins that copurify with TEX1, UAP56 and MOS11

TEX1[1]	UAP56[1]	MOS11[1]	Interactor	AGI	Complex[2]	Function[3]
3355/4	5853/4		THO2	AT1G24706	TREX	E
1922/4	3421/4		THO5A	AT5G42920	TREX	E
1815/4	1013/4		TEX1	AT5G56130	TREX	E
1534/4	2379/4		HPR1,THO1	AT5G09860	TREX	E
452/4	1684/4		THO7B	AT5G16790	TREX	E
430/4	1598/4		THO7A	AT3G02950	TREX	E
341/4	1013/4		THO6, DWA1	AT2G19430	TREX	E
339/4	1634/4		THO5B	AT1G45233	TREX	E
	2023/4	3123/4	MOS11	AT5G02770	TREX	E
492/3			ALY3	AT1G66260	TREX	E
	405/4		ALY2	AT5g02530	TREX	E
	274/4		ALY4	AT5g37720	TREX	E
131/3	4491/4	771/4	UAP56	AT5G11170	TREX	E,S
	685/4		THP1, ESSP1	AT2G19560	TREX-2	E
	629/4		SAC3C	AT3G54380	TREX-2	E
478/4	282/4		eIF4A3-1	AT3G19760	EJC	S
336/3			eIF4A3-2	AT1G51380	EJC	S
206/3			SAP18	AT2G45640	EJC	S
109/3			Y14	AT1G51510	EJC	S
166/3			Luc7b	AT5G17440	U1 snRNP	S
116/3			U1A	AT2G47580	U1 snRNP	S
500/4			SAP155	AT5G64270	17S U2 snRNP	S
337/4			SAP130a	AT3G55200	17S U2 snRNP	S
142/4			SF3b150	AT4G21660	17S U2 snRNP	S
140/3			SAP114-1a	AT1G14650	17S U2 snRNP	S
301/3			SR140-1	AT5G25060	17S U2 assoc.	S
117/3			PRP24, EMB140	AT4G24270	U4/U6 snRNP	S
768/4			U5-220/Prp8a	AT1G80070	U5 snRNP	S
222/4			U5-102KD	AT4G03430	U5 snRNP	S
934/3			U5-200-2a	AT1G20960	U5 snRNP	S
196/3			U5-40	AT2G43770	U5 snRNP	S
532/4			CDC5, MOS4	AT1G09770	Core NTC	S
441/4			SYF1	AT5G28740	Core NTC	S
297/4			MAC3B, PRP19B	AT2G33340	Core NTC	S
256/4			SKIP	AT1G77180	Core NTC	S
290/3			CRN1c	AT5G41770	Core NTC	S
186/4			U2AF35a	AT1G27650	Splice site select.	S
142/3			U2AF65b	AT1G60900	Splice site select.	S
137/3			U2AF65a	AT4G36690	Splice site select.	S
303/4			RSp32	AT2G46610	SR proteins	S
292/4			SR34a	AT3G49430	SR proteins	S
192/4			SCL30	AT3G55460	SR proteins	S
301/3			RSzp21/SRZ21	AT1G23860	SR proteins	S
160/3			RRM protein	AT4G35785	SR proteins	S
133/3			SR33/SCL33	AT1G55310	SR proteins	S
243/3			DHX15	AT3G62310	RES	S
106/3			NTR1	AT1G17070	RES	S
192/4			DHX8/PRP22	AT3G26560	RES	S
296/3			PRP16	AT5G13010	RES	S
413/4			RZ-1A	AT3G26420	hnRNP	S
213/4			RZ-1B	AT1G60650	hnRNP	S
204/4			RZ-1C	AT5G04280	hnRNP	S
230/3			hnRNP-R3	AT2G44710	hnRNP	S
206/3			NRP(B/D/E)3a	AT2G15430	RNAPII, IV, V	T
259/3			NRPB2	AT4G21710	RNAPII	T
122/3			CTR9, ELF8, VIP6	AT2G06210	PAF-C	T
755/4			ELP1, ELO2	AT5G13680	Elongator	T
444/4			ELP3; ELO3	AT5G50320	Elongator	T

[1]The numbers indicate in which affinity purifications the interactors were identified and the respective average Mascot scores are given as well the number of times the interactor was detected in four independent affinity purifications – only proteins are listed that were detected at least three-times out of four experiments.
[2]It is indicated to which protein complex or to which protein family the interactors belong.
[3]It is listed in which process (E, mRNA export; S, splicing; T, transcription) the interactors are primarily involved according to literature.

Table 2. Proteins that copurify with THP1.

THP1[1]	Interactor	AGI	Complex[2]
3571/4	SAC3A	AT2G39340	TREX-2
3517/4	SAC3B	AT3G06290	TREX-2
2079/4	THP1	AT2G19560	TREX-2
1033/4	SAC3C	AT3G54380	TREX-2
389/3	NUP133	AT2G05120	NPC
292/3	NUP155	AT1G14850	NPC
88/3	HPR1,THO1	AT5G09860	TREX
115/3	THO5A	AT5G42920	TREX
368/3	UAP56	AT5G11170	TREX
778/4	ALY4	AT5G37720	TREX
436/4	ALY2	AT5G02530	TREX

[1]The numbers indicate the respective average Mascot scores of the identified proteins as well the number of times the interactor was detected in four independent affinity purifications – only proteins are listed that were detected at least three-times out of four experiments.
[2]It is indicated to which protein complex the interactors belong.

Acknowledgements

I would like to thank all the people that during my PhD work have supported and helped me. In particular, I would like to thank my PhD supervisor Prof. Dr. Grasser for the opportunity to do my PhD in his group. I would also like to thank him for the guidance and advice given in regards to this project, as well as correcting this work.

In addition, a big thank you to PD Dr. Andreas Houben and PD Dr. Joachim Grisenbeck, who as my mentor team have been very helpful with guidance and advice in regards to the experimental part of my work, especially the establishment of affinity purification method.

I would like to thank Prof. Dr. Thomas Dresselhaus for the friendly environment in the department, and for having taken the time to be the chairperson of my examination committee.

Likewise I would like to thank Prof. Dr. Herbert Tschochner and Prof. Dr. Gernot Längst for taking the time to be the 2nd PhD Assessor and 3rd examiner at my PhD defence.

I would also like to thank Dr. Michael Meltzer, The group of Prof. Dr. Rainer Deutzmann, Judith Hauptmann and Daniele Hasler for providing their services and help with different experiments.

A big thank you to current and former colleagues in the workgroup for productive discussions, helpfulness, and the non-work related activities and generally just for having contributed to a good atmosphere. I would escpecially like to thank Simon Mortensen, Martin Antosch, Lena Dietz, Julius Dürr, Alex Pfab, Wojciech Antosz, Hans-Friedrich Ehrnsberger, Silvia Esposito, Antje Böttinger, Marion Grasser and Michael Scmitzberger.

Lastly, a heartfelt thank you goes out to my lovely wife Shuang and my son Svend, who together with parents and sister have been supporting me and providing inspiration to keep working hard.