Molecular and functional studies on the transcript elongation factor FACT and the SAGA-DUBm subunit ENY2 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

ALEXANDER PFAB

aus

MÜNCHEN

im Jahr 2017

Der Promotionsgesuch wurde eingereicht am:
17.11.2017

Die Arbeit wurde angeleitet von:
PROF. DR. KLAUS D GRASSER

Unterschrift:

ALEXANDER PFAB

Für Silli, Jari und Loki

Contents

Abbreviations

(v/v)	Volume per volume
(w/v)	Weight per volume
aa	Amino acid
AB	Antibody
ADA2b	Transcriptional adapter 2 b
Ade	Adenine
ANS	Anthocyanidin synthase
AP-MS	Affinity purification coupled to mass spectrometry
At	*Arabidopsis thaliana*
bp	base pair
CBB	Coomassie brilliant blue
cDNA	complementary DNA
CDS	Coding sequence
CHS	Chalcone synthase
CLSM	Confacal laser scanning microscopy
Col-0	Columbia-0
CTD	C-terminal domain
Cy3, Cy5	Cyanine dye 3, 5
DAPI	4',6-Diamidin-2-phenylindol
DAS	Days after stratification
DDO	Double drop out medium
DFR	Dihydroflavonol reductase
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
Ds	*Drosophila melanogaster*
DTT	Di-thiotreitol
DUBm	Deubiquitinating module
E. coli	*Escherichia coli*
e. g.	For example, latin: exempli gratia
EDTA	Ethylene diamine tetraacetic acid
eGFP	enhanced green fluorescent protein
EMSA	Electrophoretic mobility shift assay
ENY2	Enhancer of yellow 2
F3'H	Flavonone 3'hydroxylase
FACT	Facilitates chromatin transcription
FLC	Flowering locus C
FPLC	Fast performance liquid chromatography
FRAP	Fluorescence recovery after photobleaching
FRET	Förster resonance energy transfer
gDNA	Genomic DNA
GO	Gene ontology
GST	Glutathione S-transferase
GUS	β-glucuronidase
HAT	Histone acetyltransferase
His	Histidine
HL	High light
HMG	High-mobility group
Hs	*Homo sapiens*
HUB1	Histone monoubiquitination 1

HYG	Hygromycin B
IgG	Immunglobulin G
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
kDa	kilo Dalton
LB	Left border OR Luria Bertani
LD	Long day
Ler	*Landsberg erecta*
Leu	Leucine
LUT	Lookup-table
MOS4	Modifier of snc1, 4
mRNA	Messenger RNA
mRNP	Messenger ribonucleoprotein
MS	Murashige-Skoog OR mass spectroscopy
MSA	Multiple sequence alignment
N. benthamiana	*Nicotiana benthamiana*
NASC	Nottingham *Arabidopsis* stock center
NLS	Nuclear localization signal
NPC	Nuclear pore complex
nptII	Neomycin phosphotransferase II (Kanamycin resistance gene)
NTC	Nineteen complex
NTR	Nineteen complex-related proteins
NUPs	Nucleoporin proteins
o/n	Over night
OD600	Optical density measured at 600 nm
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PIC	Pre-initiation complex
PMSF	Phenylmethylsulphonyl fluoride
P-TEFb	Positive transcription elongation factor b
QDO	Quadruple drop out medium
qRT-PCR	Quantitative reverse transcription-polymerase chain reaction
RB	Right border
RNAPII	RNA polymerase II
RNA	Ribonucleic acid
RNAi	RNA interference
ROI	Region of interest
Rpm	Rounds per minute
S. cerevisiae	*Saccheromyces cerevisiae*
SAGA	Spt-Ada-Gcn5 acetyltransferase
SAIL	Syngenta arabidopsis insertion library
SBP	Streptavidin binding peptide
Sc	*Saccharomyces cerevisiae*
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser2	Serine 2 phosporylated CTD
SG	Tag consisting of SBP and 2 x Protein G domains
SGF11	SAGA-associated factor 11, DUBm subunit
SPT16	Suppressor of Ty 16, FACT subunit
SSRP1	Structure specific recognition protein 1, FACT subunit
TAF10	TATA-box binding protein associated factor 10, SAGA component
TAF13	TATA-box binding protein associated factor 13, SAGA component
T-DNA	Transfer DNA
TDO	Triple drop out medium
TEC	Transcript elongation complex
TEF	Transcription elongation factor
Ter	Terminator
TEV	Tobacco Etch Virus (TEV) protease
TF	Transcription factor
THP1	THO2/HPR1 phenotype 1, TREX-2 subunit
TREX-2	Transcription and export complex 2
Trp	Tryptophan
UBP22	Ubiquitin carboxyl-terminal hydrolase 22, DUBm subunit
UTR	Untranslated region
WT	*Wild type*
Y2H	Yeast-two-Hybrid

1 Introduction

1.1 The chromatin structure

In the nuclei of eukaryotic cells, the genomic DNA is compacted into chromatin. Nucleosomes, the building blocks of the chromatin, consist of a histone octamer (two copies of each histones H2A, H2B, H3, H4) with 147 bps of genomic DNA wrapped around. The nucleosomes are connected by the linker DNA into nucleosomal arrays like "beads on a string" that are organized into higher order structures (Luger et al. 1997, Andrews et al. 2011). The chromatin structure does not only package the entire genome into the nucleus, but as repressive barrier it has also profound regulatory implications on all DNA-depending processes including transcription, replication or repair (Li et al. 2007a, Luger et al. 2012) To make the DNA accessible and to facilitate transcription in the chromatin context, the nucleosomes must be moved, partially disassembled or modified. During transcription, this histone exchange is managed by numerous factors including energy-dependent chromatin remodellers, histone chaperones and enzymes that add post-translational modifications (PTMs) to the N-terminal tails of histones (Venkatesh et al. 2015). Histone modifications including acetylation, methylation, phosphorylation, and ubiquitylation do not only change the physical and chemical properties of chromatin, but are also important for recruiting other histone modifiers (Saunders et al. 2006).

1.2 Transcription by RNA polymerase II

DNA templates are transcribed into mRNA molecules by the RNA polymerase II (RNAPII) with the help of several auxiliary factors. The process of transcription is divided into three regulated phases known as initiation, elongation and termination (**Figure 1.1**). The repetition of these three phases (transcription cycle) over a gene determines its expression levels (Jonkers et al. 2015, Venkatesh et al. 2015). The first step of the transcription cycle starts with the assembly of the pre-initiation complex (PIC) at the promoter of a gene. The PIC consists of the RNAPII and the general transcription factors (GTFs) TFIID, TFIIB, TFIIE, TFIIF and TFIIH. The GTFs recruit and position the RNAPII near the transcription start site and dictate the precise location and direction of the transcription initiation (Sims et al. 2004, Saunders et al. 2006). For regulated transcription, general co-factors like the Mediator or the Spt-Ada-Gcn5 acetyltransferase (SAGA) complex are often required to transmit regulatory signals between gene-specific activators and the general transcription machinery (Thomas et al. 2006). To initiate RNA synthesis, the double stranded DNA is melted and the single stranded template is positioned in the active site of the RNAPII to form the open complex. In the following, the transition from transcript initiation to elongation is accompanied by structural and functional changes of the RNAPII including the phosphorylation of the C-terminal domain (CTD) of its largest subunit NRPB1. The RNAPII loses contact to the GTFs and moves along the template strand to synthesize the RNA transcript by joining nucleotides that are complement with the DNA template (Sims et al. 2004, Saunders et al. 2006). In some eukaryotes, the RNAPII can pause and accumulate at promoter-proximal regions during the initial steps of transcript elongation before productive elongation starts (Jonkers et al. 2015). During the highly regulated step of transcript elongation, several factors are recruited by the phosphorylated CTD that are important for efficient mRNA synthesis and co-transcriptional mRNA processing. Finally,

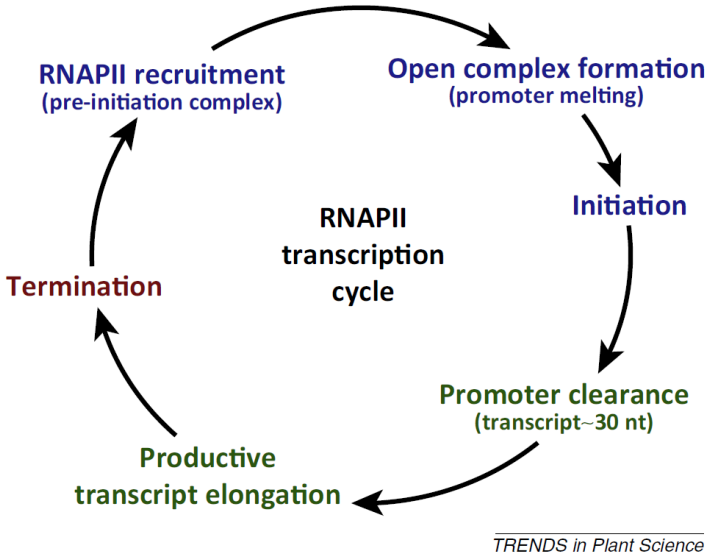


Figure 1.1 RNAPII transcription cycle
The transcription by RNAPII is characterized by a cycle of events that starts with the recruitment of the polymerase and the formation of the closed pre-initiation complex (RNAPII with GTFs) at the promoter of a gene. By melting the DNA of the core promoter, an open RNAPII complex is formed to initiate transcription. During the transition from transcript initiation to elongation, the RNAPII undergoes structural and functional changes that are collectively comprised as promoter clearance. Thereby, the RNAPII loses contact to the GTFs and establishes a stable association with the nascent transcript. Approximately 30 nucleotides downstream of the transcription start-site, promoter clearance is complete and the RNAPII becomes engaged in productive elongation. Thereby, a variety of TEFs ensures the efficient mRNA synthesis on chromatin templates. To regulate transcription, the RNAPII can pause and accumulate in some eukaryotes during the initial steps of transcript elongation (usually 20-60 nucleotides downstream of the transcription start-site). The progression of the transcription machinery is accompanied by changes in the phosphorylation pattern of the C-terminal heptapeptide repeats of the largest RNAPII subunit. For instance, the phosphorylation of serine 5 and serine 2 are characteristic for early and later transcript elongation, respectively. Thereby, the phosphorylation of the RNAPII CTD plays an important role in the coordination of the ongoing mRNA synthesis with co-transcriptional mRNA processing events. During termination, the produced transcript is released from the RNAPII. Subsequently, the polymerase dissociates from the DNA and is recycled for the next round of the transcription cycle. The early stages of the transcription cycle (depicted in blue) are well investigated, whereas much less is known about the later stages (dipected in green) that are as well dynamic and highly regulated. The figure is taken from Van Lijsebettens et al. 2014.

transcription is terminated and the transcribed RNA is released from the RNAPII for further processing and export to the cytoplasm. The transcription machinery dissociates from the DNA and is recycled for the next round of mRNA synthesis (Saunders et al. 2006, Selth et al. 2010).

1.2.1 Transcript elongation

Traditionally, the phase of transcript initiation with the recruitment of the RNAPII to the promoter was assumed as the most crucial step in the regulation of mRNA biogenesis. In contrast, during the subsequent phase of transcript elongation, the RNAPII was considered to simply behave like a machine that is quickly "reading the gene". In the recent years, it became more and more apparent that the transcript elongation is as well a dynamic and highly regulated process. Moreover, centrally located in the gene expression pathway, it coordinates the synthesis of mRNA with its maturation and export (Sims et al. 2004, Saunders et al. 2006, Selth et al. 2010, Jonkers et al. 2015).

In the early stage of transcription, the eukaryotic RNAPII exchanges the associated transcript initiation factors like GTFs with transcript elongation factors (TEFs) to form a transcript

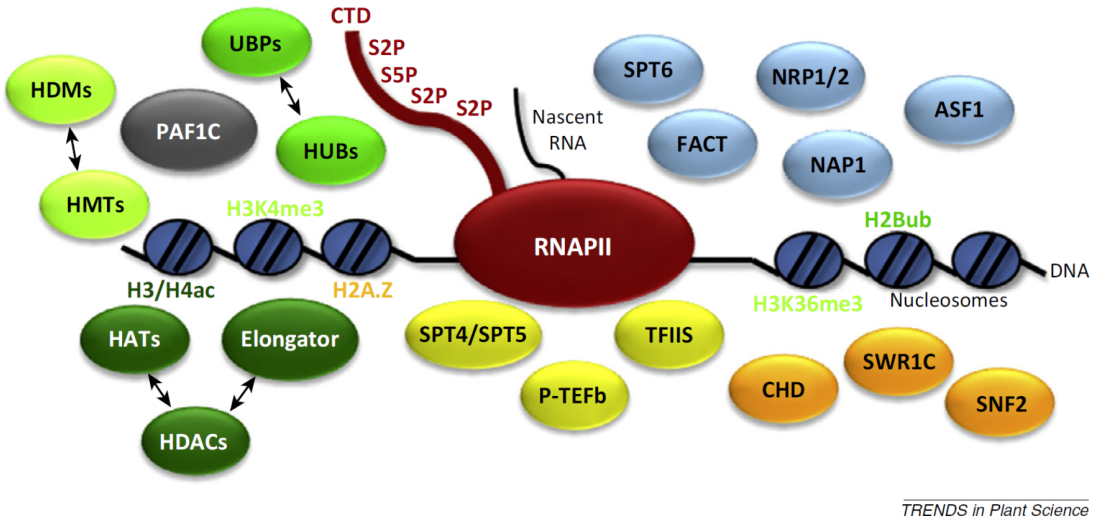


Figure 1.2 A variety of transcript elongation factors (TEFs) enables the efficient transcription of chromatin templates by RNAPII.
TEFs are a heterogenous group of proteins that serve diverse functions during transcription by RNAPII. Factors that directly modulate RNAPII properties, for instance, to allow the polymerase to counteract transcriptional pausing/arrest are depicted in yellow. Histone chaperone (depicted in blue) and ATP-dependent chromatin remodelers (depicted in orange) play important roles to facilitate transcription in the chromatin context. Additional factors like the enzymes depicted in green, can control transcription by modifying histones covalently within transcribed regions. The modifications (methylation, acetylation, and ubiquitination) can be removed by the reverse activity of other enzymes (indicated by double-headed arrows). Some activating marks like H3K4me3, H3K36me3, H3/H4ac, H2Bub are shown. The figure is taken from Van Lijsebettens et al. 2014.

elongation complex (TEC) for processive mRNA synthesis (Ehara et al. 2017). A variety of different TEFs has been identified in the last years (**Figure 1.2**). These TEFs can be associated with the TEC permanently or just transiently and can therefore affect the global expression or just the expression of a subset of genes. TEFs are a very heterogenous group of proteins that serve diverse functions to facilitate the efficient progression of the transcription machinery through the repressive chromatin. Thereby TEFs can directly modulate the catalytic properties of the processive RNAPII, can modify nucleosomal histones within the transcribed region or can function as histone chaperones and chromatin remodelers (Sims et al. 2004, Selth et al. 2010, Van Lijsebettens et al. 2014). Moreover, the TEC functions as platform for co-transcriptional nuclear transactions including pre-mRNA processing (Perales et al. 2009).

Recently, the structure of the yeast RNAPII transcript elongation complex was elucidated by X-ray crystallography and cryo-EM (Xu et al. 2017a, Ehara et al. 2017). Beside the assembly of the TEC, several studies in yeast revealed genetic/biochemical interactions between various TEFs and tracked the TEC along transcribed regions (Selth et al. 2010, Jonkers et al. 2015). In *Arabidopsis* and in higher eukaryotes, considerably less is known about the composition of the RNAPII TEC and the interplay of the TEFs among each other and the transcribing RNAPII.

Genetic and biochemical studies in *Arabidopsis* revealed that TEFs can regulate plant growth and development. For instance, the histone chaperone FACILITATES CHROMATIN TRANSCRIPTION (FACT) assists the progression of the transcribing RNA polymerase on chromatin templates by destabilizing nucleosomes (Van Lijsebettens et al. 2014). The *Arabidopsis* FACT complex was shown to be necessary for the expression of the floral repressor *FLOWERING LOCUS C* (*FLC*) and thereby for the transition from the vegetative to the reproductive phase

(flowering) (Lolas et al. 2010). The POLYMERASE-ASSOCIATED FACTOR 1 (PAF1) COMPLEX (PAF1-C) has been shown to regulate transcription by promoting specific histone modifications including the methylation of lysine residues and the monoubiquitylation of H2B K123 (Tomson et al. 2013). The *Arabidopsis* PAF1-C controls the flowering time by modifying the chromatin of floral repressors like *FLC* (He et al. 2004, Oh et al. 2004). The *Arabidopsis* POSITIVE TRANSCRIPTION ELONGATION FACTOR b (P-TEFb) complex can as well regulate the expression level of *FLC* and affect flowering time. Therefore, P-TEFb that consists of CDKC;2 and CYCT1 influences the phosphorylation status of the Ser2 residue of the RNAPII C-Terminal repeats (CTD) globally (Wang et al. 2014). The TRANSCRIPTION FACTOR S-II (TFIIS) promotes efficient transcription by RNAPII, since it assists in bypassing blocks during mRNA synthesis (Fish et al. 2002). Thereby TFIIS controls the *Arabidopsis* seed dormancy by regulating the expression of *DELAY OF GERMINATION1* (*DOG1*) (Grasser et al. 2009, Mortensen et al. 2014). The SUPPRESSOR OF TY 4/5 (SPT4/5) complex directly interacts with the RNAPII to couple chromatin modification states and RNA processing for processive transcript elongation (Hartzog et al. 2013). In *Arabidopsis*, SPT4/5 regulates transcript elongation of RNAPII with particular impact on the expression of certain auxin-related genes (Dürr et al. 2014).

1.3 The histone chaperone FACT

1.3.1 The FACT architecture

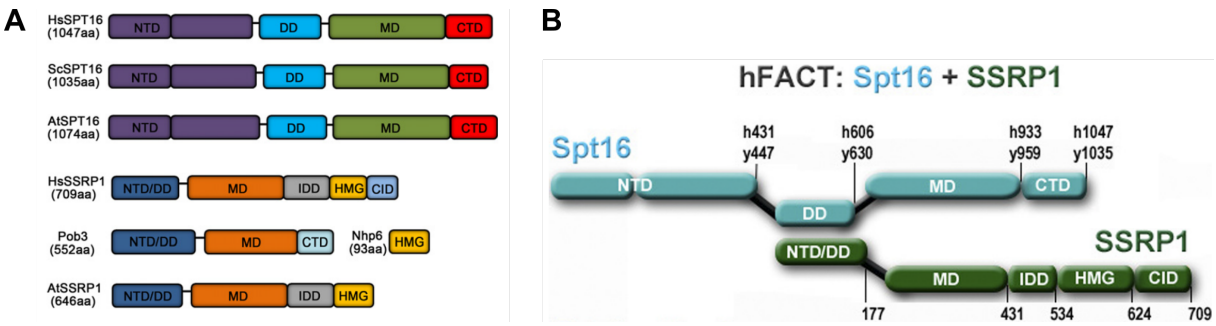


Figure 1.3 Conserved domain organization and structural alignment of the FACT subunits.
A) Schematic illustration of the conserved domain organization of human, yeast and *Arabidopsis* FACT subunits. SPT16 is highly conserved from yeast to humans, while the C-terminus of SSRP1/Pob3 shows species-specific characteristics. CID = C-terminal intrinsically disordered domain, CTD = C-terminal domain, DD = dimerization domain, HMG = high mobility group domain, IDD = intrinsically disordered domain, MD = middle domain, NTD = N-terminal domain. The figure is taken from Zhou et al. 2015. B) Dimerization of both FACT subunits is accomplished through specific interactions between the centrally and the N-terminally located DDs of Spt16 and SSRP1, respectively. The figure is taken from Winkler et al. 2011a.

The histone chaperone FACT is evolutionarily conserved in eukaryotes including the plant kingdom (Formosa 2008, Lolas et al. 2010). The heterodimer FACT consists of the SUPPRESSOR OF TY 16 (SPT16) and the STRUCTURE-SPECIFIC RECOGNITION PROTEIN 1 (SSRP1) in metazoans and plants or Spt16 and Pob3 in yeast and fungi (Orphanides et al, 1998). SPT16/Spt16 is highly conserved from yeast to higher eukaryotes, whereas SSRP1/Pob3 shows a more varied degree of sequence conservation and domain organization (Zhou et al. 2015) as shown in **Figure 1.3**. Most strikingly, the metazoan/plant SSRP1 features a high mobility

group (HMG)-box domain at the C-terminus which is missing in the yeast Pob3. In return, the yeast Spt16/Pob3 complex is loosely associated with the small HMGB protein Nhp6, which provides the HMG-box function and mediates the interaction with DNA and nucleosomes (Brewster et al. 2001, Formosa et al. 2001). Nhp6 seems to have a general role in promoting chromatin dynamics in yeast as it is important for the activities of several other chromatin factors (McCullough et al. 2011). The genes coding for SPT16 and SSRP1/Pob3 are essential for viability in yeast and metazoans (Winkler et al. 2011a, Formosa 2012). In contrast, the double knockout of the two Nhp6 genes is viable, but results in slow growth of the yeast cells (Costigan et al. 1994, Stillman 2010).

1.3.2 The FACT mechanism

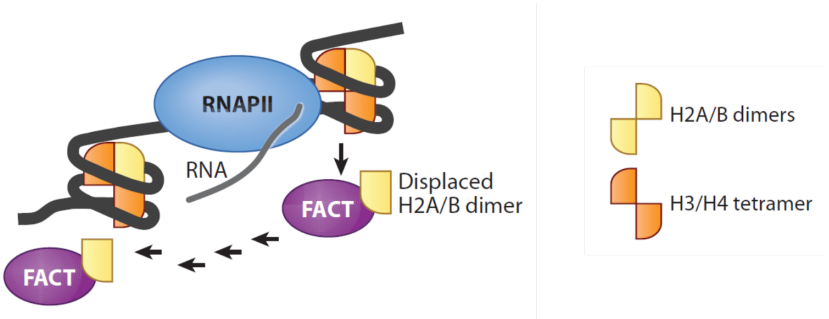


Figure 1.4 Nucleosome reorganization by FACT during transcript elongation.
During transcript elongation, the FACT complex enables the progression of the elongating RNAPII through the chromatin template by destabilizing nucleosomes in the path of the transcription machinery. Afterwards in the wake of the progressing polymerase, the normal chromatin structure is maintained by the reverse action of the histone chaperone. The nucleosome reorganization may occur without complete displacement of the histone proteins from the DNA. The figure is taken from Selth et al. 2010.

The FACT complex was first identified as a human factor that facilitated productive transcript elongation through nucleosomes *in vitro* (Orphanides et al. 1999). In eukaryotes, FACT plays an important role in many genomic DNA-dependent processes like transcription, replication and repair by both establishing and overcoming the repressive chromatin barrier. Thereby, FACT can temporarily promote the access to DNA by reorganizing nucleosomes. Afterwards FACT can also restrict the access to DNA again by restoring the original chromatin state. During transcript elongation (**Figure 1.4**), FACT binds to H2A-H2B dimers and facilitates the passage of the TEC through the chromatin template by destabilizing the octameric nucleosomes in the path of the transcribing RNAPII (Winkler et al. 2011a, Formosa 2012). After passage of the transcription machinery, FACT most likely restores the normal nucleosome structure by the reverse action (Jamai et al. 2009).

Currently, two main models exist about the mechanism by which FACT performs its functions. According to the "dimer eviction" model, FACT re-organizes nucleosomes and promotes DNA accessibility by displacing a single H2A-H2B dimer from a nucleosome (Belotserkovskaya et al. 2003, Orphanides et al. 1999, Reinberg et al. 2006). The "global access" model suggests that FACT loosens internal contacts of the nucleosome to obtain a more open configuration that allows access to the DNA. In this model, the loss of the H2A-H2B dimer is not essential, but just an optional outcome (Xin et al. 2009, Formosa 2008). To what extend the "global access" model

and/or "dimer eviction" models apply to FACT function in plants requires further investigations (Zhou et al. 2015). Kemble et al. 2015 revealed that conserved residues in the acidic regions near the C-terminus of both yeast FACT subunits are necessary to bind H2A-H2B dimers. Thereby, the Spt16-Pob3 heterodimer can bind two H2A-H2B dimers simultaneously, because the binding sites of both FACT subunits are overlapping on the histone dimer. All components of the reorganized nucleosome are tethered together by FACT (**Figure 1.5**). One interesting option is that the interaction of the FACT complex with the H2A-H2B dimer might dependent on the prior association of the C-terminal HMG-box domain with the nucleosomal DNA.

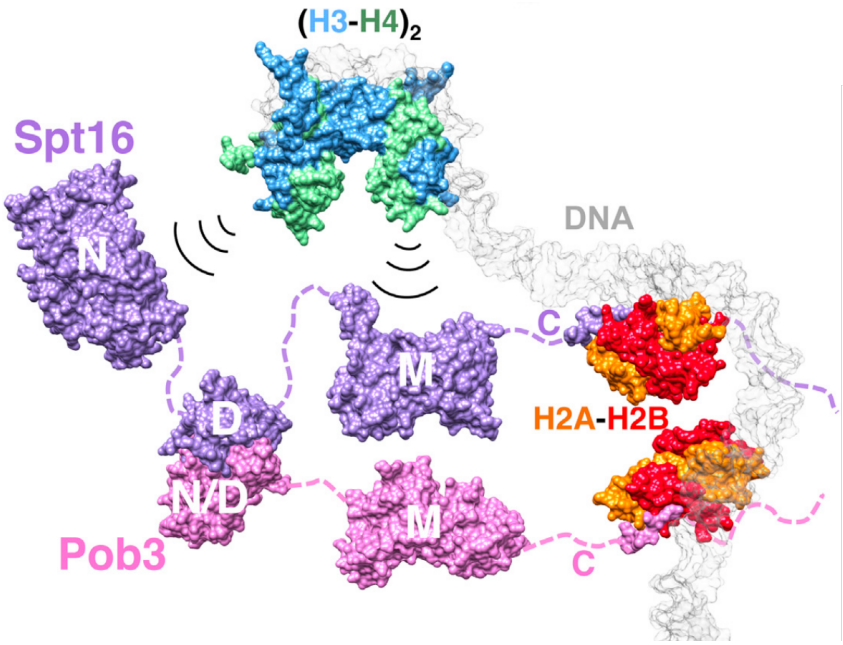


Figure 1.5 C-Terminal peptides of Spt16 and Pob3 are required for H2A-H2B binding and nucleosome reorganization.
The model illustrates the binding of FACT to the components of a nucleosome. Two H2A-H2B dimers are bound by C-terminal peptides of Spt16 and Pob3. The H3-H4 histones are bound by the N-terminal and the middle domains of Spt16. The figure is taken from Kemble et al. 2015.

1.3.3 The FACT complex in *Arabidopsis*

An *Arabidopsis* FACT complex consisting of SPT16 and SSRP1 has been identified by Duroux et al. 2004. The complete loss of SSRP1 is critical for plant viability, whereas reduced amounts of SSRP1 and SPT16 display various defects in the vegetative and reproductive development including early flowering, a "bushy" phenotype and a reduced seed set. In accordance to the early transition from the vegetative to the reproductive phase, SSRP1- and SPT16-depleted plants show a reduced expression of the key floral repressor FLC (Lolas et al. 2010). But not much is known about the impact of FACT-depletion on the genome-wide gene expression in *Arabidopsis*.

In line with its role as TEF, the *Arabidopsis* FACT complex localizes to the de-condensed euchromatin and associates there with actively transcribed genes like the above mentioned FLC (Duroux et al. 2004). Moreover, FACT was found to interact genetically with HUB1 that catalyses the mono-ubiquitination of the histone H2B, to regulate various developmental processes

(Lolas et al. 2010, Van Lijsebettens et al. 2014). Additionally, SSRP1 was shown to be required for DNA methylation and genomic imprinting (Ikeda et al. 2011). But interactions of the *Arabidopsis* FACT complex with other TEFs of the RNAPII TEC complex were not investigated yet and would provide new insights into the transcript elongation process of higher eukaryotes.

1.4 Co-transcriptional processing of mRNA

During transcription by RNAPII, the growing pre-mRNA is maturated by processing factors and packaged into an export competent messenger ribonucleoprotein particle (mRNP) that is exported to the cytoplasm for translation (Moore et al. 2009, Saunders et al. 2006, Bentley 2014, Saldi et al. 2016). The processing of pre-mRNA occurs mostly co-transcriptional at the gene. The nascent transcript that is protruding out of the transcribing RNAPII is immediately a substrate for the mRNA processing factors. The maturation of most pre-mRNA molecules includes 5′ capping, splicing, 3′ cleavage and polyadenylation as well as RNA editing. The first processing event (5′ capping) is the attachment of a 7-methylguanosine cap to the 5′ end of the nascent transcript to convert the 5′ pppA to a 5′ 7meGpppA end. Next, the spliceosome is assembled on the growing pre-mRNA, to excise the introns and to ligate the exons in a two-step transesterification reaction. Certain exons can be included or excluded from the mature mRNA. This so-called alternative splicing can result in the production of multiple protein isoforms from a single gene and increases the protein diversity (Jonkers et al. 2015).

Moreover, the transcript elongation rate is highly dynamic and varies between genes (intergenic) as well as between regions within a gene (intragenic), what can affect the outcome of co-transcriptional processes like alternative splicing (Bentley 2014). Finally, when the polymerase transcribes the poly(A) signals, the emerging pre-mRNA sequences are recognized by the polyA complex. This protein machinery catalyses the cleavage of the growing transcript and polyadenylates the 3′end by addition of an adenosine monophosphate (Moore et al. 2009, Saunders et al. 2006).

1.5 Coupling of the RNAPII and mRNA processing factors

The transcription and mRNA processing as well as the transport of export competent mRNPs through the nuclear pore complex (NPC) to the cytoplasm depends on the coordinated action of several multiprotein complexes. The functional and physical coupling of the numerous factors of the gene expression machinery and their correct interplay ensures the proper production of proteins in the cell. Several transcription factors were found to influence processing factors and vice versa (Maniatis et al. 2002, Köhler et al. 2007, Komili et al. 2008).

The processing factors are recruited directly by the nascent RNA, the proteins of the TEC or specially by the CTD of the RNAPII that acts as a flexible and versatile "landing pad" for nuclear factors. The phosphorylation patterns on the CTD repeats change during the transcription cycle and determine which factors can bind or not (Phatnani et al. 2006, Bentley 2014).

Beside the linkage of the processing factors to the transcript-elongation machinery, these factors are also highly connected to each other. This coupling occurs not only between factors of sequential mRNA processing steps but also between the earliest and latest, which revises the image of the gene expression pathway as a simple linear assembly line (Maniatis et al. 2002).

1.6 The coupling factor Sus1/ENY2

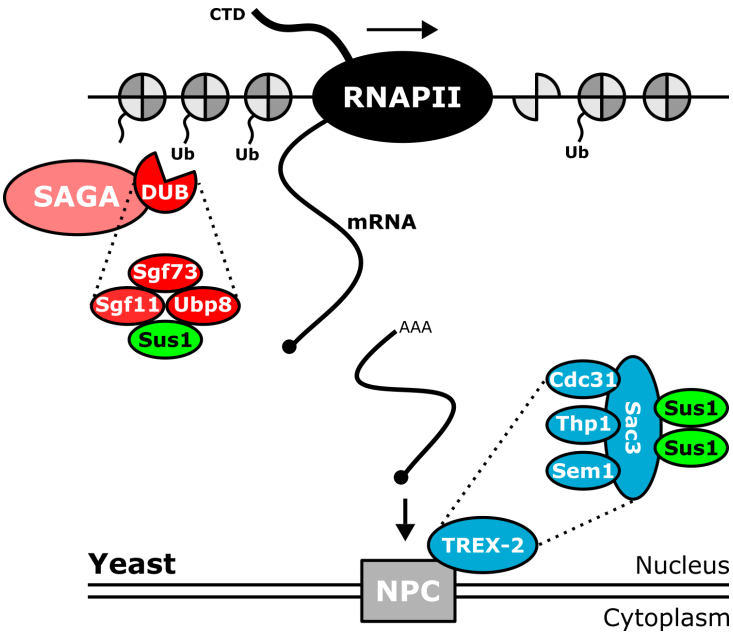


Figure 1.6 The small adaptor protein Sus1/ENY2 links transcript initiation and mRNA export.
In yeast, fruit-fly and humans, Sus1/ENY2 is part of the transcriptional co-activator SAGA and the NPC-associated TREX-2 complex. To gain more clarity, only the protein names of *S.cerevisae* are indicated in the model. As part of the SAGA-DUB module, Sus1 is important for transcript initiation by de-ubiquitination of the histone H2B. The TREX-2 complex that facilitates the transport of export competent mRNP from the nucleus to the cytoplasm contains two copies of Sus1 that bind to the scaffold protein Sac3. Moreover the yeast TREX-2 complex contains additionally Thp1, Cdc31 and Sem1. Thereby, the conserved Sus1/ENY2 protein bridges the beginning (transcript initiation) and the end (mRNA export) of the gene expression pathway in eukaryotic nuclei (Pascual-García et al. 2009, Kopytova et al. 2010a).

Sus1/ENY2 (yeast/higher eukaryotes) is a small evolutionary conserved protein with approximately 11-kDa in yeast, fruit-flies and humans (Georgieva et al. 2001, Rodríguez-Navarro et al. 2004, Zhao et al. 2008). In agreement with its high sequence conservation, the function of Sus1/ENY2 as an important coupling factor in the gene expression pathway seems to be conserved as well. In yeast and metazoan, Sus1/ENY2 is a shared component of the two multi-protein complexes SAGA and TREX-2 (transcription and export complex 2)(**Figure 1.6**). Sus1/ENY2 provides a physical link between the promoter-bound transcriptional co-activator SAGA and the nuclear pore associated mRNA export factor TREX-2. Thereby, Sus1/ENY2 is bridging the ends of the nuclear gene expression pathway, the transcription initiation by RNAPII with the export of the mature mRNA throught the nuclear pore complex (NPC) (Rodríguez-Navarro et al. 2004, Kurshakova et al. 2007, Zhao et al. 2008, Pascual-García et al. 2009, Kopytova et al. 2010a). Additionally, Sus1/ENY2 plays a role during transcript elongation that differs between yeast and *Drosophila*. The *Drosophila* ENY2 is important for the co-transcriptional recruitment of the THO complex on the nascent mRNA to form export-competent mRNPs, but does not directly associate with the elongating RNAPII (Kopytova et al. 2010b). In addition, ENY2-THO functions independent of SAGA or AMEX (TREX-2). The yeast Sus1 was directly found on coding regions (chromatin) and was associated with the elongating RNAPII in a SAGA and TREX-2 dependent manner (Pascual-García et al. 2008). Under specific circumstances Sus1/ENY2 (yeast and *Drosophila*) was observed to a lesser extend in the cytoplasm,

most likely involved in the cytoplasmic mRNA metabolism (Cuenca-Bono et al. 2010, Kopytova et al. 2010a).

1.7 The transcriptional co-activator SAGA

As mentioned earlier (Section 1.2), the transcription is initiated by PIC assembly (RNAPII and GTFs) at the transcription start site (Sims et al. 2004, Saunders et al. 2006). Transcription factors that are binding gene-specific DNA elements upstream of the core promoter can recruit co-activators like SAGA to the PIC to promote the basal transcription (Thomas et al. 2006). In the last years, the transcriptional co-activator SAGA was the prime example to study gene activation in eukaryotes (Koutelou et al. 2010). The well characterized SAGA complex is an approximately 2 MDa multi-protein complex that is structurally and functionally conserved from yeast to humans (Baker et al. 2007, Rodríguez-Navarro 2009, Koutelou et al. 2010, Spedale et al. 2012). In yeast, the SAGA complex regulates the transcription of approximately 10 % of the genes genome-wide (Baker et al. 2007, Samara et al. 2011, Weake et al. 2012). The proteins of the SAGA complex are arranged into four functional modules to regulate transcription (**Figure 1.7**). The HAT (histone H3 acetyltransferase) and the DUB (histone H2B de-ubiquitylation) modules possess enzymatic activities to modify chromatin. Whereas, the TAF (TBP-associated factor) and SPT (suppressor of Ty) modules mediate the interactions of SAGA with the activators and the PIC (Rodríguez-Navarro 2009, Koutelou et al. 2010, Samara et al. 2011). In yeast (y), *Drosophila* (d) and humans (h), the HAT module consists of the following four subunits: The acetyltransferase yGcn5/dGCN5/hGCN5 or hPCAF, Ada2/dADA2b/hTADA2b, yAda3/dTADA3/hTADA3 and Sgf29/dSGF29/hSGF29. The HAT module can acetylate the histone H3 at lysine K9 and K14 (Koutelou et al. 2010, Spedale et al. 2012). The DUB module comprises the four following components: The deubiquitinating enzyme yUbp8/dNonstop/hUSP22, the two adaptor proteins ySgf11/dSgf11/hATXN7L3 and ySus1/dENY2/hENY2 as well as the anchoring protein ySgf73/dSgf73/hATXN7 (Köhler et al. 2008, Weake et al. 2008a, Zhao et al. 2008, Lee et al. 2009, Samara et al. 2010, Gurskiy et al. 2012). The crystal structure of the complete DUB module of the yeast SAGA complex revealed that Ubp8, Sus1, Sgf11 and Sgf73 form a highly interconnected complex that is organized into the "catalytic lobe" and the "assembly lobe". Each protein is thereby physically connected the other three (Köhler et al. 2010, Samara et al. 2010). The crystal structure of the yeast DUB module bound to an ubiquitinated nucleosome revealed that the DUB module primarily contacts the acidic patch of H2A/H2B with the zinc finger domain (ZnF) of Sgf11 (Morgan et al. 2016). The DUB module catalyses the cleavage of monoubiquitin from K123 (yeast) or K120 (*Drosophila* and human) of histone H2B (Daniel et al. 2004, Henry et al. 2003, Weake et al. 2008a). All components of the DUB module are required for the activity of de-ubiquitinating enzyme Ubp8 (Köhler et al. 2008).

At the onset of transcription by RNAPII, the co-activator SAGA is recruited to the transcriptional start site. The acetylation of histones by the HAT module leads to de-compaction of the chromatin structure, which facilitates transcript initiation by the basal transcription machinery (Sterner et al. 1999, Nagy et al. 2007). The transition from transcript initiation to transcript elongation is accompanied by changes in the phosphorylation pattern of the RNAPII CTD. Im-

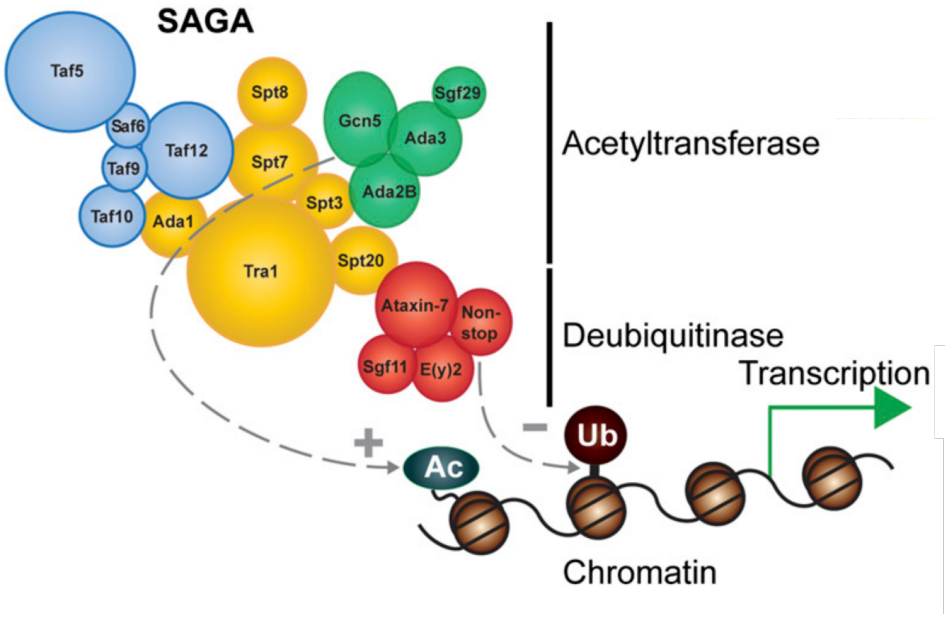


Figure 1.7 Composition of the chromatin modifying SAGA complex in *Drosophila*.
Schematic illustration of the modular structure of the multi-protein SAGA complex. The TAF (depicted in blue) and the SPT module (depicted in yellow) mediate the interactions of the SAGA complex with activators and the RNAPII pre-initiation complex. The HAT module (depicted in green) possesses a histone H3 acetyltransferase activity, while the DUB module (depicted in red) harbours a histone H2B de-ubiquitinase activity. The figure is taken from Mohan et al. 2014.

mediately after transcript initiation, the heptapeptides repeats of the CTD are phosphorylated at serine 5 and 7 to recruit factors for the early steps of transcript elongation. Subsequent these CTD repeats are phosphorylated at serine 2 to recruit factors for the following productive elongation phase (Buratowski 2009, Workman 2016). The monoubiquitin at the N-terminal tail of H2B blocks the recruitment of the kinase Ctk1, which phosphorylates the serine 2 of the RNAPII CTD. Therefore, the de-ubiquitylation of H2B by the SAGA DUB module promotes transcript elongation by removing the repressive histone mark (Wyce et al. 2007, Workman 2016).

Little is known about the SAGA complex in plants. The potential *Arabidopsis* counterparts of the yeast and human SAGA components were predicted bioinformatically (Moraga et al. 2015, Srivastava et al. 2015). Mutant studies on *Arabidopsis* GCN5 and ADA2b, the homologs of the yeast HAT components Gcn5 and Ada2, revealed that these proteins are critical for plant growth and development (Vlachonasios et al. 2003). The interaction of *Arabidopsis* ADA2b with GCN5 was shown *in vitro* by pulldown and yeast two-hybrid assays (Stockinger et al. 2001, Mao et al. 2006). ADA2b can stimulate the HAT activity of GCN5 on nucleosomal histones *in vitro* (Mao et al. 2006). In ADA2b or GCN5 depleted plants, the expression of 5 % of the 8200 investigated genes was affected (Vlachonasios et al. 2003). Another study showed that 40 % of the investigated promoters are associated with the *Arabidopsis* GCN5 protein (Benhamed et al. 2008). Furthermore, the *Arabidopsis* GCN5 was shown to be required for the homeostasis of the recently discovered histone modification H3K36ac, which is conserved in plants and highly enriched in euchromatin (Mahrez et al. 2016a).

1.8 Aims of the thesis

1.8.1 The histone chaperone FACT

FACT, a conserved heterodimer of SSRP1 and SPT16, facilitates the progression of the transcription machinery through the chromatin template by destabilizing nucleosomes in the path of the elongating RNAPII. The histone chaperone FACT reorganizes nucleosomes and promotes access to the genomic DNA by binding to H2A-H2B dimers. One attractive possibility is that the association of FACT with chromatin depends, in the first place, on the association of the SSRP1 HMG-box domain with the nucleosomal DNA. To gain more knowledge about the SSRP1 HMG-box domain in higher eukaryotes, the proposed DNA- and nucleosome-binding properties of the *Arabidopsis* SSRP1 HMG-box domain will be analysed by Electrophoretic Mobility Shift Assays (EMSA) *in vitro*. In the following, FRAP experiments will reveal if removing of the C-terminal HMG-box domain alters the kinetics of SSRP1 in living *Arabidopsis* cells, which is a good indication if the HMG-box domain is required for the binding of FACT to chromatin. Additionally, immunoblotting analysis will reveal if the plant SSRP1 lacking the HMG-box domain still associates with the transcriptionally active RNAPII *in vivo*. Moreover, the generation and phenotypic analysis of *Arabidopsis* HMG-box-deficiency mutants will reveal if the DNA-binding domain of SSRP1 is important for proper plant growth and development. Collectively, we will gain new insights about the role of the SSRP1 HMG-box domain in Arabidopsis and whether the DNA-binding domain is indispensable for the association of FACT with nucleosomes.

During transcript elongation, a heterogenous group of TEFs including the histone chaperone FACT enables efficient mRNA synthesis. A proteomic approach using reciprocal tagging in combination with affinity purification and mass spectrometry will reveal the composition of the *Arabidopsis* transcript elongation complex. Furthermore, the effect of TEF-depletion in *Arabidopsis* plants is ranging from mild phenotypes (e.g. TFIIS) to severe and lethal phenotypes (e.g. SPT5). The phenotypic analysis of *Arabidopsis* double mutants that are defective in different combinations of TEFs will reveal if there are genetic interactions between the genes encoding subunits of FACT, TFIIS and PAF1-C.

The reduced expression of either *SSRP1* or *SPT16* results in similar pleiotropic phenotypes affecting plant growth and development like early flowering and a reduced seed set. SSRP1- and SPT16-depleted plants will be analysed by genome-wide transcript profiling compared to *wild-type*, to identify differentially expressed genes.

1.8.2 ENY2, a coupling factor of transcription and mRNA processing

ENY2, an evolutionary conserved adaptor protein, bridges both ends of the nuclear gene expression pathway as part of the transcriptional co-activator SAGA and the mRNA export complex TREX-2 in yeast and metazoa. The small protein with its important function in controlling eukaryotic gene expression is well described in yeast, fruit-fly and humans, but nothing is known about ENY2 in plants. As well little is known about the chromatin-modifying SAGA complex in plants, although its counterparts in other organism play critical roles during transcription by RNAPII. To identify the interaction network of *Arabidopsis* ENY2 and its putative interactors, the SAGA complex and the TREX-2 complex, a proteomic approach using reciprocal tagging in combination with affinity purification and mass spectrometry will be applied. Additionally, to gain more knowledge about the function of ENY2 in plants, the following experiments will be conducted. The ENY2 promoter activity will be analysed in *Arabidopsis* by GUS staining, to characterize the spatio-temporal ENY2 gene expression throughout the entire life cycle of a plant. The subcellular localization and the dynamics of ENY2 in *Arabidopsis* root nuclei will be investigated in transgenic plants expressing GFP-tagged ENY2 by in-detail CLSM analysis. Moreover, a reverse genetics approach using RNAi and CRISPR/Cas9 and the phenotypic characterization of the generated knockdown/knockout mutants will give insights about the function of ENY2 during plant growth and development.

2 Results: Analysis of the HMG-box domain of *Arabidopsis* SSRP1

2.1 *Arabidopsis* SSRP1 features a conserved C-terminal HMG-box domain

Among eukaryotes, the high mobility group (HMG)-box is a common features of various chromosomal proteins that are important for the regulation of DNA-dependent processes. To bind DNA, the 75 aa HMG-box domain forms a L-shaped molecule consisting of three alpha-helices (Thomas et al. 2001, Štros et al. 2007, Malarkey et al. 2012). In humans and *Arabidopsis*, the essential histone chaperone FACT features a conserved HMG-box domain at the C-terminus of SSRP1 (**Figure 2.1**) (Duroux et al. 2004). This HMG-box domain may help FACT to recognize, bind and reorganize chromatin by binding to nucleosomal DNA (Winkler et al. 2011b). The yeast ortholog Pob3 lacks the the HMG-box domain. However, a third protein called Nhp6A/B is loosely associated with the yeast Pob3/Spt16 heterodimer. This small HMGB protein binds to DNA and provides the HMG-box function. Whereas SSRP1/Pob3 is an essential protein, Nhp6 is not required for viability, but important for normal cell growth (Stillman 2010, Formosa 2012).

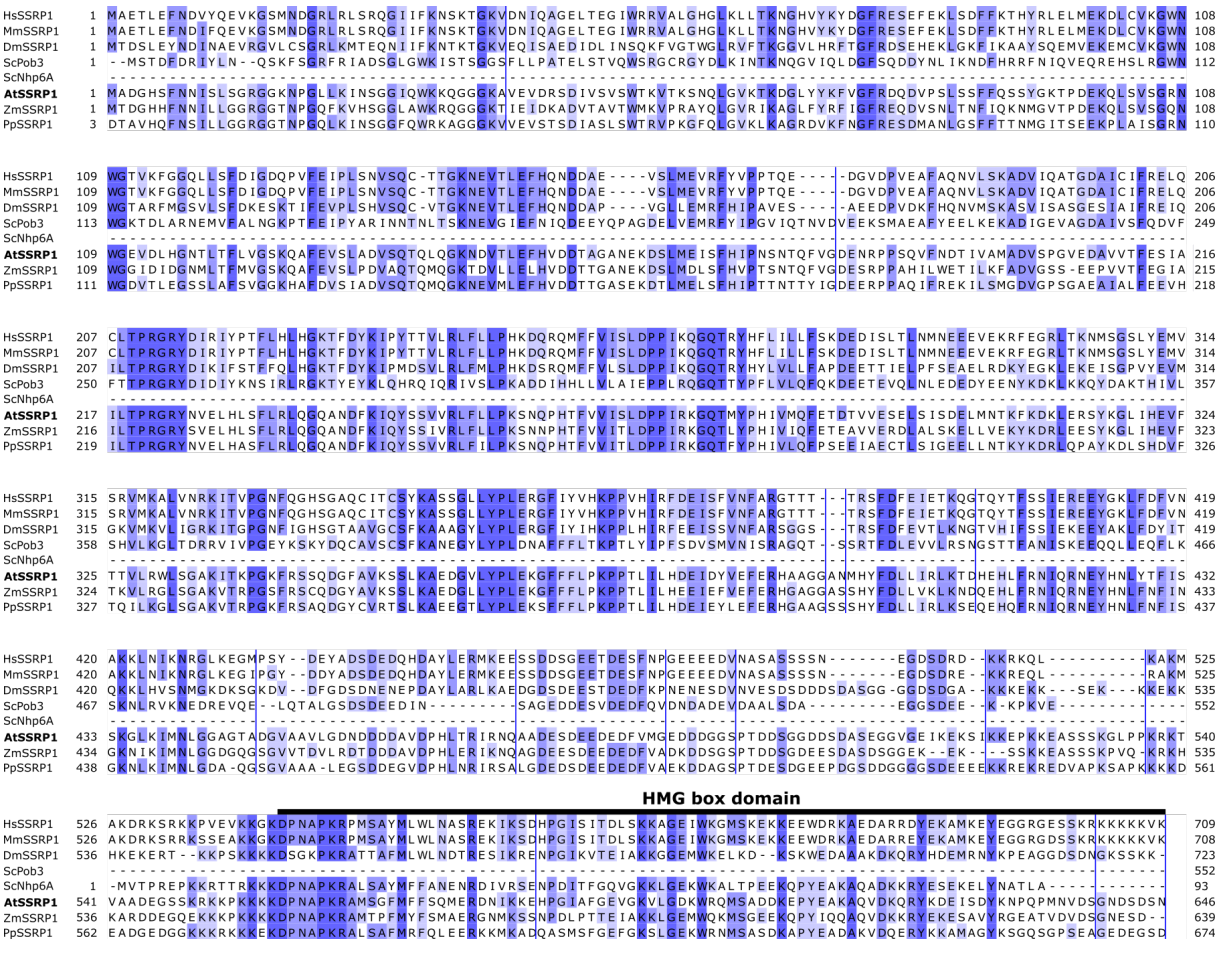


Figure 2.1 The amino acid sequence of *Arabidopsis* SSRP1 is highly conserved among different species.
The SSRP1 protein sequence was aligned to its orthologs from the following other species: *Homo sapiens* (*Hs*), *Mus musculus* (*Mm*), *Drosophila melanogaster* (*Dm*), *Saccharomyces cerevisiae* (*Sc*), *Zea mays* (*Zm*) and *Physcomitrella patens* (*Pp*) as representatives for mammals (*Ms*, *Mm*), invertebrates (*Dm*), fungi (*Sc*), dicotyledons (*At*), monocotyledons (*Zm*) and mosses (*Pp*). The yeast ortholog of AtSSRP1 is Pob3 that is lacking the C-terminal HMG-box domain. Nhp6 provides the HMG-box function for the yeast FACT complex. The multiple sequence alignment was generated using Clustal Omega (Sievers et al. 2011) and the conservation of aa residues was highlighted with blue coloration using JalView (Waterhouse et al. 2009).

2.2 Truncated SSRP1△HMG showed reduced DNA- and nucleosome-binding

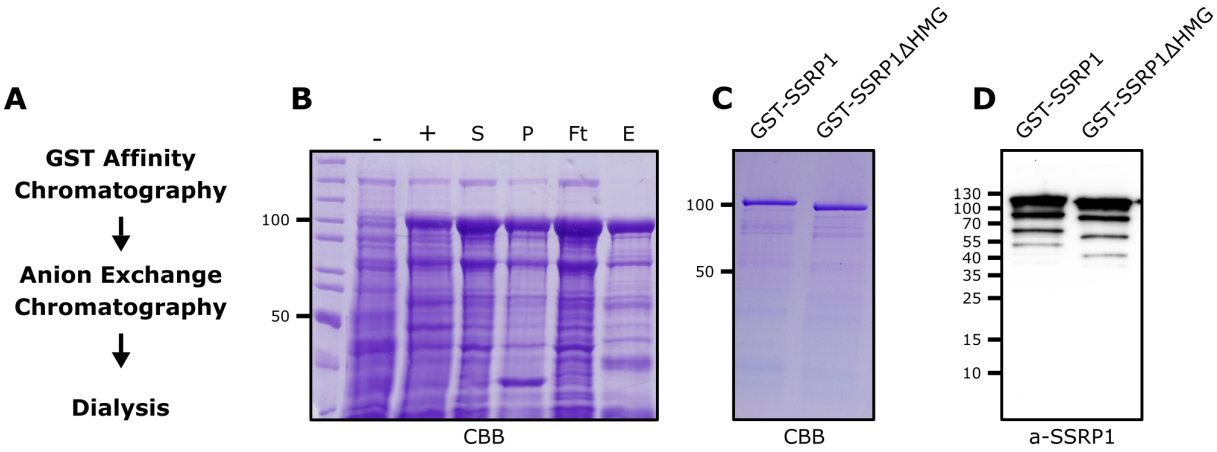


Figure 2.2 Two-step purification of the recombinant GST-SSRP1 and GST-SSRP1△HMG proteins.
A) Overview of the two-step chromatography with adjacent dialysis of the recombinant proteins. The GST affinity purification was followed by an anion exchange chromatography and dialysis. B) First, the GST-tagged SSRP1 proteins were expressed in *E. coli* BL21 cells and affinity purified using Glutathione Sepharose. All steps were monitored by SDS-PAGE and Coomassie staining (CBB). A representative gel for the expression/purification of the full length SSRP1 was depicted. The whole cell extracts, before (-) and two hours after (+) IPTG induction are shown. Additionally, the samples of the supernatant (S), the pellet (P), the flow-through (FT) and the eluate (E) were analyzed. Following affinity purification, the recombinant SSRP1 proteins were subjected to an anion exchange chromatography and dialysis. The two-step purified proteins were first analyzed by SDS-PAGE and Coomassie staining (C) and secondly by SDS-PAGE and immunoblot analysis using SSRP1 antisera (D).

In yeast/human it has been shown that the HMG-box domain of Nhp6/SSRP1 binds nucleosomal DNA (Winkler et al. 2011b). This was also demonstrated in plants by investigations of the maize SSRP1 HMG-box domain (Röttgers et al. 2000, Lichota et al. 2001). To get more information about the HMG-box domain of *Arabidopsis* SSRP1 (AT3G28730), electrophoretic mobility shifts assays (EMSA) were performed. Therefore, recombinant full length and truncated SSRP1 proteins were expressed in *E. coli* with an N-terminal GST (Glutathione S-Transferase) tag and purified by two-step chromatography (**Figure 2.2 A**). Thereby, the affinity purification of the GST-fusion proteins was followed by an anion exchange chromatography to obtain recombinant SSRP1 and SSRP1△HMG proteins in higher purity. Subsequently, the purified proteins were desalted by dialysis and subjected to the EMSA experiments. All steps were monitored by SDS-PAGE and Coomassie staining (**Figure 2.2 B-C**) as well as by Western Blot analysis with SSRP1 specific (Duroux et al. 2004) antiserum (**Figure 2.2 D**). This verified the high purity of the recombinant SSRP1 fusion proteins. In the following, the interaction of SSRP1 and SSRP1△HMG with mononucleosomes and DNA was examined. The mononucleosomes that were either lacking or possessing linker DNA were kindly provided by Prof. Dr. Gernot Längst. Recombinant *Drosophila* histones were reconstituted into nucleosomes by salt dialysis with DNA fragments of 147 bps (without linker DNA) or 198 bps (with linker DNA) that were containing the 601 nucleosome position sequence. The mononucleosomes or the free DNA were incubated with increasing amounts of recombinant SSRP1 and SSRP1△HMG proteins, respectively. The formation of complexes was analysed by electrophoretic mobility shift assays using agarose gels. In contrast to the truncated SSRP1△HMG, the full length SSRP1 showed a high affinity for the Cy5-labeled free DNA (**Figure 2.3 A-B**). The smear indicated that multiple different DNA-

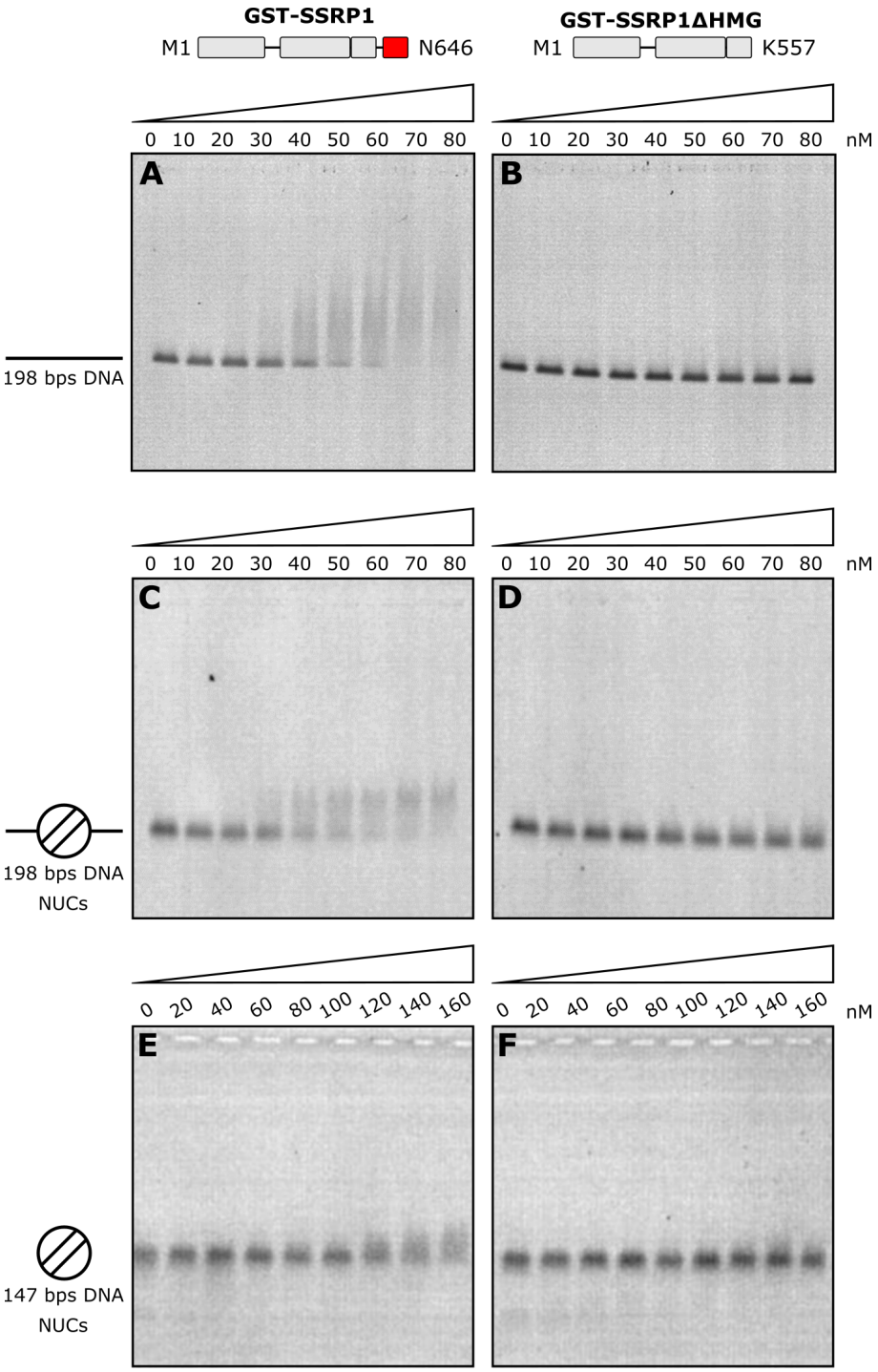


Figure 2.3 EMSA analysis revealed that the HMG-box domain of SSRP1 was crucial for the binding to DNA and mononucleosomes with linker DNA.
Increasing amounts of the recombinant GST-SSRP1(M1-N646) or GST-SSRP1(M1-K557) proteins were incubated with A-B) free DNA, C-D) mononucleosomes with linker DNA (198 bps) E-F) or mononucleosomes with only core nucleosomal DNA (147 bps). The formation of the SSRP1/DNA or SSRP1/mononucleosome complexes was analyzed by agarose gel electrophoresis. The fluorescent Cy5-labeled DNA was detected by scanning the gels with the Typhoon FLA9500.

SSRP1 complexes were formed and SSRP1 bound DNA a non sequence-specific manner. In addition, full length SSRP1 formed a specific complex with linker-DNA containing nucleosomes (**Figure 2.3 C**). Half of the nucleosomes were shifted with 50 nM of SSRP1 protein. In contrast, SSRP1 without HMG-box domain showed no binding at these concentrations and approximately twice the amount of SSRP1△HMG protein was needed to shift the nucleosomes (**Figure 2.3 D**). On the other hand, similar concentrations of full length SSRP1 ($\sim$ 120 nM) and truncated SSRP1△HMG ($\sim$ 140 nM) were needed to start shifting core nucleosomes without linker DNA (**Figure 2.3 E-F**). Compared to nucleosomes with linker DNA, the affinity of SSRP1 was lower and 160 nM recombinant SSRP1 protein was needed to shift half of the nucleosomes. The affinity of truncated SSRP1△HMG was similar for nucleosomes with and without linker DNA.

Taken together, the ability of SSRP1 to bind DNA is mediated by the HMG-box domain. Full length SSRP1 had a high affinity for DNA and nucleosomes with linker DNA, whereas the affinity for nucleosomes with core DNA was reduced. In contrast, SSRP1 lacking the HMG-box domain did not bind to DNA and showed a reduced binding affinity to nucleosomes with linker-DNA in comparison to SSRP1.

2.3 Full-length SSRP1 and truncated SSRP1△HMG showed similar protein dynamics

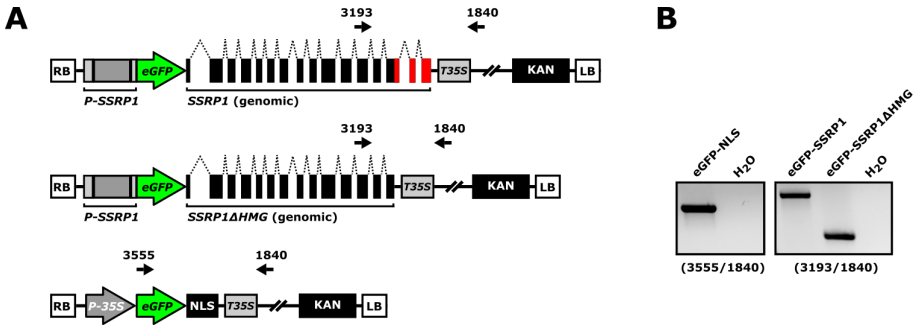


Figure 2.4 Generation of transgenic *Arabidopsis* cell lines -expressing eGFP-SSRP1 and eGFP-SSRP1△HMG. A) Schematic illustration of the eGFP-SSRP1, eGFP-SSRP1△HMG and eGFP-NLS transgenes. The genomic sequences of the full length or the truncated *SSRP1* were expressed as N-terminal translational fusions with eGFP under the native *SSRP1* promoter. Free eGFP-NLS was expressed under control of the CaMV 35S promoter. The primers for the genotyping PCR are indicated as arrows (Black bars = exons, dark grey bars = promoter, dotted lines = introns, green bar = eGFP coding sequence, light grey bars = UTR, KAN = Kanamycin resistance marker, RB/LB = Right Border/Left Border). B) The genotyping PCR of *wild type* and transgenic cell lines. The following primer pairs were used to detect the transgenes: 3193/1840 (eGFP-SSRP1 fusions) and 3555/1840 (eGFP-NLS).

The previously described EMSA experiments showed that the HMG-box domain of SSRP1 was crucial for the binding of SSRP1 to nucleosomes *in vitro*. In a next step, the role of the SSRP1 HMG-box domain in mediating SSRP1 – chromatin interactions was examined by fluorescence recovery after photo bleaching (FRAP) experiments. To characterize the dynamics of SSRP1 and SSRP1△HMG in living cells, transgenic *Arabidopsis* suspension cell lines-expressing eGFP-SSRP1 fusion proteins were generated. Therefore, transgenes were created to drive the expression of either genomic SSRP1 or SSRP1△HMG with eGFP as N-terminal translational fusion under the SSRP1 promoter, a 614 bps region upstream of the translational start site (**Figure 2.4**). A plasmid driving the expression of eGFP-NLS under the CaMV 35S promoter was generated as control. The constructs were introduced into the genome of cultured *Ara-*

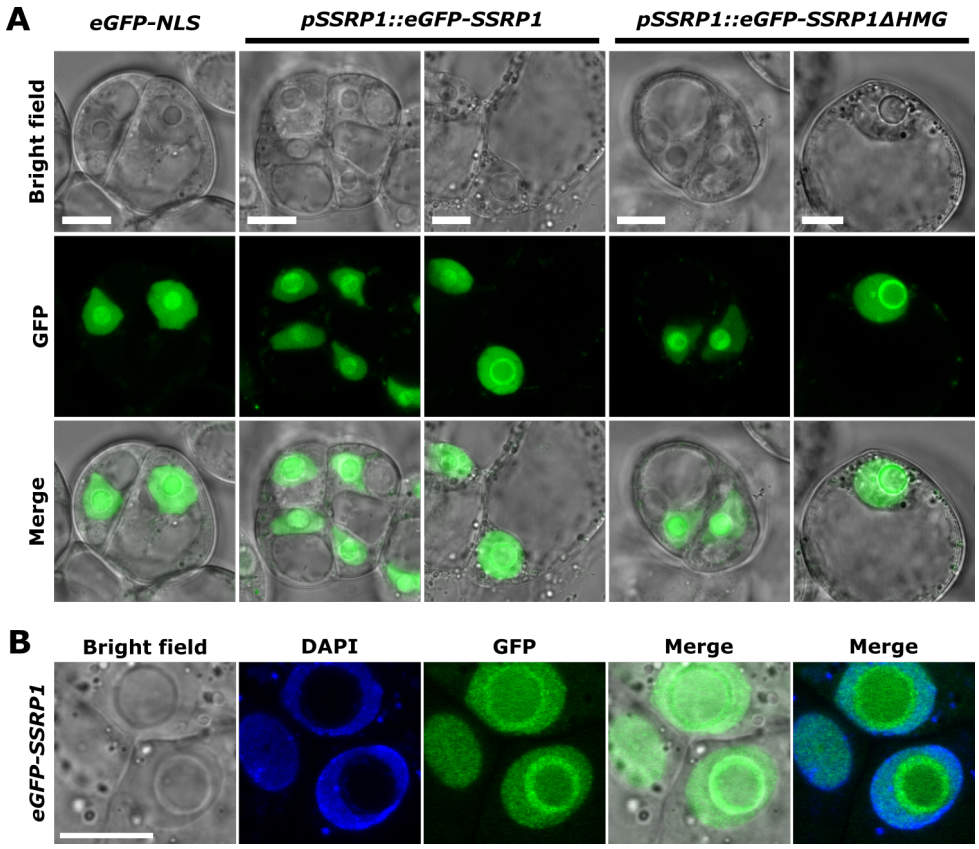
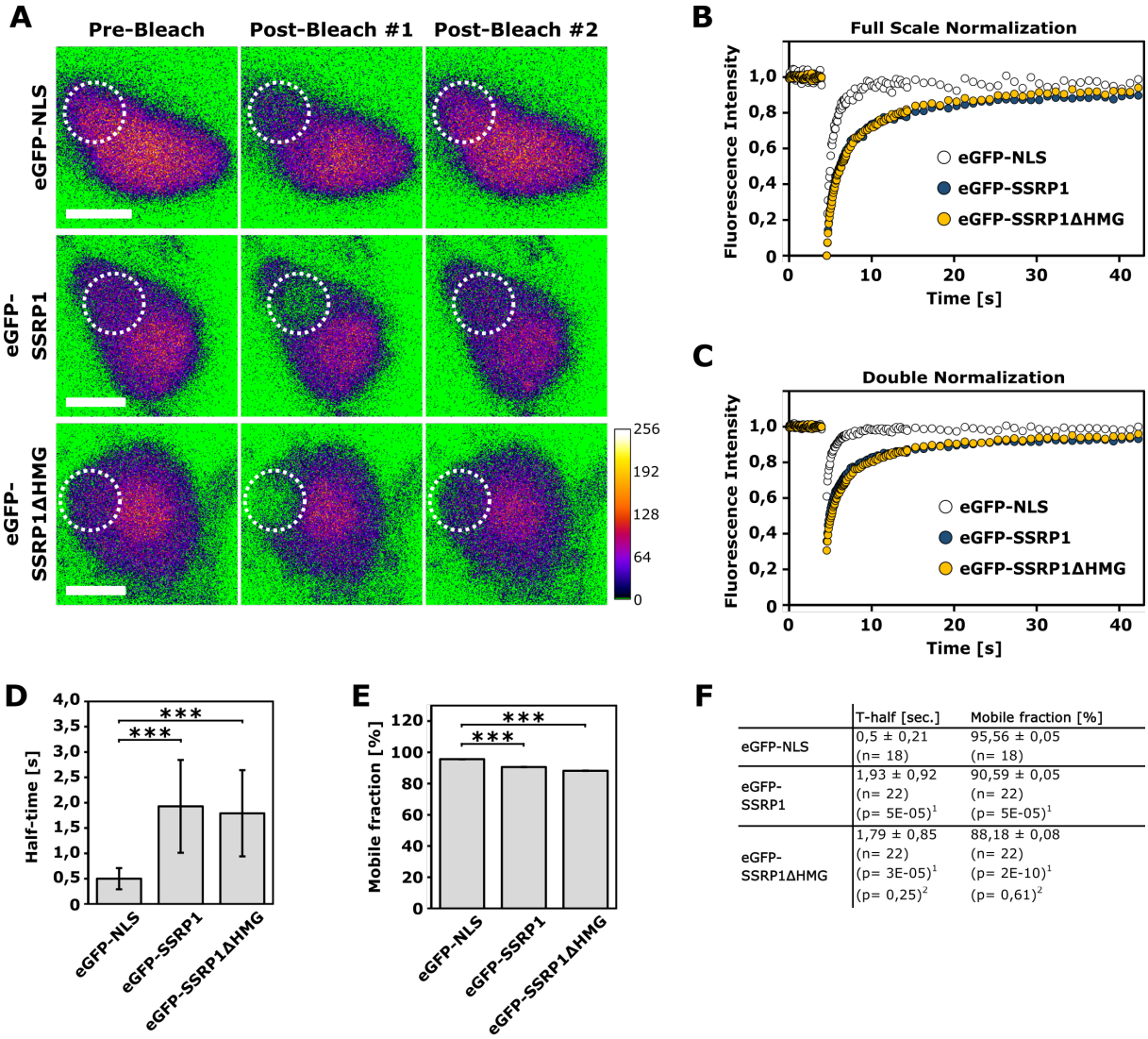


Figure 2.5 Full length and truncated SSRP1 showed the same nuclear localization in living cells.
A) Transgenic *Arabidopsis* cell lines -expressing either eGFP-SSRP1, eGFP-SSRP1△HMG or eGFP-NLS were generated by co-cultivation of PSB-D cells with transgenic *Agrobacterium*. Three days old cells were investigated by confocal microscopy. One or two representative images were selected for cell line -expressing eGFP-NLS or eGFP-SSRP1 fusion proteins, respectively. The bar indicates 10 μm. B) Cells -expressing eGFP-SSRP1 were fixed using 1 % formaldehyde and stained with DAPI. The bar indicates 10 μm.

bidopsis cells (PSB-D) by co-cultivation with transgenic *Agrobacterium* (Van Leene et al. 2011, Pfab et al. 2017) and the generated cell lines were analyzed by confocal microscopy. The control cell line showed a homogenous distribution of eGFP-NLS in the nucleus. The signal in the nucleolus was more intense (**Figure 2.5 A**). The GFP signals of both, the full length and the truncated SSRP1 were detected in the nucleoplasm and nucleolus. In approximately 5 % of the cells-expressing either full length or truncated eGFP-SSRP1, a ring-like structure surrounding the nucleolus was observable. DAPI staining of fixed transgenic cells showed that the eGFP fusion proteins accumulated at the periphery of the nucleolus (**Figure 2.5 B**). Taken together, the removal of the HMG-box domain did not change the subcellular localization of SSRP1.

In a following step, the transgenic cell lines were subjected to fluorescence recovery after photobleaching (FRAP) experiments using time-lapse confocal microscopy. The GFP signal was bleached in a defined region of interest (ROI) in the nucleoplasm and the fluorescence recovery was monitored over time (**Figure 2.6 A-C**). The time of the fluorescence recovery after the bleaching was not significantly different for the full length ($t_{1/2}$ = 1.93 s) and truncated SSRP1 ($t_{1/2}$ = 1.79 s) (**Figure 2.6 D,F**). Both versions of SSRP1 showed a similar high mobility. But both proteins recovered significantly slower than free eGFP-NLS ($t_{1/2}$ = 0.5 s). This indicated that the full length and the truncated SSRP1 were not freely diffusing in



F

	T-half [sec.]	Mobile fraction [%]
eGFP-NLS	0,5 ± 0,21 (n= 18)	95,56 ± 0,05 (n= 18)
eGFP-SSRP1	1,93 ± 0,92 (n= 22) (p= 5E-05)[1]	90,59 ± 0,05 (n= 22) (p= 5E-05)[1]
eGFP-SSRP1ΔHMG	1,79 ± 0,85 (n= 22) (p= 3E-05)[1] (p= 0,25)[2]	88,18 ± 0,08 (n= 22) (p= 2E-10)[1] (p= 0,61)[2]

Figure 2.6 Full length and truncated SSRP1 showed the same kinetics in living cells.
A) The FRAP of transgenic eGFP-SSRP1, eGFP-SSRP1△HMG and eGFP-NLS -expressing cells is shown. The circular 3 μm ROIs (indicated by doted circles) were photobleached and the recovery of the fluorescence intensity was measured over time in these areas. The Pre-Bleach indicates the first timepoint of the series (t = 0 s), the Post-Bleach the first timepoint after the bleaching (t = 4.56 s) and the Post-Bleach 2 the last time point of the series (t = 43.3 s). Pseudo-coloured images (modified fire LUT) with respective colour calibration bar are shown. B) The full scale- and C) double- normalized mean fluorescence recovery curves for eGFP-SSRP1 (n = 16), eGFP-SSRP1△HMG (n = 16) and eGFP-NLS (n = 12) are shown. This plot shows the normalized GFP intensities at 50x pre-bleach and 90x post-bleach time points. The white dots indicate eGFP-NLS, the blue dots indicate eGFP-SSRP1, and the yellow dots indicate eGFP-SSRP1△HMG. D) The Half-time ($t_{1/2}$) of eGFP-NLS (0.5 s), eGFP-SSRP1 (1.93 s) and eGFP-SSRP1△HMG (1.79 s). E) The mobile fraction of eGFP-NLS (95.6 %), eGFP-SSRP1 (90.6 %) and eGFP-SSRP1△HMG (88.2 %). F) The exact half-time and mobile fraction values of the indicated cell lines. P-Value (P) of (1) eGFP-SSRP1 or eGFP-SSRP1△HMG versus eGFP-NLS and (2) eGFP-SSRP1 versus eGFP-SSRP1△HMG. All data are means ± SD. The significance was tested by the Student's T-Test.

the nucleoplasm, but were transiently interacting with other proteins/structures to a similar extend. Additionally, the mobile fractions of SSRP1 (90.6%) and SSRP1△HMG (88.2%) were comparable, but both were significantly lower than the mobile fraction of free eGFP (95.6%) (**Figure 2.6 E,F**). This indicated that a small percentage of SSRP1 proteins was immobilized, for instance, at large structure like chromatin. In summary, SSRP1 lacking its HMG-box domain showed similar kinetics as *wild type* SSRP1. But the mobility of both, SSRP1 and SSRP1△HMG was significantly different to the free diffusion of eGFP-NLS. This suggested that the loss of the SSRP1 HMG-box domain, that mediates the binding to the nucleosomal DNA, did not impair the interaction of SSRP1 and chromatin in living cells.

2.4 HMG-box-deficient SSRP1 interacted with SPT16 and the transcriptionally active RNAPII

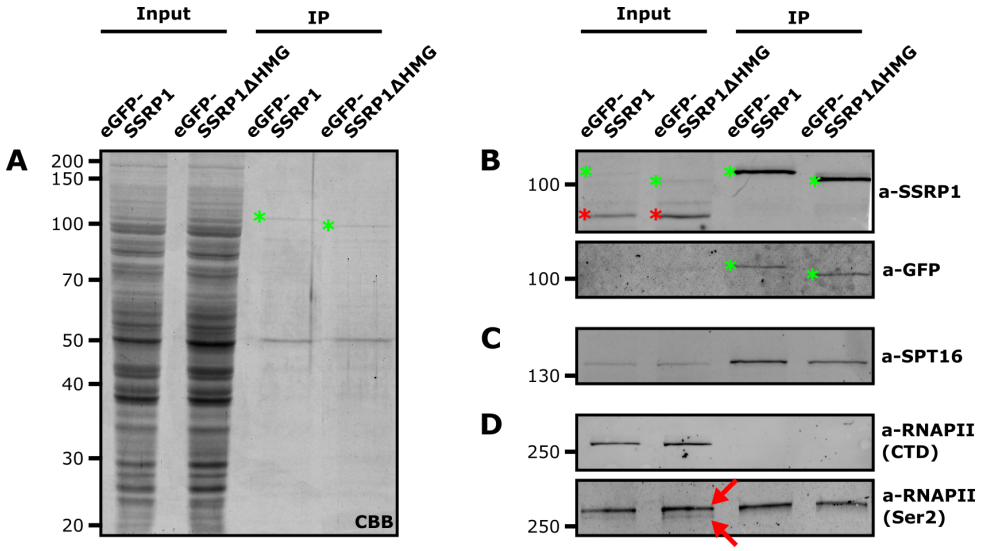


Figure 2.7 SSRP1 lacking the C-terminal HMG-box domain interacted with SPT16 and the elongating RNAPII. Immunoprecipitation of eGFP-SSRP1 and eGFP-SSRP1△HMG proteins from transgenic *Arabidopsis* cell lines using GFP-Trap™. Whole cell extracts were used for the affinity purifications. Co-purified proteins were analysed by SDS-PAGE and CBB staining (A). Endogenous SSRP1 is always indicated by red asterics and eGFP-SSRP1 fusions by green asterics. The bait proteins showed the electrophoretic mobility that is consistent with theirs predicted masses of 98.6 kDa (eGFP-SSRP1) and 88.4 kDa (eGFP-SSRP1△HMG). Immunoblot analysis of co-purified proteins (IP) versus input is shown. Different antibodies were used as following: (B) a-SSRP1 (Duroux et al. 2004) and a-GFP (Chromotek, 3H9), (C) a-SPT16 (Duroux et al. 2004), (D) a-RNAPII (CTD) (Abcam, ab817) specifically binding the non-phosphorylated C-terminal domain (CTD) of the largest subunit of RNAPII (NRPB1) and a-RNAPII (Ser2) (Abcam, ab5095) binding specifically the Ser2 phosphorylated CTD of NRPB1. Two bands appeared in the input sample by using the a-RNAPII (Ser2) antibody (Indicated by red arrows). The band migrating at a lower molecular weight represents the hypophosphorylated (NRPB1A) form of the RNAPII, the upper band the hyperphosphorylated (NRPB1O) form.

The FRAP experiments revealed that the loss of the HMG-box domain did not change the kinetics of SSRP1 and its putative interaction with chromatin. In *Arabidopsis*, the FACT complex was detected over the entire transcribed region of actively transcribed genes (Duroux et al. 2004). This raised the question if SSRP1△HMG is still part of FACT and is associated with the elongating RNAPII? Therefore, the proteins co-purifying with either the full length SSRP1 or the truncated SSRP1 lacking the HMG domain were analyzed by Western Blotting. For this purpose, the transgenic *Arabidopsis* suspension cell lines-expressing either eGFP-SSRP1 or

eGFP-SSRP1△HMG that were described before in section 2.3 were used. The total protein extracts of the transgenic PSB-D cells were used for affinity purification using GFP-Trap^{TM}beads. The co-purified proteins were first analysed by SDS-PAGE and Coomassie staining of the gel (**Figure 2.7 A**). The size difference between the full length and the truncated SSRP1 proteins was visible. Both, eGFP-SSRP1 and eGFP-SSRP1△HMG showed the electrophoretic mobility that is consistent with the predicted masses of 98.6 kDa and 88.4 kDa, respectively. No obvious differences in the pattern of the proteins that co-purified with eGFP-SSRP1 or eGFP-SSRP1△HMG were detectable on the Coomassie-stained gel.

In both cell lines, the endogenous SSRP1 and the plasmid-derived eGFP-SSRP1 fusions (full length or truncated) were detected (**Figure 2.7 B**) by Western Blotting using SSRP1 specific (Duroux et al. 2004) antiserum. The endogenous SSRP1 (indicated by red asterics) was detectable in both input samples as band migrating at 71.6 kDa, but was not enriched in the IP samples. The eGFP-SSRP1 and eGFP-SSRP1△HMG proteins (indicated by green asterics) were seen in the input samples as very faint slower migrating bands in comparison to the endogenous SSRP1. Both eGFP fusion proteins showed comparable signal intensities and a difference in their sizes. In contrast to the endogenous SSRP1, the fusion proteins were strongly enriched in the IP samples. This SSRP1 expression was validated by Western Blotting using a GFP antibody (AB) (**Figure 2.7 B**). Thereby, the endogenous SSRP1 could not be detected in the input or the IP samples. But both eGFP-SSRP1 fusion proteins were strongly enriched in the IP samples. Moreover, no free eGFP was detected on the Western Blot (not shown).

The immunoblotting using the SPT16 antiserum (Duroux et al. 2004) revealed that SPT16 was co-purified with the full length and the truncated SSRP1△HMG (**Figure 2.7 C**). This is consistent with findings that the N-terminal domain of yeast and human SSRP1 is required for the dimerization with SPT16 (Winkler et al. 2011a). Therefore, the lack of the C-terminal HMG-box domain had no effect on the binding of SSRP1 to SPT16 and the formation of the FACT complex.

To examine the SSRP1 – RNAPII interaction, a Western blot analysis using antisera directed against the non-phosphorylated C-terminal domain (CTD) of RNAPII and the Ser2-phosphorylated RNAPII-CTD was performed (**Figure 2.7 D**). The unphosphorylated form of the RNAPII was detected in the input, but not in the IP samples. With the antibody directed against the Ser2-phosphorylated RNAPII-CTD, a hypo- (NRPB1A) and a hyper-phosphorylated (NRPB1O) form of the largest RNA polymerase II subunit were detected. The hyperphosphorylated form with the slightly lower electrophoretic mobility (higher running band) was enriched by SSRP1 and SSRP1△HMG affinity purification. These results demonstrated that the Ser2-phosphorylated RNAPII associated with SSRP1 in a HMG-box independent manner, whereas the non phosophorylated RNAPII did not. This indicated that the HMG-box-deficient FACT complex is part of the active RNAPII transcript elongation complex.

2.5 SSRP1 HMG-box domain is not important for proper plant development

On the one hand the HMG-box domain is essential for SSRP1 to bind DNA and nucleosomes *in vitro* (Section 11.2.5), on the other hand HMG-box-deficient SSRP1 still retains chromatin interaction properties *in vivo* (Section 2.3) and is part of the active RNAPII transcript elonga-

tion complex (Section 2.4). In yeast, the HMG-box function of the heterodimer Spt16/Pob3 is provided by loosely associated Nhp6, which is not essential for viability (Stillman 2010, Winkler et al. 2011a, Formosa 2012). To get more information about the role of the *Arabidopsis* SSRP1 HMG-box domain, mutant plants lacking this domain were characterized. The *ssrp1-1* mutant harbours a transposon (GT7431) insertion in the first exon of *SSRP1*, which is homozygous lethal (Lolas et al. 2010; (**Figure 2.8 A**)). The genetic background of this GeneTrap mutant (Sundaresan et al. 1995) is ecotype Landsberg erecta-0 (Ler-0). The HMG-box-deficiency mutants were generated by complementation of the *ssrp1-1* mutant with transgenes expressing full length SSRP1 (control) and truncated SSRP1△HMG under its native promoter.

2.5.1 SSRP1 HMG-box deficiency mutants

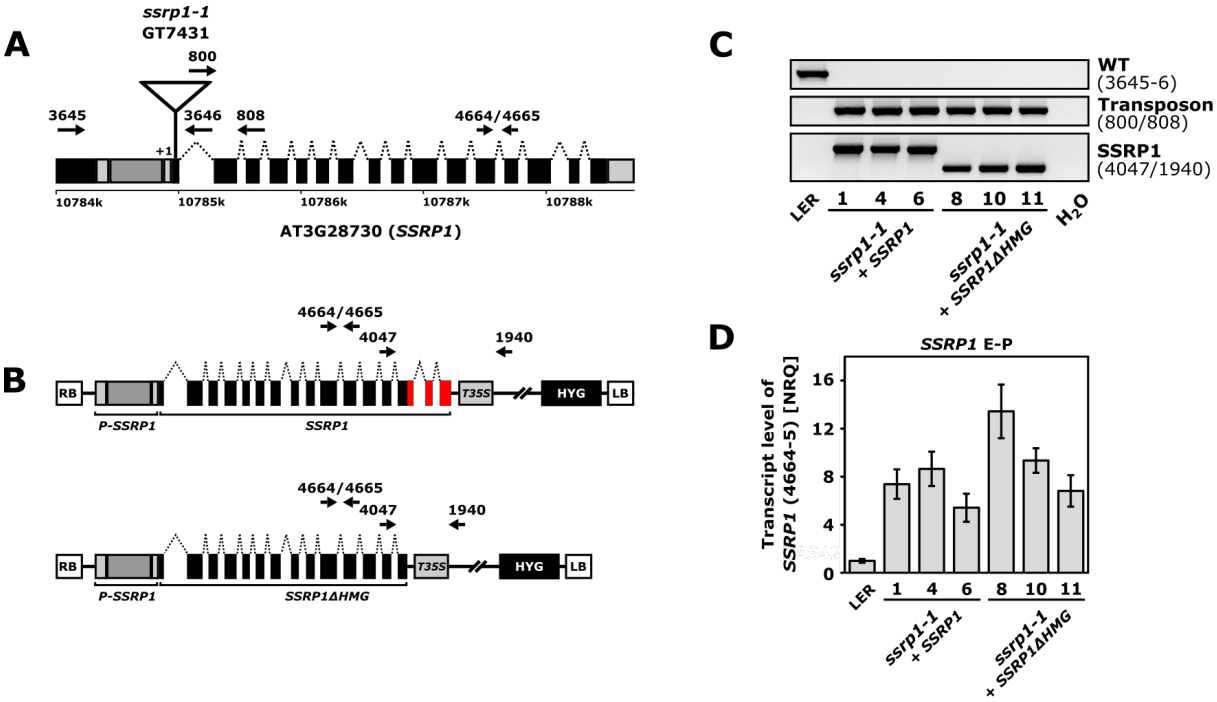


Figure 2.8 Molecular characterization of *ssrp1-1* complementation lines.
A) Schematic illustration of the *SSRP1* locus (AT3G28730) with a transposon insertion (GT7431) in the first exon. B) Schematic illustration of transgenes driving the expression of either *wild type* or truncated SSRP1 lacking the HMG-box domain under the native promoter. The C-terminal HMG-box domain is depicted in red. The primers used for genotyping and expression analysis are indicated by arrows (Black bars = exons, dark grey bars = promoter, dotted lines = introns, HYG = hygromycin resistance marker, light grey bars = UTR). C) Genotyping PCR of *wild type* Ler-0 and *ssrp1-1* complementation lines. The following primer pairs were used to detect the *wild type* and the mutant *SSRP1* allele, 3645/3646 and 800/808 respectively. D) Relative expression of *SSRP1* in *wild type* Ler-0 and complementation lines was determined by qRT-PCR. The primer pair (4664/4665) was spanning the intron 12 of *SSRP1* to detect the endogenous and the plasmid-derived *SSRP1* transcripts (*SSRP1* E-P). The normalized relative quantities (NRQ) ± the normalized relative standard error (SE[NRQ]) are shown. The relative quantity of *SSRP1* mRNA was normalized to the relative quantities of *GAPC*, *PP2AA3* and *UBI10* mRNA (*** indicates P-Value<0.001, Student's T-Test).

Therefore, either the *wild type* genomic *SSRP1* locus (4.16 kb) or the genomic *SSRP1* locus lacking the HMG-box domain (3.63 kb) were cloned into pGreen0179 plasmids (**Figure 2.8 B**). The expression of *SSRP1* was thereby driven by the SSRP1 promoter, a 614-nt region upstream of the translational start site (including the *SSRP1* 5′UTR and the 3′UTR of the upstream gene).

Due to the low efficiency of a direct transformation of heterozygous *ssrp1-1* mutants, both

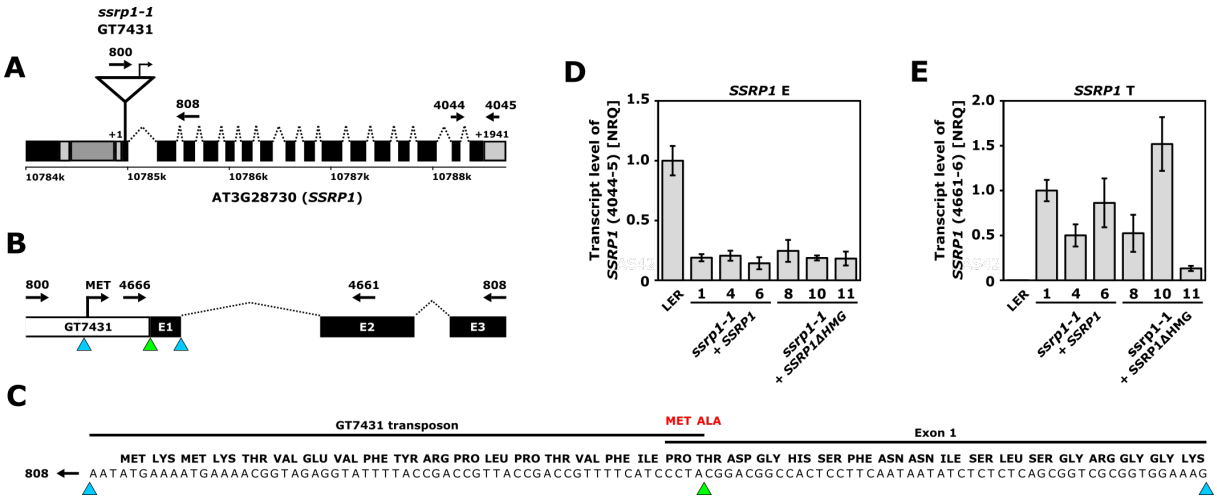


Figure 2.9 *ssrp1-1* transposon insertion mutants showed a weak expression of an aberrant *SSRP1* transcript that was not sufficient for plant viability
A) Schematic illustration of the *SSRP1* locus (AT3G28730) with a transposon insertion (GT7431) in the first exon. B) Schematic illustration of the PCR-amplicon to determine the position of the transposon insertion and potential transposon-derived ORFs by DNA sequencing. The DNA fragment was amplified by the primers 800 and 808 that were binding to the transposon and the third exon of *SSRP1*, respectively. The putative translational start site of an alternative ORF within the transposon is indicated by an arrow (MET). C) The sequencing of the PCR-amplicon (800/808) with the primer 808 revealed the exact position of the transposon insertion site (green triangle) within the first exon of *SSRP1* and the start codon of an alternative ORF within the transposon (58 nt upstream of the insertion site). The primary structure of the aberrant (black) and the *wild type* (red) SSRP1 protein are shown. The transposon insertion results in an abberant *SSRP1* transcript that is potentially translated into a mutated SSRP1. Thereby, the first two aa (MET, ALA) of the WT SSRP1 are replaced by 20 transgene-derived aa. D-E) The relative expression of *SSRP1* in *wild type* Ler-0 and the complementation lines was determined by qRT-PCR. D) The primer pair (4044/4045) was binding to the last exon of *SSRP1* and the 3′UTR to detect just the endogenous *SSRP1* (*SSRP1* E). The normalized relative quantities (NRQs) were normalized to *wild type* Ler-0 E) The primer pair (4661/4666) that was binding to the transposon and the third exon of *SSRP1* was used to detect the aberrant *SSRP1* transcript (*SSRP1* T). The NRQs were normlized to the complementation line #1, because no transcript was produced in the *wild type* Ler-0 plants

transgenes were first integrated into *wild type* Ler-0 plants (T0) by *Agrobacterium*-mediated transformation. For each transgene, three independent primary-transformants (T1) were selected on medium containing hygromycin and subsequently crossed with heterozygous *ssrp1-1* mutants. In the T3 generation, the complementation lines expressing *wild type SSRP1* or *SSRP1△HMG* in the putative *SSRP1* knockout background (homozygous for *ssrp1-1*) were selected. In total, three full length *SSRP1* (#1, #4, #6) and three truncated *SSRP1* (#8, #10, #11) lines were obtained. The homozygous insertion of the *ssrp1-1* transposon in the first exon of *SSRP1* and the integration of the complementation construct in the genome was shown by genotyping PCR using the primers depicted in **Figure 2.8 A, B, and C**.

Next, the transcript levels of the plasmid-derived *SSRP1*/*SSRP1△HMG* and the endogenous *SSRP1* were analysed in the generated *SSRP1* complementation lines in comparison to *wild type* Ler-0. Therefore, the total RNA was extracted from 10-days old seedlings and transcribed into cDNA as template for quantitative Real-Time PCR (qRT-PCR). The detection of the endogenous and the transgene-derived *SSRP1* mRNA (*SSRP1* E-P) was carried out by using primers binding to the second exon of *SSRP1* (**Figure 2.8 A**). Compared to *wild type*, all six complementation lines showed increased levels of the *SSRP1* transcript as following: SSRP1 #1 (7.4 fold), #4 (8.6 fold), #6 (5.4 fold), SSRP1△HMG #8 (13.4 fold), #10 (9.3 fold), #11 (6.8 fold) (**Figure 2.8 D**).

The complementation lines were homozygous for the transposon insertion in the first exon of *SSRP1*. According to Lolas et al. 2010, this results most likely in a complete knockout of *SSRP1* as homozygous *ssrp1-1* mutants are embryonic lethal. To confirm that no endogenous *SSRP1* was expressed in the *SSRP1* complementation lines, specific primers (Binding to exon 15 and 3′UTR as depicted in **Figure 2.9 A**) were used to detect just the endogenous *SSRP1* mRNA (*SSRP1* -E). Unexpectedly the six complementation lines showed low expression of the endogenous *SSRP1* compared to *wild type*: SSRP1 #1 (18.9 %), #4 (20.6 %), #6 (14.3 %), SSRP1△HMG #8 (24.6 %), #10 (18.7 %), #11 (18.2 %)(**Figure 2.9 D**).

One possible explanation for the detection of *SSRP1* transcript in the *ssrp1-1* background is the occurrence of an lower expressed transcript that starts within the transposon insertion. This abberant *SSRP1* transcript could potentially generate a N-terminally truncated SSRP1. To test this hypothesis, the exact position of the transposon insertion in *SSRP1* was verified by PCR coupled to DNA sequencing of the amplicon. The primers used for this purpose, binding to the genomic *SSRP1* sequence as well as to the GT7431-insertion, were depicted in **Figure 2.9 A,B**. The insertion was confirmed in the first exon of *SSRP1* between +4-adenine and +5-cytosine (the translational start site is +1; **Figure 2.9 C**) as already shown before in Lolas et al. 2010. Moreover, the DNA sequencing revealed that an alternative open reading frame started within the transposon-insertion and was in-frame with the first exon of *SSRP1*. In the following, qRT-PCR analysis was performed to test, if this alternative ORF was expressed and if the putative transcript was correctly spliced as the endogenous *SSRP1*. Hence primers were used to amplify specifically the putative abberant *SSRP1* transcript (*SSRP1* T) and not the endogenous *SSRP1* as well as to span the first intron of *SSRP1*. The qRT-PCR analysis revealed that the transposon-derived ORF was expressed and the intron was correctly spliced out (**Figure 2.9 B**). These data suggested that homozygous *ssrp1-1* mutant plants and therefore also the generated SSRP1 complementation lines were expressing low levels of an aberrant N-terminally mutated *SSRP1* version (**Figure 2.9 E**). Whether the mutated *SSRP1* transcript is really translated is not known, but it seems not sufficient for plant viability as homoygous *ssrp1-1* mutants were embryonic lethal.

2.5.2 Phenotypic analysis of SSRP1 HMG-box deficiency mutants

In the following, the generated HMG-box deficiency mutants were phenotypically analysed in comparison to the full length SSRP1 complementation lines and *wild type* Ler-0 plants (**Figure 2.10**). The plants were grown on soil under long day conditions to monitor their morphology and development. The following plant characteristics were examined: The time of bolting (Elongation of the first internode), the rosette diameter and the number of leaves at bolting, the plant height and the rosette diameter at DAS35, the plant height and the primary inflorescences at DAS42, the flowers and the seed set. This experiment showed that the homozygous lethal *ssrp1-1* mutant plants could be rescued by the expression of genomic SSRP1 and SSRP1△HMG (**Figure 2.10, Table S14**). The three independent lines expressing full length SSRP1 and as well the three independent lines expressing the truncated SSRP1 showed essentially *wild type* appearance like Ler-0 plants. Moreover, both SSRP1 complementation lines showed normal *wild type* flowers and fully elongated siliques (**Figure 2.11**).

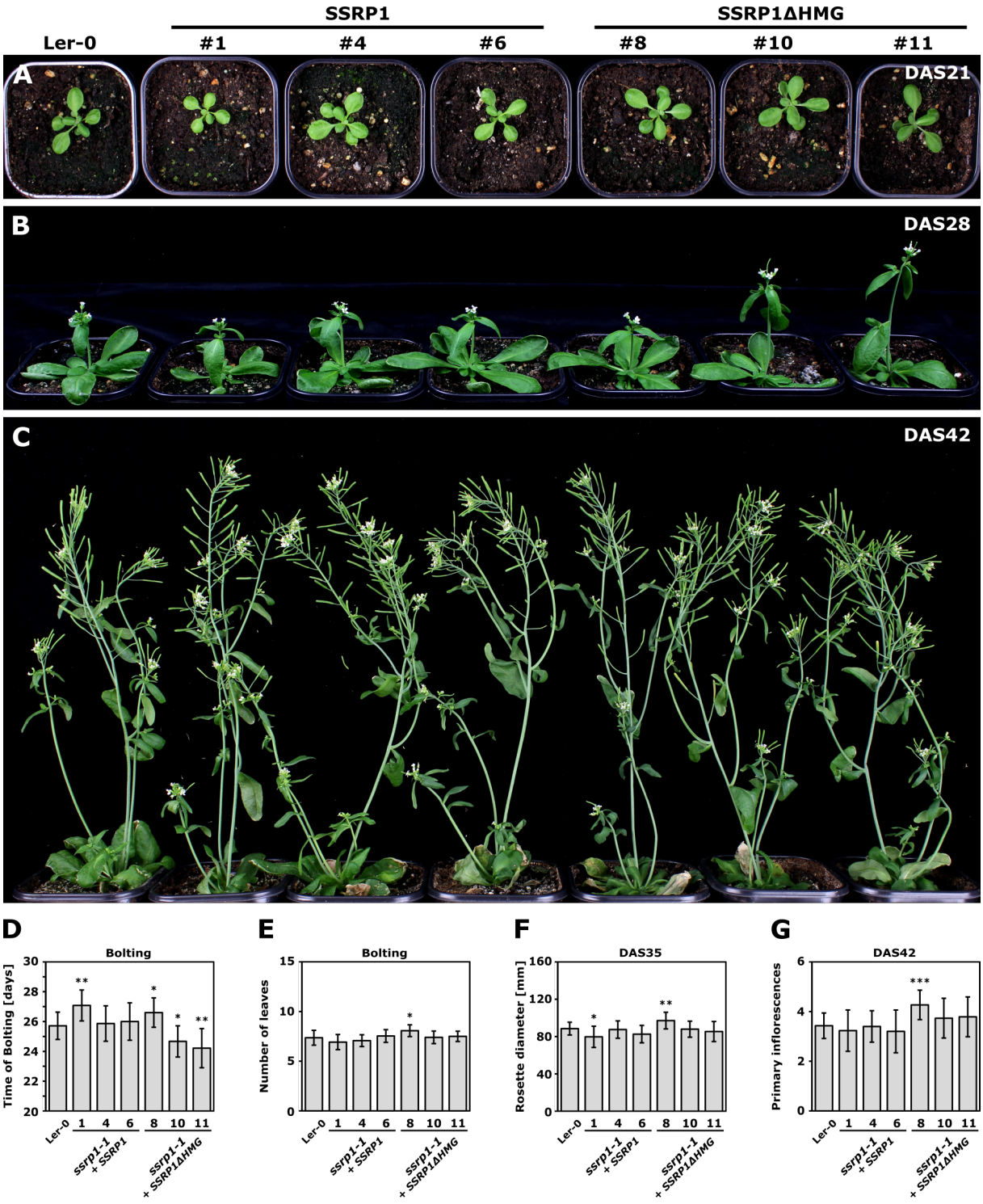


Figure 2.10 The HMG-deficiency mutants showed no obvious phenotype.
Phenotypical analysis of the *ssrp1-1* complementation lines in comparison to *wild type* Ler-0 plants grown under LD conditions. Representative individuals are shown at various developmental stages A) DAS21 (days after stratification) B) DAS28 C) DAS42 D-G) The following plant features were statistically evaluated: (D) Time of bolting (Elongation of the first internode), (E) number of leaves at bolting, (F) Rosette diameter at DAS35 and (G) primary inflorescences at DAS42 (All data are means ± SD, * indicates P-Value<0.05, ** indicates P-Value <0.01, *** indicates P-Value<0.001, Student's T-Test).

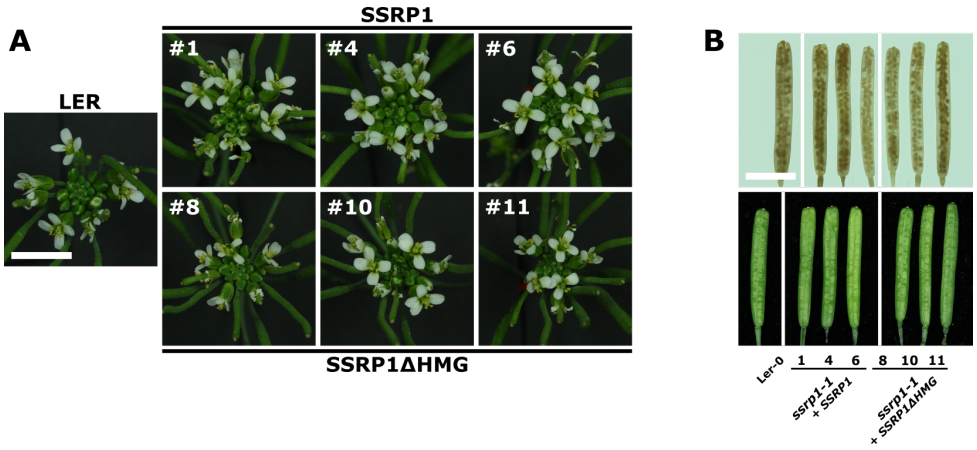


Figure 2.11 The HMG-deficiency mutants showed normal flower architecture and seed set.

Phenotypical analysis of the *ssrp1-1* complementation lines in comparison to *wild type* Ler-0 plants grown under LD conditions. A) Flowers buds B) Bleached and non-bleached siliques

In summary, the *wild type* phenotype of *ssrp1-1* mutants could be rescued by the genomic sequence of SSRP1 as well as SSRP1△HMG. The lack of the SSRP1 HMG-box domain had no effect on proper plant development.

2.6 Overexpression of SSRP1 and SSRP1△HMG had no dominant negative effect on plant development

As protein overexpression can cause mutant phenotypes (Herskowitz 1987), *wild type* and truncated SSRP1△HMG were overexpressed in *Arabidopsis* Ler-0 to see if this had a dominant negative effect on plant growth and development. Therefore, transgenes expressing either SSRP1 (CDS) or SSRP1△HMG (CDS) under control of the strong *UBIQUITIN-10* (UBI10) promoter were integrated mediated by *Agrobacterium* into the genome of *wild type* Ler-0 plants. For both transgenes, three independent overexpression lines (*ssrp1-OE*#1,4,6 and *ssrp1△HMG-OE*#8,10,11) were selected on MS plates supplemented with kanamycin and genotyped using the primers depicted in **Figure 2.12 A**. The transcript levels of *SSRP1* were determined in T2 plants that were homozygous for the integrated *ssrp1-OE* transgenes. Therefore, the total RNA was extracted from 10-days old seedlings and converted into cDNA as template for qRT-PCR. To detect the endogenous and the transgene-derived *SSRP1* mRNA, primers were used that were binding to the second exon of *SSRP1* (**Figure 2.12 A**). Compared to *wild type* Ler-0 plants, all six *ssrp1-OE* lines showed a strong upregulation of the *SSRP1* transcript as follows (**Figure 2.12 B**): 54-, 11- and 20-fold for the *ssrp1-OE*#2, 3 and 5 lines as well as 16-, 9- and 11-fold for the *ssrp1△HMG-OE*#11, 13 and 14 lines. In a next step, SSRP1 overexpression lines were phenotypically analysed in comparison to *wild type* Ler-0 plants. The plants were grown on soil under long day conditions to monitor their morphology and development regarding several plant characteristics depicted in **Figure 2.12 D-L, Table S15**. This revealed that the overexpression of full length *SSRP1* caused only slight, but measureable effects. Plants overexpressing *SSRP1△HMG* showed *wild type* appearance as well, except for a slightly reduced rosette size and number of primary inflorescences. Taken together, neither the overexpression of *SSRP1* nor *SSRP1△HMG* showed a dominant negative effect on plant growth and development.

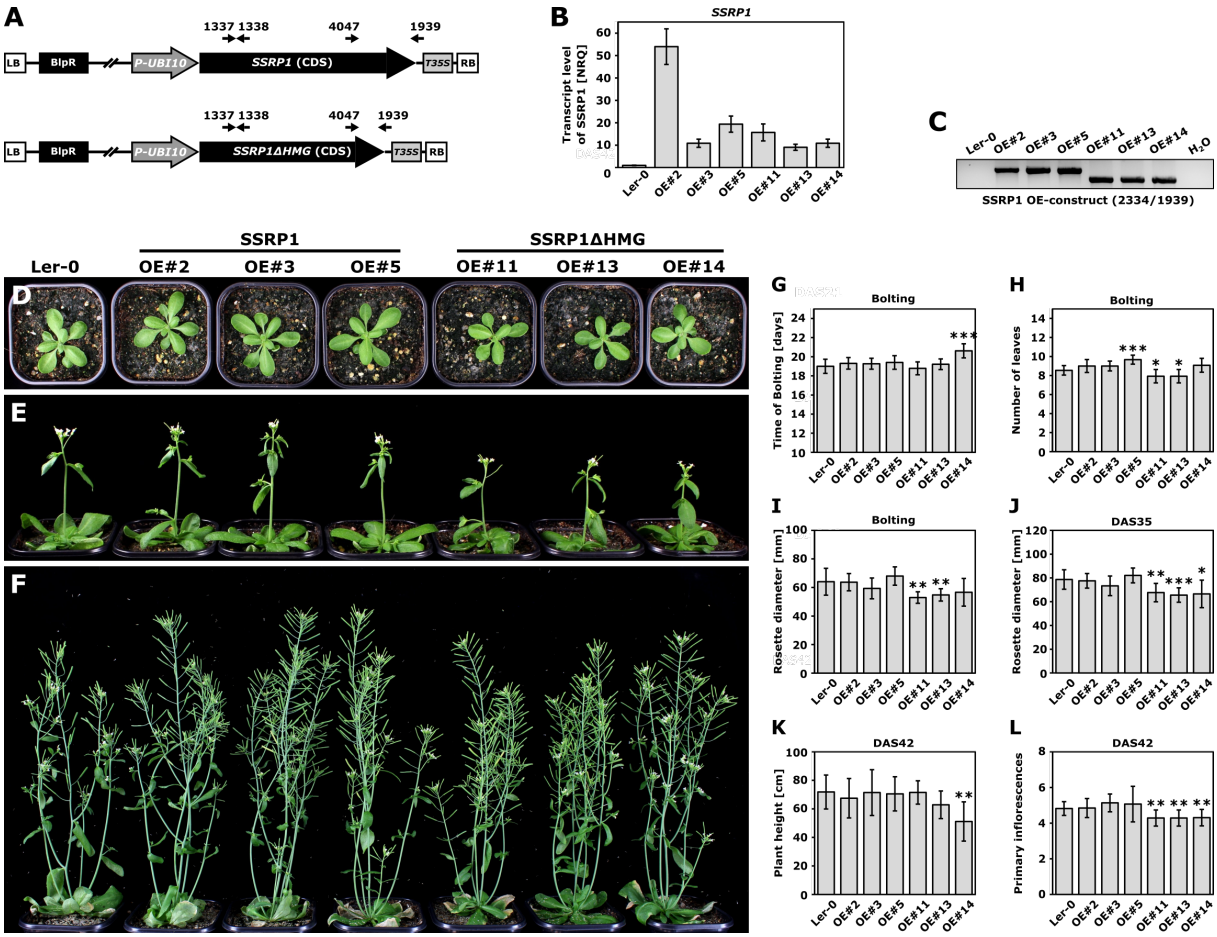


Figure 2.12 Overexpression of *SSRP1* or *SSRP1△HMG* showed no dominant negative effect on plant development.

A) Schematic illustration of the transgenes for *SSRP1* and *SSRP1△HMG* overexpression. The expression of the full length or the truncated coding sequence of *SSRP1* was driven by the *UBI10* promoter. The primers used for genotyping and expression analysis were indicated as arrows (Black bars = exons, BlpR = Basta resistance marker, dark grey bars = promoter, light grey bars = UTR, white bars = left/right border of T-DNA). B) The relative expression of *SSRP1* in *wild type* Ler-0 and *SSRP1* overexpression lines was determined by qRT-PCR. For the detection of the endogenous and the transgene-derived *SSRP1* mRNA, the primer pair 1337/1338 (binding to the second exon of *SSRP1*) was used. The normalized relative quantities (NRQ) ± normalized relative standard error (SE[NRQ]) are shown. The relative quantity of *SSRP1* mRNA was normalized to the relative quantities of *GAPC*, *PP2AA3* and *UBI10* mRNA. The significance was tested by the Student's T-Test (*** indicates $P<0.001$). C) Genotyping PCR of *wild type* Ler-0 and *SSRP1* overexpression lines. The primer pair 2334/1939 was used to detect the overexpression transgenes that were integrated in the genome. D-F) Phenotypical analysis of the SSRP1 overexpression lines in comparison to *wild type* Ler-0 plants grown under LD conditions. Representative individuals were shown at various developmental stages (D) DAS21 (days after stratification) (E) DAS28 (F) DAS42 G-L) The following plant features were statistically evaluated: (G) The time of bolting (Elongation of the first internode), (I) the rosette diameter and (H) the number of leaves at bolting, (J) the rosette diameter at DAS35, (K) the plant height and (L) the primary inflorescences at DAS42 (All data are means ± SD, Significance was tested by Student's T-Test, *** indicates $P<0.001$)

3 Results: The role of FACT in anthocyanin biosynthesis

3.1 Transcript profiling of *ssrp1-2* and *spt16-1* mutants

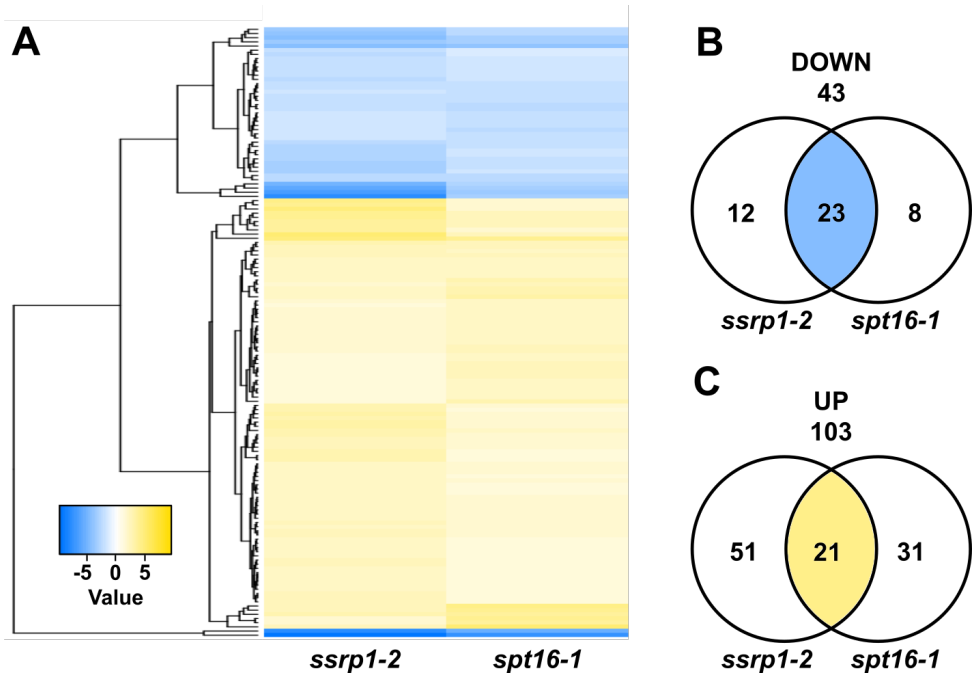


Figure 3.1 *ssrp1-2* and *spt16-1* mutants shared most of the differnetially expressed genes.
A) The microarray analysis on *ssrp1-2* and *spt16-1* mutants relative to *wild type* Col-0 plants. Differentially ($\geq$ 2-fold with $p < 0.05$) expressed genes were hierarchical clustered and visualized as heatmap. In total, 146 genes were included and represented as horizontal lines. Like shown in the color calibration bar (log2 fold changes), the different shades of yellow/blue indicated the upregulated and the downregulated genes, respectively. B) The Venn diagrams show the up- or down-regulated ($\geq$ 2-fold with $p < 0.05$) genes in *ssrp1-2* and *spt16-1* mutants compared to *wild type* Col-0 plants.

The conserved FACT complex plays a central role during RNAPII-mediated transcript elongation. By destabilizing the nucleosome barrier, the histone chaperone FACT enables together with other factors the passage of the transcription machinery through the chromatin template (Belotserkovskaya et al. 2004, Formosa 2012). In the two FACT knockdown mutants, *ssrp1-2* and *spt16-1*, the reduced expression of the disrupted genes results in similar pleiotropic phenotypes including a "bushy" appearance, early flowering and a reduced seed set (Lolas et al. 2010). The T-DNA insertions are located either in the last exon of *SSRP1* or the 5′UTR of *SPT16*. To identify FACT-dependent alteration in the genome-wide gene expression, the transcriptome profiles of *ssrp1-2* and *spt16-1* mutants were analyzed by DNA microarray hybridization. The transcriptomic profiles of *ssrp1-2* and *spt16-1* mutants were analyzed comparatively to *wild type* Col-0 plants. Therefore, both FACT mutants and *wild type* Col-0 plants were grown on MS plates under long day conditions for 10 days to isolate the total RNA from their aerial parts. The experiment was done in three biological replicates. The sample processing and the Affymetrix microarray hybridization were carried out at a genomics core facility: Center of Excellence for Fluorescent Bioanalytics (KFB, University of Regensburg, Germany).

The analysis of this experiment revealed that in total, 146 genes were significantly differentially ($\geq$2-fold, $P<0.05$) expressed in *ssrp1-2* mutant plants, *spt16-1* mutant plants or both in comparison to *wild type* Col-0 (**Tables S21 and S22**). These expression data were clustered and visualized in form of a heatmap (Babicki et al. 2016) that is shown in **Figure 3.1 A**.

Relative to Col-0, 43 genes (0.2%) were down-regulated and 103 genes (0.5%) were upregulated in at least one of the two FACT mutants (**Figure 3.1 B**). In detail, 72 and 52 genes were at least 2-fold upregulated (P<0.05) as well as 52 and 31 genes were at least 2-fold downregulated (P<0.05) in *ssrp1-2* and *spt16-1* mutants, respectively.

This showed the remarkable similarity of the gene expression profiles of both FACT mutants, *ssrp1-2* and *spt16-1*. In general, slightly more genes were determined as significantly differentially expressed in the *ssrp1-2* compared to the *spt16-1* mutant. In *SSRP1*-depleted plants, 1.13 times and 1.38 times more genes were significantly up- and down-regulated (≥2fold; P<0.05) relativ to *SPT16*-depleted plants, respectively. But almost all of the differentially expressed genes in *ssrp1-2* and *spt16-1* showed the same trend of being up or down-regulated in both mutants. The genes that were up/down-regulated in the one mutant, showed the same tendency in the other mutant and vice versa. For example, the eight genes that were ≥2-fold downregulated in *ssrp1-2*, were also downregulated in *spt16-1*, but less strongly (1.57 - 1.92 fold) (**Figure 3.1 B**). The other way around it was the same, the 12 genes that were ≥2-fold downregulated in *spt16-1*, were also downregulated in *ssrp1-2*, but to a lesser extend and below the 2-fold treshold. This explains the higher total number of significantly ≥2-fold differentially expressed genes in *ssrp1-2* (107) compared to *spt16-1* (83).

Taken together, the same set of genes was differentially expressed in *ssrp1-2* and *spt16-1* mutants and no genes were oppositely affected. This demonstrated that the reduced expression of either *SSRP1* or *SPT16* resulted genome-wide in comparable alterations in the gene expression of a small set of genes.

3.2 GO Analysis revealed that flavonoid biosynthesis genes were downregulated in FACT mutants

To shed light on the FACT-dependent biological processes, a Gene Ontology (GO) analysis was performed on the genes that were differentially (≥2-fold; P<0.05) expressed in the *ssrp1-2* and *spt16-1* mutants using agriGO (Tian et al. 2017). The genes that were upregulated upon depletion of *SSRP1*, *SPT16* or both were mostly assigned to the GO terms "Response to stimulus" (GO: 0006950) and "Response to stress" (GO: 0050896) (**Figures S1 and S2**). Most strikingly, a substantial part of the downregulated genes was allocated to GO categories related to the flavonoid biosynthetic pathway (**Figure 3.2**). In total, 10 of the 23 genes that were significantly downregulated (≥2-fold; P<0.05) in both FACT mutants were found to be integral parts of the anthocyanin biosynthesis pathway (**Figure 3.3 and 3.4**).

The biosynthetic pathway from phenylalanine to anthocyanin can be divided into three stages (Shi et al. 2014) as shown in **Figures 3.3 B**: (1) The genes of the general phenylpropanoid pathway were not affected in the FACT mutants. (2) The genes of the early steps of the flavonoid pathway including the *CHALCONE SYNTHASE* (CHS) and the *FLAVONONE 3′HYDROXYLASE* (F3′H) as well as (3) the genes of the late anthocyanin specific pathway including the *DIHYDROFLAVONOL REDUCTASE* (DFR), the *ANTHOCYANIDIN SYNTHASE* (ANS) and also the transport gene *TRANSPARENT TESTA 19* (TT19) were strongly downregulated in *SSRP1/SPT16*-depleted plants. However, the expression of the regulatory genes including the *PRODUCTION OF ANTHOCYANIN PIGMENT 1/2* (PAP1/2) that are

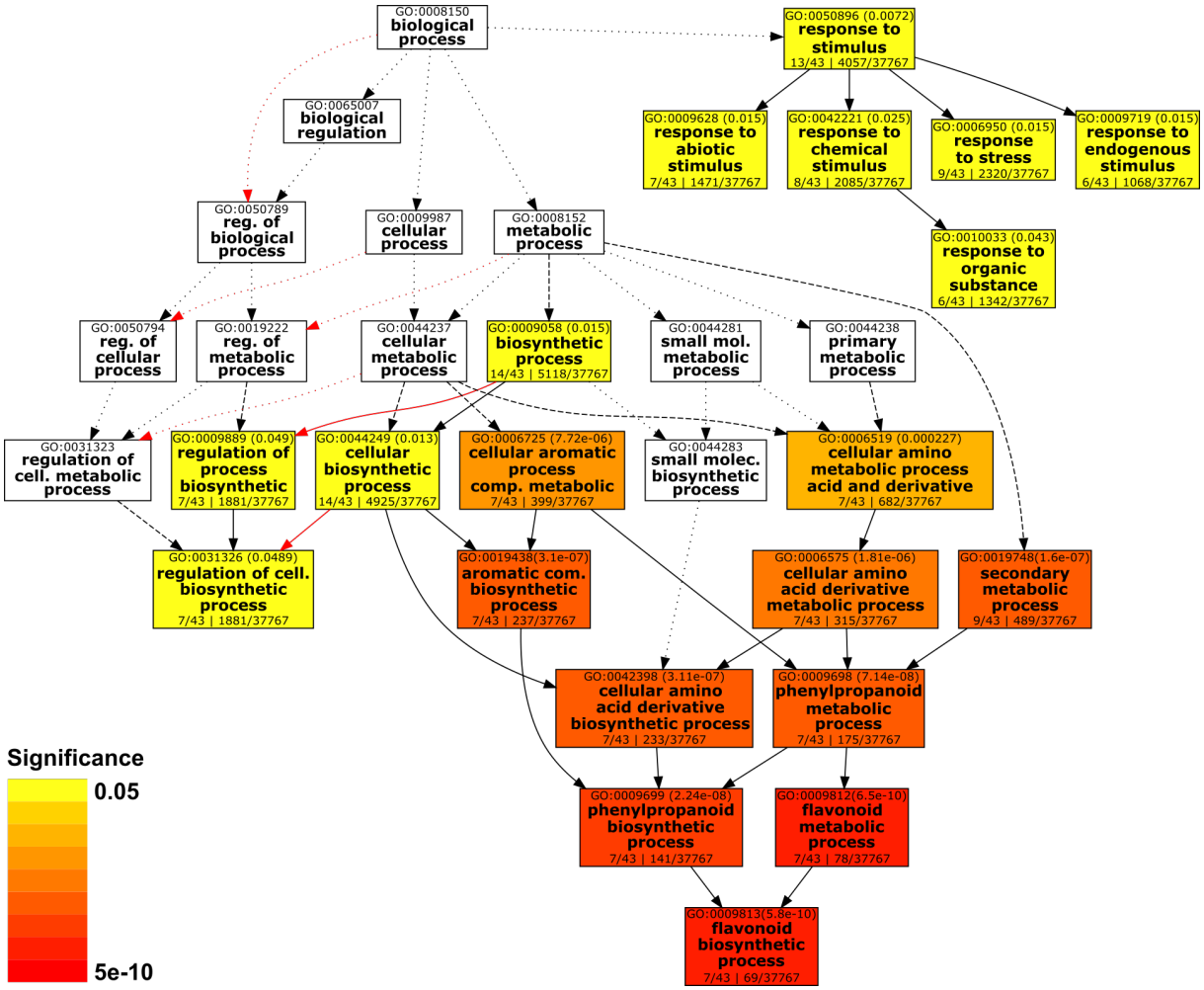


Figure 3.2 Overrepresented biological processes in the 43 genes that were downregulated in *ssrp1-2* and *spt16-1* mutants.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All 43 genes that were significantly ($p<0.5$) downregulated ($\geq$ 2-fold) in at least one of the two FACT mutants were included. Highly overrepresented GO terms are depicted in red.

controlling the expression of the anthocyanin pathway genes were not altered in a FACT-dependent manner. The expression of the affected anthocyanin biosynthesis genes was significantly ($p<0.01$) downregulated in *ssrp1-2* as well as *spt16-1* mutants.

To confirm the FACT-dependent downregulation of the anthocyanin synthesis genes as determined by the microarray-based transcriptomic data, the expression level of *CHS*, *F'3H*, *DFR*, *ANS* was analyzed by qRT-PCR in *ssrp1-2* and *spt16-1* mutants in comparison to *wild type* plants. Therefore, *wild type* plants and *ssrp1-2* as well as *spt16-1* mutants were grown under the same constant light and temperature conditions as the plants used for the microarray experiment before. The total RNA was isolated from the aerial parts of 10-days old seedlings and transcribed into cDNA as template for the qRT-PCR.

Relative to Col-0, the transcript levels of all four selected representative anthocyanin synthesis genes were strongly reduced in the *ssrp1-2* and *spt16-1* mutants to the same extend as shown by the transcript profiling approach (**Figure 3.5**). E.g. the transcript level of *CHS* in the *ssrp1-2* mutant was determined as 2.6-fold and 2.8-fold downregulated by the microarray

analysis and the qRT-PCR, respectively. In the *spt16-1* mutant, the transcript level of *CHS* was determined as 2.7-fold (microarray analysis) and 3.0-fold (qRT-PCR) downregulated, respectively. The high correlation between the microarray and the qRT-PCR results validated the data obtained by the genome-wide transcript profiling and suggested a role for SSRP1/SPT16 in the biosynthesis of anthocyanins.

In addition to the anthocyanin biosynthesis genes, all four members of a basic Helix-Loop-Helix (bHLH) transcription factor family that plays an important role in the iron homeostasis of *Arabidopsis* plants (Yuan et al. 2008, Wang et al. 2013) were among the 23 down-regulated genes in *ssrp1-2* and *spt16-1*. All four genes (*bHLH038*, *bHLH039*, *bHLH100*, and *bHLH101*) showed a very strong downregulation in both FACT mutants, ranging from 3.6- to 9.7-fold in *ssrp1-2* and from 3.0- to 7.7-fold in *spt16-1* as shown in the discussion (**Figure 7.1**).

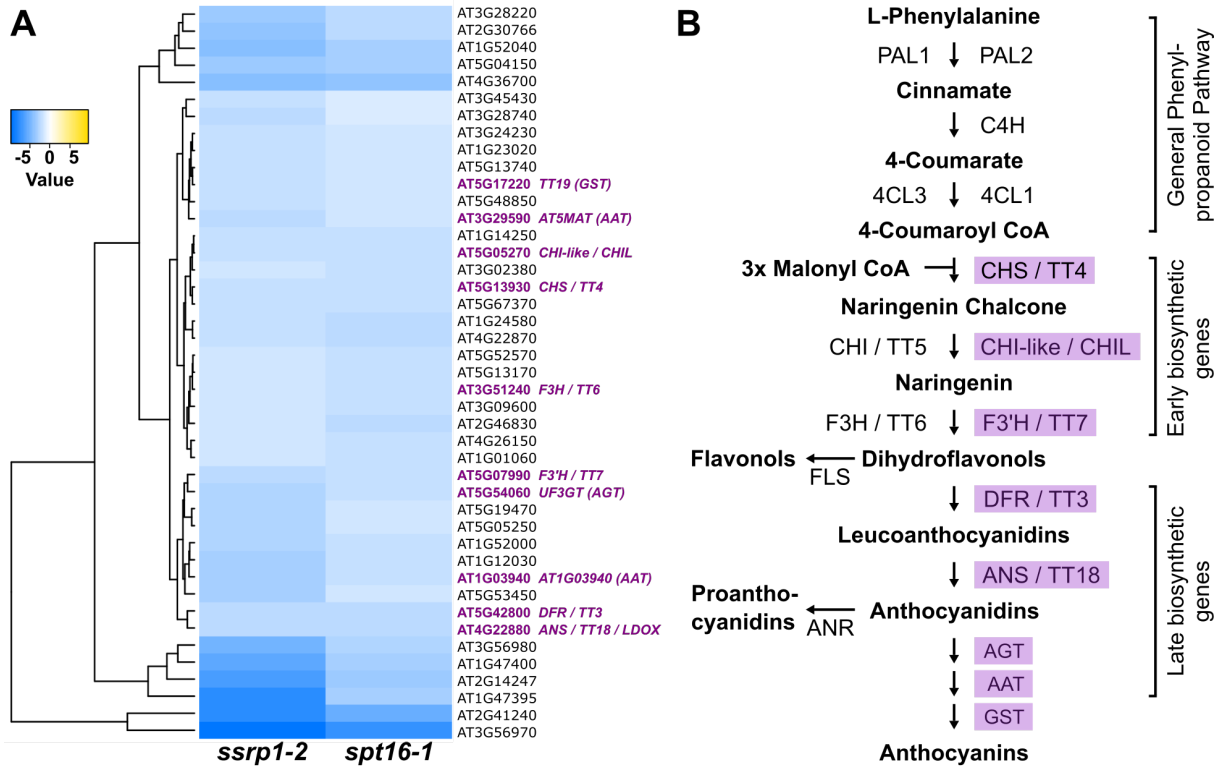


Figure 3.3 Early and late anthocyanin biosynthetic genes were strongly downregulated in the FACT mutants. A) Genes downregulated (≥2-fold;P<0.05) in *ssrp1-2* and/or *spt16-1* mutants were hierarchically clustered and visualized as heatmap. Different shades of blue indicate the reduced expression level of the 43 genes. The corresponding colour calibration bar is shown. The anthocyanin biosynthetic genes were depicted in bold purple. B) The schematic illustration of the anthocyanin biosynthetic pathway was modified according to Pérez-García et al. 2015. The genes with a significantly (p<0.05) reduced (≥2-fold) expression in *ssrp1-2* and *spt16-1* mutants were depicted in purple.

3.3 FACT-depleted plants showed defects in the light-induced accumulation of anthocyanin

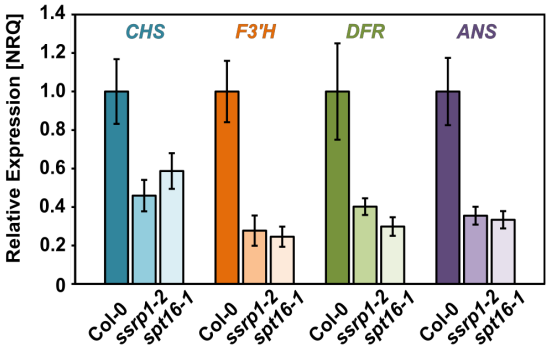
Anthocyanins are a group of plant pigments with several important physiological and ecological functions, including the protection from high light irradiation (Shi et al. 2014). The accumulation of anthocyanin in plants can be triggered by many biotic and abiotic factors including high light (HL) and sucrose (Saito et al. 2013, Shi et al. 2014). In response to HL stress,

	AGI	Gene	*ssrp1*	p-Value	*spt16*	p-Value
Regulatory Genes	AT5G35550	*TT2*	-1.1615	0.3590	-1.1039	0.6137
	AT4G09820	*TT8*	-1.3636	0.0079	-1.2532	0.0706
	AT5G24520	*TTG1*	-1.1111	0.2690	1.0015	0.9397
	AT1G56650	*PAP1*	-1.4874	0.1388	-1.4237	0.0964
	AT1G66390	*PAP2*	-1.2955	0.4142	-1.0111	0.8853
	AT5G41315	*GL3*	-1.1717	0.1660	-1.0395	0.7110
	AT1G63650	*EGL3*	-1.0353	0.5879	-1.0936	0.2203
Biosynthetic Genes	AT3G53260	*PAL2*	-1.3475	0.0029	-1.2394	0.0053
	AT2G37040	*PAL1*	-1.4495	0.0228	-1.4059	0.0402
	AT2G30490	*C4H*	-1.1738	0.0025	-1.1948	0.1139
	AT1G65060	*4CL3*	-1.1626	0.1562	-1.1922	0.2017
	AT1G51680	*4CL1*	-1.4532	0.0023	-1.2265	0.0139
	AT5G13930	*CHS / TT4*	-2.0757	0.0004	-2.0379	0.0012
	AT5G05270	*CHI-like / CHIL*	-2.0115	2.17E-05	-2.1126	0.0002
	AT3G55120	*CHI / TT5*	-1.2771	0.0014	-1.4191	0.0004
	AT5G07990	*F3'H / TT7*	-2.5929	0.0026	-2.0540	0.0045
	AT3G51240	*F3H / TT6*	-1.7723	0.0003	-2.2395	0.0001
	AT5G42800	*DFR / TT3*	-2.5563	0.0009	-2.4136	0.0004
	AT4G22880	*ANS / TT18 / LDOX*	-2.6219	0.0001	-2.6903	0.0002
	AT5G08640	*FLS*	-1.0437	0.4653	-1.1431	0.1805
	AT1G61720	*ANR*	-1.0703	0.3739	-1.0421	0.7897
	AT5G17050	*UGT78D2 (AGT)*	-1.1384	0.4498	-1.1560	0.3924
	AT4G14090	*AT4G14090 (AGT)*	-1.6921	0.0214	-1.6787	0.0062
	AT5G54060	*UF3GT (AGT)*	-2.8032	0.0044	-2.2328	0.0011
	AT3G21560	*UGT84A2 (AGT)*	-1.4610	0.0055	-1.2696	0.0089
	AT2G29750	*UGT71C1 (AGT)*	1.1477	0.4314	-1.1443	0.4671
	AT4G01070	*GT72B1 (AGT)*	-1.0737	0.3326	-1.0484	0.2875
	AT1G03940	*AT1G03940 (AAT)*	-3.1166	0.0009	-2.0378	0.0010
	AT3G29590	*AT5MAT (AAT)*	-2.4382	0.0001	-1.7776	0.0022
	AT1G03495	*AT1G03495 (AAT)*	-1.8482	0.0045	-1.8628	0.0005
Transport Genes	AT5G17220	*TT19 (GST)*	-2.2878	0.0009	-1.7738	0.0019

Figure 3.4 Anthocyanin biosynthetic genes were downregulated in *ssrp1-2* and *spt16-1* mutants, whereas regulatory genes were not affected.
Regulatory, biosynthetic and transport genes of the anthocyanin biosynthesis pathway were shown together with the microarray expression data of the FACT mutants in comparison to *wild type* plants. The genes that were downregulated more than 1.5-fold / 2.0-fold were depicted in light/dark blue, respectively. The non-significant values ($p>0.05$) were depicted in red.

wild type plants accumulate anthocyanin in the vegetative tissues and their leaves turn purple (Hatier et al. 2008, Kovinich et al. 2014). Thereby, the plant pigments function most likely as antioxidants and protect the plant from reactive oxygen species (ROS) that are generated during photosynthesis under conditions of photoinhibition (typically under HL conditions)(Albert et al. 2009, Kovinich et al. 2014). Mutations in most of the regulatory or biosynthesis genes of the anthocyanin pathway result in an impaired anthocyanin synthesis. In response to HL stress, these mutants show defects in the anthocyanin accumulation and their leaves turn yellow, which is called a senescence phenotype (Xu et al. 2017b). The transcript profiling analysis in this study revealed that FACT was important for the expression of anthocyanin biosynthesis pathway genes. This raised the question if SSRP1/SPT16-depleted plants show defects in the light-stimulated induction of anthocyanin production. To visualize the putative defects in the accumulation of anthocyanin, the FACT mutants were grown in two different experimental setups under high light (HL) stress conditions. In the first setup, the plants were grown on 0.5x MS plates containing 2 % sucrose for 14 days under moderate high light (160 μmol $^{-2}$ s^{-1})

A



B

AGI	Gene	*ssrp1-2*		*spt16-1*	
		fold change microarray	fold change qRT-PCR	fold change microarray	fold change qRT-PCR
AT5G13930	CHS	-2.622	-2.818	-2.690	-2.998
AT5G07990	F'3H	-2.593	-3.608	-2.054	-4.072
AT5G42800	DFR	-2.076	-2.178	-2.038	-1.704
AT4G22880	ANS	-2.556	-2.488	-2.414	-3.351

Figure 3.5 The microarray dataset was validated by qRT-PCR.
The relative expression of *CHS*, *F3'H*, *DFR* and *ANS* was determined in *wild type* Col-0 and *ssrp1-2* as well as *spt16-1* mutants by qRT-PCR. . The three genotypes were grown on 0.5x MS plates at 100 μmol m^{-2} s^{-1}. The normalized relative quantities (NRQ) $\pm$ the normalized relative standard errors (SE[NRQ]) are shown. The relative quantity of the analysed mRNA was normalized to the relative quantities of *GAPC*, *PP2AA3* and *UBI10* mRNA.

and for comparison also under low light conditions (60 μmol $^{-2}$ s^{-1}) (**Figure 3.6 A**). In the second setup, the seedlings were grown on 0.5x MS plates containing 1 % sucrose for 14 days under low light conditions (60 μmol m^{-2} s^{-1}) and subsequently for 3 days under strong high light conditions (600 μmol $^{-2}$ s^{-1}). The control plants were kept under the low light conditions (**Figure 3.6 B**). Under both low light conditions, Col-0 plants as well as both FACT mutants showed no obvious accumulation of anthocyanin and their leaves appeared green (**Figure 3.6 D,F**). Under both HL stress conditions (moderate and strong), the *wild type* plants showed a clear accumulation of anthocyanin as their leaves turned purple. In contrast, *ssrp1-2* and *spt16-1* mutants showed defects in the light-induced accumulation of anthocyanin. Relative to Col-0, the leaves of SSRP1/SPT16-depleted plants showed less purple coloring in response to both HL stresses. In order to quantify these observations, the anthocyanin content of the stressed versus non-stressed plants was determined photometrically in three (strong high light) and five (moderate high light) biological replicates (**Figure 3.6 C,E**). Col-0 as well as FACT mutants that were grown under low light conditions showed a similar low anthocyanin content. Upon moderate and strong HL stress, the anthocyanin content in *wild type* leaves was strongly increased. In contrast, the anthocyanin accumulation in reponse to both HL stresses was significantly reduced in *ssrp1-2* and *spt16-1* mutants (Moderate HL: 50-51% of the Col-0 content; Strong HL: 22-43% of the Col-0 content). These quantitative measurements confirmed the phenotypic observations that the FACT-depleted plants showed defects in the light-induced accumulation of anthocyanin and suggested a role for SSRP1 and SPT16 in the anthocyanin synthesis in response to light stimulus.

3.4 FACT was required for light-induced upregulation of anthocyanin biosynthesis genes

In response to HL stress, both FACT mutants showed an anthocyanin-deficiency phenotype. This raised the question if the light-induced upregulation of anthocyanin biosynthesis genes is impaired in *ssrp1-2* and *spt16-1* mutants, which would explain the observed absent accumulation of purple pigments in these mutants. Therefore, qRT-PCR analysis was performed to

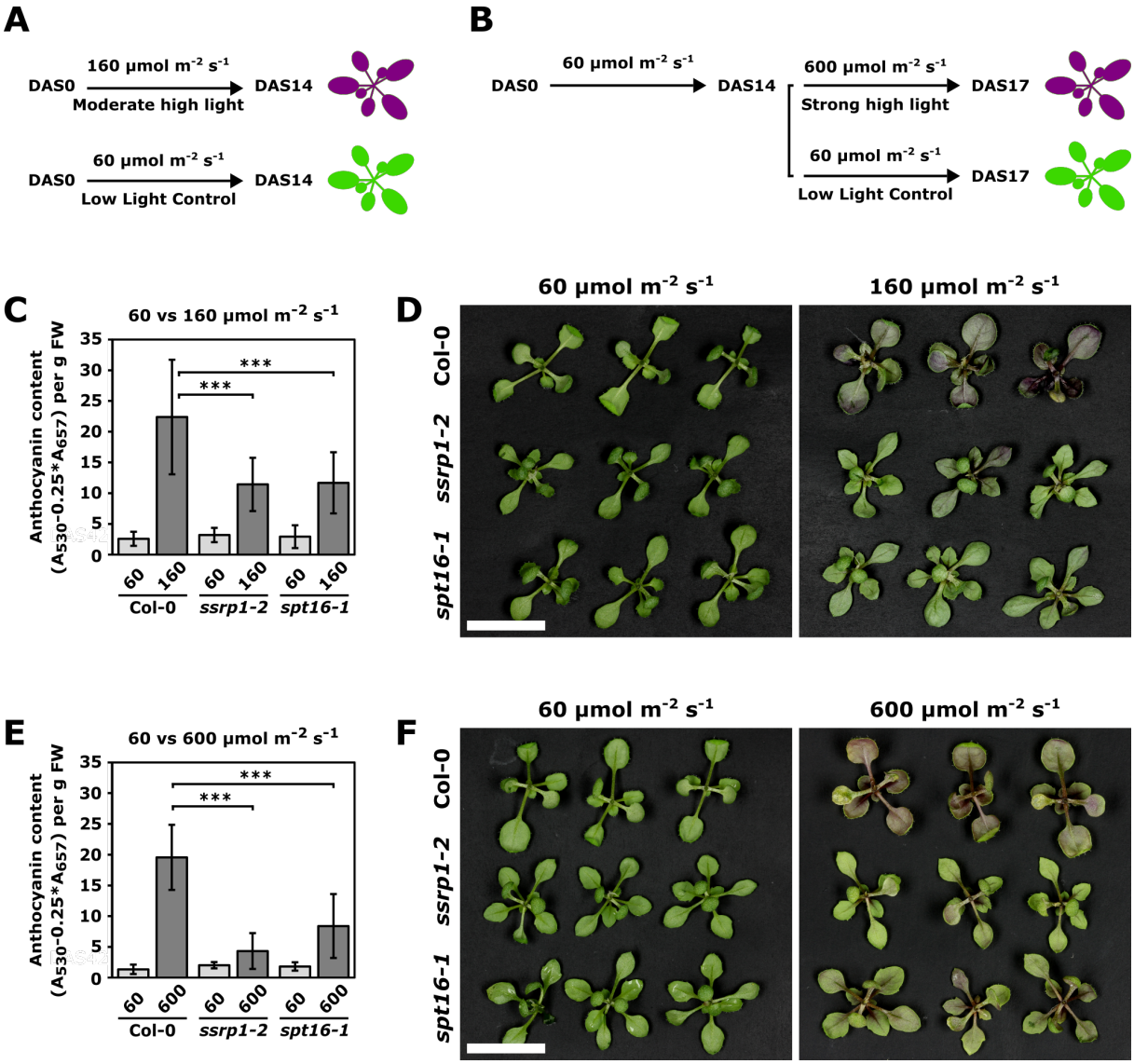


Figure 3.6 FACT mutants showed an anthocyanin-deficiency phenotype in response to high light stress. A) Schematic illustration of the plant growth setup to test for defects in the anthocyanin accumulation upon moderate high light stress. The seedlings of the three genotypes were grown on 0.5x MS plates supplemented with 1% sucrose for 14 days either under moderate high light (160 μmol m^{-2} s^{-1}) or as control plants under low light (60 μmol m^{-2} s^{-1}). B) Schematic illustration of the plant growth setup to test for defects in the anthocyanin accumulation upon strong high light stress. The seedlings of the three genotypes were grown on 0.5x MS plates supplemented with 2% sucrose for 14 days under low light (60 μmol m^{-2} s^{-1}). Subsequently, plants were either shifted to strong high light (600 μmol m^{-2} s^{-1}) for 3 days or kept the same time under low light as controls. C) The anthocyanin content in moderate HL-stressed plants and non-stressed control plants was measured photometrically in three biological replicates (n = 13). The light grey bars indicate the control conditions, whereas the dark grey bars indicate the HL-treatment. All data are means ± SD, the significance was tested by Student's T-Test, *** indicates $p < 0.001$. D) Phenotypical analysis of moderate high light stress plants versus non-stressed control plants. Three representative plants of each genotype were shown upside down (abaxial) to visualize the anthocyanin accumulation in the leaves. Scale bar indicates 1 cm. E) The anthocyanin content in strong HL-stressed plants and non-stressed control plants was measured photometrically in five biological replicates (n=29). The light grey bars indicate the control conditions, whereas the dark grey bars indicate the HL-treatment. All data are means ± SD, the significance was tested by Student's T-Test, *** indicates $p < 0.001$. F) Phenotypical analysis of strong high light stress plants versus non-stressed control plants. Three representative plants of each genotype are shown upside down (abaxial) to visualize anthocyanin accumulation in the leaves. Scale bar indicates 1 cm.

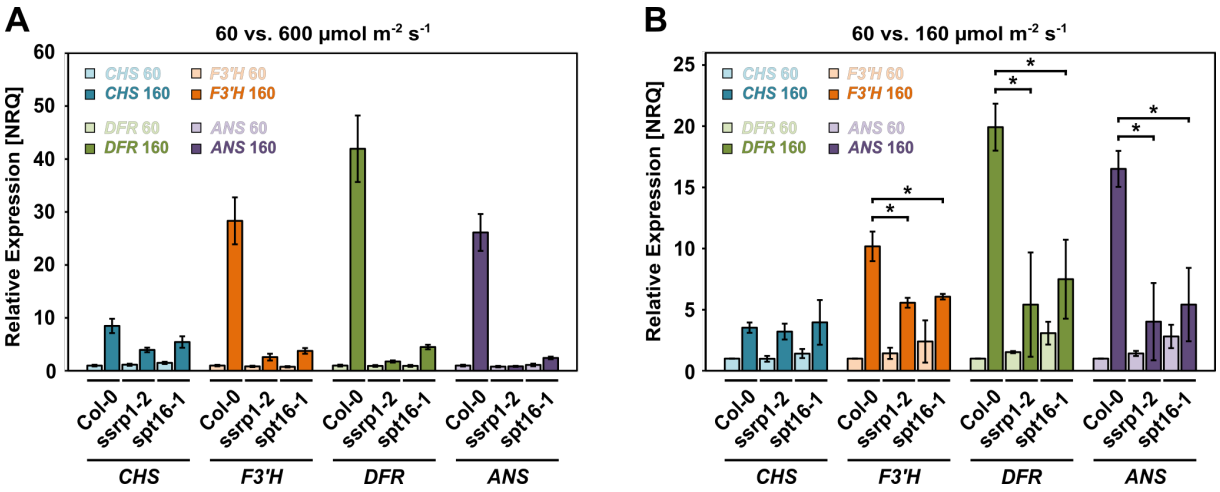


Figure 3.7 The HL-induced expression of anthocyanin biosynthetic genes was impaired in both FACT mutants. In *wild type* plants and *ssrp1-2* as well as *spt16-1* mutants, the relative transcript levels of early (*CHS*, *F3′H*) and late (*DFR*, *ANS*) anthocyanin biosynthetic pathway genes were examined in response to (A) strong and (B) moderate high light by qRT-PCR. (A) The normalized relative quantities (NRQ) ± the normalized relative standard errors (SE[NRQ]) of one biological replicate are shown. (B) The mean normalized relative quantities (NRQ) ± the standard deviation of two biological replicate are shown. (A-B) The relative quantities of *CHS*, *F3′H*, *DFR* and *ANS* mRNA were normalized to the relative quantities of *GAPC*, *PP2AA3* and *UBI10* mRNA. Relative to *wild type* Col-0, the light/dark colours indicate the expression of the specific gene under low/high light conditions, respectively.

examine the changes in the expression levels of two early (*CHS*, *F3′H*) and two late (*DFR*, *ANS*) anthocyanin biosynthetic genes in response to HL stress in FACT mutants compared to *wild type* plants (**Figure 3.7**). To induce anthocyanin synthesis, strong and moderate high light were used as described before in section 3.3. To quantify the light-induced upregulation of anthocyanin biosynthsis genes, total RNA was isolated from Col-0 and *ssrp1-2* as well as *spt16-1* mutants that were grown under control and high light conditions. Three reference genes (*GAPC*, *PP2AA3* and *UBI10*) were used for qRT-PCR data normalization (Hellemans et al. 2007). Upon moderate and strong high light stress, the *ssrp1-2* and *spt16-1* mutants showed a clearly reduced induction of *F3′H*, *DFR* and *ANS* expression in comparison to Col-0. Relative to the other three tested genes, the light-induced upregulation of *CHS* was less strongly and to a similar extent in all three tested genotypes in both experimental setups. These results confirmed the previously characterized light-induced anthocyanin-deficiency phenotype of FACT mutants on a molecular level. The HL-induced upregulation of the anthocyanin biosynthetic genes was strongly impaired in the FACT mutants.

3.5 *SSRP1* and *SPT16* gene expression was strongly upregulated in response to high light stress

In response to high light, the accumulation of anthocyanin and the up-regulation of anthocyanin biosynthesis genes were impaired in *ssrp1-2* and *spt16-1* mutants. This suggested that *SSRP1* and *SPT16* are important factors for the light-induced anthocyanin production. If this is the case, possibly the expression of both FACT subunits is as well upregulated by the high light treatment. To test this hypothesis the expression levels of *SSRP1* and *SPT16* upon HL stress were examined by GUS-Reporter lines. GUS reporter lines for both FACT subunits were generated by *Agrobacterium*-mediated transformation of *Arabidopsis* Col-0 plants with trans-

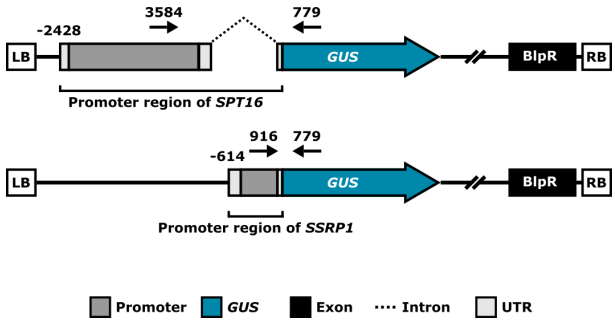


Figure 3.8 Schematic illustration of the *SPT16*- and *SSRP1-GUS* reporter transgene constructs.
Expression of the beta-glucuronidase gene (*GUS*) is driven by the promoters of either *SSRP1* or *SPT16*. As putative *SPT16* and *SSRP1* promoters, the regions 2428-nt and 614-nt upstream of the translational start sites were selected, respectively (Black bars = exons, blue bar = β-glucuronidase (*GUS*) gene, dotted lines = introns, dark grey bars = promoter, light grey bars = UTR, BlpR = Basta resistance marker).

genes driving the expression of beta-glucuronidase (*GUS*) under the promoters of either *SSRP1* or *SPT16* (**Figure 3.8**), the regions 614-nt (*SSRP1*) and 2428-nt (*SPT16*) upstream of theirs translational start sites. Two representative lines were selected to study the gene expression of *SSRP1* and *SPT16* in response to high light stress histochemically by GUS staining (**Figure 3.9**). The transgenic GUS lines were stressed with either strong or moderate high light, while the control plants were kept under low light conditions. In both experimental setups, the non-stressed control plants revealed that the promoters of *SSRP1* and *SPT16* were widely active in the aerial parts under low light conditions, especially in the leave hydathodes. Strikingly, the expression patterns of both FACT subunits were highly overlapping. A strong increase in the *SSRP1* and *SPT16* promoter activity was observed upon stressing the transgenic GUS-lines with moderate or strong high light. This indicated that the expression of both FACT subunits was highly upregulated in response to high light treatment. In general, the SSRP1- and SPT16-GUS lines that were stressed with moderate high light for a long period, showed more indigo blue staining, than the GUS lines stressed with strong high light for a shorter period.

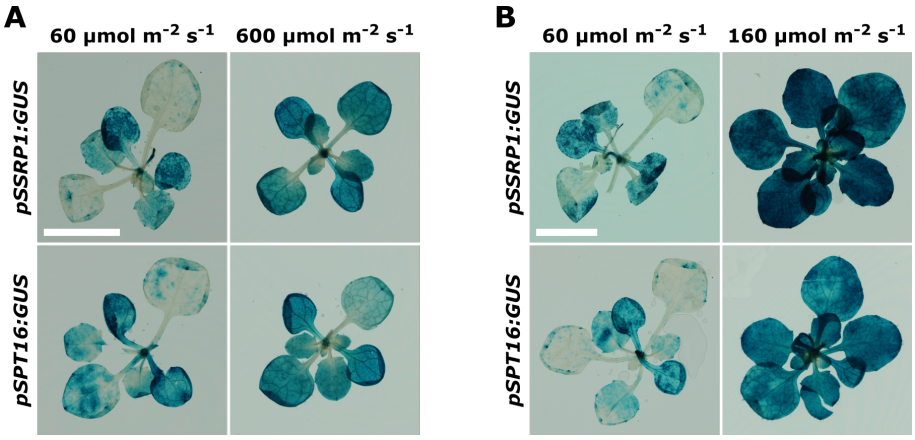


Figure 3.9 The ***SSRP1* and *SPT16* promoter activities were strongly increased upon high light stress.**
Histochemical GUS staining of transgenic *SSRP1* and *SPT16* GUS-Reporter lines. A) Transgenic plants were grown under normal (60 μmol m^{-2} s^{-1}) and strong (600 μmol m^{-2} s^{-1}) high light conditions B) Transgenic plants were grown under normal (60 μmol m^{-2} s^{-1}) and moderate (160 μmol m^{-2} s^{-1}) high light conditions. Scale bar indicates 1 cm.

Additional to the promoter activities of both FACT genes by GUS staining, the transcript levels of *SSRP1* and *SPT16* were analysed by qRT-PCR in response to high light stress. *Wild*

type Arabidopsis plants were stressed with either strong or moderate high light, while control plants were kept under low light conditions. Total RNA was isolated from stressed and non-stressed plants and transcribed into cDNA as template for qRT-PCR using *SSRP1* (4044/4045) and *SPT16* (1785/1786) specific primers. In both experimental setups, *SSRP1* and *SPT16* were 1.2- and 1.5-fold upregulated upon HL stress, respectively (**Figure 3.10**).

Taken together, the expression of *SSRP1* and *SPT16* in *Arabidopsis* was upregulated in a concerted manner in response to high light treatment.

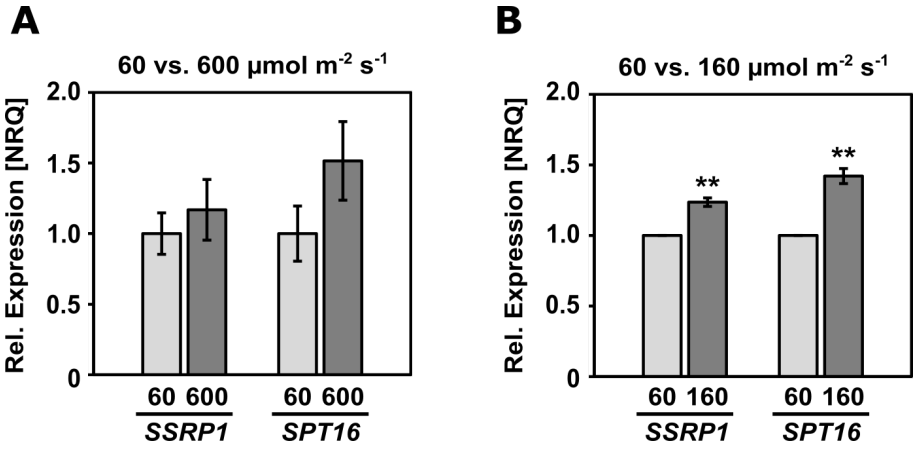


Figure 3.10 The transcript levels of *SSRP1* and *SPT16* were upregulated in response to high light stress. The changes in the relative expression of *SSRP1* and *SPT16* upon high light stress were examined by qRT-PCR using *SSRP1* (4044/4045) and *SPT16* (1785/1786) specific primers. A) *Wild type* plants were grown under normal (60 μmol m^{-2} s^{-1}) and strong (600 μmol m^{-2} s^{-1}) high light conditions. The normalized relative quantities (NRQ) $\pm$ normalized relative standard error (SE[NRQ]) of one biological replicate are shown. B) *Wild type* plants were grown under normal (60 μmol m^{-2} s^{-1}) and moderate (160 μmol m^{-2} s^{-1}) high light conditions. The normalized relative quantities (NRQ) $\pm$ standard deviation of two biological replicates are shown. (A-B) The relative quantities of *SSRP1* and *SPT16* mRNA were normalized to the relative quantity of *GAPC, PP2AA3* and *UBI10* mRNA. The significance was tested by Student's T-Test. ** indicates $p < 0.01$.

3.6 Proanthocyanin synthesis was not impaired in seeds of FACT mutants

As shown before, *ssrp1-2* and *spt16-1* mutants showed a reduced expression of genes of the core flavonoid biosynthesis pathway. Next to anthocyanin, the other two phenolic compounds flavonol glycosides and proanthocyanin (PA) are produced by this biosynthesis pathway (Saito et al. 2013). PA accumulates in the seed coat and appears brown in its oxidized polymeric form. Therefore, the mature seeds of most knockout mutants of flavonoid biosynthesis genes show no brown coloration due to lack of PA production (Lepiniec et al. 2006, Appelhagen et al. 2014, Shi et al. 2014). This raised the question if the accumulation of PAs is affected in the FACT mutants. First, the seed colour of *ssrp1-2* and *spt16-1* mutants looked like the one of *wild type* seeds. The brown pigmentation indicated that PAs were accumulated in the seed coat of FACT mutants. Additionally, seeds of *wild type* and FACT mutants were stained with PA-specific DMACA and acidic vanillin to detect smaller changes in the pigmentation. Neither DMACA- (data not shown) nor vanillin-staining showed differences between the seeds of *wild type* and *ssrp1-2* as well as *spt16-1* mutants. Following vanillin staining, PA in the innermost cell layer of the seed coat (endothelium), the chalaza and the micropyle appeared red. In these cells, PA is stored in the vacuoles (Lepiniec et al. 2006) as seen in **Figure 3.11**. Worth to mention is

that the seed coats of *ssrp1-2* and especially *spt16-1* mutants showed general structural defects as depicted in **Figure 3.11**, but these were unlikely related to defects in flavonoid biosynthesis. In summary, the FACT mutants showed no defect in PA accumulation in seeds.

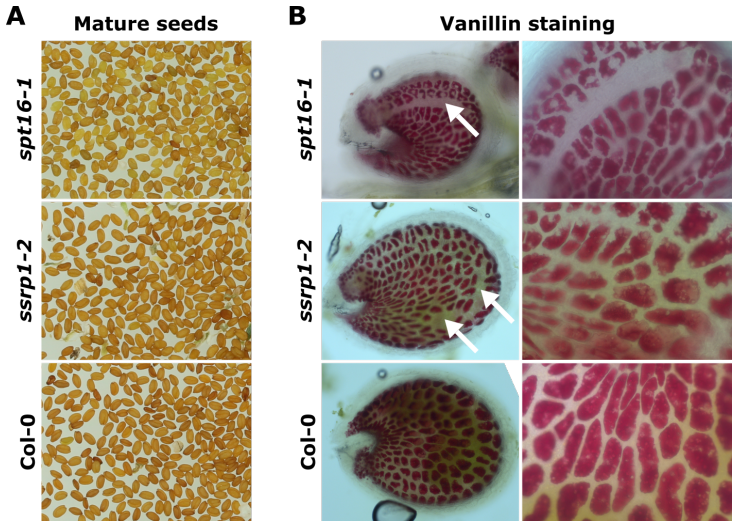


Figure 3.11 Proanthocyanin accumulation was not impaired in the seed coat of *ssrp1-2* and *spt16-1* mutants. A) Seeds of FACT mutants showed the same brown pigmentation as *wild type* seeds. Structural defects in the seed coat are indicated by white arrows. B) Whole mount staining of seeds from *wild type* plants and FACT mutants with acidic vanillin. PAs were present in the *wild type* and mutant endothelium.

4 Results: The composition of the *Arabidopsis* transcript elongation complex

In the context of this thesis, the characterization of the *Arabidopsis* RNAPII TEC represents rather a side aspect and this work was published recently in Plant Cell (Antosz et al. 2017). Therefore, this aspect is only briefly summarized here. Alexander Pfab contributed to this work by the affinity purification coupled to mass spectrometry (AP-MS) analysis of SG-tagged TEFs including the molecular cloning, the generation of transgenic PSB-D lines, the affinity purification, the sample preparation for LC-MS/MS and the bioinformatic analysis of the MS data. Moreover, the phenotypic analysis of the TEF double mutants was performed by A.P. and Wojciech Antosz.

The components of the *Arabidopsis* RNAPII elongation complex were experimentally identified by AP-MS analysis. Both FACT subunits (SSRP1 and SPT16), two PAF1-C components (ELF7 and CDC73), the SPT4/5 subunit SPT4-2, TFIIS, the CDKC;2 component of P-TEFb and the largest subunit (NRPB1) of the RNAPII were selected for this approach. The SG-tagged (Streptavidin binding peptide and 2x protein G domains) bait proteins were expressed in transgenic PSB-D suspension cultured cells (ecotype Landsberg erecta) and the GS fusion proteins as well as the unfused GS control were affinity purified from the cell extracts using IgG-coupled magnetic beads (Van Leene et al. 2015). The co-purified prey proteins were identified after tryptic digestion by mass spectrometry. The proteomic analyses revealed that the *Arabidopsis* transcript elongation factors FACT, PAF1-C, TFIIS and SPT4/5, but not P-TEFb frequently co-purified with each other and the elongating RNAPII. Moreover, TEFs like the histone chaperone SPT6L (Gu et al. 2012) and the Elongator with its histone acetyl-transferase activity (Woloszynska et al. 2016) were co-purified with the TEFs. Additional factors like the members of the NAP1 histone chaperone family, chromatin remodelling complexes (CRCs) and histone deacetylases (HDACs) were repeatedly co-purified with the TEC and may contribute to efficient transcript elongation. Next to the composition of the *Arabidopsis* RNAPII TEC, the reciprocal tagging approach demonstrated an extensive interplay between the TEFs and mRNA processing factors. Amongst others, many components of splicing complexes including the U1, U2, U5, Sm and NTC were co-purified with the examined TEFs, what suggests that different assembly stages of the spliceosome interact with the TEC. This suggests that the RNAPII TEC represents a platform for different TEFs as well as mRNA processing factors to coordinate efficient transcript elongation with simultaneous mRNA maturation.

TEF-depleted *Arabidopsis* plants show a great variety of growth and development defects that are ranging from mild to severe/lethal phenotypes, like for instance in the case of TFIIS (Grasser et al. 2009) and SPT5-2 (Dürr et al. 2014), respectively. *Arabidopsis* double mutants that were deficient in different combinations of TEFs were phenotypically analysed along with the respective single mutants and Col-0 *wild type* plants. Consequences on plant growth and development as well as possible genetic interactions between the genes encoding different TEFs were examined for the following TEF combinations: FACT/TFIIS (SSRP1/TFIIS, SPT16/TFIIS), FACT/PAF1-C (SSRP1/ELF7, SPT16/ELF7), and TFIIS/PAF1-C (TFIIS/ELF7). *Arabidopsis* mutants with a reduced expression of one of the FACT subunits (SSRP1 or SPT16) show multiple similar growth and developmental defects (Lolas et al. 2010), whereas TFIIS-depleted

plants have *wild type* appearance (Grasser et al. 2009). The leave vein patterning of TFIIS mutants looks like in *wild type* Col-0 plants, whereas the leave venation of SSRP1- and SPT16-depleted plants displays mild and strong alterations, respectively (Lolas et al. 2010). Regarding rosette diameter and number of primary inflorescences, SSRP1/TFIIS double mutants were synergistically affected causing for example plant sterility, while SPT16/TFIIS mutants looked like SPT16 single mutants. Accordant to these phenotypes, TFIIS/SSRP1 double mutants are synergistically affected, while SPT16 acts epistatically to TFIIS in SPT16/TFIIS mutants. However, regarding bolting time and leave venation, both FACT/TFIIS double mutants showed the phenotype of the respective FACT single mutant. The analysis of the interaction between FACT and PAF1-C revealed that SSRP1/ELF7 double mutants are lethal, while SPT16/ELF double mutants are viable, but show a strongly reduced growth. In accordance with these findings, there are several lines of evidence that suggest a close cooperation of FACT and PAF1-C in plants, possibly involving transcription-related histone ubiquitination (Adelman et al. 2006, Pavri et al. 2006, Squazzo et al. 2002, Lolas et al. 2010). SSRP1- and SPT16-depleted plants are phenotypically similar (Lolas et al. 2010), but in combination with mutations in TFIIS or ELF7, the SSRP1 version of the double mutant is more severely affected than the SPT16 version. This suggests that the FACT subunit SSRP1 is more critical for plant growth and development. Regarding most of the investigated plant characteristics, TFIIS/ELF7 double mutants were synergistically affected what accumulated in plant sterility. Therefore, the analysis of different TEF double mutants revealed distinct genetic interactions between the genes encoding the transcript elongation factors FACT, PAF1-C and TFIIS. In the future, further analyses are required to make use of all the opportunities that are provided by the generated plant lines.

5 Results: The *Arabidopsis* SAGA-DUBm component ENY2

5.1 *ENY2* gene expression in *Arabidopsis thaliana*

To study *ENY2* (At3G27100) in *Arabidopsis thaliana*, the GUS (β-glucuronidase) Reporter system was used. For this purpose, an *ENY2-GUS* reporter transgene was generated (**Figure 5.1**). This construct is controlling the expression of the *GUS* gene by the putative native *ENY2* promoter, the region 1970-nt upstream of the *ENY2* transcription start site. The *ENY2-GUS* reporter construct was integrated into *wild type* Col-0 plants by *Agrobacterium*-mediated transformation. Three independent transgenic lines (#4, #11, #19) were identified by resistance to the herbicide Basta® as well as by genotyping PCR. In the next generation, individuals homozygous for the *ENY2-GUS* T-DNA were obtained for all three independent lines. To analyse the authentic *ENY2* expression pattern, these three transgenic lines were examined by GUS staining. The *ENY2* promoter activity was thereby visualized histochemically and cells expressing *ENY2* were stained indigo blue. If not otherwise stated, all three independent *ENY2-GUS* reporter lines showed an overlapping expression pattern with just minor differences in the intensity but not in the localization. A complete comparison of all three independent *ENY2-GUS* lines is shown in the **Supplemental Figures S3, S4, S5**. In the following, the representative line #11 is shown.

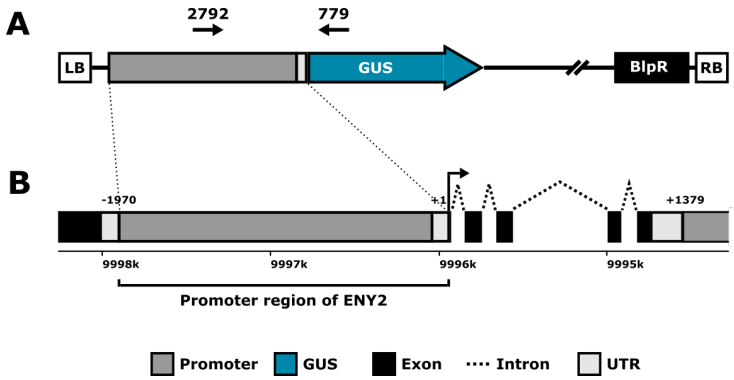


Figure 5.1 Schematic illustration of the *ENY2-GUS* reporter transgene.
A) The *ENY2* promoter is controlling the expression of the β-glucuronidase (*GUS*) gene. The primers used for genotyping PCR are depicted as arrows B) The genomic locus of *ENY2* (AT3G27100), including the *ENY2* promoter region (1970 nucleotides upstream of *ENY2* translational start site). Black bars = exons, BlpR = Basta® resistance marker, blue bar = β-glucuronidase (*GUS*) gene, dark grey bars = promoter, dotted lines = introns, light grey bars = UTRs, LB/RB = right/left border.

The GUS staining revealed that young seedlings (DAS3) show a wide expression of the reporter gene in the roots, the hypocotyls and the cotyledons (**Figures 5.2 A and S3**). Thereby, the vasculature of the roots and the cotyledons showed a particularly strong staining as well as the root-hypocotyl transition regions, whereas no GUS signal was detectable in the primary root tip and the shoot tip.
Taking a closer look at the meristematic regions of young seedlings (DAS7-10) revealed that *ENY2* expression could not be found in the shoot (**Figure 5.2 B,C**) and root apical meristems (**Figure 5.2 I**). Only the shoot and leaf primordia of line #4 showed a weak *ENY2* promoter activity (**Figure S3 J**). But line #4 is in general not completely in accordance with the other two lines and showed additional GUS signals in tissus (shoot meristem and reproductive organs) that were not stained in the other two lines.

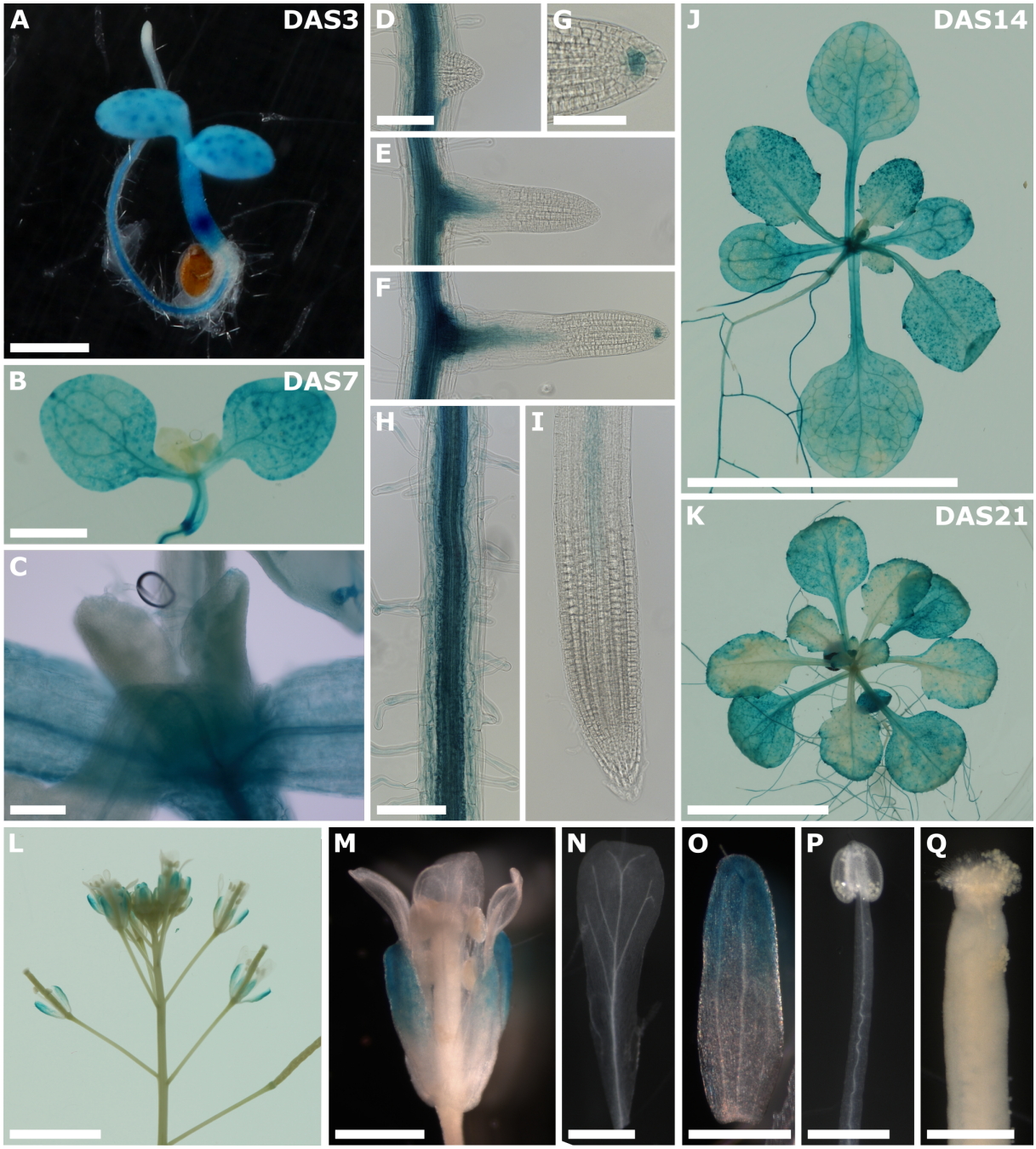


Figure 5.2 *ENY2* is widely expressed in *Arabidopsis*
The *ENY2* Promoter-GUS activity was monitored in three independent transgenic plant lines. The representative line #11 is shown here. A) Overview of young seedling (DAS3). Bar indicates 1 mm. B) Aerial parts of young seedling (DAS7). Bar indicates 1 mm. C) Close up of shoot apical meristem and young leaf primordia (DAS10). Bar indicates 100 μm. D-F) Three developmental stages of lateral root formation are shown (DAS7). Bar indicates 100 μm. G) Close up of lateral root tip. Bar indicates 50 μm. H-I) Differentiated (H) and meristematic (I) region of the primary root. Bar indicates 100 μm. J) Young plantlets at DAS14. Bar indicates 1 cm. K) Young plantlets at DAS21. Bar indicates 1 cm. L) Flower buds. Bar indicates 5 mm M) Flowers. Bar indicates 1 mm. N) Petals. Bar indicates 0.5 mm. O) Sepals. Bar indicates 0.5 mm. P) Stamen (Anther and filament). Bar indicates 0.5 mm. Q) Stigma and style of carpel. Bar indicates 0.5 mm.

The primary roots showed no *ENY2* promoter activity in the meristematic zone, but very high activity in the vasculature (**Figure 5.2 H,I**). A similar pattern was observable in the lateral roots. At early stages of the lateral root formation, no GUS signal was visible in the enlarging primordium and meristem (**Figure 5.2 D, E**). At later stages, the GUS signal became visible in the emerging vasculature and meristematic cells (**Figure 5.2 F, G**).
The investigation of later developmental stages (DAS14-21) showed that *ENY2* is widely expressed in cotyledons and leaves, especially in leave hydathodes (**Figures 5.2 J,K and S4**). As observed before in young seedlings, a strong GUS signal could be detected in the root vasculature, but not in the shoot and the primary root meristem. The GUS pattern looked the same in all three transgenic lines, but the line #4 showed the highest and the line #19 the lowest GUS signal intensity.
The GUS staining of reproductive tissues showed differences in the three independent ENY2-GUS lines (**Figures 5.2 L-Q and S5**). In the lines #11 and #19, the ENY2 promoter activity was restricted to the petals. In contrast, the line #4 showed additionally GUS signals in sepals, anthers and the stigma.
These results showed that the *ENY2* promoter is widely active in *Arabidopsis* plants, especially strong in the plant vasuclature and less prominent in the meristematic regions.

5.2 Localization of eGFP-ENY2 fusion protein in *Arabidopsis thaliana*

To visualize ENY2 in living cells, *Arabidopsis* plant lines were generated that were stably expressing eGFP-ENY2 under its native promoter. The transgene consists of 3.35 kb genomic locus of *ENY2* with an *eGFP* sequences as N-terminal translational fusion (**Figure 5.3**). The construct was integrated into *wild type* Col-0 plants by *Agrobacterium*-mediated transformation and three independent lines that were homozygous for the eGFP-ENY2 T-DNA were identified by resistance to the antibiotic kanamycin as well as by genotyping PCR.

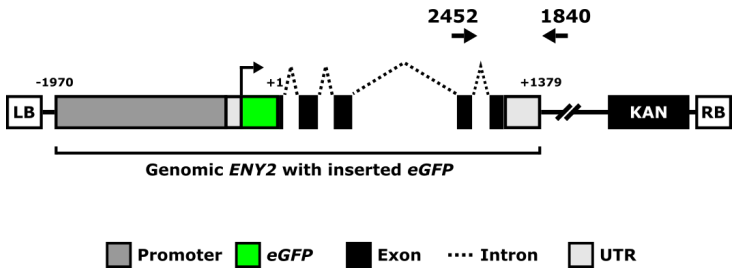


Figure 5.3 Schematic illustration of the *eGFP-ENY2* transgene construct.
The sequence of *eGFP* is inserted in the genomic locus of *ENY2* as N-terminal translational fusion (Black bars = exons, BlpR = Basta® resistance marker, dark grey bars = promoter, dotted lines = introns, green bar = *eGFP* coding sequence, light grey bars = UTRs, LB/RB = right/left border).

These three transgenic plant lines were used for the following experiments: First, the integrity of the eGFP-ENY2 fusion protein was verified by affinity purification coupled to immunoblotting (**Section 5.2.1**) and mass spectrometry. Second, the localization (**Section 5.2.2**) as well as the subcellular localization (**Section 5.2.3**) of eGFP-ENY2 in roots and root nuclei was analyzed by confocal laser scanning microscopy (CLSM). Third, the dynamics of eGFP-ENY2 were examined by fluorescence recovery after photobleaching (FRAP) experiments (**Section 5.2.4**).

Three independent plant lines expressing free eGFP-NLS under constitutive CaMV 35S

promoter were generated the same way as described above. According to CLSM analysis of root nuclei (data not shown) these lines showed a strong nuclear and almost no cytoplasmic fluorescence signal. Line #1 was identified as the most representative eGFP-NLS line and was used as control line for the following experiments.

None of these transgenic plants expressing either eGFP-ENY2 or eGFP-NLS showed any obvious phenotype (**Figure S6**), indicating that the overexpression of these fluoresent proteins had no influence on the plant development.

5.2.1 Verification of eGFP-ENY2 fusion protein integrity

The integrity of the eGFP fusion proteins was tested by GFP-Trap immunoprecipitation followed by Western Blotting (**Figure 5.4**). Total protein extracts of transgenic eGFP-ENY2 and eGFP-NLS (control) plants were used for this procedure. GFP fusion proteins were enriched by affinity purification using GFP-Trap® beads and visualized by subsequent Western Blotting using α-GFP AB. Both, eGFP-ENY2 and eGFP-NLS showed the electrophoretic mobility that was consistent with the predicted masses of 40.3 kDa and 28.1 kDa, respectively. Compared to eGFP-ENY2, the expression of eGFP-NLS was driven by the strong CaMV 35S promoter which explains most likely the more intense Western Blot signal. Furthermore, the identify of the eGFP-ENY2 fusion protein was confirmed by mass spectrometry (data not shown).

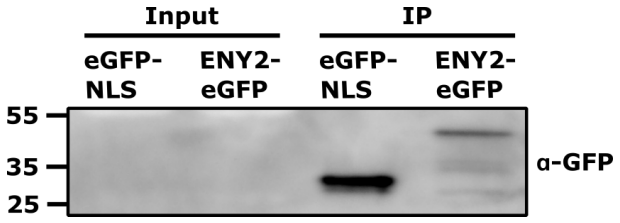


Figure 5.4 Detection of eGFP-ENY2 protein.
A) Immunoprecipitation of eGFP-ENY2 fusion protein using GFP-Trap® beads followed by Western Blot analysis. Total protein extracts of transgenic plants expressing eGFP-ENY2 and eGFP-NLS (control) were used for the IP. The immunoprecipitations were separated by SDS-PAGE and detected by Western Blotting using anti-GFP antibodies. The mass of the protein ladder is shown in kDa. The predicted masses of eGFP-NLS and eGFP-ENY2 are 28.1 kDa and 40.3 kDa, respectively.

5.2.2 In *Arabidopsis* roots, eGFP-ENY2 is localized in the nuclei

Both, eGFP-ENY2 and eGFP-NLS were visualized in *Arabidopsis* roots using Confocal Laser Scanning Microscopy (**Figure 5.5**). A prominent GFP signal was detectable in the root vasculature, but almost absent of the root meristem. In detail, the CLSM-analysis of the primary root tip (**Figure 5.5C**) revealed that line #36 showed a weak GFP signal in the stele near the Quiescent Center. Whereas the GFP signal in lines #33 and #34 became only later visible at more differentiated cells of the vasculature. In comparison, eGFP-NLS was equally expressed in all nuclei of the root meristem. This eGFP-ENY2 fluorescence pattern was similar to the *ENY2* promoter activity as shown before by GUS staining (**Figure 5.2**). Meaning that *ENY2* transcript and ENY2 protein share a similar spatio-temporal distribution in the root tissues of young seedlings.

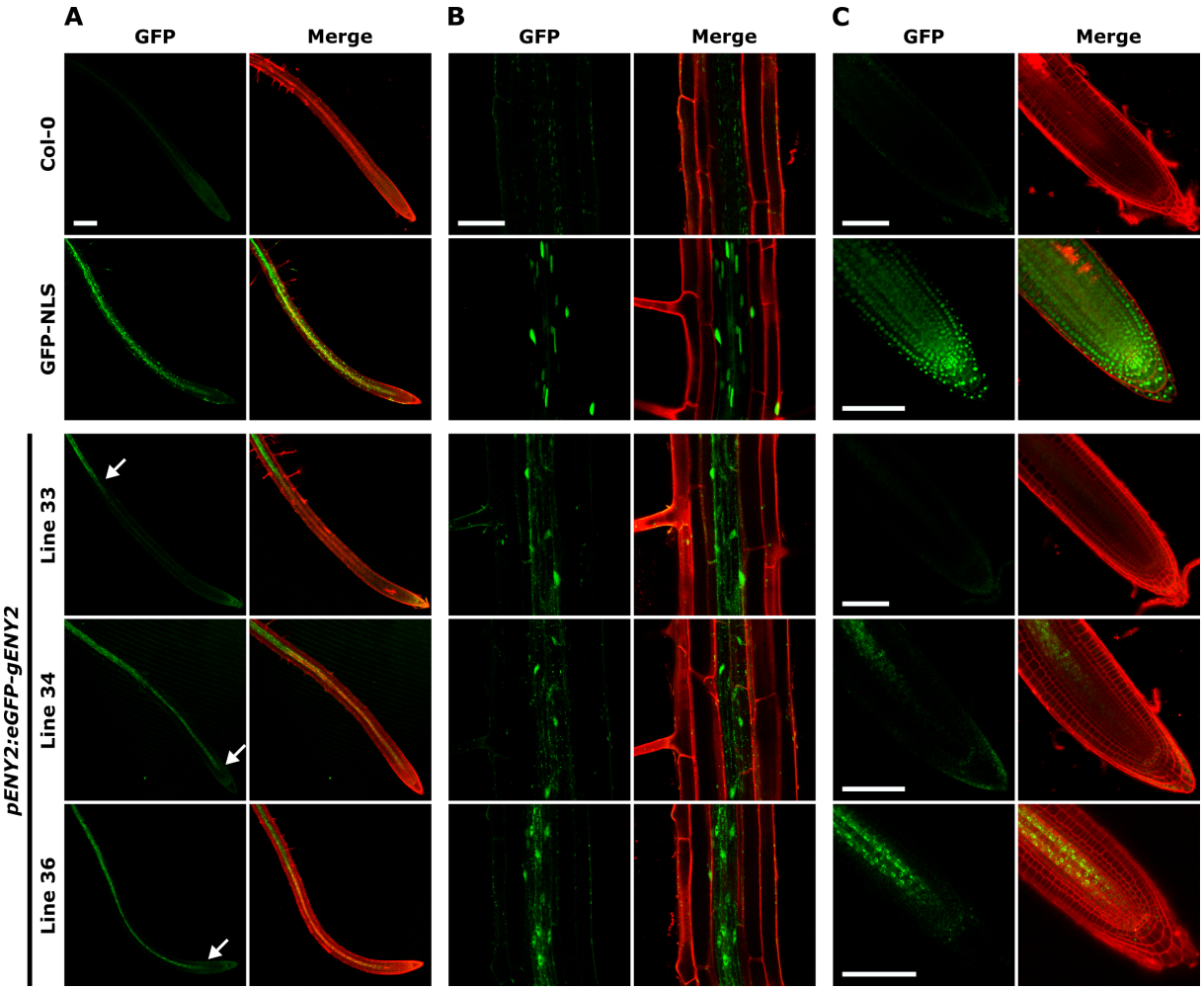


Figure 5.5 In *Arabidopsis* roots, eGFP-ENY2 was visualized in the vasculature.
Five days old plants of the three transgenic eGFP-ENY2 lines, the eGFP-NLS line (positive control) and *wild type* Col-0 (negative control) were subjected to confocal microscopy to investigate the eGFP-tagged proteins in *Arabidopsis* roots. Cell walls were stained with propidium iodide (PI; red signal). A) Overview of the *Arabidopsis* root spanning from the root tip to the differentiation zone. Bar indicates 200 μm. B) Close up of an optical-section through the differentiated root cells. Bar indicates 50 μm. B) Close up of optical-section through the root meristem. Bar indicates 100 μm.

5.2.3 In the nucleoplasm, eGFP-ENY2 is forming speckle-like structures

Optical sections through differentiated roots clearly showed that ENY2 was localized mainly nuclear (**Figure 5.5 B**). In comparison to eGFP-NLS-expressing and *wild type* plants, some cytoplasmic signal was detected. Moreover, GFP-expressing cells were more frequent in the root vasculature than in the outer cell layers of the root. To monitor the subcellular localization of eGFP-tagged ENY2 a higher magnification was used (**Figure 5.6**). This revealed that eGFP-ENY2 was forming speckle-like structures in the nucleoplasm and was absent of the nucleolus. Which was in contrast to eGFP-NLS-expressing plants that showed a homogenous distribution of the fluorescence signal throughout the nucleoplasm, with a slight enrichment in the nucleolus. As previously shown, the cytoplasmic fraction of eGFP-ENY2 was very low. All three independent transgenic eGFP-ENY2 lines (#33, #34, #36) confirmed these observations.

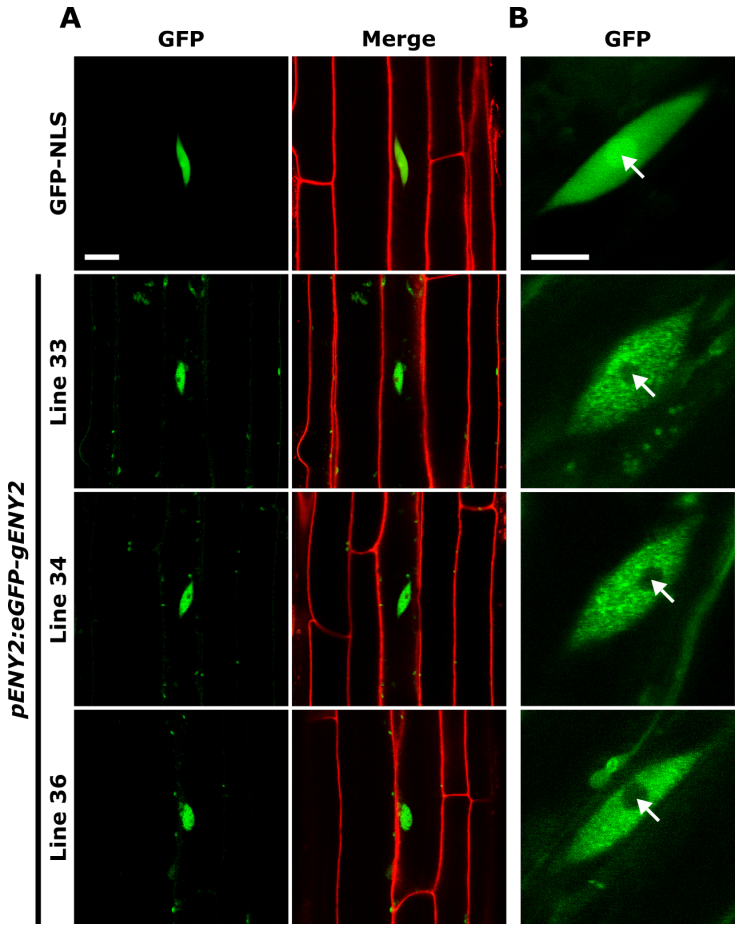
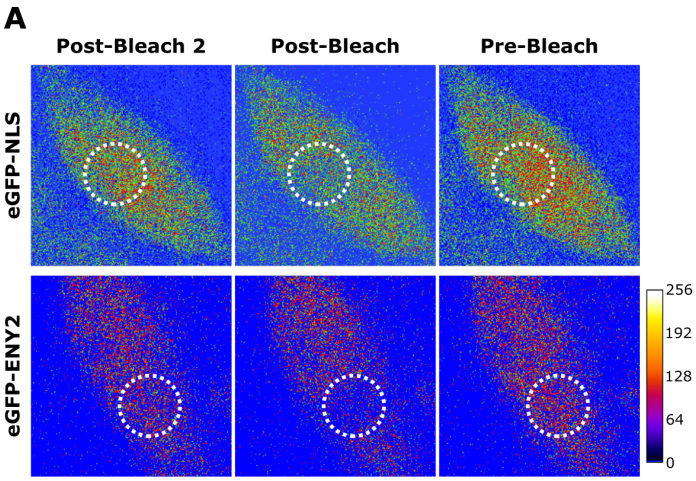


Figure 5.6 In the nucleoplasm of living cells, eGFP-ENY2 fusion proteins were forming speckle-like structures. The three transgenic eGFP-ENY2 lines and the representative eGFP-NLS line (positive control) were subjected to confocal microscopy to investigate the subcellular localization of the GFP-tagged proteins. The cell walls were stained with propidium iodide (red signal). A) Differentiated root epidermis cells are shown. Bars indicate 10 μm. B) Nuclei of differentiated root epidermal cells are shown. White arrows indicate the nucleolus. Bars indicate 10 μm.

5.2.4 High mobility of eGFP-ENY2 in the nucleoplasm

The root epidermal nuclei of transgenic eGFP-ENY2 and eGFP-NLS plants were analysed by FRAP to examine the mobility and *in vivo* binding properties of eGFP-tagged ENY2 in comparision to free eGFP-NLS. These experiments were done in three biological replicates. Both proteins, eGFP-ENY2 and eGFP-NLS recovered the bleached spot in the nucleoplasm within seconds (**Figure 5.7**). With a half-time ($t_{1/2}$) for fluorescence recovery of 2.28 seconds, eGFP-ENY2 was significantly less mobile than eGFP-NLS with 0.32 seconds. This quantitative analysis indicated short term bindings of ENY2 to proteins. The recovery curves of both eGFP fusion proteins reached a very high plateau, at 95% (eGFP-ENY2) and 97.8% (eGFP-NLS) of the initial pre-bleach fluorescence intensity. Almost the whole population of both GFP fusion proteins was mobile. This indicated no long term bindings to slow or immobile structures, for instance, like chromatin or the nuclear envelope.



B

	T-half [sec.]	Mobile fraction [%]
eGFP-NLS	0.32 ± 0.18 (n= 13)	97.77 ± 3.22 (n= 13)
eGFP-ENY2	2.28 ± 0.99 (n= 13) (p= 2.9E-06)	94.92 ± 5.98 (n= 13) (p= 0.17)

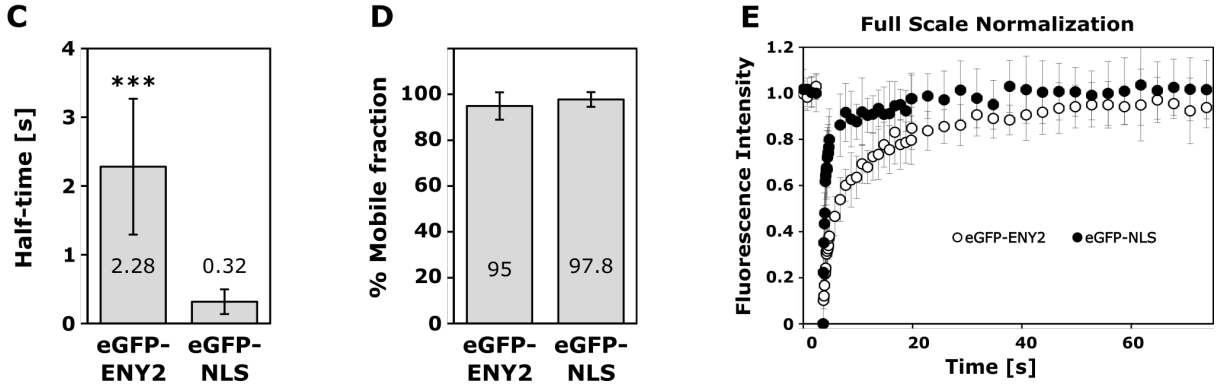


Figure 5.7 The characterization of eGFP-ENY2 dynamics in living cells.
A) FRAP of transgenic eGFP-NLS and eGFP-ENY2 epidermal cells is shown. The circular 3 μm ROIs (indicated by doted circles) were photobleached and the recovery of the fluorescence intensity was measured over time in these areas. Pre-Bleach indicates the first timepoint of the series (t=0 s), Post-Bleach the first timepoint after bleaching (t=3.7 s) and Post-Bleach 2 the last time point of the series (t=76.8 s). Pseudo-coloured images (modified fire LUT) with respective colour calibration bar are shown. B) Half-time ($t_{1/2}$) and mobile fraction values (with standard deviations) of eGFP-NLS and eGFP-ENY2 measurements. The significance was tested by Student's T-Test. C) The half-time of eGFP-ENY2 and eGFP-NLS was 2.28 s and 0.32 s, respectively. The significance was tested by Student's T-Test. D) The mobile fraction of eGFP-ENY2 and eGFP-NLS was 95% and 97.8%, respectively. E) The mean fluorescence recovery curves (Full scale normalization) for eGFP-ENY2 (n= 13) and eGFP-NLS (n= 13) are shown. This plot shows the normalized GFP intensities of 40x pre-bleach and 50x post-bleach time points. The white dots indicate eGFP-ENY2 and the black dots indicate eGFP-NLS. The standard deviations are shown as light grey bars.

5.3 Affinity purification coupled to mass spectrometry to identify the interaction network of ENY2 in plants

In other organisms like yeast and *Drosophila*, Sus1/ENY2 is a shared component of several multi-protein complexes and links the different steps of the gene expression pathway with each other. As part of both, the SAGA-DUB module and the TREX-2 complex, ENY2 is connecting transcript initiation with mRNA export, respectively (Rodríguez-Navarro et al. 2004 Kurshakova et al. 2007). To unravel a potential role of ENY2 in plant gene expression, putative interaction partners of ENY2 were identified by affinity purification coupled to mass spectrometry (AP-MS).

ENY2 was N-terminally fused to a SG-Tag that consists of a streptavidin binding peptide and two protein G domains (Van Leene et al. 2008; **Figure 5.8 A**). The transgenes are driving the expression of ENY2-SG fusion protein or unfused SG-Tag (negative control) under control of

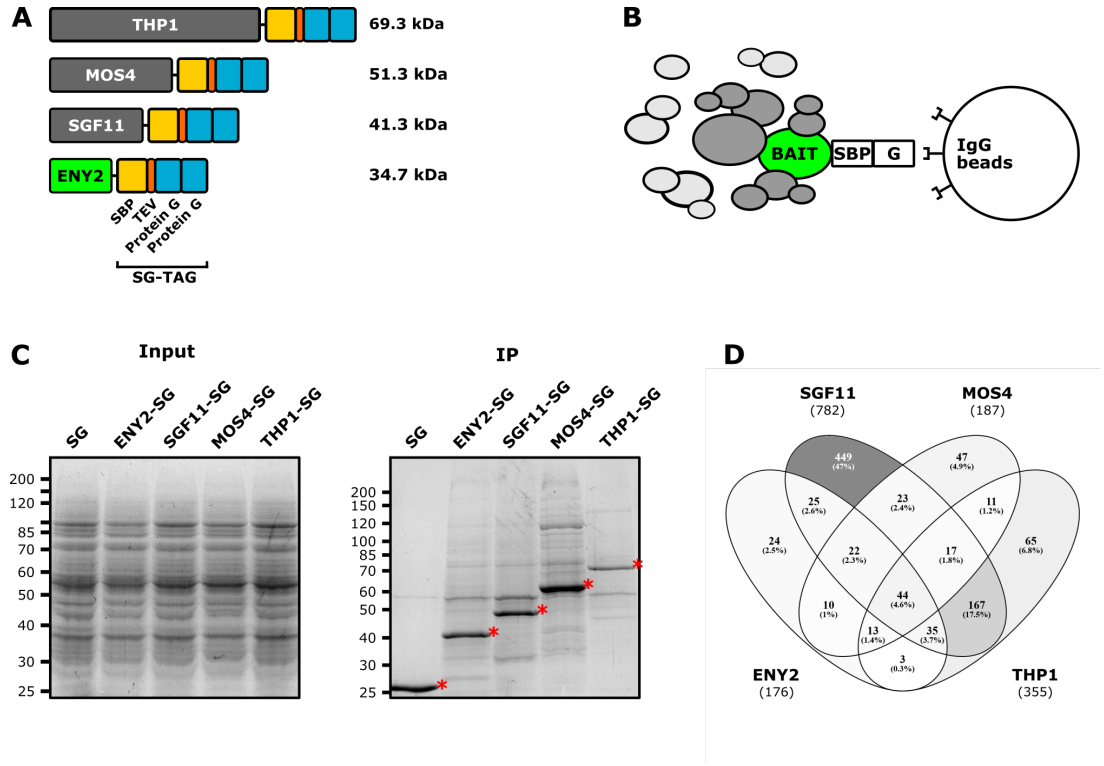


Figure 5.8 AP-MS reveals association of *Arabidopsis* ENY2 with transcription and splicing, but not mRNA export.
A) Schematic illustration of bait proteins C-terminally fused to SG tag (Streptavidin-binding protein and two Protein G domains linked by TEV cleavage site). Molecular weight of fusion proteins is computed by ExPaSy pI/Mw tool. B) Schmatic illustration of AP purification of SG-tagged proteins with IgG-coupled magnetic beads. C) Total protein extracts (Input, 0.01 % of total) of transgenic cells expressing bait proteins and SG-tag only. Eluates (50 % of total) of one-step affinity purifications (AP) using IgG coupled magnetic beads. Proteins were separated by 9 % SDS-PAGE and gels were stained with Coomassie Blue. Red asterisks indicate the bands corresponding to unfused SG-tag and bait proteins. Sizes of molecular weight marker proteins are shown in kDa. D) Overlap of proteins co-purified with four bait proteins were plotted in Venn diagram.

CaMV 35S promoter. The constructs were introduced into the genome of cultured *Arabidopsis* cells (PSB-D) by co-cultivation with transgenic *Agrobacterium* (Van Leene et al. 2011, Van Leene et al. 2015). The transformed cells were selected, upscaled to 10 liters and harvested for IgG affinity purification. Following one-step affinity purification (AP) using IgG coupled magnetic beads, the ENY2-SG or the unfused SG-Tag with corresponding co-purifying proteins were separated by SDS-PAGE and detected by staining with Coomassie Blue (**Figure 5.8 C**). The two bait proteins, ENY2-SG and unfused SG were clearly visible as most prominent bands with 26 and 34.7 kDa, respectively. The candidate interactors of ENY2-SG appeared with lower intensities (substoichiometric amounts) over the whole migration length. In comparison, almost no proteins were co-purified with the unfused SG-tag (Negative control). The entire gel lane was cut into several pieces for in-gel tryptic digest and the identification of putative interactors by mass spectrometry. Reliable identified proteins in at least two out of three AP-MS replicates were taken into consideration. Proteins found in the empty SG control were subtracted to remove unspecific background and SG-tag binding proteins. To remove interactors that are less relevant for further investigations, co-purified proteins were compared with a list of nonspecific proteins (Van Leene et al. 2015). This list based on the occurrence of proteins in 543 TAP experiments from PSB-D cells using 115 different bait proteins. Additionally, the

experimentally identified proteins were compared to a list of expected interactors (**Table S23**) to get more information about their association to protein complexes. In total 176 proteins were co-purified with ENY2 (**Table ??**). The entire experimental procedure of affinity purification coupled to mass spectrometry and processing of the MS data is described in detail in Pfab et al. 2017.

Table 5.1 Transcription-related proteins co-purifying with ENY2-SG, SGF11-SG, MOS4-SG and THP1-SG.
The numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent APs. The proteins that were detected in less than two out of three AP are not listed.

AGI	Protein	Complex	Process	MOS4	ENY2	SGF11	THP1
AT3G27100	ENY2	SAGA-DUBm	Transcription	402 / 3	783 / 3	1262 / 3	
AT5G58575	SGF11	SAGA-DUBm	Transcription		542 / 3	1747 / 3	
AT5G10790	UBP22	SAGA-DUBm	Transcription		113 / 3	209 / 3	
AT2G17930	TRA1a	SAGA_SPT	Transcription			270 / 2	
AT4G36080	TRA1b	SAGA_SPT	Transcription			160 / 2	
AT4G04920	Med16	Mediator	Transcription			171 / 2	
AT4G35800	NRPB1	RNAPII	Transcription			328 / 3	
AT4G21710	NRPB2	RNAPII	Transcription			330 / 3	
AT4G10710	SPT16	FACT	Transcription			215 / 2	
AT5G63670	SPT4-2	SPT4/SPT5	Transcription			137 / 2	
AT4G08350	SPT5-2	SPT4/SPT5	Transcription			327 / 3	152 / 2
AT1G65440	SPT6L	SPT6	Transcription			209 / 2	

As expected from literature, the other SAGA-DUB module components SGF11 and UBP22 were co-purified with ENY2-SG, but surprisingly no other subunits of the multi-protein SAGA complex (**Table 5.1**). Also unexpectedly, no components of the TREX-2 complex as well as no other nuclear pore complex (NPC) related proteins were found (**Table 5.2**). This suggests an asscociation of *Arabidopsis* ENY2 with of the SAGA-DUB module, but not with the TREX-2 complex. Additionally, several components of the THO/TREX complex (ALY1-4 and THO5/7) were co-purified with ENY2-SG (**Table S23**). Moreover, several components of the splicing machinery co-purified with ENY2-SG including all nine core components of the NineTeen Complex (NTC)(**Table 5.3**). During splicing, the NTC complex is important for the catalytic activation of the spliceosome. Furthermore, NTC plays a role in transcript elongation in yeast and in the genome maintenance in higher eukaryotes (Chanarat et al. 2013). This striking association between ENY2 and the NTC complex was unexpected and so far not reported for ENY2 orthologs in other species.

Table 5.2 TREX-2 proteins co-purifying with ENY2-SG, SGF11-SG, MOS4-SG and THP1-SG.
The numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent APs. The proteins that were detected in less than two out of three AP are not listed.

AGI	Protein	Complex	Process	MOS4	ENY2	SGF11	THP1
AT2G19560	THP1	TREX-2	Export				2078 / 3
AT2G39340	SAC3A	TREX-2	Export				3571 / 3
AT3G06290	SAC3B	TREX-2	Export				3517 / 3
AT3G54380	SAC3C	TREX-2	Export				1033 / 3

To confirm the results of the initial AP-MS analysis of ENY2-SG, the reverse experiments were conducted. Therefore, representatives of the SAGA-DUB module (SGF11), the NTC complex (MOS4) and the TREX-2 complex (THP1) were subjected to AP-MS (**Figure 5.8**) as described above for ENY2. All four proteins (ENY2, SGF11, MOS4 and THP1) are ubiqui-

tously co-expressed (**Figure S7**) in *Arabidopsis* plants throughout development, which allows for hypothetical interaction. The MS data (**Table 5.1, 5.2, 5.3**) of the reverse experiments revealed that ENY2 co-purified with SGF11-SG and MOS4-SG, but not with THP1-SG. This confirmed the conclusion of the ENY2-SG experiment that *Arabidopsis* ENY2 is interacting with components of the SAGA-DUB module, but not with proteins of the TREX-2 complex. This indicates that the role of ENY2 in plants differs partly from that of ENY2 in other organisms. In plants, ENY2 seems to be associated with transcription (SAGA-DUBm) and splicing (NTC), instead of transcription (SAGA-DUBm) and mRNA export (TREX-2).

In total, after processing of the MS data, 176 proteins were co-purified with ENY2-SG, 782 proteins with SGF11-SG, 187 with MOS4-SG and 355 with THP1-SG (**Figure 5.8 D**). The Venn diagram showed that several proteins were co-purified with more than one bait protein. With SGF11-SG (449 Proteins; 47 %) and THP1-SG (65 Proteins; 6.8 %) the most proteins were identified that only co-purified with one of the four bait proteins. The largest intersection of two bait proteins was between SGF11 and THP1 with 167 proteins (17.5 %). The smallest intersection of two bait proteins was between ENY2 and THP1 with just 3 proteins (0.3 %). 44 proteins (4.6 %) were co-purified with all four bait proteins.

GO analysis was carried out to assess if bait proteins act in common pathways (**SFig S8**). Overrepresented GO terms with p-Values < 0.05 were identified by AgriGO using Singular enrichment analysis (SEA) (Tian et al. 2017) and redundant terms were removed by REViGO (Supek et al. 2011). Most GO terms were overepresented in the AP-MS data of all four bait proteins, including the most promiment term "RNA processing" (Splicing) or less prominent terms like "post-embryonic development" and "cellular process". A few GO terms were just overepresented in individual AP-MS datasets like the "Nucleocytoplasmic transport" in THP1-SG, which is in accordance with the well known role of THP1 in mRNA export.

In the ENY2 and SGF11 AP-MS, all predicted *Arabidopsis* DUB module components (ENY2, SGF11, UBP22) were found, but no components of the other SAGA modules (**Table 5.1**). This raised the questions if the SAGA complex really exists in *Arabidopsis* and if so whether the DUB module is a part of it. Furthermore, with MOS4-SG just ENY2 was co-purified and no other SAGA-DUB or SAGA components. This indicated that not SAGA-DUBm is associated with NTC, but just ENY2 independently of SGF11 and UBP22. Moreover, several chromatin-related factors like the histones H2A and H2B and transcription-related proteins like the RNAPII and transcript elongation factors including PAF-C, SPT4/5 and SPT16 were co-purified with SGF11-SG (**Table 5.1**), what is in line with its putative conserved function in the deubiquitination of histone H2B. Surprisingly, these chromatin and transcription-related factors were not found in the AP-MS data of ENY2-SG.

All three variants of mRNA export factor SAC3 (SAC3A-C) were exclusively identified in the eluates of the THP1-SG affinity purification (**Table 5.2**). No additional subunits of the putative *Arabidopsis* TREX-2 like DSS1 or CEN1/2 were detected. SAC3 (GANP) represents the core component of human and yeast TREX-2 around which the other subunits assemble (Ellisdon et al. 2012, Jani et al. 2012). Moreover, several components of the THO/TREX complex including UAP56 and five NPC-related nucleoporins co-purified with THP1-SG (**Figure ??**). The THO/TREX proteins also co-purified with ENY2-SG, SGF11-SG and MOS4-SG, but

Table 5.3 Splicing-related proteins co-purifying with ENY2-SG, SGF11-SG, MOS4-SG and THP1-SG.
The numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent APs. The proteins that were detected in less than two out of three AP are not listed.

AGI	Protein	Complex	MOS4	ENY2	SGF11	THP1
AT3G50670	U1-70K	U1 snRNP		166 / 3	210 / 3	194 / 3
AT2G47580	U1A	U1 snRNP	364 / 2	342 / 3	472 / 2	511 / 2
AT1G44910	PRP40A	U1 snRNP-rel.	313 / 2			113 / 2
AT1G60200	AT1G60200	U1 snRNP-rel.	207 / 2			148 / 2
AT1G09760	atU2A-	U2 snRNP	1127 / 2	418 / 3	489 / 3	304 / 3
AT2G30260	atU2B-b	U2 snRNP	568 / 2	423 / 3	304 / 3	
AT5G12190	P14	U2 snRNP	348 / 3	267 / 3	330 / 2	244 / 2
At1g14650	SAP114-1a	U2 snRNP	935 / 3	350 / 2		
AT3G55200	SAP130a	U2 snRNP	1734 / 2	637 / 3	859 / 3	550 / 2
AT5G64270	SAP155	U2 snRNP	2867 / 2	866 / 3	902 / 3	906 / 3
AT2G18510	SAP49a	U2 snRNP	144 / 2	139 / 2	154 / 3	232 / 2
AT5G06160	SAP61	U2 snRNP	648 / 2	205 / 2		135 / 2
AT2G32600	SAP62	U2 snRNP	260 / 2	199 / 2		132 / 2
AT4G21660	SF3b150	U2 snRNP	749 / 3	157 / 2	243 / 2	
AT1G30480	SPF45	U2 assoc.				163 / 2
AT5G25060	SR140	U2 assoc.			253 / 3	246 / 2
AT2G41500	EMB2776	U4/U6 snRNP	249 / 2		115 / 2	
AT1G28060	SAP90-1	U4/U6 snRNP	436 / 2			
AT2G38730	Tri-20	U4/U6 snRNP	415 / 2	294 / 2	229 / 2	193 / 2
AT1G60170	EMB1220	U4/U6 snRNP		97 / 2		
AT5G57370	Tri-27	U4/U6.U5 snRNP			132 / 2	
AT5G16780	MDF	U4/U6.U5 snRNP	443 / 2			
AT2G33730	U5-100KD	U5 snRNP	381 / 2			
AT4G03430	EMB2770	U5 snRNP	781 / 2	190 / 3	695 / 3	
AT5G08290	YLS8	U5 snRNP	119 / 2			
AT5G61140	U5-200-1	U5 snRNP			453 / 2	
AT1G20960	EMB1507	U5 snRNP	5413 / 3	3102 / 3	1650 / 3	1134 / 3
AT2G43770	U5-40	U5 snRNP	423 / 2	172 / 3	284 / 2	300 / 2
AT5G42820	U2AF35	Splice site sel.			150 / 2	222 / 2
AT1G27650	U2AF35a	Splice site sel.			125 / 2	
AT4G36690	U2AF65a	Splice site sel.	219 / 3		127 / 2	314 / 2
AT1G60900	U2AF65b	Splice site sel.	297 / 3	122 / 2	180 / 2	274 / 2
AT3G19590	BUB3.1	A complex-assoc.			156 / 2	
AT1G67580	AT1G67580	A complex-assoc.	252 / 2			
AT5G38840	AT5G38840	A complex-assoc.	281 / 2			
AT3G54230	SUA	A complex-assoc.	223 / 2			
AT4G09980	EMB1691	B complex-assoc.			106 / 2	
AT3G12300	AT3G12300	Bact complex	170 / 2			
AT1G05460	SED3	Bact complex			338 / 3	180 / 2
AT3G25840	PRP4K-1	Bact complex	184 / 2			
At5g09880	At5g09880	Bact complex				123 / 2
At2g16940	At2g16940	Bact complex	548 / 2	130 / 2	186 / 3	286 / 3
AT4G18465	DDX35	C complex	835 / 3	166 / 2	112 / 3	
AT3G63400	AT3G63400	C complex	499 / 3			
AT3G44600	CYP71	C complex	930 / 2			
AT5G51280	AT5G51280	C complex	711 / 2	181 / 2	85 / 2	
AT1G03910	AT1G03910	C complex	278 / 2			
AT1G61620	AT1G61620	C complex	620 / 2	130 / 2	193 / 2	
AT1G01940	AT1G01940	C complex	345 / 2			
AT2G21150	XCT	C complex		102 / 2		
AT5G23080	TGH	C complex	949 / 2			152 / 2
AT5G64730	AT5G64730	C complex		106 / 2	95 / 2	105 / 2
AT1G18080	ATARCA	C complex			250 / 2	
AT2G20330	AT2G20330	C complex-assoc.	711 / 2			

Table 5.3 (Continuation)

AGI	Protein	Protein complex	MOS4	ENY2	SGF11	THP1
AT5G13010	EMB3011	RES complex	601 / 2	104 / 3	279 / 3	255 / 2
AT3G26560	Prp22-1	RES complex	1431 / 3	469 / 3		408 / 2
AT1G27900	Prp22-3	RES complex	135 / 2			
AT3G62310	Prp43-2a	RES complex	819 / 3	252 / 2	221 / 3	
AT2G47250	Prp43-2b	RES complex	704 / 3		147 / 3	
AT3G20550	DAWDLE	RES complex			171 / 3	
AT1G31870	CWC26	RES complex	292 / 2			
AT1G17070	AT1G17070	RES complex	728 / 3			92 / 2
AT1G09770	CDC5	Core NTC	4211 / 3	613 / 3	279 / 2	495 / 2
AT5G41770	CRN1c	Core NTC	1176 / 3	346 / 3		
AT3G18790	ISY1	Core NTC	442 / 2	288 / 3		128 / 2
AT1G04510	MAC3A	Core NTC	2304 / 3	435 / 2	297 / 2	365 / 2
AT2G33340	MAC3B	Core NTC	3137 / 3	339 / 3	269 / 2	349 / 2
AT3G18165	MOS4	Core NTC	1849 / 3	174 / 3	127 / 2	
AT4G15900	PRL1	Core NTC	1296 / 3	277 / 3		248 / 2
At1g77180	SKIP	Core NTC	1243 / 3	280 / 3	247 / 2	306 / 3
AT5G28740	SYF1	Core NTC	2124 / 3	613 / 3		
AT2G16860	SYF2	NTR		108 / 2		
AT2G38770	EMB2765	NTR	2481 / 3	1072 / 3	554 / 3	468 / 3
AT4G21110	Bud31	NTR	190 / 2			
AT3G02710	CTNNNBL1	NTR	912 / 2			
AT4G34870	ROC5	NTR	128 / 2			
AT1G07360	ECM2-1a	NTR	592 / 3	315 / 3	299 / 3	
AT2G29580	ECM2-1b	NTR	376 / 2	178 / 3	126 / 2	151 / 2
AT1G10580	PRP17-1	NTR	608 / 2	155 / 2		111 / 2
AT1G32490	PRP2a	NTR	925 / 3	221 / 2	97 / 2	107 / 2
AT3G23900	SRM300-l	NTR	247 / 2			
AT3G13200	EMB2769	NTR	297 / 2	142 / 2		175 / 2
AT3G05070	CWF18	NTR	241 / 2	191 / 2		
AT1G56290	CWFJ-like	NTR	581 / 2			
AT2G36130	AT2G36130	NTR	1279 / 2	280 / 3	137 / 2	213 / 2
AT1G80930	CWC22	NTR	862 / 3	385 / 3		
AT3G19840	PRP40C	NTR	222 / 2			
AT5G56900	AT5G56900	NTR	452 / 2			
AT4G20440	SmB-b	Sm core		323 / 2		
AT3G07590	SmD1-a	Sm core	699 / 3	524 / 3	439 / 2	350 / 3
AT4G02840	SmD1-b	Sm core	691 / 3	565 / 2	398 / 2	345 / 3
AT3G62840	SmD2-b	Sm core	660 / 2	238 / 3	276 / 3	312 / 2
AT1G76300	SmD3-a	Sm core	430 / 3	244 / 3	276 / 2	192 / 2
AT1G20580	SmD3-b	Sm core	448 / 2	298 / 2	233 / 2	
AT4G30330	SmE-a	Sm core	519 / 3	424 / 2	288 / 2	
AT2G18740	SmE-b	Sm core	558 / 3	331 / 3	341 / 2	285 / 3
AT3G11500	SmG-b	Sm core	133 / 2			
AT1G03330	LSM2	Lsm core	319 / 2		115 / 2	
At1g21190	LSM3a	Lsm core	130 / 2	88 / 2		
AT1G76860	LSM3b	Lsm core	138 / 2	85 / 2	105 / 2	
AT5G27720	EMB1644	Lsm core	137 / 2			
AT3G59810	LSM6a	Lsm core	123 / 2			
AT2G43810	LSM6b	Lsm core	175 / 2		94 / 2	
AT2G03870	LSM7	Lsm core	144 / 2	97 / 2		128 / 2
AT1G65700	LSM8	Lsm core	103 / 2			
AT3G61860	RSp31	SR proteins			250 / 3	268 / 3
AT2G46610	RSp32	SR proteins	486 / 2	348 / 3	166 / 3	212 / 3
AT5G52040	RSp41	SR proteins		114 / 2	127 / 2	
AT1G23860	RSzp21	SR proteins		204 / 3	124 / 3	201 / 2

Table 5.3 (Continuation)

AGI	Protein	Protein complex	MOS4	ENY2	SGF11	THP1
AT4G31580	RSZp22	SR proteins	289 / 3	424 / 2	391 / 3	343 / 3
AT2G24590	RSZp22a	SR proteins	303 / 2	264 / 3	262 / 3	302 / 2
AT5G64200	SC35	SR proteins			122 / 2	
AT5G18810	atSCL28	SR proteins	148 / 2	163 / 3		119 / 2
AT3G55460	SCL30	SR proteins	225 / 3	215 / 3	196 / 3	181 / 2
AT3G49430	SR34a	SR proteins	148 / 3	225 / 2		
AT3G44850	SRPK2c	SR proteins		121 / 2		
AT2G29210	AT2G29210	SR-related	158 / 2			
AT1G18630	GRBP1a	GRPs			442 / 2	
AT5G04280	hnRNP-G1	hnRNP family	210 / 2			162 / 2
AT1G60650	hnRNP-G3	hnRNP family		137 / 3		
AT5G54900	ATRBP45A	hnRNP family			135 / 2	624 / 3
AT3G19130	AtRBP47b	hnRNP family				305 / 2
AT1G47500	AtRBP47c	hnRNP family				120 / 3
AT3G13224	RNP_N1	hnRNP family			108 / 2	326 / 3
AT4G26650	RNPA	hnRNP family				189 / 2
AT5G40490	RNPA/B_8a	hnRNP family				357 / 2
AT3g26420	ATRZ-1A	hnRNP family	278 / 2	272 / 3	166 / 3	309 / 2
AT1G54080	AtUBP1a	hnRNP family				197 / 3
AT2G44710	hnRNP-R3	hnRNP family	413 / 2		415 / 3	276 / 3
AT3G07810	AT3G07810	hnRNP family				417 / 3
AT3G20890	AtRNPH	hnRNP family				96 / 2
AT5G58470	AT5G58470	hnRNP family			174 / 2	119 / 2
AT3G15010	UBA2c	hnRNP family			190 / 2	563 / 3
AT3G58570	AT3G58570	EJC/mRNP	362 / 2	219 / 2	477 / 3	430 / 3
AT1G02140	HAP1	EJC/mRNP	499 / 2	431 / 2		
AT2G45640	SAP18	EJC/mRNP	623 / 2	337 / 3	285 / 2	243 / 2
AT1G16610	SR45	EJC/mRNP				174 / 2
AT1G51510	Y14	EJC/mRNP	452 / 2	251 / 2		225 / 3

to a smaller extend. Beside THP1-SG, the NPC-related nucleoporins were just identified in the affinity purification of SGF11-SG.

All nine components of the splicing NTC complex were co-purified with ENY2-SG and the NTC subunit MOS4-SG (**Table 5.3**). Moreover, fitting to the role of the NTC complex in the catalytic activation of the spliceosome, several U2 and U5 snRNP proteins as well as NTC-related (NTR) and C complex associated proteins were co-purified with both bait proteins. The striking overlap between the splicing factors that co-purified with ENY2-SG and MOS4-SG indicated an association of ENY2 with the NTC complex. As the process of mRNA splicing is a central step in the gene expression pathway, splicing factors including the NTC complex were also co-pufied with SGF11-SG and THP1-SG, but in comparison to MOS4-SG and ENY2-SG to a lower extend.

Taken together, the potential interaction partners of ENY2, SGF11, MOS4 and THP1 were determined by AP-MS and a network of protein-protein interactions could be identified (**Figure 5.9**). This showed the expected interaction of ENY2 with the SAGA-DUB module (transcription), but not with TREX-2 (mRNA export). Moreover, ENY2 showed a strong association with the splicing complex NTC.

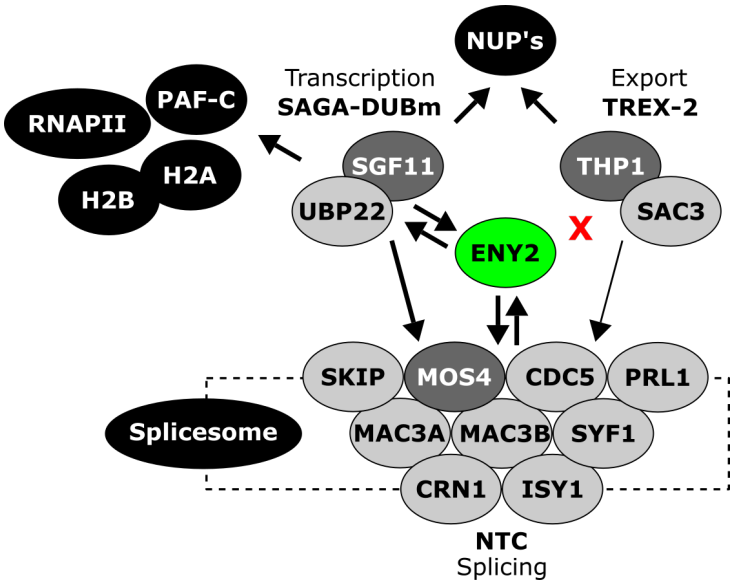


Figure 5.9 Protein interaction network of ENY2, SGF11, MOS4, and THP1.
Schematic interaction network revealed by AP-MS analysis of SG-fusion proteins (ENY2-SG, SGF11-SG, MOS4-SG and THP1-SG). Bait proteins are depicted in dark grey and green. Arrows indicate co-purification. Red cross indicates no co-purification.

5.4 The SAGA-DUB module in plants

The AP-MS experiments using reciprocal tagging showed that all three DUBm components co-purified with each other in *Arabidopsis* (**Section 5.3**). In the following, these interactions were further analysed by Yeast Two-Hybrid (Y2H) and Förster Resonance Energy Transfer (FRET) experiments.

5.4.1 Y2H revealed PPIs between SGF11 and ENY2 as well as SGF11 and UBP22

The direct PPIs of the putative DUB components (ENY2, SGF11, UBP22) were studied using the Matchmaker^{TM}GAL4 yeast-two hybrid (Y2H) system (Clonetech). For each test combination, transgenic yeast cells expressing one candidate fused to the GAL4 DNA-binding domain (DNA-BD) and the other one to the GAL4 activation domain (AD) were grown under selective conditions to identify positive interactions (Fields et al. 1989, Chien et al. 1991). Therefore, pGBKT7 (DNA-BD; bait) and pGADT7 (AD; prey) plasmids containing the CDS of the respective genes (*ENY2*, *SGF11* and *UBP22*) were generated and competent AH109 cells were co-transformed with all test combinations using the PEG/LiAc method. The transgenic yeast cells expressing both the bait and the prey plasmids were identified by growth on double dropout media (DDO) lacking leucine (-Leu) and tryptophan (-Trp). Positive interactions of bait and prey proteins were assessed by growth on triple (TDO; -Leu/-Trp/-His) and quadruple (QDO; -Leu/-Trp/-His/-Ade) dropout plates. Cells expressing DNA-BD/murine p53 and AD/SV40 large T-antigen functioned as positive control showing growth on DDO, TDO and QDO (Li et al. 1993, Iwabuchi et al. 1993). Cells expressing DNA-BD/Lamin and AD/T were used as negative control with no growth on TDO and QDO media. All bait and prey proteins showed no autoactivation or toxicity, as the following combinations showed no growth on TDO and QDO media: DNA-BD/bait proteins with AD alone and vice versa DNA-BD alone with

AD/prey proteins.

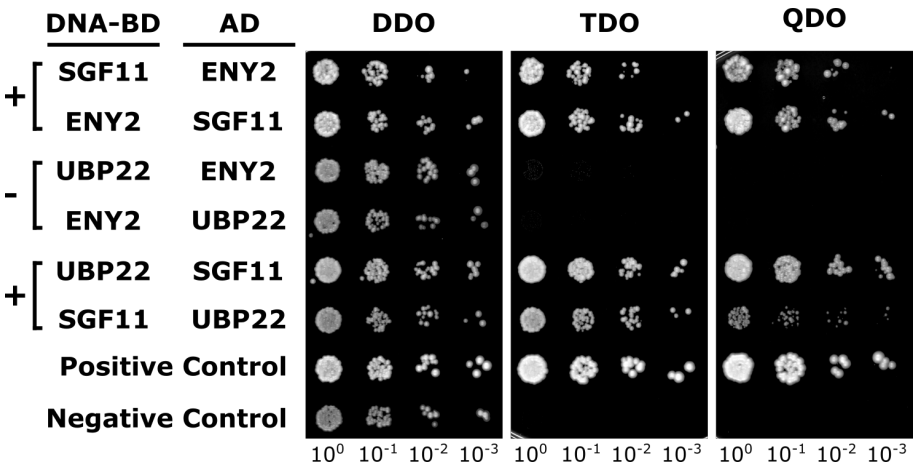


Figure 5.10 Y2H assay revealed interactions between ENY2 and SGF11 as well as SGF11 and UBP22. Different combinations of DNA-BD/bait and AD/prey fusion proteins were spotted on DDO (SD/ -LEU -TRP), TDO (SD/- LEU -TRP -HIS) and QDO (SD/ -LEU -TRP -HIS -ADE) plates in serial dilutions. DNA-BD/murine p53 and AD/SV40 large T-antigen functioned as positive control. DNA-BD/Lamin and AD/T was used as negative control.

DNA-BD/ENY2 and AD/SGF11 as well as DNA-BD/SGF11 and AD/ENY2 showed growth on all three selective media (**Figures 5.10 and S9**). The same growth pattern could be observed for both combinations to test the interaction between SGF11 and UBP22. Whereby cells expressing DNA-BD/SGF11 and AD/UBP22 showed reduced growth on TDO and QDO media, compared to the DNA-BD/UBP22 – AD/SGF11 combination. Yeast cells expressing ENY2 bait and UBP22 prey or vice versa showed no growth on TDO and QDO media. In summary, the Y2H assay revealed direct protein-protein interactions of ENY2 and SGF11 as well as SGF11 and UBP22. However, no PPI between ENY2 and UBP22 was observed in yeast.

5.4.2 FRET revealed PPIs between SGF11 and ENY2 as well as SGF11 and UBP22

The protein-protein interactions of ENY2, SGF11 and UBP22 were validated by *in vivo* FRET (Fluorescence/Foerster Resonance Energy Transfer) measurements in plants. FRET in a nutshell: First, both putative interactors are fused to fluorescent proteins (FP). One is tagged with a donor FP (shorter wavelength range) and the other one with an acceptor FP (Longer wavelength range). The emission spectrum of the donor FP and the excitation spectrum of the acceptor FP need to overlap. In case of a protein-protein interaction, both FP are brought in close proximity to each other (< 10 nm). This allows an energy transfer (FRET) from the excited donor to the acceptor by a nonradiative dipole-dipole coupling mechanism that can be detected (Förster 1948, Piston et al. 2007, Chudakov et al. 2010).

In this approach, eGFP (donor FP) and mCherry (acceptor FP) were chosen as FRET partners. This FP pair shows a very good spectral separation, but at the same time the donor emission and the acceptor excitation show a high overlap (Förster Radius R0=5.1 nm) (**Figure S10**). This makes it very suitable for quantitative FRET efficiency measurements as shown by other groups (Tramier et al. 2006, Hazelwood et al. 2008, Albertazzi et al. 2009).

For the FRET efficiency measurements in plants, first a vector system was generated by restriction-ligation cloning including the following (**Figure S11**): Two plasmids to fuse the donor

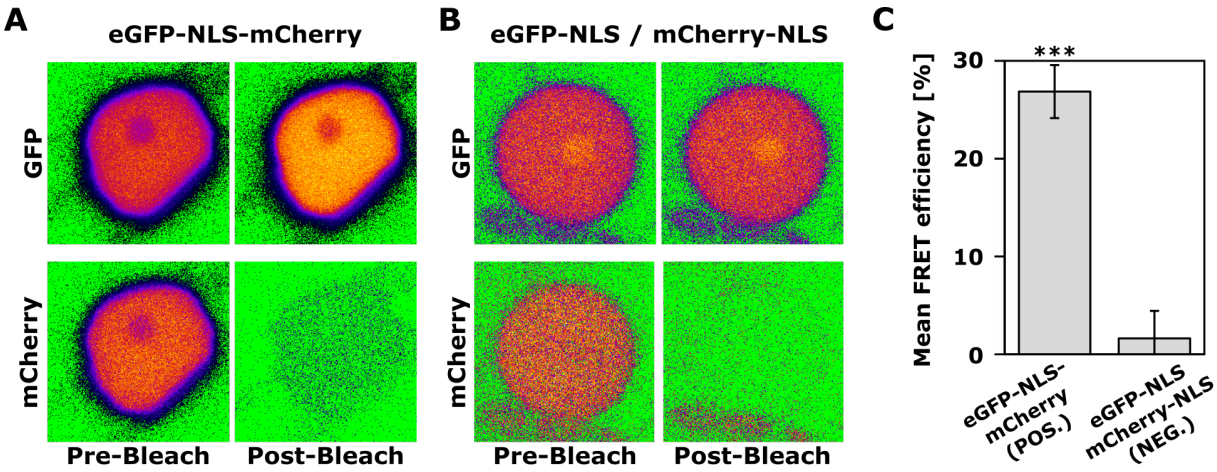


Figure 5.11 Testing the FRET system in plants.
(A) The positive (eGFP-NLS-mCherry fusion) and (B) negative controls (free eGFP-NLS and mCherry-NLS) were transiently expressed in *N. benthamiana* leaves. The transformed epidermal cells were used for the Acceptor-Photo-Bleaching (APB) experiment. "Pre-Bleach" and "Post-Bleach" indicate the time points immediately before and after the bleaching of the acceptor fluorophore. In both controls, the acceptor signal (mCherry) was irreversibly removed after the bleaching. In the positive control (A), the donor signal (eGFP) was increased after the bleaching (= FRET). Whereas the donor signal in the negative control (B) was unchanged (= No FRET). The fluorescence intensities were pseudo-coloured (modified fire LUT) with ImageJ. The respective colour calibration bar is shown. C) The mean FRET efficiencies of the positive and the negative control ± SD are shown

FP (eGFP) or acceptor FP (mCherry) N-terminally to the protein of interest. Two plasmids for the expression of either free eGFP-NLS or mCherry-NLS as negative controls. One plasmid for the expression of an eGFP-NLS-mCherry fusion protein, as positive control. All constructs were based on pCambia2300 or pGreen0179 backbones and expression of the fluorescent proteins was driven and terminated by the CaMV 35S promoter and poly(A)signal, respectively.

To test the system, transgenic *Agrobacterium* cells harbouring the positive (eGFP-NLS-mCherry fusion) or negative control (free eGFP-NLS and mCherry-NLS) plasmids were both infiltrated into different *N. benthamiana* leaves. In this first approach, the epidermal cells that were transiently expressing the fluorescent proteins were used for the FRET analysis using confocal microscopy. The efficiency of the energy transfer was measured using the Acceptor Photo-Bleaching (APB) method (Weidtkamp-Peters et al. 2017, Karpova et al. 2003). As seen in the positive control, there was an increase of the donor (eGFP) fluorescence upon the irreversible bleaching of the acceptor (mCherry)(**Figure 5.11 A,C**). In this fusion protein, eGFP transfers its excitation energy to the physically linked mCherry. The close proximity of both fluorescent proteins could be measured as mean FRET efficiency of 25%. In contrast, cells expressing the not linked eGFP-NLS and mCherry-NLS (negative control) showed no increase of the donor fluorescence after the bleaching of the acceptor (**Figure 5.11 B,C**). The mean FRET efficiency of this donor/acceptor pair was approximately 3% which showed that there was no interaction between both freely diffusing proteins in the nucleus. This initial experiment verified the functionality of the established FRET system in *N. benthamiana*.

The components of the putative *Arabidopsis* DUB module were cloned into the generated FRET expression vectors to produce the following translational fusions with fluorescent proteins: eGFP-SGF11, eGFP-UBP22, mCherry-ENY2, and mCherry-SGF11. For all three test combinations (SGF11/ENY2, SGF11/UBP22, and UBP22/ENY2) the corresponding donor/acceptor

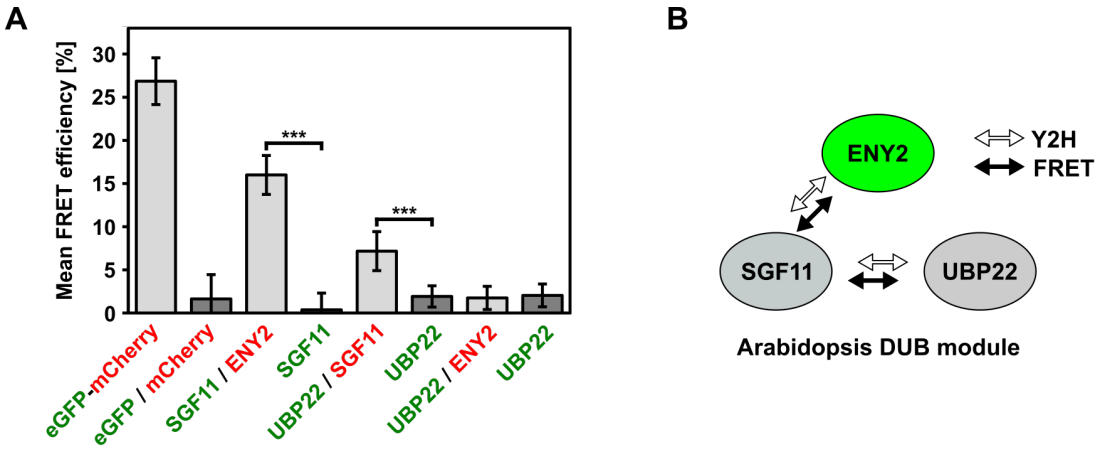


Figure 5.12 FRET revealed interactions between ENY2 and SGF11 as well as SGF11 and UBP22.
A) To test the interactions between the putative DUBm components, the corresponding donor/acceptor combinations (SGF11/ENY2, UBP22/SGF11, UBP22/ENY2) were co-infiltrated in *N. benthamiana*. The transiently transformed cells were subjected to acceptor-photobleaching FRET (APB-FRET). The mean FRET efficiencies of the three test combinations and respective controls are shown. B) The schematic illustration of the ENY2 – SGF11 – UBP22 interaction network. White/black arrows indicate the interactions determined by Y2H and FRET, respectively.

plasmids were transiently co-expressed in *N. benthamiana* leaves via *Agrobacterium*-mediated transformation. To exclude false FRET, cells expressing only the donor (eGFP-SGF11 or eGFP-UBP22) in the absence of the acceptor were used as negative controls (Karpova et al. 2003).

For the combinations eGFP-SGF11 / mCherry-ENY2 and eGFP-UBP22 / mCherry-SGF11 a positive FRET efficiency of 16% and 8% respectively was calculated. The corresponding donor only controls (eGFP-SGF11 and eGFP-UBP22) showed just a low FRET efficiency of 1% and 3% respectively. This indicated interactions between SGF11 and ENY2 as well as SGF11 and UBP22. With about 3%, the FRET efficiency for eGFP-UBP22 / mCherry-ENY2 was not higher than the negative control. This indicated that there was no direct interaction between UBP22 and ENY2. Therefore, the *in vivo* FRET analysis in plants could confirm the data obtained from the Y2H assay showing that SGF11 was interacting with ENY2 as well as with UBP22. These data indicated the formation of a putative DUB module in plants.

5.4.3 DUB module components are highly conserved

In eukaryotes, the conserved transcriptional co-activator SAGA is comprised of four subcomplexes with distinct functions. This modular structure allows the multiprotein complex to do both, modify chromatin and recognize histone modifications (Koutelou et al. 2010,Rodríguez-Navarro 2009, Samara et al. 2011). In yeast, the DUB module consists of the four proteins Ubp8, Sgf11, Sus1 and Sgf73 (Köhler et al. 2008, Lee et al. 2009). With exception of Sgf73, this module is conserved in higher plants (Moraga et al. 2015, Srivastava et al. 2015), but the proteins are not characterized so far.

Multiple sequence alignments (MSA) were generated to compare the protein sequences of ENY2, SGF11 and UBP22 with their orthologs from other species. The following species were selected as representatives for mammals (*Homo sapiens*, *Mus musculus*), invertebrates (*Drosophila melanogaster*), fungi (*Saccharomyces cerevisiae*), dicotyledons (*Glycine max*), monocotyledons (*Oryza sativa*, *Zea mays*) and mosses (*Physcomitrella patens*). The protein sequences were ob-

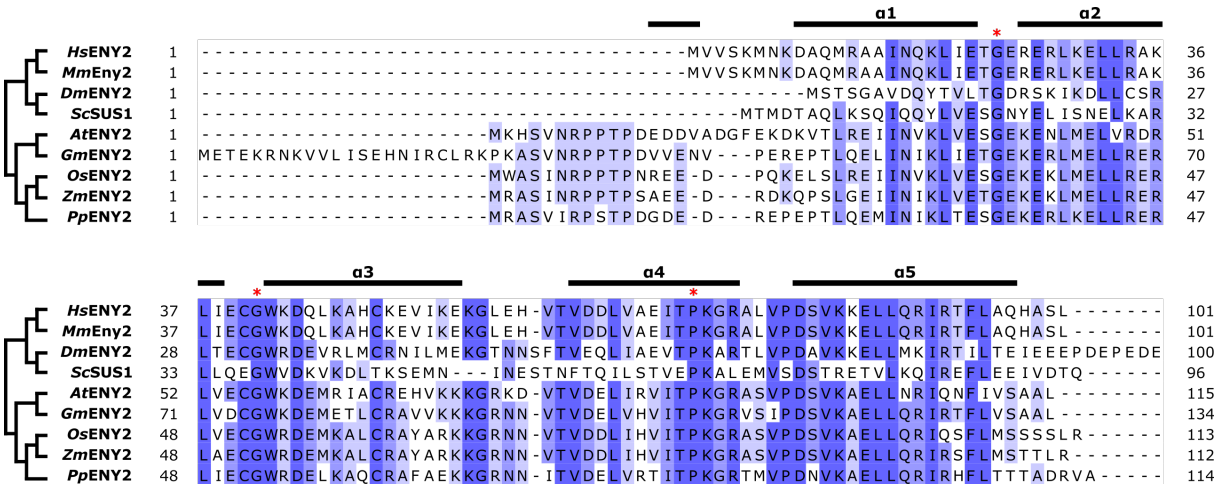


Figure 5.13 The amino acid sequence of the *Arabidopsis* ENY2 is conserved among different species.
The ENY2 protein sequence was aligned to its orthologs from the following other species: *Homo sapiens* (*Hs*), *Mus musculus* (*Mm*), *Drosophila melanogaster* (*Dm*), *Saccharomyces cerevisiae* (*Sc*), *Glycine max* (*Gm*), *Oryza sativa* (*Os*), *Zea mays* (*Zm*) and *Physcomitrella patens* (*Pp*). These species were selected as representatives for mosses (*Pp*ENY2), monocotyledons (*Zm*ENY2, *Os*ENY2), dicotyledons (*Gm*ENY2, *At*ENY2), fungi (*Sc*Sus1), invertebrates (*Dm*Eny2) and mammals (*Mm*ENY2, *Hs*ENY2). The multiple sequence alignment was generated using Clustal Omega (Sievers et al. 2011). The conservation of the aa residues was highlighted with blue coloration and the trees were calculated using JalView (Waterhouse et al. 2009). The alpha-helices predicted by I-TASSER (Yang et al. 2015) were indicated by black bars. The single gycine and proline residues that are important for the tertiary structure (Jani et al. 2009) were indicated by red asterisks.

tained from the Uniprot database (Apweiler et al. 2004) and were aligned by Clustal Omega (Sievers et al. 2011). The conserved residues were boxed in blue using the JalView software (Waterhouse et al. 2009).

*Arabidopsis At*ENY2 and all its counterparts in other species were similar in their small size of approximately 100 amino acids and showed a high level of sequence similarity (**Figure 5.13**). The local similarity program SIM (Huang et al. 1991) revealed that the amino acid sequence of *Arabidopsis* ENY2 was 26.1%, 43.4% and 56.8% identical to its orthologs from yeast, fruit-fly and humans, respectively. This indicated that the plant ENY2 had the highest sequence identity to humans and least to yeast. The N-terminus of *At*ENY2 featured a short appendix (approximately 15 aa) that is conserved in plants (mosses, monocotyledons and dicotyledons), but did not appear in yeast or animals.

Comparing *Arabidopsis At*SGF11 to its orthologs from other species revealed size-differences at the C-terminus (**Figure 5.14**). *Pp*SGF11 and *Sc*SGF11 showed a shortened C-terminus whereas human and mouse SGF11 had a large extended C-terminus. Comparing the *Arabidopsis* SGF11 aa sequence to its relatives in yeast, *Drosophila* and humans, revealed a sequence identity of 37.2%, 41.7%, 44.4%, respectively. The predicted ENY2 binding site (Lang et al. 2011) and the Zink-finger domain showed a high level of sequence similarity in all tested species. The arginine residues in the basic ZnF domain that are important for docking the human DUBm on the H2A/H2B acidic patch (Morgan et al. 2016) were conserved as well.

UBP22 was the largest protein of the DUB module (**Figure 5.15**). Size-wise the differences between the analysed species were not remarkable with the exception of *Dm*Nonstop, that showed a large extended C-terminus compared to all other tested species. The SIM analysis showed that the aa sequence of *At*UBP22 is 30.6%, 33.7%, 34.7% identical to its orthologs from

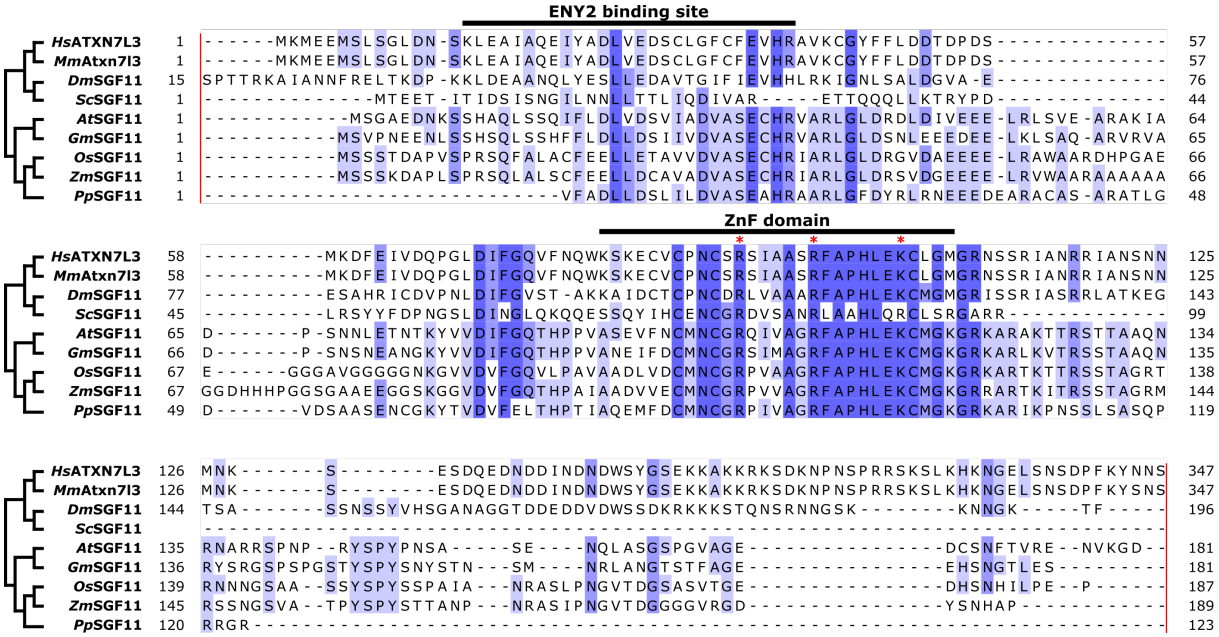


Figure 5.14 The amino acid sequence of the *Arabidopsis* SGF11 is conserved among different species. The SGF11 protein sequence was aligned to its orthologs from the following other species: *Homo sapiens* (*Hs*), *Mus musculus* (*Mm*), *Drosophila melanogaster* (*Dm*), *Saccharomyces cerevisiae* (*Sc*), *Glycine max* (*Gm*), *Oryza sativa* (*Os*), *Zea mays* (*Zm*) and *Physcomitrella patens* (*Pp*). The multiple sequence alignment was generated using Clustal Omega (Sievers et al. 2011). The conservation of the aa residues was highlighted with blue coloration and the trees were calculated using JalView (Waterhouse et al. 2009). In humans, the "ENY2 binding site" was ranging from K15 to R41 and the Zink finger domain (ZnF) from K80 to M108 (Lang et al. 2011). Both domains were indicated in the MSA by black bars. The arginine residues (R78, R84, R91) that were important for the histone H2A/H2B interaction (Morgan et al. 2016) were indicated by red asterisks. The first 15 aa of *Dm*SGF11 and the last 159 aa of *Hs*SGF11 and *Mm*SGF11 were cropped to obtain more clarity (Indicated by red lines).

yeast, *Drosophila* and humans, respectively. The predicted N-terminal Zink finger domain and the C-terminal USP domain showed a high level of sequence similarity. In particular also the residues C146, H427 and N443 that are forming the active center in yeast Ubp8 (Köhler et al. 2010).

Taken together, the subunits of the putative *Arabidopsis* DUB module were highly conserved across fungi, plants and animals. Beyond the plant kingdom, the putative *Arabidopsis* DUB components showed the highest sequence identity to higher eukaryotes and least to yeast.

5.4.4 Homology modelling of *Arabidopsis* DUB components

The protein structures of *Arabidopsis* ENY2, SGF11 and UBP22 were predicted using the I-TASSER (Iterative Threading ASSEmbly Refinement) server (Yang et al. 2015; **Figure 5.16 A**). To model the tertiary structure of *At*ENY2, the crystal structure of the human ENY2:GANP complex (Protein Data Bank: 4DHXB) at 2.1 *Å* resolution was used as structure template (Jani et al. 2012), because of the high sequence identity of the *Arabidopsis* and the human ENY2. The quality of the model was estimated as good, with a C-Score[1] of -0.6. For the modelling of *At*SGF11 and *At*UBP22, the crystal structure of the yeast Ubp8-Sgf11-Sus1-Sgf73 complex (Protein Data Bank: 3MHH) at 2.45 *Å* resolution was used as template (Samara et al. 2010). The quality of the SGF11 model was estimated as average (C-Score = -2.96) and the quality

[1]I-TASSER used the C-score to estimate the quality of a predicted model. Typically, the C-Score ranged from -5 to 2, where a high score indicated a high confidence of the model (Yang et al. 2015)

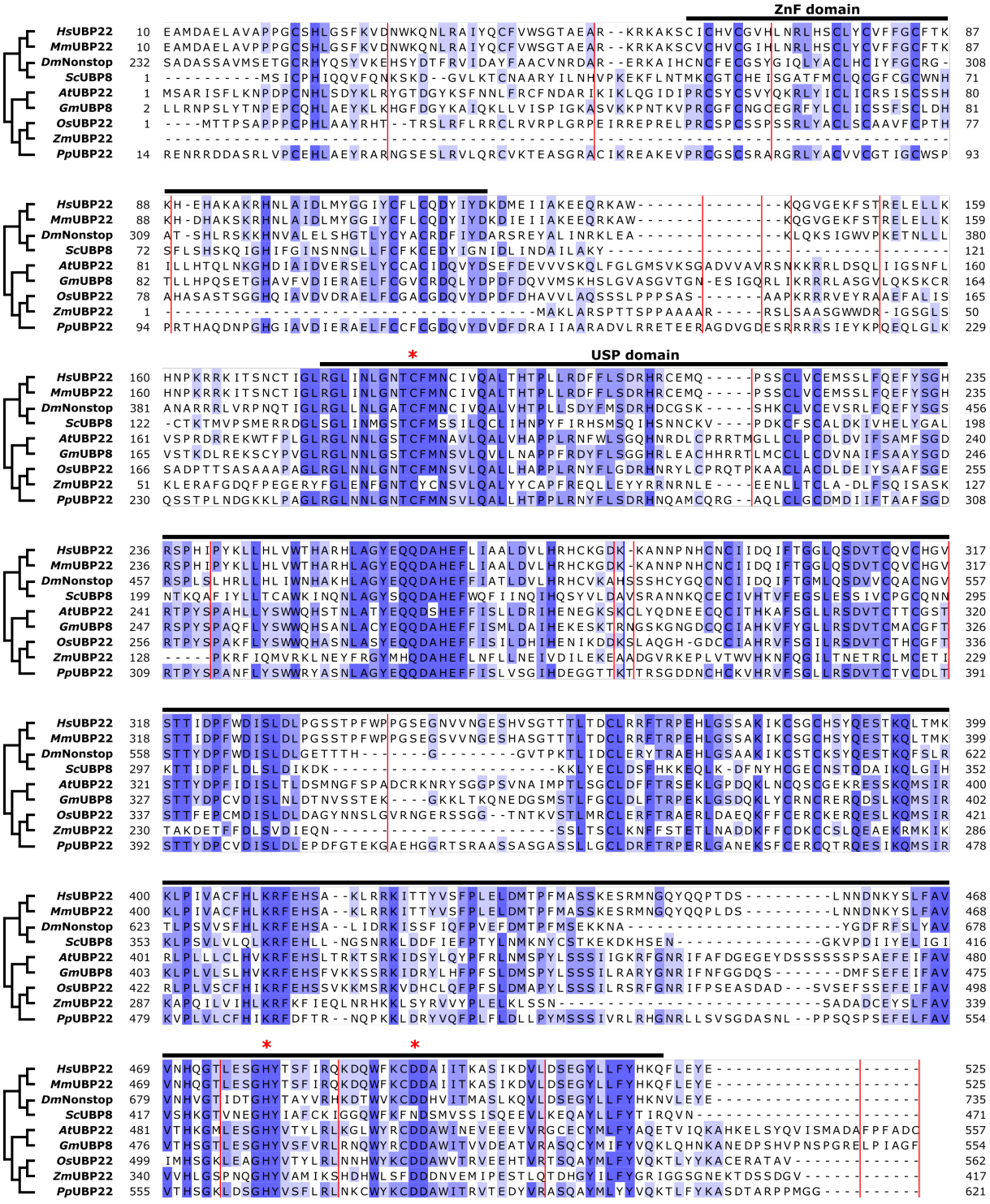


Figure 5.15 The amino acid sequence of the *Arabidopsis* UBP22 is conserved among different species.
The UBP22 protein sequence was aligned to its orthologs from the following other species: *Homo sapiens* (*Hs*), *Mus musculus* (*Mm*), *Drosophila melanogaster* (*Dm*), *Saccharomyces cerevisiae* (*Sc*), *Glycine max* (*Gm*), *Oryza sativa* (*Os*), *Zea mays* (*Zm*) and *Physcomitrella patens* (*Pp*). The multiple sequence alignment was generated using Clustal Omega (Sievers et al. 2011). The conservation of the aa residues was highlighted with blue coloration and the trees were calculated using JalView (Waterhouse et al. 2009). The "Zink finger" domain (ZnF) at the N-terminus and the "Ubiquitin specific hydrolase" domain (USP) at the C-terminus were indicated by black bars. The residues forming the active center in the yeast Ubp8 (Catalytic triad: C146, H427 and N443) were indicated (Köhler et al. 2010) by red asterisks. All columns containing gaps in the AtUBP22 sequence were hidden to obtain more clarity (Indicated by red lines).

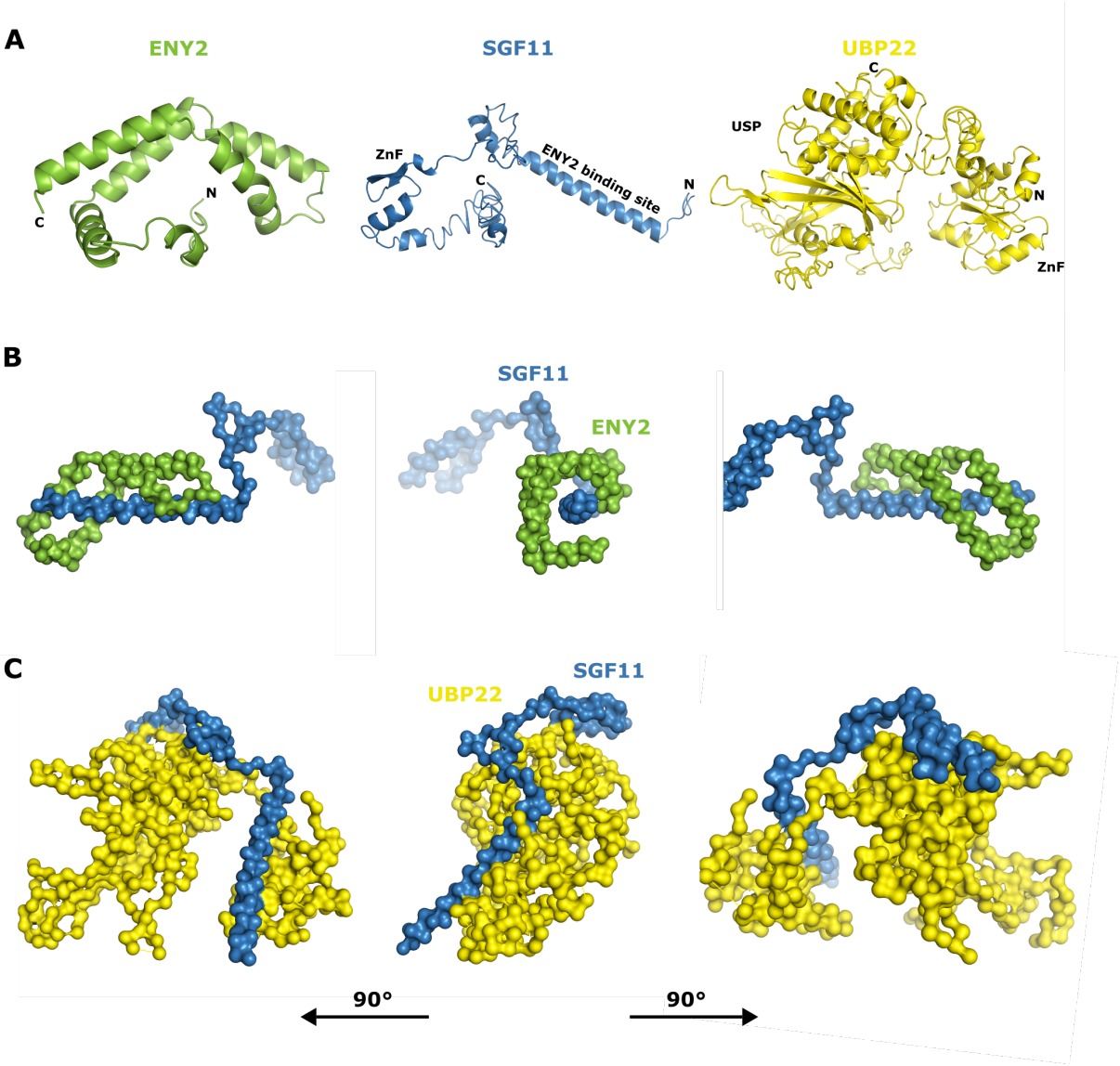


Figure 5.16 The homology models of *Arabidopsis* ENY2, SGF11 and UBP22.
A) The protein structures of *Arabidopsis* ENY2, SGF11 and UBP22 were predicted using the I-TASSER server (Yang et al. 2015). In SGF11, the "ENY2-binding site" and the ZnF domain were indicated. In UBP22, the ZnF domain and the USP domain were indicated. B) The *Arabidopsis* SGF11/ENY2 complex was predicted by SPRING (Guerler et al. 2013). ENY2 was depicted in green and SGF11 is pedicted in blue. C) The *Arabidopsis* SGF11/UBP22 complex was predicted by SPRING (Guerler et al. 2013). SGF11 was depicted in blue and UBP22 is depicted in yellow.

of the UBP22 model as good (C-Score = 0.21). The protein models showed that the tertiary structure of all three proteins resembled their well-characterized relatives in yeast and humans.

The protein-protein structure predictions of *Arabidopsis* AtSGF11 and AtENY2 as well as AtSGF11 and AtUBP22 were modelled by SPRING (Guerler et al. 2013). The model of *Arabidopsis* ENY2/SGF11 (**Figure 5.16 B**) showed that ENY2 wraps tighly around the long alpha-helix of SGF11. Thereby ENY2 binds to the predicted "ENY2 binding site" of SGF11. The structure prediction of the *Arabidopsis* SGF11/UBP8 complex (**Figure 5.16 C**) showed that both proteins were extensively interacting. SGF11 was binding to both, the USP domain as well as the ZnF domain of UBP22. The predictions for the assembly of *At*ENY2/*At*SGF11 as well as *At*SGF11/*At*UBP22 were in accordance to quaternary structures of the yeast DUB module that were determined by crystallography (Jani et al. 2009, Ellisdon et al. 2010, Köhler et al. 2010, Samara et al. 2010, Morgan et al. 2016).

5.5 Interaction of ENY2 with the NTC complex and splicing

In the previously described AP-MS experiments (**Section 5.3**), the NTC complex co-purified with the SG-tagged ENY2 and vice versa ENY2 co-purified with the SG-tagged MOS4 (NTC complex). This indicated a reciprocal interaction between ENY2 and the NTC complex in *Arabidopsis*. In the following, Y2H (**Section 5.5.1**) and FRET (**Section 5.5.2**) experiments were used to identify putative direct protein-protein interactions between ENY2 and the core NTC complex as well as some NTC-related proteins (NTR). Moreover, the asscociation of the small adaptor protein ENY2 with the spliceosome was analysed by co-localization experiments (**Section 5.5.3**).

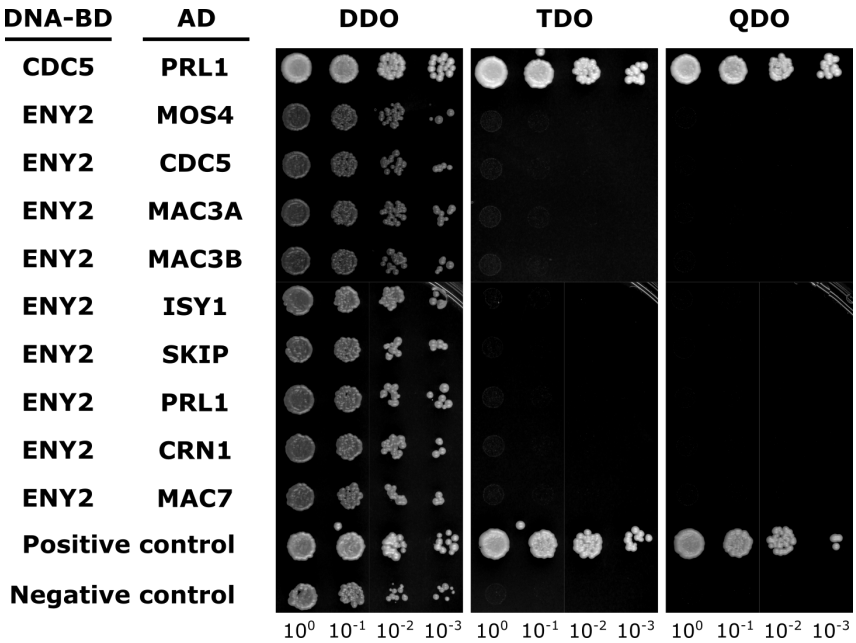


Figure 5.17 Y2H showed no direct interaction between ENY2 and NTC/NTR components.
Different combinations of DNA-BD/bait and AD/prey fusion proteins were spotted on DDO (SD/ -LEU -TRP), TDO (SD/- LEU -TRP -HIS) and QDO (SD/ -LEU -TRP -HIS -ADE) plates in serial dilutions. DNA-BD/murine p53 and AD/SV40 large T-antigen function as positive control. DNA-BD/Lamin and AD/T act as negative control.

5.5.1 No direct PPI between ENY2 and NTC/NTR components was detected by Y2H

As first approach, the yeast two-hybrid (Y2H) system was used as described in **section 5.4.1** to test direct interactions between ENY2 and splicing factors. According to this experiment, there was no direct interaction between ENY2 and the following tested NTC/NTR components (**Figure 5.17**): MOS4, CDC5, MAC3A, MAC3B, ISY1, SKIP, PRL1, CRN1, MAC7 (NTR). All these combinations were growing on DDO plates, but neither on TDO nor on QDO plates. Additionally, the interaction of NTC components among each other was tested. Therefore, yeast cells were co-transformed with DNA-BD/CDC5 and AD/PRL1 plasmids. The growth on DDO as well as TDO and QDO plates demonstrated interaction of the two NTC components CDC5 and PRL1.

5.5.2 No direct PPI between ENY2 and NTC/NTR components was detected by FRET

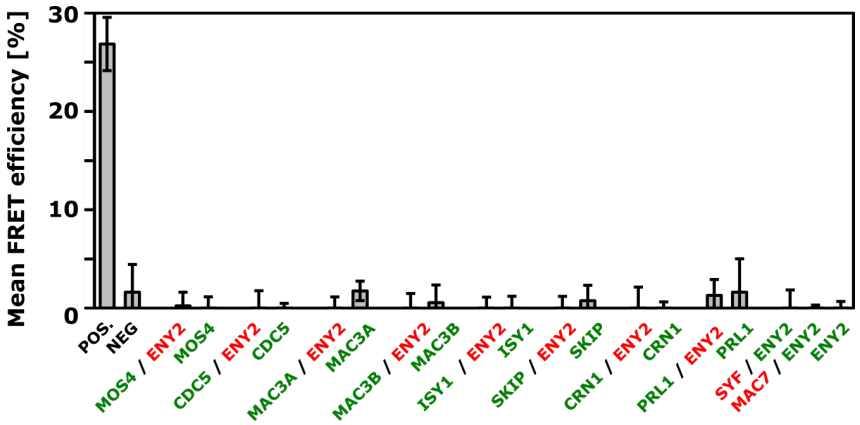


Figure 5.18 FRET revealed no direct interaction between ENY2 and components of the NTC/NTR complex. To test the interactions between ENY2 and the NTC/NTR complex, corresponding donor/acceptor (eGFP fusions/mCherry fusions) combinations were transiently co-expressed in the epidermal cells of *N. benthamiana* and FRET-APB was measured using CLSM. The mean FRET efficiencies ± SD of all test combinations and the respective controls are shown. As positive control, cells expressing eGFP-NLS-mCherry fusion proteins were used.

In the following, interactions between ENY2 and NTC/NTR components were examined in the plant system using the established FRET-APB method as described in **section 5.4.2**. Therefore, the coding sequences of the NTC/NTR components and *ENY2* were cloned into plant expression vectors to generate the following translational fusions with eGFP (MOS4, CDC5, MAC3A, MAC3B, ISY1, SKIP, CRN1, PRL1, ENY2) and mCherry (SYF1, MAC7, ENY2).

In the positive control, the bleaching of the acceptor led to an increase of the donor fluorescence, that could be measured as a mean FRET efficiency of approximately 25%. In all donor/acceptor combinations that were used to test the interaction of ENY2 with the NTC/NTR components, no FRET was detectable (**Figure 5.18 and S12**). The measured FRET efficiencies of the putative interactors were not higher than the negative controls (approximately 0-5%). This indicated that the putative interactors brought their fused fluorophores in insufficient proximity (> 10 nm) to induced FRET and to be detected. Hence, it was not possible to show an interaction between ENY2 and the NTC/NTR complex by this FRET setup.

5.5.3 ENY2 and the NTC component MOS4 co-localized in splicing-speckles in the nucleoplasm

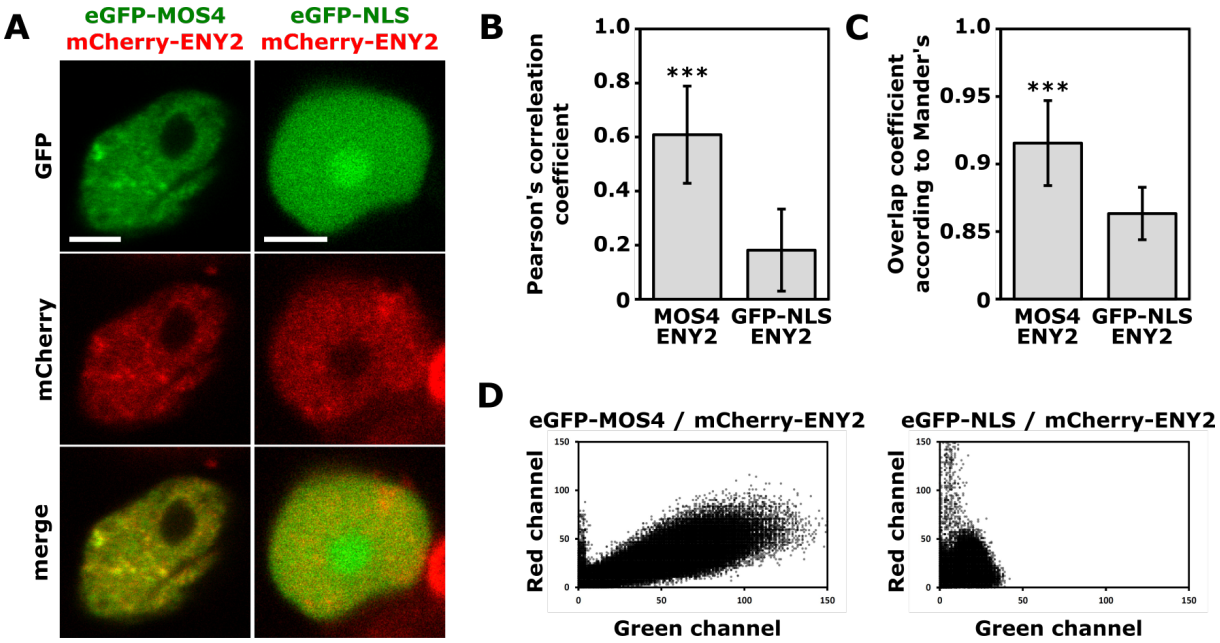


Figure 5.19 Co-localization studies indicated linkages of ENY2 with mRNA splicing.
A) The Co-Localization of mCherry-ENY2 and eGFP-MOS4 (NTC core component) fusion proteins was analysed in transiently transformed epidermal cells of *N. benthamiana* using CLSM. The cells expressing mCherry-ENY2 and eGFP-NLS functioned as negative control. From representative images, the Z-stacks of the green (eGFP) and the red (mCherry) channels as well as a merge of both are shown. The scale bars indicate 5 μm. B-C) To quantify the degree of co-localization between ENY2 and MOS4 as well as ENY2 and the negative control, coefficients were calculated: B) Pearson′s correlation coefficient (PCC) C) Overlap coefficient according to Mander′s D) To visualize the degree of co-localization between ENY2 and MOS4 as well as ENY2 and the negative control, scatterplots were generated.

Although no direct interaction was found in Y2H and FRET analyses, a possible co-localization of ENY2 and the core NTC component MOS4 was investigated. Plant expression vectors were generated, to drive the expression of ENY2 and MOS4 as translational fusions with fluorescent proteins under the 35S promoter (eGFP-MOS4 and mCherry-ENY2). Via *Agrobacterium*, the mCherry-ENY2 plasmid was transiently co-expressed with the eGFP-MOS4 or the eGFP-NLS plasmids (negative control) in *N. benthamiana* leaves. The epidermal cells that were expressing both combinations (mCherry-ENY2/eGFP-MOS4 or mCherry-ENY2/eGFP-NLS) were analysed using the confocal microscopy (**Figure 5.19 A**). Z-stacks of nuclei revealed that the fluorescent signal of the eGFP-NLS control was homogenously distributed across the whole nucleus and accumulated in the nucleolus. In contrast, both ENY2 and MOS4 were just localized in the nucleoplasm and showed speckle-like structures in this compartment. It was shown before that splicing proteins often accumulate in splicing speckles (Fang et al. 2004, Tillemans et al. 2005). Superimposing of the GFP and the mCherry fluorescence signals demonstrated that there was a clear co-localization of ENY2 and MOS4. Best visible in the overlapping nuclear speckle structures, that were clearly absent in the negative control (eGFP-NLS).

To characterize the degree of overlap between GFP and mCherry signals, coefficients were determined using the JACoP plugin for ImageJ (Bolte et al. 2006). First, the Pearson′s correlation coefficient (PCC) was calculated to describe the linear correlation of both fluorescence

signals (eGFP and mCherry). This value can range from 1 (complete correlation) to -1 (complete exclusion), with zero standing for no correlation (Manders et al. 1993). Cells, expressing fluorescently tagged ENY2 and MOS4 showed a mean PCC of 0.6 (**Figure 5.19 B**). In contrast, cells expressing eGFP-NLS and mCherry-ENY2 showed a significantly lower PCC of 0.18. This indicated a positive correlation for the mCherry and eGFP signals in the case of MOS4/ENY2.

As alternative approach to quantify the degree of co-localization between eGFP and mCherry, the Mander′s overlap coefficient (MOC) was calculated (Manders et al. 1993). The MOC is calculated like the PCC, but the mean intensity values of both channels are taken out of the expression. This value can range from 0 (No overlapping pixels) to 1 (100 % overlapping pixels). The MOC for eGFP-MOS4 and mCherry-ENY2 was 0.91 (**Figure 5.19 C**). The MOC for eGFP-NLS and mCherry-ENY2 (negative control) was significantly lower with 0.86.

Moreover, a scatterplot (Cytofluorogram) is a simple way to visualize the co-localization of two fluorescent signals (Bolte et al. 2006). For each pixel of the image, the green (X-coordinate) and red channel intensities (Y-coordinate) were plotted against each other. In case of a complete co-localization, for each pixel (shown as dot on the diagram) the signal intensities of the red and the green channels are the same and a cloud of dots will position on a straight line following the linear function f(x) = x. In case of a complete exclusion, the dots are positioned along the axes. In the scatterplot of eGFP-MOS4 and mCherry-ENY2, the dots were distributed on a straight line with a deflection towards the X-coordinate (**Figure 5.19 D**). This indicated a high degree of co-localization with some differences in the fluorescence intensity. In contrast, in case of eGFP-NLS and mCherry-ENY2, the dots were not arranged along a linear line, but were positioned at the origin of the diagram (basal noise) and along the Y-coordinate. In summary, the co-purification (**Section 5.3**) in combination with the co-localization between ENY2 and MOS4 (NTC complex) suggested an involvement of ENY2 in the splicing of mRNA.

5.6 The SAGA complex and SAGA-DUB module in plants

The SAGA complex, consists of four sub-complexes with distinct functions: The DUB module (deubiquitination), the HAT module (histone acetyltransferase), the SPT module (pre-initiation complex assembly) and the TAF module (SAGA architecture) (Koutelou et al. 2010). Surprisingly, although the existence of the *Arabidopsis* SAGA complex was predicted bioinformatically (Moraga et al. 2015, Srivastava et al. 2015), no components of the HAT, the SPT or the TAF module were co-purified with *Arabidopsis* SGF11 or ENY2 as bait proteins in AP-MS experiments (**Section 5.3**). In yeast and humans, SGF73 was identified as a linker between the DUB module and the SAGA complex (Köhler et al. 2008, Lee et al. 2009). Moreover, deletions of Sgf73 disconnected the DUB module from the co-activator (Durand et al. 2014). In *Arabidopsis*, all DUB components (ENY2, SGF11, UBP22) are conserved except for SGF73 (Moraga et al. 2015, Srivastava et al. 2015). This opens the possibility, that the DUB module is not permanently linked to the remaining SAGA complex in plants.

5.6.1 AP-MS analysis of the *Arabidopsis* SAGA complex

To verify the existence of the SAGA complex in plants biochemically and to figure out if the DUBm components co-purify with other SAGA modules, further AP-MS experiments

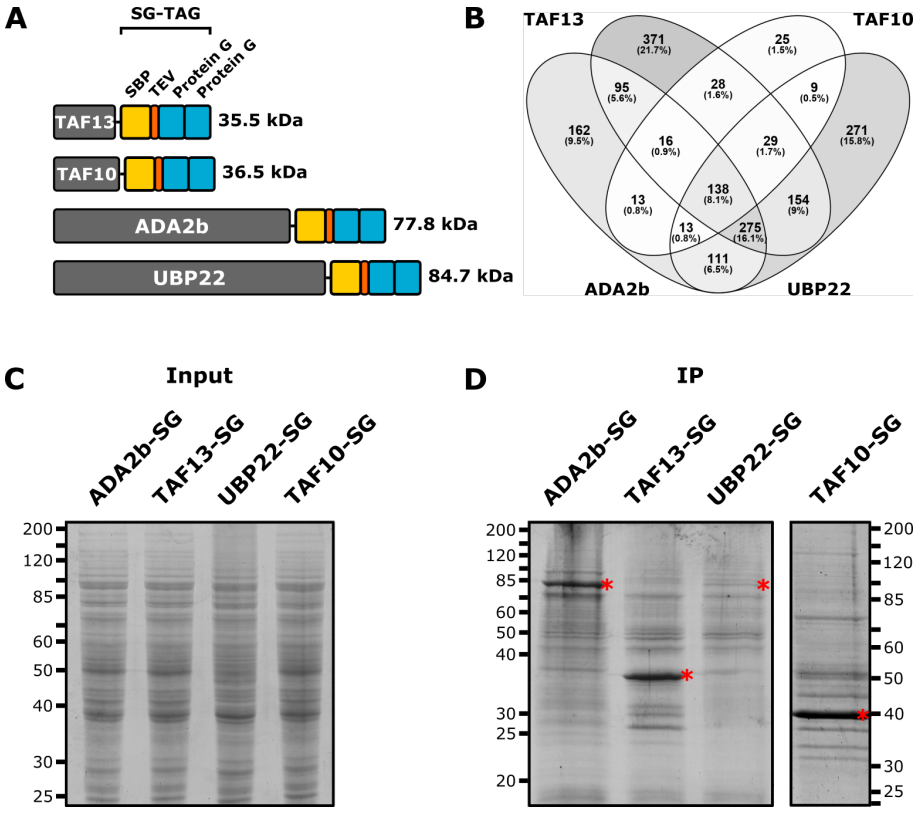


Figure 5.20 AP-MS revealed that the SAGA complex is assembled in *Arabidopsis*.
A) Schematic illustration of the bait proteins C-terminally fused to a SG tag (Streptavidin-binding protein and two Protein G domains linked by TEV cleavage site). The molecular weight of the fusion proteins was computed by the ExPaSy pI/Mw tool. B) The overlap of the proteins co-purified with the four bait proteins were plotted in a Venn diagram. C) Total protein extracts (Input, 0.01 % of total) of transgenic cells expressing the bait proteins and the SG-tag only. D) Eluates (50 % of total) of the one-step affinity purifications (AP) using IgG coupled magnetic beads. The proteins were separated by 9 % SDS-PAGE and the gels were stained with Coomassie Blue. Red asterisks indicate the bands corresponding to the unfused SG-tag and the bait proteins. The sizes of the molecular weight marker proteins are shown in kDa.

were conducted as described before (**Section 5.3**). Briefly summarized, ADA2b, SPT3/TAF13, TAF10 and UBP22 were chosen as representatives for the HAT, the SPT, the TAF and the DUB module, respectively (**Figure 5.20 A**). The SG-fusions of these proteins were expressed in *Arabidopsis* suspension cell culture and the interacting proteins were co-purified by one-step affinity purification using IgG coupled magnetic beads. To remove DNA- and RNA-mediated interactions and to enrich for purification of chromatin proteins, cell extracts were treated with benzonase. Following SDS-PAGE and Coomassie staining of the AP eluates, the bait proteins ADA2b-SG, SPT3/TAF13-SG, and TAF10-SG were clearly visible as most prominent bands (**Figure 5.20 C-D**). In contrast, UBP22-SG appeared as one of several equally strong bands on the stained gel. Further co-purified proteins appeared in all four APs with lower intensities over the whole lanes. Following in-gel trypsin digest, the co-purified candidate interactors were identified by mass spectrometry and the raw MS data were processed as described before in Section 5.3 and in Pfab et al. 2017.

In total, after processing the MS data, 823 proteins were co-purified with ADA2b-SG, 1,106 proteins with SPT3/TAF13-SG, 271 with TAF10-SG and 1,000 with UBP22-SG (**Figure 5.20 B**). In total, 8.1 % of the identified proteins were co-purified with all four SG-tagged proteins.

A substantial portion (16.1 %) was only detected in the APs of ADA2b, SPT3/TAF13, and UBP22. The amount of unique proteins that only co-purified with one of the bait proteins was prominent for SPT3/TAF13 (21.7 %), UBP22 (15.8 %), and ADA2b (9.5 %) and almost non-existent for TAF10 (1.5 %). Overall, the four bait proteins ADA2b, SPT3/TAF13, TAF10 and UBP22 shared a huge set of commonly co-purified proteins.

The analysis of the AP-MS results revealed some interesting findings about the plant SAGA complex. The DUB module co-purified very reproducible with SG-ENY2, SG-SGF11 and SG-UBP22, but almost none of the other predicted SAGA components were identified in these APs (**Table 5.4**). The proteins TRA1a, ADA2b and GCN5 were identified as interaction partners of UBP22 and the two proteins TRA1a and TRA1b as interaction partner of SGF11. On the other hand, all predicted HAT and TAF subunits as well as five of the seven SPT subunits were co-purified with ADA2b (HAT module). From the DUB module, just ENY2 co-purified in two out of three AP with ADA2b. Five of seven components of the TAF as well as the SPT module co-purified with SPT3/TAF13 (SPT module), but just one HAT-subunit (GCN5) and no DUB subunit. As well, five of seven SPT subunits and six of seven TAF subunits co-purified with TAF10 (TAF module), but only two components of the HAT module (ADA2b and GCN5) and no DUB subunit. Interestingly, the predicted SAGA-associated protein CHD1/CHR5 (AT2G13370) was not found in any of the SAGA APs, but the CHD1-like protein (AT2G25170) co-purified with SPT3/TAF13 and UBP22.

In summary, all of the bioinformatically predicted SAGA components except HAF2 (AT3G19040) were biochemically identified by this AP-MS approach, which validates the capacity of this experimental setup. The three subunits of the DUB module (ENY2, SGF11, UBP22) co-purified almost exclusively just with each other and not with the proteins of the other SAGA modules (HAT, TAF or SPT). Moreover, the complete HAT module only co-purified with ADA2b and not with components of the DUB, TAF or SPT module. This suggest that the SAGA complex is conserved in *Arabidopsis*, but its architecture and module composition may differ from its orthologs in yeast or humans. A figure showing the composition of the *Arabidopsis* SAGA complex basing on the AP-MS data is shown in the discussion (**8.1**).

As known from literature (Lawit et al. 2007, Srivastava et al. 2015), the co-activators SAGA and TFIID have several subunits in common including all seven components of the SAGA-TAF module and two components of the SAGA-SPT module (SPT3 and SPT7). Therefore, the two bait proteins SPT3/TAF13 and TAF are most likely components of the SAGA as well the TFIID complex as discussed later (**Section 8.2.4**). Next to SAGA-specific proteins and shared components of SAGA and TFIID, six TFIID-specific proteins were robustly co-purified with SPT3/TAF13 and TAF10.

Several subunits of the RNAPII and transcript elongation factors including SPT16 or SPT4/5 were co-purified with the putative *Arabidopsis* SAGA subunits ADA2b, SPT3/TAF13, SGF11 and UBP22 (**Table 5.4**). Additionally, several components of the Mediator complex, a large protein complex that globally regulates the transcription by RNAPII (Allen et al. 2015), were co-purified with ADA2b, SGF11 and especially SPT3. These findings are in accordance with the role of the yeast/human SAGA complex (Weake et al. 2012) and suggest that the plant SAGA complex may function as transcriptional coactivator.

Table 5.4 Transcription-related proteins co-purifying with ADA2b-SG, SPT3/TAF13-SG, TAF10-SG and UBP22-SG

Numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent AP′s. The proteins detected in less than two out of three AP were not listed.

AGI	Protein	Complex	ADA2b	SPT3	TAF10	ENY2	SGF11	UBP22
AT3G27100	ENY2	SAGA_DUBm	154 / 2			783 / 3	1262 / 3	611 / 3
AT5G58575	SGF11	SAGA_DUBm				542 / 3	1747 / 3	326 / 3
AT5G10790	UBP22	SAGA_DUBm				113 / 3	209 / 3	1858 / 3
AT4G16420	ADA2b	SAGA_HAT	5515 / 3		446 / 3			127 / 2
AT4G29790	ADA3	SAGA_HAT	489 / 3					
AT3G54610	GCN5	SAGA_HAT	6257 / 3	159 / 3	726 / 3			283 / 2
AT3G27460	SGF29a	SAGA_HAT	412 / 3					
AT5G40550	SGF29b	SAGA_HAT	185 / 3					
AT2G14850	ADA1a	SAGA_SPT	678 / 3		163 / 3			
AT5G67410	ADA1b	SAGA_SPT	151 / 3					
AT1G72390	SPT20	SAGA_SPT	4499 / 3	129 / 2	1606 / 3			
AT2G17930	TRA1a	SAGA_SPT	4304 / 3	1804 / 3	1124 / 3		270 / 2	482 / 3
AT4G36080	TRA1b	SAGA_SPT	2564 / 3	1498 / 2	712 / 3		160 / 2	
AT1G02680	SPT3, TAF13	SAGA_SPT / TFIID		612 / 3				
AT1G32750	SPT7, TAF1	SAGA_SPT / TFIID		5178 / 3	1727 / 3			
AT5G25150	TAF5	SAGA_TAF /TFIID	1971 / 3	2393 / 3	1786 / 3			
AT1G04950	TAF6	SAGA_TAF /TFIID	130 / 2	1573 / 3	767 / 3			
AT1G54360	TAF6b	SAGA_TAF /TFIID	865 / 3		963 / 3			
AT1G54140	TAF9	SAGA_TAF /TFIID	539 / 3	481 / 2	545 / 2			
AT4G31720	TAF10	SAGA_TAF /TFIID	378 / 3	195 / 2	1144 / 3			
AT1G17440	TAF12b	SAGA_TAF /TFIID	687 / 3					
AT3G10070	TAF12	SAGA_TAF /TFIID	189 / 3	761 / 3	273 / 3			
AT2G25170	CHD1-like	SAGA_Others		473 / 2				360 / 2
AT3G13445	TBP1	TFIID		283 / 2	211 / 2			
AT1G55520	TBP2	TFIID		354 / 2	208 / 2			
AT1G73960	TAF2	TFIID		2595 / 3	600 / 3			
AT5G43130	TAF4	TFIID		2919 / 3	1046 / 3			
AT4G34340	TAF8	TFIID		599 / 3	1080 / 3			
AT4G20280	TAF11	TFIID		1310 / 3	411 / 3			
AT3G04740	Med14	Mediator		756 / 3				
AT1G15780	Med15	Mediator		406 / 2				
AT4g04920	Med16	Mediator	265 / 2	480 / 3			171 / 2	
AT5G20170	Med17	Mediator		148 / 2				
AT2G22370	Med18	Mediator		110 / 2				
AT2G28230	Med20	Mediator		131 / 2				
AT1G23230	Med23	Mediator		615 / 2				
AT5G02850	Med4	Mediator	110 / 2					
AT4G35800	NRPB1	RNAPII	374 / 2	848 / 3			328 / 3	578 / 3
AT4G21710	NRPB2	RNAPII	504 / 3	1010 / 3			330 / 3	554 / 2
AT3G22590	CDC73	PAF-C					282 / 2	451 / 3
AT5G61150	LEO1	PAF-C	230 / 2	155 / 2			225 / 2	
AT1G79730	ELF7	PAF-C	176 / 3	184 / 2			167 / 2	210 / 3
AT4G29830	SKI8	PAF-C	215 / 3	192 / 3			417 / 3	383 / 3
AT4G10710	SPT16	FACT	136 / 2				215 / 2	
AT1G32130	IWS1a	IWS1	330 / 2					313 / 2
AT5G63670	SPT4-2	SPT4/SPT5					137 / 2	
AT4G08350	SPT5-2	SPT4/SPT5	331 / 3	456 / 3			327 / 3	214 / 3
AT1G65440	SPT6L	SPT6		662 / 2			209 / 2	249 / 2

Table 5.5 Splicing-related proteins co-purifying with ADA2b-SG, SPT3/TAF13-SG, TAF10-SG and UBP22-SG
Numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent AP's. The proteins detected in less than two out of three AP were not listed.

AGI	Protein	Process	ADA2b	SPT3/TAF13	TAF10	ENY2	SGF11	UBP22
AT1G04080	Prp39a	U1 snRNP	315 / 2	300 / 3				268 / 3
AT2G47580	atU1A	U1 snRNP	248 / 3			342 / 3	472 / 2	158 / 3
AT3G50670	atU1-70K	U1 snRNP	182 / 3	293 / 3		166 / 3	210 / 3	233 / 3
AT4G03120	U1C	U1 snRNP	108 / 3					
AT5G17440	Luc7b	U1 snRNP		160 / 2				399 / 3
At3g03340	Luc7a	U1 snRNP						239 / 3
AT5G51410	Luc7	U1 snRNP						218 / 3
AT1G60200	AT1G60200	U1 snRNP-rel.		454 / 3				269 / 3
AT1G44910	PRP40A	U1 snRNP-rel.		256 / 2				
AT2G30260	U2B-b	U2 snRNP	110 / 2	104 / 2		423 / 3	304 / 3	171 / 2
AT1G09760	atU2A-	U2 snRNP	228 / 2			418 / 3	489 / 3	352 / 3
AT2G18510	EMB2444	U2 snRNP	246 / 2	202 / 2		139 / 2	154 / 3	323 / 3
AT3G55200	SAP130a	U2 snRNP	3556 / 3	4345 / 3	3105 / 3	637 / 3	859 / 3	1428 / 3
AT5G64270	SAP155	U2 snRNP	1287 / 3	1674 / 3		866 / 3	902 / 3	1408 / 3
At1g14650	SAP114-1a	U2 snRNP	525 / 3	256 / 2		350 / 2		871 / 3
AT4G21660	SF3b150	U2 snRNP	411 / 3	365 / 3		157 / 2	243 / 2	851 / 3
AT5G06160	SAP61	U2 snRNP	208 / 3	213 / 2		205 / 2		238 / 3
AT2G32600	SAP62	U2 snRNP	149 / 3			199 / 2		251 / 3
AT5G12190	P14-1	U2 snRNP	264 / 3	143 / 2		267 / 3	330 / 2	229 / 3
AT1G30480	SPF45	U2 snRNP-assoc.	162 / 2	325 / 2				302 / 3
AT5G25060	SR140-1	U2 snRNP-assoc.	483 / 3	1302 / 3			253 / 3	635 / 3
AT2G47330	PRP5-2	U2 snRNP-assoc.	144 / 3	275 / 3				353 / 3
AT2G02570	SPF30	U2 snRNP-assoc.						246 / 2
AT2G38730	Tri-20	U4/U6 snRNP	520 / 3		236 / 2	294 / 2	229 / 2	462 / 3
AT2G41500	EMB2776	U4/U6 snRNP	700 / 3	819 / 2	200 / 3		115 / 2	1150 / 3
AT1G60170	EMB1220	U4/U6 snRNP	278 / 3	207 / 3	166 / 3	97 / 2		470 / 3
At1g28060	SAP90-1	U4/U6 snRNP		494 / 2	149 / 2			652 / 3
AT5G16780	MDF	U4/U6.U5 tri snRNP	2012 / 3	1710 / 3	626 / 3			2379 / 3
AT3G05760	AT3G05760	U4/U6.U5 tri snRNP	247 / 3	191 / 2				196 / 2
AT2G40650	AT2G40650	U4/U6.U5 tri snRNP		238 / 3				329 / 3
AT2G43770	U5-40	U5 snRNP	545 / 2	308 / 3	166 / 3	172 / 3	284 / 2	742 / 3
AT2G33730	U5-100KD	U5 snRNP	736 / 3	726 / 3				810 / 3
AT4G03430	EMB2770	U5 snRNP	629 / 3	1518 / 3	309 / 2	190 / 3	695 / 3	724 / 3
AT1G20960	EMB1507	U5 snRNP	1829 / 3	3728 / 3	307 / 3	3102 / 3	1650 / 3	4541 / 3
AT5G61140	U5-200-1	U5 snRNP		448 / 3			453 / 2	200 / 3
AT5G09390	AT5G09390	U5 snRNP		162 / 2	151 / 2			
At1g60900	U2AF65b	Splice site sel.	116 / 2	93 / 2		122 / 2	180 / 2	161 / 3
AT1G27650	U2AF35a	Splice site sel.	125 / 2	218 / 2			125 / 2	115 / 3
AT4G36690	U2AF65a	Splice site sel.	234 / 3	433 / 3			127 / 2	357 / 3
AT5G42820	U2AF35	Splice site sel.		155 / 2			150 / 2	93 / 2
AT5G59160	TOPP2	A complex-assoc.	179 / 2	319 / 2				175 / 3
AT3G19590	BUB3.1	A complex-assoc.	393 / 3	431 / 3			156 / 2	605 / 3
AT3G54230	SUA	A complex-assoc.						223 / 2
AT1G29220	AT1G29220	A complex-assoc.						320 / 3
At1g30970	SUF4	A complex-assoc.						151 / 3
AT5G38840	AT5G38840	A complex-assoc.		126 / 2				
AT1G53720	CYP59	B complex-assoc.	188 / 2	240 / 3				252 / 2
AT4G09980	EMB1691	B complex-assoc.	251 / 2	306 / 3			106 / 2	310 / 3
AT4G31120	SKB1	B complex-assoc.		174 / 2				
AT4G08580	AT4G08580	B complex-assoc.	521 / 3	343 / 2	159 / 3			675 / 3
AT5G67530	PUB49	Bact complex	211 / 3	207 / 3				257 / 3
AT3G57910	AT3G57910	Bact complex						221 / 3
AT1G05460	SED3	Bact complex		692 / 3			338 / 3	
At5g09880	At5g09880	Bact complex	200 / 2	155 / 3				241 / 3
AT2G16940	AT2G16940	Bact complex	688 / 3	797 / 3		130 / 2	186 / 3	978 / 3
AT1G73720	SMU1	Bact complex	248 / 3	134 / 2				479 / 3
AT1G18080	ATARCA	C complex	190 / 2	346 / 2			250 / 2	221 / 3
AT1G61620	AT1G61620	C complex	120 / 2	171 / 2		130 / 2	193 / 2	297 / 3
AT4G18465	DDX35	C complex	628 / 3	427 / 2		166 / 2	112 / 3	741 / 3
AT5G64730	AT5G64730	C complex	282 / 3	232 / 3		106 / 2	95 / 2	265 / 3
AT5G51280	AT5G51280	C complex	224 / 3			181 / 2	85 / 2	352 / 2
AT2G21150	XCT	C complex		123 / 2		102 / 2		234 / 3

Table 5.3 (Continuation)

AGI	Protein	Process	ADA2b	SPT3/TAF13	TAF10	ENY2	SGF11	UBP22
AT3G44600	CYP71	C complex						173 / 2
AT2G20330	AT2G20330	C complex-assoc.	183 / 2					133 / 3
AT3G20550	DAWDLE	RES complex					171 / 3	128 / 2
AT3G26560	PRP22-1	RES complex	669 / 2	383 / 2		469 / 3		503 / 2
AT5G13010	EMB3011	RES complex	828 / 3	1166 / 3		104 / 3	279 / 3	968 / 2
AT2G44200	AT2G44200	RES complex		193 / 2				156 / 2
AT1G03140	PRP18-1	RES complex						232 / 3
AT4G37120	SLU7-1b	RES complex						161 / 3
AT1G26370	PRP22-2	RES complex		114 / 2				
AT3G62310	PRP43-2a	RES complex	427 / 3	374 / 3		252 / 2	221 / 3	942 / 3
AT1G17070	AT1G17070	RES complex		145 / 2				240 / 2
AT2G47250	PRP43-2b	RES complex		300 / 2			147 / 3	766 / 3
AT3G18790	ISY1	Core NTC	371 / 2	154 / 2		288 / 3		612 / 3
AT1G09770	CDC5	Core NTC	906 / 3	482 / 3		613 / 3	279 / 2	1288 / 3
AT5G28740	SYF1	Core NTC	508 / 3	359 / 3		613 / 3		753 / 3
At1g77180	SKIP	Core NTC	451 / 3	583 / 3		280 / 3	247 / 2	856 / 3
AT2G33340	MAC3B	Core NTC	404 / 3	219 / 3		339 / 3	269 / 2	603 / 3
AT1G04510	MAC3A	Core NTC	243 / 3	173 / 3		435 / 2	297 / 2	570 / 3
AT5G41770	CRN1c	Core NTC	243 / 3	198 / 3		346 / 3		303 / 3
AT4G15900	PRL1	Core NTC	186 / 3	159 / 2		277 / 3		416 / 3
AT3G18165	MOS4	Core NTC	133 / 3			174 / 3	127 / 2	135 / 3
AT2G16860	SYF2	NTR	145 / 2			108 / 2		113 / 2
AT5G45990	CRN1a	NTR		162 / 3				
AT1G10580	PRP17-1	NTR	162 / 2			155 / 2		265 / 3
AT5G23590	AT5G23590	NTR	123 / 2					
AT4G34870	ROC5	NTR	265 / 2					
AT1G32490	PRP2a	NTR	98 / 2			221 / 2	97 / 2	279 / 2
AT2G38770	EMB2765	NTR	1238 / 3	1231 / 3	331 / 2	1072 / 3	554 / 3	1842 / 3
AT1G56290	AT1G56290	NTR	258 / 3	341 / 3	92 / 2			475 / 3
AT2G36130	AT2G36130	NTR	168 / 3			280 / 3	137 / 2	208 / 3
AT1G07360	ECM2-1a	NTR	288 / 3	447 / 3	309 / 3	315 / 3	299 / 3	547 / 3
AT4G21110	Bud31	NTR	178 / 3					177 / 3
AT4G33060	CWC27	NTR						94 / 2
AT1G20580	SmD3-b	Sm core	151 / 2			298 / 2	233 / 2	172 / 2
AT2G47640	SmD2-a	Sm core	222 / 2					507 / 2
AT3G07590	SmD1-a	Sm core	391 / 3			524 / 3	439 / 2	367 / 3
AT4G02840	SmD1-b	Sm core	424 / 3	227 / 2		565 / 2	398 / 2	403 / 3
AT1G76300	SmD3-a	Sm core	198 / 3			244 / 3	276 / 2	204 / 3
AT2G18740	SmE-b	Sm core	156 / 2			331 / 3	341 / 2	107 / 2
AT1G03330	LSM2	Lsm core					115 / 2	114 / 2
AT3G49430	SR34a	SR	110 / 2	92 / 2		225 / 2		201 / 2
AT4G31580	SRZ22	SR	159 / 3	347 / 2		424 / 2	391 / 3	189 / 3
AT3G61860	RSp31	SR	129 / 3	242 / 2			250 / 3	138 / 2
AT1G23860	RSzp21	SR	140 / 3			204 / 3	124 / 3	134 / 3
AT2G24590	RSZp22a	SR				264 / 3	262 / 3	131 / 2
AT3G55460	SCL30	SR				215 / 3	196 / 3	168 / 2
AT2G46610	RSp32	SR		131 / 2		348 / 3	166 / 3	134 / 3
At5g22840	SRPK2a	SR kinase	153 / 2					92 / 2
AT3G44850	SRPK2c	SR kinase	279 / 3	107 / 2		121 / 2		203 / 3
AT1G60650	hnRNP-G3	hnRNP family	189 / 2	195 / 3		137 / 3		243 / 3
AT2G44710	hnRNP-R3	hnRNP family	234 / 2				415 / 3	
AT3G04610	FLK	hnRNP family	91 / 2	150 / 2				
AT5G04280	hnRNP-G1	hnRNP family	139 / 2					
AT3g26420	ATRZ-1A	hnRNP family	201 / 3	253 / 2		272 / 3	166 / 3	343 / 2
AT5G54900	ATRBP45A	hnRNP family					135 / 2	137 / 2
AT3G19130	AtRBP47b	hnRNP family						188 / 2
AT3G15010	UBA2c	hnRNP family					190 / 2	96 / 2
AT2G22090	UBA1A	hnRNP family						125 / 2
AT1G02140	HAP1	EJC/mRNP	276 / 2			431 / 2		237 / 2
AT3G58570	AT3G58570	EJC/mRNP	720 / 3	832 / 3	327 / 3	219 / 2	477 / 3	644 / 3
AT2G45640	SAP18	EJC/mRNP	262 / 3			337 / 3	285 / 2	261 / 3

Strikingly, all the investigated SAGA subunits except TAF10 showed a strong association with the splicing machinery (**Table 5.5**). Most prominent, the NTC complex, several subunits of the U2/U5 snRNPs as well as the C complex were co-purifed along with the HAT, SPT and DUB module of the SAGA complex. This suggests a functional link of the *Arabidopsis* SAGA complex and the spliceosome.

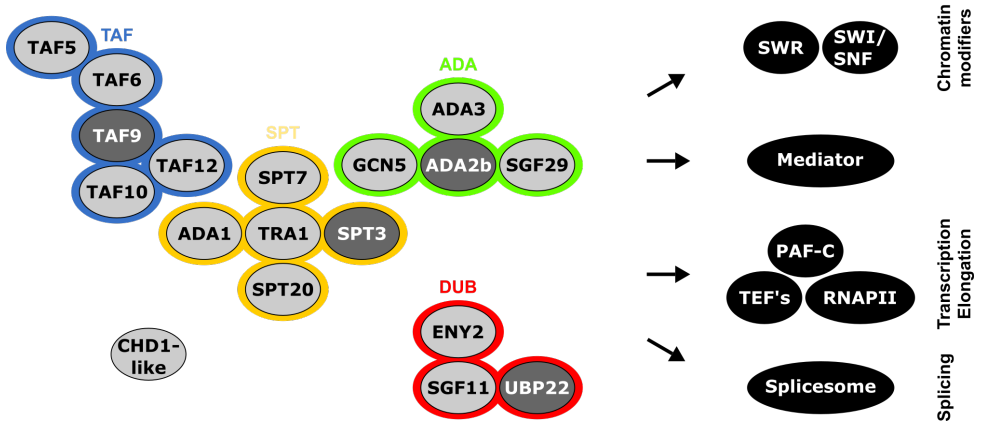


Figure 5.21 Model of the plant SAGA complex
Schematic interaction network revealed by AP-MS analysis of the SG-fusion proteins ADA2b, SPT3/TAF13, TAF10, and UBP22. The bait proteins are depicted in dark grey. The different SAGA modules are depicted in red (DUB), green (HAT), yellow (SPT), and blue (TAF). Co-purified chromatin- and trancription-related proteins are depicted in black. Arrows indicate co-purification.

5.7 Investigation of SAGA and SAGA-DUBm by AP-Superose6-Western Analysis

To shed more light on the plant SAGA complex and its interaction with the DUB module, the SAGA and co-purified protein complexes were analysed by size-exclusion chromatography followed by Western Blot analysis as well as mass spectrometry. The experimental approach was similar to the previously described AP-MS (**Section 5.3**). The SG-fusions of ENY2, SGF11 and UBP22 (DUB module) as well as ADA2b (HAT module) were expressed in *Arabidopsis* suspension cell culture. One step affinity purification was conducted using benzonase endonuclease to reduce high-molecular-mass nucleic acids and to remove RNA- or DNA-mediated interactions. Protein complexes co-purifying with the bait proteins were eluted from the IgG-coupled magnetic beads by proteolytic digest (TEV) rather than pH to keep the protein quaternary structures intact. The TEV site between the Streptavidin-Binding Protein (SBP) and the two Protein G domains allowed cleavage of the SG-Tag by the AcTEV protease. The proteolytic cleavage showed the same elution efficiency as the previously described elution with a low pH buffer. Following AP, the co-purified protein complexes were subjected to size-exclusion chromatography. The protein complexes were separated and fractionated according to their native sizes. The fractions of the Superose6 column were analysed by SDS-PAGE and Western Blotting using SBP-specific antisera to visualize the bait proteins (**Figure 5.22**). Moreover, the fractions 11 and 23 from the ADA2b and SGF11 AP were additionally analysed by mass spectrometry.

The immunoblotting showed that the four bait proteins eluted almost across all fractions ranging from high to low molecular weight. Slight differences of the signal intensities in the different fractions (peaks and gaps) were detectable. However, the bait proteins do not elute in distinct peaks, but still showed unique Western Blot pattern that were mostly overlapping with

each other.

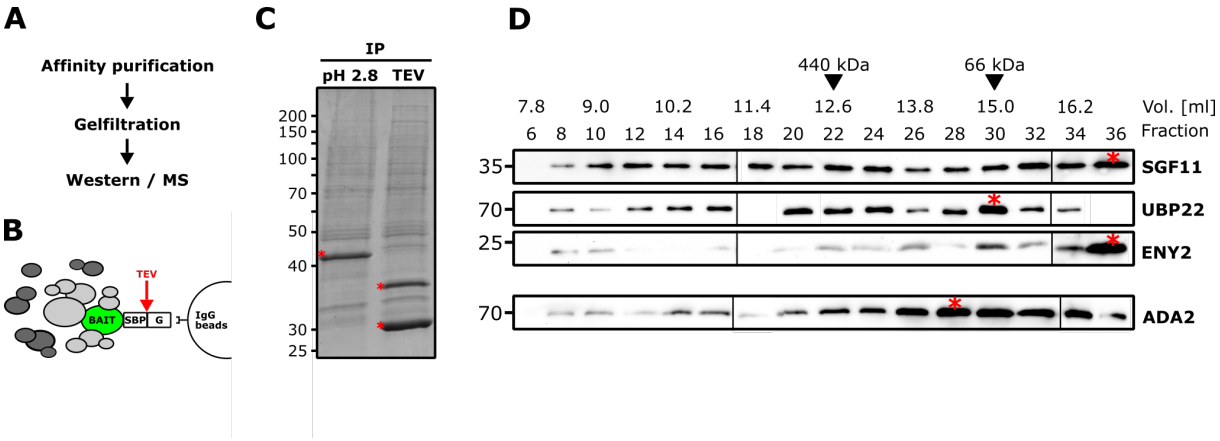


Figure 5.22 Characterization of the protein complexes co-purified with SGF11, UBP22, ENY2 and ADA2b. A)Schematic illustration showing the experimental setup. AP followed by Superose6 gelfiltration and coupled to Western Blot / MS analsis. B) Schematic illustration showing the affinity purification of a SG-tagged bait protein with IgG coupled magnetic beads. The bait protein is depicted in green, co-purifying proteins in light grey, and non-interacting proteins in dark grey. The TEV cleavage site between SBP and the two Protein G domains is indicated in red. C) The elution of co-purified proteins with TEV protease and low pH buffer. Top down, prominent bands indicated by red asterisks are SGF11-SG (41.3 kDa), SGF11-S (Protein G domains are cleaved off: 26.8 kDa) and AcTEV Protease (27 kDa). D) Protein complexes co-purified with SG-bait proteins using one-step AP were analysed by size-exclusion chromatography using Superose6 column. The collected fractions were analysed by SDS-PAGE/Western Blot using α-SBP antisera. The elution profile of the marker proteins (kDa) together with the elution volume and the fraction number are indicated at the top.

A strong Western Blot signal was detectable in the fractions that correspond to the monomeric bait proteins as following: SGF11 (fraction 36), UBP22 (fractions 28-34), ENY2 (fractions 36) and ADA2b (fractions 28-34). The signal intensity of the monomeric ENY2 was most strikingly. This phenomenon was most likely due to the overexpression of the bait proteins in the *Arabidopsis* suspension cell culture. The interacting proteins/complexes were co-purified at sub-stochiometric levels possibly resulting in free monomeric bait proteins.

The fractions 22 to 26 were corresponding to the putative DUB module ($\tilde{1}$05 kDa). The co-elution of SGF11, UBP22 and ENY2 in these fractions suggested that this module was formed in *Arabidopsis*. This was supported by the MS analysis of the fraction 23 from the SGF11 AP, where all three DUBm subunits were identified (**Table 5.4**).

The multiprotein SAGA complex with potentially approximately 2 MDa is expected in the high-molecular-mass fractions. The HAT component ADA2b as well as the DUB subunits ENY2, SGF11 and UBP22 co-eluted in these fractions (8 - 10). This suggested that the DUB module was a part of the SAGA complex.

By MS analysis, eight components of the SAGA modules HAT, SPT and TAF were identified in the high molecular weight fraction 11 of the ADA2b AP-Superose6, but no components of the DUB module (**Table 5.4**). On the other side, four subunits of the SAGA modules HAT, SPT and TAF were identified in the fraction 11 of the SGF11 AP-Superose6 and two components of the DUB module (ENY2, SGF11). This indicated that the DUB module is associated with the plant SAGA complex. But at the same time these data suggest that the HAT, SPT, and TAF modules are more strongly associated with each other than with the DUB module.

Moreover, several proteins of splicesomal complexes including the NTC, the NTR, the U2

Table 5.4 High- and medium- molecular mass fractions containing protein complexes co-purified with SGF11-SG (DUB module) and ADA2b-SG (HAT module).
Fraction 11 corresponds to an elution volume of 9.3 ml (1 MDa). Fraction 23 corresponds to an elution volume of 12.9 ml (= 400 kDa). Numbers indicate the respective MASCOT score.

				ADA2b-SG		SGF11-SG	
AGI	**Protein**	**Complex**	**Process**	**fr. 11**	**fr. 23**	**fr. 11**	**fr. 23**
AT5G13680	ELO2	Elongator	Transcription	156.2	295.8	117.2	141
AT5G50320	ELO3	Elongator	Transcription		211.8	249.9	264.5
AT3G28730	SSRP1	FACT	Transcription	178.8	97		
AT4G35800	NRPB1	RNAPII	Transcription				137
AT3G27100	ENY2	SAGA_DUBm	Transcription			168.8	191.6
AT5G58575	SGF11	SAGA_DUBm	Transcription			734	1198
AT4G16420	ADA2b	SAGA_HAT	Transcription	1555	2031	1218	1594
AT3G54610	GCN5	SAGA_HAT	Transcription	2187	2710	1902	2216
AT1G72390	SPT20	SAGA_SPT	Transcription	900.3	372.1	176.7	248.1
AT2G17930	TRA1a	SAGA_SPT	Transcription	1445	243.1		
AT4G36080	TRA1b	SAGA_SPT	Transcription	888			
AT3G10070	TAF12	SAGA_TAF	Transcription	97.9			
AT5G25150	TAF5	SAGA_TAF	Transcription	640.8	192.5	130.6	114.5
AT1G54140	TAF9	SAGA_TAF	Transcription	227.7	141.4		109.5
AT3G18790	ISY1	Core NTC	Splicing	156.2		145.5	
AT1G09770	CDC5	Core NTC	Splicing	388.5		86.5	
AT1G04510	MAC3A	Core NTC	Splicing	275.6			
AT2G33340	MAC3B	Core NTC	Splicing	329.9		142.3	
AT5G28740	SYF1	Core NTC	Splicing	167.8			119.2
AT4G15900	PRL1	Core NTC	Splicing	92.9			
AT1G77180	SKIP	Core NTC	Splicing	165.6		171.2	
AT2G38770	EMB2765	NTC-assoc.	Splicing	457.2		271.3	105.6
AT1G07360	ECM2-1a	NTC-assoc.	Splicing	82.6		231.7	99.8
AT4G18465	DDX35	C complex	Splicing	203.4			
AT1G14650	SAP114-1a	U2 snRNP	Splicing	83.6			
AT3G55200	SAP130a	U2 snRNP	Splicing	994.4	541.3	338.8	349
AT5G64270	SAP155	U2 snRNP	Splicing	118.7			
AT4G21660	SF3b150	U2 snRNP	Splicing	112.7			204.9
AT4G03430	EMB2770	U5 snRNP	Splicing	121.5		136.9	
AT1G20960	EMB1507	U5 snRNP	Splicing	1035	269.6	586.4	410.1
AT2G43770	U5-40	U5 snRNP	Splicing	94.9		109.8	113.1
AT5G09390	AT5G09390	U5 snRNP	Splicing				123.9
AT5G16780	MDF	U4/6/5 snRNP	Splicing		132.4		100.2

snRNP and the U5 snRNP were found in the fraction 11 of both APs, ADA2b and SGF11. This confirmed the assciation of the DUB module (SAGA-dependent or -independent) with the splicing machinery as described in **Section 5.3**. Therefore, the occurence of the DUB module in the high molecular weight fraction 11 could be due to the association of the DUB module with the large SAGA complex and/or the multi-megadalton splicing machinery.

Taken together, the protein complexes that co-purified with all four bait proteins showed a relatively broad elution profile. This suggested that these bait proteins may form a great variety of different-sized protein complexes. The co-elution of ENY2, SGF11 and UBP22 suggested that the *Arabidopsis* DUBm stably exists without a connection to the remaining SAGA complex. The question if the DUB module is permanently linked to the plant SAGA complex or not cannot be conclusively answered by this experimental setup, because of the huge overlap of the many co-purified complexes.

5.8 Characterization of plants with altered expression of *ENY2* or *SGF11*

For the functional analysis of the *Arabidopsis* DUB module, a reserve genetics approach was chosen.

5.8.1 *eny2-1* T-DNA line showed no downregulation of *ENY2*

To gain insights into the function of ENY2 in *Arabidopsis*, plants with altered expression of *ENY2* were characterized. At Nottingham *Arabidopsis* Stock Centre (NASC), there was one T-DNA-insertion line available for *ENY2* (AT3G27100). The annotated position of the T-DNA (SALK_045015) within the 3′UTR of *ENY2* was confirmed by PCR and subsequent DNA sequencing of the amplicon. The primers used for this purpose, binding to the genomic sequence of *ENY2* as well as to the SALK-insertion were depicted in (**Figure 5.23 A-C**). The SALK T-DNA was inserted exactly between thymidine and cytosine at position 99945621/2 of chromosome three. The left border (LB) of this SALK insertion was directing upstream towards the transcription start site of *ENY2*. The genetic background of this T-DNA insertion line SALK_045015, hereinafter named *eny2-1*, was Columbia (Col-0). The homozygous *eny2-1* mutants were selected by genotyping PCR using the primers as shown in (**Figure 5.23 A-B**).

The transcript level of *ENY2* was analysed in homozygous *eny2-1* plants by qRT-PCR. The primers used for this purpose were either spanning the fourth intron of *ENY2* or were binding to fifth exon and the 3′UTR as shown in (**Figure 5.23 A**). Compared to *wild type* Col-0 plants, the *eny2-1* mutant showed no downregulation of the *ENY2* transcript (**Figure 5.23 D-E**). Due to the unavailability of an ENY2-specific anitserum, the ENY2 protein levels were not determined in the *eny2-1* mutant plant line.

Consistent phenotypic analyses revealed that *eny2-1* mutant plants showed no obvious phenotype (**Figure 5.23 F-N and Table S16**). According to the measured features, there was no difference between the *eny2-1* mutant and the *wild type* Col-0 plants.

5.8.2 *ENY2* knockdown by RNAi showed no obvious phenotype

Since *eny2-1* mutants did not exhibit a reduction in the *ENY2* transcript level nor any obvious phenotypes, a RNA interference (RNAi) approach was pursued to knockdown *ENY2*

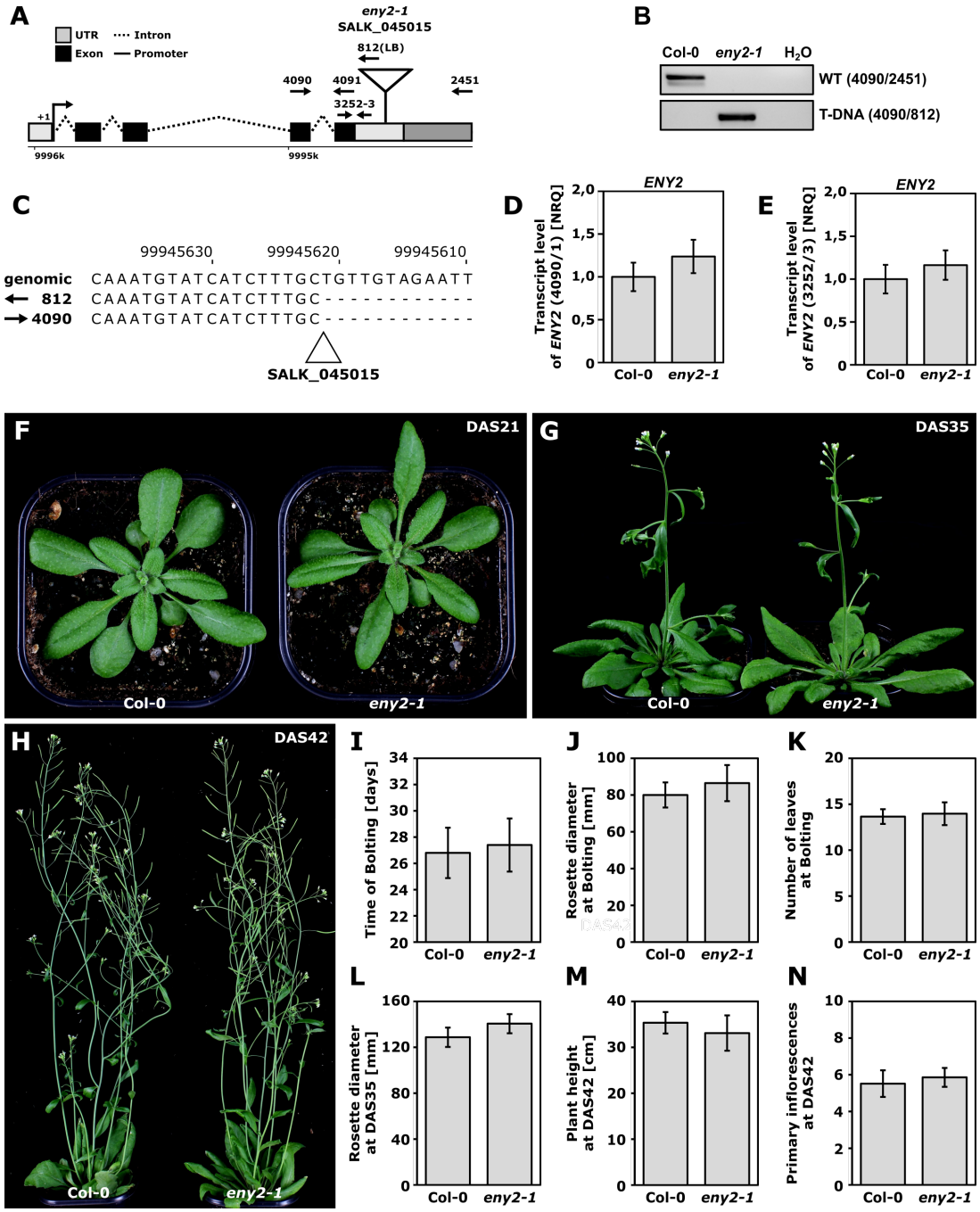


Figure 5.23 *eny2-1* mutants showed no obvious phenotype.
A) Schematic illustration of the *ENY2* locus (AT3G27100) with the T-DNA insertion (SALK_045015) in the untranslated 3′UTR. The primers used for the genotyping and the expression analysis are indicated as arrows (Black bars = exons, dark grey bars = Promoter, dotted lines = introns, light grey bars = UTR). B) The genotyping PCR of *wild type* Col-0 and mutant *eny2-1* plants. The following primer pairs were used to detect the *wild type* and the mutant *ENY2* allele, 4090/2451 and 4091/812 respectively. C) The position of the T-DNA insertion was determined by PCR using the primer pair 812/4090 and subsequent DNA sequencing of the amplicon using the primers 802 and 4090. The genomic sequence of *ENY2* was aligned with the obtained sequencing results. Mismatches indicate the T-DNA insertion site (Arrow). D-E) The relative expression of *ENY2* in *wild type* Col-0 and mutant *eny2-1* plants was determined by qRT-PCR. D) The primer pair (4090/4091) was spanning the fourth exon of *ENY2*. E) The primer pair (3252/3253) was binding to the last exon and 3′UTR of *ENY2*. The normalized relative quantities (NRQ) ± the normalized relative standard error (SE[NRQ]) are shown. The relative quantity of *ENY2* mRNA was normalized to the relative quantities of *GAPC*, *PP2AA3* and *UBI10* mRNA. The significance was tested by Student's T-Test. F-N) Phenotypical analysis of mutant *eny2-1* plants in comparison to *wild type* Col-0 plants grown under LD conditions. F-H) Representative individuals are shown at various developmental stages (F) DAS21 (days after stratification) (G) DAS35 (H) DAS42 I-N) The following plant features were statistically evaluated: (I) The time of bolting (Elongation of the first internode), (J) the rosette diameter and (K) the number of leaves at time of bolting, (L) the rosette diameter at DAS35, (M) the plant height and (N) the primary inflorescences at DAS42 (All data are means ± SD, Significance was tested by Student's T-Test)

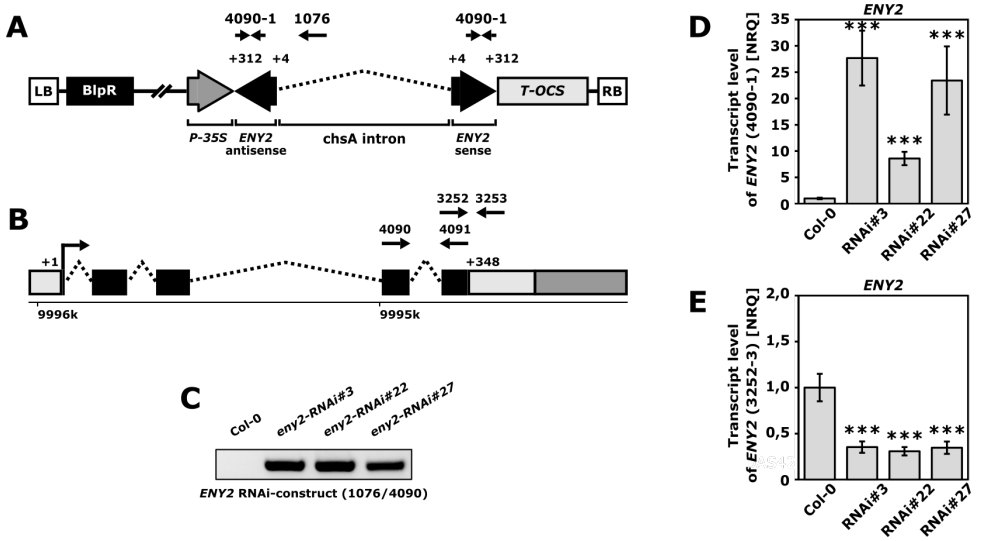


Figure 5.24 Molecular characterization of *eny2-RNAi* knockdown mutants
A) Schematic illustration of the RNAi transgene for silencing of *ENY2*. Two inverted repeats of *ENY2* fragments (corresponding to 4-312 bps of *ENY2* CDS) were separated by a chsA intron. B) Schematic illustration of the *ENY2* locus (AT3G27100). The primers used for the genotyping and the expression analysis were indicated as arrows (Black bars = exons, dark grey bars = Promoter, dotted lines = introns, light grey bars = UTR). C) Genotyping PCR of *wild type* Col-0 and three independent *eny2-RNAi* lines (#3, #22, #27). The primer pair 1076/4090 was used to detect plants carrying the *eny2-RNAi* transgene. D-E) The relative expression of *ENY2* in *wild type* Col-0 and *eny2-RNAi* mutant plants was determined by qRT-PCR. D) Strong overexpression of the RNAi construct in transgenic RNAi lines was detectable. The expression of the plasmid-derived *ENY2* dsRNA and the endogenous *ENY2* mRNA were detected by using the primer pair 4090/4091 (spanning the fourth exon of *ENY2*). E) The endogenous *ENY2* mRNA was significantly reduced in the transgenic RNAi lines. The primer pair 3252/3253 was binding to the last exon and the untranslated 3′UTR of *ENY2*. The normalized relative quantities (NRQ) ± the normalized relative standard error (SE[NRQ]) are shown. The relative quantity of ENY2 mRNA was normalized to the relative quantity of GAPC, PP2AA3 and UBI10 mRNA (*** indicates p-Value < 0.001, Student′s T-Test).

activities. To generate plants producing a potential interfering RNA molecule, *Arabidopsis* Col-0 plants were *Agrobacterium*-mediated transformed using an *ENY2* specific RNAi plasmid. This pFGC5941-derived construct drove the expression of the ENY2 sense and antisense fragments (4-312 of CDS) that were separated by a ChsA intron under the CaMV 35S promoter. Three independent transgenic lines that were homozygous for the RNAi-transgene were selected by genotyping PCR using primers as depicted in (**Figure 5.24 A - C**).

In the next step, the transcript level of *ENY2* was investigated by qRT-PCR in these plant lines (*eny2-RNAi* #3, #22, #27) to see if the expression of the double stranded RNA (dsRNA) led to silencing of the *ENY2* gene expression. First, the primer pair (4090/4091) that was spanning the fourth exon of *ENY2* as depicted in **Figure 5.24 B**, was used to detect the plasmid-derived *ENY2* dsRNA and the endogenous *ENY2* mRNA. This showed the overexpression of the RNAi construct in the mutant lines (**Figure 5.24 D**).
Primers that were binding specific to the 5′end of *ENY2* (3252/3253) as depicted in **Figure 5.24 B**, allowed the quantification of the endogenous *ENY2* mRNA level. This revealed that *ENY2* gene expression was significantly affected by the RNAi (**Figure 5.24 E**). Compared to *wild type* plants, these three RNAi lines exhibited reduced amounts of *ENY2* transcripts as follows: *eny2-RNAi* #3 (35%), *eny2-RNAi* #22 (31%) and *eny2-RNAi* #27 (35%).

In the following, these three RNAi lines, showing efficient silencing of the *ENY2* gene expression, were phenotypically analysed in comparison to *wild type* Col-0 plants. This revealed

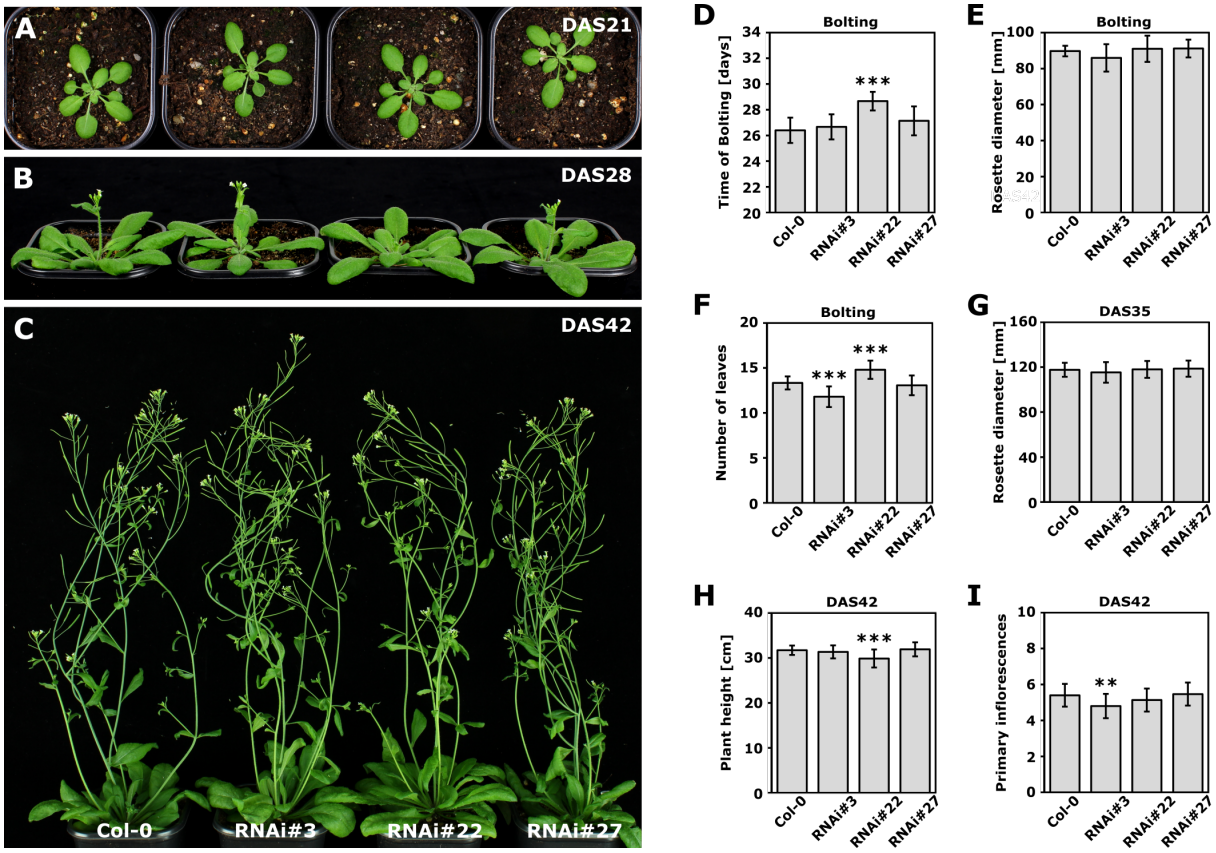


Figure 5.25 eny2-RNAi plant lines showed no obvious phenotype.
Phenotypical analysis of the mutant *eny2-RNAi* lines in comparison to *wild type* Col-0 plants grown under LD conditions. Representative individuals were shown at various developmental stages A) DAS21 (days after stratification) B) DAS28 C) DAS42 D-I) The following plant features were statistically evaluated: (D) The time of bolting (Elongation of the first internode), (E) the rosette diameter and (F) the number of leaves at bolting, (G) the rosette diameter at DAS35, (H) the plant height and (I) the primary inflorescences at DAS42 (All data are means ± SD, *** indicates p-Value < 0.001, Student's T-Test).

that plants with reduced levels of *ENY2* showed *wild type* appearance (**Figure 5.25 and Table S17**). Just one of the transgenic lines (*eny2-RNAi* #22) showed a late flowering phenotype. Apart from this, only smaller phenotypic fluctuations could be observed.

Taken together, in plants producing *ENY2* interfering RNA molecules, *ENY2* gene expression was reduced by approximately 70%. This did not lead to an obvious phenotype that could be consistently observed in all three tested RNAi lines.

5.8.3 CRISPR/Cas9-induced mutations to disrupt *ENY2*

The CRISPR/Cas9 system of Wang et al. 2015 was used to disrupt the *ENY2* gene and alter its expression. Therefore, a CRISPR construct for the co-expression of the *ENY2*-specific guide RNA (gRNA) and the non-specific CRISPR associated endonuclease (Cas9) was generated (**Figure 5.26 A**). The gRNA, a fusion of the 20 nt *ENY2*-specifc spacer sequence and the RNA scaffold (necessary for the Cas9-binding), was driven by the U6-26 promoter. The spacer sequence was predicted by CRISPR-P (Lei et al. 2014) and was targeting the second exon of *ENY2*. The expression of the endonuclease was driven by the fusion promoter of the two egg-cell-specific genes *EC1.2* and *EC1.1*. The CRISPR transgene was integrated into the genome of

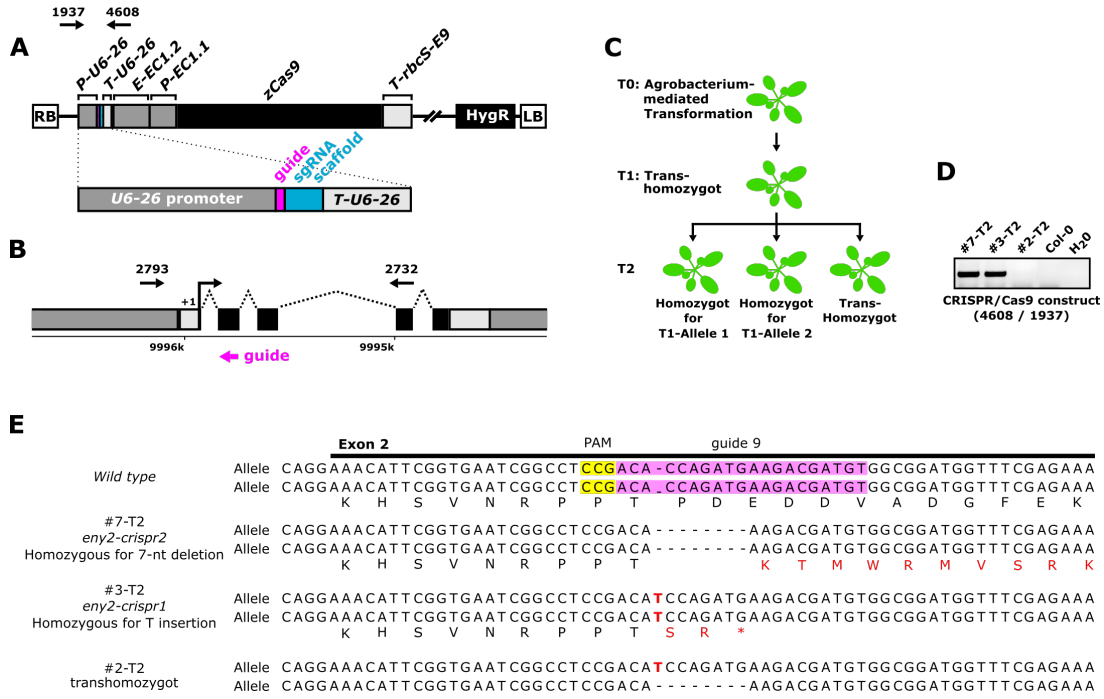


Figure 5.26 CRISPR/Cas9-induced mutations to disrupt *ENY2*.
A) Schematic illustration of the *ENY2* specific CRISPR/Cas9 transgene. The expression of *ENY2*-specific guide RNA fuse to sgRNA scaffold is driven by U6-26 promoter. The expression of the zCas9 endonuclease is driven by *EC1.2* enhancer and *EC1.2* promoter. Primers used for genotyping and the expression analysis are indicated as arrows (Black bars = exons, blue bar = sgRNA scaffold, dark grey bars = promoter, HygR = hygromycin resistance marker, light grey bars = terminator, pink bar = guide RNA, white bars = left/right border of the T-DNA). B) Schematic illustration of the *ENY2* (AT3G27100) locus with *ENY2* specific guide RNAs for CRISPR/Cas9-induced mutations in the second exon of *ENY2*. C) Schematic diagram is visualizing the outcome of the genome editing by *ENY2*-specific CRISPR/Cas9 mutagensis in the T1 and T2 generation. A transhomozygous mutant was obtained in the T1 generation, carrying a 7-nt deletion in the one *ENY2* allel and a 1-nt insertion in the other. The two different mutations were inherited to the next generation what results in the following T2 progenies: (1) homozygous for the deletion, (2) homozygous for the insertion or (3) transhomozgyous. D) Genotyping PCR with CRISPR/Cas9 transgene specific primers (4608/1937) if Cas9 is still integrated into the genome of T2 mutants. E) The CRISPR/Cas9 induced mutations were determined by PCR using the primer pair 2793/2732 and subsequent DNA sequencing of the amplicon using the primers 2793. The genomic sequence of *ENY2* was aligned with the obtained sequencing results. Mismatches indicate the insertions and/or deletions (red). The aa sequence is shown in one-letter notation (*Wild type* aa in black, mutated aa in red, and premature stop codons as asteriks).

Arabidopsis Col-0 plants by *Agrobacterium*-mediated transformation and positively transformed T1 individuals were selected by growth on MS (hygromycin) plates and by genotyping PCR using the primer pair (1937/4608). The CRISPR/Cas9-induced mutations in the transgenic T1 plants were identified by DNA sequencing of the PCR products encompassing the CRISPR target site using the primers depicted in **Figure 5.26 B**. Thereby, one trans-homozygous ENY2 mutant was obtained. The gene editing efficiency of the CRISPR transgene was 10% (1 out of 10 screened T1 plants). The bi-allelic *eny2-crispr* mutant (T1) showed two different mutations at position 9995809/10 on the paternal and maternal alleles as following (**Figure 5.26 E**): 1) Thymine (T) insertion 2) T insertion and additional 7 nt deletion. Both mutations induced a frame-shift/pre-mature stop codon and resulted most likely in a knockout of *ENY2*. The CRISPR-induced bi-allelic mutation was inherited according to Mendel′s law (**Figure 5.26 C**), which resulted in a mixed T2 population consisting of mutants i) homozygous for the T-insertion (*eny2-crispr1*) ii) homozygous for the T-insertion/7-nt deletion (*eny2-crispr2*) and iii) trans-

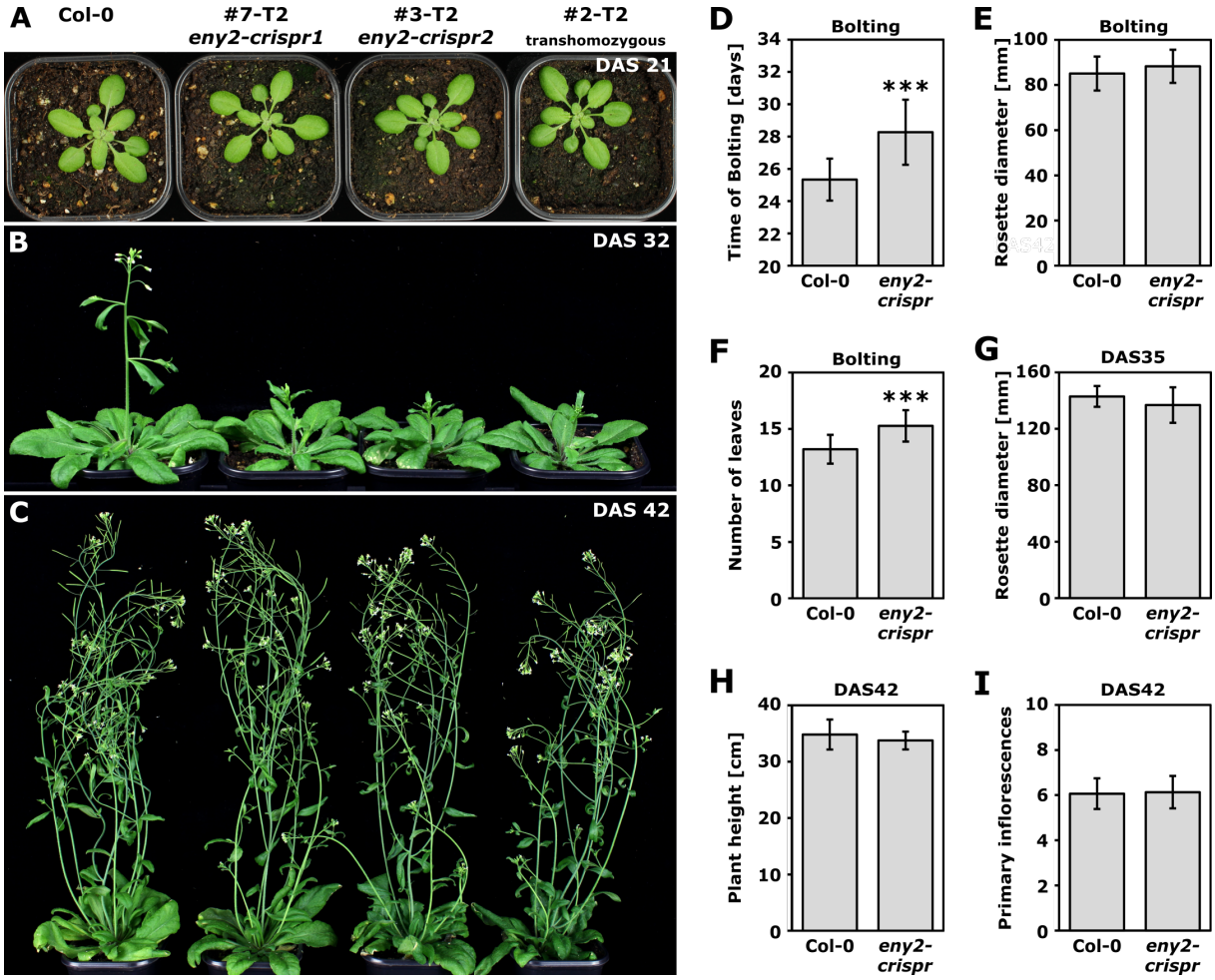


Figure 5.27 *eny2-crispr* mutants showed late flowering phenotype.
Phenotypical analysis of *eny2-crispr* lines in comparison to *wild type* Col-0 plants grown under LD conditions. Representative individuals are shown at various developmental stages (A) DAS21 (days after stratification) (B) DAS28 (C) DAS42 D-I) The following plant features were statistically evaluated: (D) The time of bolting (Elongation of the first internode), (E) the rosette diameter and (E) the number of leaves at bolting, (G) the rosette diameter at DAS35, (H) the plant height and (I) the primary inflorescences at DAS42 (All data are means ± SD, Significance was tested by Student's T-Test, *** indicates $p < 0.001$)

homozygous (*eny2-crispr*). Moreover, 10 % (n = 10) of the T2 ENY2 mutants were already transgene-free (**Figure 5.26 D**).

The segregating *eny2-crispr* line was phenotypically analysed in comparison to *wild type* Col-0 plants. The plants were grown on soil under long day conditions to monitor their morphology and development. The following characteristics were examined: The time of bolting (Elongation of the first internode), the rosette diameter and the number of leaves at bolting, the plant height and the rosette diameter at DAS35, the plant height and the primary inflorescences at DAS42. The ENY2-depleted plants bolted on average (n = 15) 2.94 days later than *wild type* Col-0 plants (**Figure 5.27 and Table S18**). Moreover, the *eny2-crispr* plants exhibited 2.07 leaves more than the control plants at the time of bolting. Therefore the phenotypic analysis revealed that the complete knockout of *ENY2* caused a late flowering phenotype.

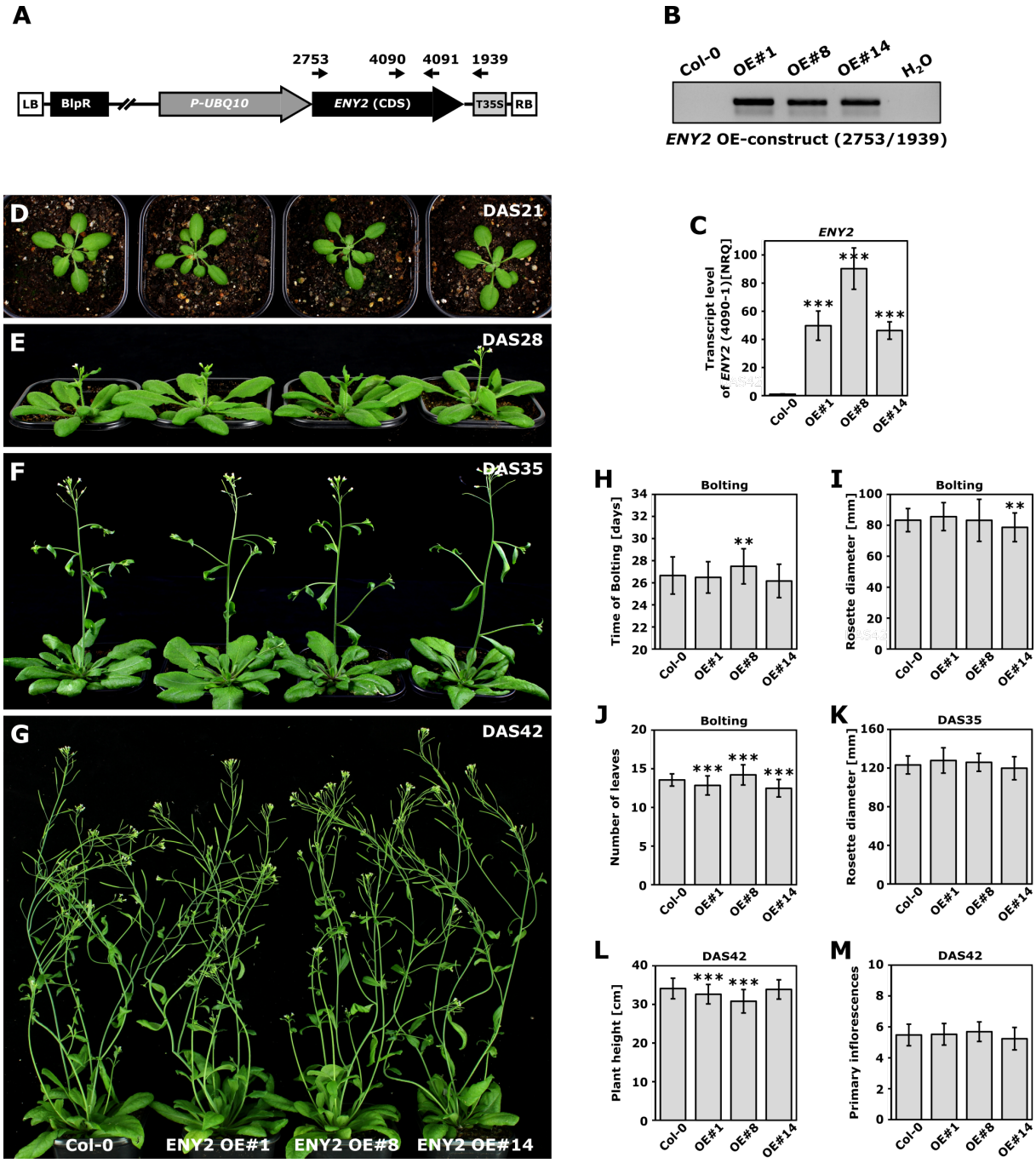


Figure 5.28 Overexpression of *ENY2* did not cause obvious phenotype.
A) Schematic illustration of the *ENY2* overexpression transgene. The expression of *ENY2* coding sequence (AT3G27100) is driven by *UBI10* Promoter. Primers used for genotyping and the expression analysis are indicated as arrows (Black bars = exons, dark grey bars = promoter, light grey bars = UTR, white bars = left/right border of T-DNA). B) Genotyping PCR of *wild type* Col-0 and *eny2-OEx* lines. The primer pair 2753/1939 was used to detect the overexpression transgene integrated in the genome. C) The relative expression of *ENY2* in *wild type* Col-0 and mutant *eny2-OEx* plants was determined by qRT-PCR. For detection of the endogenous and transgene-derived *ENY2* mRNA, the primer pair 4090/4091 (spanning the fourth exon of *ENY2*) was used. The normalized relative quantities (NRQ) ± the normalized relative standard error (SE[NRQ]) are shown. The relative quantity of *ENY2* mRNA was normalized to the relative quantities of *GAPC*, *PP2AA3* and *UBI10* mRNA. The significance was tested by Student's T-Test (*** indicates $p < 0.001$). D-G) Phenotypical analysis of the *ENY2* overexpression lines in comparison to *wild type* Col-0 plants grown under LD conditions. Representative individuals are shown at various developmental stages (D) DAS21 (days after stratification) (E) DAS28 (F) DAS35 (G) DAS42 H-M) The following plant features were statistically evaluated: (H) The time of bolting (Elongation of the first internode), (I) the rosette diameter and (J) the number of leaves at bolting, (K) the rosette diameter at DAS35, (L) the plant height and (M) the primary inflorescences at DAS42 (All data are means ± SD, Significance was tested by Student's T-Test, *** indicates $p <$ 0.001)

5.8.4 *ENY2* overexpression by the *UBI10* promoter

As parallel approach to knockout *ENY2* and study loss-of-function effects, *ENY2* was overexpressed to gain insights about its function. The overexpression of *wild type* gene products can cause mutant phenotypes (Prelich 2012). Therefore, a transgene driving the expression of *ENY2* coding sequence by *UBI10* promoter was *Agrobacterium*-mediated integrated into the genome of *Arabidopsis* Col-0 plants. Three independent *ENY2* overexpression lines (*eny2-OEx* #1,#8,#14) were selected on MS plates containing kanamycin and genotyped using primers depicted in **Figure 5.28 A,B**.

The transcript level of *ENY2* was analysed by qRT-PCR in the three selected transgenic plants lines that were homozygous for the integrated *eny2-OEx* transgene. For detection of the endogenous and the transgene-derived *ENY2* mRNA, primers spanning the fourth intron of genomic *ENY2* were used as depicted in **Figure 5.28 A**. Compared to *wild type* Col-0 plants, all three *eny2-OEx* lines showed highly increased levels of the *ENY2* transcript as follows (**Figure 5.28 C**): *eny2-OEx* #1 (49.7 fold), #8 (90.3 fold), #14 (46.3 fold).

Next, the *eny2-OEx* lines were phenotypically analysed in comparison to *wild type* Col-0 plants as described before. Thereby, plants overexpressing *ENY2* showed consistently *wild type* appearance (**Figure 5.28 and Table S19**). Only smaller phenotypic fluctuations were observable between the three independent *eny2-OEx* lines, that were most likely due to the random integration of the T-DNA.

5.8.5 Knockdown of *SGF11* showed late flowering phenoytpe

To investigate the role of *SGF11* (AT5G58575), the T-DNA insertion line SAIL_856_F11 was obtained from the Nottingham *Arabidopsis* Stock Centre (NASC). The genetic background of this T-DNA mutant was *Arabidopsis* ecotype Columbia (Col-0). In mutant plants, the position of the annotated T-DNA insertion was determined by PCR and DNA sequencing of the amplicon. The primers used for this purpose, binding to the *SGF11* genomic sequence as well as to the SAIL-insertion were depicted in **Figure 5.29 A**. The T-DNA insertion SAIL_856_F11 was mapped to the second exon of *SGF11* at position 23674751/2 of chromosome five (**Figure 5.29 C**). The left border (LB) of the T-DNA insertion was directing downstream towards the 3′ end of *SGF11*. *sgf11-1* mutant plants, homozygous for the T-DNA insertion, were selected by genotyping PCR using the primers shown in **Figure 5.29 A, D**.

The transcript level of *SGF11* was analysed by qRT-PCR in 10 day old seedlings of homozygous *sgf11-1* plants. With primers flanking the T-DNA insertion, no SGF11 transcript was detected in the *sgf11-1* mutants compared to *wild type* plants (**Figure 5.29 E**).

A second open reading frame (ORF) that was in-frame with the second exon of *SGF11* started within the T-DNA insertion. Whether this ORF is expressed and if the putative transcript is correctly spliced was analysed by qRT-PCR. Primers were designed to bind the putative transcript as well as to span the second intron of *SGF11* (**Figure 5.29 B**). This revealed that the second ORF was expressed, but much weaker than the endogenous *SGF11* (**Figure 5.29 F**). The expression of this truncated *SGF11* transcript was most likely driven by the 1′2′ promoter that was located 96 nt from the LB of the SAIL_856_F11 T-DNA insertion (Ülker et al. 2008). The melting curves of the transcripts obtained from the *wild type* and the *sgf11-1* plants were

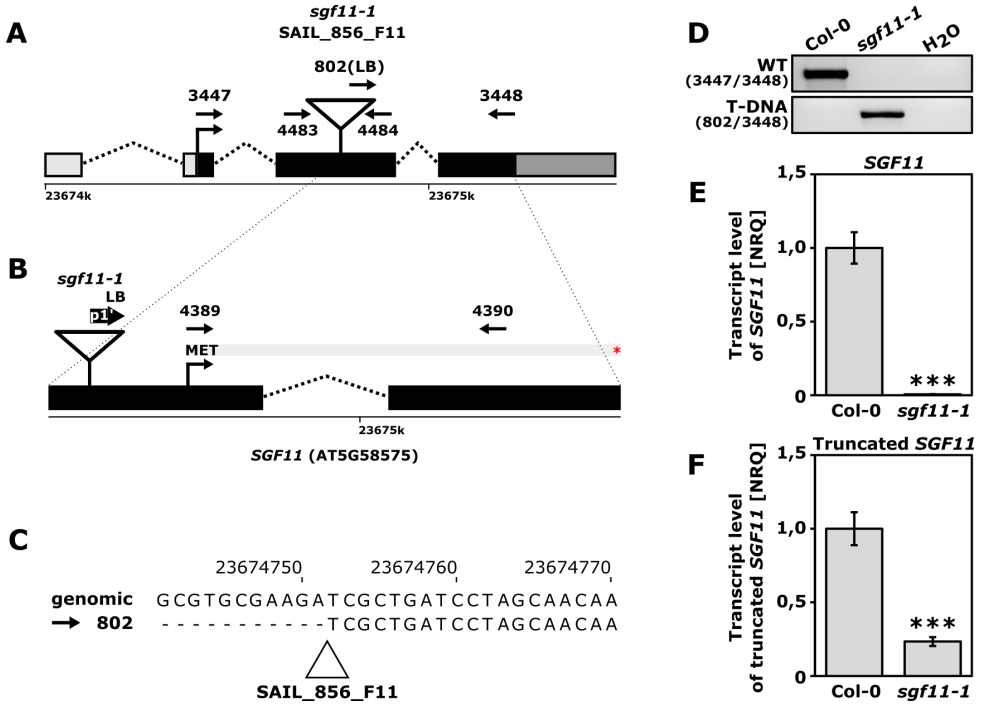


Figure 5.29 Molecular characterization of the *sgf11-1* knockdown mutant.
A) Schematic illustration of the *SGF11* locus (AT5G58575) with the T-DNA insertion (SAIL_856_F11) located in the second exon. The primers used for genotyping and expression analysis were indicated as arrows (Black bars = exons, dark grey bars = promoter, dotted lines = introns, light grey bars = UTR). B) Closeup of the *SGF11* region downstream of the T-DNA insertion site. The expression of the second open reading frame (in-frame with second exon of *SGF11*) was putatively driven by 1′2′ promoter that was located at left border (LB) of the T-DNA insertion. C) The position of the T-DNA insertion was determined by PCR, using the primer pair 802/3448, and subsequent amplicon sequencing using the primer 802. The genomic sequence of *SGF11* was aligned with the obtained sequencing result. The mismatches indicate the position of the T-DNA insertion. D) Genotyping PCR of *wild type* Col-0 and mutant *sgf11-1* plants. The following primer pairs were used to detect *wild type* and mutant *SGF11* allele, 3447/3448 and 802/3448 respectively. E-F) The relative expression of *SGF11* in *wild type* Col-0 and mutant *sgf11-1* plants was determined by qRT-PCR. E) The primer pair (4483/4484) was flanking the T-DNA insertion site. F) The primer pair (4389/4390) was spanning the second intron of *SGF11*. The normalized relative quantities (NRQ) ± the normalized relative standard errors (SE[NRQ]) are shown. The abundance of the *SGF11* mRNA in Col-0 and *sgf11-1* derived sampples was normalized to the transcript levels of *GAPC*, *PP2AA3* and *UBI10* (*** indicates p-Value < 0.001, Student′s T-Test).

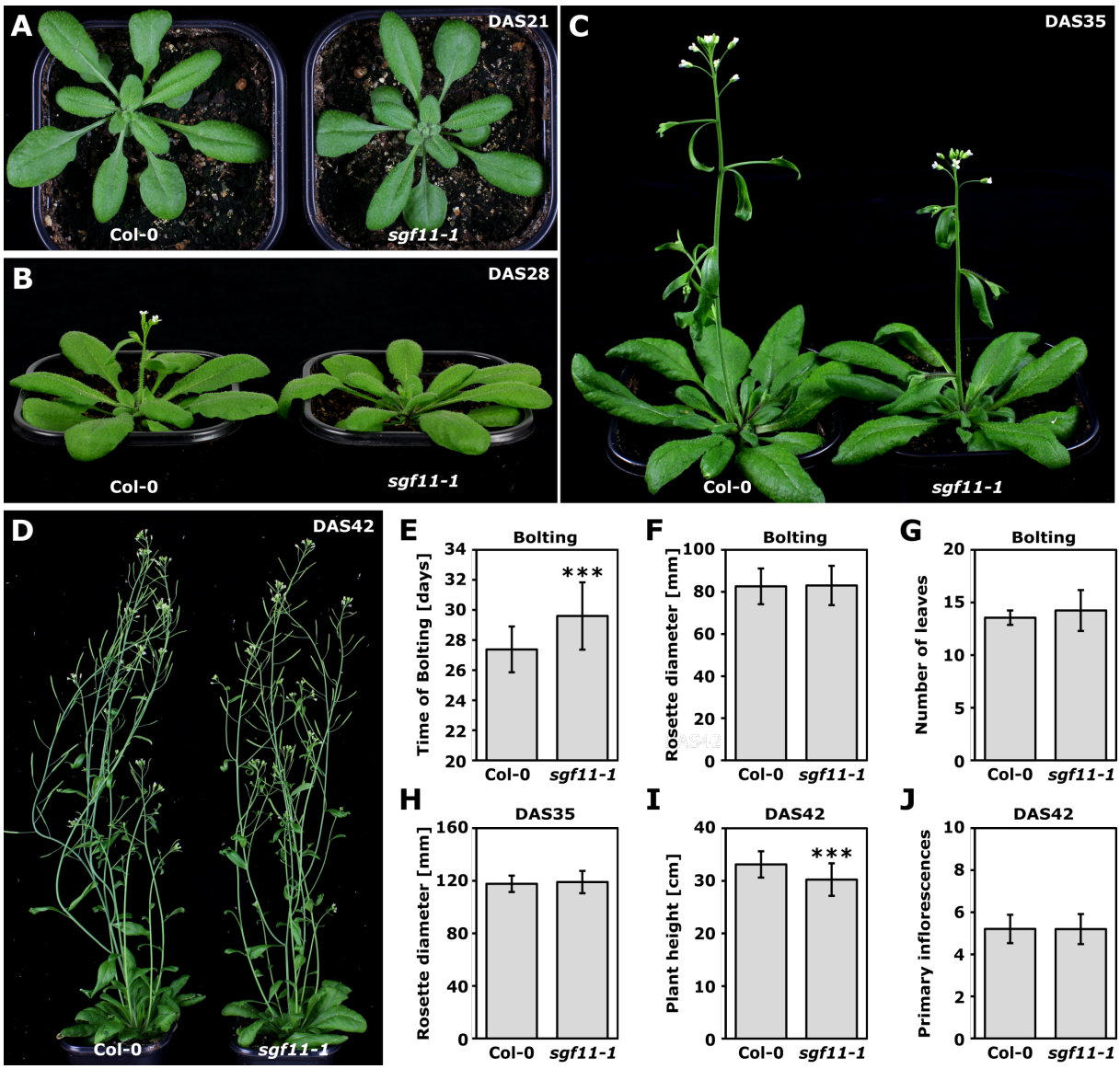


Figure 5.30 *sgf11-1* mutant plants showed late flowering phenotype.
Phenotypical analysis of *sgf11-1* mutant plants in comparison to *wild type* Col-0 plants grown under LD (long day) conditions. Representative individuals were shown at various developmental stages A) DAS21 (days after stratification) B) DAS28 C) DAS35 D) DAS42 E-J) The following plant features were statistically evaluated: (E) The time of bolting (The elongation of the first internode), (F) the rosette diameter and (G) the number of leaves at bolting, (H) the rosette diameter at DAS35, (I) the plant height and (J) the primary inflorescences at DAS42 (All data are means ± SD, *** indicates p-Value < 0.001, Student's T-Test).

identical (data not shown), indicating that the second intron was correctly spliced out.

Taken together, the T-DNA insertion resulted in a complete knockout of the endogenous *SGF11*, but promoted the weak expression of an aberrant *SGF11* transcript that is strongly truncated at the 5′end.

In the following, the *sgf11-1* mutant plants were phenotypically analysed in comparison to *wild type* Col-0 plants (**Figure 5.30 and Table S20**) as described before. This analysis revealed that the plants lacking endogenous *SGF11* showed *wild type* appearance except a slight late flowering phenotype. The mutant plants showed on average a delay of 2.2 days whose significance was determined using the Student′s T-Test. The significantly reduced plant height at DAS35 and DAS42 is most likely a consequence of the late bolting phenotype.

5.8.5.1 The level of *FLC* is upregulated in *sgf11-1* knockout plants

In *Arabidopsis*, the elongation of the first internode (bolting) indicates the transition from the vegetative to the reproductive phase. In the development of flowering plants, the timing of this crucial step is controlled by a variety of endogenous and environmental cues. These factors that are controlling the flowering time include the vernalization, the photoperiod and the autonomous pathways. The variety of sometimes competing input signals converges in the activation of a small number of floral pathway integrators (Simpson et al. 2002, Hepworth et al. 2015). In *Arabidopsis*, the *FLOWERING LOCUS C* (*FLC*) plays a central role in the repression of the flowering time (Sheldon et al. 1999, Michaels et al. 1999). In many *Arabidopsis* mutants, that are showing early or late flowering phenotypes, the level of *FLC* is reduced or increased, respectively (Mahrez et al. 2016b). The floral integrator *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*) promotes the flowering and is repressed by FLC (Michaels et al. 2005).

To test if the *FLC* and *SOC1* mRNA levels are affected in *sgf11-1* mutant plants, the total RNA of 10 days old seedlings was isolated and transcribed into cDNA as template for qRT-PCR. *FLC* and *SOC1* specific primers were used. It was shown by two biological replicates that the level of *FLC* was significantly (p=0.02) 1.64-fold increased in *sgf11-1* mutants plants compared to *wild type* Col-0 plants. The level of *SOC1* was reduced to 70%, but according to the Student′s T-Test this was not significant (p=0.19).

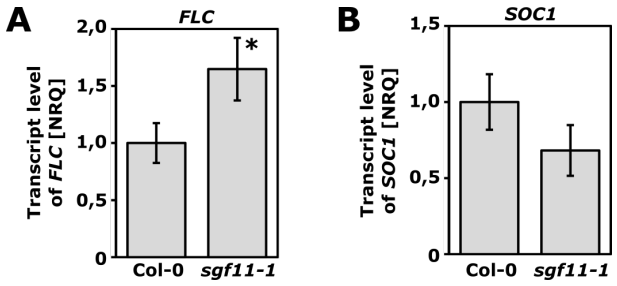


Figure 5.31 Gene expression of floral integrators was affected in *sgf11-1* mutants.
The relative expression of (A) *FLC* and (B) *SOC1* in *wild type* Col-0 and mutant *sgf11-1* plants was determined by qRT-PCR. The normalized relative quantities (NRQ) ± the normalized relative standard error (SE[NRQ]) are shown. The relative quantity of *FLC*/*SOC1* mRNA was normalized to the relative quantity of *GAPC*, *PP2AA3* and *UBI10* mRNA (* indicates p-Value < 0.05, Student′s T-Test).

5.9 Global H2Bub levels were increased in *eny2-RNAi* and *sgf11-1* mutants

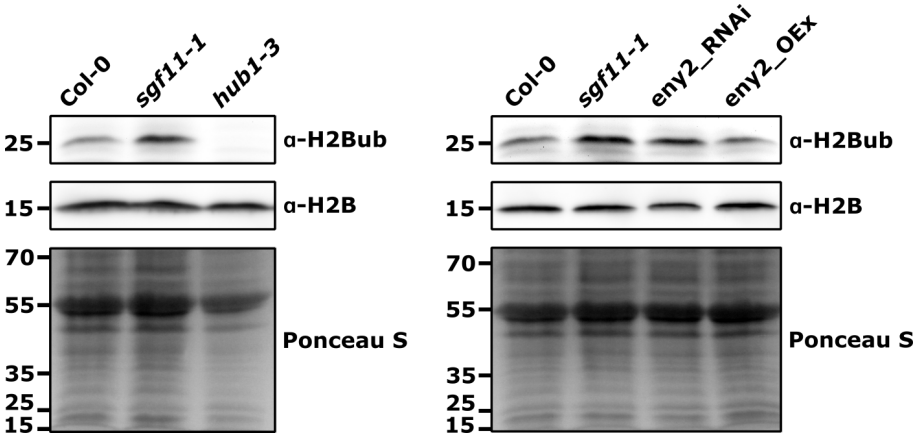


Figure 5.32 The global H2Bub levels were increased in *eny2-RNAi* and *sgf11-1* mutants
Total proteins were extracted from flowers of Col-0, *sgf11-1*, *eny2RNAi*, *eny2OEx* and *hub1-3* (neg control). Crude protein extracts were analysed by SDS-PAGE coupled to immunoblotting using H2B and H2Bub antisera. Ponceau staining of the membrane was used as additional loading control. The sizes of the molecular weight marker proteins were shown in kDa.

In yeast and animals, the DUB module of the SAGA complex is important for removing monoubiquitin from the histone H2B (H2Bub) (Samara et al. 2010, Köhler et al. 2010). The antagonist of the DUBm in yeast, the enzyme that is responsible for the transient modification of lysine K123 at H2B in the first place, is the ubiquitin conjugase (E2) Rad6 that acts in conjugation with the ubiquitin ligase (E3) Bre1 (Wood et al. 2003, Robzyk et al. 2000, Hwang et al. 2003). The monoubiquitinated H2B (H2Bub) is linked to other chromatin modifications and plays an important role, especially during transcript initiation and elongation (Weake et al. 2008a). In *Arabidopsis*, HUB1/HUB2 (HISTONE MONOUBIQUITINATION1 and 2) and UBC1/UBC2 (UBIQUITIN CARRIER PROTEIN1 and 2) are the orthologs of yeast Bre1 and Rad6, respectively and work together to monoubiquitinate lysine K143 of H2B (Fleury et al. 2007, Liu et al. 2007, Zhang et al. 2007, Cao et al. 2008). This can be observed as *hub1* and *hub2* mutants show globally reduced levels of H2Bub (Cao et al. 2008). This raised the question, if on the other hand *sgf11-1* and *eny2-RNAi* knockdown plants with a putative impaired DUB module, show increased levels of H2Bub. Therefore, H2Bub levels were examined in floral tissue protein extracts of *sgf11-1*, *eny2-RNAi* and *eny2-OEx* mutants compared to *wild type* Col-0 and *hub1-3* mutant plants *in vivo*. The total protein extracts of flowers were analysed by SDS-PAGE and Western Blot using H2B and H2Bub antisera (**Figure 5.32**). The level of the non-ubiquitinated H2B was in all samples comparable between the *wild type* and the mutant plants. Additionally, the ponceau staining of the membrane confirmed as well an equally loading of the gel. The Western Blot analysis showed a clear increase of the global H2Bub level in *sgf11-1* plants compered to *wild type* plants. As expected, the H2Bub level was strongly reduced and not detectable in *hub1-3* mutant plants. The knockdown of *ENY2* was as well accompanied by slightly increased H2Bub levels. On the other side, the overexpression of *ENY2* did not cause any alterations of the H2Bub levels compared to *wild type* plants. These results showed that SGF11 and ENY2 are required for the global deubiquitylation of H2Bub, indicating that the function of the *Arabidopsis* DUB module is conserved and its catalytic activity depends on the

two adaptor proteins SGF11 and ENY2. Moreover, these data suggest a putative connection of the previously described late-flowering phenotype of *sgf11-1* mutants (**Figure 5.30**) and the increased H2Bub level.

6 Discussion: Analysis of the HMG-box domain of *Arabidopsis* SSRP1

The FACT complex, a heterodimer of SSRP1/Pob3 (higher eukaryotes/yeast) and SPT16 is an essential histone chaperone. Both, SPT16 and SSRP1 are highly conserved from yeast to humans, except for species-specific characteristics at the C-terminus of SSRP1 (Wittmeyer et al. 1997, Brewster et al. 1998, Evans et al. 1998, Wittmeyer et al. 1999). In humans and plants, SSRP1 features an HMG-box domain at the C-terminus that is lacking in the yeast Pob3. In return, the yeast SPT16/Pob3 complex is loosely associated with Nhp6, which contains a HMGB domain and provides the DNA binding/bending functions for the yeast FACT complex (Brewster et al. 2001, Formosa et al. 2001). The genes coding for SPT16 and SSRP1/Pob3 are essential for viability in all investigated species (Winkler et al. 2011a, Formosa 2012). In contrast, the double knockout of both Nhp6 genes in yeast is viable (Costigan et al. 1994, Stillman 2010).

6.1 The SSRP1 HMG-box domain is important for FACT – nucleosome interaction

As histone chaperone, FACT is important for chromatin-dependent processes like transcription, replication and DNA repair (Wittmeyer et al. 1997, Orphanides et al. 1998, Schlesinger et al. 2000, Keller et al. 2002). Thereby, FACT reorganizes nucleosomes which allows or restricts other factors to gain access to the DNA (Orphanides et al. 1999, Xin et al. 2009). In this study, EMSA experiments revealed that *Arabidopsis* SSRP1 has DNA-binding properties that were mediated by the C-terminal HMG-box domain. SSRP1 lacking the C-terminal HMG-box domain showed no binding to DNA as well as to nucleosomes at the tested protein concentrations. These results were in accordance to studies in maize, where ZmSSRP1 showed comparable affinities for free DNA as AtSSRP1 and the DNA-binding ability was abolished by removing the C-terminal HMG-box domain (Röttgers et al. 2000). Moreover, ZmSSRP1 was shown to bind to mononucleosomes (Lichota et al. 2001) at concentrations that are comparable to AtSSRP1.

Moreover, AtSSRP1 had a higher affinity (Kd = 50 nM) for nucleosomes with flanking linker DNA (198-Nucs) than for nucleosomes with only core DNA (147-Nucs; Kd>160 nM) wrapped around, which suggests that the linker DNA is a main target of SSRP1 binding. This is in line with studies of the human SSRP1, which also requires linker DNA to stably interact with nucleosomes (Winkler et al. 2011b). Comparable to *Arabidopsis* SSRP1, the interaction of the human SSRP1 with large 207-Nucs (with linker DNA) had a Kd of 67 nM, whereas the interaction of hSSRP1 with small 147-Nucs (without linker DNA) had a much higher Kd of 1 μM (Winkler et al. 2011). Noteworthy, the affinity of the human FACT complex was therefore also higher for 207-Nucs (Kd = 22 nM) than for 147-Nucs (Kd = 64 nM) and neither Spt16 nor SSRP1 alone could bind nucleosomes as tightly as the complete FACT complex (Winkler et al. 2011b). The interaction of yFACT with nucleosomes in the presence of Nhp6 was comparable with a Kd of 50 nM. In contrast to the yeast and human FACT complex, the affinity of yeast FACT for nucleosomes was independent of the linker DNA (Formosa et al. 2001, Ruone et al. 2003). Most likely multiple *Arabidopsis* SSRP1 proteins can bind free DNA or linker DNA with their HMG-box domains. This results in multiple SSRP1 – DNA intermediates that appeared as a smear on the gel especially at high protein concentrations. The same was observed for yeast Nhp6 – DNA interactions (Ruone et al. 2003).

Recently, crystallographic studies revealed that the C-terminal acidic domains of yeast Spt16

and Pob3 are the primary determinants of H2A/B binding (Kemble et al. 2015). Both FACT subunits have overlapping binding sites on H2A-H2B, which allows the Spt16-Pob3 heterodimers to bind two H2A-H2B dimers (Kemble et al. 2015). The site on histone H2A-H2B are normally not accessible for FACT binding as they are located at the histone-DNA binding interface of nucleosomes (Hodges et al. 2017). These data suggest that the DNA binding by the HMG-box proteins SSRP1 (metazoan) or Nhp6 (yeast) in the first place, might enable the interaction of FACT with the histones H2A-H2B, that disrupts the DNA-histone interface. As the HMG-box domain of *Arabidopsis* SSRP1 mediates the binding to DNA, this hypothesis highlights the importance of this domain for FACT functionality.

6.2 Loss of the HMG-box domain did not alter the nuclear localization of SSRP1

In immunolocalization studies, *Arabidopsis* SSRP1 was previously found uniformly expressed in interphase nuclei with enrichment in euchromatic regions (Duroux et al. 2004). Moreover, C-terminal mRFP-tagged SSRP1 was found throughout the nuclei of Arabidopsis roots (Ikeda et al. 2011). In accordance, in-detail CLSM analysis of *Arabidopsis* cells expressing either eGFP-SSRP1 or eGFP-SSRP1△HMG revealed that both proteins are located in the nuclei of transgenic PSB-D cells. A few cells showed a strong ring-like accumulation of the GFP-fusion proteins at the periphery of the nucleolus, which is most likely an overexpression artefact. Since the cell lines do not arise from a single transformation event, in some cases multiple integrations of the transgene or T-DNA integrations in transcriptionally highly active regions are likely (Van Leene et al. 2015). These cells that are potentially overexpressing SSRP1 accumulate the fusion proteins in the nucleolus. Nevertheless, most importantly, there was no difference in the subcellular localization of SSRP1 and SSRP1△HMG observable.

6.3 Loss of the HMG-box domain did not alter the SSRP1 – chromatin binding properties *in vivo*

In vivo microscopy techniques allow to study the binding of nuclear proteins to chromatin in living cells (Houtsmuller et al. 2001, Lippincott-Schwartz et al. 2001). Photobleaching (FRAP) experiments shed light on the kinetics of many nuclear proteins. Strikingly, most nucleosomal histones show a low mobility/turnover due to their stable association with DNA (Kimura et al. 2001, Kimura 2005). Active processes like transcription can increase the histone turnover especially of the dimer H2A/H2B to make the DNA more accessible for other factors (Weber et al. 2014). In contrast, chromatin-associated proteins including histone chaperones show in general a high mobility and bind to chromatin only transiently (Phair et al. 2000, Pederson 2001, Dundr et al. 2002, Phair et al. 2004). Chromatin-associated proteins are scanning through the genome space to find their binding sites (Phair et al. 2004). In this study, FRAP experiments revealed that eGFP-SSRP1 is relatively mobile with a half-life time of approximately 2 seconds and showed only a small fraction (<10%) of immobile proteins. This is in accordance with the mean half-life time of chromatin-associated proteins that is typically ranging from 2 to 20 seconds (Phair et al. 2004). Thereby, eGFP-SSRP1 ($T_{1/2}$: 1.93 s) and eGFP-SSRP1△HMG ($T_{1/2}$: 1.79 s) is moving very fast through the nucleus, but still slower that the freely diffusing eGFP-NLS ($T_{1/2}$: 0.5 s). Noteworthy, due to the much smaller size, the diffusion speed of

eGFP-NLS (∼ 27 kDa) is higher than of eGFP-SSRP1 (∼ 100 kDa). The diffusion speed of the larger eGFP-SSRP1 can be estimated using the equation of Sprague et al. 2004. According to this calculation the diffusion speed of eGFP-SSRP1 would be approximately 0.8 seconds relative to the 0.5 seconds of eGFP alone. Therefore, eGFP-SSRP1 proteins moved slower in the nucleus as they could if they would freely diffuse. This suggests short term bindings of FACT to the chromatin in accordaning to its function as histone chaperone. But surprisingly, full length SSRP1 and SSRP1△HMG showed the same kinetics in the nuclei of living cells. The loss of the HMG-box domain did not alter the *in vivo* mobility of the protein. This suggests that SSRP1 can bind to chromatin even without the HMG-box domain that mediates the DNA-binding. In consequence, this suggests that other proteins may compensate the loss of the HMG-box function.

6.4 HMG-box-deficient SSRP1 was still part of the active transcript elongation complex *in vivo*

The *Arabidopsis* FACT complex consist of SSRP1 and SPT16 like in other higher eukaryotes (Duroux et al. 2004). The affinity purification of eGFP-SSRP1 and eGFP-SSRP1△HMG from whole cell extracts followed by immunoblotting revealed that SPT16 co-purified with both, the full length and the HMG-box deficient SSRP1. This suggest that SSRP1 and SSRP1△HMG can be incorporated into the FACT complex *in vivo*. This is in line with findings that both proteins dimerize with domains at the N-terminus of SSRP1 and in the middle of SPT16 as well as that the HMG-box domain is not necessary for the assembling of the heterodimer (Winkler et al. 2011b).

Transcript elongation factors (TEFs) are associated with the elongating, Ser2- phosphorylated form of RNAPII during transcription (Yoh et al. 2007, Sun et al. 2010). In accordance with the AP-MS experiments showing that the histone chaperone FACT is part of the *Arabidopsis* transcript elongation complex (TEC) (Antosz et al. 2017), the hyper-phosphorylated and elongating form of the RNAPII co-purified with eGFP-SSRP1. Strikingly, the elongating RNAPII also co-purified with eGFP-SSRP1△HMG. Even if the HMG-box domain of SSRP1 that mediates the DNA-binding is missing, the truncated SSRP1 is still involved in active transcription by RNAPII. This suggests that the SPT16/SSRP1△HMG complex is functional with most likely the help of other proteins that compensate the HMG-box function.

6.5 The HMG-box domain of SSRP1 was not important for proper plant development

This study revealed that the HMG-box domain of SSRP1 is important for mediating the binding to DNA and nucleosomes *in vitro*. Nevertheless, SSRP1 lacking the C-terminal HMG-box domain still associates with chromatin (FRAP) and with the elongating RNAPII (AP-Western) *in vivo*. This suggests that the HMG-box function can somehow be provided by another factor and it is not necessary that SSRP1 itself features the HMG-box domain. Therefore, complementation experiments were performed to figure out if the HMG-box domain must be physically fused to SSRP1 to function fully and to ensure proper plant growth and development.

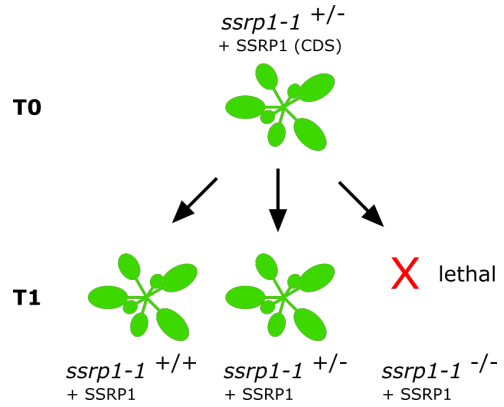


Figure 6.1 The coding sequence (CDS) of SSRP1 under its native promoter could not rescue *ssrp1-1* knockout mutants. Schematic illustration showing the segregation pattern of *ssrp1-1* plants expressing *wild type* SSRP1 (CDS). No homozygous *ssrp1-1* plants expressing the SSRP1 (CDS) complementation construct were obtained. This suggeted that the complementation constructs were not functional.

A first attempt to generate *Arabidopsis* HMG-box deficiency mutants was made by the previous PhD student Simon Mortensen. Briefly, transgenes that drive the expression of either full length SSRP1 (CDS) or truncated SSRP1△HMG (CDS) under control of the native SSRP1 promoter were introduced into *ssrp1-1* mutant plants (homozygous lethal). Phenotypic analysis revealed that full length SSRP1 could rescue the wild type phenotype (three independent lines), whereas the one analysed SSRP1△HMG mutant showed a strong dwarf phenotype (Mortensen 2010). These data suggested that the HMG-box domain of SSRP1 is essential for proper plant growth and development. However, comprehensive re-analysis of the SSRP1 complementation lines revealed that all three independent lines -expressing full length SSRP1 were lethal in the SSRP1 knockout background (*ssrp1-1*$^{-/-}$)(**Figure 6.1**). Therefore, the plants described in Mortensen 2010 were by mistake heterozygous for the *ssrp1-1* transposon insertion. Hence, the observed *wild type* phenotype of the described SSRP1 complementation lines is not due to functional complementation, but simply due to expression of endogenous SSRP1 (One allele of *SSRP1* is suffcient for WT appearance). Taken together, the observations described in Mortensen 2012 were disproven, as these findings were based upon experiments mistakenly performed in the heterozygous *ssrp1-1*$^{+/-}$ background. Since *SSRP1* (CDS) under control of its native promoter cannot rescue *ssrp1-1* mutants, this raised the question about the relevance of the one described dwarf mutant expressing truncated *SSRP1△HMG* (CDS) in the *ssrp1-1* background. The observed dwarf phenotype in this single case is most likely an artefact and not due to the lack of the SSRP1 HMG-box domain. This prompted the generation of an authentic HMG-box-deficiency mutant as performed in this study (**Section 2.5**). Thereby it was possible to show that the HMG-box domain of SSRP1 is not required for plant growth and development.

Studying the expression level of *SSRP1* in the complementation lines revealed two interesting aspects. First, the transposon insertion (GT7431; *ssrp1-1*) in the first exon of *SSRP1* did not completely knock out the expression *SSRP1*, but resulted in the low (relative to *wild type*) expression of a slightly aberrant *SSRP1* transcript. This alternative *SSRP1* transcript initiated from an ORF within the transposon, it was spliced like *WT SSRP1* and potentially generated a SSRP1 protein with an altered N-terminus. The first two amino acids of *wild type* SSRP1 (Met,

Ala) would be replaced by 20 transposon-derived aa. However, if the aberrant SSRP1 protein is produced or not, it is not sufficient for plant viability as homozygous *ssrp1-1* mutants are embryonic lethal.

Secondly, in all six SSRP1-complementation lines, the complementing transgene-derived *SSRP1* is stronger expressed ($\sim$ 8-14 fold) than the endogenous *SSRP1* in *wild type* plants. Therefore, the complemented *SSRP1* is overexpressed, even though its expression is driven by the native promoter, the region 614 nt upstream of the SSRP1 translation start site. The transgene is randomly inserted in the genome, therefore the overexpression of *SSRP1* could be due to T-DNA position effects or multiple insertions (Dean et al. 1988), but this is not very likely as the expression of *SSRP1* is upregulated consistently in all six independent complementation lines. Adjacent plasmid sequences could influence the expression of *SSRP1*, but also this is very unlikely as this was never observed for pGreen0179-derived T-DNAs (Hellens et al. 2000). Moreover, the lack of regulatory DNA sequences like silencers can influence the expression and stability of a certain transcript (Rose et al. 2008, Riethoven 2010). In the complementation lines, the transgene drives the expression of the genomic sequence of *SSRP1* under control of the putative native promoter and the CaMV 35S terminator. The reason that the expression of transgene-derived *SSRP1* is higher than the level of endogenous *SSRP1* could be that some regulatory cis-elements that are located upstream of the 614-nt promoter or in the native 3′UTR are important for proper SSRP1 expression and are missing in the transgene. Nevertheless, overexpression of neither SSRP1 nor SSRP1△HMG had a dominant negative effect on plant development.

The SSRP1 and SSRP1△HMG complementation lines looked like *wild type* plants. Just smaller phenotypic fluctuations were observed, that can be explained by T-DNA position effects (Dean et al. 1988) or minimal fluctuations of the growth conditions in the growth chamber. Considering the data of this study, the reported phenotypes of other *Arabidopsis* mutants (*ssrp1-2*, *ssrp1-3*) with truncations at the C-terminus are not linked to the loss/damage of the HMG-box domain. One mutant (*ssrp1-2*) harbours a T-DNA insertion at the end of the HMG-box domain ($\sim$ 13% of the DNA-binding domain are missing), which results in a generally knockdown of *SSRP1* expression. In view of the complementation experiments in this study, the pleiotropic phenotypes reported for *ssrp1-2* (Lolas et al. 2010) including bushy appearance and reduced seed set are most likely due to the reduced expression of *SSRP1* and not due to the small truncation of the HMG-box domain. The other mutant (*ssrp1-3*) harbours a point mutation that leads to a pre-mature stop codon and a large truncation of the C-terminus ($\sim$ 27% of SSRP1 are missing). The truncated SSRP1 is lacking the whole HMG-box domain, the adjacent NLS sequence as well as $\sim$ 65% of the acidic region. The *ssrp1-3* mutant shows a severely reduced growth and sterility (Ikeda et al. 2011). In light of the results of this study, the loss of the NLS sequence (Röttgers et al. 2000) and/or the impaired H2A/H2B binding acidic region (Kemble et al. 2015) are responsible for this phenotype and not the loss of the HMG-box domain.

The *wild type* phenotype was fully rescued by complementing the embryonic lethal *ssrp1-1* knockout mutant with full length SSRP1 and HMG-box deficient SSRP1. The *wild type* appearance of plants solely expressing the truncated SSRP1△HMG suggests that the loss of the HMG-box domain is dispensable and can be compensated somehow. This is in accordance

with cross-species experiments that showed that the knockout of Ssrp1 in the eye of *Drosophila* (causing severe defects) could be rescued by either full length ssrp1a (zebrafish) or truncated ssrp1a lacking the HMG-box domain (Koltowska et al. 2013).

6.6 Other HMGB proteins might compensate the loss of the SSRP1 HMG-box domain *in vivo*

In vitro, SSRP1 is not able to bind nucleosomes without the DNA-binding activities of the HMG-box domain. On the other hand, *in vivo*, SSRP1△HMG showed no defects in its chromatin-binding properties nor in its association with the elongating RNAPII as shown by FRAP and Immunoblotting analyses. Moreover, plants expressing HMG-box-deficient SSRP1 showed *wild type* appearance. These data suggest that the HMG-box domain of SSRP1 is dispensable *in vivo*. The yeast FACT complex consisting of SPT16/Pob3 (the equivalent to *Arabidopsis* SPT16/SSRP1△HMG) is loosely associated with the HMGB protein Nhp6, which provides the DNA-binding properties (Wittmeyer et al. 1999). Spt16/Pob3 alone without Nhp6 are also not able to bind nucleosomes (Formosa et al. 2001), but still yeast cells without Nhp6 are viable, which suggests that other HMGB proteins can compensate the loss (Stillman 2010). Moreover, the Nhp6 protein is not exclusively important for FACT-binding to nucleosomes, but also interacts with other chromatin modifiers (Stillman et al. 2010). In the genome of *Arabidopsis*, several chromosomal HMGB protein are encoded (Antosch et al. 2012), that could provide the DNA-binding activities for the SSRP1△HMG. SPT16/SSRP1△HMG could interact with a so far not identified HMGB protein to function as histone chaperone, like the yeast SPT16/Pob3 complex with Nhp6. This suggests that the mechanism by which the yeast FACT re-organizes nucleosomes with the loosely associated Nhp6 that provides the DNA-binding properties could be conserved in *Arabidopsis*. Even if the *Arabidopsis* SSRP1 is equipped with a C-terminal HMG-box domain, it is still able to function without the own physically fused DNA-binding domain and can revert on other HMGB proteins. This might be a general issue of SSRP1 proteins in higher eukaryotes. Moreover, recent studies revealed that diverse H2A/H2B histone chaperones use common mechanisms of histone binding (Kemble et al. 2015), which suggests that additional factors that are required for nucleosome binding like HMGB proteins might be shared among chromatin-modifiers like the binding mechanism. It might be interesting to figure out if, in view of evolution, the metazoan SSRP1 is a fusion of the yeast Pob3 and Nhp6 or if it is the other way around.

6.7 Outlook

This work demonstrated that the HMG-box domain of *Arabidopsis* SSRP1 was required for DNA/nucleosome binding *in vitro*. Whereas the loss of the C-terminal HMG-box domain had no effects on the investigated protein properties *in vivo* as well as on the plant growth and development. This suggests that the loss of SSRP1 HMG-box domain is dispensable *in vivo* and can be somehow compensated. One or more of the eight identified *Arabidopsis* HMGB proteins (Antosch et al. 2012) could provide the DNA-binding properties for the SPT16/SSRP1-HMG complex like the yeast Nhp6 for the Spt16/Pob3 complex. This suggests that yeast and metazoan FACT with their differences in the SSRP1 domain structure, share the same basic mechanism

to reorganize nucleosomes that are bound by HMGB proteins in the first place. Quantitative AP-MS analysis using plants cells expressing tagged SSRP1 in comparison to SSRP1-HMG could reveal candidate HMGB proteins that are enriched in the HMG-box-deficient samples. Moreover, analysis of double mutant plants defective in SSRP1 or SPT16 in combination with genes encoding the HMGB proteins could demonstrate genetic interactions that are detectable in synergistic effects on plant growth and development. In view of evolution, bioinformatic analyses could show if the metazoan SSRP1 is the result of a gene fusion of the yeast Pob3 and Nhp6 or if it happened the other way around by a gene separation.

7 Discussion: The role of the FACT complex in plant anthocyanin biosynthesis

7.1 A small subset of genes is differentially expressed in FACT-depleted plants

In eukaryotes, the transcript elongation by RNA polymerase II (RNAPII) is a dynamic and highly regulated step of the gene expression pathway. A variety of transcript elongation factors (TEFs) with diverse functions promote the progression of the transcription machinery through the chromatin template and ensure efficient mRNA synthesis of subsets of genes (Kwak et al. 2013, Selth et al. 2010, Sims et al. 2004). In *Arabidopsis*, the TEFs TFIIS, SPT4/SPT5, SPT6, PAF1-C and FACT associate with the elongating RNAPII (Antosz et al. 2017). The effect of TEF-depletion in *Arabidopsis* plants is ranging from mild phenotypes (e.g. TFIIS) (Grasser et al. 2009) to severe and lethal phenotypes (e.g. SPT5)(Dürr et al. 2014) which highlights their important roles for plant growth and development (Van Lijsebettens et al. 2014). The histone chaperone FACT facilitates transcript elongation by the disassembling of nucleosomes (Removing of H2A/B dimers) in the path of the RNAPII (Avvakumov et al. 2011, Formosa 2012). In *Arabidopsis*, the reduced expression of the FACT subunits SSRP1 or SPT16 results in similar phenotypes like early bolting (Lolas et al. 2010).

In yeast, the transcript elongation complex (TEC) is formed at the 5′ ends and disassembled at the 3′ ends of genes (Mayer et al. 2010). Thereby, the involved TEFs including the FACT complex showed a characteristic distribution over all transcribed RNAPII genes. The yeast FACT complex is associated genome-wide with all RNAPII transcribed genes and there are several lines of evidence that support the idea that the same is true for *Arabidopsis* as follows: In *Arabidopsis*, the FACT complex associates with the elongating RNAPII (Antosz et al. 2017) and was detected at transcribed regions in the euchromatin (Duroux et al. 2004, Lolas et al. 2010). Moreover, the microarray-based transcriptomic data in this study revealed that the reduced expression of *Arabidopsis SSRP1* or *SPT16* affected genome-wide only a small subset of genes. Relative to *wild type* plants, only 0.5% of the 24,000 tested genes were significantly (>2-fold; P<0.05) up- and 0.2% were downregulated in the *ssrp1-2* and *spt16-1* mutants. These observations are comparable to the depletion of other plant TEFs like TFIIS or SPT4/5 that are directly associated with the RNAPII and modulate their properties, for instance, to rescue backtrack tracked RNAPII (TFIIS) or to rendering RNAPII processivity (SPT4/5) (Van Lijsebettens et al. 2014). The inactivation of TFIIS, which has no obvious effects on plant morphology except a reduced seed dormancy, resulted in the differentially expression of 2.3% of the genome-wide tested genes (Grasser et al. 2009, Mortensen et al. 2011). The downregulation of SPT4/5 has more drastic effects on the plant growth and development, but still only 5% of the genome-wide tested genes especially involved in auxin-signalling were found to be differentially expressed (Dürr et al. 2014). Additionally, the transcript profiling of FACT-depleted plants is in accordance to studies in human cells, where the RNAi-mediated knockdown of SSRP1 and SPT16 resulted in subtle changes of the expression of a size-wise comparable small subset of genes (Li et al. 2007b). In generally, the knockdown of a particular TEF seems to alter locally the expression of a specific subset of genes than globally the expression of all genes.

Noteworthy, the FACT as well as the SPT4/5 complex are essential for plant viability (Duroux et al. 2004, Dürr et al. 2014). Therefore, the genome-wide changes of the transcriptome

were examined that occur due to a reduced expression of these TEFs and not due to a complete knockout. In the *Arabidopsis* FACT mutants, according to the microarray data, the transcript levels of *SSRP1* and *SPT16* were -1.84-fold and -1.23-fold reduced, respectively. These results are in line with formerly performed qRT-PCR analysis of *ssrp1-2* and *spt16-1* mutants (Lolas et al. 2010). The same is true for the before mentioned studies with FACT-depleted human cells (Li et al. 2007b). Therefore, the knockdown of the FACT complex reduces just the functionality of the histone chaperone and does not abolish it. Most likely, there are still sufficient amounts of FACT in the cells to ensure proper transcription in general and only the expression of a subset of genes that is highly sensitive towards FACT-depletion is affected.

Strikingly, a substantial part of the downregulated genes in SSRP1/SPT16-depleted plants was allocated to the anthocyanin biosynthesis pathway. FACT-depleted plants showed defects in the high light-mediated induction of several anthocyanin biosynthesis genes. Therefore, *ssrp1-2* and *spt16-1* mutants showed defects in the accumulation of the purple anthocyanin pigments in the leaves. Moreover, the expression of *SSRP1* and *SPT16* was upregulated in *wild type* plants in response to high light as shown by qRT-PCR and GUS reporter analysis. Most likely due to the accumulation of GUS over time (Jefferson 1987), the SSRP1/SPT16 promoter activity seems to be stronger in response to 14-days under moderate high light than upon 3 days under strong high light. However, both, moderate and strong high light induced elevated the promoter activity of *SSRP1* and *SPT16*. Taken together this suggested a novel role for the histone chaperone FACT in high light-induced expression of anthocyanin biosynthesis genes.

It seems that the expression of certain genes (or sets of genes) is more sensitive to the absence or deficiency of certain TEFs than others (Van Lijsebettens et al. 2014). Especially the expression of genes that are reactive to internal or external stimuli seems to be affected in TEF mutants (Van Lijsebettens et al. 2014). In line with this, SSRP1/SPT16-depleted plants showed an impaired light-induced expression of anthocyanin biosynthesis genes and therefore an anthocyanin-deficiency phenotype upon high light stress. Moreover, *ssrp1-2* and *spt16-1* mutants showed an early flowering phenotype. Therefore, it seems that the plants can somehow compensate the reduced amounts of FACT under standard growth conditions in the growth chamber and it has almost no effect on the transcription of most genes. But in some cases, especially by the transcription of genes in response to internal (flowering) or external (light) stimuli, the transcription machinery is more dependent on the *wild type* FACT level. In these cases, FACT-depletion is observable in the impaired expression of FLC (early flowering phenotype) or anthocyanin biosynthesis genes (anthocyanin-deficiency phenotype).

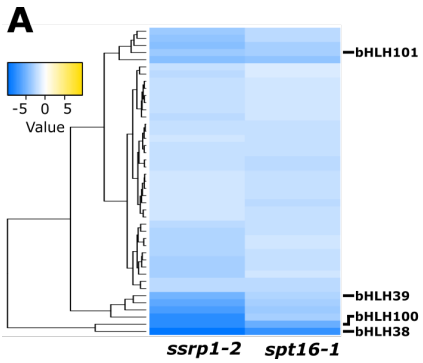
It is also unclear to which extend the histone chaperones are redundant (Elsässer et al. 2012). Next to the FACT complex there are also other H2A/H2B histone chaperones characterized in *Arabidopsis* like the NUCLEOSOME ASSEMBLY PROTEIN1 (NAP1) family that is associated with the Arabidopsis TEC or the Chz family with potentially overlapping functions. Furthermore, until now, the gene specific features or factors that determine the recruitment of TEFs to specific transcription sites are not known and will be focus of future research.

7.2 Genes are commonly differentially expressed in SSRP1/SPT16-depleted plants

Almost all affected genes in SSRP1 and SPT16-depleted plants were differentially expressed in the same way of being up- or downregulated. Both FACT mutants showed a huge overlap in up- and down-regulated genes, which is in accordance with the function of SSRP1 and SPT16 as a heterodimer. The transcriptome profiling of humans FACT-depleted cells showed the same trend of commonly up- and down-regulated genes (Li et al. 2007b). Moreover, the shared phenotypes of *ssrp1-2* and *spt16-1* (e.g. anthocyanin-deficiency, early flowering) further support that both proteins operate mainly in form of a heterodimer.

Most genes seem to be common targets of SSRP1 and SPT16, but still, slightly more genes were strongly differentially expressed (>2fold; P<0.05) in the *ssrp1-2* compared to the *spt16-1* mutant. For double mutant studies to examine the *Arabidopsis* TEC complex, FACT single mutants were crossed with plants deficient in other TEFs like TFIIS (Antosz et al. 2017). Thereby, the *ssrp1-2* version of the double mutant was always clearly more affected than the *spt16-1* version, which indicates that the *SSRP1* mutation is more critical for plant growth and development. Maybe this is because the T-DNA insertion in *ssrp1-2* disrupts the gene structure and results in a truncated SSRP1 protein, while the T-DNA in *spt16-1* is inserted in the promoter region (Lolas et al. 2010). Furthermore, *SSRP1* expression is more severely downregulated in *ssrp1-2* than *SPT16* expression in *spt16-1* as shown by microarray data as formerly performed qRT-PCR (Lolas et al. 2010).

7.3 Iron homeostasis genes are downregulated in FACT mutants



B

AGI	Gene	*ssrp1-2*	p-Value	*spt16-1*	p-Value
AT3G56970	*BHLH038*	-9.69	4.0E-05	-7.73	2.8E-04
AT3G56980	*BHLH039*	-5.26	2.9E-04	-3.00	7.2E-04
AT2G41240	*BHLH100*	-8.00	5.1E-07	-5.61	4.7E-03
AT5G04150	*BHLH101*	-3.60	3.3E-04	-3.35	2.9E-03

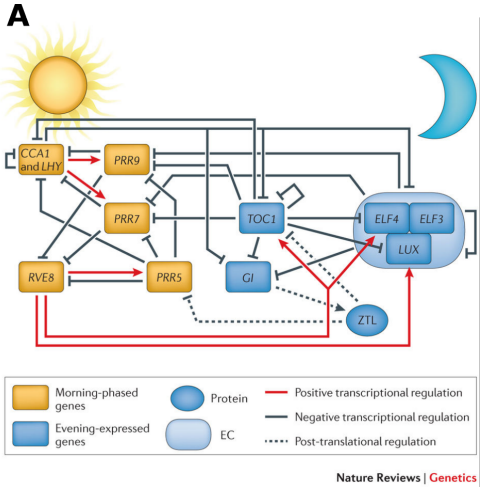
Figure 7.1 Regulatory genes for the iron hoeostasis in *Arabidopsis* were downregulated in FACT-depleted plants
A) Genes downregulated (≥2-fold; P<0.05) in *ssrp1-2* and/or *spt16-1* mutants relative to Col-0 were hierachical clustered and visualized as heatmap. All four members of the 'Ib bHLH subgroup' of transcription factors were among the 43 downregulated genes. Genes are represented as horizontal lines and different shades of blue indicate the degree of downregulation as shown in corresponding color calibration bar. B) The bHLH transcription factors are shown with corresponding microarray expression data of FACT mutants in comparison to wild type plants. Genes downregulated more than 2.0 fold are depicted in blue.

The complex regulation of iron homeostasis in *Arabidopsis* includes the combined action of several transcription factors like the basic helix-loop-helix (bHLH) protein FIT (FER-LIKE IRON DEFICIENCY INDUCED TRANSCRIPTION FACTOR) (Brumbarova et al. 2015). In *Arabidopsis*, FIT plays a central role in the iron uptake by controlling the transcription of the ferric-chelate reductase FRO2 and the iron transporter IRT1 (Colangelo et al. 2004). Thereby, FIT is forming heterodimers with one of the four 'Ib subgroup bHLH' transcription factors

bHLH038, bHLH039, bHLH100, or bHLH101 to activate the iron uptake in response to iron-deficiency (Yuan et al. 2008, Wang et al. 2013). The transcriptome profiling of In SSRP1/SPT16-depleted plants revealed that all four bHLH (38,39,100,101) transcription factors were strongly downregulated in the microarray experiment. Worth mentioning, just the aerial parts of *ssrp1-2* and *spt16-1* mutants were examined by the microarray analysis. However, the bHLH proteins (38,39,100,101) unlike FIT were not just found in the roots (site of iron uptake) but also in the leaves of *Arabidopsis* (Andriankaja et al. 2014, Brumbarova et al. 2015). These data suggest that the FACT complex is putatively involved in the iron homeostasis of *Arabidopsis*.

7.4 Circadian clock genes are differentially expressed in FACT mutants

A



B

AGI	Gene	*ssrp1-2*	P-Value	*spt16-1*	P-Value	
AT2G46830	*CCA1*	-1.76	0.0001	-2.47	0.0001	Morning expressed genes
AT1G01060	*LHY*	-1.57	0.0010	-2.02	0.0001	
AT3G09600	*RVE8*	-1.89	0.0030	-2.30	0.0003	
AT5G60100	*PRR3*	1.39	0.0969	3.50	0.0002	
AT5G24470	*PRR5*	2.53	0.0001	4.44	1.18E-05	
AT5G02810	*PRR7*	1.14	0.0152	1.21	0.0003	
AT2G46790	*PRR9*	-1.24	0.1094	-1.10	0.0268	
AT5G61380	*TOC1*	1.55	3.40E-05	1.99	2.42E-05	
AT1G22770	*GI*	1.76	2.07E-05	2.31	1.79E-05	
AT5G08330	*CHE*	1.07	0.0990	-1.01	0.8343	Evening expressed genes
AT2G25930	*ELF3*	1.17	0.0055	1.28	0.0049	
AT2G40080	*ELF4*	1.06	0.5845	1.69	0.0233	
AT3G46640	*LUX*	2.03	0.0004	2.83	0.0005	
AT5G64170	*LNK1*	-1.20	0.0162	-1.14	0.0057	
AT3G54500	*LNK2*	-1.33	0.0021	-1.61	0.0009	
AT3G12320	*LNK3*	-1.60	0.0048	-1.94	0.0008	
AT5G06980	*LNK4*	-1.21	0.0012	-1.79	0.0015	
AT3G22380	*TIC*	1.20	0.0599	1.31	0.0142	
AT1G01520	*RVE3*	-1.81	0.1108	-1.76	0.0447	

Figure 7.2 SSRP1/SPT16-depleted plants showed a downregulation of morning-phased and an upregulation of daytime/evening-phased genes of the circadian clock. A)Schematic illustration of the Arabidopsis circadian clock from Greenham et al. 2015. B) Circadian clock genes were shown with the respective microarray data of SSRP1/SPT16-depleted plants in comparison to *wild type* Col-0. Genes that were downregulated more than 1.5 fold were depicted in blue, whereas genes that were upregulated more than 1.5 fold were depicted in yellow. Non-significant values were written in red.

The circadian clock is an endogenous time-keeping mechanism that integrates multiple environmental stimuli that the sessile plants cannot avoid (Greenham et al. 2015). In the complex regulatory network of the *Arabidopsis* circadian clock, the morning expressed genes CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) and LATE ELONGATED HYPOCOTYL (LHY) repress the evening expressed genes TIMING OF CAB EXPRESSION 1 (TOC1) and PSEUDO RESPONSE REGULATOR 1 (PRR1) as well as activate the daytime expressed genes PSEUDO RESPONSE REGULATOR 9 (PRR9) and PRR7. In turn, PRR9/7 together with PRR5 repress the morning-phased genes CCA1 and LHY as well as REVEILLE 8 (RVE8) that induces the expression of the evening expressed genes, including TIMING OF CAB EXPRESSION 1 (TOC1) and LUX ARRHYTHMO (LUX). In response, the evening-phased genes repress the morning and daytime expressed genes (Greenham et al. 2015, Pérez-García et al. 2015). In SSRP1/SPT16-depleted plants, several morning-expressed genes (CCA1, LHY, RVE8) are commonly downregulated, whereas certain daytime and evening expressed genes (PRR5, TOC1, GI, LUX) are commonly up-regulated. This suggests an involvement of the FACT complex in the regulatory network of the circadian clock. Moreover, the circadian clock components RVE8 and

LNK were shown to shape the oscillation of the anthocyanin metabolic pathway (Pérez-García et al. 2015). In FACT-depleted plants, RVE8 and LNKs were downregulated. Moreover, FACT mutants showed defects in the expression of anthocyanin synthesis genes as shown in this study. Therefore, FACT could also indirectly influence anthocyanin biosynthesis by affecting the expression of circadian clock genes. FACT-induced defects in the circadian clock could result in the impaired integration of external signals as the light-induced synthesis of anthocyanin.

7.5 Outlook

Transcript profiling of SSRP1- and SPT16-depleted *Arabidopsis* plants revealed that only a small subset of genes was differentially expressed in comparison to *wild type*. Strikingly, a substantial part of the downregulated genes was allocated to the anthocyanin biosynthesis pathway and follow-up experiments indicated that FACT is a novel factor required for the accumulation of anthocyanins in response to light-induction. Following ChIP experiments of HL-treated plants in comparison to control plants using SSRP1 and SPT16 ABs could show if FACT is really recruited to the loci of anthocyanin biosynthesis genes and provide biochemical evidence for the observed light-induced anthocyanin-deficiency phenotype of FACT mutants. Moreover, ChIP-seq (chromatin immunoprecipitation [ChIP] combined with high-throughput sequencing [seq]) experiments could reveal the genome-wide distribution of FACT with chromatin and add an important piece to the puzzle how the recruitment of FACT to specific transcription sites is determined? Additionally, the transcriptomic data of SSRP1- and SPT16-depleted plants suggested an involvement of FACT in the regulatory network of the circadian clock. A research field that was recently awarded by the Nobel Prize in Physiology or Medicine (2017). Following direct protein-protein studies between FACT and components of the circadian clock as well as ChIP analysis of clock genes using SSRP1/SPT16 ABs could reveal a functional link.

8 Discussion: The *Arabidopsis* SAGA-DUBm component ENY2

8.1 The expression of *ENY2* in *Arabidopsis thaliana*

To visualize the expression level of *ENY2* in *Arabidopsis* plants, transgenic *ENY2-GUS* reporter lines were generated and analysed by GUS staining (**Figure 5.2**). To determine the putative authentic *ENY2* expression pattern, three independent transgenic *ENY2-GUS* lines were analysed simultaneously. In most tissues, the GUS expression patterns were completely overlapping and just slightly varying in their intensities. Just the line #4 showed GUS staining in additional tissues (shoot meristem and reproductive organs) that was not observed in the other two transgenic lines. The smaller differences in th GUS signal intensity and the discrepancies in GUS signal pattern of line #4 were most likely due to the different and random integration of the T-DNA in the plant genomes (Dean et al. 1988). To describe the potential authentic expression pattern of plant *ENY2*, just the GUS signals observed in at least two out of the three investigated ENY2-GUS lines were considered.
These GUS experiments revealed that *ENY2* is widely expressed, from young seedlings to mature plants. These findings are supported by the data from *Drosophila* and vertebrates, where *ENY2* is also ubiquitously expressed (Georgieva et al. 2001, Krasnov et al. 2005).
The ubiquitous expression pattern of *ENY2* in *Arabidopsis* roots was also confirmed by the analysis of transgenic plant lines expressing eGFP-ENY2 under the putative native promoter of *ENY2*. CLSM analysis showed that ENY2 is strongly expressed in the root vasculature but almost absent of the root meristem. In all three independent lines, the signal became stronger in the more differentiated cells like the epidermis or stele of the differentiation zone. The eGFP-ENY2 fluorescence signals in the *Arabidopsis* roots were matching with the ENY2-GUS pattern discussed before, what indicates that *ENY2* transcript and ENY2 protein showed the same spatio-temporal distribution in the root tissues of young seedlings.
The *ENY2* promoter is very active in the vasculature of roots. This could suggest a role for *ENY2* in root growth and architecture. Recently, plant microRNAs were identified as key regulators of root architecture by fine tuning gene expression. Thereby, the modulation of the gene expression enables plants to adapt their root system to changing environmental conditions (Stauffer et al. 2014, Couzigou et al. 2016). It is conceivable that *ENY2* plays a similar role in modulating gene expression to respond to biotic and abiotic stimuli.

8.2 The interaction network of ENY2 in *Arabidopsis thaliana*

In living *Arabidopsis* cells, ENY2 is localized in the nucleoplasm (**Figure 5.6**). In this compartment, plant ENY2 is highly mobile and could maintain several short-term bindings to other proteins as suggested by FRAP experiments (**Figure 5.7**). This is in accordance to its known role as small adaptor protein in other species. In yeast, *Drosophila* and vertebrates, ENY2 is described as a shared component of several multi-protein complexes that is involved in different steps of the gene expression pathway. As component of both the SAGA-DUB module and the TREX-2 complex, ENY2 is linking the two processes ′transcription activation′ and ′mRNA export′, respectively (Rodríguez-Navarro et al. 2004, Kurshakova et al. 2007).

8.2.1 Old friends: ENY2 and the DUB module of the transcriptional co-activator SAGA

8.2.1.1 Composition and assembly of the putative plant DUB Module is conserved

In eukaryotes, the transcriptional co-activator SAGA consists of four modules. Each subcomplex has distinct functions, allowing the multiprotein complex to do both, modify chromatin and recognize histone modifications (Koutelou et al. 2010, Rodríguez-Navarro 2009, Samara et al. 2011). The SAGA-DUB module is highly conserved and consists of four proteins as following: The deubiquitinating enzyme yUbp8/dNonstop/hUSP22 (Yeast/Drosophila/Humans), the two adaptor proteins ySgf11/dSgf11/hATXN7L3 and ySus1/dENY2/hENY2 as well as the anchoring protein ySgf73/dSgf73/hATXN7 (Köhler et al. 2008, Weake et al. 2008a, Zhao et al. 2008, Lee et al. 2009, Samara et al. 2010, Gurskiy et al. 2012). Except for Sgf73/ATXN7 that links the DUB module to the SAGA complex, the putative orthologs were predicted in higher plants (Srivastava et al. 2015).

To shed light on the SAGA-DUB module in *Arabidopsis*, AP-MS experiments were conducted (**Tables 5.1 and 5.4**). All three predicted DUB components (ENY2, SGF11 and UBP22) were co-purified with ENY2-SG, SGF11-SG, and UBP22-SG. These biochemical data indicated that a putative DUB module lacking SGF73 exists in plants. Moreover, AP followed by gel filtration and coupled to Western Blot and MS analysis (**Figure 5.22**) showed that *Arabidopsis* ENY2, SGF11 and UBP22 were forming a stable complex *in vivo*. All three subunits of the potential plant DUB module co-eluted in the same low-molecular weight fractions from the superose6 column. The putative *Arabidopsis* DUB module and its connection to the plant SAGA complex will be discussed later in section 8.2.5.

The crystal structures of the yeast SAGA DUB module revealed that Ubp8, Sus1, Sgf11 and Sgf73 are forming a highly interconnected complex, where each protein physically contacts the other three (Köhler et al. 2010, Samara et al. 2010). Thereby, the DUB module is organized into two functional lobes. The ′catalytic lobe′ consists of the USP (Ubiquitin-Specific Protease) domain of Ubp8 and the ZnF (Zink-Finger) domain of Sgf11. The ′assembly lobe′ consists of the ZnF domain of Ubp8, the N-terminal regions of Sgf11 and Sgf73 as well as Sus1. By spanning the two lobes, Sgf11 makes extensive contacts to Ubp8. Sus1 functions as molecular clamp and stabilizes the Ubp8-Sgf11 interaction. Therefore, Sus1 wraps completely around the long alpha helix of Sgf11 and contacts Ubp8 just with its distal ends (Jani et al. 2009, Ellisdon et al. 2010, Köhler et al. 2010, Samara et al. 2010).

Consistent with the spatial arrangement of the yeast DUB module, Y2H (**Figure 5.10**) and FRET (**Figure 5.12**) analysis revealed direct protein-protein interactions between AtSGF11 and AtENY2 as well as AtSGF11 and AtUBP8, but not between AtENY2 and AtUBP8. These findings were supported by immunoprecipitation experiments in metazoans that showed interactions between ATXN7L3 and USP22 as well as between ATXN7L3 and ENY2, but not between ENY2 and USP22 (Zhao et al. 2008). The measured FRET efficiencies for SGF11/ENY2 and UBP22/SGF11 were 16% and 8%, respectively. The observed differences were most likely due to the fact that the FRET efficiency strongly depends on the distance and the relative orientation of the donor-acceptor pair (Piston et al. 2007, Chudakov et al. 2010). These data together with the protein-protein complex structure prediction by SPRING (**Figure 5.16**) support that the assembly of the yeast and human DUB module is conserved in *Arabidopsis*.

Therefore, it is not surprising that multiple sequence alignments (**Figure 5.13, 5.14, 5.15**) and homology modellings (**Figure 5.16**) showed that the protein sequences and structures of the putative *Arabidopsis* DUBm subunits were highly conserved. For example, the tertiary structure of ENY2 is basing on five highly conserved a-helices separated by flexible hinge regions. The structurally important glycine residues that function as hinges between the alpha-helices $\alpha1/\alpha2$ and $\alpha2/\alpha3$ and the proline residue in alpha-helix 4 that intriduces a kink (Jani et al. 2009) are conserved in *Arabidopsis* as well (**Figure 5.13**). This hand-like structure allows ENY2 to grab long α-helices, like Sgf11/SGF11 (SAGA-DUBm) and Sac3/GANP (TREX-2 component) (Jani et al. 2009, Ellisdon et al. 2010, Samara et al. 2010, Jani et al. 2012, Morgan et al. 2016). The binding is not sequence specific, but Sus1/ENY2 recognizes the overall shape and hydrophobic surface pattern of the target helix (Köhler et al. 2010).

Taken together, these data indicate that the three highly conserved proteins ENY2, SGF11 and UBP22 are forming a stable putative DUB module in *Arabidopsis*.

8.2.1.2 De-ubiquitinating activity of the plant DUB module

As shown above, the composition and the assembly including the catalytic USP domain of the putative DUB module are highly conserved in *Arabidopsis*. As the shape of a protein complex determines its function, this suggests that also the de-ubiquitinating activity of the module is conserved.

The monoubiquitination of histone H2B (H2Bub) is a universal feature of actively transcribed genes by RNA Polymerase II (RNAPII) (Bonnet et al. 2014). In yeast, *Drosophila* and mammals, the deubiquitinating (DUB) module of the SAGA complex is an important regulator of H2Bub levels (Henry et al. 2003, Lang et al. 2011, Mohan et al. 2014, Weake et al. 2008a) and recent crystal structures revealed how the DUB module is docking on the H2A/H2B acidic patch with the ZnF domain of Sgf11 and the catalytic USP domain of Ubp8 (Morgan et al. 2016).

In accordance with its putative function in H2B de-ubiquitination, the histones H2A and H2B were co-purified with SGF11-SG and UBP22-SG in AP-MS experiments (**Tables 5.1 and 5.4**). No histones were co-purified with ENY2-SG, which could be due to the transient nature of the DUBm – histone interaction. Strikingly, the global level of H2Bub was strongly increased in *sgf11-1* knockout mutants and slightly increased in *eny2-RNAi* knockdown mutants (**Figure ??**). This is in line with the findings that the catalytic activity of Ubp8 depends on the presence of the other DUB subunits (Köhler et al. 2008, Lee et al. 2009). It is typical for de-ubiquitinating enzymes to depend on partner proteins for their full enzymatic activity (Samara NL et al. 2012). In summary, this indicates that the *Arabidopsis* DUB module has de-ubiquitinating activity. Moreover, the adaptor proteins SGF11 and ENY2 are necessary for full catalytic activity of the DUB module.

8.2.2 The composition of the *Arabidopsis* SAGA complex

As described before, the multiprotein SAGA complex is comprised of four modules (DUB, HAT, SPT and TAF). In *Arabidopsis*, 17 potential candidates for SAGA subunits were identified *in silico* (Moraga et al. 2015, Srivastava et al. 2015), but no biochemical data are available. This suggests that next to yeast, fruit-fly and mammals (Rodríguez-Navarro 2009, Koutelou et al.

2010, Spedale et al. 2012), the SAGA complex is also conserved in *Arabidopsis*.
Surprisingly, with ENY2-SG, SGF11-SG and UBP22-SG the plant DUB module (ENY2, SGF11, UBP22) co-purified very robustly, whereas almost no proteins from the other SAGA modules were identified (**Tables 5.1 and 5.4**).

This raised the questions if the predicted SAGA complex really exists in *Arabidopsis* and if yes, is the DUB module a permanent part of it? To answer the questions and to clarify the situation of the plant SAGA complex, one representative of each SAGA-module was selected for further AP-MS analysis as following: UBP22 (DUB module), ADA2b (HAT module), SPT3/TAF13 (SPT module), TAF10 (TAF module). This proteomic study confirmed the existence of 17 predicted SAGA subunits in *Arabidopsis*. The plant DUB module consists of ENY2, SGF11, UBP22, the HAT module of GCN5, ADA2b, ADA3, SGF29 (SGF29a and SGF29b), the SPT module of ADA1 (ADA1a and ADA1b), TRA1 (TRA1a and TRA1b), SPT20, SPT7, SPT3/TAF13 and the TAF module of TAF5, TAF6 (TAF6a and TAF6b), TAF9, TAF10 and TAF12. In **Figure 8.1**, the identified plant SAGA components were compared to their yeast and human counterparts.

In yeast, the chromatin remodelling protein Chd1 (CHROMO-ATPASE/HELICASE-DNA BINDING DOMAIN 1) was identified as a additional component of the SAGA complex (Pray-Grant et al. 2005). The yeast and the human Chd1 can bind to the methylated histone H3K4 with a CHROMO domain (Flanagan et al. 2005). The homologous *Arabidopsis* protein CHR5 (AT2G13370), which was assumed to be involved in the plant SAGA complex (Moraga et al. 2015, Srivastava et al. 2015), did not co-purify with ADA2b, SPT3/TAF13, TAF10 or the DUB subunits in the AP-MS experiments. A BLAST search revealed another CHD1 homologous protein in *Arabidopsis* called CHD1-like/CHR6 (AT2G25170), which co-purified with TAF13-SG and UBP22-SG. Therefore, the mass spectrometry data suggest that the CHD1-like protein might act as a reader of H3K4 modifications for the plant SAGA complex.

Therefore, the biochemical approach confirmed the conservation of the SAGA complex in *Arabidopsis*, but revealed some structural differences to the proposed models for the yeast, *Drosophila* and the human SAGA complex.

8.2.3 The evolution of the *Arabidopsis* SAGA complex

During evolution, the yeast SAGA complex has diverged into the two related complexes, SAGA and ATAC, that play distinct roles during transcription (Spedale et al. 2012). Nine ATAC-specific subunits were described in *H. sapiens* and ten in *Drosophila* (Spedale et al. 2012). In *Arabidopsis*, there was just one protein (ADA2a) with a high sequence similarity identified by BLAST search. In the AP-MS experiments, ADA2a did not co-purify with the SAGA proteins ADA2b, SPT3/TAF13 and TAF13 or with any of the DUB subunits. In *Drosophila* and humans, the ADA2a homologs were essential for normal development (Pankotai et al. 2005), whereas knockout of the *Arabidopsis* ADA2a showed no phenotype (Hark et al. 2009). In comparison ADA2b (SAGA HAT module) depleted plants showed strong pleiotropic developmental defects including dwarfism (Vlachonasios et al. 2003). The AP-MS data suggested that AtADA2a and AtADA2b were not part of the same protein complex in *Arabidopsis*. While AtADA2b seems to be part of the SAGA histone acetyltransferase (HAT) module, the role of AtADA2a is not

known. One possibility is that the *Arabidopsis* SAGA complex is diverged into two related complexes containing either ADA2a (ATAC complex) or ADA2b (SAGA complex). In this case the composition of the ADA2a-containing complex is clearly different from the *Drospophila* and human counterparts as the other ATAC proteins were not conserved in *Arabidopsis*. The other possibility is that the *Arabidopsis* SAGA complex did not diverge into two complexes and no ATAC-like complex evolved in plants.

8.2.4 The SAGA complex versus the TFIID complex

In general, the yeast TFIID complex promotes the expression of housekeeping genes with TATA-like elements, whereas the yeast SAGA complex is responsible for the expression of regulatable/responsive genes with a consensus TATA-box (Jonge et al. 2016). Both co-activators, the SAGA and the TFIID complex, shared several subunits, the so-called TATA-binding protein (TBP) - associated factors (TAF′s)(Wu et al. 2004). The putative composition of the *Arabidopsis* general transcription factor TFIID was revealed by yeast-two hybrid studies and 16 putative subunits were identified (Lawit et al. 2007).

In the literature, AtTAF13 was described as component of the plant TFIID complex (Lawit et al. 2007) and as the putative homolog of the yeast and human Spt3/SPT3 (SAGA-SPT module)(Moraga et al. 2015, Srivastava et al. 2015). A de novo BLAST search confirmed that AtTAF13 is the only *Arabidopsis* protein that shares a sequence identity with the ySpt3/hSPT3 protein. Therefore, in yeast and humans, two distinct proteins exist that are part of the SAGA-SPT module (Spt3/SPT3) and the TFIID complex (Taf13/TAF13). In contrast, in *Arabidopsis*, just one protein (AtTAF13) was identified that is potentially part of both co-activator complexes. Taken together, the SAGA-TAF component AtTAF10 and the SAGA-SPT component TAF13/SPT3 are potential shared components of the SAGA and the TFIID complex. Whereas the SAGA-HAT component ADA2b and the DUB components are putative SAGA-specific.

The AP-MS analysis using SAGA-specific bait proteins (ADA2b, ENY2, SGF11, UBP22) and SG-tagged shared components of SAGA and TFIID (SPT3/TAF13 and TAF10) revealed some interesting findings. Of the 16 predicted TFIID components (Lawit et al. 2007), 13 (TBP1, TBP2, TAF1, TAF2, TAF4, TAF5, TAF6a-b, TAF8, TAF9, TAF10, TAF11, TAF12a-b, TAF13) were biochemically identified. Therefore, all five components of the *Arabidopsis* SAGA-TAF module (TAF5, TAF6a-b, TAF9, TAF10 and TAF12a-b) as well as two proteins of the SAGA-SPT module (SPT3/TAF13 and SPT7/TAF1) that were identified by the AP-MS approach in this study were shared subunits of the SAGA and TFIID complex. Moreover, the AP-MS experiments showed that with the SAGA-specific ADA2b (HAT module) just SAGA-specific and SAGA/TFIID-shared components were co-purified. With the SAGA/TFIID-shared components SPT3/TAF13 and TAF10, SAGA-specific, shared components, and TFIID-specific components were co-purified.

The finding that TAF13 was not co-purified with the SAGA-specific ADA2b might indicate that TAF13 is no part of the SAGA complex. But on the other hand, SAGA- as well as TFIID- specific components were co-purified with TAF13. This indicates that TAF13 is most likely identical to the yeast and human Spt3/SPT3 and is therefore a shared compenent of the *Arabidopsis* TFIID and the SAGA complex.

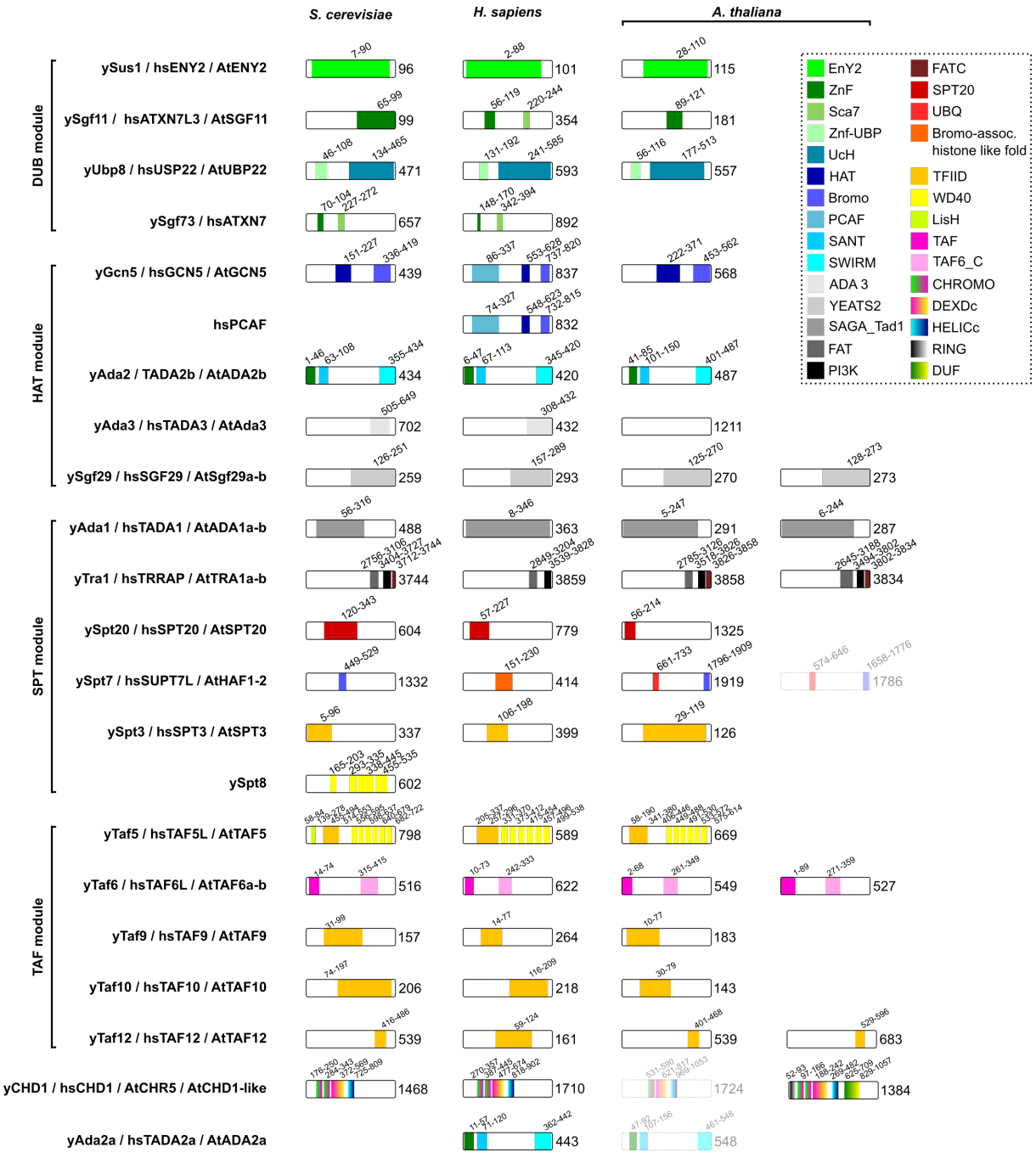


Figure 8.1 The composition of the *Arabidopsis* SAGA complex as revealed by AP-MS analysis in comparison to the published yeast and human counterparts. The bioinformatically predicted *Arabidopsis* SAGA proteins (Moraga et al. 2015, Srivastava et al. 2015) that were not identified biochemically were shown transparent. The domain organization of the *S.cerevisae* (*y*) and *H.sapiens* (*hs*) proteins is basing on the published data from Spedale et al. 2012. The *Arabidopsis* (*At*) protein domains have been assigned using SMART (Letunic et al. 2017) and UniProt (Apweiler et al. 2004).

8.2.5 Permanent relationship or temporary liaison? The *Arabidopsis* SAGA complex and its DUB module

The AP-MS analysis of ENY2-SG, SGF11-SG and UBP22 revealed that almost no subunits of the HAT, SPT or TAF module co-purified with the DUB subunits. And vice versa, no DUB subunit co-purified with SPT3/TAF13-SG (SPT) and TAF10-SG (TAF) and just ENY2 co-purified with ADA2b-SG (HAT) in two out of three replicates. The proteomic data suggest that the DUB module is no permanent module of SAGA complex, but can also exist independently of the remaining SAGA complex (HAT, SPT, TAF). In this experimental setup, the composition of the purified SAGA complex could be influenced by the accumulation of un-complexed bait proteins (stoichiometric imbalances) due to expression of the baits under control of the CaMV 35S promoter (Gibson et al. 2013, Van Leene et al. 2015). Although no obvious accumulation of the bait proteins was visible in the input samples (**Figures 5.8 C and 5.20 C**) and the levels of the bait proteins are often in the range of the endogenous protein (Van Leene et al. 2007, Van Leene et al. 2015), this cannot be completely excluded.

To further investigate the composition of plant SAGA complex, the protein complexes that co-purified with SG-tagged bait proteins were fractionated on a gel-filtration column according to their sizes. Immunoblot (**Figure 5.22**) and mass spectrometry (**Table 5.4**) analyses of low molecular weight fractions confirmed that ENY2, SGF11 and UBP22 co-eluted from the superose6 column. This indicated that the DUB module was stably assembled *in vivo* in *Arabidopsis*. However, immunoblotting of high molecular weight fractions revealed that all three DUB subunits also co-eluted there together with the SAGA protein ADA2b. The strong association of all three DUB subunits with the huge splicing machinery was shown before by AP-MS experiments (**Tables 5.1 and 5.4**). In accordance to these findings, the MS analysis of the high molecular weight fractions of the SGF11-GS and ADA2b-SG AP′s showed that especially SGF11 is strongly associated with proteins of the spliceosome. Therefore, the fractionation experiments (AP-Superose6-Western/MS) could differentiate if the DUB components were present at the high molecular weight fractions because of the association with the SAGA complex or with the spliceosome. Noteworthy, most likely due to instability of the protein complexes during the purification procedure, monomeric baits were prominently detected in the low molecular weight fractions.

In yeast, fruit-fly and humans, the conserved SAGA-DUB module consists of four proteins (Köhler et al. 2008, Weake et al. 2008a, Zhao et al. 2008, Lee et al. 2009, Samara et al. 2010, Gurskiy et al. 2012). The *Arabidopsis* DUB module is lacking the Sgf73/ATXN7 subunit. In yeast, Sgf73 is important for anchoring the DUB module to the remaining SAGA complex (Han et al. 2014). Consistently, complete or partial deletions of Sgf73 disconnect the DUB module from the SAGA complex (Durand et al. 2014). The lack of the SAGA âĂŞ DUB linking protein SGF73 in *Arabidopsis*, supports the conclusions of the AP-MS experiments that the DUB module is not permanently connected to the remaining *Arabidopsis* SAGA complex (HAT, SPT, TAF). In accordance with this, there are several lines of evidence that the DUB module has SAGA-independent functions in other organisms, too. Most strikingly, recent ChIP-seq analysis in *Drosophila* revealed that the DUB module can bind to chromatin and regulate transcription independently of the HAT module, which suggests that the *Drosophila* DUB module has SAGA-

dependent and independent functions (Li et al. 2017). In mammalian cells, ATXN7L3 (SGF11) and ENY2 orchestrate the functions of multiple DUBs: with USP22 they form the SAGA-DUB module and with USP27X and USP51 they can form two further SAGA-independent DUBs that also act on H2B K120ub1 (Atanassov et al. 2016). Another study in yeast showed that the DUB module can dissociates from the SAGA complex by interaction with the proteasome and can interact with chromatin in a SAGA-independent manner (Lim et al. 2013). In summary, the existence of catalytically active SAGA-independent DUBs in yeast, *Drosophila* and humans supports the hypothesis that the plant DUB module can act on histones independently of the remaining SAGA complex.

In accordance to the data of (Atanassov et al. 2016), plant AtENY2 and AtSGF11 could also form more than one DUB module with other deubiquitinating enzymes. In *Arabidopsis*, five different DUB (deubiquitylating enzymes) families were identified and UBP/USP (the ubiquitin-specific proteases) is largest with 27 members (Yan et al. 2000, Isono et al. 2014). The AP-MS experiments revealed that next to UBP22, also UBP5 co-purified with SGF11-SG (**Table 5.1**). This opens the possibility that ENY2 and SGF11 do not just form one DUB module. In mammalian cells, the knockout of UBP22 resulted unexpectedly in a decrease of the global H2Bub1 levels, because the loss of UBP22 could be compensated by two other USP proteins (SAGA-indepedendent). In comparison, the deletion of the adaptors, ATXN7L3 of ENY2 increased the bulk levels of H2Bub1 (Atanassov et al. 2016). In contrast, knockout of the USP22 homologs in yeast (Ubp8) and *Drosophila* (Nonstop) results in the expected increase of global H2Bub levels (Ingvarsdottir et al. 2005, Weake et al. 2008a). In *Arabidopsis*, knockout of SGF11 and knockdown of ENY2 results in globally increased H2Bub levels (**Figure 5.32**). It will be attractive to examine in future experiments how the global H2Bub level is affected in UBP22 knockout plants to learn if ENY2 and SGF11 can form more than one DUB module.

Taken together, our analyses revealed that the *Arabidopsis* SAGA complex consists of 17 subunits that are divided in four modules as following: DUB module (ENY2, SGF11, UBP22), HAT module (GCN5, ADA2a-b, ADA3, SGF29a-b), SPT module (ADA1a-b, TRA1a-b, SPT20, SPT7, SPT3/TAF13), TAF module (TAF5, TAF6a-b, TAF9, TAF10, TAF12) and the potentially associated protein CHD1-like. Moreover, the *Arabidopsis* DUB module can most likely function SAGA-dependent and independent.

8.2.6 The *Arabidopsis* SAGA complex during transcription by RNAPII

The SAGA complex with its histone-modifying activities acts as transcriptional co-activator that facilitates transcript initiation and elongation (Koutelou et al. 2010, Weake et al. 2012). The HAT module of the SAGA complex mediates the histone acetylation of gene promoters to open the repressive chromatin structure and enhance PIC formation and transcription activation (Balasubramanian et al. 2002, Mohibullah et al. 2008). Moreover, the DUB module of the SAGA complex facilitates transcript elongation by deubiquitination of histones downstream of RNAPII (Wyce et al. 2007). Additionally, transcript elongation is promoted by acetylation of histones within the coding region to evict nucleosomes (Govind et al. 2007). In accordance with these data from other species, several subunits of the RNAPII and transcript elongation factors (TEFs) including SPT16, SPT4/5 and the PAF1-C were robustly co-purified with the *Arabidopsis* SAGA

complex.

Moreover, the RNAPII preinitiation complexes (PIC) coordinates the interplay of the transcription and the chromatin machineries. Studies with human cells revealed that the PIC contains two major co-activators, the Mediator and the SAGA complex (Chen et al. 2012). Thereby, the Mediator coordinates the assembly of the RNAPII initiation factors and chromatin modifiers into a PIC, whereas the SAGA acts after the PIC formation and promotes transcription on chromatin (Chen et al. 2012) as discussed before. Additionally, the SAGA complex was identified as Mediator-interacting protein complex in a recent proteomic analysis in yeast (Uthe et al. 2017). In line with these findings, several subunits of the Mediator complex and chromatin modifiers including SWI/SNF and SWR1 were co-purified with the *Arabidopsis* SAGA complex.

8.2.7 No friends anymore: ENY2 and the mRNA export complex TREX-2

In yeast, *Drosophila* and mammals, Sus1/ENY2 is part of the mRNA export complex TREX-2/AMEX (Fischer et al. 2002, Fischer et al. 2004, Rodríguez-Navarro et al. 2004, Kurshakova et al. 2007, Jani et al. 2012, Ellisdon et al. 2012) as well of the SAGA-DUB module (Köhler et al. 2010, Samara et al. 2010, Weake et al. 2008a, Zhao et al. 2008). The yeast nuclear pore-associated TREX-2 complex is assembled around the Sac3 scaffold to which Thp1, Sem1, Cdc31 and two copies of Sus1 bind. The TREX-2 complex mediates the localization of actively transcribed genes to the nuclear periphery and is required for the formation of export-competent mRNP (Fischer et al. 2004, Köhler et al. 2007, González-Aguilera et al. 2008)

The AP-MS experiments showed that neither components of the TREX-2 complex nor other NPC-related proteins co-purified with ENY2-SG (**Table ??**). In accordance, ENY2 was not co-purified with THP1-SG in the reciprocal experiment. Several putative export factors including all three variants of SAC3 (SAC3a-c) and different nucleoporins were identified by the purification of THP1-SG. Moreover, GO analysis (**Figure S8**) assigned exclusively proteins co-purified with THP1-SG (and not ENY2-SG) to the overrepresented GO term ′Nucleocytoplasmic transport′. The putative *Arabidopsis* TREX-2 subunits CEN1/2 and DSS1 were not co-purified with THP1-SG, although their physical association was demonstrated by yeast two-hybrid and bi-molecular fluorescence complementation assays (Lu et al. 2010). *In silico* analyses of the *Arabidopsis* orthologues of SAC3 uncovered that there is only limited and local sequence conservation (Soerensen et al. 2017). The SAC3-GANP domain that is important for interaction with Thp1/PCID2 and Sem1/DSS1 in yeast/mammals is conserved in *Arabidopsis*, but the CID region that interacts with Cdc31/Centrins and Sus1/ENY2 (Jani et al. 2012) was not found in any land plant protein sequence (Soerensen et al. 2017). In accordance, *Arabidopsis* ENY2 was not found to be part of the plant TREX-2 complex (Lu et al. 2010). The differences in the composition and the assembly of the plant TREX-2 complex compared to its yeast/mammal counterparts in combination with the AP-MS data of THP1-SG and ENY2-SG, suggest that ENY2 is not associated with the TREX-2 complex in *Arabidopsis*.

In yeast, *Drosophila* and mammals, Sus1/ENY2 showed an enrichment at the nuclear envelope, which highlights its association with the NPC-associated TREX-2 complex (Rodríguez-Navarro et al. 2004, Kurshakova et al. 2007, Jani et al. 2012). In detail-CLSM analysis of the subcellular localization of ENY2 in living *Arabidopsis* cells showed that plant ENY2 was form-

ing speckle-like structures in the nucleoplasm and did not accumulate at the nuclear periphery (**Figure 5.6**), where THP1 and SAC3 of the plant TREX-2 complex were located (Lu et al. 2010). This finding support the idea that *Arabidopsis* ENY2 is no part of the mRNA export complex TREX-2.

FRAP experiments in human cells revealed that nucleoporins and TREX-2 components had a very low mobility at the nuclear periphery, which indicated the stable association of the TREX-2 complex with the nuclear pore basket (Rabut et al. 2004, Umlauf et al. 2013). In contrast, human SAGA subunits and chromatin-associated proteins in general showed a high mobility and had mainly transient associations with the chromatin (no detectable immobile fractions)(Phair et al. 2004, Vosnakis et al. 2017). In humans, the two pools of ENY2 proteins showed very different dynamics: the SAGA-associated pool in the nucleoplasm was highly dynamic, whereas the NPC-associated pool (TREX-2) had a very slow turnover. In *Arabidopsis*, ENY2 was not associated with the nuclear envelope (**Figure 5.6**). The dynamic of *Arabidopsis* ENY2 in the nucleoplasm was very high (**Figure 5.7**). The reduced mobility of ENY2 in comparison to free eGFP suggested short term bindings of ENY2 to other proteins like transient interactions with the chromatin. But almost the whole ENY2 population was mobile as observed for human SAGA subunits (Vosnakis et al. 2017), what indicated that there were no long-term bindings to slow/immobile structures like the nuclear periphery. Therefore, this analysis of the *in vivo* binding properties of ENY2 further supported the observation that the plant ENY2 was not associated with the TREX-2 complex and the nuclear envelope.

8.2.8 New friends: The SAGA – Splicing association

The spliceosome, which is responsible for pre-mRNA splicing, is a macromolecular machinery that consists of five small nuclear ribonucleoprotein particles (snRNPs), the nineteen complex (NTC) and a large variety of accessory proteins and cofactors (Will et al. 2011). The NTC complex was also characterised in *Arabidopsis* (Palma et al. 2007, Monaghan et al. 2009). In the initial AP-MS experiments, strikingly all nine core components of the NTC complex and several further proteins of the splicing machinery were co-purified with ENY2-SG. The reverse experiment confirmed this association of ENY2 with the NTC complex, thereby ENY2 was successfully co-purified with MOS4-SG, a subunit of the NTC complex. To clarify which components of the splicing machinery were co-purified with a specific bait protein, the experimentally identified proteins were compared to a list of putative interactors including splicing-related proteins (List XXX) and visualized as heatmap (**Figure 8.2**). This gives a comprehensive view of the interactions between the splicing machinery and specific bait proteins.

This demonstrates clearly the huge overlap between splicing factors that were co-purified with ENY2-SG and with the NTC component MOS4-SG. With MOS4-SG all nine core NTC components as well as several U2 snRNP, U5 snRNP and C complex proteins were co-purified. This proteomic analysis was consistent with previous purification of HA-tagged MOS4 in *Arabidopsis* plants (Monaghan et al. 2009, Koncz et al. 2012). The identified splicing proteins were in accordance with the function of the NTC complex in the catalytically activation of the spliceosome (Fabrizio et al. 2009). Surprisingly, a comparable set of splicing proteins was co-purified with ENY2-SG. Most strikingly, with both bait proteins all nine core components of the NTC

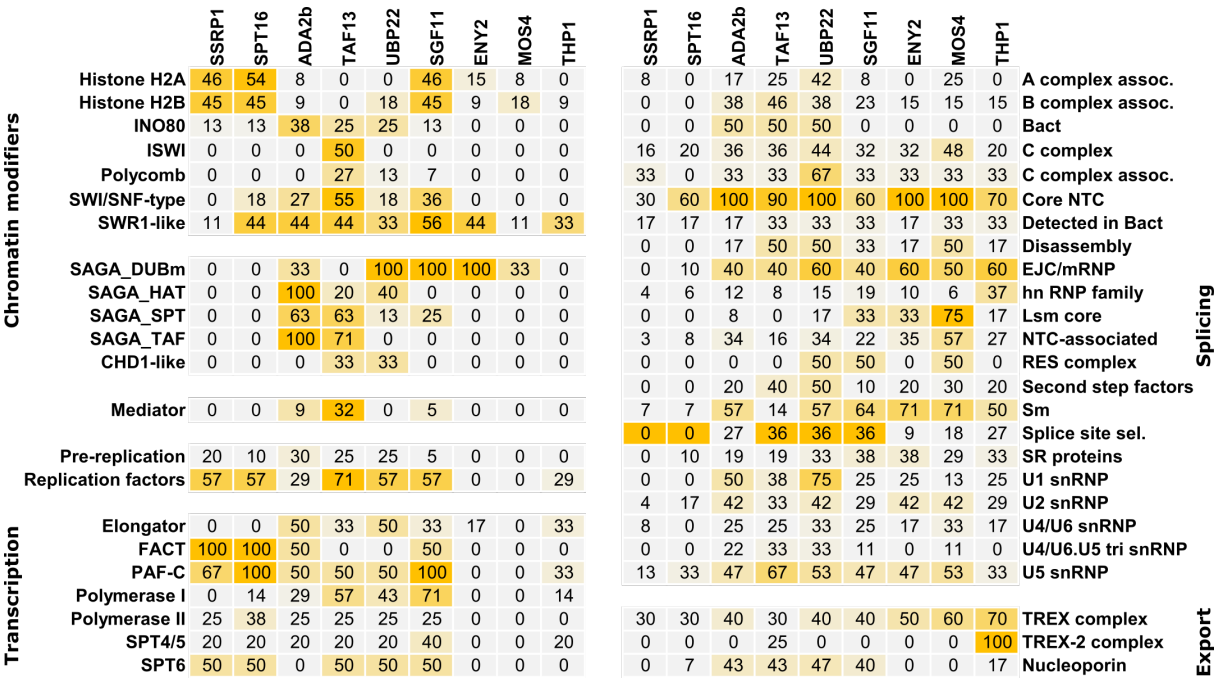


Figure 8.2 Chromatin, transcription and mRNA processing related protein complexes co-purified with different SG-fusion proteins.
Experimentally identified proteins were compared to a list of expected interactors (SUPP). Numbers in the heatmap indicate percentage of identified proteins of a certain protein complex or protein family. A) Splicing related protein complexes B) Transcription-related protein complexes C) Capping, Polyadenylation and mRNA export related protein complexes.

complex were co-purified. This suggested an association of ENY2 with the NTC complex and pre-mRNA splicing.

The three-dimensional structure of yeast spliceosome mainly containing the U2 and U5 snRNP′s as well as the NTC and NTR (NTC-related proteins) proteins was resolved by single-particle cryogenic electron microscopy (Yan et al. 2015) (**Figure 8.3 A-C**). This study revealed that these four subcomplexes were highly intertwined, whereby the centrally located RNA molecule was enclosed by the U5 snRNP at the base and the NTC (including the yeast MOS4 ortholog) as well as the NTR at the top (**Figure 8.3 A**). The *Arabidopsis* orthologs of the yeast NTC and NTR subunits were co-purified with MOS4-SG as well as ENY2-SG in almost all three replicates (**Figure 8.3 D**). Some NTC/NTR subunits were also co-purified with SGF11-SG and THP1-SG and we will come back to this point in a moment.

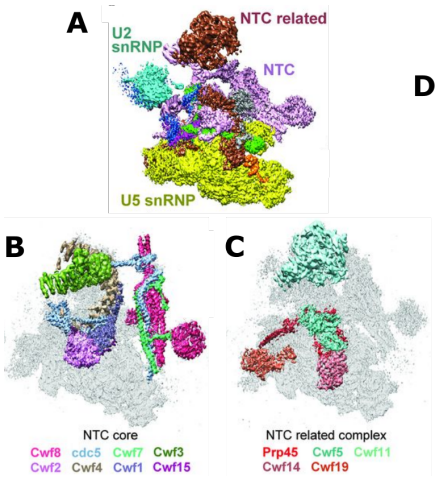
Unfortunately, no direct protein-protein interactions between ENY2 and the core NTC as well as selected NTR subunits were identified by yeast-two-hybrid (**Figure 5.10**) and FRET (**Figure 5.12**) analysis. This could indicate that the ENY2-NTC interaction is mediated by another unknown protein or that the interaction is simply not detectable due to limitations of the experimental setup. Limitations of the Y2H system that are potentially causing false negative results are reviewed in Brückner et al. 2009 and Rajagopala et al. 2012. Moreover, the detectable FRET efficiency depends strongly on the distance of the two fluorophores (Piston et al. 2007, Chudakov et al. 2010), which could be problematic in case of the NTC complex that has a large and extended architecture (Yan et al. 2015).

As mentioned above, several splicing factors were also co-purified with SGF11-SG and

THP1-SG. This is not surprisingly given the fact that the yeast assembling spliceosome has extensive crosstalk to transcription and other nuclear machineries including the export complex THO/TREX and TREX-2 (Herzel et al. 2017). As well it was shown, that the *Arabidopsis* THO component TEX1 associated with the spliceosome and could modulate certain mRNA splicing events (Soerensen et al. 2017). But in comparison to the affinity purifications of ENY2-SG and MOS4-SG, less proteins of the NTC complex, the U2 and U5 snRNP as well as the C complex were co-purified with SGF11-SG and THP1-SG. Moreover, the co-purification of these splicing proteins was less efficient with SGF11-SG and THP1-SG. Meaning, that these proteins showed lower average MASCOT scores and were identified in less replicates, in comparison to the ENY2-SG and MOS4-SG datasets. This suggests a more direct association of ENY2 with the NTC/NTR complex compared to SGF11 and THP1.

Several studies showed that splicing proteins often accumulated in nuclear splicing speckles (Tillemans et al. 2005, Fang et al. 2004. In *Arabipopsis*, ENY2 was also forming speckle-like structures in the nucleoplasm, in contrast to free eGFP that was homogenously distributed (**Figure 5.6**). Several facts supported the authenticity of this observed subcellular localization of ENY2 as follows: 1) The expression of the eGFP-ENY2 fusion protein was driven by its native promoter to avoid overexpression artefacts. 2) The identity and integrity of eGFP-ENY2 fusion protein was validated by immunoblotting (**Figure 5.4**) and mass spectrometry. 3) The N-terminal eGFP tag had no influence on the *Arabidopsis* development (**Figure S6**). The connection of ENY2 and pre-mRNA splicing was studied by co-localization experiments (**Figure 5.19**). This in detail-CLSM analysis revealed a high degree of co-localization between ENY2 and the splicing protein MOS4 (NTC complex). Superimposing of both fluorescence signals (mCherry-ENY2 / eGFP-MOS4) showed a clear overlap of the splicing speckles (**Figure 5.19 A**). The degree of overlap was further characterized by PCC and MOC (**Figure 5.19 B,C**). The calculated coefficients clearly indicated a positive correlation between ENY2 and MOS4, but not between ENY2 and free eGFP. Therefore, the measured co-localization between ENY2 and MOS4 (NTC complex) supported the data of the AP-MS analysis and suggested an involvement of ENY2 in splicing of pre-mRNA. This ENY2 – NTC interaction was not observed for other eukaryotes and therefore seemed to be plant specific.

The initial assumption that ENY2 could interact with the splicing machinery DUB- and SAGA-independent was revised after analysis of the AP-MS data of further SG-tagged SAGA subunits ADA2b, TAF13 and UBP22. As shown in **Figure 8.2**, almost the same set of splicing proteins was co-purified with the NTC component MOS4 and the tested SAGA subunits. These AP-MS data together with the co-localization of ENY2 and MOS4 in splicing speckles, demonstarted the strong asscociation of the ENY2 and the SAGA complex in pre-mRNA splicing. In accordance with these observations, the HAT activity of the yeast SAGA complex is important for the co-transcriptional recruitment of U2 snRNP and the recognition of the intron branchpoint (gunderson2009acetylation). Thereby, a genetic interaction between Gcn5 (HAT module) and the U2 snRNP′s Msl1 and Lea was detected. The *Arabidopsis* counterparts of both proteins were co-purified with SAGA components along with other snRNP proteins. Moreover, two splicing factors of the U2 snRNP (SF3B3 and SF3B5) were found to be novel subunits of the *Drosophila* SAGA complex (Stegeman et al. 2016, Acosta et al. 2017). Noteworthy, almost



D

S. pombe	*A. thaliana*		AP			
NTC/Prp19C	TAIR	Protein	MOS4	ENY2	SGF11	THP1
Cwf8	AT1G04510	MAC3A, PRP19A	2304 / 3	435 / 2	297 / 2	365 / 2
	AT2G33340	MAC3B, PRP19B	3137 / 3	339 / 3	269 / 2	349 / 2
Cwf3	AT5G28740	SYF1, MAC9	2124 / 3	613 / 3		
Cwf12	AT3G18790	ISY1, MAC8	442 / 2	288 / 3		128 / 2
Cdc5	AT1G09770	CDC5, MAC1	4211 / 3	613 / 3	279 / 2	495 / 2
Cwf1	AT4G15900	PRL1, MAC2	1296 / 3	277 / 3		248 / 2
Cwf7	AT3G18165	MOS4, MAC4	1849 / 3	174 / 3	127 / 2	
Cwf15	AT3G13200	EMB2769	297 / 2	142 / 2		175 / 2
Cwf4	AT5G41770	CRN1c, MAC10	1176 / 3	346 / 3		
Cwf2	AT1G07360	MAC5A	592 / 3	315 / 3	299 / 3	
	AT2G29580	MAC5B	376 / 2	178 / 3	126 / 2	151 / 2
Prp45	AT1G77180	SKIP, MAC6	1243 / 3	280 / 3	247 / 2	306 / 3
Cwf11	AT2G38770	EMB2765, Aquarius	2481 / 3	1072 / 3	554 / 3	468 / 3
Cwf14	AT4G21110	Bud31	190 / 2			

Figure 8.3 NTC and NTC-related (NTR) proteins co-purified with MOS4-SG as well as with ENY2-SG
A) Crystal structure of the yeast spliceosome reveals that U5 snRNP, U2 snRNP, NTC and NTC-related proteins (NTR) associate closely with one another (Yan et al. 2015). B) Relative positions of the 8 core NTC components (B) and the five NTC-related (NTR) proteins (C) within the spliceosome are shown. The subunits are dipicted in different colors. D) Composition of the yeast NTC complex in comparison to the *Arabidopsis* NTC as revealed by AP-MS. The average MASCOT score is shown and how many times the proteins were detected in three replicates. Proteins that were identified in all three replictes were depicted in blue.

no splicing proteins were co-purified with the SAGA-TAF subunit TAF10-SG.

Taken together, the SAGA modules DUB, HAT and SPT (but not TAF) seem to play a role in the splicing of pre-mRNA in *Arabidopsis*.

8.3 Reverse genetics to learn more about the function of the plant DUB module

In other species, mutations in Sus1/ENY2 and SGF11 can induce obvious phenotypes. In *Drosophila*, Sgf11 null-mutants are embryonic lethal and a four-fold reduction of the ENY2 transcript level shows diverse effects on the fly morphology (Georgieva et al. 2001, Weake et al. 2008a). Yeast cells lacking Sus1 or Sgf11 are viable, but Sus1△ strains show growth defects (Rodríguez-Navarro et al. 2004). Therefore, reverse genetics approaches were performed to learn more about the function of the DUB components ENY2 and SGF11 in *Arabidopsis*. The available *eny2-1* T-DNA line (SALK_045015) had an insertion in 3′UTR of *ENY2*, which had no effect on the *ENY2* transcript level. If the ENY2 protein level was affected in *eny2-1* mutants was not determined due to the unavailability of ENY2-specific antisera. However, the *eny2-1* mutants looked like *wild type* plants. The generated ENY2 RNAi lines showed a clear downregulation of the *ENY2* transcript level of almost 70%, but this did not cause any obvious phenotype. Noteworthy, the three independent RNAi lines were not consistent regarding their time of bolting: The line #27 bolted like WT, whereas the lines #22/#3 showed a slight early/late bolting phenotype, respectively. The small differences in the growth conditions (inherent to the growth chamber) could be responsible for the small observed phenotypic fluctuations among the three lines and could be diluted out by more biological replicates. Moreover, the random and potentially gene-disrupting insertions of the T-DNA into the genome could further result in phenotypic characteristics that were non-consistent among the three independent ENY2 RNAi lines. However, the RNAi mutants with reduced levels of ENY2 transcripts look like *wild type* plants under standard growth conditions. Therefore, ENY2 knockout mutants

were successfully generated by CRISPR/Cas9 under the control of an egg cell-specific promoter (Wang et al. 2015). With this genome editing system, stable inheritable mutations were produced at a frequency of 10%. In comparison, it was not possible to obtain CRISPR-induced mutations in *ENY2* with a ′first generation′ system (Fauser et al. 2014) that is basing on the constitutive overexpression (CaMV 35S promoter) of CRISPR/Cas9. In the trans-homozygous *ENY2* mutant of the T1 generation as well as in the homozygous *ENY2* mutants of the T2 generation (*eny2-crispr1*, *eny2-crispr2*), a frame shift after the first 30 nt resulted most likely in a complete inactivation of *ENY2*. The phenotypic analysis showed a significant late flowering phenotype that was consistent in both *ENY2* CRISPR mutants. These initial data suggested that the transition from the vegetative to the generative phase was affected by the complete knockout of *ENY2*, but not by the knockdown.

Beside ENY2, a potential knockout mutant of the second DUB adaptor protein SGF11 was analysed. The T-DNA insertion in the second exon disrupted the expression of the endogenous *SGF11*, but coincidently promoted the weak expression (20% transcript level of the WT *SGF11*) of a severely N-terminal truncated *SGF11* transcript (66% CDS of the WT *SGF11*). Due to the unavailability of SGF11 antisera, it could not be excluded that a truncated version of SGF11 is produced. However, the *sgf11-1* mutant showed a mild late flowering phenotype like the *ENY2* CRISPR mutants. Moreover, *sgf11-1* mutants showed a significantly increased expression of the floral repressor FLC and a reduced expression of the downstream target SOC1 (FLC acts as a repressor of SOC1), which supported the observed late-flowering phenotype on a molecular level. This suggested a novel role for the two DUB subunits SGF11 and ENY2 in controlling flowering time as discussed later in section 8.3.1.

It is rather unlikely that functional redundancy explains the lack of strong phenotypic alterations in *eny2-crispr* and *sgf11-1* mutants because there are no paralogs to ENY2/SGF11 in *Arabidopsis*. AtUBP25 (AT3G14400) which harbours an ′ENY2′ (pfam10163) domain at the C-terminus, was the only protein with a high similarity to ENY2 found by blastp search. Moreover, the knockout of ENY2 and SGF11 could be conditional. Thereby, even in the presence of severe physiological defects, the mutants are not showing any obvious phenotypes under standard growth conditions in the growth chamber (Bouché et al. 2001).

8.3.1 De-ubiquitination activity of the plant DUB module plays a role in flowering time control

Immunoblotting revealed that the global H2Bub level was strongly increased in the *sgf11-1* (knockout) and slightly increased in the *eny2-RNAi* (knockdown) mutants. The differences in the quantity of the H2Bub accumulation between both DUB mutants could be explained by the fact that *sgf11-1* was most likely a complete knockout, whereas *eny2-RNAi* mutants still express reduced levels of *ENY2* mRNA. In the future, the analysis of the H2Bub level in the recently generated *eny2-crispr* knockout mutants can reveal if the H2Bub accumulation is dosage-dependent on the *ENY2* transcript level and if a complete *ENY2* inactivation leads to H2Bub levels comparable to *sgf11-1*.

The ubiquitination of histones is important to trigger gene activity in yeast (Weake et al. 2008b). The *Arabidopsis* orthologs UBC1/2 and HUB1/2 of the yeast Rad6/Bre1 complex,

monoubiquitinate the histone H2B at lysine K143 (Sridhar et al. 2007, Fleury et al. 2007, Liu et al. 2007, Cao et al. 2008). In *hub1/2*, *ubc1/2* mutants, the impaired mono-ubiquitination of H2Bub correlates with defects in the regulation of key players in flowering time like the floral repressor FLC and with an earlier transition to flowering (Cao et al. 2008, Gu et al. 2009, Xu et al. 2009). In yeast and animals, the SAGA-dependent DUB module is responsible for the antagonist reaction and is removing monoubiquitin from H2B (Samara et al. 2010, Köhler et al. 2010).

In *Arabidopsis*, the DUB mutants *eny2-crispr* and *sgf11-1* showed a late-flowering phenotype that was opposite to the early flowering phenotype of *hub1/2* and *ubc1/2* mutants. Moreover, the transcript level of *FLC* was increased and the downstream target of the floral repressor was decreased in the *sgf11-1* mutant. This suggested that the DUB module in plants (SAGA-dependent or not) is also involved in controlling flowering time by de-ubiquitination of key flowering time genes.

8.3.2 ENY2, a novel link between histone de-ubiquitination and splicing to control flowering time?

Next to its role in histone de-ubiquitination, ENY2 is also associated with the spliceosome activating NTC/NTR complexes which suggests a putative functional link of both processes. How ENY2 and the DUB module could influence the expression of key proteins that control flowering time by H2B de-ubiquitination was discussed above. Additionally, there is some evidence that flowering time is also controlled by pre-mRNA splicing. First, several *Arabidopsis* NTC knockout mutants like *cdc5-1*, *mos4-2* as well as the *mac3a/mac3b* double mutant showed a delay in the transition to flowering (Palma et al. 2007, Monaghan et al. 2009), whereas the knockout of the NTR component MAC5A caused an early flowering phenotype (Monaghan et al. 2010). During activation of the spliceosome, the U5 snRNP including BRR2a and PRP8 was directly interacting with the NTC/NTR complexes (Wan et al. 2016). Recently it was shown that BRR2a affects the flowering time via FLC splicing (Mahrez et al. 2016b). Furthermore, a *prp8* mutant showed that *FLC* expression could be modulated by altered splicing of the antisense *COOLAIR* (Marquardt et al. 2014). Both examples suggest that splicing is an important mechanism to control flowering time. Therefore, one could speculate that the adaptor protein ENY2, which is associated with the DUB module and the NTC complex is linking histone de-ubiquitination and pre-mRNA splicing to control flowering time.

8.4 Outlook

A proteomic approach revealed the composition of the *Arabidopsis* SAGA complex and the ENY2-containing DUB module. If the DUB module is acting in SAGA-dependent or independent manner remains still unclear. RNA-seq analysis of *Arabidopsis* HAT mutants like *gcn5-1* and *ada2b-1* (Vlachonasios et al. 2003) and *Arabidopsis* DUB mutants like *sgf11-1* and *eny2-crispr* could reveal if both SAGA modules regulate the expression of a similar set of genes. Moreover, ChIP-seq analysis using plants -expressing tagged representatives of all four SAGA modules like, for instance, ADA2B (HAT), ENY2 (DUB), TAF13 (SPT), and TAF10 (TAF) could determine their genome-wide binding sites. This could reveal if the modules, and espe-

cially DUB, are binding to the same target sites as whole SAGA complex or if the modules have independent binding sites. Furthermore, a combination of the RNA-seq and ChIP-seq data could reveal if the genes that are bound by the DUB or the HAT module are also differentially expressed in the DUB- and the HAT- mutants. Additionally, these data will show if the DUB module can bind to specific target sites and control the expression of certain genes in a SAGA-independent manner. The CRISPR/Cas9-induced knockout of ENY2 as well as the T-DNA-mediated knockdown of SGF11 resulted in a late flowering phenotype, what suggests a role for the DUB module in controlling flowering time. The above-mentioned ChIP-seq data will also reveal if genes encoding floral integrators like FLC are occupied by ENY2 and SGF11. The expression of FLC depends on the monoubiquitination of histone H2B (Ying Cao et al. 2008), therefore the DUB module could be involved in regulating the transition from vegetative to reproductive phase by de-ubiquitination of H2Bub. ChIP analysis could show if the H2Bub level at FLC is increased in SGF11- and ENY2-defective plants. Moreover, the characterized *eny2-crispr* and *sgf11-1* mutants are a promising tool to reveal a potential functional link between the SAGA DUB module and the splicing machinery. The analysis of double mutants defective in one of the DUB subunits and in one of the components of the splicing machinery, for instance the well described NTC mutants (Palma et al. 2007, Monaghan et al. 2009), could reveal genetic interactions that are observable in synergistic effects on plant growth and development. Moreover, splicing defects caused by the lack of ENY2 or SGF11 could be determined by the above-mentioned RNA-seq experiments using *eny2-crispr* and *sgf11-1* mutants.

9 Summary

FACT, a heterodimer of SSRP1 and SPT16, is a conserved and essential histone chaperone. FACT binds to H2A-H2B dimers to reorganizes octameric nucleosomes and to make the genomic DNA accessible. For instance, FACT facilitates the progression of the transcription machinery through the chromatin template by destabilizing nucleosomes in the path of the elongating RNAPII. By the reverse action, FACT restores the chromatin structure in the wake of the RNAPII and maintains the status quo.

By *in vitro* EMSA experiments, this study shows that the DNA- and nucleosome binding properties of *Arabidopsis* SSRP1 were mediated by its C-terminal HMG-box domain. *In vivo*, the loss of the HMG-box domain did not alter the subcellular localization of SSRP1 or the nuclear protein dynamics/binding properties as shown by in-detail CLSM and FRAP analysis, respectively. Additionally, immunoblot analysis showed that HMG-box-deficient SSRP1 was still associated with SPT16 and the transcriptionally active RNAPII *in vivo*. Phenotyping of SSRP1 HMG-box deficiency mutants showed that the lack of the DNA-binding domain had no effect on *Arabidopsis* growth and development. The *in vitro* data indicate that the binding of FACT to H2A-H2B dimers in higher eukaryotes depends, in the first place, on the association of the SSRP1 HMG-box domain with nucleosomal DNA. Nevertheless, the *in vivo* data suggest that the loss of the SSRP1 HMG-box domain can be compensated in *Arabidopsis* by other unknown factors as may be HMGB proteins that provide the DNA-binding function for FACT.

During transcript elongation, the histone chaperone FACT facilitates together with other TEFs efficient mRNA synthesis by RNAPII. This study contributed to reveal the composition of the *Arabidopsis* transcript elongation complex (TEC) by a proteomic approach using reciprocal tagging in combination with affinity purification and mass spectrometry. The TEFs FACT, PAF1-C, SPT4/5, SPT6, and TFIIS co-purified robustly with each other and the elongating RNAPII, while P-TEFb was not among the interactors. Additionally, further chromatin modifying factors including NAP1 and the Elongator were repeatedly co-purified with different TEFs. The phenotypic analysis of *Arabidopsis* double mutants that are defective in different combinations of TEFs revealed genetic interactions between the genes encoding subunits of FACT, PAF1-C, and TFIIS, resulting in synergistic/epistatic effects on plant growth/development

Genome-wide transcriptome profiling of SSRP1- or SPT16-depleted plants in comparison to *wild-type* revealed that almost the identical small set of genes was differentially expressed in both mutants. Strikingly, genes encoding key anthocyanin biosynthesis enzymes were overrepresented among the genes that were downregulated in both FACT mutants. A phenotypic analysis showed that FACT-depleted plants have clear defects in the light-induced accumulation of anthocyanin in their leaves. In response to high light (HL) stress, anthocyanin biosynthesis genes were upregulated to a lesser extend in the FACT mutants compared to *wild type*. Additionally, the gene expression of *SSRP1* and *SPT16* was upregulated upon HL stress. These data suggest that FACT is novel factor required for the accumulation of anthocyanins in response to HL stress.

ENY2, an evolutionary conserved adaptor protein, links transcription by RNAPII with export of the newly synthesized mRNA to the cytoplasm by being part of the transcriptional co-activator SAGA and the NPC-associated mRNA export complex TREX-2. Histochemical GUS staining revealed that the *Arabidopsis* ENY2 promoter is widely active during plant growth and development. In the plant cells that are expressing ENY2, the protein is forming speckle-like structures in the nucleoplasm and is highly mobile as shown by in-detail CLSM and FRAP analysis, respectively.

Using reciprocal tagging of ENY2 and its putative interactors in combination with affinity purification and mass spectrometry revealed that ENY2 associates with two components /SGF11, UBP22) of the SAGA histone H2B deubiquitinase (DUB) module. Furthermore, no subunits of other SAGA modules or the TREX-2 complex were co-purified with ENY2. Additionally, several splicing complexes especially the U2, U5 and NTC/NTR were identified in the affinity purification of ENY2. In accordance with these findings, ENY2 and the NTC component MOS4 co-localized in splicing speckles, whereas no direct protein-protein (PPIs) interactions between ENY2 and NTC/NTR components were found by Y2H and FRET. Three (ENY2, SGF11, UBP22) of the four SAGA-DUB components that were identified in yeast, fruit-fly and humans were highly conserved in *Arabidopsis*, the two adaptor proteins ENY2 and SGF11 as well as enzymatically active UBP22. SGF73, the missing protein that links the DUB module to the remaining SAGA complex in other organisms had no homolog in plants. Direct PPIs between SGF11 and ENY2 as well as UBP22 could be detected by Y2H and FRET analysis.

The composition of the *Arabidopsis* SAGA complex was revealed by a proteomic approach using reciprocal tagging of one representative of each bioinformatically predicted SAGA module in combination with affinity purification and mass spectrometry. In total, 17 *Arabidopsis* SAGA subunits were biochemically identified. The DUB module did almost not co-purify with the other SAGA modules (HAT, SPT, TAF), which suggest that the plant DUB module can act in a SAGA-independent manner. Additionally, several subunits of the spliceosome were repeatedly co-purified with the DUB, HAT, and SPT modules of SAGA.

A reverse genetics approach revealed that knockdown of *ENY2* by RNAi had no obvious effect on plant growth and development, while the complete knockout of *ENY2* by CRISPR/Cas9 induced a late flowering phenotype. The overexpression of *ENY2* did not cause any obvious phenotype. The knockdown of *SGF11* by a T-DNA insertion resulted as well in a late flowering phenotype and an upregulation of the floral repressor *FLC*. Moreover, the global H2Bub levels were increased in ENY2- and SGF11-depleted plants.

Taken together, this study revealed the composition of the *Arabidopsis* SAGA complex. ENY2 is part of a histone H2B de-ubiquitinating module that can exist most likely SAGA-independent (**Figure 9.1**). DUB-defective plants show a late flowering phenotype. Additionally, the DUB as well as other SAGA modules show a strong association to the pre-mRNA splicing machinery. Surprisingly, *Arabidopsis* ENY2 is no part of the mRNA export complex TREX-2 as it was shown in other eukaryotes.

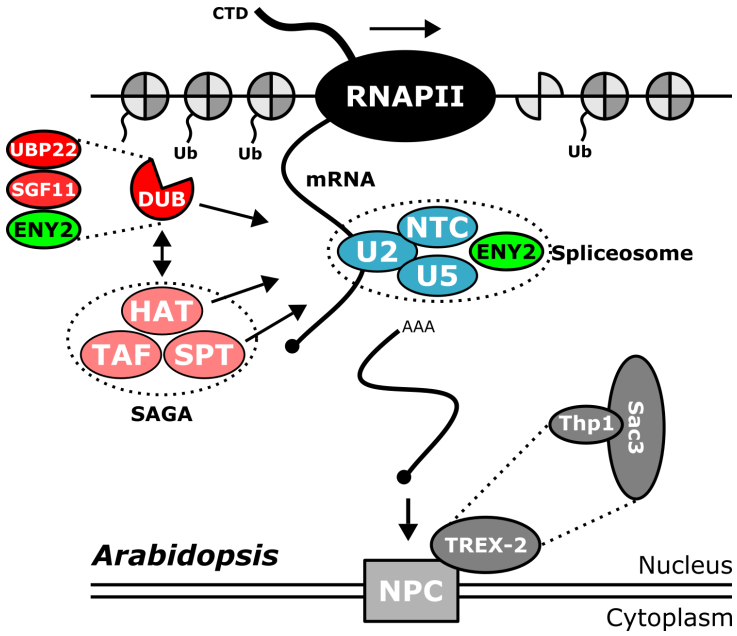


Figure 9.1 Model of ENY2 in the gene expression pathway of *Arabidopsis*.
In *Arabidopsis*, the small adaptor protein ENY2 is part of a de-ubiquitination (DUB) module. The DUB module can most act on the chromatin in a SAGA-dependent and -independent manner (Indicated by the double-headed arrow). In plants, ENY2 is no subunit of the NPC-associated mRNA export complex TREX-2. Whereas the proteomic data indicate a strong association of some SAGA modules (DUB, HAT and SPT) with the splicing machinery, especially the U2, the U5 and the NTC complexes (Indicated as arrows). Additionally, co-localization studies support the interaction of ENY2 with the splicing-associated NTC complex.

10 Materials

10.1 Instruments

Table 10.1 List of instruments

Instrument	Manufacturer /model
Blotting System	Semi-dry Blotter Maxi (Carl Roth)
Centrifuges	Sorvall Evolution RC and Sorvall LYNX 4000 equipped with SLA1500 or SS34 rotor (Thermo Fisher Scientific), Centrifuge 5417R and 5804 R (Eppendorf)
Digital camera	EOS 600D equipped with Macro lens EF-S 60 mm 1:2.8 USM or a ETS 18-55 mm objective (Canon)
FPLC System	Gradient Programmer GP250 (Pharmacia Biotech)
Homogenizer	TissueLyser II (Quiagen)
Imager	BioDocAnalyze System (Biometra), Typhoon FLA 9500 (GE Healthcare), Multiimage^{TM} FluorChem FC2 (Alpha Innotech)
Microscope	TCS SP8 (Leica), ApoTome.2 with Axiocam 503 (Zeiss), SMZ645 stereo microscope (Nikon) with KL 1500 LCD (Schott), Discovery V8 stereo with Axiocam MRc5 and KL1500 LCD (Zeiss)
Plant incubator	Plant incubator (Percival Scientific and polyklima)
Quantum Meter	Quantum Flux ML-200 (Apogee Instruments)
RT-qPCR Cycler	MastercyclerÂő ep RealPlex (Eppendorf)
Shaking Incubator	Multitron Standart and Pro (Infors HT)
Sonicator	UW2070 MS73 (Bandelin electronic)
Spektrophotometer	NanoDrop ND-1000 (PEQLAB)
Thermocycler	T3000 Thermocycler and T Gradient (Biometra)

10.2 Chemicals and enzymes

Laboratory grade chemicals and reagents were purchased from Applichem (Germany), Carl Roth (Germany), Clonetech, Duchefa (Netherlands), Fluka (Switzerland), Life Technologies (UK), Merck (Germany), Sigma Aldrich (Germany), USBiologcal (USA), and VWR (USA). Enzymes were purchased from Thermo Fisher Scientific (USA), PEQLAB/VWR (USA) and New England Biolabs (USA).

10.3 Oligonucleotides

Table 10.2 List of oligonucleotides for cloning
Oligonucleotides were obtained from MWG eurofins genomics (* = from this study, L = from lab collection, No. = lab ID, R. sites = Restriction sites, S. = Source).

Sequence 5' - 3'	Description	R. sites	S.	No.
AA TCTAGA GGTTTTCTGTTGAGAGACACGA	Insertion of pSSRP1 (-614 bps) in pCambia3300:GUS	HindIII / XbaI	*	675
AATT AAGCTT AGAATTCTAGCAATGCAGGGTAA	Insertion of pSSRP1 (-614 bps) in pCambia3300:GUS	HindIII / XbaI	*	896
CG GGATCC CCATGGCGGACGGCCACTCCTT	Insertion of SSRP1△HMG in pGEX-5X-1	BamHI / SalI	*	3050
CGC GTCGAC TTACTTCTTCTTCTTCGGCTTCTTCC	Insertion of SSRP1△HMG in pGEX-5X-1	BamHI / SalI	*	3051
GGAATTC CATATG ATGAAACATTCGGTGAATCGG CC	Insertion of ENY2 (CDS) in pGADT7 (p68) and pGBKT7 (p69)	NdeI / BamHI	*	3194
CG GGATCC TCAAAGAGCAGCTGATACAATGAAG	Insertion of ENY2 (CDS) in pGADT7 (p68), pGBKT7 (p69), pCambia:p35S::5'eGFP (p810)	NdeI / BamHI	*	3195

Table 10.2 (Continuation) List of oligonucleotides for cloning
Oligonucleotides were obtained from MWG eurofins genomics (* = from this study, L = from lab collection, No. = lab ID, R. sites = Restriction sites, S. = Source).

Sequence 5' - 3'	Description	R. sites	S.	No.
GGAATTC CATATG ATGTCTGGCGCAGAGGATAA TAAAT	Insertion of SGF11 (CDS) in pGADT7 (p68) and pGBKT7 (p69)	NdeI / BamHI	*	3196
CG GGATCC TCAGTCTCCTTTCACGTTCTCTC	Insertion of SGF11 (CDS) in pGADT7 (p68), pGBKT7 (p69), and pCambia2300:p35S::5'eGFP (p810)	NdeI / BamHI	*	3197
GC TCTAGA ATGTCTGGCGCAGAGGATAATAAATC	Insertion of SGF11 (CDS) in pCambia2300:p35S::SG	XbaI / BamHI	*	3447
CG GGATCC GTCTCCTTTCACGTTCTCTCGAA	Insertion of SGF11 (CDS) in pCambia2300:p35S::SG	XbaI / BamHI	*	3448
GC TCTAGA ATGGCGACGAACAATGGTGATGT	Insertion of MOS4 (CDS) in pCambia2300:p35S::SG	XbaI / BamHI	*	3445
CG GGATCC TTGCATTTGAAGTGGCTCGACGTT	Insertion of MOS4 (CDS) in pCambia2300:p35S::SG	XbaI / BamHI	*	3446
GC TCTAGA ATGGCGTACGTTAGTATGGGTGA	Insertion of THP1 (CDS) in pCambia2300:p35S::SG	XbaI / BamHI	*	3449
CG GGATCC TGAGCTAACAGGCTTCCCGTTTA	Insertion of THP1 (CDS) in pCambia2300:p35S::SG	XbaI / BamHI	*	3450
TCC CCCGGG AAGCTTTATATCAAGAACAAGAAG AAACAAGTCTTTG	Insertion of pSPT16 (-2428 bps) in pCambia3300:GUS	SmaI / BamHI	*	3551
CG GGATCC TCTAGACTAAAGAGTCCAGCAGCAA CCT	Insertion of pSPT16 (-2428 bps) in pCambia3300:GUS	SmaI / BamHI	*	3552
GC TCTAGA ATGGTGAGCAAGGGCGAGGA	Insertion of eGFP-NLS into pCambia2300	XbaI / EcoRI	*	3555
G GAATTC CATATGGACCTTTCTCTTCTTTTTTG GAGGC	Insertion of eGFP-NLS into pCambia2300	XbaI / EcoRI	*	3556
CG GAATTC CTAGGATCCTCGCGAGAGCTCGTCGA CCATATGCTTGTACAGCTCGTCCATGCC	Insertion of eGFP + MCS in pCambia2300:p35S	XbaI / EcoRI	*	3557
CGC GAGCTC ATGGCGGACGGCCACTC	Insertion of SSRP1 and SSRP1△HMG (genomic) in pCambia2300:p35S::eGFP (p810)	SacI / NruI	*	3576
CGC TCGCGA TTAGTTACTATCGGAATCGTTTCCTG	Insertion of SSRP1 (genomic) in pCambia2300:p35S::eGFP (p810)	SacI / NruI	*	3577
CGC TCGCGA TTACTTCTTCTTCTTCGGCTTCTTC	Insertion of SSRP1△HMG (genomic) in pCambia2300:p35S::eGFP (p810)	SacI / NruI	*	3578
CGC GAGCTC TCGCGA ATGAAGTCTCTTAATGATC TTCCTGCG	Insertion of SKIP (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI-NruI/ BamHI	*	3710
CG GGATCC TTAACGCCGGTCACTGCGTTC	pCambia2300:p35S::5'eGFP (p810) Insertion of SKIP (CDS) in	SacI-NruI/ BamHI	*	3711
CGC GAGCTC TCGCGA ATGGCTCGTAATGAAGA GAAAGCAC	Insertion of ISY1 (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI-NruI/ BamHI	*	3712
CG GGATCC CTATCTATGAATATTGAGCATGGAT TTAGC	Insertion of ISY1 (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI-NruI/ BamHI	*	3713
CGC TCGCGA ATGGCGACGAACAATGGTGATGT	Insertion of MOS4 (CDS) in pCambia2300:p35S::5'eGFP (p810)	NruI/BamHI	*	3714
CG GGATCC TCATTGCATTTGAAGTGGCTCGAC	Insertion of MOS4 (CDS) in pCambia2300:p35S::5'eGFP (p810)	NruI/BamHI	*	3715
CGC GAGCTC TCGCGA ATGAATTGTGCAATTTCC GGCGAA	Insertion of MAC3A (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI-NruI/ BamHI	*	3716
CG GGATCC TCATGAATCTTGTGCTGAATCTTCAG	Insertion of MAC3A (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI-NruI/ BamHI	*	3717
CGC GAGCTC ATGAACTGTGCAATTTCAGGAGAAG	Insertion of MAC3B (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI / BamHI	*	3718
CG GGATCC TCACGAGTCTTGCGCAGAGT	Insertion of MAC3B (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI / BamHI	*	3719
CGC TCGCGA ATGAGGATTATGATTAAGGGAGGTG	Insertion of CDC5 (CDS) in pCambia2300:p35S::5'eGFP (p810)	NruI / BamHI	*	3720
CG GGATCC TTATGCAGAAGCTTCCATGGCTAT	Insertion of CDC5 (CDS) in pCambia2300:p35S::5'eGFP (p810)	NruI / BamHI	*	3721
CGC TCGCGA ATGTCTGGCGCAGAGGATAATAAAT	Insertion of SGF11 (CDS) in pCambia2300:p35S::5'eGFP (p810)	NruI / BamHI	*	3755
CGC TCGCGA ATGAAACATTCGGTGAATCGGCC	Insertion of ENY2 (CDS) in pCambia2300:p35S::5'eGFP (p810)	NruI / BamHI	*	3754
CCC AAGCTT ACAAGCGTTTGCCGTCCACAAC	Insertion of pENY2 (-1970 bps) in pCambia3300:GUS	HindIII / XbaI	*	3832
GC TCTAGA GTGTTTGAATCGATAGATTCGTTTC AAGA	Insertion of pENY2 (-1970 bps) in pCambia3300:GUS and pCambia2300:5'eGFP	HindIII / XbaI	*	3833

Table 10.2 (Continuation) List of oligonucleotides for cloning
Oligonucleotides were obtained from MWG eurofins genomics (* = from this study, L = from lab collection, No. = lab ID, R. sites = Restriction sites, S. = Source).

Sequence 5' - 3'	Description	R. sites	S.	No.
CAAAGCGCCTCCAAAAAAGAAGAGAAAGGTCCCC	Insertion of NLS in pGreen0179:p35S::mCherry	SacI / SmaI	*	3834
GGGGACCTTTCTCTTCTTTTTTGGAGGCGCTTT GAGCT	Insertion of NLS in pGreen0179:p35S::mCherry	SacI / SmaI	*	3835
CGC TCGCGA TCTAGA ATGGTGAGCAAGGGCGA GGA	Insertion of mCherry in pGreen0179:p35S and pGreen0179:p35S::eGFP-NLS	NruI - XbaI / SacI	*	3873
CG GGATCC GAGCTC CATATGCTTGTACAGCTCGT CCATGCC	Insertion of mCherry in pGreen0179:p35S and pGreen0179:p35S::eGFP-NLS	NruI / SacI	*	3874
CGC TCGCGA GCGGCCGCAGAATTCTAGCAATGCA GGGTAATTTATA	Insertion of pSSRP1::SSRP1/SSRP1△HMG (genomic) in pGreen0179	NruI / SacI	*	3975
CGC GAGCTC TTACTTCTTCTTCTTCGGCTTCTTC	Insertion of pSSRP1::SSRP1△HMG (genomic) in pGreen0179	NruI / SacI	*	3976
CGC GAGCTC TTAGTTACTATCGGAATCGTTTCCTG	Insertion of pSSRP1::SSRP1 (genomic) in pGreen0179	NruI / SacI	*	3977
AATTCTCGCGATAAGAGCTCTAAG	Insertion of altered NLS in pGreen0179:p35S::mCherry	EcoRI	*	3978
AATTCTTAGAGCTCTTATCGCGAG	Insertion of altered NLS in pGreen0179:p35S::mCherry	EcoRI	*	3979
CGC GAGCTC ATGTATGTATTAGACAATCTTATTCC CTTC	Insertion of ENY2 (genomic) in pCambia2300:pENY2-5'eGFP	SacI / BamHI	*	4015
CG GGATCC TCGCGAGAGGATGGCAACGCTTTGA TTT	Insertion of ENY2 (genomic) in pCambia2300:pENY2-5'eGFP	SacI / BamHI	*	4016
CGC GAGCTC ATGGCGATTTCCAAAGATCTCTATC	Insertion of SYF1 (CDS) in pGreen0179:p35S::5'mCherry (p920)	SacI / BamHI	*	4028
CG GGATCC TCACTGATTAAGCTTCTGTCTCTTG	Insertion of SYF1 (CDS) in pGreen0179:p35S::5'mCherry (p920)	SacI / BamHI	*	4029
CGC GAGCTC TCGCGAGAATTCATGGCTTCCGGC GGCAAAGA	Insertion of CRN1c (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI / BamHI	*	4030
CG GGATCC TCAATCATCCTCAGAAGCAGCAAC	Insertion of CRN1c (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI / BamHI	*	4031
CGC GAGCTC TCGCGAATGCCGGCTCCGACGACG	Insertion of PRL1 (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI / EcoRI	*	4032
CCG GAATTC TTAGAAGCGCCTAATCTCCTTTGGTG	Insertion of PRL1 (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI / EcoRI	*	4033
CGC CCCGGG T TCGCGA ATGACGAAGGTCTATGG AACTGG	Insertion of MAC7 (CDS) in pGreen0179:p35S::5'mCherry (p920)	SmaI(NruI)/ EcoRI	*	4036
CCGC TCGAGT CGCGACTAATTCTTCTCATCAGCCT TTCCATT	Insertion of MAC7 (CDS) in pGreen0179:p35S::5'mCherry (p920)	SmaI(NruI)/ EcoRI	*	4037
GC TCTAGA ATGTCCGCGAGGATTTCATTTCTG	Insertion of UBP22 (CDS) in pCambia2300:p35S::SG	XbaI / SacI	*	4278
CGC GAGCTC GCAATCAGCAAAGGGAAATGCATC	Insertion of UBP22 (CDS) in pCambia2300:p35S::SG	XbaI / SacI	*	4279
CGC GTCGAC ATGGGTCGCTCTCGAGGGAA	Insertion of ADA2b (CDS) in pCambia2300:p35S::SG	SalI / BamHI	*	4280
CG GGATCC AAGTTGAGCAATACCCTTCTTCACAAG	Insertion of ADA2b (CDS) in pCambia2300:p35S::SG	SalI / BamHI	*	4281
CG GAATTC ATGTCCGCGAGGATTTCATTTCTG	Insertion of UBP22 (CDS) in pGADT7 (p68) and pGBKT7 (p69)	EcoRI / SacI	*	4375
CGC GAGCTC GCGGCCGC TCAGCAATCAGCAAAG GGAAATGC	Insertion of UBP22 (CDS) in pGADT7 (p68) and pGBKT7 (p69)	EcoRI / SacI (NotI)	*	4376
GC TCTAGA ATGAGTAACACACCAGCAGCGG	Insertion of TAF13 (CDS) in pCambia2300:p35S::SG	XbaI / SacI	*	4378
CGC GAGCTC ATCAACGAGTTCCTTTTCGTCGACAT	Insertion of TAF13 (CDS) in pCambia2300:p35S::SG	XbaI / SacI	*	4379
ATTGCATCGTCTTCATCTGGTGT	Insertion of ENY2 CRISPR guide 9 in pHEE401	BsaI	*	4488
AAACACACCAGATGAAGACGATG	Insertion of ENY2 CRISPR guide 9 in pHEE401	BsaI	*	4489
CGC GAGCTC ATGTCCGCGAGGATTTCATTTCTG	Insertion of UBP22 (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI / EcoRI	*	4545
CG GAATTC TCAGCAATCAGCAAAGGGAAATGC	Insertion of UBP22 (CDS) in pCambia2300:p35S::5'eGFP (p810)	SacI / EcoRI	*	4546
CG GGATCC TTCGTCCCTTGTTGCAGGGTC	Insertion of TAF10 (CDS) in pCambia2300:p35S::SG	BamHI / XbaI	*	4625
GC TCTAGA ATGAATCACGGCCAACAATCTGGT	Insertion of TAF10 (CDS) in pCambia2300:p35S::SG	BamHI / XbaI	*	4626

Table 10.3 List of oligonucleotides for genotyping
Oligonucleotides were obtained from MWG eurofins genomics (* = from this study, L = from lab collection, No. = lab ID, S. = Source).

Sequence 5' - 3'	Description	S.	No.
CGATCCAGACTGAATGCCCA	Genotyping, T-DNA insertion GUS (pSSRP1::GUS, pSPT16::GUS, and pENY2::GUS)	L	779
ACCCGACCGGATCGTATCGGT	Genotyping, Transposon insertion ssrp1-1 (GT7431)	L	800
GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC	Genotyping, T-DNA insertion SAIL LB (spt16-1, sgf11-1)	L	802
GATAATTGCTTCTCATCCGGTGT	Genotyping, Transposon insertion ssrp1-1 (GT7431)	L	808
CCCTCATCTTACGCGTATCAGA	Genotyping, T-DNA insertion ssrp1-2 (SALK_001283)	L	810
GTTGCCCGTCTCACTGGTGA	Genotyping, T-DNA insertion SALK LBb1.3 (ssrp1-2, eny2-1)	L	812
AATTAAGCTTAGTTACTATCGGAATCGTTTCCT	Genotyping, T-DNA insertion ssrp1-2 (SALK_001283)	L	900
AATTAAGCTTAGAATTCTAGCAATGCAGGGTAA	Genotyping, T-DNA insertion pSSRP1::GUS	L	916
GAAGAGCCAATTAAGATAAAACGTTGAATGTA	Genotyping, T-DNA insertion eny2-RNAi	L	1076
CTATCTCTGCATTGCCTCTTAGC	Genotyping, T-DNA insertion spt16-1 (SAIL_392_G06)	L	1123
TACTTGTCTAACGCAGCGAAATC	Genotyping, T-DNA insertion spt16-1 (SAIL_392_G06)	L	1124
GCGGATAACAATTTCACACAGGAAACAG	Genotyping, T-DNA insertion pCambia (pSSRP1::eGFP-SSRP1, p35S::eGFP-NLS, and pENY2::eGFP-ENY2)	L	1840
GTTGTAAAACGACGGCCAGTG	Genotyping, T-DNA insertion eny2-crispr	*	1937
ATTTGTAGAGAGAGACTGGTG	Genotyping, T-DNA insertion pGreen (pUBI10::SSRP1, pUBI10::SSRP1△HMG, and pUBI::ENY2)	L	1939
GTTTACCCGCCAATATATCCT	Genotyping, T-DNA insertion pGreen (pSSRP1::SSRP1 and pSSRP1::SSRP1△HMG)	L	1940
CATGTTAGTTTGATTTCTTTG	Genotyping, T-DNA insertion eny2-1 (SALK_045015)	L	2451
GGAGCATGTAAAGAAGAAAGG	Genotyping, T-DNA insertion pENY2::eGFP-ENY2	L	2452
TTAGGAGTGATCACTCGGATCAGTTC	Screening for CRISPR/Cas9-induced mutations	*	2732
GCTCTAGAATGAAACATTCGGTGAATCGGCC	Genotyping, T-DNA insertion pUBI::ENY2	*	2753
CGAGTTGAAGGCAATAACAAACGAGA	Genotyping, T-DNA insertion pENY2::GUS	*	2792
GTAGAGTCCATTTAAAACGGCACGTA	Screening for CRISPR/Cas9-induced mutations	*	2793
CGACGCGTATCAGAAACCAAGCTG	Genotyping, T-DNA insertion pSSRP1::eGFP-SSRP1 / pSSRP1::eGFP-SSRP1△HMG	*	3193
GCTCTAGAATGTCTGGCGCAGAGGATAATAAATC	Genotyping, T-DNA insertion SAIL LB (spt16-1)	*	3447
CGGGATCCGTCTCCTTTCACGTTCTCTCGAA	Genotyping, T-DNA insertion sgf11-1 (SAIL_856_F11)	*	3448
GCTCTAGAATGGTGAGCAAGGGCGAGGA	Genotyping, T-DNA insertion p35S::eGFP-NLS	*	3555
AAACAGTGTTAGGCATTAAGCGTACAT	Genotyping, T-DNA insertion pSPT16::GUS	*	3584
ACGGGTAGAAGCGAGTCTCA	Genotyping, Transposon insertion ssrp1-1 (GT7431)	*	3645
AGGGTTTTCGAAGGAAGAGG	Genotyping, Transposon insertion GT7431 (ssrp1-1)	*	3646
TGGTTCACCAACTGATGATTCT	Genotyping, T-DNA insertion pSSRP1::SSRP1, pSSRP1::SSRP1△HMG, pUBI10::SSRP1, and pUBI10::SSRP1△HMG	*	4047
GGAGCATGTAAAGAAGAAAGGG	Genotyping, T-DNA insertion eny2-1 (SALK_045015), eny2-RNAi	*	4090
TCGGTGCCACTTTTTCAAGTTGATA	Genotyping, T-DNA insertion eny2-crispr	*	4608

Table 10.4 List of oligonucleotides for expression analysis
Oligonucleotides were obtained from MWG eurofins genomics (* = from this study, L = from lab collection, No. = lab ID, S. = Source).

Sequence 5' - 3'	Description	S.	No.
AGGTGGTGGAAAAGCTGTGGAA	qRT-PCR SSRP1 (AT3G28730)	L	1337
TGATCTCGAAATCCAACGAACTTGT	qRT-PCR SSRP1 (AT3G28730)	L	1338
GAGGGGCTCGGGCATTACCAT	qRT-PCR SPT16 (AT4G10710)	L	1785
CCAAAACCGCCTTTGTGTAAAGCT	qRT-PCR SPT16 (AT4G10710)	L	1786
CTCTCAGTGCTTTGTGATGCTGAAG	qRT-PCR SOC1	L	1819
AGAAACCGGTTTGGTGCTGACTC	qRT-PCR SOC1	L	1820
TCTGTGAAGGCAGAGCTGTTGAAC	qRT-PCR ENY2 (AT3G27100)	*	3252
TGGCTGCTATCTGGTTAGTGTTTGG	qRT-PCR ENY2 (AT3G27100)	*	3253
AGCCAAGAAGACCGAACTCA	qRT-PCR FLC	L	3376
TTTGTCCAGCAGGTGACATC	qRT-PCR FLC	L	3377
GGGAAAGGTGCTTGGAGATA	qRT-PCR SSRP1 (AT3G28730)	*	4044
TCCAACGTTTACTACATGCCA	qRT-PCR SSRP1 (AT3G28730)	*	4045
GGAGCATGTAAAGAAGAAAGGG	qRT-PCR ENY2 (AT3G27100)	*	4090
CAACAGCTCTGCCTTCACAG	qRT-PCR ENY2 (AT3G27100)	*	4091
ATGAACTGTGGGCGACAAAT	qRT-PCR SGF11 (AT5G58575)	*	4389
AAGGAGAATATCGTGGATTGG	qRT-PCR SGF11 (AT5G58575)	*	4390
TGCGTATCCTGAAGAGAAGAG	qRT-PCR ANS (AT4G22880)	*	4473
GACGGTCAGGCTCTAAACCT	qRT-PCR ANS (AT4G22880)	*	4474
TGATATTGTTGTGGGCCGTG	qRT-PCR F3âĂŹH (AT3G51240)	*	4475
CCGTTGATCTCACAGCTCTC	qRT-PCR F3âĂŹH (AT3G51240)	*	4476
GGAAGAGAAGATGAGGGCGA	qRT-PCR CHS (AT5G13930)	*	4477
AACAAGACACCCCACTCCAA	qRT-PCR CHS (AT5G13930)	*	4478
TCAGGCCAAAATACCCCGAA	qRT-PCR DFR (AT5G42800)	*	4479
ATGTCCGTCAGCTTCTTGGA	qRT-PCR DFR (AT5G42800)	*	4480
TCTGAGTGTCATCGGGTAGC	qRT-PCR SGF11 (AT5G58575)	*	4483
CTACAGGAGGGTGAGTCTGC	qRT-PCR SGF11 (AT5G58575)	*	4484
AGTGTTAGTTGGACGAAAGTGAC	qRT-PCR SSRP1 (AT3G28730)	*	4661
TGAAAACTGACCATGAACATCTG	qRT-PCR SSRP1 (AT3G28730)	*	4664
CATTATCCCCAAGAACAGCAG	qRT-PCR SSRP1 (AT3G28730)	*	4665
CGTTACCGACCGTTTTCATC	qRT-PCR SSRP1 (AT3G28730)	*	4666
TGGGAAAGTGTTGCCATCC	qRT-PCR GAP (AT1G13440)	*	TH526
CTTCATTTTGCCTTCAGATTCCTC	qRT-PCR GAP (AT1G13440)	*	TH527
ACCCTTGAAGTGGAAAGCTCC	qRT-PCR UBI10 (AT4G05320)	*	TH528
TTCCAGCGAAGATGAGACGC	qRT-PCR UBI10 (AT4G05320)	*	TH529
AACGTGGCCAAAATGATGC	qRT-PCR PP2AA3 (AT1G13320)	*	TH646
CACATTGTCAATAGATTGGAGAGC	qRT-PCR PP2AA3 (AT1G13320)	*	TH647

10.4 Plasmids

Table 10.5 List of plasmids
(* = from this study, A = Ampicillin, B = Basta, Exp. = Use for the following experiments, H = Hygromycin, K = Kanamycin, P. = Project: 1 = Analysis of SSRP1 HMG-box domain, 2 = FACT and anthocyanin synthesis, 3 = ENY2 project, L = from lab collection, No. = lab ID, R. = Plant resistance marker, R. sites = Restriction sites, S. = Source).

Plasmid	P.	Exp.	Description (Insert, plasmid, primers)	R. sites	R.	S.	No.
pGEX-5X-1:GST-SSRP1	1	AP	SSRP1 (CDS) in pGEX-5X-1	NotI	A	L	680
pGEX-5X-1:GST-SSRP1△HMG	1	AP	SSRP1△HMG (CDS) in pGEX-5X-1, 3050 and 3051	BamHI/SalI	A	*	812
pCambia2300:p35S::eGFP-NLS	1	CLSM FRAP	eGFP-NLS in pCambia2300, 3555 and 3556	XbaI/EcoRI	K	*	824
pCambia2300:p35S::eGFP-SSRP1	1	CLSM, FRAP	SSRP1 into p810, 3576 and 3577	SacI/NruI	K	*	827
pCambia2300:p35S::eGFP-SSRP1âĹĘHMG	1	CLSM FRAP	SSRP1âĹĘHMG in p810, 3576 and 3578	SacI/NruI	K	*	828
pCambia2300:pSSRP1::eGFP-gSSRP1	1	CLSM, FRAP	pSSRP1 in p827	HindIII/XbaI	K	*	840
pCambia2300:pSSRP1::eGFP-gSSRP1△HMG	1	CLSM FRAP	pSSRP1 in p828	HindIII/XbaI	K	*	841
pGreen0179:pSSRP1::gSSRP1	1	Compl.	pSSRP1::gSSPR1 in pGreen0179, 3975 and 3977	EcoRV/SacI	H	*	916
pGreen0179:pSSRP1::gSSRP1△HMG	1	Compl	pSSRP1::gSSPR1âĹĘHMG in pGreen0179, 3975 and 3976	EcoRV/SacI	H	*	918
pCambia3300:UBI10::SSRP1	1	OEx	SSRP1 (CDS) in pCambia2300:UBI10	XhoI	B	L	671
pCambia3300:UBI10::SSRP1△HMG	1	OEx	SSRP1âĹĘHMG (CDS) in pCambia2300:UBI10	XhoI	B	L	672
pCambia3300:pSSRP1::GUS	2	GUS	pSSRP1 (-614 bps) in pCambia3300, 896 and 675	HindIII/XbaI	B	*	830
pCambia3300:pSPT16::GUS	2	GUS	pSPT16 (-2428 bps) in pCambia3300, 3551 and 3552	SmaI/BamHI	B	*	823
pCambia3300:pENY2::GUS	3	GUS	pENY2 (-1970 bps) in pCambia3300, 3832 and 3833	HindIII/XbaI	B	*	919
pCambia2300:pENY2::5'eGFP	3	CLSM FRAP	pENY2 (-1970 bps) in p810, 3832 and 3833	HindIII/XbaI	K	*	926
pCambia2300:pENY2::5'eGFP-gENY2	3	CLSM, FRAP	ENY2 (genomic) in p926, 4015 and 4016	SacI/BamHI	K	*	927
pCambia2300:p35S::SG-ENY2	3	AP	ENY2 (CDS) in pCambia2300:p35S::SG	XbaI/BamHI	K	L	634
pCambia2300:p35S::SG-SGF11	3	AP	SGF11 (CDS) in pCambia2300:p35S::SG	XbaI/BamHI	K	*	805
pCambia2300:p35S::SG-MOS4	3	AP	MOS4 (CDS) in pCambia2300:p35S::SG	XbaI/BamHI	K	*	804
pCambia2300:p35S::SG-THP1	3	AP	THP1 (CDS) in pCambia2300:p35S::SG	XbaI/BamHI	K	*	806
pCambia2300:p35S::SG	3	AP	Empty SG Tag	-	K	L	728
pGBKT7	3	Y2H	Bait vector for Y2H	-	K	L	68
pGADt7	3	Y2H	Prey vector for Y2H	-	A	L	69
pGADT7:ENY2	3	Y2H	ENY2 (CDS) in p69, 3194 and 3195	NdeI/BamHI	A	*	815
pGBKT7-ENY2	3	Y2H	ENY2 (CDS) in p68, 3194 and 3195	NdeI/BamHI	K	*	813
pGADT7-SGF11	3	Y2H	SGF11 (CDS) in p69, 3196 and 3197	NdeI/BamHI	A	*	816
pGBKT7-SGF11	3	Y2H	SGF11 (CDS) in p68, 3196 and 3197	NdeI/BamHI	K	*	814
pGADT7:UBP22	3	Y2H	UBP22 (CDS) in p69, 4375 and 4376	EcoRI/SacI	A	*	1155
pGBKT7:UBP22	3	Y2H	UBP22 (CDS) from p1155 in p68	EcoRI/NotI	K	*	1189
pGreen0179:p35S::5'mCherry	3	FRET	mCherry in pGreen0179:p35S, 3873 and 3874	XbaI/SacI	K	*	920

Table 10.5 (Continuation) List of plasmids
(* = from this study, A = Ampicillin, B = Basta, Exp. = Use for the following experiments, H = Hygromycin, K = Kanamycin, P. = Project: 1 = Analysis of SSRP1 HMG-box domain, 2 = FACT and anthocyanin synthesis, 3 = ENY2 project, L = from lab collection, No. = lab ID, R. = Plant resistance marker, R. sites = Restriction sites, S. = Source).

Plasmid	P.	Exp.	Description (Insert, plasmid, primers)	R. sites	R.	S.	No.
pGreen0179:p35S::mCherry-NLS	3	FRET	NLS in p920, 3834 and 3835	EcoRI	K	*	921
pCambia2300:p35S::eGFP-NLS-mCherry	3	FRET	mCherry in p923, 3873 and 3874	NruI/SacI	K	*	966
pCambia2300:p35S::eGFP-NLS	3	FRET	Altered NLS in p824, 3978 and 3979	EcoRI	K	*	923
pCambia2300:p35S::eGFP-SGF11	3	FRET	SGF11 (CDS) in p810, 3755 and 3197	NruI/BamHI	K	*	932
pGreen0179:p35S::mCherry-ENY2	3	FRET	ENY2 (CDS) from p929 (NruI/BamHI) in p920 (SmaI/BamHI)	NruI-SmaI/ BamHI	K	*	931
pCambia2300:p35S::eGFP-ENY2	3	FRET	ENY2 (CDS) in p810, 3754 and 3195	NruI/BamHI	K	*	929
pCambia2300:p35S::eGFP-UBP22	3	FRET	UBP22 (CDS) in p810, 4545 and 4546	SacI/EcoRI	K	*	1300
pGreen0179:p35S::mCherry-SGF11	3	FRET	SGF11 (CDS) from p932 in p920	SacI/BamHI	K	*	1301
pGADT7:CDC5	3	Y2H	CDC5 (CDS) from p910 in p69	BamHI/NdeI	A	*	900
pGADT7:MAC3A	3	Y2H	MAC3A (CDS) from p911 in p69	BamHI/NdeI	A	*	901
pGADT7:MAC3B	3	Y2H	MAC3B (CDS) from p912 in p69	EcoRI/NdeI	A	*	902
pGADT7:SKIP	3	Y2H	SKIP (CDS) from p913 in p69	BamHI/NdeI	A	*	903
pGADT7:MOS4	3	Y2H	MOS4 (CDS) from p914 in p69	BamHI/NdeI	A	*	904
pGADT7:ISY	3	Y2H	ISY1 (CDS) from p915 in p69	BamHI/NdeI	A	*	905
pGADT7:MAC7	3	Y2H	MAC7 (CDS) from p1089 in p69	SmaI/XhoI	A	*	1085
pGADT7:CRN1	3	Y2H	CRN1 (CDS) in p69, 4030 and 4031	EcoRI/BamHI	A	*	1086
pGBKT7:CDC5	3	Y2H	CDC5 (CDS) from p900 in p68	NdeI/BamHI	A	*	1087
pGADT7:PRL1	3	Y2H	PRL1 (CDS) from p967 in p69	NdeI/EcoRI	A	*	968
pCambia2300:p35S::eGFP-CDC5	3	FRET	CDC5 (CDS) in p810, 3720 and 3721	NruI / BamHI	K	*	910
pCambia2300:p35S::eGFP-MAC3A	3	FRET	MAC3A (CDS) in p810, 3716 and 3717	NruI / BamHI	K	*	911
pCambia2300:p35S::eGFP-MAC3B	3	FRET	MAC3B (CDS) in p810, 3718 and 3719	NruI / BamHI	K	*	912
pCambia2300:p35S::eGFP-SKIP	3	FRET	SKIP (CDS) in p810, 3710 and 3711	NruI / BamHI	K	*	913
pCambia2300:p35S::eGFP-MOS4	3	FRET	MOS4 (CDS) in p810, 3714 and 3715	NruI / BamHI	K	*	914
pCambia2300:p35S::eGFP-ISY1	3	FRET	ISY1 (CDS) in p810, 3712 and 3713	NruI / BamHI	K	*	915
pGreen0179:p35S::mCherry-MAC7	3	FRET	MAC7 (CDS) in p920, 4036 and 4037	SmaI-NruI/ EcoRI	K	*	1089
pCambia2300:p35S::eGFP-PRL1	3	FRET	PRL1 (CDS) in p810, 4032 and 4033	SacI/EcoRI	K	*	967
pCambia2300:p35S::eGFP-CRN1	3	FRET	CRN1 (CDS) in p810, 4030 and 4031	SacI/BamHI	K	*	972
pGreen0179:p35S::mCherry-SYF1	3	FRET	SYF1 (CDS) in p920, 4028 and 4029	SacI/BamHI	K	*	1091
pCambia2300:p35S::eGFP	3	FRET	eGFP in pCambia2300:p35S, 3555 and 3557	XbaI/EcoRI	K	*	810
pCambia2300:p35S::ADA2b-SG	3	AP	ADA2b in pCambia2300:p35S::SG (p805), 4280 and 4281	SalI/BamHI	K	*	1151
pCambia2300:p35S::TAF13-SG	3	AP	TAF13 in pCambia2300:p35S::SG (p805), 4378 and 4379	XbaI/SacI	K	*	1152
pCambia2300:35S::UBP22-SG	3	AP	UBP22 in pCambia2300:p35S::SG (p805), 4278 and 4279	XbaI/SacI	K	*	1153
pCambia2300:35S::TAF10-SG	3	AP	TAF10 in pCambia2300:p35S::SG (p805), 4625 and 4626	BamHI/XbaI	K	*	1302
pFGC5941:ENY2	3	RNAi	ENY2 (4-312) sense (s) and antisense (as) in pFGC5941, by 2662/2663 and 2678/2679, respectively.	BamHI/XbaI(s) NcoI/XhoI(as)	B	*	629
pHEE401	3	CRISP	CRISPR/Cas9 (Wang et al. 2015)		K	*	94
pHEE401 guide 9 (ENY2)	3	CRISPR	ENY2 guide 9 (20 nt-spacer sequence) in pHEE401 (p94), 4488 and 4489	BsaI	K	*	1149
pGreen:pUBI10::ENY2	3	OEx	ENY2 (CDS) in pGreen0179:UBI10	BamHI/blunt	K	*	628

10.5 Organisms

Table 10.6 List of T-DNA lines

Name	T-DNA insertion	AGI	Source
eny2-1	SALK_045015	AT3G27100	NASC
sgf11-1	SAIL_856_F11	AT5G58575	NASC
ssrp1-1	GT7431	AT3G28730	Cold Spring Harbor Loratories
ssrp1-2	SALK_001283	AT3G28730	NASC
spt16-1	SAIL_392_G06	AT4G10710	NASC
hub1-3	GABI_276D08	AT2G44950	GABI-Kat

Table 10.7 List of bacteria and yeast strains

Organism	Name	Resistance	Purpose	Company
A. tumefaciens	GV3101::pMP90	Gentamycin, Rifampicin	Plant transformation	DSMZ
A. tumefaciens	GV3101::pMP90 + pSOUP	Gentamycin, Rifampicin, Tetracyclin	Plant transformation	DSMZ
E. coli	BL21 Rosetta	Tetracyclin	Protein expression	Novagen
E. coli	XL1blue	Tetracyclin	Plasmid amplification	Stratagene
S. cerevisiae	AH109	-Ade -His -Leu -Trp	Y2H	Clontech

10.6 Databases, Online Tools, Software

> AgriGO (`http://bioinfo.cau.edu.cn/agriGO/`)

> Excel2Latex (`https://ctan.org/tex-archive/support/excel2latex/`)

> Gene Expression Omnibus (`https://www.ncbi.nlm.nih.gov/geo/`)

> Geneinvestigator (`https://genevestigator.com/gv/`)

> **ImageJ 1.49d** (`https://imagej.nih.gov/ij/`)

> **Inkscape** (`https://inkscape.org/en/`)

> I-Tasser (`https://zhanglab.ccmb.med.umich.edu/I-TASSER/`)

> LAS AF V 3.1.0 build 8587 (Leica Microsystems)

> Mendeley (`https://www.mendeley.com/`)

> **Microsoft Excel 2016** (`https://www.microsoft.com/`)

> **MiKTEX**(`https://miktex.org/`)

> Needle (EMBOSS) (`https://www.ebi.ac.uk/Tools/psa/emboss\_needle/`)

> Pfam database v30.0 (`http://pfam.xfam.org/`)

> Primer3 v 0.4.0 (`http://bioinfo.ut.ee/primer3-0.4.0/primer3/input.htm`)

> Smart database (`http://smart.embl-heidelberg.de/`)

> **Snapgene v2.3.2** (`http://www.snapgene.com/`)

> SPRING (`https://zhanglab.ccmb.med.umich.edu/spring/`)

> **TEXmaker** (`http://www.xm1math.net/texmaker/`)

> **The Arabidopsis Information Resource v10** (`https://www.arabidopsis.org/`)

> **UniProt** (`http://www.uniprot.org/`)

> Venny v2.1 (`http://bioinfogp.cnb.csic.es/tools/venny/`)

11 Methods

11.1 Nucleic acid based methods

11.1.1 Isolation of genomic DNA from *Arabidopsis* leaves

The method for the isolation of genomic DNA is based on Edwards et al. 1991. One to two small leaves of *Arabidopsis* plants were harvested in an Eppendorf tube, frozen with two glass beads in liquid nitrogen and homogenized using the Tissue Lyser II (Qiagene) with a frequency 30 Hz for 1 min. 400 μL freshly prepared Edward buffer (200 mM Tris pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5 % (w/v) SDS) was added to the ground tissue. The sample was briefly vortexed and centrifuged for 5 min. (12000 g, RT). To precipitate the DNA, 300 μL of the supernatant were mixed with an equal volume of 100 % isopropanol, briefly vortex and incubated at RT for 2 min. After centrifugation for 5 min. (12000 g, RT), the DNA pellet was washed once with 70 % ethanol (v/v), air dried and resuspended in 50 μL H_2O.

11.1.2 Isolation of RNA from *Arabidopsis* leaves

For the isolation of RNA, 50-100 mg of *Arabidopsis* leave tissue was homogenized using the Tissue Lyser II (Qiagene) with a frequency 30 Hz for 1 minute. The RNA extraction was performed using the TRIzol^{TM}reagent (Invitrogen) or the RNeasy® Mini Plant kit (Qiagen) according to the manufacturer′s instructions. The purified RNA was dissolved in 20 (TRIZol) or 50 (RNeasy) μL H_2O. To remove DNA contaminations, 3 μg extracted RNA was incubated with 2 U of DNaseI (NEB) for 80 minutes at 37 °C according to the manufacturer′s instructions.

11.1.3 Reverse Transcription (cDNA synthesis)

Isolated and DNaseI-treated RNA was transcribed into cDNA using RevertAid^{TM}H Minus M-MuLV Reverse Transcriptase (Thermo Fisher Scientific). In a total volume of 11 μL, 2 μg RNA were incubated with 0.5 μg oligo-dT primers for 5 min. at 70 °C and subsequently cooled down to 4 °C. Reaction buffer (1x), dNTP (1mM) and 20 U RNase Inhibitor (Thermo Fisher Scientific) were added to a total volume of 19 μL and the mixture was incubated for 5. min at 37 °C. To synthesize cDNA, 200 U of RevertAid^{TM}H Minus M-Mul V Reverse Transcriptase were added and the sample was incubated for 42 min. at 60 °C and then for 10 min at 70 °C. The reaction was stopped by heating the samples at 70 °C for 10 min.

11.1.4 Polymerase chain reaction (PCR)

In general, Taq DNA Polymerase (PEQLAB) was used for Genotyping-PCR, Colony-PCR and other standard PCR based validations, while the Hercules II Fusion DNA Polymerase was used for cloning due to their proofreading function. The following standard PCR reaction conditions were used: The Taq reaction mix contains 1 x reaction buffer Y, 0.2 μM of each primer, 0.3 mM dNTP, template DNA as needed, 1.25 U Taq Polymerase and H_2O up to 25 μL. The Hercules II reaction mix contains 1 x Hercules II reaction buffer, 0.2 μM of each primer, 0.3 mM dNTP, template DNA as needed, 0.5 U Hercules II Fusion DNA Polymerase and H_2O up to 50 μL. The amount of template DNA was depending on the type of DNA: 1-10 ng of plasmid DNA, 4 μL of extracted genomic DNA (0.1 - 1 μg), and 2 μL of synthesized cDNA (0.1 - 2 μg)

were used. The PCR reaction was performed using a T1 Thermocycler or T3000 Thermocycler (Biometra GmbH, Göttingen). The PCR settings were depending on the DNA polymerase, the primers and the size of the amplified DNA fragment and were depicted in Table 11.1. The PCR products were mixed with 10 x loading buffer (50 mM Tris-HCl pH 7.5, 0.5 mg/mL bromphenol blue, 0.5 mg/mL xylene cyanol, 0.6 mL/mlL glycerol) and were analyzed on a 1 - 2 % agarose gel depending on the size of the fragment.

Table 11.1 Cycling conditions for PCR reactions using Taq or Hercules II DNA Polymerase.
A) Settings to amplify genomic or plasmid targets B) Settings to amplify cDNA targets

	Taq		Hercules II		
Step	**Temp. [°C]**	**Time [sec.]**	**Temp. [°C]**	**Time [sec.]**	**Cycles**
Initial Denaturation	95	300	95	120	1
Denaturation **Annealing** **Extension**	95 Primer Tm -5 °C 72	30 30 60 sec. per 1 kb	95 Primer Tm -5 °C 68^{A} / 72^{B}	20 20 30 sec. per 1 kb^{A} / 60 sec. per 1 kb^{B}	34
Final extension	72	300	68^{A} / 72^{B}	300	1
Storage	4	PAUSE	4	PAUSE	

11.1.5 Real time quantitative PCR (qRT-PCR)

The qPCR reactions with a total volume of 10 μL were performed using KAPA^{TM} SYBR® FAST QPCR MasterMix Universal (PEQLAB), G003-SF stripes (Kisker Biotech GmbH and Co KG) and the Mastercycler epgradient S realplex^2 with realplex software v2.2 (Eppendorf AG) according to the manufacturer′s instructions. For expression analysis, total RNA was isolated with RNeasy Mini Plant kit (Qiagen) (Section 11.1.2) and transcribed into cDNA (11.1.3) as template for the qRT-PCR. Targets were quantified with specific primer pairs (Table XXX) that were design with the web application primer3 (Untergasser et al. 2012). The following cycling program was used: 1) Initial Denaturation (2 sec. at 98°C), 2) Two step cycling (40 x 5 sec. at 98°C followed by 15 sec. at 59°C) and 3) Melting curve for quality control. The normalised relative quantities (NRQ) were calculated according to (Hellemans et al. 2007) using the three references genes GAPC, PP2AA3 and UBI10 (Kudo et al. 2016). The standard error (SE) of the NRQ was calculated without taking the SE derived of the oligo efficiency determination into consideration. Primer efficiency for the specific primer pairs was calculated using a three-step dilution of the cDNA template.

11.1.6 Restriction-Ligation-based Cloning

Restriction-Ligation based cloning was performed according to Sambrook et al. 1989.

11.1.6.1 Agarose gel electrophoresis

For gel electrophoresis, 1-2 % (w/v) agarose gels supplemented with 0,005 % (v/v) ethidium bromide were cast and run (140 V) using 1x TAE buffer (40 mM Tris, 20 mM acetic acid, and 1 mM EDTA). DNA/RNA samples were mixed with 10 x DNA loading buffer (50 mM Tris-HCl pH 7.5, 0.5 mg/ml bromphenol blue, 0.5 mg/mL xylene cyanol, 0.6 mL/mL glycerol) to a final concentration of 1x. Nucleic acids were visualized by excitation at 256 nm using the BioDoc Analyser (Biometra GmbH, Göttingen) or for preparative gels at 354 nm using a UV lamp.

11.1.6.2 DNA extraction from agarose gels and PCR clean up

For the DNA extraction from agarose gels and the clean-up of PCR samples, the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel) was used according to the manufacturer′s instructions.

11.1.6.3 Phosphorylation and Annealing of DNA-Oligos

In a total volume of 40 μL, equal amounts of DNA oligonucleotides (200 μM) were incubated with 10 U of T4 Polynucleotide Kinase (NEB) in 1x T4 Ligase buffer (NEB) in the T3000 Thermocycler (Biometra GmbH, GÃűttingen) using the following program: 30 min at 37 °C, 5 min. at 95 °C, the temperature decreased stepwise 5 °C every 30 sec. to RT.

11.1.6.4 Restriction digestion, Dephosphorylation and Klenow fragment

PCR fragments or plasmids were digested with restriction enzymes from NEB or Fermentas according to the manufacturer′s instructions. Digestions were performed over night.

To prevent self-ligation, 5′ - phosphate groups from digested plasmids were removed by incubation with 5 U of Antarctic Phosphatase (NEB) in 1 x Antarctic Phosphatase buffer for 30 - 60 min. at 37 °C.

Blunt ends were formed by incubating the DNA fragmnent with 1 U Klenow Fragment (Fermentas) in 1 x T4 Ligation buffer (NEB) for 30. min at 37 °C. The enzyme was heat inactivated by incubating the samples for 10. min at 75 °C.

11.1.6.5 Ligation

For standard ligation reactions, the insert and the plasmid DNA (4 : 1 molar ration) were incubated in 1x T4 DNA Ligase Buffer (NEB) with 400 U of T4 DNA Ligase o/n at 4 °C in a thermos canister.

For oligo-plasmid ligation reactions, 100 ng plasmid DNA was mixed with 20 μM phosphorylated and annealed DNA oligos (Section 11.1.6.3), 400 U T4 ligase (NEB) and 20 U BsaI-HF (NEB) in 0.75 x cutsmart (NEB) and 0.75 x T4 Ligase (NEB) buffer. The ligase reaction was performed using the T3000 Thermocycler (Biometra GmbH, Göttingen) with the following program: 1) 20x Two-step cycling (2 min. at 37 °C and 2 min. at 16 °C), 2) 10 min. at 80 °C.

11.1.6.6 Isolation of plasmid DNA from *E. coli*

For minipreparation of plasmid DNA, 4 mL selective LB-medium were inoculated with a positive transformed *E. coli* colony and incubated o/n at 37 °C and 200 rpm. Cells of 3 ml culture were pelleted by centrifugation for 3 min. at 2000 g and subsequently resuspended in 200 μL P1 buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 μg/mL RNase A). To lyse the cells, 300 μL of P2 buffer (0.2 M NaOH, 1 % (w/v) SDS) was added to the sample and the mixture was incubated for 5 min. at RT. To stop the cell lysis, 300 μL of P3 buffer (3 M potassium acetate, pH 4.8) were added and the sample was incubated for 10 min on ice before subsequently centrifugation (12000 g, 10 min, RT). The supernatant was transfered to a new Eppendorf tube. An equal amount of 100 % (v/v) isopropanol was added and the mixture was incubated for 5 min. at RT to precipitate the plasmid DNA. After centrifugation (12000 g, 10

min., RT), the pellet was washed with 70 % (v/v) ethanol, air dried and re-dissolved in 50 μL H_2O.

Midipreparation of plasmid DNA was performed using the NucleoBond® Xtra Midi Kit (Macherey Nagel) according to the manufacturer's instructions.

11.1.6.7 Sequencing

Sequencing of purified plasmid DNA (11.1.6.6) or PCR (Section 11.1.6.2) products was performed by the TubeSeq Service of Eurofins MWG Operon (Ebersberg). DNA samples and sequencing primers were prepared according to instructions of the provider (https://www.eurofinsgenomics.eu).

11.2 Protein based methods

11.2.1 Protein Extraction from *Arabidopsis* flowers

To extract total proteins, 100 mg *Arabidopsis* flowers were harvested in a 2 ml Eppendorf tube, frozen with one metal bead in liquid nitrogen and homogenized using the Tissue Lyser II (Qiagene) with a frequency 30 Hz for 1 min. 500 μL extraction buffer (25 mM Hepes pH 7.4, 0.05 % IGEPAL CA-630, 1 mM DTT, 2 mM $MgCl_2$, 5 mM EGTA, 10% glycerol, cOmplete^{TM}EDTA free proteinase inhibitor tablets (Sigma-Aldrich), 1 mM PMSF dissolved in 2-propanol) were added to the ground tissue and samples were mixed using the Tissue Lyser II with a low frequency for 1 min. $MgCl_2$ was added to a final concentration of 5 mM as well as 50 U/mL Benzonase. The mixture was incubated for 30 minutes at 4 °C on a rotating wheel. The cell debris was pelleted by centrifugation (16.000 g, 10 min., 4 °C). To analyze the sample by SDS-PAGE, 6x SDS loading dye (250 mM Tris-HCl pH 7.5, 10 % (w/v) SDS, 30 % (v/v) glycerol, 0.5 M DTT, 0.1 % (w/v) Bromophenol Blue) was added to the supernanat to get a final concentration of 1x and the sample was heated for 5 minutes at 90 °C.

11.2.2 Protein purification

11.2.2.1 Affinity Purification

GST-tagged proteins

The bacteria pellet of an upscaled and IPTG-induced 2 l culture was resuspended in 35 ml lysis buffer (50 mM Tris-HCl pH 7.6, 500 mM NaCl, 0.05 % IGEPAL CA-630, 10 % glycerol, 2 mM DTT, 2 mM EDTA, 2 mM EGTA, 0.5 mM PMSF dissolved in 2-propanol, cOmplete^{TM}EDTA free proteinase inhibitor tablets (Sigma-Aldrich); adjusted to pH 7.6). To disrupt cells and shear the nucleic acids, the cell suspension was sonicated with six 15-seconds bursts at 30 % intensity followed by 60-seconds intervals for cooling using a UW2070 MS73 (Bandelin) Sonicator. The sample was centrifuged with a SS34 rotor (Sorvall® Superspeed RC2-B) for 60 min. at 40000 g. 2 mL washed Glutathione cellulose (Carl Roth GmbH and Co. KG) was added to the supernatant and the mixture was incubated for 80 min. at 4 °C on a rotating wheel. This mixture was filled in an empty PD-10 column, to remove the flow through by gravity. The cellulose were washed with 100 mL lysis buffer. Bound proteins were eluted twice by shaking the cellulose each time for 15 min. in 5 mL elution buffer (50 mM Tris-HCl pH 8.0, 20 mM reduced glutathione, 1 mM DTT). All steps of the protein purification were monitored by SDS-PAGE analysis. All

steps including the centrifugations were performed at 4 °C or on ice. To remove further impurities, the recombinant proteins were subsequently purified by anion exchange chromatography (Section 11.2.2.3).

GS-tagged proteins

The affinity purification of SG-tagged proteins that utilizes suspension cultured cells as experimental system was performed as described in Pfab et al. 2017. Briefly, 15 g of transformed PSB-D cells were frozen in liquid nitrogen and ground to a fine powder using mortar and pestle. The homogenised cell material was divided into two 50 mL Falcon tubes and mixed each with 10 mL prechilled extraction Buffer (25 mM HEPES-KOH pH 7.4, 100 mM NaCl, 0.05 % IGEPAL CA-630, 1 mM DTT, 2 mM $MgCl_2$, 5 mM EGTA, 10 % glycerol, cOmplete^{TM}EDTA free proteinase inhibitor tablets (Sigma-Aldrich), 1 mM PMSF dissolved in 2-propanol). The semi-thawed slurry of both Falcon tubes was pooled (Total volume of 30ml) and kept on ice. To disrupt the cells and shear the nucleic acids the cell suspension was sonicated with five 30-seconds bursts at 30 % intensity followed by 60-seconds intervals for cooling using a UW2070 MS73 (Bandelin) Sonicator. $MgCl_2$ (to a final concentration of 5 mM) and 50 U/mL Benzonase were added to the mixture. The sample was incubated for 30 minutes at 4 °C on a rotating wheel to allow degradation of nucleic acids. Cell debris was removed by centrifugation at 40.000 x g at 4 °C for 60 minutes. The supernatant was filtered through a 0.45 μm syringe filter. The protein concentration of the cell extract was determined using the Bradford protein assay and 150 mg of total proteins were adjusted with extraction buffer to a final volume of 30 mL. 50 μL were kept as input sample. 100 μL washed (three times with extraction buffer) magnetic beads were added to the protein extract and the mixture was incubated for 60 minutes at 4 °C on the rotating wheel. After centrifugation (2.000 g, 15 min., at 4 °C), the supernatant was discared and the beads were transfered into an Eppendorf tube. The beads were washed three times with 1 mL extraction buffer using the magnetic rack. Proteins were eluted by adding 300 μL elution buffer (0.1 M glycine-HCl, adjusted to pH 2.7) to the washed beads and incubating the mixture for 5 minutes at RT shaking at 700 rpm. After pelleting the beads with the magnetic rack, the supernatant that contains the purified proteins was transfered to new Eppendorf tube.

GFP-tagged proteins

For the affinity purification of GFP-tagged proteins, 4 g of transformed PSB-D cells were frozen in liquid nitrogen and ground to a fine powder using mortar and pestle. 8 mL prechilled extraction buffer (25 mM HEPES-KOH pH 7.4, 100 mM NaCl, 0.05 % IGEPAL CA-630, 1 mM DTT, 2 mM $MgCl_2$, 5 mM EGTA, 10 % glycerol, cOmplete^{TM}EDTA free proteinase inhibitor tablets (Sigma-Aldrich), 1 mM PMSF dissolved in 2-propanol) were added to the homogenised cell material and the samples were kept on ice. To disrupt the cells and shear the nucleic acids the cell suspension was sonicated with four 30-seconds bursts at 2x 20 % and 2x 10 % intensity followed by 60-seconds intervals for cooling using a UW2070 MS73 (Bandelin) Sonicator. $MgCl_2$ (to a final concentration of 5 mM) and 50 U/mL benzonase were added to the mixture. The sample was incubated for 60 minutes at 4 °C on a rotating wheel to allow degradation of nucleic acids. Cell debris was removed by slow centrifugation at 1.800 x g at 4 °C for 10 minutes to avoid pelleting of chromatin-bound proteins. The supernatant was filtered through a 0.45

μm syringe filter. In case of several samples, the protein concentration of each cell extract was determined using the Bradford protein assay and the concentrations of all samples were adjusted with extraction buffer. 50 μL of each sample were kept as input sample. 25 μL washed (three times with extraction buffer) GFP-Trap® agarose beads (Chromotek) were added to the protein extract and the mixture was incubated for 120 minutes at 4 °C on the rotating wheel. After centrifugation (2.000 g, 5 min., at 4 °C), the supernatant was discared and the beads were transfered to an Eppendorf tube. The beads were washed three times with 1 mL extraction buffer. The proteins were eluted by adding 100 μL 2x SDS loading dye and heating the sample at 90 °C for 5 min.

11.2.2.2 Size-exclusion chromatography

Size-exclusion chromatography was performed on a Pharmacia FPLC system (Gradient Programmer GP250). Protein sample with a volume of 600 μl was applied on a Superose^{TM}6 column (GE Healthcare Life Sciences, Sweden) equilibrated with sterile filtered Tris buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.5 mM EDTA, 1 mM DTT, 1 mM PMSF dissolved in 2-propanol). The eluted proteins were collected in 300 μL fractions at a flow rate of 0.5 mL/min. The elution profile was monitored spectrometrically at 280 nm. To estimate the masses of the eluted proteins, the column was calibrated with thyroglobulin (670 kDa), apoferritin (443 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa) and ovalbumin (43 kDa).

11.2.2.3 Ion-Exchange chromatography

The anion exchange chromatography was performed in a batch procedure. The purified GST-tagged proteins (Total volume of 10 ml) (Section 11.2.2.1) were mixed with 200 μL washed ResourceQ particles and diluted with 45 mL 0M-salt-buffer D (10 mM phosphate buffer pH 7.5, 1 mM EDTA, 1 mM DTT, 1 mM PMSF dissolved in 2-propanol, , cOmplete^{TM}EDTA free proteinase inhibitor tablets (Sigma-Aldrich)) in a 50 mL Falcon tube to a final salt concentration of 125 mM. Sample was incubated for 30 min. at 4 °C on a rotating wheel. After centrifugation (500 g, 5 min., 4°C) the ResourceQ particles were transfered to an Eppendorf tube. ResourceQ pellet was washed with 1 ml 0M-salt-buffer D and two time with 1 ml 200mM-salt-buffer D (10 mM phosphate buffer pH 7.5, 200 mM NaCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF dissolved in 2-propanol, cOmplete^{TM}EDTA free proteinase inhibitor tablets (Sigma-Aldrich)). Proteins were eluted twice with 400 μL 600mM-salt-buffer D (10 mM phosphate buffer pH 7.5, 600 mM NaCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF dissolved in 2-propanol, cOmplete^{TM}EDTA free proteinase inhibitor tablets (Sigma-Aldrich)). For desalting, the recombinant proteins were subjected to o/n dialysis (Section 11.2.2.4).

11.2.2.4 Dialysis

Before use, dialysis tubes were boiled for 15 min. in 100 mM $NaHCO_3$ and 1 mM EDTA. After washing with H_2O, the dialysis tubes were filled with 1 ml of purified proteins (Section 11.2.2.3), sealed with clamps and incubated o/n in 1 L buffer D (10 mM phosphate buffer pH 7.5, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF dissolved in 2-propanol) at 4 °C on a magnetic stirrer. The purified and desalted recombinant proteins ($\sim$ 70 ng/μL) were stored

in 50% glycerol at -20Âř for EMSA experiments (Section 11.2.5).

11.2.2.5 Acetone precipitation

To precipitate proteins, the sample was mixed with ice-cold acetone to a final concentration of (v/v) 20 %, vortexed and incubated o/n at -20 °C. After centrifugation (20.000 g, 20 min., 4 °C), the precipitated proteins were washed twice with 500 μL ice cold acetone. Complete drying of the pellet was avoided in any case. The protein pellet was resuspended in 25 μL 1x PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.47 mM KH_2PO_4, adjusted to a final pH of 7.4 with 1 M HCl).

11.2.3 Protein detection (non-specific)

11.2.3.1 Bradford Assay

10 μL protein sample was filled up with H_2O to a total volume of 200 μL H_2O and mixed with 1 mL Bradford reagent (0.01 % Coomassie Blue G-250, 5 % (v/v) ethanol, 10 % (v/v) phosphoric acid) in a polystyrol cuvette (Sarstedt AG & Co, Germany). After incubation of 10 minutes, the absorbance was measured at 595 nm with the BioPhotometer® (Eppendorf AG, Hamburg). Protein concentration was estimated by comparison of the measured extinction with a BSA calibration curve.

11.2.3.2 SDS-PAGE

Depending on the size of proteins to be analyzed, different resolving gels were prepared with 9 %, 12 % or 18 % (w/v) acrylamide: bisacrylamide (30 : 0.15), 0.75 M Tris pH 8.8, 0.2 % (w/v) SDS, 0.1 % (w/v) ammonium persulfate (APS) and 0.02 % TEMED (v/v) in a Bio-RAD Mini-Protean® 3 Multicaster system and left to polymerize covered with water. Stacking gel solution (5 % (v/v) acrylamide mix Gel 30 (5 : 1), 140 mM Tris pH 6.8, 0.23 % (w/v) SDS, 0.11 % (w/v) APS and 0.06 % (v/v) TEMED) was poured on top of the polymerized separation gel and combs were inserted. Samples to be anazlyzed were mixed with an appropriate amount of 6 x SDS loading buffer (50 mM Tris pH 6.8, 0.002 % (w/v) bromophenol blue, 2.5 % (w/v) glycerol, 1 % (w/v) SDS and 143 mM β-mercaptoethanol) to a final concentration of 1x and heated at 90 °C for 5 min. The SDS polyacrylamide gel electrophoresis (SDS-PAGE) was performed in a Bio-RAD Mini-Protean® 3 running chamber using Laemmli running buffer (0.1% SDS (w/v), 3.03 g/L Tris, and 14.41 g/L glycine). The gels were run at 200 V. The PAGE Rulers ′unstained protein ladder′ (#26614) or ′prestained protein ladder′ (#26616) from ThermoFisher were used as markers.

11.2.3.3 Coomassie Brilliant Blue (CBB) staining

The proteins were visualized by gently shaking the polyacrylamide gel for 30 min. in Coomassie Brilliant Blue (CBB) solution (0.2 % (w/v) CBB G-250, 30 % (v/v) ethanol and 10 % (v/v) acetic acid) and subsequently destaining with 7.5 % (v/v) ethanol and 5 % (v/v) acetic acid o/n.

11.2.3.4 Ponceau staining

The blotted proteins were stained by gently shaking the membrane in 10 mL Ponceau S solution (Sigma Aldrich) for 30 min. and repeatedly washing of the membrane in 50 mL H_2O till the bands become visible.

11.2.4 Protein detection (specific)

11.2.4.1 Western blotting

For Western Blotting, the protein samples were first separated by SDS-PAGE (Section 11.2.3.2) and then transferred onto Amersham Hybond LFP 0.2 PVDF membrane using a Semidry Blotter Maxi (Roth, Germany). Therefore, the blotting membrane was activated by incubation for 30 sec. in 100 % methanol and equilibrated together with the Whatman paper (Biometra) for 10 min. in blotting buffer (20 % (v/v) methanol, 200 mM glycine, 20 mM Tris, 0.01 % (w/v) SDS). The following sandwich was assembled in the blotter from the bottom to the top: three pieces of Whatman paper, blotting membrane, SDS-gel and three pieces of Whatman paper. The proteins were blotted to the membrane by applying 50 mA per gel for three hours. Unspecific binding sites were blocked by incubating the membrane for one hour in 15 mL blocking buffer (5 % (w/v) skimmed milk powder, 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05 % (v/v) Tween 20) on the rotating wheel at 4 °C. The primary antibody was added in a 1:2000 dilution and incubated o/n. The membrane was washed three times for 10 min. with washing buffer (0.1 % (v/v) Triton X-100, 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05 % (v/v) Tween 20). Then, the membrane was incubated for two hours in secondary binding buffer (5 % (w/v) skimmed milk powder, 0.047 % (w/v) Thimerosol, 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05 % (v/v) Tween 20) with respective secondary antibody. For chemiluminescent detection, 1 : 5000 Anti-Rabbit or Anti-Mouse IgG peroxidase coupled antibody (Sigma Aldrich) was added. For fluorescent detection, 1 : 5000 ECL Plex Goat-α-Rabbit IgG-Cy3 or ECL Plex Goat-α-Mouse IgG-Cy5 (Amersham) was added. After the incubation, the membrane was washed again three times with washing buffer. The blotting as well as all incubation and washing steps were performed at 4 °C. For chemiluminescent detection, the membrane was incubated for 10 min. in SuperSignal® West Pico Chemiluminescent substrate (Thermo Scientific) or, in case of weak signals, for 1 - 5 min. in SuperSignal^{TM}West Femto Maximum Sensitivity Substrate for 1-5 min. Chemiluminescene was detected by MultiimageTM FlurChem FC2 imager (Alpha Innotech). For fluorescent detection, the membrane was rinsed with H2O, completely air-dried in the dark and imaged with the laser scanner Typhoon FLA 9500 (GE Healthcare).

11.2.4.2 Trypsin digestion and Mass spectrometry

In-gel digestion of purified proteins

This method was modified according to Pfab et al. 2017. The proteins to be analyzed were separated by SDS-PAGE electrophoresis (Section 11.2.3.2) using a 9 % gel. The gel was stained with with Coomassie Brilliant Blue (Section 11.2.3.3). Using a scarpel and a glass plate, the whole lane was cut out of the gel and was divided into $\sim$ 8 gel pieces. Each piece was cut in 1 mm stripes and was transfered into a 2 mL Eppendorf tube. The gel pieces were washed four times

at RT as follows: For the first washing step add 1 mL Washing Buffer 1 (50 mM NH_4HCO_3), incubate samples for 60 minutes while gently shaking and remove the supernatant by pipetting. For washing steps 2 – 4, repeat this procedure successively by using 1 mL Washing buffers 2 (Mix 50 mM NH_4HCO_3 and acetonitrile in the ratio 3:1) for 60 min., 1 mL washing buffer 3 (Mix 50 mM NH_4HCO_3 and acetonitrile in the ratio 1:1) for 30 min. and 200 μL washing buffer 4 (acetonitrile) for 10 min. Lyophilise the samples for 1 hour to dehydrate gel slices. To reduce cysteins, 200 μL of 1 mg/ml DTT dissolved in 50 mM NH_4HCO_3 were added to the samples and incubated for 35 min. at 56 °C. For carbamidomethylation of the cysteins, remove the supernatant and add 200 μL of 5 mg/ml iodoacetamide dissolved in 50 mM NH_4HCO_3 to the samples. Incubate the samples at RT in the dark for 35 min. Repeat the four washing steps as described above and lyophize the samples again. Transfer lyophilized gel slices from each tube to new 0.5 mL safe-lock Eppendorf tubes. Add 10-20 μL of trypsin mix (0.04 μg/μL trypsin in 50 mM NH_4HCO_3) to the gel pieces and let them soak for 10 min. Cover the transparent gel pieces with 40 μL 50 mM NH_4HCO_3 and incubate the sample o/n at 37 °C. To extract the digested proteins from the gel transfer the supernatant that contains the extracted peptides to a 0.5 mL safe-lock Eppendorf tube (collection tube). Add 40 μL extraction buffer to each tube, incubate the samples for 1 - 2 hours at 39 °C while gently shaking the samples every 10 minutes and transfer the supernatant to the collection tube. Repeat the extraction step once with extraction buffer 1 and once with extraction buffer 3 at 30 °C. The tube with the pooled peptides of all three extraction steps was lyophilize o/n.

Mass spectrometry

The mass spectrometry was performed in the lab of Dr. Astrid Bruckmann (Department of Biochemistry I, University of Regensburg) as described in Antosz et al. 2017.

Briefly, the peptides were separated by reverse-phase chromatography on an UltiMate 3000 RSLCnano System (Thermo Scientific) using a Reprosil-Pur Basic C18 nano column and applying a linear 90-min gradient of 4 to 40% acetonitrile in 0.1% formic acid. The LC system was coupled online to a maXis plus UHR-QTOF system (Bruker Daltonics) via a nanoflow electrospray source (Bruker Daltonics). Data-dependent acquisition of tandem mass spectrometry (MS/MS) spectra by CID fragmentation was performed using a dynamic method with a fixed cycle time of 3 s (Compass 1.7; Bruker Daltonics). Protein Scape 3.1.3 (Bruker Daltonics) in connection with Mascot 2.5.1 (Matrix Science) facilitated database searching of the NCBInr database. Mascot peptide ion score cutoff was set to 25. A protein score of minimum 80 and at least two peptides found with an individual ion score of 25 were considered as criteria for reliable protein identification.

11.2.5 Protein-DNA interactions by EMSA

For detection of protein binding to DNA or nucleosomes, different concentrations of purified recombinant proteins were incubated with 7.5 nM linear DNA or 7.5 nM mononucleosomes for 10 min. at RT in a total of 15 μL of emsa buffer (10 mM phosphate buffer pH 7.5, 80 mM KCl, 1.6 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, 0.2 μg/μL BSA, 1 mM PMSF dissolved in 2-propanol). Orange G was used as loading dye. Complex formation was analyzed by electrophoresis in 0.8

% (w/v) agarose gels in 0.5x TBE buffer, by applying 130 V for 70 min. at RT in the dark. The Cy5-labeled DNA was visualized by the laser scanner Typhoon FLA 9500 (GE Healthcare).

11.3 Cell based methods

11.3.1 Cultivation of bacteria

Luria Bertani (LB) medium (5 g/L Yeast extract, 10 g/L NaCl, 10 g/L tryptone) was used as standard growth medium for all bacterial strains used in this study. To get solid media, 1.5 % (w/v) agar was added prior to autoclaving. For bacterial selection, the LB medium was supplemented with sterile-filtered appropriate antibiotics at the following final concentrations: 100 μg/ml ampicillin dissolved in H_2O, 50 μg/ml gentamycin dissolved in H_2O, 50 μg/ml (*E. coli*) or 25 μg/ml (*A. tumefaciens*) dissolved in H_2O, 50 μg/ml rifampicin dissolved in methanol, 12 μg/ml tetracyclin dissolved in ethanol. *E. coli* and *A. tumefaciens* bacteria strains were grown at 37 °C or 30 °C, respectively. Liquid LB media was aerated by agitation at 200 rpm.

11.3.1.1 Preparation of chemically competent cells

As starter culture, 10 mL LB medium with appropriate antibiotics was inoculated with a single colony of *E. coli* or *A. tumefaciens* bacteria and incubated o/n agitating at 37 °C or 30 °C, respectively. 100 mL selective LB media was inoculated with the required amount of pre-culture to reach an OD_{600} of 0.1 and was grown to an OD_{600} of 0.75. Then, the cells were harvested by centrifugation for 10 min. at 4000 g and 4 °C. The pellet was re-suspended in 30 mL sterile filtered cold TBF1 buffer (100 mM RbCl, 10 mM $CaCl_2$, 50mM $MnCl_2$, 30mM NaOAc; adjusted to pH 5.8 with acetic acid) and incubated on ice for 90 min. The cells were centrifugated for 10 min. at 3000 g and 4 °C. The pellet was re-suspended in 4 mL of sterile filtered cold TBF2 buffer (10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2$ and 15% (v/v) glycerol) and the chemically competent cells were stored in 50 μL aliquots at -80 °C.

11.3.1.2 Transformation by heat shock

Transformation of competent *E. coli* cells

An aliquot of 50 μL chemically competent *E. coli* cells (Section 11.3.1.1) was thawed on ice and mixed with 50 ng plasmid DNA or 5 μL ligation product (Section 11.1.6.5). Before and after the heat shock (2 min., 42 °C), the cells were incubated on ice for 20 and 2 min., respectively. Afterwards, the sample was mixed with 1 mL LB medium without selection and incubated for one hour at 200 rpm at 37 °C. The transformed cells were pelleted by centrifugation (2000 g, 5 min) and re-suspended in 50 μL LB medium. 10 μL of the bacteria suspension were spread on LB plates with the appropriate selection and incubated o/n at 37 °C.

Transformation of competent *A. tumefaciens* cells

An aliquot of 50 μL chemically competent *A. tumefaciens* cells (Section 11.3.1.1) was thawed on ice, mixed with 50 ng plasmid DNA and incubated as following: 1) 5 min. on ice 2) 5 min. in liquid nitrogen 3) 5 min. at 37 °C 4) 5 min. on ice. After the heat shock, the sample was mixed with 1 mL LB medium without selection and incubated for 2-3 hours at 200 rpm at 28-30 °C.

The transformed cells were pelleted by centrifugation (2000 g, 5 min) and re-suspended in 50 μL LB medium. 10 μL of the bacteria suspension were spread on LB plates with the appropriate selection and incubated for 2-3 days at 30 °C.

11.3.2 Protein expression in *E. coli*

As starter culture, 100 mL LB medium supplemented with the appropriate antibiotics were inoculated with a single positive-selected colony from freshly transformed (Section 11.3.1.2) *E. coli* Rosetta^{TM}cells and incubated o/n at 37 °C. Two liter selective LB medium were mixed with the required amount of starter-culture to reach an OD_{600} of 0.1 and were grown to an OD_{600} of 0.8. As non-induced control, 100 μL bacterial culture was centrifuged (4000 g, 1 min.) and the pellet was re-suspended in 1x SDS loading buffer as well as heated at 90 °C for 5 min. The expression of the recombinant proteins was induced by addition of IPTG to a final concentration of 1 mM. The cells were grown for another 2 h at 37 °C. As induced control, 100 μL bacterial culture was collected and treated like the non-induced control above. The cells were harvested by centrifugation (Sorvall® Superspeed RC2-B with SLA1500 rotor) for 15 min. at 6000 g and 4 °C. The bacterial pellet was frozen in liquid nitrogen and stored at -80 °C.

11.3.3 Preparation of chemically competent yeast cells

Yeast AH109 cells from a glycerol stock were streaked on YPAD plates (2.2 % (w/v) micro agar, 2% (w/v) tryptone, 1% (w/v) yeast extract, 2% (w/v) glucose and 0.004% (w/v) adenine hemisulfate) and incubated at for 3 days at 30 °C. As starter culture, 3 mL YPAD (2% (w/v) tryptone, 1% (w/v) yeast extract, 2% (w/v) glucose and 0.004% (w/v) adenine hemisulfate) medium were inoculated with a single colony of AH109 yeast cells and incubated o/n under agitation at 200 rpm at 30 °C. 50 mL YPAD medium was mix with the required amount of starter-culture to reach an OD_{600} of 0.1. The yeast suspension was grown (200 rpm, 30 °C) to an OD_{600} of 0.5 - 1.0. The cells were pelleted by centrifugation for 5 min. at 500 g and washed as following: 1) with 25 mL sterile H_2O 2) with 5 mL sterile filtered SORB buffer (100 mM LiOAc, 10 mM Tris, 1 mM EDTA, 1 M sorbitol; adjust to pH 8.0) 3) with 500 μL SORB buffer. Finally, the washed pellet was re-suspended in 360 μL SORB buffer and mixed with 40 μL ice-cold denatured single stranded 10 mg/mL salmon sperm DNA. For denaturing, the salmon sperm DNA was heated at 90 °C for 5 min. and immediately cooled on ice. 50 μL aliquots of the competent yeast cells were stored at -80 °C.

11.3.4 Co-transformation of yeast cells by heat shock

An aliquot of 50 μL chemically competent yeast cells (Section 11.3.3) was thawed on ice, mixed with each 500 ng of bait (pGBKT7-derivate) and prey (pGADT7-derivate) plasmid DNA as well as with 300 μL sterile filtered PEG solution (100 mM LiOAc, 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 40% PEG3350; adjusted to pH 8.0). The yeast suspension was incubated for 30 min. at RT. The sample was mixed with 40 μL DMSO and incubated for 15 min. at 42 °C. After the heat shock, the cells were centrifugated for 2 min. at 500 g and the pellet was re-suspended in 200 μL H_2O. The total volume was plated on double dropout (DDO) plates (2% (w/v) glucose, 0.67% (w/v) yeast nitrogen base w/o amino acids, 2.2% (w/v) micro agar, 0.064%

(w/v) -Leu/-Trp DO supplement (Clonetech: #630417); adjusted to pH 5.8 and autoclaved for 15min.). DDO plates were incubated for 3-4 days at 30 °C to select for yeast cells with bait and prey cassettes.

11.3.5 Yeast-2-Hybrid Assay

After co-transformation of bait and prey plasmids (Section 11.3.4), a single positive-selected yeast colony was picked from DDO plates and was resuspended in 200 μL H_2O. One half of the yeast suspension was used to determine the OD_{600}. The other half of the yeast suspension was adjusted to an OD_{600} of 1.0. A 1:10 dilution series of the yeast solution was prepared as following: 10^0, 10^{-1}, 10^{-2}, 10^{-3}). The dilution series was plotted on DDO, triple dropout TDO (2% (w/v) glucose, 0.67% (w/v) yeast nitrogen base w/o amino acids, 2.2% (w/v) micro agar, 0.062% (w/v) -His/-Leu/-Trp DO supplement (Clonetech: #630419); adjusted to pH 5.8 and autoclaved for 15min.) and quadrubel dropout plates QDO (2% (w/v) glucose, 0.67% (w/v) yeast nitrogen base w/o amino acids, 2.2% (w/v) micro agar, 0.06% (w/v) -Ade/-His/-Leu/-Trp DO supplement (Clonetech: #630428); adjusted to pH 5.8 and autoclaved for 15min.) using a frogger. The plates were incubated for 3-4 days at 30 °C.

11.4 Plant based methods

11.4.1 Cultivation of *Arabidopsis* plants

Arabidopsis seeds were sown on soil (80% (v/v) Einheitserde Typ ED 73, 10% (v/v) sand and 10% Isoself® from Knauf Perlite) that was pre-soaked in water containing 0.03% (v/v) confidor WG70 (Bayer) and 3 g/L fertiliser Osmocote Start (The Scotts Company). The pots with the seeds were stratified for 48 - 72 hours at 4 °C in the dark and subsequently transferred into the plant growth chamber. Plants were grown under long-day conditions (LD; 16 hour light a day) at 22 °C and 120 μmol m^{-2} s^{-1} light. The plants were watered weekly from the bottom. Seeds of the T-DNA insertion lines *eny2-1*, *sgf11-1* and *hub1-3* were obtained from the European *Arabidopsis* stock centre (http://www.arabidopsis.info/).

For selection of transgenic plants harbouring a bar gene casette in the T-DNA, young seedlings (DAS7) were sprayed three times every two days with a BASTA® solution (100 mg/L glufosinate ammonium (Bayer Crop Science), 200 μL/L Silwet®).

For plant growth under sterile conditions, the surface of *Arabidopsis* seeds was sterilized with chloric gas (40 mL 12.5% hypochloric acid (w/v) and 2 mL 37% HCl (v/v)) in an exsiccator. *Arabidopsis* seeds were sown on solid 0.5x MS plates (2.15 g/L Murashige and Skoog media including vitamins, 1% sucrose, 0.8% phyto agar (w/v); adjusted to pH 5.8). After stratification (2-3 days at 4 °C in the dark), the plates were kept in a plant incubator (Percival Scientific or XXX) under LD conditions at 22 °C and 100 μmol m^{-2} s^{-1}.

For selection of transgenic plants harbouring a nptII (kanamycin resistance marker) or hygR (hygromycin resistance marker) gene cassette in the T-DNA, plants were grown on 0.5x MS supplemented with 50 μg/ml kanamycin or 25 μg/ml hygromycin B, respectively.

For CLSM analysis of *Arabidopsis* roots, seedling were grown on 0.5x MS plates with 1% micro agar that were kept vertically in the plant incubator.

11.4.1.1 Transformation of *Arabidopsis* plants by floral dipping

Arabidopsis plants were stably transformed using the ′Floral Dip′ method described in Clough et al. 1998. Briefly, approximately 20 *Arabidopsis* plants with 10-15 cm high inflorescences were used for each transformation. *A. tumefaciens* cells were transformed with plasmid DNA (Section 11.3.1.2). For transformations of pGreen-derived plasmids the bacteria strain GV3101 + pSoup was used. Selected bacteria colonies were tested by genotyping PCR. Positive tested colonies were grown o/n in 5 mL selective LB at 30 °C and 200 rpm. 500 mL selective LB was inoculated with 500 μL of the starter culture and incubated o/n. The cells were harvested by centrifugation for 15 min. at 5000 g and were re-suspended in 500 mL infiltration medium (5% (w/v) sucrose, 10 mM $MgCl_2$, 0.02% (v/v) Silvet L77 and 0.01 mM acetosyringone dissolved in ethanol). The aerial parts of the plants were dipped into the infiltration medium for 1 min. and were left o/n covered with plastic foil. To harvest the transgenic seeds, the dipped plants were grown to maturity in the plant growth chamber.

11.4.1.2 Crossing of *Arabidopsis* plants

For crossing of *Arabidopsis* plants, one inflorescence of the ′mother plant′(Genotype A) was fixed with tape and all open flowers as well as the meristem were removed with a crossing tweezer. 3-5 flower buds have the right size and were emasculated by gently removing all sepals, petals and stamens. The pollen from the ′father plant′ (Genotype B) was transfered on the stigma of the ′mother plant′ and the inflorescence was labeled. After 14-21 days in the plant growth chamber, the silique with the hybrid seeds was harvested.

11.4.1.3 Soil-based phenotypic analysis

For phenotypic analysis, the plants were grown on soil under LD conditions. The plants were rotated every 2 days in shelf of the growth chamber to avoid position effects. The morphology and development of *Arabidopsis* plants was monitored by documenting specific plant characteristics including the number/shape of leaves or the rosette diameter at several defined developmental stages as described in Boyes et al. 2001. Siliques were bleached by o/n incubation in 75 % ethanol and 25 % acetic acid. All pictures were taken with a Zeiss Discovery V8 stereo microscope or a Canon EOS 600D equipped with a Macro lens EF-S 60 mm 1:2.8 USM (Canon) or a ETS 18-55 mm objective (Canon).

11.4.1.4 GUS staining and clearing of roots

For histochemical GUS staining, *Arabidopsis* plants were incubated in staining solution (50 mM $NaHPO_4$ pH 7.2, 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$, 1% Triton X-100 and 2 mM X-Gluc) in microtiter plates for 6 hours at 37 °C. After the staining, the plants were washed in 100% ethanol until the leaves were cleared. For taking pictures, whole plants were mounted on a microscope slide with a cover slip. All pictures were taken with a Zeiss Discovery V8 stereo microscope or a Canon EOS 600D equipped with a Macro lens EF-S 60 mm 1:2.8 USM (Canon).

11.4.1.5 Determination of anthocyanin content in *Arabidopsis* plants

The extraction and the photometric determination of anthocyanins in the aerial parts of *Arabidopsis* plants was performed as previously described (Yin et al. 2012) with minor modifications. Pools (3-5) of *Arabidopsis* plants were frozen in liquid nitrogen and homogenised using the Tissue Lyser II (Qiagen). 500 μL extraction buffer (acidic methanol, 1 % HCl (v/v)) was added to 10-20 mg ground plant material and the samples were moderately agitated for 15 minutes at 4 °C in the dark. The homogenate was clarified by centrifugation and the supernatant was used for photometric measurements. To quantify the anthocyanin content the following equation was used: (A_{530} - 0.25 x A_{657}) x M^{-1}[g] = relative units of anthocyanin (with A_{530} and A_{657} = absorption at the indicated wavelengths, M = plant fresh weight).

11.4.2 Cultivation of *Arabidopsis* PSB-D cells

The *Arabidopsis landsberg erecta* PSB-D suspension cells (Arabidopsis Biological Resource Center) were cultured and transformed under sterile conditions according to Van Leene et al. 2011. The cells were cultivated in darkness at 25 °C while shaking at 130 rpm. Every week, the cells were diluted by transferring 7 mL cells into a 100 mL Erlenmeyer flask containing 43 mL MSMO medium (0.443 % Murashige and Skoog Salt mixture (US Biological), 3 % sucrose, 0.5 mg/L NAA dissolved in 100 mM NaOH, 100 mg/L myo-inositol, 0.05 mg/L kinetin dissolved in DMSO, 0.4 mg/L thiamine, 25 μg/mL kanamycin if indicated, adjusted to pH 5.7 with 1 M KOH).

11.4.3 Cryopreservation and re-initiation of *Arabidopsis* PSB-D cells

For long term storage of *Arabidopsis* suspension cells, cryo-stocks were generated. 10 mL of 7-days old cell culture were transfered into 100 mL Erlenmeyer flasks containing 40 mL pre-freeze medium (0.443 % Murashige and Skoog Salt mixture (US Biological), 3 % sucrose, 2.7 % mannitol, 0.5 mg/L NAA dissolved in 100 mM NaOH, 100 mg/L myo-inositol, 0.05 mg/L kinetin dissolved in DMSO, 0.4 mg/L thiamine; adjusted to pH 5.7 with 1 M KOH and autoclave) and was incubated for two days at 25 °C and 200 rpm. The pre-frezze culture was transfered into a 50 mL Falcon tube and pelleted by centrifugation (1 min., 130 g). The supernatant was removed and the cell pellet ($\sim$ 2-3 mL) was mixed with 6 mL cryoprotective media (0.443 % Murashige and Skoog Salt mixture (US Biological), 8.56 % sucrose, 1.15 % glycerol, 0.25 % proline, 0.5 mg/L NAA dissolved in 100 mM NaOH, 100 mg/L myo-inositol, 0.05 mg/L kinetin dissolved in DMSO, 0.4 mg/L thiamine; adjusted to pH 7.0 with 1 M KOH and filter sterilize). The mixture was incubated for one hour on ice. The cells in cryoprotective media were aliquoted (2 mL) in ice cold 2 mL Nalgene tubes. The tubes were placed for 100 min. at -80 °C, then frozen in liquid nitrogen (LN2) and stored in a cryogenic container filled with LN2.

To re-initiate *Arabidopsis* cells from the cryo-stock, a frozen Nalgene tube was thawed in a water bath for 2 min. at 45 °C and the cells were spread on MSMO plates (MSMO medium with 0.8 % micro agar) supplemented with appropriate antibiotics if necessary. After incubation of 2 - 3 week at 25 °C, the callus was scraped off the plate, chopped and transfered into a 100 mL Erlenmayer flask containing 30 mL MSMO. After another week of incubation at 25 °C in the shaking incubator, the re-initiated cell culture was diluted once a week as described in Section

11.4.2.

11.4.4 Transformation of *Arabidopsis* PSB-D cells

The *Arabidopsis* suspension cells were transformed by co-cultivation with *A. tumefaciens* harboring plant expression vector. As starter culture, 2 mL LB medium supplemented with kanamycin and rifampicin was inoculated with 3-4 selected *A. tumefaciens* colonies (Section 11.3.1.2) and incubated o/n at 30 °C at 200 rpm. The starter culture (2 mL) was transfered into a 100 mL flask containing 20 mL LB medium with antibiotics and was grown o/n at 30 °C at 200 rpm. The cells were pelleted by centrifugation for 15 min. at 3000 g and re-suspended in 40 mL sterile MSMO by vortexing. This washing step was repeated and the OD600 of the cell suspension was adjusted to 1.0. For co-cultivation, 3 mL of 3-days old *Arabidopsis* suspension cells (OD600: 1.2 - 1.3), 200 μl of the *Agrobacterium* solution and 6 μL 100 mM acetosyringone dissovled in ethanol were pipetted into one well of a 6-well plate. The multiwell plate was taped with Micropore surgical tape and incubated for 3 days in a shaking incubator at 130 rpm and 25 °C.

For selection in liquid medium, the ∼3 mL of transformed cells were transfered into a 25 mL Erlenmayer flask containing 8 mL MSMO supplemented with 50 μL/mL kanamycin (For plasmid selection), 500 μL/mL vancomycin and 500 μL/mL carbenicillin (To kill the *Agrobactrium* cells). The cell culture was incubated for 8 days in a shaking incubator at 130 rpm and 25 °C. The ∼10 mL of transformed cells were transfered into a 100 mL Erlenmayer flask containing 25 mL MSMO supplemented with kanamycin, vancomycin and carbenicillin and the diluted culture was incubated for 7 days as before. Transfer as much cells as possible into a 100 mL Erlenmayer flask containing 35 mL MSMO supplemented with kanamycin and incubate the cell culture for 7 days as before. The transformed cell culture was now diluted every week as described in Section 11.4.2.

For selection on solid medium, the ∼3 mL of transformed cells were washed in a 50 mL Falcon tube with 40 mL MSMO medium. After centrifugation (5 min., 500 g), 2-3 mL of the cell pellet was spread on MSMO plates supplemented with 50 μL/mL kanamycin, 500 μL/mL vancomycin and 500 μL/mL carbenicillin. The MSMO plates were incubated for 2 - 3 week at 25 °C. Using a scarpel, the grown callus was scraped off the plate, chopped and transfered into a 100 mL Erlenmayer flask containing 30 mL MSMO supplemented with 50 μL/mL kanamycin, 500 μL/mL vancomycin and 500 μL/mL carbenicillin. After one week of incubation at 25 °C in the shaking incubator, the transformed cell were diluted regularly once a week as described in Section 11.4.2.

11.4.5 Upscaling of transformed *Arabidopsis* PSB-D cells

The transformed *Arabidopsis* suspension cells (Section 11.4.4) were gradually upscaled to 5 L by the following dilutions: 50 mL of 7-days old cell culture was transfered into 500 mL Erlenmeyer flasks containing 180 mL MSMO medium with 25 μL kanamycin. After one week of incubation at 25 °C and 200 rpm., the cell culture was used to transfer 50 mL of the cell suspension into 5x 500 mL Erlenmeyer flasks containing 180 mL MSMO medium. After another week of incubation at 25 °C and 200 rpm., each of the five cultures was transfered into 2 L

Erlenmeyer flask containing 800 mL MSMO medium. After 3 days of incubation, the cells were harvested by filtering the cell suspension through two layers of Miracloth. Portions of 15 g cells were frozen in liquid nitrogen and stored at -80 °C.

11.4.5.1 Tobacco Infiltration

For transient expression of fluorescent proteins in *Nicotiana benthamiana*, 10 mL LB medium with appropriate antibiotics were inoculated with 2 - 3 *A. tumefaciens* colonies harbouring the plant expression plasmid and grown o/n at 28 °C and 200 rpm. The bacteria were pelleted by centrifugation for 10 min. at 4000 g and were re-suspended in 10 mL infiltration medium (10 mM MES-KOH pH 5.7, 10 mM $MgCl_2$, 0.1 mM acetosyringone dissolved in ethanol). Using a syringe, 1 - 2 mL of the bacteria suspension was infiltrated in the abaxial side of leaves from 2 - 4 weeks old *N. benthamiana* plants. The treated plants were kept for further 3 days in the greenhouse before analysis by confocal microscopy.

11.5 Microscopy

11.5.1 Confocal Laser Scanning Microscopy

Confocal laser scanning microscopy (CLSM) was performed using a Leica SP8 microscope, equipped with a 10X NA 0,3, 40X Oil NA 1,3 or 63X Glycerol NA 1.3 objective. DAPI was excited with a 405 nm laser, GFP was excited using an Argon laser at 488 nm and mCherry/PI was exited using an DPSS laser at 561 nm. The emission of GFP / mCherry was detected with Hybrid detectors at 500 - 550 nm or 570-620 nm, respectively. The emission of DAPI or PI was detected with PMTs at 410-495 nm or 570-620 nm, respectively. Roots or suspension cells were mounted in H_2O on objective slides with cover slips.

11.5.1.1 Fluorescence recovery after photobleaching (FRAP)

The FRAP experiments were performed using a SP8 (Leica,Wetzlar, Germany) confocal laser scanning microscope (CLSM) equipped with a 63X Glycerol NA 1.3 objective. Beforehand, the best parameters for ROI size, time of bleaching and bleaching intensity were empirically determined as conditions are highly instrument dependent (data not shown). Before and after the bleaching pulse of six iterations (6 x 79 ms) at 100 % laser power (488 nm), 50 pre-bleach and 90 post-bleach images were aquired with 2 % laser power. The imaging was performed with the following settings: 256 x 256 pixel, 1800 Hz, bi-directional scanning mode, no line averaging. The images were processed using the ImageJ software version 1.49m. With the easyFRAP software (Rapsomaniki et al. 2012), the raw fluorescence intensity measurements of the bleached area (circular 9 μm), the whole nucleus and the background (circular 9 μm) were processed for ′double′ or ′double full scale′ normalization. Normalized values were used to calculate the half life time and the mobile fraction. Significance was tested by Student′s T-Test.

11.5.1.2 Förster resonance energy transfer (FRET)

The measuring of the FRET acceptor photobleaching (FRET-APB) and the calculation of the FRET efficiencies was done as described in Weidtkamp-Peters et al. 2017.

Briefly, cut out a square (0.5 x 0.5 cm) of an infiltrated *Nicotiana benthamiana* leaf (Section 11.4.5.1). The piece of leaf was mounted in H_2O on a objective slide with the abaxial side facing up. The cover slip was fixed with tape. The leaf was flatten and became translucent, by applying moderate pressure with a scarpel. Images were acquired using a SP8 (Leica,Wetzlar, Germany) confocal laser scanning microscope (CLSM) equipped with a 63X Glycerol NA 1.3 objective. GFP was excited using an Argon laser at 488 nm and mCherry was excited / bleached using an DPSS laser at 561 nm. Images were taken with the following settings: 256 x 256 pixel, 1400 Hz, no line averaging, sequential scan mode, pinhole 3, 10x zoom, PMT detector gain of 800 V. For bleaching, a circular area of 9 μm was bleached at 100% laser power, for 60 iterations. Ten pre-bleach and ten post-bleach images were analysed by ImageJ software version 1.49m. If the fluorescent intensity dropped more than 5 % in the first ten pre-bleached images, the meassurement was discarded. The mean FRET efficiency was calculated by the following formular: ((I_{POST} - I_{PRE}) / I_{PRE}) x 100. With I_{PRE} / I_{PRE} = mean flurorescence intensity of 10 pre-bleached / post-bleached frames.

11.5.2 Cell stainings

11.5.2.1 DAPI

Before CLSM analysis, suspension cells were fixed for 2 min. in 1 % paraformaldehyde, quenched for 2 min. in 25 mM glycine and stained for 10 min. in DAPI solution (1x PBS pH 7.5, 0.2 μg/mL 4′,6-diamidino-2-phenylindole (DAPI), 0.1 % Triton X-100).

11.5.2.2 Propidiumiodid

Arabidopsis roots were mounted in 15 μM propidium iodide on an ojective slide with a cover slip and were incubated for 5 min. before CLSM analysis.

12 Supplements

	GO ID	Biological process	Query item / Query total	Bg item / Bg total	P-Value	FDR
Down-regulated	GO:0009813	flavonoid biosynthetic process	7 / 43	69 / 37767	3.00E-12	6.70E-10
	GO:0009812	flavonoid metabolic process	7 / 43	78 / 37767	6.80E-12	7.50E-10
	GO:0009699	phenylpropanoid biosynthetic process	7 / 43	141 / 37767	3.50E-10	2.60E-08
	GO:0009698	phenylpropanoid metabolic process	7 / 43	175 / 37767	1.50E-09	8.10E-08
	GO:0019748	secondary metabolic process	9 / 43	489 / 37767	4.20E-09	1.80E-07
	GO:0042398	cellular amino acid derivative biosynthetic process	7 / 43	233 / 37767	1.00E-08	3.50E-07
	GO:0019438	aromatic compound biosynthetic process	7 / 43	237 / 37767	1.10E-08	3.50E-07
	GO:0006575	cellular amino acid derivative metabolic process	7 / 43	315 / 37767	7.50E-08	2.10E-06
	GO:0046148	pigment biosynthetic process	5 / 43	112 / 37767	2.30E-07	5.60E-06
	GO:0006725	cellular aromatic compound metabolic process	7 / 43	399 / 37767	3.60E-07	7.90E-06
	GO:0042440	pigment metabolic process	5 / 43	134 / 37767	5.40E-07	1.10E-05
	GO:0009753	response to jasmonic acid stimulus	5 / 43	215 / 37767	5.10E-06	9.40E-05
	GO:0006519	cellular amino acid and derivative metabolic process	7 / 43	682 / 37767	1.20E-05	0.0002
	GO:0050896	response to stimulus	13 / 43	4057 / 37767	0.00041	0.0065
	GO:0009416	response to light stimulus	5 / 43	596 / 37767	0.00058	0.0085
	GO:0009314	response to radiation	5 / 43	613 / 37767	0.00066	0.0091
	GO:0044249	cellular biosynthetic process	14 / 43	4925 / 37767	0.00079	0.01
	GO:0006950	response to stress	9 / 43	2320 / 37767	0.001	0.013
	GO:0003700	transcription factor activity	9 / 43	2173 / 37767	0.00066	0.01
Up-regulated	GO:0006950	response to stress	29 / 103	2320 / 37767	1.00E-12	2.00E-10
	GO:0050896	response to stimulus	38 / 103	4057 / 37767	7.10E-13	2.00E-10
	GO:0042221	response to chemical stimulus	24 / 103	2085 / 37767	7.30E-10	9.40E-08
	GO:0009743	response to carbohydrate stimulus	10 / 103	240 / 37767	1.20E-09	1.20E-07
	GO:0009266	response to temperature stimulus	12 / 103	485 / 37767	7.50E-09	5.80E-07
	GO:0009409	response to cold	10 / 103	328 / 37767	2.20E-08	1.40E-06
	GO:0010033	response to organic substance	17 / 103	1342 / 37767	8.40E-08	4.60E-06
	GO:0009628	response to abiotic stimulus	16 / 103	1471 / 37767	1.50E-06	7.40E-05
	GO:0007623	circadian rhythm	5 / 103	75 / 37767	2.30E-06	8.30E-05
	GO:0048511	rhythmic process	5 / 103	75 / 37767	2.30E-06	8.30E-05
	GO:0009605	response to external stimulus	9 / 103	429 / 37767	2.40E-06	8.30E-05
	GO:0006979	response to oxidative stress	8 / 103	332 / 37767	3.30E-06	0.00011
	GO:0010200	response to chitin	6 / 103	151 / 37767	3.80E-06	0.00011
	GO:0006955	immune response	8 / 103	367 / 37767	6.70E-06	0.00018
	GO:0002376	immune system process	8 / 103	368 / 37767	6.80E-06	0.00018
	GO:0009611	response to wounding	6 / 103	197 / 37767	1.60E-05	0.00039
	GO:0006952	defense response	10 / 103	766 / 37767	3.70E-05	0.00084
	GO:0045087	innate immune response	7 / 103	347 / 37767	4.20E-05	0.0009
	GO:0009719	response to endogenous stimulus	11 / 103	1068 / 37767	0.00012	0.0025
	GO:0048583	regulation of response to stimulus	5 / 103	188 / 37767	0.00016	0.0031
	GO:0006970	response to osmotic stress	6 / 103	408 / 37767	0.00078	0.014
	GO:0016020	membrane	22 / 103	4068 / 37767	0.00076	0.017
	GO:0044464	cell part	56 / 103	15217 / 37767	0.00079	0.017
	GO:0005623	cell	56 / 103	15217 / 37767	0.00079	0.017

Figure S1 Flavonoid biosynthesis pathway genes are strongly downregulated in *ssrp1-2* and *spt16-1*.
Gene ontology analysis (AgriGO) was performed on differentially ($\geq$ 2-fold with $p<0.05$) expressed genes in at least one of the two FACT mutants. Overrepresented GO biological processes are shown. GO terms ascribed to the flavonoid biosynthetic pathway are depicted in purple. The GO terms âĂIJtranscription factor activityâĂİ and âĂIJCircadian rhythm are depicted in yellow (discussed later in xxx). Fisher's test: P-value<0.05, FDR<0.05.

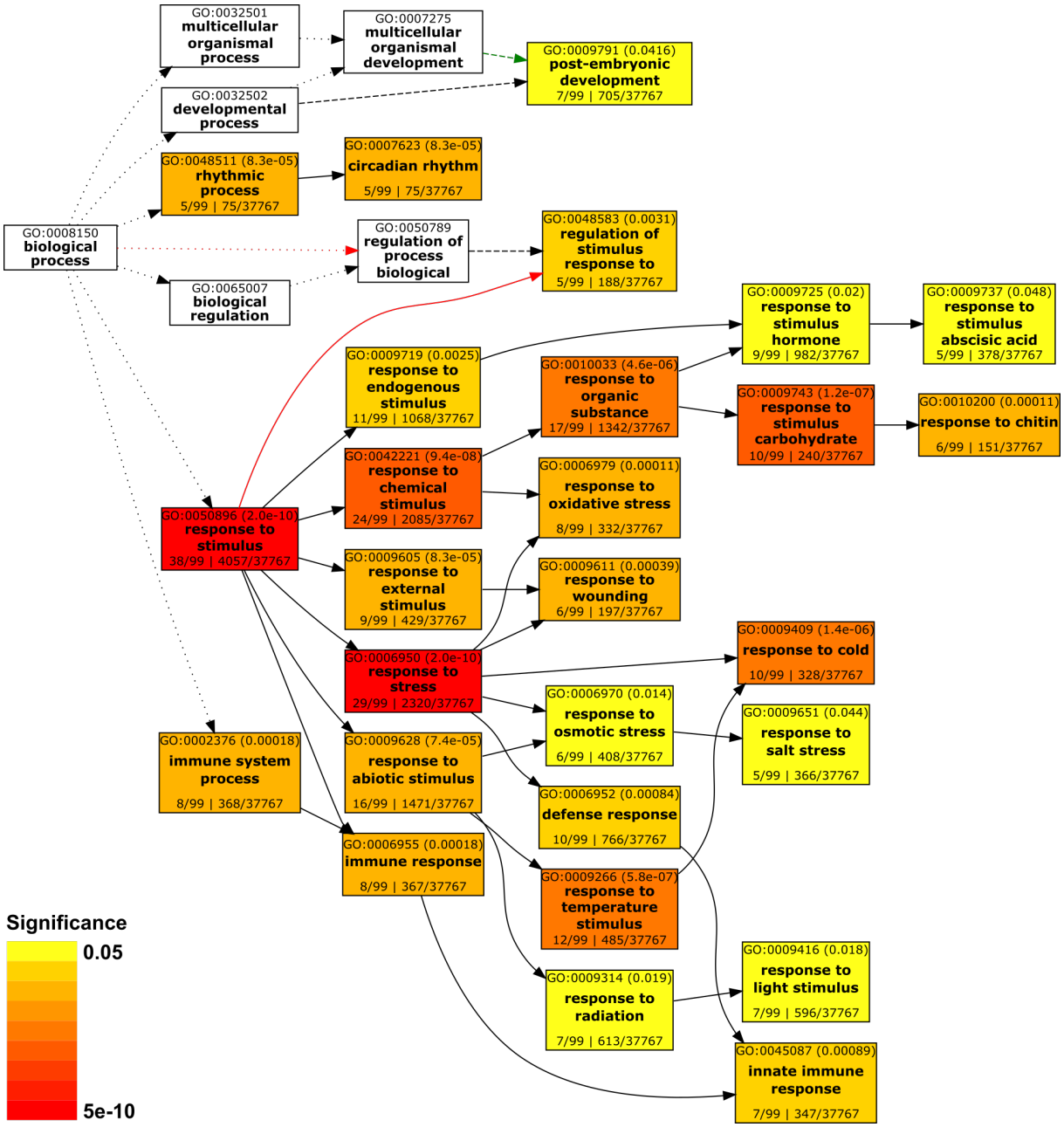


Figure S2 Biological processes identified by an overrepresented number of genes up-regulated in *ssrp1-2* and/or *spt16-1*.

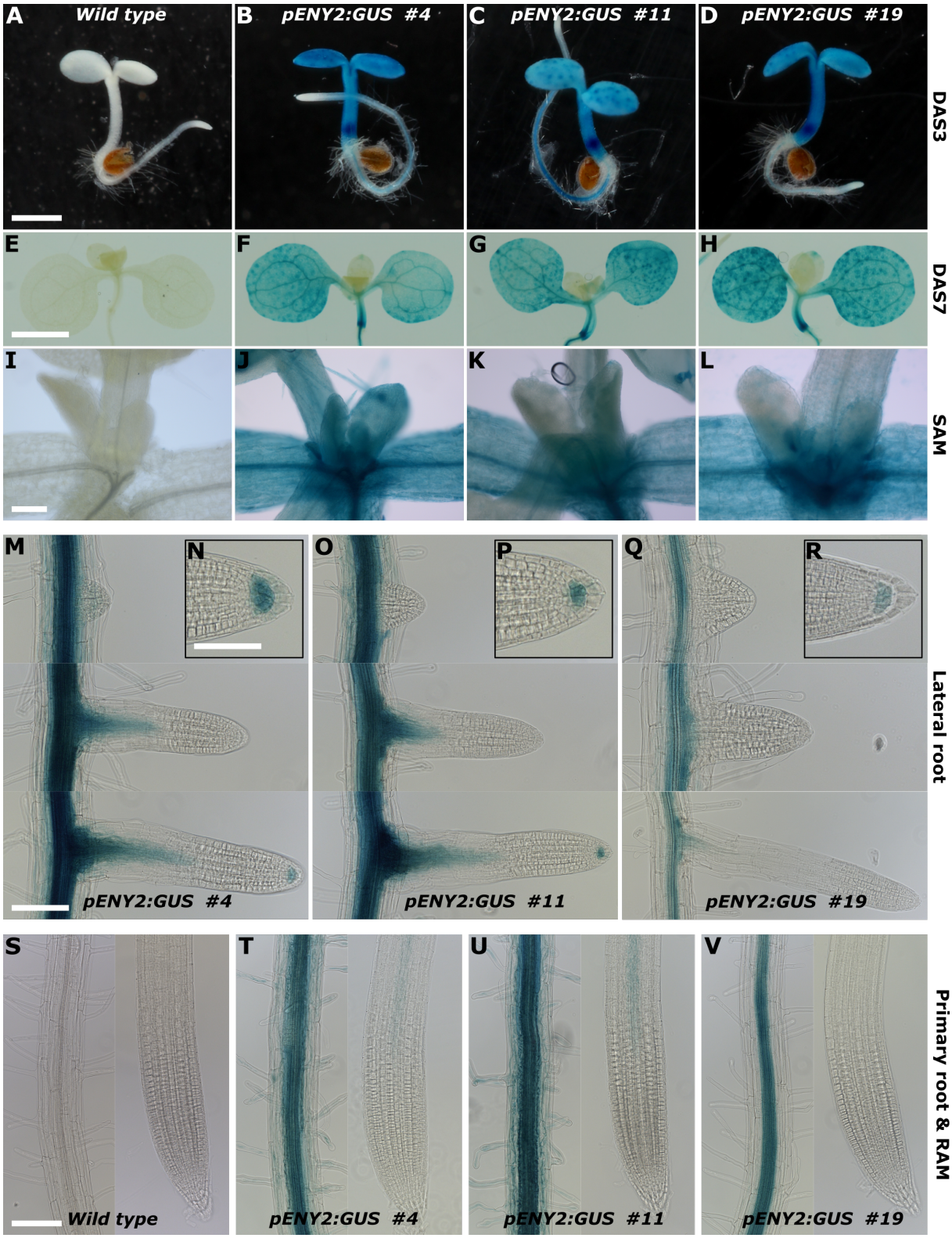


Figure S3 The *ENY2* Promoter (GUS staining) is highly active in young *Arabidopsis* seedlings, except meristematic regions.
The *ENY2* Promoter-GUS activity was monitored in three independent transgenic plant lines. *Wild type Arabidopsis* plants were used as negative control. A-D) Overview of young seedlings (DAS3). Bars indicate 1 mm. E-H) Aerial parts of young seedlings (DAS7). Bars indicate 1 mm. I-L) Close up of shoot apical meristem and young leaf primordia (DAS10). Bars indicate 100 μm. M, O, Q) Lateral root formation of transgenic lines #4, #11, #19 (from left to right) (DAS7). Three developmental stages are shown (Top down). Bars indicate 100 μm. N, P, R) Close up of lateral root tip. Bars indicate 100 μm. Bars indicate 50 μm. S-V) Differentiated and meristematic region of primary root. Bars indicate 100 μm.

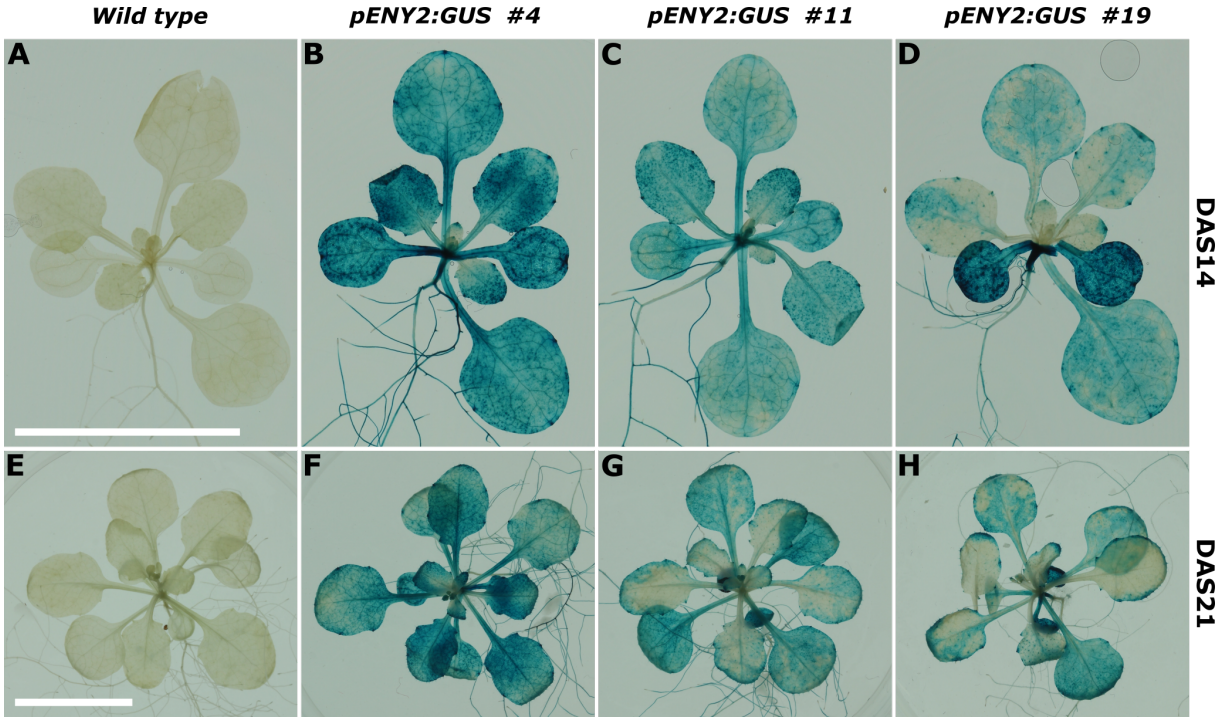


Figure S4 The *ENY2* Promoter (GUS staining) is highly active in *Arabidopsis* plantlets, except meristematic regions.
The *ENY2* Promoter-GUS activity was monitored in whole plants of three independent transgenic lines. *Wild type Arabidopsis* was used as negative control. A-D) Young plantlets at DAS14. E-H) Plantlets at DAS21. Bars indicate 1 cm.

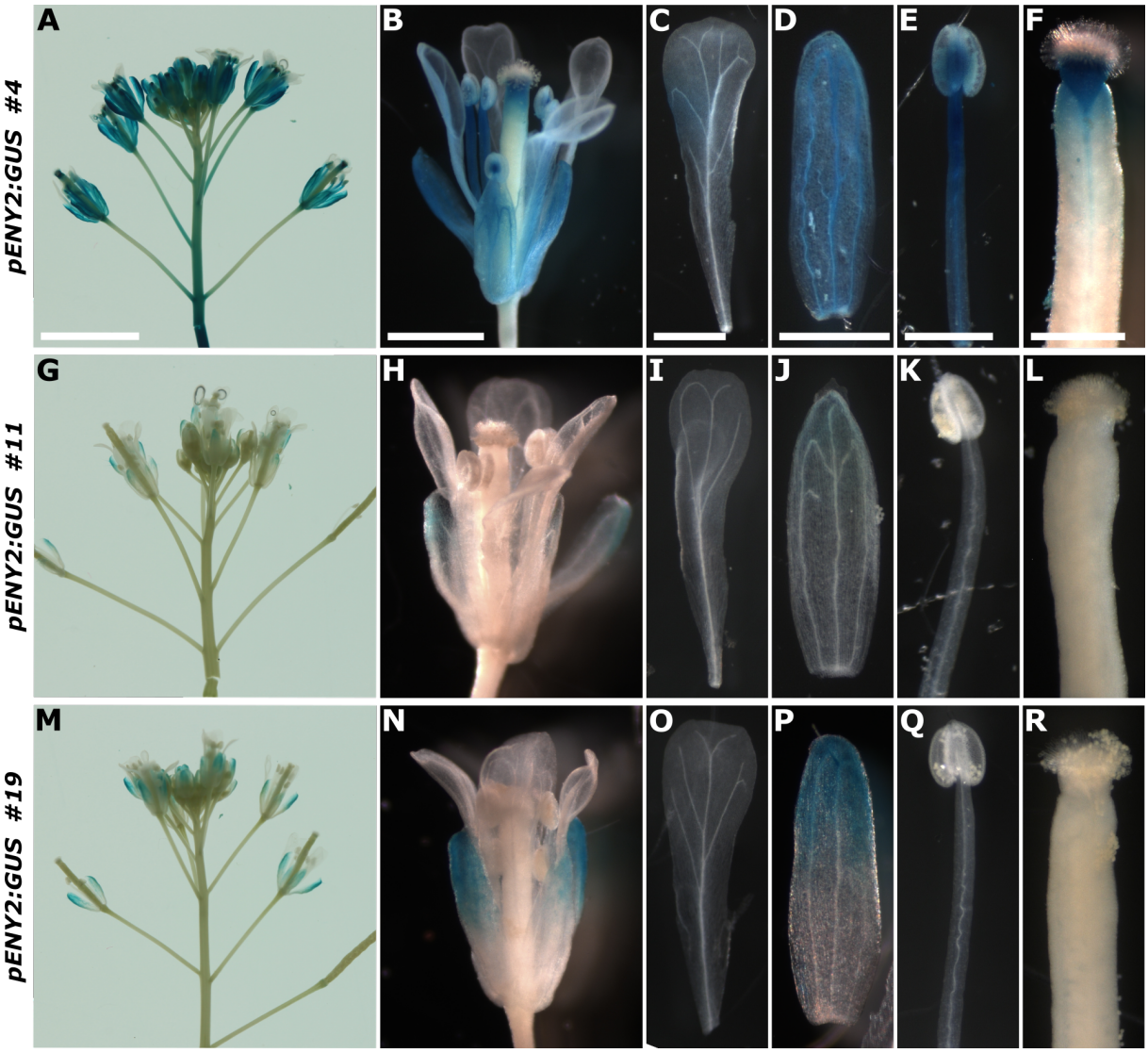


Figure S5 The *ENY2* promoter activity (GUS staining) in reproductive organs.
ENY2 Promoter-GUS activity was monitored in flowers of three independent transgenic lines. A-F) Transgenic line #4 G-L) Transgenic line #11 M-R) Transgenic line #19 A,G,M) Flower buds. Bar indicates 5 mm. G,H,N) Flowers. Bar indicates 1 mm. C,I,O) Petals. Bar indicates 0.5 mm. D,J,P) Sepals. Bar indicates 0.5 mm. E,K,Q) Stamen (Anther and filament). Bar indicates 0.5 mm. F,L,R) Stigma and style of carpel. Bar indicates 0.5 mm.

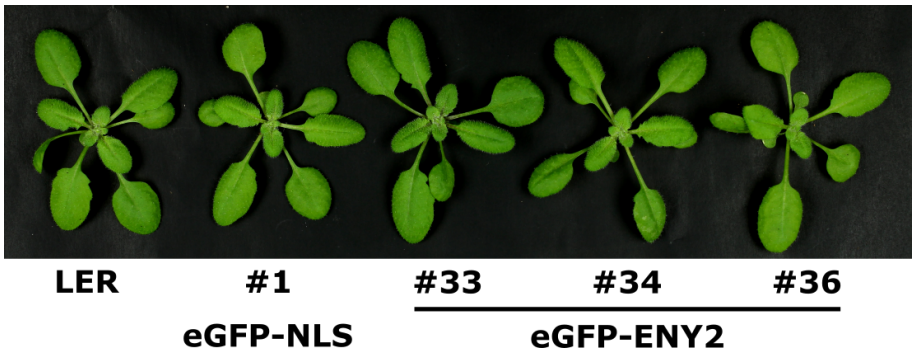


Figure S6 Transgenic plants expressing eGFP-NLS or eGFP-ENY2 show no obvious phenotype.
Plants (DAS21) were grown under normal LD conditions

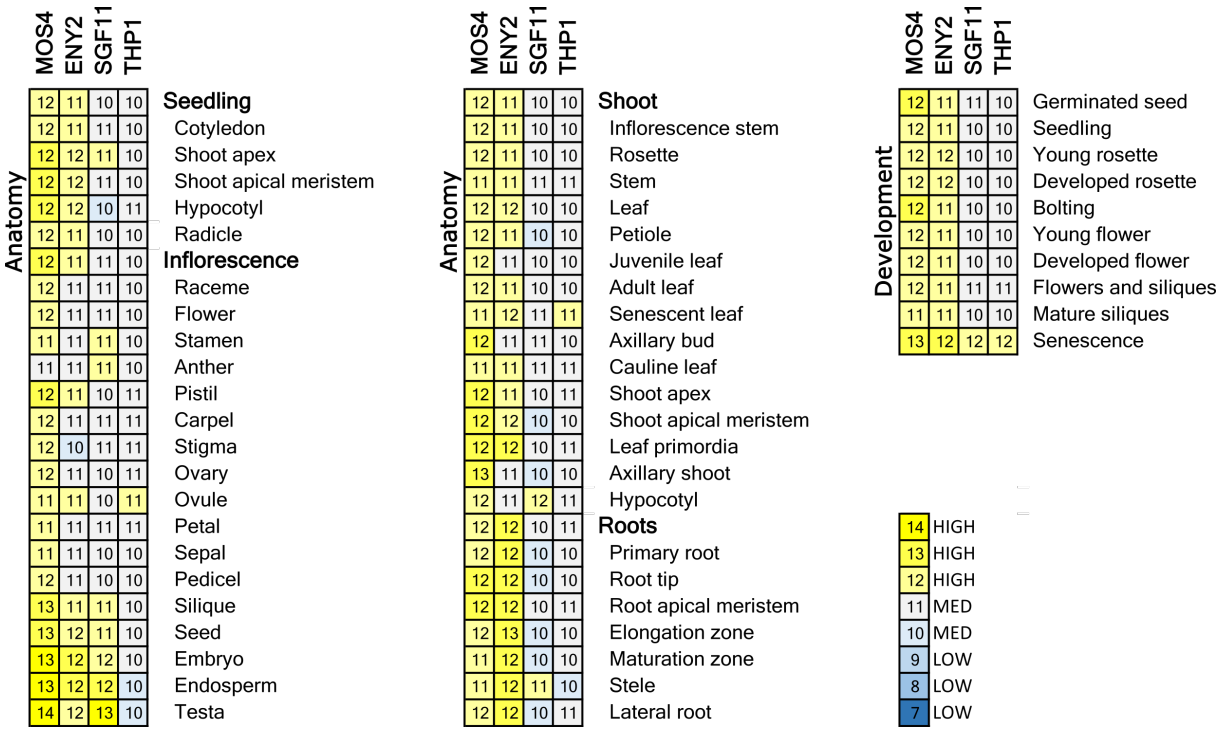


Figure S7 Co-Expression of ENY2, SGF11, MOS4 and THP1 in different *Arabidopsis* tissues throughout development was visualized by analysing public microarray data using Gene investigator.

ENY2

ID	Description	Frequency [%]	-log$_{10}$(*P*)
GO:0006396	RNA processing	3,21	15,03
GO:0009791	post-embryonic development	0,16	1,54
GO:0006950	response to stress	4,58	0,96
GO:0009987	cellular process	63,78	0,96
GO:0050896	response to stimulus	12,21	0,96
GO:0008152	metabolic process	75,39	0,82
GO:0000003	reproduction	0,77	0,68
GO:0040007	growth	0,32	0,68
GO:0051704	multi-organism process	0,75	0,59
GO:0032501	multicellular organismal process	2,37	0,32

MOS4

ID	Description	Frequency [%]	-log$_{10}$(*P*)
GO:0006396	RNA processing	3,21	19,82
GO:0042254	ribosome biogenesis	1,42	6,96
GO:0009987	cellular process	63,78	4,28
GO:0008152	metabolic process	75,39	4,24
GO:0009791	post-embryonic development	0,16	3,68
GO:0000003	reproduction	0,77	2,46
GO:0032501	multicellular organismal process	2,37	0,89
GO:0009058	biosynthetic process	31,61	0,89
GO:0009409	response to cold	0,04	0,68
GO:0032502	developmental process	2,81	0,55

SGF11

ID	Description	Frequency [%]	-log$_{10}$(*P*)
GO:0009987	cellular process	63,78	20,28
GO:0009791	post-embryonic development	0,16	19,23
GO:0006396	RNA processing	3,21	13,37
GO:0000003	reproduction	0,77	10,64
GO:0032501	multicellular organismal process	2,37	10,64
GO:0032502	developmental process	2,81	10,28
GO:0008152	metabolic process	75,39	4,31
GO:0009615	response to virus	0,12	3,92
GO:0022403	cell cycle phase	0,00	1,33
GO:0000279	M phase	0,00	1,21

THP1

ID	Description	Frequency [%]	-log$_{10}$(*P*)
GO:0006396	RNA processing	3,21	9,64
GO:0006913	Nucleocytoplasmic transport	0,24	8,60
GO:0009791	Post-embryonic development	0,16	5,72
GO:0009615	Response to virus	0,12	4,03
GO:0009987	Cellular process	63,78	2,85
GO:0032502	Developmental process	2,81	2,70
GO:0032501	Multicellular organismal process	2,37	2,62
GO:0040007	Growth	0,32	2,42
GO:0000003	Reproduction	0,77	2,34
GO:0051179	Localization	18,50	0,68

Figure S8 Proteins co-purified with SG-fusion proteins were analysed by Gene Ontology analysis (AgriGO) to identify overrepresented biological processes.
Redundant proteins were removed by REViGO. Frequency indicate the percentage of this GO term in the whole UniProt database. P = P-Values.

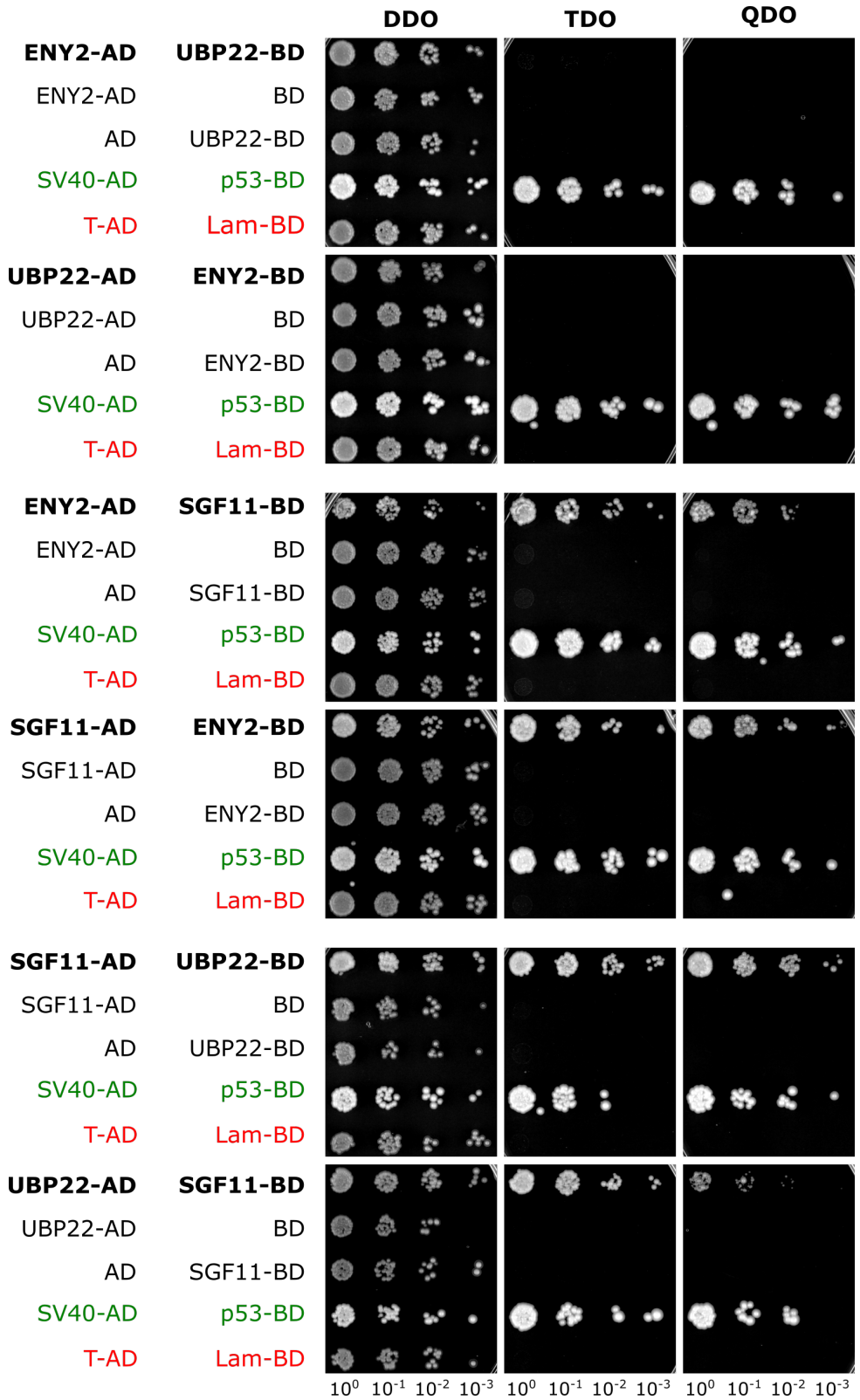


Figure S9 Y2H Assay to test interactions between ENY2, SGF11 and UBP22 with respective controls.
In addition to test combinations of DNA-BD/bait and AD/prey fusion proteins, also corresponding controls to test for autoactivation or toxicity were spotted in serial dilutions on DDO, TDO, QDO plates: DNA-BD/bait proteins with AD alone and vice versa DNA-BD alone with AD/prey proteins. DNA-BD/murine p53 and AD/SV40 large T-antigen function as positive control. DNA-BD/Lamin and AD/T act as negative control.

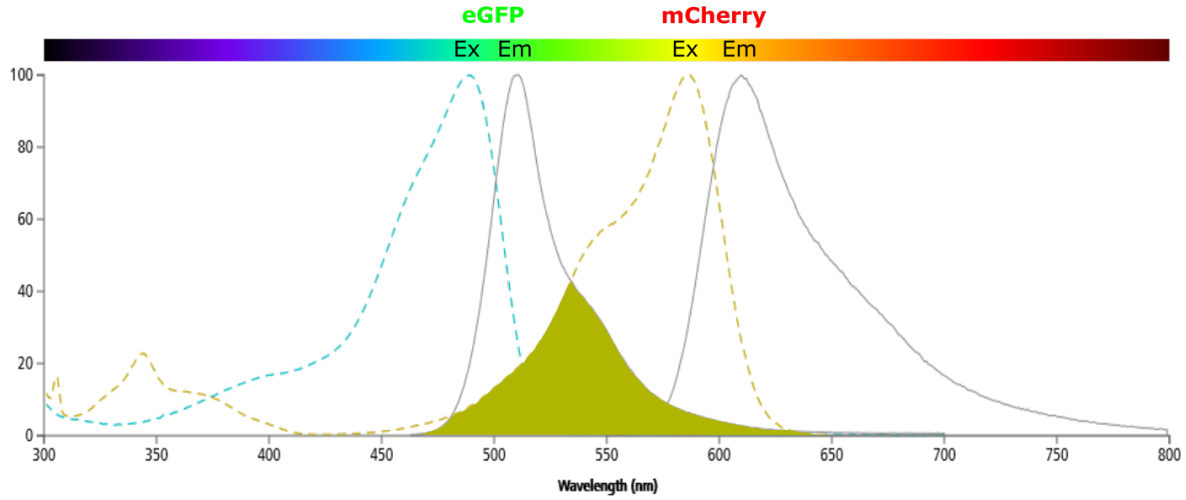


Figure S10 Plant FRET vector system.
Expression of fluorescent fusion proteins is driven by CaMV 35S promoter A) Donor construct with eGFP as translational fusion at N-terminus of protein of interest. B) Acceptor construct with mCherry as translational fusion at N-terminus of protein of interest. Negative Controls are expressing free eGFP-NLS (C) and mCherry-NLS (D). Positive control is expressing eGFP-NLS-mCherry fusion protein (E). All constructs are basing on pCambia2300 (A,C,D) or pGreen0179 (B,D) backbones.

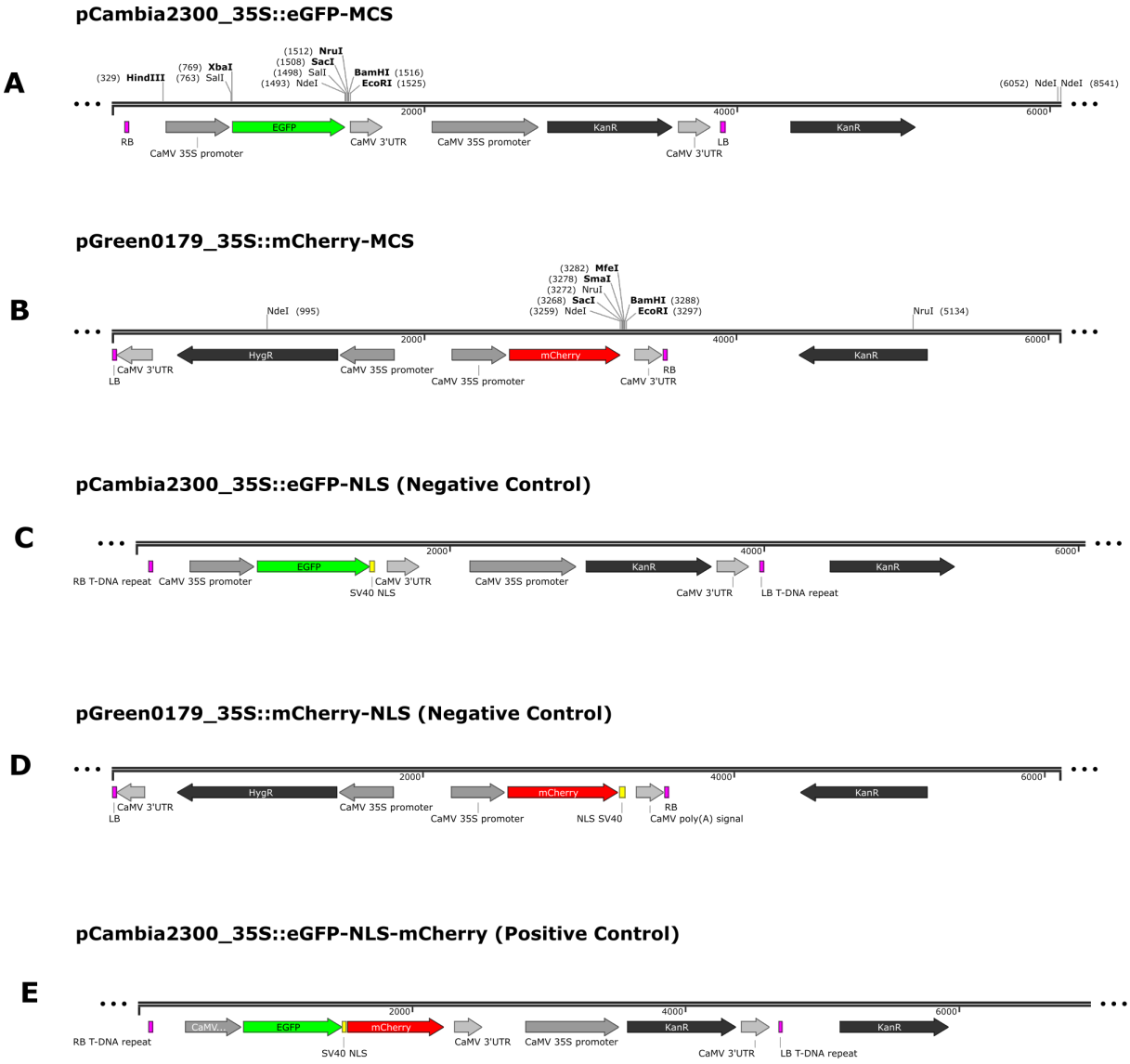


Figure S11 Spectral overlap of FRET pair eGFP/mCherry.
Absorption (Dotted line) and emission spectra (Normal line) of both fluorophores are generated by by Fluorescence Spectrum Viewer (BioScience). Exitation (Ex) and emission (Em) maxima are indicated. Good spectral overlap of eGFP emission and mCherry absorption (Fully coloured) is important for FRET measurements.

Donor / Acceptor	Mean ± SD
eGFP-NLS-mCherry Fusion	26,85 ± 2,71
eGFP-NLS / mCherry-NLS	1,63 ± 2,81
eGFP-MOS4 / mCherry-ENY2	0,23 ± 1,4
eGFP-MOS4	-0,21 ± 1,46
eGFP-CDC5 / mCherry-ENY2	-0,46 ± 2,23
eGFP-CDC5	-0,4 ± 0,89
eGFP-MAC3A / mCherry-ENY2	-0,56 ± 1,7
eGFP-MAC3A	1,76 ± 0,99
eGFP-MAC3B / mCherry-ENY2	-0,05 ± 1,54
eGFP-MAC3B	0,55 ± 1,81
eGFP-ISY1 / mCherry-ENY2	-0,33 ± 1,46
eGFP-ISY1	-0,22 ± 1,43
eGFP-SKIP / mCherry-ENY2	-0,42 ± 1,62
eGFP-SKIP	0,76 ± 1,56
eGFP-CRN1 / mCherry-ENY2	-0,28 ± 2,42
eGFP-CRN1	-2,33 ± 2,97
eGFP-PRL1 / mCherry-ENY2	1,3 ± 1,61
eGFP-PRL1	1,61 ± 3,4
mCherry-SYF1 / eGFP-ENY2	-0,24 ± 2,09
mCherry-MAC7 / eGFP-ENY2	-1,64 ± 1,96
eGFP-ENY2	-1,45 ± 2,13

Figure S12 FRET reveal no direct interaction between ENY2 and components of the NTC/NTR complex. Mean FRET efficiencies ± SD of all test combinations and the respective controls are shown.

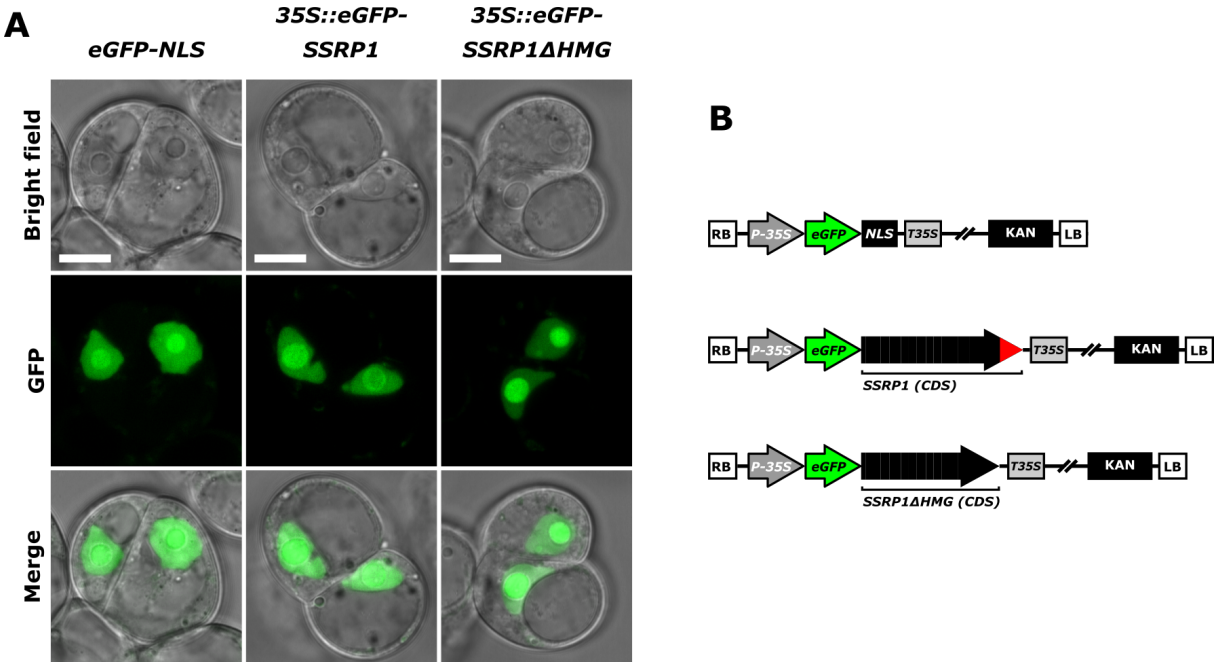


Figure S13 Full length and truncated SSRP1 show the same nuclear localization in living cells. A) Transgenic *Arabidopsis* cell lines expressing either eGFP-SSRP1, eGFP-SSRP1ΔHMG or eGFP-NLS under CaMV35S promoter were subjected to confocal microscopy. Bar indicates 10 μm. B) Schematic illustration of eGFP-SSRP1, eGFP-SSRP1ΔHMG and eGFP-NLS transgenes. Coding sequences of full length or truncated SSRP1 as well as free eGFP-NLS are expressed as N-terminal translational fusions with eGFP under CaMV 35S promoter. HMG box domain is depicted in red (Black bars = exons, dark grey bars = Promoter, dotted lines = introns, green bar = eGFP coding sequence, light grey bars = UTR, nptII = Kan resistance marker, RB/LB = Right border/Left Border).

12.1 Phenotypic data

Table S14 HMG-deficiency mutants show no obvious phenotype.
Phenotypical analysis of *ssrp1-1* complementation lines in comparison to *wild type* Col-0 plants grown under LD conditions. Representative individuals are shown at various developmental stages A) DAS21 (days after stratification) B) DAS28 C) DAS42 D-G) The following plant feature were statistically evaluated: (D) Time of bolting (Elongation of the first internode), (E) number of leaves at bolting, (F) Rosette diameter at DAS35 and (G) primary inflorescences at DAS42 (All data are means ± SD, * indicates P-Value<0.05, ** indicates P-Value<0.01, *** indicates P-Value<0.001, Student's T-Test).

Genotype	Bolting [days]	Rosette ∅ at bolting [mm]	No. of leaves at bolting	Rosette ∅ at DAS35 [mm]	Plant Height at DAS42 [cm]	Pr. Inflores. at DAS42
LER	25.71 ± 0.91 (n= 14)	69.36 ± 5.47 (n= 14)	7.36 ± 0.74 (n= 14)	88.5 ± 6.85 (n= 14)	20.07 ± 1.25 (n= 14)	3.43 ± 0.51 (n= 14)
Full SRRP1 #1.4	27.08 ± 1.04 ** (n= 13) (p= 2.4E-03)	60.15 ± 9.88 ** (n= 13) (p= 3.5E-03)	6.92 ± 0.76 (n= 13) (p= 0.14)	79.69 ± 11.39 * (n= 13) (p= 2.3E-02)	23.31 ± 1.28 *** (n= 13) (p= 1.0E-06)	3.23 ± 0.83 (n= 13) (p= 0.57)
Full SRRP1 #4.6	25.87 ± 1.19 (n= 15) (p= 0.60)	64.73 ± 7.43 (n= 15) (p= 0.08)	7.07 ± 0.59 (n= 15) (p= 0.28)	87.47 ± 9.3 (n= 15) (p= 0.87)	23.77 ± 1.15 *** (n= 15) (p= 2.0E-08)	3.4 ± 0.63 (n= 15) (p= 1)
Full SRRP1 #6.12	26 ± 1.25 (n= 15) (p= 0.72)	62.93 ± 7.56 * (n= 15) (p= 0.02)	7.53 ± 0.64 (n= 15) (p= 0.59)	82.6 ± 9.31 (n= 15) (p= 0.05)	22 ± 1.69 *** (n= 15) (p= 4.0E-04)	3.2 ± 0.86 (n= 15) (p= 0.29)
SSRP1△HMG #8.20	26.6 ± 0.99 * (n= 15) (p= 0.03)	73.53 ± 6.14 * (n= 15) (p= 0.03)	8.07 ± 0.59 * (n= 15) (p= 0.01)	97.07 ± 8.92 ** (n= 15) (p= 1.6E-03)	21.9 ± 1.62 ** (n= 15) (p= 1.2E-03)	4.27 ± 0.59 *** (n= 15) (p= 4.5E-04)
SSRP1△HMG #10.97	24.67 ± 1.05 * (n= 15) (p= 0.01)	63.13 ± 5.33 ** (n= 15) (p= 8.0E-03)	7.4 ± 0.63 (n= 15) (p= 0.78)	87.87 ± 8.59 (n= 15) (p= 0.77)	23.43 ± 1.33 *** (n= 15) (p= 4.0E-07)	3.73 ± 0.8 (n= 15) (p= 0.17)
SSRP1△HMG #11.29	24.21 ± 1.31 ** (n= 15) (p= 1.5E-03)	69.71 ± 6.67 (n= 15) (p= 0.87)	7.5 ± 0.52 (n= 15) (p= 0.47)	85.36 ± 10.75 (n= 15) (p= 0.27)	23 ± 1.32 *** (n= 15) (p= 6.0E-06)	3.79 ± 0.8 (n= 15) (p= 0.29)

Table S15 Phenotypic evaluation of SSRP1 and SSRP1△HMG overexpression lines compared to *wild type* Ler-0 plants.
All data are means ± SD. Statistical analyses were performed with Student's T-Test (* indicates P<0.05, ** indicates P<0.01, *** indicates P<0.001).

Genotype	Bolting [days]	Rosette ∅ at bolting [mm]	No. of leaves at bolting	Rosette ∅ at DAS35 [mm]	Plant Height at DAS42 [cm]	Pr. Inflores. at DAS42
LER	19 ± 0.74 (n= 12)	64 ± 9.41 (n= 12)	8.55 ± 0.5 (n= 12)	78.64 ± 8.19 (n= 12)	71.82 ± 11.89 (n= 12)	4.82 ± 0.39 (n= 12)
SSRP1 OE#2	19.31 ± 0.61 (n= 13) (p= 0.207)	63.69 ± 6.04 (n= 13) (p= 0.96)	9 ± 0.68 (n= 13) (p= 0.052)	77.54 ± 6.16 (n= 13) (p= 0.747)	67.46 ± 13.8 (n= 13) (p= 0.429)	4.85 ± 0.53 (n= 13) (p= 0.66)
SSRP1 OE#3	19.27 ± 0.57 (n= 15) (p= 0.254)	59.33 ± 7.29 (n= 15) (p= 0.16)	9 ± 0.52 (n= 15) (p= 0.124)	73.33 ± 8.24 (n= 15) (p= 0.089)	71.4 ± 16.05 (n= 15) (p= 0.454)	5.13 ± 0.5 (n= 15) (p= 0.195)
SSRP1 OE#5	19.4 ± 0.71 (n= 15) (p= 0.381)	68 ± 6.41 (n= 15) (p= 0.4)	9.7 ± 0.47 *** (n= 15) (p= 1.78E-05)	82.07 ± 6.23 (n= 15) (p= 0.268)	70.53 ± 11.99 (n= 15) (p= 0.761)	5.07 ± 1 (n= 15) (p= 0.462)
SSRP1△HMG OE#11	18.79 ± 0.67 (n= 14) (p= 0.776)	52.93 ± 4.04 ** (n= 14) (p= 0.005)	7.93 ± 0.7 * (n= 14) (p= 0.027)	67.64 ± 7.74 ** (n= 14) (p= 0.003)	71.5 ± 8.18 (n= 14) (p= 0.601)	4.29 ± 0.45 ** (n= 14) (p= 0.007)
SSRP1△HMG OE#13	19.21 ± 0.56 (n= 14) (p= 0.546)	54.79 ± 4.33 ** (n= 14) (p= 0.008)	7.93 ± 0.7 * (n= 14) (p= 0.027)	65.6 ± 6.07 *** (n= 14) (p= 2E-04)	62.86 ± 9.67 (n= 14) (p= 0.057)	4.29 ± 0.45 ** (n= 14) (p= 0.007)
SSRP1△HMG OE#14	20.6 ± 0.74 *** (n= 13) (p= 4.44E-05)	56.62 ± 9.68 (n= 13) (p= 0.07)	9.08 ± 0.73 (n= 13) (p= 0.052)	66.54 ± 11.59 * (n= 13) (p= 0.011)	51.15 ± 13.75 ** (n= 13) (p= 0.0012)	4.31 ± 0.46 * (n= 13) (p= 0.01)

Table S16 Phenotypic evaluation of mutant *eny2-1* plants compared to wild type Col-0 plants.
All data are means ± SD. Statistical analyses were performed with Student's T-Test (* indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$).

Genotype	Time of Bolting [days]	Rosette ∅ at bolting [mm]	No. of leaves at bolting	Rosette ∅ at DAS35 [mm]	Plant Height at DAS42 [cm]	Pr. Inflores. DAS42
Col-0	26.79 ± 1.92 (n= 29)	80.03 ± 6.83 (n= 29)	13.66 ± 0.8 (n= 29)	84.67 ± 4.8 (n= 15)	35.34 ± 2.34 (n= 29)	5.52 ± 0.72 (n= 29)
eny2-1	27.39 ± 2.02 (n= 30) (p= 0.26)	86.46 ± 9.83 ** (n= 30) (p= 0.01)	13.96 ± 1.24 (n= 30) (p= 0.27)	81.79 ± 8.35 *** (n= 14) (p= 1.8E-27)	33.11 ± 3.85 * (n= 30) (p= 0,01)	5.86 ± 0.52 (n= 30) (p= 0,05)

Table S17 Phenotypic evaluation of mutant *eny2-RNAi* plant lines compared to wild type Col-0 plants.
All data are means ± SD. Statistical analyses were performed with Student's T-Test (* indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$).

Genotype	Time of Bolting [days]	Rosette ∅ at bolting [mm]	No. of leaves at bolting	Rosette ∅ at DAS35 [mm]	Plant Height at DAS42 [cm]	Pr. Inflores. at DAS42
Col-0	26.4 ± 0.99 (n= 15)	89.73 ± 2.94 (n= 15)	13.33 ± 0.72 (n= 15)	11.77 ± 0.62 (n= 15)	31.7 ± 1.05 (n= 15)	5.4 ± 0.63 (n= 15)
ENY2 RNAi 3	26.67 ± 0.98 (n= 15) (p= 0.46)	85.93 ± 7.61 (n= 15) (p= 0.08)	11.8 ± 1.15 *** (n= 15) (p= 1.5E-04)	11.53 ± 0.92 (n= 15) (p= 0.42)	31.33 ± 1.45 (n= 15) (p= 0.43)	4.8 ± 0.68 * (n= 15) (p= 0.02)
ENY2 RNAi 22	28.67 ± 0.72 *** (n= 15) (p= 0)	91 ± 7.29 (n= 15) (p= 0.54)	14.8 ± 1.01 *** (n= 15) (p= 9.2E-05)	11.8 ± 0.75 (n= 15) (p= 0.9)	29.87 ± 2 ** (n= 15) (p= 3.9E-03)	5.13 ± 0.64 (n= 15) (p= 0.26)
ENY2 RNAi 27	27.13 ± 1.13 (n= 15) (p= 0.07)	91.13 ± 4.96 (n= 15) (p= 0.35)	13.07 ± 1.1 (n= 15) (p= 0.44)	11.87 ± 0.72 (n= 15) (p= 0.69)	31.9 ± 1.56 (n= 15) (p= 0.68)	5.47 ± 0.64 (n= 15) (p= 0.78)

Table S18 Phenotypic evaluation of ENY2 CRISPR knockout lines compared to *wild type* Col-0 plants.
All data are means ± SD. Statistical analyses were performed with Student's T-Test (*** indicates $p < 0.001$).

Genotype	Bolting [days]	Rosette ∅ at bolting [mm]	No. of leaves at bolting	Rosette ∅ at DAS35 [cm]	Plant Height at DAS42 [cm]	Pr. Infloresc. at DAS42
Col-0	25.33 ± 1.3 (n= 15)	85.07 ± 7.52 (n= 15)	13.2 ± 1.28 (n= 15)	13.93 ± 1.07 (n= 15)	35.73 ± 1.84 (n= 15)	6 ± 0.65 (n= 15)
ENY2_crispr	28.27 ± 2.02 *** (n= 15) (p= 8.8E-05)	88.27 ± 7.33 (n= 15) (p= 0.26)	15.27 ± 1.39 *** (n= 15) (p= 3.2E-04)	13.5 ± 0.63 (n= 15) (p= 0.21)	34.2 ± 3.15 (n= 15) (p= 0.13)	6.13 ± 0.72 (n= 15) (p= 0.80)

Table S19 Phenotypic evaluation of ENY2 overexpression lines compared to wild type Col-0 plants. All data are means ± SD. Statistical analyses were performed with Student's T-Test (* indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$).

Genotype	Bolting [days]	Rosette ∅ at at bolting [mm]	No. of leaves at bolting	Rosette ∅ at DAS35 [mm]	Plant Height at DAS42 [cm]	Pr. Inflores. at DAS42
Col-0	26.66 ± 1.68 (n= 44)	83.34 ± 7.48 (n= 44)	13.55 ± 0.79 (n= 44)	12.32 ± 0.93 (n= 30)	34.1 ± 2.66 (n= 44)	5.48 ± 0.7 (n= 44)
ENY2 OEx 1	26.49 ± 1.42 (n= 44) (p = 5.3E-01)	85.6 ± 9.05 (n= 44) (p = 1.8E-01)	12.84 ± 1.23 ** (n= 44) (p = 1.8E-03)	12.78 ± 1.31 (n= 30) (p = 1.2E-01)	32.63 ± 2.52 ** (n= 44) (p = 6.4E-03)	5.52 ± 0.7 (n= 44) (p = 7.0E-01)
ENY2 OEx 8	27.49 ± 1.59 * (n= 45) (p = 1.5E-02)	83.2 ± 13.59 (n= 45) (p = 7.6E-01)	14.2 ± 1.31 ** (n= 45) (p = 4.0E-03)	12.59 ± 0.93 (n= 30) (p = 2.6E-01)	30.8 ± 3.03 *** (n= 45) (p = 6.8E-07)	5.69 ± 0.63 (n= 45) (p = 1.5E-01)
ENY2 OEx 14	26.17 ± 1.51 (n= 42) (p = 1.2E-01)	78.69 ± 9.3 * (n= 42) (p = 3.0E-02)	12.48 ± 1.13 *** (n= 42) (p= 4.8E-06)	11.98 ± 1.19 (n= 29) (p = 1.7E-01)	33.86 ± 2.5 (n= 42) (p = 4.0E-01)	5.24 ± 0.73 (n= 42) (p = 7.8E-02)

Table S20 Phenotypic evaluation of mutant *sgf11-1* plants compared to *wild type* Col-0 plants.
All data are means ± SD. Statistical analyses were performed with Student's T-Test (** indicates p < 0.01, *** indicates p < 0.001).

Genotype	Bolting [days]	Rosette ∅ at bolting [mm]	No. of leaves at bolting	Rosette ∅ at DAS35 [mm]	Plant Height at DAS42 [cm]	Pr. Inflores. at DAS42
Col-0	27,38 ± 1,52 (n= 29)	82,66 ± 8,49 (n= 29)	13,55 ± 0.69 (n= 29)	11,77 ± 0.62 (n= 15)	33.12 ± 2.5 (n= 29)	5.21 ± 0.68 (n= 29)
sgf11-1	29,6 ± 2,24 *** (n= 30) (p = 5.0E-05)	83,07 ± 9.33 (n= 30) (p = 9.8E-01)	14.23 ± 1.94 (n= 30) (p = 8.1E-02)	11.9 ± 0.85 (n= 15) (p = 6.2E-01)	30.25 ± 3.09 *** (n= 30) (p= 1.8E-04)	5.2 ± 0.71 (n= 30) (p= 1)

12.2 Microarray data

Table S21 Microarray data of the 43 significantly (p < 0.05) downregulated (≥ 2-fold) genes in *ssrp1-2* and/or *spt16-1* mutants in comparison to *wild type* Col-0.

Fold Change ssrp1-2 vs Col-0	p-value ssrp1-2 vs Col-0	Fold Change spt16-1 vs Col-0	p-value spt16-1 vs Col-0	TAIR Numbers	Gene Symbol
-2.00	0.00065	-2.08	0.00048	AT1G14250	AT1G14250
-2.01	0.00002	-2.11	0.00016	AT5G05270	AT5G05270
-2.08	0.00045	-2.04	0.00118	AT5G13930	TT4
-2.15	0.00006	-2.43	0.00049	AT1G24580	AT1G24580
-2.18	0.00142	-2.36	0.00001	AT4G22870	AT4G22870
-2.56	0.00091	-2.41	0.00039	AT5G42800	DFR
-2.59	0.00260	-2.05	0.00452	AT5G07990	TT7
-2.62	0.00012	-2.69	0.00019	AT4G22880	LDOX
-2.80	0.00438	-2.23	0.00113	AT5G54060	UF3GT
-3.03	0.00024	-2.27	0.00073	AT1G52000	AT1G52000
-3.12	0.00087	-2.04	0.00100	AT1G03940	AT1G03940
-3.21	0.00281	-2.21	0.01043	AT1G12030	AT1G12030
-3.60	0.00033	-3.35	0.00286	AT5G04150	BHLH101
-3.75	0.00003	-2.71	0.00011	AT3G28220	AT3G28220
-4.22	0.00003	-2.67	0.00012	AT2G30766	AT2G30766
-4.36	0.00148	-3.16	0.00188	AT1G52040	MBP1
-4.58	0.01191	-4.22	0.01144	AT4G36700	AT4G36700
-5.26	0.00029	-3.00	0.00072	AT3G56980	BHLH039
-5.88	0.00003	-3.26	0.00084	AT1G47400	AT1G47400
-6.62	0.00032	-3.50	0.00035	AT2G14247	AT2G14247
-7.78	0.00024	-3.41	0.00002	AT1G47395	AT1G47395
-8.00	0.00000	-5.61	0.00470	AT2G41240	BHLH100
-9.69	0.00004	-7.73	0.00028	AT3G56970	BHLH038
-2.07	0.00202	-1.40	0.00067	AT3G45430	AT3G45430
-2.10	0.00978	-1.71	0.00583	AT3G24230	AT3G24230
-2.15	0.00029	-1.99	0.00128	AT5G67370	AT5G67370
-2.18	0.00024	-1.72	0.00048	AT1G23020	FRO3
-2.26	0.00010	-1.70	0.00010	AT5G13740	ZIF1
-2.28	0.00306	-1.90	0.00154	AT5G48850	ATSDI1
-2.29	0.00094	-1.77	0.00187	AT5G17220	GSTF12
-2.37	0.00013	-1.25	0.02251	AT3G28740	CYP81D1
-2.44	0.00007	-1.78	0.00222	AT3G29590	AT5MAT
-2.84	0.00014	-1.81	0.00528	AT5G19470	NUDT24
-2.95	0.00105	-1.81	0.01126	AT5G05250	AT5G05250
-3.28	0.00004	-1.83	0.00208	AT5G53450	ORG1
-1.57	0.00096	-2.02	0.00015	AT1G01060	LHY
-1.92	0.00406	-2.12	0.00345	AT3G02380	COL2
-1.57	0.00462	-2.16	0.00316	AT4G26150	CGA1
-1.69	0.00104	-2.22	0.00287	AT5G52570	BETA-OHASE 2
-1.77	0.00025	-2.24	0.00010	AT3G51240	F3H
-1.70	0.00671	-2.28	0.00094	AT5G13170	SAG29
-1.89	0.00300	-2.30	0.00025	AT3G09600	AT3G09600
-1.76	0.00015	-2.47	0.00008	AT2G46830	CCA1

Table S22 Microarray data of the 103 significantly ($p < 0.05$) upregulated ($\geq$ 2-fold) genes in *ssrp1-2* and/or *spt16-1* mutants in comparison to *wild type* Col-0.

Fold Change ssrp1-2 vs Col-0	p-value ssrp1-2 vs Col-0	Fold Change spt16-1 vs Col-0	p-value spt16-1 vs Col-0	TAIR Numbers	Gene Symbol
5.23	0.00000	2.42	0.00008	AT2G20142	AT2G20142
4.73	0.00133	4.07	0.00151	AT2G19800	MIOX2
4.17	0.00004	2.33	0.00115	AT1G27730	STZ
3.93	0.00086	2.40	0.00110	AT1G73805	AT1G73805
3.74	0.00086	2.35	0.00167	AT3G30775	ERD5
3.57	0.00171	2.35	0.00254	AT3G59790	MPK10
3.09	0.00255	2.00	0.02743	AT1G07135	AT1G07135
2.78	0.00234	3.56	0.00103	AT1G11070	AT1G11070
2.77	0.00006	2.44	0.00006	AT5G35735	AT5G35735
2.70	0.00075	2.39	0.00137	AT2G33830	AT2G33830
2.61	0.00043	2.66	0.00124	AT1G15125	AT1G15125
2.60	0.00032	5.36	0.00000	AT5G23240	AT5G23240
2.57	0.00024	2.28	0.00034	AT3G23170	AT3G23170
2.53	0.00005	4.44	0.00001	AT5G24470	PRR5
2.48	0.00005	3.96	0.00002	AT1G56300	AT1G56300
2.31	0.00247	2.29	0.00026	AT3G43670	AT3G43670
2.25	0.00008	2.04	0.00029	AT1G20630	CAT1
2.17	0.00667	2.30	0.01010	AT5G45095	AT5G45095
2.11	0.00019	2.05	0.00006	AT2G15890	MEE14
2.06	0.00053	3.19	0.00035	AT2G05915	AT2G05915
2.03	0.00036	2.83	0.00053	AT3G46640	PCL1
45.06	0.00001	1.82	0.00735		—
4.93	0.00016	1.82	0.00735	AT5G20230	BCB
4.44	0.00034	1.85	0.04137	AT3G56210	AT3G56210
4.21	0.00007	1.80	0.00525	AT2G24600	AT2G24600
3.71	0.00005	1.90	0.00001	AT3G01290	AT3G01290
3.43	0.00004	1.65	0.00543	AT1G80840	WRKY40
3.37	0.00000	1.77	0.00266	AT1G25400	AT1G25400
3.22	0.00128	1.59	0.11464	AT1G76650	CML38
3.08	0.00005	1.60	0.00377	AT3G50480	HR4
3.05	0.00008	1.47	0.00111	AT1G72416	AT1G72416
2.99	0.00187	1.72	0.02198	AT3G52400	SYP122
2.80	0.00289	1.20	0.07212	AT3G27940	LBD26
2.77	0.00000	1.17	0.00700	AT3G47340	ASN1
2.75	0.01686	1.53	0.02462	AT5G51190	AT5G51190
2.75	0.00070	1.71	0.00455	AT5G41080	AT5G41080
2.71	0.00001	1.71	0.00021	AT5G20250	DIN10
2.68	0.00023	1.77	0.00416	AT4G30270	XTH24
2.60	0.00210	1.88	0.00963	AT3G48360	BT2
2.43	0.00068	1.52	0.00525	AT5G41750	AT5G41750
2.43	0.00003	1.51	0.00278	AT2G37540	AT2G37540
2.41	0.00556	1.71	0.00894	AT1G02770	AT1G02770
2.41	0.00095	1.73	0.00258	AT3G15630	AT3G15630
2.39	0.00016	1.36	0.00008		—
2.37	0.00213	1.44	0.04311	AT2G41100	TCH3
2.36	0.00017	1.65	0.00434	AT3G59350	AT3G59350
2.35	0.00007	1.32	0.00976	AT2G34930	AT2G34930
2.30	0.00056	1.63	0.02628	AT4G17490	ERF6
2.28	0.00001	1.22	0.03135	AT4G37770	ACS8
2.27	0.00086	1.34	0.01675	AT1G35350	AT1G35350
2.24	0.00009	1.24	0.01122	AT3G45970	EXLA1
2.24	0.03337	1.44	0.01593	AT4G10695	AT4G10695
2.23	0.00184	1.41	0.04150	AT5G54710	AT5G54710
2.23	0.00050	1.43	0.00831	AT4G16563	AT4G16563
2.21	0.00051	1.51	0.00002	AT1G24530	AT1G24530
2.20	0.00070	1.43	0.00443	AT3G61190	BAP1
2.20	0.00065	1.22	0.03699	AT2G40000	HSPRO2
2.20	0.00212	1.61	0.02553	AT5G51720	AT5G51720
2.14	0.00075	1.23	0.00314	AT1G19020	AT1G19020
2.14	0.00658	1.80	0.00921	AT5G45840	AT5G45840
2.12	0.00072	1.39	0.00961	AT4G37610	BT5
2.11	0.01256	1.56	0.02750	AT5G56870	BGAL4
2.06	0.01124	1.46	0.04911	AT2G26560	PLA2A
2.05	0.00020	1.82	0.00053	AT4G38550	AT4G38550
2.05	0.01781	1.60	0.01081	AT1G51270	AT1G51270
2.04	0.01626	1.73	0.01236	AT2G45510	CYP704A2
2.04	0.00614	1.69	0.01878	AT2G15880	AT2G15880
2.03	0.00059	1.63	0.00366	AT2G30250	WRKY25
2.02	0.03788	1.40	0.02571	AT4G40020	AT4G40020
2.01	0.00524	1.79	0.00628	AT1G69890	AT1G69890
2.01	0.01018	1.75	0.00940	AT1G67265	RTFL21
2.00	0.00121	1.53	0.00339	AT2G36885	AT2G36885
2.00	0.00005	1.23	0.00482	AT1G08930	ERD6
1.02	0.84964	8.37	0.00002	AT2G18193	AT2G18193

Table S22 (Continuation) Microarray data of the 103 significantly ($p < 0.05$) upregulated ($\geq$ 2-fold) genes in *ssrp1-2* and/or *spt16-1* mutants in comparison to *wild type* Col-0.

Fold Change ssrp1-2 vs Col-0	p-value ssrp1-2 vs Col-0	Fold Change spt16-1 vs Col-0	p-value spt16-1 vs Col-0	TAIR Numbers	Gene Symbol
1.12	0.34822	8.36	0.00011	AT5G65320	AT5G65320
1.26	0.16412	6.72	0.00003	AT5G51330	SWI1
1.11	0.64170	6.09	0.00017		—
3.60	0.13353	4.11	0.00155	AT1G74870	AT1G74870
1.78	0.00471	3.90	0.00007	AT4G33930	AT4G33930
1.77	0.00006	3.77	0.00000	AT2G21660	CCR2
1.39	0.09694	3.50	0.00021	AT5G60100	PRR3
1.21	0.14379	3.12	0.00311	AT3G22231	PCC1
-1.03	0.73031	2.88	0.00096	AT5G53230	AT5G53230
1.98	0.00001	2.86	0.00000	AT4G04330	AT4G04330
1.10	0.25187	2.86	0.00048	AT5G60250	AT5G60250
1.94	0.00351	2.73	0.00007	AT5G54960	PDC2
1.41	0.00027	2.72	0.00005	AT4G30650	AT4G30650
-1.11	0.40518	2.68	0.00114	AT4G05370	AT4G05370
1.93	0.00085	2.67	0.00028	AT4G34950	AT4G34950
1.53	0.01274	2.49	0.00085	AT2G19450	TAG1
1.51	0.01453	2.43	0.00100	AT4G19120	ERD3
1.69	0.01408	2.43	0.00593	AT4G33070	AT4G33070
1.52	0.02828	2.42	0.00453	AT2G39920	AT2G39920
1.18	0.23883	2.40	0.00584	AT1G70440	SRO3
1.72	0.00878	2.37	0.00218	AT1G51090	AT1G51090
1.41	0.00598	2.36	0.00006	AT5G48250	AT5G48250
1.76	0.00002	2.31	0.00002	AT1G22770	GI
1.66	0.00426	2.28	0.00181	AT4G12290	AT4G12290
1.23	0.01887	2.27	0.00008	AT5G52310	LTI78
1.70	0.00024	2.27	0.00009	AT1G49720	ABF1
1.42	0.01573	2.25	0.00132	AT4G16146	AT4G16146
1.27	0.25243	2.24	0.01182	AT1G49490	AT1G49490
1.80	0.00038	2.24	0.00097	AT5G37600	GSR 1
1.47	0.00135	2.23	0.00004	AT1G08890	AT1G08890
1.78	0.00452	2.21	0.00138	AT5G40180	AT5G40180
-1.23	0.40467	2.19	0.00193	AT2G04050	AT2G04050
1.02	0.87118	2.19	0.00139	AT1G17960	AT1G17960
1.10	0.59895	2.17	0.00184	AT2G21640	AT2G21640
1.60	0.03108	2.17	0.00727	AT1G65330	PHE1
1.05	0.39820	2.16	0.00656	AT5G57730	AT5G57730
1.17	0.55298	2.11	0.02018	AT4G33980	AT4G33980
1.43	0.00455	2.10	0.00318	AT1G11210	AT1G11210
1.54	0.03806	2.09	0.01151	AT2G22450	AT2G22450
1.65	0.01026	2.08	0.00307	AT1G21130	AT1G21130
1.24	0.19450	2.08	0.02237	AT1G68050	FKF1
1.20	0.31343	2.07	0.00734	AT1G76790	AT1G76790
1.22	0.01084	2.06	0.00158	AT2G42530	COR15B
1.18	0.16206	2.05	0.00112	AT5G15970	KIN2
1.12	0.50455	2.05	0.02133	AT3G54340	AP3
1.10	0.54530	2.04	0.01949	AT5G25110	CIPK25
1.39	0.00333	2.04	0.00079	AT5G51440	AT5G51440
1.17	0.30247	2.03	0.00725	AT3G15720	AT3G15720
1.58	0.00147	2.02	0.00005	AT2G28900	OEP16-1
1.63	0.00408	2.02	0.00089	AT1G51610	AT1G51610
1.98	0.00326	2.02	0.00295	AT1G71890	SUC5
1.24	0.01303	2.01	0.00005	AT5G24280	GMI1
1.81	0.00357	2.01	0.00035	AT5G57630	CIPK21

12.3 Mass spectrometry

All mass spectrometry data can be found on the CD attached to the back cover of this thesis.

Table S23 List of expected interactors
Proteins that co-purified with ADA2b, TAF13, TAf10, UBP22, ENY2, SGF11, MOS4, and THP1 were identified by mass spectrometry and compared to a list of expected interactors to get more information about their association to protein complexes. Proteins that co-purified with the unfused SG-tag were removed. The average MASCOT score of the identified proteins is shown and how many times the protein was detected out of three independent affinity purifications.

TAIR	Protein	Complex	Process	ADA2b	TAF13	TAF10	UBP22	ENY2	SGF11	MOS4	THP1
AT5G54640	HTA1/H2A.1	Canonical H2A	Chromatin					103 / 2	442 / 2	287 / 2	
AT1G51060	HTA10/H2.10	Canonical H2A	Chromatin					150 / 2			
AT3G20670	HTA13/H2A.13	Canonical H2A	Chromatin						324 / 2		
AT4G27230	HTA2/H2A.2	Canonical H2A	Chromatin								
AT1G08880	HTA5/H2A.X.5	Canonical H2B	Chromatin								
AT1G07790	HTB1/H2B.1	Canonical H2B	Chromatin								
AT1G07790	HTB1/H2B.1	Canonical H2B	Chromatin								
AT5G02570	HTB10/H2B.10	Canonical H2B	Chromatin			205 / 2			545 / 2		
AT3G46030	HTB11/H2B.11	Canonical H2B	Chromatin					203 / 2	663 / 2	595 / 2	
AT5G22880	HTB2/H2B.2	Canonical H2B	Chromatin	689 / 2			336 / 2		945 / 3	643 / 2	533 / 2
AT2G28720	HTB3/H2B.3	Canonical H2B	Chromatin								
AT5G59910	HTB4/H2B.4	Canonical H2B	Chromatin								
AT2G37470	HTB5/H2B.5	Canonical H2B	Chromatin								
AT3G53650	HTB6/H2B.6	Canonical H2B	Chromatin				433 / 2		282 / 2		
AT3G09480	HTB7/H2B.7	Canonical H2B	Chromatin						114 / 2		
AT1G08170	HTB8/H2B.8	Canonical H2B	Chromatin								
AT3G45980	HTB9/H2B.9	Canonical H2B	Chromatin								
AT5G65360	HTR1/H3.1	Canonical H3	Chromatin								
AT5G10390	HTR13/H3.1	Canonical H3	Chromatin								
AT1G09200	HTR2/H3.1	Canonical H3	Chromatin						181 / 2		
AT3G27360	HTR3/H3.1	Canonical H3	Chromatin								
AT5G10400	HTR9/H3.1	Canonical H3	Chromatin								
AT1G01370	HTR12/CENH3	Centromeric H3	Chromatin								
AT1G06760	HON1/H1.1	H1	Chromatin								
AT2G30620	HON2/H1.2	H1	Chromatin						181 / 2		
AT2G18050	HON3/H1.3	H1	Chromatin								
AT3G46320	HFO1	H4	Chromatin								
AT5G59690	HFO2	H4	Chromatin								
AT2G28740	HFO3	H4	Chromatin								
AT1G07820	HFO4	H4	Chromatin								
AT3G53730	HFO5	H4	Chromatin								
AT5G59970	HFO6	H4	Chromatin								
AT3G45930	HFO7	H4	Chromatin								
AT1G07669	HFO8	H4	Chromatin								
AT1G19890	HTR10/H3.10	Unusual H3	Chromatin								
AT1G75600	HTR14/H3.14	Unusual H3	Chromatin								
AT1G13370	HTR6/H3.6	Unusual H3	Chromatin								
AT1G75610	HTR7/H3.7	Unusual H3	Chromatin								
AT1G08880	H2A.X.5	Variant H2A	Chromatin								
AT3G54560	HTA11/H2A.Z.11	Variant H2A	Chromatin								
AT5G02560	HTA12/H2A.W.12	Variant H2A	Chromatin								
AT1G54690	HTA3/H2A.X.3	Variant H2A	Chromatin						455 / 2		
AT4G13570	HTA4/H2A.Z.4	Variant H2A	Chromatin								
AT5G59870	HTA6/H2A.W.6	Variant H2A	Chromatin						171 / 2		
AT5G27670	HTA7/H2A.W.7	Variant H2A	Chromatin						130 / 2		
AT2G38810	HTA8/H2A.Z.8	Variant H2A	Chromatin								
AT1G52740	HTA9/H2A.Z.9	Variant H2A	Chromatin	120 / 3					188 / 2		
AT4G40030	HTR4/H3.3	Variant H3	Chromatin								166 / 2
AT4G40040	HTR5/H3.3	Variant H3	Chromatin								
AT5G10980	HTR8/H3.3	Variant H3	Chromatin								
AT4G17380	MSH4	DNA repair	DNA Repair								
AT2G41460	DNA lyase	DNA repair	DNA Repair								
AT2G41460	DNA- lyase	DNA repair	DNA Repair								
AT3G48425	DNA- lyase	DNA repair	DNA Repair								
AT4G36050	DNA lyase 2	DNA repair	DNA Repair								
AT4G09140	MLH1	DNA repair	DNA Repair								
AT4G35520	MLH3	DNA repair	DNA Repair								
AT3G24320	MSH1	DNA repair	DNA Repair		286 / 2						
AT3G18524	MSH2	DNA repair	DNA Repair								
AT4G25540	MSH3	DNA repair	DNA Repair								
AT3G20475	MSH5	DNA repair	DNA Repair								
AT4G02070	MSH6	DNA repair	DNA Repair								
AT3G24495	MSH7	DNA repair	DNA Repair								
AT4G02460	PMS1	DNA repair	DNA Repair								
AT3G26680	SNM1	DNA repair	DNA Repair								
AT3G22880	DMC1	RecA family	DNA Repair								
AT1G79050	ecA	RecA family	DNA Repair	372 / 2	840 / 3	297 / 3	523 / 3	315 / 3	690 / 3		286 / 2
AT5G20850	RAD51	RecA family	DNA Repair								
AT2G28560	RAD51	RecA family	DNA Repair								
AT1G07745	RAD51	RecA family	DNA Repair								
AT2G45280	RAD51	RecA family	DNA Repair								
AT3G10140	recA	RecA family	DNA Repair								
AT2G19490	recA	RecA family	DNA Repair	238 / 2	477 / 3	256 / 3	305 / 3		364 / 3		
AT3G32920	recA	RecA family	DNA Repair								
AT5G64520	XRCC2	RecA family	DNA Repair								
AT5G57450	XRCC3	RecA family	DNA Repair								
AT3G60500	CER7	Exosome	Exosome								
AT2G17510	EMB2763	Exosome	Exosome	368 / 3	305 / 3		273 / 3		256 / 2		
AT3G12990	PH45A	Exosome	Exosome								
AT3G61620	RRP41	Exosome	Exosome								
AT3G46210	RRP46	Exosome	Exosome								
AT5G35910	RRP6	Exosome	Exosome	223 / 2	312 / 3		293 / 3		156 / 3		
AT1G54440	RRP6-like	Exosome	Exosome								
AT1G75660	XRN3	Exosome	Exosome	341 / 2	278 / 3		133 / 2		159 / 2		
AT1G66740	ASF1A	Chromatin assembly	Replication								
AT5G38110	ASF1B	Chromatin assembly	Replication								
AT1G65470	FAS1	Chromatin assembly	Replication								
AT5G64630	FAS2	Chromatin assembly	Replication								
AT3G44530	HIRA	Chromatin assembly	Replication								
AT5G58230	MSI1	Chromatin assembly	Replication		223 / 2		235 / 3		258 / 2		
AT4G26110	NAP1-1	Chromatin assembly	Replication	335 / 3	254 / 3	219 / 3	464 / 3	109 / 2	522 / 3		228 / 2
AT2G19480	NAP1-2	Chromatin assembly	Replication	539 / 3	439 / 3	369 / 3	715 / 3		654 / 3		252 / 2
AT5G56950	NAP1-3	Chromatin assembly	Replication	238 / 2		181 / 2	375 / 2		471 / 3		162 / 3
AT3G13782	NAP1-4	Chromatin assembly	Replication								
AT1G74560	NRP1	Chromatin assembly	Replication								
AT1G18800	NRP2	Chromatin assembly	Replication								
AT2G29680	CDC6A	Pre-replication	Replication								
AT1G07270	CDc6b	Pre-replication	Replication								
AT2G31270	CDT1A	Pre-replication	Replication								
AT3G54710	CDT1B	Pre-replication	Replication								

Table S23 (Continuation) List of expected interactors
Proteins that co-purified with ADA2b, TAF13, TAf10, UBP22, ENY2, SGF11, MOS4, and THP1 were identified by mass spectrometry and compared to a list of expected interactors to get more information about their association to protein complexes. Proteins that co-purified with the unfused SG-tag were removed. The average MASCOT score of the identified proteins is shown and how many times the protein was detected out of three independent affinity purifications.

TAIR	Protein	Complex	Process	ADA2b	TAF13	TAF10	UBP22	ENY2	SGF11	MOS4	THP1
AT5G64610	HAM1	Pre-replication	Replication								
AT5G09740	HAM2	Pre-replication	Replication								
AT1G44900	MCM2	Pre-replication	Replication	469 / 3	362 / 3	133 / 3	356 / 2				
AT5G46280	MCM3	Pre-replication	Replication	238 / 2	135 / 2						
AT2G16440	MCM4	Pre-replication	Replication	473 / 3	911 / 3	144 / 3	340 / 3		285 / 3		
AT2G07690	MCM5	Pre-replication	Replication	467 / 3	218 / 3		228 / 3				
AT5G44635	MCM6	Pre-replication	Replication	381 / 2	398 / 3		267 / 2				
AT4G02060	MCM7	Pre-replication	Replication	174 / 2			105 / 2				
AT2G14050	MCM9	Pre-replication	Replication								
AT4G14700	ORC1A	Pre-replication	Replication								
AT4G12620	ORC1B	Pre-replication	Replication								
AT2G37560	ORC2	Pre-replication	Replication								
AT5G16690	ORC3	Pre-replication	Replication								
AT2G01120	ORC4	Pre-replication	Replication								
AT4G29910	ORC5	Pre-replication	Replication								
AT1G26840	ORC6	Pre-replication	Replication								
AT1G21690	AT1G21690	Replication factors	Replication		288 / 3		254 / 3		405 / 2		
AT1G24290	AT1G24290	Replication factors	Replication	98 / 2	232 / 3	145 / 2	253 / 2		233 / 3		241 / 2
AT1G63160	AT1G63160	Replication factors	Replication	310 / 2	514 / 3		281 / 3		353 / 2		99 / 2
AT1G77470	AT1G77470	Replication factors	Replication		200 / 3		158 / 3		251 / 2		
AT1G77620	AT1G77620	Replication factors	Replication								
AT5G22010	AT5G22010	Replication factors	Replication								
AT5G27740	AT5G27740	Replication factors	Replication		98 / 2						
AT3G25100	CDC45	Replication fork	Replication								
AT5G26680	FEN1	Replication fork	Replication								
AT1G08130	LIG1	Replication fork	Replication		132 / 2						
AT2G20980	MCM10	Replication fork	Replication								
AT3G09660	MCM8	Replication fork	Replication								
AT1G07370	PCNA1	Replication fork	Replication								
AT2G29570	PCNA2	Replication fork	Replication								
AT5G67100	POLA1	Replication fork	Replication		269 / 2						
AT1G67630	POLA2	Replication fork	Replication								
AT1G67320	POLA3/PRI1	Replication fork	Replication								
AT5G41880	POLA4/PRI2	Replication fork	Replication				181 / 2				
AT5G63960	POLD1	Replication fork	Replication		338 / 3						
AT2G42120	POLD2	Replication fork	Replication								
AT1G78650	POLD3	Replication fork	Replication								
AT1G09815	POLD4	Replication fork	Replication								
AT1G08260	POLE1a	Replication fork	Replication		340 / 2						
AT2G27120	POLE1b	Replication fork	Replication								
AT5G22110	POLE2	Replication fork	Replication								
AT1G80190	PSF1	Replication fork	Replication								
AT3G12530	PSF2	Replication fork	Replication								
AT1G19080	PSF3	Replication fork	Replication								
AT5G22010	RFC1	Replication fork	Replication								
AT1G63160	RFC2	Replication fork	Replication	310 / 2	514 / 3		281 / 3		353 / 2		99 / 2
AT5G27740	RFC3	Replication fork	Replication		98 / 2						
AT1G21690	RFC4	Replication fork	Replication		288 / 3		254 / 3		405 / 2		
AT1G77470	RFC5	Replication fork	Replication		200 / 3		158 / 3		251 / 2		
AT5G06510	RPA1a	Replication fork	Replication								
AT5G08020	RPA1b	Replication fork	Replication								
AT5G45400	RPA1c	Replication fork	Replication								
AT5G61000	RPA1d	Replication fork	Replication								
AT2G24490	RPA2a	Replication fork	Replication								
AT3G02920	RPA2b	Replication fork	Replication								
AT3G52630	RPA3a	Replication fork	Replication								
AT4G18590	RPA3b	Replication fork	Replication								
AT5G49010	SLD5	Replication fork	Replication								
AT2G27040	AGO4	AGO	Transcription								
AT2G25170	PKL	CHD1-like	Transcription		473 / 2		360 / 2				
AT5G44800	PKR1	CHD1-like	Transcription								
AT4G31900	PKR2	CHD1-like	Transcription								
AT5G67380	CKA1	CK2	Transcription	139 / 2			408 / 2				
AT3G50000	CKA2	CK2	Transcription	243 / 2	477 / 3		328 / 3		397 / 3		164 / 2
AT2G23080	CKA3	CK2	Transcription		474 / 2		275 / 2				
AT2G23070	CKA4	CK2	Transcription		353 / 3		337 / 3		265 / 2		
AT5G47080	CKB1	CK2	Transcription								
AT4G17640	CKB2	CK2	Transcription								
AT3G60250	CKB3	CK2	Transcription								
AT2G44680	CKB4	CK2	Transcription								
AT3G44660	HDA10	Deacetylase	Transcription								
AT4G33470	HDA14	Deacetylase	Transcription								
AT3G18520	HDA15	Deacetylase	Transcription	142 / 3	276 / 2						
AT3G44490	HDA17	Deacetylase	Transcription								
AT4G38130	HDA19	Deacetylase	Transcription	135 / 3	143 / 2		208 / 2		97 / 2		
AT5G61070	HDA19	Deacetylase	Transcription								
AT5G26040	HDA2	Deacetylase	Transcription								
AT5G61060	HDA5	Deacetylase	Transcription								
AT5G63110	HDA6	Deacetylase	Transcription								
AT5G35600	HDA7	Deacetylase	Transcription								
AT1G08460	HDA8	Deacetylase	Transcription								
AT3G44680	HDA9	Deacetylase	Transcription								
AT3G44750	HDT1	Deacetylase	Transcription								
AT5G22650	HDT2	Deacetylase	Transcription		169 / 2		193 / 3		206 / 2		195 / 3
AT5G03740	HDT3	Deacetylase	Transcription				189 / 3		170 / 2		160 / 3
AT2G27840	HDT4	Deacetylase	Transcription				130 / 2				
AT5G55760	SRT1	Deacetylase	Transcription								
AT5G09230	SRT2	Deacetylase	Transcription								
AT5G46030	ELF-1	ELF1	Transcription								
AT5G13680	ELP1, ELO2	Elongator	Transcription	1493 / 3	2438 / 3	413 / 3	1449 / 3		892 / 3		338 / 2
AT1G49540	ELP2	Elongator	Transcription	208 / 2							
AT5G50320	ELP3; ELO3	Elongator	Transcription	672 / 3	652 / 3	339 / 3	753 / 3	245 / 2	338 / 3		142 / 2
AT3G11220	ELP4; ELO1	Elongator	Transcription								
AT2G18410	ELP5	Elongator	Transcription				197 / 2				
AT4G10090	ELP6	Elongator	Transcription								
AT4G10710	SPT16	FACT	Transcription	136 / 2					215 / 2		
AT3G28730	SSRP1	FACT	Transcription								
AT5G38110	AtASF1L1	histone chaperones	Transcription								
AT1G66740	AtASF1L2	histone chaperones	Transcription								
AT1G65470	AtCAF1AL	histone chaperones	Transcription								

Table S23 (Continuation) List of expected interactors
Proteins that co-purified with ADA2b, TAF13, TAf10, UBP22, ENY2, SGF11, MOS4, and THP1 were identified by mass spectrometry and compared to a list of expected interactors to get more information about their association to protein complexes. Proteins that co-purified with the unfused SG-tag were removed. The average MASCOT score of the identified proteins is shown and how many times the protein was detected out of three independent affinity purifications.

TAIR	Protein	Complex	Process	ADA2b	TAF13	TAF10	UBP22	ENY2	SGF11	MOS4	THP1
AT5G64630	AtCAF1BL	histone chaperones	Transcription								
AT5G58230	AtCAF1CL1	histone chaperones	Transcription		223 / 2		235 / 3		258 / 2		
AT4G35050	AtCAF1CL2	histone chaperones	Transcription								
AT2G16780	AtCAF1CL3	histone chaperones	Transcription								
AT2G19520	AtCAF1CL4	histone chaperones	Transcription				278 / 2		238 / 2		
AT4G29730	AtCAF1CL5	histone chaperones	Transcription								
AT2G19540	AtCAF1CL6	histone chaperones	Transcription	331 / 3	283 / 2	162 / 3	494 / 3		111 / 2		
AT3G44530	AtHIRAL	histone chaperones	Transcription								
AT4G26110	AtNAPL1	histone chaperones	Transcription	335 / 3	254 / 3	219 / 3	464 / 3	109 / 2	522 / 3		228 / 2
AT2G19480	AtNAPL2	histone chaperones	Transcription	539 / 3	439 / 3	369 / 3	715 / 3		654 / 3		252 / 2
AT5G56950	AtNAPL3	histone chaperones	Transcription	238 / 2		181 / 2	375 / 2		471 / 3		162 / 3
AT3G13782	AtNAPL4	histone chaperones	Transcription								
AT1G18800	AtNAPL5	histone chaperones	Transcription								
AT1G74560	AtNAPL6	histone chaperones	Transcription								
AT4G37210	AtNASPL	histone chaperones	Transcription		86 / 2						
AT2G44950	HUB1	HUB1/2	Transcription								
AT1G55255	HUB2	HUB1/2	Transcription								
AT1G55250	HUB2	HUB1/2	Transcription	170 / 2							
AT3G12380	ARP5	INO80	Transcription								
AT3G60830	ARP7	INO80	Transcription	362 / 2	197 / 3		365 / 3		142 / 2		
AT5G43500	ARP9	INO80	Transcription	123 / 2							
AT3G57300	INO80	INO80	Transcription	92 / 2							
AT5G13950	NFRKB	INO80	Transcription								
AT1G65650	UCH2	INO80	Transcription								
AT4G06634	YY1	INO80	Transcription								
AT5G16310	UCH2	INO80	Transcription		111 / 2		156 / 2				
AT3G06400	CHR11	ISWI	Transcription								
AT5G18620	CHR17	ISWI	Transcription		762 / 3						
AT1G32130	IWS1a	IWS1	Transcription	330 / 2			313 / 2				
AT4G19000	IWS1b	IWS1	Transcription								
AT5G41910	Med10af	Mediator	Transcription								
AT1G26665	Med10bf	Mediator	Transcription								
AT3G01435	Med11	Mediator	Transcription								
AT3G04740	Med14	Mediator	Transcription		756 / 3						
AT1G15780	Med15	Mediator	Transcription		406 / 2						
AT4G04920	Med16	Mediator	Transcription	265 / 2	480 / 3				171 / 2		
AT5G20170	Med17	Mediator	Transcription		148 / 2						
AT2G22370	Med18	Mediator	Transcription		110 / 2						
AT2G28230	Med20	Mediator	Transcription		131 / 2						
AT4G04780	Med21	Mediator	Transcription								
AT1G16430	Med22ah	Mediator	Transcription								
AT1G07950	Med22bh	Mediator	Transcription								
AT1G23230	Med23	Mediator	Transcription		615 / 2						
AT1G25540	Med25	Mediator	Transcription								
AT3G52860	Med28	Mediator	Transcription								
AT5G19910	Med31	Mediator	Transcription								
AT5G02850	Med4	Mediator	Transcription	110 / 2							
AT3G21350	Med6	Mediator	Transcription								
AT5G03220	Med7ad	Mediator	Transcription								
AT5G03500	Med7bd	Mediator	Transcription								
AT2G03070	Med8	Mediator	Transcription								
AT1G55080	Med9	Mediator	Transcription								
AT3G22590	CDC73	PAF-C	Transcription				451 / 3		282 / 2		
AT2G06210	CTR9, ELF8, VIP6	PAF-C	Transcription						530 / 2		221 / 3
AT5G61150	LEO1, VIP4	PAF-C	Transcription	230 / 2	155 / 2				225 / 2		
AT1G79730	PAF1, ELF7	PAF-C	Transcription	176 / 3	184 / 2		210 / 3		167 / 2		
AT1G61040	RTF1, VIP5	PAF-C	Transcription						159 / 3		
AT4G29830	SKI8, VIP3	PAF-C	Transcription	215 / 3	192 / 3		383 / 3		417 / 3		196 / 2
AT3G63270	ALP1	Polycomb	Transcription								
AT3G23980	BLI	Polycomb	Transcription				182 / 2				
AT2G23380	CLF	Polycomb	Transcription								
AT5G51230	EMF2	Polycomb	Transcription								
AT3G20740	FIE, FIE1, FIS3	Polycomb	Transcription		99 / 2						
AT5G17690	LHP1	Polycomb	Transcription								
AT1G02580	MEA	Polycomb	Transcription								
AT5G58230	MSI1	Polycomb	Transcription		223 / 2		235 / 3		258 / 2		
AT5G44280	RING 1A	Polycomb	Transcription		234 / 2						
AT1G03770	RING 1B	Polycomb	Transcription								
AT4G02020	SWN	Polycomb	Transcription		125 / 2						
AT4G28190	ULT1	Polycomb	Transcription								
AT5G57380	VIN3	Polycomb	Transcription								
AT4G16845	VRN2	Polycomb	Transcription								
AT3G24440	VRN5	Polycomb	Transcription								
AT3G57660	NRPA1	Pol I	Transcription	164 / 2	856 / 3		462 / 2		600 / 3		
AT3G13940	NRPA13	Pol I	Transcription		193 / 2		269 / 2		105 / 2		
AT5G64680	NRPA14	Pol I	Transcription								
AT1G29940	NRPA2	Pol I	Transcription		229 / 2				135 / 2		
AT1G60850	NRPA3	Pol I	Transcription		212 / 2		317 / 3		312 / 3		146 / 2
AT1G75670	NRPA7	Pol I	Transcription								
AT3G25940	NRPA9	Pol I	Transcription	122 / 3					87 / 2		
AT3G22320	NRP(A/B/C/D)5	Pol I, II, III, IV	Transcription	250 / 3	123 / 2	97 / 2	173 / 2		265 / 3		150 / 2
AT1G11475	NRP(A/B/C/D/E)10	Pol I, II, III, IV, V	Transcription								
AT5G41010	NRP(A/B/C/D/E)1	Pol I, II, III, IV, V	Transcription								
AT1G54250	NRP(A/B/C/D/E)8a	Pol I, II, III, IV, V	Transcription								
AT3G59600	NRP(A/B/C/D/E)8	Pol I, II, III, IV, V	Transcription								
AT2G29540	NRP(A/C)11	Pol I, III	Transcription						93 / 2		
AT4G35800	NRPB1	Pol II	Transcription	374 / 2	848 / 3		578 / 3		328 / 3		
AT1G61700	NRPB10-like	Pol II	Transcription								
AT1G53690	NRPB12-like	Pol II	Transcription								
AT4G21710	NRPB2	Pol II	Transcription	504 / 3	1010 / 3		554 / 2		330 / 3		
AT5G09920	NRPB4	Pol II	Transcription								
AT5G57980	NRPB5-like	Pol II	Transcription								
AT5G59180	NRPB7	Pol II	Transcription								
AT4G14520	NRPB7-like	Pol II	Transcription								
AT2G04630	NRP(B/C6b)	Pol II, III	Transcription								
AT5G51940	NRP(B/C/D/E)6a	Pol II, III, IV, V	Transcription								
AT3G52090	NRP(B/D/E)11	Pol II, IV, V	Transcription	119 / 3		121 / 2	121 / 2		154 / 2		
AT2G15430	NRP(B/D/E)3a	Pol II, IV, V	Transcription	536 / 2	402 / 3	355 / 3	523 / 3		283 / 3		233 / 2
AT3G16980	NRP(B/D/E)9a	Pol II, IV, V	Transcription	216 / 3		153 / 2	158 / 3	121 / 3	162 / 2		
AT4G16265	NRP(B/D/E)9b	Pol II, IV, V	Transcription	210 / 2		120 / 2	185 / 3		117 / 2		

Table S23 (Continuation) List of expected interactors
Proteins that co-purified with ADA2b, TAF13, TAf10, UBP22, ENY2, SGF11, MOS4, and THP1 were identified by mass spectrometry and compared to a list of expected interactors to get more information about their association to protein complexes. Proteins that co-purified with the unfused SG-tag were removed. The average MASCOT score of the identified proteins is shown and how many times the protein was detected out of three independent affinity purifications.

TAIR	Protein	Complex	Process	ADA2b	TAF13	TAF10	UBP22	ENY2	SGF11	MOS4	THP1
AT5G60040	NRPC1	Polymerase III	Transcription		188 / 2				190 / 2		
AT3G49000	NRPC13	Polymerase III	Transcription								
AT4G25180	NRPC14a	Polymerase III	Transcription								
AT5G09380	NRPC14b	Polymerase III	Transcription								
AT5G49530	NRPC15	Polymerase III	Transcription								
AT5G23710	NRPC16	Polymerase III	Transcription								
AT4G01590	NRPC17	Polymerase III	Transcription								
AT4G35680	NRPC17-like	Polymerase III	Transcription								
AT5G45140	NRPC2	Polymerase III	Transcription								
AT1G60620	NRPC3	Polymerase III	Transcription				133 / 2				
AT5G62950	NRPC4	Polymerase III	Transcription						189 / 2		
AT3G28956	NRPC4-like	Polymerase III	Transcription								
AT1G06790	NRPC7	Polymerase III	Transcription								
AT4G07950	NRPC9a	Polymerase III	Transcription								
AT1G01210	NRPC9b	Polymerase III	Transcription								
AT1G63020	NRPD1	Polymerase IV	Transcription		202 / 2						
AT3G18090	NRPD2b	Polymerase IV	Transcription	174 / 2							
AT3G22900	NRPD7	Polymerase IV	Transcription								
AT3G23780	NRP(D/E)2	Polymerase IV & V	Transcription		437 / 2				188 / 2		
AT4G15950	NRP(D/E)4	Polymerase IV & V	Transcription								
AT2G40030	NRPE1	Polymerase V	Transcription								
AT2G15400	NRPE3b	Polymerase V	Transcription						249 / 2		243 / 2
AT3G57080	NRPE5	Polymerase V	Transcription								
AT2G41340	NRPE5-like	Polymerase V	Transcription								
AT3G54490	NRPE5-like	Polymerase V	Transcription								
AT4G14660	NRPE7	Polymerase V	Transcription								
AT5G10270	CDKC;1	P-TEFb	Transcription		199 / 3				178 / 2		97 / 2
AT5G64960	CDKC;2	P-TEFb	Transcription				157 / 2				
AT4G19560	CYCT1;2	P-TEFb	Transcription								
AT1G27630	CYCT1;3	P-TEFb	Transcription								
AT4G19600	CYCT1;4	P-TEFb	Transcription								
AT5G45190	CYCT1;5	P-TEFb	Transcription	102 / 2	235 / 3		178 / 2		135 / 2		179 / 2
AT3G27100	ENY2	SAGA_DUBm	Transcription	154 / 2			611 / 3	783 / 3	1262 / 3	402 / 3	
AT5G58575	SGF11	SAGA_DUBm	Transcription				326 / 3	542 / 3	1747 / 3		
AT5G10790	UBP22	SAGA_DUBm	Transcription				1858 / 3	113 / 3	209 / 3		
AT4G16420	ADA2b	SAGA_HAT	Transcription	5515 / 3		446 / 3	127 / 2				
AT4G29790	ADA3	SAGA_HAT	Transcription	489 / 3							
AT3G54610	GCN5	SAGA_HAT	Transcription	6257 / 3	159 / 3	726 / 3	283 / 2				
AT3G27460	SGF29a	SAGA_HAT	Transcription	412 / 3							
AT5G40550	SGF29b	SAGA_HAT	Transcription	185 / 3							
AT2G13370	CHR5	SAGA_Other	Transcription								
AT2G14850	ADA1a	SAGA_SPT	Transcription	678 / 3		163 / 3					
AT5G67410	ADA1b	SAGA_SPT	Transcription	151 / 3							
AT1G32750	HAF1	SAGA_SPT	Transcription		5178 / 3	1727 / 3					
AT3G19040	HAF2	SAGA_SPT	Transcription								
AT1G72390	SPT20	SAGA_SPT	Transcription	4499 / 3	129 / 2	1606 / 3					
AT1G02680	TAF13	SAGA_SPT	Transcription		612 / 3						
AT2G17930	TRA1a	SAGA_SPT	Transcription	4304 / 3	1804 / 3	1124 / 3	482 / 3		270 / 2		
AT4G36080	TRA1b	SAGA_SPT	Transcription	2564 / 3	1498 / 2	712 / 3			160 / 2		
AT4G31720	TAF10	SAGA_TAF	Transcription	378 / 3	195 / 2	1144 / 3					
AT3G10070	TAF12	SAGA_TAF	Transcription	189 / 3	761 / 3	273 / 3					
AT1G17440	TAF12b	SAGA_TAF	Transcription	687 / 3							
AT5G25150	TAF5	SAGA_TAF	Transcription	1971 / 3	2393 / 3	1786 / 3					
AT1G04950	TAF6	SAGA_TAF	Transcription	130 / 2	1573 / 3	767 / 3					
AT1G54360	TAF6b	SAGA_TAF	Transcription	865 / 3		963 / 3					
AT1G54140	TAF9	SAGA_TAF	Transcription	539 / 3	481 / 2	545 / 2					
AT1G77300	SDG8, ASHH2	SDG8	Transcription								
AT5G08565	SPT4-1	SPT4/SPT5	Transcription								
AT5G63670	SPT4-2	SPT4/SPT5	Transcription						137 / 2		
AT2G34210	SPT5-1	SPT4/SPT5	Transcription								
AT4G08350	SPT5-2	SPT4/SPT5	Transcription	331 / 3	456 / 3		214 / 3		327 / 3		152 / 2
AT5G04290	SPT5L	SPT4/SPT5	Transcription								
AT1G65440	SPT6-1, SPT6L	SPT6	Transcription		662 / 2		249 / 2		209 / 2		
AT1G63210	SPT6-2	SPT6	Transcription								
AT2G46020	BRM	SWI/SNF-type	Transcription		714 / 3		366 / 3		270 / 3		
AT3G17590	BSH	SWI/SNF-type	Transcription	187 / 3	186 / 2				247 / 2		
AT3G06010	MINU1	SWI/SNF-type	Transcription								
AT5G19310	MINU2	SWI/SNF-type	Transcription								
AT2G47620	SWI3A	SWI/SNF-type	Transcription								
AT2G33610	SWI3B	SWI/SNF-type	Transcription								
AT1G21700	SWI3C	SWI/SNF-type	Transcription		102 / 2						
AT4G34430	SWI3D	SWI/SNF-type	Transcription	119 / 2	257 / 3				263 / 2		
AT3G01890	SWP73A	SWI/SNF-type	Transcription								
AT5G14170	SWP73B	SWI/SNF-type	Transcription	307 / 2	192 / 3	161 / 3	169 / 2		207 / 2		
AT2G28290	SYD	SWI/SNF-type	Transcription		372 / 2						
AT3G33520	ARP6	SWR1/NuA4	Transcription								
AT3G24880	AtEAF1A	SWR1/NuA4	Transcription								
AT3G24870	AtEAF1B	SWR1/NuA4	Transcription								
AT2G47210	AtSWC4	SWR1/NuA4	Transcription								
AT5G45600	AtYAF9A	SWR1/NuA4	Transcription								
AT2G18000	AtYAF9B	SWR1/NuA4	Transcription								
AT4G14385	EAF6	SWR1/NuA4	Transcription								
AT1G26470	EAF7	SWR1/NuA4	Transcription								
AT1G16690	EPL1A	SWR1/NuA4	Transcription								
AT1G79020	EPL1B	SWR1/NuA4	Transcription								
AT5G64610	HAM1	SWR1/NuA4	Transcription								
AT1G54390	ING2	SWR1/NuA4	Transcription								
AT4G37280	MRG1	SWR1/NuA4	Transcription								
AT3G12810	PIE1	SWR1/NuA4	Transcription								
AT2G36740	SWC2	SWR1/NuA4	Transcription								
AT5G37055	SWC6	SWR1/NuA4	Transcription								
AT5G67630	RVB21	SWR1/NuA4, INO80	Transcription	1021 / 2	992 / 3	345 / 3	1164 / 3	259 / 3	800 / 3	237 / 2	408 / 2
AT3G49830	RVB22	SWR1/NuA4, INO80	Transcription								
AT5G22330	RVB1	SWR1/NuA4, INO80	Transcription	1060 / 2	926 / 3	762 / 3	1293 / 3	220 / 3	612 / 3	204 / 2	398 / 2
AT1G18450	AtARP4	SWR1/NuA4, INO80	Transcription	308 / 3	237 / 3	139 / 3	307 / 3		183 / 2		
AT2G37620	ACT1	SWR1-like	Transcription	815 / 3	859 / 3	1377 / 3	1016 / 3	576 / 3	1080 / 3		
AT3G12110	ACT11	SWR1-like	Transcription	1002 / 2	899 / 2	1148 / 2			1214 / 2		
AT3G46520	ACT12	SWR1-like	Transcription								
AT3G18780	ACT2	SWR1-like	Transcription					259 / 2	465 / 3		
AT3G53750	ACT3	SWR1-like	Transcription								1100 / 2

Table S23 (Continuation) List of expected interactors
Proteins that co-purified with ADA2b, TAF13, TAf10, UBP22, ENY2, SGF11, MOS4, and THP1 were identified by mass spectrometry and compared to a list of expected interactors to get more information about their association to protein complexes. Proteins that co-purified with the unfused SG-tag were removed. The average MASCOT score of the identified proteins is shown and how many times the protein was detected out of three independent affinity purifications.

TAIR	Protein	Complex	Process	ADA2b	TAF13	TAF10	UBP22	ENY2	SGF11	MOS4	THP1
AT5G59370	ACT4	SWR1-like	Transcription								
AT5G09810	ACT7	SWR1-like	Transcription	1230 / 3	1348 / 3	2055 / 3	1559 / 3	997 / 3	1328 / 3	423 / 2	776 / 3
AT1G49240	ACT8	SWR1-like	Transcription	808 / 3	1117 / 2	1210 / 2	1127 / 3	658 / 3	922 / 2		894 / 2
AT5G30490	SWC5	SWR1-like	Transcription								
AT1G07480	TFIIA-L1	TFIIA	Transcription								
AT1G07470	TFIIA-L2	TFIIA	Transcription								
AT5G59230	TFIIA-L3	TFIIA	Transcription								
AT4G24440	TFIIA-S	TFIIA	Transcription								
AT2G41630	TFIIB1	TFIIB	Transcription								
AT3G10330	TFIIB2	TFIIB	Transcription								
AT3G29380	TFIIB3	TFIIB	Transcription								
AT3G57370	TFIIB4	TFIIB	Transcription								
AT4G36650	TFIIB5	TFIIB	Transcription								
AT4G10680	TFIIB6	TFIIB	Transcription								
AT1G32750	TAF1	TFIID	Transcription		5178 / 3	1727 / 3					
AT4G31720	TAF10	TFIID	Transcription	378 / 3	195 / 2	1144 / 3					
AT4G20280	TAF11	TFIID	Transcription		1310 / 3	411 / 3					
AT1G20000	TAF11b	TFIID	Transcription								
AT3G10070	TAF12	TFIID	Transcription	189 / 3	761 / 3	273 / 3					
AT1G17440	TAF12b	TFIID	Transcription	687 / 3							
AT1G02680	TAF13	TFIID	Transcription		612 / 3						
AT2G18000	TAF14	TFIID	Transcription								
AT5G45600	TAF14b	TFIID	Transcription								
AT1G50300	TAF15	TFIID	Transcription								
AT5G58470	TAF15b	TFIID	Transcription						174 / 2		119 / 2
AT3G19040	TAF1b	TFIID	Transcription								
AT1G73960	TAF2	TFIID	Transcription		2595 / 3	600 / 3					
AT5G43130	TAF4	TFIID	Transcription		2919 / 3	1046 / 3					
AT1G27720	TAF4b	TFIID	Transcription								
AT5G25150	TAF5	TFIID	Transcription	1971 / 3	2393 / 3	1786 / 3					
AT1G04950	TAF6	TFIID	Transcription	130 / 2	1573 / 3	767 / 3					
AT1G54360	TAF6b1	TFIID	Transcription	865 / 3		963 / 3					
AT1G54360	TAF6b2	TFIID	Transcription	865 / 3		963 / 3					
AT1G54360	TAF6b3	TFIID	Transcription	865 / 3		963 / 3					
AT1G54360	TAF6b4	TFIID	Transcription	865 / 3		963 / 3					
AT1G55300	TAF7	TFIID	Transcription								
AT4G34340	TAF8	TFIID	Transcription		599 / 3	1080 / 3					
AT1G54140	TAF9	TFIID	Transcription	539 / 3	481 / 2	545 / 2					
AT3G13445	TBP1	TFIID	Transcription		283 / 2	211 / 2					
AT1G55520	TBP2	TFIID	Transcription		354 / 2	208 / 2					
AT1G03280	TFIIEa1	TFIIE	Transcription								
AT4G20340	TFIIEa2	TFIIE	Transcription								
AT4G20810	TFIIEa3	TFIIE	Transcription								
AT4G21010	TFIIEb1	TFIIE	Transcription								
AT4G20330	TFIIEb2	TFIIE	Transcription								
AT4G12610	TFIIF	TFIIF	Transcription	226 / 2	343 / 3		172 / 3				
AT1G75510	TFIIF1	TFIIF	Transcription	144 / 3	177 / 2		195 / 2		110 / 2		
AT3G52270	TFIIF2	TFIIF	Transcription								
AT2G38560	TFIIS	TFIIS	Transcription								
AT1G14400	UBC1	UBC1/2	Transcription								
AT2G02760	UBC2	UBC1/2	Transcription								
AT2G32780	UBP1	UBP	Transcription								
AT4G10590	UBP10	UBP	Transcription								
AT1G32850	UBP11	UBP	Transcription								
AT5G06600	UBP12	UBP	Transcription	514 / 2	565 / 2						
AT3G11910	UBP13	UBP	Transcription	648 / 2	579 / 3		317 / 3				
AT3G20630	UBP14	UBP	Transcription	216 / 3	136 / 2		136 / 3				
AT1G17110	UBP15	UBP	Transcription								
AT4G24560	UBP16	UBP	Transcription								
AT5G65450	UBP17	UBP	Transcription								
AT4G31670	UBP18	UBP	Transcription								
AT2G24640	UBP19	UBP	Transcription								
AT1G04860	UBP2	UBP	Transcription	173 / 2	371 / 3		168 / 2				
AT4G17890	UBP20	UBP	Transcription			195 / 3	137 / 3				
AT5G46740	UBP21	UBP	Transcription								
AT5G57990	UBP23	UBP	Transcription								
AT4G30890	UBP24	UBP	Transcription			131 / 2					
AT3G14400	UBP25	UBP	Transcription								
AT3G49600	UBP26	UBP	Transcription		291 / 3						
AT4G39370	UBP27	UBP	Transcription								
AT4G39910	UBP3	UBP	Transcription		142 / 2						
AT2G22310	UBP4	UBP	Transcription								
AT2G40930	UBP5	UBP	Transcription	227 / 2	379 / 3				235 / 2		
AT1G51710	UBP6	UBP	Transcription	520 / 2	507 / 3	214 / 3	453 / 3				
AT3G21280	UBP7	UBP	Transcription								
AT5G22030	UBP8	UBP	Transcription								
AT4G10570	UBP9	UBP	Transcription								
AT3G09620	atPrp5-1a	17S U2 associated	Splicing								
AT1G20920	atPrp5-1b	17S U2 associated	Splicing								
AT2G47330	atPrp5-2	17S U2 associated	Splicing	144 / 3	275 / 3		353 / 3				
AT2G02570	atSPF30	17S U2 associated	Splicing				246 / 2				
AT1G30480	atSPF45	17S U2 associated	Splicing	162 / 2	325 / 2		302 / 3				163 / 2
AT5G25060	atSR140-1	17S U2 associated	Splicing	483 / 3	1302 / 3		635 / 3		253 / 3		246 / 2
AT5G12190	atP14-1	17S U2 snRNP	Splicing	264 / 3	143 / 2		229 / 3	267 / 3	330 / 2	348 / 3	244 / 2
AT2G14870	atP14-2	17S U2 snRNP	Splicing								
AT1G14650	atSAP114-1a	17S U2 snRNP	Splicing	525 / 3	256 / 2		871 / 3	350 / 2		935 / 3	
AT1G14640	atSAP114-1b	17S U2 snRNP	Splicing								
AT5G06520	atSAP114-2	17S U2 snRNP	Splicing								
AT4G16200	atSAP114-3	17S U2 snRNP	Splicing								
AT4G15580	atSAP114p	17S U2 snRNP	Splicing								
AT3G55200	atSAP130a	17S U2 snRNP	Splicing	3556 / 3	4345 / 3	3105 / 3	1428 / 3	637 / 3	859 / 3	1734 / 2	550 / 2
AT3G55220	atSAP130b	17S U2 snRNP	Splicing								
AT5G64270	atSAP155	17S U2 snRNP	Splicing	1287 / 3	1674 / 3		1408 / 3	866 / 3	902 / 3	2867 / 2	906 / 3
AT2G18510	atSAP49a	17S U2 snRNP	Splicing	246 / 2	202 / 2		323 / 3	139 / 2	154 / 3	144 / 2	232 / 2
AT2G14550	atSAP49b	17S U2 snRNP	Splicing								
AT5G06160	atSAP61	17S U2 snRNP	Splicing	208 / 3	213 / 2		238 / 3	205 / 2		648 / 2	135 / 2
AT2G32600	atSAP62	17S U2 snRNP	Splicing	149 / 3			251 / 3	199 / 2		260 / 2	132 / 2
AT1G07170	atSF3b_14b	17S U2 snRNP	Splicing								
AT2G30000	atSF3b_14b	17S U2 snRNP	Splicing								
AT4G21660	atSF3b150	17S U2 snRNP	Splicing	411 / 3	365 / 3		851 / 3	157 / 2	243 / 2	749 / 3	

Table S23 (Continuation) List of expected interactors
Proteins that co-purified with ADA2b, TAF13, TAf10, UBP22, ENY2, SGF11, MOS4, and THP1 were identified by mass spectrometry and compared to a list of expected interactors to get more information about their association to protein complexes. Proteins that co-purified with the unfused SG-tag were removed. The average MASCOT score of the identified proteins is shown and how many times the protein was detected out of three independent affinity purifications.

TAIR	Protein	Complex	Process	ADA2b	TAF13	TAF10	UBP22	ENY2	SGF11	MOS4	THP1
AT1G11520	atSF3b150p	17S U2 snRNP	Splicing								
AT1G09760	atU2A	17S U2 snRNP	Splicing	228 / 2			352 / 3	418 / 3	489 / 3	1127 / 2	304 / 3
AT1G06960	atU2Ba	17S U2 snRNP	Splicing								
AT2G30260	atU2Bb	17S U2 snRNP	Splicing	110 / 2	104 / 2		171 / 2	423 / 3	304 / 3	568 / 2	
AT5G16260	ELF9	17S U2 snRNP	Splicing								
AT4G14342	SF3b10a	17S U2 snRNP	Splicing								
AT3G23325	SF3b10b	17S U2 snRNP	Splicing								
AT1G29220	AT1G29220	A complex assoc.	Splicing				320 / 3				
AT1G30970	AT1G30970	A complex assoc.	Splicing				151 / 3				
AT1G49910	AT1G49910	A complex assoc.	Splicing								
AT3G19590	AT3G19590	A complex assoc.	Splicing	393 / 3	431 / 3		605 / 3		156 / 2		
AT3G46820	AT3G46820	A complex assoc.	Splicing								
AT3G52120	AT3G52120	A complex assoc.	Splicing								
AT3G54230	AT3G54230	A complex assoc.	Splicing				223 / 2			223 / 2	
AT5G38840	AT5G38840	A complex assoc.	Splicing		126 / 2					281 / 2	
AT5G47790	AT5G47790	A complex assoc.	Splicing								
AT5G59160	AT5G59160	A complex assoc.	Splicing	179 / 2	319 / 2		175 / 3				
AT1G67580	AT1G67580	A complex assoc.	Splicing							252 / 2	
AT1G67580	AT1G67580	A complex assoc.	Splicing							252 / 2	
AT5G63370	AT5G63370	A complex assoc.	Splicing								
AT5G63370	AT5G63370	A complex assoc.	Splicing								
AT3G02860	AT3G02860	Bact complex	Splicing								
AT5G67530	AT5G67530	Bact complex	Splicing	211 / 3	207 / 3		257 / 3				
AT1G01940	AT1G01940	C complex	Splicing							345 / 2	
AT1G02330	AT1G02330	C complex	Splicing								
AT1G03910	AT1G03910	C complex	Splicing							278 / 2	
AT1G18080	AT1G18080	C complex	Splicing	190 / 2	346 / 2		221 / 3		250 / 2		
AT1G48630	AT1G48630	C complex	Splicing	343 / 3	363 / 3		335 / 3	174 / 3	245 / 3	124 / 2	152 / 3
AT1G61620	AT1G61620	C complex	Splicing	120 / 2	171 / 2		297 / 3	130 / 2	193 / 2	620 / 2	
AT2G21150	AT2G21150	C complex	Splicing		123 / 2		234 / 3	102 / 2			
AT3G06455	AT3G06455	C complex	Splicing								
AT3G07790	AT3G07790	C complex	Splicing								
AT3G09440	AT3G09440	C complex	Splicing	3285 / 3	3943 / 3	3146 / 3	3255 / 3	1138 / 3	2960 / 3	1657 / 3	1506 / 3
AT3G18130	AT3G18130	C complex	Splicing	386 / 3	461 / 2		460 / 2				
AT3G63400	AT3G63400	C complex	Splicing							499 / 3	
AT4G01000	AT4G01000	C complex	Splicing								
AT4G02720	AT4G02720	C complex	Splicing								
AT4G15030	AT4G15030	C complex	Splicing								
AT4G18465	AT4G18465	C complex	Splicing	628 / 3	427 / 2		741 / 3	166 / 2	112 / 3	835 / 3	
AT4G33370	AT4G33370	C complex	Splicing								
AT5G23080	AT5G23080	C complex	Splicing							949 / 2	152 / 2
AT5G49400	AT5G49400	C complex	Splicing								
AT5G51280	AT5G51280	C complex	Splicing	224 / 3			352 / 2	181 / 2	85 / 2	711 / 2	
AT5G64730	AT5G64730	C complex	Splicing	282 / 3	232 / 3		265 / 3	106 / 2	95 / 2		105 / 2
AT3G44600	AT3G44600	C complex	Splicing				173 / 2			930 / 2	
AT3G44600	AT3G44600	C complex	Splicing				173 / 2			930 / 2	
AT1G49590	At1g49590	B complex assoc.	Splicing								
AT1G53720	AT1G53720	B complex assoc.	Splicing	188 / 2	240 / 3		252 / 2				
AT1G55928	AT1G55928	B complex assoc.	Splicing								
AT2G27280	AT2G27280	B complex assoc.	Splicing								
AT2G27285	AT2G27285	B complex assoc.	Splicing								
AT2G42520	AT2G42520	B complex assoc.	Splicing	981 / 3	1056 / 3	471 / 3	846 / 3	215 / 3	645 / 3	442 / 2	482 / 3
AT3G54670	AT3G54670	B complex assoc.	Splicing								
AT3G58510	AT3G58510	B complex assoc.	Splicing	711 / 3	723 / 3	383 / 3	590 / 3	255 / 3	493 / 3	323 / 2	405 / 3
AT4G09980	AT4G09980	B complex assoc.	Splicing	251 / 2	306 / 3		310 / 3		106 / 2		
AT4G31120	AT4G31120	B complex assoc.	Splicing		174 / 2						
AT5G17900	AT5G17900	B complex assoc.	Splicing								
AT5G67320	AT5G67320	B complex assoc.	Splicing								
AT4G08580	AT4G08580	B complex assoc.	Splicing	521 / 3	343 / 2	159 / 3	675 / 3				
AT4G08580	AT4G08580	B complex assoc.	Splicing	521 / 3	343 / 2	159 / 3	675 / 3				
AT2G20330	AT2G20330	C complex assoc.	Splicing	183 / 2			133 / 3			711 / 2	
AT5G25754	AT5G25754	C complex assoc.	Splicing		486 / 2	214 / 2	256 / 2				
AT5G25757	AT5G25757	C complex assoc.	Splicing					153 / 3	238 / 2		150 / 2
AT3G51110	atCRN2	Core NTC	Splicing								
AT1G09770	CDC5	Core NTC	Splicing	906 / 3	482 / 3		1288 / 3	613 / 3	279 / 2	4211 / 3	495 / 2
AT5G45990	CRN1a	Core NTC	Splicing		162 / 3						
AT3G13210	CRN1b	Core NTC	Splicing								
AT5G41770	CRN1c	Core NTC	Splicing	243 / 3	198 / 3		303 / 3	346 / 3		1176 / 3	
AT3G18790	ISY1	Core NTC	Splicing	371 / 2	154 / 2		612 / 3	288 / 3		442 / 2	128 / 2
AT1G04510	MAC3A	Core NTC	Splicing	243 / 3	173 / 3		570 / 3	435 / 2	297 / 2	2304 / 3	365 / 2
AT2G33340	MAC3B	Core NTC	Splicing	404 / 3	219 / 3		603 / 3	339 / 3	269 / 2	3137 / 3	349 / 2
AT3G18165	MOS4	Core NTC	Splicing	133 / 3			135 / 3	174 / 3	127 / 2	1849 / 3	
AT4G15900	PRL1	Core NTC	Splicing	186 / 3	159 / 2		416 / 3	277 / 3		1296 / 3	248 / 2
AT3G16650	PRL2	Core NTC	Splicing								
AT1G77180	SKIP	Core NTC	Splicing	451 / 3	583 / 3		856 / 3	280 / 3	247 / 2	1243 / 3	306 / 3
AT5G28740	SYF1	Core NTC	Splicing	508 / 3	359 / 3		753 / 3	613 / 3		2124 / 3	
AT1G05460	AT1G05460	Detected in Bact	Splicing		692 / 3				338 / 3		180 / 2
AT1G61780	AT1G61780	Detected in Bact	Splicing								
AT3G12300	AT3G12300	Detected in Bact	Splicing							170 / 2	
AT3G52250	At3g52250	Detected in Bact	Splicing								
AT3G57910	AT3G57910	Detected in Bact	Splicing				221 / 3				
AT5G26742	AT5G26742	Detected in Bact	Splicing	621 / 3	445 / 3		674 / 3	486 / 3	1945 / 3	1259 / 3	1316 / 3
AT1G02140	AT1G02140	EJC/mRNP	Splicing	276 / 2			237 / 2	431 / 2		499 / 2	
AT1G15200?	AT1G15200?	EJC/mRNP	Splicing								
AT1G16610	AT1G16610	EJC/mRNP	Splicing								174 / 2
AT1G51380	At1g51380	EJC/mRNP	Splicing		241 / 3		270 / 3	124 / 2	263 / 3		
AT1G51510	AT1G51510	EJC/mRNP	Splicing					251 / 2		452 / 2	225 / 3
AT2G45640	AT2G45640	EJC/mRNP	Splicing	262 / 3			261 / 3	337 / 3	285 / 2	623 / 2	243 / 2
AT3G19760	AT3G19760	EJC/mRNP	Splicing	396 / 3	299 / 3		559 / 3	403 / 3	365 / 3	801 / 3	594 / 3
AT3G58570	AT3G58570	EJC/mRNP	Splicing	720 / 3	832 / 3	327 / 3	644 / 3	219 / 2	477 / 3	362 / 2	430 / 3
AT1G18630	At1g18630	GR RNA binding	Splicing	147 / 2			162 / 2		442 / 2		
AT1G74230	At1g74230	GR RNA binding	Splicing								
AT2G16260	At2g16260	GR RNA binding	Splicing								
AT2G21660	AT2G21660	GR RNA binding	Splicing	337 / 2			187 / 2		108 / 2		392 / 3
AT3G23830	At3g23830	GR RNA binding	Splicing								
AT4G13850	At4g13850	GR RNA binding	Splicing								
AT4G39260	AT4G39260	GR RNA binding	Splicing	216 / 2					124 / 2		159 / 3
AT5G06210	AT5G06210	GR RNA binding	Splicing								
AT5G61030	AT5G61030	GR RNA binding	Splicing								
AT1G03457	AT1G03457	hnRNP family	Splicing								

Table S23 (Continuation) List of expected interactors
Proteins that co-purified with ADA2b, TAF13, TAf10, UBP22, ENY2, SGF11, MOS4, and THP1 were identified by mass spectrometry and compared to a list of expected interactors to get more information about their association to protein complexes. Proteins that co-purified with the unfused SG-tag were removed. The average MASCOT score of the identified proteins is shown and how many times the protein was detected out of three independent affinity purifications.

TAIR	Protein	Complex	Process	ADA2b	TAF13	TAF10	UBP22	ENY2	SGF11	MOS4	THP1
AT1G11650	At1g11650	hnRNP family	Splicing								
AT1G17370	At1g17370	hnRNP family	Splicing								
AT1G17640	At1g17640	hnRNP family	Splicing								
AT1G47490	At1g47490	hnRNP family	Splicing								
AT1G47500	At1g47500	hnRNP family	Splicing								120 / 3
AT1G49600	At1g49600	hnRNP family	Splicing								
AT1G50300	At1g50300	hnRNP family	Splicing								
AT1G54080	At1g54080	hnRNP family	Splicing								197 / 3
AT1G58470	At1g58470	hnRNP family	Splicing								
AT1G60650	At1g60650	hnRNP family	Splicing	189 / 2	195 / 3		243 / 3	137 / 3			
AT2G18830	AT2G18830	hnRNP family	Splicing								
AT2G19380	At2g19380	hnRNP family	Splicing								
AT2G22090	At2g22090	hnRNP family	Splicing				125 / 2				
AT2G22100	At2g22100	hnRNP family	Splicing								
AT2G33410	AT2G33410	hnRNP family	Splicing								
AT2G41060	AT2G41060	hnRNP family	Splicing								
AT2G44710	At2g44710	hnRNP family	Splicing	234 / 2					415 / 3	413 / 2	276 / 3
AT2G47310	At2g47310	hnRNP family	Splicing								
AT3G04610	At3g04610	hnRNP family	Splicing	91 / 2	150 / 2						
AT3G07810	AT3G07810	hnRNP family	Splicing								417 / 3
AT3G13224	At3g13224	hnRNP family	Splicing						108 / 2		326 / 3
AT3G14100	At3g14100	hnRNP family	Splicing								
AT3G15010	At3g15010	hnRNP family	Splicing				96 / 2		190 / 2		563 / 3
AT3G19130	At3g19130	hnRNP family	Splicing				188 / 2				305 / 2
AT3G20890	At3g20890	hnRNP family	Splicing								96 / 2
AT3G26420	At3g26420	hnRNP family	Splicing	201 / 3	253 / 2		343 / 2	272 / 3	166 / 3	278 / 2	309 / 2
AT3G52660	At3g52660	hnRNP family	Splicing								
AT3G56860	At3g56860	hnRNP family	Splicing								
AT4G00830	AT4G00830	hnRNP family	Splicing								
AT4G03110	At4g03110	hnRNP family	Splicing								
AT4G14300	AT4G14300	hnRNP family	Splicing								
AT4G16280	At4g16280	hnRNP family	Splicing								
AT4G16830	At4g16830	hnRNP family	Splicing					174 / 2	252 / 2		273 / 3
AT4G17520	At4g17520	hnRNP family	Splicing		191 / 2		298 / 2	278 / 3	798 / 3		558 / 3
AT4G26000	AT4G26000	hnRNP family	Splicing								
AT4G26650	At4g26650	hnRNP family	Splicing								189 / 2
AT4G27000	At4g27000	hnRNP family	Splicing								205 / 2
AT5G04280	AT5G04280	hnRNP family	Splicing	139 / 2						210 / 2	162 / 2
AT5G28390	AT5G28390	hnRNP family	Splicing								
AT5G40490	At5g40490	hnRNP family	Splicing								357 / 2
AT5G46840	AT5G46840	hnRNP family	Splicing								
AT5G47210	At5g47210	hnRNP family	Splicing	96 / 2			260 / 3	318 / 3	724 / 3		767 / 3
AT5G47620	At5g47620	hnRNP family	Splicing								
AT5G54900	At5g54900	hnRNP family	Splicing				137 / 2		135 / 2		624 / 3
AT5G55550	At5g55550	hnRNP family	Splicing								
AT5G58470	AT5G58470	hnRNP family	Splicing						174 / 2		119 / 2
AT5G66010	At5g66010	hnRNP family	Splicing								
AT5G19350	AT5G19350	hnRNP family	Splicing								
AT5G19350	At5g19350	hnRNP family	Splicing								
AT1G19120	atLSM1a	Lsm core proteins	Splicing								
AT3G14080	atLSM1b	Lsm core proteins	Splicing								
AT1G03330	atLSM2	Lsm core proteins	Splicing				114 / 2		115 / 2	319 / 2	
AT1G21190	atLSM3a	Lsm core proteins	Splicing					88 / 2		130 / 2	
AT1G76860	atLSM3b	Lsm core proteins	Splicing					85 / 2	105 / 2	138 / 2	
AT5G27720	atLSM4	Lsm core proteins	Splicing							137 / 2	
AT5G48870	atLSM5	Lsm core proteins	Splicing								
AT3G59810	atLSM6a	Lsm core proteins	Splicing							123 / 2	
AT2G43810	atLSM6b	Lsm core proteins	Splicing						94 / 2	175 / 2	
AT2G03870	atLSM7	Lsm core proteins	Splicing					97 / 2		144 / 2	128 / 2
AT1G65700	atLSM8	Lsm core proteins	Splicing							103 / 2	
AT1G01350	AT1G01350	NTC-associated	Splicing								
AT1G07360	AT1G07360	NTC-associated	Splicing	288 / 3	447 / 3	309 / 3	547 / 3	315 / 3	299 / 3	592 / 3	
AT1G10580	AT1G10580	NTC-associated	Splicing	162 / 2			265 / 3	155 / 2		608 / 2	111 / 2
AT1G25682	AT1G25682	NTC-associated	Splicing								
AT1G32490	AT1G32490	NTC-associated	Splicing	98 / 2			279 / 2	221 / 2	97 / 2	925 / 3	107 / 2
AT1G33520	AT1G33520	NTC-associated	Splicing								
AT1G56290	AT1G56290	NTC-associated	Splicing	258 / 3	341 / 3	92 / 2	475 / 3			581 / 2	
AT1G80930	AT1G80930	NTC-associated	Splicing					385 / 3		862 / 3	
AT2G16600	At2g16600	NTC-associated	Splicing	312 / 3			213 / 3				
AT2G21130	At2g21130	NTC-associated	Splicing								
AT2G29580	AT2G29580	NTC-associated	Splicing					178 / 3	126 / 2	376 / 2	151 / 2
AT2G35340	AT2G35340	NTC-associated	Splicing								
AT2G36130	AT2G36130	NTC-associated	Splicing	168 / 3			208 / 3	280 / 3	137 / 2	1279 / 2	213 / 2
AT2G38770	AT2G38770	NTC-associated	Splicing	1238 / 3	1231 / 3	331 / 2	1842 / 3	1072 / 3	554 / 3	2481 / 3	468 / 3
AT2G41020	At2g41020	NTC-associated	Splicing								
AT3G02710	AT3G02710	NTC-associated	Splicing							912 / 2	
AT3G05070	AT3G05070	NTC-associated	Splicing					191 / 2		241 / 2	
AT3G12580	At3g12580	NTC-associated	Splicing	3187 / 3	3821 / 3	2608 / 3	3052 / 3	1202 / 3	2749 / 3	1376 / 3	1319 / 3
AT3G13200	AT3G13200	NTC-associated	Splicing					142 / 2		297 / 2	175 / 2
AT3G19840	AT3G19840	NTC-associated	Splicing							222 / 2	
AT3G23900	At3g23900	NTC-associated	Splicing							247 / 2	
AT3G29390	AT3G29390	NTC-associated	Splicing								
AT4G16680	At4g16680	NTC-associated	Splicing								
AT4G25020	AT4G25020	NTC-associated	Splicing								
AT4G33060	AT4G33060	NTC-associated	Splicing				94 / 2				
AT4G34870	At4g34870	NTC-associated	Splicing	265 / 2						128 / 2	
AT4G38740	At4g38740	NTC-associated	Splicing								
AT5G02490	AT5G02490	NTC-associated	Splicing	3027 / 3	3342 / 3	2752 / 3	2795 / 3	772 / 3	2207 / 3	1468 / 3	1266 / 3
AT5G02500	AT5G02500	NTC-associated	Splicing	4218 / 3	4789 / 3	3751 / 3	3954 / 3	1763 / 3	3984 / 3	2170 / 3	2440 / 3
AT5G06420	AT5G06420	NTC-associated	Splicing								
AT5G07060	AT5G07060	NTC-associated	Splicing								
AT5G23590	AT5G23590	NTC-associated	Splicing	123 / 2							
AT5G54520	At5g54520	NTC-associated	Splicing								
AT5G56900	AT5G56900	NTC-associated	Splicing							452 / 2	
AT4G21110	AtBud31	NTC-associated	Splicing	178 / 3			177 / 3			190 / 2	
AT2G16860	SYF2	NTC-associated	Splicing	145 / 2			113 / 2	108 / 2			
AT1G13350	AT1G13350	Recruited to Bact	Splicing								
AT1G55150	AT1G55150	Recruited to Bact	Splicing								
AT1G55460	AT1G55460	Recruited to Bact	Splicing								
AT1G59760	AT1G59760	Recruited to Bact	Splicing								

Table S23 (Continuation) List of expected interactors

Proteins that co-purified with ADA2b, TAF13, TAf10, UBP22, ENY2, SGF11, MOS4, and THP1 were identified by mass spectrometry and compared to a list of expected interactors to get more information about their association to protein complexes. Proteins that co-purified with the unfused SG-tag were removed. The average MASCOT score of the identified proteins is shown and how many times the protein was detected out of three independent affinity purifications.

TAIR	Protein	Complex	Process	ADA2b	TAF13	TAF10	UBP22	ENY2	SGF11	MOS4	THP1
AT1G73720	AT1G73720	Recruited to Bact	Splicing	248 / 3	134 / 2		479 / 3				
AT2G16940	AT2G16940	Recruited to Bact	Splicing	688 / 3	797 / 3		978 / 3	130 / 2	186 / 3	548 / 2	286 / 3
AT3G25840	AT3G25840	Recruited to Bact	Splicing							184 / 2	
AT3G53640	AT3G53640	Recruited to Bact	Splicing								
AT5G09880	AT5G09880	Recruited to Bact	Splicing	200 / 2	155 / 3		241 / 3				123 / 2
AT1G31870	AT1G31870	RES complex	Splicing							292 / 2	
AT3G20550	AT3G20550	RES complex	Splicing				128 / 2		171 / 3		
AT1G03140	AT1G03140	RES, 2nd step	Splicing				232 / 3				
AT1G26370	At1g26370	RES, 2nd step	Splicing		114 / 2						
AT1G27900	At1g27900	RES, 2nd step	Splicing							135 / 2	
AT1G54590	At1g54590	RES, 2nd step	Splicing								
AT1G65660	AT1G65660	RES, 2nd step	Splicing								
AT2G44200	AT2G44200	RES, 2nd step	Splicing		193 / 2		156 / 2				
AT3G26560	AT3G26560	RES, 2nd step	Splicing	669 / 2	383 / 2		503 / 2	469 / 3		1431 / 3	408 / 2
AT3G45950	At3g45950	RES, 2nd step	Splicing								
AT4G37120	AT4G37120	RES, 2nd step	Splicing				161 / 3				
AT5G13010	AT5G13010	RES, 2nd step	Splicing	828 / 3	1166 / 3		968 / 2	104 / 3	279 / 3	601 / 2	255 / 2
AT1G17070	AT1G17070	RES, Disassembly	Splicing		145 / 2		240 / 2			728 / 3	92 / 2
AT2G42330	AT2G42330	RES, Disassembly	Splicing								
AT2G47250	AT2G47250	RES, Disassembly	Splicing		300 / 2		766 / 3		147 / 3	704 / 3	
AT3G62310	AT3G62310	RES, Disassembly	Splicing	427 / 3	374 / 3		942 / 3	252 / 2	221 / 3	819 / 3	
AT4G31770	AT4G31770	RES, Disassembly	Splicing								
AT5G14900	At5g14900	RES, Disassembly	Splicing								
AT2G14285	AT2G14285	Sm core proteins	Splicing								
AT5G44500	atSmB-a	Sm core proteins	Splicing	459 / 3	422 / 2	159 / 2	405 / 3	173 / 3	397 / 3	276 / 3	280 / 3
AT4G20440	atSmB-b	Sm core proteins	Splicing					323 / 2			
AT3G07590	atSmD1-a	Sm core proteins	Splicing	391 / 3			367 / 3	524 / 3	439 / 2	699 / 3	350 / 3
AT4G02840	atSmD1-b	Sm core proteins	Splicing	424 / 3	227 / 2		403 / 3	565 / 2	398 / 2	691 / 3	345 / 3
AT2G47640	atSmD2-a	Sm core proteins	Splicing	222 / 2			507 / 2				
AT3G62840	atSmD2-b	Sm core proteins	Splicing					238 / 3	276 / 3	660 / 2	312 / 2
AT1G76300	atSmD3-a	Sm core proteins	Splicing	198 / 3			204 / 3	244 / 3	276 / 2	430 / 3	192 / 2
AT1G20580	atSmD3-b	Sm core proteins	Splicing	151 / 2			172 / 2	298 / 2	233 / 2	448 / 2	
AT4G30330	AtSmE-a	Sm core proteins	Splicing					424 / 2	288 / 2	519 / 3	
AT4G30220	atSmF	Sm core proteins	Splicing								
AT2G23930	atSmG-a	Sm core proteins	Splicing	188 / 2			145 / 2	208 / 2	154 / 2	557 / 2	188 / 3
AT3G11500	atSmG-b	Sm core proteins	Splicing							133 / 2	
AT2G18740	atSmE-b	Sm core proteins	Splicing	156 / 2			107 / 2	331 / 3	341 / 2	558 / 3	285 / 3
AT2G18740	atSmE-b	Sm core proteins	Splicing	156 / 2			107 / 2	331 / 3	341 / 2	558 / 3	285 / 3
AT5G53180	At5g53180	Splice site selec.	Splicing								
AT1G43190	atPTB1	Splice site selec.	Splicing								
AT3G01150	atPTB2a	Splice site selec.	Splicing								
AT5G51300	atSF1	Splice site selec.	Splicing								
AT5G42820	atU2AF35	Splice site selec.	Splicing		155 / 2		93 / 2		150 / 2		222 / 2
AT1G27650	atU2AF35a	Splice site selec.	Splicing	125 / 2	218 / 2		115 / 3		125 / 2		
AT4G36690	atU2AF65a	Splice site selec.	Splicing	234 / 3	433 / 3		357 / 3		127 / 2	219 / 3	314 / 2
AT1G60900	atU2AF65b	Splice site selec.	Splicing	116 / 2	93 / 2		161 / 3	122 / 2	180 / 2	297 / 3	274 / 2
AT2G33440	atULrp	Splice site selec.	Splicing								
AT1G10320	atUrp	Splice site selec.	Splicing								
AT1G60830	AUL3p	Splice site selec.	Splicing								
AT2G17530	AT2G17530	SR protein kinase	Splicing								
AT3G44850	AT3G44850	SR protein kinase	Splicing	279 / 3	107 / 2		203 / 3	121 / 2			
AT3G53030	AT3G53030	SR protein kinase	Splicing								
AT3G53570	At3g53570	SR protein kinase	Splicing								
AT4G24740	AT4G24740	SR protein kinase	Splicing								
AT4G32660	AT4G32660	SR protein kinase	Splicing								
AT4G35500	At4g35500	SR protein kinase	Splicing								
AT5G22840	AT5G22840	SR protein kinase	Splicing	153 / 2			92 / 2				
AT1G02840	AT1G02840	SR proteins	Splicing								
AT1G07350	At1g07350	SR proteins	Splicing								
AT1G09140	AT1G09140	SR proteins	Splicing								
AT1G23860	At1g23860	SR proteins	Splicing	140 / 3			134 / 3	204 / 3	124 / 3		201 / 2
AT1G55310	AT1G55310	SR proteins	Splicing								
AT2G24590	AT2G24590	SR proteins	Splicing				131 / 2	264 / 3	262 / 3	303 / 2	302 / 2
AT2G37340	At2g37340	SR proteins	Splicing								
AT2G46610	At2g46610	SR proteins	Splicing		131 / 2		134 / 3	348 / 3	166 / 3	486 / 2	212 / 3
AT3G13570	AT3G13570	SR proteins	Splicing								
AT3G49430	AT3G49430	SR proteins	Splicing	110 / 2	92 / 2		201 / 2	225 / 2		148 / 3	
AT3G53500	At3g53500	SR proteins	Splicing								
AT3G55460	At3g55460	SR proteins	Splicing				168 / 2	215 / 3	196 / 3	225 / 3	181 / 2
AT3G61860	At3g61860	SR proteins	Splicing	129 / 3	242 / 2		138 / 2		250 / 3		268 / 3
AT4G02430	AT4G02430	SR proteins	Splicing								
AT4G25500	AT4G25500	SR proteins	Splicing								
AT4G31580	AT4G31580	SR proteins	Splicing	159 / 3	347 / 2		189 / 3	424 / 2	391 / 3	289 / 3	343 / 3
AT4G35785	AT4G35785	SR proteins	Splicing								
AT5G18810	At5g18810	SR proteins	Splicing					163 / 3		148 / 2	119 / 2
AT5G52040	At5g52040	SR proteins	Splicing					114 / 2	127 / 2		
AT5G64200	At5g64200	SR proteins	Splicing						122 / 2		
AT1G10890	AT1G10890	SR-related proteins	Splicing								
AT2G29210	AT2G29210	SR-related proteins	Splicing							158 / 2	
AT5G13340	AT5G13340	SR-related proteins	Splicing								
AT5G22330	AT5G22330	SWR1/NuA4, INO80	Splicing	1060 / 2	926 / 3	762 / 3	1293 / 3	220 / 3	612 / 3	204 / 2	398 / 2
AT3G03340	atLuc7a	U1 snRNP	Splicing				239 / 3				
AT5G17440	atLuc7b	U1 snRNP	Splicing		160 / 2		399 / 3				
AT5G51410	atLuc7-rl	U1 snRNP	Splicing				218 / 3				
AT1G04080	atPrp39a	U1 snRNP	Splicing	315 / 2	300 / 3		268 / 3				
AT5G46400	atPrp39b	U1 snRNP	Splicing								
AT3G50670	atU1-70K	U1 snRNP	Splicing	182 / 3	293 / 3		233 / 3	166 / 3	210 / 3		194 / 3
AT2G47580	atU1A	U1 snRNP	Splicing	248 / 3			158 / 3	342 / 3	472 / 2	364 / 2	511 / 2
AT4G03120	atU1C	U1 snRNP	Splicing	108 / 3							
AT1G09230	At1g09230	U11/U12 specific	Splicing								
AT2G43370	At2g43370	U11/U12 specific	Splicing								
AT3G07860	At3g07860	U11/U12 specific	Splicing								
AT3G10400	At3g10400	U11/U12 specific	Splicing								
AT1G60200	AT1G60200	U1snRNP related	Splicing		454 / 3		269 / 3			207 / 2	148 / 2
AT1G44910	atPRP40A	U1snRNP related	Splicing		256 / 2					313 / 2	113 / 2
AT3G19670	atPRP40B	U1snRNP related	Splicing								
AT2G41500	atSAP60	U4/U6 snRNP	Splicing	700 / 3	819 / 2	200 / 3	1150 / 3		115 / 2	249 / 2	
AT1G28060	atSAP90-1	U4/U6 snRNP	Splicing		494 / 2	149 / 2	652 / 3			436 / 2	
AT3G55930	atSAP90-2	U4/U6 snRNP	Splicing								
AT3G56790	atSAP90-3	U4/U6 snRNP	Splicing								

Table S23 (Continuation) List of expected interactors
Proteins that co-purified with ADA2b, TAF13, TAf10, UBP22, ENY2, SGF11, MOS4, and THP1 were identified by mass spectrometry and compared to a list of expected interactors to get more information about their association to protein complexes. Proteins that co-purified with the unfused SG-tag were removed. The average MASCOT score of the identified proteins is shown and how many times the protein was detected out of three independent affinity purifications.

TAIR	Protein	Complex	Process	ADA2b	TAF13	TAF10	UBP22	ENY2	SGF11	MOS4	THP1
AT2G38730	atTri-20	U4/U6 snRNP	Splicing	520 / 3		236 / 2	462 / 3	294 / 2	229 / 2	415 / 2	193 / 2
AT5G20160	atU4/U6-15.5a	U4/U6 snRNP	Splicing						169 / 2	108 / 2	127 / 3
AT4G12600	atU4/U6-15.5b	U4/U6 snRNP	Splicing								
AT4G22380	atU4/U6-15.5c	U4/U6 snRNP	Splicing								
AT4G24270	PRP24, EMB140	U4/U6 snRNP	Splicing								
AT3G56070	ROC2	U4/U6 snRNP	Splicing								
AT1G60170	U4/U6-61a1	U4/U6 snRNP	Splicing	278 / 3	207 / 3	166 / 3	470 / 3	97 / 2			
AT3G60610	U4/U6-61b	U4/U6 snRNP	Splicing								
AT2G40650	AT2G40650	U4/U6.U5	Splicing		238 / 3		329 / 3				
AT3G05760	AT3G05760	U4/U6.U5	Splicing	247 / 3	191 / 2		196 / 2				
AT4G22285	AT4G22285	U4/U6.U5	Splicing								
AT5G37370	ATSRL1	U4/U6.U5	Splicing								
AT5G57370	atTri-27 kD	U4/U6.U5	Splicing						132 / 2		
AT4G22350	atTri65a	U4/U6.U5	Splicing								
AT4G22290	atTri65b	U4/U6.U5 tri snRNP	Splicing								
AT4G22410	atTri65c	U4/U6.U5	Splicing								
AT5G16780	DOT2	U4/U6.U5	Splicing	2012 / 3	1710 / 3	626 / 3	2379 / 3			443 / 2	
AT2G33730	atU5-100KD	U5 snRNP	Splicing	736 / 3	726 / 3		810 / 3			381 / 2	
AT4G03430	atU5-102KD	U5 snRNP	Splicing	629 / 3	1518 / 3	309 / 2	724 / 3	190 / 3	695 / 3	781 / 2	
AT1G06220	atU5-116-1a	U5 snRNP	Splicing	2427 / 3	2695 / 3	1577 / 3	1910 / 3	2214 / 3	2982 / 3	3734 / 3	1174 / 3
AT5G25230	atU5-116-1b	U5 snRNP	Splicing								
AT1G56070	atU5-116-2	U5 snRNP	Splicing	1829 / 3	2536 / 3	896 / 3	1540 / 3	831 / 3	1264 / 3	665 / 2	1147 / 3
AT3G22980	atU5-116-3	U5 snRNP	Splicing								
AT5G08290	atU5-15	U5 snRNP	Splicing							119 / 2	
AT5G61140	atU5-200-1	U5 snRNP	Splicing		448 / 3		200 / 3		453 / 2		
AT1G20960	atU5-200-2a	U5 snRNP	Splicing	1829 / 3	3728 / 3	307 / 3	4541 / 3	3102 / 3	1650 / 3	5413 / 3	1134 / 3
AT2G42270	atU5-200-2b	U5 snRNP	Splicing								
AT3G27730	atU5-200-3	U5 snRNP	Splicing								
AT1G80070	atU5-220/Prp8a	U5 snRNP	Splicing	1443 / 3	4151 / 3	284 / 2	3938 / 3	2275 / 3	1580 / 3	4227 / 3	1649 / 2
AT4G38780	atU5-220/Prp8b	U5 snRNP	Splicing		3359 / 2			176 / 3			
AT2G43770	atU5-40	U5 snRNP	Splicing	545 / 2	308 / 3	166 / 3	742 / 3	172 / 3	284 / 2	423 / 2	300 / 2
AT5G09390	CD2 antigen	U5 snRNP	Splicing		162 / 2	151 / 2					
AT1G49760	AT1G49760	mRNA binding	Polyadenylation	153 / 2			125 / 2	209 / 3	474 / 3	844 / 2	417 / 2
AT2G23350	AT2G23350	mRNA binding	Polyadenylation	252 / 2	163 / 3		234 / 2		274 / 3	502 / 2	360 / 3
AT2G27100	AT2G27100	mRNA binding	Polyadenylation	900 / 3	1237 / 3	640 / 3	1270 / 3	713 / 3	1179 / 3	1315 / 2	670 / 3
AT4G32720	AT4G32720	mRNA binding	Polyadenylation						130 / 2		303 / 2
AT4G34110	AT4G34110	mRNA binding	Polyadenylation						134 / 2		207 / 2
AT5G63120	AT5G63120	mRNA binding	Polyadenylation	650 / 3	657 / 3	377 / 3	804 / 3	122 / 3	608 / 2		348 / 3
AT5G58040	FIPS5	Polyadenylation	Polyadenylation				183 / 2			223 / 2	
AT4G29820	CFIS1	Polyadenylation	Polyadenylation								
AT4G25550	CFIS2	Polyadenylation	Polyadenylation				123 / 2	102 / 3	127 / 2	225 / 2	150 / 2
AT3G04680	CLPS3	Polyadenylation	Polyadenylation								
AT5G39930	CLPS5	Polyadenylation	Polyadenylation								
AT5G23880	CPSF100	Polyadenylation	Polyadenylation	307 / 2	224 / 3		228 / 2		126 / 2		
AT5G51660	CPSF160	Polyadenylation	Polyadenylation	307 / 3	641 / 3		430 / 3		593 / 2		
AT1G30460	CPSF30,	Polyadenylation	Polyadenylation								
AT1G61010	CPSF73-I	Polyadenylation	Polyadenylation		223 / 2		130 / 2				
AT2G01730	CPSF73-II, FEG	Polyadenylation	Polyadenylation								
AT5G60940	CSTF50	Polyadenylation	Polyadenylation								
AT1G71800	CSTF64	Polyadenylation	Polyadenylation								
AT1G17760	CSTF77	Polyadenylation	Polyadenylation								
AT5G01400	ESP4, SYM5	Polyadenylation	Polyadenylation		214 / 3		133 / 2		188 / 3		
AT3G66652	FIPS3	Polyadenylation	Polyadenylation								
AG5G58040	FIPS5	Polyadenylation	Polyadenylation								
AT5G13480	FY	Polyadenylation	Polyadenylation		324 / 2	88 / 2			139 / 3		
AT5G65260	PABN1	Polyadenylation	Polyadenylation								
AT5G51120	PABN2	Polyadenylation	Polyadenylation						127 / 2		
AT5G10350	PABN3	Polyadenylation	Polyadenylation								
AT1G17980	PAPS1	Polyadenylation	Polyadenylation								
AT2G25850	PAPS2	Polyadenylation	Polyadenylation								
AT3G06560	PAPS3	Polyadenylation	Polyadenylation								
AT4G32850	PAPS4	Polyadenylation	Polyadenylation								
AT2G31320	PARP2	Polyadenylation	Polyadenylation								
AT1G66500	PCFS1	Polyadenylation	Polyadenylation								
AT4G04885	PCFS4	Polyadenylation	Polyadenylation								
AT5G43620	PCFS5	Polyadenylation	Polyadenylation								
AT1G27590	SYM1	Polyadenylation	Polyadenylation								
AT1G27595	SYM2	Polyadenylation	Polyadenylation		626 / 3		128 / 2		129 / 2		
AT2G13540	AT2G13540	CAP-binding	Capping								
AT5G44200	AT5G44200	CAP-binding	Capping								
AT5G16310	UCH1	INO80	Export		111 / 2		156 / 2				
AT3G56900	ALADIN	Nucleoporin	Export								
AT2G39810	Eyls/HOS1	Nucleoporin	Export	316 / 3	387 / 3		319 / 3		321 / 3		
AT1G13120	GLE1	Nucleoporin	Export								
AT5G40480	gp210	Nucleoporin	Export						170 / 2		
AT1G75340	NLP1, CGI	Nucleoporin	Export								
AT3G14120	Nup107	Nucleoporin	Export	217 / 2	156 / 3						
AT2G05120	Nup133	Nucleoporin	Export	490 / 3	1044 / 3		418 / 2		436 / 3		389 / 2
AT3G10650	Nup136	Nucleoporin	Export	332 / 3	696 / 3	234 / 2	407 / 3		146 / 3		
AT1G14850	Nup155	Nucleoporin	Export	817 / 3	956 / 3		916 / 2		219 / 3		292 / 2
AT1G33410	Nup160	Nucleoporin	Export	265 / 3	621 / 3		431 / 2		372 / 2		293 / 2
AT5G51200	Nup205	Nucleoporin	Export	210 / 3	689 / 2		264 / 2		203 / 2		
AT1G55540	NUP214	Nucleoporin	Export		107 / 2						
AT3G16310	Nup35	Nucleoporin	Export				119 / 2				
AT4G30840	Nup43	Nucleoporin	Export								
AT1G52380	Nup50a	Nucleoporin	Export								
AT3G15970	Nup50b	Nucleoporin	Export								
AT1G24310	Nup54	Nucleoporin	Export								
AT4G37130	Nup58	Nucleoporin	Export								
AT2G45000	Nup62	Nucleoporin	Export								
AT4G32910	NUP85	Nucleoporin	Export	349 / 2	128 / 3		184 / 2				
AT5G05680	Nup88	Nucleoporin	Export	140 / 2	152 / 2		87 / 2				
AT2G41620	Nup93a	Nucleoporin	Export	186 / 2					110 / 2		
AT3G57350	Nup93b	Nucleoporin	Export								
AT1G80680	Nup96	Nucleoporin	Export								
AT1G10390	Nup98a	Nucleoporin	Export	216 / 2			210 / 2		91 / 2		
AT1G59660	Nup98b	Nucleoporin	Export								
AT1G80670	RAE1	Nucleoporin	Export		152 / 2		201 / 3		108 / 2		
AT2G30050	Sec13	Nucleoporin	Export	543 / 3	422 / 3	448 / 3	323 / 3		360 / 3		283 / 2
AT1G64350	Seh1	Nucleoporin	Export				162 / 3				

Table S23 (Continuation) List of expected interactors
Proteins that co-purified with ADA2b, TAF13, TAf10, UBP22, ENY2, SGF11, MOS4, and THP1 were identified by mass spectrometry and compared to a list of expected interactors to get more information about their association to protein complexes. Proteins that co-purified with the unfused SG-tag were removed. The average MASCOT score of the identified proteins is shown and how many times the protein was detected out of three independent affinity purifications.

TAIR	Protein	Complex	Process	ADA2b	TAF13	TAF10	UBP22	ENY2	SGF11	MOS4	THP1
AT1G79280	Tpr/NUA	Nucleoporin	Export	405 / 2	936 / 2		723 / 3		339 / 3		409 / 2
AT3G09250	NTF2-Like 1	pot. export factor	Export								
AT1G69250	NTF2-Like 10	pot. export factor	Export		97 / 2				202 / 2		
AT1G13730	NTF2-Like 11	pot. export factor	Export								
AT2G03640	NTF2-Like 12	pot. export factor	Export						205 / 3		152 / 2
AT5G43960	NTF2-Like 13	pot. export factor	Export	122 / 2	157 / 2				216 / 3	152 / 2	269 / 2
AT3G25150	NTF2-Like 14	pot. export factor	Export								
AT5G60980	NTF2-Like 15	pot. export factor	Export	232 / 2	196 / 2	137 / 2	197 / 3	150 / 3	388 / 3	164 / 2	311 / 3
AT1G11570	NTF2-Like 16	pot. export factor	Export								
AT1G27310	NTF2-Like 17	pot. export factor	Export								
AT1G27970	NTF2-Like 18	pot. export factor	Export								
AT5G48650	NTF2-Like 19	pot. export factor	Export								207 / 2
AT4G10925	NTF2-Like 2	pot. export factor	Export								
AT2G46100	NTF2-Like 3	pot. export factor	Export								
AT3G04890	NTF2-Like 4	pot. export factor	Export								
AT5G04830	NTF2-Like 5	pot. export factor	Export								
AT1G71480	NTF2-Like 6	pot. export factor	Export								
AT5G41470	NTF2-Like 7	pot. export factor	Export								
AT3G07250	NTF2-Like 8	pot. export factor	Export								
AT3G55540	NTF2-Like 9	pot. export factor	Export								
AT5G02530	atALY-1a	TREX complex	Export	245 / 3	255 / 3		239 / 3	272 / 3	691 / 3	432 / 2	435 / 3
AT5G59950	atALY-1b	TREX complex	Export	251 / 3	197 / 2		291 / 2	244 / 3	466 / 3	211 / 2	549 / 3
AT5G37720	atALY-2a	TREX complex	Export	393 / 3	420 / 3		293 / 3	301 / 3	769 / 3	448 / 2	778 / 3
AT1G66260	atALY-2b	TREX complex	Export	102 / 2			102 / 2	110 / 2	246 / 2		269 / 2
AT5G11200	atUAP56a	TREX complex	Export								
AT5G11170	atUAP56b	TREX complex	Export		255 / 3		194 / 3				368 / 3
AT4G10970	atUIF1	TREX complex	Export	167 / 2							
AT4G23910	atUIF2	TREX complex	Export								
AT2G19430	DWA1	TREX complex	Export								
AT5G56130	TEX1	TREX complex	Export								
AT5G09860	THO1, HPR1	TREX complex	Export							293 / 2	87 / 2
AT1G24706	THO2	TREX complex	Export							625 / 2	229 / 2
AT5G42920	THO5	TREX complex	Export					187 / 2		325 / 2	114 / 2
AT1G45233	THO5	TREX complex	Export								
AT3G02950	THO7	TREX complex	Export					160 / 2			
AT5G16790	THO7	TREX complex	Export								
AT5G02770	atMOS11	TREX complex	Export								
AT2G19560	AtTHP1, ESSP1	TREX-2 complex	Export		104 / 2						2078 / 3
AT3G50360	CEN1	TREX-2 complex	Export								
AT4G37010	CEN2	TREX-2 complex	Export								
AT5G45010	DSS1(V)	TREX-2 complex	Export								
AT2G39340	SAC3A	TREX-2 complex	Export								3571 / 3
AT3G06290	SAC3B	TREX-2 complex	Export								3517 / 3
AT3G54380	SAC3C	TREX-2 complex	Export								1033 / 3

13 Publications

Publications related to this work:

- **Pfab A.**, Antosz W., Holzinger P., Bruckmann A., Griesenbeck J., Grasser KD. (2017) Analysis of In Vivo Chromatin and Protein Interactions of *Arabidopsis* Transcript Elongation Factors. Methods Mol Biol., 1629:105-122

- Antosz W., **Pfab A.**, Ehrnsberger HF., Holzinger P., Köllen K., Mortensen SA., Bruckmann A., Schubert T., Längst G., Griesenbeck J., Schubert V., Grasser M., Grasser KD. (2017). The Composition of the *Arabidopsis* RNA Polymerase II Transcript Elongation Complex Reveals the Interplay between Elongation and mRNA Processing Factors. Plant Cell, 29(4):854-870

- **Pfab A.**, Breindl M., Grasser KD. (2017) The *Arabidopsis* histone chaperone FACT is required for stress-induced expression of anthocyanin biosynthetic genes.
  Manuscript under revision.
  The manuscript is available in this section.

- **Pfab A.**, Bruckmann A., Nazet J., Merkl R., Grasser KD. (2017) The adaptor protein ENY2 is a component of the deubiquitination module of the *Arabidopsis* SAGA (transcriptional co-activator) complex but not of the TREX-2 complex
  In preparation

- **Pfab A.**, Grasser KD. (2017) Role of the SSRP1 HMG-box domain in *Arabidopsis*
  In preparation

Other publications:

- Sörensen BB., Ehrnsberger HF., Esposito S., **Pfab A.**, Bruckmann A., Hauptmann J., Meister G., Merkl R., Schubert T., Längst G., Melzer M., Grasser M., Grasser KD. (2017). The *Arabidopsis* THO/TREX component TEX1 functionally interacts with MOS11 and modulates mRNA export and alternative splicing events. Plant Mol Biol., 93(3):283-298

The *Arabidopsis* histone chaperone FACT is required for stress-induced expression of anthocyanin biosynthetic genes

Alexander Pfab, Matthias Breindl and Klaus D. Grasser*

Cell Biology & Plant Biochemistry, Biochemistry Centre Regensburg,
University of Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany

*To whom correspondence should be addressed:
Tel: +49-941-9433032; Fax: +49-941-9433352; Email: Klaus.Grasser@ur.de

Abstract

The histone chaperone FACT consists of the SSRP1 and SPT16 proteins and associates with transcribing RNAPII (RNAPII) along the transcribed region of genes. FACT can promote transcriptional elongation by destabilising nucleosomes in the path of RNA polymerase II, thereby facilitating efficient transcription of chromatin templates. Transcript profiling of *Arabidopsis* plants depleted in SSRP1 or SPT16 demonstrates that only a small subset of genes is differentially expressed relative to wild type. The majority of these genes is either up- or down-regulated in both the *ssrp1* and *spt16* plants. Among the down-regulated genes, those encoding enzymes of the biosynthetic pathway of the plant secondary metabolites termed anthocyanins (but not regulators of the pathway) are overrepresented. Upon exposure to moderate high-light stress several of these genes are up-regulated to a lesser extent in *ssrp1/spt16* compared to wild type plants, and accordingly the mutant plants accumulate lower amounts of anthocyanin pigments. Moreover, the expression of *SSRP1* and *SPT16* is induced under these conditions. Therefore, our findings indicate that FACT is a novel factor required for the accumulation of anthocyanins in response to light-induction.

Keywords: SSRP1; SPT16; chromatin; histones; gene transcription;

Abbreviations: FACT, facilitates chromatin transcription; SSRP1, structure-specific recognition protein 1; SPT16, Suppressor of Ty 16; RNAPII, RNA polymerase II; GUS, beta-glucuronidase; NRQ, normalised relative quantities;

Introduction

In response to changing environmental conditions plants adapt the expression levels of many mRNAs, which is accomplished at various stages including transcriptional and post-transcriptional mechanisms. Transcript synthesis is regulated by the combinatorial action of transcription factors to adjust the frequency of RNA polymerase II (RNAPII) transcriptional initiation to the cellular requirements [1;2]. In addition to controlling the initiation stage of transcription it became apparent in recent years that also the transcriptional elongation is a dynamic and highly regulated phase of transcript synthesis. Efficient transcript elongation by RNAPII on chromatin templates requires the coordinated action of a variety of so-called transcript elongation factors including modifiers of polymerase activity, histone modifiers and histone chaperones [3-5]. Among these factors is the FACT histone chaperone, a heterodimer consisting of the SSRP1 and SPT16 proteins (Pob3 and SPT16 in yeast), which are encoded by essential genes [6;7]. FACT promotes RNAPII transcript elongation on chromatin templates by destabilising nucleosomes in the path of the enzyme without requirement for ATP and the mechanism is somewhat different in yeast and mammals [8-10]. Interestingly, FACT is also important for the reassembly of nucleosomes following passage of RNAPII, maintaining the original chromatin state and preventing cryptic transcript initiation from within coding sequences [11;12]. Upon down-regulation of SSRP1 or SPT16 in human lung carcinoma (H1299) cells (siRNA mediated knock-down) rather subtle changes of global transcript levels were detected affecting only a small subset of genes [13].

Both SSRP1 and SPT16 are conserved in higher and lower plants, showing a clearly higher degree of similarity to their metazoan orthologues compared to their yeast counterparts [14;15]. In *Arabidopsis* nuclei, FACT along with other transcript elongation factors (including SPT6, TFIIS, SPT4/SPT5, and PAF1-C) assembles preferentially with elongating (Ser2-phosphorylated) RNAPII forming the transcript elongation complex [16]. Accordingly, SSRP1 and SPT16 co-localise to the euchromatin and associate with the transcribed region of genes in a transcription-dependent manner [17]. The analysis of *Arabidopsis* mutant plants (*ssrp1, spt16*) revealed that loss of SSRP1 is lethal, but mutant lines depleted in FACT subunits are viable. These plants are phenotypically similar, displaying various defects in vegetative and reproductive development such as increased number of inflorescences, early bolting, abnormal leaf architecture and reduced seed set [18]. Here, we have systematically analysed mRNA levels in FACT depleted plants relative to wild type, to identify biochemical processes that require FACT activity. Our experiments show that FACT is necessary for proper expression of genes encoding anthocyanin biosynthetic enzymes under normal growth conditions as well as for the stress-induced expression of these genes.

Materials and Methods

Plant material

Arabidopsis thaliana Col-0 plants as well as the T-DNA insertion lines were propagated on soil in a phytochamber at 16 h photoperiods (120 µmol m^{-2} s^{-1}) at 21°C [16;19]. For light stress experiments plants were grown on solid 0.5x MS medium (supplemented with 2% sucrose) in a plant incubator (Poly Klima GmbH) at 21°C and different light intensities (60 - 600 µmol m^{-2} s^{-1}) at 16 h photoperiods. The insertion mutant lines *ssrp1-2* and *spt16-1* used here, were described before [18].

Transcript profiling and qRT-PCR

Total RNA was isolated from the aerial part of the 10-day old seedlings grown on solid MS at 16 h photoperiods (100 µmol m^{-2} s^{-1}) with RNeasy Mini Plant kit (Qiagen) and quality checked with a 2100 Bioanalyser (Agilent Technologies). Sample processing and microarray hybridisation were carried out at the genomics core facility, Center of Excellence for Fluorescent Bioanalytics (KFB, University of Regensburg, Germany; http://www.kfb-regensburg.de/) using the ATH1 chip (Affymetrix) representing 22 800 probe sets. The experimental design comprised three replicates of each genotype, corresponding to one RNA extraction of an independent pool of plants. Raw data (CEL files) were summarised with the robust multi-chip analysis followed by quality control with principal component analysis [20]. Transcripts showing a significantly differential expression were identified by pair-wise comparison using a Student's t-test ($P<0.01$). Microarray data have been deposited in the Gene Expression Omnibus repository (www.ncbi.nlm.nih.gov/geo) under accession number GSE103623. Analysis of significantly overrepresented gene ontology (GO) categories among up- and down-regulated genes ($P<0.05$), was performed using Singular Enrichment Analysis (SEA) of agriGO [21]. Differentially expressed genes (≥2-fold; $P<0.05$) were hierarchically clustered (average linkage; euclidean) and visualised as heatmap using the Heatmapper [22]. For qRT-PCR analyses random hexamer-primed complementary DNA was prepared from total RNA that was used for qPCR with KAPA SYBR FAST Universal reagents (PEQLAB) and a Mastercycler ep realplex2 (Eppendorf) as previously described [23]. Targets were quantified with specific primer pairs (Table S1). The normalised relative quantities (NRQ) were calculated according to [24] using the three reference genes *GAPC, PP2AA3* and *UBI10* [25]. The standard error (SE) of the NRQ was calculated without taking the SE derived of the oligo efficiency determination into consideration. P-values (Student's t-test) were calculated with the software Excel (Microsoft).

Anthocyanin analysis

Extraction and photometric determination of anthocyanins of *Arabidopsis* leaves were performed as previously described [26] with minor modifications. Pools of *Arabidopsis* plants were frozen in liquid nitrogen and homogenised using Tissue Lyser II (Qiagene). To the homogenised plant material extraction buffer (acidic methanol, 1% (v/v) HCl) was added (50 µl/mg) and the samples were incubated for 15 minutes with moderate rotation at 4°C in the dark. The homogenate was clarified by centrifugation and the supernatant was used for photometric measurements. To quantify anthocyanins in the samples the following equation was used: $(A_{530} - 0.25 \times A_{657}) \times M^{-1}[g]$ = relative units of anthocyanin (with A_{530} and A_{657} = absorption at the indicated wavelengths, M = plant fresh weight).

GUS reporter assays

Arabidopsis Col-0 plants harbouring GUS-reporter constructs driven by the *SSRP1*- and *SPT16*-promoters (Table S1) were generated by *Agrobacterium*-mediated transformation as previously described [16;19]. Histochemical GUS staining of *Arabidopsis* plants using staining solution (50 mM $NaHPO_4$, pH 7.2; 0.5 mM $K_3Fe(CN)_6$; 0.5 mM $K_4Fe(CN)_6$; 1% Triton X-100 and 2 mM X-Gluc) was performed as previously described [23;27] and images were taken with an EOS600 camera (Canon).

Results

Transcript profiling of plants depleted in SSRP1 and SPT16

In view of the transcription-related function of FACT, genome-wide transcript profiling was performed to identify possible FACT-dependent alterations in gene expression. RNA isolated from 10-day old *ssrp1*, *spt16* and Col-0 wild type plants was comparatively examined by microarray hybridisation. Analysis of this experiment revealed that 146 genes exhibited significantly differential expression (≥2-fold, P<0.05) in *ssrp1* and/*or spt16* relative to Col-0. Of the differentially expressed genes 43 (0.2%) were down-regulated and 103 (0.5%) were up-regulated, and basically all genes show the same trend of being up- or down-regulated in both mutants (Fig. 1A). When considering only genes that are regulated ≥2-fold in both mutants, 23 (53.5%) were down-regulated in both *ssrp1* and *spt16*, while only 21 (20.4%) were up-regulated in both mutants (Fig. 1B). To gain insight into the biological processes, in which the differentially expressed genes are involved, the data was analysed for gene ontology (GO). The categories termed “response to stimulus/stress” were most prominent among the up-regulated genes (Fig. S1). Since FACT is known to promote transcription by RNAPII, the genes down-regulated in SSRP1/SPT16-depleted plants were analysed in detail. Most strikingly, genes encoding enzymes of the flavonoid biosynthesis [28] were

identified as differentially expressed (Fig. 2). Further analysis demonstrated that genes encoding enzymes of the anthocyanin biosynthesis pathway are particularly overrepresented (Fig.3A). These include genes encoding early (e.g. chalcone synthase, CHS; flavanone 3´-hydroxylase, F´3H) and late biosynthetic enzymes (e.g. dihydroflavonol reductase, DFR; anthocyanidin synthase, ANS) (Fig. 3B). These genes are significantly down-regulated in both the *ssrp1* and *spt16* mutants, whereas regulatory genes of the anthocyanin pathway are expressed at similar levels as in Col-0 (Fig. S2A). The reduced expression of four genes was validated by qRT-PCR and this experiment confirmed the down-regulation of the *CHS, F3´H, DFR* and *ANS* genes in both mutant lines relative to Col-0 (Fig. S2B). In addition to the genes involved in anthocyanin biosynthesis, the microarray experiment revealed reduced transcript levels of some genes encoding bHLH transcription factors (Fig. S3). These bHLH factors were reported to be involved in iron-deficiency responses and uptake in *Arabidopsis* [29;30], and therefore it might be interesting to examine in future also a possible link between FACT and iron homeostasis.

FACT is required for light-induced anthocyanin accumulation

Anthocyanins are plant pigments that belong to the flavonoid-type of secondary plant metabolites and they originate from the phenylpropanoid pathway [28]. Anthocyanin biosynthesis can be induced by a variety of abiotic stresses including UV or high light [31-33]. Under these conditions anthocynanins are suggested to protect the plant photosynthesis machinery from damage [34]. To examine whether plants depleted in FACT are affected in the light-stimulated induction of anthocyanin production, plants were grown under different light conditions. Col-0, *ssrp1* and *spt16* plants were grown for 14 d under moderate high-light (160 µmol m^{-2} s^{-1}) and for comparison under low-light conditions (60 µmol m^{-2} s^{-1}). Under low-light conditions no obvious anthocyanin accumulation is seen, while under moderate high-light the leaves of Col-0 plants show purple colouring indicative of anthocyanins (Fig. 4A). This colouring is clearly less with the *ssrp1* and *spt16* plants. The isolation and spectrometric quantification of anthocyanins confirms these observations. Plants of the three genotypes grown at low-light have similar anthocyanin contents, but under moderate high-light approximately half the amount of anthocyanins accumulates in *ssrp1* and *spt16* plants compared to Col-0 (Fig. 4B). To examine whether the light-induced anthocyanin accumulation is associated with an up-regulation of genes of the biosynthesis pathway, transcript levels of four genes were measured by qRT-PCR. For this purpose RNA was analysed of the three genotypes grown either under low-light or moderate high-light conditions. With the *ssrp1* and *spt16* plants a reduced induction of the *F´3H*, *DFR* and *ANS* transcripts upon moderate high-light treatment relative to Col-0 was observed, while the *CHS* induction was similar with the three genotypes (Fig. 4C). The accumulation of anthocyanins

was additionally examined upon induction by severe high-light stress. 14-d old plants of the three genotypes were grown for three days under severe high-light (600 µmol m^{-2} s^{-1}) and for comparison under low-light (60 µmol m^{-2} s^{-1}). Also this treatment resulted in a decreased anthocyanin accumulation with the *ssrp1* and *spt16* plants compared to Col-0 (Fig. S4). To determine whether perhaps *SSRP1* and *SPT16* are regulated by the light treatment, the transcript level of the two genes was measured comparatively by qRT-PCR under low-light and moderate high-light conditions. The transcript levels of both *SSRP1* and *SPT16* are slightly increased at moderate high-light (Fig. 5A). In addition, transgenic Col-0 plants harbouring the *GUS* reporter gene under control of the *SSRP1* or *SPT16* promoters were examined under the same light conditions. Histochemical analysis of 14-d old plants demonstrated higher reporter gene activity with plants grown under moderate high-light compared to low-light (Fig. 5B). Due to the accumulation over time of the GUS protein more of the blue chloro-bromoindigo cleavage product of the beta-glucuronidase reaction with the X-Gluc substrate might be produced [27], and therefore the difference between the two light conditions appears rather strong. Still both the transcript levels measured by qRT-PCR and the promoter-GUS reporter assays show an increased expression of *SSRP1* and *SPT16* upon moderate high-light treatment.

Discussion

The mammalian histone chaperone FACT can facilitate RNAPII transcription on nucleosomal templates [8;35], but it appears to be critical for efficient transcription of only a small subset of genes in human cells [13]. Differential expression of a relatively small fraction of genes is also observed in *Arabidopsis* plants deficient in various transcript elongation factors such as TFIIS, SPT4 and Elongator [23;36;37]. In line with that our transcript profiling analysis of FACT depleted *Arabidopsis* plants revealed also only minor changes of the transcriptome. One should keep in mind that in both the experiments with human cells [13] and *Arabidopsis* plants, because the genes encoding FACT are essential, reduced expression of FACT rather than loss-of-function samples were analysed. Therefore, it is possible that the results of these experiments underestimate the number of FACT-dependent genes. In human cells, the vast majority of genes were commonly regulated by SSRP1 and SPT16, although very few genes appeared to be regulated by SSRP1 independent from SPT16 [13]. Similarly, with the genes differentially expressed in *Arabidopsis* a marked trend was observed that the genes were either up- or down-regulated both in *ssrp1* and *spt16* plants. This trend was even more prominent with the down-regulated genes, which one would expect to comprise genes directly affected by SSRP1 and SPT16 acting as a heterodimer that promotes chromatin transcription [6;7]. Strikingly, genes encoding anthocyanin biosynthesis enzymes are overrepresented among the genes down-regulated in *Arabidopsis* upon FACT depletion

under normal growth conditions, while the expression of regulatory genes of the anthocyanin pathway is not affected. Moreover, compared to the wild type control the induction of these genes by moderate high-light is decreased and consistently the accumulation of anthocyanins is reduced in the *ssrp1* and *spt16* plants. Since the expression of the *SSRP1* and *SPT16* genes is up-regulated by moderate high-light treatment, FACT contributes to the induced transcription of genes encoding anthocyanin biosynthetic enzymes. Therefore, the FACT histone chaperone is a novel factor involved in the light-stimulated accumulation of anthocyanin pigments in *Arabidopsis*.

In the yeast genome, SPT16 is detected along the transcribed region of all genes with ongoing RNAPII transcription [38]. Similarly, FACT was detected in a transcription-dependent manner within the transcribed region of all tested *Arabidopsis* genes that comprised house-keeping as well as inducible genes [16-18;39]. This is likely due to the association of FACT and other factors with RNAPII in the transcript elongation complex [16]. It remains obscure, why the transcription of only a small subset of genes such as those encoding anthocyanin biosynthetic enzymes is affected upon FACT depletion. Likewise, only a minor fraction of genes appears to be particularly sensitive to the depletion/loss of other transcript elongation factors, while the majority of genes is normally expressed [5]. Currently, it is rather unclear which feature(s) of genes (e.g. sequence, expression level, chromatin structure, co-transcriptional mRNA processing) determine that certain transcript elongation factors are required for efficient transcription, which will be a focus of future research.

Acknowledgements

We thank Thomas Stempfl for help with the microarray analyses. This work was supported by the German Research Foundation (DFG) through Grant Gr1159/14-1 and the EC Research Training Network CHIP-ET, FP7-PEOPLE-2013-ITN607880) to K.D.G.

Author contributions

KDG conceived and supervised the study; AP and KDG designed experiments; AP and MB performed experiments; all authors analysed data; KDG wrote the manuscript; all authors reviewed and edited the manuscript.

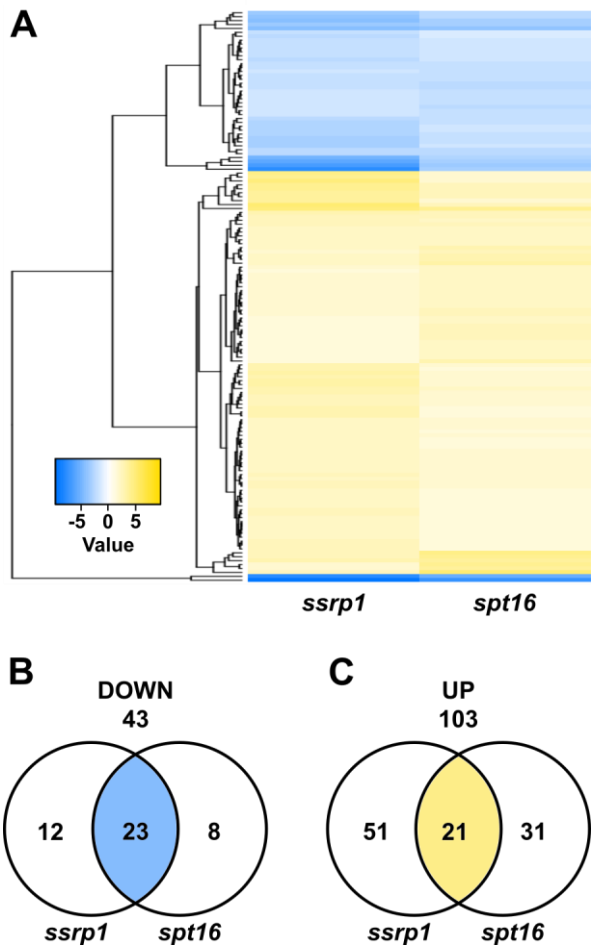


Fig. 1. A relatively small number of genes is differentially expressed in *ssrp1* and *spt16* plants relative to Col-0. A) Transcript levels were determined by comparative microarray hybridisation analysis of RNA isolated from *ssrp1* and *spt16* relative to Col-0 plants. Genes differentially expressed in *ssrp1* and/or *spt16* (≥2-fold, P<0.05) were hierarchically clustered and visualised as heatmap. In total, 146 genes are included represented as horizontal lines. As indicated different shades of yellow/blue represent up- or down-regulated genes, respectively. B,C, Venn diagrams show the fraction of genes up- or down-regulated ≥2-fold in both mutants.

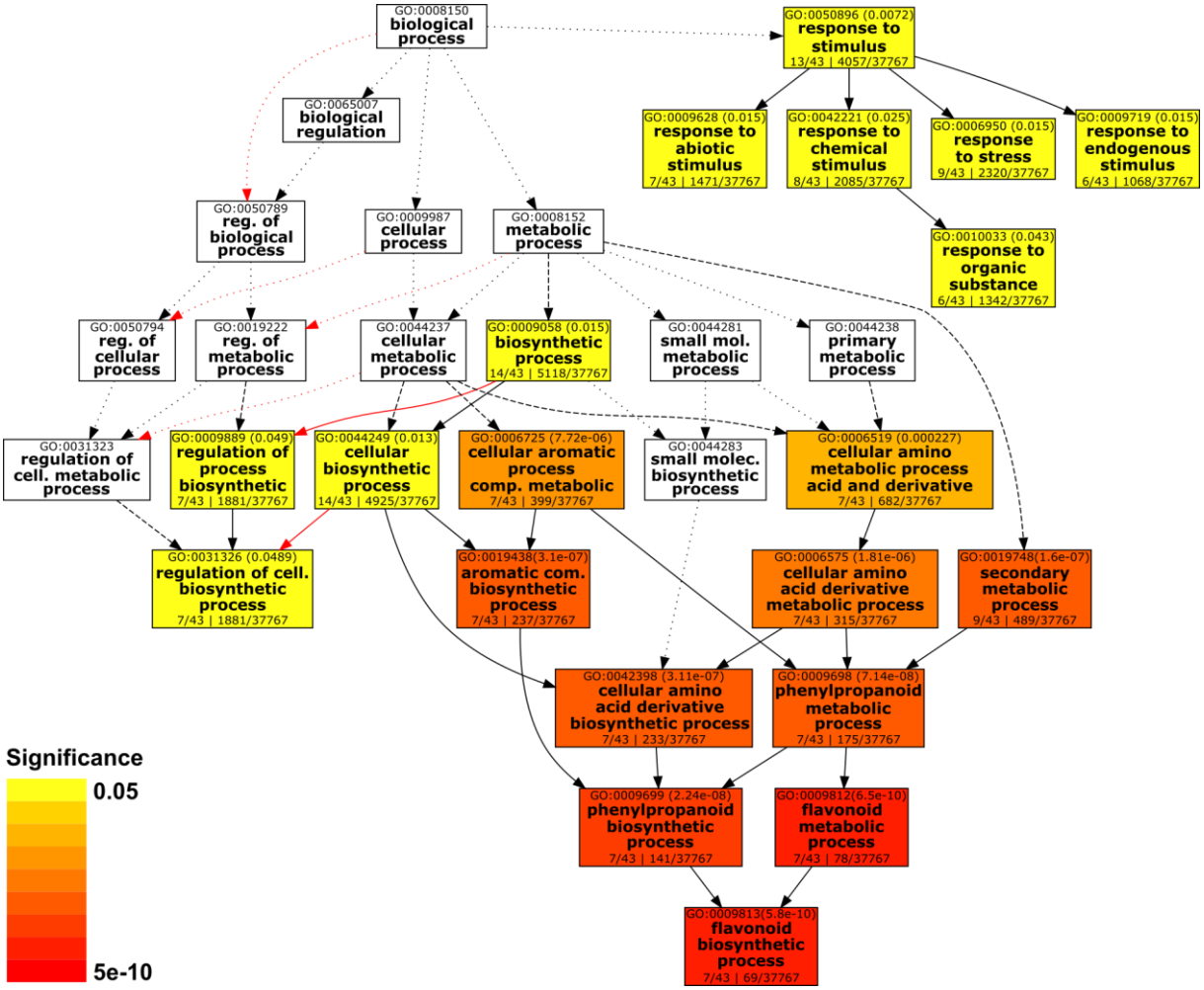


Fig. 2. Biological processes identified by an overrepresented number of genes down-regulated in *ssrp1* and/or *spt16*. Overrepresented GO terms among the significantly down-regulated genes (≥2-fold, P<0.05) in at least one of the two mutants was performed using agriGO [21]. The significance of overrepresented GO terms is colour-coded as indicated.

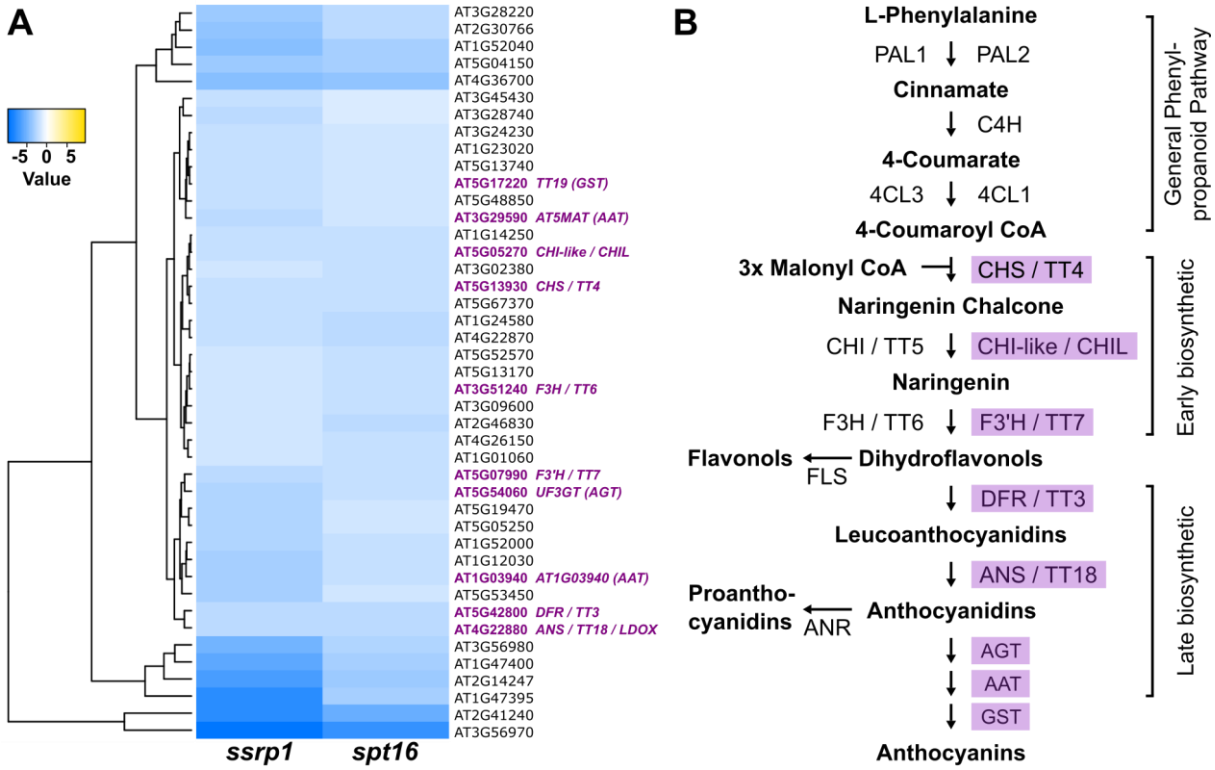


Fig. 3. Anthocyanin biosynthetic genes are prominently down-regulated in *ssrp1*/spt16. A, Genes down-regulated (≥2-fold, P<0.05) in *ssrp1* and/or *spt16* were hierarchically clustered and visualised as heatmap. Different shades of blue represent the reduced expression level of the genes as indicated. Anthocyanin biosynthetic genes are depicted in purple. B, Schematic illustration of anthocyanin biosynthetic pathway (modified from [40]). Genes with significantly reduced expression (≥2-fold, P<0.05) in *ssrp1-2* and *spt16* are highlighted in purple.

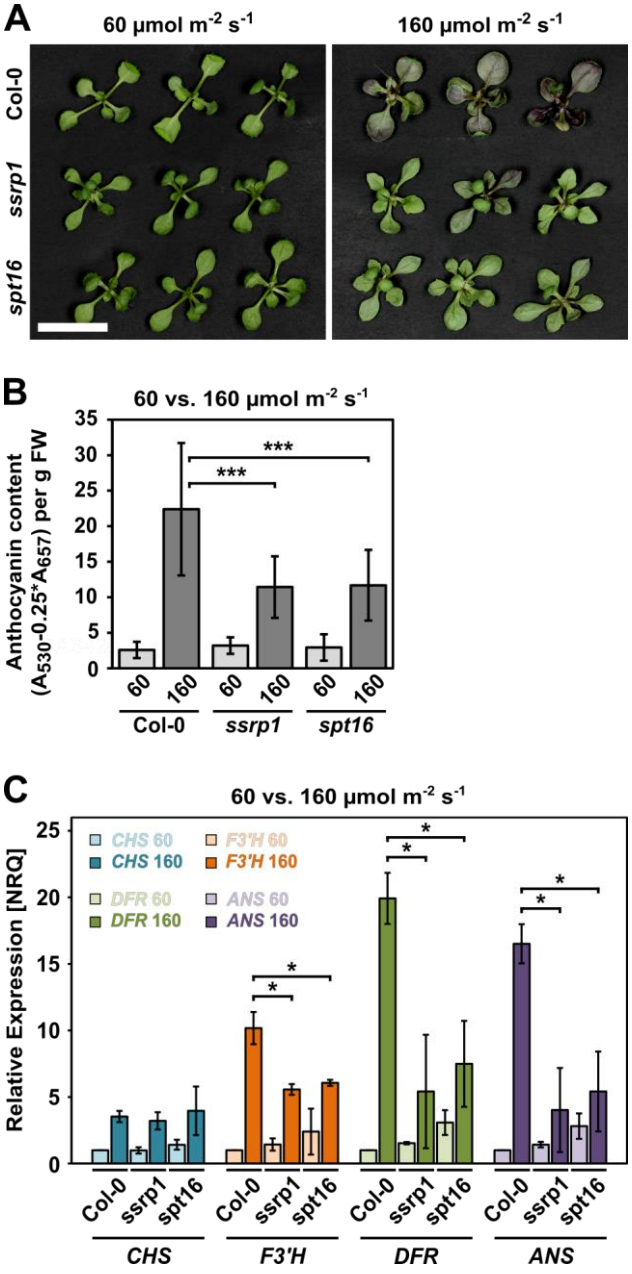


Fig. 4. *ssrp1/spt16* plants exhibit reduced anthocyanin levels in response to moderate high-light stress. A, Three representative 14-d plants each (*ssrp1, spt16,* Col-0; abaxial view) grown under low-light (60 µmol m^{-2} s^{-1}) or moderate high-light (160 µmol m^{-2} s^{-1}) are shown. Scale bar: 1 cm. B, Anthocyanin content of plants of the three genotypes grown under the two light conditions was measured photometrically each in three biological replicates. Data are means ± SD and significance was tested by Student's t-test (*** indicating P<0.001). C, RNA isolated from plants of the three genotypes grown under the two light conditions was analysed by qRT-PCR measuring the transcript level of four anthocyanin biosynthetic genes (*CHS, F´3H, DFR* and *ANS*). The mean normalised relative quantities (NRQ) ± SD are shown of two biological replicates. Significance was analysed by Student´s t-test (* indicating p<0.05).

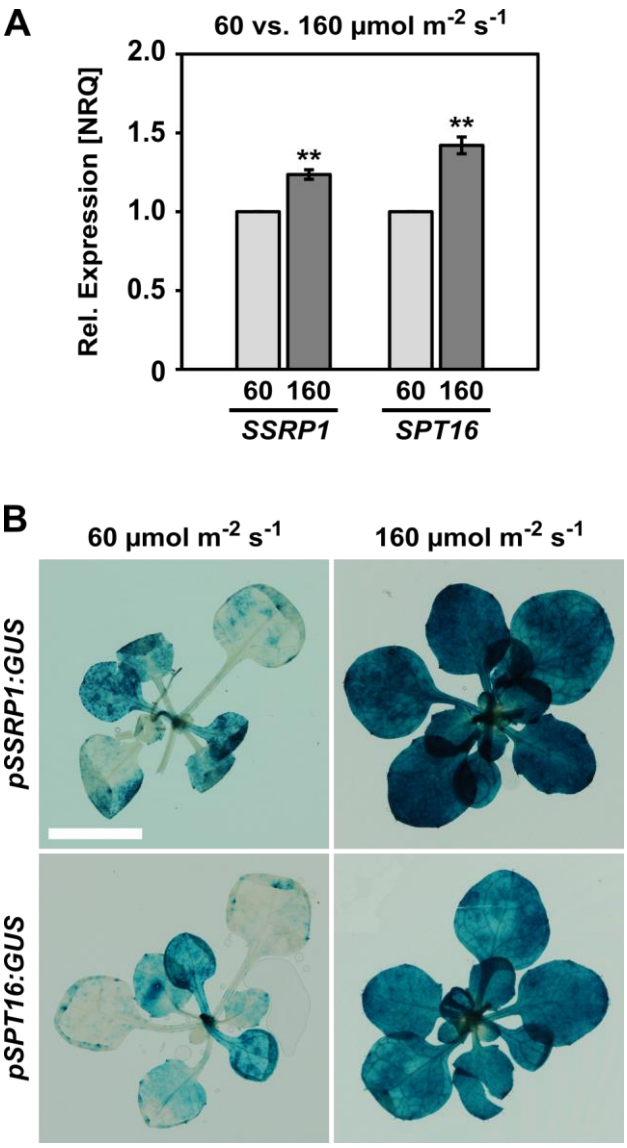


Fig. 5. Expression of *SSRP1* and *SPT16* is up-regulated by moderate high-light treatment. A, Transcript levels of *SSRP1* and *SPT16* in 14-d Col-0 plants grown under low-light (60 µmol m^{-2} s^{-1}) or moderate high-light (160 µmol m^{-2} s^{-1}) were measured by qRT-PCR. The mean NRQs ± SD are shown of two biological replicates. Significance was analysed by Student´s t-test (** indicating P<0.01). B, *SSRP1* and *SPT16* promoter activity determined by histochemical GUS staining of transgenic Col-0 plants harbouring *pSSRP1*- and *pSPT16-GUS* reporter constructs grown under the two light conditions. Scale bar: 1 cm.

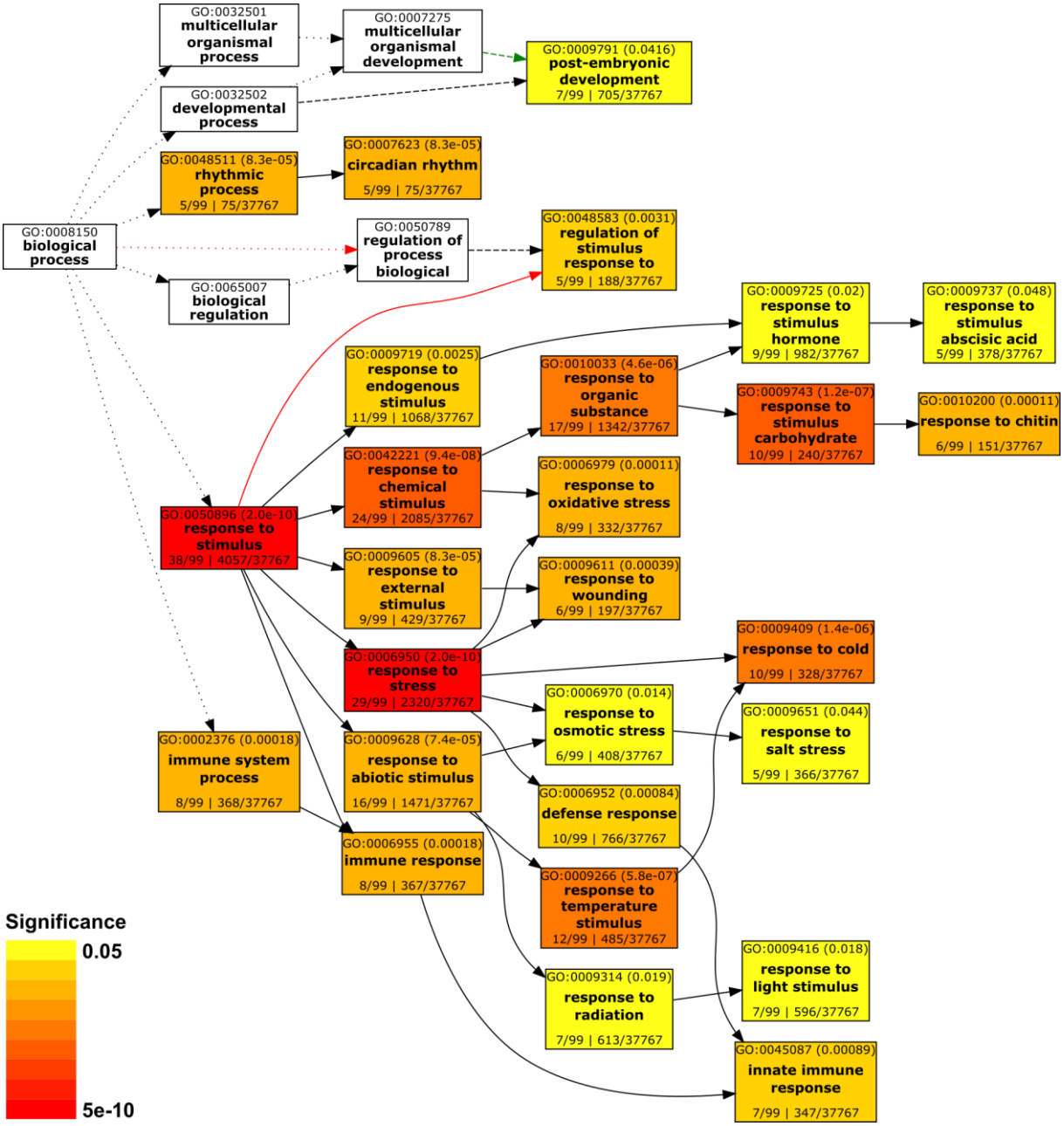


Fig. S1. Biological processes identified by an overrepresented number of genes up-regulated in *ssrp1* and/or *spt16*.

A

	AGI	Gene	*ssrp1*	p-Value	*spt16*	p-Value
Regulatory Genes	AT5G35550	*TT2*	-1.1615	0.3590	-1.1039	0.6137
	AT4G09820	*TT8*	-1.3636	0.0079	-1.2532	0.0706
	AT5G24520	*TTG1*	-1.1111	0.2690	1.0015	0.9397
	AT1G56650	*PAP1*	-1.4874	0.1388	-1.4237	0.0964
	AT1G66390	*PAP2*	-1.2955	0.4142	-1.0111	0.8853
	AT5G41315	*GL3*	-1.1717	0.1660	-1.0395	0.7110
	AT1G63650	*EGL3*	-1.0353	0.5879	-1.0936	0.2203
Biosynthetic Genes	AT3G53260	*PAL2*	-1.3475	0.0029	-1.2394	0.0053
	AT2G37040	*PAL1*	-1.4495	0.0228	-1.4059	0.0402
	AT2G30490	*C4H*	-1.1738	0.0025	-1.1948	0.1139
	AT1G65060	*4CL3*	-1.1626	0.1562	-1.1922	0.2017
	AT1G51680	*4CL1*	-1.4532	0.0023	-1.2265	0.0139
	AT5G13930	*CHS / TT4*	-2.0757	0.0004	-2.0379	0.0012
	AT5G05270	*CHI-like / CHIL*	-2.0115	2.17E-05	-2.1126	0.0002
	AT3G55120	*CHI / TT5*	-1.2771	0.0014	-1.4191	0.0004
	AT5G07990	*F3'H / TT7*	-2.5929	0.0026	-2.0540	0.0045
	AT3G51240	*F3H / TT6*	-1.7723	0.0003	-2.2395	0.0001
	AT5G42800	*DFR / TT3*	-2.5563	0.0009	-2.4136	0.0004
	AT4G22880	*ANS / TT18 / LDOX*	-2.6219	0.0001	-2.6903	0.0002
	AT5G08640	*FLS*	-1.0437	0.4653	-1.1431	0.1805
	AT1G61720	*ANR*	-1.0703	0.3739	-1.0421	0.7897
	AT5G17050	*UGT78D2 (AGT)*	-1.1384	0.4498	-1.1560	0.3924
	AT4G14090	*AT4G14090 (AGT)*	-1.6921	0.0214	-1.6787	0.0062
	AT5G54060	*UF3GT (AGT)*	-2.8032	0.0044	-2.2328	0.0011
	AT3G21560	*UGT84A2 (AGT)*	-1.4610	0.0055	-1.2696	0.0089
	AT2G29750	*UGT71C1 (AGT)*	1.1477	0.4314	-1.1443	0.4671
	AT4G01070	*GT72B1 (AGT)*	-1.0737	0.3326	-1.0484	0.2875
	AT1G03940	*AT1G03940 (AAT)*	-3.1166	0.0009	-2.0378	0.0010
	AT3G29590	*AT5MAT (AAT)*	-2.4382	0.0001	-1.7776	0.0022
	AT1G03495	*AT1G03495 (AAT)*	-1.8482	0.0045	-1.8628	0.0005
Transport Genes	AT5G17220	*TT19 (GST)*	-2.2878	0.0009	-1.7738	0.0019

B

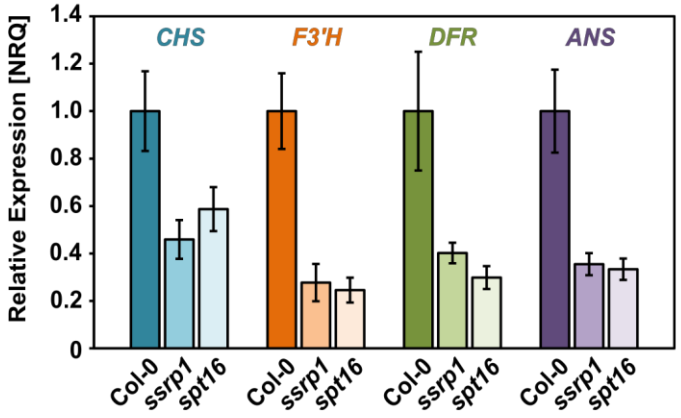


Fig. S2. Genes of the anthocyanin biosynthetic pathway are down-regulated in *ssrp1 and spt16*, while regulatory genes are not affected.

AGI	Gene	*ssrp1*	p-Value	*spt16*	p-Value	
AT3G56970	*BHLH038*	-9.6918	4.01E-05	-7.7266	2.81E-04	Iron homeo-stasis
AT3G56980	*BHLH039*	-5.2627	2.94E-04	-3.0012	7.16E-04	
AT2G41240	*BHLH100*	-7.9985	5.12E-07	-5.6102	4.70E-03	
AT5G04150	*BHLH101*	-3.5967	3.28E-04	-3.3528	2.86E-03	

Fig. S3. Genes encoding bHLH transcription factors are down-regulated in *ssrp1* and *spt16*.

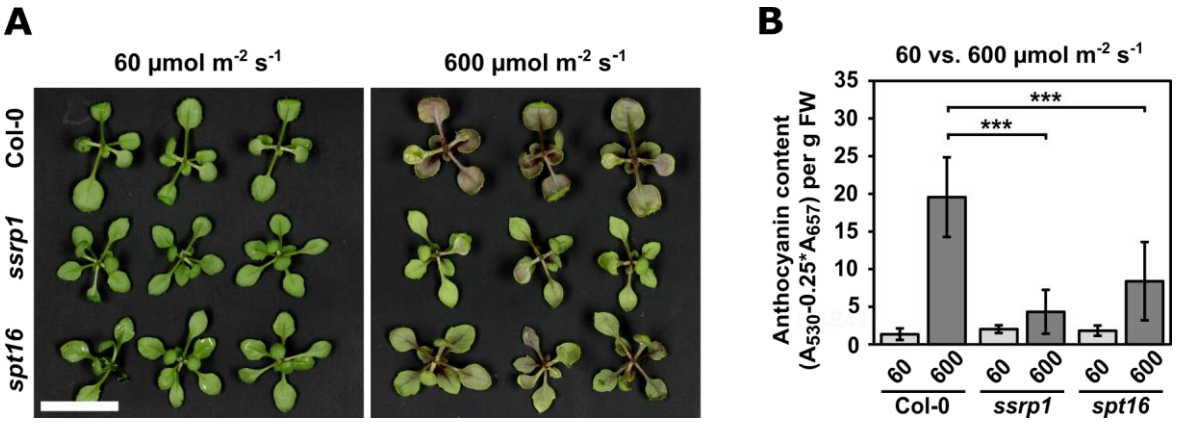


Fig. S4. *ssrp1/spt16* plants exhibit reduced anthocyanin levels in response to severe high-light stress.

Supplemental Table. Oligonucleotide primers used in this study and construction of plasmids primer

	use	plasmid	restr. site
AATTAAGCTTAGAATTCTAGCAATGCAGGGTAA	Insertion of pSSRP1 in pCambia3300-GUS	pCambia3300:pSSRP1::GUS	HindIII
AATCTAGAGGTTTTCTGTTGAGAGACACGA	Insertion of pSSRP1 in pCambia3300-GUS	pCambia3300:pSSRP1::GUS	XbaI
TCCCCCGGGAAGCTTTATATCAAGAACAAGAAGAAACAAGTCTTTG	Insertion of pSPT16 in pCambia3300-GUS	pCambia3300:pSPT16::GUS	SmaI
CGGGATCCTCTAGACTAAAGAGTCCAGCAGCAACCT	Insertion of pSPT16 in pCambia3300-GUS	pCambia3300:pSPT16::GUS	BamHI
TGAAAACTGACCATGAACATCTG	qRT-PCR *SSRP1 (AT3G28730)*		
CATTATCCCCAAGAACAGCAG	qRT-PCR *SSRP1 (AT3G28730)*		
GAGGGGCTCGGGCATTACCAT	qRT-PCR *SPT16 (AT4G10710)*		
CCAAAACCGCCTTTGTGTAAAGCT	qRT-PCR *SPT16 (AT4G10710)*		
GGAAGAGAAGATGAGGGCGA	qRT-PCR *CHS (AT5G13930)*		
AACAAGACACCCCACTCCAA	qRT-PCR *CHS (AT5G13930)*		
TGATATTGTTGTGGGCCGTG	qRT-PCR *F3'H (AT3G51240)*		
CCGTTGATCTCACAGCTCTC	qRT-PCR *F3'H (AT3G51240)*		
TCAGGCCAAAATACCCCGAA	qRT-PCR *DFR (AT5G42800)*		
ATGTCCGTCAGCTTCTTGGA	qRT-PCR *DFR (AT5G42800)*		
TGCGTATCCTGAAGAGAAGAG	qRT-PCR *ANS (AT4G22880)*		
GACGGTCAGGCTCTAAACCT	qRT-PCR *ANS (AT4G22880)*		
ACCCTTGAAGTGGAAAGCTCC	qRT-PCR *UBI10 (At4g05320)*		
TTCCAGCGAAGATGAGACGC	qRT-PCR *UBI10 (At4g05320)*		
AACGTGGCCAAAATGATGC	qRT-PCR *PP2AA3 (At1g13320)*		
CACATTGTCAATAGATTGGAGAGC	qRT-PCR *PP2AA3 (At1g13320)*		
TGGGAAAGTGTTGCCATCC	qRT-PCR *GAP (At1g13440)*		
CTTCATTTTGCCTTCAGATTCCTC	qRT-PCR *GAP (At1g13440)*		
CCCTCATCTTACGCGTATCAGA	Genotyping, T-DNA insertion *ssrp1-2* (SALK_001283)		
AATTAAGCTTAGTTACTATCGGAATCGTTTCCT	Genotyping, T-DNA insertion *ssrp1-2* (SALK_001283)		
GTTGCCCGTCTCACTGGTGA	Genotyping, T-DNA insertion SALK LBb1.3 (*ssrp1-2*)		
CTATCTCTGCATTGCCTCTTAGC	Genotyping, T-DNA insertion *spt16-1* (SAIL_392_G06)		
TACTTGTCTAACGCAGCGAAATC	Genotyping, T-DNA insertion *spt16-1* (SAIL_392_G06)		
GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC	Genotyping, T-DNA insertion SAIL LB (*spt16-1*)		
AAT TAA GCT TAG AAT TCT AGC AAT GCA GGG TAA	Genotyping, T-DNA insertion *pSSRP1::GUS*		
AAA CAG TGT TAG GCA TTA AGC GTA CAT	Genotyping, T-DNA insertion *pSPT16::GUS*		
CGA TCC AGA CTG AAT GCC CA	Genotyping, T-DNA insertion GUS (*pSSRP1::GUS*,*pSPT16::GUS*)		

1. Brkljacic,J. and Grotewold,E. 2017. Combinatorial control of plant gene expression. Biochim Biophys Acta 1860: 31-40.
2. Reiter,F., Wienerroither,S., and Stark,A. 2017. Combinatorial function of transcription factors and cofactors . Curr Opin Genet Dev 43: 73-81.
3. Jonkers,I. and Lis,J.T. 2015. Getting up to speed with transcription elongation by RNA polymerase II. Nat Rev Mol Cell Biol 16: 167-177.
4. Sims,R.J., Belotserkovskaya,R., and Reinberg,D. 2004. Elongation by RNA polymerase II: the short and the long of it. Genes Dev 18: 2437-2468.
5. Van Lijsebettens,M. and Grasser,K.D. 2014. Transcript elongation factors: shaping transcriptomes after transcript initiation. Trends Plant Sci 19: 717-726.
6. Formosa,T. 2012. The role of FACT in making and breaking nucleosomes. Biochim Biophys Acta 1819: 247-255.
7. Reinberg,D. and Sims,R.J. 2006. de FACTo nucleosome dynamics. J Biol Chem 281: 23297-23301.
8. Belotserkovskaya,R., Oh,S., Bondarenko,V.A., Orphanides,G., Studitsky,V.M., and Reinberg,D. 2003. FACT facilitates transcription-dependent nucleosome alteration. Science 301: 1090-1093.
9. Valieva,M.E., Armeev,G.A., Kudryashova,K.S., Gerasimova,N.S., Shaytan,A.K., Kulaeva,O.I., McCullough,L.L., Formosa,T., Georgiev,P.G., Kirpichnikov,M.P., Studitsky,V.M., and Feofanov,A.V. 2016. Large-scale ATP-independent nucleosome unfolding by a histone chaperone. Nat Struct Mol Biol 23: 1111-1116.
10. Xin,H., Takahata,S., Blanksma,M., McCullough,L., Stillman,D.J., and Formosa,T. 2009. yFACT induces global accessibility of nucleosomal DNA without H2A-H2B displacement. Mol Cell 35: 365-376.
11. Cheung,V., Chua,G., Batada,N.N., Landry,C.R., Michnick,S.W., Hughes,T.R., and Winston,F. 2008. Chromatin- and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome. PLoS Biology 6: e277.
12. Mason,P.B. and Struhl,K. 2003. The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation *in vivo*. Mol Cell Biol 23: 8323-8333.
13. Li,G., Zeng,S.X., Landais,I., and Lu,H. 2007. Human SSRP1 has Spt16-dependent and -independent roles in gene transcription. J Biol Chem 282: 6936-6945.
14. Antosch,M., Mortensen,S.A., and Grasser,K.D. 2012. Plant proteins containing high mobility group box DNA-binding domains modulate different nuclear processes. Plant Physiol 159: 875-883.

15. Zhou,W., Zhu,Y., Dong,A., and Shen,W.-H. 2015. Histone H2A/H2B chaperones: from molecules to chromatin-based functions in plant growth and development. Plant J 83: 78-95.
16. Antosz,W., Pfab,A., Ehrnsberger,H.F., Holzinger,P., Köllen,K., Mortensen,S.A., Bruckmann,A., Schubert,T., Längst,G., Griesenbeck,J., Schubert,V., Grasser,M., and Grasser,K.D. 2017. Composition of the *Arabidopsis* RNA polymerase II transcript elongation complex reveals the interplay between elongation and mRNA processing factors. Plant Cell 29: 854-870.
17. Duroux,M., Houben,A., Ruzicka,K., Friml,J., and Grasser,K.D. 2004. The chromatin remodelling complex FACT associates with actively transcribed regions of the *Arabidopsis* genome. Plant J 40: 660-671.
18. Lolas,I.B., Himanen,K., Grønlund,J.T., Lynggaard,C., Houben,A., Melzer,M., Van Lijsebettens,M., and Grasser,K.D. 2010. The transcript elongation factor FACT affects *Arabidopsis* vegetative and reproductive development and genetically interacts with *HUB1/2*. Plant J 61: 686-697.
19. Antosch,M., Schubert,V., Holzinger,P., Houben,A., and Grasser,K.D. 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.
20. Bolstad,B.M., Irizarry,R.A., Astrand,M., and Speed,T.P. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 22: 185-193.
21. Tian,T., Liu,Y., Yan,H., You,Q., Yi,X., Du,Z., Xu,W., and Su,Z. 2017. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res 45: W122-W129.
22. Babicki,S., Arndt,D., Marcu,A., Liang,Y., Grant,J.R., Maciejewski,A., and Wishart,D.S. 2017. Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res 44: W147-W153.
23. Dürr,J., Lolas,I.B., Sørensen,B.B., Schubert,V., Houben,A., Melzer,M., Deutzmann,R., Grasser,M., and Grasser,K.D. 2014. The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in *Arabidopsis*. Nucleic Acids Res 42: 4332-4347.
24. Hellemans,J., Mortier,G., De Paepe,A., Speleman,F., and Vandesompele,J. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol 8: R19.
25. Kudo,T., Sasaki,Y., Terashima,S., Matsuda-Imai,N., Takano,T., Saito,M., Kanno,M., Ozaki,S., Suwabe,K., Suzuki,G., Watanabe,M., Matsuoka,M., Takayama,S., and Yano,K. 2016. Identification of reference genes for quantitative expression analysis

using large-scale RNA-seq data of *Arabidopsis thaliana* and model crop plants. Genes Genet Syst 91: 111-125.

26. Yin,R., Messner,B., Faus-Kessler,T., Hoffmann,T., Schwab,W., Hajirezaei,M.R., von Saint Paul,V., Heller,W., and Schäffner,A.R. 2012. Feedback inhibition of the general phenylpropanoid and flavonol biosynthetic pathways upon a compromised flavonol-3-O-glycosylation. J Exp Bot 63: 2465-2478.
27. Jefferson,R.A. 1987. Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol Biol Rep 5: 387-405.
28. Saito,K., Yonekura-Sakakibara,K., Nakabayashi,R., Higashi,Y., Yamazaki,M., Tohge,T., and Fernie,A.R. 2013. The flavonoid biosynthetic pathway in *Arabidopsis*: structural and genetic diversity. Plant Physiol Biochem 72: 21-34.
29. Wang,H.Y., Klatte,M., Jacoby,M., Bäumlein,H., Weisshaar,B., and Bauer,P. 2007. Iron deficiency-mediated stress regulation of four subgroup Ib BHLH genes in *Arabidopsis thaliana*. Planta 226: 897-908.
30. Wang,N., Cui,Y., Liu,Y., Fan,H., Du,J., Huang,Z., Yuan,Y., Wu,H., and Ling,H.Q. 2013. Requirement and functional redundancy of Ib subgroup bHLH proteins for iron deficiency responses and uptake in *Arabidopsis thaliana*. Mol Plant 6: 503-513.
31. Li,S., Wang,W., Gao,J., Yin,K., Wang,R., Wang,C., Petersen,M., Mundy,J., and Qiu,J.L. 2016. MYB75 phosphorylation by MPK4 Is required for light-induced anthocyanin accumulation in *Arabidopsis*. Plant Cell 28: 2866-2883.
32. Lotkowska,M.E., Tohge,T., Fernie,A.R., Xue,G.P., Balazadeh,S., and Mueller-Roeber,B. 2015. The *Arabidopsis* transcription factor MYB112 promotes anthocyanin formation during salinity and under high light stress. Plant Physiol 169: 1862-1880.
33. Vanderauwera,S., Zimmermann,P., Rombauts,S., Vandenabeele,S., Langebartels,C., Gruissem,W., Inzé,D., and Van Breusegem,F. 2005. Genome-wide analysis of hydrogen peroxide-regulated gene expression in *Arabidopsis* reveals a high light-induced transcriptional cluster involved in anthocyanin biosynthesis. Plant Physiol 139: 806-821.
34. Albert,N.W., Lewis,D.H., Zhang,H., Irving,L.J., Jameson,P.E., and Davies,K.M. 2009. Light-induced vegetative anthocyanin pigmentation in *Petunia*. J Exp Bot 60: 2191-2202.
35. Orphanides,G., Wu,W.-H., Lane,W.S., Hampsey,M., and Reinberg,D. 1999. The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. Nature 400: 284-288.
36. Grasser,M., Kane,C.M., Merkle,T., Melzer,M., Emmersen,J., and Grasser,K.D. 2009. Transcript elongation factor TFIIS is involved in *Arabidopsis* seed dormancy. J Mol Biol 386: 598-611.

37. Nelissen,H., DeGroeve,S., Fleury,D., Neyt,P., Bruno,L., Bitonti,M.B., Vandenbussche,F., Van Der Straeten,D., Yamaguchi,T., Tsukaya,H., Witters,E., de Jaeger,G., Houben,A., and Van Lijsebettens,M. 2010. Plant Elongator regulates auxin-related genes during RNA polymerase II-mediated transcription elongation. Proc Natl Acad Sci USA 107: 1678-1683.
38. Mayer,A., Lidschreiber,M., Siebert,M., Leike,K., Söding,J., and Cramer,P. 2010. Uniform transitions of the general RNA polymerase II transcription complex. Nat Struct Mol Biol 17: 1272-1278.
39. Perales,M. and Más,P. 2007. A functional link between rhythmic changes in chromatin structure and the *Arabidopsis* biological clock. Plant Cell 19: 2111-2123.
40. Pérez-García,P., Ma,Y., Yanovsky,M.J., and Mas,P. 2015. Time-dependent sequestration of RVE8 by LNK proteins shapes the diurnal oscillation of anthocyanin biosynthesis. Proc Natl Acad Sci USA 112: 5249-5253.

Acknowledgements

Ich möchte mich an dieser Stelle bei allen bedanken, die mich bei der Anfertigung meiner Doktorarbeit unterstützt haben bzw. diese erst ermöglicht haben.

Zu aller erst möchte ich mich bei Prof. Dr. Klaus Grasser bedanken, der es mir ermöglicht hat an diesem Projekt zu arbeiten und in seiner Arbeitsgruppe zu promovieren. Danke für die Begutachtung meiner Doktorarbeit und die Unterstützung durch Ratschläge und Diskussionen.

Bei Prof. Dr. Gernot Längst möchte ich mich für die Betreuung als Mentor und die Zweitbegutachtung meiner Arbeit bedanken.

Bei Prof. Dr. Herbert Tschochner und Prof. Dr. Stephan Schneuwly möchte ich mich, für die Bereitschaft als Drittprüfer beziehungsweise als Prüfungsvorsitzender zu fungieren, bedanken.

Bei allen aktuellen und ehemaligen Mitgliedern der AG Grasser bedanke ich mich für die freundschaftliche und entspannte Atmosphäre (Vor allem zur Weihnachtszeit:-). Es war eine sehr schöne Zeit mit euch im Labor! (Chronologisch:-) Marion, Schmitzi, Martin, Julius, Brian, Lena, Wojciech, Hans, Silvia, Philipp, Antje, Irene, Tina, Ines, und allen anderen die ich vergessen haben sollte.

Ich möchte mich darüber hinaus bei allen Mitgliedern des Lehrstuhls bedanken, für Hilfestellungen, Tipps und die schönen Ausflüge;) Dabei geht ein besonderer Dank an meinen Kletterpartner Tom!

Ich möchte mich bei der Armee an Studenten bedanken die mich unterstützt haben. Im Besonderen bei Marcell, Nico, Philipp, Korbinian, und Matthias.

Ein besonderer Dank geht an Astrid und Edi für die MS Analysen und an Prof. Dr. Rainer Merkl für alle bioinformatischen Fragestellungen.

Ein ganz besonderer Dank geht an meine fantastische Frau Silli! Für deine Unterstützung, dein Verständnis und weil mit dir einfach alles mehr Spaß macht!

Schlussendlich möchte ich mich bei meiner Familie bedanken, für die Unterstützung und da ohne sie das alles nicht möglich gewesen wäre!! Vielen Dank!!

List of Figures

List of Tables

Bibliography

Acosta, E. E. C. and V. Weake (2017). "Determination of Pair-wise Protein Interactions between Spt7 and Drosophila melanogaster SAGA Subunits". In: *The FASEB Journal* 31.1 Supplement, pp. 601–13.

Adelman, K., W. Wei, M. B. Ardehali, J. Werner, B. Zhu, D. Reinberg, and J. T. Lis (2006). "Drosophila Paf1 modulates chromatin structure at actively transcribed genes". In: *Molecular and cellular biology* 26.1, pp. 250–260.

Albert, N. W., D. H. Lewis, H. Zhang, L. J. Irving, P. E. Jameson, and K. M. Davies (2009). "Light-induced vegetative anthocyanin pigmentation in Petunia". In: *Journal of experimental botany* 60.7, pp. 2191–2202.

Albertazzi, L., D. Arosio, L. Marchetti, F. Ricci, and F. Beltram (2009). "Quantitative FRET Analysis With the E0GFP-mCherry Fluorescent Protein Pair". In: *Photochemistry and photobiology* 85.1, pp. 287–297.

Allen, B. L. and D. J. Taatjes (2015). "The Mediator complex: a central integrator of transcription". In: *Nature reviews. Molecular cell biology* 16.3, p. 155.

Andrews, A. J. and K. Luger (2011). "Nucleosome structure (s) and stability: variations on a theme". In: *Annual review of biophysics* 40, pp. 99–117.

Andriankaja, M. E. et al. (2014). "Transcriptional coordination between leaf cell differentiation and chloroplast development established by TCP20 and the subgroup Ib bHLH transcription factors". In: *Plant molecular biology* 85.3, pp. 233–245.

Antosch, M., S. A. Mortensen, and K. D. Grasser (2012). "Plant proteins containing high mobility group box DNA-binding domains modulate different nuclear processes". In: *Plant physiology* 159.3, pp. 875–883.

Antosz, W. et al. (2017). "The composition of the Arabidopsis RNA polymerase II transcript elongation complex reveals the interplay between elongation and mRNA processing factors". In: *The Plant Cell*, tpc–00735.

Appelhagen, I., K. Thiedig, N. Nordholt, N. Schmidt, G. Huep, M. Sagasser, and B. Weisshaar (2014). "Update on transparent testa mutants from Arabidopsis thaliana: characterisation of new alleles from an isogenic collection". In: *Planta* 240.5, pp. 955–970.

Apweiler, R. et al. (2004). "UniProt: the universal protein knowledgebase". In: *Nucleic acids research* 32.suppl_1, pp. D115–D119.

Atanassov, B. S. et al. (2016). "ATXN7L3 and ENY2 coordinate activity of multiple H2B deubiquitinases important for cellular proliferation and tumor growth". In: *Molecular cell* 62.4, pp. 558–571.

Avvakumov, N., A. Nourani, and J. Côté (2011). "Histone chaperones: modulators of chromatin marks". In: *Molecular cell* 41.5, pp. 502–514.

Babicki, S., D. Arndt, A. Marcu, Y. Liang, J. R. Grant, A. Maciejewski, and D. S. Wishart (2016). "Heatmapper: web-enabled heat mapping for all". In: *Nucleic acids research* 44.W1, W147–W153.

Baker, S. and P. Grant (2007). "The SAGA continues: expanding the cellular role of a transcriptional co-activator complex". In: *Oncogene* 26.37, p. 5329.

Balasubramanian, R., M. G. Pray-Grant, W. Selleck, P. A. Grant, and S. Tan (2002). "Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation". In: *Journal of Biological Chemistry* 277.10, pp. 7989–7995.

Belotserkovskaya, R., S. Oh, V. A. Bondarenko, G. Orphanides, V. M. Studitsky, and D. Reinberg (2003). "FACT facilitates transcription-dependent nucleosome alteration". In: *Science* 301.5636, pp. 1090–1093.

Belotserkovskaya, R. and D. Reinberg (2004). "Facts about FACT and transcript elongation through chromatin". In: *Current opinion in genetics & development* 14.2, pp. 139–146.

Benhamed, M. et al. (2008). "Genome-scale Arabidopsis promoter array identifies targets of the histone acetyltransferase GCN5". In: *The Plant Journal* 56.3, pp. 493–504.

Bentley, D. L. (2014). "Coupling mRNA processing with transcription in time and space". In: *Nature Reviews Genetics* 15.3, pp. 163–175.

Bolte, S. and F. Cordelieres (2006). "A guided tour into subcellular colocalization analysis in light microscopy". In: *Journal of microscopy* 224.3, pp. 213–232.

Bonnet, J., D. Devys, and L. Tora (2014). "Histone H2B ubiquitination: signaling not scrapping". In: *Drug Discovery Today: Technologies* 12, e19–e27.

Bouché, N. and D. Bouchez (2001). "Arabidopsis gene knockout: phenotypes wanted". In: *Current opinion in plant biology* 4.2, pp. 111–117.

Boyes, D. C., A. M. Zayed, R. Ascenzi, A. J. McCaskill, N. E. Hoffman, K. R. Davis, and J. Görlach (2001). "Growth stage–based phenotypic analysis of Arabidopsis a model for high throughput functional genomics in plants". In: *The Plant Cell* 13.7, pp. 1499–1510.

Brewster, N. K., G. C. Johnston, and R. A. Singer (1998). "Characterization of the CP complex, an abundant dimer of Cdc68 and Pob3 proteins that regulates yeast transcriptional activation and chromatin repression". In: *Journal of Biological Chemistry* 273.34, pp. 21972–21979.

Brewster, N. K., G. C. Johnston, and R. A. Singer (2001). "A bipartite yeast SSRP1 analog comprised of Pob3 and Nhp6 proteins modulates transcription". In: *Molecular and Cellular Biology* 21.10, pp. 3491–3502.

Brückner, A., C. Polge, N. Lentze, D. Auerbach, and U. Schlattner (2009). "Yeast two-hybrid, a powerful tool for systems biology". In: *International journal of molecular sciences* 10.6, pp. 2763–2788.

Brumbarova, T., P. Bauer, and R. Ivanov (2015). "Molecular mechanisms governing Arabidopsis iron uptake". In: *Trends in plant science* 20.2, pp. 124–133.

Buratowski, S. (2009). "Progression through the RNA polymerase II CTD cycle". In: *Molecular cell* 36.4, pp. 541–546.

Cao, Y., Y. Dai, S. Cui, and L. Ma (2008). "Histone H2B monoubiquitination in the chromatin of FLOWERING LOCUS C regulates flowering time in Arabidopsis". In: *The Plant Cell* 20.10, pp. 2586–2602.

Chanarat, S. and K. Sträßer (2013). "Splicing and beyond: the many faces of the Prp19 complex". In: *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* 1833.10, pp. 2126–2134.

Chen, X.-F. et al. (2012). "Mediator and SAGA have distinct roles in Pol II preinitiation complex assembly and function". In: *Cell reports* 2.5, pp. 1061–1067.

Chien, C.-T., P. L. Bartel, R. Sternglanz, and S. Fields (1991). "The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest." In: *Proceedings of the National Academy of Sciences* 88.21, pp. 9578–9582.

Chudakov, D. M., M. V. Matz, S. Lukyanov, and K. A. Lukyanov (2010). "Fluorescent proteins and their applications in imaging living cells and tissues". In: *Physiological reviews* 90.3, pp. 1103–1163.

Clough, S. J. and A. F. Bent (1998). "Floral dip: a simplified method forAgrobacterium-mediated transformation ofArabidopsis thaliana". In: *The plant journal* 16.6, pp. 735–743.

Colangelo, E. P. and M. L. Guerinot (2004). "The essential basic helix-loop-helix protein FIT1 is required for the iron deficiency response". In: *The Plant Cell* 16.12, pp. 3400–3412.

Costigan, C., D. Kolodrubetz, and M. Snyder (1994). "NHP6A and NHP6B, which encode HMG1-like proteins, are candidates for downstream components of the yeast SLT2 mitogen-activated protein kinase pathway." In: *Molecular and cellular biology* 14.4, pp. 2391–2403.

Couzigou, J.-M. and J.-P. Combier (2016). "Plant microRNAs: key regulators of root architecture and biotic interactions". In: *New Phytologist* 212.1, pp. 22–35.

Cuenca-Bono, B., V. García-Molinero, P. Pascual-García, E. García-Oliver, A. Llopis, and S. Rodríguez-Navarro (2010). "A novel link between Sus1 and the cytoplasmic mRNA decay machinery suggests a broad role in mRNA metabolism." In: *BMC cell biology* 11, p. 19.

Daniel, J. A., M. S. Torok, Z.-W. Sun, D. Schieltz, C. D. Allis, J. R. Yates, and P. A. Grant (2004). "Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription". In: *Journal of Biological Chemistry* 279.3, pp. 1867–1871.

Dean, C., J. Jones, M. Favreau, P. Dunsmuir, and J. Bedbrook (1988). "Influence of flanking sequences on variability in expression levels of an introduced gene in transgenic tobacco plants". In: *Nucleic acids research* 16.19, pp. 9267–9283.

Dundr, M., U. Hoffmann-Rohrer, Q. Hu, I. Grummt, L. I. Rothblum, R. D. Phair, and T. Misteli (2002). "A kinetic framework for a mammalian RNA polymerase in vivo". In: *Science* 298.5598, pp. 1623–1626.

Durand, A., J. Bonnet, M. Fournier, V. Chavant, and P. Schultz (2014). "Mapping the deubiquitination module within the saga complex". In: *Structure* 22.11, pp. 1553–1559.

Duroux, M., A. Houben, K. Ruzicka, J. Friml, and K. D. Grasser (2004). "The chromatin remodelling complex FACT associates with actively transcribed regions of the Arabidopsis genome". In: *The Plant Journal* 40.5, pp. 660–671.

Dürr, J., I. B. Lolas, B. B. Sørensen, V. Schubert, A. Houben, M. Melzer, R. Deutzmann, M. Grasser, and K. D. Grasser (2014). "The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in Arabidopsis". In: *Nucleic acids research* 42.7, pp. 4332–4347.

Edwards, K., C. Johnstone, and C. Thompson (1991). "A simple and rapid method for the preparation of plant genomic DNA for PCR analysis." In: *Nucleic acids research* 19.6, p. 1349.

Ehara, H., T. Yokoyama, H. Shigematsu, S. Yokoyama, M. Shirouzu, and S.-i. Sekine (2017). "Structure of the complete elongation complex of RNA polymerase II with basal factors". In: *Science* 357.6354, pp. 921–924.

Ellisdon, A. M., L. Dimitrova, E. Hurt, and M. Stewart (2012). "Structural basis for the assembly and nucleic acid binding of the TREX-2 transcription-export complex". In: *Nature structural & molecular biology* 19.3, pp. 328–336.

Ellisdon, A. M., D. Jani, A. Köhler, E. Hurt, and M. Stewart (2010). "Structural basis for the interaction between yeast Spt-Ada-Gcn5 acetyltransferase (SAGA) complex components Sgf11 and Sus1". In: *Journal of Biological Chemistry* 285.6, pp. 3850–3856.

Elsässer, S. J. and S. DArcy (2012). "Towards a mechanism for histone chaperones". In: *Biochimica Et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1819.3, pp. 211–221.

Evans, D. R., N. K. Brewster, Q. Xu, A. Rowley, B. A. Altheim, G. C. Johnston, and R. A. Singer (1998). "The yeast protein complex containing cdc68 and pob3 mediates core-promoter repression through the cdc68 N-terminal domain". In: *Genetics* 150.4, pp. 1393–1405.

Fabrizio, P., J. Dannenberg, P. Dube, B. Kastner, H. Stark, H. Urlaub, and R. Lührmann (2009). "The Evolutionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome". In: *Molecular Cell* 36.4, pp. 593–608.

Fang, Y., S. Hearn, and D. L. Spector (2004). "Tissue-specific expression and dynamic organization of SR splicing factors in Arabidopsis". In: *Molecular biology of the cell* 15.6, pp. 2664–2673.

Fauser, F., S. Schiml, and H. Puchta (2014). "Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana". In: *The Plant Journal* 79.2, pp. 348–359.

Fields, S. and O.-k. Song (1989). "A novel genetic system to detect protein–protein interactions". In: *Nature* 340.6230, pp. 245–246.

Fischer, T., S. Rodríguez-Navarro, G. Pereira, A. Rácz, E. Schiebel, and E. Hurt (2004). "Yeast centrin Cdc31 is linked to the nuclear mRNA export machinery". In: *Nature cell biology* 6.9, pp. 840–848.

Fischer, T., K. Sträßer, A. Rácz, S. Rodriguez-Navarro, M. Oppizzi, P. Ihrig, J. Lechner, and E. Hurt (2002). "The mRNA export machinery requires the novel Sac3p–Thp1p complex

to dock at the nucleoplasmic entrance of the nuclear pores". In: *The EMBO journal* 21.21, pp. 5843–5852.

Fish, R. N. and C. M. Kane (2002). "Promoting elongation with transcript cleavage stimulatory factors". In: *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* 1577.2, pp. 287–307.

Flanagan, J. F., M. Li-Zhi, M. Chruszcz, M. Cymborowski, K. L. Clines, Y. Kim, W. Minor, F. Rastinejad, and S. Khorasanizadeh (2005). "Double chromodomains cooperate to recognize the methylated histone H3 tail". In: *Nature* 438.7071, p. 1181.

Fleury, D. et al. (2007). "The Arabidopsis thaliana homolog of yeast BRE1 has a function in cell cycle regulation during early leaf and root growth". In: *The Plant Cell* 19.2, pp. 417–432.

Formosa, T. (2008). "FACT and the reorganized nucleosome". In: *Molecular BioSystems* 4.11, pp. 1085–1093.

Formosa, T. (2012). "The role of FACT in making and breaking nucleosomes". In: *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1819.3, pp. 247–255.

Formosa, T., P. Eriksson, J. Wittmeyer, J. Ginn, Y. Yu, and D. J. Stillman (2001). "Spt16–Pob3 and the HMG protein Nhp6 combine to form the nucleosome-binding factor SPN". In: *The EMBO journal* 20.13, pp. 3506–3517.

Förster, T. (1948). "Zwischenmolekulare energiewanderung und fluoreszenz". In: *Annalen der physik* 437.1-2, pp. 55–75.

Georgieva, S., E. Nabirochkina, F. J. Dilworth, H. Eickhoff, and P. Becker (2001). "The Novel Transcription Factor e (y) 2 Interacts with TAF II 40 and Potentiates Transcription Activation on Chromatin Templates". In: *Molecular and Cellular Biology* 21.15, pp. 5223–5231.

Gibson, T. J., M. Seiler, and R. A. Veitia (2013). "The transience of transient overexpression". In: *Nature methods* 10.8, pp. 715–721.

González-Aguilera, C., C. Tous, B. Gómez-González, P. Huertas, R. Luna, and A. Aguilera (2008). "The THP1-SAC3-SUS1-CDC31 complex works in transcription elongation-mRNA export preventing RNA-mediated genome instability". In: *Molecular biology of the cell* 19.10, pp. 4310–4318.

Govind, C. K., F. Zhang, H. Qiu, K. Hofmeyer, and A. G. Hinnebusch (2007). "Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions". In: *Molecular cell* 25.1, pp. 31–42.

Grasser, M., C. M. Kane, T. Merkle, M. Melzer, J. Emmersen, and K. D. Grasser (2009). "Transcript elongation factor TFIIS is involved in Arabidopsis seed dormancy". In: *Journal of molecular biology* 386.3, pp. 598–611.

Greenham, K. and C. R. McClung (2015). "Integrating circadian dynamics with physiological processes in plants". In: *Nature Reviews. Genetics* 16.10, p. 598.

Gu, X., D. Jiang, Y. Wang, A. Bachmair, and Y. He (2009). "Repression of the floral transition via histone H2B monoubiquitination". In: *The Plant Journal* 57.3, pp. 522–533.

Gu, X.-L., H. Wang, H. Huang, and X.-F. Cui (2012). "SPT6L encoding a putative WG/GW-repeat protein regulates apical–basal polarity of embryo in Arabidopsis". In: *Molecular plant* 5.1, pp. 249–259.

Guerler, A., B. Govindarajoo, and Y. Zhang (2013). "Mapping monomeric threading to protein–protein structure prediction". In: *Journal of chemical information and modeling* 53.3, pp. 717–725.

Gurskiy, D., A. Orlova, N. Vorobyeva, E. Nabirochkina, A. Krasnov, Y. Shidlovskii, S. Georgieva, and D. Kopytova (2012). "The DUBm subunit Sgf11 is required for mRNA export and interacts with Cbp80 in Drosophila." In: *Nucleic acids research* 40.21, pp. 1–12.

Han, Y., J. Luo, J. Ranish, and S. Hahn (2014). "Architecture of the Saccharomyces cerevisiae SAGA transcription coactivator complex". In: *The EMBO journal* 33.21, pp. 2534–2546.

Hark, A. T., K. E. Vlachonasios, K. A. Pavangadkar, S. Rao, H. Gordon, I.-D. Adamakis, A. Kaldis, M. F. Thomashow, and S. J. Triezenberg (2009). "Two Arabidopsis orthologs of

the transcriptional coactivator ADA2 have distinct biological functions". In: *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1789.2, pp. 117–124.

Hartzog, G. A. and J. Fu (2013). "The Spt4–Spt5 complex: a multi-faceted regulator of transcription elongation". In: *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1829.1, pp. 105–115.

Hatier, J.-H. B. and K. S. Gould (2008). "Foliar anthocyanins as modulators of stress signals". In: *Journal of Theoretical Biology* 253.3, pp. 625–627.

Hazelwood, K. L., E. B. Ramko, A. P. Ozarowska, S. G. Olenych, P. N. Worthy, A. Guan, C. S. Murphy, and M. W. Davidson (2008). "Searching the fluorescent protein color palette for new FRET pairs". In: *Proc. of SPIE Vol.* Vol. 6868, pp. 68680C–1.

He, Y., M. R. Doyle, and R. M. Amasino (2004). "PAF1-complex-mediated histone methylation of FLOWERING LOCUS C chromatin is required for the vernalization-responsive, winter-annual habit in Arabidopsis". In: *Genes & development* 18.22, pp. 2774–2784.

Hellemans, J., G. Mortier, A. De Paepe, F. Speleman, and J. Vandesompele (2007). "qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data". In: *Genome biology* 8.2, R19.

Hellens, R. P., E. A. Edwards, N. R. Leyland, S. Bean, and P. M. Mullineaux (2000). "pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation". In: *Plant molecular biology* 42.6, pp. 819–832.

Henry, K. W. et al. (2003). "Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8". In: *Genes & development* 17.21, pp. 2648–2663.

Hepworth, J. and C. Dean (2015). "Flowering Locus Cs lessons: conserved chromatin switches underpinning developmental timing and adaptation". In: *Plant physiology* 168.4, pp. 1237–1245.

Herskowitz, I. (1987). "Functional inactivation of genes by dominant negative mutations". In: *Nature* 329.6136, pp. 219–222.

Herzel, L., D. S. M. Ottoz, T. Alpert, and K. M. Neugebauer (2017). "Splicing and transcription touch base: co-transcriptional spliceosome assembly and function". In: *Nature Reviews Molecular Cell Biology*.

Hodges, A. J., L. M. Gloss, and J. J. Wyrick (2017). "Residues in the Nucleosome Acidic Patch Regulate Histone Occupancy and Are Important for FACT Binding in Saccharomyces cerevisiae". In: *Genetics*, genetics–117.

Houtsmuller, A. B. and W. Vermeulen (2001). "Macromolecular dynamics in living cell nuclei revealed by fluorescence redistribution after photobleaching". In: *Histochemistry and cell biology* 115.1, pp. 13–21.

Huang, X. and W. Miller (1991). "A time-efficient, linear-space local similarity algorithm". In: *Advances in Applied Mathematics* 12.3, pp. 337–357.

Hwang, W. W., S. Venkatasubrahmanyam, A. G. Ianculescu, A. Tong, C. Boone, and H. D. Madhani (2003). "A conserved RING finger protein required for histone H2B monoubiquitination and cell size control". In: *Molecular cell* 11.1, pp. 261–266.

Ikeda, Y., Y. Kinoshita, D. Susaki, Y. Ikeda, M. Iwano, S. Takayama, T. Higashiyama, T. Kakutani, and T. Kinoshita (2011). "HMG domain containing SSRP1 is required for DNA demethylation and genomic imprinting in Arabidopsis". In: *Developmental cell* 21.3, pp. 589–596.

Ingvarsdottir, K., N. J. Krogan, N. T. Emre, A. Wyce, N. J. Thompson, A. Emili, T. R. Hughes, J. F. Greenblatt, and S. L. Berger (2005). "H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the Saccharomyces cerevisiae SAGA complex". In: *Molecular and cellular biology* 25.3, pp. 1162–1172.

Isono, E. and M.-K. Nagel (2014). "Deubiquitylating enzymes and their emerging role in plant biology". In: *Frontiers in plant science* 5.

Iwabuchi, K., B. Li, P. Bartel, and S. Fields (1993). "Use of the two-hybrid system to identify the domain of p53 involved in oligomerization." In: *Oncogene* 8.6, pp. 1693–1696.

Jamai, A., A. Puglisi, and M. Strubin (2009). "Histone chaperone spt16 promotes redeposition of the original h3-h4 histones evicted by elongating RNA polymerase". In: *Molecular cell* 35.3, pp. 377–383.

Jani, D., S. Lutz, E. Hurt, R. a. Laskey, M. Stewart, and V. O. Wickramasinghe (2012). "Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export." In: *Nucleic acids research* 40.10, pp. 4562–73.

Jani, D., S. Lutz, N. J. Marshall, T. Fischer, A. Köhler, A. M. Ellisdon, E. Hurt, and M. Stewart (2009). "Sus1, Cdc31, and the Sac3 CID region form a conserved interaction platform that promotes nuclear pore association and mRNA export". In: *Molecular cell* 33.6, pp. 727–737.

Jefferson, R. A. (1987). "Assaying chimeric genes in plants: the GUS gene fusion system". In: *Plant molecular biology reporter* 5.4, pp. 387–405.

Jonge, W. J. de, E. O'Duibhir, P. Lijnzaad, D. van Leenen, M. J. G. Koerkamp, P. Kemmeren, and F. C. Holstege (2016). "Molecular mechanisms that distinguish TFIID housekeeping from regulatable SAGA promoters". In: *The EMBO journal*, e201695621.

Jonkers, I. and J. T. Lis (2015). "Getting up to speed with transcription elongation by RNA polymerase II". In: *Nature reviews Molecular cell biology* 16.3, pp. 167–177.

Karpova, T., C. Baumann, L. He, X. Wu, A. Grammer, P. Lipsky, G. Hager, and J. McNally (2003). "Fluorescence resonance energy transfer from cyan to yellow fluorescent protein detected by acceptor photobleaching using confocal microscopy and a single laser". In: *Journal of microscopy* 209.1, pp. 56–70.

Keller, D. M. and H. Lu (2002). "p53 serine 392 phosphorylation increases after UV through induction of the assembly of the CK2· hSPT16· SSRP1 complex". In: *Journal of Biological Chemistry* 277.51, pp. 50206–50213.

Kemble, D. J., L. L. McCullough, F. G. Whitby, T. Formosa, and C. P. Hill (2015). "FACT disrupts nucleosome structure by binding H2A-H2B with conserved peptide motifs". In: *Molecular cell* 60.2, pp. 294–306.

Kimura, H. (2005). "Histone dynamics in living cells revealed by photobleaching". In: *DNA repair* 4.8, pp. 939–950.

Kimura, H. and P. R. Cook (2001). "Kinetics of core histones in living human cells". In: *The Journal of cell biology* 153.7, pp. 1341–1354.

Köhler, A. and E. Hurt (2007). "Exporting RNA from the nucleus to the cytoplasm". In: *Nature reviews. Molecular cell biology* 8.10, p. 761.

Köhler, A., M. Schneider, G. G. Cabal, U. Nehrbass, and E. Hurt (2008). "Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export". In: *Nature cell biology* 10.6, p. 707.

Köhler, A., E. Zimmerman, M. Schneider, E. Hurt, and N. Zheng (2010). "Structural basis for assembly and activation of the heterotetrameric SAGA histone H2B deubiquitinase module". In: *Cell* 141.4, pp. 606–617.

Koltowska, K., H. Apitz, D. Stamataki, E. M. Hirst, H. Verkade, I. Salecker, and E. A. Ober (2013). "Ssrp1a controls organogenesis by promoting cell cycle progression and RNA synthesis". In: *Development* 140.9, pp. 1912–1918.

Komili, S. and P. A. Silver (2008). "Coupling and coordination in gene expression processes: a systems biology view". In: *Nature reviews. Genetics* 9.1, p. 38.

Koncz, C. et al. (2012). "The spliceosome-activating complex: molecular mechanisms underlying the function of a pleiotropic regulator". In: *Frontiers in plant science* 3.

Kopytova, D. V., A. N. Krasno, A. V. Orlova, D. Y. Gurskiy, E. N. Nabirochkina, S. G. Georgieva, and Y. V. Shidlovskii (2010a). "Eny2: couple, triple... more?" In: *Cell Cycle* 9.3, pp. 479–481.

Kopytova, D. V., A. V. Orlova, A. N. Krasnov, D. Y. Gurskiy, J. V. Nikolenko, E. N. Nabirochkina, Y. V. Shidlovskii, and S. G. Georgieva (2010b). "Multifunctional factor ENY2 is associated with the THO complex and promotes its recruitment onto nascent mRNA". In: *Genes and Development* 24.1, pp. 86–96.

Koutelou, E., C. L. Hirsch, and S. Y. Dent (2010). "Multiple faces of the SAGA complex". In: *Current opinion in cell biology* 22.3, pp. 374–382.

Kovinich, N., G. Kayanja, A. Chanoca, K. Riedl, M. S. Otegui, and E. Grotewold (2014). "Not all anthocyanins are born equal: distinct patterns induced by stress in Arabidopsis". In: *Planta* 240.5, pp. 931–940.

Krasnov, A. N., M. M. Kurshakova, V. E. Ramensky, P. V. Mardanov, E. N. Nabirochkina, and S. G. Georgieva (2005). "A retrocopy of a gene can functionally displace the source gene in evolution." In: *Nucleic acids research* 33.20, pp. 6654–61.

Kudo, T. et al. (2016). "Identification of reference genes for quantitative expression analysis using large-scale RNA-seq data of Arabidopsis thaliana and model crop plants". In: *Genes & genetic systems* 91.2, pp. 111–125.

Kurshakova, M. M. et al. (2007). "SAGA and a novel Drosophila export complex anchor efficient transcription and mRNA export to NPC." In: *The EMBO journal* 26.24, pp. 4956–65.

Kwak, H. and J. T. Lis (2013). "Control of transcriptional elongation". In: *Annual review of genetics* 47, pp. 483–508.

Lang, G., J. Bonnet, D. Umlauf, K. Karmodiya, J. Koffler, M. Stierle, D. Devys, and L. Tora (2011). "The tightly controlled deubiquitination activity of the human SAGA complex differentially modifies distinct gene regulatory elements". In: *Molecular and cellular biology* 31.18, pp. 3734–3744.

Lawit, S. J., K. O-Grady, W. B. Gurley, and E. Czarnecka-Verner (2007). "Yeast two-hybrid map of Arabidopsis TFIID". In: *Plant molecular biology* 64.1-2, pp. 73–87.

Lee, K. K., S. K. Swanson, L. Florens, M. P. Washburn, and J. L. Workman (2009). "Yeast Sgf73/Ataxin-7 serves to anchor the deubiquitination module into both SAGA and Slik (SALSA) HAT complexes". In: *Epigenetics & chromatin* 2.1, p. 2.

Lei, Y., L. Lu, H.-Y. Liu, S. Li, F. Xing, and L.-L. Chen (2014). "CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants". In: *Molecular plant* 7.9, pp. 1494–1496.

Lepiniec, L., I. Debeaujon, J.-M. Routaboul, A. Baudry, L. Pourcel, N. Nesi, and M. Caboche (2006). "Genetics and biochemistry of seed flavonoids". In: *Annu. Rev. Plant Biol.* 57, pp. 405–430.

Letunic, I. and P. Bork (2017). "20 years of the SMART protein domain annotation resource". In: *Nucleic Acids Research*.

Li, B. and S. Fields (1993). "Identification of mutations in p53 that affect its binding to SV40 large T antigen by using the yeast two-hybrid system." In: *The FASEB Journal* 7.10, pp. 957–963.

Li, B., M. Carey, and J. L. Workman (2007a). "The role of chromatin during transcription". In: *Cell* 128.4, pp. 707–719.

Li, X., C. W. Seidel, L. T. Szerszen, J. J. Lange, J. L. Workman, and S. M. Abmayr (2017). "Enzymatic modules of the SAGA chromatin-modifying complex play distinct roles in Drosophila gene expression and development". In: *Genes & Development* 31.15, pp. 1588–1600.

Li, Y., S. X. Zeng, I. Landais, and H. Lu (2007b). "Human SSRP1 has Spt16-dependent and-independent roles in gene transcription". In: *Journal of Biological Chemistry* 282.10, pp. 6936–6945.

Lichota, J. and K. D. Grasser (2001). "Differential chromatin association and nucleosome binding of the maize HMGA, HMGB, and SSRP1 proteins". In: *Biochemistry* 40.26, pp. 7860–7867.

Lim, S., J. Kwak, M. Kim, and D. Lee (2013). "Separation of a functional deubiquitylating module from the SAGA complex by the proteasome regulatory particle". In: *Nature communications* 4, p. 2641.

Lippincott-Schwartz, J., E. Snapp, and A. Kenworthy (2001). "Studying protein dynamics in living cells". In: *Nature reviews. Molecular cell biology* 2.6, p. 444.

Liu, Y., M. Koornneef, and W. J. Soppe (2007). "The absence of histone H2B monoubiquitination in the Arabidopsis hub1 (rdo4) mutant reveals a role for chromatin remodeling in seed dormancy". In: *The Plant Cell* 19.2, pp. 433–444.

Lolas, I. B., K. Himanen, J. T. Grønlund, C. Lynggaard, A. Houben, M. Melzer, M. Van Lijsebettens, and K. D. Grasser (2010). "The transcript elongation factor FACT affects Arabidopsis vegetative and reproductive development and genetically interacts with HUB1/2". In: *The Plant Journal* 61.4, pp. 686–697.

Lu, Q. et al. (2010). "Arabidopsis homolog of the yeast TREX-2 mRNA export complex: components and anchoring nucleoporin." In: *The Plant journal : for cell and molecular biology* 61.2, pp. 259–70.

Luger, K., M. L. Dechassa, and D. J. Tremethick (2012). "New insights into nucleosome and chromatin structure: an ordered state or a disordered affair?" In: *Nature reviews Molecular cell biology* 13.7, pp. 436–447.

Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond (1997). "Crystal structure of the nucleosome core particle at 2.8 angstrom resolution". In: *Nature* 389.6648, p. 251.

Mahrez, W., M. S. T. Arellano, J. Moreno-Romero, M. Nakamura, H. Shu, P. Nanni, C. Köhler, W. Gruissem, and L. Hennig (2016a). "H3K36ac is an evolutionary conserved plant histone modification that marks active genes". In: *Plant physiology*, pp–01744.

Mahrez, W. et al. (2016b). "BRR2a affects flowering time via FLC splicing". In: *PLoS genetics* 12.4, e1005924.

Malarkey, C. S. and M. E. Churchill (2012). "The high mobility group box: the ultimate utility player of a cell". In: *Trends in biochemical sciences* 37.12, pp. 553–562.

Manders, E., F. Verbeek, and J. Aten (1993). "Measurement of co-localization of objects in dual-colour confocal images". In: *Journal of microscopy* 169.3, pp. 375–382.

Maniatis, T. and R. Reed (2002). "An extensive network of coupling among gene expression machines". In: *Nature* 416.6880, pp. 499–506.

Mao, Y., K. A. Pavangadkar, M. F. Thomashow, and S. J. Triezenberg (2006). "Physical and functional interactions of Arabidopsis ADA2 transcriptional coactivator proteins with the acetyltransferase GCN5 and with the cold-induced transcription factor CBF1". In: *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* 1759.1, pp. 69–79.

Marquardt, S., O. Raitskin, Z. Wu, F. Liu, Q. Sun, and C. Dean (2014). "Functional consequences of splicing of the antisense transcript COOLAIR on FLC transcription". In: *Molecular cell* 54.1, pp. 156–165.

Mayer, A., M. Lidschreiber, M. Siebert, K. Leike, J. Söding, and P. Cramer (2010). "Uniform transitions of the general RNA polymerase II transcription complex". In: *Nature structural & molecular biology* 17.10, pp. 1272–1278.

McCullough, L., R. Rawlins, A. Olsen, H. Xin, D. J. Stillman, and T. Formosa (2011). "Insight into the mechanism of nucleosome reorganization from histone mutants that suppress defects in the FACT histone chaperone". In: *Genetics* 188.4, pp. 835–846.

Michaels, S. D. and R. M. Amasino (1999). "FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering". In: *The Plant Cell* 11.5, pp. 949–956.

Michaels, S. D., E. Himelblau, S. Y. Kim, F. M. Schomburg, and R. M. Amasino (2005). "Integration of flowering signals in winter-annual Arabidopsis". In: *Plant Physiology* 137.1, pp. 149–156.

Mohan, R. D., G. Dialynas, V. M. Weake, J. Liu, S. Martin-Brown, L. Florens, M. P. Washburn, J. L. Workman, and S. M. Abmayr (2014). "Loss of Drosophila Ataxin-7, a SAGA subunit, reduces H2B ubiquitination and leads to neural and retinal degeneration". In: *Genes & development* 28.3, pp. 259–272.

Mohibullah, N. and S. Hahn (2008). "Site-specific cross-linking of TBP in vivo and in vitro reveals a direct functional interaction with the SAGA subunit Spt3". In: *Genes & development* 22.21, pp. 2994–3006.

Monaghan, J., F. Xu, M. Gao, Q. Zhao, K. Palma, C. Long, S. Chen, Y. Zhang, and X. Li (2009). "Two Prp19-like U-box proteins in the MOS4-associated complex play redundant roles in plant innate immunity". In: *PLoS pathogens* 5.7, e1000526.

Monaghan, J., F. Xu, S. Xu, Y. Zhang, and X. Li (2010). "Two putative RNA-binding proteins function with unequal genetic redundancy in the MOS4-associated complex". In: *Plant physiology* 154.4, pp. 1783–1793.

Moore, M. J. and N. J. Proudfoot (2009). "Pre-mRNA processing reaches back totranscription and ahead to translation". In: *Cell* 136.4, pp. 688–700.

Moraga, F. and F. Aquea (2015). "Composition of the SAGA complex in plants and its role in controlling gene expression in response to abiotic stresses". In: *Frontiers in plant science* 6.

Morgan, M. T., M. Haj-Yahya, A. E. Ringel, P. Bandi, A. Brik, and C. Wolberger (2016). "Structural basis for histone H2B deubiquitination by the SAGA DUB module". In: *Science* 351.6274, pp. 725–728.

Mortensen, S. (2012). "Molecular analysis of the transcript elongation factor TFIIS and the histone chaperone FACT in Arabidopsis thaliana". PhD thesis. University of Regensburg.

Mortensen, S. A. and K. D. Grasser (2014). "The seed dormancy defect of Arabidopsis mutants lacking the transcript elongation factor TFIIS is caused by reduced expression of the DOG1 gene". In: *FEBS letters* 588.1, pp. 47–51.

Mortensen, S. A., M. Sønderkær, C. Lynggaard, M. Grasser, K. L. Nielsen, and K. D. Grasser (2011). "Reduced expression of the DOG1 gene in Arabidopsis mutant seeds lacking the transcript elongation factor TFIIS". In: *FEBS letters* 585.12, pp. 1929–1933.

Musinova, Y. R., O. M. Lisitsyna, S. A. Golyshev, A. I. Tuzhikov, V. Y. Polyakov, and E. V. Sheval (2011). "Nucleolar localization/retention signal is responsible for transient accumulation of histone H2B in the nucleolus through electrostatic interactions". In: *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* 1813.1, pp. 27–38.

Nagy, Z. and L. Tora (2007). "Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation". In: *Oncogene* 26.37, p. 5341.

Oh, S., H. Zhang, P. Ludwig, and S. van Nocker (2004). "A mechanism related to the yeast transcriptional regulator Paf1c is required for expression of the Arabidopsis FLC/MAF MADS box gene family". In: *The Plant Cell* 16.11, pp. 2940–2953.

Orphanides, G., G. LeRoy, C.-H. Chang, D. S. Luse, and D. Reinberg (1998). "FACT, a factor that facilitates transcript elongation through nucleosomes". In: *Cell* 92.1, pp. 105–116.

Orphanides, G., W. Wei-Hua, W. S. Lane, M. Hampsey, and D. Reinberg (1999). "The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins". In: *Nature* 400.6741, p. 284.

Palma, K., Q. Zhao, Y. T. Cheng, D. Bi, J. Monaghan, W. Cheng, Y. Zhang, and X. Li (2007). "Regulation of plant innate immunity by three proteins in a complex conserved across the plant and animal kingdoms". In: *Genes & Development* 21.12, pp. 1484–1493.

Pankotai, T., O. Komonyi, L. Bodai, Z. Újfaludi, S. Muratoglu, A. Ciurciu, L. Tora, J. Szabad, and I. Boros (2005). "The homologous Drosophila transcriptional adaptors ADA2a and ADA2b are both required for normal development but have different functions". In: *Molecular and cellular biology* 25.18, pp. 8215–8227.

Pascual-García, P., C. K. Govind, E. Queralt, B. Cuenca-Bono, A. Llopis, S. Chavez, A. G. Hinnebusch, and S. Rodríguez-Navarro (2008). "Sus1 is recruited to coding regions and functions during transcription elongation in association with SAGA and TREX2". In: *Genes and Development* 22.20, pp. 2811–2822.

Pascual-García, P. and S. Rodríguez-Navarro (2009). "A tale of coupling, Sus1 function in transcription and mRNA export". In: *RNA biology* 6.2, pp. 141–144.

Pavri, R., B. Zhu, G. Li, P. Trojer, S. Mandal, A. Shilatifard, and D. Reinberg (2006). "Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II". In: *Cell* 125.4, pp. 703–717.

Pederson, T. (2001). "Protein mobility within the nucleus. what are the right moves?" In: *Cell* 104.5, pp. 635–638.

Perales, R. and D. Bentley (2009). "Cotranscriptionality: the transcription elongation complex as a nexus for nuclear transactions". In: *Molecular cell* 36.2, pp. 178–191.

Pérez-García, P., Y. Ma, M. J. Yanovsky, and P. Mas (2015). "Time-dependent sequestration of RVE8 by LNK proteins shapes the diurnal oscillation of anthocyanin biosynthesis". In: *Proceedings of the National Academy of Sciences* 112.16, pp. 5249–5253.

Pfab, A., W. Antosz, P. Holzinger, A. Bruckmann, J. Griesenbeck, and K. D. Grasser (2017). "Analysis of In Vivo Chromatin and Protein Interactions of Arabidopsis Transcript Elongation Factors". In: *Methods in molecular biology (Clifton, NJ)* 1629, p. 105.

Phair, R. D. and T. Misteli (2000). "High mobility of proteins in the mammalian cell nucleus". In: *Nature* 404.6778, p. 604.

Phair, R. D. et al. (2004). "Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins". In: *Molecular and cellular biology* 24.14, pp. 6393–6402.

Phatnani, H. P. and A. L. Greenleaf (2006). "Phosphorylation and functions of the RNA polymerase II CTD". In: *Genes & development* 20.21, pp. 2922–2936.

Piston, D. W. and G.-J. Kremers (2007). "Fluorescent protein FRET: the good, the bad and the ugly". In: *Trends in biochemical sciences* 32.9, pp. 407–414.

Pray-Grant, M. G., J. A. Daniel, D. Schieltz, J. R. Yates III, and P. A. Grant (2005). "Chd1 chromodomain links histone H3 methylation with SAGA-and SLIK-dependent acetylation". In: *Nature* 433.7024, p. 434.

Prelich, G. (2012). "Gene overexpression: uses, mechanisms, and interpretation". In: *Genetics* 190.3, pp. 841–854.

Rabut, G., V. Doye, and J. Ellenberg (2004). "Mapping the dynamic organization of the nuclear pore complex inside single living cells". In: *Nature cell biology* 6.11, p. 1114.

Rajagopala, S. V., P. Sikorski, J. H. Caufield, A. Tovchigrechko, and P. Uetz (2012). "Studying protein complexes by the yeast two-hybrid system". In: *Methods* 58.4, pp. 392–399.

Rapsomaniki, M. A., P. Kotsantis, I.-E. Symeonidou, N.-N. Giakoumakis, S. Taraviras, and Z. Lygerou (2012). "easyFRAP: an interactive, easy-to-use tool for qualitative and quantitative analysis of FRAP data". In: *Bioinformatics* 28.13, pp. 1800–1801.

Reinberg, D. and R. J. Sims (2006). "de FACTo nucleosome dynamics". In: *Journal of Biological Chemistry* 281.33, pp. 23297–23301.

Riethoven, J.-J. M. (2010). "Regulatory regions in DNA: promoters, enhancers, silencers, and insulators". In: *Computational Biology of Transcription Factor Binding*, pp. 33–42.

Robzyk, K., J. Recht, and M. A. Osley (2000). "Rad6-dependent ubiquitination of histone H2B in yeast". In: *Science* 287.5452, pp. 501–504.

Rodríguez-Navarro, S. (2009). "Insights into SAGA function during gene expression". In: *EMBO reports* 10.8, pp. 843–850.

Rodríguez-Navarro, S., T. Fischer, M. J. Luo, O. Antúnez, S. Brettschneider, J. Lechner, J. E. Pérez-Ortín, R. Reed, and E. Hurt (2004). "Sus1, a Functional Component of the SAGA

Histone Acetylase Complex and the Nuclear Pore-Associated mRNA Export Machinery". In: *Cell* 116.1, pp. 75–86.

Rose, A. B., T. Elfersi, G. Parra, and I. Korf (2008). "Promoter-proximal introns in Arabidopsis thaliana are enriched in dispersed signals that elevate gene expression". In: *The Plant Cell* 20.3, pp. 543–551.

Röttgers, K., N. M. Krohn, J. Lichota, C. Stemmer, T. Merkle, and K. D. Grasser (2000). "DNA-interactions and nuclear localisation of the chromosomal HMG domain protein SSRP1 from maize". In: *The Plant Journal* 23.3, pp. 395–405.

Ruone, S., A. R. Rhoades, and T. Formosa (2003). "Multiple Nhp6 molecules are required to recruit Spt16-Pob3 to form yFACT complexes and to reorganize nucleosomes". In: *Journal of Biological Chemistry* 278.46, pp. 45288–45295.

Saito, K., K. Yonekura-Sakakibara, R. Nakabayashi, Y. Higashi, M. Yamazaki, T. Tohge, and A. R. Fernie (2013). "The flavonoid biosynthetic pathway in Arabidopsis: structural and genetic diversity". In: *Plant Physiology and Biochemistry* 72, pp. 21–34.

Saldi, T., M. A. Cortazar, R. M. Sheridan, and D. L. Bentley (2016). "Coupling of RNA polymerase II transcription elongation with pre-mRNA splicing". In: *Journal of molecular biology* 428.12, pp. 2623–2635.

Samara, N. L., A. B. Datta, C. E. Berndsen, X. Zhang, T. Yao, R. E. Cohen, and C. Wolberger (2010). "Structural insights into the assembly and function of the SAGA deubiquitinating module". In: *Science* 328.5981, pp. 1025–1029.

Samara, N. L. and C. Wolberger (2011). "A new chapter in the transcription SAGA". In: *Current opinion in structural biology* 21.6, pp. 767–774.

Sambrook, J., E. F. Fritsch, T. Maniatis, et al. (1989). *Molecular cloning: a laboratory manual.* Ed. 2. Cold spring harbor laboratory press.

Saunders, A., L. J. Core, and J. T. Lis (2006). "Breaking barriers to transcription elongation". In: *Nature reviews. Molecular cell biology* 7.8, p. 557.

Schlesinger, M. B. and T. Formosa (2000). "POB3 is required for both transcription and replication in the yeast Saccharomyces cerevisiae". In: *Genetics* 155.4, pp. 1593–1606.

Selth, L. A., S. Sigurdsson, and J. Q. Svejstrup (2010). "Transcript elongation by RNA polymerase II". In: *Annual review of biochemistry* 79, pp. 271–293.

Sheldon, C. C., J. E. Burn, P. P. Perez, J. Metzger, J. A. Edwards, W. J. Peacock, and E. S. Dennis (1999). "The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation". In: *The Plant Cell* 11.3, pp. 445–458.

Shi, M.-Z. and D.-Y. Xie (2014). "Biosynthesis and metabolic engineering of anthocyanins in Arabidopsis thaliana". In: *Recent patents on biotechnology* 8.1, pp. 47–60.

Sievers, F. et al. (2011). "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega". In: *Molecular systems biology* 7.1, p. 539.

Simpson, G. G. and C. Dean (2002). "Arabidopsis, the Rosetta stone of flowering time?" In: *Science* 296.5566, pp. 285–289.

Sims, R. J., R. Belotserkovskaya, and D. Reinberg (2004). "Elongation by RNA polymerase II: the short and long of it". In: *Genes & development* 18.20, pp. 2437–2468.

Soerensen, B. B. et al. (2017). "The Arabidopsis THO/TREX component TEX1 functionally interacts with MOS11 and modulates mRNA export and alternative splicing events". In: *Plant Molecular Biology* 93.3, pp. 283–298.

Spedale, G., H. T. M. Timmers, and W. P. Pijnappel (2012). "ATAC-king the complexity of SAGA during evolution". In: *Genes & development* 26.6, pp. 527–541.

Sprague, B. L., R. L. Pego, D. A. Stavreva, and J. G. McNally (2004). "Analysis of binding reactions by fluorescence recovery after photobleaching". In: *Biophysical journal* 86.6, pp. 3473–3495.

Squazzo, S. L., P. J. Costa, D. L. Lindstrom, K. E. Kumer, R. Simic, J. L. Jennings, A. J. Link, K. M. Arndt, and G. A. Hartzog (2002). "The Paf1 complex physically and function-

ally associates with transcription elongation factors in vivo". In: *The EMBO journal* 21.7, pp. 1764–1774.

Sridhar, V. V., A. Kapoor, K. Zhang, J. Zhu, T. Zhou, P. M. Hasegawa, R. A. Bressan, and Z. Jian-Kang (2007). "Control of DNA methylation and heterochromatic silencing by histone H2B deubiquitination". In: *Nature* 447.7145, p. 735.

Srivastava, R., K. M. Rai, B. Pandey, S. P. Singh, and S. V. Sawant (2015). "Spt-Ada-Gcn5-Acetyltransferase (SAGA) complex in plants: Genome wide identification, evolutionary conservation and functional determination". In: *PLoS ONE* 10.8, pp. 1–30.

Stauffer, E. and A. Maizel (2014). "Post-transcriptional regulation in root development". In: *Wiley Interdisciplinary Reviews: RNA* 5.5, pp. 679–696.

Stegeman, R., P. J. Spreacker, S. K. Swanson, R. Stephenson, L. Florens, M. P. Washburn, and V. M. Weake (2016). "The Spliceosomal Protein SF3B5 is a Novel Component of Drosophila SAGA that Functions in Gene Expression Independent of Splicing". In: *Journal of molecular biology* 428.18, pp. 3632–3649.

Sterner, D. E., P. A. Grant, S. M. Roberts, L. J. Duggan, R. Belotserkovskaya, L. A. Pacella, F. Winston, J. L. Workman, and S. L. Berger (1999). "Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction". In: *Molecular and cellular biology* 19.1, pp. 86–98.

Stillman, D. J. (2010). "Nhp6: a small but powerful effector of chromatin structure in Saccharomyces cerevisiae". In: *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1799.1, pp. 175–180.

Stockinger, E. J., Y. Mao, M. K. Regier, S. J. Triezenberg, and M. F. Thomashow (2001). "Transcriptional adaptor and histone acetyltransferase proteins in Arabidopsis and their interactions with CBF1, a transcriptional activator involved in cold-regulated gene expression". In: *Nucleic acids research* 29.7, pp. 1524–1533.

Štros, M., D. Launholt, and K. D. Grasser (2007). "The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins". In: *Cellular and Molecular Life Sciences* 64.19, pp. 2590–2606.

Sun, M., L. Larivière, S. Dengl, A. Mayer, and P. Cramer (2010). "A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA polymerase II C-terminal repeat domain (CTD)". In: *Journal of Biological Chemistry* 285.53, pp. 41597–41603.

Sundaresan, V., P. Springer, T. Volpe, S. Haward, J. D. Jones, C. Dean, H. Ma, and R. Martienssen (1995). "Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements." In: *Genes & development* 9.14, pp. 1797–1810.

Supek, F., M. Bošnjak, N. Škunca, and T. Šmuc (2011). "REVIGO summarizes and visualizes long lists of gene ontology terms". In: *PloS one* 6.7, e21800.

Thomas, J. O. and A. A. Travers (2001). "HMG1 and 2, and related architectural DNA-binding proteins". In: *Trends in biochemical sciences* 26.3, pp. 167–174.

Thomas, M. C. and C.-M. Chiang (2006). "The general transcription machinery and general cofactors". In: *Critical reviews in biochemistry and molecular biology* 41.3, pp. 105–178.

Tian, T., Y. Liu, H. Yan, Q. You, X. Yi, Z. Du, W. Xu, and Z. Su (2017). "agriGO v2. 0: a GO analysis toolkit for the agricultural community, 2017 update". In: *Nucleic Acids Research*.

Tillemans, V., L. Dispa, C. Remacle, M. Collinge, and P. Motte (2005). "Functional distribution and dynamics of Arabidopsis SR splicing factors in living plant cells". In: *The Plant Journal* 41.4, pp. 567–582.

Tomson, B. N. and K. M. Arndt (2013). "The many roles of the conserved eukaryotic Paf1 complex in regulating transcription, histone modifications, and disease states". In: *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1829.1, pp. 116–126.

Tramier, M., M. Zahid, J.-C. Mevel, M.-J. Masse, and M. Coppey-Moisan (2006). "Sensitivity of CFP/YFP and GFP/mCherry pairs to donor photobleaching on FRET determination

by fluorescence lifetime imaging microscopy in living cells". In: *Microscopy research and technique* 69.11, pp. 933–939.

Ülker, B. et al. (2008). "Getting the most out of publicly available T-DNA insertion lines". In: *The Plant Journal* 56.4, pp. 665–677.

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

Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, and S. G. Rozen (2012). "Primer3.new capabilities and interfaces". In: *Nucleic acids research* 40.15, e115–e115.

Uthe, H., J. T. Vanselow, and A. Schlosser (2017). "Proteomic Analysis of the Mediator Complex Interactome in Saccharomyces cerevisiae". In: *Scientific Reports* 7.

Van Leene, J., D. Eeckhout, G. Persiau, E. Van De Slijke, J. Geerinck, G. Van Isterdael, E. Witters, and G. De Jaeger (2011). "Isolation of transcription factor complexes from Arabidopsis cell suspension cultures by tandem affinity purification". In: *Plant Transcription Factors: Methods and Protocols*, pp. 195–218.

Van Leene, J., E. Witters, D. Inzé, and G. De Jaeger (2008). "Boosting tandem affinity purification of plant protein complexes". In: *Trends in plant science* 13.10, pp. 517–520.

Van Leene, J. et al. (2007). "A tandem affinity purification-based technology platform to study the cell cycle interactome in Arabidopsis thaliana". In: *Molecular & Cellular Proteomics* 6.7, pp. 1226–1238.

Van Leene, J. et al. (2015). "An improved toolbox to unravel the plant cellular machinery by tandem affinity purification of Arabidopsis protein complexes". In: *Nature protocols* 10.1, pp. 169–187.

Van Lijsebettens, M. and K. D. Grasser (2014). "Transcript elongation factors: shaping transcriptomes after transcript initiation". In: *Trends in plant science* 19.11, pp. 717–726.

Venkatesh, S. and J. L. Workman (2015). "Histone exchange, chromatin structure and the regulation of transcription". In: *Nature reviews. Molecular cell biology* 16.3, p. 178.

Vlachonasios, K. E., M. F. Thomashow, and S. J. Triezenberg (2003). "Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect Arabidopsis growth, development, and gene expression". In: *The Plant Cell* 15.3, pp. 626–638.

Vosnakis, N., M. Koch, E. Scheer, P. Kessler, Y. Mély, P. Didier, and L. Tora (2017). "Coactivators and general transcription factors have two distinct dynamic populations dependent on transcription". In: *The EMBO Journal*, e201696035.

Wan, R., C. Yan, R. Bai, G. Huang, and Y. Shi (2016). "Structure of a yeast catalytic step I spliceosome at 3.4 A resolution". In: *Science* 353.6302, pp. 895–904. arXiv: `arXiv:1011.1669v3`.

Wang, N., Y. Cui, Y. Liu, H. Fan, J. Du, Z. Huang, Y. Yuan, H. Wu, and H.-Q. Ling (2013). "Requirement and functional redundancy of Ib subgroup bHLH proteins for iron deficiency responses and uptake in Arabidopsis thaliana". In: *Molecular plant* 6.2, pp. 503–513.

Wang, Z.-P., H.-L. Xing, L. Dong, H.-Y. Zhang, C.-Y. Han, X.-C. Wang, and Q.-J. Chen (2015). "Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation". In: *Genome biology* 16.1, p. 144.

Wang, Z.-W., Z. Wu, O. Raitskin, Q. Sun, and C. Dean (2014). "Antisense-mediated FLC transcriptional repression requires the P-TEFb transcription elongation factor". In: *Proceedings of the National Academy of Sciences* 111.20, pp. 7468–7473.

Waterhouse, A. M., J. B. Procter, D. M. Martin, M. Clamp, and G. J. Barton (2009). "Jalview Version 2 - a multiple sequence alignment editor and analysis workbench". In: *Bioinformatics* 25.9, pp. 1189–1191.

Weake, V. M., K. K. Lee, S. Guelman, C.-H. Lin, C. Seidel, S. M. Abmayr, and J. L. Workman (2008a). "SAGA-mediated H2B deubiquitination controls the development of neuronal connectivity in the Drosophila visual system". In: *The EMBO journal* 27.2, pp. 394–405.

Weake, V. M. and J. L. Workman (2008b). "Histone ubiquitination: triggering gene activity". In: *Molecular cell* 29.6, pp. 653–663.

Weake, V. M. and J. L. Workman (2012). "SAGA function in tissue-specific gene expression". In: *Trends in Cell Biology* 22.4, pp. 177–184. arXiv: NIHMS150003.

Weber, C. M. and S. Henikoff (2014). "Histone variants: dynamic punctuation in transcription". In: *Genes & development* 28.7, pp. 672–682.

Weidtkamp-Peters, S. and Y. Stahl (2017). "The Use of FRET/FLIM to Study Proteins Interacting with Plant Receptor Kinases". In: *Plant Receptor Kinases: Methods and Protocols*, pp. 163–175.

Will, C. L. and R. Lührmann (2011). "Spliceosome structure and function". In: *Cold Spring Harbor perspectives in biology* 3.7, a003707.

Winkler, D. D. and K. Luger (2011a). "The histone chaperone FACT: structural insights and mechanisms for nucleosome reorganization". In: *Journal of Biological Chemistry* 286.21, pp. 18369–18374.

Winkler, D. D., U. M. Muthurajan, A. R. Hieb, and K. Luger (2011b). "Histone chaperone FACT coordinates nucleosome interaction through multiple synergistic binding events". In: *Journal of Biological Chemistry* 286.48, pp. 41883–41892.

Wittmeyer, J. and T. Formosa (1997). "The Saccharomyces cerevisiae DNA polymerase alpha catalytic subunit interacts with Cdc68/Spt16 and with Pob3, a protein similar to an HMG1-like protein." In: *Molecular and Cellular Biology* 17.7, pp. 4178–4190.

Wittmeyer, J., L. Joss, and T. Formosa (1999). "Spt16 and Pob3 of Saccharomyces cerevisiae form an essential, abundant heterodimer that is nuclear, chromatin-associated, and copurifies with DNA polymerase α". In: *Biochemistry* 38.28, pp. 8961–8971.

Woloszynska, M., S. Le Gall, and M. Van Lijsebettens (2016). "Plant Elongator-mediated transcriptional control in a chromatin and epigenetic context". In: *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1859.8, pp. 1025–1033.

Wood, A. et al. (2003). "Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter". In: *Molecular cell* 11.1, pp. 267–274.

Workman, J. L. (2016). "It takes teamwork to modify chromatin". In: *Science* 351.6274, p. 667.

Wu, P.-Y. J., C. Ruhlmann, F. Winston, and P. Schultz (2004). "Molecular architecture of the S. cerevisiae SAGA complex". In: *Molecular cell* 15.2, pp. 199–208.

Wyce, A. et al. (2007). "H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex". In: *Molecular cell* 27.2, pp. 275–288.

Xin, H., S. Takahata, M. Blanksma, L. McCullough, D. J. Stillman, and T. Formosa (2009). "yFACT induces global accessibility of nucleosomal DNA without H2A-H2B displacement". In: *Molecular cell* 35.3, pp. 365–376.

Xu, L., R. Ménard, A. Berr, J. Fuchs, V. Cognat, D. Meyer, and W.-H. Shen (2009). "The E2 ubiquitin-conjugating enzymes, AtUBC1 and AtUBC2, play redundant roles and are involved in activation of FLC expression and repression of flowering in Arabidopsis thaliana". In: *The Plant Journal* 57.2, pp. 279–288.

Xu, Y., C. Bernecky, C.-T. Lee, K. C. Maier, B. Schwalb, D. Tegunov, J. M. Plitzko, H. Urlaub, and P. Cramer (2017a). "Architecture of the RNA polymerase II-Paf1C-TFIIS transcription elongation complex". In: *Nature communications* 8.

Xu, Z., K. Mahmood, and S. J. Rothstein (2017b). "ROS Induces Anthocyanin Production Via Late Biosynthetic Genes and Anthocyanin Deficiency Confers the Hypersensitivity to ROS-generating Stresses in Arabidopsis". In: *Plant and Cell Physiology*, pcx073.

Yan, C., J. Hang, R. Wan, M. Huang, C. C. Wong, and Y. Shi (2015). "Structure of a yeast spliceosome at 3.6-angstrom resolution". In: *Science* 349.6253, pp. 1182–1191.

Yan, N., J. H. Doelling, T. G. Falbel, A. M. Durski, and R. D. Vierstra (2000). "The ubiquitin-specific protease family from Arabidopsis. AtUBP1 and 2 are required for the resistance to the amino acid analog canavanine". In: *Plant Physiology* 124.4, pp. 1828–1843.

Yang, J., R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang (2015). "The I-TASSER Suite: protein structure and function prediction". In: *Nature methods* 12.1, pp. 7–8.

Yin, R., B. Messner, T. Faus-Kessler, T. Hoffmann, W. Schwab, M.-R. Hajirezaei, V. von Saint Paul, W. Heller, and A. R. Schäffner (2012). "Feedback inhibition of the general phenylpropanoid and flavonol biosynthetic pathways upon a compromised flavonol-3-O-glycosylation". In: *Journal of experimental botany* 63.7, pp. 2465–2478.

Yoh, S. M., H. Cho, L. Pickle, R. M. Evans, and K. A. Jones (2007). "The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export". In: *Genes & development* 21.2, pp. 160–174.

Yuan, Y., H. Wu, N. Wang, J. Li, W. Zhao, J. Du, D. Wang, and H.-Q. Ling (2008). "FIT interacts with AtbHLH38 and AtbHLH39 in regulating iron uptake gene expression for iron homeostasis in Arabidopsis". In: *Cell research* 18.3, p. 385.

Zhang, K., V. V. Sridhar, J. Zhu, A. Kapoor, and J.-K. Zhu (2007). "Distinctive core histone post-translational modification patterns in Arabidopsis thaliana". In: *PloS one* 2.11, e1210.

Zhao, Y. et al. (2008). "A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing." In: *Molecular cell* 29.1, pp. 92–101.

Zhou, W., Y. Zhu, A. Dong, and W.-H. Shen (2015). "Histone H2A/H2B chaperones: from molecules to chromatin-based functions in plant growth and development". In: *The Plant Journal* 83.1, pp. 78–95.