The role of *Arabidopsis* TFIIS in regulating transcript elongation: molecular and functional characterisation

DISSERTATION ZUR ERLANGUNG DES

DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT)

DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN

DER UNIVERSITÄT REGENSBURG

vorgelegt von

Wojciech Antosz

aus Krakau, Polen

im Februar 2019

Das Promotionsgesuch wurde eingereicht am: 15.02.2019

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

Unterschrift:

Wojciech Antosz

Dla Marianny i Izabeli

Contents

Abbreviations

A. thaliana	*Arabidopsis thaliana*
A. tumefaciens	*Agrobacterium tumefaciens*
ABS	Absolute value
Ala (A)	Alanine
AP-MS	Affinity purification coupled to mass spectrometry
Asp (D)	Aspartic acid
BH	Bridge helix
bp	Base pair
CaMV	Cauliflower mosaic virus
CBB	Coomassie brilliant blue
CCR4-NOT	Carbon catabolite repression 4-Negative on TATA
CDK9	Cyclin-dependent kinase 9
cDNA	Complementary DNA
CDS	Coding sequence
ChIP	Chromatin immunoprecipitation
CLSM	Confocal laser scanning microscopy
Col-0	Columbia-0
Cy3	Cyanine dye 3
DAPI	4',6-Diamidin-2-phenylindol
DAS	Days after stratification
DDO	Double drop out medium
DEG	Differentially expressed gene
DNA	Deoxyribonucleic acid
DOG1	DELAY OF GERMINATION 1
DSB	DNA double-strand break
DSIF	DRB-sensitivity inducing factor
ΔTFIIS	Truncated TFIIS
DTT	Di-thiotreitol
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetraacetic acid
EEC	Early elongation complex
eGFP	Enhanced green fluorescent protein
ELF7	EARLY FLOWERING 7
ELL	Lysine-rich in leukaemia
FACS	Fluorescence activated cell sorting
FACT	Facilitates chromatin transcription
FC	Fold change
FDR	False discovery rate
FLC	Flowering locus C
FPKM	Fragments per kilobase of exon per million fragments mapped
FRAP	Fluorescence recovery after photobleaching
FRET	Förster resonance energy transfer
GFP	Green fluorescent protein
Glu (E)	Glutamic acid
GO	Gene ontology
GRO-seq	Global run-on sequencing
GTF	General transcription factor
GUS	β-glucuronidase
HAT	Histone acetyltransferase
His	Histidine
HMT	Histone methyltransferase
HRP	Horseradish peroxidase

HU	Hydroxyurea
HygR	Hygromycin B resistance
IgG	Immunglobulin G
IP	Immunoprecipitation
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Initially transcribing complex
KanR	Kanamycin resistance
kDa	Kilo Dalton
LB	Luria Bertani LB
LD	Long day
LexA BD	Binding domain of LexA operon
LUC	Luciferase
mRNA	Messenger RNA
MS	Murashige-Skoog/mass spectroscopy
N. benthamiana	*Nicotiana benthamiana*
NAC	Nucleotide addition cycle
NASC	Nottingham Arabidopsis stock centre
NCP	Nucleosome core particle
ncRNA	Non-coding RNA
NELF	Negative elongation factor
NLS	Nuclear localization signal
NoAb	No antibody
NRPB	Nuclear RNA polymerase II
NTP	Nucleoside triphosphate
OD600	Optical density measured at 600 nm
ORF	Open reading frame
ORI	Origin of replication
P-TEFb	Positive transcription elongation factor b
PAF1-C	Polymerase-associated factor 1 complex
PAGE	Polyacrylamide gel electrophoresis
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction
PI	Pausing index
PIC	Pre-initiation complex
PMSF	Phenylmethylsulphonyl fluoride
polyA	Polyadenylation site
PPI	Protein-protein interaction
Pro	Proline
QDO	Quadruple drop out medium
qPCR	Quantitative polymerase chain reaction
R^2	Coefficient of determination
RB	Right border
RdDM	RNA-directed DNA methylation pathway
RNA	Ribonucleic acid
RNAP	RNA polymerase
RNAPII-CTD	carboxy terminal domain of RNAPII largest subunit
RNAPII-S2P	RNAPII-CTD phosphorylated at serine 2 position
RNAPII-S5P	RNAPII-CTD phosphorylated at serine 5 position
ROI	Region of interest
Rpm	Rounds per minute
RT	Room temperature
RT-PCR	Reverse transcription-polymerase chain reaction
SBP	Streptavidin binding peptide
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEA	Single enrichment analysis
Ser	Serine

Ser2	Serine 2 position of RNAPII-CTD
Ser5	Serine 5 position of RNAPII-CTD
SG/GS	Tag consisting of SBP and 2 x Protein G domains
SHR	Somatic homologous recombination
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
SPT16	Suppressor of Ty 16, FACT subunit
SPT5	Suppressor of Ty 5
SSRP1	Structure specific recognition protein 1, FACT subunit
t35S	CaMV 35S terminator
T-DNA	Transfer DNA
TDO	Triple drop out medium
TEC	Transcript elongation complex
TEF	Transcription elongation factor
TEF-seq	TEF-associated nascent elongating transcript sequencing
TES	Transcription termination site
TEV	Tobacco Etch Virus TEV-protease
TFIIS	Transcription factor IIS
TFIISmut	TFIIS with point mutations
Thr	Threonine
TL	Trigger loop
TPM	Transcripts per kilobase million
TRC	Transcription-replication conflict
tRNA	Transfer RNA
Tyr	Tyrosine
UTR	Untranslated region
v/v	Volume per volume
VP16	Acidic transactivation domain of human VP16
w/v	Weight per volume
WT	Wild type
XVE	LexA X, VP16 V and estrogen receptor E
Y2H	Yeast-two-Hybrid

1. Introduction

1.1 DNA transcription by RNA polymerases

The gene expression is fundamental process in eukaryotes primarily regulated at the level of mRNA synthesis on DNA template during transcription (Alberts et al., 2002). The genetic information can be accessed by transcribing DNA template into different RNA classes. This process is controlled by the action of various DNA dependent RNA polymerases I-III (RNAPI-III) which are multisubunit enzymes present in three domains of life (Vannini and Cramer, 2012; Werner and Grohmann, 2011). Each polymerase transcribes different classes of RNA with some minor overlaps. While RNAPI transcribes the 25S, 18S, and 5.8S rRNA, RNAPII drives the transcription of the messenger RNA (mRNA) microRNAs (miRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) and most non-coding RNAs. Finally, RNAPIII synthesises 5S rRNA, tRNA, 7SL RNA, U6 snRNA and a few other small stable RNAs (Arimbasseri et al., 2013; Dieci et al., 2007; Paule and White, 2000). All three RNA Polymerases have a common accessor and exhibit mechanistic and structural homology of their catalytic cores with some additional subunits identical for all three polymerases. In total, RNAP I, II and III contain 14, 12, and 17 subunits, respectively (Werner and Grohmann, 2011). Additionally, RNAP I-III share some of the general transcription factors (GTF) including TFIIB, TFIIE, and TFIIF and several other proteins related to basic transcriptional activity. Notably, each RNA polymerase additionally possess its specific interactome (Vannini and Cramer, 2012). Interestingly, two additional plant specific RNA polymerases, RNAPIV and RNAPV, have been identified (Pontier et al., 2005). RNAPIV and RNAPV may be characterised with structural and functional divergence from RNAPI-III and they are primarily involved in siRNA-mediated gene silencing by RNA-directed DNA methylation pathway (RdDM) (Köllen et al., 2015; Landick, 2009).

This study focuses on the regulation of gene expression by RNAPII during productive transcript elongation and these aspects will be detailly discussed in the following chapters.

1.2 DNA transcription by RNAPII

mRNA synthesis by RNAPII may be divided into three main phases known as initiation, elongation and termination. Each step is subjected to precise regulatory mechanisms which together comprise the so called RNAPII transcription cycle (Figure 1) (Sims et al., 2004; Van Lijsebettens and Grasser, 2014). During transcription initiation RNAPII is being recruited together with GTFs to form a pre-initiation complex (PIC) at the promoter region. The transcription is being initiated after aligning RNAPII on the DNA template and subsequent promoter melting. Next, RNAPII escapes the early elongation complex during the process of promoter clearance to enter productive transcript elongation (Jonkers and Lis, 2015; Saunders et al., 2006). Following this process (detailly discussed in 1.2.1) termination of transcription occurs together with mRNA cleavage and polyadenylated. Proceeded mRNA is subsequently transported to cytoplasm and may be translated into protein (Moore and Proudfoot, 2009). Following mRNA release from RNAPII complex, the components of transcript elongation complex (TEC) dissociate from the DNA template and may recycle to start a new round of transcription cycle (Shandilya and Roberts, 2012) (Figure 1).

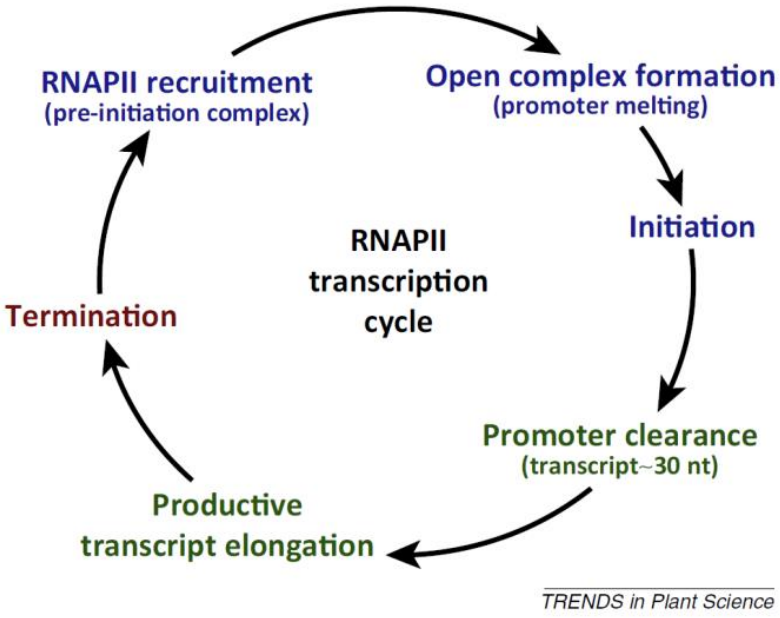


Figure 1. RNAPII transcription cycle.
The conserved steps of RNAPII transcription cycle are shown. Blue indicate early elongation events related to initiation, whereas green reflects late elongation events considered as transcription elongation. Transcription termination (depicted in red) is followed by the re-initiation of entire cycle. Picture is from Van Lijsebettens et al. 2014.

1.2.1 Transcript elongation

Transcript elongation has been initially considered as a simple process of nucleotide addition to the growing RNA chain. However, over last decades it has emerged as a precisely controlled and dynamically regulated process (Kwak and Lis, 2013; Sims et al., 2004). Highly compacted chromatin structure within the nucleus requires extend post-translational modifications to allow RNAPII recruitment to gene promoters and its subsequent passage through transcription units. Thus, the action of numerous transcript elongation factors (TEFs) is fundamental during transcript elongation to ensure the progression of RNAPII through different obstacles created by either DNA structure of DNA bound proteins (Kwak and Lis, 2013; Van Lijsebettens and Grasser, 2014). The precise discrimination of transcriptional events following transcription initiation is still elusive in plants while in other organisms comprises three distinct stages: promoter escape, RNAPII pausing and productive elongation (Figure 2) (Saunders et al., 2006; Van Lijsebettens and Grasser, 2014).

1.2.1.1 Promotor escape and early elongation complex formation

Promoter escape, also referred as promoter clearance, is considered the transition phase between transcription initiation and elongation. During this process PIC undergoes structural and functional maturation, accompanied with the escape from promotor region and transformation into so called early elongation complex (EEC) (Jonkers and Lis, 2015; Shandilya and Roberts, 2012). The switch between respective transcriptional stages is strongly associated with the length of synthetized nascent RNA (Dvir, 2002). Initially, RNAPII shows a tendency for the production of short RNA products accompanied by abortive initiation and

transcript slippage (Cai and Luse, 1987; Holstege et al., 1997). The synthesis of 8 - 9 nucleotides coincidence with the transition into the early elongation complex and sudden collapse of transcription bubble (Holstege et al., 1997). With growing length of nascent RNA, EEC undergoes further adjustments until it reaches considerable stability at about +23 nucleotides, accompanied by significant decrease of transcript slippage (Pal and Luse, 2003). EEC is very prone to backtracking and arrest before synthetizing ~ 30-nuleotide long nascent RNA (Pal et al., 2005). EECs arrested before that stage can be rescued by the transcription factor TFIIS (Figure 2 B; Saunders et al., 2006), which stimulates the intrinsic RNA cleavage activity of RNAPII so that a newly generated 3'-OH of nascent RNA is being aligned with RNAPII active site (Fish and Kane, 2002; Kettenberger et al., 2003).

Despite some distinct steps during early stages of transcription cycle it is difficult to set precise boundaries for initiation-to-elongation transition. The relationship between the distance of the EEC from the transcription start site (TSS) and the formation of a mature elongation complex is not fully understood (Chen et al., 2018a; Saunders et al., 2006). However, the maturation of EEC is corelated with the length of nascent RNA and is often accompanied by promotor-proximal pausing of RNAPII in metazoans (Adelman and Lis, 2012; Uptain et al., 1997).

1.2.1.2 Promoter-proximal pausing

RNAPII may be subjected to promoter-proximal pausing before becoming a part of fully mature elongation complex and entering productive transcript elongation (Adelman and Lis, 2012). This regulatory process is a widespread phenomenon in metazoans (Core et al., 2008; Muse et al., 2007) while its occurrence in plants in not uniform (Hetzel et al., 2016; Lozano et al., 2018). Promoter-proximal pausing of RNAPII has been first described in *Drosophila* (Gilmour and Lis, 1986). This process functions as a checkpoint before entering productive elongation and consist a key rate-limiting step in the regulation of RNAPII transcript elongation *in vivo* (Kwak and Lis, 2013; Selth et al., 2010). RNAPII pausing in the promotor-proximal regions has been also suggested to prevent nucleosomes entry into nucleosome-free region of open promoters (Gilchrist et al., 2010).

The exact molecular mechanisms underlying establishment and regulation of promoter-proximal pausing are still not fully understood, although high resolution studies revealed RNAPII pausing at several sites from +20 to +60 bz downstream TSS (Chen et al., 2018a; Core et al., 2008; Jonkers and Lis, 2015). Many features influencing RNAPII pausing in promoter-proximal regions has been identified including the elements of core promoter, RNAPII kinetic or physical barriers such as nucleosomes (Buckley et al., 2014; Lee et al., 2008; Mavrich et al., 2008). Nonetheless, the best understood model comprises the recruitment of some transcription factors and their direct influence on RNAPII entry and release from promotor-proximal pausing. Those effects are primarily imposed by DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF). Following RNAPII binding, NELF and DSIF act cooperatively to induce transcriptional pausing and to stabilize it (Figure 2 A) (Hartzog et al., 1998; Lee et al., 2008; Wada et al., 1998). DSIF is a heterodimer comprising SPT4 and SPT5 transcription factors. While SPT5 is conserved among three domains of life, SPT4 is absent in bacteria (Dürr et al., 2014; Hartzog and Fu, 2013).

NELF consist of four subunits: NELF-1, B, C/D and E and is conserved between mammals and *Drosophila* while its counterparts were not identified in yeast and *Arabidopsis* (Narita et al., 2003).

RNAPII pausing determined by NELF and DSIF is being eventually broken by the action of the positive transcription elongation factor-b (P-TEFb) complex (Peterlin and Price, 2006). P-TEFb comprises cyclin T1 and cyclin-dependent kinase 9 (CDK9) and is being recruited to promoters through direct and/or indirect interactions with the components of paused complex. Following its recruitment, P-TEFb drives the phosphorylation of NELF, DSIF and RNAPII-CTD at serine 2 position (Jonkers and Lis, 2015; Kwak and Lis, 2013). While NELF is evicted from the complex upon phosphorylation, DSIF is being transformed into positive elongation factor (Figure 2 B) (Saunders et al., 2006; Wada et al., 1998; Yamada et al., 2006).

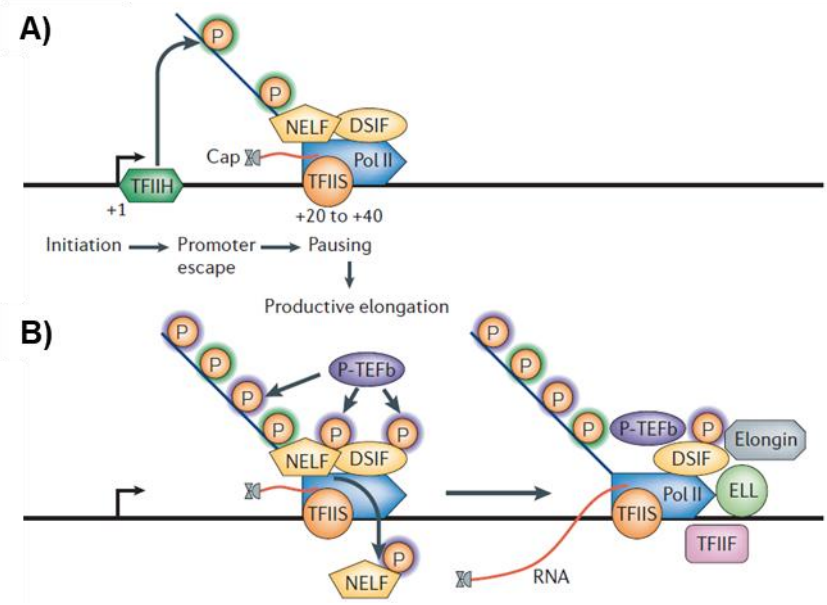


Figure 2. The transition from transcript initiation to productive transcript elongation in metazoans.
(A) TFIIH-mediated phosphorylation of Ser5 and Ser7 of the carboxy-terminal domain (CTD) of RNA polymerase II (RNAPII-CTD) occurs before promoter-proximal pausing in metazoans. DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) facilitate RNAPII pausing in the promoter-proximal region, and TFIIS also associates with the paused polymerase. Capping enzyme are being recruited and the nascent RNA (red line) becomes capped (Cap). (B) RNAPII is released from promoter-proximal pausing (left) into productive transcript elongation (right). Positive transcription-elongation factor-b (P-TEFb) mediates phosphorylation of DSIF, NELF and RNAPII-CTD at Ser2. TFIIS facilitates release of RNAPII from the pause site by stimulating RNA cleavage. NELF dissociates from the transcription complex and DSIF, TFIIS and P-TEFb track with RNAPII along gene body. TFIIF, eleven-nineteen lysine-rich in leukemia (ELL), and Elongin may additionally associate with the elongation complex. Picture is from Saunders et al., 2006.

The activity of TFIIS has been also demonstrated as necessary for the efficient release of paused RNAPII from promoter-proximal sites (Figure 2) (Adelman et al., 2005; Saunders et al., 2006). TFIIS stimulatory effects on RNA cleavage within paused RNAPII allow many rounds of transcription, pausing and backtracking until eventually RNAPII is being released from promoter-proximal region to carry out productive transcript elongation (Nechaev et al., 2010; Weber et al., 2014).

Taken together, a precisely coordinated balance between pausing inducers (such as NELF, DSIF, the +1 nucleosome and the core promoter elements) and activating factors (P-TEFb, TFIIS) may largely determine the level of RNAPII pausing before its entry into productive transcript elongation phase (Jonkers and Lis, 2015).

1.2.1.3 Productive transcript elongation

Following RNAPII release from promoter-proximal region, RNAPII complex gains stability and is being transformed into transcript elongation complex (TEC) also referred as mature elongation complex (Kwak and Lis, 2013; Selth et al., 2010). TEC plays a crucial role in driving productive transcript elongation and its regulation (Kwak and Lis, 2013; Sims et al., 2004).

RNAPII transition into productive elongation phase is accompanied by the gradual exchange of transcript initiation factors with elongation factors to finally form a fully mature TEC (Ehara et al., 2017). A growing number of TEFs has been identified over last decades by various genomic and biochemical studies (Kwak and Lis, 2013; Van Lijsebettens and Grasser, 2014). During the transcript elongation TEFs ensure progression of RNAPII through different obstacles created by either DNA structures or DNA bound proteins (Sims et al., 2004). The maintenance of TEC encompassing the RNA:DNA hybrid is a critical aspect of RNAPII processivity since disruption of this heteroduplex may result in premature termination of transcription (Ardehali and Lis, 2009). The stability and progression of elongating machinery is constantly being challenged by many features of transcriptional environment such as certain DNA sequences, exons-intron junctions or nucleosomes (Jonkers and Lis, 2015). Additionally, TEFs play an important role coordinating transcript elongation with ongoing co-transcriptional processes (Bentley, 2014; Van Lijsebettens and Grasser, 2014).

TEFs are very heterogenous group of proteins which associated with TEC either permanently or transiently (Sims et al., 2004). Many TEFs were shown to have their counterparts in *Arabidopsis* (Dürr et al., 2014; Grasser et al., 2009; Lolas et al., 2010). Functionally, TEFs can be divided into several groups including factors that directly modulate the catalytic activity of RNAPII, facilitate progression through chromatin or impose certain histone modification within transcribed regions (Figure 3) (Jonkers and Lis, 2015; Sims et al., 2004; Van Lijsebettens and Grasser, 2014). TEC composition as well as the recruitment of particular TEFs during transcription cycle and their mutual coordination of transcript elongation is still poorly understood in plants (Van Lijsebettens and Grasser, 2014). Notably, carboxy terminal domain of RNAPII largest subunit (RNAPII-CTD) has been demonstrated as a very important and conserved docking platform for the recruitment of many TEFs in *Arabidopsis* as well as other organisms (Hajheidari et al., 2013; Heidemann et al., 2013).

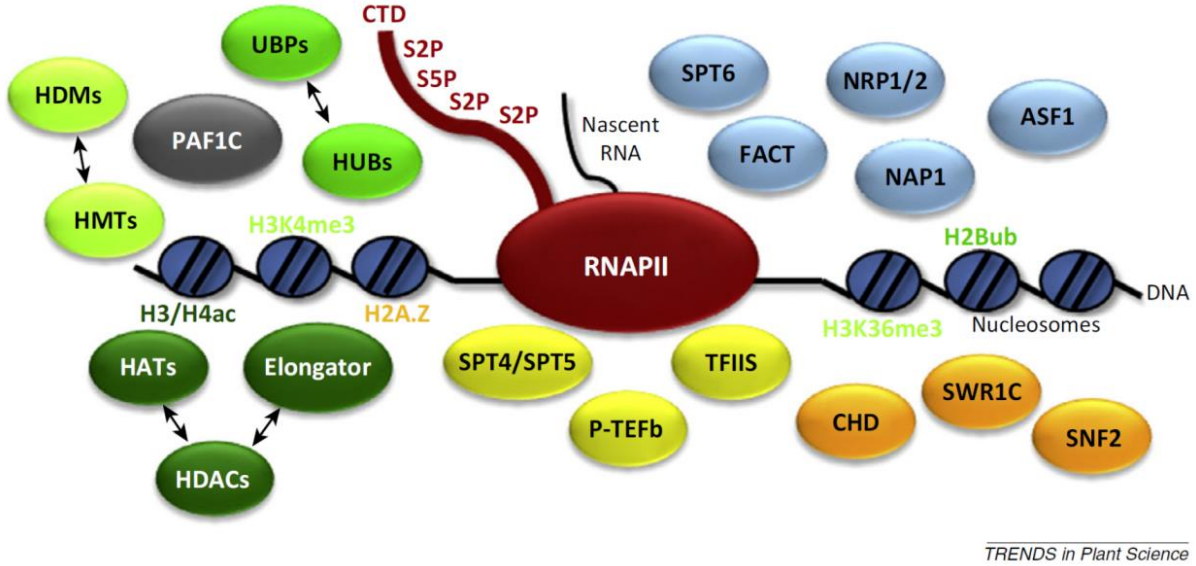


Figure 3. A variety of transcript elongation factors (TEFs) influencing transcript elongation in *Arabidopsis*. TEFs may determine RNAPII progression through the transcription cycle in various ways. Some of the factors may directly modulate RNAPII properties (depicted in yellow). TEFs regulating the chromatin structure: histone chaperones and ATP-dependent chromatin remodelers are depicted in blue and orange, respectively. Other factors (depicted in green) control transcription by reversibly imposed/removed covalent histones modification (indicated by double-headed arrows). Some of the activating histone marks are shown. Picture is from Van Lijsebettens et al. 2014.

1.2.1.4 RNAPII-CTD phosphorylation cycle

RNAPII-CTD is a flexible part of largest RNAPII subunit absent in other eukaryotic RNA polymerases (Liu et al., 2010; Werner and Grohmann, 2011). RNAPII-CTD contains a tandem repetition of Tyr1–Ser2–Pro3–Thr4–Ser5–Pro6–Ser7 consensus motif with 26, 52 and 34 repeats in yeast, mammals and *Arabidopsis*, respectively (Heidemann et al., 2013; Werner and Grohmann, 2011). Although CTD has been shown to be dispensable for RNAPII catalytic activity its presence is essential for organism viability (Serizawa et al., 1993; West and Corden, 1995). RNAPII-CTD residues may be subjected to various post-translational modifications including peptidyl–prolyl isomerization (Pro), glycosylation (Ser, Thr) and phosphorylation (Tyr, Ser, Thr) (Heidemann et al., 2013; Li et al., 2007; Sims et al., 2011). RNAPII-CTD phosphorylation at serine 2 (RNAPII-S2P) and serine 5 (RNAPII-S5P) positions are best studied modifications in the context of productive elongation (Dronamraju and Strahl, 2014; Heidemann et al., 2013; Yu et al., 2015). RNAPII-CTD phosphorylation status may link transcription, mRNA processing and chromatin remodelling by acting as a "docking station" for various proteins (Fong et al., 2017; Komarnitsky et al., 2000). In line with that, differently phosphorylated RNAPII-CTD reveals distinct interactomes although the precise boundaries between particular RNAPII-CTD stages are rather diffused (Harlen et al., 2016; Heidemann et al., 2013).

Numerous phosphatases and kinases modifying RNAPII-CTD have been identified, in line with dynamic changes in RNAPII-CTD phosphorylation status accompanying RNAPII progression throughout transcription cycle (Hajheidari et al., 2013; Heidemann et al., 2013). Initially, RNAPII is being recruited to gene promoters in a hypo-phosphorylated form (Lu et al., 1991; Vinayachandran et al., 2018). The subsequent changes of RNAPII-CTD phosphorylation status

corelates with promoter escape and/or promotor-proximal pausing. Accordingly, shortly after entering gene body, RNAPII-CTD is being phosphorylated at Ser5 and Ser7 positions by the cyclin dependent kinase CDK7, the component of TFIIH (Hajheidari et al., 2013; Liu et al., 2004). Further major changes in terms of RNAPII-CTD phosphorylation are the consequence of RNAPII release from promoter-proximal pausing, when RNAPII-CTD is being phosphorylated at Ser2 position by P-TEFb (Ramanathan et al., 2001). RNAPII release into the gene body results in gradual decrease of RNAPII-S7P and RNAPII-S5P modifications accompanied by the accumulation of RNAPII-S2P form towards gene end (Figure 4) (Heidemann et al., 2013; Mayer et al., 2010). Recent high-resolution studies have suggested the precise switch from RNAPII-S5P to RNAPII-S2P phosphorylation at ~ 600 – 700 bp downstream TSS (Vinayachandran et al., 2018). In accordance with their genome-wide profiles (Figure 4), RNAPII-S5P modification has been associated with promoter clearance, 5' capping of nascent RNA and RNAPII transition into productive transcript elongation, whereas RNAPII-S2P is being linked with productive transcript elongation and co-transcriptional processes (Komarnitsky et al., 2000; Morris and Greenleaf, 2000; Perales and Bentley, 2009). RNAPII-CTD phosphorylation at Tyr and Thr generally accumulate towards 3' end (Tyr1) or downstream polyadenylation site (Thr4), while their influence on productive transcript elongation remains poorly understood (Figure 4) (Heidemann et al., 2013).

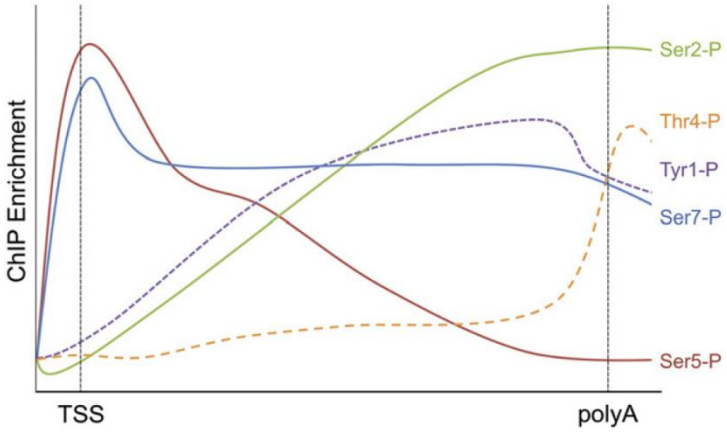


Figure 4. RNPII-CTD phosphorylation throughout transcription cycle.
Average profiles of RNAPII-CTD phosphorylation marks across yeast and metazoan genes revealed by chromatin immunoprecipitation (ChIP) experiments (Heidemann et al., 2013 and references therein). Coloured lines show schematic representation of genome-wide occupancy profiles for each CTD phosphorylation marks. TSS: transcription start site. polyA: polyadenylation site. Picture is from Heidemann et al., 2013.

1.2.1.5 Nucleotide addition cycle and RNAPII backtracking

Apart from promotor-proximal pausing, RNAPII has been suggested to experience backtracking-mediated pausing throughout entire DNA template due to various barriers including DNA-bound proteins or DNA lesions (Churchman and Weissman, 2011; Nudler, 2012; Steurer and Marteijn, 2017). ~ 2×10^5 pause sites have been detected in yeast genome and ~75% of them have been associated with RNAPII backtracking (Churchman and Weissman, 2011). RNAPII backtracking is not simply a side effects of complex chromatin environment but it is rather believed to play an important role in the regulation of many

molecular processes including proximal-pausing release, transcription termination, transcriptional fidelity, RNA processing, and genome stability (James et al., 2017; Nudler, 2012) (McKay and Cabrita, 2013; Steurer and Marteijn, 2017).

During the process of RNA synthesis, RNAPII has been suggested to oscillate between forward and backward movement driven by thermal energy in accordance to Brownian ratchet model (Bar-Nahum et al., 2005). RNAPII oscillation between pre- and post-translocation states depends on the correct incorporation of incoming NTP determined by two components of the RNAPII catalytic centre, the bridge helix (BH) and the trigger loop (TL) (Figure 5) (Cheung and Cramer, 2011; Nudler, 2009). Consequently, RNAPII moves rapidly forward on average but can also perform retrograde motion when its energetically favourable (Bar-Nahum et al., 2005; Mejia et al., 2015). Accordingly, *in vitro* transcript elongation has been demonstrated as highly discontinuous process, with frequent backtracking, pausing and transcriptional arrest (Reines et al., 1999; Svejstrup, 2007).

NTP addition during mRNA synthesis by RNAPII have been largely resolved structurally and mechanistically and is referred as nucleotide addition cycle (NAC) (Brueckner et al., 2009; Erie et al., 1992; Nudler, 2009). The catalytic centre of RNAPII includes the binding sites for the RNA 3' end (i site) and the insertion site for the incoming NTP (Figure 5) (Bochkareva et al., 2012; Nudler, 2009). Initially, RNAPII is in the post-translocation state, which is characterized by an empty active site and open trigger loop (TL). Binding of incoming NTP (at i+1 site) triggers TL closure which in turn stabilizes the incorporation of correct NTP (Brueckner et al., 2009; Fouqueau et al., 2013). After phosphodiester bond formation and the release of pyrophosphate accompanied by the pre-translocation stage of elongation complex, TL and BH may oscillate between various conformation, forcing the movement of the RNA-DNA hybrid back along the catalytic cleft (Figure 5). As a result, elongation complex enters back again the post-translocation state with free active site for NTP binding (Nudler, 2009; Zhang et al., 2016). Complex translocation following NTP binding is generally smooth process but it may be compromised by certain DNA sequences and DNA-bound proteins imposing translocation barrier (Imashimizu et al., 2013). Such barrier may result in RNAP pausing and its properties has been resolved by high-resolution atomic structures (Cheung and Cramer, 2011; Wang et al., 2009). Upon pausing RNAPII is likely to undergo backtracking reaching the depth from +1 to over 20 nucleotides which leads to the misalignment of RNA 3’ end from RNAPII active centre (Imashimizu et al., 2015). Backtracked RNA interacts with RNAPII within the secondary channel resulting in titled DNA–RNA hybrid and locked trigger loop oscillation which consequently renders backtracked complexes very stable (Cheung and Cramer, 2011; Forde et al., 2002). Thus, backtracked 3’ end of nascent RNA must be realigned to RNAPII active site in order to resume transcription. It is being achieved by RNAPII diffusion along DNA (Galburt et al., 2007; Hodges et al., 2009) or more commonly by generating new 3’ end through RNA cleavage activity of RNAPII which may be largely stimulated by TFIIS (Fish and Kane, 2002; Izban and Luse, 1992).

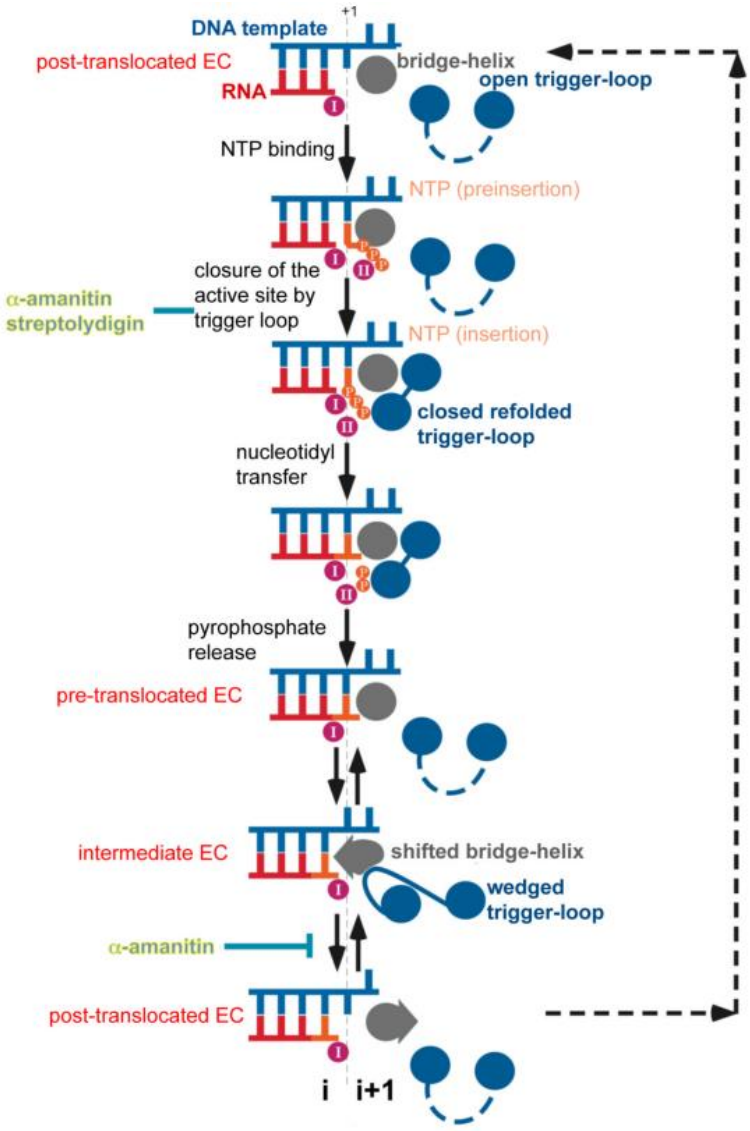


Figure 5. Nucleotide addition cycle (NAC).
Schematic diagram depicting each step of nucleotide addition cycle (NAC). The binding of correct NTP (orange) in post-translocated stage triggers the oscillation of bridge helix (BH) and trigger loop (TL) domains as well as subsequent elongation complex (EC) transition into the pre-translocation stage. TL/BH may oscillate between the unfolded (open) conformation and intermediate state, forcing the "push" against the RNA-DNA hybrid, thereby inducing EC translocation into the post-translocated state. α-amanitin and steptolydigin (both depicted in yellow) have been suggested to interfere with NAC by inhibiting substrate loading or RNAP translocation. DNA and RNA are depicted in blue and red, respectively. BH is depicted in grey and TL in blue. Catalytic Mg-A (I) and Mg-B (II) are shown as magenta circles. i: the binding sites for the RNA 3' end. i+1: binding site of incoming NTP. Picture is from Nudler, 2009.

1.2.1.6 Nascent RNA cleavage

Backtracked RNAPII remains catalytically inactive due to 3' end of nascent RNA misaligned from RNAPII active centre (Wang et al., 2009). Nascent RNA cleavage by RNAPII generates a new 3' end proximal to the catalytic centre which allows continued NTP incorporation for catalytic elongation of the transcript (Cheung and Cramer, 2011; Nudler, 2009). The coordination of these processes is crucial for cells viability since the inhibition of intrinsic RNAPII cleavage activity has been demonstrated lethal in yeast (Sigurdsson et al., 2010). Nascent RNA cleavage is conserved mechanisms in many DNA-dependent RNA polymerases and within RNAPI-III complexes is being primary stimulated by homologous A12.2, Rpb9, and

C11 subunits, respectively (Chédin et al., 1998; Vannini and Cramer, 2012; Walmacq et al., 2009). However, intrinsic RNAPII cleavage activity is much weaker then observed in RNAPI and RNAPIII. Unlike A12.2 and C11, RPB9 subunit possess altered C-ribbon domain absents from RNAPII catalytic centre what results in only weak, allosteric stimulation of RNA cleavage (Koyama et al., 2007; Ruan et al., 2011). In contrast to RNAPI and RNAPIII, intrinsic cleavage activity of RNAPII may be largely stimulated by a group of TFIIS-like cleavage factors, including bacterial GreA/B, eukaryotic TFIIS and archaeal TFS (Borukhov et al., 1993; Izban and Luse, 1992; Lange and Hausner, 2004).

The generation of a new RNA 3' end proximal to RNAPII active site has been proposed to occur within stalled ternary complexes by the hydrolytic cleavage of nascent transcript (Izban and Luse, 1992; Weilbaecher et al., 2003). Intrinsic cleavage activity of RNAPII seems to be particularly favourable within +1 backtracked complexes (Cheung and Cramer, 2011) but it was also shown to release complexes from much deeper backtracking (Lisica et al., 2016; Sigurdsson et al., 2010).

Intrinsic RNAPII cleavage activity is catalysed by TL domain and requires the positioning of metal cofactor in RNAPII active centre (Čabart et al., 2014; Miropolskaya et al., 2017; Yuzenkova and Zenkin, 2010). Metal ion coordination within RNAPII active centre has been proposed as the mechanism for the "remodeling" of the active site from RNA synthesis to transcript cleavage (Kettenberger et al., 2004; Opalka et al., 2003; Svetlov and Nudler, 2013). In line with that, the stimulatory effects of TFIIS-like factors on intrinsic RNAPII cleavage activity associates with modified coordination of metal ions in RNAPII active centre (chapter 1.2.2.2) (Cheung and Cramer, 2011; Wang et al., 2009; Weilbaecher et al., 2003). The deposition of C-ribbons structure of TFIIS-like factors in the proximity of RNAPII active site additionally locks the TL away, switching off the relatively slow TL-dependent intrinsic transcript hydrolysis (Cheung and Cramer, 2011; Da et al., 2016; Roghanian et al., 2011).

1.2.2 TFIIS

Transcription factor IIS (TFIIS) is one of the best studied transcript elongation factors directly affecting RNAPII properties, yet its stimulatory role during productive transcript elongation *in vivo* is not fully understood (Fish and Kane, 2002). Eukaryotic TFIIS has been extensively demonstrated to stimulate intrinsic endonuclease cleavage activity of RNAPII allowing the realignment of nascent RNA 3' end within RNAPII active centre (Fish and Kane, 2002; Izban and Luse, 1992). In line with its molecular function, TFIIS allows RNAPII release from paused stage and promote RNAPII read-through of various barriers to transcript elongation (Adelman et al., 2005; Reines et al., 1989). TFIIS-like cleavage factors are present in three domains of life, with functionally homologous GreA/B in bacteria and TFS cleavage factors in archaea (Borukhov et al., 1993; Lange and Hausner, 2004; Zenkin and Yuzenkova, 2015). Surprisingly, TFIIS has been shown dispensable for organisms growth under normal conditions in plants (Grasser et al., 2009) and yeast (Sigurdsson et al., 2010; Williams and Kane, 1996). Interestingly, the lack of functional TFIIS in mice embryos results in their lethality due to impaired hematopoiesis (Ito et al., 2006).

1.2.2.1 TFIIS structural and functional analysis

TFIIS is composed of three distinct domains conserved between yeast, mammals and *Arabidopsis* (Figure 6) (Booth et al., 2000; Fish and Kane, 2002; Grasser et al., 2009). N-terminal domain I is the least conserved part of the protein and it forms a highly flexible four-helix bundle in yeast which has been suggested to protrude from RNAPII complex (Booth et al., 2000; Kettenberger et al., 2003). Domain I has been demonstrated to play a role in TFIIS nuclear targeting (Ling et al., 2006) as well as in PIC assembly (Kim et al., 2007; Prather et al., 2005). Domain I may mediate the integration with SAGA and Mediator components and thus has been suggested as important for transcription regulation independent from RNAPII pausing (Guglielmi et al., 2007; Wery et al., 2004). Indeed, domain I is dispensable for transcriptional stimulation related to RNA cleavage within RNAPII (Awrey et al., 1998; Fish and Kane, 2002). Accordingly, domains II-III are sufficient for the stimulatory effects of TFIIS on intrinsic RNAPII cleavage activity (Awrey et al., 1998; Wind and Reines, 2000). Domains II and III are also the most conserved parts of TFIIS (Figure 6) (Grasser et al., 2009; Kettenberger et al., 2003). Domain II forms a three-helix bundle and is crucial for TFIIS binding to RNAPII (Awrey et al., 1998). Domain II connects with domain III through a flexible linker which integrity is necessary for protein activity (Awrey et al., 1998; Kettenberger et al., 2003;

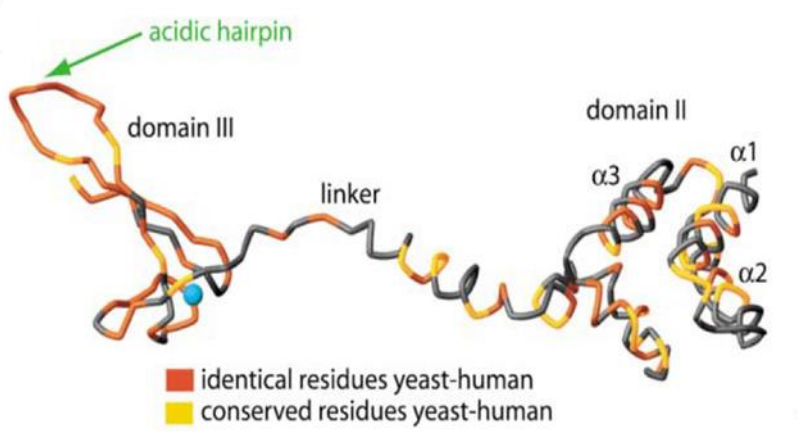


Figure 6. TFIIS structure and conservation.
The side view at the structure of yeast TFIIS containing domains II and III. Residues identical and conserved between yeast and human TFIIS are depicted in red and orange, respectively. Acidic hairpin containing invariant aspartic acid and glutamic acid residues is indicated with green arrow. Picture is from Kettenberger et al., 2003.

Olmsted et al., 1998). Domain III is composed of three antiparallel β-sheets that form a zinc ribbon structure (Kettenberger et al., 2003). It is the most conserved part of the protein containing RSADE motif within acidic hairpin (Figure 6) (Fish and Kane, 2002; Grasser et al., 2009). Invariant aspartic and glutamic acid residues (D290 and E291, referring to the yeast protein) within this motif are critical for TFIIS stimulation on intrinsic RNAPII cleavage and consequently cell viability (Awrey et al., 1998; Sigurdsson et al., 2010)

The details regarding TFIIS structure has been further resolved by its crystallisation in the complex with yeast RNAPII (Kettenberger et al., 2003). The three-dimensional structure of TFIIS-RNAPII complex, lacking flexible TFIIS N-terminal domain, revealed TFIIS occupancy on polymerase surface extending from one of the jaws to the active centre (Figure 7) (Kettenberger et al., 2003; Martinez-Rucobo and Cramer, 2013). In line with previous finding, domain II binds to the jaw domain of RNAPII (Awrey et al., 1998; Kettenberger et al., 2003). Upon RNAPII binding, linker fragment between domain II and III forms an α-helix and reaches through crevice into RNAPII funnel, also called a secondary channel (Kettenberger et al., 2003; Martinez-Rucobo and Cramer, 2013). Domain III has been previously demonstrated as highly flexible in free TFIIS (Qian et al., 1993) as well as dispensable for RNAPII binding (Awrey et al., 1998). It makes, however, many contacts with RNAPII at the entrance to the pore upon TFIIS recruitment and it inserts further into the funnel where it approaches the polymerase active site (Figure 7) (Kettenberger et al., 2003; Martinez-Rucobo and Cramer, 2013).

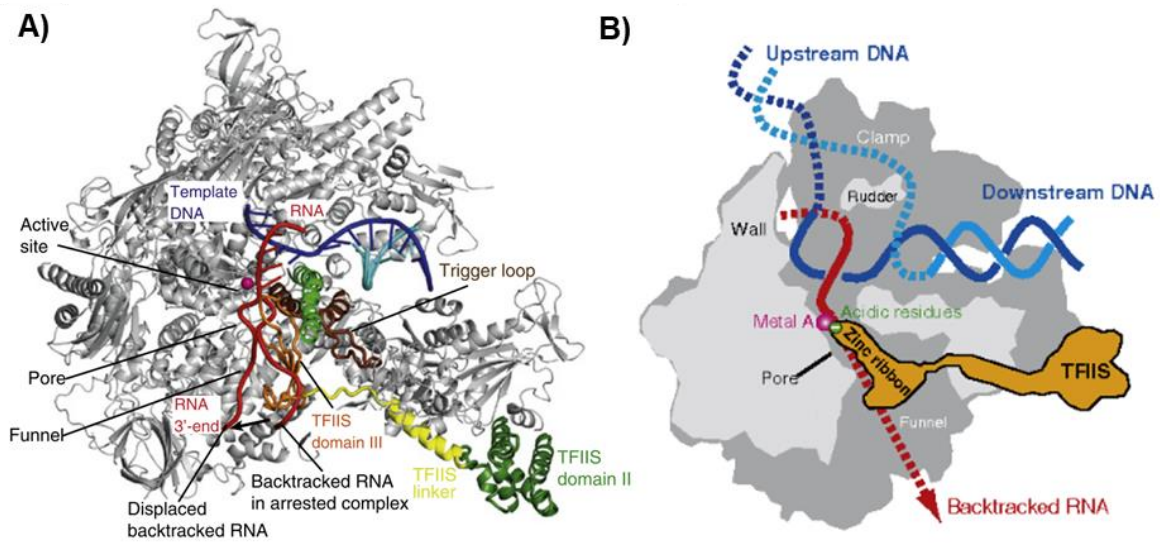


Figure 7. The structure of yeast TFIIS-RNAPII complex.
Side view of TFIIS-RNAPII complex in cartoon representation (A) and in a schematic cutaway view (B). (A) Structure of RNAPII reactivation intermediate with TFIIS bound and displaced backtracked RNA (red). A second backtracked RNA from the arrested complex clashing with TFIIS domain III (orange) was additionally modelled into the structure. (A-B) TFIIS reach RNAPII active site with an acidic hairpin for the stimulation of RNA cleavage. Picture is modified from Martinez-Rucobo and Cramer, 2013.

TFIIS binding to RNAPII triggers broad structural changes in the mobile part of yeast RNAPII complex inducing jaws, clamp, cleft and foot domains (Kettenberger et al., 2003). Those structural changes result in a coordinated repositioning of about one third of the polymerase mass and correspond to the mobile part of RNAPII (Cramer et al., 2001; Kettenberger et al., 2003). In additional to overall structural changes in RNAPII complex, TFIIS induces local remodeling of RNAPII active centre including BH and TL domains (Figure 7 A) (Kettenberger et al., 2004; Martinez-Rucobo and Cramer, 2013). Additionally, TFIIS binding to the elongation complex realigns RNA position in RNAPII active centre (Kettenberger et al., 2004).

1.2.2.2 The mechanisms of TFIIS-stimulated transcript cleavage

The resolution of TFIIS-RNAPII structure revealed a detailed mechanism of TFIIS-dependent RNAPII reactivation involving the endonucleolytic cleavage of backtracked RNA (Cheung and Cramer, 2011; Wang et al., 2009). The insertion of TFIIS domain III into the pore has been demonstrated crucial to complete RNAPII active site next to backtracked RNA (Kettenberger et al., 2003). The stimulation of intrinsic RNAPII cleavage activity by TFIIS involves the correct positioning of two metal ions and a water molecule in RNAPII active centre (Figure 8) (Cheung and Cramer, 2011; Zhang et al., 2010). First metal ion (metal A) is immobile and persistently bound to the active site in order to align the scissile phosphodiester bond. TFIIS-stimulated cleavage determines the positioning of second metal ion (metal B) and a water molecule, which acts as the nucleophile (Cheung and Cramer, 2011). The positioning of metal B is being coordinated by the acidic residue D290 and E291 located at the tip of acidic hairpin of domain III (Cheung and Cramer, 2011; Sosunov et al., 2003). The coordination of metal B allows proton subtraction from the water molecule and subsequent proton donation to the RNA 3' end (Cheung and Cramer, 2011). Additionally, TFIIS residue R287 reaches into RNAPII catalytic site in order to stabilize the transition state during RNA cleavage (Cheung and Cramer, 2011). Following RNA cleavage and formation of a new RNA 3' end proximal to the active site, transcript elongation may resume (Awrey et al., 1998; Cheung and Cramer, 2011). This mechanism explains the molecular bases for RNA cleavage inhibition upon the mutation of invariant D290 and E291 residues in yeast (Sigurdsson et al., 2010).

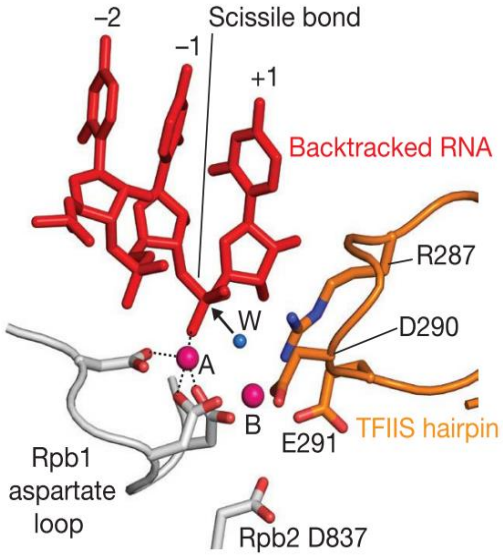


Figure 8. Molecular bases of TFIIS-stimulated RNA cleavage.
Model for RNAPII active site geometry during TFIIS-stimulated RNA cleavage. Metal A (magenta sphere) is coordinated by RNAPII aspartate loop and the RNA phosphate. The side chains of TFIIS hairpin (orange) completes RNAPII active centre. Metal B (B) is coordinated by TFIIS residues D290 and E291 (orange, referring to yeast protein) and allows the nucleophilic attack (indicates with an arrow) on RNA scissile bond by the water molecule (W, blue sphere). Picture is from Cheung and Cramer, 2011.

Apart from its role in stimulating RNA cleavage, TFIIS may also contribute to the release of backtracked complexes by inducing the changes in RNAPII structure as well as by displacing backtracked RNA from the funnel (Cheung and Cramer, 2011; Kettenberger et al., 2004). The stimulatory effects of TFIIS-like cleavage factors on RNAPII release may also occur through the

selection of RNA cleavage site although the underlying mechanism remains elusive (Weilbaecher et al., 2003).

1.2.2.3 TFIIS role *in vitro* and *in vivo*

TFIIS stimulatory effects on intrinsic RNAPII cleavage activity have been broadly demonstrated *in vitro* in the context of RNAPII release from pausing (Adelman et al., 2005; Ishibashi et al., 2014; Koyama et al., 2007; Nock et al., 2012). Consequently, TFIIS has been primary characterised as a transcription factor positively regulating productive transcript elongation and facilitating RNAPII read-through of various blocks to transcript elongation (Fish and Kane, 2002). Apart from its stimulatory effects on RNA cleavage, TFIIS has been also shown to regulate transcription initiation process (Kim et al., 2007; Prather et al., 2005).

In line with its role in regulating productive transcript elongation, TFIIS absence in yeast has been related to many transcriptional defects despite unaffected growth properties. Accordingly, the molecular analysis of yeast lacking functional TFIIS revealed compromised nascent transcription (Gutiérrez et al., 2017) as well as modified backtracking properties of RNAPII (Churchman and Weissman, 2011). Additionally, TFIIS absence associates with defects in transcription fidelity (James et al., 2017), inhibited release of promoter-proximally paused RNAPII (Adelman et al., 2005), increased nucleosome fuzziness (Gutiérrez et al., 2017) and enhanced exon inclusion (Howe et al., 2003). From these studies TFIIS emerges as a general regulator of any transcriptional process related to RNAPII pausing, in line with its molecular function (Cheung and Cramer, 2011; Fish and Kane, 2002). Additionally, TFIIS deficiency results in increased sensitivity to 6-azauridine (6-AU) (Archambault et al., 1992; Williams and Kane, 1996).

The examination of TFIIS absence in *Arabidopsis* has revealed compromised seed dormancy and overall wild type-like growth properties (Grasser et al., 2009). Additionally, TFIIS removal in *Arabidopsis* results in the missregulation of many genes at later developmental stages (Grasser et al., 2009).

Based on the mutagenesis and crystallisation studies, the residues crucial for TFIIS activity has been identified (Awrey et al., 1998; Cheung and Cramer, 2011). Accordingly, the replacement of invariant D290 and E291 residues render TFIIS as a negative dominant form which inhibits intrinsic RNAPII cleavage activity both *in vitro* and vivo (Nock et al., 2012; Sigurdsson et al., 2010). In contrast to TFIIS absence, its dominant negative form (referred to as “TFIISmut” in this study) strongly inhibits growth when expressed Col-0 (Figure 9) (Dolata et al., 2015). TFIISmut expression in yeast has been shown to inhibit intrinsic RNAPII cleavage activity resulting in organism lethality (Sigurdsson et al., 2010). Molecular consequences of TFIISmut expression are however less understood. In yeast the expression of TFIISmut has been demonstrated to inhibits transcript elongation *in vivo* resulting in RNAPII polyubiquitination (Sigurdsson et al., 2010). In *Arabidopsis* enhanced exon inclusion has been observed in the presence of mutated TFIIS (Dolata et al., 2015)

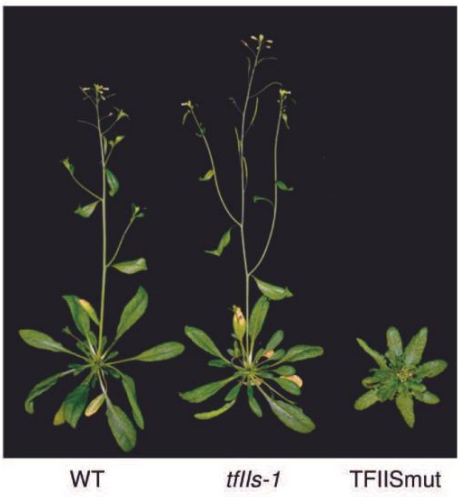


Figure 9. TFIISmut expression in Col-0 results in severe developmental defects.
The phenotypic analysis of plants expressing TFIISmut in Col-0 in comparison to WT (Col-0) and *tfIIs-1*. Picture is from Dolata et al., 2015 and was taken at bolting time of TFIISmut mutant for plants grown on soil in long day conditions.

1.2.2.4 The interaction between TFIIS with other transcription-related factors

TFIIS binding to RNAPII allows the cooperative regulation of transcription between TFIIS and other transcriptionally related proteins (Dutta et al., 2015; Guglielmi et al., 2007). Thus, TFIIS role is not simply limited to the stimulation of RNA cleavage but likely associates with a large network of interactions within transcriptionally active complexes (Fish and Kane, 2002; Wind and Reines, 2000).

Several studies have demonstrated the genetic interaction between yeast TFIIS and other TEFs including SPT4-STP5 (Wada et al., 1998) SPT6 (Swanson and Winston, 1992) and Elongator (Otero et al., 1999). Additionally, genetic interaction has been observed between TFIIS and mutants with perturbated RNAPII-CTD properties or with CTD modulators deficiency (Lindstrom and Hartzog, 2001). Interestingly, TFIIS has been shown to interact physically via its N-terminal domain with the components of SAGA and Mediator complex (Nock et al., 2012; Wery et al., 2004). TFIIS may also play a role in the recruitment and/or affinity of other TEFs to RNAPII, including CCR4-Not (Dutta et al., 2015) and PAF1-C (Xu et al., 2017).

In this study the genetic interaction between *Arabidopsis* TFIIS and PAF1-C was analysed in detail, thus PAF1-C properties will be described in following chapters.

1.2.3 PAF1-C

Polymerase-associated factor 1 complex (PAF1-C) was first identified as a novel RNAPII-interacting complex in *Saccharomyces cerevisiae* (Wade et al., 1996). PAF1-C is a conserved complex in eukaryotes where it comprises five (yeast) to six (human, *Drosophila*) subunits (Rondón et al., 2004; Tomson and Arndt, 2013).

PAF1-C is generally considered to act during entire transcription cycle in line with its role in regulating gene expression (Tomson and Arndt, 2013). It has been first characterised as enriched at actively transcribed open reading frames (ORFs) (Pokholok et al., 2002). More detailed studies revealed PAF1-C entry downstream TSS with its subsequent dissociation from

TEC at the polyadenylation site in yeast and human (Chen et al., 2015; Fischl et al., 2017; Mayer et al., 2010). PAF1-C recruitment to TEC requires direct association with RNAPII but also additional contacts with other TEFs including TFIIS and SPT4-SPT5 (Mayekar et al., 2013; Xu et al., 2017). Interestingly, PAF1-C interaction with nascent RNA has been demonstrated (Dermody and Buratowski, 2010). This broad interaction network between PAF1-C and other transcriptional components may consequently determine complex entry and exit point during transcription cycle (Van Oss et al., 2017).

1.2.3.1 Diverse functions of PAF1-C

PAF1-C has been suggested to regulate gene expression in yeast and metazoans though the processes related to transcript elongation (Moore and Proudfoot, 2009). Accordingly, it has been demonstrated as a regulator of co-transcriptional mRNA maturation (Sheldon et al., 2005) and polyadenylation of mRNAs (Kowalik et al., 2015; Yang et al., 2016). Moreover PAF1-C has multiple roles in diverse regulatory mechanisms linking transcription elongation to chromatin structure by controlling various histone modification cascades (Van Oss et al., 2017; Verrier et al., 2015) and the maintenance of heterochromatin (Sadeghi et al., 2015). Yeast PAF1-C has been shown to directly interact with chromatin remodeller Chd1 (Simic et al., 2003) and to regulate histone H3 methylation (Krogan et al., 2003). PAF1-C plays also a crucial role in the recruitment of ubiquitylation factors and subsequent ubiquitination of histone H2B (Ng et al., 2003; Wood et al., 2003; Xiao et al., 2005). Additionally, PAF1-C may stimulate H3 histone methylation associated with its direct binding to histone H3 tail (Wu and Xu, 2012). Through its broad influence on histone mark deposition, PAF1-C may determine overall nucleosome dynamic and their traversal by RNAPII (Tomson and Arndt, 2013).

In recent years, new roles for PAF1-C have been identified in yeast including regulation of promoter-proximal pausing (Chen et al., 2015; Lu et al., 2016) and the resolution of transcription-replication conflicts (Poli et al., 2016).

1.2.3.2 PAF1-C in *Arabidopsis*

PAF1-C shows high conservation across organisms and is composed from five subunits in yeast and six subunits in human and *Drosophila* (Tomson and Arndt, 2013). Some counterpart of PAF1-C components were first identified in *Arabidopsis* during genetic screening for early flowering mutants. Accordingly, plants deficient in ELF7 or ELF8 subunit show early flowering phenotype with many additional pleiotropic developmental defects which has been connected to reduced transcript levels of the floral repressors FLC and MAF (He et al., 2004; Oh et al., 2004)

Similarly to observed in yeast and metazoans, *Arabidopsis* PAF1-C is being involved into the deposition of histone methylation marks (Van Lijsebettens and Grasser, 2014). CDC73 subunit of PAF1-C determines the deposition of H3K27me (Park et al., 2010; Yu and Michaels, 2010) while the absence of functional VIP3 subunit results in severally affected distribution of H3K36me2 and H3K27me3 modifications (Oh et al., 2008). Additionally, PAF1-C ensures correct distribution of H3K4me3 mark at target genes (He et al., 2004).

Together, those studies underline diverse and largely conserved roles of PAF1-C in regulating gene expression through chromatin structure, RNAPII properties and other molecular processes related to productive transcript elongation (Van Lijsebettens and Grasser, 2014; Van Oss et al., 2017).

Aim of the study

TFIIS is one of the best characterised transcript elongation factors (TEFs) which can directly affect RNAPII properties. Its molecular role in stimulating intrinsic RNAPII cleavage activity has been demonstrated *in vitro*, yet details regarding its molecular role *in vivo* remain elusive.

Considering TFIIS redundancy for *Arabidopsis* growth under normal condition, this study will utilise the dominant negative version of TFIIS (TFIISmut). Following the mutagenesis of invariant residues within TFIIS acidic loop (Asp309 and Glu310), TFIISmut will be expressed in plants lacking functional TFIIS. Additionally, system with conditionally controlled TFIISmut expression by β-estradiol induction will be developed. Inducible expression of TFIISmut could allow to overcome likely lethal consequences of its expression at early developmental stages. Additionally, more direct determination of molecular and morphological consequences of TFIISmut expression could be achieved. Obtained mutants will be further analysed phenotypically to unravel over-time morphological defects triggered by TFIISmut.

The main focus of this study will be the analysis of molecular consequences of TFIISmut expression, shedding light on its role in regulating transcript elongation *in vivo*. It will be attempted to determine active RNAPII occupancy in *Arabidopsis* since RNAPII properties are most likely target of mutated TFIIS. To better understand the properties of transcript elongation in plants, the analysis will be applied genome-wide by optimising and utilizing ChIP-seq. Additionally, the resolution of over-time transcriptome rearrangement upon TFIISmut expression by RNA-seq could provide further insight into the biological role of TFIIS. The properties of transcriptionally engaged RNAPII associated with TFIISmut will be further evaluated to better understand the dynamic of their association in the chromatin context.

TFIIS has been demonstrated as a component of transcript elongation complex (TEC), which composition in *Arabidopsis* remains elusive. The combination of affinity purification coupled with mass spectrometry (AP-MS) will be adopted to determine *Arabidopsis* TEC, with particular focus on TFIIS interactome. Several transcription-related proteins have been shown to directly interact with TFIIS including PAF1-C, Mediator and SAGA components. Those finding will be evaluated in the context of *Arabidopsis* TEC. Identified interaction between TFIIS and other TEFs will be further studied biochemically and genetically to better understand their mutual contribution into the regulation of transcript elongation and other molecular processes.

The control of transcript elongation rate is considered an important determinant of transcriptional outcome. A variety of methods allowed the determination of elongation rate in several organisms, while this information is still elusive in plants. Additionally, transcript elongation rate determinants are still poorly understood, including for instance the role of TEFs. Thus, an additional aim of this study is to develop novel molecular tool for the determination of transcript elongation rate *in vivo* in various genomic background taking the advantage of β-estradiol inducible system.

Taken together, this study aims to deepen the knowledge regarding the properties and the regulation of transcript elongation in plants, particularly related to TFIIS molecular function.

2. Results: TFIIS mutation reveals its importance in regulating transcript elongation in *Arabidopsis*

Organisms lacking functional TFIIS perform well under normal growth conditions despite its seemingly crucial role in regulating transcript elongation (Fish and Kane, 2002). Also in *Arabidopsis thaliana* TFIIS absence does not result in any clear morphological defects apart from compromised seed dormancy (Grasser et al., 2009). Intriguingly, the mutation of invariant acidic residues in TFIIS acidic hairpin leads to lethality in yeast (Sigurdsson et al., 2010) and severe growth defects when expressed in Col-0 (Dolata et al., 2015) implying the importance of TFIIS-stimulated RNA cleavage activity of *Arabidopsis* RNAPII similarly to observed in yeast. Therefore, in the first part of this thesis the dominant negative version of TFIIS was used as a molecular tool to shed light on TFIIS importance in regulating transcript elongation in higher eukaryotes.

2.1 Constitutive expression of mutated TFIIS in *tfIIs-1*

The constitutive expression of mutated TFIIS in Col-0 results in a severe growth defects in comparison with Col-0 and *tfIIs-1* (Dolata et al., 2015). Since in the study by Dolata et al., 2015 TFIISmut was integrated into Col-0 background, morphological defects are the result of simultaneously expressed WT TFIIS and transgene-derived TFIISmut. Accordingly, observed phenotype may only partially reflect the consequences of TFIIS mutation and the interpretation of future molecular data obtained from these transgenic plants could have been very challenging. Therefore, the first step in this study was to create a system comprising TFIISmut expression in plants lacking functional TFIIS. The vector containing mutated version of *Arabidopsis* TFIIS, generated by Simon Arnold Mortensen, was initially used. The genomic sequence of TFIIS driven by its native promoter has been mutated using overlapping PCR, leading to the replacement of conserved glutamic and aspartic acid residues to alanines in the positions 309 and 310 of *Arabidopsis* TFIIS (Asp309 -> Ala and Glu310 -> Ala) (Mortensen, 2012). Vector containing mutated TFIIS was next incorporated into *tfIIs-1* since in this T-DNA line no expression of functional TFIIS has been determined (Grasser et al., 2009). However, despite extensive screening no transgenic lines carrying mutated TFIIS (TFIISmut) in *tfIIs-1* could be identified after *Agrobacterium*-mediated transformation in the course of this study. Thus, it was decided to introduce vector containing TFIISmut into plants heterozygous for T-DNA insertion within endogenous TFIIS (*tfIIs* $^{+/-}$) to possibly obtain *tfIIs-1* knockout plants in the subsequent generations by segregation. Accordingly, after *Agrobacterium*-mediated transformation several individuals with transgene-derived TFIISmut in *tfIIs* $^{+/-}$ background were selected. Their progeny (T2 generation) was genotyped in order to identify individuals carrying TFIISmut in *tfIIs-1*. The progeny of two independent transgenic lines was screened (~ 50 plants each) however no individuals carrying TFIISmut in *tfIIs-1* could be identified according to PCR-based genotyping (as exemplified on Figure 10).

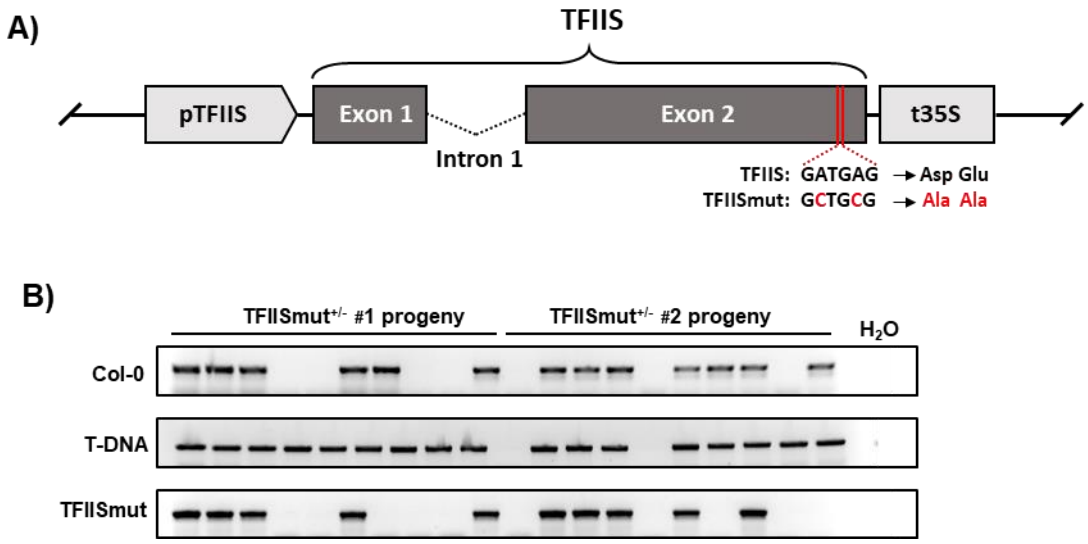


Figure 10. Full length TFIISmut design and transgenic lines genotyping.
(A) Schematic illustration of the transgene with mutated TFIIS (TFIISmut) driven by its native promoter (Mortensen, 2012). Red horizontal lines: point mutations within TFIIS. Red font: changes in DNA and amino acid sequence. pTFIIS: native promoter of TFIIS; t35S: CaMV 35S terminator; Asp: aspartic acid; Glu: glutamic acid; Ala: alanin. (B) PCR-based genotyping of the progeny of plants heterozygous for TFIISmut transgene and T-DNA within endogenous TFIIS (TFIISmut$^{+/-}$). Representative results for 10 individual seedlings of each line are shown. 50 individuals per line were screened.

The inability to obtain transgenic lines expressing TFIISmut in *tfIIs-1* could be caused by those plants lethality as suggested previously (Dolata et al., 2015). Therefore, another approach was applied based on the inducible expression of TFIISmut using β-estradiol system.

2.2 Inducible TFIISmut expression in *tfIIs-1*: design, optimisation and validation

2.2.1 TFIISmut incorporation into β-estradiol inducible system

The β-estradiol inducible system used in this study was designed and created based on the previously described two-component system (Brand et al., 2006). For the purpose of this study a novel one-component system was obtained by fusing pGreen0179 vector together with pMDC150 (containing the activator unit) and pMDC221 (containing the responsive unit) vectors resulting in single plasmid DNA comprising both activator and responsive units (Figure 69). Like in the two-component system, the activation unit is a fusion of the DNA-binding domain of the bacterial repressor LexA, the acidic transactivation domain of VP16 and the regulatory region of the human estrogen receptor. The expression of activator unit was put under the control of UBQ10 promotor as described previously (Dürr et al., 2014) to ensure the ubiquitous expression in all tissues. Continuously expressed activator unit can be further selectively activated by the addition of β-estradiol which allow activator binding to the responsive unit comprising LexA operon. Consequently, responsive unit drives the expression of downstream gene of interest (Zuo et al., 2000) (Figure 11 A).

In this study the mutated version of TFIIS was created by overlapping PCR using Col-0 cDNA as a template. It was decided to use TFIIS coding sequence in order to avoid any splicing defects of transgene-derived TFIISmut. Obtained TFIISmut (Asp309->Ala, Glu310->Ala) was

additionally N-terminally fused with GFP to allow the detection of TFIIS expression by monitoring GFP fluorescence and by using antibodies directed against GFP. GFP-TFIISmut was subsequently introduced into created inducible system by gateway cloning (Figure 11 A). Finally, wild type coding sequence of TFIIS was fused with GFP and placed into inducible system to serve as a control in performed experiments.

Both vectors containing inducible GFP-TFIIS or GFP-TFIISmut were integrated into *tfIIs-1* genome using *Agrobaterium*-mediated transformation. Primary-transformants were identified using hygromycin selection and transgene incorporation into *tfIIs-1* genome was confirmed by PCR-based genotyping (Figure 11 B). Three independent transgenic lines carrying either GFP-TFIIS or GFP-TFIISmut were selected and preliminary analysed. These lines are referred as "iGFP-TFIIS#1-3" and "iGFP-TFIISmut#1-3", respectively.

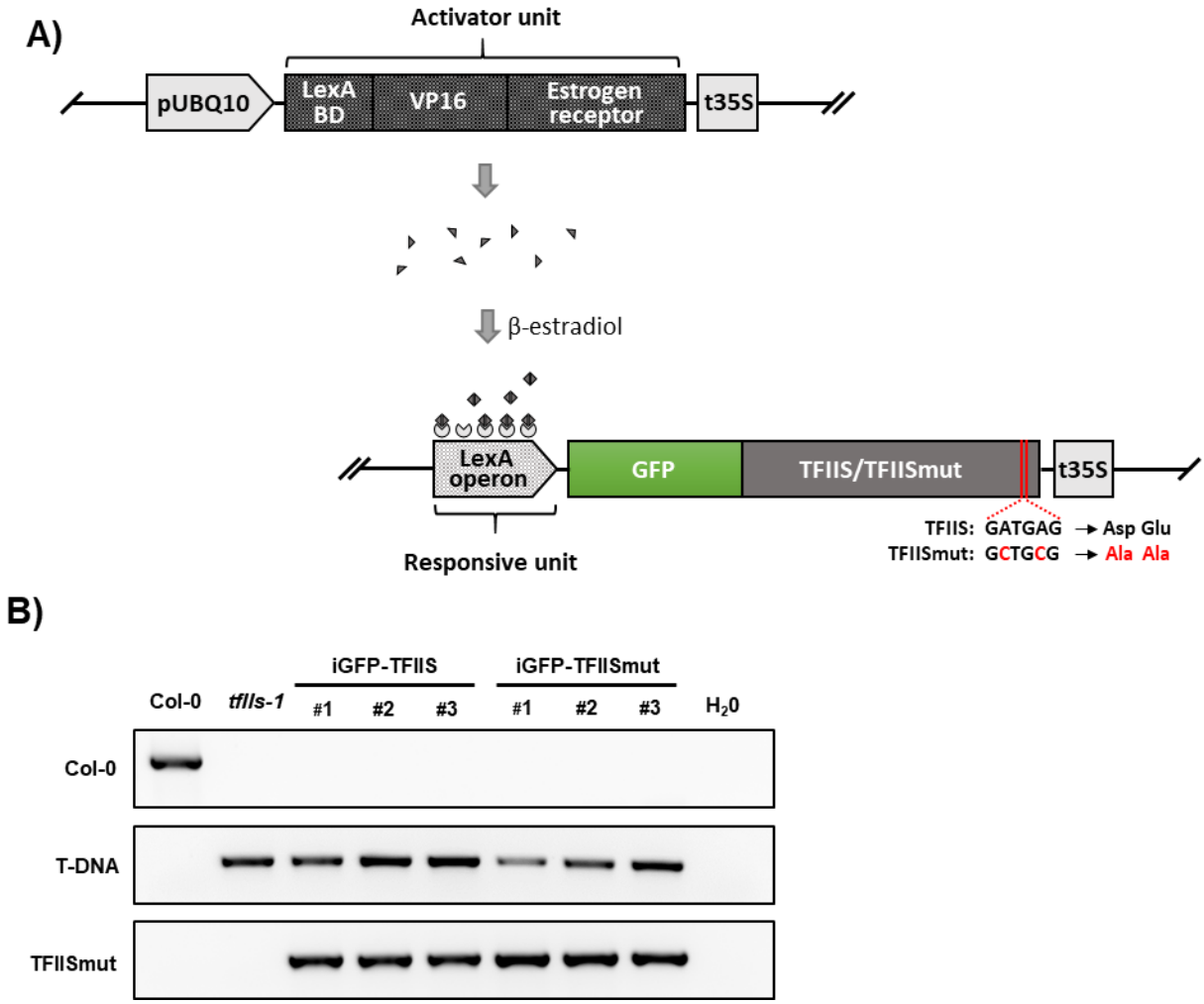


Figure 11. β-estradiol system for inducible expression of GFP-TFIIS and GFP-TFIISmut in *tfIIs-1*.
(A) Schematic illustration of created β-estradiol inducible system containing either GFP-TFIIS or GFP-TFIISmut. Constitutive expression of activator unit (grey triangles) bind to the responsive unit in the presence of β-estradiol determining its activation. pUBQ10: native promoter of *Arabidopsis* Ubiquitin 10; LexA BD: binding domain of LexA operon; VP16: acidic transactivation domain of human VP16; t35S: CaMV 35S terminator. Red horizontal lines: point mutations within TFIIS. Red font: changes in DNA and amino acid sequence. (B) PCR-based genotyping of inducible GFP-TFIIS and GFP-TFIISmut transgenes introduced into *tfIIs-1* (iGFP-TFIIS and iGFP-TFIISmut lines, respectively). DNA was extracted from 8DAS seedlings.

2.2.2 The validation of β-estradiol inducible system

2.2.2.1 Inducible system allows ubiquities expression of target proteins

The kinetic of β-estradiol inducible system has been previously determined showing the highest accumulation of inducible transcript within 24 - 48h after β-estradiol application (Zuo et al., 2000). Similarly controlled β-estradiol inducible system was further successfully used in other studies (Brand et al., 2006; Dürr et al., 2014). Therefore, comparable induction conditions were applied in the pilot experiments to test the functionality of the modified system in selected transgenic lines. *Arabidopsis* seedlings were grown on the MS medium for 5 days and subsequently transferred on MS medium containing 2 µM β-estradiol. After additional 24h of continuous growth in the presence of β-estradiol, GFP expression was determined in *Arabidopsis* roots using illumination microscopy performed with ApoTome system. Several individuals of three independent iGFP-TFIIS and iGFP-TFIISmut transgenic lines were analysed showing ubiquitous expression of target proteins specifically in the nuclei (Figure 12). The expression of both inducible transgenic proteins was primarily detected in the elongation zone (Figure 12, middle panels) and in the root tips with somewhat weaker expression in the meristematic zone (Figure 12, top panels). Importantly, no obvious differences in terms of inducibility or subcellular localisation could be observed between iGFP-TFIIS and iGFP-TFIISmut lines (Figure 12).

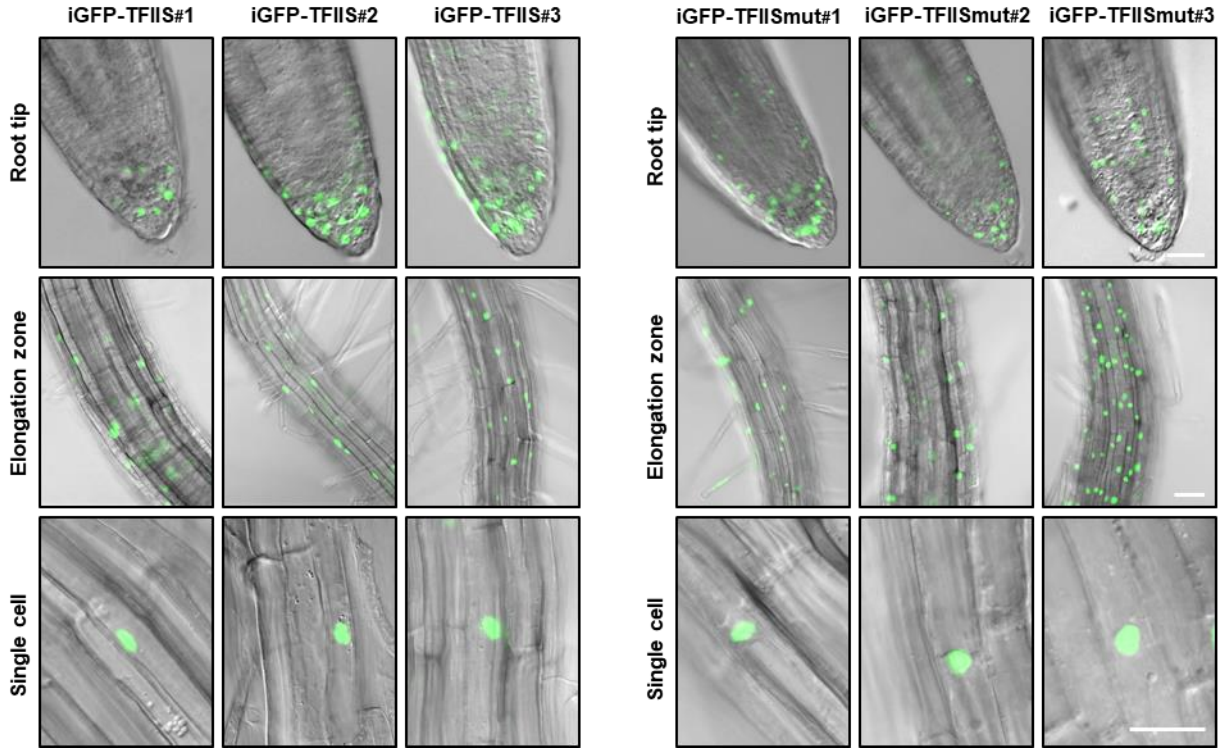


Figure 12. Inducible GFP-TFIIS and GFP-TFIISmut show comparable inducibility and localization in *tfIIs-1* roots.
The inducibility and localisation of inducible GFP-TFIIS and GFP-TFIISmut was studied in respective transgenic lines. 5 DAS seedlings were exposed to β-estradiol for 24h prior to documentation with ApoTome system. In the upper and middle panel, optical sections of the root tip and GFP signal (depicted in green) in roots are shown. A 20X/0.8 objective was used. At the lower panels A 40X/1.4 objective was used. White bars indicate 50 µm.

2.2.2.2 Target proteins are expressed within 3h following β-estradiol application

The big advantage of β-estradiol inducible system application is the temporally controlled expression of the target protein (Brand et al., 2006). In the context of TFIISmut it could potentially limit the secondary effects of arrested transcription, focusing on the early molecular consequences of TFIISmut expression. Previously determined accumulation of inducible transcript at ~ 24 - 48h after β-estradiol application was accompanied by the saturation or even decrease of inducible transcript level over longer induction (Zuo et al., 2000). Therefore, is has been decided not to extend the 24 hours induction, at which point inducible GFP-TFIIS and GFP-TFIISmut proteins were already ubiquitously expressed in selected transgenic lines (Figure 12). Further optimisation of induction conditions was applied to ensure fast and relatively homogenous induction across studied seedling material. To this end, the system was further optimised in terms of β-estradiol concentration, application method as well as plant age. Confocal microscopy (CLSM) was used to monitor GFP expression in *Arabidopsis* roots and leaves, revealing the most robust induction after exposing entire 5 - 9 DAS *Arabidopsis* seedling to 2 µM β-estradiol dissolved in liquid MS media supplemented with 1% sucrose under the vacuum pressure (data not shown). The induction kinetic was subsequently studied using those preoptimized condition.

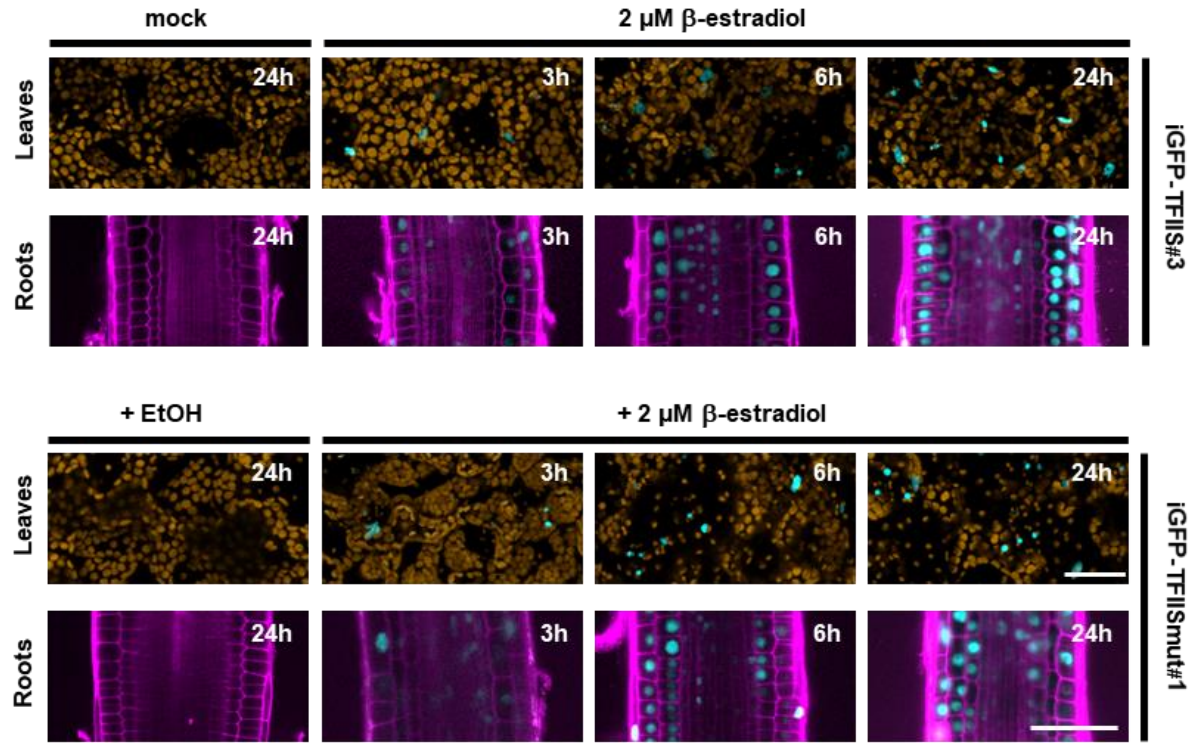


Figure 13. Transgenic iGFP-TFIIS and iGFP-TFIISmut lines show similar induction kinetic.
The expression of transgenic GFP-TFIIS and GFP-TFIISmut was studied in *Arabidopsis* roots and leaves by Confocal laser scanning microscopy (CLSM) following the exposure of 5DAS seedlings to β-estradiol or EtOH (mock). Z-stacking of leaf tissue and optical-section through the differentiated root cells were acquired. GFP signal is shown in cyan. Autofluorescence detected in leaves is shown in brown. Cell wall in roots was counterstained with propidium iodide (magenta). Numbers on the pictures indicate induction time. White bars indicate 100 µm.

In view of comparable inducibility within selected transgenic lines, iGFP-TFIIS#3 and iGFP-TFIISmut#1 were chosen to study the induction kinetic in details based on their similar expression pattern in root tip and elongation zone (Figure 12). 5DAS *Arabidopsis* seedling were exposed to 2 µM β-estradiol and GFP signal emergence was examined by confocal microscopy every 30 min in the main root tip and cotyledon leaves. When analysing different induction timepoints a clear expression of both GFP-TFIIS and GFP-TFIISmut could be observed within ~ 3 hours following β-estradiol application (Figure 13, second panels from the left). The number of cells expressing inducible proteins increased rapidly during subsequent 3 hours and more steadily afterwards (Figure 13, third and fourth panels from the left), similarly to previously determined kinetic (Zuo et al., 2000). Importantly, no expression of neither GFP-TFIIS nor GFP-TFIISmut was detected after 24h mock (EtOH) induction (Figure 13, first panels from the left).

2.2.2.3 Inducible system allows for precisely controlled expression

Western Blot assay was performed to confirm the correctness of inducible fusion proteins with no subproducts as well as to exclude any background expression in mock-induced plants. 9DAS iGFP-TFIIS#3 and iGFP-TFIISmut#1 were exposed to 2 µM β-estradiol or mock induction for 24h. Induced plants were flash-frozen in the liquid nitrogen and the whole protein extracts were studied using Western Blot. A single band corresponding to either GFP-TFIIS or GFP-TFIISmut could be detected at the expected size of 70 kDa (Figure 14, top panels) when using antibodies directed against GFP (α-GFP). No other bands were detected confirming that the previously observed nuclear signal may be attributed to either GFP-TFIIS or GFP-TFIISmut expression (data not shown). Importantly, no bands were detected in the whole protein extract from mock-induced seedlings (Figure 14, top panels). Antibodies against UAP56 were used to ensure comparable amount of total protein extract between analysed samples indicating somewhat stronger expression of inducible target protein in iGFP-TFIIS#3 then in iGFP-TFIISmut#1 line (Figure 14). Overall, those findings support the functionality of created inducible system and are in line with the data obtained by confocal microscopy.

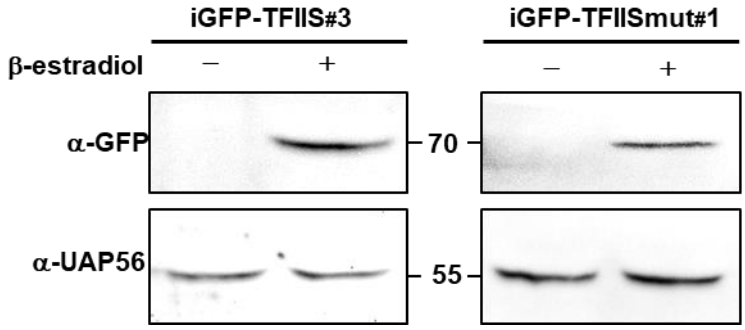


Figure 14. GFP-TFIIS and GFP-TFIISmut are expressed only upon β-estradiol induction.
Immunoblot analysis of whole protein extracts obtained from 10DAS Arabidopsis seedlings with the antibodies against GFP and UAP56 (loading control). Seedlings were exposed to either 2 µM β-estradiol (+) or mock (-) for 24 hours prior to material harvesting. Numbers indicate molecular weight in kDa.

2.2.2.4 TFIISmut expression in *tfIIs-1* results in severe growth defects

The constitutive expression of TFIISmut under its native promoter in plants lacking functional TFIIS has been previously suggested to be lethal (Dolata et al., 2015a) which could presumably explain the inability to obtain those mutants in the course of this study. By using the created β-estradiol inducible system it was possible to mimic the constitutive expression of mutated TFIIS in *tfIIs-1* and evaluate the effect of prolonged TFIISmut expression on *Arabidopsis* growth and development. Hence, the growth of three independent iGFP-TFIIS and iGFP-TFIISmut lines was analysed in comparison to Col-0 and *tfIIs-1* on the MS medium supplemented with either 2 µM β-estradiol or EtOH (mock induction). Plant growth was documented at 7DAS revealing very severe defects for all iGFP-TFIISmut lines exposed to β-estradiol whereas Col-0, *tfIIs-1* and iGFP-TFIIS#1-3 lines developed normally (Figure 15, left panel). The growth of all analysed lines exposed to mock induction was comparable (Figure 15, right panel).

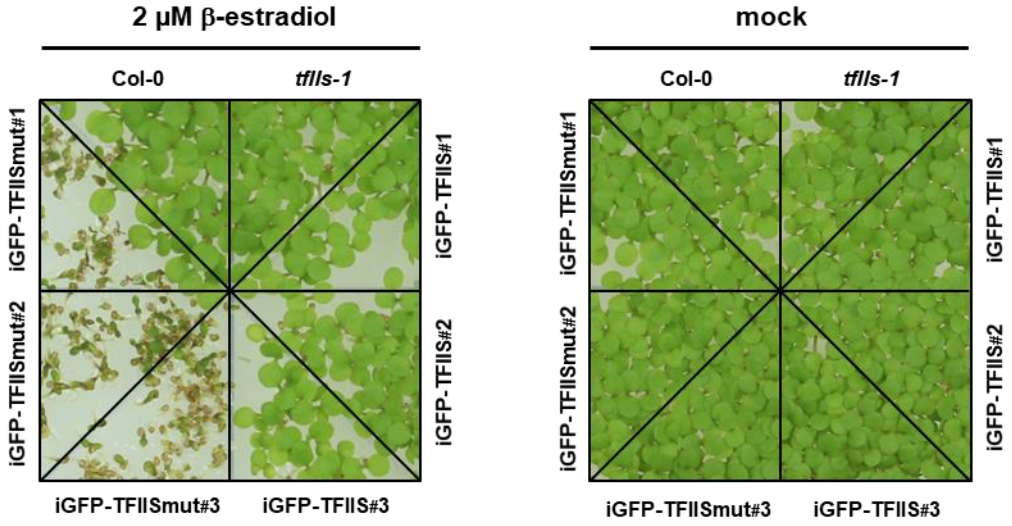


Figure 15. The expression of mutated TFIIS in *tfIIs-1* inhibits *Arabidopsis* growth.
Transgenic lines harbouring inducible GFP-TFIIS or GFP TFIISmut transgenes were grown together with Col-0 and *tfIIs-1* on the MS medium supplemented with 2 µM β-estradiol or EtOH (mock). Pictures were taken at 7DAS.

2.2.3 TFIISmut expression inhibits roots elongation within < 6 hours

In view of temporally controlled expression of inducible target proteins it was expected that observed growth defects triggered by TFIIS mutation would also occur in the time-dependent manner. Thus, the growth of inducible transgenic lines was analysed over time for 8 - 10DAS *Arabidopsis* seedling in terms of main root elongation to allow better determination of inducible system kinetic reflected by morphological aberrations. The kinetic of main root elongation was evaluated since roots growth is linear and symplastic whereas leaf growth is much more complex (Asl et al., 2011). The induction time in this experiment was extended to 48 hours since the morphological respond to TFIISmut expression was expected to take longer than induction itself. Independent seedling of iGFP-TFIIS#3 and iGFP-TFIISmut#1 lines (n = 6) were grown vertically on the MS medium for 8 days and subsequently transferred on MS medium containing either 2 µM β-estradiol or EtOH (mock) The tip of the main root was aligned to the single horizontal line and its elongation was documented over time by taking pictures at 0, 6, 12, 24 and 48 hours after the transfer (Figure 16 A). As a result, iGFP-TFIIS#3

seedlings exposed for 48h to either β-estradiol or mock induction showed comparable elongation rate of the main root (Figure 16 B). In contrary, main root elongation of iGFP-TFIISmut#1 was strongly inhibited after 48h exposure to β-estradiol in comparison with mock-induced plants (Figure 16 C) as well as with iGFP-TFIIS#3 line (Figure 16 B-C).

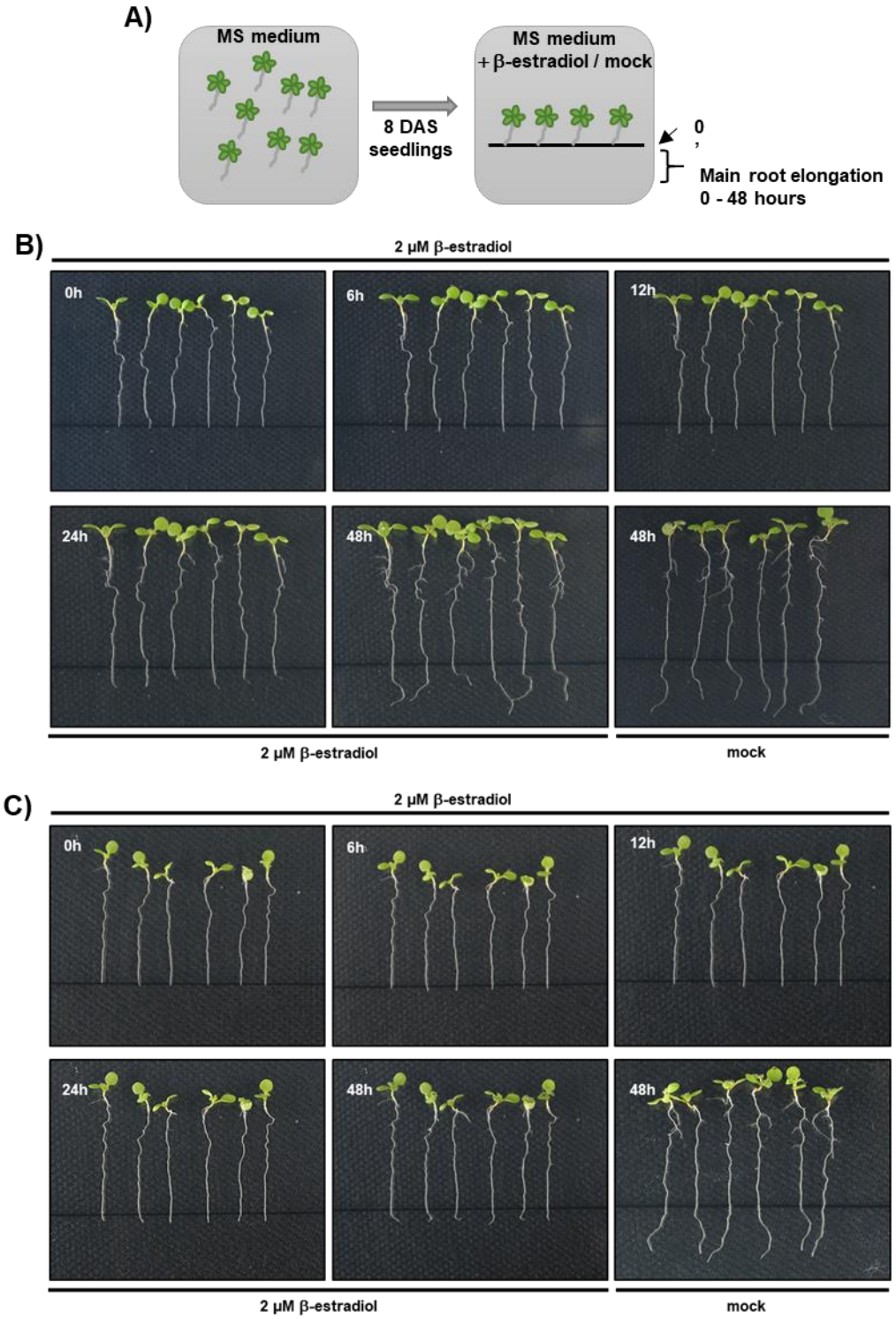


Figure 16. Over-time inhibition of *Arabidopsis* main root growth upon TFIISmut expression.
(A) Schematic illustration of the experiment designed for studding main root elongation kinetic. (B, C) Pictures were taken 0, 6 ,12, 24 and 48 hours after transferring individual iGFP-TFIIS#3 (A) or iGFP-TFIISmut#1 (B) seedlings on MS medium containing either β-estradiol (0-48h) or EtOH (mock; 48h). All plants were grown vertically throughout the assay and transferred between plates using a tweezer.

The length of main root elongation was further measured by ImageJ for each plant to precisely determine their elongation kinetic. iGFP-TFIIS#3 line could be characterised with the linear growth of main root over time upon both β-estradiol and mock induction (Figure 17 A). Although main root elongation of iGFP-TFIISmut#1 line exposed to β-estradiol was also linear, the kinetic was much slower relatively to mock-induced plants as well as iGFP-TFIIS#3 line (Figure 17 A). Additionally, the measurements distribution for β-estradiol-induced iGFP-TFIISmut#1 line showed better fit to polynomial regression then linear regression, indicating stronger inhibition of main root elongation over time (R^2 = 0,97 vs R^2 = 0,84, respectively; Supplementary Figure S 1). The average relative main root elongation (β-estradiol vs mock) was further analysed for each timepoint and statistically analysed between iGFP-TFIIS#3 and iGFP-TFIISmut#1. As a result, significantly compromised main root elongation could be observed for iGFP-TFIISmut#1 already after 6h exposure to β-estradiol relatively to iGFP-TFIIS#3 (Figure 17 B).

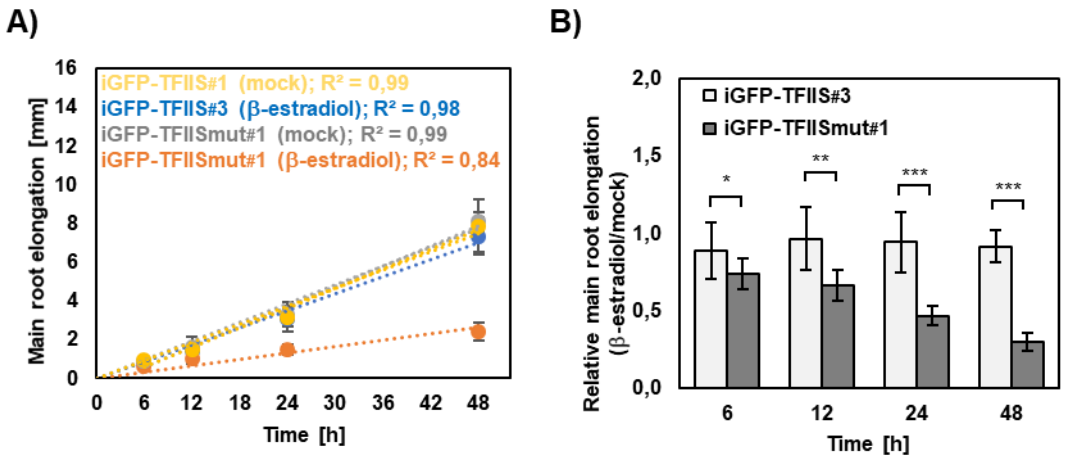


Figure 17. Main root growth kinetic in inhibited within 6 h following β-estradiol application.
Main root elongation was measured by ImageJ in order to calculate the absolute (A) and the relative (B) elongation rate. (A) Dots indicate mean values (n = 6) ± SD (error bars). Dotted lines reflect linear regression; R^2: the coefficient of determination. (B) Bars indicate mean values (n = 6) ± SD (error bars). The outcome the Student's T-Test: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001.

Observed data suggests that the morphological defects triggered by mutated TFIIS follow the expression of inducible protein within the range of few hours. Since observed growth defects are likely a consequence of TFIISmut affecting RNAPII properties, TFIIS and TFIISmut association with RNAPII was further studied in detail by affinity purification coupled to mass spectrometry (AP-MS).

2.2.4 TFIIS and TFIISmut associate with similarly composite TEC

Although TFIIS association and direct interaction with RNAPII is well-established (Awrey et al., 1998; Kettenberger et al., 2003; Xu et al., 2017), the properties of inducibly expressed TFIIS association with transcription machinery has not been addressed. It additionally remains unclear how the mutation within TFIIS acidic loop could affect TEC composition *in vivo* although it has been demonstrated to be redundant for RNAPII binding *in vitro* (Awrey et al., 1998; Kettenberger et al., 2004).

To study association of inducible TFIIS and TFIISmut with *Arabidopsis* TEC, GS-affinity purification coupled to mass spectrometry (AP-MS) was applied (Pfab et al., 2017; Van Leene et al., 2011, 2015). This method has been previously successfully used to characterize other nuclear protein complexes (Dürr et al., 2014; Nelissen et al., 2010). To determine the interactomes of inducible TFIIS and TFIISmut, their coding sequence were framed into β-estradiol inducible system together with N-terminally fused GS-tag (Figure 18 A).

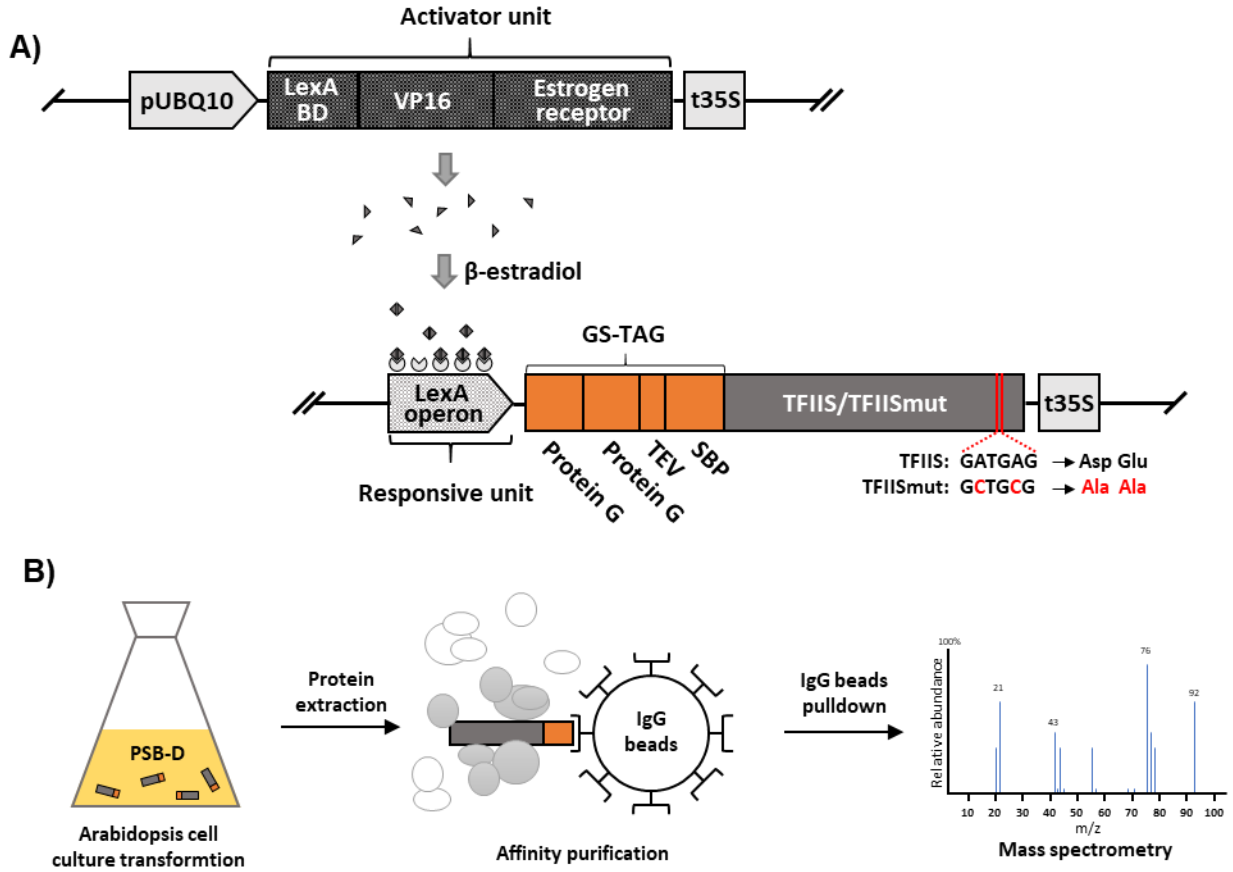


Figure 18. The design of AP-MS assay comprising β-estradiol inducible system.
A) Schematic illustration of β-estradiol system created for inducible expression of GS-TFIIS and GS-TFIISmut in *Arabidopsis* PSB-D cells. Constitutively expressed activator unit (grey triangles) bind to the responsive unit in the presence of β-estradiol determining its activation. pUBQ10: native promoter of *Arabidopsis* Ubiquitin 10; LexA BD: binding domain of LexA operon; VP16: acidic transactivation domain of human VP16; ProteinG: IgG-binding domain of protein G; TEV: tobacco etch virus (TEV) protease cleavage site; SBP, streptavidin-binding peptide. t35S: CaMV 35S terminator. Red horizontal lines: point mutations within TFIIS. Red font: changes in DNA and amino acid sequence. (B) Schematic illustration of affinity purification coupled with mass spectrometry (AP-MS) workflow. Bait proteins were purified together with their putative interactors from the whole protein extract of transgenic PSB-D line using IgG coupled magnetic beads. Copurified proteins were further identified by mass spectrometry. PSB-D: cell culture of *Arabidopsis Landsberg erecta*; IgG: immunoglobulin G.

Obtained vectors were subsequently introduced into *Arabidopsis* cells culture system (PSB-D) by *Agrobacterium*-mediated transformation (Pfab et al., 2017). Transgenic cell culture carrying either GS-TFIIS or GS-TFIISmut ("iGS-TFIIS" and "iGS-TFIISmut" transgenic line, respectively) were upscaled and supplemented with 2 µM β-estradiol for 24h prior to cell culture harvesting. Inducibly expressed GS-TFIIS and GS-TFIISmut (referred to as "bait") were next purified from the whole protein extract together with their putative interactors by one-step affinity purification (AP) using IgG coupled magnetic beads. The eluates were further separated by SDS-PAGE and stained with Coomassie Blue. CBB stained gels were next trypsin digested in-gel followed by peptides extraction and proteins identification mass spectrometry (Figure 18 B).

The samples of whole protein extracts (Input) obtained from transgenic iGS-TFIIS and iGS-TFIISmut PSB-D cells were resolved by SDS-PAGE and Coomassie stained revealing similar band patterns between analysed samples. In the AP eluates bands corresponding to inducibly expressed GS-TFIIS or GS-TFIISmut fusion protein (~ 70 kDa) were clearly visible together with some additional faint bands indicating substoichiometric purification (Figure 19 A). The association of TFIIS and TFIISmut with RNAPII was further studied comparatively by Western Blot using the antibodies against largest RNAPII subunit (NRPB1) phosphorylated at the CTD position Ser2 (α-CTD-S2P) and Ser5 (α-CTD-S5P) as well as non-phospho specific antibodies (α-CTD).

With the antibodies directed against phosphorylated RNAPII-CTD (α-CTD-S2P and α-CTD-S5P), a hypo- (NRPB1A) and a hyper-phosphorylated (NRPB1O) form of the largest RNAPII subunit were detected in both GS-TFIIS or GS-TFIISmut AP eluates (Figure 19 B). In contrary, when using α-CTD antibodies only a single band likely corresponding to unphosphorylated and/or hypo-phosphorylated form of NRPB1 was detected (Figure 19 B). RNAPII-S5P and unphosphorylated RNAPII associated with both TFIIS and TFIISmut at similar ratio whereas RNAPII-S2P was ~3-fold enriched in GS-TFIISmut pulldown relative to GS-TFIIS (Figure 19 B).

AP eluates were further subjected for mass spectrometry analysis. Identified proteins were subsequently proceeded to increase interactomes reliability. Only proteins with a protein score higher than 80 and at least two peptides with an individual peptide score > 25, found in at least two out of three purifications were selected for further analysis. Additionally, proteins routinely copurified with empty GS used as a "bait" (Antosz et al., 2017) were subtracted to remove unspecifically bound proteins, resulting in 280 and 256 nuclear proteins immunoprecipitated together with GS-TFIIS and GS-TFIISmut, respectively (Figure 19 C). Identified interactomes overlapped only partially (~60%) so that distinct subgroups specific for each bait could be identified (Figure 19 D).

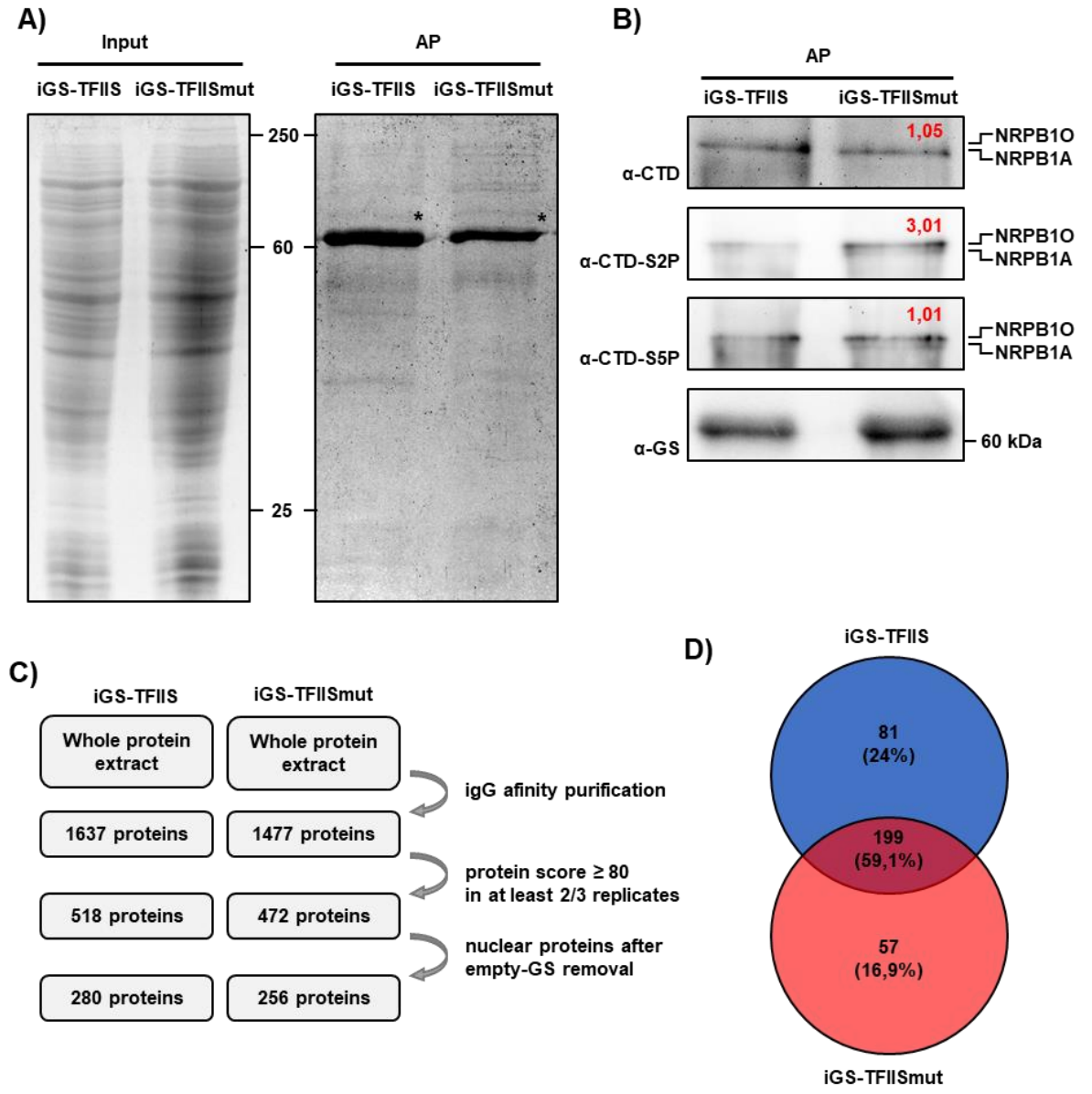


Figure 19. Identification of TFIIS and TFIISmut interactomes following AP-MS assay.
(A) Total protein extracts of transgenic *Arabidopsis* PSD-B cells expressing inducible GS-TFIIS or GS-TFIISmut and the eluates of their affinity purifications (AP) were separated by 9% SDS-PAGE and stained with Coomassie Blue. Black asterisks indicate the bands corresponding to bait proteins. Transgenic PSD-B cells were supplemented with 2 μM β–estradiol 24h prior to harvesting. Numbers indicate molecular weight in kDa. (B) Immunoblot analysis of GS-TFIIS and GS-TFIISmut AP eluates with the antibodies against the CTD part of NRPB1 phosphorylated at Ser2 (α-CTD-S2P) and Ser5 position (α-CTD-S5P), non-phospho specific antibodies (α-CTD) as well as GS tag (α-GS, loading control). NRPB1A indicates a hypo- and NRPB1O a hyper-phosphorylated forms of NRPB1. Red font indicates signal intensity ratios of NRPB1A + NRPB1O in GS-TFIISmut relatively to GS-TFIIS, both normalised to loading control (α-GS). Intensities were measured by ImageJ. (C) The workflow of data proceeding following mass spectrometry. (D) Overlap between GS-TFIIS or GS-TFIISmut interactomes depicted on Venn diagram.

GO enrichment analysis was next performed to identify overrepresented groups in terms of biological function within GS-TFIIS and GS-TFIISmut interactomes. Most significantly overrepresented GO terms (FDR < 0.001) fully overlapped between GS-TFIIS and GS-TFIISmut interactomes and mostly referred to RNAPII-driven transcription and other co-transcriptional processed (Table 1).

Table 1. The overlap between overrepresented GO terms across GS-TFIIS and GS-TFIISmut interactomes.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO for interactomes specified at Figure 19C. All significantly enriched GO terms (false discovery rate (FDR) < 0.05) are shown in the table. FE: fold enrichment.

		GS-TFIIS			GS-TFIISmut		
Term	**GO accession**	**log_2FE**	**p-value**	**FDR**	**log_2FE**	**p-value**	**FDR**
transcription elongation from RNA polymerase II promoter	GO:0006368	4,30	1,51E-06	1,50E-05	4,66	4,06E-08	4,25E-07
regulation of translation	GO:0006417	4,31	4,25E-10	5,71E-09	3,94	8,48E-07	7,59E-06
rRNA metabolic process	GO:0016072	3,61	1,41E-14	3,32E-13	3,20	4,83E-09	5,67E-08
nuclear transport	GO:0051169	3,00	1,05E-06	1,09E-05	3,14	4,33E-07	4,07E-06
RNA splicing, via transesterification reactions	GO:0000375	2,91	1,33E-09	1,47E-08	3,05	3,24E-10	4,69E-09
mRNA splicing, via spliceosome	GO:0000398	2,85	7,43E-10	8,72E-09	2,99	1,66E-10	2,84E-09
mRNA processing	GO:0006397	2,82	1,06E-12	1,81E-11	2,96	1,38E-13	5,18E-12
translation	GO:0006412	3,11	2,64E-17	1,24E-15	2,91	9,71E-13	2,61E-11
transcription from RNA polymerase II promoter	GO:0006366	1,92	4,78E-06	4,28E-05	2,37	2,14E-09	2,68E-08
RNA metabolic process	GO:0016070	1,95	1,02E-20	1,93E-18	2,07	1,92E-22	3,61E-20
cellular component biogenesis	GO:0044085	2,16	9,25E-14	1,74E-12	1,94	1,18E-09	1,59E-08

Both interactomes were further examined in context of transcription-related factors representing the set of 24 and 34 proteins immunoprecipitated with GS-TFIIS and GS-TFIISmut, respectively (Table 2). GS-TFIIS and GS-TFIISmut interactomes strongly overlapped in terms of isolated complexes and individual transcription-related proteins including diverse subunits of RNAPII, 5 subunits of PAF1-C, various TEFs (SPT5-2, SPT6L, FACT and Elongator) as well as NAP1 proteins and HDACs (Table 2) resembling the interactome of constitutively expressed TFIIS (Supplementary Table 1). Surprisingly ELF8 subunit of *Arabidopsis* PAF1-C was not identified among the interactomes of inducibly expressed baits although it was reproducibly copurified with GS-TFIIS driven by its native promoter (Supplementary Table 1). Additionally, putative components of *Arabidopsis* INO80 complex were identified together with inducible GS-TFIIS and GS-TFIISmut as well as some transcription-related factor copurified specifically with GS-TFIISmut including IWS1a and HAF1 (SAGA complex) and RIN1 (Table 2).

Table 2. Transcription-related proteins copurified with GS-TFIIS and GS-TFIISmut.
The list of transcription-related proteins copurified with GS-TFIIS and GS-TFIISmut during AP-MS followed by data proceeding (Figure 19 C). The numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent experiments.

TFIIS	TFIISmut	Interactor	Complex	Process	AGI
3748 / 3	2882 / 3	TFIIS	TFIIS	Transcription	AT2G38560
1487 / 3	1763 / 3	NRPB1	Polymerase II	Transcription	AT4G35800
1378 / 3	1319 / 3	NRPB2	Polymerase II	Transcription	AT4G21710

TFIIS	TFIISmut	Interactor	Complex	Process	AGI
466 / 3	551 / 3	NRP(B/D/E)3a	Polymerase II	Transcription	AT2G15430
462 / 3	474 / 3	NRP(A/B/C/D)5	Polymerase II	Transcription	AT3G22320
216 / 3	149 / 2	NRPB7	Polymerase II	Transcription	AT5G59180
	105 / 2	NRP(B/D/E)9a	Polymerase II	Transcription	AT3G16980
904 / 3	859 / 3	CDC73	PAF-C	Transcription	AT3G22590
504 / 3	416 / 3	LEO1, VIP4	PAF-C	Transcription	AT5G61150
553 / 3	408 / 3	PAF1, ELF7	PAF-C	Transcription	AT1G79730
438 / 3	408 / 2	SKI8, VIP3	PAF-C	Transcription	AT4G29830
327 / 3	245 / 3	RTF1, VIP5	PAF-C	Transcription	AT1G61040
342 / 2	546 / 3	SPT6-1, SPT6L	SPT6	Transcription	AT1G65440
	256 / 3	IWS1a	IWS1	Transcription	AT1G32130
364 / 3	649 / 3	SPT5-2	SPT4/SPT5	Transcription	AT4G08350
435 / 2	544 / 3	SPT16	FACT	Transcription	AT4G10710
211 / 2	261 / 3	SSRP1	FACT	Transcription	AT3G28730
218 / 2	171 / 2	ELP1, ELO2	Elongator	Transcription	AT5G13680
	548 / 3	RIN1	SWR1/NuA4, INO80	Transcription	AT5G22330
300 / 2	432 / 3	RVB21	SWR1/NuA4, INO80	Transcription	AT5G67630
151 / 2	198 / 3	AtARP4	SWR1/NuA4, INO80	Transcription	AT1G18450
	146 / 3	AtSPT7 / HAF1	SAGA_SPT putative	Transcription	AT1G32750
131 / 2	212 / 2	AtNAPL1	NAP1	Transcription	AT4G26110
308 / 3	236 / 3	AtNAPL2	NAP1	Transcription	AT2G19480
506 / 2	422 / 3	HTB2	Histone H2B family	Transcription	AT5G22880
262 / 3	188 / 2	HDT2	Deacetylase	Transcription	AT5G22650
411 / 3	173 / 3	HDT3	Deacetylase	Transcription	AT5G03740

Taken together, inducibly expressed GS-TFIIS and GS-TFIISmut associate with similarly composite TEC in *Arabidopsis*. Importantly, those interactomes also strongly resemble the TEC copurified with constitutively expressed GS-TFIIS. It suggests that both GS-TFIIS and GS-TFIISmut are being incorporated into functional TECs shortly after being expressed by β-estradiol inducible system. Moreover, observed molecular and morphological consequences of TFIISmut expression are likely the result of its direct influence on TEC machinery unbiased by the inducible system.

2.3 Direct molecular consequences of TFIISmut expression

TFIIS biological function and its importance in regulating transcript elongation *in vivo* remain elusive although TFIISmut expression leads to severe growth defects in yeast and *Arabidopsis*. In yeast TFIISmut was shown to inhibits transcription through pause sites (Sigurdsson et al., 2010). In view of that, it has been hypothesised that spacial determination of RNAPII properties following TFIISmut expression could shed light on TFIIS functions *in vivo*. Thus, the molecular consequences of TFIISmut expression will be demonstrated in the following chapters with the focus on transcriptome rearrangement coupled with the changes in active RNAPII occupancy. iGFP-TFIIS#3 and iGFP-TFIISmut#1 lines were chosen for further experiments based on their previous characterisation and referred thereafter as “iGFP-TFIIS” and “iGFP-TFIISmut”, respectively.

2.3.1 DEGs upon TFIISmut expression: RNA-seq

To better understand the biological importance of TFIIS, the genome-wide analysis of transcriptomic changes upon TFIISmut expression was performed by RNA-seq. Based on

induction kinetic determination (chapters 2.2.2.2 and 2.2.3), 10DAS iGFP-TFIIS and iGFP-TFIISmut seedlings were studied following their exposure to 24h β-estradiol or mock induction. Additionally, both lines were analysed in the context of 6h induction with β-estradiol to monitor the early transcriptomic changes triggered by TFIISmut.

TFIISmut expression leads to broad transcriptomic changes

Total RNA was isolated from iGFP-TFIIS and iGFP-TFIISmut seedlings exposed to either β-estradiol (6h or 24h) or mock treatment (24h) using RNeasy R Mini Plant kit (Qiagen). cDNA libraries preparation and sequencing were performed by Kompetenzzentrum Fluoreszente Bioanalytik (KFB; Regensburg). cDNA libraries were created using TruSeq Stranded mRNA Sample Preparation kit (Illumina). Three biological replicates for each line/condition were sequenced on Illumina HiSeq 1000, generating a total number of ~ 27 - 37 million high quality reads per sample. The initial analysis of RNA-seq data was performed by Dr. Kevin Begcy. Obtained reads were aligned onto the *Arabidopsis* reference transcriptome assembly (TAIR10) and Fragments Per Kilobase Of Exon Per Million Fragments Mapped (FPKM) were calculated for all annotated *Arabidopsis* loci among analysed samples. The data distribution across individual biological replicates of each line/condition was highly reproducible as determined by principal component analysis (PCA; Supplementary Figure S 2). Next, the functionality of inducible system was evaluated by analysing the changes in TFIIS transcript level upon β-estradiol induction relatively to mock treatment in iGFP-TFIIS and iGFP-TFIISmut. In both transgenic lines there was a clear increase in TFIIS transcript level already 6h after β-estradiol application with stronger induction in iGFP-TFIIS line relatively to iGFP-TFIISmut ($log_2FC = 5,17$ vs $log_2FC = 1,56$, respectively; Figure 20). Longer exposure to β-estradiol (24h) resulted in further increase of TFIIS transcript level in iGFP-TFIISmut ($log_2FC = 2,31$) whereas in iGFP-TFIIS line TFIIS transcript level remained comparable ($log_2FC = 5,14$) to the value determined after 6h, indicating the saturation of inducible system.

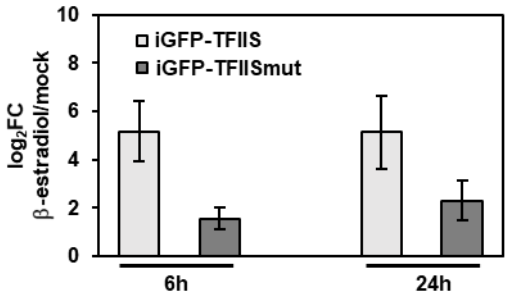


Figure 20. The increase of TFIIS and TFIISmut transcript levels upon β-estradiol induction.
Fragments Per Kilobase Of Exon Per Million Fragments Mapped (FPKM) were calculated for iGFP-TFIIS and iGFP-TFIISmut transgenes upon β-estradiol (6h or 24h) and mock (24h) induction based on the result from RNA-seq. Bars reflect mean log_2 fold change (FC; β-estradiol vs mock) for three biological replicates. Error bars reflect SD.

Using FPKM values, differentially expressed genes (DEGs; $|log2FC| > 1$, p-value < 0.001 and FDR < 0.001) were determined between analysed lines and/or conditions. First, iGFP-TFIIS and iGFP-TFIISmut lines exposed to 24h β-estradiol or mock induction were analysed comparatively. Over a thousand of DEGs were identified in iGFP-TFIISmut line induced with β-estradiol when compared with respective controls (Figure 21 A). These significant

transcriptomic changes were dominated by gene upregulation upon TFIISmut expression (Figure 21 A). The transcriptomic changes between analysed controls were less profound (Figure 21 A).

To better understand transcriptome rearrangement upon TFIISmut expression, a heatmap of all transcriptionally active genes across analysed lines/conditions (FPKM ≥ 5 in at least one line/condition) was generated. As seen on the Figure 21 B-C, the expression of GFP-TFIISmut led to broad transcriptomic changes when compared with control samples. In line with the overall expression patterns seen at the heatmap, hierarchical clustering of analysed samples revealed clear separation of β-estradiol treated iGFP-TFIISmut from respective controls (Figure 21 D). Mock-treated iGFP-TFIIS and iGFP-TFIISmut lines were closest in the distance and clustered together with slightly separated iGFP-TFIIS induced with β-estradiol (Figure 21 D).

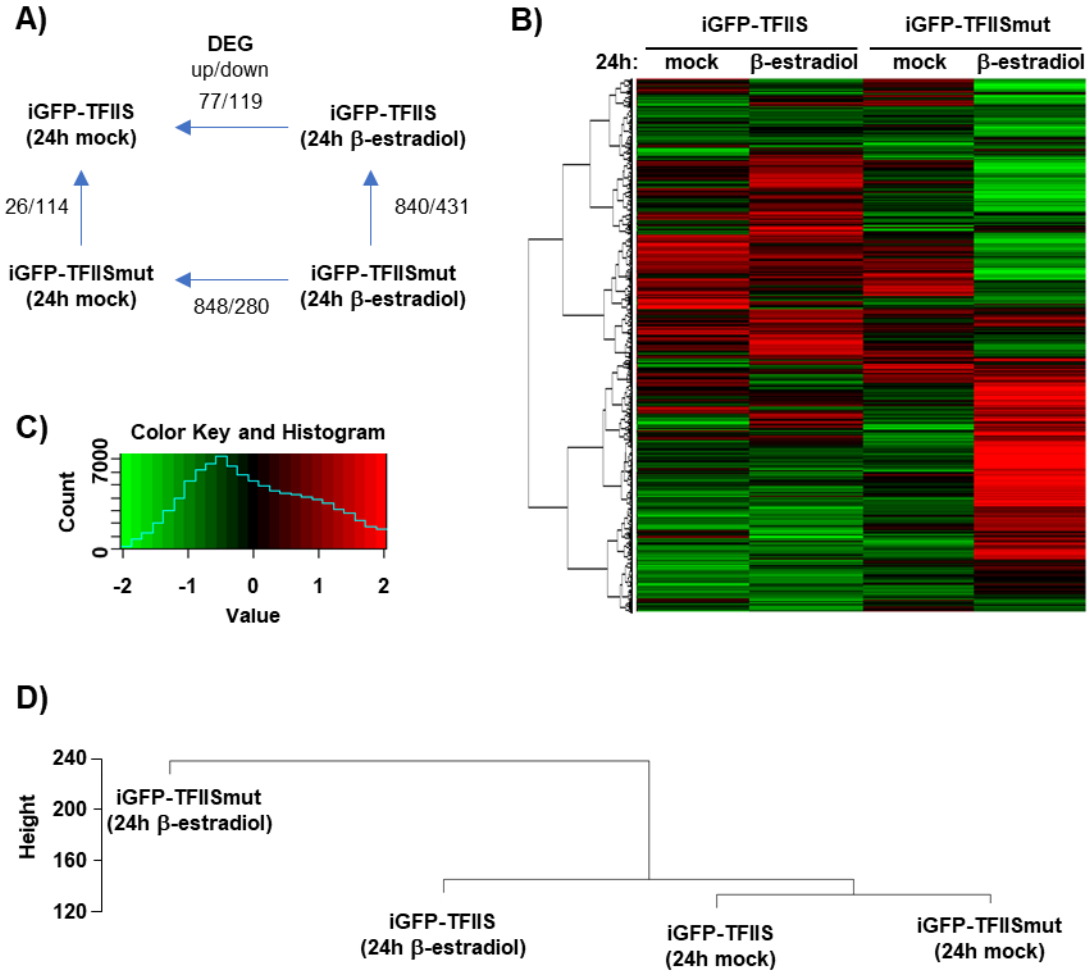


Figure 21. The expression of mutated TFIIS leads to broad transcriptomic changes.
A) Schematic illustration of differentially expressed genes (DEGs) analysis between analysed lines/conditions. Arrows indicate the direction of the comparison and number show DEG with $|\log_2 FC| > 1$, p-value < 0.001 and false discovery rate (FDR) < 0.001. (B-C) Heatmap visualisation of analysed lines/conditions. Only genes with FPKM ≥ 5 in at least one line/condition were considered to build the heatmap (n = 15836). Hierarchical clustering is shown on the left side. (C) Heatmap legend. Red-green gradient indicated the $\log_2$ FPKM in the [-2; 2] colour range ("Value"). Blue line indicates the number of values in the given colour range. (D) Dendrogram of RNA-seq analysis performed for the averaged biological replicates (n = 3). The cluster analysis was performed through complete linkage and Euclidean distance as a measure of similarity. Figures B-D were created by Dr. Kevin Begcy.

Transcriptomic changes in iGFP-TFIISmut were next analysed relatively to iGFP-TFIIIS following 6 hours β-estradiol induction. This analysis revealed the number of 178 DEGs between studied lines (82 up- and 96 downregulated genes; Supplementary Figure S 3 A). When compared with respective lines exposed to either β-estradiol or mock for 24h, several hundreds of DEGs were identified in both iGFP-TFIIIS and iGFP-TFIISmut (Supplementary Figure S 3 A). Heatmap visualisation of all transcriptionally active genes across analysed lines/conditions revealed surprisingly similar transcriptome profile between iGFP-TFIISmut and iGFP-TFIIS after exposure to β-estradiol for 6h (Supplementary Figure S 3 B-C). These profiles were clearly distinct from the samples induced for 24h (Supplementary Figure S 3 B-C). This apparent inconsistency was further supported by hierarchical clustering. A close distance between samples exposed to β-estradiol for 6h was identified showing their subsequent clustering together with iGFP-TFIISmut exposed to β-estradiol for 24h and a clear separation from remaining control samples (Supplementary Figure S 3 D). In line with hierarchical clustering, distinct transcriptome profile in iGFP-TFIIS upon 6h β-estradiol induction seemed to be rearranged after longer exposure to β-estradiol, consequently resembling the profile observed after 24h mock induction, unlike in iGFP-TFIISmut (Supplementary Figure S 3 B-D). It may suggest that observed genome-wide transcriptomic changes upon 6h β-estradiol induction reflect certain oscillation in genes expression rather than specific response to TFIIS and TFIISmut expression.

TFIISmut triggers stress and defence response

840 DEGs unregulated specifically upon TFIISmut expression (iGFP-TFIISmut vs iGFP-TFIIIS; 24h β-estradiol induction) were further analysed in terms of GO enrichment to unravel biological processes behind transcriptomic changes. Upregulated DEGs identified between iGFP-TFIISmut and iGFP-TFIIIS exposed to 24h mock induction were subtracted from the analysis resulting in 835 genes subjected for GO enrichment analysis in terms of biological processes. Among many identified processes, GO analysis revealed the upregulation of genes involved in immune and defence response (Table 3) and ultimately cell death (Supplementary Table 2 and 3) following TFIISmut expression. Similar molecular response could be observed in iGFP-TFIISmut already after 6h β-estradiol induction (vs 24h mock induction), but not in iGFP-TFIIS (Supplementary Table 8). These data suggest broad transcriptional induction of various biological pathways to counteract the consequences of TFIISmut expression likely interfering with transcriptional regulation of crucial cellular processes.

Table 3. TFIISmut expression triggers stress, defence and immune response.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO among DEGs significantly upregulated upon GFP-TFIISmut expression relatively to GFP-TFIIS induction. GO terms with the frequency < 10% and $\log_{10}$ FDR < -10 are show in the table. Redundant proteins were removed by REViGO. Frequency indicate the percentage of each GO term in the whole UniProt database. All overrepresented GO terms are shown in Supplementary Table 2.

GO term ID	description	Frequency [%]	$\log_{10}$ FDR
GO:0010200	response to chitin	0,5	-24,64
GO:0006952	defense response	6,1	-24,57
GO:0009607	response to biotic stimulus	5,2	-22,31
GO:0051707	response to other organism	5,0	-21,00

GO term ID	description	Frequency [%]	log_{10} FDR
GO:0009743	response to carbohydrate	0,5	-20,28
GO:0051704	multi-organism process	3,4	-16,41
GO:0006955	immune response	1,4	-12,22
GO:0002376	immune system process	1,5	-12,22

Cell metabolisms is compromised upon TFIISmut expression

Next, 431 DEGs downregulated upon TFIISmut expression were categorised into enriched categories according to GO analysis. In terms of biological processes, TFIISmut expression triggered the downregulation of many genes involved in lipid localisation, photosynthesis, cell redox homeostasis as well as other stimulus and stress response processes (Table 4; Supplementary Table 8. Supplementary Table 3 and 4).

Table 4. TFIISmut expression compromises various biological processes following 24h induction.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO among DEGs significantly downregulated upon GFP-TFIISmut expression relatively to GFP-TFIIS induction. GO terms with the frequency < 10% and log_{10} FDR < -10 are show in the table. Redundant proteins were removed by REViGO. Frequency indicate the percentage of each GO term in the whole UniProt database. All overrepresented GO terms are shown in Supplementary Table 3.

GO term ID	description	Frequency [%]	log_{10} FDR
GO:0010876	lipid localization	0,7	-8,85
GO:0015979	photosynthesis	1,1	-3,77
GO:0045454	cell redox homeostasis	0,6	-3,33

"Lipid localisation" and "Photosynthesis" GO terms were additionally found enriched among downregulated DEGs in iGFP-TFIISmut already after 6h exposure to β-estradiol (vs 6h iGFP-TFIIS) (Table 5). Moreover, several GO categories related to cell redox homeostasis (enriched after 24h β-estradiol induction) were also identified among those genes. That finding may imply generally compromised cellular respiration and energy production in the presence of TFIISmut (Table 5).

Table 5. TFIISmut expression compromises various biological processes following 6 h induction.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO among DEGs significantly downregulated upon GFP-TFIISmut expression relatively to GFP-TFIIS induction. GO terms with the frequency < 10% and log_{10} FDR < -10 are show in the table. Redundant proteins were removed by REViGO. Frequency indicate the percentage of each GO term in the whole UniProt database. All overrepresented GO terms are shown in Supplementary Table 7.

GO term ID	description	Frequency [%]	log_{10} FDR
GO:0006091	generation of precursor metabolites and energy	1,5	-10,59
GO:0010876	lipid localization	0,7	-6,52
GO:0015979	photosynthesis	1,1	-6,52
GO:0055114	oxidation-reduction process	7,5	-4,70
GO:0045333	cellular respiration	0,5	-4,31
GO:0006979	response to oxidative stress	1,9	-3,64

"Photosynthesis" and "Photosynthetic electron transport chain" GO terms were also identified as enriched among downregulated DEGs in both iGFP-TFIISmut and iGFP-TFIIIS after

6h exposure in β-estradiol (vs 24h mock). Importantly, plants exposed to β-estradiol for 6h were harvested at the different time of day-light cycle then plants induced for 24h (2pm vs 8am, respectively) Thus, observed fluctuations in transcript level of photosynthesis-related genes are not surprising. Photosynthesis-related genes could be further found among upregulated DEGs in iGFP-TFIIS after 24h β-estradiol induction (vs 6h β-estradiol) but not in iGFP-TFIISmut (Supplementary Table 9). In contrary, transcriptomic changes upon prolonged β-estradiol induction (24h vs 6h) in iGFP-TFIISmut were dominated by a stress and defence response (Supplementary Table 8 and 9).

These finding may reflect certain transcriptomic oscillations freely rearranged in the presence of TFIIS but strongly compromised by TFIISmut expression. In line with that hypothesis transcriptomic changes following longer β-estradiol induction in GFP-TFIISmut become more extreme in comparison with mock induction, whereas transcriptional profiles in iGFP-TFIIS become more alike (Supplementary Figure S 3).

2.3.2 Active RNAPII occupancy in the presence of mutated TFIIS

To better understand observed transcriptomic changes upon TFIISmut expression, chromatin immunoprecipitation (ChIP) was employed to study GFP-TFIIS and GFP-TFIISmut association with actively transcribe genes accompanied with the rearrangement of active RNAPII occupancy.

2.3.2.1 TFIIS vs TFIISmut distribution over actively transcribed genes

TFIIS occupancy in the chromatin context has been shown to strongly resemble RNAPII distribution in yeast (Ghavi-Helm et al., 2008; Pinskaya et al., 2014) while TFIIS distribution over transcription units in plants as well as TFIISmut association with chromatin remain uncharacterised. Therefore, GFP-TFIISmut occupancy was first analysed comparatively to GFP-TFIIS in transgenic lines exposed to 24h β-estradiol or mock induction. GFP-trap (ChromoTek) was used to immunoprecipitate GFP-TFIIS or GFP-TFIISmut proteins crosslinked with their proximal DNA fragments extracted from 10DAS iGFP-TFIIS and iGFP-TFIISmut line, respectively. For comparison, ChIP experiments were additionally performed with an antibody directed against histone H3 and without any antibody (NoAb) as a negative control.

GFP-TFIIS and GFP-TFIISmut occupancy was quantified by qPCR at both 5’ and 3’ end of a long *Arabidopsis* genes At3g02260 and At1g48090 (~17,5 kb and 26,4 kb, respectively) ensuring good spacial resolution between analysed fragments (“2” and “3” fragments; Figure 22 A). The enrichment was additionally calculated at putative promoter region (“1”; Figure 22 A) and at the transcriptionally inactive region downstream transcription termination site (TES) (“4”; Figure 22A). Noteworthy, according to RNA-seq the transcript level of both At3g02260 and At1g48090 was comparable between iGFP-TFIIS and iGFP-TFIISmut transgenic lines regardless applied induction (data not shown).

Both GFP-TFIIS and GFP-TFIISmut were significantly enriched at the putative promotor region and transcribed regions but not at the transcriptionally inactive region of analysed loci (Figure 22 B-C). Additionally, GFP-TFIISmut occupancy was significantly increased relatively to GFP-TFIIS over At3g02260 and At1g48090 transcribed region but not at the promoter

(Figure 22 B-C). The distribution of Histone H3 over analysed genomic region was not significantly altered between analysed samples suggesting their comparability (Supplementary Figure S 4). Thus, higher level of mutated TFIIS could likely reflect some perturbation in RNAPII distribution since TFIIS has been suggested to bind RNAPII in the stoichiometric amounts (Awrey et al., 1998; Xu et al., 2017).

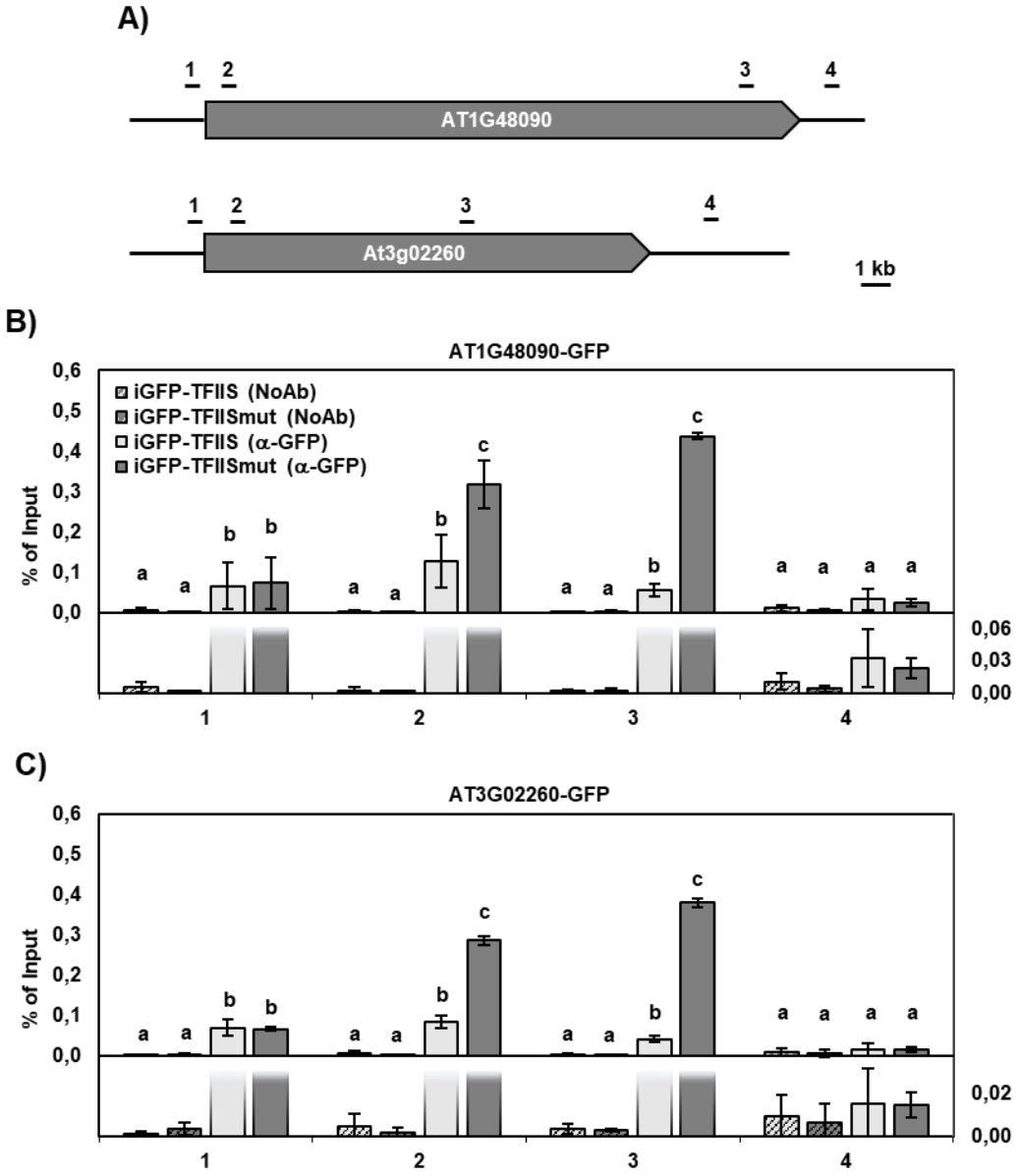


Figure 22. TFIISmut accumulates over transcriptionally active regions.
(A) Schematic representation of At3g02260 and At1g48090 loci with the transcribed (grey boxes) and transcriptionally inactive regions (black lines). The bars above indicate the relative positions of the regions analysed by qPCR. (B-C) ChIP analysis of At1g48090 (B) and At3g02260 (C) with α-GFP as well as without any antibody (NoAb). For ChIP experiments percentage of Input was determined by qPCR and statistically analysed between samples using one-way ANOVA. The letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05). Error bars indicate SD of at least two biological and two technical replicates.

2.3.2.2 Active RNAPII accumulates at transcribed units upon TFIISmut expression

The molecular consequences of TFIIS mutation were further studied in the context of active RNAPII behaviour by analysing the occupancy of NRPB1 phosphorylated at the CTD position Ser2 (RNAPII-S2P) and Ser5 (RNAPII-S5P). Changes in active RNAPII occupancy were determined comparatively between iGFP-TFIIS and iGFP-TFIISmut lines exposed to 24h β-estradiol or mock induction.

Following DNA immunoprecipitation with antibodies directed against RNAPII-S2P or RNAPII-S5P, active RNAPII occupancy was quantified by qPCR at 5' and 3' ends of *Arabidopsis* genes described in the previous chapter (fragments "2" and "3" respectively; Figure 22 A). When looking at active RNAPII occupancy over At3g02260, both RNAPII-S2P and RNAPII-S5P were significantly enriched at 5´end upon GFP-TFIISmut expression relatively to respective controls (Figure 23 A, top panels). RNAPII-S2P and RNAPII-S5P occupancy was also enriched at the 3' end of At3g02260 in iGFP-TFIISmut line exposed to β-estradiol relatively to respective controls, however the differences in RNAPII-S5P level were not statistically significant (Figure 23 A, bottom panels).

Both RNAPII-S2P and RNAPII-S5P were also significantly enriched over At1g48090 at the 5'end upon GFP-TFIISmut expression (Figure 23 B, top panels). The differences in active RNAPII level at the 3'end of At1g48090 were not statistically significant between analysed samples, althogh RNAPII-Ser2P and RNAPII-Ser5P mean occupancy was highest in iGFP-TFIISmut line exposed to β-estradiol (Figure 23 B, bottom panels).

As shown in the previous chapter the association of Histone H3 at analysed loci was not significantly altered between analysed samples (Supplementary Figure S 4) neither specific enrichment was observed when analysing DNA copurified without any antibody (Figure 22 B-C).

Obtained results revealed strong accumulation of active RNAPII at transcribed regions upon TFIIS mutation. Therefore, it has been decided to subject DNA immunoprecipitated with α-CTD-S2P and α-CTD-S5P to Illumina Deep Sequencing to shed light on active RNAPII behaviour upon TFIISmut expression genome-wide. Based on the preliminary ChIP-qPCR results as well as RNA-seq data it has been decided to further study changes in active RNAPII occupancy within iGFP-TFIISmut line exposed to 24h β-estradiol induction relatively to mock treatment.

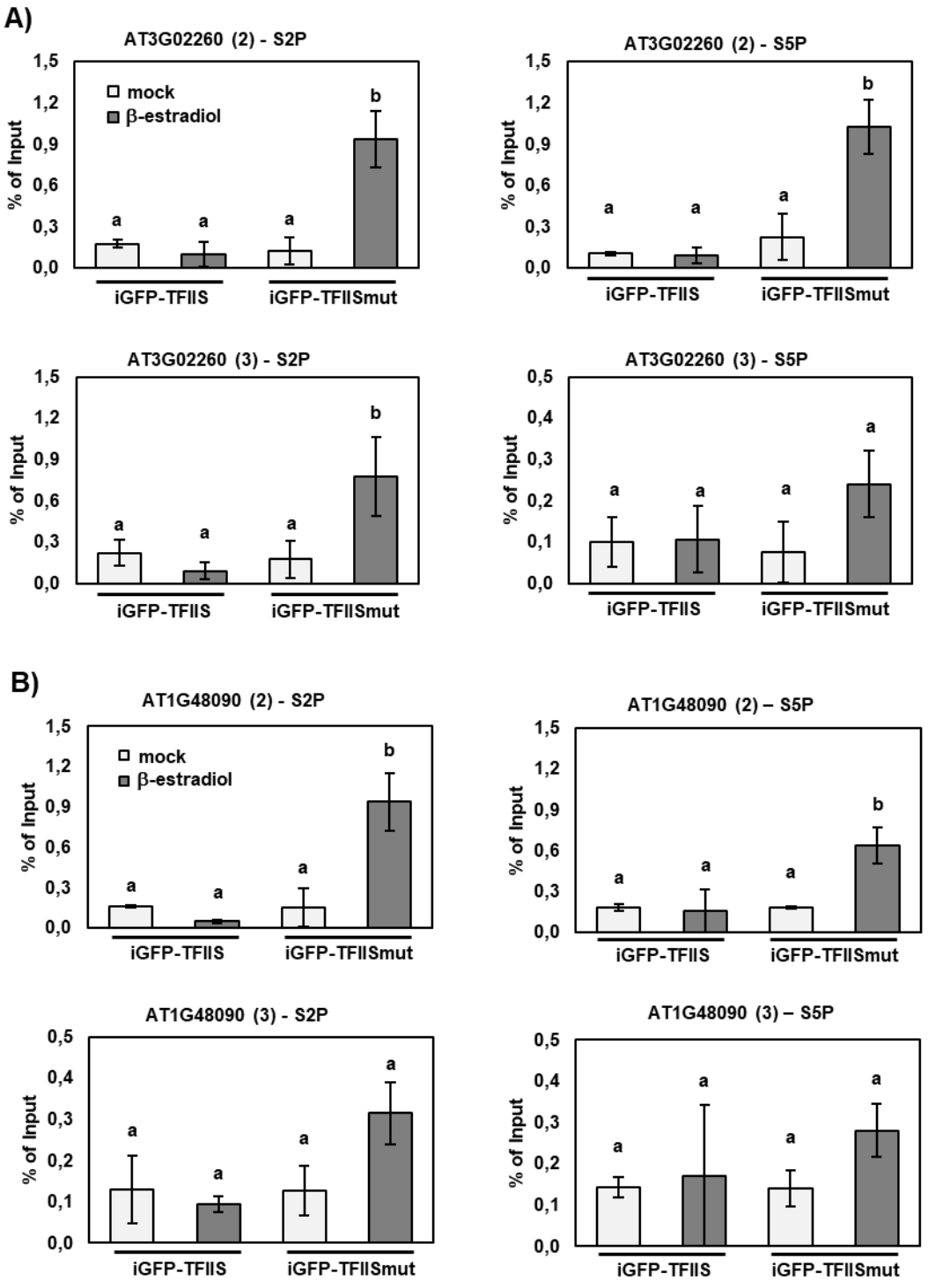


Figure 23. TFIIS mutation results in elevated level of active RNAPII over transcriptionally active regions.
ChIP analyses of At3g02260 (A) and At1g48090 (B) using antibodies against RNAPII-CTD phosphorylated at Ser2 (RNAPII-S2P; "S2P") and Ser5 (RNAPII-S5P; "S5P") position. For ChIP experiments percentage of Input was determined by qPCR and statistically analysed between samples using one-way ANOVA. Error bars indicate SD of at least two biological and two technical replicates. The letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05).

2.3.2.3 Changes in active RNAPII occupancy upon TFIISmut expression

TFIISmut expression leads to broad redistribution and accumulation of active RNAPII

Remaining DNA material immunoprecipitated isolated from three biological replicates of β-estradiol or mock treated iGFP-TFIISmut was sent for IIumina Deep Sequencing. DNA libraries were created with DNA SMART ChIP-Seq kit (Clontech) and sequenced on Illumina HiSeq 2500 by the Lausanne Genomic Technologies Facility (University of Lausanne). The bioinformatical analysis of ChIP-seq data was performed by Dr. Jules Deforges (University of Lausanne). Obtained reads were aligned onto the *Arabidopsis* reference genome assembly (TAIR10). Although the reads coverage after mapping was rather low (~ 0,3 - 4,3 mln reads per sample, ~ 4 - 25% reads mapped). The principal component analysis (PCA) of individual replicates revealed distinct clustering of β-estradiol vs mock induced samples (Supplementary Figure S 5). Additionally, reads were specifically enriched over transcribed relatively to transcriptionally inactive regions indicating good technical quality among individual replicates (Supplementary Figure S 6 and 7).

Profiles of active RNAPII occupancy genome-wide were further created using DeepTools suit (see "Methods"). RNAPII occupancy was visualised around transcription start site (TSS [-2000;2000]), transcription termination site (TES [-2000;2000]) as well as with the relative distribution over transcribed loci (TSS-TES). The analysis of RNAPII-S2P distribution in iGFP-TFIISmut upon mock induction revealed its accumulation immediately downstream TSS with a clear tendency for being enriched towards 3'end (Figure 24 A-B; left panel, black line). This finding is in accordance with previously shown RNAPII-S2P accumulation in Col-0 at individual loci (Ding et al., 2011; Dürr et al., 2014) as well as with genome-wide RNAPII-S2P distribution in other organisms (Hajheidari et al., 2013). Additionally, a sharp accumulation of RNAPII-S2P could be observed around TES (Figure 24 B-C; left panel, black line) similarly to previously observed for transcriptionally engaged RNAPII in *Arabidopsis* (Hetzel et al., 2016).

The analysis of RNAPII-S5P distribution revealed distinct profile with immediate enrichment downstream TSS (Figure 24 A-B; right panel, black line) as well as clear accumulation around TES (Figure 24 B-C; right panel, black line). Surprisingly, RNAPII-S5P was also enriched towards the 3'end, however to much lesser degree then RNAPII-S2P. In contrast to findings in this study, RNAPII-S5P accumulation near TSS is followed by a clear decrease towards the 3'end in yeast (Lidschreiber et al., 2013) and mammals (Nojima et al., 2015). This contradicting observation could reflect distinct genome-wide distribution of total RNAPII characteristic for *Arabidopsis* (Hetzel et al., 2016; Zhang et al., 2015c) or be accounted for altered RNAPII-S5P distribution in *tfIIs-1* background.

The analysis of RNAPII-S2P and RNAPII-S5P occupancy in iGFP-TFIISmut upon β-estradiol treatment revealed very broad changes in their genome-wide profiles relatively to mock induction (Figure 24 A-C). In the presence of TFIISmut, RNAPII-S2P distribution was clearly shifted towards the 5'end with a sharp peak ~ 150 bp downstream TSS followed by gradual decrease towards the 3'end (Figure 24 A-B, left panel, red line). RNAPII-S2P occupancy around TES was decreased although the distribution seemed unaffected (Figure 24 B-C, left panel, red line).

Mutation in TFIIS also strongly affected RNAPII-S5P occupancy leading to its accumulation near TSS (Figure 24 A-B, right panel, red line) although observed redistribution was not as striking as for RNAPII-S2P. A gradual decrease in RNAPII-S5P occupancy towards TES could be observed following ~ 750 bz region downstream TSS. RNAPII-S5P level upon GFP-TFIISmut expression was clearly increased over transcribed units except the regions very proximal to TES. Similarly to RNAPII-S2P, the distribution of RNAPII-S5P around TES was not affected but the occupancy was much lower (Figure 24 B-C, right panel, red line).

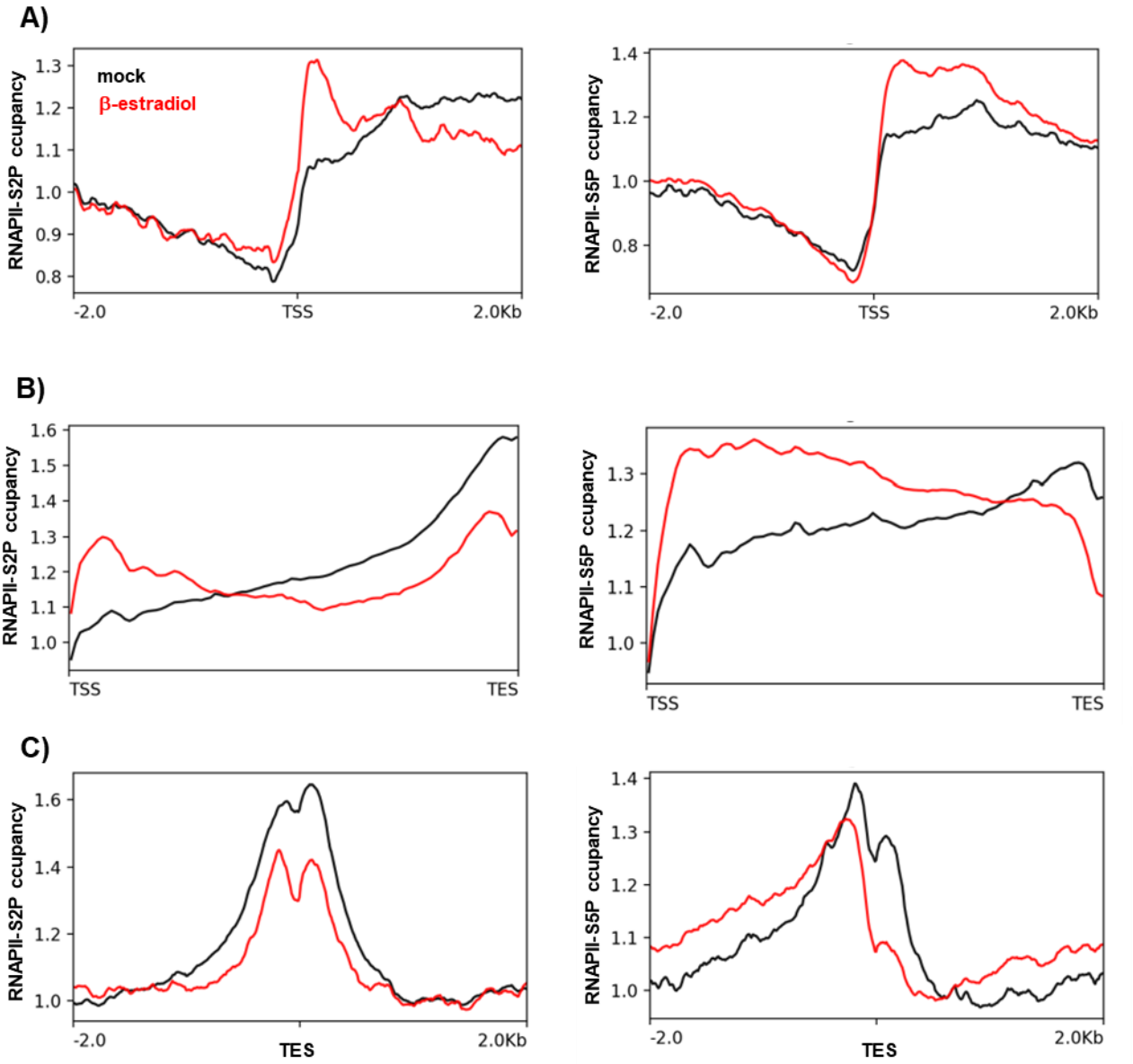


Figure 24. Genome-wide occupancy profiling of active RNAPII using ChIP-seq.
(A) Gene-averaged profiles for RNAPII-S2P and RNAPII-S5P around TSS (A), gene bodies (B) and TES (C) in iGFP-TFIISmut line exposed to β-estradiol (red line) or mock induction (black line). Occupancies and relative position within a gene are given on the y and y axes, respectively. The number of loci combining profiles of RNAPII-S2P = 33574, 33594, 33574 and of RNAPII-S5P = 33560, 33590, 33564 for (A), (B) and (C), respectively. Figures were created by Dr. Jules Deforges.

Changes in RNAPII occupancy are correlated with gene expression

The level of total RNAPII was shown to be correlated with gene expression in *Arabidopsis* (Zhang et al., 2015c). Thus, the changes in active RNAPII occupancy were further analysed in the context of transcriptome rearrangement. By looking at the single gene profiles it could be seen that either decrease or increase in active RNAPII occupancy was reflected by the changes in expression level (Figure 25 A-B).

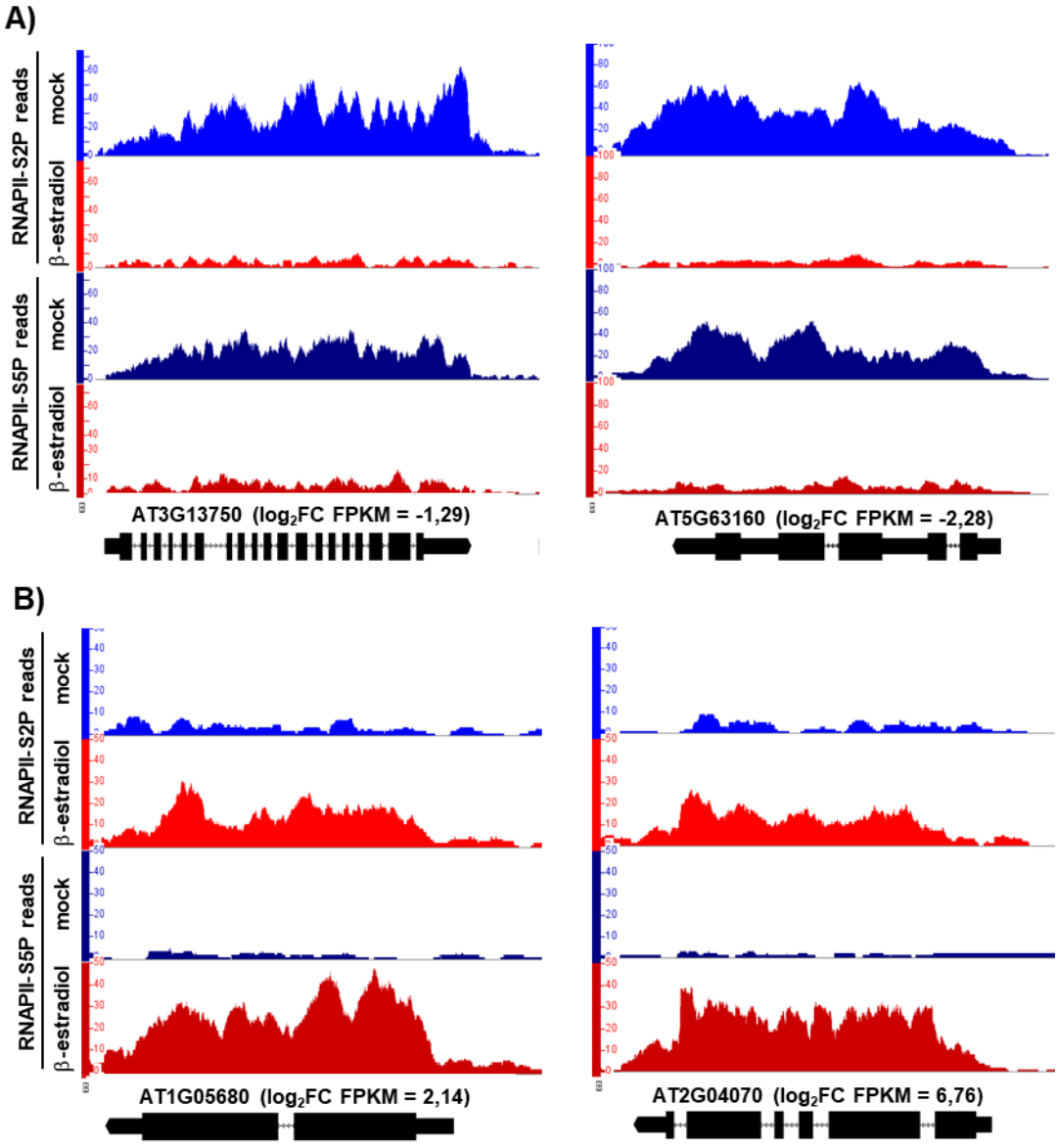


Figure 25. The correlation between active RNAPII occupancy and expression level changes at single genes. ChIP-seq tracks showing RNAPII-S2P and RNAPII-S5P reads over significantly down- (A) and upregulated (B) genes upon TFIISmut expression. Plots were generated in Integrated Genome Browser (IGB) using representative biological replicates. Gene model = thin black bars: UTRs; thick black bars: exons; black line: introns. log_2FC FPKM values were calculated based on the RNA-seq.

The correlation observed at single genes was further evaluated genome-wide. As a result, there was a weak correlation identified between the changes in both RNAPII-S2P and RNAPII-S5P occupancy and transcriptomic changes (Pearson correlation coefficient (r) = 0,37

and 0,38, respectively). In proposed linear regression model the coefficient of determination (R^2) was low (0,17 and 0,15 for RNAPII-S2P and RNAPII-S5P, respectively) indicating that the changes in either RNAPII-S2P or RNAPII-S5P occupancy were weakly related to transcriptomic changes.

These results suggest that active RNAPII redistribution are largely independent from transcriptome rearrangement in line with ChIP-qPCR (Figure 23) and similarly to recently reported in yeast (Vinayachandran et al., 2018). Intriguingly, the changes in RNAPII-S2P and RNAPII-S5P occupancy upon TFIISmut expression correlated well with each other despite rather distinct profiles genome-wide (Supplementary Figure S 8 A).

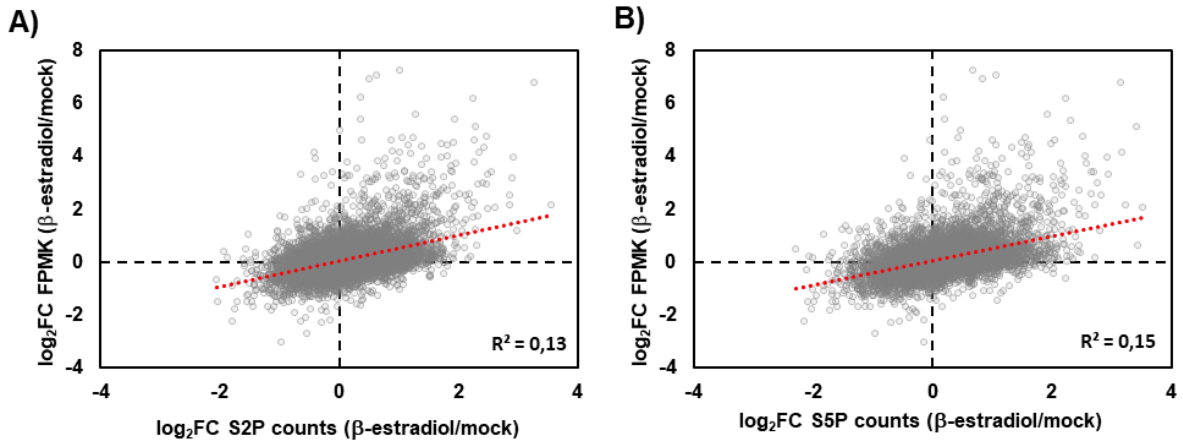


Figure 26. The correlation between active RNAPII occupancy and expression level changes genome-wide. Scatter plot of changes in RNAPII-S2P (A) and RNAPII-S5P (B) enrichment as well as FPKM fold change upon GFP-TFIISmut expression. All transcriptionally active genes comprise the plot (FPKM ≥ 5, n = 11723). Dotted red line reflects the best-fit linear regression with a positive slope. R^2 = coefficient of determination.

TFIISmut expression leads to promotor proximal enrichment of active RNAPII (PPEP)

A very clear consequence of GFP-TFIISmut expression was the accumulation of active RNAPII immediately downstream TSS (Figure 24 A-B). Particularly for RNAPII-S2P it was accompanied by subsequent decrease in occupancy over gene bodies resulting in clear redistribution towards the 5' end (Figure 24 B). In view of that, the promoter proximal enrichment of active RNAPII (PPEP) was calculated for individual genes in RNAPII-S2P (PPEP-S2P) and RNAPII-S5P (PPEP-S5P) context in order to quantitatively examine active RNAPII redistribution upon GFP-TFIISmut expression. For PPEP calculation the counts within TSS proximal region [0; 500] were compared relatively to the counts within a gene body (+500 bp to gene end) and normalised to the gene length (see "Methods"). TSS proximal region was chosen based on the position of RNAPII-S2P and RNAPII-S5P peaks near TSS upon GFP-TFIISmut expression (Figure 24).

PPEP-S2P and PPEP-S5P were successfully calculated for ~17000 genes with significantly increased PPEP upon GFP-TFIISmut expression for 731 and 659 genes in RNAPII-S2P and RNAPII-S5P context, respectively (z-score < -2). At the same time no genes with significantly decreased PPEP were identified for RNAPII-S2P and only 2 genes had decreased PPEP in RNAPII-S5P context (z-score > 2). This strong tendency for PPEP increase upon GFP-TFIISmut expression correlates with the changes in RNAPII-S2P and RNAPII-S5P profiles upon β-estradiol induction (Figure 24 A-B). Two representative genes with increased PPEP upon GFP-TFIISmut expression (AT5G11090 and AT2G20562) were further visulised by Integrated Genome Browser (IGB) showing strong redistribution of acive RNAPII towards the 5'end in accordance with calculated PPEP (Figure 27).

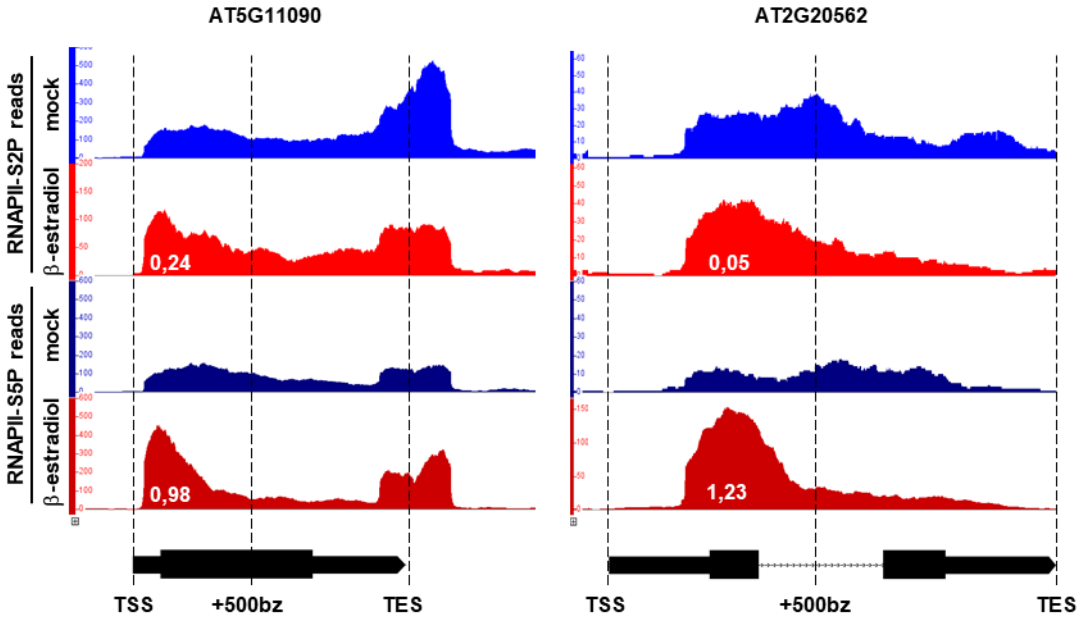


Figure 27. Active RNAPII occupancy enrichment towards TSS upon TFIISmut expression at single genes.
ChIP-seq tracks showing RNAPII-S2P and RNAPII-S5P reads over AT5G11090 (A) and AT2G20562 (B) genes with increased PPEP upon GFP-TFIISmut expression (β-estradiol vs mock). Plots were generated in Integrated Genome Browser (IGB) using representative biological replicates. Gene model = thin black bars: UTRs; thick black bars: exons; black line: introns. Numbers indicate $\log_2$ fold change PPEP (β-estradiol vs mock).

The correlation between PPEP-S2P and PPEP-S5P increases upon TFIISmut expression

PPEP-S2P correlation with PPEP-S5P was further examined to identify any general RNAPII behaviour in response to TFIIS mutation regardless CTD phosphorylation status. There was a well genome-wide correlation identified between differentially phosphorylated RNAPII in terms of PPEP value upon mock induction (Pearson correlation coefficient (r) = 0,76). Calculated correlation was also clearly seen after visualising both datasets with a scatter plot (Figure 28 A). The correlation between PPEP-S2P and PPEP-S5P was further increased upon TFIISmut expression (r = 0,83; Figure 28 B). Accordingly, there was a weak correlation between the changes in PPEP-S2P and PPEP-S5P upon GFP-TFIISmut expression (r = 0,20). However, the changes in PPEP-S5P were a poor predictor of the changes in PPEP-S2P upon GFP-TFIISmut expression (coefficient of determination (R^2) = 0,04) suggesting rather independent

mechanisms of PPEP establishment for differentially phosphorylated RNAPII at most genes (Supplementary Figure S 8 B).

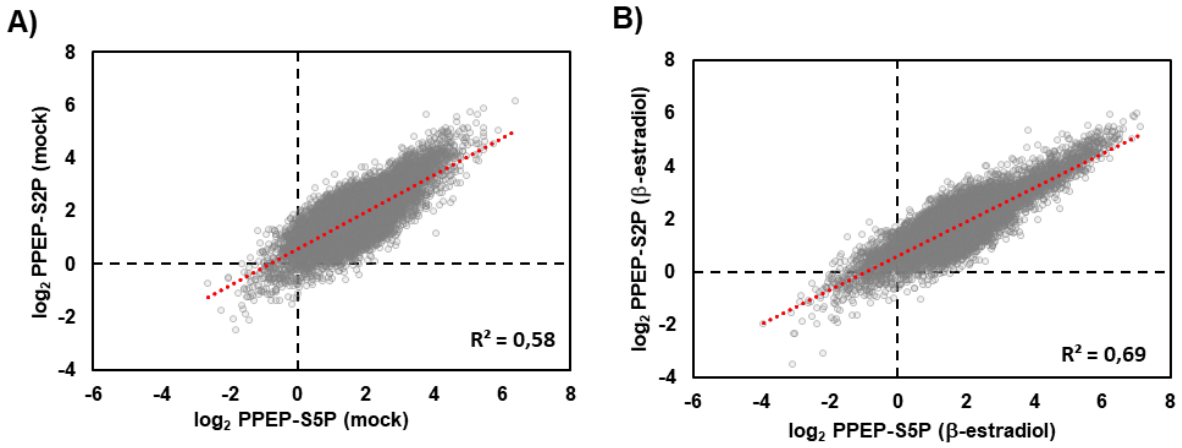


Figure 28. PPEP-S2P correlation with PPEP-S5P increases upon TFIISmut expression.
Pair-wise scatter plot analysis of PPEP-S2P vs PPEP-S5P in iGFP-TFIISmut upon mock (A) and β-extradiol (B). Dotted red line reflects the best-fit linear regression with a positive slope (n=16482). R^2 = coefficient of determination.

PPEP-S2P level may prime its subsequence enrichment upon TFIISmut expression

It has been attempted to determine whether active RNAPII establishment prior to TFIISmut expression could influence further RNAPII accumulation near TSS. To this end, PPEP-S2P and PPEP-S5P determined upon mock induction were compered to their respective PPEP values upon β-estradiol induction genome-wide. Analysed datasets were visualised by a dispersion plot in order to determine the slope of regression line. As seen on the Figure 29, the relationship PPEP-S5P datasets (β-estradiol vs mock) was almost perfectly linear (a = 0,99), whereas clear tendency for increased PPEP values in the presence of TFIISmut was observed

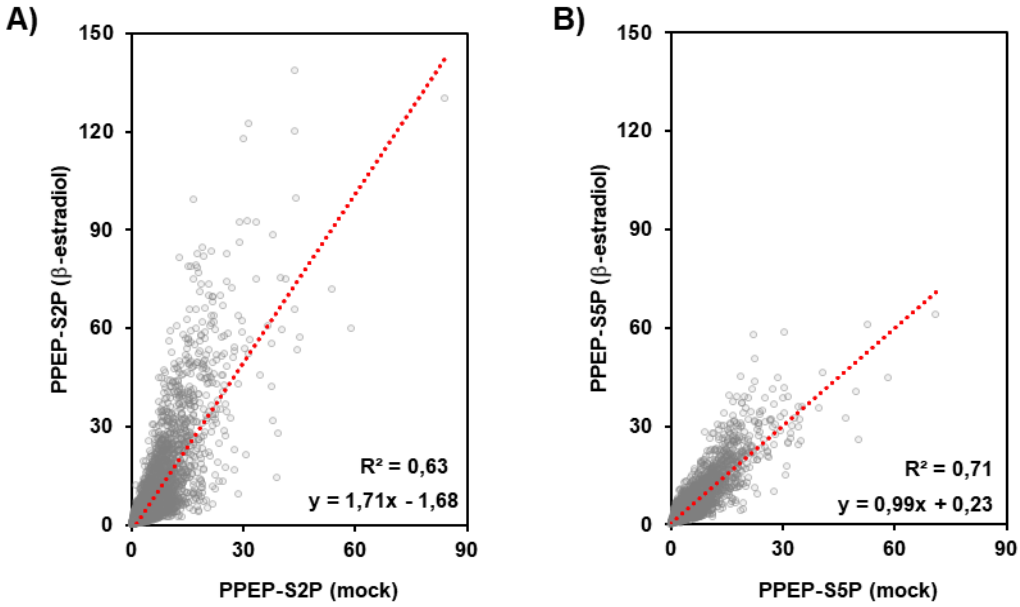


Figure 29. PPEP-S2P may prime subsequence enrichment of RNAPII-S2P upon TFIISmut expression
Dispersion plots showing the correlation between PPEP establishment before (mock) and after (β-estradiol) TFIISmut expression for RNAPII-S2P (A) and RNAPII-S5P (B). Dotted red line indicates the best-fit linear regression model. (n=16482). R^2 = coefficient of determination.

in the context of PPEP-S2P (a = 1,71). It may suggest that the establishment of PPEP-S2P could prime subsequence enrichment of RNAPII-S2P upon TFIISmut expression at most genes (R^2 = 0,63), unlike RNAPII-S5P related PPEP.

PPEP increases only at actively transcribed genes upon TFIISmut expression

Changes in PPEP triggered by TFIISmut were further evaluated in the context of expression level. All genes with determined PPEP were clustered into subgroups regarding their expression level under mock induction according to RNA-seq (chapter 2.3.1). Based on the $\log_{10}$ FPKM value, genes were grouped with high (≥ 2,5), medium (2,5 - 1,3), low (1,3 - 0,7) or no expression (< 0,7). The changes in PPEP-S2P and PPEP-S5P within these subgroups were further evaluated upon GFP-TFIISmut expression (β-estradiol vs mock) and visualised using a whisker-box plots. As seen on the Figure 30 both PPEP-S2P and PPEP-S5P increased significantly upon in the presence of mutated TFIIS among highly, medium and lowly transcribed genes but not in the group of transcriptionally inactive genes (Figure 30 A). The changes in PPEP were the highest among highly and moderately expressed genes for PPEP-S2P (16,5% and 13,4% increase, respectively) and PPEP-S5P (6,5% and 6,9% increase, respectively). Among lowly expressed genes the average PPEP increase was less profound (10,9% and 3,7% for PPEP-S2P and PPEP-S5P, respectively). Additionally, PPEP establishment

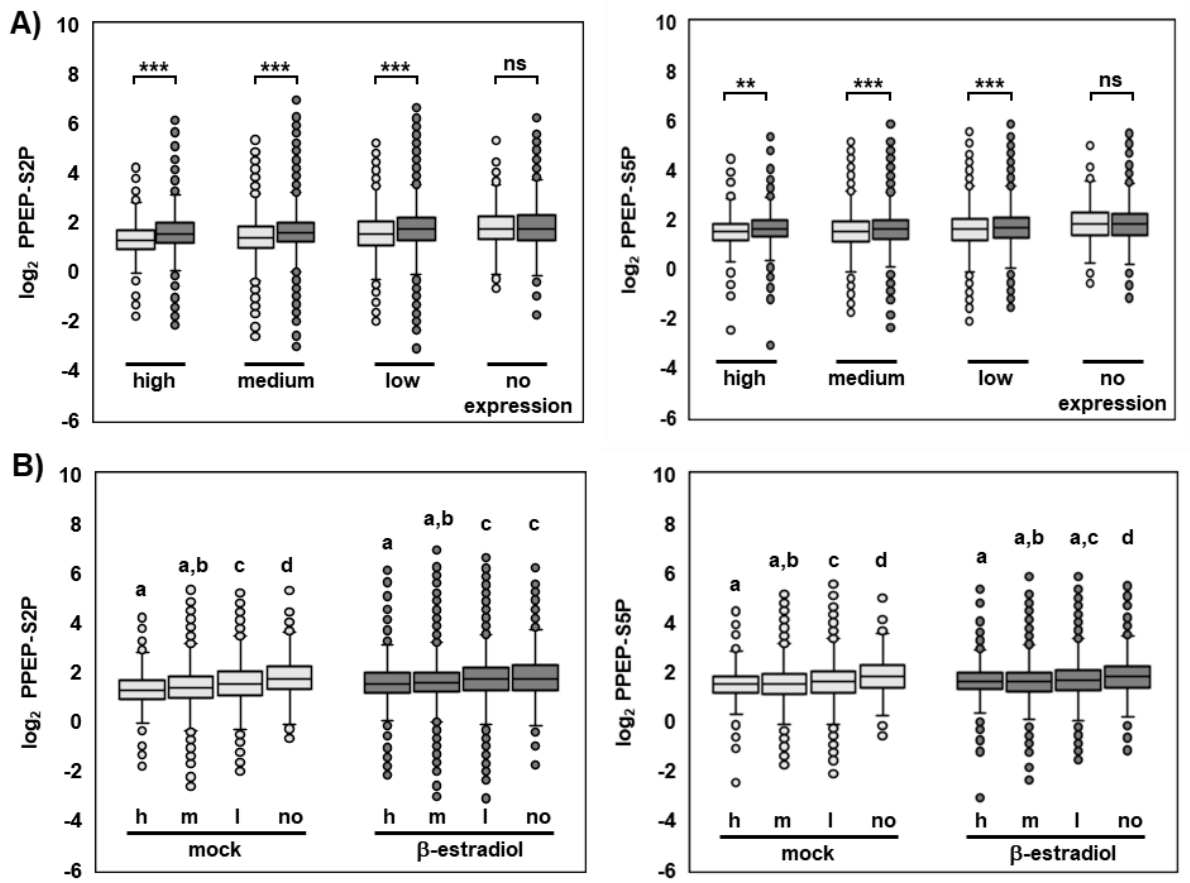


Figure 30. Changes in PPEP depending on the expression level.
Changes in PPEP upon GFP-TFIISmut expression were examined among genes with no (508), low (5707), medium (5802) or high (430) expression based on FPKM reads in mock treated iGFP-TFIISmut and visualised by whisker-box plots (light grey = mock; dark grey = b-estradiol) The significance was tested by Student's T-Test (A; ** p-value < 0.01, *** p-value < 0.01) or using one-way ANOVA (B). The letters above the bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05). h = high, l = low, m = medium, no = no expression.

showed a tendency to decrease with higher expression level under mock induction whereas these differences were less obvious upon β-estradiol induction (Figure 30 B).

PPEP has a non-linear effect on gene expression

PPEP dependency on expression level observed in the previous chapter was further examined in detail by analysing Pearson correlation coefficient between PPEP establishment and expression level. Similarly to shown above, PPEP level was associated with lower expression upon mock induction although the linear correlation was weak for both PPEP-S2P and PPEP-S5P (r = 0,21 and r = 0,17, respectively). Observed correlations were further decreased upon β-estradiol induction (r = 0,087 and r = 0,090, respectively). These findings were further evaluated by looking at the global distribution of PPEP values in the context of genes expression level. In line with observed weak correlation, PPEP distribution was largely independent from expression level (Figure 31 A-B), being a very weak predictor of transcriptomic changes especially upon TFIISmut expression (R^2 values on Figure 31 A-B). PPEP-S5P distribution did not change much upon TFIISmut expression relatively to mock induction, while PPEP-S2P distribution could be characterised by more profound changes extending its PPEP maxima horizontally in both directions, especially in the group of moderately expressed genes (Figure 31 A-B).

Next, the changes in PPEP and genes expression level upon TFIISmut expression were analysed comparatively revealing no obvious tendency neither among DEG nor genome-wide (Figure 31 C). The most profound changes in PPEP were associated with only minor effects on transcript level, while highly missregulated genes were rather randomly distributed in terms of PPEP modifications (Figure 31 C).

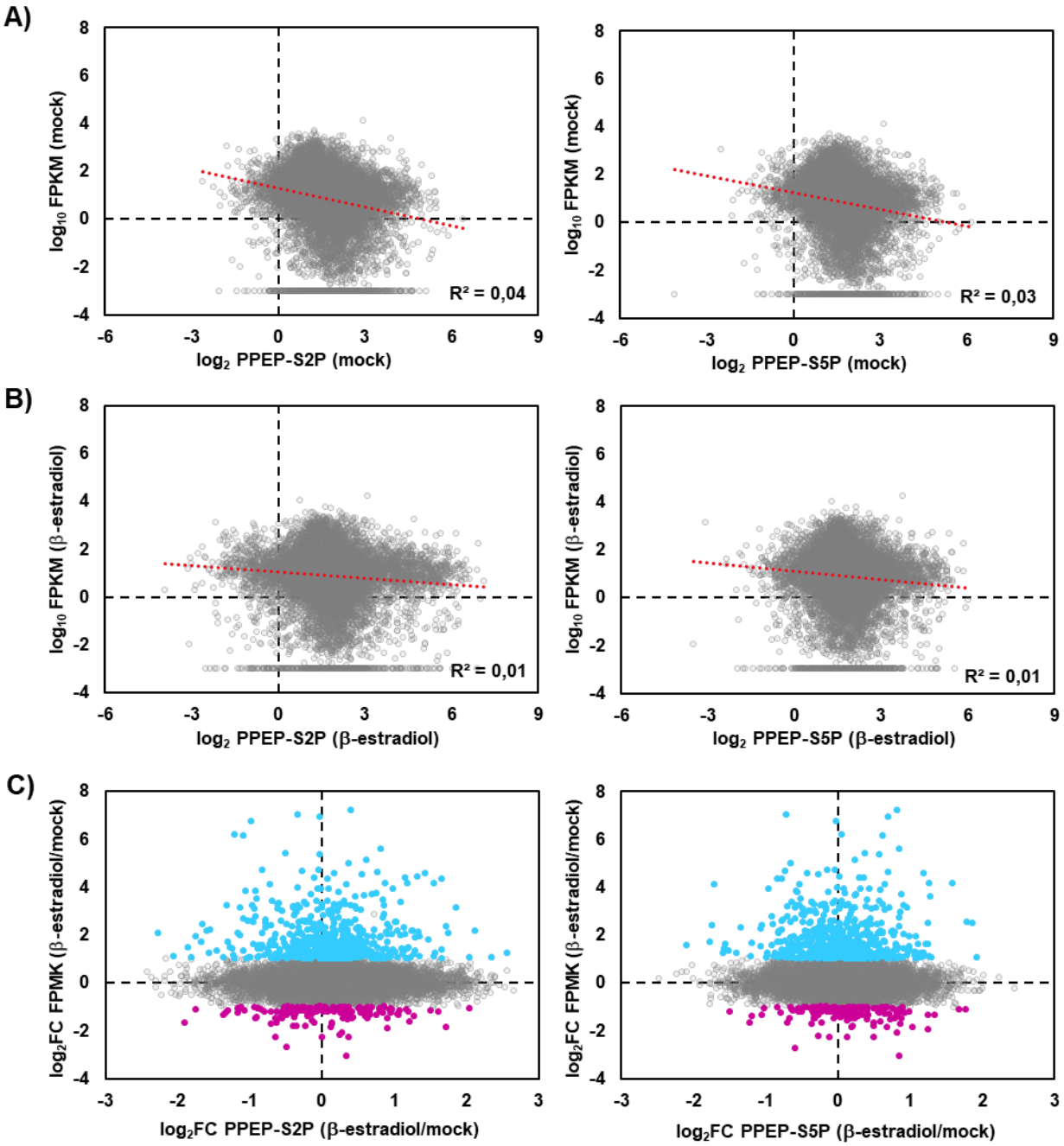


Figure 31. PPEP has a non-linear effect on gene expression.
(A-B) Scatter plot for PPEP and expression level before (A) and after GFP-TFIISmut expression (B). Dotted red line reflects the best-fit linear regression with a negative slope. All genes with measurable PPEP combines the plots (n = 16483). (C) Pair-wise scatter plot analysis of PPEP vs FPMK fold changes upon β-estradiol induction relatively to mock. DEGs are depicted in cyan (upregulated) or magenta (downregulated). All expressed genes with measurable PPEP combine the plots (n=11723). R^2 = coefficient of determination.

PPEP-S5P increases upon TFIISmut expression among downregulated genes

Those PPEP-associated observations were further evaluated quantitatively within the group of DEG. The average PPEP fold change (β-estradiol vs mock) was calculated among downregulated DEG and compared relatively to the group of significantly upregulated genes. PPEP-S2P average fold change was not altered when comparing up- vs downregulated DEG although the distribution of PPEP-S2P fold change was broader in the group of downregulated genes (Figure 32). In contrary, PPEP-S5P average fold change was significantly higher in the group of downregulated DEG upon TFIISmut expression (Figure 32).

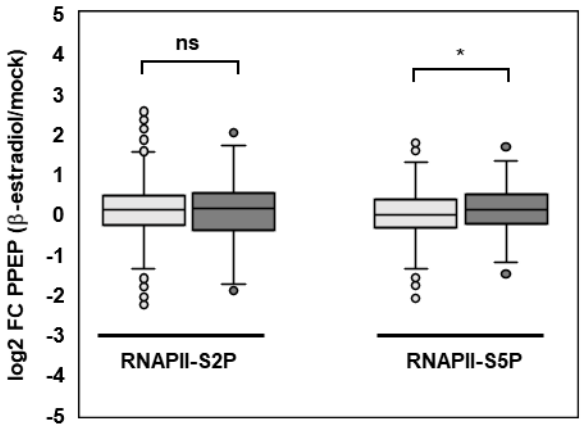


Figure 32. PPEP-S5P increases significantly among DEGs downregulated upon GFP-TFIISmut expression.
Changes in PPEP upon TFIISmut expression (log2FC) were examined among DEG downregulated (n = 162) in comparison to upregulated DEG (n = 584). The significance was tested by Student's T-Test: * $P < 0.05$).

TFIIS-dependent establishment of PPEP may influence certain biological processes

PPEP establishment upon GFP-TFIISmut expression appeared highly heterogenous in terms of gene expression (chapter 2.3.2.3). Consequently, genes downregulation was presumably triggered by compromised PPEP regulation only for a fraction of genes. It has been hypothesised that some of these genes could be involved in the regulation of crucial biological processes. Thus, GO enrichment analysis was performed to identify compromised biological processes shared between downregulated and PPEP-responsive genes upon TFIISmut expression. GO terms enriched among downregulated DEG upon GFP-TFIISmut expression were already characterised in this study (Supplementary Table 5), however here more relaxed analysis was performed to find a better correlation with PPEP-related response (Zheng and Wang, 2008) (see "Methods"). The GO analysis was next performed among PPEP-responsive genes defined as having increased PPEP upon GFP-TFIISmut expression (log_2FC PPEP > 0; β-estradiol vs mock), resulting in clear PPEP establishment upon β-estradiol induction (log_2PPEP > 2). Several GO terms were identified as enriched among both downregulated DEG and PPEP-responsive genes upon GFP-TFIISmut expression including GO terms related to lipid and macromolecule localisation as well as protein complex biogenesis and assembly (Table 6). Additionally, "photosynthesis" and "homeostatic processes" GO terms were mutually overrepresented among DEG and PPEP-S2P responsive genes, whereas response to osmotic and salt stress were identified in the PPEP-S5P context as well as among DEG (Table 6). These

findings imply that perturbation of some biological processes in the presence of mutated TFIIS may be a consequence of RNAPII arrest triggered by TFIISmut.

Table 6. The correlation between overrepresented GO terms among downregulated DEGs and PPEP-responsive genes.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) with hypergeometic testing and non-adjusted multitesting (AgriGO) among PPEP-responsive genes and DEGs downregulated upon GFP-TFIISmut expression. All significantly GO terms enriched among downregulated DEG (p-value < 0.01) and frequency < 10% are show in the table. Respective p-values from GO terms analysis for PPEP-responsive genes are shown in the table. All GO terms overrepresented among PPEP-responsive genes are shown in Supplementary Table 10 and Supplementary Table 11.

		p-value < 0.01		
GO term ID	**description**	**DEG**	**PPEP-S2P**	**PPEP-S5P**
GO:0010876	lipid localization	1,30E-12	5,60E-04	1,10E-04
GO:0070271	protein complex biogenesis	3,20E-03	4,80E-06	8,50E-06
GO:0006461	protein complex assembly	3,20E-03	4,80E-06	8,50E-06
GO:0033036	macromolecule localization	6,70E-03	3,40E-06	1,20E-04
GO:0015979	photosynthesis	4,80E-09	9,80E-04	ns
GO:0042592	homeostatic process	7,60E-07	9,00E-03	ns
GO:0006970	response to osmotic stress	8,60E-04	ns	9,20E-05
GO:0009651	response to salt stress	1,50E-03	ns	9,30E-05
GO:0009733	response to auxin stimulus	2,10E-09	ns	ns
GO:0045454	cell redox homeostasis	6,50E-09	ns	ns
GO:0019725	cellular homeostasis	9,80E-08	ns	ns
GO:0019684	photosynthesis, light reaction	1,30E-06	ns	ns
GO:0009767	photosynthetic electron transport chain	3,10E-05	ns	ns
GO:0006869	lipid transport	3,20E-05	ns	ns
GO:0006091	generation of precursor metabolites and energy	2,70E-04	ns	ns
GO:0009664	plant-type cell wall organization	3,30E-04	ns	ns
GO:0022900	electron transport chain	5,30E-04	ns	ns
GO:0009739	response to gibberellin stimulus	1,20E-03	ns	ns
GO:0009751	response to salicylic acid stimulus	3,60E-03	ns	ns
GO:0009753	response to jasmonic acid stimulus	5,10E-03	ns	ns

Since PPEP establishment is seemingly independent from transcriptomic changes upon GFP-TFIISmut expression, GO enrichment analysis was further performed regardless gene expression level. To better imitate the putative changes in total RNAPII distribution, RNAPII-S2P and RNAPII-S5P counts were averaged and PPEP-S2P&S5P was calculated upon β-estradiol and mock induction. Only genes with significantly increased PPEP (z-score > 2) upon GFP-TFIISmut expression (β-estradiol vs mock) resulting in PPEP establishment upon β-estradiol induction ($\log_2$ PPEP > 1, β-estradiol) were considered. Performed GO enrichment analysis for those genes revealed their involvement into the establishment of seed dormancy and localisation within the cell (Figure 33). Intriguingly, the regulation of seed dormancy has been previously demonstrated as compromised in the absence of *Arabidopsis* TFIIS (Grasser et al., 2009).

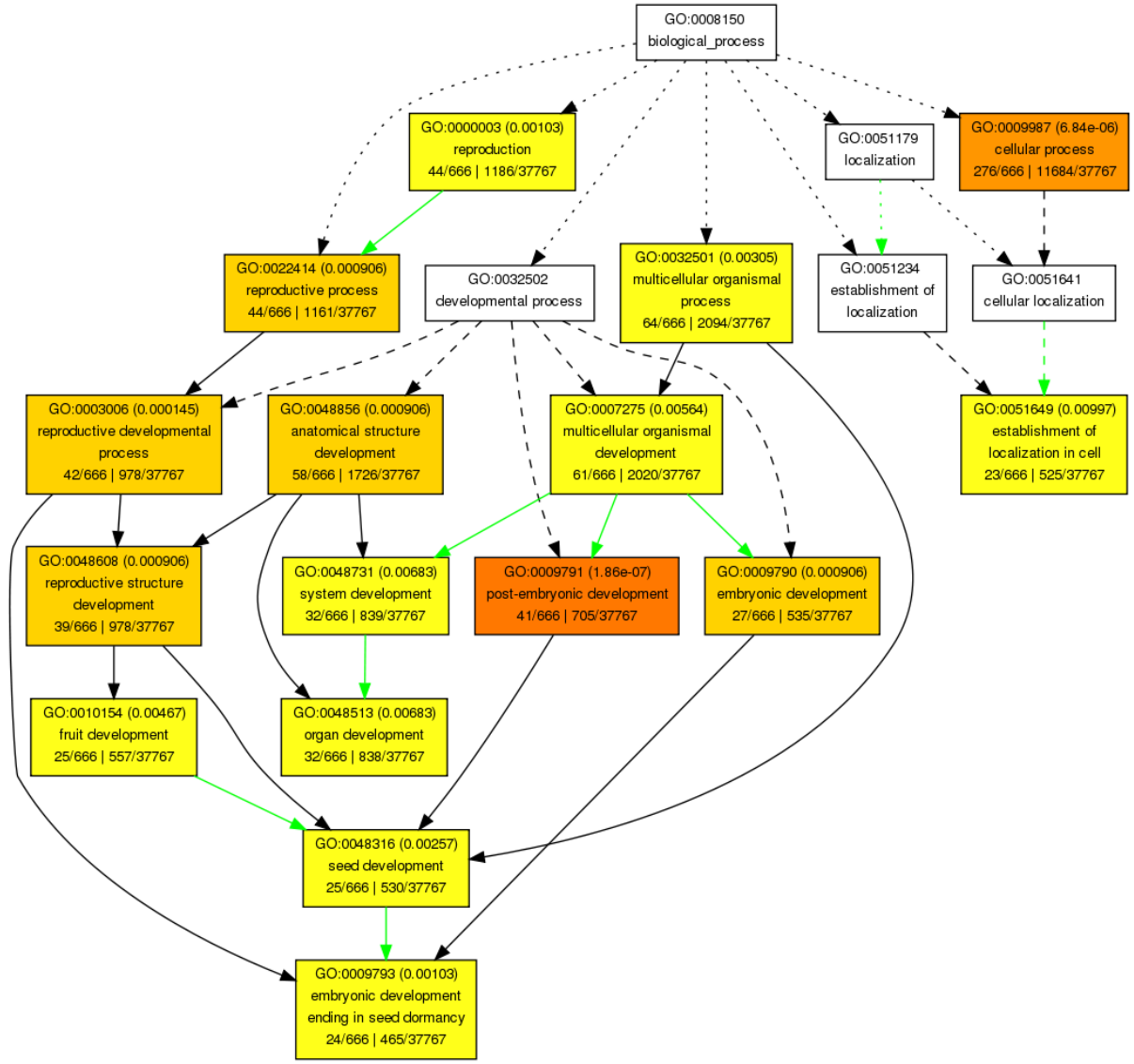


Figure 33. PPEP establishment may involve the regulation of seed dormancy and localisation processes.
The Gene Ontology (GO) analysis was performed and visualised using the single enrichment analysis (SEA) for genes with significantly increased PPEP (merged S2P&S5P; z-score > 2) upon GFP-TFIISmut expression (β-estradiol vs mock) resulting in PPEP establishment upon β-estradiol induction (log_2PPEP > 1, β-estradiol). Overrepresented GO terms with FDR < 0.01 are shown on the graph. All enriched GO terms are listed in Supplementary Table 12.

The occupancy of +1 nucleosome contributes to PPEP establishment

Nucleosome occupancy and local sequence composition near TSS are considered an important determinants of RNAPII enrichment in the promotor proximal region (Kwak and Lis, 2013). Given the role of TFIIS in passaging nucleosomes *in vitro* (Gaykalova et al., 2015; Ishibashi et al., 2014) as well as observed accumulation of active RNAP ~ 150 bz downstream TSS upon β-estradiol induction, the +1 nucleosome positioning was hypothesised to influence active RNAPII accumulation in the presence of TFIISmut. Thus, using the publicly available MNAse-seq data (Li et al., 2014), nucleosome occupancy in Col-0 seedlings was determined and profiled relatively to active RNAPII occupancy in iGFP-TFIISmut upon β-estradiol and mock induction. As seen on the Figure 34 the position of the +1 nucleosome strongly overlapped with the peak of both RNAPII-S2P and RNAPII-S5P upon GFP-TFIISmut expression (Figure 34). Whereas RNAPII-S2P reaches its local maximum slightly upstream the +1 nucleosome (Figure 34, left panel), RNAPII-S5P seems to peak somewhat downstream the +1 nucleosome (Figure 34, right panel). These data suggest that the position of +1 nucleosome may contribute to accumulation of active RNAPII near TSS when nascent RNA cleavage is blocked by TFIISmut.

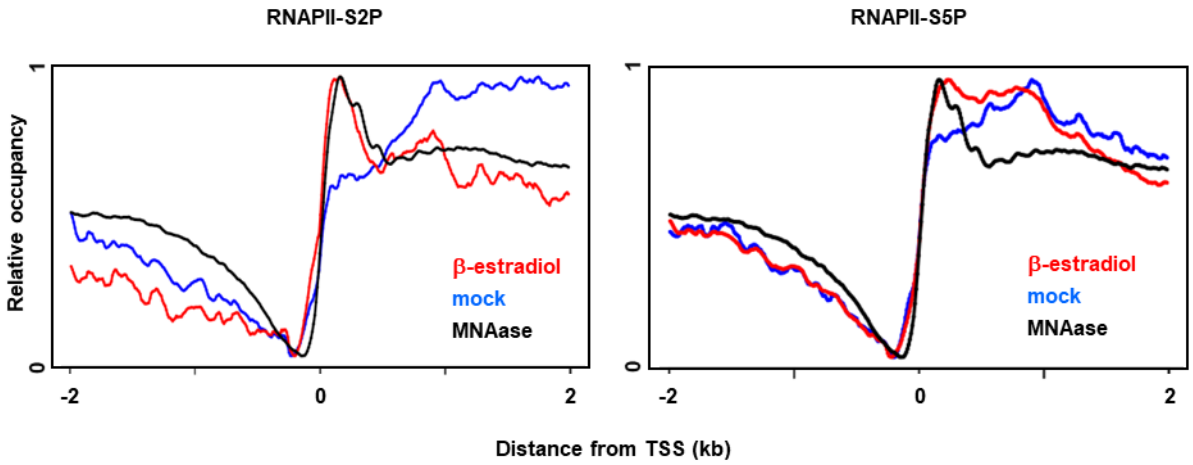


Figure 34. The overlap between active RNAPII occupancy and +1 nucleosome.
Frequency distribution of gene-averaged active RNAPII occupancy before (mock; blue line) and after (β-estradiol, red line) GFP-TFIISmut expression determined by ChIP-seq. Black line represents genome-wide nucleosome occupancy in *Arabidopsis* seedling at 14 DAS based on the MNAse-seq (obtained from Li et al., 2014). Plots are separately scaled to 1 for each data set.

Sequence composition may influence increase in PPEP and RNAPII occupancy

Local sequence composition near TSS has been shown to influence promoter-proximal enrichment of RNAPII in *Drosophila* as well as across diverse mammalian cell types (Gaertner and Zeitlinger, 2014; Gout et al., 2017). Thus, it was attempted to identify DNA sequence motifs enriched in TSS proximal region [-150;150] among genes with significantly increased PPEP upon GFP-TFIISmut expression (z-score > 2). De novo motif discovery was performed with MEME Suite 5.0.2 in the discriminative mode using genes with decreased/unmodified PPEP (z-score < 0) as a background. Following MEME analysis in PPEP-S2P or PPEP-S5P context, various combinations of TC-repeats and simple polyA-repeats were identified as the

most prevalent DNA sequence patterns (Figure 35). Additionally, GA- and GAA-repeats were commonly observed together with some specific sequence motifs, for instance “CGN”, “TGGGC” or “GCCCA” (Figure 35). Taken together, the promoter-proximal sequence among genes with significantly increased PPEP upon GFP-TFIISmut expression seems to be dominated by tandem purines-pyrimidines repeats and simple polyA-repeats.

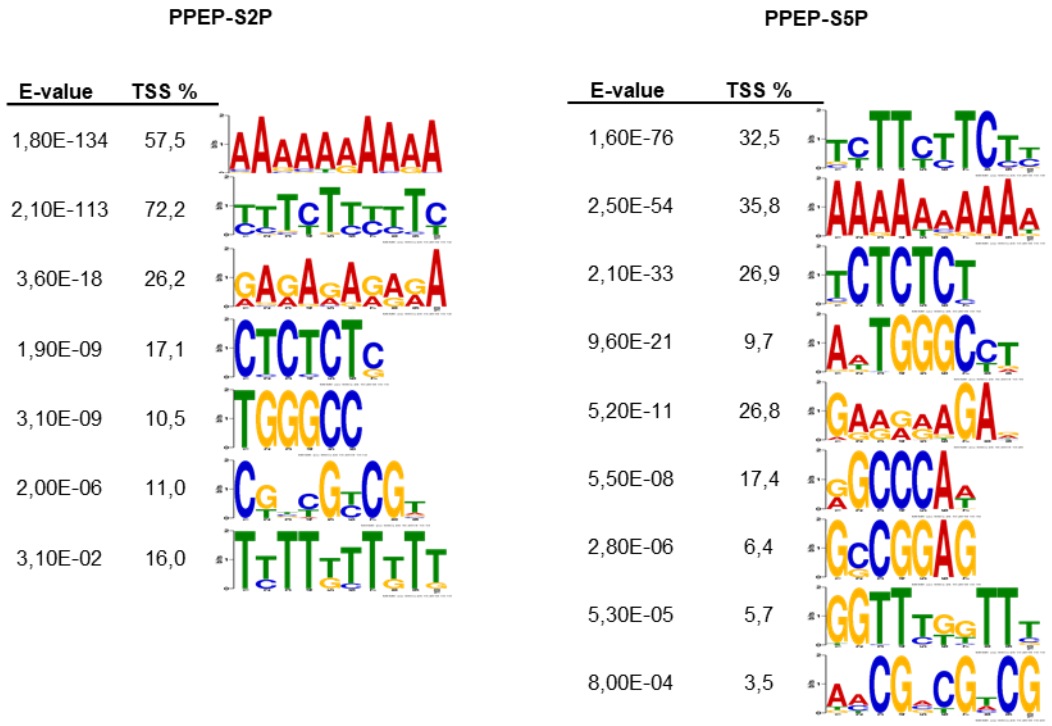


Figure 35. Sequence motives discovered among PPEP-responsive genes.
De novo motif analysis of the proximal-promoter region from −150 to +150 with respect to the TSS using MEME Suite. Significantly enriched motifs (E-value < 0.05) were screen among genes with significanly enriched PPEP (z-score > 2; n =724 and n = 648, for PPEP-S2P and PPEP-S5P, respectively) using genes with decreased/unmodified PPEP (z-score < -1; n = 1597 and n = 2029 PPEP-S2P and PPEP-S5P, respectively) as a background. All significantly enriched motifs (coloured logos) are shown along with the percentage of TSSs containing the motif (TSS %) and the significance level (E-value).

Apart from clear enrichment in PPEP establishment, another striking consequence of GFP-TFIISmut expression was the accumulation of RNAPII-S5P oven transcriptionally active units (Figure 24). RNAPII accumulation, not reflected in increased expression level, has been previously connected with defects in transcript elongation (Dürr et al., 2014; Saunders et al., 2006). Ongoing transcript elongation is also known to be affected by sequence composition (Deighan et al., 2011; Klopf et al., 2018) where certain trinucleotide repeats may impede the progress of RNA polymerase II (McIvor et al., 2010; Morris and Greenleaf, 2000). It was therefore expected that observed RNAPII-S5P accumulation over transcribed units upon GFP-TFIISmut expression could coincidence with the enrichment of certain trinucleotide repeats. To test this hypothesis, genes exhibiting significantly increased RNAPII-S5P occupancy upon GFP-TFIISmut expression (log_2FC RNAPII-S5P > 2) not accompanied with increased expression level (log_2FC FMKP < 0) (n = 368) were analysed in the context of trinucleotides sequence composition. The frequency of each trinucleotide repeat (n = 64) was

compared relatively to their frequency among control genes. Only genes with comparable increased RNAPII-S5P occupancy (log_2 FC RNAPII-S5P > 2) followed with a strong increase in gene expression level (log_2 FC FPKM > 2) (n = 319) were considered as control genes. Trinucleotides frequency was next analysed for each gene with R software using "seqinr" package. As a result, several trinucleotide combinations comprising thymine residue surrounded with any other nucleotides were found significantly enriched among analysed genes relatively to control genes (Table 7, left panel). Intriguingly, enriched trinucleotides resemble to some extend motifs identified by MEME analysis (Figure 35) as well as DNA sequences identified with high transcriptional error rate in yeast upon TFIIS deficiency (James et al., 2017). Additionally, many trinucleotides were found depleted among analysed genes in comparison with control genes (Table 7, right panel).

Table 7. Trinucleotide frequency among genes with significantly increased RNAPII-S5P occupancy.
Trinucleotides frequency was analysed among genes with significantly increased RNAPII-S5P occupancy upon GFP-TFIISmut expression (log_2 FC RNAPII-S5P > 2) not accompanied with increased expression level (log_2 FC FPKM < 0; n = 368) in comparison to control genes (log_2 FC RNAPII-S5P > 2; log_2 FC FPKM > 2; n = 319). Trinucleotides frequency was determined using R software with "DNAstat" package. p-value reflects the outcome of Student's T-test.

Increased frequency			Decreased frequency		
Sequence	**Fold change**	**p-value**	**Sequence**	**Fold change**	**p-value**
CTT	0,15	1,60E-03	AAG	-0,24	1,15E-04
ATT	0,12	9,92E-04	AGA	-0,17	6,68E-03
CTG	0,12	2,66E-02	ATG	-0,10	3,66E-02
CTC	0,05	1,23E-03	ACG	-0,10	9,82E-04
GTT	0,05	2,57E-02	GAG	-0,09	2,01E-06
CCC	0,04	1,04E-02	TGG	-0,07	1,81E-07
GTA	0,03	2,57E-02	CGG	-0,06	2,40E-02
GCC	0,01	1,25E-02	CGT	-0,06	2,92E-02
AGT	0,01	4,28E-03	GAC	-0,05	1,25E-05
			ATA	-0,02	6,79E-03
			ACC	-0,02	1,55E-02
			AGC	-0,01	4,27E-02
			AGG	-0,01	1,13E-05

2.4 Indirect genome-wide consequences of TFIIS mutation

The profound changes in active RNAPII distribution accompanied with broad transcriptomic changes and severe growth defects were observed in this study as a molecular and morphological consequences of mutation within TFIIS acetic loop. Those observations were likely caused by RNA cleavage inhibition in the presence of TFIISmut leading to RNAPII arrest, similarly to observed in yeast (Sigurdsson et al., 2010). Thus, the genome-wide consequences of RNAPII arrest were addressed in the following chapters. Accordingly, TEC mobility, RNAPII degradation and its collision with replication machinery have been studied upon TFIISmut expression.

2.4.1 TFIISmut association lowers TEC mobility

Transcription is a very dynamic process and RNAPII itself was shown to very rapidly travel between various transcriptional states (Steurer et al., 2018; Van Lijsebettens and Grasser, 2014). In the presence of TFIISmut triggering RNAPII arrest, TEC was expected to persist bound to the chromatin with compromised dynamic, similarly to reported in yeast upon α-amanitin treatment (Steurer et al., 2018). Since inducible GFP-TFIIS and GFP-TFIISmut were shown to be part of TEC in *Arabidopsis* (Table 2), TEC mobility was analysed in the context of GFP-TFIISmut association with chromatin in comparison to GFP-TFIIS. Additionally, NRPB1 mobility was studied upon GFP-TFIISmut expression relatively to GFP-TFIIS.

Inducible GFP-TFIIS and GFP-TFIISmut transgenes were introduced into PSB-D *Arabidopsis* cell culture genome by co-cultivation with *Agrobacterium tumefaciens* (Pfab et al., 2017). GFP-TFIIS and GFP-TFIISmut mobility was further analysed within transgenic PSB-D cells by Fluorescence Recovery After Photobleaching (FRAP) using time-lapse confocal microscopy. Transgenic cell cultures were exposed to β-estradiol for 24h prior to the measurements. Detected GFP signal was bleached in a defined region of interest (ROI) in the nucleoplasm (Figure 36). Based on the over-time fluorescence recovery measurements, the mobile fraction and fluorescence recovery time after photobleaching ($t_{½}$) were calculated. As a result, the mobility of mutated TFIIS (72%) was significantly compromised relatively to its wild-type counterpart (98%, Figure 36 B-D). Additionally, GFP-TFIISmut recovery half-time after photobleaching was significantly lower than for GFP-TFIIS ($t_{½}$ = 0,72 s vs $t_{½}$ = 0,26 s, respectively), indicating longer association with chromatin for TEC containing mutated TFIIS (Figure 36 B-D).

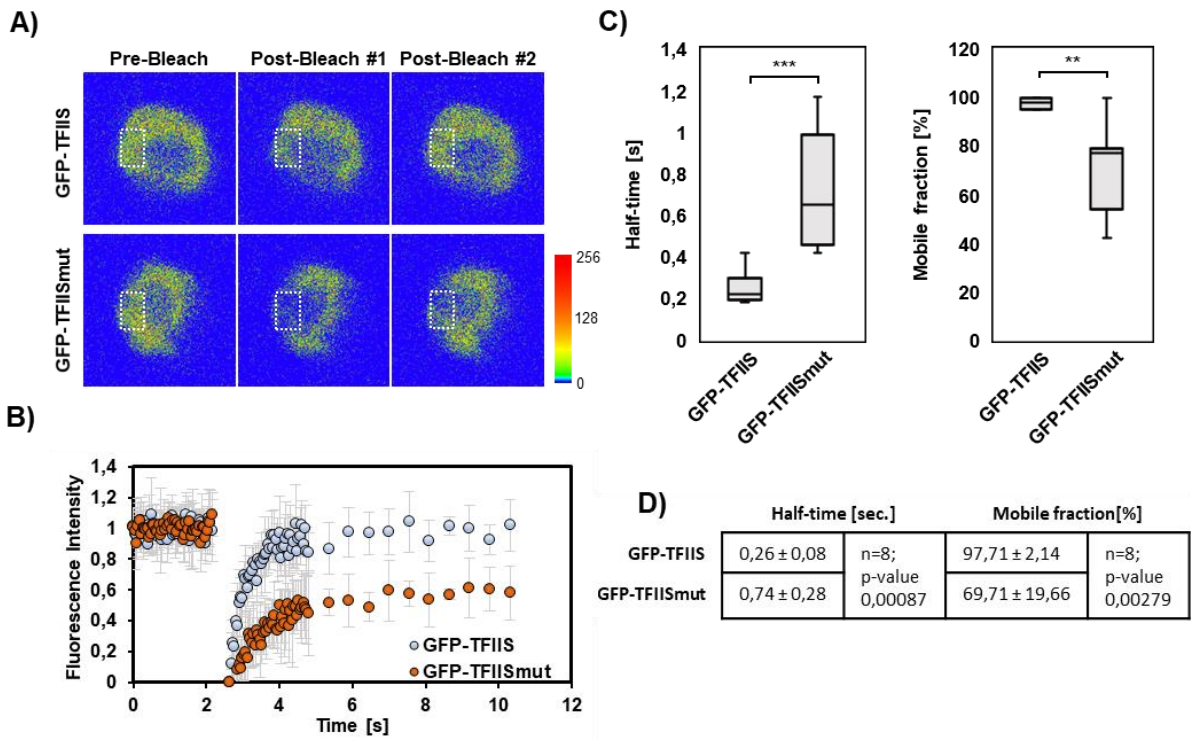
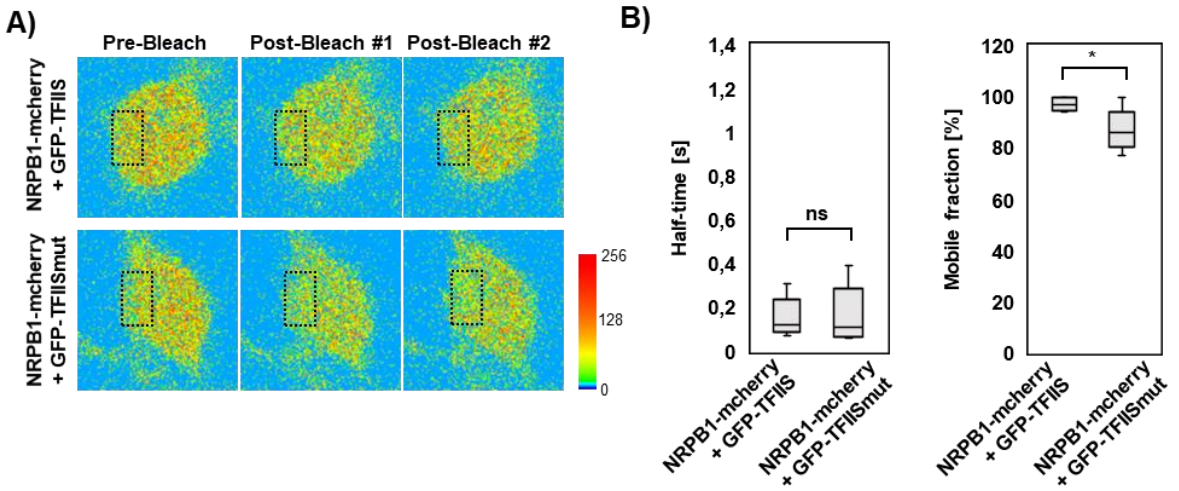


Figure 36. GFP-TFIISmut shows significanly lower mobility than GFP-TFIIS.
(A) Representative pictures of *Arabidopsis* nuclei in transgenic PSB-D cells exposed to 24h β-estradiol induction. Pictures were taken during FRAP experiment performed for GFP-TFIIS and GFP-TFIISmut. The region of interest (ROI; doted white line) were photobleached and the recovery of the GFP fluorescence intensity was measured over-time by confocal microscopy (CLSM). Pre-Bleach indicates the first timepoint of the series (t = 0 s), Post-Bleach#1 the first timepoint after bleaching (t = 2,7 s) and Post-Bleach#2 the last time point of the series (t = 10,3 s). Pseudo-coloured images (modified fire LUT) with respective colour calibration bar are shown. (B) The mean fluorescence recovery curves after full scale normalization for GFP-TFIIS and GFP-TFIISmut are shown with 40x pre-bleach and 50x post-bleach time points. The standard deviations of individual measurements (n=8) at each timepoint are shown as light grey bars. (C-D) Mobile fraction and recovery half-time after photobleaching ($t_{½}$) values were calculated for GFP-TFIIS and GFP-TFIISmut using easyFRAP and visualised using a whisker-box plot (C) or given in the table together with the significance level (D). The significance was tested by Student's T-Test: ** p-value < 0.01, *** p-value < 0.001.

NRPB1 mobility in the presence of TFIISmut was further analysed to confirm that obtained results were consequence of lowered TEC mobility. To this end, NRPB1 driven by CaMV 35S promoter was N-terminally fused with mCherry and transiently co-expressed with inducible GFP-TFIIS or GFP-TFIISmut transgenes in *Nicotiana benthamiana*. Following *Agrobacterium*-mediated leaves infiltration and their exposure to β-estradiol for 24h, co-transformed epidermal cells were identified (Supplementary Figure S 9 A). FRAP measurements were first performed for inducible GFP-TFIIS and GFP-TFIISmut to confirm results obtained in PSB-D. Similarly to observed in the cull culture system, GFP-TFIISmut showed significantly lower mobility (89% vs 99%) and increased recovery time after photobleaching relatively to GFP-TFIIS ($t_{½}$ = 0,55 s vs $t_{½}$ = 0,16 s, respectively) (Supplementary Figure S 9 B-E). Next, NRPB1 mobility was examined in co-transformed epidermal cells using FRAP by studding mCherry recovery after photobleaching. NRPB1 mobile fraction was significantly lowered when co-expressed with GFP-TFIISmut in comparison to co-expression with GFP-TFIIS (87% vs 97%, respectively; Figure 37 B-C), suggesting its immobilisation within arrested TEC. Noteworthy, NRPB1 recovery time after photobleaching was unaffected

regardless the co-expression with GFP-TFIISmut (Figure 37 B-C) which could reflect the fraction of free NRPB1 still dynamically recruited to chromatin (Steurer et al., 2018).



C)

	Half-time [sec.]		Mobile fraction [%]	
NRPB1-mcherry + GFP-TFIIS	0,15 ± 0,09	n=5; p-value: 0,916	97,20 ± 2,77	n=5; p-value: 0,032
NRPB1-mcherry + GFP-TFIISmut	0,16 ± 0,14		87,00 ± 8,37	

Figure 37. NRPB1 mobile fraction decreases in the presence of mutated TFIIS.
(A) Representative pictures of nuclei in *N. benthamiana* leaf epidermal cells taken during FRAP performed for NRPB1-mCherry transiently co-expressed with inducible GFP-TFIIS in GFP-TFIISmut (24h β-estradiol induction). The region of interest (ROI; doted white line) were photobleached and the recovery of the mCherry fluorescence intensity was measured over-time by confocal microscopy (CLSM). Pre-Bleach indicates the first timepoint of the series (t = 0 s), Post-Bleach#1 the first timepoint after bleaching (t = 2,7 s) and Post-Bleach#2 the last time point of the series (t = 10,3 s). Pseudo-coloured images (modified fire LUT) with respective colour calibration bar are shown. (C-D) Mobile fraction and recovery half-time after photobleaching ($t_{½}$) values were calculated for NRPB1-mCherry using easyFRAP and visualised using a whisker-box plot (C) or given in the table together with the significance level (D). The significance was tested by Student's T-Test: * p-value < 0.05.

2.4.1 TFIIS mutation triggers recruitment of proteasomal components

RNAPII backtracked in the absence of functional TFIIS has been previously shown to be subjected for polyubiquitination followed by proteasomal degradation (Karakasili et al., 2014). Similarly, TFIIS mutation results in yeast NRPB1 polyubiquitination presumably followed by proteasomal degradation (Sigurdsson et al., 2010) Thus, the fate of *Arabidopsis* NRPB1 in the presence of mutated TFIIS was examined. NRPB1 polyubiquitination was first evaluated by Western Blot using the antibodies directed against ubiquitin (α-UBQ) incubated with membrane containing GS-TFIIS and GS-TFIISmut AP eluates isolated from transgenic PSB-D cells (chapter 2.2.4). As seen on the Figure 38 A, strongly smeared signal with some faint bands appeared in the region where NRPB1 protein would be expected to travel on the polyacrylamide gel. Detected signal was somewhat stronger in GS-TFIISmut pulldown when compared relatively to GS-TFIIS AP eluate (Figure 38 A).

Following AP-MS approach performed with iGS-TFIIS and iGS-TFIISmut lines (chapter 2.2.4), proteins copurified specifically with GS-TFIISmut were identified by using GS-TFIIS pulldown

as a background. Accordingly, 21 nuclear proteins specifically copurified with GS-TFIISmut bait could be identified (Figure 38 B, Supplementary Table 13).

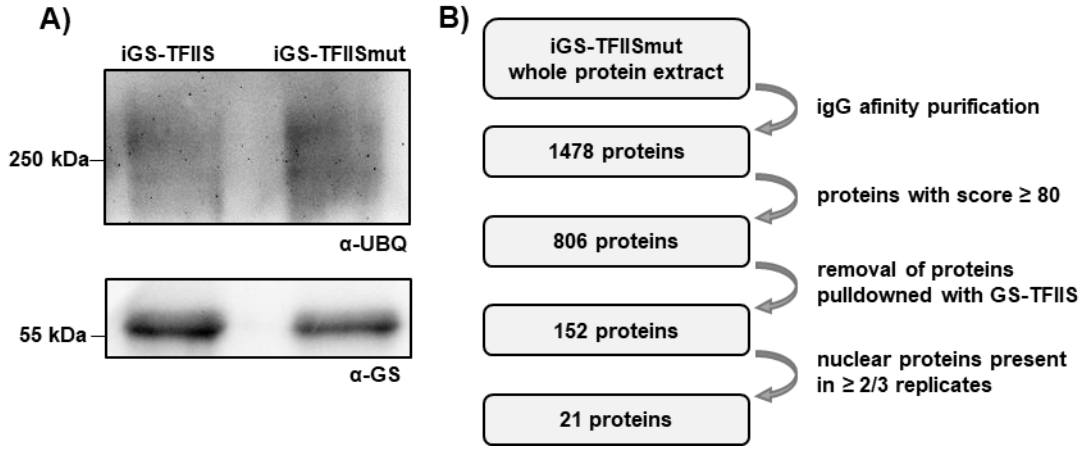


Figure 38. NRPB1 ubiquitination in the presence of mutated TFIIS.
(A) Immunoblot with α-UBQ and α-GS (loading control) of GS-TFIIS and GS-TFIISmut AP eluates affinity-purified from the whole protein extract of transgenic PSB-D lines. (B) Data proceeding workflow following AP-MS performed with transgenic PSD-B cultures harbouring GS-TFIIS and GS-TFIISmut (2.2.4).

Identified 21 proteins were further subjected to GO enrichment unravel the cellular components specifically associated with TEC containing mutated TFIIS. As a result, several GO categories related to proteasome as well as ubiquitin ligase complex were overrepresented among the interactome specific for GS-TFIISmut (Table 8). These findings could suggest the targeting of arrested RNAPII for proteasomal degradation similarly to previously reported in yeast.

Table 8. GO terms overrepresented among proteins specifically copurified with GS-TFIISmut.
The Gene Ontology (GO) analysis was performed with PANTHER using the single enrichment analysis (SEA) in the context of cellular components for GS-TFIISmut interactome specified at Supplementary Table 13. All enriched GO terms (FDR < 0.05) are shown in the table sorted by fold enrichment level. Cellular component related to proteasome and ubiquitin ligase complex related are highlighted in yellow.

GO term ID	description	Fold enrichment	p-value	FDR
GO:0008540	proteasome regulatory particle, base subcomplex	> 100	1,15E-04	8,24E-03
GO:0022624	proteasome accessory complex	74,84	3,63E-04	1,69E-02
GO:0005838	proteasome regulatory particle	74,84	3,63E-04	1,62E-02
GO:0032040	small-subunit processome	53,45	6,91E-04	2,85E-02
GO:0080008	Cul4-RING E3 ubiquitin ligase complex	31,94	1,17E-04	7,39E-03
GO:0005730	nucleolus	20,19	3,36E-08	5,13E-06
GO:0031461	cullin-RING ubiquitin ligase complex	17,70	6,43E-04	2,75E-02
GO:0044445	cytosolic part	14,97	1,35E-04	7,63E-03
GO:0000151	ubiquitin ligase complex	13,79	1,31E-03	4,84E-02
GO:0031981	nuclear lumen	13,30	8,95E-10	2,40E-07
GO:0070013	intracellular organelle lumen	12,00	2,69E-10	2,88E-07
GO:0031974	membrane-enclosed lumen	12,00	2,69E-10	1,44E-07
GO:0043233	organelle lumen	12,00	2,69E-10	9,60E-08
GO:1990904	ribonucleoprotein complex	10,59	2,48E-06	2,05E-04
GO:0044428	nuclear part	9,76	1,70E-08	3,03E-06
GO:1902494	catalytic complex	8,40	1,12E-05	8,58E-04
GO:1990234	transferase complex	8,39	1,18E-03	4,51E-02

GO term ID	description	Fold enrichment	p-value	FDR
GO:0043232	intracellular non-membrane-bounded organelle	7,19	1,48E-06	1,76E-04
GO:0043228	non-membrane-bounded organelle	7,19	1,48E-06	1,59E-04
GO:0032991	protein-containing complex	5,13	4,12E-07	5,52E-05
GO:0044446	intracellular organelle part	3,52	2,33E-06	2,27E-04
GO:0044422	organelle part	3,51	2,37E-06	2,11E-04
GO:0005886	plasma membrane	3,42	2,23E-04	1,20E-02
GO:0071944	cell periphery	2,93	8,05E-04	3,19E-02
GO:0005634	nucleus	2,68	1,41E-08	3,02E-06
GO:0043231	intracellular membrane-bounded organelle	1,57	1,16E-04	7,77E-03
GO:0043227	membrane-bounded organelle	1,56	1,22E-04	7,26E-03
GO:0043226	organelle	1,53	2,54E-04	1,30E-02

2.4.2 Arrested NRPB1 is targeted for proteasomal degradation

Considering likely RNAPII degradation following RNA cleavage inhibition, the stability of NRPB1 in the presence of mutated TFIIS was further studied by Western Blot. The whole protein extracts were obtained from 10DAS iGFP-TFIIS and iGFP-TFIISmut seedlings exposed to 24h β-estradiol or mock induction prior to material harvesting. All plants were simultaneously treated with 10 mM cycloheximide to inhibit novel protein synthesis (Kurepa et al., 2010). NRPB1 level was examined by Western Blot using non-phospho specific antibodies directed against CTD fragment of NRPB1 (α-CTD). NRPB1 level was then comparatively analysed between iGFP-TFIIS and iGFP-TFIISmut relatively to the level of UAP56 protein used as a loading control (α-UAP56). As a result, NRPB1 protein level in iGFP-TFIISmut induced with β-estradiol was reduced by ~ 50% relatively to control samples (Figure 39 A-B).

To confirm that observed NRPB1 degradation upon TFIISmut expression is mediated through proteasomal pathway, iGFP-TFIISmut seedling were additionally treated with MG132 to inhibit proteasomal degradation (Croager, 2004). As expected, upon 10 µM MG132 addition NRPB1 level in iGFP-TFIISmut was restored to the one observed in iGFP-TFIIS exposed to β–estradiol (Figure 39 A-B).

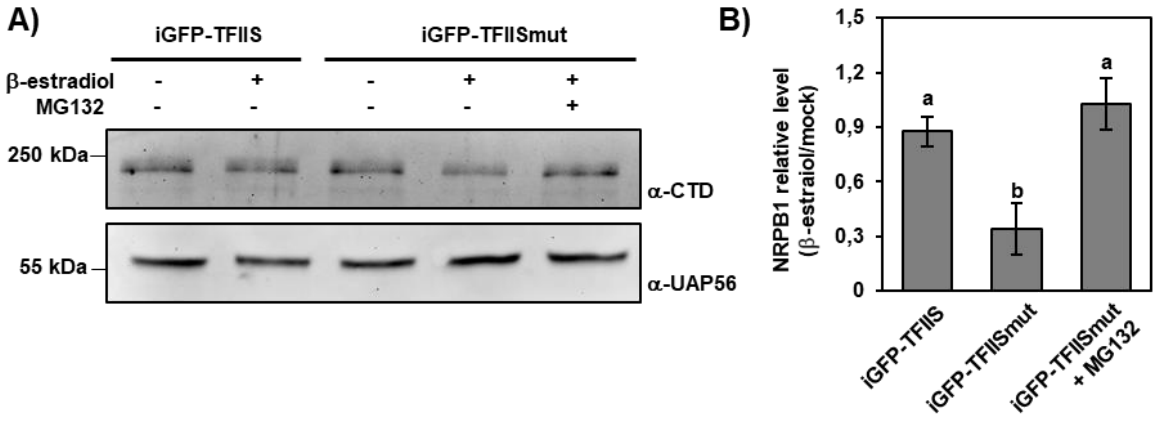


Figure 39. NRPB1 is subjected to proteasomal degradation upon TFIISmut expression.
(A) Immunoblot with α-CTD (NRPB1) and α-UAP56 (loading control) of whole protein extracts from 10DAS iGFP-TFIIS and iGFP-TFIISmut seedlings exposed to 24h induction with 2 µM β-estradiol or EtOH (mock). All seedlings were co-treated with 10 mM cycloheximide for 24h. iGFP-TFIISmut seedlings were additionally supplemented with 10 µM MG132 for 24h (last line). (B) Band intensities detected by Western Blot were measured by ImageJ. Error bars indicate SD of three independent experiments. Results were statistically analysed by one-way ANOVA. The letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05).

2.4.3 Arrested RNAPII may collide with cell cycle progression

Arrested TECs are very stable thus may present a serious obstacle to any ongoing cellular processes on DNA template including DNA replication (García-Muse and Aguilera, 2016). In view of that, the progression through the mitotic cycle as well as endoreduplication was examined in plants expressing mutated TFIIS.

CYCB1;1-GFP reported line was used to study the progression through mitotic cycle as described previously (Colón-Carmona et al., 1999; Dürr et al., 2014). First, GS-TFIIS and GS-TFIISmut transgenes used in AP-MS approach (chapter 2.2.2.3) were integrated into *tfIIs-1* genome by *Agrobacterium*-mediated transformation. The inducibility of selected primary transformants (T1) was validated by RT-PCR showing expression of target inducible transgenes upon 24h β-estradiol induction with no background expression (mock induction, Figure 40 A). These transgenic lines were further crossed with CYCB1;1-GFP reporter line to visualise the cells at the G2-M phase of the cell cycle (Colón-Carmona et al., 1999). Several individuals of F1 progeny following crossing ("iGS-TFIIS$^{+/-}$ "and "iGS-TFIISmut$^{+/-}$ ", respectively)

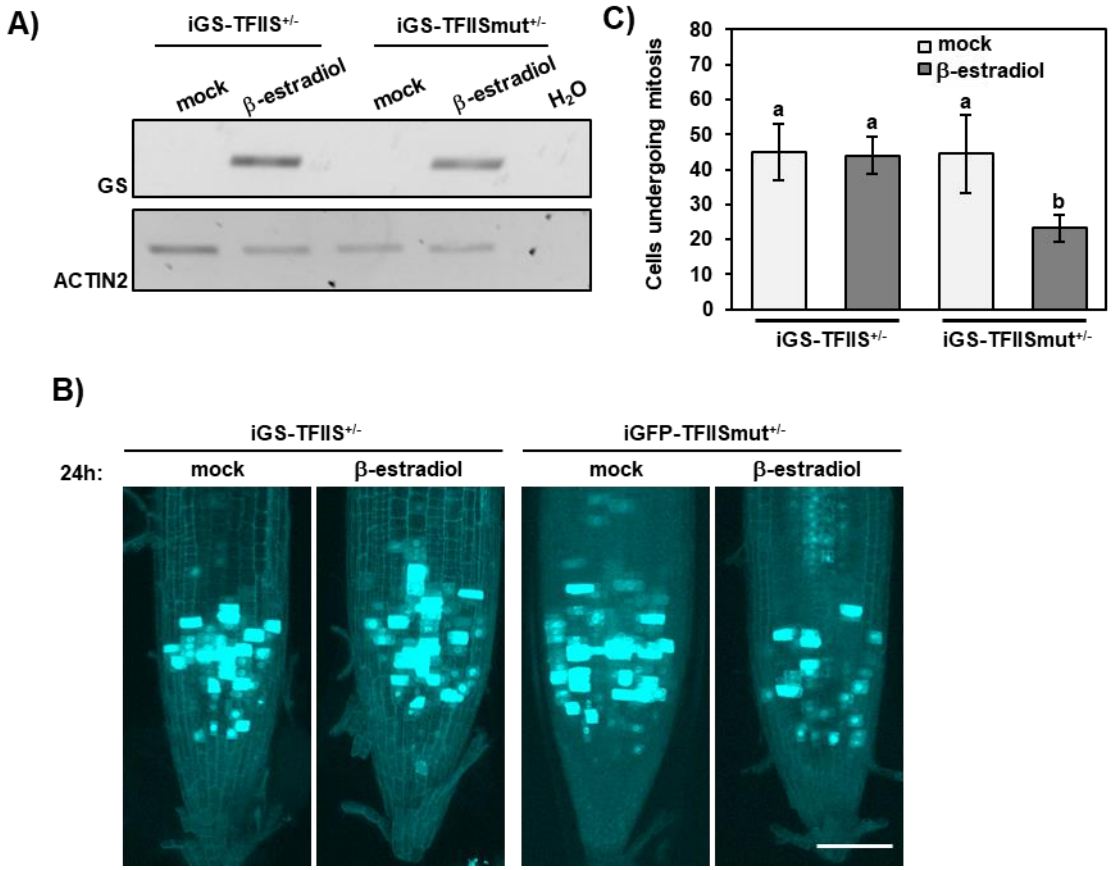


Figure 40. Arrested RNAPII may interfere with mitotic cell cycle progression
(A) RT-PCR validation of transgenes inducibility in iGS-TFIIS and iGS-TFIISmut lines. Primer targeting GS tag and ACTIN2 were used for the detection of inducible transgenes and reference gene, respectively. RNA was isolated from 10DAS seedlings. (B) Representative pictures of CLSM z-stack imaging of primary root tips. Seedlings harbouring either GS-TFIIS or GS-TFIISmut transgene in combination with pCYCB1;1-CYCB1;1-GFP reporter transgene were subjected for the analysis. GFP and autofluorescence signal are shown in cyan. White bars indicate 100 µm. Representative pictures taken at 10DAS are shown. (C) Mean number of cells undergoing mitosis (GFP-positive) was counted for each line/condition (n = 5). Error bars indicate SD. Statistical analysis was performed using one-way ANOVA. The letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05).

were exposed to 24h β-estradiol or mock induction and GFP expression was studied by confocal microscopy (CLSM). The number of GFP-stained cells was determined revealing significant reduction of cells undergoing mitosis in iGS-TFIISmut$^{+/-}$ genomic background after exposure to β-estradiol in comparison to respective controls (Figure 40 B-C). Observed compromised cell division upon in the presence of mutated TFIIS could explain the inhibition of main root elongation in iGFP-TFIISmut seedings exposed to β-estradiol (Figure 17).

Plant cells possess the ability to modify their classical cell cycle into a partial cell cycle where DNA synthesis occurs independently from mitotic division resulting in cell polyploidisation (Barow and Meister, 2003; Joubès and Chevalier, 2000). The process called "endoreduplication" is widespread among eukaryotes and occurs very commonly in *Arabidopsis* (Ullah et al., 2009; Yin et al., 2014). Since the progression through mitotic cycle was compromised in the presence of mutated TFIIS (Figure 40 B-C), the endoreduplication process was also examined in that context. Cell polyploidisation was used as a measure for successful endoreduplication (Lermontova et al., 2006). Measurements were expressed as a "cycle value" which indicates the mean number of endocycles per nucleus (Barow and Meister, 2003). Cell ploidy was determined by using Fluorescence Activated Cell Sorting (FACS) in 10DAS iGFP-TFIIS#3 and iGFP-TFIISmut#1 seedlings exposed to 24h β-estradiol or mock induction. FACS measurements and initial data analysis were performed by Dr. Jörg Fuchs (IPK, Gatersleben). Nuclei were extracted from individual seedlings by tissue chopping in extraction buffer. Obtained nuclei extracts were stained with DAPI and loaded into Flow Cytometer. After reads collection (~ 5000) cell ploidy was determined for each measurable nucleus of iGFP-TFIIS#3 and iGFP-TFIISmut#1 regardless GFP expression level. As a result, no significant differences were observed between analysed genotypes in terms of ploidy level nor cycle value (Figure 41 A-B).

Next, remaining nuclei were fluorescently sorted according to the GFP signal. The baseline GFP signal intensity was established based on the background signal in iGFP-TFIIS#3 and iGFP-TFIISmut#1 seedlings exposed to mock induction (Supplementary Figure S 10). Cell ploidy was subsequently calculated across the population of GFP-positive cells (GFP+) revealing significantly decreased cycle value in the presence of GFP-TFIISmut relatively to GFP-TFIIS expression (Figure 41 D). Accordingly, cell ploidy in iGFP-TFIISmut#1 could be characterised with lower ploidy states (2C = 21,6% and 4C = 32,4%) when compared with iGFP-TFIIS#3 (2C = 14,3% and 4C = 21,6%) (Figure 41 C).

Similar effects were observed among plants exposed to 72h β-estradiol induction whereas mock-induced iGFP-TFIISmut#1 (either 24h or 72h) had comparable ploidy to iGFP-TFIIS#3 line (data not shown).

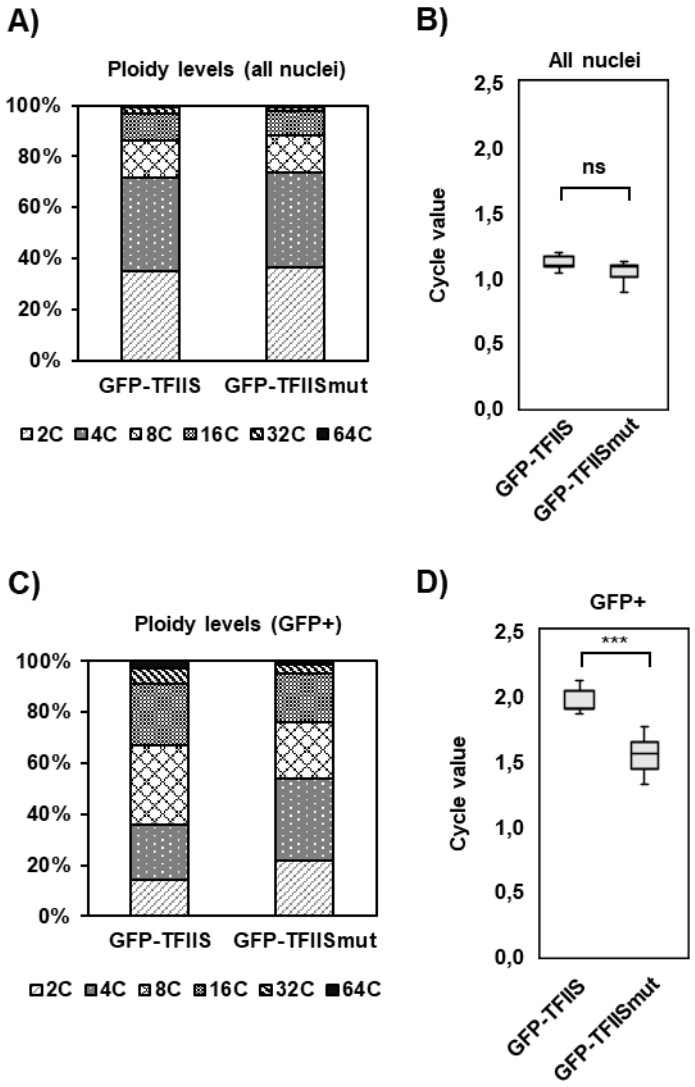


Figure 41. Arrested RNAPII may interfere with endoreduplication.
Ploidy levels (A, C) and cycle values (B, D) determined by FACS for the nuclei (~5000) extracted from 10DAS iGFP-TFIIS and iGFP-TFIISmut seedlings exposed to 24h β-estradiol or mock induction. Ploidy was determined regardless GFP expression (A-B) or only for nuclei with detectable GFP signal (C-D).

Additionally, GFP-directed sorting of iGFP-TFIIS#3 and iGFP-TFIISmut#1-derived nuclei allowed the estimation of induction efficiency by calculating the fraction of cells expressing target inducible protein in studied 10DAS seedlings. Accordingly, transgenic GFP-TFIIS was expressed in 25,8% cells whereas GFP-TFIISmut was present in 33,1% cells in iGFP-TFIIS#3 and iGFP-TFIISmut#1 seedling upon 24h β-estradiol induction, respectively (Supplementary Table 14). The number of GFP-positive cells decreased in studied transgenic lines upon 72h β-estradiol induction (21,1% and 17,9%, respectively) suggesting system saturation over longer induction, similarly to observed previously (Zuo et al., 2000).

Taken together, the indirect molecular consequences of TFIIS mutations are reflected in lower TEC mobility as well as compromised progression through the cell cycle. Following RNAPII arrest, *Arabidopsis* NRPB1 is likely being targeted for polyubiquitination and proteasomal degradation.

2.5 Determination of transcript elongation rate

Recent advances in sequencing techniques measuring nascent transcripts or RNAPII position in single nucleotide resolution has lead to the development of diverse methods detecting elongation rates over single genes (Fukaya et al., 2017; Jonkers and Lis, 2015; Lavigne et al., 2017). However, most of these methods require high-throughput sequencing and their usage is still technically very challenging in *Arabidopsis* (Hetzel et al., 2016).

In view of that, it has been attempted to develop a novel molecular tool allowing the detection of newly transcribed mRNA as well as the determination of transcript elongation rate in different genomic backgrounds. The molecular system for transcript elongation rate determination was designed utilizing β-estradiol inducible system described in the previous chapters.

2.5.1 Elongation rate system design

Yeast VPS13 (~ 9,5 kb) was chosen as a reporter gene to allow the determination of newly synthesised mRNA allowing high spacial resolution without background expression. This gene has been also previously used for studying RNAPII elongation rate in yeast (Santisteban et al., 2011). VPS13 was additionally C-terminally fused with GFP and Luciferase (LUC) to allow its over-time detection on the protein level. VPS13-GFP-LUC fusion transgene was subsequently

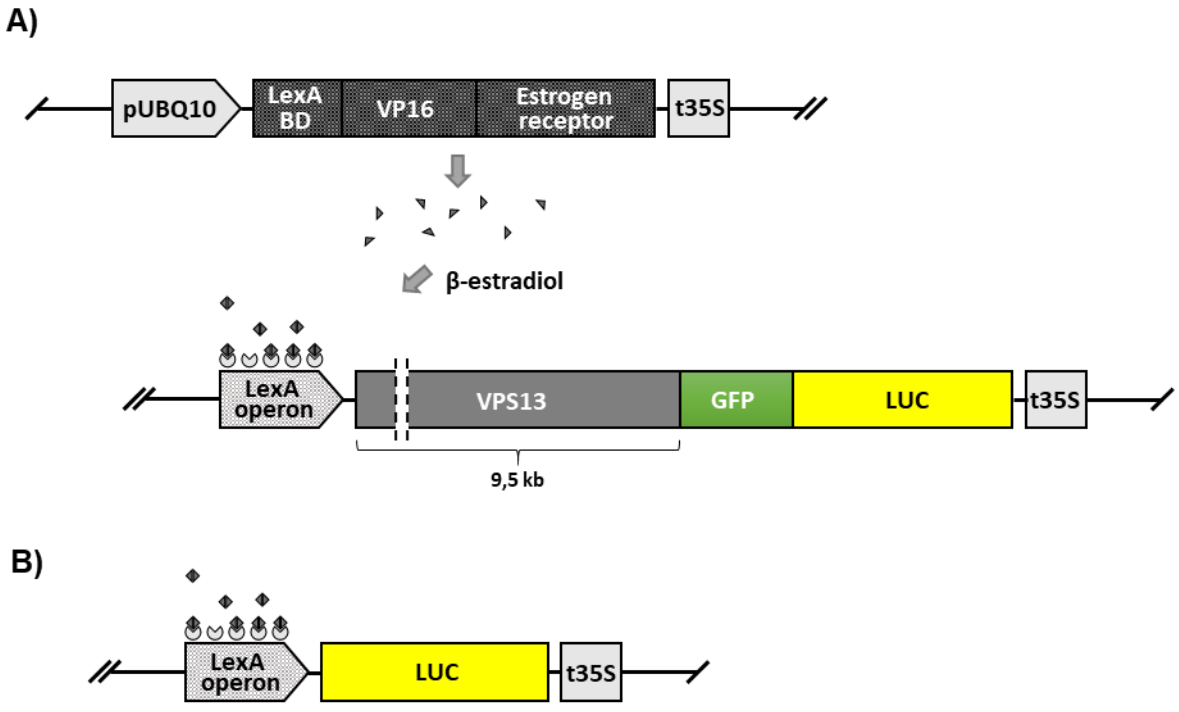


Figure 42. Inducible transgenes design for the determination of transcript elongation rate.
(A-B) Schematic illustration of transgenes used for the determination of elongation rate framed into β-estradiol inducible system. Inducible transgenes contain VGL13-GFP-LUC (A) or LUC alone (B). (A) Constitutive expression of activator unit (grey triangles) bind to the responsive unit in the presence of β-estradiol determining its activation. pUBQ10: native promoter of *Arabidopsis* Ubiquitin 10; LexA BD: binding domain of LexA operon; VP16: acidic transactivation domain of human VP16; t35S: CaMV 35S terminator. (B) Only responsive unit of β-estradiol inducible system containing LUC alone is shown.

framed into created β-estradiol inducible system (Figure 42 A) and incorporated into Col-0 genome by *Agrobacterium*-mediated transformation.

Since production of a fusion protein expressed in heterogeneous system could hamper the functionality of reporter genes (Snapp, 2005), independent reporter system comprising exclusively LUC gene under β-estradiol system was created in parallel (Figure 42 B). Obtained transgene was also incorporated into Col-0 genome by *Agrobacterium*-mediated transformation.

Several independent transgenic lines carrying inducible VPS13-GFP-LUC or LUC transgene in Col-0 (referred as "iVGL" and "iLUC", respectively) were identified and further analysed to determine transgene copy number.

2.5.2 Genetic validation of elongation rate system

As a result of *Agrobacterium*-mediated transformation, carried T-DNA is being randomly inserted in the genome with multiple insertions (Dean et al. 1988). T-DNA position in the genome could affect transgene inducibility for instance due to the lack of regulatory DNA sequences like silencers (Rose et al. 2008, Riethoven 2010). Therefore, it was desired that each comparatively analysed line contains single insertion of reporter transgene in the same genomic position. To this end, it was attempted to select transgenic lines containing single insertion in Col-0 background which could be further crossed with any T-DNA line deficient in chosen TEFs. Such approach would ensure comparable inducibility between analysed transgenic lines.

Southern Blot assay was applied to determine transgene copy number in iVGL and iLUC lines. Following the digestion of genomic DNA with chosen restriction enzymes (EcoRI and XbaI), radiolabelled probe would allow the detection of T-DNA fragment of reporter transgene. Probe was designed to complement the part of T-DNA sequence proximal to left border and adjacent to the restriction site recognised by chosen enzymes. While one cut occurs within T-DNA (~ 2,1 kb away from left border), the position of the second cut would occur in genomic DNA in the position dependent on the sequence adjacent to T-DNA (Figure 43 A). As a result, DNA fragment/s of unknow size (> ~ 2000 bz) would be hybridised by radiolabelled probe. Consequently, single T-DNA insertion in the genome could only produce single detectable fragment, while multiple T-DNA insertion would likely result in the detection the several DNA fragments of different size.

Southern Blot was performed using genomic DNA extracted from 10DAS seedling of several independent iVGL and iLUC lines. The usage of EcoRI and XbaI restriction enzymes ensured robust digestion of entire genomic DNA (Supplementary Figure S 11). Digested DNA was transferred onto Hybond N membrane and incubated with 32P-labelled DNA probe complementary to transgenic reporter sequence. After visualisation distinct band patterns could be observed across individual iVGL and iLUC lines (Figure 43 B). As expected, all DNA fragments hybridised with radiolabelled probe were > 2 kb suggesting its specificity. In several lines including iVGL#3, 5 and 8 as well as iLUC#1 and 6 only single band was observed regardless applied restriction enzyme whereas remaining lines contained several detectable DNA fragments (Figure 43 B). Lines identified with unique DNA fragment hybridised with

radiolabelled DNA probe presumably contained a single T-DNA insertion and were further analysed in terms of system functionality.

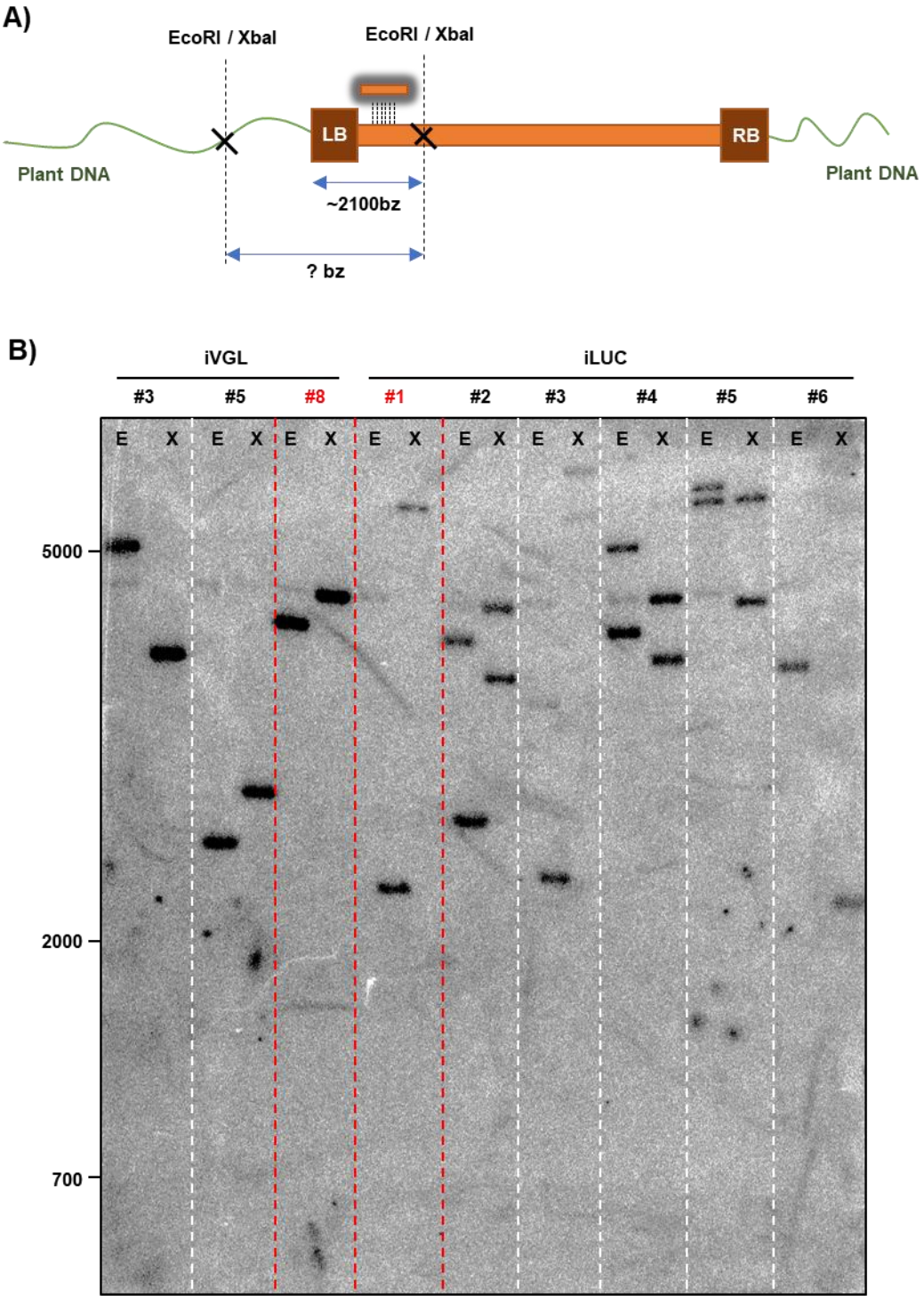


Figure 43. Determination of inducible transgene copy number in iVGL and iLUC lines using Southern Blot.
(A) Schematic illustration demonstrating the principles underlying performed Southern Blot analysis and expected outcome (for details see text). Inducible transgene (orange boxes) and adjacent DNA genomic sequence (green line) are shown. Orange box surrounded with shadow reflect DNA probe and its hybridisation site is indicated. Known and putative restriction sites are shown as "x". LB: left border. RB: right border. (B) The outcome of Southern Blot assay. Following genomic DNA digestion with chosen restriction enzymes, DNA was transferred on the membrane and hybridised with ^{32}P-labelled DNA probe. Red colour indicates transgenic lines with single insertion selected for further analysis. E: EcoRI. X: XbaI. Numbers on the left indicate size in bp.

2.5.3 Functional validation of elongation rate system

The inducibility of pre-selected iVGL lines was tested on transcript level. To this end total RNA was extracted from 6DAS seedlings exposed to 20 µM β-estradiol for 0h (no induction), 1h or 24h prior to material harvesting. VPS13 transcript level was further evaluated in each line/condition by RT-PCR using UBQ5 as a reference gene. As seen on the Figure 44, clear induction was observed for iVGL#8 line whereas VPS13 transcript level was not enriched for iVGL#3 and iVGL#5 lines even after 24h β-estradiol induction. Importantly, no background expression was detected in iVGL#8 unlike in iVGL#3 and iVGL#5. Therefore iVGL#8 line was chosen for further experiments.

System inducibility within iVGL#8 line was further studied by using confocal microscopy (CLSM) and microplate luminometer (Centro XS³ LB 960) to determine GFP expression and LUC activity, respectively. However, despite the efforts neither GFP expression nor LUC activity could be detected following 24h β-estradiol induction (data not shown). In line with RT-PCR data, neither GFP expression nor LUC activity were detected in iVGL#3 and iVGL#5 lines (data not shown).

In view of that, iLUC lines were analysed in terms of LUC activity to evaluate the influence of N-terminal tagging with VGP13-GFP on LUC inactivity. 6DAS iLUC#1 and iLUC#6 lines identified with single T-DNA insertion of reporter transgene were exposed to 20 µM β-estradiol. Following β-estradiol addition to the liquid MS media all seedlings were vacuum infiltrated for 10 min. D-Luciferin (Promega) was next added to each sample up to total concentration of 60 µM. First bioluminescence measurement from individual seedlings was taken 20 min after β-estradiol application followed by continuous measurements every 10 min over the period of 2 hours. Each measurement was normalised to the average background bioluminescence signal coming from uninduced plants. As a result, clear bioluminescence could be detected ~30 min following β-estradiol application in both iLUC#1 and iLUC#6 with signal intensity growing exponentially over 2 hours (Figure 44 B). Detected signal was clearly stronger in iLUC#1 relatively to iLUC#6. Thus, iLUC#1was chosen for further experiments due to better inducibility. The lack of LUC activity in iVGL lines was therefore likely caused by incorrect folding of VGL fusion protein.

It was next attempted to determine whether iLUC system could reflect the differences in transcript elongation rate on the protein level. Thus, the accumulation of bioluminescence signal in iLUC#1 was examined in the presence of α-amanitin, a well-established inhibitor of RNAPII (Bensaude, 2011). α-amanitin was applied simultaneously with β-estradiol to the final concentration of 10 µM or 100 µM. Consequently, the kinetic of LUC activity in iLUC#1 exposed to α-amanitin was strongly compromised in the dose-dependent manner in comparison to non-treated plants (Figure 44 C). Since RNAPII & RNAPIII are only known

targets of α-amanitin (Bensaude, 2011), observed effects are presumably the consequence of compromised transcript elongation rate.

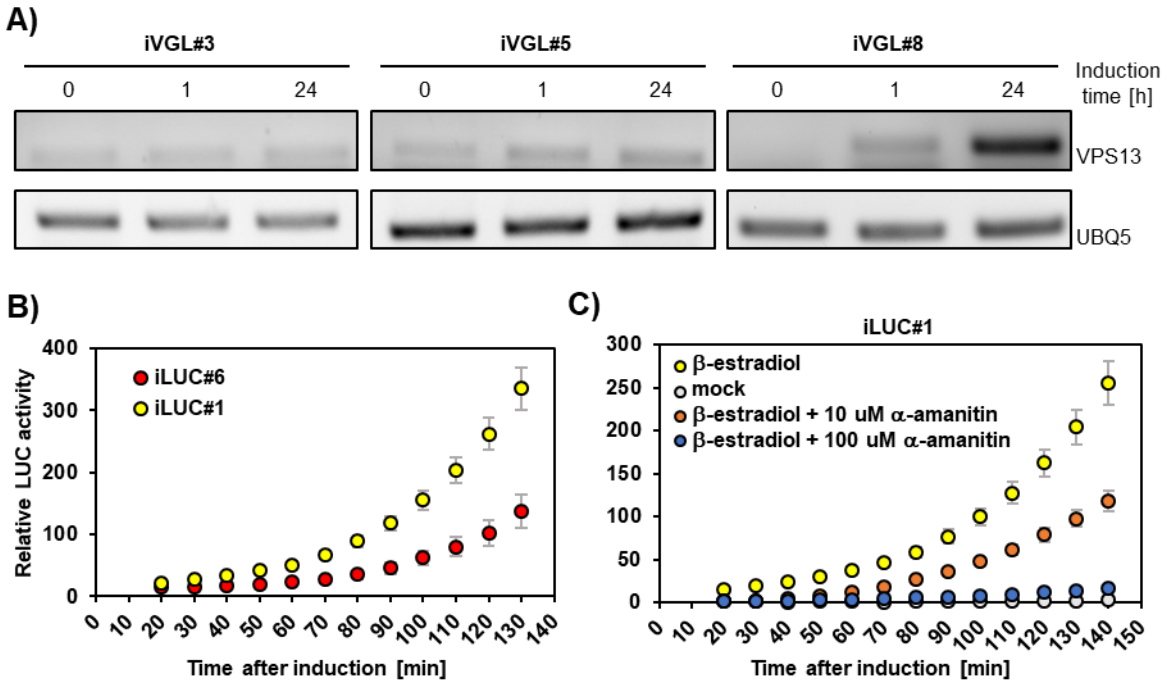


Figure 44. The validation of elongation rate systems inducibility.
(A) Transcript level of inducible VGL transgene was determined by semi quantitative RT-PCR with UBQ5 as a reference gene. Total RNA was extracted from 6DAS seedlings exposed to 20 μM β-estradiol induction for 0, 1 and 24 hours. (B-C) Bioluminescence signal measured from 6DAS *Arabidopsis* seedlings (n = 6 for each line/condition) following β-estradiol application. LUC activity was monitored every 10 minutes over the period of 2 hours using a microplate luminometer LB-960. Dots indicate mean signal ± SD (error bars) normalised to the signal from uninduced plants. (C) iLUC#1 seedlings were additionally supplemented with α-amanitin applied simultaneously with β-estradiol.

The optimisation of induction conditions as well as measurements normalisation were next addressed for iLUC#1 to ensure reliable comparison between individual seedlings regardless plant size. To this end, LUC activity was measured separately within roots and leaf tissue dissected from 6DAS iLUC#1 seedlings. Signal intensities for each seedling were then normalised to fresh weight of analysed roots or leaf area, respectively. Measurements performed for iLUC#1 line revealed clear signal detection in roots similarly to observed before for whole 6DAS seedlings, while relatively weak signal was detected in leaves (Supplementary Figure S 12). In view of that, following experiments in terms of elongation rate system were performed using root tissue.

Taken together, iVGL and iLUC reporter lines harbouring single T-DNA insertion in Col-0 were obtained. While iVGL showed good inducibility on transcript level, iLUC allowed the measurements of perturbated transcription on the protein level. Thus, it has been decided to use both iVGL and iLUC reporter system to measure the effect of TFIIS deficiency on transcript elongation rate on either transcript or protein level, respectively.

2.5.4 Transcript elongation rate in plants lacking functional TFIIS

TFIIS has been recently demonstrated as a major stimulus of *in vivo* elongation rate in human (Sheridan et al., 2019). To assess the role of *Arabidopsis* TFIIS in the regulation of transcript elongation rate, iVGL#8 and iLUC#1 line were crossed with *tfIIs-1* to introduce inducible iVGL and iLUC transgenes into genomic background lacking functional TFIIS. Their presence in *tfIIs-1* was further confirmed by PCR-based genotyping following the segregation of progeny lines (data not shown).

Elongation rate was first evaluated on the transcript level using iVGL reporter transgene. 6DAS Col-0 and *tfIIs-1* plants carrying iVGL transgene were vacuum infiltrated for 10 min with the liquid MS medium containing 20 µM β-estradiol. Roots were dissected from individual seedling at flash frozen in liquid nitrogen after 20 min and 60 min following β-estradiol induction. Uninduced plants were used as a control for background expression (0 min). Total RNA was isolated for each genotype/condition and iVGL transcript level was determined by qPCR at both 5' and 3' ends (Figure 45) using GAPC, PP2AA3 and UBQ10 as reference genes (Kudo et al. 2016). Calculated iVGL amounts were next normalized relatively to their respective values determined in Col-0 exposed to 60 min β-estradiol induction. As a result, iVGL transcript level at the 5' end was significantly decreased in *tfIIs-1* relatively to Col-0 after 60 min induction, whereas the difference after 20 min was not statistically different (Figure 45, left panel). Towards the 3' end of the reporter, slower accumulation of iVGL transcript in *tfIIs-1* could be observed already 20 min following β-estradiol application (Figure 45, right panel). The determination of iVGL transcript in uninduced plants showed no expression in neither Col-0 nor *tfIIs-1* (Figure 45 B, "0").

The effects of compromised elongation rate in *tfIIs-1* were next evaluated in the context of protein accumulation. To this end, 6DAS Col-0 and *tfIIs-1* plants carrying iLUC transgene were exposed to β-estradiol induction and LUC activity was determined as described the chapter above. The analysis revealed comparable LUC activity between Col-0 and *tfIIs-1*, with initially higher signal in *tfIIs-1* (~ 70 - 90 min) followed by elevated signal accumulation in Col-0 (~ 120 min) (Figure 45 C). Those results do not reflect compromised transcript elongation rate observed in *tfIIs-1*. However, due to a small size of iLUC reported the spatial resolution of those measurements could be limited.

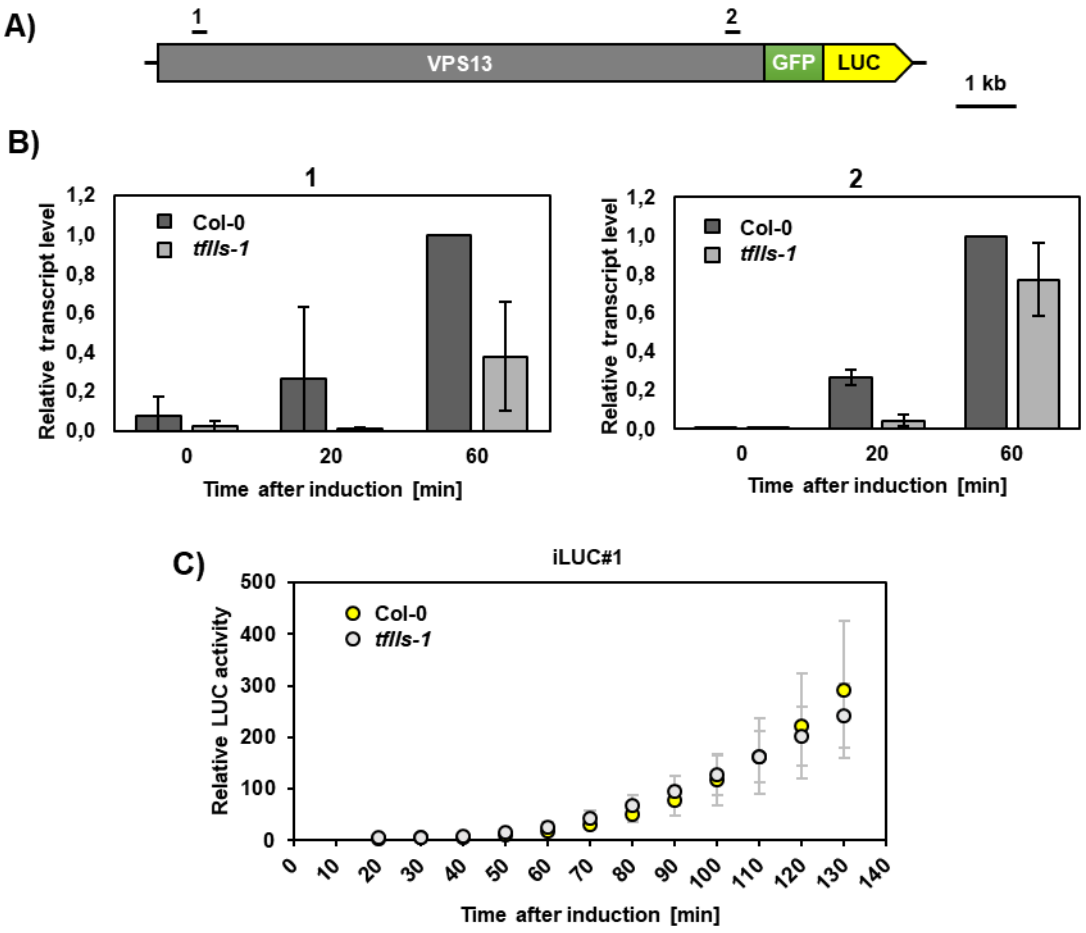


Figure 45. Transcript elongation rate in compromised in *tfIIs-1.*
(A) Schematic illustration of VGL transgene with the transcriptionally active (boxes) and inactive regions (black lines). The bars above indicate the relative positions of the regions analysed by qPCR. (B) VGL transcript level measured by qPCR at the regions indicated at (A) and determined relatively to GAPC, PP2AA3 and UBQ10 reference genes (Kudo et al. 2016) Each value was subsequently normalised to VGL transcript level in Col-0 after 60 min β-estradiol induction. (C) Bioluminescence signal measured from 6DAS *Arabidopsis* seedlings (n = 6 for each line) following β-estradiol application. LUC activity was monitored every 10 minutes over the period of 2 hours using a microplate luminometer LB-960. Dots indicate mean signal ± SD (error bars) normalised to the signal from uninduced plants. Measurements were adjusted to the respective fresh weight of individual roots.

3. Results: Genetic interaction between TFIIS and ELF7

TFIIS can be characterised as a one of the key regulators of transcript elongation as well as integral component of the transcript elongation complex. Accordingly, a broad spectrum of genetic interaction between TFIIS and other transcription-related factors has been demonstrated (see "Introduction"). The genetic interaction between *Arabidopsis* TFIIS and ELF7 has been studied in the second part of this study to better understand the mutual contribution of various TEFs into the regulation of transcript elongation but also other molecular processes.

3.1 The interaction between TFIIS and ELF7 in *Arabidopsis*

TFIIS interactome in *Arabidopsis* was initially determined by using affinity purification coupled with mass spectrometry (AP-MS) with TFIIS as a bait protein. This experiment was a part of project aiming to determine to composition of *Arabidopsis* TEC (Antosz et al., 2017). The finding related to TFIIS and ELF7 will be presented in detail and elaborated in the following chapters.

3.1.1 TFIIS and ELF7 efficiently copurify with each other as a part of TEC

In order to determine putative interaction partners of constitutively expressed TFIIS, its coding sequence was fused with GS tag (Figure 46 A) and used as a "bait" protein in the AP-MS approach. GS-TFIIS transgene driven by TFIIS native promoter was integrated into *Arabidopsis* cell culture (PSB-D) system by *Agrobacterium*-mediated transformation. Following the selection of transformed cells and their upscaling, GS-TFIIS bait protein was immunoprecipitated together with putative interactors from whole protein extract by one-step affinity purification (AP) using IgG coupled magnetic beads. The eluate was further subjected to SDS-PAGE followed by Coomassie Blue stanning, in-gel trypsin digestion and protein identification by mass spectrometry. Unfused GS (GS) under the control of CaMV 35S promoter was used as a control to identify proteins unspecifically bound to GS tag.

Coomassie staining of whole protein extracts (Input) resolved by SDS-PAGE revealed no obvious differences between GS-TFIIS and GS transgenic lines. In the GS-TFIIS AP eluate a band corresponding to TFIIS fused with GS taq (~ 70 kDa) was clearly visible. Additional faint bands could be seen indicating substoichiometric purification of TFIIS interactome. In the GS pulldown a strong band corresponding to the bait protein (~ 25 kDa) was detected, whereas very few additional bands could be seen (Figure 46 B). AP eluates were further subjected for mass spectrometry and the proteins copurified with baits were identified. Only proteins with a score higher than 80 and at least two peptides with an individual peptide score > 25 were considered for further analysis. Proteins immunoprecipitated with GS were subtracted from TFIIS interactome to remove unspecific background, resulting in 182 nuclear proteins copurified specifically with GS-TFIIS in at least two out of three independent experiments (Figure 46 C).

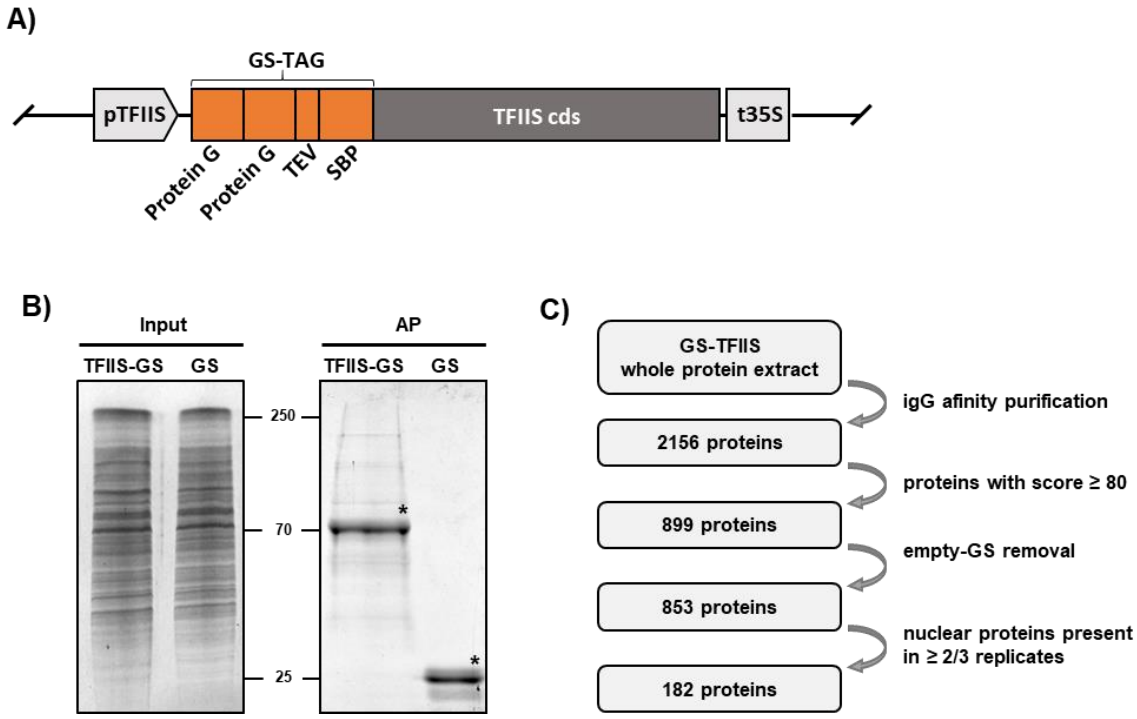


Figure 46. Identification of TFIIS interactome using AP-MS approach.
(A) Schematic illustration of TFIIS bait protein N-terminally fused to GS tag. ProteinG: IgG-binding domain of protein G; TEV: tobacco etch virus (TEV) protease cleavage site; SBP, streptavidin-binding peptide. (B) Total protein extracts of PSD-B cells expressing GS-TFIIS or unfused GS-tag (Input) and the eluates of affinity purifications (AP) were separated by 9 % SDS-PAGE and gels were stained with Coomassie Blue. Black asterisks indicate the bands corresponding to GS-TFIIS and unfused GS-tag. Number indicate molecular weight in kDa. (C) The workflow of MS data proceeding.

Proteins involved into the regulation of gene expression were further identified among TFIIS interactome. Not surprisingly most of them were related to "Transcription" although proteins involved into "Splicing", "Export", "Polyadenylation" and "Replication" were also identified. Within the set of 35 transcription-related proteins, diverse subunits of RNAPII as well as individual TEFs could be identified, including SPT5-2, SPT6L, FACT, Elongator and all subunits of *Arabidopsis* PAF1-C. Additionally, NAP1 proteins, HDACs and various histone variants were copurified together with TFIIS (Table 9).

Table 9. Transcription-related proteins co-purifying with constitutively expressed GS-TFIIS.
The list of transcription-related proteins immunoprecipitated with GS-TFIIS after data proceeding (Figure 40 C, bottom panel). The numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent APs.

GS-TFIIS	Interactor	Complex	Process	AGI
2576 / 3	TFIIS	TFIIS	Transcription	AT2G38560
2177 / 3	NRPB1	RNAPII	Transcription	AT4G35800
838 / 3	NRPB2	RNAPII	Transcription	AT4G21710
489 / 3	NRP(A/B/C/D)5	RNAPII	Transcription	AT3G22320
189 / 3	NRP(B/C/D/E)6a	RNAPII	Transcription	AT5G51940
471 / 3	NRPB7	RNAPII	Transcription	AT5G59180
160 / 2	NRP(A/B/C/D/E)8a	RNAPII	Transcription	AT1G54250
305 / 2	NRP(A/B/C/D/E)8b	RNAPII	Transcription	AT3G59600

GS-TFIIS	Interactor	Complex	Process	AGI
102 / 2	NRP(B/D/E)9b	RNAPII	Transcription	AT4G16265
207 / 3	NRP(B/D/E)11	RNAPII	Transcription	AT3G52090
1064 / 3	CTR9, ELF8, VIP6	PAF-C	Transcription	AT2G06210
1062 / 3	LEO1, VIP4	PAF-C	Transcription	AT5G61150
917 / 3	PAF1, ELF7	PAF-C	Transcription	AT1G79730
603 / 3	RTF1, VIP5	PAF-C	Transcription	AT1G61040
495 / 3	CDC73	PAF-C	Transcription	AT3G22590
422 / 3	SKI8, VIP3	PAF-C	Transcription	AT4G29830
864 / 3	SPT6-1, SPT6L	SPT6	Transcription	AT1G65440
225 / 3	SPT5-2	SPT4/SPT5	Transcription	AT4G08350
222 / 2	ELP3; ELO3	Elongator	Transcription	AT5G50320
155 / 2	SPT16	FACT	Transcription	AT4G10710
864 / 3	SPT6-1, SPT6L	SPT6	Transcription	AT1G65440
342 / 2	H2A.F/Z 3/HTA9	Histone H2A family	Transcription	AT1G52740
275 / 2	HTA6	Histone H2A family	Transcription	AT5G59870
182 / 2	HTA7	Histone H2A family	Transcription	AT5G27670
1399 / 2	HTB6	Histone H2B family	Transcription	AT3G53650
129 / 2	Histone H1.2	Histone H1 family	Transcription	AT2G30620
1030 / 3	HDT3	Deacetylase	Transcription	AT5G03740
557 / 2	HDT4	Deacetylase	Transcription	AT2G27840
602 / 2	AtNAPL2	histone chaperones	Transcription	AT2G19480
418 / 2	AtNAPL3	histone chaperones	Transcription	AT5G56950
137 / 2	atP14-1	17S U2 snRNP	Splicing	AT5G12190
233 / 2	atSF3b150	17S U2 snRNP	Splicing	AT4G21660
118 / 2	atSAP49a	17S U2 snRNP	Splicing	AT2G18510
106 / 2	atSR140-1	17S U2 associated	Splicing	AT5G25060
168 / 3	SKIP	Core NTC	Splicing	AT1G77180
131 / 2	SAP18	EJC/mRNP	Splicing	AT2G45640
129 / 2	atRSZp22/atSRZ22	SR proteins	Splicing	AT4G31580
344 / 3	SERRATE	mRNA binding proteins	Polyadenylation	AT2G27100
279 / 2	Nup136	Nucleoporin	Export	AT3G10650
225 / 2	Nup98a	Nucleoporin	Export	AT1G10390
124 / 2	Nup155	Nucleoporin	Export	AT1G14850
398 / 2	atALY-2a	TREX complex	Export	AT5G37720

To better understand the composition of *Arabidopsis* TEC containing TFIIS, its prevalence was screened among the AP of NRPB1 (performed by Karin Köllen) as well as other TEFs copurified with TFIIS (SPT5, SPT16 (by Alex Pfab), CDC73 and ELF7 (both by Hans Ehrnsberger). The reciprocal analysis revealed the presence of TFIIS in NRPB1, ELF7 and SPT5 pulldowns, however TFIIS could be identified in all biological replicates (3/3) only when using ELF7-SG as a bait protein. TFIIS and ELF7 interactomes were therefore analysed comparatively in the context of transcription-related proteins revealing their robust copurification together with some RNAPII subunits as well as other TEFs (Table 10).

Table 10. The overlap between transcription-related proteins copurified with GS-TFIIS and ELF7-SG.
The comparison of transcription-related proteins immunoprecipitated with GS-TFIIS and ELF7-SG. The numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent APs. ELF7-SG affinity purification was performed by Hans Ehrnsberger.

GS-TFIIS	ELF7-SG	Interactor	Complex	Process	AGI
2576 / 3	541 / 3	TFIIS	TFIIS	Transcription	AT2G38560
1064 / 3	8944 / 3	CTR9, ELF8, VIP6	PAF-C	Transcription	AT2G06210
1062 / 3	3141 / 3	LEO1, VIP4	PAF-C	Transcription	AT5G61150
917 / 3	3196 / 3	PAF1, ELF7	PAF-C	Transcription	AT1G79730
603 / 3	2269 / 3	RTF1, VIP5	PAF-C	Transcription	AT1G61040
495 / 3	2892 / 3	CDC73	PAF-C	Transcription	AT3G22590
422 / 3	2274 / 3	SKI8, VIP3	PAF-C	Transcription	AT4G29830
2177 / 3	1117 / 3	NRPB1	RNAPII	Transcription	AT4G35800
838 / 3	1017 / 3	NRPB2	RNAPII	Transcription	AT4G21710
489 / 3	364 / 2	NRP(A/B/C/D)5	RNAPII	Transcription	AT3G22320
189 / 3		NRP(B/C/D/E)6a	RNAPII	Transcription	AT5G51940
471 / 3	154 / 2	NRPB7	RNAPII	Transcription	AT5G59180
160 / 2		NRP(A/B/C/D/E)8a	RNAPII	Transcription	AT1G54250
305 / 2		NRP(A/B/C/D/E)8b	RNAPII	Transcription	AT3G59600
102 / 2		NRP(B/D/E)9b	RNAPII	Transcription	AT4G16265
207 / 3		NRP(B/D/E)11	RNAPII	Transcription	AT3G52090
864 / 3	1556 / 3	SPT6-1, SPT6L	SPT6	Transcription	AT1G65440
225 / 3	1444 / 3	SPT5-2	SPT4/SPT5	Transcription	AT4G08350
222 / 2	794 / 3	ELP3; ELO3	Elongator	Transcription	AT5G50320
155 / 2	4220 / 3	SPT16	FACT	Transcription	AT4G10710
	1696 / 3	SSRP1	FACT	Transcription	AT3G28730

3.1.2 TFIIS and ELF7 associate with active RNAPII

To determine whether TFIIS and ELF7 are part of actively transcribing TEC, both pulldowns were followed with Western Blot analysis using antibodies directed against various forms of NRPB1, the largest subunit of *Arabidopsis* RNAPII. The association of NRPB1 with either TFIIS or ELF7 was studied comparatively between Input samples and AP eluates. When using antibodies directed against the CTD part of NRPB1 phosphorylated at the position Ser2 (α–CTD-S2P), a hypo- (NRPB1A) and a hyper-phosphorylated (NRPB1O) forms of the largest RNAPII subunit were detected in both Input and AP eluates of TFIIS and ELF7 (Figure 47). NRPB1A showed a slightly higher electrophoretic mobility than hyperphosphorylated NRPB1O as shown previously (Fontrodona et al., 2013). The analysis with non-phospho specific antibodies directed against RNAPII-CTD (α-CTD) revealed a single band likely corresponding to NRPB1A and/or unphosphorylated form present in both Input and AP eluates of TFIIS and ELF7. Western Blot analysis also revealed that both TFIIS and ELF7 predominantly associated with the elongating, hyper-phosphorylated form of RNAPII (RNAPII-S2P) (Figure 47, the ratios between “AP” and “Input”).

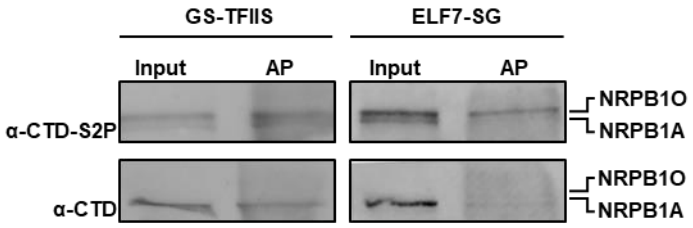


Figure 47. TFIIS and ELF7 associate with transcriptionally engaged RNAPII.
Immunoblot analysis for the AP eluates of GS-TFIIS and ELF7-SG affinity purifications with antibodies against the CTD part of NRPB1 phosphorylated at the position Ser2 (α-CTD-S2P) as well as non-phospho specific antibodies(α-CTD). NRPB1A indicates a hypo- and NRPB1O a hyper-phosphorylated form of NRPB1.

3.1.3 The direct interaction between *Arabidopsis* TFIIS and ELF7

Robust copurification of *Arabidopsis* TFIIS and ELF7 could imply their direct interaction within the TEC, similarly to observed in yeast where ELF7 was shown to residue in the proximity N-Terminal part of TFIIS within RNAPII-PAF1-C-TFIIS complex (Xu et al., 2017). Additionally, TFIIS was shown to directly interact with human Paf1 (ortholog of *Arabidopsis* ELF7) by using affinity purification approach (Kim et al., 2010).

To test protein-protein interactions (PPI) between *Arabidopsis* TFIIS and ELF7, the Matchmaker TM GAL4 yeast-two hybrid (Y2H) system (Clonetech) was applied. TFIIS and ELF7 coding sequences were introduced into pGADT7 and pGBKT7 and integrated with DNA-binding domain (DNA-BD, “bait”) and GAL4 activation domain (AD, “prey”), respectively. Vectors were further co-transfected into the auxotrophic yeast strain AH109 using the PEG/LiAc method. The transgenic yeast cells expressing both bait and prey plasmids were identified by selection on double dropout media (DDO) lacking leucine (-Leu) and tryptophan (-Trp). Double positive cells were next transferred and assessed on triple (TDO; -Leu/-Trp/-His) and quadruple (QDO; -Leu/-Trp/-His/-Ade) dropout plates showing no growth under applied condition and thus no interaction between *Arabidopsis* TFIIS and ELF7 in the studied system (Figure 48). Cells expressing DNA-BD/murine p53 and AD/SV40 large T-antigen served as a positive control exhibiting normal growth on DDO, TDO and QDO plates (Iwabuchi et al., 1993). Cells expressing DNA-BD/Lamin and AD/SV40 large T-antigen were used as negative control with no growth on TDO and QDO media. As an additional negative control, cells co-expressing DNA-BD/TFIIS with AD/ SV40 large T-antigen and DNA-BD/murine p53 with AD/ELF7 were studied showing no background detection on TDO nor QDO plates.

Noteworthy, the interaction between TFIIS and other five subunit of *Arabidopsis* PAF1-C identified in GS-TFIIS pulldown was studied using Y2H system by Irene Fuhrmann as a part of her bachelor projects in connection with this work. Similarly to findings presented in this chapter, no interaction could be detected between *Arabidopsis* TFIIS and other individual PAF1-C subunits.

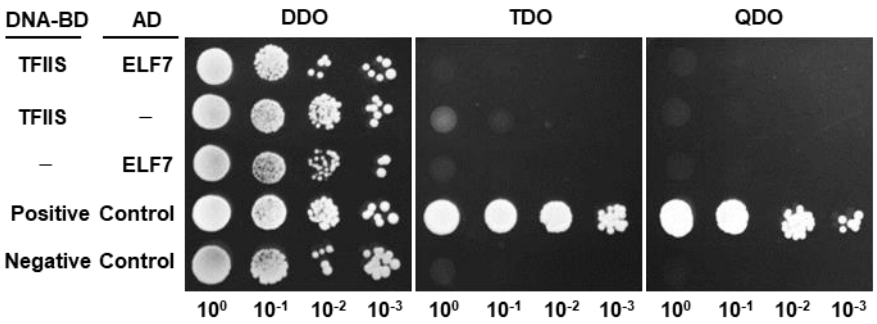


Figure 48. Y2H assay for studding direct interaction between *Arabidopsis* TFIIS and ELF7.
The growth of serially dilute yeast cells co-transformed with different combinations of DNA-BD/bait and AD/prey fusion proteins was evaluated on DDO (SD/ -LEU -TRP), TDO (SD/- LEU -TRP -HIS) and QDO (SD/ -LEU -TRP -HIS -ADE). Positive control: DNA-BD/murine p53 and AD/SV40 large T-antigen. Negative control: DNA-BD/Lamin and AD/ SV40 large T-antigen. DNA-BD: DNA binding domain. AD: activation domain.

In view of those observations, another approach for PPI determination *in vivo* was applied using Foerster Resonance Energy Transfer (FRET). eGFP and mCherry reporter genes were used as FRET partners since they have been characterised with good spectral separation as well as high overlap between donor (eGFP) emission and acceptor (mCherry) excitation (Förster Radius R0 = 5,1 nm) (Albertazzi et al., 2009; Tramier et al., 2006). As a positive control vector for the expression of eGFP-NLS-mCherry fusion protein driven by CaMV 35S promoter was used (provided by Alex Pfab).

The coding sequences of either TFIIS or ELF7 were fused with eGFP and mCherry, respectively and placed under the CaMV 35S promoter. Additionally, the interaction between eGFP-TFIIS and free mCherry fused with NLS was studied as a negative control (provided by Alex Pfab). Described combinations of fusion proteins were transiently co-expressed in *Nicotiana benthamiana* leaves by using *Agrobacterium*-mediated infiltration. Co-transformed epidermal cells were identified by confocal microscopy and subjected for FRET analysis. The efficiency of the energy transfer was measured using the Acceptor Photo-Bleaching method (Weidtkamp-Peters and Stahl, 2017).

Similarly to shown previously (Pfab et al., 2018a), eGFP physically linked to mCherry ensured close proximity of both fluorescent proteins reaching FRET efficiency of 18% (Figure 49 B). Next, the interaction between TFIIS and ELF7 was studied. In contrast to positive control, cells co-expressing eGFP-TFIIS and ELF7-mCherry showed no increase of the donor fluorescence after the acceptor bleaching (Figure 49 A-B). The mean FRET efficiency of eGFP-TFIIS/ELF7-mCherry pair was approximately 3% and was not statistically different from the negative control (eGFP-TFIIS/NLS-mCherry) indicating no interaction between *Arabidopsis* TFIIS and ELF7 in applied system.

The results generated with Y2H and FRET assays may suggest no direct interaction between *Arabidopsis* TFIIS and ELF7. However, obtained data could be a consequence of those methods limitation as well as reflect the high complexity and/or transiency of studied PPI.

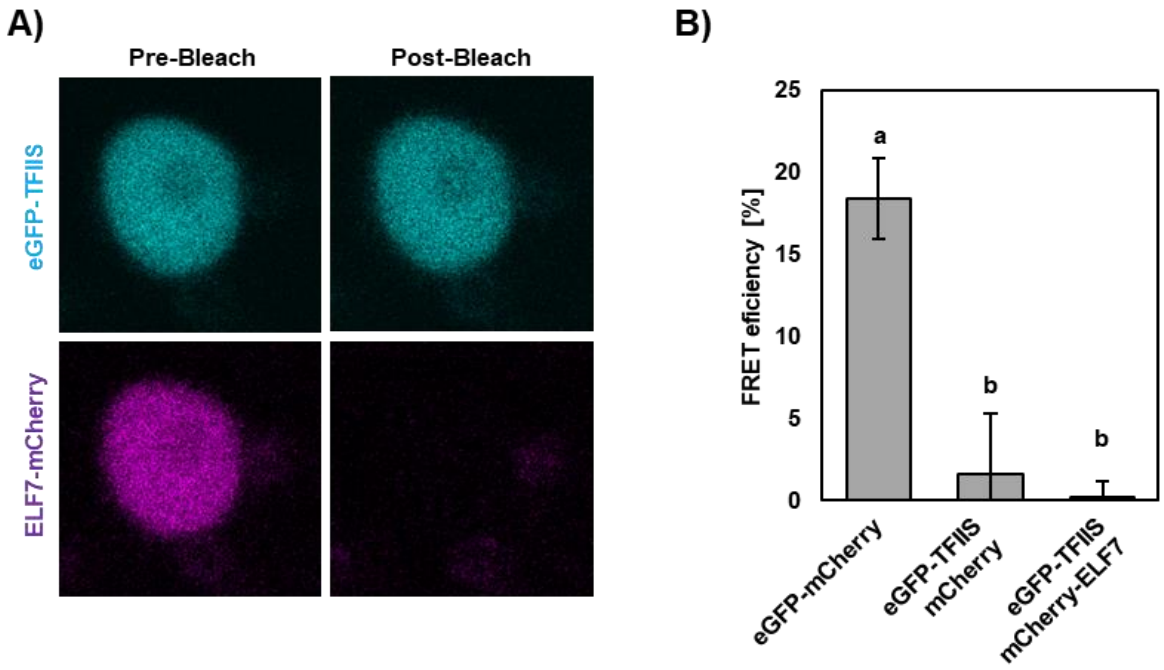


Figure 49. FRET assay for studding direct interaction between *Arabidopsis* TFIIS and ELF7.
(A) Co-expression of eGFP-TFIIS (cyan) and ELF7-mCheryy (magenta) in the epidermal cells of *N. benthamiana* leaves. Pictures were taken using confocal microscopy (CLMS) during FRET assay. "Pre-Bleach" and "Post-Bleach" indicate the time points immediately before and after mCherry (acceptor) bleaching. (B) The mean FRET efficiencies comprising the result of eight individual cells for each combination further analysed statistically using one-way ANOVA. Error bars reflect SD and letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05).

3.1.4 Full length vs truncated TFIIS interactome

In yeast N-terminal domain of TFIIS has been shown to play a role in PAF1-C recruitment to RNAPII (Xu et al., 2017), although this part of TFIIS remains unstructured in the crystallography experiments. Thus, it was expected that investigation of TFIIS interactome in comparison with its truncated counterpart could reveal the influence of N-terminal TFIIS on *Arabidopsis* TEC composition.

Truncated TFIIS (ΔTFIIS) lacking N-terminal domain I (Grasser et al., 2009) was created by PCR-based amplification of TFIIS coding sequence. ΔTFIIS was subsequently N-terminally fused with GS tag and placed under TFIIS native promoter. Since N-terminal part of TFIIS was suggested to play a role in nuclear targeting, additional NLS sequence was introduced between GS and ΔTFIIS (Figure 50 B). Indeed, in a pilot experiment ΔTFIIS lacking additional NLS sequence showed a certain degree of accumulation in the cytoplasm whereas NLS-ΔTFIIS localised specifically in the nucleus (Supplementary Figure S 14).

Next, AP-MS was performed as described in chapter 3.1.1 using GS-TFIIS and GS-NLS-ΔTFIIS transgenic PSB-D cells. Unfused-GS (GS) and GS fused with the NLS sequence (GS-NLS) under the control of CaMV 35S promoter were used as a control to identify proteins immunoprecipitated unspecifically with GS-TFIIS and GS-ΔTFIIS, respectively. The whole protein extracts from PSD-B cells expressing transgenic bait proteins were first resolved by SDS-PAGE. Coomassie staining revealed no obvious differences between analysed samples in

terms of global protein expression (Figure 50 C, left panel). Next, Coomassie staining of gel containing respective AP eluates revealed clear bands corresponding to GS-TFIIS (~ 70 kDa), GS-ΔTFIIS (~ 60 kDa), GS (~ 25 kDa) and GS-NLS (~ 30 kDa) (Figure 50 C, right panel). Additional bands detected in GS-TFIIS and GS-ΔTFIIS pulldown showed similar pattern although some minor differences could be seen. Interestingly, few additional faint bands could be seen above the bait protein in GS-NLS but not in GS pulldown.

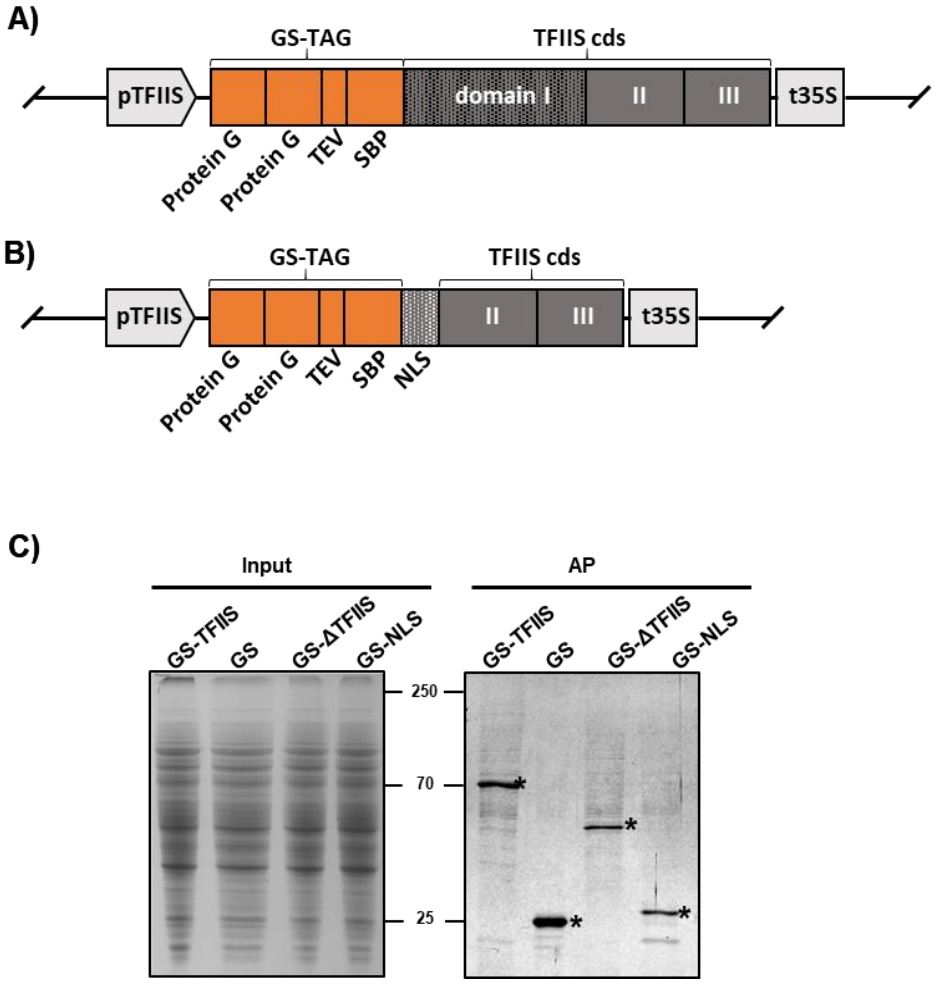


Figure 50. AP-MS approach performed with full length or truncated TFIIS as a bait protein.
Schematic illustration of TFIIS (A) and ΔTFIIS (B) bait proteins N-terminally fused with GS tag. pTFIIS: native promoter of TFIIS; t35S: terminator of 35S promoter; ProteinG: IgG-binding domain of protein G; TEV: tobacco etch virus (TEV) protease cleavage site; SBP, streptavidin-binding peptide. (C) Whole protein extracts of transgenic PSD-B cells (Input) harbouring GS-TFIIS, ΔTFIIS, GS-NLS or unfused GS-tag (Input) and the AP eluates of their affinity purifications (AP) were separated by 9 % SDS-PAGE and stained with Coomassie Blue. Black asterisks indicate the bands corresponding to bait proteins. Numbers indicate molecular weight in kDa.

Individual eluates were subjected for mass spectrometry and immunoprecipitated proteins were identified. Further data proceeding was performed to increase data robustness similarly to described in chapter 3.1.1 (Figure 51 A). Proceeded TFIIS and ΔTFIIS interactomes were compared using Venn diagram revealing a weak occupancy between them (~ 29%; Figure 51 B).

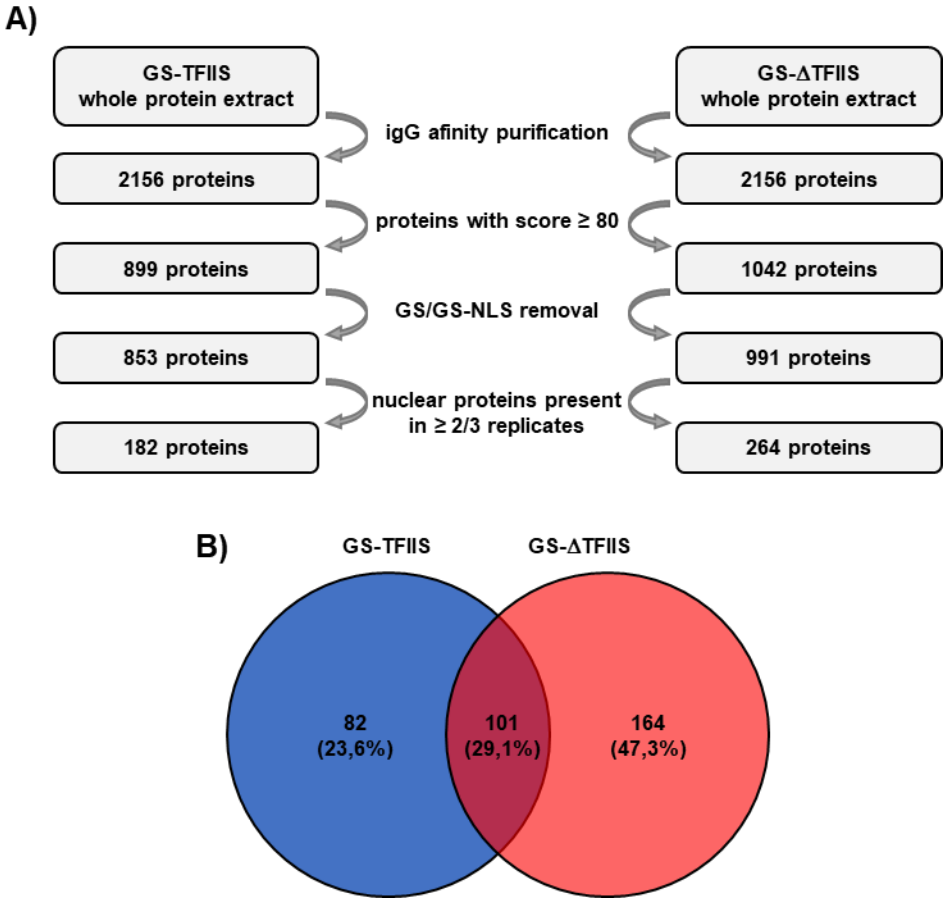


Figure 51. Data proceeding and initial analysis of TFIIS and ΔTFIIS interactomes following AP-MS approach. (A) The workflow of MS data proceeding following affinity purification with GS-TFIIS and GS-ΔTFIIS. (B) Overlap between proteins co-purified with TFIIS and ΔTFIIS depicted by Venn diagram.

To better understand the consequences of N-terminal TFIIS absence on TEC composition, transcription-related factors were comparatively analysed between GS-TFIIS and GS-ΔTFIIS interactomes. Similarly to GS-TFIIS interactome described before (3.1.1), the analysis of GS-ΔTFIIS pulldown revealed the presence of diverse RNAPII complex subunit as well as some PAF1-C subunits (Table 11). Noteworthy, ΔTFIIS interactome contained less subunit of both RNAPII and PAF1-C when compared with TFIIS interactome (5 vs 11 and 2 vs 6, respectively) and all present subunits had lower Mascot score. TFIIS and ΔTFIIS interactomes also shared some TEFs (SPT16, ELO3), whereas others copurified specifically with the full length (SPT6L, SPT5-2) or truncated (SSPR1) TFIIS.

TFIIS and ΔTFIIS pulldowns did not differ much in terms of H2A/H2B histone variants, NAPs or HATD. ΔTFIIS interactome contained some additional proteins including CAF1, MINU1, CHR17 as well as putative components of INO80 complex.

Table 11. The overlap between transcription-related proteins copurified with GS-TFIIS and GS-ΔTFIIS.
The comparison of transcription-related proteins immunoprecipitated with GS-TFIIS and GS-ΔTFIIS. The numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent APs.

GS-TFIIS	GS-ΔTFIIS	Interactor	Complex	Process	AGI
2576 / 3	661 / 3	AT2G38560	TFIIS	TFIIS	Transcription
2176 / 3	939 / 3	AT4G35800	NRPB1	Polymerase II	Transcription
838 / 3	837 / 3	AT4G21710	NRPB2	Polymerase II	Transcription
656 / 3	374 / 3	AT2G15430	NRP(B/D/E)3a	Polymerase II	Transcription
489 / 3	212 / 2	AT3G22320	NRP(A/B/C/D)5	Polymerase II	Transcription
471 / 3		AT5G59180	NRPB7	Polymerase II	Transcription
160 / 2		AT1G54250	NRP(A/B/C/D/E)8a	Polymerase II	Transcription
305 / 2		AT3G59600	NRP(A/B/C/D/E)8b	Polymerase II	Transcription
189 / 3		AT5G51940	NRP(B/C/D/E)6a	Polymerase II	Transcription
168 / 3	115 / 2	AT3G16980	NRP(B/D/E)9a	Polymerase II	Transcription
102 / 2		AT4G16265	NRP(B/D/E)9b	Polymerase II	Transcription
207 / 3		AT3G52090	NRP(B/D/E)11	Polymerase II	Transcription
1064 / 3	414 / 2	AT2G06210	CTR9, ELF8, VIP6	PAF-C	Transcription
1062 / 3	189 / 2	AT5G61150	LEO1, VIP4	PAF-C	Transcription
917 / 3		AT1G79730	PAF1, ELF7	PAF-C	Transcription
603 / 3		AT1G61040	RTF1, VIP5	PAF-C	Transcription
495 / 3		AT3G22590	CDC73	PAF-C	Transcription
422 / 3		AT4G29830	SKI8, VIP3	PAF-C	Transcription
155 / 2	1331 / 2	AT4G10710	SPT16	FACT	Transcription
	349 / 2	AT3G28730	SSRP1	FACT	Transcription
	265 / 2	AT2G23070	CKA4	CK2	Transcription
222 / 2	196 / 2	AT5G50320	ELP3; ELO3	Elongator	Transcription
863 / 3		AT1G65440	SPT6-1, SPT6L	SPT6	Transcription
225 / 3		AT4G08350	SPT5-2	SPT4/SPT5	Transcription
129 / 2	215 / 2	AT2G30620	Histone H1.2	Histone H1 family	Transcription
688 / 2	836 / 2	AT5G54640	H2A.6 (HTA1)	Histone H2A family	Transcription
498 / 2	486 / 2	AT1G54690	H2AXb (HTA3)	Histone H2A family	Transcription
342 / 2	292 / 2	AT1G52740	H2A.F/Z 3 (HTA9)	Histone H2A family	Transcription
275 / 2	505 / 2	AT5G59870	H2A.7 (HTA6)	Histone H2A family	Transcription
182 / 2	254 / 2	AT5G27670	H2A.5 (HTA7)	Histone H2A family	Transcription
	327 / 2	AT3G20670	H2A.2 (HTA13)	Histone H2A family	Transcription
1399 / 2	1357 / 2	AT3G53650	H2B.8 (HTB6)	Histone H2B family	Transcription
1119 / 2		At5g22880	H2B.10 (HTB2)	Histone H2B family	Transcription
	118 / 2	AT3G09480	H2B.5 (HTB7)	Histone H2B family	Transcription
	659 / 2	AT5G02570	H2B.9 (HTB10)	Histone H2B family	Transcription
602 / 2	715 / 3	AT2G19480	AtNAPL2	putative histone chaperones	Transcription
418 / 2	580 / 2	AT5G56950	AtNAPL3	putative histone chaperones	Transcription
	602 / 2	AT4G26110	AtNAPL1	putative histone chaperones	Transcription
	133 / 2	AT2G19540	AtCAF1CL6	putative histone chaperones	Transcription
	108 / 2	AT2G19520	AtCAF1CL4	putative histone chaperones	Transcription
1030 / 3	1366 / 2	AT5G03740	HDT3	Deacetylase	Transcription
557 / 2	435 / 2	AT2G27840	HDT4	Deacetylase	Transcription
491 / 3	1623 / 2	AT5G22650	HDT2	Deacetylase	Transcription

GS-TFIIS	GS-ΔTFIIS	Interactor	Complex	Process	AGI
	238 / 2	AT3G44750	HDT1	Deacetylase	Transcription
	154 / 2	AT3G06010	MINU1	SWI/SNF-type	Transcription
	674 / 2	AT5G18620	CHR17	ISWI	Transcription
	487 / 2	AT5G67630	RVB21	INO80	Transcription
	483 / 2	AT5G22330	RIN1	INO80	Transcription
	106 / 2	AT3G60830	ARP7	INO80	Transcription

3.1.5 PAF1-C in being depleted in ΔTFIIS interactome

Label-free quantitative mass spectrometry approach was subsequently applied to qualitatively compare TFIIS and ΔTFIIS interactome, particularly in terms of TEC composition. Remaining AP eluates of GS-TFIIS and GS-ΔTFIIS immunoprecipitation (Figure 50 C, right panel) were again subjected to mass spectrometry to identify peptides allowing quantitative comparison of immunoprecipitated proteins. The mass spectrometry was again conducted in the lab of Dr. Astrid Bruckmann and the label-free quantification was performed by Dr. Rasha ElBashir. As a result, the abundance of ~2000 proteins could be determined quantitatively between TFIIS and ΔTFIIS interactomes based on two biological replicates. Obtained dataset was further analysed in the context of *Arabidopsis* TEC to reveal the differences between TFIIS or ΔTFIIS interactomes. The relative abundance of RNAPII subunits together with TEFs identified as a part of *Arabidopsis* TEC (Antosz et al., 2017) was further analysed. Proteins abundance was compared relatively between TFIIS and ΔTFIIS interactomes (log_2 AP-MS) and plotted together with determined significance level (-10*log_{10}(p-value)). Resulting volcano plot distribution revealed clear depletion of five out of six PAF1-C subunits in ΔTFIIS pulldown relatively to other TEC components in comparison to TFIIS interactome (Figure 52). VIP3 was the only subunit of PAF1-C not depleted in ΔTFIIS pulldown. These findings indicate that the

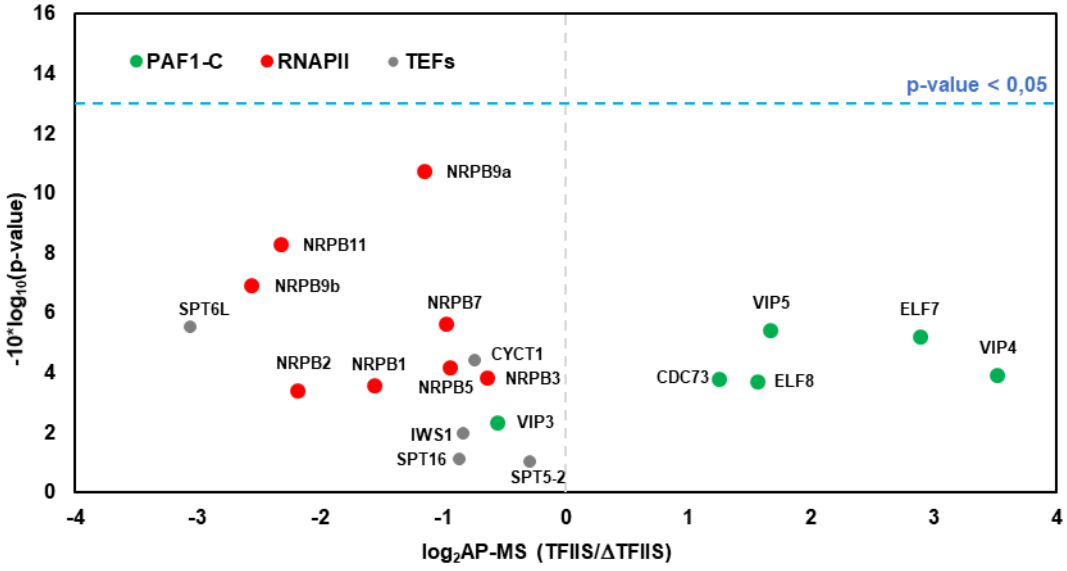


Figure 52. The relative composition of *Arabidopsis* TEC copurified with TFIIS or ΔTFIIS.
The abundance of *Arabidopsis* TEC components (Antosz et al 2017) was compared relatively between TFIIS and ΔTFIIS pulldown analysed by label-free quantitative mass spectrometry. Relative abundance (log_2AP-MS) was plotted together with determined significance level (-10*log_{10}(p-value). Blue line indicates the bottom baseline for statistical significance (p-value < 0.05).

N-terminal part of *Arabidopsis* TFIIS plays a role in PAF1-C recruitment to RNAPII and/or enhances its affinity to RNAPII, similarly to previously suggested in yeast (Xu et al., 2017).

3.1.6 TFIIS and ΔTFIIS associate with differentially phosphorylated NRPB1

N-terminal part of *Arabidopsis* TFIIS was shown to be redundant for RNAPII binding in the previous chapter in line with the data obtained in yeast (Awrey et al., 1998). To better understand its role in regulating transcript elongation, NRPB1 immunoprecipitated with either full length or truncated TFIIS was examined in terms of CTD phosphorylation status. Western blot analysis was performed using α-CTD-S2P, α-CTD-S5P as well as α-CTD antibodies described in chapter 3.1.2. When antibodies directed against phosphorylated RNAPII-CTD were applied, NRPB1A and NRPB1O forms of the largest RNA polymerase II subunit could be detected in the Input samples of both TFIIS and ΔTFIIS (Figure 53, middle and bottom panel). An additional band could be observed above NRPB1O in the TFIIS AP eluate which could reflect another state of hyperphosphorylated NRPB1 (NRPB1O'). This band was absent in ΔTFIIS AP where only a single band of hyperphosphorylated NRPB1O could be seen after blotting the membrane with α-CTD-S2P or α-CTD-S5P. When applying the non-phospho specific antibodies (α-CTD) two bands could be detected in ΔTFIIS pulldown unlike in TFIIS AP eluate sample nor in Input samples where only a single lower band was observed (Figure 53, top panels).

Western blot analysis revealed that TFIIS and ΔTFIIS differ in terms of association with various forms of NRPB1. The identification of NRPB1O specifically in ΔTFIIS AP eluate may imply its delayed recruitment during transcription cycle and thus preferential association with active RNAPII. This interpretation is in line with the role of N-terminal part of TFIIS in transcription initiation (Kim et al., 2007). Different phosphorylation patterns of RNAPII-CTD in TFIIS and ΔTFIIS pulldowns could additionally indicate the importance of N-terminal TFIIS in the progression through the transcription cycle accompanied with the modifications in RNAPII-CTD phosphorylation pattern.

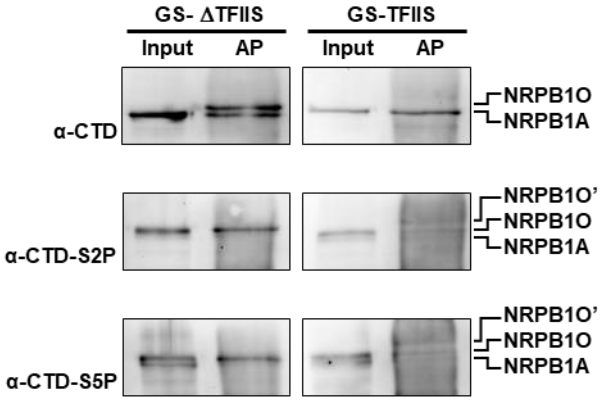


Figure 53. The association of TFIIS and ΔTFIIS with differentially phosphorylated NRPB1.
Immunoblot analysis of GS-TFIIS and GS-ΔTFIIS AP eluates with the antibodies against the CTD part of NRPB1 phosphorylated at Ser2 (α-CTD-S2P) and Ser5 position (α-CTD-S5P) as well as non-phospho specific antibodies (α–CTD). NRPB1A indicates a hypo- and NRPB1O as well as NRPB1O' a hyper-phosphorylated form of NRPB1.

3.1.7 Double mutants deficient in TFIIS and ELF7 show synergistic growth defects

In view of efficient copurification of TFIIS and PAF1-C, the double mutant deficient in TFIIS and ELF7 was further studied in detail. *tfIIs elf7* double mutant was generated by crossing *tfIIs-1* with *elf7-3* by former PhD student Simon Arnold Mortensen. The phenotypic analysis of *tfIIs elf7* revealed a severe growth defects in comparison with respective single mutants and Col-0. While *tfIIs-1* appearance was comparable to Col-0 at 21DAS, *elf7-3* showed compromised growth which was further reduced in *tfIIs elf7* (Figure 54 A). Observed growth differences were even more extreme at 42DAS (Figure 54 B). The quantitative analysis of various developmental traits confirms observed phenotypes revealing smaller rosette diameter and early bolting of *elf7-3* in comparison with both *tfIIs-1* and Col-0 (Figure 55 A-B) and similarly to reported previously (He et al., 2004). The bolting time of *tfIIs elf7* was comparable to *elf7-3*, however the rosette diameter was synergistically reduced in

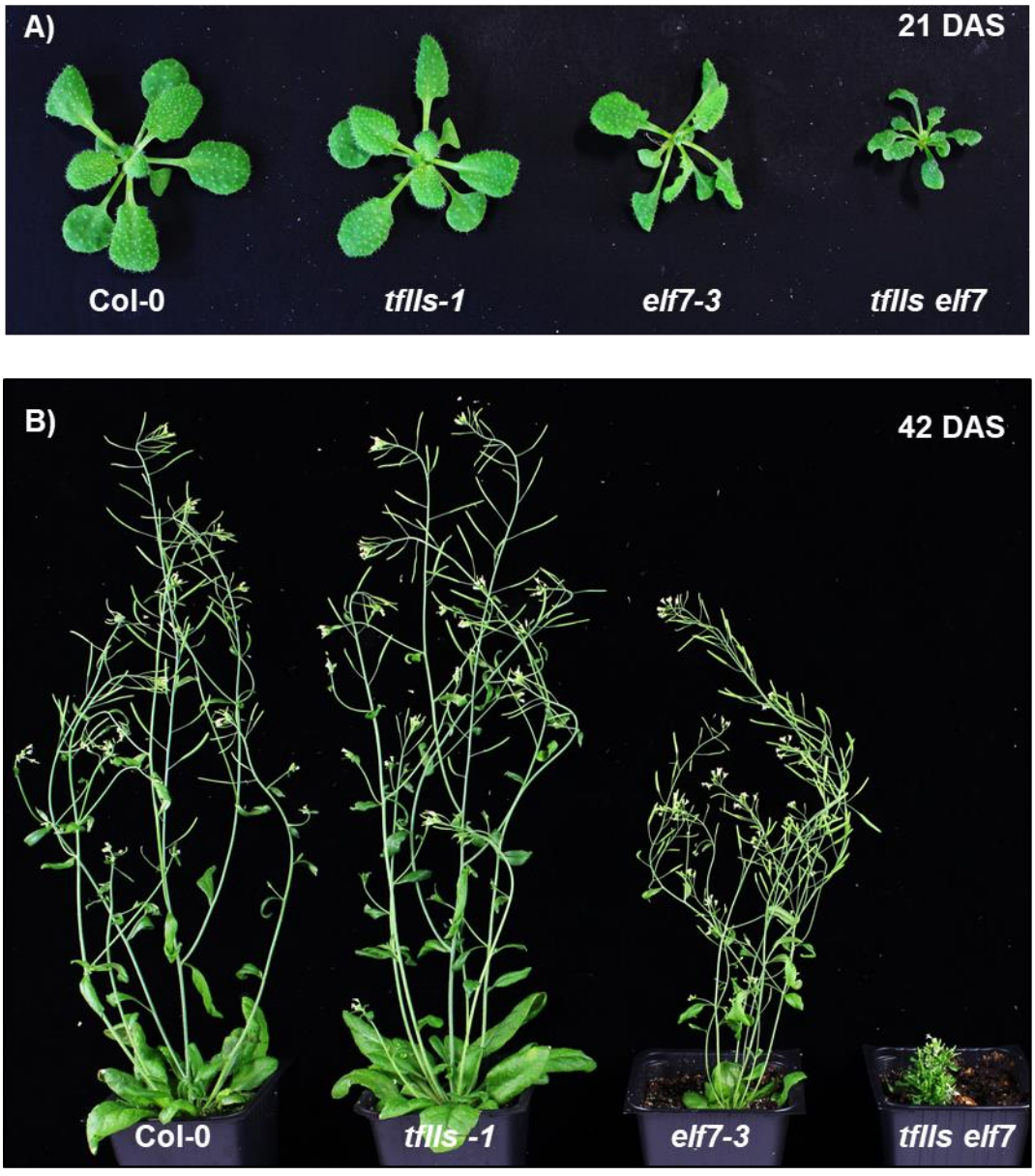


Figure 54. Phenotypic analysis of plants lacking functional TFIIS and/or ELF7.
tfIIs elf7 plants were phenotypically analysed in comparison with Col-0 and respective single mutants. Pictures were taken at 21DAS (A) and 42 DAS (B). Representative individuals of each genotype are shown.

comparison with Col-0 and respective single mutants (Figure 55 A-B). At 42DAS, *tfIIs elf7* was synergistically affected in terms of both plant height and number of primary inflorescences relatively to Col-0 and respective single mutants (Figure 55 C-D). Additionally, double mutant was sterile whereas *elf7-3* showed a reduced seed set and seed production in *tfIIs-1* was comparable to that of Col-0. To conclude, the examination of *tfIIs elf7* double mutants revealed distinct genetic interactions between the genes encoding *TFIIS* and *ELF7*.

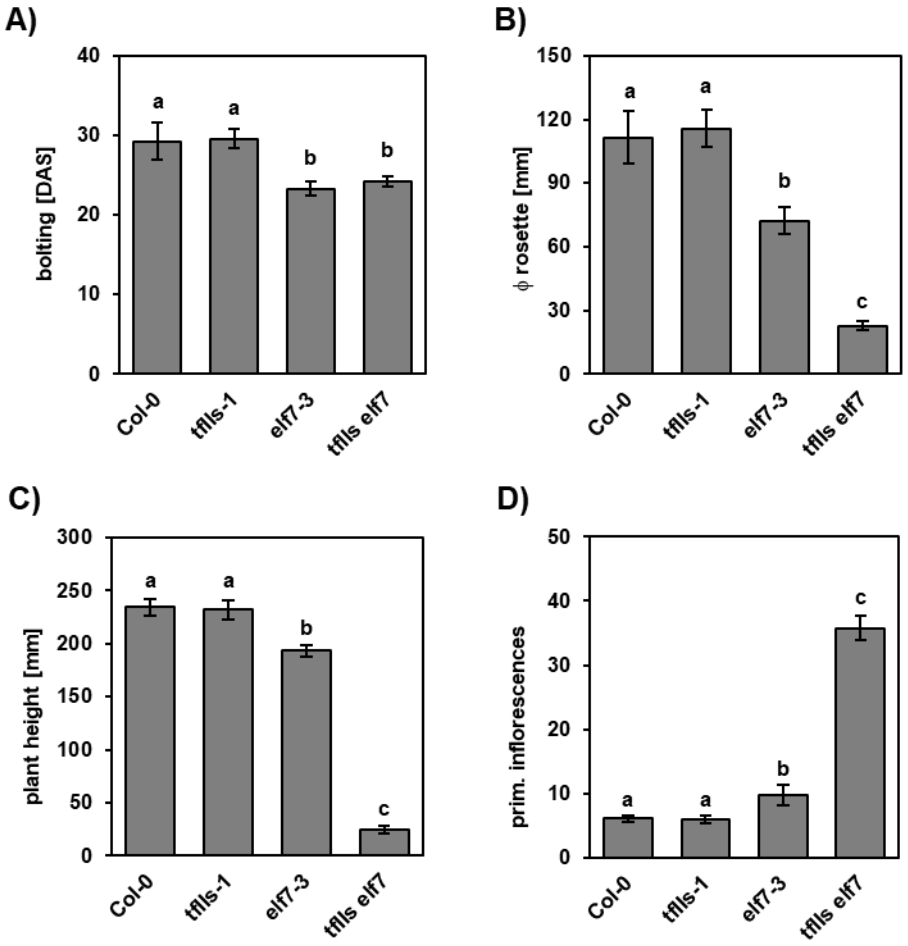


Figure 55. The quantification of developmental traits affected upon TFIIS and/or ELF7 deficiency.
tfIIs elf7 plants were phenotypically analysed in comparison with Col-0 and respective single mutants in terms of (A) bolting time (A), rosette diameter (B) plant height (C) and the number of primary inflorescences (D). Data comprising the measurements of 13 individual plants were next analysed by two-way ANOVA. Error bars reflect SD and the letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05).

3.1.8 The lack of functional TFIIS and ELF7 results in broad transcriptomic changes

To better understand the genetic interaction between the genes encoding *TFIIS* and *ELF7*, RNA-seq was performed for *tfIIs elf7* in comparison to Col-0 and respective single mutants. To this end, total RNA was isolated from aerial parts of 21DAS seedling (Figure 54 A) using RNeasy R Mini Plant kit (Qiagen). cDNA libraries preparation and sequencing were performed by Kompetenzzentrum Fluoreszente Bioanalytik (KFB; Regensburg). cDNA libraries were created using TruSeq Stranded mRNA Sample Preparation kit (Illumina). Three biological replicates for each line/condition were sequenced on Illumina HiSeq 1000 yielding a total numer of ~28-40 mln high quality reads for each replicate of analysed genotypes. The initial

proceeding and analysis of sequencing data was performed by Dr. Kevin Begcy. Sequencing reads were aligned onto the *Arabidopsis* reference transcriptome assembly (TAIR10). Next, Transcripts Per Kilobase Million (TPM) was calculated for all mappable loci and differentially expressed genes (DEGs) were determined. Hundreds of DEGs were identified when comparing individual genotypes (Figure 56). 398 upregulated and 991 genes downregulated were determined in *tfIIs elf7* double mutant relatively to Col-0. Significantly downregulated genes were further subjected to hierarchical clustering to unravel the correlations between genes downregulation in *tfIIs elf7* relatively to Col-0 and respective single mutants. As a result,

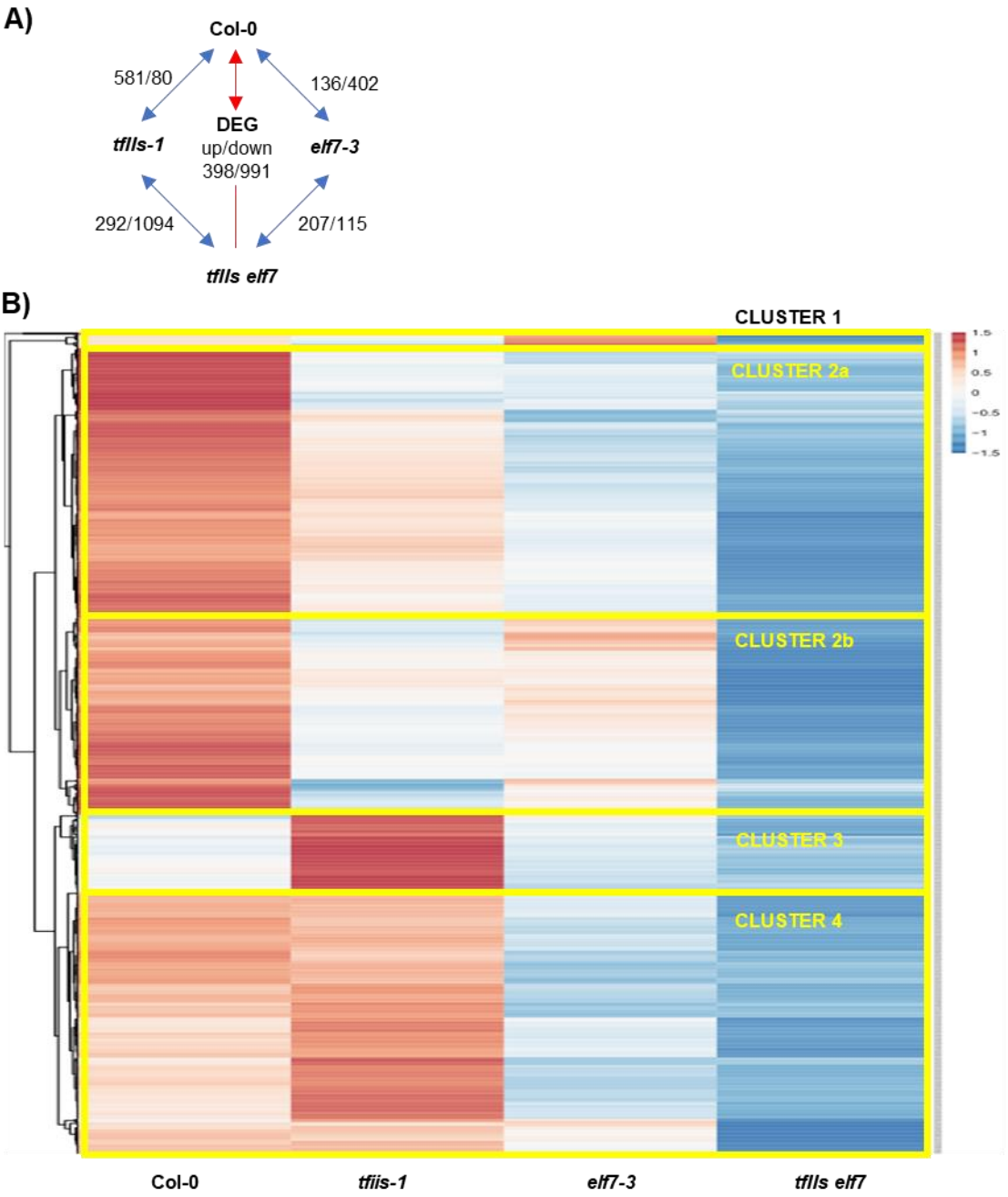


Figure 56. The analysis of transcriptomic changes upon TFIIS and/or ELF7 deficiency.
(A) Differentially expressed genes (DEGs; $log_2FC > 1$, p-value < 0.05, FDR < 0.05) identified between analysed lines. (B) Heatmap created for the 991 DEGs downregulated in *tfIIs elf7* in comparison with Col-0. Yellow boxes indicated the clusters manually defined based on the hierarchical clustering.

the synergy for many genes was observed in line with the phenotypic analysis. As seen on the Figure 56 B, certain subgroups of genes could be determined based on the hierarchical clustering and the relative expression level in Col-0, giving a rise to CLUSTER 1: genes downregulated in *tfIIs-1* and upregulated in *elf7-3* (~ 1,5% of all 991 genes), CLUSTER 2: genes downregulated in both *tfIIs-1* and *elf7-3* (~ 57%), CLUSTER 3: genes upregulated in *tfIIs-1* and unchanged in *elf7-3* (~ 9,4%), CLUSTER 4: genes unchanged in *tfIIs-1* and downregulated in *elf7-3* (~ 32%). CLUSTER 2 was additionally subdivided into CLUSTER 2a: genes with higher expression in *tfIIs-1* then *elf7-3* and CLUTER2b: genes with higher expression in *elf7-3* then *tfIIs-1*.

Identified clusters were further subjected to GO enrichment analysis to reveal the biological processes among the genes missregulated upon mutual TFIIS and ELF7 deficiency. Genes in the smallest clusters categorised into "response to stimuli" (CLUSTER 1 and 3) and "metabolic processes" (CLUSTER 1) terms (Table 12). In the biggest CLUSTER 2 only three enriched GO terms were identified, namely "lipid localization", "photosynthesis" and "heterocycle biosynthetic process". Many GO terms were enriched in CLUSTER 4 comprising various response processes and regulation of post-embryonic development in line with ELF7 role in the regulation of plant flowering.

Table 12. GO terms enriched among DEGs downregulated in *tfIIs elf7*.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO for the subgroups of genes identified by hierarchical clustering of DEGs downregulated in *tfIIs elf7* in comparison with Col-0 (Figure 56).

		log10(FRD)			
GO term	**description**	**CLUSTER 1**	**CLUSTER 2**	**CLUSTER 3**	**CLUSTER 4**
GO:0050896	response to stimulus	-1,33		-1,40	-4,34
GO:0008152	metabolic process	-1,33			
GO:0010876	lipid localization		-2,60		
GO:0018130	heterocycle biosynthetic process		-2,60		
GO:0015979	photosynthesis		-1,74		
GO:0009409	response to cold				-5,68
GO:0009628	response to abiotic stimulus				-4,89
GO:0006950	response to stress				-3,31
GO:0042221	response to chemical				-3,28
GO:0048580	regulation of post-embryonic development				-1,85
GO:0009719	response to endogenous stimulus				-1,62
GO:0010033	response to organic substance				-1,54

Since CLUSTER 2 was highly heterogenous regarding observed changes in gene expression, it was further sub-divided into two separated subclusters (CLUSTER 2a and CLUSTER 2b). GO enrichment analysis within determined subclusters revealed that lipid localization and photosynthesis related genes were enriched within CLUSTER 2b (Table 13). These genes showed somewhat stronger downregulation in *tfIIs-1* than in *elf7-3* in comparison to Col-0 and their expression is being further significantly downregulated in the double mutant. Intriguingly, "lipid localization" and "photosynthesis" GO terms were previously identified as a highly enriched among genes significantly downregulated upon TFIISmut expression.

It additionally suggests presumable importance of TFIIS in the regulation of those biological processes and PAF1-C could play a role in counteracting the transcriptional consequences of TFIIS absence.

Table 13. GO terms enriched among DEGs downregulated in *tfIIs elf7* in CLUSTER 2.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO for the genes in the CLUSTER2 further subdivided into additional clusters based on the transcriptomic changes in *tfIIs-1* and *elf7-3* (Figure 56).

		log10(FRD)	
GO term	**description**	**CLUSTER 2a**	**CLUSTER 2b**
GO:0009628	response to abiotic stimulus	-1,70	
GO:0010876	lipid localization		-3,34
GO:0015979	photosynthesis		-2,54

3.1.9 Truncated TFIIS fails to rescue *tfIIs elf7* phenotype

Considering the role of N-terminal part of TFIIS in the regulation of PAF1-C level within the TEC, observed genetic interaction was hypothesis to be the consequence of TFIIS-dependent PAF1-C recruitment to TEC. It was therefore attempted to express either full length or truncated TFIIS in *tfIIs elf7* to assess their ability to restore *elf7-3* phenotype. To this end TFIIS and ΔTFIIS were N-terminally fused with GFP and placed under TFIIS native promoter (Dolata et al., 2015) (Figure 57 A-B). Since *tfIIs elf7* plants produce no seeds, created vectors were introduced into *tfIIs-1*$^{-/-}$ *elf7-3*$^{-/+}$ genomic background by *Agrobacterium*-mediated transformation. Three individual transgenic lines carrying either GFP-TFIIS or GFP-ΔTFIIS transgene were selected and further analysed.

The progeny of plants homozygous for either GFP-TFIIS or GFP-ΔTFIIS transgene in *tfIIs-1*$^{-/-}$ *elf7-3*$^{-/+}$ background was next subjected to segregation to obtain the individuals lacking functional ELF7. Accordingly, transgenic lines carrying either GFP-TFIIS or GFP-ΔTFIIS transgene in *tfIIs-1*$^{-/-}$ *elf7-3*$^{-/-}$ background (referred as “TFIIS#1-3” and “ΔTFIIS#1-3”, respectively) were identified by PCR-based genotyping. Transgenic lines showed comparable expression level of transgenic TFIIS and ΔTFIIS relatively to TFIIS expression level observed in Col-0 (Figure 57 C). Despite the efforts no GFP signal could be detected in selected transgenic lines neither by confocal microscopy nor by western blot likely due to insufficient sensitivity of these methods (data not shown).

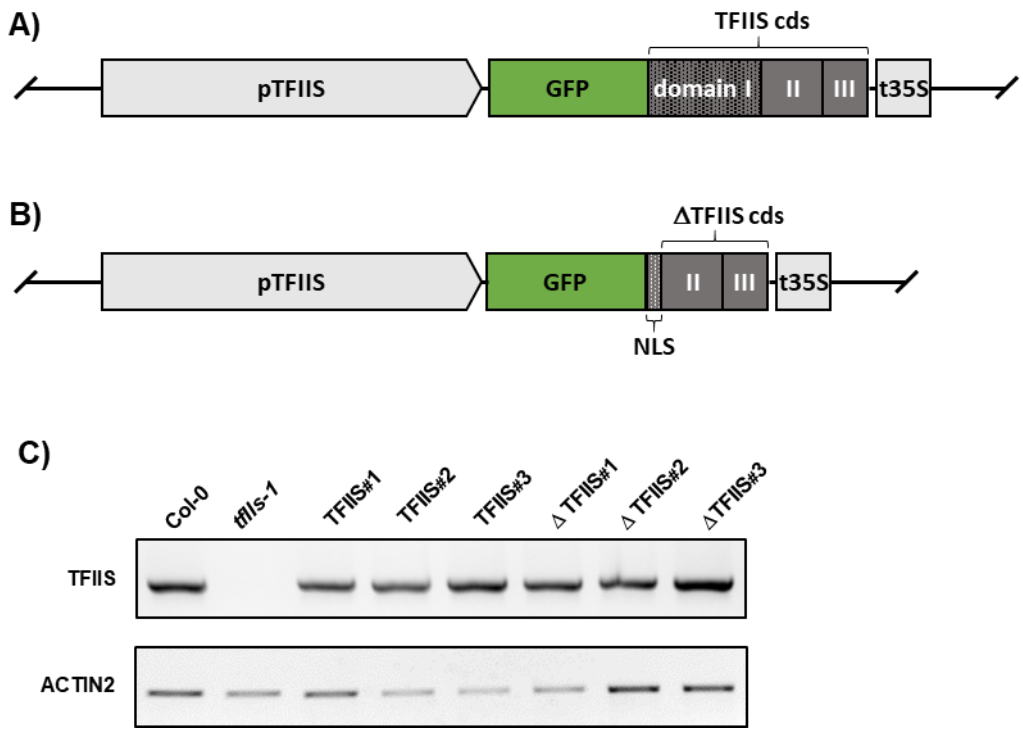


Figure 57. Transgenic TFIIS and ΔTFIIS design and expression level in *tfIIs elf7*.
Schematic illustration of GFP-TFIIS (A) and GFP-ΔTFIIS (B) transgenes integrated into *tfIIs elf7* genome. pTFIIS: native promoter of TFIIS; t35S: CaMV 35S terminator; NLS: nuclear localization sequence. (C) TFIIS expression level was examined by RT-PCR in 10DAS Col-0, *tfIIs-1*, TFIIS#1-3 and ΔTFIIS#1-3 seedling grown on MS medium. ACTIN2 was used as a reference gene.

Six individual plants of each line were analysed phenotypically in comparison to Col-0, *tfIIs elf7* and respective single mutants. Severe developmental defects of *tfIIs elf7* double mutant were previously shown to accumulate at 42DAS (chapter 3.1.7), thus all lines were also assessed comparatively at that timepoint. Phenotypic analysis revealed comparable plants height and number of secondary inflorescences between *elf7-3* and individual TFIIS#1-3 lines whereas those developmental traits were significantly different from observed for Col-0, *tfIIs-1* and *tfIIs elf7* (Figure 58 and Figure 59). In contrary, ΔTFIIS#1-3 lines showed significantly reduced plant height and increased number of primary inflorescences relatively to Col-0, *tfIIs-1* and *elf7-3*. However, in comparison with *tfIIs elf7,* ΔTFIIS#1-3 lines had significantly increased plant height and decreased number of primary inflorescences. Importantly, plants expressing GFP-TFIIS were able to produce a small seed set similarly to that of *elf7-3* whereas no seeds could be obtained from ΔTFIIS#1-3 lines. Taken together, the constitutive expression of full length TFIIS in *tfIIs elf7* could restore *elf7-3* phonotype whereas the presence of truncated TFIIS resulted in only partial complementation.

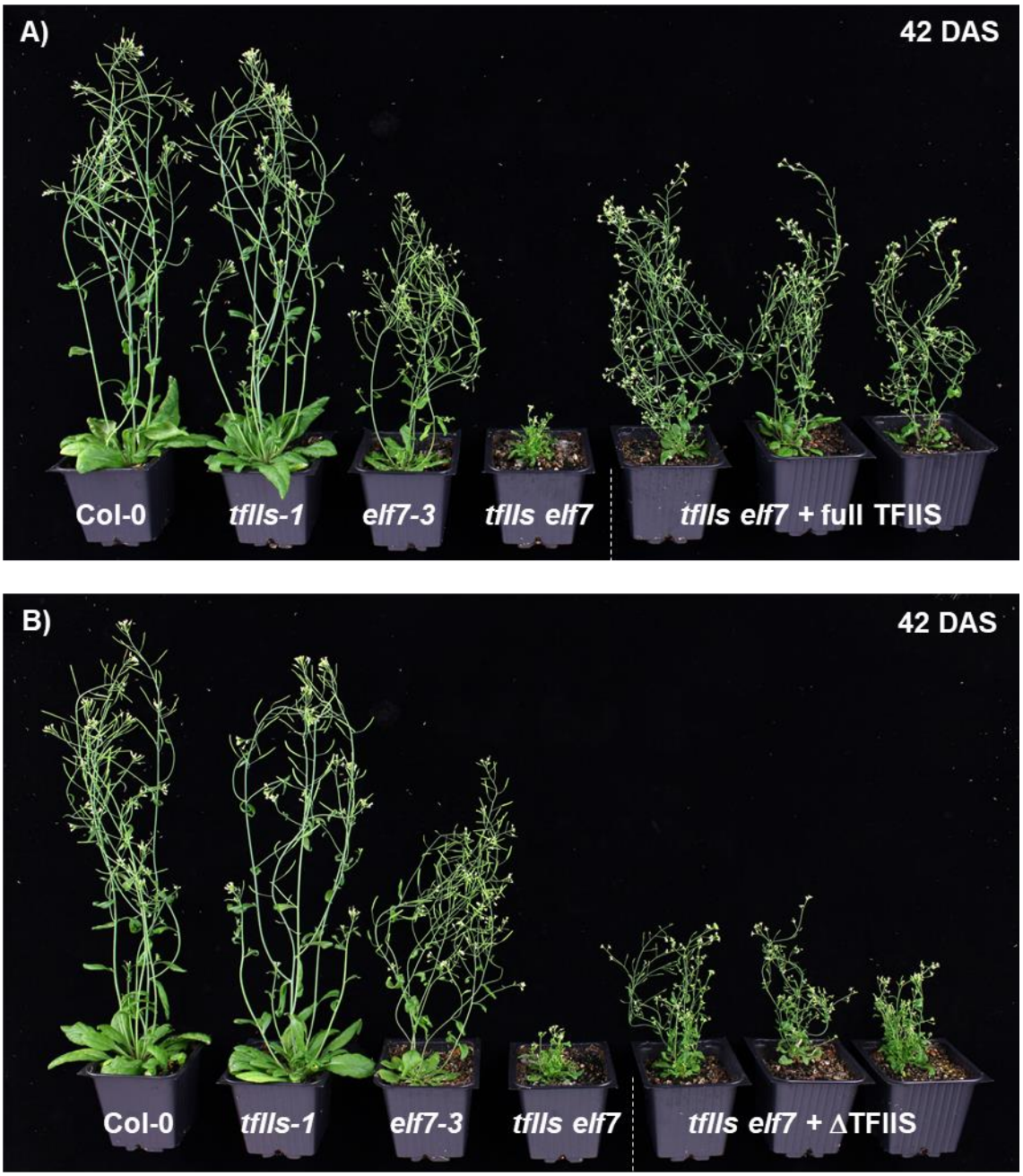


Figure 58. Phenotypic analysis of *tfIIs elf7* harbouring full length or truncated TFIIS transgene.
tfIIs elf7 plants were phenotypically analysed in comparison with Col-0, respective single mutants and three individual transgenic lines harbouring either full length (TFIIS#1-3; A) or truncated (ΔTFIIS#1-3; B) TFIIS in *tfIIs elf7* background. Pictures were taken at 42DAS.

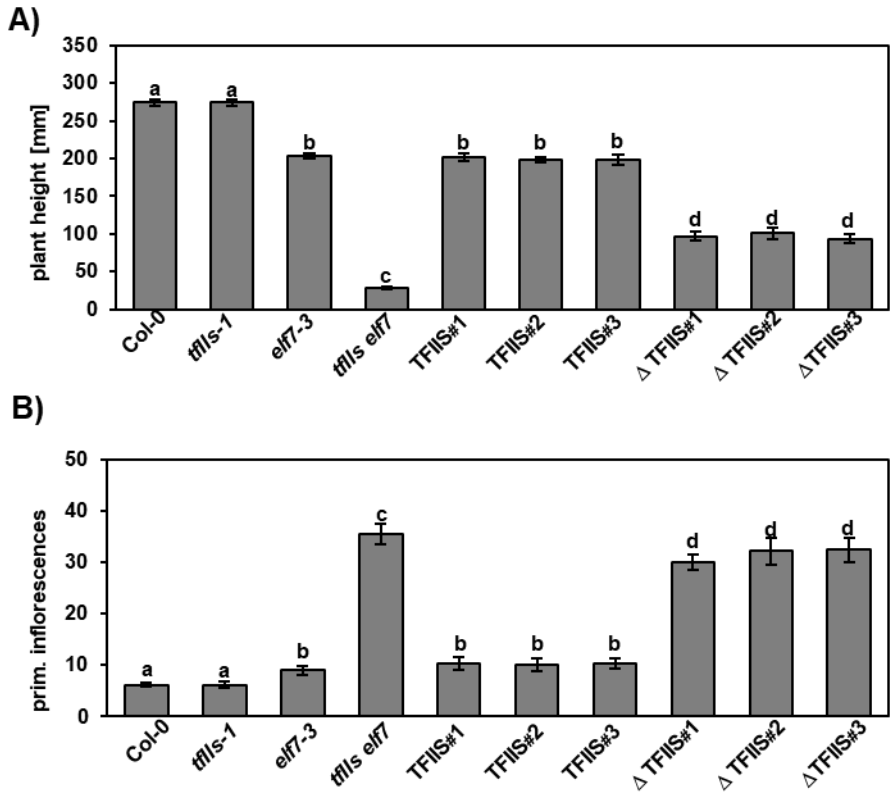


Figure 59. The quantification of developmental traits restored upon full length or truncated TFIIS expression in *tfIIs elf7.*

TFIIS#1-3 and ΔTFIIS#1-3 complementation lines were phenotypically analysed in comparison with Col-0, *tfIIs elf7* and respective single mutants in terms of plant height (A) and the number of primary inflorescences (B). Data comprising the measurements of eight individual plants were next statistically analysed by one-way ANOVA. Error bars reflect SD and the letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05).

3.2 The role of *Arabidopsis* ELF7 is resolving transcription-replication conflict

PAF1-C has been characterised with broad range of molecular functions related to transcription (Tomson and Arndt, 2013). Additionally, yeast Paf1 (ortholog of *Arabidopsis* ELF7) was recently reported to play a role in the resolution of transcription-replication conflict (TRC) (Poli et al., 2016). Interestingly, GreA/B, the bacterial counterparts of TFIIS, has been shown to prevent the conflict between DNA replication and transcription machinery (Tehranchi et al., 2010) and similar role was suggested for yeast TFIIS (Dutta et al., 2015). Therefore, the role of *Arabidopsis* ELF7 was analysed in the context and DNA replication as well as TRC resolution to better understand the genetic interaction between TFIIS and ELF7.

3.2.1 *Arabidopsis* PAF1-C may association with replication machinery and INO80

To shed light on PAF1-C involvement into the cellular processes related to DNA replication in *Arabidopsis*, GO enrichment analysis was performed among proteins immunoprecipitated with ELF7 from PSB-D *Arabidopsis* cell culture (performed by Hans Ehrnsberger; Antosz et al., 2017). Out of many enriched biological processes (Supplementary Table 15) several GO terms related to DNA replication and cell division could be identified (Table 14). Additionally, GO terms related to DNA damage and DNA repair were enriched implying the involvement of *Arabidopsis* PAF1-C into the regulation of those processes.

Table 14. GO terms presumably related to DNA replication enriched among proteins copurified with ELF7.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO for the proteins copurified with ELF7-SG. Overrepresented GO terms (FDR < 0.05) presumably related and/or influenced by DNA replication are listed below. All overrepresented GO terms can be found in Supplementary Table 15. ELF7-SG affinity purification was performed by Hans Ehrnsberger.

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	pvalue	FDR
GO:0006268	DNA unwinding during replication	5 / 535	11 / 37767	2,00E-06	6,10E-05
GO:0032392	DNA geometric change	5 / 535	16 / 37767	8,80E-06	2,20E-04
GO:0032508	DNA duplex unwinding	5 / 535	16 / 37767	8,80E-06	2,20E-04
GO:0033205	cytokinesis during cell cycle	5 / 535	26 / 37767	6,60E-05	1,40E-03
GO:0051301	cell division	8 / 535	99 / 37767	1,30E-04	2,70E-03
GO:0006261	DNA-dependent DNA replication	5 / 535	60 / 37767	2,20E-03	3,00E-02
GO:0006281	DNA repair	9 / 535	214 / 37767	4,40E-03	5,70E-02
GO:0006974	response to DNA damage stimulus	9 / 535	221 / 37767	5,30E-03	6,60E-02
GO:0022402	cell cycle process	7 / 535	149 / 37767	6,50E-03	8,00E-02
GO:0006260	DNA replication	6 / 535	117 / 37767	7,80E-03	8,90E-02

3.2.2 Plants lacking functional ELF7 exhibit elevated level of homologous recombination

Yeast PAF1 has been demonstrated to resolve TRCs in the cooperation with Mec1 and INO80 complexes (Poli et al., 2016). Interestingly, several putative components of *Arabidopsis* INO80 copurified with ELF7 (Table 15) while no components of *Arabidopsis* Mec1 complex could be identified.

Table 15. Putative components of *Arabidopsis* INO80 complex identified among ELF7 interactome.
ELF7-SG interactome was obtained from (Antosz et al., 2017). The numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent APs. ELF7-SG affinity purification was performed by Hans Ehrnsberger.

ELF7-SG	Interactor	Complex	Process	AGI
543 / 3	RIN1	SWR1/NuA4, INO80	Transcription	AT5G22330
516 / 3	RVB21	SWR1/NuA4, INO80	Transcription	AT5G67630
263 / 3	AtARP4	SWR1/NuA4, INO80	Transcription	AT1G18450
201,8 / 2	ARP7	INO80	Transcription	AT3G60830
200 / 2	INO80	INO80	Transcription	AT3G57300

Arabidopsis INO80 protein (AT3G57300) is the chromatin remodeling factor which has been previously shown to control the somatic homologues recombination (SHR) and genome stability maintenance (Fritsch et al., 2004; Zhang et al., 2015a), similarly to its mammalian and yeast orthologs (Lademann et al., 2017; Wu et al., 2007). INO80 protein was identified in ELF7 pulldown but it was not found across the interactomes of other *Arabidopsis* TEFs (Antosz et al., 2017). In view of that, it was hypothesised that *Arabidopsis* PAF1-C may be involved into the regulation of homologous recombination in cooperation in INO80. To test this possibility, plants lacking functional ELF7 were crossed with the DGU.US-8 reported line allowing the determination of HR evens in somatic cells by monitoring the restoration of the β-glucuronidase (GUS) marker gene upon somatic homologous recombination (SHR) repair mechanism triggered by DNA double-strand break (DSB) (Orel et al., 2003). Accordingly, further histochemical analysis of whole leaves would allow the determination of recombination events by scoring the number of GUS positive areas (sectors).

After crossing *elf7-3* with DGU.US-8, their progeny was analysed by PCR-based genotyping to identify lines homozygous for SHR reporter transgene. Despite the efforts no plants homozygous for DGU.US transgene could be identified in *elf7-3* background. Thus, the analysis was performed for the Col-0 and *elf7-3* mutant carrying single copy of DGU.US. In order to detect SHR events, 7DAS seedlings grown on the MS medium were transferred into the GUS stanning solution for 24h, washed with the clearing solution and the picture of individual leaves were taken. For the quantitative analysis, Sectors were scored in one cotyledon and one first leaf from three individual plants for each genotype. As seen on the Figure 60 A-B, the number of SHR was elevated in *elf7-3* relatively to Col-0. Additionally, SHR events were analysed in *tfIIs-1* showing no increase in comparison to Col-0.

These data imply that the molecular prevention of DNA damage resolved by SHR may involve *Arabidopsis* PAF1-C and it is independent from the absence of functional TFIIS.

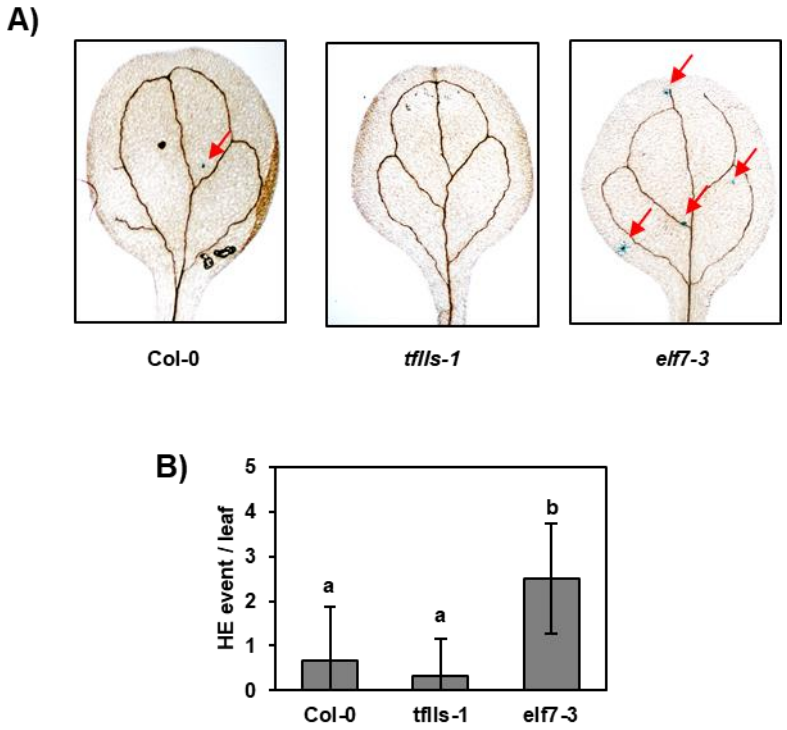


Figure 60. ELF7 deficiency results in elevated homologous recombination (HR).
HR events visualised by GUS staining (sectors) in Col-0, *tfIIs-1* and *elf-7*. One cotyledon and one first leaf were analysed in three independent 8DAS plants of each line grown on MS media. (A) Representative cotyledons are shown. Red arrows indicate identified sectors. Size bars indicate 1 mm. (B) Quantification of sector per leaf. Collected data was analysed by one-way ANOVA. Error bars reflect SD and the letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05).

3.2.3 ELF7 plays a role in the response to replication stress

In view of increased SHR in the absence of functional ELF7, its importance was further studied in the context of DNA replication stress. To this end, *elf7-3* growth was evaluated upon Hydroxyurea (HU) treatment in comparison to Col-0, to determine the morphological response to chemically induced replication stress (Schuermann et al., 2009) upon ELF7 deficiency. Additionally, the growth of *tfIIs-1* and *tfIIs elf7* double mutant was analysed to assess the effect of additional transcriptional stress in the absence of TFIIS. Individual plants of all genotypes were first grown vertically on the MS medium lacking HU. 5DAS plants were next transferred on the media containing either 2 µM HU or a mock treatment and their growth was documented at 10DAS. As seen on the Figure 61, *elf7-3* growth was strongly inhibited in the presence of HU in comparison with both Col-0 and *tfIIs-1*. Interestingly, the growth of *tfIIs elf7* was further inhibited showing the synergy between transcriptional stress in *tfIIs-1* and HU induced replication stress in the absence of functional ELF7, as determined by two-way ANOVA (Figure 61 C). No growth retardation of *tfIIs-1* in comparison to Col-0 suggests that TFIIS-mediated PAF1-C recruitment does not play a role in the context of DNA replication stress.

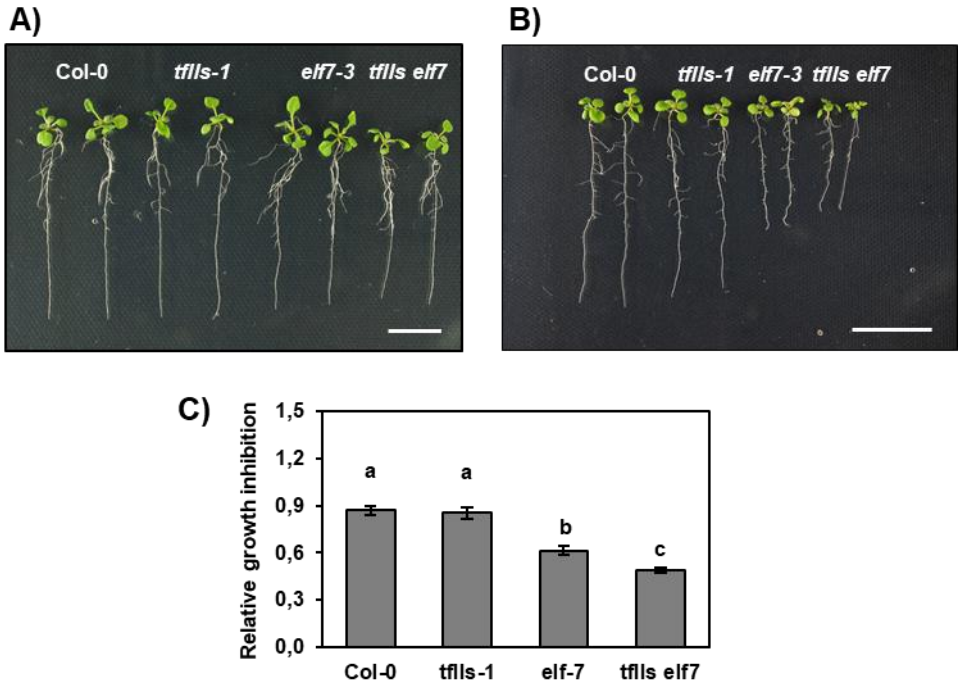


Figure 61. *elf7-3* shows hypersensitivity to hydroxyurea (HU).
Arabidopsis seedlings were grown on MS medium for 5 days and subsequently transferred on MS medium supplemented with mock (A) or 2 µM HU (B). Pictures were taken at 10DAS. (C) Quantification of roots growth retardation. Roots length was determined using ImageJ. Data significance was tested by two-way ANOVA. Error bars reflect SD and the letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05).

3.2.4 PAF1-C enables cell cycle progression upon compromised transcription

DNA replication stress has been shown to orchestrate the arrest or delay of cell cycle progression and subsequent fork restart ensured by homologous recombination (Gelot et al., 2015). In view of *elf7-3* hypersensitivity to replication stress further increased in the absence of TFIIS, it has been asked whether cell cycle progression particularly affected in *tfIIs elf7*. To answer this question, cells endoreduplication was studied in these transgenic lines. Endoreduplication occurs very commonly in terminally differentiated cells in *Arabidopsis* as a process of doubling chromosomal DNA without mitotic division which consequently determines cell polyploidisation (Joubès and Chevalier, 2000; Schuermann et al., 2009).

The aerial parts of 21DAS seedlings grown on the MS medium were used to evaluate presumable perturbations in cell polyploidisation in *tfIIs elf7* in comparison to Col-0 and respective single mutants (Figure 54 A). FACS measurements and initial data analysis was performed by Dr. Jörg Fuchs as described in chapter 2.4.3. In brief, nuclei extracted from individual plants were stained with DAPI and nuclei extracts were loaded into the Flow Cytometer. After reads collection (~ 5000), cell ploidy was determined for each measurable nucleus and the average ploidy level was calculated for each genotype. As a result, a broad range of ploidy level could be identified in all genotypes, ranging from 2C to 128C (Figure 62 A). The majority of cell showed either diploid (2C) or tetraploid (4C) states, similarly to reported previously (Schuermann et al., 2009). The higher ploidy states (> 4C) comprised much smaller fraction (~ 30%) which seemed to differ between analysed genotypes. Therefore "> 4C" cells were analysed statistically in detail. As seen on the Figure 62 B, "> 4C"

fraction was significantly decreased in *tfIIs elf7* in comparison with Col-0 and *elf7-3*. Although *tfIIs-1* did not differ statistically form neither genotype in terms of "> 4C" fraction, a slight decreased in comparison to Col-0 and *elf7-3* could be seen, however to the lesser extend then in the double mutant. No ploidy defects in *elf-7* suggest that under normal growth conditions PAF1-C integrity is not crucial to proceed through the endoreduplication process and the effects observed in *tfIIs-1* are not caused by lower PAF1-C recruitment to TEC but rather by generally compromised transcription. The significant decrease of higher ploidy level in *tfIIs elf7* in comparison to Col-0 and *elf7-3* may imply that the presence of functional PAF1-C complex is necessary to ensure efficient progression through endoreduplication upon additional transcriptional stress which could in turn trigger TRC and replication stress, as suggested previously (García-Muse and Aguilera, 2016; Gelot et al., 2015).

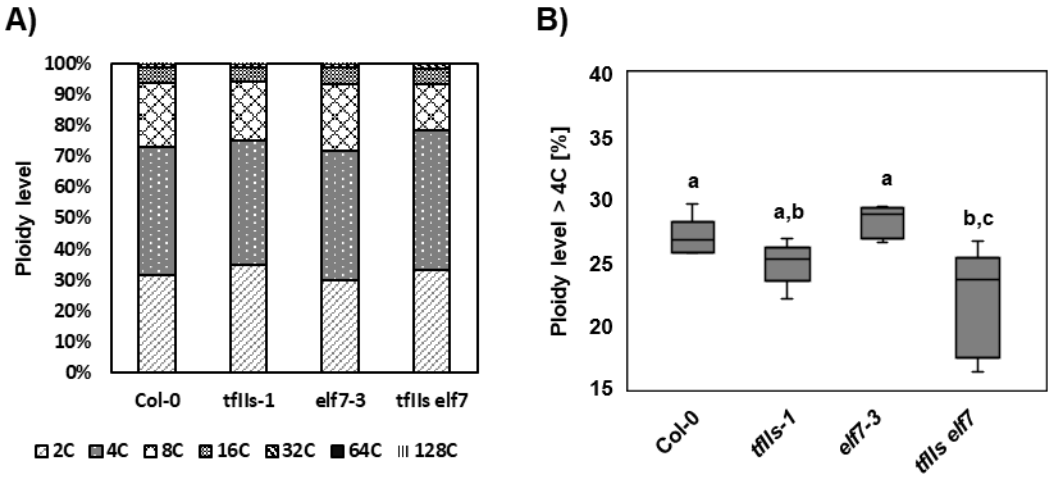


Figure 62. The analysis of cell ploidy in plants lacking functional TFIIS and/or ELF7.
Cell ploidy was measured for ~5000 nuclei extracted from the aerial parts of 21DAS Col-0, *tfIIs-1 elf7-3* and the double mutants. (A) The bar plot showing the percentage of each ploidy level identified across analysed lines. (B) Whisker-box plots showing the proportion of cells with ploidy higher then tetraploid. Data was collected for five individual plants of each line and analysed by two-way ANOVA. Error bars reflect SD and the letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05).

3.2.5 TRC-related transcriptional regulation in *tfIIs elf7*

The regulation of TRC resolution involves the coordination on many molecular mechanisms (Gelot et al., 2015). To better understand the regulation of biological processes potentially related to *Arabidopsis* TRCs, genes expression patterns were detailly studied in *tfIIs elf7* in comparison to Col-0 and respective single mutants. Based on the GO terms enrichment among ELF7 interactome (Table 14), the comparative gene expression analysis was performed in the subgroups of genes related to "DNA replication", "DNA damage" and "DNA repair". GO terms potentially related to these biological processes were extracted from all Gene Ontology annotations available in TAIR database and can be found in the Supplementary Table 16. The changes in expression level between studied genotypes within defined GO subgroups were further comparatively analysed and visualised using heatmaps. As seen on the Figure 63 A-C, broad transcriptomic changes could be observed across studied lines. The expression pattern seen in Col-0 seemed to differ particularly from *tfIIs elf7,* where the upregulation of many genes could be observed. To validate this observation, the absolute value of the expression

level fold change in the logarithmic scale (ABS(log_2FC)) was calculated between Col-0 and studied mutant lines. As seen on the Figure 63 D-F, the highest median of ABS(log_2FC) was indeed observed for *tfIIs elf7* double mutant in every analysed GO term subgroup. However, observed differences were not statistically significant between *tfIIs elf7* and respective single mutants within the subgroup of genes related to DNA damage nor between *tfIIs elf7* and *tfIIs-1* among DNA repair related genes.

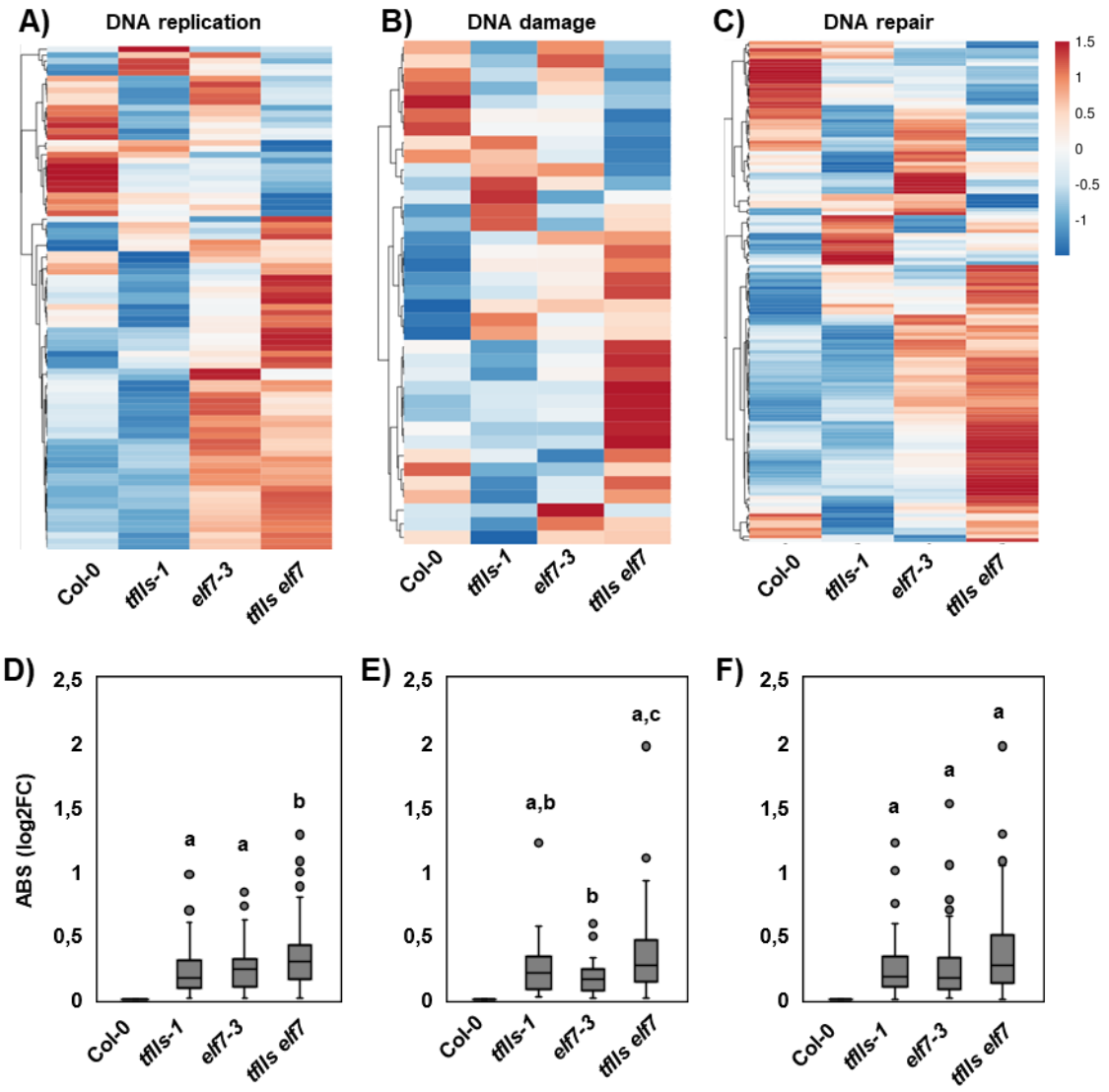


Figure 63. TRC-related transcriptome rearrangement in plants lacking functional TFIIS and/or ELF7.
(A-C) Heatmap visualisation of the expression pattern in Col-0, *tfIIs-1, elf7-3* and *tfIIs elf7* across genes related to DNA replication (A), DNA damage (B) or DNA repair (C) presumably participating in the control of transcription-replication conflict (TRC). Transcript levels were determined by RNA-seq and obtained values represent Transcripts Per Kilobase Million (TPM; chapter 3.1.8). Heatmaps were generated using ClustVis. The range of scaled fold change in transcript level (TPM) between lines (-1,5 to 1,5) with respective colour code shown on the right. (D-F) The quantification of average fold change between lines within defined subgroups. The absolute value of the expression level fold change in the logarithmic scale (ABS(log2FC)) was calculated for each mutant line relatively to Col-0 and depicted by whisker-box plots. Data significance was tested by one-way ANOVA. Error bars reflect SD and the letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05).

The analysis of expression level among genes potentially related to TRC revealed broad transcriptomic changes in the analysed mutants in comparison to Col-0. These changes were most striking in the double mutant lacking functional TFIIS and ELF7, in line with the highest hypersensitivity to HU as well as compromised endoreduplication in *tfIIs elf7*. Taken together, PAF1-C may play an important role in preventing the replication stress in *Arabidopsis* also in the context of transcription-replication conflict.

4. Discussion: TFIIS mutation reveals its importance in regulating transcript elongation in *Arabidopsis*

The molecular role of TFIIS in stimulating intrinsic RNAPII cleavage activity of misplaced nascent mRNA is well-established *in vitro* (Jeon et al., 1994; Kettenberger et al., 2003) yet its biological importance remains elusive. Considering TFIIS redundancy for plant growth and development (Grasser et al., 2009) its dominant negative form (TFIISmut) was used in this study as a molecular tool to shed light on the regulation of transcript elongation in *Arabidopsis*.

4.1 β-estradiol inducible system implementation

The functionality of inducible system

β-estradiol inducible system was successfully used in this study to express TFIISmut in plants lacking functional TFIIS. The system comprising single vector was created based on the two-component system described previously (Brand et al., 2006) simplifying its incorporation in different genomic backgrounds as well as antibiotic-based selection of created T-DNA lines. One potential risk of placing both activator and responsive units on one vector is system leakiness due to their proximity but it was not observed for analysed transgenic lines. Overall, system inducibility was comparable to previous reports in *Arabidopsis* reaching highest activity within 24h after β-estradiol application (Brand et al., 2006; Zuo et al., 2000).

The optimisation of induction kinetic was of main interest in this study in order to minimise TFIISmut toxicity at the early developmental stages as well as the accumulation of secondary effects following its long-term expression. Accordingly, inducible GFP-TFIIS/GFP-TFIISmut could be detected using confocal microscopy within ~ 3 h following β-estradiol application. Induction kinetic was also precisely determined in this study in the context of elongation rate system revealing inducible transcript accumulation within ~ 20 min similarly to previously reported (Zuo et al., 2000) as well as the luciferase activity within ~ 30 – 40 min following induction. Although those inconsistencies likely reflect the different sensitivity of applied detection systems, further improvements of GFP-TFIIS/GFP-TFIISmut induction kinetic could be likely achieved for instance by modifying β-estradiol concentration and/or application method (Brand et al., 2006; Zuo et al., 2000). In this study induction kinetic was largely improved by β-estradiol application in the liquid MS media upon vacuum infiltration. Nevertheless studding direct molecular consequences of TFIISmut expression would likely require major shortening of induction time since broad transcriptomic rearrangement may occur within few minutes in response to an external stimuli (Vinayachandran et al., 2018). Such shortening could be very challenging since homogenous β-estradiol delivery into complex organisms as well as production of the activator unit prior to target protein expression present serious limitations for rapid response.

In all experiments performed in this study transgenic TFIIS/TFIISmut was N-terminally fused with either GFP of GS tag. N-terminal fusion of TFIIS has been previously reported neutral for protein functionality and has been routinely used in many studies (Dutta et al., 2015; Ghavi-

Helm et al., 2008; Prather et al., 2005) while C-terminal tagging may impair TFIIS function leading to mutant phenotype (Prather et al., 2005) in accordance with the crystal structure of TFIIS-RNAPII complex (Kettenberger et al., 2003). In this study, TFIIS/TFIISmut N-terminal tagging with either GS or GFP did not affect protein incorporation into TEC as shown by AP-MS and ChIP-qPCR. Considering observed molecular and morphological consequences of TFIIS mutation within the conserved acetic loop, those findings suggest a similarly structured TFIIS-RNAPII complex in *Arabidopsis* in line with TFIIS sequence similarity in other organisms (Grasser et al., 2009).

In this study robust TFIIS/TFIISmut expression was desired rather than tissue-specific activation in accordance with comparable TFIIS expression level across different *Arabidopsis* tissues (Grasser et al., 2009). Thus, β-estradiol system was put under the control of *Arabidopsis* Ubiquitin 10 promoter (UBQ10) previously characterised to drive strong and constitutive expression of transgenes in stably transformed plants (Grefen et al., 2010) as well as in the context of β-estradiol inducible system (Dürr et al., 2014; Schlücking et al., 2013). Based on the GFP detection, transgenic TFIIS/TFIISmut expression could be observed across many cell types with its accumulation in the nucleoplasm in line with previous findings (Grasser et al., 2009). However, the expression of inducible proteins was not detected in all cells and the expression level varied across different cell types which could be likely caused by limited cells accessibility to β-estradiol (Brand et al., 2006; Schlücking et al., 2013). Alternatively, those findings could reflect modified inducibility and/or stability of inducible system components in certain cell types.

The morphological effects of TFIISmut expression

Constitutive TFIISmut expression in Col-0 background has been previously shown to affect plant growth and development (Dolata et al., 2015a). Although those findings cannot be directly compared with the phenotypes observed in this study, the morphological consequences of TFIISmut expression in plants lacking functional TFIIS seem more severe since no seedling developed upon 7 days induction of TFIISmut expression in *tfIIs-1*. Thus far no organisms with constitutive TFIISmut expression in the absence of functional TFIIS could be obtained neither in *Arabidopsis* (Dolata et al., 2015a) nor in yeast (Sigurdsson et al., 2010) while the co-expression of wild type TFIIS with transgene-derived TFIISmut has rather moderate effects on organism growth (Dolata et al., 2015a; Parsa et al., 2018; Sheridan et al., 2019). It suggests that wild-type TFIIS may be diluting negative dominant effect of TFIISmut and a fraction of functional TFIIS-RNAPII complexes ensure a certain degree of transcription correctness.

Considering comparable association of TFIIS and TFIISmut with RNAPII in this study, seemingly unaffected by TFIIS mutation (Awrey et al., 1998; Kettenberger et al., 2003), observed growth defects are likely the direct consequence of RNAPII arrest similarly to observed in yeast (Sigurdsson et al., 2010) supporting the biological functionality of created β-estradiol inducible system.

4.2 The molecular consequences of TFIIS mutation

The characteristics of transcript elongation in Arabidopsis

The molecular consequences of TFIIS mutation were analysed in this study in the context of active RNAPII occupancy using the antibodies directed against CTD part of NRPB1 phosphorylated at either serine 2 (RNAPII-S2P) or serine 5 (RNPII-S5P) position. RNAPII-S2P and RNAPII-S5P occupancies determined in this study represents the first genome-wide analysis performed in *Arabidopsis* thus far. Although active RNAPII profiling was performed in *tfIIs-1* as a control it presumably largely reflects active RNAPII occupancy in Col-0 since the absence of TFIIS was shown to only moderately modify the occupancy of transcriptionally engaged RNAPII in yeast (Churchman and Weissman, 2011). Despite the efforts no antibody ensuring robust immunoprecipitation of total RNAPII in *Arabidopsis* could be identified. In line with that, the genome-wide data representing total RNAPII occupancy in *Arabidopsis* are very limited and comprise a single study (Zhang et al., 2015c) utilizing the data obtained by ChIP-chip approach (Chodavarapu et al., 2010).

RNPII-S2P and RNPII-S5P occupancies have been precisely profiled in other organisms showing the overall tendency for 3' and 5'end enrichment, respectively. RNAPII-S2P profile upon normal growth condition in this study resembles findings in yeast and metazoans, whereas RNAPII-S5P distribution is clearly different from its previously determined profile in other organisms due to the lack of 5'end specific enrichment (Chen et al., 2018a; Hajheidari et al., 2013; Vinayachandran et al., 2018). Importantly, the RNAPII-S5P profile demonstrated in this study resembles total RNAPII occupancy determined by ChIP-chip in *Arabidopsis* (Zhang et al., 2015c) as well as the profile of transcriptionally engaged *Arabidopsis* RNAPII resolved by GRO-seq (Hetzel et al., 2016). Thus, *Arabidopsis* RNAPII-S5P occupancy determined in this study is presumably well correlated with total RNAPII occupancy, similarly to observed in human (Chen et al., 2018a).

Several characteristics of transcriptional regulation in *Arabidopsis* distinct from yeast and humans have been determined recently including lack of enhancer RNAs, promoter-proximal pausing and divergent transcription (Hetzel et al., 2016). While enhancer RNAs where not addressed in this study, performed ChIP-seq confirmed the lack of promoter-proximal pausing and divergent transcription. *Arabidopsis* and *maize* genes have been additionally shown to accumulate RNAPII in close proximity to polyadenylation sites (Hetzel et al., 2016; Lozano et al., 2018). In line with these findings, clear accumulation of active RNAPII could be seen around TES. Thus, previously reported plant-specific characteristics of transcript elongation are in accordance with the findings in this study supporting their biological relevance.

TFIIS association with RNAPII

The mutation of canonical amino acids within TFIIS acetic loop has been shown to inhibit RNA cleavage in yeast (Cheung and Cramer, 2011; Sigurdsson et al., 2010). Although TFIIS was shown to strongly resemble RNAPII distribution in yeast (Ghavi-Helm et al., 2008; Sheridan et al., 2019), it remains unclear which genomic regions and/or processes are primarily affected by TFIISmut.

TFIIS association with RNAPII has been well characterised to be mediated by domain II independently from domain III (Awrey et al., 1998; Kettenberger et al., 2003). Their direct interaction seems to occur in stochiometric amounts in accordance with determined TFIIS-RNAPII structure (Xu et al., 2017). TFIIS was additionally shown to be recruited at the subset of gene promotors playing a role in initiation independent from its activity in stimulating transcript cleavage (Guglielmi et al., 2007; Kim et al., 2007; Prather et al., 2005).

In this study the largest subunits of *Arabidopsis* RNAPII complex were identified among the most prominent TFIIS interactors as determined by AP-MS. A similar approach has led to comparable findings in yeast (Cojocaru et al., 2011). Here, TFIIS was shown to predominantly associate with *Arabidopsis* TEC considering lack of transcription initiation factors in TFIIS pulldown. In contrary, the prominent presence of initiation and elongation factors has been identified in NRPB1 pulldown (Antosz et al., 2017) indicating the isolation of functionally different RNAPII forms given the high turnover of initiation to elongation factors throughout transcription cycle progression (Pokholok et al., 2002). However, TFIIS association at some genomic loci during initiation cannot be excluded since TFIIS enrichment was observed by ChIP-qPCR at the genomic regions upstream TSS. It was not possible to determine whether these finding reflect the precise position of TFIIS or are rather caused by the low resolution of applied ChIP-qPCR (Rhee and Pugh, 2011).

Despite well characterised TFIIS-RNAPII interaction, its dynamics within the TEC remains elusive. It has not been determined whether the TFIIS molecule remains bound to RNAPII throughout the entire transcription cycle or rather dynamically associates and dissociates from RNAPII and stimulates the enzyme intrinsic RNA cleavage activity when needed. Strong overlap between TFIIS and RNAPII occupancy genome-wide (Ghavi-Helm et al., 2008; Sheridan et al., 2019) might suggest their permanent association. However, as determined by TFIIS-RNAPII crystal structure the permanent presence of TFIIS within RNAPII complex could sterically interfere with NTP diffusion (Wang et al., 2009; Zhang et al., 2015b).

Intriguingly, TFIIS has been shown to change its conformation from the “open” to the “closed” form upon RNAPII binding *in vitro* (Eun et al., 2014) so that multiple rounds of backtracking and transcript cleavage following permanent TFIIS association with RNAPII have been suggested.

In contrast to RNAPII, RNAPI and RNAPIII possess strong intrinsic RNA cleavage activity attributed to the action of A12.2 and C11 subunits, respectively (Vannini and Cramer, 2012). Their evolutionary relationship with TFIIS has been demonstrated revealing A12.2 and C11 C-ribbon correspondence to the TFIIS acidic loop (Ruan et al., 2011). Additionally, A12.2 and C11 C-ribbon domains were shown capable of swinging between surface and pore locations (Eun et al., 2014; Ruan et al., 2011). Interestingly, the bacterial TFIIS homolog GreA (Erie et al., 1993), has been also observed in at least two conformations (Laptenko et al., 2006). These findings imply a certain degree of TFIIS structural flexibility within the TEC in line with suggested in Eun et al., 2014, allowing permanent association between TFIIS and RNAPII.

TFIIS vs RNAPII occupancy: RNAPII backtracking triggered by TFIISmut

In this study TFIIS was found enriched over transcribed regions but not over transcriptionally inactive regions. TFIIS occupancy was further significantly enriched upon its mutation, especially at the 5' end of analysed genes. A similar observation was demonstrated in yeast where genome-wide distribution of TFIISmut was strongly shifted towards the 5'end relative to wild type TFIIS distribution (Sheridan et al., 2019). In this study *Arabidopsis* TFIIS and TFIISmut were shown to associate with similarly composed TEC in line with TFIIS acidic loop redundancy in RNAPII binding (Awrey et al., 1998; Kettenberger et al., 2003). Thus, detected differences in their occupancy over transcribed regions are likely the result of TEC accumulation rather than increased TFIISmut affinity. Considering TFIIS-RNAPII association genome-wide (Ghavi-Helm et al., 2008) as well as RNA cleavage inhibition by TFIISmut (Sigurdsson et al., 2010), RNAPII accumulation in the presence of TFIISmut could likely reveal certain genomic positions and/or processes regulated by RNAPII pausing in a TFIIS-dependent manner.

Accordingly, active RNAPII enrichment in the promotor proximal region (PPEP) was one of the most striking molecular consequences of TFIISmut expression in this study. This finding is in accordance with genome-wide observations in yeast (Parsa et al., 2018) and human (Sheridan et al., 2019) showing inhibited RNAPII escape from 5' pause sites in the presence of TFIISmut. However, the accumulation of either total or Ser2 phosphorylated RNAPII towards the 5'end upon TFIISmut expression was not observed in Sheridan et al., 2019. Their study assessed the effects of transgenic mouse TFIISmut and endogenous human TFIIS co-expression, thus reflecting only partial inhibition of RNA cleavage activity (Sheridan et al., 2019). Another contradicting finding was strong RNAPII accumulation downstream polyadenylation sites observed upon TFIISmut expression in Sheridan et al., 2019, which was absent in this study. Since timing of both assays was comparable (24h induction) those results could reflect specific role of human TFIIS in regulating RNAPII pausing downstream polyadenylation sites. In yeast, an elevated level of promoter-proximally paused RNAPII upon TFIISmut expression could be observed by using NET-seq (Parsa et al., 2018) although the analysis was also performed upon concerted expression of wild-type and mutated TFIIS. In contrary, total RNAPII decrease over transcribed regions upon TFIISmut expression was detected by using ChIP-qPCR in Sigurdsson et al. 2010, however the analysis was only performed for three individual genes.

Remarkably, the mutated version of archaeal TFIIS homolog (TFS4) was recently shown to fully inhibit cleavage *in vitro* leading to organisms growth retardation although the effect on RNAPII redistribution remains unknown (Fouqueau et al., 2017). Additionally, compromised RNAPII release from promoter-proximal stall sites has been also demonstrated in the absence of functional TFIIS in *Drosophila* (Adelman et al., 2005).

Taken together, TFIISmut triggered RNAPII redistribution in this study corelates with findings in other organisms supporting the conserved molecular function of TFIIS. The effects observed here could be more profound since mutated TFIIS was the only expressed form of the protein in contrast to other genome-wide studies.

RNAPII arrest and its consequences

The molecular effects observed in this study upon TFIISmut expression are likely the consequence of inhibited RNA cleavage within backtracked TFIISmut-RNAPII complexes leading to RNAPII stalling and accumulation. Indeed, the direct connection between RNAPII accumulation and increased backtracking frequency upon TFIISmut have been shown in human (Sheridan et al., 2019).

Different terms have been used to describe stalled RNAPII complex depending on its ability to resume RNA synthesis. While paused RNAPII can be readily induced to resume transcription (Figure 64 top), arrested elongation complexes require additional stimuli like for instance TFIIS (Figure 64middle) to avoid their removal from DNA template as terminating complexes (Figure 64 botton) (Adelman and Lis, 2012).

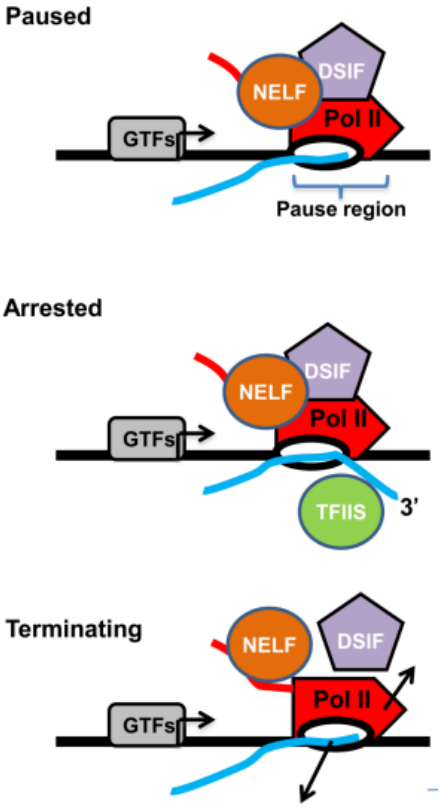


Figure 64. Terms describing stalled RNAPII.
Transcriptionally engaged complexes with stalled RNAPII may be described as paused, arrested or terminating. The promoter region and the transcription start site (TSS) as marked with an arrow. RNAPII is illustrated as a red rocket. The general transcription factors (GTFs) are shown as a grey oval. NELF (orange oval), DSIF (purple pentagon) and TFIIS (green circle) association is indicated. The nascent RNA transcript is shown in blue. Modified from: Adelman and Lis, 2012.

The distinction between paused, arrested and terminating RNAPII complexes is not possible with footprinting methods while “run-on” assays like GRO-seq allow the visualisation of paused RNAPII unlike for arrested or terminating complexes (Core et al., 2008; Min et al., 2011). Since decreased RNAPII occupancy in the promoter proximal region has been observed

in yeast upon TFIISmut expression by using GRO-seq but not with ChIP-seq (Sheridan et al., 2019) TFIIS mutation presumably traps RNAPII in arrested TEC being eventually transformed into terminating complex.

Several molecular strategies have been characterised to counteract the consequences of RNAPII arrest, including backtracking reversion by RNA cleavage or polymerase forward movement as well as RNAPII eviction by ubiquitylation-mediated proteasomal degradation (Gómez-Herreros et al., 2012; Svejstrup, 2003). Accordingly, mutations simultaneously affecting TFIIS-driven RNA cleavage and NRPB1 ubiquitylation are synthetic lethal (Somesh et al., 2007) although other molecular mechanisms for rescuing arrested RNAPII have been reported (Mao et al., 2014). Thus, RNAPII degradation could be the main mechanisms to rescue arrested complexes upon inhibited RNA cleavage in the presence of TFIISmut. In line with that assumption clear polyubiquitination of yeast NRPB1 was observed upon TFIISmut expression (Sigurdsson et al., 2010). Remarkably, elevated NRPB1 polyubiquitination in the absence of functional TFIIS was also found in yeast despite their growth properties being unaffected (Karakasili et al., 2014). NRPB1 polyubiquitination upon TFIISmut expression could not be convincingly determined in the course of this study due to the lack of reliable antibodies against polyubiquitinated NRPB1. Although an elevated level of UBQ was seen in TFIISmut pulldown it remains to be determined whether this corresponded to NRPB1-attached ubiquitin. Intriguingly, some components of proteasomal machinery and ubiquitin ligase complex were copurified specifically with TFIISmut as determined by affinity purification coupled with mass spectrometry. In line with that, the reduction of NRPB1 protein level by ~ 50% was observed in the presence of TFIISmut while NRPB1 level was comparable to wild type upon inhibition of proteasomal degradation with MG132. Thus, the rescue of arrested RNAPII upon TFIISmut expression in *Arabidopsis* may involve similar molecular mechanisms as reported in yeast.

The molecular consequences of stalled transcription have also been studied for other inducers of RNAPII arrest such as UV radiation and α-amanitin. UV radiation is one on the DNA-damaging agents which may trigger RNAPII arrest resulting in ubiquitination and proteasomal degradation of human NRPB1 (Ratner et al., 1998). α-amanitin is a well-established inhibitor of RNAPII (Bensaude, 2011) and its dose-dependent effects on *Arabidopsis* RNAPII progression were demonstrated in the course of this study. RNAPII inhibited with α-amanitin fails to be rescued by TFIIS-dependent mRNA cleavage and consequently is being targeted for NRPB1 polyubiquitination and degradation (Nguyen et al., 1996; Szeberenyi, 2006). Those similar molecular consequences to the once observed upon TFIISmut expression support the idea of RNAPII arrest in the presence of mutated TFIIS.

Arrested RNAPII has been implicated in many studies to present a serious barrier to other DNA-based processes like replication (García-Muse and Aguilera, 2016; Poli et al., 2016). In this study TFIISmut expression was showed to interfere with cell cycle progression. It was, however, not attempted to determine the direct influence of TFIISmut-RNAPII complexes on the progression of the replication machinery. The mobility of TEC containing TFIISmut was addressed in this study revealing a strongly compromised dynamic, while highly mobile RNAPII was shown to rapidly travel between various transcriptional stages in yeast

(Darzacq et al., 2007; Steurer et al., 2018) similarly to other chromatin-associated proteins (Phair et al. 2000, Pederson 2001, Dundr et al. 2002, Phair et al. 2004). Those data may suggest longer persistence of arrested TFIISmut-RNAPII on chromatin in line with high stability of arrested complexes (Cheung and Cramer, 2011), similarly to stable association of most nucleosomal histones with DNA resulting in their low mobility (Kimura et al. 2001, Kimura 2005).

Together, the conserved molecular consequences of RNAPII arrest across various organisms are in line with observations in this study supporting the idea of inhibited RNA cleavage in the presence of mutated TFIIS followed by RNAPII arrest.

4.3 TFIIS role *in vivo*

Nucleosome traversal by RNAPII

Although the molecular role of TFIIS in promoting RNA cleavage is well-established its implications *in vivo* remain elusive (Fish and Kane, 2002; Sheridan et al., 2019). In this study active RNAPII accumulation in the promotor proximal region (PPEP) was revealed upon TFIISmut expression showing a strong overlap with the position of the +1 nucleosome.

Nucleosomes are very frequent obstacles for transcriptionally engaged RNAPII on a DNA template and could consequently lead to RNAPII arrest (Gómez-Herreros et al., 2012). Accordingly, RNAPII backtracking has been detected in *Drosophila* at the +1 nucleosome (Weber et al., 2014) and nucleosome-induced pausing has been demonstrated as a major barrier to transcript elongation in yeast (Churchman and Weissman, 2011).

Nucleosome occupancy determines DNA accessibility and its level has been shown predictive of gene expression level in *Arabidopsis* (Liu et al., 2015; Zhang et al., 2015c). Nonetheless the molecular processes underlying nucleosome traversal by RNAPII *in vivo* are still poorly understood (Krajewski et al., 2018; Luse and Studitsky, 2011). Even a single nucleosomes promotes backtracking *in vitro* by presenting a strong barrier to transcript elongation (Gómez-Herreros et al., 2012) hence efficient and rapid nucleosome traversal by RNAPII requires the assistance of various transcript elongation factors (Luse and Studitsky, 2011; Nock et al., 2012). Recently the structure of transcribing RNAPII-NCP (nucleosome core particle) has been resolved in yeast revealing the involvement of several TEFs into nucleosome traversal and chromatin re-establishment after RNAPII passage including TFIIS, DSIF, NELF, SPT6, and PAF1 (Farnung et al., 2018).

TFIIS has been extensively studied in the context of nucleosome traversal considering its role in promoting transcript elongation (Izban and Luse, 1992). Accordingly, many *in vitro* studies revealed TFIIS importance in promoting RNAPII passage thought nucleosome barrier (Gaykalova et al., 2015; Ishibashi et al., 2014; Nock et al., 2012). TFIIS ability to relieve nucleosome-induced RNAPII arrest was assigned to its role in stimulating RNA cleavage (Kireeva et al., 2005; Luse et al., 2011; Nock et al., 2012). In line with that, RNA cleavage inhibition by TFIISmut eradicates its stimulatory effects on transcription across the nucleosome-containing template (Nock et al., 2012). Moreover, TFIIS knockdown in *Drosophila* results in increased RNAPII stalling within the +1 nucleosome (Weber et al., 2014)

whereas high RNAPII pause density at each of the first four nucleosomes has been detected in yeast lacking functional TFIIS (Churchman and Weissman, 2011). Significantly altered nucleosome occupancy and positioning have been recently reported as a consequence of TFIIS absence in yeast (Gutiérrez et al., 2017). Importantly, transcriptionally engaged RNAPII was shown to primary accumulate at the +1 nucleosome in the absence of functional TFIIS in Gutiérrez et al., 2017.

In this study active RNAPII accumulated upon TFIISmut expression at the region occupied by the +1 nucleosome and additionally overlapped with the position of +2 and +3 nucleosomes. The precise determination of relative position between arrested RNAPII and the first three nucleosomes would, however, require the application of more precise methods then ChIP-seq (Rhee and Pugh, 2011). Overall, the entry site of +1 nucleosome was characterized as a major barrier to RNAPII at most genes, while gene body nucleosomes present rather low barriers (Weber et al., 2014) in line with RNAPII profiles in the absence of functional TFIIS (Gutiérrez et al., 2017). Intruignly, in the study by Gutiérrez et al., 2017 transcriptionally engaged RNAPII has been shown to primary accumulate at the +1 nucleosome in the absence of functional TFIIS, similarly to active RNAPII accumulation observed in this study upon TFIISmut expression (Figure 65).

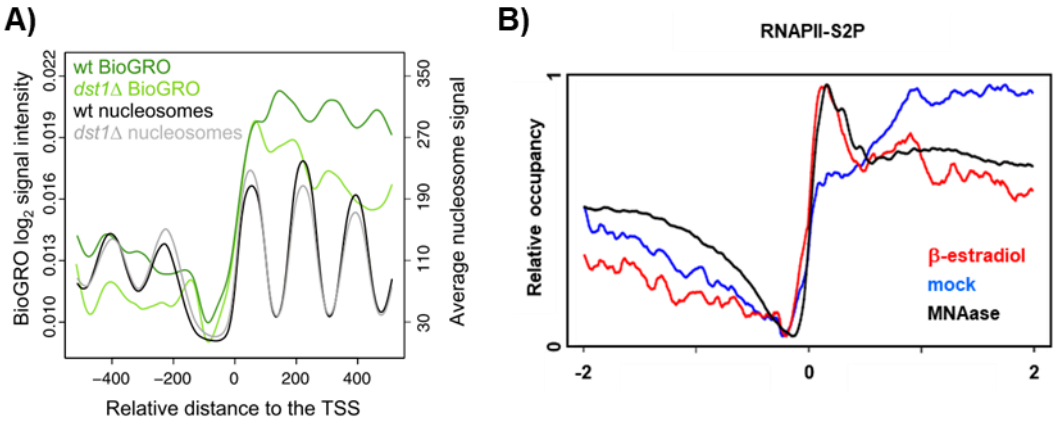


Figure 65. RNAPII accumulates at the +1 nucleosome upon perturbated TFIIS-stimulatory effects.
(A) Figure from Gutiérrez et al., 2017 depicting GRO-seq and nucleosome occupancy at TATA-like genes in yeast lacking functional TFIIS relatively to wild type cells. (B) RNAPII-S2P profiled together with nucleosomes occupancy performed in this study (2.3.2.3).

The increase in pause density at nucleosomes was shown comparable to the increase in nucleosome occupancy in yeast (Weiner et al., 2010). In this study the increase in PPEP in the presence of TFIISmut was not directly correlated with +1 nucleosome density nor with the first three nucleosomes (data not shown). Importantly, publicly available MNAse data analysed in this study were obtained using different plant material and developments stage as well as not-comparable growth conditions (Li et al., 2014). That could significantly bias analysed relationship since nucleosome occupancy has been proven very dynamic and tissue-specific in *Arabidopsis* (Liu et al., 2015; Zhang et al., 2015c). Alternatively, other nucleosome characteristics such as positioning or associated sequences motifs could have an impact on their traversal by RNAPII (Luse and Studitsky, 2011; Zhang et al., 2015c). The comparative genome-wide characterization of TFIISmut-RNAPII occupancy and nucleosome

properties under comparable experimental conditions could provide better insight in TFIIS-dependent nucleosome traversal by RNAPII *in vivo*.

In this study no major differences regarding TEC composition has been identified by AP-MS, however *Arabidopsis* IWS1a was demonstrated to copurify with GS-TFIISmut while it was absent in the GS-TFIIS interactome. IWS1a role in nucleosome traversal have not been extensively studied, although a switching mechanism has been proposed in which the association of SPT6 with nucleosomes is regulated by the binding of IWS1a in yeast (McDonald et al., 2010). Yeast SPT6 has been characterized with the periodic enrichment reflecting nucleosomes periodicity (Fischl et al., 2017). Interestingly, SPT6 peaks downstream of the SPT16 subunit of the FACT complex, a conserved regulator of nucleosomes passaging (Chen et al., 2018b; Pfab et al., 2018b), while FACT and RNAPII periodic occupancy was shown to overlap (Vinayachandran et al., 2018). These findings may imply that TFIISmut-TEC accumulates at the early steps of nucleosome traversal not reaching the stage where SPT6 interacts with H3. Interestingly, the phenotypic analysis of plants lacking functional TFIIS and FACT components revealed synergistic (*tfIIs ssrp1*) and epistatic (*tfIIs spt16*) effects on plant growth and development (Antosz et al., 2017).

Taken together, those findings may support the importance of TFIIS in nucleosome traversal by RNAPII accompanied with coordinated changes in TEC composition. When RNA cleavage activity of RNAPII is being blocked by TFIISmut, TEC might accumulate at nucleosome entry sites being depleted at nucleosome bodies.

Promoter-proximal pausing

One of the best studied characteristic of transcriptional regulation in metazoans is promoter-proximal pausing (Adelman and Lis, 2012; Gaertner and Zeitlinger, 2014). While its existence in plants has long been debated (Hajheidari et al., 2013) recent genome-wide studies revealed prominent promoter-proximal pausing in *Manihot esculenta* and to some degree in *maize* (Lozano et al., 2018) but not in *Arabidopsis* (Hetzel et al., 2016; Lozano et al., 2018). In accordance with that finding no enrichment of active RNAPII was observed in *tfIIs-1* at the promoter-proximal region in this study. Nonetheless clear PPRP increase could be seen upon TFIISmut expression with its local maximum ~ 150 bz downstream TSS and subsequent gradual decrease of RNAPII occupancy. These characteristics of active RNAPII distribution upon TFIISmut expression could presumably reflect an impaired release of promoter-proximally paused RNAPII, not-detectable upon normal growth conditions.

The role of TFIIS-stimulated RNA cleavage in RNAPII release from promoter-proximal pausing has been demonstrated in *Drosophila* (Adelman et al., 2005). Accordingly, RNA cleavage inhibition by TFIISmut has been shown to strongly compromise the release of promoter-proximally paused RNAPII in yeast (Parsa et al., 2018) and human (Sheridan et al., 2019).

The degree of proximal-pausing has been routinely determined by ChIP-seq subsequently supported with “run-on” sequencing methods (Gaertner and Zeitlinger, 2014; Mayer et al., 2017). Accordingly, the position of paused RNAPII in the promoter-proximal region has been estimated ~ 50 bz downstream TSS. Due to the low resolution of CHIP-seq method in this

study it was not possible to precisely determine the position of arrested RNAPII although it seems to peak ~ 150 bz downstream TSS.

In most studies "pausing index" (PI) is being used to quantitatively describe the level of promoter-proximal pausing and it refers to the ratio of promoter to gene body RNAPII density (Day et al., 2016) similarly to PPEP values calculated in this study. However, the determination of "paused" genes is highly variable among different studies (Adelman and Lis, 2012). Notably, the usage of "run on" based analysis such as GRO-seq seems to provide more consistency in terms of RNAPII pausing revealing a fraction of ~30% paused genes among all human, mouse and *Drosophila* genes(Core et al., 2008; Larschan et al., 2011; Min et al., 2011).

In this study the proportion of genes with presumable promotor-proximal pausing was estimated considering genes with highest PPEP values ($log_2 > 2$) as paused. Accordingly, their frequency was determined as ~ 15% of all genes under normal growth conditions and further increased upon TFIISmut expression could be seen (Table 16).

Table 16. RNAPII proximal pausing increases in the presence of TFIISmut.
Genes with log_2 PPEP > 2 were considered as paused and compared relative to all genes (n=33486). Promoter proximal enrichment of active RNAPII (PPEP) was determined in iGFP-TFIISmut line using ChIP-seq (2.3.2.3).

	RNAPII-S2P	RNAPII-S5P
β-estradiol	16,8 %	16,1 %
mock	14,6 %	15,5 %

RNAPII accumulation upon TFIISmut expression observed in this study was more prominent in the context of RNAPII-S2P. This finding was accompanied with elevated level of RNAPII-S2P in TFIISmut pulldown relatively to TFIIS. While TFIIS recruitment to RNAPII does not involve the CTD (Awrey et al., 1998), genetic interactions between TFIIS and RNAPII-CTD as well as CTD modifying enzymes have been identified in yeast (Lindstrom and Hartzog, 2001). The determination of RNAPII phospho-CTD specific interactomes did not reveal higher TFIIS affinity for RNAPII-S2P (Harlen et al., 2016) thus RNAPII-S2P accumulation upon TFIISmut expression could reflect the post-transcriptional modification imposed in the presence of TFIISmut. The CTD phosphorylation in the Ser2 position in one the key determinants for proximally-paused RNAPII release into the gene body. This process is regulated by cyclin-dependent kinase 9 (CDK9) which is a part of P-TEFb complex (Jonkers and Lis, 2015; Peterlin and Price, 2006). In *Arabidopsis*, CDKC;2 was identified as a CDK9 homolog (Wang, 2014) and accordingly P-TEFb complex was determined as a part of *Arabidopsis* TEC (Antosz et al., 2017). Thus, RNAPII-S2P accumulation could reflect an attempt to release proximally-paused RNAPII impeded by inhibited RNA cleavage in the presence of TFIISmut. Accordingly, the accumulation of RNAPII-S2P could be a consequence of not mutually exclusive longer "dwelling time" in the promoter-proximal region (discussed in the next chapter) or the role of TFIIS in transient recruitment of CTD kinases. In the latter scenario prolonged TFIISmut-RNAPII persistence in promoter-proximal region could result in increased phosphorylation of RNAPII-CTD at Ser2 position. Notably, TFIIS presence was shown to increase CDK9 recruitment and consequently RNAPII-S2P level in human (Cojocaru 2011).

The regulation of RNAPII promoter-proximal pausing is a very complex process including many regulator factors (Jonkers and Lis, 2015). In *Arabidopsis*, there are no known counterparts for components of the NELF complex, snRNA 7SK, HEXIM1, and HEXIM2 (Hajheidari et al., 2013). Noteworthy, NELF and TFIIS have been reported to occupy overlapping sites within the RNAPII complex, suggesting additional level of TFIIS-mediated regulation of RNAPII proximal-pausing (Vos et al., 2018a) in line with previously reported inhibition of TFIIS activity by NELF in yeast (Palangat et al., 2005).

Active RNAPII accumulation observed in this study upon TFIISmut expression near TSS may once again raise a question regarding RNAPII promoter-proximal pausing existence in plants. The lack of many important regulators of promoter-proximal pausing in *Arabidopsis* may imply different strategies and/or modified dynamics of promoter-proximal pausing with no clear genome-wide accumulation of RNAPII near TSS upon normal growth conditions.

Transcript elongation rate

Following regulatory mechanisms at the 5' end, RNAPII proceeds within genes with high variety of transcript elongation rates (Danko et al., 2013; Fuchs et al., 2014). RNAPII elongation rate has been addressed *in vivo* by a broad range of methods and has been characterised as dependent on multiple factors such as histone marks or certain DNA sequences (Jonkers and Lis, 2015). However, the determination of transcript elongation rate is still elusive in plants. In this study a system for measuring elongation rate was created based on the β-estradiol inducible system. Although precise determination of transcript elongation rate over the analysed single reporter gene could not be achieved due to insufficient temporal resolution, mRNA emergence was significantly slower in *tfIIs-1* relatively to Col-0. This result could be likely explained by modified pausing dynamics in the absence of functional TFIIS since overall RNAPII elongation rate is being consider as a combination of pausing properties as well as pause-free velocity (Schweikhard et al., 2014). Accordingly, TFIIS has been shown to stimulate transcript elongation *in vitro* by shortening the durations of transcriptional pauses, without affecting the pause-free velocity (Ishibashi et al., 2014). In yeast the absence of functional TFIIS results in the defects in nascent transcription (Gutiérrez et al., 2017). In line with that, TFIIS has been recently demonstrated as a major determinant of *in vivo* elongation rate in human since the expression of TFIISmut decreases overall transcript elongation rate by ~ 50% (Sheridan et al., 2019).

The effects of compromised transcript elongation rate on RNAPII occupancy have been detailly studied in yeast. While relative total RNAPII density within gene bodies is higher in mutants expressing "slow" RNAPII, no major changes regarding total RNAPII occupancy at 5'end have been observed (Fong et al., 2017). In contrary, a major shift towards the 5'end could be detected for RNAPII-S2P (Figure 66 A) but not for other CTD phosphorylations in response to slower transcription (Fong et al., 2017). Remarkably, similar RNAPII-S2P accumulation towards the 5'end was observed in this study upon TFIISmut expression (Figure 66 B). Observed redistribution towards the 5'end was less prominent for RNAP-S5P where overall accumulation over genes bodies was detected unlike for RNAPII-S2P. These observations could imply total RNAPII accumulation near TSS with additional RNAPII-S2P hyperphosphorylation resulting from a longer window of opportunity for its deposition.

Consequently, *Arabidopsis* TFIIS could play a major role in ensuring correct elongation rate in line with recently suggested in human (Sheridan et al., 2019), resulting in longer "dwelling time" of arrested TFIISmut-RNAPII complexes at the 5'end of the genes.

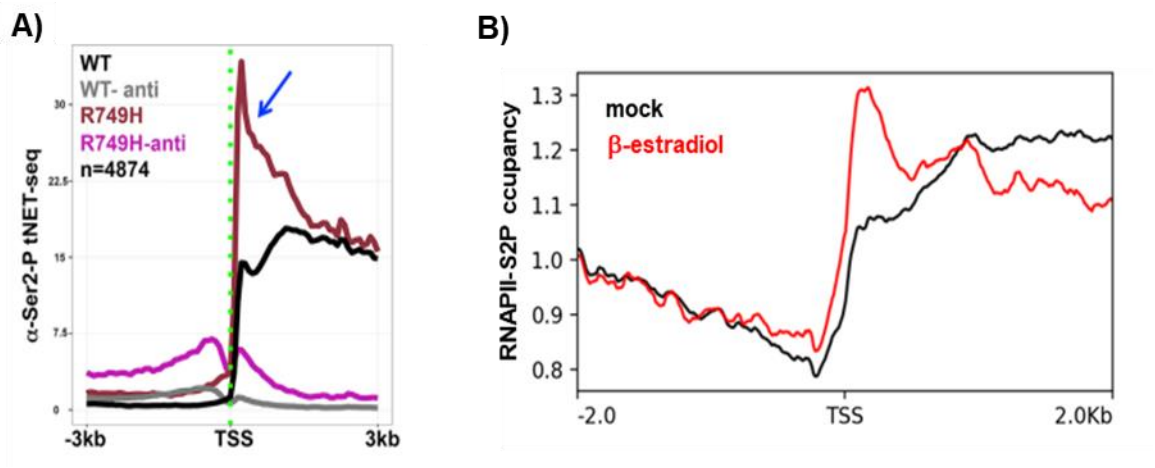


Figure 66. RNAPII-S2P is being enriched near TSS in response to slowed down elongation rate.
(A) Figure from Fong et al., 2017 showing slow RNAPII mutant (R749H; Fong et al., 2014) occupancy (red line) determined by using α-S2P NET-seq in comparison to wild type RNAPII (WT; black line). Purple and grey lines represent divergent transcription absent in *Arabidopsis* (B) RNAPII-S2P profiling performed in this study (2.3.2.3).

Transcription fidelity

Despite its importance in proper gene expression, RNAPII driven transcription may be characterised with a relatively high elongation rate error in comparison to other DNA-based processes (Gamba and Zenkin, 2018). Overall correctness of the final mRNA product has been estimated with an error rate of $\sim 10^{-3} - 10^{-5}$ (Gout et al., 2017; Imashimizu et al., 2015; Yuzenkova et al., 2010). The correctness of synthetized mRNA depends on the accuracy of nucleotide incorporation largely regulated by trigger loop oscillations (Yuzenkova et al., 2010) as well as the molecular mechanisms for transcriptional error correction (Gamba and Zenkin, 2018). Following single nucleotide misincorporation, RNAPII may undergo backtracking by 1 base pair or more, resulting in long-lived pausing (Sydow et al., 2009; Yuzenkova et al., 2010). Accordingly, misincorporation by RNAPII was shown to be a major source of transcriptional pausing in yeast, suggesting the direct link between transcription fidelity and pausing *in vivo* (Gamba et al., 2017; James et al., 2017).

TFIIS role in maintaining transcriptional fidelity has been early suggested (Jeon et al., 1994) and further supported by several *in vitro* studies involving yeast TFIIS (Irvin et al., 2014; Koyama et al., 2007; Thomas et al., 1998). In the study by Irvin et al., 2014, comprising random mutagenesis of yeast NRPB1, the integrity of TFIIS binding domain has been identified to be critical for transcriptional correctness *in vivo* together with trigger loop and bridge helix domains. In line with these results, yeast cells lacking functional TFIIS exhibit synthetic lethality when combined with mutations within trigger loop increasing elongation error rate (Thomas et al., 1998). Importantly, sequencing-based studies have revealed 7-fold (James et al., 2017) and 10-fold (Gout et al., 2017) increase of transcriptional error rate genome-wide when lacking functional TFIIS in yeast and human, respectively. In both studies particular DNA

sequence compositions have been identified as error-prone including “TGT” or “CGT” trinucleotides, suggesting the transition between purines and pyrimidines as when TFIIS is absent (Gout et al., 2017; James et al., 2017). Similar sequences were also found error-prone in bacteria (Imashimizu et al., 2015; Traverse and Ochman, 2016) and *Caenorhabditis elegans* (Gout et al., 2013). Intriguingly, the role of bacterial and archaeal RNA cleavage factors, GreA/B and TFS, in ensuring transcriptional fidelity has also been well-established (Erie et al., 1993; Lange and Hausner, 2004) implying the existence of conserved mechanisms of transcriptional mutagenesis across the tree of life.

Although transcriptional fidelity has not been addressed in this study, the sequence composition of loci exhibiting elevated RNAPII-S5P upon TFIISmut expression was analysed in the context of trinucleotide frequency. Notably, there is a striking similarity between enriched trinucleotides identified in this study and DNA sequences motif associated with a high error rate in yeast in the context of U > C misincorporation (James et al., 2017; Figure 67).

More detailed resolution of DNA motifs associated with TFIISmut-RNAPII arrested complexes would allow a valuable assessment of the degree of RNAPII pausing triggered by nucleotide misincorporation in the presence of mutated TFIIS. However, it would require the application of genome-wide sequencing methods with single-nucleotide resolution such as GRO-seq.

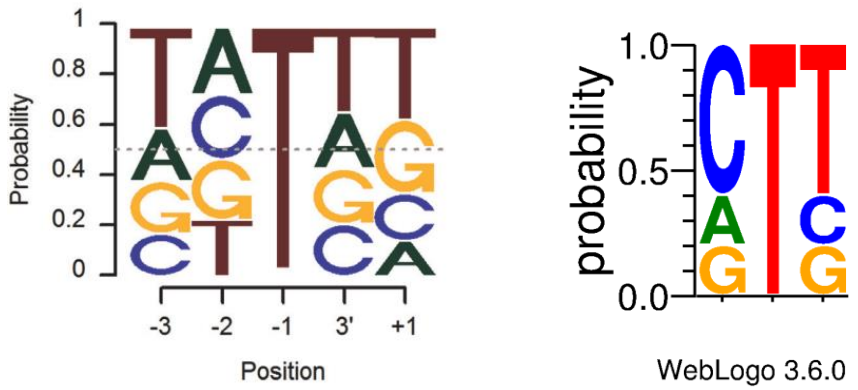


Figure 67. Sequence logos presumably related with TFIIS-dependent transcriptional fidelity.
(A) The sequences surrounding the U > C misincorporations at the − 1 positions yeast lacking functional TFIIS (James et al., 2017). (B) Sequence logos of trinucleotides sequences enriched in the genes exhibiting elevated RNAPII-S5P level upon TFIISmut expression in this study (2.3.2.3).

One of the biological consequences of high transcriptional error rate is the generation of proteotoxic stress which leads to significantly reduced cellular lifespan (Vermulst et al., 2015). Accordingly, error-prone yeast cells have been characterised with greatly increased sensitivity to MG-132, an inhibitor of proteasomal degradation (Vermulst et al., 2015). In view of that, *tfIIs-1* transgenic line was also tested in terms of MG132 sensitivity in the course of this study. As a result, plants lacking functional TFIIS showed strong hypersensitivity to MG132 relatively to Col-0 (Supplementary Figure S 16). In Vermulst et al., 2015 an increase in transcription error rate has been also linked to cells aging. Interestingly, the expression level of *Arabidopsis* TFIIS may be characterised with a clear increase during senescence (Supplementary Figure S 15).

Taken together, the correction mechanisms of missincorporated nucleotides could potentially consists additional molecular process controlled by *Arabidopsis* TFIIS as seen in other

organisms. However, the degree of this phenomena is still elusive, hence the usage of TFIISmut inducible system coupled with nucleotide-resolution sequencing methods could shed light on the degree of TFIIS-mediated transcription fidelity.

Transcriptome rearrangement

In this study, over-time transcriptome rearrangement upon TFIISmut expression was analysed to unravel critical biological processes regulated by TFIIS. Considering mRNA half-life time determined as > 6h for most *Arabidopsis* genes (Narsai et al., 2007), transcripts particularly sensitive to RNA cleavage inhibition were expected to gradually decrease their level in the presence of TFIISmut. However, observed transcriptome rearrangement was rather dominated by gene overexpression. This is in accordance with presumably much faster genome reprogramming which may occur within few minutes following the external stimuli, as determined in yeast (Vinayachandran et al., 2018). Accordingly, a broad defence and immune response were observed as a result of TFIISmut expression, suggesting RNA cleavage inhibition being recognised as a major interference with cell homeostasis.

In this study several biological processes could be identified as compromised already after 6h following β-estradiol application, including lipid localisation, photosynthesis and redox homeostasis. While lipid transport and macromolecule localisation have been previously reported in human as one of the crucial processes regulated by RNAPII proximal-pausing (Day et al., 2016), perturbated redox homeostasis is likely the direct consequence of compromised photosynthesis as determined previously (Liu et al., 2015; Scheibe et al., 2005).

Following the day/night cycle in plants, photosynthesis is being precisely and dynamically regulated on the transcription level and the majority of photosynthesis-related genes are encoded by the nuclear genome (Surpin et al., 2002; Timm et al., 2013). However, the regulation of photosynthesis gene expression may also be triggered by environmental stimuli such as nutrient availability or drought (Wang et al., 2017). Thus, downregulation of photosynthesis-related genes observed in this study could reflect missregulated transcriptome rearrangement in the day/night cycle and/or generally compromised metabolisms by TFIISmut-inhibited expression of some critical genes. Indeed, there is a growing number of evidence showing an extensive role for transcriptional control of metabolic network in response to developmental and environmental stimuli in plants (Gaudinier et al., 2015). Interestingly, TFIISmut has been recently demonstrated to hinder normal metabolic response to hypoxic stress in human (Sheridan et al., 2019).

Observed broad transcriptomic rearrangement upon TFIISmut expression could be alternatively a consequence of perturbated miRNA homeostasis. In human, many miRNA are being derived from RNAPII promoter-proximal paused regions (Zamudio et al., 2014) and miRNA synthesis was shown to be highly sensitive to α-amanitin, a potent RNAPII inhibitor (Lee et al., 2004). The perturbations in CTD phosphorylation pattern in *Arabidopsis* has been associated with changes in the level of some miRNAs (Hajheidari et al., 2012). Notably, plant miRNAs have been implicated into broad range of physiological and metabolic adaptations including photosynthesis, lipid synthesis and stress-response as well as morphological adaptation such as seed dormancy (Khraiwesh et al., 2012).

Outlook

In this study a novel approach for the examination of TFIIS importance *in vivo* was applied utilising inducible expression of TFIIS dominant negative form in a genomic background lacking functional TFIIS. The created system allowed robust and specific expression of inducible target proteins in a temporally controlled manner. However, further improvements would be necessary to ensure more homogenous and rapid expression. That could be achieved by modifying induction conditions or alternatively by focusing the analysis on a single cell level or a fraction of GFP-positive cells preselected by FACS. Although technically challenging, such approach could significantly improve data consistency and robustness, allowing more simplified interpretation.

Performed conditional expression of *Arabidopsis* TFIISmut likely results in RNA cleavage inhibition followed by RNAPII arrest in accordance with observations in other organisms. As a result, a broad range of molecular consequences within the period of > 24h following β-estradiol application were demonstrated, eventually leading to the inhibition of plant growth. Promoter proximal enrichment of active RNAPII (PPEP) upon TFIISmut expression was one of the most prominent findings in this study. Intriguingly, PPEP accumulation was shown to highly overlap with the position of the +1 nucleosome, suggesting the role of *Arabidopsis* TFIIS in nucleosome traversal by RNAPII. Possible TFIIS involvement into other, not mutually exclusive, molecular processes including regulation of promoter-proximal pausing, transcript elongation rate and fidelity as well as transcriptomic rearrangement cannot be excluded. Better spatial and temporal resolution of TFIISmut molecular consequences would be necessary to directly connect its role with those processes and could be achieved for instance by conditional nascent RNA sequencing. At the moment this method is still technically challenging in *Arabidopsis* and requires considerable big amount of plant material (~ 20 g) which makes it problematic in the context of any inducible system. Alternatively, *Arabidopsis* PSD-B cell culture system could be used as a material followed by TEF-seq with TFIIS as a bait.

Overall, using negative dominant version of TFIIS framed into an inducible system comprises a highly attractive molecular tool for studding TFIIS role *in vivo*. Further improvements in terms of system inducibility and sequencing data resolution could provide valuable data regarding regulation of transcript elongation in a plant-specific as in a general context.

5. Discussion: genetic interaction between TFIIS and PAF1-C

Although the positioning of TFIIS within RNAPII complex is well understood, the data explaining its dynamic association with other TEC components is still obscure (Kettenberger et al., 2003; Vos et al., 2018a; Xu et al., 2017). However, there is a lot of examples showing the genetic interaction between TFIIS and other transcriptionally-related factors (Fish and Kane, 2002). In this study plants lacking functional TFIIS and ELF7 were analysed and the genetic interaction has been identified between genes encoding *Arabidopsis TFIIS* and *ELF7*. It has been further attempted to understand observed negative synergy by analysing TFIIS-dependent affinity of PAF1-C to TEC as well as by studding their mutual involvement into various molecular processes.

The identification of *Arabidopsis* PAF1-C components using AP-MS approach (performed by Hans Ehrnsberger; (Antosz et al., 2017)) revealed the complex composed of six subunits, similarly to observed in human and *Drosophila melanogaster* (Table 17).

Table 17.The homologous components of PAF1-C in different organisms.

Arabidopsis (Antosz et al., 2017)	Yeast (Jaehning, 2010)	Human and *Drosophila* (Tomson and Arndt, 2013)
CDC73	Cdc73	Cdc73
ELF7	Paf1	Paf1
VIP3	-	Ski8/Wdr61
VIP4	Leo1	Leo1
VIP5	Rtf1	Rtf1
VIP6/ELF8	Ctr9	Ctr9

5.1 TFIIS-PAF1-C interaction

Who comes first?

The temporal resolution of TFIIS and PAF1-C recruitment during transcription cycle is crucial for understanding their interplay within the TEC. While TFIIS may be recruited already during transcription initiation (chapter 4.2), PAF1-C is generally considered to enter TEC downstream of the transcription start site (Mayer et al., 2010). Although total RNAPII occupancy show a positive correlation with human PAF1-C (Chen et al., 2015), more detailed studies reveal high similarity between the occupancy of individual PAF1-C subunits with human Paf1 and Cdc73 occupancy additionally correlated with promotor-proximally paused RNAPII (Lu et al., 2016). Additionally, the analysis of PAF-1C occupancy in yeast using ChIP clearly showed the accumulation of PAF1-C subunits downstream of TFIIS and RNAPII (Kim et al., 2010) as well as other TEFs (Fischl et al., 2017). In line with these findings, human PAF1-C was recently suggested to replace NELF in the process of RNAPII release from proximal-pausing (Vos et al., 2018b). Notably, PAF1-C has been shown relatively enriched at 5'end of analysed reporter gene while TFIIS level was comparable regardless the relative position over the gene (Harlen and Churchman, 2017), suggesting more dynamic association of PAF1-C with TEC.

In this study both TFIIS and ELF7 were shown to predominantly associate with the TEC. Further determination of their distribution over transcribed units in *Arabidopsis* could potentially reveal the relative timing of their incorporation into TEC, however it was not attempted in the course of the study. Notably, TFIIS could be identified among NRPB1 interactome while no PAF1-C components copurified with NRPB1 (performed by Karin Köllen; Antosz et al., 2017). Considering possible TFIIS association with RNAPII already during initiation as well as NRPB1 AP-MS approach revealing functionally different RNAPII forms, PAF1-C lack in NRPB1 pulldown could reflect its absence at RNAPII during early phases of transcription cycle. Together with the findings in other organisms, those data may suggest TFIIS recruitment prior to PAF1-C association during *Arabidopsis* RNAPII transcription cycle.

Relative positioning within the TEC

The direct binding of PAF1-C to RNAPII has been shown in yeast (Shi et al., 1996) and human (Kim et al., 2010) in line with recently resolved crystal structures of PAF1-C-RNAPII complex in those organisms (Vos et al., 2018b; Xu et al., 2017). Presented structures revealed a trilobal architecture of yeast and human PAF1-C deposited on the surface of RNAPII with Cdc73 and Paf1-Leo1 heterodimer on the opposite ends, bridged by Ctr9 (for *Arabidopsis* orthologs see Table 17). In Xu et al., 2017 a ternary structure of PAF1-RNAPII associated with TFIIS has been additionally resolved. While TFIIS extends from a polymerase jaw to the active site (Kettenberger et al., 2003; Xu et al., 2017), PAF1-C has been located over the outer RNAPII surface, reaching near the rim of the funnel opposite the active centre cleft (Figure 68; Xu et al., 2017).

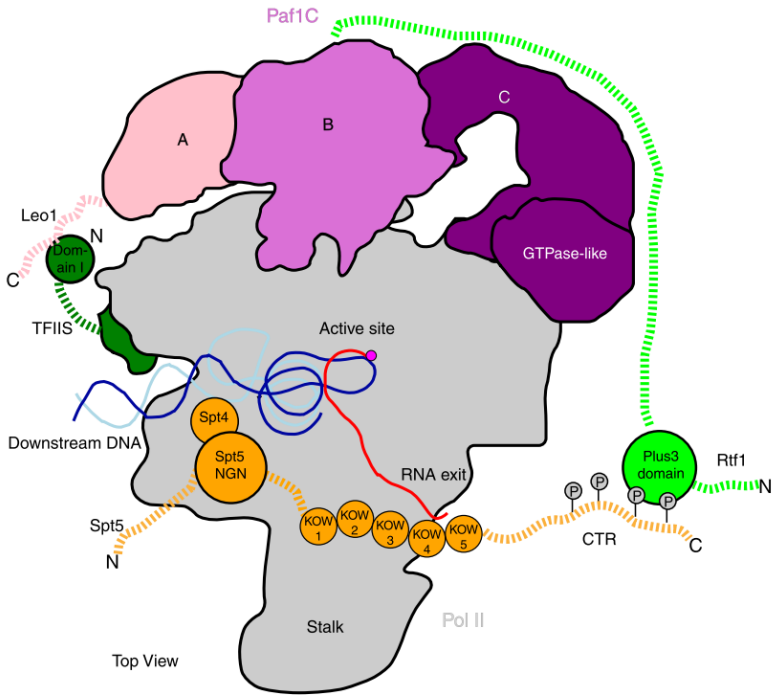


Figure 68. Model of transcriptionally engaged RNAPII with some TEFs.
The location and interactions between transcript elongation factors (TEFs) on RNAPII surface, including TFIIS, SPT4-SPT5 and PAF1-C. Dashed lines indicate flexible regions participating in protein-protein interactions (PPI). Yeast Paf1-Leo1 heterodimer may contact the N-terminal domain I of TFIIS. Other established PPI as also indicated. Picture is from Xu et al., 2017.

PAF1-C binding to RNAPII was also shown to be determined by the binding of yeast Bur1-Bur2 kinase (Qiu et al., 2012), SPT4-SPT5 (Mayekar et al., 2013), RNAPII-CTD (Cao et al., 2015) and nascent RNA (Dermody and Buratowski, 2010). All those findings suggest very complex and potentially highly regulated and dynamic interaction between PAF1-C and RNAPII.

TFIIS influence on PAF1-C level: direct or allosteric?

The obtained TFIIS-PAF1-C-RNAPII structure revealed the flexibility of the N-terminal TFIIS domain as well as the N- and C-terminal domains of yeast Paf1 and Leo1, explaining the possibility for their direct interaction near RNAPII jaw-lobe module (Xu et al., 2017). Accordingly, the direct interaction between yeast TFIIS and Leo1 has been revealed using chemical crosslinking coupled to mass spectrometry (Xu et al., 2017) while TFIIS has been shown to directly interact with human Paf1 and Leo1 by using pulldown approach (Kim et al., 2010).

In this study, no direct interaction between *Arabidopsis* TFIIS and ELF7 could be determined using Y2H and FRET. Additionally, TFIIS did not interact with remaining subunits of *Arabidopsis* PAF1-C as determined using Y2H assay (performed by Irene Fuhrmann). However, all PAF1-C subunits were identified among TFIIS interactome in this study while TFIIS copurified with ELF7 but not with CDC73 as determined by reciprocal tagging (performed by Hans Ehrnsberger; Antosz et al., 2017). These findings may reflect the proximity of ELF7 to TFIIS in the complex with RNAPII while CDC73 is located more distally as determined in yeast and human (Vos et al., 2018b; Xu et al., 2017). The N-terminal domain of TFIIS has been previously shown to play a role in the recruitment of yeast PAF1-C *in vitro* (Xu et al., 2017). Importantly, in this study the binding of all corresponding subunits od *Arabidopsis* PAF1-C to RNAPII was depleted upon N-terminal TFIIS removal *in vivo*. Only plant specific VIP3 subunit was unaffected by N-terminal TFIIS removal while ELF7 and VIP4 subunits showed the strongest depletion in ΔTFIIS pulldown in line with the presumable direct interaction between their yeast and human orthologs and N-terminal TFIIS (Kim et al., 2010; Xu et al., 2017). These results indicate an important role of *Arabidopsis* N-terminal TFIIS in determining PAF1-C level within the TEC presumably regulated by modified recruitment and/or affinity of PAF1-C to RNAPII. Consequently, no direct interaction detected in this study could account for methods limitation not allowing the detection of presumably transient or gene specific interaction.

Alternatively, the lack of directs interaction between TFIIS and ELF7 with simultaneously observed PAF1-C depletion in ΔTFIIS pulldown could reflect, not mutually exclusive, allosterically missregulated PAF1-C recruitment and/or affinity by truncated TFIIS. Interestingly, TFIIS binding to RNAPII triggers broad structural changes in the RNAPII complex inducing a coordinated repositioning of about one third of the polymerase mass including RNAPII jaws, clamp, cleft and foot domains of yeast RNAPII (Kettenberger et al., 2003). There are also evidences for conformational isomerization of RNAPII upon the transition from initiation to elongation as well as for distinct conformational states of elongating RNAPII (Erie, 2002; Palangat and Landick, 2001).

In this study ΔTFIIS was shown to associate with a similarly composed TEC relatively to full length TFIIS. However, the potential lower affinity of ΔTFIIS to the RNAPII complex as well as

the impact of N-terminal TFIIS absence on TEC properties remains to be determined. Interestingly, ΔTFIIS association with RNAPII could be characterised with higher enrichment of hyperphosphorylated RNAPII-CTD as well as distinct RNAPII-CTD phosphorylation patterns in comparison with its full-length form. Considering presumable role of N-terminal TFIIS in transcription initiation (Kim et al., 2007), the data in this study may suggest delayed recruitment of ΔTFIIS during transcription cycle which would consequently trigger a reduced PAF1-C level at the early transcriptional stages independently from its affinity to TFIIS N-terminal domain.

Although the mechanism underlying the described observations remain unclear, the absence of full length TFIIS results in reduced level of *Arabidopsis* PAF1-C within the TEC *in vivo*, in line with the suggested mechanism in yeast (Xu et al., 2017).

5.2 The consequences of a reduced PAF1-C level

Synergistic growth defects of tfIIs elf7

The genetic interaction between genes encoding *Arabidopsis TFIIS* and *ELF7* was identified in this study, resulting in strong synergistic defects in *tfIIs elf7* growth and development. Assuming reduced PAF1-C level within the TEC in *tfIIs-1,* observed phenotypes could be explained as a consequence of synergistically lowered PAF1-C level over transcribed units. It remains unclear whether the lack of functional ELF7 in *elf7-3* results in the compromised formation of the remaining five subunits into a functional *Arabidopsis* PAF1-C complex and its association with RNAPII in the chromatin context. It has been previously shown that human Paf1 is crucial for PAF1-C stability *in vitro* since its deletion may result in significantly diminished levels of an additional subunit (Kim et al., 2010). Surprisingly, the human Leo1 level at transcribed units remains largely unchanged upon Paf1 knockdown *in vivo* despite their direct interaction (Chen et al., 2015). Although in Xu et al., 2017 yeast Ctr9 has been shown as a scaffolding protein of PAF1-C, Paf1 subunit was also demonstrated as a part of core PAF1-C complex similarly to observations in human (Vos et al., 2018b). Thus, it remains to be determined to which extend *Arabidopsis* ELF7 removal results in the perturbated association of remaining PAF1-C subunits in order to better understand the molecular basis of observed genetic interaction.

Is it only the recruitment?

Remarkably, relative PAF1-C level within the TEC has been shown to dictate the fate of the nascent transcripts independent of transcription elongation leading to mRNA nuclear retention (Fischl et al., 2017). Thus, synergistically perturbated PAF1-C level within the TEC in the absence of functional TFIIS in *tfIIs elf7* could account for the general defects of mRNA homeostasis, consequently leading to observed synergistic growth defects. To test this hypothesis, it was attempted to restore *elf7-3* phenotype by introducing either full length or truncated transgenic TFIIS into *tfIIs elf7*. GFP-TFIIS expression under native TFIIS promoter in *tfIIs elf7* resulted in *elf7-3* phenotype suggesting protein functionality in line with TFIIS N-terminal tagging being redundant for its activity (Prather et al., 2005). In contrary, only partial complementation was observed upon GFP-ΔTFIIS expression. This important finding could suggest that observed growth defects in *tfIIs elf7* may be a consequence of reduced

PAF1-C level within the TEC with a certain degree of gene specificity and/or additional molecular processes synergistically missregulated by the mutual absence of functional TFIIS and ELF7. In line with that hypothesis no bolting phenotype, which is a hallmark of *Arabidopsis* mutant lines lacking functional PAF1-C subunits (He et al., 2004; Yu and Michaels, 2010), was observed for *tfIIs-1* in this study. Additionally, transcriptome profiling performed in this study revealed rather distinct transcriptomic profiles in plants lacking functional TFIIS or ELF7.

However, there might be some other potential explanations for observed partial complementation upon ΔTFIIS expression in *tfIIs elf7* independent from the role of N-terminal part of TFIIS in determining the PAF1-C level within the TEC, including: 1) compromised ΔTFIIS functionality and 2) modified ΔTFIIS-TEC properties such as: lower affinity and/or delayed recruitment of ΔTFIIS (resulting in allosteric effects discussed above). While ΔTFIIS functionality as well as ΔTFIIS-TEC properties would require further determination in *Arabidopsis*, mutual co-regulation of various molecular processes by TFIIS and ELF7 was addressed to some extent in the course of this study and will be discussed in the following chapters.

5.3 Other molecular levels of TFIIS PAF1-C negative interaction

Transcription-replication conflict in Arabidopsis

Being a multisubunit complex, PAF1-C has been implicated in many aspects of gene regulation (Tomson and Arndt, 2013; Yang et al., 2016). Recently its role in resolving transcription-replication conflict (TRC) by stimulating RNAPII eviction has been demonstrated in yeast (Poli et al., 2016). Since TFIIS is well characterised to regulate RNAPII backtracking and RNAPII arrest has been associated with elevated level of TRCs (García-Muse and Aguilera, 2016), it has been hypothesized that insufficient eviction of backtracked RNAPII in *tfIIs elf7* could contribute to the elevated level of TRCs and a certain degree of genome instability contributing to the observed synergistic phenotype. It has been shown that bacterial counterparts of TFIIS, GreA/B, are crucial for replication progression under substantial transcription activity (Tehranchi et al., 2010) and a similar role has been proposed for yeast TFIIS (Dutta et al., 2015). Notably, TFIISmut expression has been recently reported to increase RNAPII proximal pausing and R-loop formation leading to increased genomic instability in yeast (Zatreanu et al., 2019; under review).

In this study ELF7 association with a putative *Arabidopsis* INO80 complex as well as the importance of ELF7 for organism growth upon replication stress conditions was demonstrated, similar to observations in yeast (Poli et al., 2016). Accordingly, an elevated homologues recombination level was observed in *elf7-3* and many genes involved in DNA replication, DNA damage and DNA repair were identified to be particularly missregulated in *tfIIs elf7* relatively to Col-0. These findings could reflect an elevated level of TRCs in *tfIIs elf7* resulting for instance in cell cycle arrest (Deepak et al., 2015). While cell cycle progression was somewhat affected in *tfIIs elf7* as determined by endoreduplication, observed differences were rather minor. In line with that, the genomic distribution of *Arabidopsis* DNA replication origins (ORIs) in (Vergara et al., 2017) does not seem to directly determine gene downregulation as observed in this study in *tfIIs elf7* (Supplementary Figure S 17.).

Data obtained in the course of this study suggest a role of *Arabidopsis* ELF7 in ensuring efficient DNA replication upon stress conditions, presumably by driving RNAPII eviction as observed in yeast. Additionally, the insufficient resolution of TRCs in *Arabidopsis* could contribute to observed negative synergy between TFIIS and ELF7, although the understanding of TRC regulation still remains obscure (Chen et al., 2018c).

Nucleosome traversal by RNAPII

A presumable role of TFIIS in nucleosome traversal by RNAPII has been demonstrated and discussed in this thesis (chapter 4.3) while various regulatory roles of PAF1-C in stimulating nucleosome passaging and positioning have been demonstrated in yeast (Tomson and Arndt, 2013). Accordingly, the structure of yeast RNAPII-NCP complex has been recently resolved and superimposed with known PAF1-C structure (Farnung et al., 2018) demonstrating its precise accommodation on transcribing RNAPII in the presence of oriented NCP. In line with that, the association of PAF1-C with transcriptionally engaged RNAPII shows striking periodicity in yeast, resembling the positioning of phased nucleosomes (Fischl et al., 2017; Jiang and Pugh, 2009). Human PAF1-C has been suggested to enhanced nucleosome traversal by RNAPII via the recruitment of H2B ubiquitylation factors (Kim et al., 2009, 2010), linking transcription and H2B ubiquitylation which is one of the crucial post-translational histone modifications in passaging nucleosomes (Krajewski et al., 2018).

In Fischl et al., 2017 transcriptionally engaged RNAPII-PAF1-C has been shown absent at the +1 nucleosome, emerging at the +2 nucleosome and highest periodic accumulation around the +5 nucleosome. In contrary, other TEFs involved in nucleosome traversal (SPT16 and SPT6) have been found equally enriched across the first five nucleosomes (Fischl et al., 2017; Jiang and Pugh, 2009). Similarly, high RNAPII pause density has been detected at the first four nucleosomes in yeast lacking functional TFIIS (Churchman and Weissman, 2011). These findings may imply the importance of PAF1-C in nucleosome traversal across the gene body in line with their relatively easy passaging, in contrast to the +1 nucleosome presenting a major barrier to RNAPII (Weber et al., 2014). Spatially separated mechanisms in regulating nucleosome passaging by PAF1-C and TFIIS could be also concluded from those studies. However, PAF1-C occupancy has been demonstrated in human to accumulate immediately downstream of the TSS (Chen et al., 2015), contradicting the findings in yeast. This inconsistency may account for different methodology applied in those studies but most likely reflects distinct RNAPII properties in human related to promoter-proximal pausing of RNAPII. Thus, the genome-wide occupancy of *Arabidopsis* PAF1-C must be resolved in order to determine its mutual involvement into nucleosome traversal by RNAPII together with TFIIS.

PAF1-C has been additionally suggested to impact nucleosome dynamics and occupancy by promoting transcription-coupled histone modifications (Marton and Desiderio, 2008; Pruneski et al., 2011) or by direct interaction with yeast Chd1 (Simic et al., 2003) and human H3 (Marazzi et al., 2012). PAF1-C role in the regulation of histone H3 modifications and nucleosome levels have also been demonstrated in *Arabidopsis* (Oh et al., 2008).

Together those studies suggest an important and conserved role of PAF1-C in nucleosome traversal by promoting transcription-coupled histone modifications and impacting

nucleosome dynamics. The regulation of nucleosome properties by PAF1-C could consequently determine the requirement for TFIIS action in nucleosome traversal by RNAPII resulting in the mutual regulation of this process by TFIIS and PAF1-C.

Promoter-proximal pausing

Several TEFs play crucial role in regulating RNAPII promote-proximal pausing including DSIF, NELF and P-TEFb (Adelman and Lis, 2012) but other factors such as PAF1-C and SPT6 has also been demonstrated to influence this process (Andrulis et al., 2000; Jaehning, 2010). Recent structural resolution of the activated transcription complex has confirmed PAF1-C and SPT6 importance for the release of paused RNAPII *in vitro* (Vos et al., 2018b) and PAF1-C has been suggested to displace NELF from paused complexes (Farnung et al., 2018; Vos et al., 2018b) Interestingly, NELF and TFIIS binding to RNAPII funnel have been recently reported mutually exclusive (Vos et al., 2018a). Considering direct interaction between human TFIIS and PAF1-C (Kim et al., 2010), their proximal binding within the TEC (Vos et al., 2018b) could presumably coordinate NELF-dependent release of paused RNAPII in metazoans.

In several studies PAF1-C has been characterised as a major molecular regulator of promoter-proximal pausing by RNAPII while its influence on this process has been shown as both negative and positive (Chen et al., 2015; Yu et al., 2015). Accordingly, P-TEFb dependent phosphorylation of RNAPII-CTD in Ser2 position has been characterised as critical for the PAF1-C recruitment to target genes in yeast and human (Liu et al., 2009; Yu et al., 2015) whereas in other studies PAF1-C has been shown to drive P-TEFb recruitment (Lu et al., 2016) and S2P phosphorylation level (Dronamraju and Strahl, 2014). Those apparent inconsistencies could be interpreted as a positive feedback between PAF1-C and RNAPII-CTD in order to control differential levels of RNAPII-associated PAF1-C across transcribed units (Fischl et al., 2017). While the genome-wide occupancy of individual PAF1-C subunit is highly similar, yeast Paf1 and Cdc73 show certain tendency to overlap with the promoter-proximal pausing of RNAPII (Lu et al., 2016). Interestingly, the distribution of human Leo1 is clearly reduced only in the promoter-proximal region upon Paf1 knockdown (Chen et al., 2015).

Arabidopsis P-TEFb has been suggested to rather dynamically interact with TEC (Antosz et al., 2017). While CYCT1;5 subunit has been copurified together with *Arabidopsis* ELF7, no PAF1-C subunits were identified in CDKC;2 pulldown (Antosz et al., 2017), contradicting a comparable assay performed in human (Yu et al., 2015). More detailed studies are necessary to evaluate the role of *Arabidopsis* PAF1-C in regulating RNAPII-CTD properties in order to correctly assess its potential influence on TFIIS-dependent RNAPII promoter-proximal pausing. However, the available data from other organisms may suggest their mutual involvement into the regulation of paused RNAPII release. Similar mutual regulation of RNAPII pausing by *Arabidopsis* TFIIS and PAF1-C cannot be excluded, potentially contributing to observed synergistic growth defects in *tfIIs elf7*.

Outlook

The molecular principals driving the genetic interaction between *Arabidopsis* genes encoding *TFIIS* and *ELF7* have been addressed in this study. Observed negative synergy could be likely explained by a reduced PAF1-C level within the plant TEC lacking functional TFIIS. A reduced PAF1-C level within the TEC was attributed to the absence of the N-terminal part of TFIIS although some alternative explanations cannot be excluded. Thus, more detailed determination of direct interaction between *Arabidopsis* TFIIS and PAF1-C as well as PAF1-C level within the TEC in *tfIIs-1* could further support the proposed explanation of synergistic growth defects in *tfIIs elf7*.

In this study, TFIIS and PAF1-C were demonstrated as a part of *Arabidopsis* TEC, being likely involved in various molecular processes associated with the DNA template. The synergistic perturbation of those processes in *tfIIs elf7* could additionally contributed to the observed genetic interaction. The involvement of *Arabidopsis* ELF7 into TRCs resolution was addressed in this study to some extent. Further determination of ELF7-driven eviction of arrested RNAPII colliding with the replication machinery could be assessed upon TFIISmut expression, providing an additional interesting angle in the interpretation of molecular and morphological defects in *tfIIs elf7.*

PAF1-C involvement into nucleosome traversal by RNAPII and promoter-proximal pausing of RNAPII is well-established in other organisms. These processes are still poorly understood in *Arabidopsis* although the results in this study may suggest their dependence on TFIIS. Therefore, the mutual regulation of those processes by *Arabidopsis* TFIIS and PAF1-C cannot be excluded. The determination of *Arabidopsis* PAF1-C occupancy in the chromatin context could provide further insight into the mutual regulation of nucleosome traversal and promoter-proximal pausing by TFIIS and PAF1-C.

6. Summary

The regulation of gene expression is a fundamental process in eukaryotes encompassing many molecular mechanisms. Transcript elongation has emerged over the last decades as a highly regulated and very dynamic step in determining the transcriptional outcome. Accordingly, precisely controlled gene expression may be attributed to the properties of transcriptionally engaged RNAPII properties as well as the coordination with ongoing co-transcriptional processes, both largely influenced by the action of so-called transcript elongation factors (TEFs). However, these regulatory processes are still poorly understood in plants.

In this study the molecular and functional characterisation of TFIIS has been performed by using its dominant negative version (TFIISmut), revealing TFIIS importance in regulating transcript elongation *in vivo*. Additionally, the mutual impact of various TEFs on gene expression and consequently plant development has been addressed.

TFIIS is a well characterised TEF which directly influences RNAPII properties and its molecular function has been extensively studied *in vitro*. Accordingly, TFIIS has been demonstrated to stimulate intrinsic RNAPII endonucleolytic cleavage activity, allowing the control of RNAPII pausing and read-through. Still, the details regarding TFIIS molecular role *in vivo* as well as its biological importance remain largely unknown and those aspects have been addressed in this study by using the inducibly expressed TFIISmut.

The β-estradiol inducible system was successfully adapted in this study to allow temporally controlled expression of TFIISmut in plants lacking functional TFIIS, leading to severe growth defects of created mutants. Observed morphological defects could be subsequently linked to generally compromised transcription triggered by TFIISmut expression, supporting the biological functionality of the created inducible system. Accordingly, the replacement of invariant Asp309 and Glu310 residues renders *Arabidopsis* TFIIS as a negative dominant form which presumably inhibits intrinsic RNAPII cleavage activity *in vivo*. TFIISmut was additionally demonstrated to associate with actively transcribing RNAPII in the chromatin context as a component of the *Arabidopsis* TEC.

In this study several previously unknown characteristic of plant-specific transcript elongation were demonstrated in the context of active RNAPII occupancy. Mutation of TFIIS was subsequently shown to significantly influence genome-wide distribution of active RNAPII leading to its clear enrichment in the promoter-proximal region. Observed RNAPII accumulation has been linked with RNAPII arrest which consequently lowered TEC mobility, impeded cell cycle progression and resulted in the proteasomal degradation of NRPB1. Spatial determination of RNAPII arrest in the presence of mutated TFIIS revealed strong overlap with the position of the +1 nucleosome. Thus, the molecular role of TFIIS *in vivo* may primarily encompass nucleosome traversal by RNAPII but also other pausing-related mechanisms, such as promoter-proximal pausing, transcription fidelity and transcript elongation rate. Thus, TFIIS emerges from this study as a general regulator of many processes related to RNAPII pausing, supporting the previously suggested role of RNAPII backtracking in the regulation of many

cellular processes. Biologically, TFIIS was suggested in this study as important for the adjustment of transcriptomic programs and metabolic adaptation when facing the external stimuli or transition through various developmental stages. Thus, the usage of a negative dominant version of TFIIS framed into an inducible system consists a valuable molecular tool for studding the role of TFIIS *in vivo* as well as the properties of conditionally controlled transcript elongation.

In the second part of the thesis the composition of *Arabidopsis* TEC has been studied in detail in order to better understand the mutual regulation of transcript elongation by different TEFs. The composition of the *Arabidopsis* TEC was determined with the contribution of other lab members. The aspects related to mutual regulation of transcript elongation by TFIIS and PAF1-C were further elaborated in this study. Accordingly, the genetic interaction between TFIIS and ELF7 was demonstrated and could be subsequently associated with the reduced PAF1-C level within the TEC in the absence of N-terminal TFIIS. *Arabidopsis* ELF7 was additionally demonstrated to play a role in the response to replication stress and presumably in the resolution of transcription-replication conflicts (TRCs), similarly to observations in yeast. The potential influence of TRCs accumulation on observed genetic interaction between TFIIS and ELF7 was evaluated in this study. Additionally, mutual TFIIS and ELF7 contributions into other molecular processes have been proposed, including the regulation of nucleosome traversal and RNAPII and promoter-proximal pausing. Consequently, the data in this study suggest that the identified genetic interaction may be a combination of perturbations in PAF1-C level within the TEC as well as other molecular processes synergistically affected by the mutual absence of TFIIS and ELF7. Those findings imply many levels of interaction between various TEFs during transcript elongation.

Taken together, this study encompasses the optimisation and successful implementation of novel molecular tools and approaches in combination with conventional molecular assays and reverse genetic approaches. As a result, TFIIS importance in regulating transcript elongation *in vivo* as well as the mutual contribution of various TEFs into correctly regulated gene expression and plant development could be extensively demonstrated. Thus, findings presented in this study present a valuable insight into the regulation of transcript elongation in plants and additionally elucidate the general principles of gene expression in higher eukaryotes.

7. Materials

7.1 Instruments

Table 18. Instruments used in this study.

Instrument	Manufacturer /model
Immunoblotting system	Semi-dry Blotting System (Carl Roth)
Centrifuges	Sorvall ™ Evolution RC and Sorvall ™ LYNX 4000 with SLA1500 or SS34 rotors (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 ETS 18-55 mm objective (Canon)
FACS	Canto II flow cytometer (BD Biosciences)
Homogenizer	TissueLyser II (Quiagen)
Hybridization Oven	Hybridisierungsöfen (Uniequip)
Imager	BioDocAnalyze System (Biometra), Multimage TM FluorChem FC2 (Alpha Innotech) Typhoon FLA 9500 (GE Healthcare)
Luminometer	LB-960 (Berthold Technologies)
Microscopes	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)
Phosphoscreen	Cyclone Storage Phospho Screen (Packard Instruments Co.)
Phosphor imager	Cyclone™ Storage phosphor imager (Canberra Packard)
Plant incubator	Plant incubator (Percival Scientific)
qPCR cycler	Mastercycler® ep RealPlex (Eppendorf)
Shaking Incubator	Multitron Standart (Infors HT)
Sonicator	Bioruptor® Pico (Diagenode), UW2070 MS73 (Bandelin electronic)
Spektrophotometer	NanoDrop ND-1000 (Peqlab)
Thermocycler	T3000 and T Gradient thermocyclers (Biometra)

7.2 Chemicals and enzymes

Laboratory grade chemicals and reagents were purchased from Applichem (Germany), Carl Roth (Germany), Clonetech, Duchefa (Netherlands), Life Technologies (UK), Merck (Germany), Sigma Aldrich (Germany), US Biologcal (USA) and VWR (USA). Enzymes were purchased from Thermo Fisher Scientific (USA), PEQLAB/VWR (USA) and New England Biolabs (USA). Phosphorus-32 was obtained from Hartman Analytic (Germany).

7.3 Oligonucleotides

Table 19. Oligonucleotides used in this study obtained from MWG eurofins genomics (Germany).

Nr/purpose	Usage	Sequence	Target plasmid/gene
2135/ cloning	Amplification of GFP	TCCCCCGGGATGAGTAAAGGAGAAGAA	pENTR1A:GFP-TFIIS/TFIISmut pGreen0229::pTFIIS:GFP-TFIIS/NLS-dTFIIS pCambia2300::35S:eGFP-TFIIS/dTFIIS/NLS-dTFIIS
2455/ cloning	Amplification of GFP	CATTCCCGGGCCTTTGTATAGTTCATCCAT GCC	pENTR1A:GFP-TFIIS/TFIISmut pGreen0229::pTFIIS:GFP-TFIIS/NLS-dTFIIS pCambia2300::35S:eGFP-TFIIS/dTFIIS/NLS-dTFIIS
3467/ cloning	Amplification of TFIIS (CDS)	CGGGATCCATGGAGAGTGATTTGATTGAT TTG	pENTR1A:GFP-TFIIS/TFIISmut pGreen0229::pTFIIS:GFP-TFIIS/NLS-dTFIIS pCambia2300::35S:eGFP-TFIIS/dTFIIS/NLS-dTFIIS pENTR1A:GS-TFIIS/TFIISmut pGADT7:TFIIS
3468/ cloning	Amplification of TFIIS (CDS)	AAGAATTCCTCAACAGAACTTCCAGTGGT TG	pENTR1A:GFP-TFIIS/TFIISmut pGreen0229::pTFIIS:GFP-TFIIS/NLS-dTFIIS

7. Materials

Nr/purpose	Usage	Sequence	Target plasmid/gene
			pCambia2300::35S:eGFP-TFIIS/dTFIIS/NLS-dTFIIS pENTR1A:GS-TFIIS/TFIISmut pGADT7:TFIIS
3471/ cloning	Amplification of GS tag	GCTCTAGAATGGGCACCCCCGCAGTCA	pGreenII0179:pTFIIS::GS-TFIIS pENTR1A:GS-TFIIS/TFIISmut pGreenII0179:pTFIIS::GS-TFIIS/NLS-dTFIIS
3472/ cloning	Amplification of GS tag	AATCTAGACGGCTCGCGCTGCCCCT	pGreenII0179:pTFIIS::GS-TFIIS pENTR1A:GS-TFIIS/TFIISmut pGreenII0179:pTFIIS::GS-TFIIS/NLS-dTFIIS
2302/ cloning	Amplification of TFIIS promoter	AATTCTCGAGTTTGTGAAAAGCCCATCAA ACTTTGG	pGreenII0179:pTFIIS::GS-TFIIS/NLS-dTFIIS
3470/ cloning	Amplification of TFIIS promoter	AACTCGAGACGTTCCGACAATCCCTAGCT CA	pGreenII0179:pTFIIS::GS-TFIIS/NLS-dTFIIS
4251/ cloning	Amplification of TFIIS promoter	AAA GGA TCC CTC GTC CGC CTG TGA AGC TCT GTG C	pGreen0229::pTFIIS:GFP-TFIIS/NLS-dTFIIS
4252/ cloning	Amplification of TFIIS promoter	AAC CCG GGC GTT CCG ACA ATC CCT AGC TCA AAA AAC	pGreen0229::pTFIIS:GFP-TFIIS/NLS-dTFIIS
3895/ cloning	Amplification of NLS sequence	GATCCAAAGCGCCTCCAAAAAAGAAGAG AAAGGTGAT	pCambia2300::35S:eGFP-NLS-dTFIIS pGreen0229 pTFIIS::GFP-NLS-dTFIIS pGreenII0179:pTFIIS::GS-NLS-dTFIIS
4406/ cloning	Amplification of NLS sequence	CTAGATCACCTTTCTCTTCTTTTTTGGAGG CGCTTTG	pCambia2300::35S:eGFP-NLS-dTFIIS pGreen0229 pTFIIS::GFP-NLS-dTFIIS pGreenII0179:pTFIIS::GS-NLS-dTFIIS
3893/ cloning	Amplification of truncated TFIIS	AACCCGGGGATATCAAAACTAACTGCAAT GCTCAA	pCambia2300::35S:eGFP-NLS-dTFIIS pGreen0229 pTFIIS::GFP-NLS-dTFIIS pGreenII0179:pTFIIS::GS-NLS-dTFIIS pCambia2300::35S:eGFP-dTFIIS
3589/ cloning	Amplification of VPS13	ATCTCGAGAGTGCACCATGTTAGAGTCTT TAGCT	XVE-LexA: VPS-GFP-LUC
3592/ cloning	Amplification of VPS13	TTCCCGGGTAGGATAGCTTCACAGTACTT ATTG	XVE-LexA: VPS-GFP-LUC
3811/ cloning	Amplification of LUC	AACCCGGGGCGGCCGCAATGGAAGACGC CAAAAACATAA	XVE-LexA: VPS-GFP-LUC
3882/ cloning	Amplification of LUC	GGTCTAGATTACACGGCGATCTTTCCGCC CTTCTTGGC	XVE-LexA: VPS-GFP-LUC
2316/ cloning	Amplification of GFP	AATTGAGCTCTCATTTGTATAGTTCATCCA	XVE-LexA: VPS-GFP-LUC
2135/ cloning	Amplification of GFP	TCCCCCGGGATGAGTAAAGGAGAAGAA	XVE-LexA: VPS-GFP-LUC
3467/ cloning	Amplification of TFIIS (CDS)	CGGGATCCATGGAGAGTGATTTGATTGAT TTG	pCambia2300::35S:eGFP-TFIIS
3468/ cloning	Amplification of TFIIS (CDS)	AAGAATTCCTCAACAGAACTTCCAGTGGT TG	pCambia2300::35S:eGFP-TFIIS
3040/ cloning	Amplification of ELF7 (CDS)	GCTCTAGAATGGCGTCGTACCG	pCambia2300::35S:ELF7-mCherry
1680/ cloning	Amplification of ELF7 (CDS)	AATTCCCGGGTCATTCAGAATAATCATCC TCATT	pCambia2300::35S:ELF7-mCherry
4251/ cloning	Amplification of TFIIS promoter	AAA GGA TCC CTC GTC CGC CTG TGA AGC TCT GTG C	pGreen0229::pTFIIS:GFP-TFIIS/dTFIIS
4252/ cloning	Amplification of TFIIS promoter	AAC CCG GGC GTT CCG ACA ATC CCT AGC TCA AAA AAC	pGreen0229::pTFIIS:GFP-TFIIS
4149/ cloning	Amplification of NRPB1 (CDS)	AGGGTACCATGGATACGAGGTTTCCGTTC TCT	pGreen0229::pNRPB1:NRPB1-TagRFP
3058/ cloning	Amplification of NRPB1 (CDS)	CGCGTCTAGAAGGGTTGCCTTTATCATCC TTAC	pGreen0229::pNRPB1:NRPB1-TagRFP
3926/ cloning	Amplification of NRPB1 promoter	CCCTCGAGAGTTTGAAGAATCCTATTGAG CGATCT	pGreen0229::pNRPB1:NRPB1-TagRFP
3927/ cloning	Amplification of NRPB1 promoter	AACATATGGGCGGCTAAGCTCCGATCAAA GACGAAT	pGreen0229::pNRPB1:NRPB1-TagRFP
2300/ cloning	TFIIS point mutations	AGAAGTGCTGCTGCGCCAATGA	TFIIS (AT2G38560)
2301/ cloning	TFIIS point mutations	TCATTGGCGCAGCAGCACTTCT	TFIIS (AT2G38560)
3633/ expression	qPCR, RT-PCR	TTGCATGTCCTCTCCAGTAT	VPS13 (SGD:S000003963)
3634/ expression	qPCR, RT-PCR	CCAAGTCCACATTTTGTTTT	VPS13 (SGD:S000003963)
3641/ expression	qPCR	TACAAATTGAAAACTGCCTC	VPS13 (SGD:S000003963)

Nr/purpose	Usage	Sequence	Target plasmid/gene
3642/ expression	qPCR	TGAGATACCAACTCCTTTGA	VPS13 (SGD:S000003963)
TH526/ expression	qPCR	TGGGAAAGTGTTGCCATCC	GAP (AT1G13440)
TH527/ expression	qPCR	CTTCATTTTGCCTTCAGATTCCTC	GAP (AT1G13440)
TH528/ expression	qPCR	ACCCTTGAAGTGGAAAGCTCC	UBI10 (AT4G05320)
TH529/ expression	qPCR	TTCCAGCGAAGATGAGACGC	UBI10 (AT4G05320)
TH646/ expression	qPCR	AACGTGGCCAAAATGATGC	PP2AA3 (AT1G13320)
TH647/ expression	qPCR	CACATTGTCAATAGATTGGAGAGC	PP2AA3 (AT1G13320)
3155/ expression	ChIP-qPCR	CTCTCTCGAGTCTCTTGCTTCTTCTC	At1g48090
3400/ expression	ChIP-qPCR	GTGTATTAGCTGTTAGGTTTGCACA	At1g48090
3156/ expression	ChIP-qPCR	GATCCGAGAAACGAACCGATTCAAC	At1g48091
3401/ expression	ChIP-qPCR	CTCGGTTCTCATGATACAAACATCCT	At1g48091
3159/ expression	ChIP-qPCR	GCTTCTCCCTGACACTTCTAGATGG	At1g48092
4577/ expression	ChIP-qPCR	GCCTTGCGCTCTACCAACTGAGC	At1g48092
3160/ expression	ChIP-qPCR	GACAGTTGTGTCAAAGAGAGGAGTCA	At1g48093
4578/ expression	ChIP-qPCR	TTTTATAGAAAAAAATCGTACC	At1g48093
3147/ expression	ChIP-qPCR	ACCGCCTTTCCCTCTTTGTCGT	At3g02260
3328/ expression	ChIP-qPCR	CCAACACATTTACGTTCACACAAACC	At3g02260
3148/ expression	ChIP-qPCR	GTCTCAAAGCGTAGCTTGCCAGA	At3g02261
3329/ expression	ChIP-qPCR	TTAATTCAAACGCGACGAGTTTAACCAT	At3g02261
3153/ expression	ChIP-qPCR	CAAACCCAAGAAACCGGTCCACA	At3g02262
3399/ expression	ChIP-qPCR	TCCATGTGTTATTCTAATGATGTGCT	At3g02262
3154/ expression	ChIP-qPCR	GAGTTGATTTCGTCGAGCCACGA	At3g02263
3398/ expression	ChIP-qPCR	TTGGGAAATACTGTAATAAGCTTCCT	At3g02263
2780/ expression	RT-PCR	GCTGGAATCCACGAGACAAC	ACTIN2 (AT3G18780)
2781/ expression	RT-PCR	AAGCCTTTGATCTTGAGAGCTT	ACTIN2 (AT3G18780)
1354/ expression	RT-PCR	GAAGGCGAAGATCCAAGACAAGGAA	UBQ5 (AT3G62250)
1355/ expression	RT-PCR	GGAGGACGAGATGAAGCGTCGA	UBQ5 (AT3G62250)
3471/ expression	RT-PCR, genotyping	GCTCTAGAATGGGCACCCCCGCAGTCA	XVE-LexA:GS-TFIIS/TFIISmut
3472/ expression	RT-PCR, genotyping	AATCTAGACGGCTCGCGCTGCCCCT	XVE-LexA:GS-TFIIS/TFIISmut
1184/ genotyping	genotyping	ATCCTCTGGAATGTTGATAGT	T-DNA insertion *tfIIs-1* (SALK_056755)
1185/ genotyping	genotyping	TTTCCTCTTGTCACTTGCCAT	T-DNA insertion *tfIIs-1* (SALK_056755)
2368/ genotyping	genotyping	ATTTTGCCGATTTCGGAAC	T-DNA insertion SALK LBb1.3 (*tfIIs-1*)
1499/ genotyping	genotyping; RT-PCR	TTTCGAAAGATCCCAACGAAA	XVE-LexA:GFP-TFIIS/TFIISmut
1303/ genotyping	genotyping; RT-PCR	TTTGCAAGAGACCTCAGCTTC	XVE-LexA:GFP-TFIIS/TFIISmut
1618/ genotyping	genotyping	TTGGACCCTTCAATTCGTGATG	T-DNA insertion *elf7-3* (SALK_019433)

Nr/purpose	Usage	Sequence	Target plasmid/gene
1619/ genotyping	genotyping	CCTGGCCCTTTTCTTCCTCA	T-DNA insertion *elf7-3* (SALK_019433)
812/ genotyping	genotyping	GTTGCCCGTCTCACTGGTGA	T-DNA insertion SALK LBb1.3 (*elf7-3*)
2314/ Other	DNA probe for Southern Blot	ATCGGGAAACTACTCACACA	XVE-LexA:VPS13-GFP-LUC, XVE-LexA:LUC
1753/ Other	DNA probe for Southern Blot	AGACCTGCCTGAAACCGAACT	XVE-LexA:VPS13-GFP-LUC, XVE-LexA:LUC

7.4 Vectors

Table 20. List of vectors created in the course of this study.

Nr	Vector	Additional information	Experiment
883	pCambia2300::35S::eGFP-NLS-dTFIIS	Truncated TFIIS CDS	CLSM
882	pCambia2300:35S::eGFP-dTFIIS	Truncated TFIIS CDS	CLSM
1225	pCambia2300:35S::eGFP-TFIIS	TFIIS CDS	FRET
1142	pCambia2300:35S::ELF7-mCherry	ELF7 CDS	FRET
847	pENTR1A	entry vector for gateway cloning	gateway cloning
872	pENTR1A:GFP-TFIIS	TFIIS CDS; donor vector for gateway cloning	gateway cloning
873	pENTR1A:GFP-TFIISmut	TFIIS CDS with piont mutations; donor vector for gateway cloning	gateway cloning
1148	pENTR1A:GS-TFIIS	TFIIS CDS; donor vector for gateway cloning	gateway cloning
1144	pENTR1A:GS-TFIISmut	TFIIS CDS with piont mutations; donor vector for gateway cloning	gateway cloning
864	pENTR1A:LUC	Firefly Luciferase; donor vector for gateway cloning	gateway cloning
866	pENTR1A:VPS-GFP-LUC	Yeast VPS13 gene (SGD:S000003963); donor vector for gateway cloning	gateway cloning
850	pGADT7:TFIIS	TFIIS CDS	Y2H
880	pGreen0179:pTFIIS::GS-NLS-dTFIIS	Truncated TFIIS CDS	AP-MS
844	pGreen0179:pTFIIS::GS-TFIIS	TFIIS CDS	AP-MS
1204	pGreen0229:pNRPB1::NRPB1-TagRFP	NRPB1 CDS	FRAP
1222	pGreen0229:pTFIIS::GFP-NLS-dTFIIS	Truncated TFIIS CDS	*tfIIs elf7* complementation
1221	pGreen0229:pTFIIS::GFP-TFIIS	TFIIS CDS	*tfIIs elf7* complementation
867	XVE-LexA (Figure 69)	destination vector for gateway cloning	β-estradiol inducible system
876	XVE-LexA:GFP-TFIIS	TFIIS CDS; expression vector of gateway cloning	inducible TFIIS expression in *tfIIs-1*, AP-MS, FRAP
877	XVE-LexA:GFP-TFIISmut	TFIIS CDS with piont mutations; expression vector of gateway cloning	inducible TFIISmut expression in *tfIIs-1*, AP-MS, FRAP
1146	XVE-LexA:GS-TFIIS	TFIIS CDS; expression vector of gateway cloning	inducible TFIISmut expression in *tfIIs-1*, AP-MS
1147	XVE-LexA:GS-TFIISmut	TFIIS CDS with piont mutations; expression vector of gateway cloning	inducible TFIISmut expression in *tfIIs-1*, AP-MS
1201	XVE-LexA:LUC	Firefly Luciferase; expression vector of gateway cloning	transcript elongation rate system
1097	XVE-LexA:VPS-GFP-LUC	Yeast VPS13 gene (SGD:S000003963); expression vector of gateway cloning	transcript elongation rate system

Table 21. List of vectors used in this study from lab collection.

Nr	Vector	Additional information	Experiment
810	pCambia2300:35S::eGFP	-	cloning
966	pCambia2300:35S::eGFP-NLS-mCherry	-	FRAP
787	pGBKT7:ELF7	ELF7 CDS	Y2H
921	pGreen0179:35S::NLS-mCherry	-	FRAP, cloning
676	pGreenII0229:pTFIIS::TFIISmut	TFIIS genomic sequence with point mutations	constitutive TFIISmut (genomic sequence) expression in *tfIIs-1*
654	pGreen0179:35S	-	components for single vector comprising β-estradiol inducible system
758	pMDC150:pUBQ	-	
72	pMDC220	-	

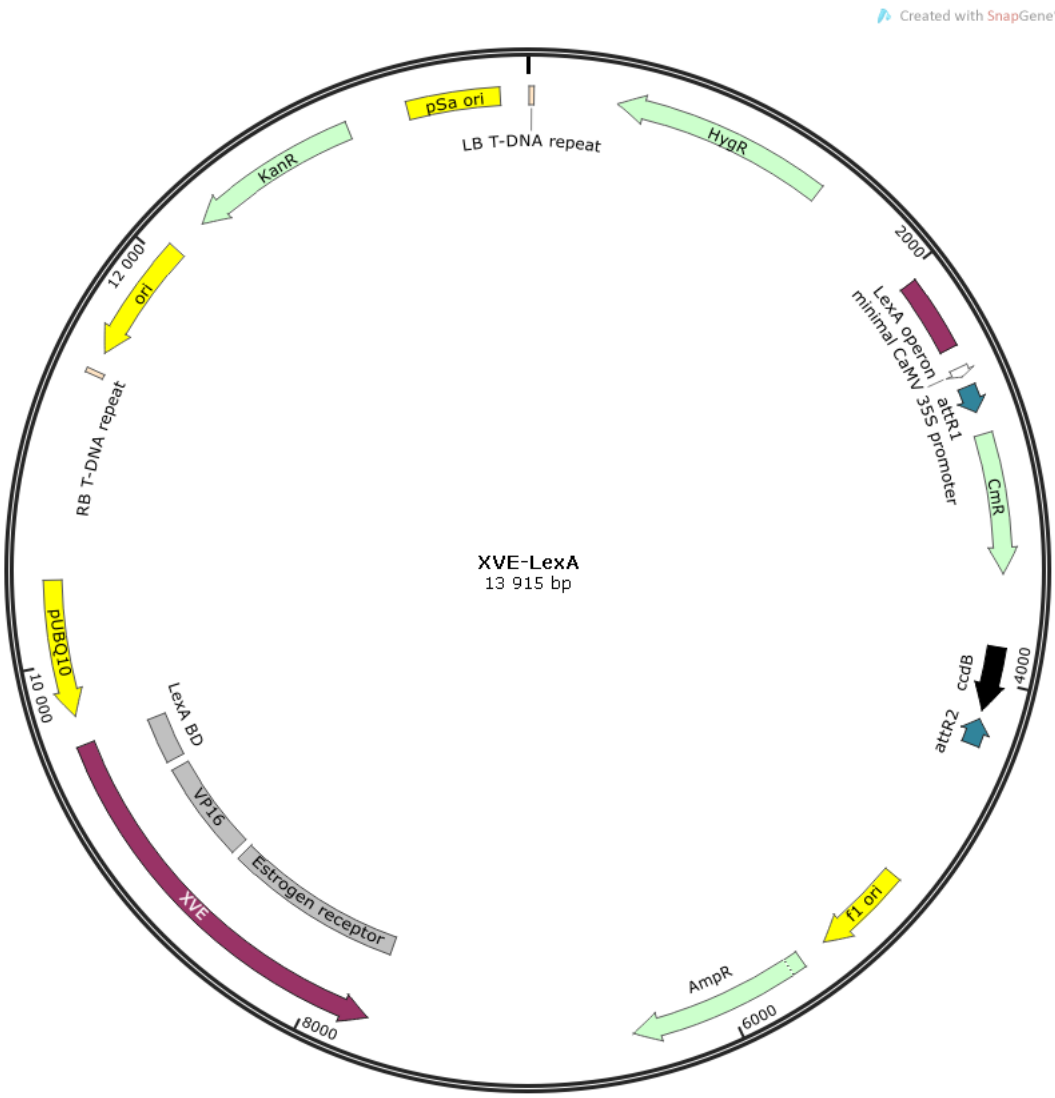


Figure 69. β-estradiol inducible system created in this study (XVE-LexA).
Inducible vector used in this study was obtained by merging pGreen0179:35S, pMDC150:pUBQ and pMDC220 vectors. In brief, pMDC150:pUBQ was cut with KpnI and fragment containing activator unit was aligned with pGreen0179:35S linearized with KpnI. Resulting vector was cut with NdeI and Eco105I and aligned with the responsive unit from pMDC220 cut out with NdeI and Eco136II. pUBQ10: native promoter of *Arabidopsis* Ubiquitin 10; XVE: activator unit. LexA BD: binding domain of LexA operon; VP16: acidic transactivation domain of human VP16; t35S: CaMV 35S terminator; ORI: origin of replication; HygR: hygromycin resistance; AmpR: ampicillin resistance; KanR: kanamycin resistance; CmR: chloramphenicol resistance; LB: left border; RB: right border; attR1/2 and ccdB: elements of Gateway cloning system. Figure was created with SnapGene 2.3.2.

7.5 T-DNA lines and established reporter lines

Table 22. T-DNA lines used in this study.

Name	AGI	T-DNA insertion	Source
tfIIs-1	AT2G38560	SALK_056755	NASC
elf7-3	AT1G79730	SALK_019433	NASC

Table 23. Reporter lines used in this study.

Name	Vector/aim	Genomic background/ref	Provided by
-	pCYCB1;1::CYCB1;1-GFP visualisation of cells at the G2-M phase	Col-0/ (Ubeda-Tomás et al., 2009)	Dr. Peter Doerner
DGU.US-8	DGU.US visualisation of homologous recombination	Col-0/ (Orel et al., 2003)	Prof. Dr. Holger Puchta

7.6 Bacterial and yeast strains

Table 24. List of bacteria and yeast strains.

Organisms	Strain	Resistance	experiment	
E. coli	XL1blue	Tetracycline	Plasmid amplification	Stratagene
A. tumefaciens	GV3101:pMP90 + pSOUP	Gentamycin, Tetracycline	Plant transformation	DSMZ
S. cerevisiae	AH109	-Ade -His -Leu -Trp	Y2H	Clontech

7.7 Databases, Online Tools and Softwares

Table 25. Databases, Online Tools and Softwares.

AgriGO	http://bioinfo.cau.edu.cn/agriGO/)
ClustVis	https://biit.cs.ut.ee/clustvis/
Geneinvestigator	https://genevestigator.com/gv/)
ImageJ 1.49	https://imagej.nih.gov/ij/)
Integrative Genomics Browser https://bioviz.org/	https://bioviz.org/
Leica Application Suite X	Leica Microsystems)
MEME Suite 5.0.4 http://meme-suite.org/	http://meme-suite.org/
Microsoft Excel 2016	https://www.microsoft.com/)
Needle (EMBOSS)	https://www.ebi.ac.uk/Tools/psa/emboss_needle/)
OligoCalc	http://biotools.nubic.northwestern.edu/OligoCalc.html)
PANTHER 14.0	http://go.pantherdb.org/webservices/go/overrep.jsp
PLAZA 4.0	https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v4_dicots/
R environment 3.2.2	https://www.r-project.org)
Saccharomyces Genome Database	https://www.yeastgenome.org/)
SnapGene v2.3.2	http://www.snapgene.com/)
The *Arabidopsis* Information Resource v10	https://www.*arabidopsis*.org/)
UniProt	http://www.uniprot.org/)
Venny v2.1	http://bioinfogp.cnb.csic.es/tools/venny/)
Zotero 5.0.60	https://www.zotero.org/

8. Methods

8.1 Plant work

8.1.1 Cultivation of *Arabidopsis thaliana* on soil

Arabidopsis seeds were sown out on soil (80% (v/v) Einheitserde Typ ED 73, 10% (v/v) supplemented with sand and 10% Isoself R from Knauf Perlite. Mixture was soaked with water containing 0.03% (v/v) confidor WG70 (Bayer) and 3 g/L fertiliser Osmocote Start (The Scotts Company) prior to sowing out. Seed in pots were stratified in the dark conditions for 48 hours at 4 °C. Pots were next transferred into the growing chamber and plants were grown under long-day (LD) conditions (16 h light and 8 h dark) at 22 °C and light intensity 120 µmol m^{-2} s^{-1}. Pots were regularly shuffled on the shelfs to ensure comparable conditions.

8.1.2 Stable transformation of *Arabidopsis* thaliana

Stable transformation of *Arabidopsis* was performed by using the "Floral Dip" method (Clough and Bent, 1998). In brief, plasmids for transformation were introduced into *Agrobacterium tumefaciens* by heat stress induced transformation (8.2.3.2). Colonies grown under respective antibiotic selection were picked from the plates and tested by PCR-based genotyping. 5 mL Luria Bertani (LB) medium with selection was initially inoculated and culture was grown overnight. 500 mL LB medium with selection was next inoculated with overnight culture. After 24 h growth *Agrobacterium* cells were pelleted by centrifugation at 5000 g for 15 min. Cells were re-suspended in 500 mL infiltration medium (5% sucrose, 10mM $MgCl_2$, 0.02% Silwet L-77 and 10 µM acetosyringone). The aerial parts of plant were dipped in the infiltration medium for 1 min and were left overnight wrapped in plastic foil. Dipped plants were transferred into the growth chamber for transgenic seeds production. For the selection of transgenic plants harbouring a bar gene cassette, plants at 7DAS were sprayed once with glufosinate ammonium solution (100 mg/L Basta®, 200 µL/L Silwet® in H_2O) and again after additional three days.

8.1.3 Crossing of *Arabidopsis thaliana*

Plants with different genetic backgrounds were used for crossing to obtain double mutants. Sepals, petals and stamen were gently removed with a tweezer from unopened flowers of plants with one genotype. The remaining carpel was brushed with the pollen from the plant of the second genotype. Plants were transferred into the growth chamber for siliques development. Mature siliques were harvested after 2 - 3 weeks.

8.1.4 Soil-based phenotyping

Plants were cultivated as described above. The morphology of analysed individuals was documented in the context of specific plant characteristics, including bolting time (the emergence of the flower bud), the rosette diameter (at bolting), the number of secondary inflorescence (at 42 DAS) and plant height (at 42 DAS). All pictures were taken with EOS 600D equipped with Macro lens EF-S 60 mm 1:2.8 USM (Canon) or ETS 18-55 mm objective (Canon).

8.1.5 Cultivation of *Arabidopsis thaliana* on plates

For plant growth under sterile conditions seeds were surface sterilized in the exsiccator with a chloric gas (40 mL 12.5% hypochloric acid (w/v) and 2 mL 37% HCl (v/v)). Following sterilisation seeds were immediately sown out on solid 1x MS plates (4.3 g/L Murashige and Skoog media including vitamins, 1% sucrose, 0.8% phyto agar (w/v); adjusted to pH 5.8). Plates were secured with Micropore surgical tape and kept for 2 days in dark conditions at 4 °C for seeds stratification. Plates were next transferred into the plant incubator (Percival Scientific) under LD conditions at 22 °C with light intensity 100 µmol m^{-2} s^{-1}. For the selection of transgenic plants harbouring Hygromycin resistance, MS medium was supplemented with 30 µg/mL Hygromycin B.

In all experiments except plants selection and long-term phenotyping (chapters 2.2.2.4, 3.1.7 and 3.1.9) seedling were grown vertically in the plant incubator on MS plates with 1% phyto agar.

8.1.6 Plants on plates: β-estradiol induction

β-estradiol (Sigma Aldrich) was dissolved in 100% ethanol with the stock concentration 20 mM and kept in -20 °C. Working concentration of β-estradiol used in this study was 20 µM in the experiments related to transcript elongation rate (chapter 2.5). In all remaining experiment β-estradiol was used in the working concentration of 2 µM.

MS plates were prepared as described above. Following MS medium sterilisation, the solution was cooled down and supplemented with β-estradiol to the required working concentration. For studding long-term effects of TFIISmut expression (chapter 2.2.2.4) seeds were sown out on the MS medium supplemented with 2 µM β-estradiol and plates were put horizontally in plant incubator. In all remaining experiments plants were initially grown vertically in the plant incubator on MS plates with 1% phyto agar prior to β-estradiol induction. For the induction plants were transferred with tweezers on MS medium supplemented with β-estradiol. Additionally, β-estradiol stock solution was diluted to the required working concentration in sterile liquid MS medium (as described above but without the addition of phyto agar). Transferred plants were covered with sterile metal grids and submerged in liquid MS medium supplemented with β-estradiol. Opened plates were immediately subjected to the vacuum infiltration for 10-20 min. The excess of MS medium was subsequently removed, and plants were immediately subjected for further analysis or transferred back to the plant incubator for desired induction time.

8.1.7 Cultivation of *Arabidopsis* PSB-D cell culture

Arabidopsis landsberg erecta PSB-D suspension cells culture (Arabidopsis Biological Resource Center) was grown in 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, adjusted to pH 5.7 with 1 M KOH). Cells were cultivated under dark conditions with agitation at 130 rpm and 23 °C. Cells were diluted once a week by transferring 7 mL of old culture into 43 mL of fresh MSMO medium.

8.1.8 Transformation of *Arabidopsis* PSB-D cell culture

Plasmids for transformation were introduced into *Agrobacterium tumefaciens* by heat stress induced transformation (8.2.3.2). 2 mL LB medium with selection was initially inoculated and culture was grown overnight. 20 mL LB medium with selection was next inoculated with overnight culture. After 24 h growth *Agrobacterium* cells were pelleted by centrifugation at 3000 g for 15 min. Cells were re-suspended in 40 mL sterile MSMO. Washing was repeated twice and the OD600 was adjusted to 1.0. *Arabidopsis* PSD-B cells was next co-cultivated with *A. tumefaciens* cells. 3 mL of 3-days old *Arabidopsis* PSD-B culture (OD600: 1.2 - 1.3) was transferred into sterile 6-well plate and mixed with 200 µL of the *A. tumefaciens* solution and 6 µL of 100 mM acetosyringone. Plate was secured with Micropore surgical tape and transferred into the shaking incubator. After 3 days transformed cells were transferred into a 25 mL Erlenmeyer flask containing 8 mL of fresh MSMO supplemented with 20 µg/mL Hygromycin B (for plasmid selection). Additionally, Vancomycin and Carbenicillin were added the final concentration of 500 µg/mL each (negative *A. tumefaciens* selection). *Arabidopsis* PSD-B culture was subsequently grown in the shaking incubator for 7 days. Next, sedimented cells were transferred into 100 mL Erlenmeyer flask containing 35 mL MSMO with selection. Transformed cells were routinely diluted every week as described above with the addition of MSMO with selection.

8.1.9 Upscaling and induction of transformed *Arabidopsis* PSB-D cell culture

Transformed *Arabidopsis* PSB-D cells were cultivated as described above. Cells were gradually upscaled by transferring into increasing volume of MSMO medium every week. First week 50 mL of 1-week old cell culture was transferred into 500 mL Erlenmeyer flasks containing 180 mL MSMO medium with selection. After 7 days cell suspension was divided into 5x 500 mL Erlenmeyer flasks each containing 180 mL MSMO medium. Finally, cell culture from each flask was transferred into 2 L Erlenmeyer flask containing 800 mL MSMO medium. Following 3 days of cultivation in a shaking incubator cells were harvested by filtering the cell suspension through the double layer of Miracloth. Cells were subsequently frozen in liquid nitrogen and stored at -80 °C in the portions of 15 g. *Arabidopsis* PSB-D cells harbouring inducible system was supplemented with β-estradiol to the final concentration of 2 µM 24 hours prior to harvesting.

8.1.10 *Nicotiana benthamiana* infiltration and induction

Plasmids for transformation were introduced into *Agrobacterium tumefaciens* by heat stress induced transformation (8.2.3.2). Colonies grown under respective antibiotic selection were picked from the plates and tested by PCR-based genotyping. 5 mL LB medium with selection was initially inoculated and culture was grown overnight. After 24 h *Agrobacterium* cells were pelleted by centrifugation at 5000 g for 15 min. Cells were re-suspended in 10 mL of infiltration medium (5% sucrose, 10mM $MgCl_2$, 0.02% Silwet L-77 and 10 µM acetosyringone). *A. tumefaciens* suspension was next infiltrated with the syringe in the abaxial side of 2 - 4 weeks old leaves from *N. benthamiana* plants. Plants were transferred back to the greenhouse and analysed by confocal microscopy 2-3 days after infiltration. For the

experiments utilising inducible system, infiltration medium was additionally supplemented with β-estradiol to its final concentration of 2 µM.

8.1.11 GUS staining

For histochemical GUS staining 2-weeks old plants were grown on MS medium as described above. Transgenic plants harbouring GUS-reporter constructs were incubated overnight at 37 °C 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). Following staining plants were washed in 100% ethanol until the leaves were cleared. Leaves were next dissected from the whole plants and mounted on a microscope slide with a cover slip. All pictures were taken with a Zeiss Discovery V8 stereo microscope.

8.2 Microbial work

8.2.1 Cultivation of bacteria

All bacterial strains used in this study were grown on Luria Bertani (LB) medium sterilized by autoclaving (5 g NaCl, 5 g yeast-extract and 10 g trypton). All strains were grown in liquid LB medium under agitation at 200 rpm. For plates preparation LB medium was supplemented with 1.5% w/v agar prior to autoclaving. *E. coli* and *A. tumefaciens* bacteria strains were grown at 37 °C and 30 °C, respectively. For selection, sterile filtered antibiotics was added to the final concentration of 100 µg/mL ampicillin, 50 µg/mL kanamycin, 50 µg/mL gentamycin or 12 µg/mL tetracycline. Antibiotic stock solutions were prepared in H_20 expect tetracycline dissolved in ethanol.

8.2.2 Production of chemically competent *E. coli* and *A. tumefaciens*

5 mL of liquid LB medium containing appropriate antibiotics was inoculated with *E. coli* or *A. tumefaciens* stocks and grown overnight at 37 °C and 30 °C, respectively. 100 mL LB medium supplemented with antibiotics was next inoculated with overnight culture until OD 600 of 0.1 was reached. The pre-culture was subsequently grown to an OD600 of 0.5 - 0.75 and cells were harvested by centrifugation at 4000 g and 4 °C for 10 min. Pelleted cells were next resuspended in 30 mL buffer TBF1 (100 mM RbCl, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 30 mM NaOAc, 15 % (v/v) glycerol, adjusted to pH 5,8 with acetic acid) and incubated for 90 min on ice. The solution was next centrifugated at 2000 g and 4 °C for 10 min and pelleted cells were resuspended in 3 mL of buffer TFB2 (10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2$, 15 % (v/v) glycerol, pH 5,8). Aliquots of 100 µL were frozen in liquid nitrogen for further storage of chemically competent cells at -80 °C.

8.2.3 Transformation by heat shock

8.2.3.1 Transformation of chemically competent *E. coli* cells

Competent *E. coli* cells were thawed on ice and gently mixed with plasmid (~ 500 ng) or the product of ligation reaction. Following 20 min incubation on ice the heat shock was applied at 42 °C for 2 min. Cells were subsequently cooled on ice for 15 min. 1 mL LB medium was added and cells were incubated in a shaking incubator at 37 °C for 1 h. Cells were pelleted by centrifugation at 2000 g for 1min and re-suspended in 25 µL of LB medium. Suspended cells

were next plated out on LB plates containing required antibiotics and incubated overnight at 37 °C.

8.2.3.2 Transformation of chemically competent *A. tumefaciens* cells

Competent *A. tumefasciens* cells were thawed on ice and gently mixed with plasmid (~ 1 µg). Following 5 min incubation in liquid nitrogen the heat shock was applied at 37 °C for 5 min. Cells were subsequently cooled on ice for 10 min. 1 mL LB medium was added and cells were incubated in a shaking incubator at 28 °C for 2 - 4 h. Cells were pelleted by centrifugation at 2000 g for 1min and re-suspended in 25 µL of LB medium. Suspended cells were next plated out on LB-plates containing respective antibiotics and incubated for 48 h at 28 °C.

8.2.4 Cultivation and production of chemically competent yeast

AH109 yeast strain used in this study was grown on plates of solid YPAD medium (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose and 40 mg/L adenine sulfate) at 30 °C for 3 - 4 days. To produce chemically competent yeast cells, a starter culture of 3 mL YPAD inoculated with a single colony of AH109 yeast cells was incubated overnight at 30°C with agitation. 50 mL YPAD medium was mix with the starter culture until OD600 of 0.1 was reached. Cells were subsequently grown in a shaking incubator at 30 °C and 200 rpm until an OD600 of 0.5 - 1. Cells were centrifugated at 500 g for 5 min and washed sequentially with 25 mL sterile H_2O, 5 mL sterile SORB buffer (100 mM LiOAc, 10 mM Tris, 1 mM EDTA, 1 M sorbitol; pH 8.0) and next with 500 µL SORB buffer. Each washing step was followed by the centrifugation at 500 g for 5 min. Pelleted cells were next re-suspended in 360 µL SORB buffer. 40 µL ice-cold denatured single stranded 10 mg/mL salmon sperm DNA was added to the suspended cells. The salmon sperm was denatured at 95 °C and cooled down prior to addition. Aliquots of 50 µL were frozen in liquid nitrogen and stored at -80°C.

8.2.5 Transformation of chemically competent yeast

Competent AH109 yeast cells were thawed on ice and gently mixed with plasmid (~ 500 ng) with addition of 300 µl sterile filtered PEG (100mM LiOAc, 10mM Tris, 1mM EDTA, 40% PEG3350; adjusted to pH 8). The solution was incubated for 30 min at room temperature followed by the addition of 40 µl DMSO. Heat shock was subsequently applied at 42°C for 15 min. Cells were pelleted at 500 g for 2 min and re-suspended in 200 µL H2O. The solution was plated out on selective plates and grown for 3-4 days at 30°C until colonies were visible. Selection was done on synthetic dropout medium (SD) either double dropout (DDO) (2% w/v Glucose, 0.67% w/v Yeast nitrogen base w/o amino acids, 2.2% w/v agar, 0.064% w/v-Leu/-Trp DO supplement; adjusted to pH 5.8 and autoclaved for 15min) or quadruple dropout (QDO) (2% w/v Glucose, 0.67% w/v Yeast nitrogen base w/o amino acids, 2.2% w/v agar,0.064% w/v Ade/-His/-Leu/-Trp/ DO supplement; adjusted to pH 5.8 and autoclaved for 15min).

8.2.6 Yeast-2-Hybrid assay

AH109 yeast cells were co-transformed with required plasmids as described above. Next, single colony identified on double dropout (DD) plate was resuspended in 200 µL H_2O. OD600 was determined with one half of the solution. Remaining yeast suspension was adjusted to

an OD600 of 1.0 and subjected to 1:10 dilution series (10^0, 10^{-1}, 10^{-2}, 10^{-3}). Respective dilutions were placed on DDO and QDO plates using a frogger. Plates were incubated for 3 - 4 days at 30 °C until colonies were visible

8.3 Molecular biology methods

8.3.1 Nuclei acids

8.3.1.1 Isolation of genomic DNA from *Arabidopsis*

Small fragment of *Arabidopsis* leaf was harvested and frozen in liquid nitrogen inside Eppendorf tube with metal bead. Plant material was next homogenized using the Tissue Lyser II (Qiagene) and dissolved in 400 µL of Edward buffer (200 mM Tris HCl, 250 mM NaCl, 0.5% SDS (w/v) and 25 mM EDTA). The sample was briefly vortexed and centrifuged at 12000 g for 5 min. For DNA precipitation, 300 µL of the supernatant were mixed with an equal volume of 100% isopropanol and incubated at room temperature (RT) for 2 min. Precipitated DNA was pelleted by centrifugation at 12000 g for 5 min and washed once with 70 % ethanol (v/v). After ethanol removal DNA was dried and resuspended in 40 µL H_20.

8.3.1.2 Polymerase chain reaction (PCR)

Taq DNA Polymerase (Peqlab) was used for any standard PCR-based validations, including for instance genotyping, expression level studies or colony PCR. KAPA HiFi (Peqlab) DNA Polymerase was used for cloning due to its proofreading activity. PCR cycle program and reagents set-up used for the amplification are indicated in Table 26 and Table 27, respectively. Amplified fragments were analysed on 0,7 – 2 % agarose gels depending on fragment size.

Table 26. PCR cycle programs for Taq and KAPA HiFi.

	Taq		Kapa HiFi		
Step	Temperature	Time	Temperature	Time	Cycles
Initial denaturation	95°C	30 sec	98°C	30 sec	1
Denaturation	95°C	30 sec	98°C	20 sec	1
Annealing	~ 55 - 65°C	30 sec	~ 55 - 72°C	20 sec	25-35
Extension	72°C	1 min/1000bp	72°C	30 sec /1000bp	1
Final elongation	72°C	300sec	72°C	300sec	1

Table 27. Reagents used for Taq and KAPA HiFi.

Reagent	Taq	Kapa HiFi
Buffer	1x Taq reaction buffer (Peqlab)	1x KAPA HiFi Buffer (Peqlab)
dNTP mix	0.2 mM of each dNTP	0.3 mM of each dNTP
Forward primer	0.5 µM	0.25 µM
Reverse primer	0.5 µM	0.25 µM
Polymerase	0.5 U Taq DNA polymerase (Peqlab)	1U KAPA HiFi (Peqlab)
MiliQ water	Up to 25µl	Up to 50µl

8.3.1.3 Agarose gel electrophoresis

0,7 - 2% w/v agarose was dissolved in TAE buffer (40mMTrispH8.0,1mMEDTA) by boiling. Gels were supplemented with 0,005 % (v/v) ethidium bromide. 6x loading dye (250mM Tris 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

DNA or RNA samples to a final concentration of 1x. Samples were loaded into the wells of agarose gel and run at 140 V using 1x TAE buffer. DNA and RNA fragments were visualized using the BioDoc Analyser (Biometra).

8.3.1.4 Cloning

Cloning was performed in accordance to Sambrook et al., 1989. PCR fragments or plasmids were digested with commercial restriction enzymes according to the manufacturer's manual (NEB or Fermentas). Digested insert was mixed with plasmid DNA in 4 : 1 molar ratio with the addition of 5U T4 DNA ligase and 1x T4 Ligase buffer in total volume of 20 µL. Reagents mixture was incubated overnight at 4 °C.

Subcloning of PCR fragments into the pENTR1A® vector for performed with standard restriction enzymes (NEB or Fermentas). Gateway® reactions between entry clone and destination vector were performed according to the manufacturer's manual (Life Technologies).

DNA clean-up was performed with the NucleoSpin® Gel and PCR Clean-up kit according to the manufacturer's manual (Macherey Nagel).

8.3.1.5 Restriction analysis and plasmid DNA amplification in *E. coli*

For restriction analysis of cloned plasmid DNA, 3 mL LB medium with selection was inoculated with transformed *E. coli* colony and incubated overnight at 200 rpm and 37 °C. 1.5 mL of *E. coli* culture was harvested by centrifugation in Eppendorf tube at 2000 g for 3 min. Pelleted cells were resuspended in 200 µL P1 buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 µg/mL RNase A). 300 µL of P2 neutralization buffer (0.2 M NaOH, 1 % (w/v) SDS) was subsequently added to the sample for cells lysis and the mixture was incubated for 5 min at RT. Next, 300 µL of P3 buffer (3 M potassium acetate, pH 4.8) were added to stop the lysis and the sample was incubated for 10 min on ice followed by centrifugation at 12000 g and RT for 10 min. 750 µL of the supernatant was transferred to a new Eppendorf tube containing an equal amount of 100 % (v/v) isopropanol. Mixture was incubated at RT for 5 min. Precipitated plasmid DNA was pelleted by centrifugation at 12000 g and RT for 10 min and washed once with 70% ethanol (v/v). After ethanol removal DNA was dried and resuspended in 100 µL H_20.

For plasmid DNA amplification remaining culture of corresponding *E. coli* colony was used to inoculate 100 mL LB medium. Midi preparation of plasmid DNA was performed using the NucleoBond R Xtra Midi Kit according to the manufacturer's manual (Macherey Nagel).

8.3.1.6 Sequencing

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

8.3.1.7 Southern blot

DNA digested with desired restriction enzymes was separated on 0.7 % agarose gel overnight at 50 V. Resolved agarose gel was stained with ethidium bromide and documented using the BioDoc Analyser (Biometra).The gel was next incubated in 0.25 M HCl for 10 min and

subsequently washed under gentle agitation in denaturation solution (0.5 M NaOH, 1.5 M NaCl) for 30 min followed by washing in neutralisation solution (0.5 M Tris-HCl pH 7.2, 1.5 NaCl, 1 mM EDTA) for another 30 min. Hybond-N membrane (GE Healthcare) was prepared and assembled with the agarose gel according to manufacturer's manual (GE Healthcare). DNA transfer was performed overnight at RT. Following DNA transfer the membrane was rinsed in 2 x SSC and DNA was fixed by UV-crosslinking.

A DNA template of ~ 700 bp was generated by PCR reaction and purified as described before. [α-^{32}P]dATP (3000 Ci/mmol) was incorporated into DNA probe using the Prime-It II Random Primer Labeling Kit (Stratagene) according to manufacturer's manual. Radiolabelled probe was next purified using G50 Sephadex Cloumn (Roche) according to manufacturer's manual. 100 µg/mL salmon sperm was added to the probe and incubated at 95 °C for 5 min.

Membrane was placed in a hybridization tube containing 20 mL of QuikHyb ® Hybridisation solution (Stratagene) and incubated under rotation in the hybridization oven for 30 min at 68 °C. Next, radiolabelled probe was added to the Hybridisation solution and incubated overnight at 68 °C. Membrane was rinsed for 30 min in 2xSSC containing 0.1 % SDS (w/v) at 60 °C. The membrane was assembled with phosphor storage screen in a light excluding cassette and incubated for 24 h. The screen was visualised using a Cyclone TM phosphor imager.

8.3.1.8 Isolation of RNA from *Arabidopsis*

50-100 mg of *Arabidopsis* tissue was harvested and frozen in liquid nitrogen inside Eppendorf tube with metal bead. Plant material was next homogenized using the Tissue Lyser II (Qiagene). The RNA extraction was performed using RNeasy R Mini Plant kit (Qiagen) according to the manufacturer's manual.

8.3.1.9 Reverse Transcription (cDNA synthesis)

2 - 4 µg of extracted RNA was incubated with 2 U of DNaseI (NEB) for 100 minutes at 37 °C according to the manufacturer's manual. 1.5 µg of RNA was next used for cDNA synthesis with RevertAid TM H Minus M-MuLV Reverse Transcriptase and random hexamer primers according to manufacturer's manual (Thermo scientific). In RT-PCR experiments cDNA was synthetized using the Superscript® IV reverse transcriptase and Oligo d(T) 18 primers were used according to manufacturer's description (Thermo scientific). All incubation steps were performed in a thermocycler.

8.3.1.10 Real time quantitative PCR (qPCR)

qPCR reactions were routinely performed with KAPA TM SYBR R 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 manual. Each reaction was performed in the total volume of 10 µL using the program parameters indicated in Table 28.

Table 28. qPCR cycle program.

Step	Temperature	Time	Cycles
Initial denaturation	95°C	180 sec	1
Denaturation	95°C	3 sec	
Annealing	60°C	20 sec	40
Extension	72°C	8 sec	
	95°C	15 sec	1
Melting curve	60-95°C (gradient)	20 min	1
	95°C	15 sec	1

For expression analysis, the normalised relative quantities (NRQ) were calculated using GAPC, PP2AA3 and UBQ10 references genes (Kudo et al., 2016) according to Hellemans et al., 2007. Primer efficiency was calculated using a serial dilutions of cDNA template.

For ChIP analysis, 2 µL of immunoprecipitated DNA (following desired dilution with H_2O) was analysed with locus specific primers. Data was analysed using the $2^{-\Delta\Delta CT}$ method and normalized to the input samples.

8.3.2 Proteins

8.3.2.1 Protein extraction

Whole protein extracts from *Arabidopsis* seedlings were obtained as described in Tsugama et al., 2011. In brief, 2 - 5 seedlings were harvested and frozen in liquid nitrogen inside Eppendorf tube with metal bead. Plant material was next homogenized using the Tissue Lyser II (Qiagene). 200 - 500 µL of extraction buffer (0.1M EDTA (pH 8.0), 4% w/v SDS, 10% v/v β-mercaptoethanol, 5% v/v glycerol, 0.005w/v bromophenol blue) was added to the ground material and the mixture was incubated at 95 °C for 10 min. The solution was stored at -20 °C and directly used for SDS-PAGE.

8.3.2.2 SDS-PAGE

The SDS polyacrylamide gel electrophoresis (SDS-PAGE) was performed with polyacrylamide gels made by pouring the stacking gel on the top of the resolving gel (9%). Resolving gel was made by mixing acrylamide:bisacrylamide (30:0.15), 0.75 M Tris-HCl pH 8.8, 0.2% SDS (w/v), 0.1% ammonium persulfate (APS) and 0.02% N,N,N‟,N‟-tetramethylethylenediamine (TEMED (v/v)) in a Bio-RAD Mini-Protean® 3 Multicaster system (Bio-Rad). The stacking gels were made of 10% acrylamide:bisacrylamide (30:0.8) (w/v), 0.14 M Tris-HCl pH 6.8, 0.23% SDS (w/v), 0.11% APS (w/v) and 0.06% TEMED (v/v). SDS-PAGE was run 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) at 160-200 V.

8.3.2.3 Coomassie Brilliant Blue (CBB) staining

The proteins were visualized by polyacrylamide gel incubation with Coomassie Brilliant Blue (CBB) solution (0.2 % (w/v) CBB G-250, 30% (v/v) ethanol and 10% (v/v) acetic acid) for 30 min with agitation. The excess CBB was removed from the gel by overnight incubation in destaining solution (7.5% v/v ethanol and 5% v/v acetic acid) with agitation. Gels were documented with a digital camera.

8.3.2.4 Western Blot

Prior to protein transfer from polyacrylamide gel onto Hybond LFP 0.2 PVDF (Amersham), blotting membrane was activated by rinsing with 100 % methanol for 30 sec and subsequently equilibrated together with the Whatman paper (Biometra) in blotting buffer (20 % (v/v) methanol, 200 mM glycine, 20 mM Tris, 0.01 % (w/v) SDS). Following SDS-PAGE proteins were transferred in blotting buffer onto PVDF membrane using a Semidry Blotter Maxi (Roth) at 50 mA for 2 - 3 hours according to manufacturer's manual. The membrane was next incubated in blocking buffer (5 % (w/v) skimmed milk powder, 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05 % (v/v) Tween 20) for 1 h at 4 °C with agitation. The primary antibody was added directly to blotting buffer in a 1:2000 dilution and incubated overnight at 4 °C with agitation. α-GFP (ChromoTeq), α-GS (Sigma Aldrich), α-UAP56 (Kammel et al.,2013), α-UBQ, α-CTD, α-CTD-S2P and α-CTD-S5P (Abcam) were used in this study from the lab collection. Following incubation with primary antibodies, the membrane was washed 3x 5 min using 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). Respective secondary antibody was mixed in 1:5000 dilution with blotting buffer and incubated with the membrane for 1 - 2 hours at 4 °C with agitation. For chemiluminescent detection, α-Rabbit, α-Mouse or α-Rat IgG antibodies coupled with horseradish peroxidase (HRP) (Sigma Aldrich) were used in this study. For fluorescent detection, ECL Plex α-Rabbit IgG-Cy3 (Amersham) was used. After the incubation, the membrane was washed again 3x 5 min using washing buffer. The blotting as well as all incubation and washing steps were performed at 4 °C. For chemiluminescent detection SuperSignal R West Pico Chemiluminescent substrate was used according to manufacturer's manual (Thermo Scientific). Chemiluminescence detection was performed using MultiimageTM FlurChem FC2 imager (Alpha Innotech). Fluorescent detection was performed using the laser scanner Typhoon FLA 9500 (GE Healthcare) according to manufacturer's manual.

8.3.2.5 AP-MS

Affinity purification coupled with mass spectrometry (AP-MS) was performed as described in Pfab et al., 2017 with some minor modifications.

GS-affinity purification

15 g of transformed PSB-D cell culture was ground to a fine powder in liquid nitrogen using mortar and pestle. 10 mL extraction buffer (25 mM HEPES-KOH pH 7.4, 100 mM NaCl, 0.05 % IGEPAL CA-630, 1 mM DTT, 2 mM MgCl2, 5 mM EGTA, 10 % glycerol, cOmplete TM EDTA free proteinase inhibitor tablets (Sigma-Aldrich), 1 mM PMSF in 2-propanol) was added to each tube and left rotating in 4 °C for 10 - 20 min. Samples were next sonicated on ice (5x 30 sec at 30% of power with 60 sec intervals) with a Bandeln Sonoplus HD 2070 and a MS 73 tip. Following sociation samples were supplemented with MgCl2 (up to 5 mM) and 50 U/mL Benzonase and incubated for 30 minutes at 4 °C with agitation. The extract was next centrifuged at 40000 g and 4 °C for 1 h and the supernatant was filtered through a 0.45 µm filter. 50 µL of filtered solution was kept as an input sample. 100 µL of IgG-beads pre-washed with extraction buffer was added to each sample and mixture was incubated for 1.5 h at 4 °C with agitation. Beads were pelleted by centrifugation at 2000 g and 4 °C for 5 min and

subsequently transferred into new Eppendorf tube after supernatant removal. Beads were washed three times with 1 mL extraction buffer and proteins were eluted with 300 µL elution buffer (0.1 M glycine-HCl, adjusted to pH 2.7) by agitation at 700 rpm and RT for 5 min. The eluate was precipitated with ice-cold acetone and resuspended in 24 µL 1x PBS. 6 µl of 6x SDS loading buffer (150 mM Tris-HCl pH 7.0, 150 mM DTT, 5% SDS (w/v), 25% glycerol (v/v), and 0.1% bromophenol blue (w/v)) was added and samples were desaturated at 95 °C for 10 min. Obtained AP eluates were stored in -20 °C.

In-gel digestion of purified proteins

AP eluates were resolved with 9 % gel SDS-PAGE and stained with CBB as described above. Each lane was cut out of the gel and divided into 4-8 gel pieces using a scalpel. Each piece was cut into small stripes and transferred into 2 mL Eppendorf tube. The gel pieces were washed sequentially with 50 mM NH_4HCO_3 (60 min), 50 mM NH_4HCO_3/acetonitrile (3/1), 50 mM NH_4HCO_3/acetonitrile (1/1) (30 min) and 50% acetonitrile (w/v) (10 min), respectively. Gel pieces were next dried by 1 h lyophilisation and proteins were in-gel digested by the addition of trypsin (Roche), 0.04 µg/1 µL in 50 mM NH_4HCO_3). Digestion was performed overnight at 37 °C. Peptides were next eluted from the gel by 2-step incubation with agitation: first with 100 mM NH_4HCO_3 at 39 °C for 1 h, and next with 50 mM NH_4HCO_3/acetonitrile (1/1) at 39 - 30 °C for 1 -2 h). The tube with combined extracts was lyophilized overnight.

Mass spectrometry

Mass spectrometry was performed in the laboratory of Dr. Astrid Bruckmann (Department of Biochemistry I, University of Regensburg). The procedure was detailly described in Antosz et al., 2017 In brief, peptides obtained by trypsin digestion were resolved by reversed phase chromatography using UltiMate 3000 RSLCnano System (Thermo Scientific) and Reprosil-Pur Basic C18 nano column. Linear gradient of 4 to 40% acetonitrile in 0.1% formic acid was applied for 90 min. The HPLC-system was coupled to a Q-TOF mass spectrometer (MaXis plus) via a nanoflow electrospray source (Bruker Daltonics). The mass spectrometer was operated in the data dependent mode with MS and MS/MS scans. Searching of NCBI database was acquired with Mascot (v2.3.02) using ProteinScape software (Bruker Daltonics). The label-free quantification was performed by Dr. Rasha ElBashir (University of Regensburg) Peaks Studio® X software (Bioinformatics Solutions, Waterloo, Canada) with standard settings.

8.3.3 Detection of reporter proteins

8.3.3.1 Microscopy

8.3.3.1.1 Structured illumination microscopy and CLSM

Arabidopsis roots (5 - 10 DAS), fragments of *Arabidopsis* leaves (5 DAS), PSD-B suspension cells (3 days old) or fragments of *N. benthamiana* leaves (3 – 4 weeks old) were mounted in H_2O on objective slides with cover slips.

Zeiss Imager M2 ApoTome microscope with 20X/0.8 and 40X/1.4 apochrome objective was used for visualising GFP expression in *Arabidopsis* roots (5 DAS). Pictures were adjusted with Zen 2 blue software. Pictures in this study were generated using ImageJ software version 1.49.

Confocal laser scanning microscopy (CLSM) was performed using Leica SP8 microscope with 40X/1.3 Oil or 63X/1.3 Glycerol objective. For roots imaging Propidium Iodide (PI) was used as a counterstaining according to manufacturer's manual. GFP was excited using an Argon laser at 488 nm and mCherry/TagRFP/PI were exited using a DPSS laser at 561 nm. GFP emission was detected with Hybrid detector at 500 - 550 nm. mCherry/TagRFP/PI emission was detected with PMT detector at 580-630 nm. Pictures this study were generated using X-Las software.

8.3.3.1.2 FRET

FRET measurements and the calculation of FRET efficiencies was performed as described in Weidtkamp-Peters and Stahl, 2017. Fragments of *N. benthamiana* leaves (3 - 4 weeks old) were mounted in H_2O on objective slides with cover slips. FRET was performed using Leica SP8 microscope with 63X/1.3 Glycerol objective. GFP was excited using an Argon laser at 488 nm and mCherry was excited and bleached using a DPSS laser at 561 nm. A circular ROI area of 10 µm was defined for bleaching performed at 100% laser power with 60-80 iterations. Pictures in this study were generated using X-Las software. All images were acquired with experimentally pre-defined settings: 128 x 128 pixel, 1200 Hz, no line averaging, sequential scan mode, pinhole 3, 8 – 10 x zoom, PMT detector gain of 800 - 1000 V.

The mean FRET efficiency was calculated as a ratio of average post-bleach subtracted with averaged pre-bleach intensity (10 iterations) subsequently divided by average pre-bleached intensity (10 iterations). Pictures in this study were generated using X-Las software.

8.3.3.1.3 FRAP

FRAP was performed using Leica SP8 microscope with 63X/1.3 Glycerol objective. GFP was excited and bleached using an Argon laser at 488 nm and TagRFP was excited and bleached using a DPSS laser at 561 nm. A rectangle ROI area of ~ 9 μm^2 was defined for bleaching performed at 100% laser power with 6 - 8 iterations. 50 pre-bleach and 100 post-bleach images were taken with 1 - 2 % laser power. All images were acquired with experimentally pre-defined settings: 128 x 128 pixel, 1800 Hz, no line averaging, bi-directional scanning mode, pinhole 3, 8 – 10 x zoom, PMT detector gain of 800 - 1000 V. The raw measurements were processed using the easyFRAP software with double full-scale normalization (Rapsomaniki et al., 2012). Obtained normalized values were used to determine half life time and the mobile fraction. Pictures in this study were generated using X-Las software.

8.3.3.2 FACS

Fluorescence-activated cell sorting (FACS) and initial data analysis was performed by Dr. Jörg Fuchs (IPK, Gatersleben).

Arabidopsis seedling used for the analysis (10 - 21 DAS) were grown as described above (8.1.5). β-estradiol induction as acquired as described above (8.1.6). Entire 10DAS seedlings (2.4.3) or aerial parts of 21DAS seedlings (3.2.4) were chopped with a razor blade in a nuclei isolation buffer (45 mM $MgCl_2$, 30 mM sodium citrate, 20 mM MOPS and 0.3% Triton X-100, pH 7.0) as described in Galbraith et al., 2011. Before sorting, the nuclei solution was stained with DAPI (4',6-diamidino-2-phenylindole, 2 µg/mL). Samples were analysed in a Canto II flow

cytometer (BD Biosciences). Samples were sorted at a pressure of 35 psi using a FACSVantage SE (BD Biosciences) equipped with a UV laser (Coherent Enterprise II 621, Laser Innovations). Nuclei with the expression of inducible proteins were identified and sorted according to GFP signal revealed by Alexa 488-azide in the B530 detector. DNA content was determined by DAPI fluorescence (FL5 detector 505SP filter, BD Biosciences)

8.3.3.3 Bioluminescence

Transgenic plants carrying the luciferase gene under the control of β-estradiol inducible system were grown as described above (8.1.5). 6DAS seedlings were transferred to a 96-well plate supplemented with liquid MS medium containing 60 µM liquid solution of D-luciferine (Promega) and 60 µM β-estradiol (Sigma). Entire plant or dissected root tissue were submerged in medium and vacuum infiltrated for 10 min. Bioluminescence signal was monitored using a microplate luminometer LB-960 (Berthold Technologies) and the software MikroWin 2000, version 4.34 (Mikrotek Laborsysteme) in the laboratory of Dr. Jan Medenbach. Signal was sequentially collected from each well for 3 sec and repetitive measurement was set-up every 10 min over the period of 2 hours. The raw data containing the absolute values was exported from the software and further analysed. For signal normalization the fresh weight of roots was determined with micro scale. Leaves area was determined with with ImageJ 1.49 following their documentation with digital camera.

8.4 Sequencing-based methods

8.4.1 RNA-seq

8.4.1.1 Total RNA isolation

iGFP-TFIIS and iGFP-TFIISmut lines (2.3.1).

Total RNA was obtained from 10DAS seedlings grown vertically on sterile plates with MS medium as described above (8.1.5). For the induction plants were exposed to β-estradiol for 24 hours prior to material harvesting as descried above (8.1.6). Total RNA was extracted using RNeasy R Mini Plant kit (Qiagen) according to the manufacturer's manual.

Col-0, *tfIIs-1, elf7-3* and *tfIIs elf7* (3.1.8).

Total RNA was obtained from aerial parts of 3-weeks old plants grown on soil under long days conditions as describe above (8.1.1). Total RNA was extracted using RNeasy R Mini Plant kit (Qiagen) according to the manufacturer's manual.

8.4.1.2 RNA-seq data analysis

cDNA libraries preparation and sequencing were performed by The Kompetenzzentrum Fluoreszente Bioanalytik (KFB, University of Regensburg). cDNA libraries were created using TruSeq Stranded mRNA Sample Preparation kit (Illumina) and sequenced on Illumina HiSeq 1000. The analysis of RNA-seq data was performed by Dr. Kevin Begcy (University of Regensburg, currently at the University of Florida).

Reads were mapped to the *Arabidopsis* reference transcriptome assembly (TAIR10) using TopHat 2.1.1 and Bowtie 2.3.4. The total mapped read numbers for each transcript were

determined and normalized using Cufflinks 2.2.1 to detect Fragments Per Kilobase Of Exon Per Million Fragments Mapped (FPKM; iGFP-TFIIS/iGFP-TFIISmut) or Transcripts Per Kilobase Million (TPM; Col-0, *tfIIs-1, elf7-3, tfIIs elf7*). Genes with FPKM ≥ 5 or TPM ≥ 3 were considered transcriptionally active. For the analysis of differentially expressed genes (DEG), the values of $\log_2$ FPKM and $\log_2$ TPM) were calculated. For pairwise comparisons, differentially expressed genes were identified using DEseq2 by analyzing the number of reads aligned to the genes. The thresholds for differential expression were set at fold change greater than 2 and p-value < 0.001 (after the false discovery rate (FDR) adjustment for multiple testing < 0.001) for the null hypothesis. Hierarchical clustering analysis was performed using complete linkage and Euclidean distance as a measure of similarity to display the expression patterns. All DEG analysis and hierarchical clustering were conducted using R environment 3.2.2.

8.4.2 ChIP-seq

8.4.2.1 Material preparation

10DAS seedlings grown vertically on sterile plates with MS medium as described above (8.1.5) were used for ChIP assay. For the induction plants were exposed to β-estradiol for 24 hours prior to material harvesting as descried above (8.1.6). Plant material was harvested in the amount of 1 g and subjected to DNA crosslinking with 1% formaldehyde as described in Pfab et al., 2017. Immunoprecipitated DNA was obtained as described in Pfab et al., 2017 with some minor modifications.

Chromatin Preparation

Plant material was transferred into the Tissue Lyser II adapters (Qiagene) prechilled in liquid nitrogen and samples were homogenize with frequency 30 Hz for 1 min. Ground tissue was added to 30 mL Extraction Buffer 1 in a 50 mL Falcon tube and incubated on a rotating wheel at 4 °C for 10 - 20 min. Solution was filtered through a double layer of Miracloth into a new 50 mL Falcon tube and centrifuge at 3000 g and 4 °C for 20 min. Pellet was resuspended in 1 mL Extraction Buffer 2, transferred into 1.5 mL tube and centrifuge at 12000 g and 4 °C for 10 min. Washing was repeated 2 - 3 times. Pellet was next resuspended in 400 µL Extraction Buffer 3 and another 400 µL of Extraction Buffer 3 was added to a new 1.5 mL tube and overlaid with the pellet from the previous step. Following centrifugation at 16000 g and 4 °C for 1 h pellet was resuspended in 400 µL Nuclei Lysis Buffer and incubated on ice for 30 min. Each sample was sonicated using Bioruptor® Pico (Diagenode) with 10 cycles of 30 sec on/30 sec off. Chromatin solution was centrifuged twice at 14.000 rpm and 4 °C for 10 min. Supernatant was transferred into new 1.5 mL tube containing agarose beads (Sigma Aldrich) pre-washed in ChIP Dilution Buffer according to manufacturer's manual. The mixture was incubated at 4 °C with agitation for 2 hours to reduce the background signal. Beads were routinely pelleted by centrifugation at 2000 g for 1 min.

Immunoprecipitation (IP)

The supernatant was transferred into a new 13 mL tube and diluted 10x with ChIP Dilution Buffer. Chromatin solution was divided into several 1.5 Eppendorf tubes with 500 - 1000 µL per IP. Desired antibodies were added including α-H3, α-CTD-S2P, α-CTD-S5P (Abcam, 5 µg

each) or GFP-trap (Chromoteq, 30 μL). No antibodies were added as a negative control (NoAb). Samples were left rotating overnight at 4 °C.

Washing and Elution

Chromatin solution with antibodies was added to Protein A agarose beads (Millipore) pre-washed in ChIP Dilution Buffer according to manufacturer's manual. Following 1 - 2 hours rotation at 4 °C beads were pelleted and sequentially washed with 0.5 mL of washing buffers (buffers 6-9 below). After each washing step samples were left rotating at 4 °C for 10 min. Following last washing step beads were pelleted by centrifugation and DNA was eluted by adding 50 μL freshly made Elution Buffer following supernatant removal. Samples were nect incubated at 65 °C for 15 min with gentle rotation and the supernatant above pelted bead was carefully transferred into new 1.5 mL tube.

DNA Extraction

Crosslinking was reversed by adding 4 μL of 5 M NaCl to the 100 μL of eluate followed by overnight incubation at 65 °C with agitation. Next, 2 μL 0.5 M EDTA, 1.5 μL 3 M Tris-HCl pH 6.8 and 2 mg proteinase K was added to the eluate and samples were incubated at 45°C for 3 h with agitation. DNA was purified with NucleoSpin® Gel and PCR Clean-up kit according to the manufacturer's manual for DNA clean-up of samples containing SDS (Macherey Nagel).

Buffers used during ChIP procedure:

1) Extraction Buffer 1: 0.4 M Sucrose, 10 mM HEPES pH 8, 5 mM β-mercaptoethanol,
2) Extraction Buffer 2: 0.25 M Sucrose, 10 mM HEPES pH 8, 1 % Triton X-100, 10 mM $MgCl_2$, 5 mM β-mercaptoethanol
3) Extraction Buffer 3: 1.7 M Sucrose, 10 mM HEPES pH 8, 0.15 % Triton X-100, 2 mM $MgCl_2$, 5 mM β-mercaptoethanol
4) Nuclei Lysis Buffer: 50 mM HEPES pH 8, 10 mM EDTA, 0.1 % SDS
5) ChIP Dilution Buffer: 1.1 % Triton X-100, 1.2 mM EDTA, 16.7 mM HEPES pH 8, 167 mM NaCl
6) High Salt: 500 mM NaCl, 0.1 % SDS, 1 % Triton X-100, 2 mM EDTA, 20 mM HEPES pH 8
7) Low Salt: 150 mM NaCl, 0.1 % SDS, 1 % Triton X-100, 2 mM EDTA, 20 mM HEPES pH 8
8) LiCl Wash: 50 mM HEPES pH 8, 0.25 M LiCl, 1 % NP-40, 1 % sodium deoxy., 1 mM EDTA
9) TE Buffer: 10 mM Tris-HCl pH 8, 1 mM EDTA
10) Elution Buffer 1 % SDS, 0.1 M $NaHCO_3$

Buffers 1-5 were additionally supplemented with β-Glycerophosphate (Sigma-Aldrich), cOmplete™ EDTA free proteinase inhibitor tablets (Sigma-Aldrich) and PhosSTOP ™ (Sigma-Aldrich) according to according to the manufacturer's manual.

8.4.2.2 Sequencing and data analysis

Libraries preparation and sequencing was performed by the Lausanne Genomic Technologies Facility (Switzerland). Initial ChIP-seq data analysis was performed by Dr. Jules Deforges (University of Lausanne).

DNA sequencing

Libraries were created using DNA SMART ChIP-Seq kit (Clontech) and subsequently sequenced on Illumina HiSeq 2500. Reads were trimmed to remove adaptor contaminants with

Trimmomatic and mapped to the *Arabidopsis* reference genome assembly (TAIR10) using Bowtie2.

RNAPII density plots

The reads from 3 biological replicates were merged and read density per nucleotide was calculated with bamCoverage from the DeepTools suit. Meta-analysis of RNAPII distribution around transcription start site (TSS; [-2000;2000]), transcription termination site (TES; [-2000;2000]) as well as the relative distribution over transcribed loci (TSS-TES) was performed using computeMatrix from the DeepTools suit with maxThreshold cutoff = 500.

RNAPII occupancy

RNAPII read counts within gene body were determined for each gene in the *Arabidopsis* reference genome (TAIR10) using HTSeq and raw read counts were normalized using Deseq2. Calculated values were averaged for three biological replicates and the fold change (FC) between applied conditions (β-estradiol vs mock) was subsequently calculated and expressed as $\log_2$ FC. Gene were considered with significantly different occupancy if adjusted p-value < 0.1 and $\log_2$FC > 1. Differential analysis was performed using DESeq2 package from R environment. Principal component analysis (PCA) was performed with plotPCA and ggplot R packages. ChIP-seq tracks showing RNAPII-S2P and RNAPII-S5P reads over single genes were generated using Integrative Genomics Browser (IGB).

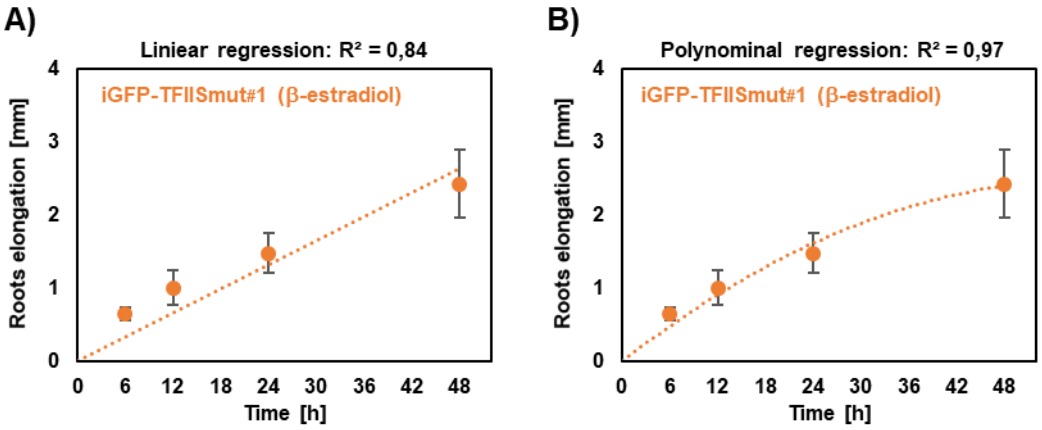
PPEP calculation

RNAPII read counts from three biological replicates were summed up for each condition. The number of tags was subsequently determined for each gene in the promoter-proximal region (0 to +500 bp with respect to the transcription start site) and gene body (+500 bp to gene end). Genes with 10 or more tags in the promoter region were selected and the number of tags per 500 bp in the gene body was determined. Subsequently, the ratio of the number of tags in the promoter and in the gene body was calculated and the $\log_2$ ratio was calculated between samples exposed to β-estradiol vs mock ($\log_2$ PEPP). Z-score of $\log_2$ PEPP was calculated in accordance to Juntawong et al., 2014 and genes were considered significantly different if |z-score| > 2.

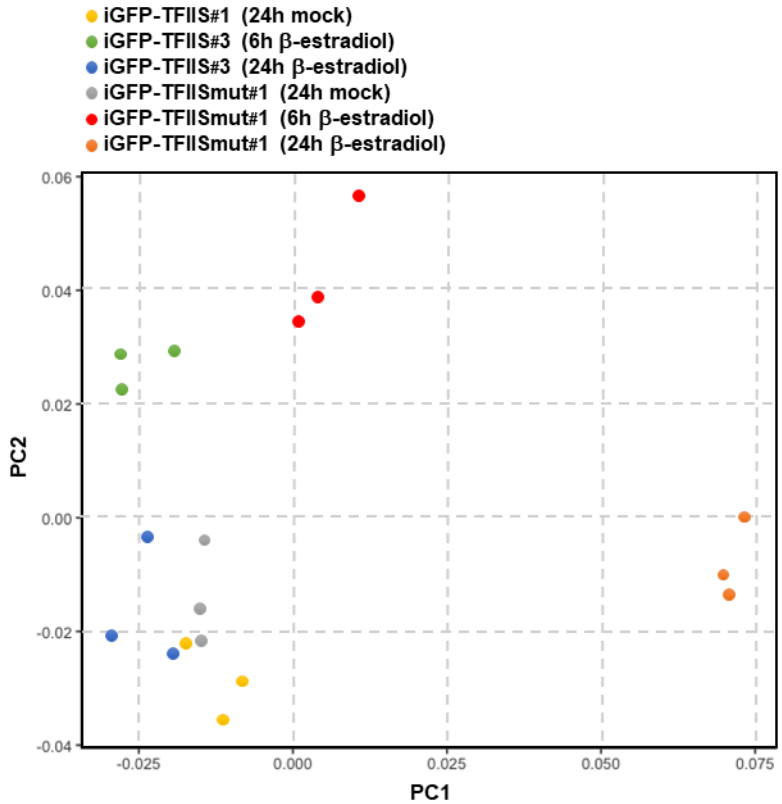
Nucleosome occupancy

Nucleosome occupancy was determined and visualized using computeMatrix (DeepTools suit) based on the publicly available MNAse-seq obtained from Li et al., 2014. Data reflects nucleosome occupancy in aerial parts of 14DAS Col-0 seedling. For the comparison with RNAPII profiles obtained in this study, plots were separately scaled to 1 for each data set.

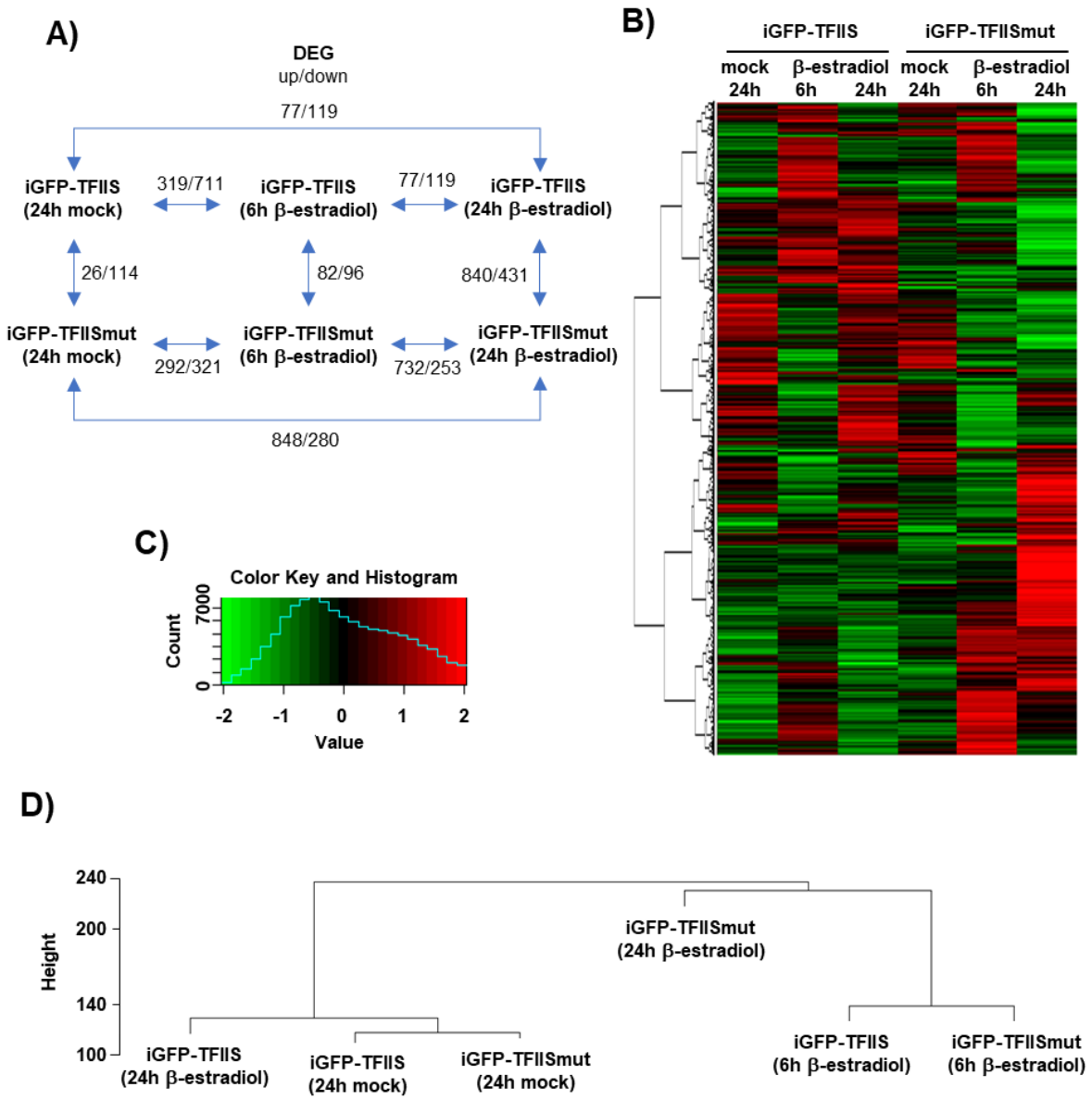
9. Supplements



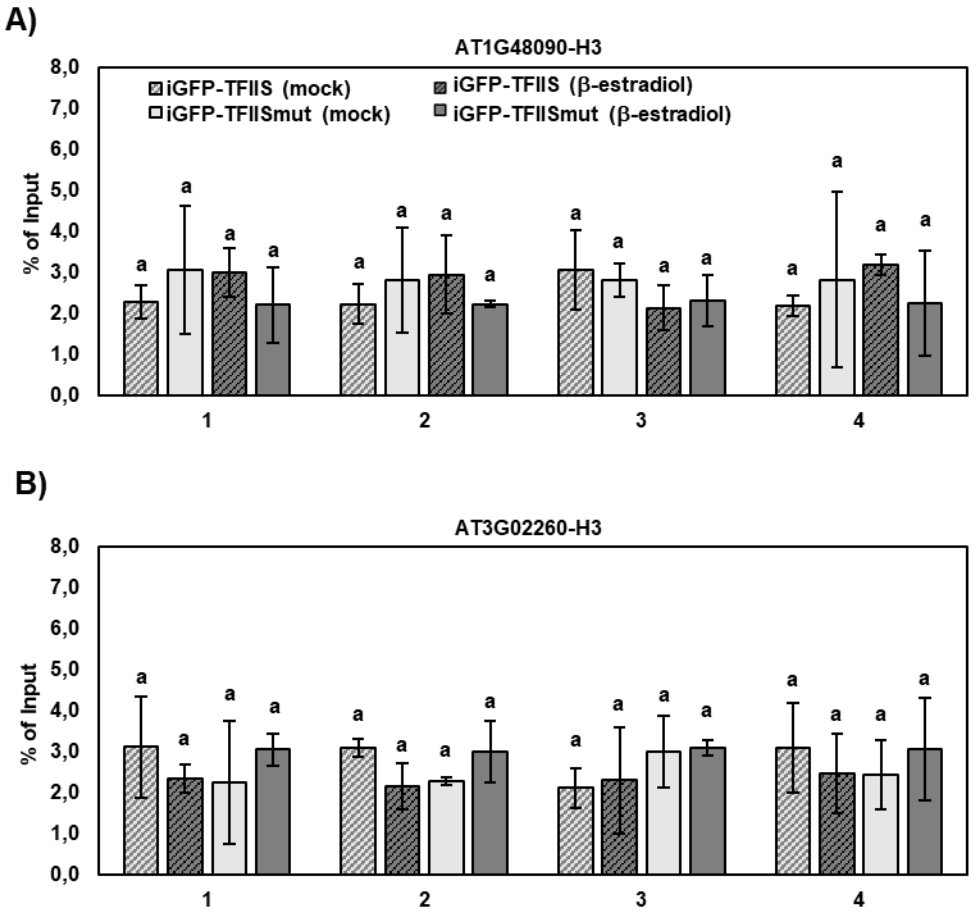
Supplementary Figure S 1. The regression analysis of main root growth kinetic in iGFP-TFIISmut exposed to β-estradiol.
Main root growth kinetic in iGFP-TFIISmut exposed to β-estradiol is best predicted by polynomial regression. The length of main root elongation depicted on Figure 16 was measured by ImageJ in order to calculate the absolute elongation over time. Dotted lines reflect linear (A) or polynomial (B) regression; R^2: the coefficient of determination; n = 6 for each timepoint. Dots indicate mean values ± SD (error bars).



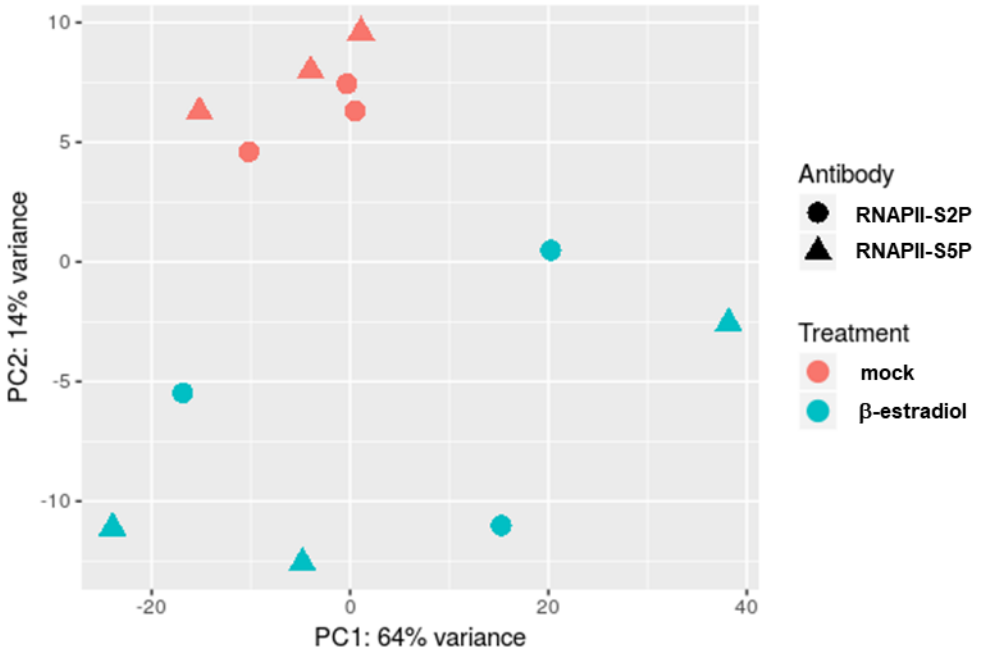
Supplementary Figure S 2. PCA analysis of datasets obtained by RNA-seq (chapter 2.3.1).
The analysis of individual biological replicates was performed by Dr. Kevin Begcy in R environment.



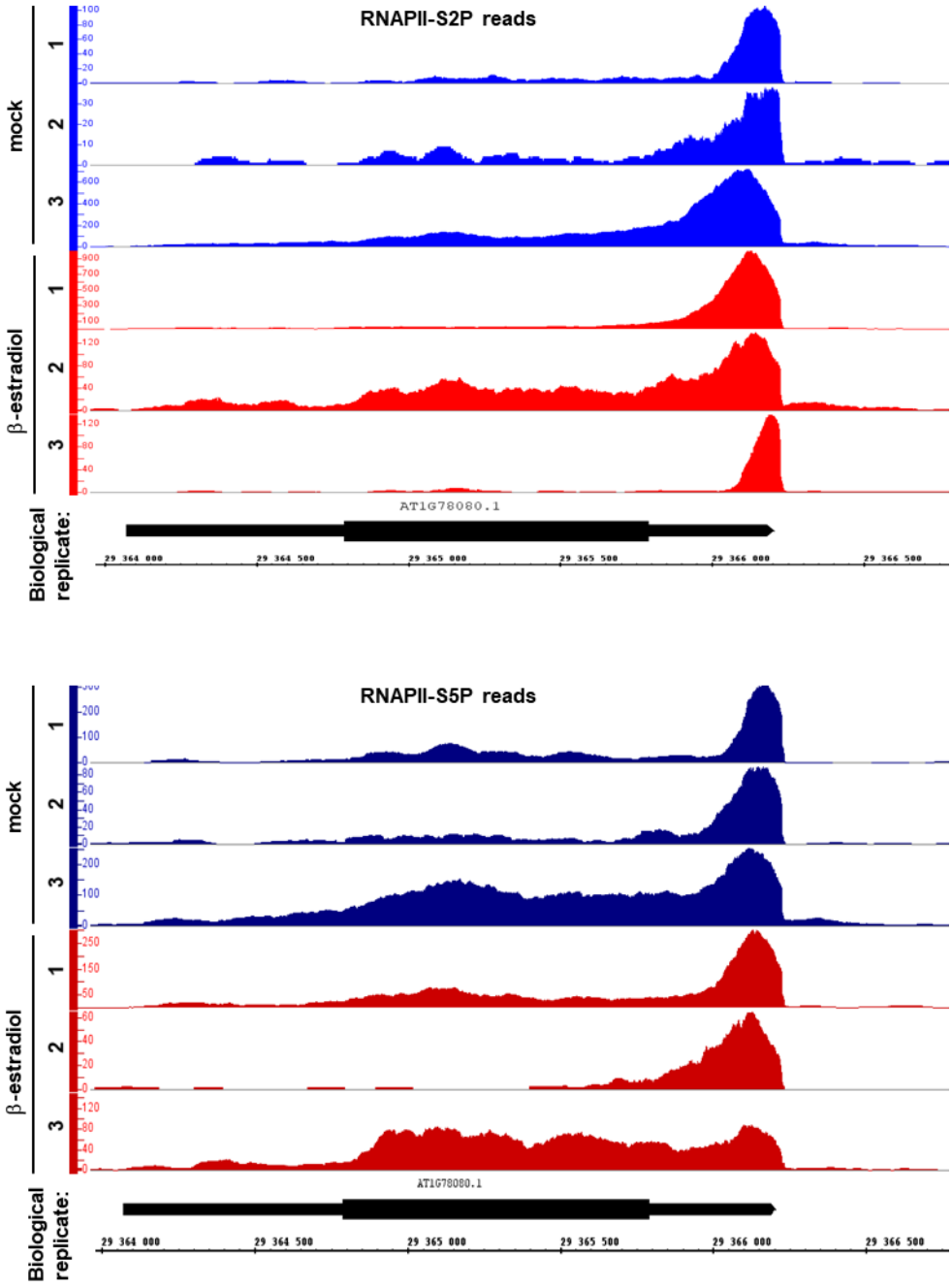
Supplementary Figure S 3. The analysis of transcriptomic changes upon TFIISmut expression over time.
(A) Schematic illustration of DEG analysis between studied lines/conditions with additional 6h β-estradiol induction. Number indicate DEGs ($|log_2FC| > 1$, p-value < 0.001 and FDR < 0.001). (B-C) Heatmap visualisation of analysed lines/conditions. Only transcriptionally active genes with FPKM ≥ 5 in at least one line/condition were considered to build the heatmap (n=15836). Hierarchical clustering is shown on the left. (C) Heatmap legend. Red-green gradient indicated the log_2FPKM in the [-2; 2] colour range (Value). Blue line indicates the number of values in the given colour range. (D) Dendrogram for RNA-seq analysis performed for the averaged biological replicates (n = 3). The cluster analysis was performed through a complete linkage clustering. Figures B-D were created by Dr. Kevin Begcy.



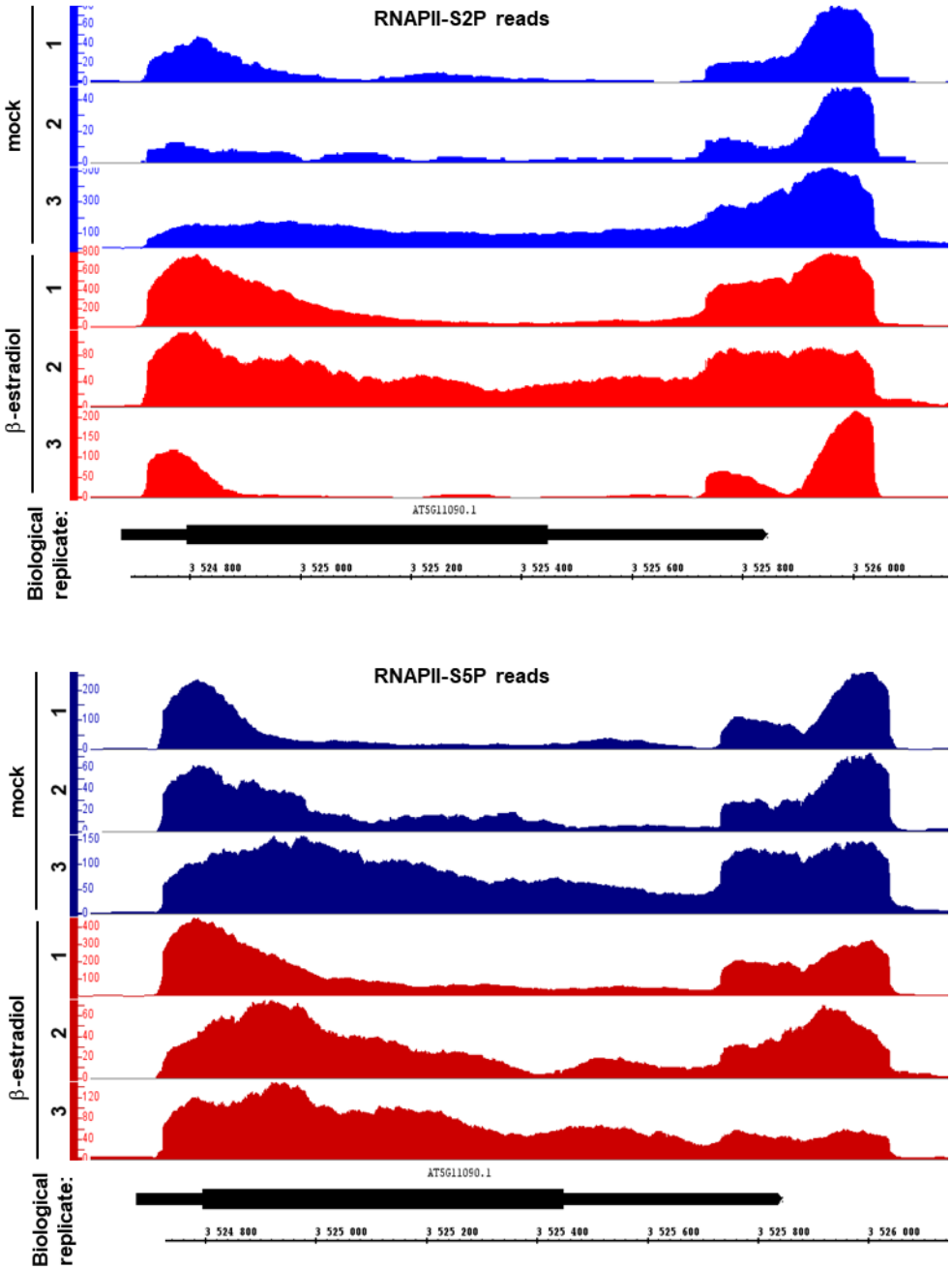
Supplementary Figure S 4. ChIP-qPCR analysis in terms of H3 level across analysed loci.
(A-B) ChIP analyses reveals H3 level at At1g48090 (A) and At3g02260 (B) loci determined in iGFP-TFIISmut line. For the ChIP experiments percentage of Input was determined by qPCR and statistically analysed between samples using one-way ANOVA. The letters above the histogram bars indicate the outcome of a multi comparisons Tukey's test (p-value < 0.05). Error bars indicate SD of at least two biological and two technical replicates. Numbers on x-axis corresponds to the relative positions of the regions analysed by qPCR depicted on Figure 22.



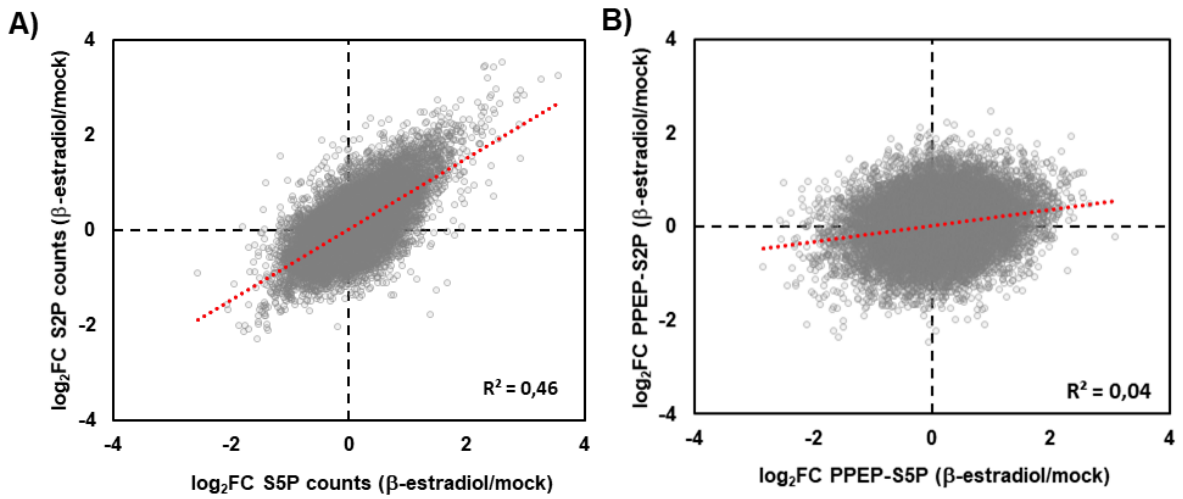
Supplementary Figure S 5. PCA analysis of datasets obtained by ChIP-seq for individual biological replicates. The distribution pattern of active RNAPII upon TFIISmut expression is reproducible across individual samples representing biological replicates exposed to β-estradiol (n = 3) and cluster separately from mock-induced samples (n = 3). The analysis was performed using R environment.



Supplementary Figure S 6. ChIP-seq tracks of individual biological replicates over AT1G78080.
ChIP-seq tracks showing RNAPII-S2P and RNAPII-S5P reads over AT1G78080 gene. Plots were generated in Integrated Genome Browser (IGB). Gene model = thin black bars: UTRs; thick black bars: exons; black line: introns. Number on the left indicate absolute signal intensity. Number below indicate the relative position in the genome.

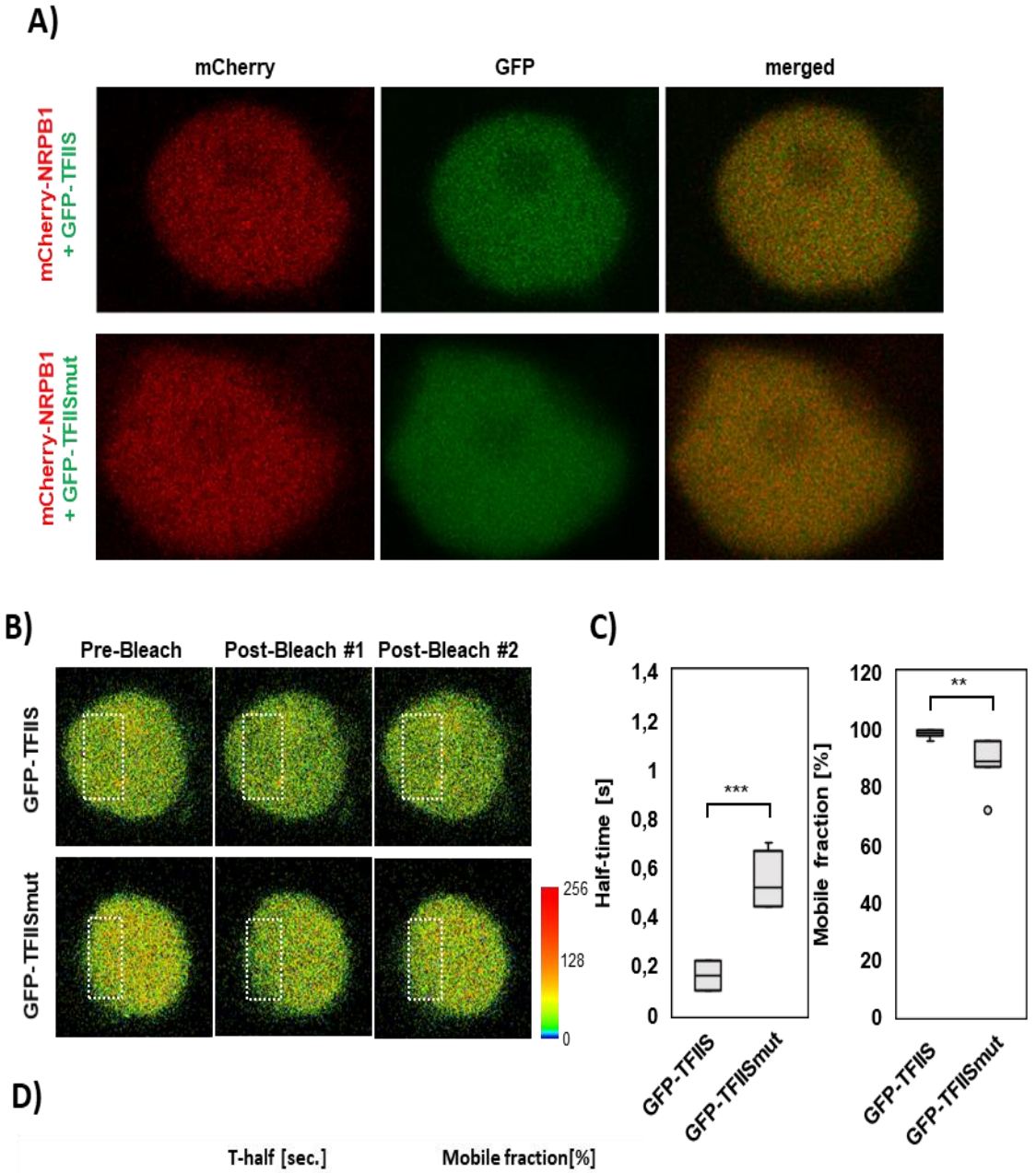


Supplementary Figure S 7. ChIP-seq tracks of individual biological replicates over AT5G11090.
ChIP-seq tracks showing RNAPII-S2P and RNAPII-S5P reads over AT5G11090 gene. Plots were generated in Integrated Genome Browser (IGB). Gene model = thin black bars: UTRs; thick black bars: exons; black line: introns. Number on the left indicate absolute signal intensity. Number below indicate the relative position in the genome.



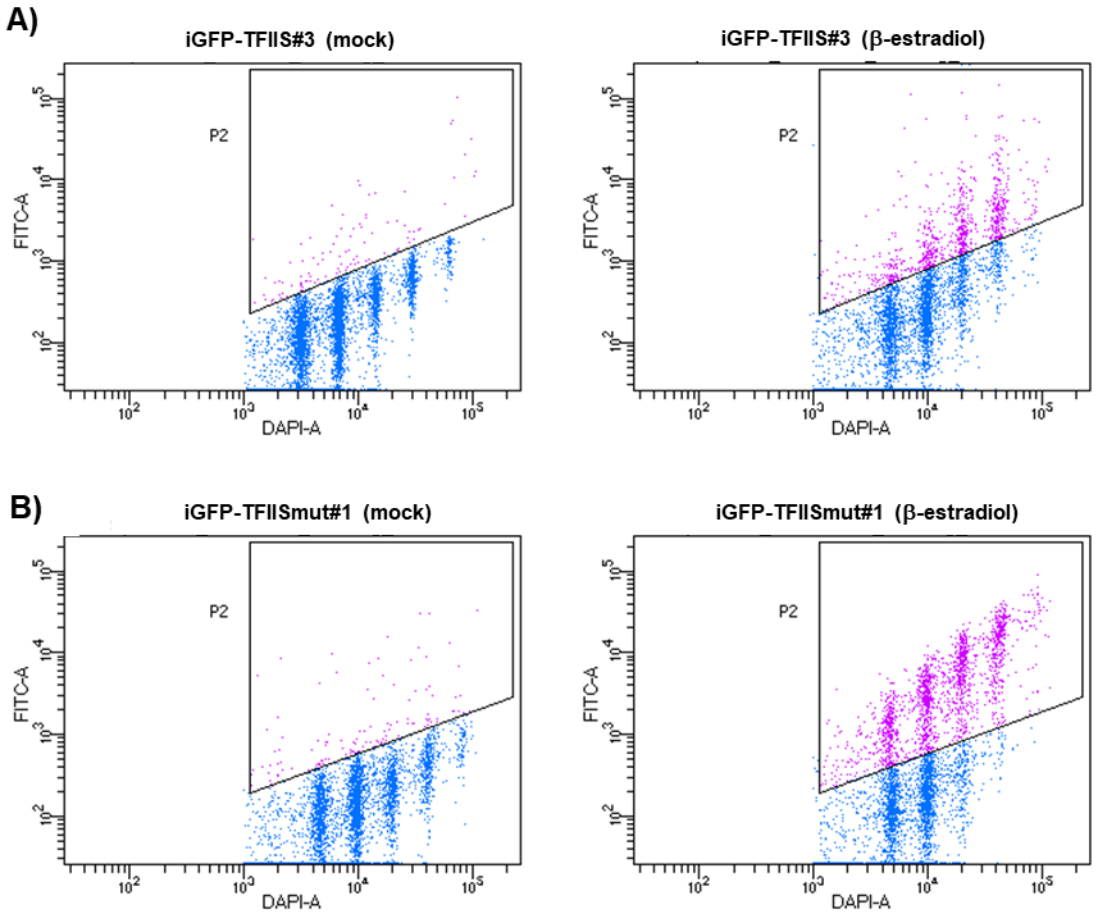
Supplementary Figure S 8. The correlation between RNAPII-S2P and RNAPII-S5P occupancy changes upon TFIISmut expression.

Pair-wise scatter plot analysis of RNAPII-S5P vs RNAPII-S2P counts (A) and PPEP (B) fold changes upon β-estradiol induction relatively to mock. Dotted red line is the best-fit linear regression with a positive slope (n=16482, B). R^2 = coefficient of determination. PPEP: promoter proximal enrichment of active RNAPII.

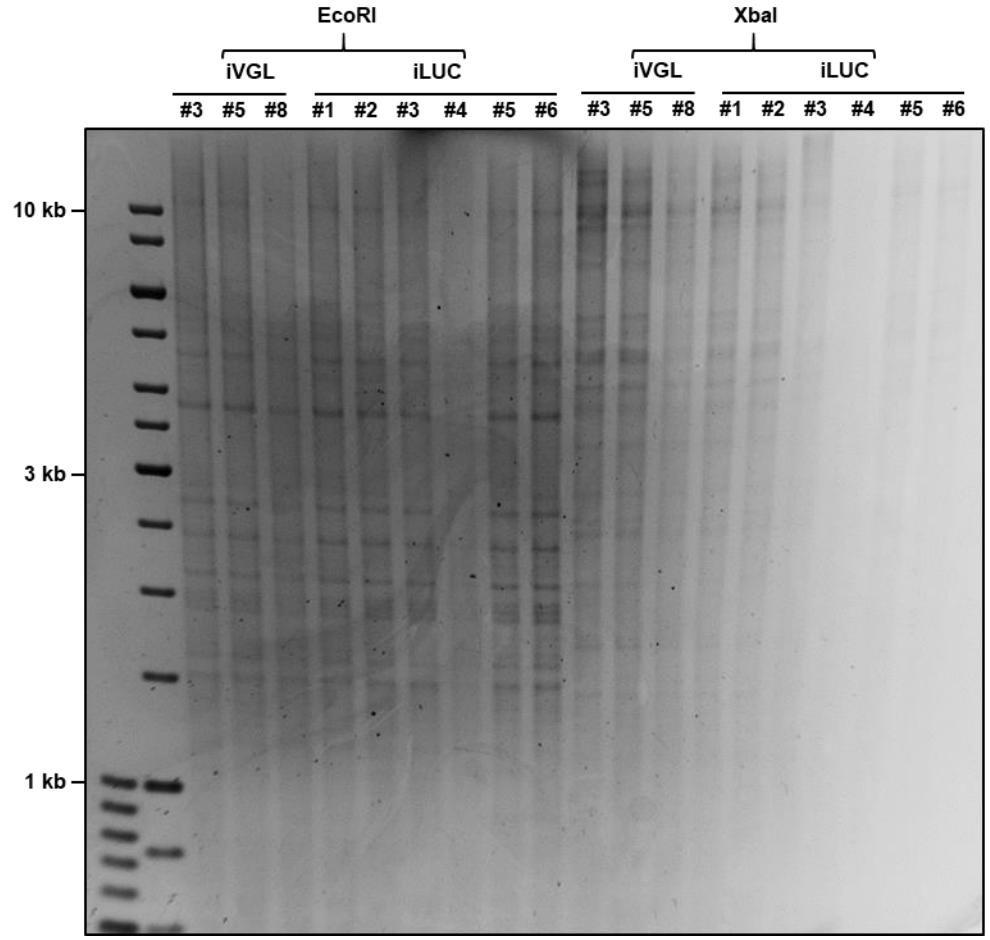


	T-half [sec.]		Mobile fraction[%]	
GFP-TFIIS	0,16 ± 0,05	n=5; p-value: < 0,0001	98,71 ± 1,50	n=7; p-value: 0,0078
GFP-TFIISmut	0,55 ± 0,12		88,71 ± 8,16	

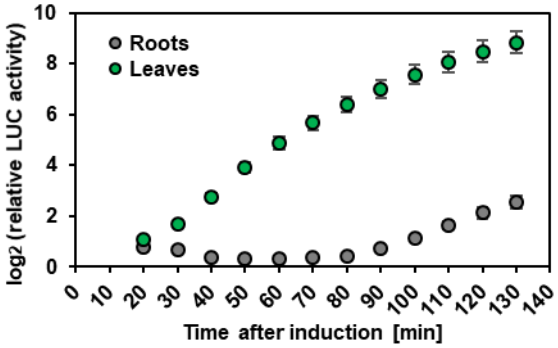
Supplementary Figure S 9. Determination of TFIISmut vs TFIIS mobility in *Nicotiana benthamiana.*
(A) The colocalization of NRPB1-mcherry and GFP-TFIIS or GFP-TFIISmut expression in *Nicotiana benthamiana* leaf epidermal cells. Representative pictures of nuclei from co-transformed cells obtained by confocal microscopy (CLSM) are shown. (B) Representative pictures on nuclei taken during FRAP experiment performed for GFP-TFIIS and GFP-TFIISmut. The region of interest (ROI; doted white line) were photobleached and the recovery of the GFP fluorescence intensity was measured over-time by confocal microscopy (CLSM). Pre-Bleach indicates the first timepoint of the series (t = 0 s), Post-Bleach#1 the first timepoint after bleaching (t = 2,7 s) and Post-Bleach#2 the last time point of the series (t = 10,3 s). Pseudo-coloured images (modified fire LUT) with respective colour calibration bar are shown. (C-D) Mobile fraction and recovery half-time after photobleaching ($t_{½}$) values were calculated for GFP-TFIIS and GFP-TFIISmut using easyFRAP and visualised using a whisker-box plot (C) or given in the table together with the significance level (D). The significance was tested by Student's T-Test: ** p-value < 0.01, *** p-value < 0.001.
TFIIS and GFP-TFIISmut using easyFRAP and visualised using a whisker-box plot (C) or given in the table together with the significance level (D). The significance was tested by Student's T-Test: ** p-value < 0.01, *** p-value < 0.001.



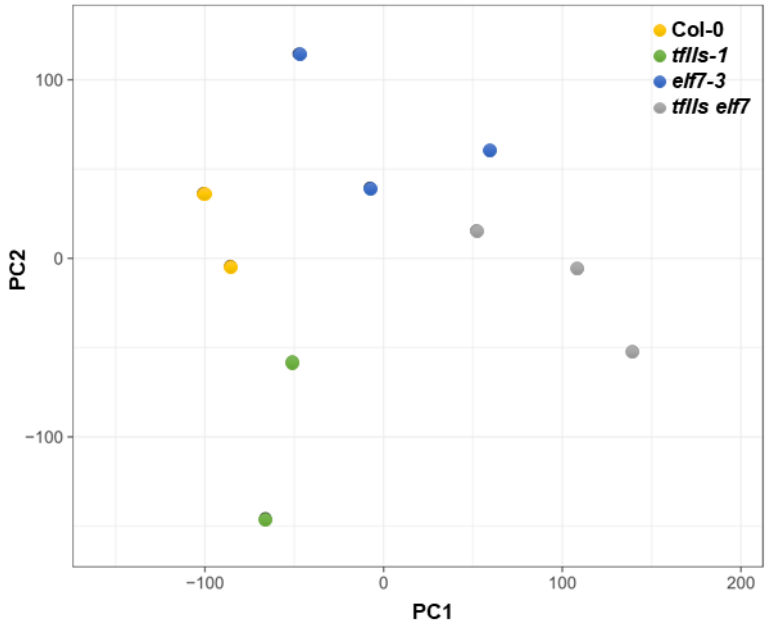
Supplementary Figure S 10. The gating strategy for GFP signal determination during FACS.
Single nuclei plotted in a FITC-A (AF488; GFP fluorescence) versus DAPI-A dot plot with the gates visualisation. The gates for GFP expression (P2) were determined based on the GFP signal intensity detected in FITC-A channel with nuclei extract from mock-induced iGFP-TFIIS (A) and iGFP-TFIISmut (B) lines. Blue dots indicate nuclei with no expression while magenta dots reflect cell considered as GFP positive. FITC: Fluorescein isothiocyanate. Figures were created by Dr. Jörg Fuchs.



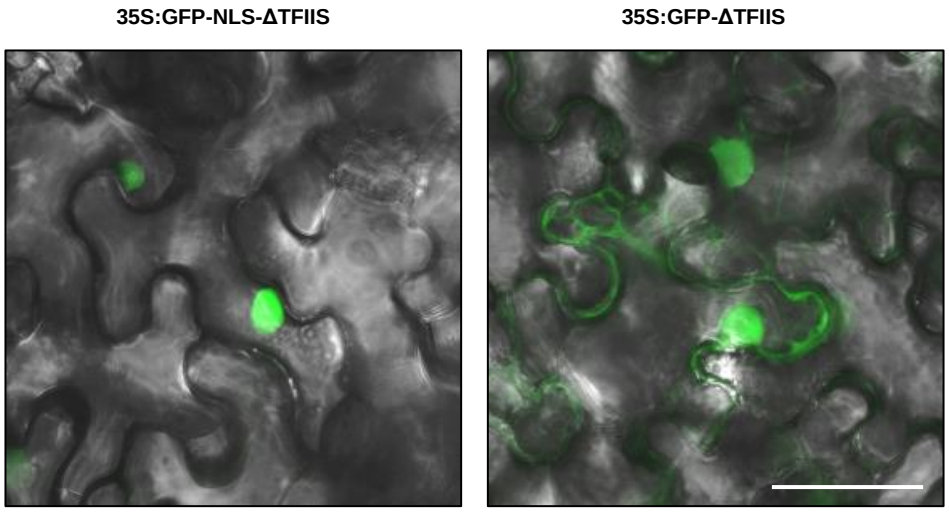
Supplementary Figure S 11. Total DNA digestion with restriction enzymes prior to Southern Blot assay.
Genomic DNA was extracted from 10DAS seedling of several independent iVGL and iLUC lines. DNA digestion was performed using EcoRI and XbaI restriction enzymes. A fraction of digested genomic DNA (1 µL out of 20 µL) of each sample was resolved on the 0,7% agarose gel, stanned with ethidium bromide and documented with BioDocAnalyze System.



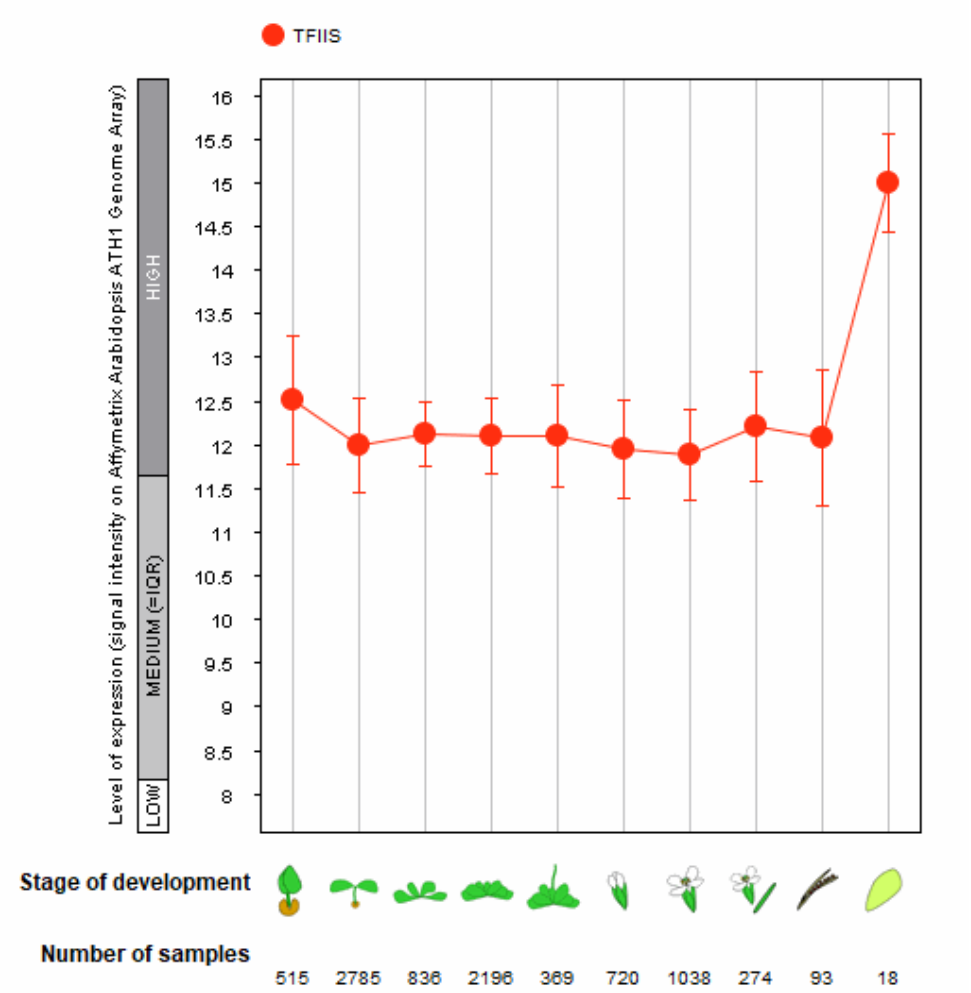
Supplementary Figure S 12. The optimisation of bioluminescence measurements.
Following β-estradiol application, bioluminescence signal from 6DAS seedlings of transgenic iLUC#1 line (n = 6) was measured separately for roots and one first leaf. LUC activity was monitored every 10 minutes over the period of 2 hours using a microplate luminometer LB-960. Dots indicate mean signal ± SD (error bars) normalised to the signal from uninduced plants. Measurement were adjusted to the relative fresh weight of individual roots or leaf area.



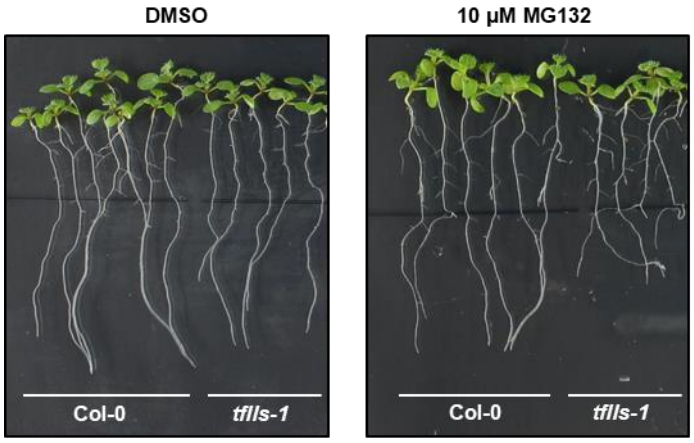
Supplementary Figure S 13. PCA analysis of datasets obtained by RNA-seq (chapter 3.1.8).
The analysis of individual biological replicates was performed using ClustVis.



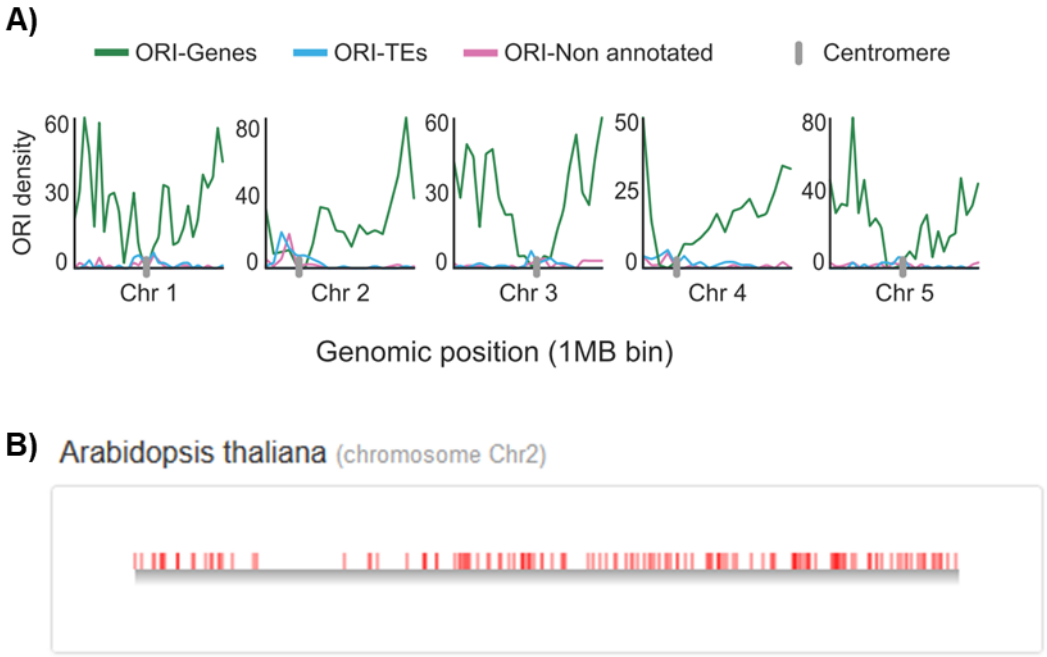
Supplementary Figure S 14. The subcellular localisation of truncated TFIIS.
GFP-ΔTFIIS and GFP-NLS-ΔTFIIS under the control of CaMV 35S promoter were transiently expressed in *Nicotiana benthamiana* following *Agrobacterium*-mediated transformation. GFP signal was monitored using confocal microscopy (CLSM). The representative pictures of leaf epidermis layer are shown. GFP signal is depicted in green. GFP channel shown as merged with the bright field channel. Bars indicate 100 µm.



Supplementary Figure S 15. TFIIS transcript level in *Arabidopsis thaliana*.
TFIIS expression in different *Arabidopsis* tissues throughout development was visualized by analysing public microarray data using Genevestigator.



Supplementary Figure S 16. *tfiis-1* shows hypersensitivity to MG132.
Col-0 and *tfIIs-1* plants were grown vertically for 5 day on MS medium and subsequently transferred on MS medium containing either 10 µM MG132 or mock treatment (DMSO). Representative picture taken at 10DAS are shown.



Supplementary Figure S 17. The genomic distribution of *Arabidopsis* ORI and DEGs in *tfIIs elf7*.
(A) The distribution of *Arabidopsis* origins of replication (ORI) over five chromosomes. Number of ORIs (per 1 Mb bin) in genes (green), TEs (blue) and non-annotated regions (pink) are indicated. Grey bar depicts centromere localization. Picture is from Vergara et al., 2017 (B) The genomic distribution of DEGs in *tfIIs elf7* (relatively to Col-0) visualised by PLAZA software. Grey bar reflects chromosome. Each red line reflects the position of single DEG. The distribution over 2nd chromosome is shown as an example.

Supplementary Table 1. Transcription-related proteins copurified with GS-TFIIS and GS-TFIISmut.
The list of transcription-related proteins copurified with inducible GS-TFIIS and GS-TFIISmut (2.2.4) as well as constitutively expressed GS-TFIIS during AP-MS approach (3.1.1). The numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent experiments.

TFIIS	TFIISmut	TFIIS*	Protein	Complex	Process	AGI
3748 / 3	2882 / 3	2576 / 3	TFIIS	TFIIS	Transcription	AT2G38560
1487 / 3	1763 / 3	2176 / 3	NRPB1	Polymerase II	Transcription	AT4G35800
1378 / 3	1319 / 3	838 / 3	NRPB2	Polymerase II	Transcription	AT4G21710
466 / 3	551 / 3	656 / 3	NRP(B/D/E)3a	Polymerase II	Transcription	AT2G15430
462 / 3	474 / 3	489 / 3	NRP(A/B/C/D)5	Polymerase II	Transcription	AT3G22320
		189 / 3	NRP(B/C/D/E)6a	Polymerase II	Transcription	AT5G51940
216 / 3	149 / 2	471 / 3	NRPB7	Polymerase II	Transcription	AT5G59180
		160 / 2	NRP(A/B/C/D/E)8a	Polymerase II	Transcription	AT1G54250
		305 / 2	NRP(A/B/C/D/E)8b	Polymerase II	Transcription	AT3G59600
	105 / 2	168 / 3	NRP(B/D/E)9a	Polymerase II	Transcription	AT3G16980
		102 / 2	NRP(B/D/E)9b	Polymerase II	Transcription	AT4G16265
		207 / 3	NRP(B/D/E)11	Polymerase II	Transcription	AT3G52090
		1064 / 3	CTR9, ELF8, VIP6	PAF-C	Transcription	AT2G06210
904 / 3	859 / 3	495 / 3	CDC73	PAF-C	Transcription	AT3G22590
504 / 3	416 / 3	1062 / 3	LEO1, VIP4	PAF-C	Transcription	AT5G61150
553 / 3	408 / 3	917 / 3	PAF1, ELF7	PAF-C	Transcription	AT1G79730
438 / 3	408 / 2	422 / 3	SKI8, VIP3	PAF-C	Transcription	AT4G29830
327 / 3	245 / 3	603 / 3	RTF1, VIP5	PAF-C	Transcription	AT1G61040
342 / 2	546 / 3	863 / 3	SPT6-1, SPT6L	SPT6	Transcription	AT1G65440
	256 / 3		IWS1a	IWS1	Transcription	AT1G32130
364 / 3	649 / 3	225 / 3	SPT5-2	SPT4/SPT5	Transcription	AT4G08350
435 / 2	544 / 3	155 / 2	SPT16	FACT	Transcription	AT4G10710
211 / 2	261 / 3		SSRP1	FACT	Transcription	AT3G28730
218 / 2	171 / 2		ELP1, ELO2	Elongator	Transcription	AT5G13680
		222 / 2	ELP3; ELO3	Elongator	Transcription	AT5G50320
	548 / 3		RIN1	SWR1/NuA4, INO80	Transcription	AT5G22330
300 / 2	432 / 3		RVB21	SWR1/NuA4, INO80	Transcription	AT5G67630
151 / 2	198 / 3		AtARP4	SWR1/NuA4, INO80, SWI/SNF-type	Transcription	AT1G18450
	146 / 3		AtSPT7 / HAF1	SAGA_SPT putative	Transcription	AT1G32750
131 / 2	212 / 2		AtNAPL1	NAP1	Transcription	AT4G26110
308 / 3	236 / 3	602 / 2	AtNAPL2	NAP1	Transcription	AT2G19480
		418 / 2	AtNAPL3	NAP1	Transcription	AT5G56950
506 / 2	422 / 3	1119 / 2	HTB2	Histone H2B family	Transcription	AT5G22880
		1399 / 2	HTB6	Histone H2B family	Transcription	AT3G53650
		688 / 2	HTA1	Histone H2A family	Transcription	AT5G54640
		498 / 2	HTA3	Histone H2A family	Transcription	AT1G54690
		342 / 2	H2A.F/Z 3	Histone H2A family	Transcription	AT1G52740
		275 / 2	HTA6	Histone H2A family	Transcription	AT5G59870
		182 / 2	HTA7	Histone H2A family	Transcription	AT5G27670
262 / 3	188 / 2	491 / 3	HDT2	Deacetylase	Transcription	AT5G22650
411 / 3	173 / 3	1030 / 3	HDT3	Deacetylase	Transcription	AT5G03740
		557 / 2	HDT4	Deacetylase	Transcription	AT2G27840

Supplementary Table 2. Overrepresented GO terms among DEGs upregulated in iGFP-TFIISmut relatively to iGFP-TFIIS line upon 24 h induction.

The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with false discovery rate (FDR) < 0.05 are shown in the table.

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0050896	response to stimulus	227 / 813	4057 / 37767	1,10E-40	1,60E-37
GO:0006950	response to stress	150 / 813	2320 / 37767	3,20E-32	2,40E-29
GO:0010200	response to chitin	40 / 813	151 / 37767	4,50E-28	2,30E-25
GO:0006952	defense response	78 / 813	766 / 37767	7,00E-28	2,70E-25
GO:0042221	response to chemical stimulus	130 / 813	2085 / 37767	2,10E-26	6,30E-24
GO:0009607	response to biotic stimulus	68 / 813	638 / 37767	2,00E-25	4,90E-23
GO:0051707	response to other organism	64 / 813	599 / 37767	4,80E-24	1,00E-21
GO:0009743	response to carbohydrate stimulus	42 / 813	240 / 37767	2,80E-23	5,20E-21
GO:0009617	response to bacterium	39 / 813	247 / 37767	2,60E-20	4,40E-18
GO:0010033	response to organic substance	89 / 813	1342 / 37767	9,80E-20	1,50E-17
GO:0051704	multi-organism process	65 / 813	776 / 37767	2,90E-19	3,90E-17
GO:0006955	immune response	39 / 813	367 / 37767	4,80E-15	6,00E-13
GO:0002376	immune system process	39 / 813	368 / 37767	5,20E-15	6,00E-13
GO:0045087	innate immune response	36 / 813	347 / 37767	1,00E-13	1,10E-11
GO:0042742	defense response to bacterium	26 / 813	177 / 37767	2,40E-13	2,40E-11
GO:0009620	response to fungus	21 / 813	158 / 37767	2,50E-10	2,40E-08
GO:0042435	indole derivative biosynthetic process	13 / 813	47 / 37767	3,10E-10	2,70E-08
GO:0019438	aromatic compound biosynthetic process	25 / 813	237 / 37767	4,40E-10	3,70E-08
GO:0009751	response to salicylic acid stimulus	23 / 813	200 / 37767	4,80E-10	3,80E-08
GO:0042434	indole derivative metabolic process	13 / 813	53 / 37767	1,10E-09	7,70E-08
GO:0042430	indole and derivative metabolic process	13 / 813	53 / 37767	1,10E-09	7,70E-08
GO:0031347	regulation of defense response	14 / 813	66 / 37767	1,30E-09	8,70E-08
GO:0009266	response to temperature stimulus	35 / 813	485 / 37767	2,40E-09	1,50E-07
GO:0080134	regulation of response to stress	15 / 813	83 / 37767	2,30E-09	1,50E-07
GO:0019748	secondary metabolic process	34 / 813	489 / 37767	9,50E-09	5,60E-07
GO:0009404	toxin metabolic process	12 / 813	53 / 37767	1,00E-08	5,60E-07
GO:0009407	toxin catabolic process	12 / 813	53 / 37767	1,00E-08	5,60E-07
GO:0009408	response to heat	19 / 813	161 / 37767	1,00E-08	5,60E-07
GO:0006979	response to oxidative stress	27 / 813	332 / 37767	1,60E-08	8,40E-07
GO:0006725	cellular aromatic compound metabolic process	29 / 813	399 / 37767	4,80E-08	2,40E-06
GO:0050832	defense response to fungus	15 / 813	108 / 37767	5,40E-08	2,60E-06
GO:0009814	defense response, incompatible interaction	17 / 813	143 / 37767	5,60E-08	2,60E-06
GO:0010876	lipid localization	8 / 813	24 / 37767	2,50E-07	1,20E-05
GO:0042398	cellular amino acid derivative biosynthetic process	20 / 813	233 / 37767	5,30E-07	2,30E-05
GO:0006568	tryptophan metabolic process	8 / 813	29 / 37767	8,40E-07	3,50E-05
GO:0006586	indolalkylamine metabolic process	8 / 813	29 / 37767	8,40E-07	3,50E-05
GO:0050776	regulation of immune response	9 / 813	41 / 37767	9,10E-07	3,60E-05
GO:0002682	regulation of immune system process	9 / 813	41 / 37767	9,10E-07	3,60E-05
GO:0006575	cellular amino acid derivative metabolic process	23 / 813	315 / 37767	1,10E-06	4,10E-05
GO:0048583	regulation of response to stimulus	17 / 813	188 / 37767	1,90E-06	7,20E-05
GO:0009737	response to abscisic acid stimulus	25 / 813	378 / 37767	2,00E-06	7,20E-05
GO:0000162	tryptophan biosynthetic process	7 / 813	25 / 37767	3,80E-06	1,30E-04
GO:0012501	programmed cell death	19 / 813	244 / 37767	3,80E-06	1,30E-04
GO:0009628	response to abiotic stimulus	60 / 813	1471 / 37767	3,80E-06	1,30E-04
GO:0046219	indolalkylamine biosynthetic process	7 / 813	25 / 37767	3,80E-06	1,30E-04
GO:0006790	sulfur metabolic process	18 / 813	220 / 37767	3,60E-06	1,30E-04
GO:0045088	regulation of innate immune response	8 / 813	38 / 37767	4,80E-06	1,50E-04
GO:0010035	response to inorganic substance	20 / 813	279 / 37767	6,80E-06	2,10E-04
GO:0009627	systemic acquired resistance	9 / 813	54 / 37767	6,70E-06	2,10E-04
GO:0034050	host programmed cell death induced by symbiont	9 / 813	55 / 37767	7,70E-06	2,30E-04
GO:0009626	plant-type hypersensitive response	9 / 813	55 / 37767	7,70E-06	2,30E-04
GO:0016265	death	20 / 813	286 / 37767	9,60E-06	2,70E-04
GO:0008219	cell death	20 / 813	286 / 37767	9,60E-06	2,70E-04
GO:0006576	cellular biogenic amine metabolic process	9 / 813	62 / 37767	1,80E-05	5,10E-04
GO:0009816	defense response to bacterium, incompatible interaction	7 / 813	35 / 37767	2,60E-05	7,00E-04
GO:0009719	response to endogenous stimulus	45 / 813	1068 / 37767	3,00E-05	8,00E-04
GO:0009072	aromatic amino acid family metabolic process	9 / 813	68 / 37767	3,50E-05	9,30E-04
GO:0009753	response to jasmonic acid stimulus	16 / 813	215 / 37767	3,60E-05	9,30E-04
GO:0042401	cellular biogenic amine biosynthetic process	8 / 813	52 / 37767	3,70E-05	9,30E-04

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0006970	response to osmotic stress	23 / 813	408 / 37767	5,40E-05	1,40E-03
GO:0046417	chorismate metabolic process	8 / 813	57 / 37767	6,60E-05	1,60E-03
GO:0009073	aromatic amino acid family biosynthetic process	8 / 813	57 / 37767	6,60E-05	1,60E-03
GO:0052482	cell wall thickening during defense response	5 / 813	17 / 37767	8,00E-05	1,90E-03
GO:0034641	cellular nitrogen compound metabolic process	26 / 813	506 / 37767	7,80E-05	1,90E-03
GO:0052544	callose deposition in cell wall during defense response	5 / 813	17 / 37767	8,00E-05	1,90E-03
GO:0009651	response to salt stress	21 / 813	366 / 37767	8,90E-05	2,00E-03
GO:0009863	salicylic acid mediated signaling pathway	7 / 813	45 / 37767	1,10E-04	2,40E-03
GO:0052542	callose deposition during defense response	5 / 813	19 / 37767	1,30E-04	2,70E-03
GO:0052543	callose deposition in cell wall	5 / 813	19 / 37767	1,30E-04	2,70E-03
GO:0052386	cell wall thickening	5 / 813	20 / 37767	1,50E-04	3,30E-03
GO:0006468	protein amino acid phosphorylation	39 / 813	946 / 37767	1,50E-04	3,30E-03
GO:0009725	response to hormone stimulus	40 / 813	982 / 37767	1,60E-04	3,30E-03
GO:0009409	response to cold	19 / 813	328 / 37767	1,70E-04	3,50E-03
GO:0009611	response to wounding	14 / 813	197 / 37767	1,70E-04	3,50E-03
GO:0033037	polysaccharide localization	5 / 813	21 / 37767	1,90E-04	3,70E-03
GO:0052545	callose localization	5 / 813	21 / 37767	1,90E-04	3,70E-03
GO:0031348	negative regulation of defense response	5 / 813	22 / 37767	2,30E-04	4,40E-03
GO:0006519	cellular amino acid and derivative metabolic process	30 / 813	682 / 37767	3,10E-04	6,00E-03
GO:0001666	response to hypoxia	5 / 813	24 / 37767	3,20E-04	6,00E-03
GO:0070482	response to oxygen levels	5 / 813	24 / 37767	3,20E-04	6,00E-03
GO:0044272	sulfur compound biosynthetic process	10 / 813	115 / 37767	3,30E-04	6,10E-03
GO:0048584	positive regulation of response to stimulus	7 / 813	62 / 37767	6,20E-04	1,10E-02
GO:0043648	dicarboxylic acid metabolic process	8 / 813	82 / 37767	6,40E-04	1,20E-02
GO:0009605	response to external stimulus	21 / 813	429 / 37767	6,70E-04	1,20E-02
GO:0033036	macromolecule localization	22 / 813	462 / 37767	7,10E-04	1,30E-02
GO:0046686	response to cadmium ion	12 / 813	178 / 37767	7,60E-04	1,30E-02
GO:0006796	phosphate metabolic process	43 / 813	1178 / 37767	8,30E-04	1,40E-02
GO:0045089	positive regulation of innate immune response	5 / 813	31 / 37767	9,00E-04	1,40E-02
GO:0002684	positive regulation of immune system process	5 / 813	31 / 37767	9,00E-04	1,40E-02
GO:0002252	immune effector process	5 / 813	31 / 37767	9,00E-04	1,40E-02
GO:0050778	positive regulation of immune response	5 / 813	31 / 37767	9,00E-04	1,40E-02
GO:0002218	activation of innate immune response	5 / 813	31 / 37767	9,00E-04	1,40E-02
GO:0002253	activation of immune response	5 / 813	31 / 37767	9,00E-04	1,40E-02
GO:0006793	phosphorus metabolic process	43 / 813	1179 / 37767	8,40E-04	1,40E-02
GO:0016310	phosphorylation	40 / 813	1079 / 37767	9,40E-04	1,50E-02
GO:0031349	positive regulation of defense response	5 / 813	34 / 37767	1,30E-03	2,00E-02
GO:0043687	post-translational protein modification	44 / 813	1248 / 37767	1,40E-03	2,20E-02
GO:0009699	phenylpropanoid biosynthetic process	10 / 813	141 / 37767	1,40E-03	2,20E-02
GO:0046483	heterocycle metabolic process	21 / 813	460 / 37767	1,50E-03	2,30E-02
GO:0009615	response to virus	6 / 813	55 / 37767	1,80E-03	2,70E-02
GO:0009723	response to ethylene stimulus	12 / 813	199 / 37767	1,90E-03	2,80E-02
GO:0050794	regulation of cellular process	98 / 813	3375 / 37767	1,90E-03	2,80E-02
GO:0009309	amine biosynthetic process	13 / 813	229 / 37767	2,00E-03	3,00E-02
GO:0008652	cellular amino acid biosynthetic process	12 / 813	202 / 37767	2,10E-03	3,00E-02
GO:0010038	response to metal ion	13 / 813	238 / 37767	2,80E-03	4,00E-02
GO:0044271	cellular nitrogen compound biosynthetic process	18 / 813	394 / 37767	3,20E-03	4,50E-02

Supplementary Table 3. Overrepresented GO terms among DEGs downregulated in iGFP-TFIISmut relatively to iGFP-TFIIS line upon 24 h induction.

The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with false discovery rate (FDR) < 0.05 are shown in the table.

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0010876	lipid localization	10 / 406	24 / 37767	1,70E-12	1,40E-09
GO:0015979	photosynthesis	12 / 406	162 / 37767	4,10E-07	1,70E-04
GO:0050896	response to stimulus	77 / 406	4057 / 37767	7,50E-07	2,10E-04
GO:0045454	cell redox homeostasis	8 / 406	72 / 37767	2,30E-06	4,70E-04
GO:0006950	response to stress	50 / 406	2320 / 37767	3,30E-06	5,50E-04
GO:0019684	photosynthesis, light reaction	8 / 406	103 / 37767	2,60E-05	3,60E-03
GO:0019725	cellular homeostasis	10 / 406	174 / 37767	3,10E-05	3,70E-03
GO:0042221	response to chemical stimulus	43 / 406	2085 / 37767	4,60E-05	4,80E-03

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0006869	lipid transport	9 / 406	163 / 37767	1,00E-04	8,50E-03
GO:0006091	generation of precursor metabolites and energy	12 / 406	285 / 37767	9,30E-05	8,50E-03
GO:0005982	starch metabolic process	5 / 406	41 / 37767	1,30E-04	9,60E-03
GO:0042592	homeostatic process	10 / 406	216 / 37767	1,70E-04	1,20E-02
GO:0009719	response to endogenous stimulus	25 / 406	1068 / 37767	3,20E-04	2,10E-02
GO:0006979	response to oxidative stress	12 / 406	332 / 37767	3,60E-04	2,10E-02
GO:0010033	response to organic substance	28 / 406	1342 / 37767	8,40E-04	4,70E-02

Supplementary Table 4. Overrepresented GO terms among DEGs upregulated in iGFP-TFIISmut upon 24 h induction relatively mock treatment.

The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with false discovery rate (FDR) < 0.05 are shown in the table.

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0050896	response to stimulus	234 / 817	4057 / 37767	8,00E-44	1,30E-40
GO:0006950	response to stress	151 / 817	2320 / 37767	1,70E-32	1,30E-29
GO:0010200	response to chitin	42 / 817	151 / 37767	4,80E-30	2,50E-27
GO:0006952	defense response	80 / 817	766 / 37767	4,20E-29	1,70E-26
GO:0042221	response to chemical stimulus	134 / 817	2085 / 37767	3,20E-28	1,00E-25
GO:0009743	response to carbohydrate stimulus	44 / 817	240 / 37767	5,80E-25	1,50E-22
GO:0009607	response to biotic stimulus	67 / 817	638 / 37767	1,30E-24	2,50E-22
GO:0051707	response to other organism	65 / 817	599 / 37767	1,30E-24	2,50E-22
GO:0010033	response to organic substance	95 / 817	1342 / 37767	1,00E-22	1,80E-20
GO:0009617	response to bacterium	41 / 817	247 / 37767	6,40E-22	1,00E-19
GO:0051704	multi-organism process	66 / 817	776 / 37767	9,20E-20	1,30E-17
GO:0006955	immune response	44 / 817	367 / 37767	1,60E-18	2,20E-16
GO:0002376	immune system process	44 / 817	368 / 37767	1,80E-18	2,20E-16
GO:0045087	innate immune response	41 / 817	347 / 37767	3,90E-17	4,50E-15
GO:0042742	defense response to bacterium	28 / 817	177 / 37767	6,20E-15	6,60E-13
GO:0031347	regulation of defense response	15 / 817	66 / 37767	1,50E-10	1,50E-08
GO:0042435	indole derivative biosynthetic process	13 / 817	47 / 37767	3,20E-10	3,00E-08
GO:0019438	aromatic compound biosynthetic process	25 / 817	237 / 37767	4,90E-10	4,30E-08
GO:0009751	response to salicylic acid stimulus	23 / 817	200 / 37767	5,30E-10	4,40E-08
GO:0042434	indole derivative metabolic process	13 / 817	53 / 37767	1,10E-09	7,90E-08
GO:0042430	indole and derivative metabolic process	13 / 817	53 / 37767	1,10E-09	7,90E-08
GO:0009404	toxin metabolic process	13 / 817	53 / 37767	1,10E-09	7,90E-08
GO:0009407	toxin catabolic process	13 / 817	53 / 37767	1,10E-09	7,90E-08
GO:0009814	defense response, incompatible interaction	19 / 817	143 / 37767	1,90E-09	1,30E-07
GO:0080134	regulation of response to stress	15 / 817	83 / 37767	2,40E-09	1,60E-07
GO:0019748	secondary metabolic process	35 / 817	489 / 37767	3,20E-09	2,00E-07
GO:0009620	response to fungus	19 / 817	158 / 37767	8,60E-09	5,10E-07
GO:0006725	cellular aromatic compound metabolic process	30 / 817	399 / 37767	1,50E-08	8,80E-07
GO:0010876	lipid localization	9 / 817	24 / 37767	2,00E-08	1,10E-06
GO:0050776	regulation of immune response	10 / 817	41 / 37767	1,00E-07	5,20E-06
GO:0002682	regulation of immune system process	10 / 817	41 / 37767	1,00E-07	5,20E-06
GO:0034050	host programmed cell death induced by symbiont	11 / 817	55 / 37767	1,30E-07	6,50E-06
GO:0009626	plant-type hypersensitive response	11 / 817	55 / 37767	1,30E-07	6,50E-06
GO:0042398	cellular amino acid derivative biosynthetic process	21 / 817	233 / 37767	1,40E-07	6,50E-06
GO:0006468	protein amino acid phosphorylation	48 / 817	946 / 37767	1,40E-07	6,50E-06
GO:0009737	response to abscisic acid stimulus	27 / 817	378 / 37767	2,10E-07	9,30E-06
GO:0016265	death	23 / 817	286 / 37767	2,40E-07	1,00E-05
GO:0008219	cell death	23 / 817	286 / 37767	2,40E-07	1,00E-05
GO:0009266	response to temperature stimulus	31 / 817	485 / 37767	2,70E-07	1,10E-05
GO:0043687	post-translational protein modification	57 / 817	1248 / 37767	2,70E-07	1,10E-05
GO:0012501	programmed cell death	21 / 817	244 / 37767	2,90E-07	1,10E-05
GO:0050832	defense response to fungus	14 / 817	108 / 37767	3,40E-07	1,30E-05
GO:0006575	cellular amino acid derivative metabolic process	24 / 817	315 / 37767	3,30E-07	1,30E-05
GO:0045088	regulation of innate immune response	9 / 817	38 / 37767	5,40E-07	2,00E-05
GO:0006979	response to oxidative stress	24 / 817	332 / 37767	8,00E-07	2,80E-05
GO:0006568	tryptophan metabolic process	8 / 817	29 / 37767	8,80E-07	3,00E-05
GO:0006586	indolalkylamine metabolic process	8 / 817	29 / 37767	8,80E-07	3,00E-05
GO:0009611	response to wounding	18 / 817	197 / 37767	9,00E-07	3,00E-05
GO:0016310	phosphorylation	50 / 817	1079 / 37767	1,00E-06	3,30E-05
GO:0006796	phosphate metabolic process	53 / 817	1178 / 37767	1,10E-06	3,50E-05
GO:0006793	phosphorus metabolic process	53 / 817	1179 / 37767	1,10E-06	3,60E-05

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0048583	regulation of response to stimulus	17 / 817	188 / 37767	2,10E-06	6,30E-05
GO:0009628	response to abiotic stimulus	61 / 817	1471 / 37767	2,20E-06	6,70E-05
GO:0006464	protein modification process	61 / 817	1474 / 37767	2,30E-06	6,90E-05
GO:0006970	response to osmotic stress	26 / 817	408 / 37767	2,60E-06	7,40E-05
GO:0009816	defense response to bacterium, incompatible interaction	8 / 817	35 / 37767	2,90E-06	8,40E-05
GO:0009719	response to endogenous stimulus	48 / 817	1068 / 37767	3,70E-06	1,00E-04
GO:0000162	tryptophan biosynthetic process	7 / 817	25 / 37767	3,90E-06	1,10E-04
GO:0046219	indolalkylamine biosynthetic process	7 / 817	25 / 37767	3,90E-06	1,10E-04
GO:0009408	response to heat	15 / 817	161 / 37767	5,80E-06	1,50E-04
GO:0009627	systemic acquired resistance	9 / 817	54 / 37767	7,00E-06	1,80E-04
GO:0052542	callose deposition during defense response	6 / 817	19 / 37767	1,10E-05	2,90E-04
GO:0009651	response to salt stress	23 / 817	366 / 37767	1,20E-05	3,00E-04
GO:0043412	macromolecule modification	63 / 817	1636 / 37767	1,40E-05	3,50E-04
GO:0009605	response to external stimulus	25 / 817	429 / 37767	1,70E-05	4,10E-04
GO:0033037	polysaccharide localization	6 / 817	21 / 37767	1,80E-05	4,30E-04
GO:0052545	callose localization	6 / 817	21 / 37767	1,80E-05	4,30E-04
GO:0006576	cellular biogenic amine metabolic process	9 / 817	62 / 37767	1,90E-05	4,40E-04
GO:0009725	response to hormone stimulus	43 / 817	982 / 37767	2,10E-05	4,90E-04
GO:0009072	aromatic amino acid family metabolic process	9 / 817	68 / 37767	3,70E-05	8,30E-04
GO:0042401	cellular biogenic amine biosynthetic process	8 / 817	52 / 37767	3,80E-05	8,50E-04
GO:0009753	response to jasmonic acid stimulus	16 / 817	215 / 37767	3,80E-05	8,50E-04
GO:0006790	sulfur metabolic process	16 / 817	220 / 37767	5,00E-05	1,10E-03
GO:0046417	chorismate metabolic process	8 / 817	57 / 37767	6,80E-05	1,40E-03
GO:0009073	aromatic amino acid family biosynthetic process	8 / 817	57 / 37767	6,80E-05	1,40E-03
GO:0052482	cell wall thickening during defense response	5 / 817	17 / 37767	8,20E-05	1,70E-03
GO:0052544	callose deposition in cell wall during defense response	5 / 817	17 / 37767	8,20E-05	1,70E-03
GO:0007165	signal transduction	48 / 817	1228 / 37767	1,10E-04	2,20E-03
GO:0009863	salicylic acid mediated signaling pathway	7 / 817	45 / 37767	1,10E-04	2,20E-03
GO:0045089	positive regulation of innate immune response	6 / 817	31 / 37767	1,20E-04	2,20E-03
GO:0002684	positive regulation of immune system process	6 / 817	31 / 37767	1,20E-04	2,20E-03
GO:0050778	positive regulation of immune response	6 / 817	31 / 37767	1,20E-04	2,20E-03
GO:0002218	activation of innate immune response	6 / 817	31 / 37767	1,20E-04	2,20E-03
GO:0002253	activation of immune response	6 / 817	31 / 37767	1,20E-04	2,20E-03
GO:0052543	callose deposition in cell wall	5 / 817	19 / 37767	1,30E-04	2,40E-03
GO:0033036	macromolecule localization	24 / 817	462 / 37767	1,30E-04	2,50E-03
GO:0052386	cell wall thickening	5 / 817	20 / 37767	1,60E-04	2,90E-03
GO:0031349	positive regulation of defense response	6 / 817	34 / 37767	1,80E-04	3,30E-03
GO:0034641	cellular nitrogen compound metabolic process	25 / 817	506 / 37767	2,00E-04	3,60E-03
GO:0031348	negative regulation of defense response	5 / 817	22 / 37767	2,30E-04	4,10E-03
GO:0051716	cellular response to stimulus	35 / 817	840 / 37767	3,00E-04	5,20E-03
GO:0001666	response to hypoxia	5 / 817	24 / 37767	3,30E-04	5,60E-03
GO:0070482	response to oxygen levels	5 / 817	24 / 37767	3,30E-04	5,60E-03
GO:0006519	cellular amino acid and derivative metabolic process	30 / 817	682 / 37767	3,40E-04	5,70E-03
GO:0009851	auxin biosynthetic process	5 / 817	28 / 37767	6,10E-04	1,00E-02
GO:0048584	positive regulation of response to stimulus	7 / 817	62 / 37767	6,40E-04	1,10E-02
GO:0043648	dicarboxylic acid metabolic process	8 / 817	82 / 37767	6,60E-04	1,10E-02
GO:0002252	immune effector process	5 / 817	31 / 37767	9,20E-04	1,50E-02
GO:0009409	response to cold	17 / 817	328 / 37767	1,20E-03	2,00E-02
GO:0009699	phenylpropanoid biosynthetic process	10 / 817	141 / 37767	1,50E-03	2,40E-02
GO:0050794	regulation of cellular process	99 / 817	3375 / 37767	1,50E-03	2,40E-02
GO:0010035	response to inorganic substance	15 / 817	279 / 37767	1,70E-03	2,60E-02
GO:0033554	cellular response to stress	19 / 817	399 / 37767	1,70E-03	2,60E-02
GO:0009723	response to ethylene stimulus	12 / 817	199 / 37767	1,90E-03	3,00E-02
GO:0009698	phenylpropanoid metabolic process	11 / 817	175 / 37767	2,20E-03	3,30E-02

Supplementary Table 5. Overrepresented GO terms among DEGs downregulated in iGFP-TFIISmut upon 24 h induction relatively mock treatment.

The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with false discovery rate (FDR) < 0.05 are shown in the table.

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0010876	lipid localization	9 / 268	24 / 37767	1,20E-12	7,40E-10
GO:0009733	response to auxin stimulus	17 / 268	360 / 37767	2,00E-09	5,90E-07
GO:0015979	photosynthesis	12 / 268	162 / 37767	4,60E-09	9,20E-07
GO:0045454	cell redox homeostasis	9 / 268	72 / 37767	6,30E-09	9,30E-07

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0019725	cellular homeostasis	11 / 268	174 / 37767	9,40E-08	1,10E-05
GO:0050896	response to stimulus	58 / 268	4057 / 37767	2,00E-07	2,00E-05
GO:0042592	homeostatic process	11 / 268	216 / 37767	7,30E-07	6,20E-05
GO:0009725	response to hormone stimulus	23 / 268	982 / 37767	9,00E-07	6,70E-05
GO:0009719	response to endogenous stimulus	24 / 268	1068 / 37767	1,00E-06	6,90E-05
GO:0019684	photosynthesis, light reaction	8 / 268	103 / 37767	1,30E-06	7,70E-05
GO:0009773	photosynthetic electron transport in photosystem I	5 / 268	23 / 37767	1,40E-06	7,80E-05
GO:0010033	response to organic substance	26 / 268	1342 / 37767	5,20E-06	2,60E-04
GO:0042221	response to chemical stimulus	33 / 268	2085 / 37767	1,80E-05	8,00E-04
GO:0006869	lipid transport	8 / 268	163 / 37767	3,10E-05	1,20E-03
GO:0009767	photosynthetic electron transport chain	5 / 268	46 / 37767	3,00E-05	1,20E-03
GO:0006950	response to stress	34 / 268	2320 / 37767	6,00E-05	2,20E-03
GO:0009628	response to abiotic stimulus	24 / 268	1471 / 37767	1,70E-04	5,90E-03
GO:0006091	generation of precursor metabolites and energy	9 / 268	285 / 37767	2,60E-04	8,60E-03
GO:0009664	plant-type cell wall organization	5 / 268	79 / 37767	3,30E-04	1,00E-02
GO:0022900	electron transport chain	5 / 268	88 / 37767	5,20E-04	1,60E-02
GO:0009739	response to gibberellin stimulus	6 / 268	159 / 37767	1,20E-03	3,30E-02

Supplementary Table 6. Overrepresented GO terms among DEGs upregulated in iGFP-TFIISmut relatively to iGFP-TFIIS line upon 6 h β-estradiol induction.

The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with false discovery rate (FDR) < 0.05 are shown in the table.

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0010200	response to chitin	11 / 78	151 / 37767	3,80E-14	1,20E-11
GO:0009743	response to carbohydrate stimulus	11 / 78	240 / 37767	4,60E-12	7,30E-10
GO:0050896	response to stimulus	27 / 78	4057 / 37767	1,90E-08	2,00E-06
GO:0010033	response to organic substance	15 / 78	1342 / 37767	1,00E-07	8,00E-06
GO:0006950	response to stress	19 / 78	2320 / 37767	2,00E-07	1,20E-05
GO:0006952	defense response	11 / 78	766 / 37767	5,70E-07	3,00E-05
GO:0050832	defense response to fungus	5 / 78	108 / 37767	3,90E-06	1,70E-04
GO:0042221	response to chemical stimulus	16 / 78	2085 / 37767	5,00E-06	2,00E-04
GO:0009266	response to temperature stimulus	8 / 78	485 / 37767	8,30E-06	2,90E-04
GO:0009620	response to fungus	5 / 78	158 / 37767	2,30E-05	7,20E-04
GO:0042742	defense response to bacterium	5 / 78	177 / 37767	3,90E-05	1,10E-03
GO:0009628	response to abiotic stimulus	12 / 78	1471 / 37767	5,00E-05	1,30E-03
GO:0009611	response to wounding	5 / 78	197 / 37767	6,30E-05	1,50E-03
GO:0009409	response to cold	6 / 78	328 / 37767	6,80E-05	1,50E-03
GO:0009414	response to water deprivation	5 / 78	229 / 37767	1,30E-04	2,70E-03
GO:0009415	response to water	5 / 78	240 / 37767	1,60E-04	3,10E-03
GO:0009617	response to bacterium	5 / 78	247 / 37767	1,80E-04	3,30E-03
GO:0051707	response to other organism	7 / 78	599 / 37767	2,60E-04	4,50E-03
GO:0009605	response to external stimulus	6 / 78	429 / 37767	2,80E-04	4,70E-03
GO:0009607	response to biotic stimulus	7 / 78	638 / 37767	3,80E-04	5,90E-03
GO:0033036	macromolecule localization	6 / 78	462 / 37767	4,20E-04	6,30E-03
GO:0019748	secondary metabolic process	6 / 78	489 / 37767	5,60E-04	8,10E-03
GO:0009737	response to abscisic acid stimulus	5 / 78	378 / 37767	1,20E-03	1,60E-02
GO:0051704	multi-organism process	7 / 78	776 / 37767	1,20E-03	1,60E-02
GO:0006970	response to osmotic stress	5 / 78	408 / 37767	1,70E-03	2,10E-02

Supplementary Table 7. Overrepresented GO terms among DEGs downregulated in iGFP-TFIISmut relatively to iGFP-TFIIS line upon 6 h β-estradiol induction.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with false discovery rate (FDR) < 0.05 are shown in the table.

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0006091	generation of precursor metabolites and energy	13 / 86	285 / 37767	1,80E-13	2,60E-11
GO:0015979	photosynthesis	8 / 86	162 / 37767	5,50E-09	3,00E-07
GO:0010876	lipid localization	5 / 86	24 / 37767	6,10E-09	3,00E-07
GO:0055114	oxidation reduction	7 / 86	203 / 37767	5,40E-07	2,00E-05
GO:0045333	cellular respiration	5 / 86	85 / 37767	2,00E-06	4,90E-05
GO:0015980	energy derivation by oxidation of organic compounds	5 / 86	85 / 37767	2,00E-06	4,90E-05
GO:0019684	photosynthesis, light reaction	5 / 86	103 / 37767	5,00E-06	1,00E-04
GO:0006979	response to oxidative stress	7 / 86	332 / 37767	1,30E-05	2,30E-04
GO:0006869	lipid transport	5 / 86	163 / 37767	4,20E-05	6,80E-04
GO:0044237	cellular metabolic process	33 / 86	8722 / 37767	1,10E-03	1,60E-02
GO:0006810	transport	11 / 86	1846 / 37767	3,20E-03	3,90E-02
GO:0051234	establishment of localization	11 / 86	1851 / 37767	3,30E-03	3,90E-02
GO:0033036	macromolecule localization	5 / 86	462 / 37767	4,30E-03	4,20E-02
GO:0008152	metabolic process	36 / 86	10614 / 37767	4,30E-03	4,20E-02
GO:0051179	localization	11 / 86	1922 / 37767	4,30E-03	4,20E-02

Supplementary Table 8. Overrepresented GO terms among DEGs upregulated in iGFP-TFIISmut and iGFP-TFIIS depending on applied induction conditions.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with false discovery rate (FDR) < 0.05 are shown in the table. Redundant proteins were removed by REViGO. Frequency indicate the percentage of each GO term in the whole UniProt database. 6h β/24h β: 6 or 24 hours β-estradiol induction; 24h: 24 hours of mock induction (EtOH). Numbers reflect $\log_{10}$ FDR.

		iGFP-TFIIS		iGFP-TFIISmut	
GO term ID	description	6h β vs 24h m	24h β vs 6h β	6h β vs 24h m	24h β vs 6h β
GO:0045333	cellular respiration	-4,64			
GO:0006091	generation of precursor metabolites and energy	-4,48			
GO:0006364	rRNA processing	-4,00			
GO:0050896	response to stimulus	-3,80	-7,72	-17,80	-33,89
GO:0042221	response to chemical	-3,59		-17,80	-17,60
GO:0009733	response to auxin	-3,57			
GO:0006811	ion transport	-3,19			
GO:0009719	response to endogenous stimulus	-3,15		-7,85	-3,68
GO:0015979	photosynthesis		-11,82		
GO:0010876	lipid localization		-8,96		-7,96
GO:0006950	response to stress		-8,42	-11,03	-26,64
GO:0009767	photosynthetic electron transport chain		-5,80		
GO:0009743	response to carbohydrate			-18,10	
GO:0010200	response to chitin			-17,80	
GO:0010033	response to organic substance			-17,10	
GO:0006952	defense response			-5,85	-19,72
GO:0002376	immune system process			-5,11	-5,41
GO:0009605	response to external stimulus			-5,06	-5,41
GO:0009611	response to wounding			-4,82	-6,10
GO:0045087	innate immune response			-4,77	-5,41
GO:0009628	response to abiotic stimulus			-3,24	-3,68
GO:0009617	response to bacterium			-3,22	
GO:0009607	response to biotic stimulus				-19,82

		iGFP-TFIIS		iGFP-TFIISmut	
GO term ID	description	6h β vs 24h m	24h β vs 6h β	6h β vs 24h m	24h β vs 6h β
GO:0051707	response to other organism				-17,77
GO:0051704	multi-organism process				-13,60
GO:0019748	secondary metabolic process				-9,52
GO:0009407	toxin catabolic process				-8,34
GO:0019438	aromatic compound biosynthetic process				-8,21
GO:0006979	response to oxidative stress				-6,23
GO:0006725	cellular aromatic compound metabolic process				-5,55
GO:0006575	cellular modified amino acid metabolic process				-5,30
GO:0010035	response to inorganic substance				-5,27
GO:0042398	cellular modified amino acid biosynthetic process				-5,04
GO:0006790	sulfur compound metabolic process				-3,85
GO:0042435	indole-containing compound biosynthetic process				-3,85
GO:0006520	cellular amino acid metabolic process				-3,80
GO:0034641	cellular nitrogen compound metabolic process				-3,64
GO:0042430	indole-containing compound metabolic process				-3,55

Supplementary Table 9. Overrepresented GO terms among DEGs downregulated in iGFP-TFIISmut and iGFP-TFIIS depending on applied induction conditions.

The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with false discovery rate (FDR) < 0.05 are shown in the table. Redundant proteins were removed by REViGO. Frequency indicate the percentage of each GO term in the whole UniProt database. 6h β/24h β: 6 or 24 hours β-estradiol induction; 24h: 24 hours of mock induction (EtOH). Numbers reflect $\log_{10}$ FDR.

		iGFP-TFIIS		iGFP-TFIISmut	
GO term ID	description	6h β vs 24h m	24h β vs 6h β	6h β vs 24h m	24h β vs 6h β
GO:0050896	response to stimulus	-35,47		-13,24	-9,32
GO:0006950	response to stress	-35,47		-13,38	
GO:0015979	photosynthesis	-23,68		-11,32	
GO:0009628	response to abiotic stimulus	-19,00		-5,09	
GO:0042221	response to chemical	-17,82		-6,19	-8,17
GO:0009414	response to water deprivation	-16,38		-3,80	
GO:0019684	photosynthesis, light reaction	-12,09		-5,00	
GO:0006952	defense response	-11,89			
GO:0010876	lipid localization	-10,28		-13,38	-3,68
GO:0019748	secondary metabolic process	-9,00			
GO:0009719	response to endogenous stimulus	-9,00	-3,85		-11,00
GO:0006091	generation of precursor metabolites and energy	-7,41		-3,85	
GO:0009607	response to biotic stimulus	-7,41			
GO:0051707	response to other organism	-6,46			
GO:0009605	response to external stimulus	-6,09		-3,80	
GO:0009611	response to wounding	-5,21			
GO:0051704	multi-organism process	-4,37			
GO:0009407	toxin catabolic process	-4,36			
GO:0006979	response to oxidative stress	-4,18		-7,23	
GO:0044262	cellular carbohydrate metabolic process	-3,74		-3,85	
GO:0044272	sulfur compound biosynthetic process	-3,39			
GO:0042254	ribosome biogenesis		-6,92		
GO:0006396	RNA processing		-5,46		
GO:0034660	ncRNA metabolic process		-4,51		
GO:0009725	response to hormone		-4,35		
GO:0044085	cellular component biogenesis		-4,26		

		iGFP-TFIIS		iGFP-TFIISmut	
GO term ID	description	6h β vs 24h m	24h β vs 6h β	6h β vs 24h m	24h β vs 6h β
GO:0009451	RNA modification		-3,49		
GO:0042592	homeostatic process			-3,46	
GO:0009991	response to extracellular stimulus			-3,11	
GO:0009733	response to auxin				-16,89
GO:0009664	plant-type cell wall organization				-6,09
GO:0006857	oligopeptide transport				-4,18
GO:0009416	response to light stimulus				-4,18
GO:0006351	transcription, DNA-templated				-3,82

Supplementary Table 10. Overrepresented GO terms among PPEP-S2P responsive genes in iGFP-TFIISmut.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with p-value < 0.01 are shown in the table. PPEP-responsive genes were defined as having increased PPEP upon GFP-TFIISmut expression (log_2FC PPEP > 0; β-estradiol vs mock), resulting in clear PPEP establishment upon β-estradiol induction (log_2PPEP > 2).

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	pvalue
GO:0009791	post-embryonic development	188 / 4136	705 / 37767	7,20E-24
GO:0009987	cellular process	1560 / 4136	11684 / 37767	9,10E-19
GO:0043687	post-translational protein modification	253 / 4136	1248 / 37767	1,10E-17
GO:0006464	protein modification process	280 / 4136	1474 / 37767	2,70E-16
GO:0050896	response to stimulus	617 / 4136	4057 / 37767	3,90E-15
GO:0043412	macromolecule modification	294 / 4136	1636 / 37767	2,10E-14
GO:0008152	metabolic process	1397 / 4136	10614 / 37767	2,50E-14
GO:0006468	protein amino acid phosphorylation	193 / 4136	946 / 37767	5,30E-14
GO:0006950	response to stress	382 / 4136	2320 / 37767	1,70E-13
GO:0044238	primary metabolic process	1193 / 4136	8995 / 37767	1,10E-12
GO:0016310	phosphorylation	205 / 4136	1079 / 37767	3,00E-12
GO:0006793	phosphorus metabolic process	219 / 4136	1179 / 37767	3,50E-12
GO:0006796	phosphate metabolic process	218 / 4136	1178 / 37767	5,30E-12
GO:0044237	cellular metabolic process	1147 / 4136	8722 / 37767	3,00E-11
GO:0032501	multicellular organismal process	330 / 4136	2094 / 37767	7,20E-10
GO:0051716	cellular response to stimulus	157 / 4136	840 / 37767	2,80E-09
GO:0007275	multicellular organismal development	309 / 4136	2020 / 37767	3,50E-08
GO:0043170	macromolecule metabolic process	927 / 4136	7127 / 37767	3,80E-08
GO:0044260	cellular macromolecule metabolic process	840 / 4136	6447 / 37767	1,70E-07
GO:0065007	biological regulation	571 / 4136	4188 / 37767	1,90E-07
GO:0003006	reproductive developmental process	165 / 4136	978 / 37767	4,30E-07
GO:0016570	histone modification	26 / 4136	66 / 37767	5,50E-07
GO:0016569	covalent chromatin modification	27 / 4136	71 / 37767	5,90E-07
GO:0006952	defense response	135 / 4136	766 / 37767	6,40E-07
GO:0070887	cellular response to chemical stimulus	89 / 4136	452 / 37767	1,20E-06
GO:0009628	response to abiotic stimulus	227 / 4136	1471 / 37767	1,40E-06
GO:0050789	regulation of biological process	503 / 4136	3697 / 37767	1,40E-06
GO:0022414	reproductive process	186 / 4136	1161 / 37767	1,60E-06
GO:0009886	post-embryonic morphogenesis	18 / 4136	35 / 37767	1,60E-06
GO:0051179	localization	283 / 4136	1922 / 37767	2,20E-06
GO:0051234	establishment of localization	274 / 4136	1851 / 37767	2,20E-06
GO:0006810	transport	272 / 4136	1846 / 37767	3,30E-06
GO:0010035	response to inorganic substance	61 / 4136	279 / 37767	3,40E-06
GO:0016568	chromatin modification	30 / 4136	95 / 37767	3,40E-06
GO:0016571	histone methylation	16 / 4136	30 / 37767	4,30E-06
GO:0000003	reproduction	186 / 4136	1186 / 37767	5,20E-06
GO:0019538	protein metabolic process	534 / 4136	4009 / 37767	5,80E-06
GO:0032502	developmental process	327 / 4136	2304 / 37767	6,10E-06
GO:0070271	protein complex biogenesis	36 / 4136	134 / 37767	8,50E-06
GO:0006461	protein complex assembly	36 / 4136	134 / 37767	8,50E-06
GO:0050794	regulation of cellular process	456 / 4136	3375 / 37767	8,80E-06
GO:0048608	reproductive structure development	156 / 4136	978 / 37767	1,40E-05
GO:0048856	anatomical structure development	252 / 4136	1726 / 37767	1,40E-05
GO:0006479	protein amino acid methylation	16 / 4136	34 / 37767	1,50E-05

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	pvalue
GO:0008213	protein amino acid alkylation	16 / 4136	34 / 37767	1,50E-05
GO:0009743	response to carbohydrate stimulus	52 / 4136	240 / 37767	2,10E-05
GO:0009605	response to external stimulus	80 / 4136	429 / 37767	2,10E-05
GO:0042221	response to chemical stimulus	295 / 4136	2085 / 37767	2,20E-05
GO:0008219	cell death	59 / 4136	286 / 37767	2,20E-05
GO:0016265	death	59 / 4136	286 / 37767	2,20E-05
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	430 / 4136	3198 / 37767	2,40E-05
GO:0012501	programmed cell death	52 / 4136	244 / 37767	3,10E-05
GO:0044262	cellular carbohydrate metabolic process	77 / 4136	417 / 37767	4,00E-05
GO:0033554	cellular response to stress	74 / 4136	399 / 37767	5,00E-05
GO:0007242	intracellular signaling cascade	110 / 4136	659 / 37767	5,00E-05
GO:0007165	signal transduction	184 / 4136	1228 / 37767	5,80E-05
GO:0044267	cellular protein metabolic process	460 / 4136	3487 / 37767	6,20E-05
GO:0045087	innate immune response	66 / 4136	347 / 37767	6,50E-05
GO:0010200	response to chitin	36 / 4136	151 / 37767	7,30E-05
GO:0006970	response to osmotic stress	74 / 4136	408 / 37767	9,20E-05
GO:0009651	response to salt stress	68 / 4136	366 / 37767	9,30E-05
GO:0006955	immune response	68 / 4136	367 / 37767	1,00E-04
GO:0002376	immune system process	68 / 4136	368 / 37767	1,10E-04
GO:0010876	lipid localization	12 / 4136	24 / 37767	1,10E-04
GO:0048518	positive regulation of biological process	52 / 4136	259 / 37767	1,20E-04
GO:0033036	macromolecule localization	81 / 4136	462 / 37767	1,20E-04
GO:0006325	chromatin organization	39 / 4136	175 / 37767	1,30E-04
GO:0048316	seed development	90 / 4136	530 / 37767	1,40E-04
GO:0009793	embryonic development ending in seed dormancy	81 / 4136	465 / 37767	1,40E-04
GO:0051276	chromosome organization	45 / 4136	216 / 37767	1,60E-04
GO:0006915	apoptosis	36 / 4136	159 / 37767	1,80E-04
GO:0009790	embryonic development	90 / 4136	535 / 37767	1,80E-04
GO:0006996	organelle organization	104 / 4136	640 / 37767	1,90E-04
GO:0032870	cellular response to hormone stimulus	60 / 4136	321 / 37767	2,00E-04
GO:0009755	hormone-mediated signaling pathway	60 / 4136	321 / 37767	2,00E-04
GO:0007049	cell cycle	53 / 4136	275 / 37767	2,40E-04
GO:0051704	multi-organism process	121 / 4136	776 / 37767	2,60E-04
GO:0010154	fruit development	92 / 4136	557 / 37767	2,70E-04
GO:0006807	nitrogen compound metabolic process	492 / 4136	3826 / 37767	2,70E-04
GO:0048584	positive regulation of response to stimulus	19 / 4136	62 / 37767	2,80E-04
GO:0048580	regulation of post-embryonic development	41 / 4136	200 / 37767	4,00E-04
GO:0006629	lipid metabolic process	128 / 4136	841 / 37767	4,20E-04
GO:0010033	response to organic substance	191 / 4136	1342 / 37767	4,60E-04
GO:0016043	cellular component organization	170 / 4136	1179 / 37767	5,50E-04
GO:0009620	response to fungus	34 / 4136	158 / 37767	5,70E-04
GO:0032259	methylation	23 / 4136	90 / 37767	6,10E-04
GO:0005975	carbohydrate metabolic process	130 / 4136	866 / 37767	6,20E-04
GO:0006730	one-carbon metabolic process	26 / 4136	111 / 37767	8,50E-04
GO:0022607	cellular component assembly	49 / 4136	265 / 37767	9,00E-04
GO:0034641	cellular nitrogen compound metabolic process	82 / 4136	506 / 37767	9,10E-04
GO:0008610	lipid biosynthetic process	73 / 4136	439 / 37767	9,20E-04
GO:0009611	response to wounding	39 / 4136	197 / 37767	9,70E-04
GO:0009736	cytokinin mediated signaling pathway	15 / 4136	48 / 37767	9,90E-04
GO:0051641	cellular localization	90 / 4136	569 / 37767	1,00E-03
GO:0009607	response to biotic stimulus	99 / 4136	638 / 37767	1,00E-03
GO:0006508	proteolysis	121 / 4136	810 / 37767	1,10E-03
GO:0006457	protein folding	50 / 4136	275 / 37767	1,10E-03
GO:0007166	cell surface receptor linked signaling pathway	35 / 4136	172 / 37767	1,10E-03
GO:0043414	macromolecule methylation	21 / 4136	83 / 37767	1,10E-03
GO:0050832	defense response to fungus	25 / 4136	108 / 37767	1,20E-03
GO:0007167	enzyme linked receptor protein signaling pathway	30 / 4136	140 / 37767	1,20E-03
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	30 / 4136	140 / 37767	1,20E-03
GO:0043933	macromolecular complex subunit organization	44 / 4136	235 / 37767	1,30E-03
GO:0048583	regulation of response to stimulus	37 / 4136	188 / 37767	1,40E-03
GO:0051707	response to other organism	93 / 4136	599 / 37767	1,40E-03
GO:0022402	cell cycle process	31 / 4136	149 / 37767	1,60E-03
GO:0065003	macromolecular complex assembly	40 / 4136	210 / 37767	1,60E-03
GO:0051649	establishment of localization in cell	83 / 4136	525 / 37767	1,60E-03
GO:0009311	oligosaccharide metabolic process	15 / 4136	52 / 37767	1,90E-03
GO:0010029	regulation of seed germination	11 / 4136	31 / 37767	2,00E-03
GO:0006350	transcription	255 / 4136	1923 / 37767	2,20E-03

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	pvalue
GO:0051169	nuclear transport	18 / 4136	71 / 37767	2,40E-03
GO:0006913	nucleocytoplasmic transport	18 / 4136	71 / 37767	2,40E-03
GO:0048522	positive regulation of cellular process	32 / 4136	163 / 37767	2,90E-03
GO:0042742	defense response to bacterium	34 / 4136	177 / 37767	3,00E-03
GO:0051239	regulation of multicellular organismal process	48 / 4136	277 / 37767	3,20E-03
GO:0019222	regulation of metabolic process	287 / 4136	2210 / 37767	3,30E-03
GO:0051301	cell division	22 / 4136	99 / 37767	3,60E-03
GO:0009889	regulation of biosynthetic process	247 / 4136	1881 / 37767	3,80E-03
GO:0031326	regulation of cellular biosynthetic process	247 / 4136	1881 / 37767	3,80E-03
GO:0048513	organ development	120 / 4136	838 / 37767	4,00E-03
GO:0000904	cell morphogenesis involved in differentiation	16 / 4136	63 / 37767	4,10E-03
GO:0048731	system development	120 / 4136	839 / 37767	4,10E-03
GO:0010038	response to metal ion	42 / 4136	238 / 37767	4,10E-03
GO:0031323	regulation of cellular metabolic process	265 / 4136	2036 / 37767	4,20E-03
GO:0015893	drug transport	18 / 4136	76 / 37767	4,50E-03
GO:0009555	pollen development	29 / 4136	148 / 37767	4,50E-03
GO:0080134	regulation of response to stress	19 / 4136	83 / 37767	4,90E-03
GO:0006259	DNA metabolic process	64 / 4136	405 / 37767	5,10E-03
GO:0042493	response to drug	18 / 4136	77 / 37767	5,10E-03
GO:0009719	response to endogenous stimulus	147 / 4136	1068 / 37767	5,70E-03
GO:0090056	regulation of chlorophyll metabolic process	6 / 4136	12 / 37767	6,00E-03
GO:0046907	intracellular transport	71 / 4136	463 / 37767	6,20E-03
GO:0009314	response to radiation	90 / 4136	613 / 37767	6,50E-03
GO:0016036	cellular response to phosphate starvation	12 / 4136	43 / 37767	6,50E-03
GO:0006855	multidrug transport	17 / 4136	73 / 37767	6,60E-03
GO:0048519	negative regulation of biological process	68 / 4136	442 / 37767	6,80E-03
GO:0010468	regulation of gene expression	258 / 4136	2001 / 37767	6,90E-03
GO:0000059	protein import into nucleus, docking	7 / 4136	17 / 37767	6,90E-03
GO:0048609	reproductive process in a multicellular organism	11 / 4136	38 / 37767	7,20E-03
GO:0010556	regulation of macromolecule biosynthetic process	239 / 4136	1843 / 37767	7,30E-03
GO:0016051	carbohydrate biosynthetic process	46 / 4136	277 / 37767	7,40E-03
GO:0009735	response to cytokinin stimulus	20 / 4136	94 / 37767	7,90E-03
GO:0048523	negative regulation of cellular process	40 / 4136	234 / 37767	7,90E-03
GO:0051193	regulation of cofactor metabolic process	6 / 4136	13 / 37767	8,10E-03
GO:0034968	histone lysine methylation	6 / 4136	13 / 37767	8,10E-03
GO:0080090	regulation of primary metabolic process	251 / 4136	1952 / 37767	8,50E-03
GO:0009653	anatomical structure morphogenesis	92 / 4136	637 / 37767	8,70E-03
GO:0048878	chemical homeostasis	26 / 4136	136 / 37767	9,00E-03
GO:0006811	ion transport	65 / 4136	427 / 37767	9,70E-03
GO:0009845	seed germination	18 / 4136	83 / 37767	9,70E-03

Supplementary Table 11. Overrepresented GO terms among PPEP-S5P responsive genes in iGFP-TFIISmut.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with p-value < 0.01 are shown in the table. PPEP-responsive genes were defined as having increased PPEP upon GFP-TFIISmut expression (log_2FC PPEP > 0; β-estradiol vs mock), resulting in clear PPEP establishment upon β-estradiol induction (log_2PPEP > 2).

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	pvalue
GO:0009791	post-embryonic development	179 / 3698	705 / 37767	9,90E-26
GO:0009987	cellular process	1408 / 3698	11684 / 37767	9,30E-19
GO:0043687	post-translational protein modification	221 / 3698	1248 / 37767	6,50E-15
GO:0050896	response to stimulus	557 / 3698	4057 / 37767	1,20E-14
GO:0006468	protein amino acid phosphorylation	174 / 3698	946 / 37767	3,30E-13
GO:0006464	protein modification process	237 / 3698	1474 / 37767	4,90E-12
GO:0016310	phosphorylation	185 / 3698	1079 / 37767	1,20E-11
GO:0006952	defense response	143 / 3698	766 / 37767	1,70E-11
GO:0032501	multicellular organismal process	309 / 3698	2094 / 37767	2,20E-11
GO:0006950	response to stress	335 / 3698	2320 / 37767	2,90E-11
GO:0043412	macromolecule modification	252 / 3698	1636 / 37767	4,70E-11
GO:0007275	multicellular organismal development	296 / 3698	2020 / 37767	1,20E-10
GO:0006793	phosphorus metabolic process	192 / 3698	1179 / 37767	2,20E-10
GO:0006796	phosphate metabolic process	191 / 3698	1178 / 37767	3,40E-10
GO:0065007	biological regulation	534 / 3698	4188 / 37767	1,70E-09
GO:0008152	metabolic process	1211 / 3698	10614 / 37767	2,10E-09

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	pvalue
GO:0003006	reproductive developmental process	159 / 3698	978 / 37767	8,80E-09
GO:0032502	developmental process	317 / 3698	2304 / 37767	9,90E-09
GO:0044238	primary metabolic process	1032 / 3698	8995 / 37767	2,60E-08
GO:0034641	cellular nitrogen compound metabolic process	95 / 3698	506 / 37767	3,30E-08
GO:0044237	cellular metabolic process	1002 / 3698	8722 / 37767	3,70E-08
GO:0048856	anatomical structure development	244 / 3698	1726 / 37767	8,00E-08
GO:0006955	immune response	73 / 3698	367 / 37767	1,80E-07
GO:0002376	immune system process	73 / 3698	368 / 37767	2,00E-07
GO:0008219	cell death	61 / 3698	286 / 37767	2,40E-07
GO:0016265	death	61 / 3698	286 / 37767	2,40E-07
GO:0051234	establishment of localization	255 / 3698	1851 / 37767	2,70E-07
GO:0051179	localization	263 / 3698	1922 / 37767	2,90E-07
GO:0006810	transport	254 / 3698	1846 / 37767	3,10E-07
GO:0022414	reproductive process	173 / 3698	1161 / 37767	3,70E-07
GO:0045087	innate immune response	69 / 3698	347 / 37767	3,90E-07
GO:0000003	reproduction	174 / 3698	1186 / 37767	8,20E-07
GO:0048608	reproductive structure development	148 / 3698	978 / 37767	1,20E-06
GO:0050789	regulation of biological process	455 / 3698	3697 / 37767	1,30E-06
GO:0012501	programmed cell death	52 / 3698	244 / 37767	1,80E-06
GO:0006915	apoptosis	39 / 3698	159 / 37767	2,00E-06
GO:0009743	response to carbohydrate stimulus	51 / 3698	240 / 37767	2,40E-06
GO:0009790	embryonic development	91 / 3698	535 / 37767	2,60E-06
GO:0033036	macromolecule localization	81 / 3698	462 / 37767	3,40E-06
GO:0007165	signal transduction	175 / 3698	1228 / 37767	3,60E-06
GO:0000904	cell morphogenesis involved in differentiation	22 / 3698	63 / 37767	3,70E-06
GO:0006461	protein complex assembly	34 / 3698	134 / 37767	4,80E-06
GO:0070271	protein complex biogenesis	34 / 3698	134 / 37767	4,80E-06
GO:0043170	macromolecule metabolic process	811 / 3698	7127 / 37767	4,80E-06
GO:0051704	multi-organism process	120 / 3698	776 / 37767	4,90E-06
GO:0051716	cellular response to stimulus	127 / 3698	840 / 37767	6,90E-06
GO:0010035	response to inorganic substance	55 / 3698	279 / 37767	7,00E-06
GO:0009793	embryonic development ending in seed dormancy	80 / 3698	465 / 37767	7,10E-06
GO:0048316	seed development	88 / 3698	530 / 37767	8,60E-06
GO:0042221	response to chemical stimulus	270 / 3698	2085 / 37767	9,40E-06
GO:0070887	cellular response to chemical stimulus	77 / 3698	452 / 37767	1,40E-05
GO:0044260	cellular macromolecule metabolic process	734 / 3698	6447 / 37767	1,50E-05
GO:0010154	fruit development	90 / 3698	557 / 37767	1,70E-05
GO:0010200	response to chitin	35 / 3698	151 / 37767	1,90E-05
GO:0048513	organ development	124 / 3698	838 / 37767	2,10E-05
GO:0048731	system development	124 / 3698	839 / 37767	2,20E-05
GO:0050832	defense response to fungus	28 / 3698	108 / 37767	2,30E-05
GO:0050794	regulation of cellular process	407 / 3698	3375 / 37767	2,60E-05
GO:0065008	regulation of biological quality	102 / 3698	665 / 37767	3,20E-05
GO:0006807	nitrogen compound metabolic process	452 / 3698	3826 / 37767	5,20E-05
GO:0051707	response to other organism	92 / 3698	599 / 37767	7,30E-05
GO:0009620	response to fungus	34 / 3698	158 / 37767	8,70E-05
GO:0009607	response to biotic stimulus	96 / 3698	638 / 37767	1,00E-04
GO:0010033	response to organic substance	178 / 3698	1342 / 37767	1,10E-04
GO:0048367	shoot development	60 / 3698	355 / 37767	1,40E-04
GO:0044085	cellular component biogenesis	87 / 3698	571 / 37767	1,40E-04
GO:0022607	cellular component assembly	48 / 3698	265 / 37767	1,60E-04
GO:0009628	response to abiotic stimulus	191 / 3698	1471 / 37767	1,70E-04
GO:0022621	shoot system development	60 / 3698	358 / 37767	1,70E-04
GO:0032870	cellular response to hormone stimulus	55 / 3698	321 / 37767	1,90E-04
GO:0009755	hormone-mediated signaling pathway	55 / 3698	321 / 37767	1,90E-04
GO:0065003	macromolecular complex assembly	40 / 3698	210 / 37767	2,20E-04
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	377 / 3698	3198 / 37767	2,60E-04
GO:0043933	macromolecular complex subunit organization	43 / 3698	235 / 37767	2,80E-04
GO:0016043	cellular component organization	156 / 3698	1179 / 37767	3,00E-04
GO:0048518	positive regulation of biological process	46 / 3698	259 / 37767	3,10E-04
GO:0006996	organelle organization	93 / 3698	640 / 37767	3,70E-04
GO:0048583	regulation of response to stimulus	36 / 3698	188 / 37767	4,00E-04
GO:0070727	cellular macromolecule localization	56 / 3698	341 / 37767	4,30E-04
GO:0006886	intracellular protein transport	52 / 3698	311 / 37767	4,70E-04
GO:0010876	lipid localization	10 / 3698	24 / 37767	5,60E-04
GO:0048519	negative regulation of biological process	68 / 3698	442 / 37767	5,70E-04
GO:0016070	RNA metabolic process	207 / 3698	1657 / 37767	5,70E-04
GO:0007242	intracellular signaling cascade	94 / 3698	659 / 37767	5,90E-04

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	pvalue
GO:0051649	establishment of localization in cell	78 / 3698	525 / 37767	5,90E-04
GO:0040007	growth	55 / 3698	340 / 37767	6,60E-04
GO:0045184	establishment of protein localization	60 / 3698	381 / 37767	7,00E-04
GO:0015031	protein transport	60 / 3698	381 / 37767	7,00E-04
GO:0046907	intracellular transport	70 / 3698	463 / 37767	7,20E-04
GO:0051169	nuclear transport	18 / 3698	71 / 37767	7,80E-04
GO:0006913	nucleocytoplasmic transport	18 / 3698	71 / 37767	7,80E-04
GO:0008104	protein localization	63 / 3698	408 / 37767	8,10E-04
GO:0044262	cellular carbohydrate metabolic process	64 / 3698	417 / 37767	8,50E-04
GO:0034613	cellular protein localization	52 / 3698	322 / 37767	9,40E-04
GO:0051641	cellular localization	82 / 3698	569 / 37767	9,70E-04
GO:0015979	photosynthesis	31 / 3698	162 / 37767	9,80E-04
GO:0006605	protein targeting	30 / 3698	155 / 37767	1,00E-03
GO:0042742	defense response to bacterium	33 / 3698	177 / 37767	1,00E-03
GO:0019538	protein metabolic process	455 / 3698	4009 / 37767	1,00E-03
GO:0048522	positive regulation of cellular process	31 / 3698	163 / 37767	1,10E-03
GO:0048584	positive regulation of response to stimulus	16 / 3698	62 / 37767	1,30E-03
GO:0048827	phyllome development	43 / 3698	258 / 37767	1,40E-03
GO:0009605	response to external stimulus	64 / 3698	429 / 37767	1,60E-03
GO:0010817	regulation of hormone levels	31 / 3698	168 / 37767	1,60E-03
GO:0044106	cellular amine metabolic process	65 / 3698	438 / 37767	1,60E-03
GO:0006520	cellular amino acid metabolic process	64 / 3698	430 / 37767	1,70E-03
GO:0009886	post-embryonic morphogenesis	11 / 3698	35 / 37767	2,00E-03
GO:0009725	response to hormone stimulus	127 / 3698	982 / 37767	2,20E-03
GO:0009653	anatomical structure morphogenesis	87 / 3698	637 / 37767	2,80E-03
GO:0033554	cellular response to stress	59 / 3698	399 / 37767	2,80E-03
GO:0010016	shoot morphogenesis	34 / 3698	198 / 37767	2,80E-03
GO:0046394	carboxylic acid biosynthetic process	61 / 3698	417 / 37767	3,00E-03
GO:0016053	organic acid biosynthetic process	61 / 3698	417 / 37767	3,00E-03
GO:0010038	response to metal ion	39 / 3698	238 / 37767	3,00E-03
GO:0055082	cellular chemical homeostasis	20 / 3698	97 / 37767	3,50E-03
GO:0016049	cell growth	42 / 3698	265 / 37767	3,70E-03
GO:0008652	cellular amino acid biosynthetic process	34 / 3698	202 / 37767	3,70E-03
GO:0044267	cellular protein metabolic process	392 / 3698	3487 / 37767	3,90E-03
GO:0009926	auxin polar transport	14 / 3698	58 / 37767	4,20E-03
GO:0009966	regulation of signal transduction	24 / 3698	128 / 37767	4,30E-03
GO:0009719	response to endogenous stimulus	134 / 3698	1068 / 37767	4,30E-03
GO:0048878	chemical homeostasis	25 / 3698	136 / 37767	4,50E-03
GO:0051239	regulation of multicellular organismal process	43 / 3698	277 / 37767	4,60E-03
GO:0009308	amine metabolic process	72 / 3698	521 / 37767	4,70E-03
GO:0060918	auxin transport	14 / 3698	59 / 37767	4,70E-03
GO:0051273	beta-glucan metabolic process	6 / 3698	13 / 37767	4,90E-03
GO:0051274	beta-glucan biosynthetic process	6 / 3698	13 / 37767	4,90E-03
GO:0006074	1,3-beta-glucan metabolic process	6 / 3698	13 / 37767	4,90E-03
GO:0007033	vacuole organization	6 / 3698	13 / 37767	4,90E-03
GO:0006075	1,3-beta-glucan biosynthetic process	6 / 3698	13 / 37767	4,90E-03
GO:0016570	histone modification	15 / 3698	66 / 37767	5,00E-03
GO:0006855	multidrug transport	16 / 3698	73 / 37767	5,10E-03
GO:0009617	response to bacterium	39 / 3698	247 / 37767	5,20E-03
GO:0009914	hormone transport	14 / 3698	60 / 37767	5,40E-03
GO:0016192	vesicle-mediated transport	42 / 3698	272 / 37767	5,40E-03
GO:0048580	regulation of post-embryonic development	33 / 3698	200 / 37767	5,50E-03
GO:0010646	regulation of cell communication	24 / 3698	131 / 37767	5,50E-03
GO:0009314	response to radiation	82 / 3698	613 / 37767	5,70E-03
GO:0048366	leaf development	37 / 3698	233 / 37767	5,90E-03
GO:0008361	regulation of cell size	42 / 3698	274 / 37767	6,10E-03
GO:0050801	ion homeostasis	20 / 3698	103 / 37767	6,20E-03
GO:0006873	cellular ion homeostasis	19 / 3698	96 / 37767	6,30E-03
GO:0045935	positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	15 / 3698	69 / 37767	7,00E-03
GO:0055080	cation homeostasis	18 / 3698	90 / 37767	7,10E-03
GO:0010604	positive regulation of macromolecule metabolic process	16 / 3698	76 / 37767	7,10E-03
GO:0015893	drug transport	16 / 3698	76 / 37767	7,10E-03
GO:0044271	cellular nitrogen compound biosynthetic process	56 / 3698	394 / 37767	7,10E-03
GO:0032535	regulation of cellular component size	42 / 3698	277 / 37767	7,10E-03
GO:0090066	regulation of anatomical structure size	42 / 3698	277 / 37767	7,10E-03
GO:0048468	cell development	32 / 3698	197 / 37767	7,50E-03
GO:0009611	response to wounding	32 / 3698	197 / 37767	7,50E-03
GO:0048825	cotyledon development	9 / 3698	31 / 37767	7,60E-03

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	pvalue
GO:0045941	positive regulation of transcription	14 / 3698	63 / 37767	7,70E-03
GO:0010628	positive regulation of gene expression	14 / 3698	63 / 37767	7,70E-03
GO:0031325	positive regulation of cellular metabolic process	18 / 3698	91 / 37767	7,80E-03
GO:0051173	positive regulation of nitrogen compound metabolic process	15 / 3698	70 / 37767	7,80E-03
GO:0030003	cellular cation homeostasis	17 / 3698	84 / 37767	7,90E-03
GO:0031399	regulation of protein modification process	5 / 3698	10 / 37767	7,90E-03
GO:0042493	response to drug	16 / 3698	77 / 37767	7,90E-03
GO:0009416	response to light stimulus	79 / 3698	596 / 37767	8,00E-03
GO:0009967	positive regulation of signal transduction	7 / 3698	20 / 37767	8,00E-03
GO:0009910	negative regulation of flower development	13 / 3698	57 / 37767	8,40E-03
GO:0009789	positive regulation of abscisic acid mediated signaling pathway	6 / 3698	15 / 37767	8,40E-03
GO:0051093	negative regulation of developmental process	23 / 3698	129 / 37767	8,50E-03
GO:0009893	positive regulation of metabolic process	18 / 3698	92 / 37767	8,50E-03
GO:0016569	covalent chromatin modification	15 / 3698	71 / 37767	8,70E-03
GO:0042592	homeostatic process	34 / 3698	216 / 37767	9,00E-03
GO:0051276	chromosome organization	34 / 3698	216 / 37767	9,00E-03
GO:0010118	stomatal movement	13 / 3698	58 / 37767	9,40E-03
GO:0034637	cellular carbohydrate biosynthetic process	29 / 3698	177 / 37767	9,50E-03
GO:0017038	protein import	17 / 3698	86 / 37767	9,50E-03
GO:0007049	cell cycle	41 / 3698	275 / 37767	9,90E-03

Supplementary Table 12. Overrepresented GO terms among genes with significantly increased PPEP-S2P&S5P establishment upon GFP-TFIISmut expression.

The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with false discovery rate < 0.01 are shown in the table. RNAPII-S2P and RNAPII-S5P read counts from ChIP-seq were merged for this analysis. Only genes with significantly increased PPEP (z-score > 2) upon GFP-TFIISmut expression (β-estradiol vs mock) resulting in PPEP establishment upon β-estradiol induction (log_2PPEP > 1, β–estradiol) were considered for the analysis.

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0009791	post-embryonic development	40 / 653	705 / 37767	2,40E-10	3,40E-07
GO:0009987	cellular process	274 / 653	11684 / 37767	2,60E-09	1,90E-06
GO:0003006	reproductive developmental process	41 / 653	978 / 37767	4,40E-07	2,20E-04
GO:0009790	embryonic development	26 / 653	535 / 37767	5,10E-06	1,20E-03
GO:0022414	reproductive process	43 / 653	1161 / 37767	5,00E-06	1,20E-03
GO:0048856	anatomical structure development	57 / 653	1726 / 37767	4,40E-06	1,20E-03
GO:0048608	reproductive structure development	38 / 653	978 / 37767	6,40E-06	1,30E-03
GO:0000003	reproduction	43 / 653	1186 / 37767	8,30E-06	1,50E-03
GO:0009793	embryonic development ending in seed dormancy	23 / 653	465 / 37767	1,30E-05	2,20E-03
GO:0032501	multicellular organismal process	63 / 653	2094 / 37767	2,30E-05	3,30E-03
GO:0048316	seed development	24 / 653	530 / 37767	3,40E-05	4,60E-03
GO:0048513	organ development	32 / 653	838 / 37767	4,70E-05	5,30E-03
GO:0048731	system development	32 / 653	839 / 37767	4,80E-05	5,30E-03
GO:0007275	multicellular organismal development	60 / 653	2020 / 37767	5,10E-05	5,30E-03
GO:0010154	fruit development	24 / 653	557 / 37767	7,20E-05	7,00E-03
GO:0051649	establishment of localization in cell	23 / 653	525 / 37767	8,00E-05	7,30E-03
GO:0009653	anatomical structure morphogenesis	26 / 653	637 / 37767	8,60E-05	7,40E-03
GO:0034641	cellular nitrogen compound metabolic process	22 / 653	506 / 37767	1,30E-04	1,00E-02
GO:0008152	metabolic process	227 / 653	10614 / 37767	1,40E-04	1,10E-02
GO:0045184	establishment of protein localization	18 / 653	381 / 37767	2,00E-04	1,30E-02
GO:0000904	cell morphogenesis involved in differentiation	7 / 653	63 / 37767	1,90E-04	1,30E-02
GO:0015031	protein transport	18 / 653	381 / 37767	2,00E-04	1,30E-02
GO:0051641	cellular localization	23 / 653	569 / 37767	2,50E-04	1,60E-02
GO:0016043	cellular component organization	38 / 653	1179 / 37767	2,80E-04	1,60E-02
GO:0046907	intracellular transport	20 / 653	463 / 37767	2,70E-04	1,60E-02
GO:0006810	transport	53 / 653	1846 / 37767	3,10E-04	1,70E-02
GO:0032502	developmental process	63 / 653	2304 / 37767	3,10E-04	1,70E-02
GO:0051234	establishment of localization	53 / 653	1851 / 37767	3,30E-04	1,70E-02
GO:0006807	nitrogen compound metabolic process	94 / 653	3826 / 37767	4,10E-04	2,10E-02
GO:0008104	protein localization	18 / 653	408 / 37767	4,30E-04	2,10E-02
GO:0065007	biological regulation	101 / 653	4188 / 37767	4,60E-04	2,20E-02

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0006886	intracellular protein transport	15 / 653	311 / 37767	5,30E-04	2,30E-02
GO:0006996	organelle organization	24 / 653	640 / 37767	5,10E-04	2,30E-02
GO:0006605	protein targeting	10 / 653	155 / 37767	5,60E-04	2,40E-02
GO:0033036	macromolecule localization	19 / 653	462 / 37767	6,80E-04	2,80E-02
GO:0034613	cellular protein localization	15 / 653	322 / 37767	7,40E-04	3,00E-02
GO:0051179	localization	53 / 653	1922 / 37767	7,70E-04	3,00E-02
GO:0006418	tRNA aminoacylation for protein translation	7 / 653	84 / 37767	9,40E-04	3,40E-02
GO:0043038	amino acid activation	7 / 653	84 / 37767	9,40E-04	3,40E-02
GO:0043039	tRNA aminoacylation	7 / 653	84 / 37767	9,40E-04	3,40E-02
GO:0009734	auxin mediated signaling pathway	5 / 653	42 / 37767	1,20E-03	4,20E-02
GO:0070727	cellular macromolecule localization	15 / 653	341 / 37767	1,30E-03	4,50E-02

Supplementary Table 13. The list of nuclear proteins copurified specifically with GS-TFIISmut.
The list of transcription-related proteins copurified with GS-TFIISmut but not with GS-TFIIS during AP-MS approach (Figure 38). The numbers indicate the respective average MASCOT score and how many times the proteins were detected in three independent experiments.

AGI	Interactor	GS-TFIIS
AT1G32130	IWS1a	256 / 3
AT5G27770	Ribosomal L22e protein family	254 / 3
AT3G05560	Ribosomal L22e protein family	212 / 3
AT1G52980	GTP-binding family protein	178 / 3
AT1G09100	26S proteasome AAA-ATPase subunit RPT5B	178 / 2
AT3G53350	ROP interactive partner 4	169 / 2
AT3G56720	pre-mRNA-splicing factor	168 / 3
AT5G19990	regulatory particle triple-A ATPase 6A	161 / 2
AT2G19540	AtCAF1CL6	153 / 2
AT1G32750	AtSPT7 / HAF1	146 / 3
AT5G63260	Zinc finger C-x8-C-x5-C-x3-H type family protein	143 / 2
AT5G14460	Pseudouridine synthase family protein	134 / 2
AT4G10840	Tetratricopeptide repeat (TPR)-like superfamily protein	132 / 3
AT4G28450	WD repeat and SOF domain-containing protein 1	130 / 2
AT5G08670	ATP synthase alpha/beta family protein	130 / 3
AT1G04510	MOS4-associated complex 3A (MAC3A), PRP19A	129 / 2
AT2G17250	CCAAT-binding factor	127 / 3
AT4G34670	Ribosomal protein S3Ae	127 / 2
AT4G32610	copper ion binding protein	120 / 2
AT5G03040	IQ-domain 2	112 / 2
AT1G60650	at-hnRNP-G3 RNA recognition motif-containing protein	107 / 2

Supplementary Table 14. The efficiency of β-estradiol induction in iGFP-TFIIS and iGFP-TFIISmut.
The fraction of nuclei with GFP expression in iGFP-TFIIS and iGFP-TFIISmut seedlings was determined as defined on Supplementary Figure S 10.

	iGFP-TFIIS		iGFP-TFIISmut	
	24h	72h	24h	72h
% nuclei GFP positive	25,8 %	17,9 %	33,1 %	21,1 %
nuclei total	4833	5214	4985	5286
nuclei GFP positive	1233	938	1649	1121
nuclei GFP negative	3498	4160	3295	4099

Supplementary Table 15. Overrepresented GO terms among proteins copurified with ELF7.
The Gene Ontology (GO) analysis was performed using the single enrichment analysis (SEA) of AgriGO. All overrepresented GO terms with false discovery rate < 0.01 are shown in the table. ELF7-SG affinity purification was performed by Hans Ehrnsberger.

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0006412	translation	122 / 535	1445 / 37767	1,20E-55	1,90E-52
GO:0044260	cellular macromolecule metabolic process	243 / 535	6447 / 37767	2,30E-51	1,80E-48
GO:0043170	macromolecule metabolic process	251 / 535	7127 / 37767	3,50E-48	1,80E-45
GO:0009987	cellular process	330 / 535	11684 / 37767	9,20E-48	3,60E-45
GO:0010467	gene expression	177 / 535	3962 / 37767	3,20E-44	1,00E-41
GO:0044267	cellular protein metabolic process	161 / 535	3487 / 37767	2,20E-41	5,70E-39
GO:0034645	cellular macromolecule biosynthetic process	163 / 535	3661 / 37767	4,60E-40	1,00E-37
GO:0009059	macromolecule biosynthetic process	163 / 535	3685 / 37767	1,00E-39	2,00E-37
GO:0044237	cellular metabolic process	263 / 535	8722 / 37767	8,40E-39	1,50E-36
GO:0019538	protein metabolic process	166 / 535	4009 / 37767	4,30E-37	6,80E-35
GO:0044238	primary metabolic process	262 / 535	8995 / 37767	5,90E-36	8,40E-34
GO:0044249	cellular biosynthetic process	179 / 535	4925 / 37767	2,80E-33	3,60E-31
GO:0009058	biosynthetic process	179 / 535	5118 / 37767	3,00E-31	3,60E-29
GO:0008152	metabolic process	274 / 535	10614 / 37767	5,20E-29	5,80E-27
GO:0044085	cellular component biogenesis	44 / 535	571 / 37767	1,10E-18	1,20E-16
GO:0042254	ribosome biogenesis	29 / 535	241 / 37767	2,60E-17	2,60E-15
GO:0022613	ribonucleoprotein complex biogenesis	29 / 535	253 / 37767	8,40E-17	7,80E-15
GO:0009791	post-embryonic development	45 / 535	705 / 37767	3,50E-16	3,00E-14
GO:0003006	reproductive developmental process	45 / 535	978 / 37767	1,80E-11	1,50E-09
GO:0007275	multicellular organismal development	69 / 535	2020 / 37767	3,50E-11	2,80E-09
GO:0032501	multicellular organismal process	70 / 535	2094 / 37767	6,30E-11	4,70E-09
GO:0048856	anatomical structure development	61 / 535	1726 / 37767	1,50E-10	1,10E-08
GO:0022414	reproductive process	46 / 535	1161 / 37767	1,10E-09	7,60E-08
GO:0032502	developmental process	71 / 535	2304 / 37767	1,40E-09	8,80E-08
GO:0000003	reproduction	46 / 535	1186 / 37767	2,10E-09	1,30E-07
GO:0048316	seed development	28 / 535	530 / 37767	8,10E-09	4,90E-07
GO:0009793	embryonic development ending in seed dormancy	26 / 535	465 / 37767	9,40E-09	5,50E-07
GO:0009790	embryonic development	28 / 535	535 / 37767	9,80E-09	5,50E-07
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	86 / 535	3198 / 37767	1,10E-08	6,10E-07
GO:0016043	cellular component organization	44 / 535	1179 / 37767	1,40E-08	7,50E-07
GO:0048608	reproductive structure development	39 / 535	978 / 37767	1,80E-08	9,30E-07
GO:0010154	fruit development	28 / 535	557 / 37767	2,20E-08	1,10E-06
GO:0006996	organelle organization	30 / 535	640 / 37767	3,10E-08	1,50E-06
GO:0006886	intracellular protein transport	20 / 535	311 / 37767	5,80E-08	2,70E-06
GO:0006325	chromatin organization	15 / 535	175 / 37767	8,70E-08	3,90E-06
GO:0034613	cellular protein localization	20 / 535	322 / 37767	9,90E-08	4,30E-06
GO:0044265	cellular macromolecule catabolic process	24 / 535	465 / 37767	1,40E-07	6,10E-06
GO:0046164	alcohol catabolic process	11 / 535	89 / 37767	1,60E-07	6,80E-06
GO:0006457	protein folding	18 / 535	275 / 37767	2,10E-07	8,20E-06
GO:0051276	chromosome organization	16 / 535	216 / 37767	2,10E-07	8,20E-06
GO:0070727	cellular macromolecule localization	20 / 535	341 / 37767	2,40E-07	9,00E-06
GO:0045184	establishment of protein localization	21 / 535	381 / 37767	3,10E-07	1,10E-05
GO:0015031	protein transport	21 / 535	381 / 37767	3,10E-07	1,10E-05
GO:0006096	glycolysis	9 / 535	57 / 37767	3,50E-07	1,20E-05
GO:0006007	glucose catabolic process	10 / 535	83 / 37767	7,40E-07	2,60E-05
GO:0019320	hexose catabolic process	10 / 535	84 / 37767	8,20E-07	2,70E-05
GO:0046365	monosaccharide catabolic process	10 / 535	84 / 37767	8,20E-07	2,70E-05
GO:0008104	protein localization	21 / 535	408 / 37767	9,00E-07	2,90E-05
GO:0006006	glucose metabolic process	10 / 535	86 / 37767	1,00E-06	3,20E-05
GO:0051649	establishment of localization in cell	24 / 535	525 / 37767	1,10E-06	3,50E-05
GO:0006807	nitrogen compound metabolic process	90 / 535	3826 / 37767	1,70E-06	5,10E-05
GO:0006268	DNA unwinding during replication	5 / 535	11 / 37767	2,00E-06	6,10E-05
GO:0022607	cellular component assembly	16 / 535	265 / 37767	2,60E-06	7,80E-05
GO:0006259	DNA metabolic process	20 / 535	405 / 37767	2,90E-06	8,60E-05
GO:0009266	response to temperature stimulus	22 / 535	485 / 37767	3,50E-06	1,00E-04
GO:0044275	cellular carbohydrate catabolic process	11 / 535	125 / 37767	3,60E-06	1,00E-04
GO:0051641	cellular localization	24 / 535	569 / 37767	4,20E-06	1,20E-04
GO:0016052	carbohydrate catabolic process	11 / 535	128 / 37767	4,50E-06	1,20E-04
GO:0033036	macromolecule localization	21 / 535	462 / 37767	5,70E-06	1,50E-04
GO:0046907	intracellular transport	21 / 535	463 / 37767	5,90E-06	1,50E-04
GO:0000904	cell morphogenesis involved in differentiation	8 / 535	63 / 37767	6,80E-06	1,70E-04
GO:0032392	DNA geometric change	5 / 535	16 / 37767	8,80E-06	2,20E-04

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0032508	DNA duplex unwinding	5 / 535	16 / 37767	8,80E-06	2,20E-04
GO:0000059	protein import into nucleus, docking	5 / 535	17 / 37767	1,10E-05	2,80E-04
GO:0006396	RNA processing	20 / 535	453 / 37767	1,40E-05	3,40E-04
GO:0016192	vesicle-mediated transport	15 / 535	272 / 37767	1,50E-05	3,60E-04
GO:0065003	macromolecular complex assembly	13 / 535	210 / 37767	1,80E-05	4,10E-04
GO:0019318	hexose metabolic process	10 / 535	126 / 37767	2,30E-05	5,30E-04
GO:0044248	cellular catabolic process	26 / 535	746 / 37767	4,10E-05	9,40E-04
GO:0048588	developmental cell growth	8 / 535	84 / 37767	4,60E-05	1,00E-03
GO:0006606	protein import into nucleus	6 / 535	41 / 37767	4,80E-05	1,10E-03
GO:0043933	macromolecular complex subunit organization	13 / 535	235 / 37767	5,30E-05	1,10E-03
GO:0017038	protein import	8 / 535	86 / 37767	5,30E-05	1,10E-03
GO:0051170	nuclear import	6 / 535	42 / 37767	5,40E-05	1,10E-03
GO:0033205	cytokinesis during cell cycle	5 / 535	26 / 37767	6,60E-05	1,40E-03
GO:0006413	translational initiation	8 / 535	90 / 37767	7,20E-05	1,50E-03
GO:0034504	protein localization in nucleus	6 / 535	45 / 37767	7,60E-05	1,60E-03
GO:0030154	cell differentiation	16 / 535	355 / 37767	8,00E-05	1,60E-03
GO:0051301	cell division	8 / 535	99 / 37767	1,30E-04	2,70E-03
GO:0009292	genetic transfer	5 / 535	32 / 37767	1,60E-04	3,00E-03
GO:0009294	DNA mediated transformation	5 / 535	32 / 37767	1,60E-04	3,00E-03
GO:0040007	growth	15 / 535	340 / 37767	1,70E-04	3,20E-03
GO:0006950	response to stress	55 / 535	2320 / 37767	1,80E-04	3,40E-03
GO:0009653	anatomical structure morphogenesis	22 / 535	637 / 37767	1,80E-04	3,40E-03
GO:0048468	cell development	11 / 535	197 / 37767	1,80E-04	3,40E-03
GO:0051704	multi-organism process	25 / 535	776 / 37767	1,90E-04	3,40E-03
GO:0009846	pollen germination	5 / 535	34 / 37767	2,00E-04	3,70E-03
GO:0034622	cellular macromolecular complex assembly	11 / 535	200 / 37767	2,10E-04	3,70E-03
GO:0005996	monosaccharide metabolic process	10 / 535	168 / 37767	2,20E-04	3,90E-03
GO:0009910	negative regulation of flower development	6 / 535	57 / 37767	2,50E-04	4,40E-03
GO:0008380	RNA splicing	8 / 535	115 / 37767	3,50E-04	6,00E-03
GO:0034621	cellular macromolecular complex subunit organization	11 / 535	224 / 37767	5,20E-04	8,80E-03
GO:0000910	cytokinesis	5 / 535	43 / 37767	5,40E-04	9,20E-03
GO:0016568	chromatin modification	7 / 535	95 / 37767	5,90E-04	9,80E-03
GO:0009408	response to heat	9 / 535	161 / 37767	6,90E-04	1,10E-02
GO:0051169	nuclear transport	6 / 535	71 / 37767	7,40E-04	1,20E-02
GO:0006913	nucleocytoplasmic transport	6 / 535	71 / 37767	7,40E-04	1,20E-02
GO:0048519	negative regulation of biological process	16 / 535	442 / 37767	8,20E-04	1,30E-02
GO:0070271	protein complex biogenesis	8 / 535	134 / 37767	9,00E-04	1,40E-02
GO:0006461	protein complex assembly	8 / 535	134 / 37767	9,00E-04	1,40E-02
GO:0048589	developmental growth	11 / 535	242 / 37767	9,50E-04	1,50E-02
GO:0009116	nucleoside metabolic process	5 / 535	50 / 37767	1,00E-03	1,60E-02
GO:0016071	mRNA metabolic process	8 / 535	139 / 37767	1,10E-03	1,70E-02
GO:0009409	response to cold	13 / 535	328 / 37767	1,10E-03	1,70E-02
GO:0009932	cell tip growth	6 / 535	79 / 37767	1,20E-03	1,90E-02
GO:0048581	negative regulation of post-embryonic development	6 / 535	80 / 37767	1,30E-03	1,90E-02
GO:0033365	protein localization in organelle	6 / 535	80 / 37767	1,30E-03	1,90E-02
GO:0046686	response to cadmium ion	9 / 535	178 / 37767	1,30E-03	1,90E-02
GO:0006333	chromatin assembly or disassembly	6 / 535	82 / 37767	1,50E-03	2,10E-02
GO:0009615	response to virus	5 / 535	55 / 37767	1,50E-03	2,20E-02
GO:0048869	cellular developmental process	17 / 535	520 / 37767	1,70E-03	2,30E-02
GO:0009860	pollen tube growth	5 / 535	58 / 37767	1,90E-03	2,60E-02
GO:0006261	DNA-dependent DNA replication	5 / 535	60 / 37767	2,20E-03	3,00E-02
GO:0006605	protein targeting	8 / 535	155 / 37767	2,20E-03	3,00E-02
GO:0006066	alcohol metabolic process	11 / 535	270 / 37767	2,20E-03	3,00E-02
GO:0009628	response to abiotic stimulus	35 / 535	1471 / 37767	2,50E-03	3,40E-02
GO:0010035	response to inorganic substance	11 / 535	279 / 37767	2,80E-03	3,70E-02
GO:0016070	RNA metabolic process	38 / 535	1657 / 37767	3,10E-03	4,10E-02
GO:0006091	generation of precursor metabolites and energy	11 / 535	285 / 37767	3,20E-03	4,30E-02
GO:0006281	DNA repair	9 / 535	214 / 37767	4,40E-03	5,70E-02
GO:0009057	macromolecule catabolic process	25 / 535	982 / 37767	4,50E-03	5,80E-02
GO:0000902	cell morphogenesis	10 / 535	259 / 37767	4,90E-03	6,30E-02
GO:0009909	regulation of flower development	7 / 535	141 / 37767	4,90E-03	6,30E-02
GO:0010605	negative regulation of macromolecule metabolic process	9 / 535	219 / 37767	5,00E-03	6,40E-02
GO:0009553	embryo sac development	6 / 535	107 / 37767	5,20E-03	6,50E-02
GO:0006974	response to DNA damage stimulus	9 / 535	221 / 37767	5,30E-03	6,60E-02
GO:0016049	cell growth	10 / 535	265 / 37767	5,70E-03	7,00E-02
GO:0022402	cell cycle process	7 / 535	149 / 37767	6,50E-03	8,00E-02
GO:0008361	regulation of cell size	10 / 535	274 / 37767	7,00E-03	8,50E-02
GO:0009913	epidermal cell differentiation	7 / 535	151 / 37767	7,00E-03	8,50E-02
GO:0090066	regulation of anatomical structure size	10 / 535	277 / 37767	7,50E-03	8,80E-02

GO term ID	description	queryitem / querytotal	bgitem / bgtotal	p-value	FDR
GO:0032535	regulation of cellular component size	10 / 535	277 / 37767	7,50E-03	8,80E-02
GO:0051239	regulation of multicellular organismal process	10 / 535	277 / 37767	7,50E-03	8,80E-02
GO:0008544	epidermis development	7 / 535	153 / 37767	7,50E-03	8,80E-02
GO:0007398	ectoderm development	7 / 535	153 / 37767	7,50E-03	8,80E-02
GO:0006260	DNA replication	6 / 535	117 / 37767	7,80E-03	8,90E-02
GO:0043414	macromolecule methylation	5 / 535	83 / 37767	7,90E-03	8,90E-02
GO:0050896	response to stimulus	76 / 535	4057 / 37767	7,90E-03	8,90E-02
GO:0009892	negative regulation of metabolic process	9 / 535	236 / 37767	7,90E-03	8,90E-02
GO:0010038	response to metal ion	9 / 535	238 / 37767	8,30E-03	9,30E-02
GO:0010629	negative regulation of gene expression	8 / 535	197 / 37767	8,50E-03	9,50E-02

Supplementary Table 16. GO terms associated with DNA replication, DNA damage or DNA repair.
All GO terms potentially related to "DNA replication", "DNA damage" and "DNA repair" processes were manually extracted from all Gene Ontology annotations available in TAIR database.

GO term ID	description	Process
GO:0006260	DNA replication	DNA replication
GO:0006270	DNA replication initiation	DNA replication
GO:0033567	DNA replication, Okazaki fragment processing	DNA replication
GO:0006269	DNA replication, synthesis of RNA primer	DNA replication
GO:0006335	DNA replication-dependent nucleosome assembly	DNA replication
GO:0006271	DNA strand elongation involved in DNA replication	DNA replication
GO:0006265	DNA topological change	DNA replication
GO:0006261	DNA-dependent DNA replication	DNA replication
GO:0033314	mitotic DNA replication checkpoint	DNA replication
GO:1902979	mitotic DNA replication termination	DNA replication
GO:0045740	positive regulation of DNA replication	DNA replication
GO:0006275	regulation of DNA replication	DNA replication
GO:0031297	replication fork processing	DNA replication
GO:0006974	cellular response to DNA damage stimulus	DNA damage
GO:0000077	DNA damage checkpoint	DNA damage
GO:2001022	positive regulation of response to DNA damage stimulus	DNA damage
GO:2001020	regulation of response to DNA damage stimulus	DNA damage
GO:1902504	regulation of signal transduction involved in mitotic G2 DNA damage checkpoint	DNA damage
GO:0051103	DNA ligation involved in DNA repair	DNA repair
GO:0006281	DNA repair	DNA repair
GO:0006302	double-strand break repair	DNA repair
GO:0000724	double-strand break repair via homologous recombination	DNA repair
GO:0006303	double-strand break repair via nonhomologous end joining	DNA repair
GO:0006282	regulation of DNA repair	DNA repair
GO:2000779	regulation of double-strand break repair	DNA repair
GO:0010569	regulation of double-strand break repair via homologous recombination	DNA repair
GO:0006283	transcription-coupled nucleotide-excision repair	DNA repair

List of figures

List of tables

Publications

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., **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., Deforges J., Begcy K., Poirier Y., Bruckmann A., Dresselhaus T., Grasser KD. (2019). Dominant negative version of TFIIS reveals the importance of RNAPII pausing in *Arabidopsis*.
In preparation

Antosz W., Begcy K., ElBashir R., Bruckmann A., Dresselhaus T., Grasser KD. (2019). Molecular properties underlying the genetic interaction between *Arabidopsis* TFIIS and PAF1-C.
In preparation

Bibliography

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat. Rev. Genet. *13*, 720–731.

Adelman, K., Marr, M.T., Werner, J., Saunders, A., Ni, Z., Andrulis, E.D., and Lis, J.T. (2005). Efficient Release from Promoter-Proximal Stall Sites Requires Transcript Cleavage Factor TFIIS. Mol. Cell *17*, 103–112.

Albertazzi, L., Arosio, D., Marchetti, L., Ricci, F., and Beltram, F. (2009). Quantitative FRET analysis with the EGFP-mCherry fluorescent protein pair. Photochem. Photobiol. *85*, 287–297.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). From DNA to RNA. Mol. Biol. Cell 4th Ed.

Andrulis, E.D., Guzmán, E., Döring, P., Werner, J., and Lis, J.T. (2000). High-resolution localization of Drosophila Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. Genes Dev. *14*, 2635–2649.

Antosz, W., Pfab, A., Ehrnsberger, H.F., Holzinger, P., Köllen, K., Mortensen, S.A., Bruckmann, A., Schubert, T., Längst, G., Griesenbeck, J., et al. (2017). The Composition of the Arabidopsis RNA Polymerase II Transcript Elongation Complex Reveals the Interplay between Elongation and mRNA Processing Factors. Plant Cell *29*, 854–870.

Archambault, J., Lacroute, F., Ruet, A., and Friesen, J.D. (1992). Genetic interaction between transcription elongation factor TFIIS and RNA polymerase II. Mol. Cell. Biol. *12*, 4142–4152.

Ardehali, M.B., and Lis, J.T. (2009). Tracking rates of transcription and splicing in vivo. Nat. Struct. Mol. Biol. *16*, 1123–1124.

Arimbasseri, A.G., Rijal, K., and Maraia, R.J. (2013). Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. Transcription *5*.

Asl, L.K., Dhondt, S., Boudolf, V., Beemster, G.T.S., Beeckman, T., Inzé, D., Govaerts, W., and Veylder, L.D. (2011). Model-Based Analysis of Arabidopsis Leaf Epidermal Cells Reveals Distinct Division and Expansion Patterns for Pavement and Guard Cells. Plant Physiol. *156*, 2172–2183.

Awrey, D.E., Shimasaki, N., Koth, C., Weilbaecher, R., Olmsted, V., Kazanis, S., Shan, X., Arellano, J., Arrowsmith, C.H., Kane, C.M., et al. (1998). Yeast transcript elongation factor (TFIIS), structure and function. II: RNA polymerase binding, transcript cleavage, and read-through. J. Biol. Chem. *273*, 22595–22605.

Bar-Nahum, G., Epshtein, V., Ruckenstein, A.E., Rafikov, R., Mustaev, A., and Nudler, E. (2005). A ratchet mechanism of transcription elongation and its control. Cell *120*, 183–193.

Barow, M., and Meister, A. (2003). Endopolyploidy in seed plants is differently correlated to systematics, organ, life strategy and genome size. Plant Cell Environ. *26*, 571–584.

Bensaude, O. (2011). Inhibiting eukaryotic transcription. Transcription *2*, 103–108.

Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. Nat. Rev. Genet. *15*, 163–175.

Bochkareva, A., Yuzenkova, Y., Tadigotla, V.R., and Zenkin, N. (2012). Factor-independent transcription pausing caused by recognition of the RNA–DNA hybrid sequence. EMBO J. *31*, 630–639.

Booth, V., Koth, C.M., Edwards, A.M., and Arrowsmith, C.H. (2000). Structure of a Conserved Domain Common to the Transcription Factors TFIIS, Elongin A, and CRSP70. J. Biol. Chem. *275*, 31266–31268.

Borukhov, S., Sagitov, V., and Goldfarb, A. (1993). Transcript cleavage factors from *E. coli*. Cell *72*, 459–466.

Brand, L., Hörler, M., Nüesch, E., Vassalli, S., Barrell, P., Yang, W., Jefferson, R.A., Grossniklaus, U., and Curtis, M.D. (2006). A Versatile and Reliable Two-Component System for Tissue-Specific Gene Induction in Arabidopsis. Plant Physiol. *141*, 1194–1204.

Brueckner, F., Ortiz, J., and Cramer, P. (2009). A movie of the RNA polymerase nucleotide addition cycle. Curr. Opin. Struct. Biol. *19*, 294–299.

Buckley, M.S., Kwak, H., Zipfel, W.R., and Lis, J.T. (2014). Kinetics of promoter Pol II on Hsp70 reveal stable pausing and key insights into its regulation. Genes Dev. *28*, 14–19.

Čabart, P., Jin, H., Li, L., and Kaplan, C.D. (2014). Activation and reactivation of the RNA polymerase II trigger loop for intrinsic RNA cleavage and catalysis. Transcription *5*, e28869.

Cai, H., and Luse, D.S. (1987). Transcription initiation by RNA polymerase II in vitro. Properties of preinitiation, initiation, and elongation complexes. J. Biol. Chem. *262*, 298–304.

Cao, Q.-F., Yamamoto, J., Isobe, T., Tateno, S., Murase, Y., Chen, Y., Handa, H., and Yamaguchi, Y. (2015). Characterization of the Human Transcription Elongation Factor Rtf1: Evidence for Nonoverlapping Functions of Rtf1 and the Paf1 Complex. Mol. Cell. Biol. *35*, 3459–3470.

Chédin, S., Riva, M., Schultz, P., Sentenac, A., and Carles, C. (1998). The RNA cleavage activity of RNA polymerase III is mediated by an essential TFIIS-like subunit and is important for transcription termination. Genes Dev. *12*, 3857–3871.

Chen, F.X., Woodfin, A.R., Gardini, A., Rickels, R.A., Marshall, S.A., Smith, E.R., Shiekhattar, R., and Shilatifard, A. (2015). PAF1, a Molecular Regulator of Promoter-Proximal Pausing by RNA Polymerase II. Cell *162*, 1003–1015.

Chen, F.X., Smith, E.R., and Shilatifard, A. (2018a). Born to run: control of transcription elongation by RNA polymerase II. Nat. Rev. Mol. Cell Biol. *19*, 464–478.

Chen, P., Dong, L., Hu, M., Wang, Y.-Z., Xiao, X., Zhao, Z., Yan, J., Wang, P.-Y., Reinberg, D., Li, M., et al. (2018b). Functions of FACT in Breaking the Nucleosome and Maintaining Its Integrity at the Single-Nucleosome Level. Mol. Cell *71*, 284-293.e4.

Chen, Y.-H., Keegan, S., Kahli, M., Tonzi, P., Fenyo, D., Huang, T.T., and Smith, D.J. (2018c). Transcription drives DNA replication initiation and termination in human cells. BioRxiv 324079.

Cheung, A.C.M., and Cramer, P. (2011). Structural basis of RNA polymerase II backtracking, arrest and reactivation. Nature *471*, 249–253.

Chodavarapu, R.K., Feng, S., Bernatavichute, Y.V., Chen, P.-Y., Stroud, H., Yu, Y., Hetzel, J.A., Kuo, F., Kim, J., Cokus, S.J., et al. (2010). Relationship between nucleosome positioning and DNA methylation. Nature *466*, 388–392.

Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature *469*, 368–373.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. Cell Mol. Biol. *16*, 735–743.

Cojocaru, M., Bouchard, A., Cloutier, P., Cooper, J.J., Varzavand, K., Price, D.H., and Coulombe, B. (2011). Transcription Factor IIS Cooperates with the E3 Ligase UBR5 to Ubiquitinate the CDK9 Subunit of the Positive Transcription Elongation Factor B. J. Biol. Chem. *286*, 5012–5022.

Colón-Carmona, A., You, R., Haimovitch-Gal, T., and Doerner, P. (1999). Technical advance: spatio-temporal analysis of mitotic activity with a labile cyclin-GUS fusion protein. Plant J. Cell Mol. Biol. *20*, 503–508.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845–1848.

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science *292*, 1863–1876.

Croager, E. (2004). Plant cell biology: Protein degradation protects plants. Nat. Rev. Mol. Cell Biol. *5*, 85.

Da, L.-T., Pardo-Avila, F., Xu, L., Silva, D.-A., Zhang, L., Gao, X., Wang, D., and Huang, X. (2016). Bridge helix bending promotes RNA polymerase II backtracking through a critical and conserved threonine residue. Nat. Commun. *7*, 11244.

Danko, C.G., Hah, N., Luo, X., Martins, A.L., Core, L., Lis, J.T., Siepel, A., and Kraus, W.L. (2013). Signaling Pathways Differentially Affect RNA Polymerase II Initiation, Pausing, and Elongation Rate in Cells. Mol. Cell *50*, 212–222.

Darzacq, X., Shav-Tal, Y., de Turris, V., Brody, Y., Shenoy, S.M., Phair, R.D., and Singer, R.H. (2007). In vivo dynamics of RNA polymerase II transcription. Nat. Struct. Mol. Biol. *14*, 796–806.

Day, D.S., Zhang, B., Stevens, S.M., Ferrari, F., Larschan, E.N., Park, P.J., and Pu, W.T. (2016). Comprehensive analysis of promoter-proximal RNA polymerase II pausing across mammalian cell types. Genome Biol. *17*, 120.

Deepak, K., Kj, Y., Tn, S., Ba, V., and Sa, D. (2015). The Conflict of Transcription-Replication Coordination: Understanding in the Perspective of Genome Integrity. Cloning Transgenesis *4*, 1–5.

Deighan, P., Pukhrambam, C., Nickels, B.E., and Hochschild, A. (2011). Initial transcribed region sequences influence the composition and functional properties of the bacterial elongation complex. Genes Dev. *25*, 77–88.

Dermody, J.L., and Buratowski, S. (2010). Leo1 subunit of the yeast paf1 complex binds RNA and contributes to complex recruitment. J. Biol. Chem. *285*, 33671–33679.

Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., and Pagano, A. (2007). The expanding RNA polymerase III transcriptome. Trends Genet. TIG *23*, 614–622.

Ding, Y., Avramova, Z., and Fromm, M. (2011). Two Distinct Roles of ARABIDOPSIS HOMOLOG OF TRITHORAX1 (ATX1) at Promoters and within Transcribed Regions of ATX1-Regulated Genes. Plant Cell *23*, 350–363.

Dolata, J., Guo, Y., Kołowerzo, A., Smoliński, D., Brzyżek, G., Jarmołowski, A., and Świeżewski, S. (2015). NTR1 is required for transcription elongation checkpoints at alternative exons in Arabidopsis. EMBO J. *34*, 544–558.

Dronamraju, R., and Strahl, B.D. (2014). A feed forward circuit comprising Spt6, Ctk1 and PAF regulates Pol II CTD phosphorylation and transcription elongation. Nucleic Acids Res. *42*, 870–881.

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.

Dutta, A., Babbarwal, V., Fu, J., Brunke-Reese, D., Libert, D.M., Willis, J., and Reese, J.C. (2015). Ccr4-Not and TFIIS Function Cooperatively To Rescue Arrested RNA Polymerase II. Mol. Cell. Biol. *35*, 1915–1925.

Dvir, A. (2002). Promoter escape by RNA polymerase II. Biochim. Biophys. Acta BBA - Gene Struct. Expr. *1577*, 208–223.

Ehara, H., Yokoyama, T., Shigematsu, H., Yokoyama, S., Shirouzu, M., and Sekine, S.-I. (2017). Structure of the complete elongation complex of RNA polymerase II with basal factors. Science *357*, 921–924.

Erie, D.A., Yager, T.D., and von Hippel, P.H. (1992). The Single-Nucleotide Addition Cycle in Transcription: a Biophysical and Biochemical Perspective. Annu. Rev. Biophys. Biomol. Struct. *21*, 379–415.

Erie, D.A., Hajiseyedjavadi, O., Young, M.C., and von Hippel, P.H. (1993). Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. Science *262*, 867–873.

Eun, C., Ortiz-Sánchez, J.M., Da, L., Wang, D., and McCammon, J.A. (2014). Molecular dynamics simulation study of conformational changes of transcription factor TFIIS during RNA polymerase II transcriptional arrest and reactivation. PloS One *9*, e97975.

Farnung, L., Vos, S.M., and Cramer, P. (2018). Structure of transcribing RNA polymerase II-nucleosome complex. BioRxiv 437574.

Fischl, H., Howe, F.S., Furger, A., and Mellor, J. (2017). Paf1 Has Distinct Roles in Transcription Elongation and Differential Transcript Fate. Mol. Cell *65*, 685-698.e8.

Fish, R.N., and Kane, C.M. (2002). Promoting elongation with transcript cleavage stimulatory factors. Biochim. Biophys. Acta *1577*, 287–307.

Fong, N., Saldi, T., Sheridan, R.M., Cortazar, M.A., and Bentley, D.L. (2017). RNA Pol II Dynamics Modulate Co-transcriptional Chromatin Modification, CTD Phosphorylation, and Transcriptional Direction. Mol. Cell *66*, 546-557.e3.

Fontrodona, L., Porta-de-la-Riva, M., Morán, T., Niu, W., Díaz, M., Aristizábal-Corrales, D., Villanueva, A., Schwartz, S., Reinke, V., and Cerón, J. (2013). RSR-2, the Caenorhabditis elegans ortholog of human spliceosomal component SRm300/SRRM2, regulates development by influencing the transcriptional machinery. PLoS Genet. *9*, e1003543.

Forde, N.R., Izhaky, D., Woodcock, G.R., Wuite, G.J.L., and Bustamante, C. (2002). Using mechanical force to probe the mechanism of pausing and arrest during continuous elongation by Escherichia coli RNA polymerase. Proc. Natl. Acad. Sci. *99*, 11682–11687.

Fouqueau, T., Zeller, M.E., Cheung, A.C., Cramer, P., and Thomm, M. (2013). The RNA polymerase trigger loop functions in all three phases of the transcription cycle. Nucleic Acids Res. *41*, 7048–7059.

Fouqueau, T., Blombach, F., Hartman, R., Cheung, A.C.M., Young, M.J., and Werner, F. (2017). The transcript cleavage factor paralogue TFS4 is a potent RNA polymerase inhibitor. Nat. Commun. *8*.

Fritsch, O., Benvenuto, G., Bowler, C., Molinier, J., and Hohn, B. (2004). The INO80 protein controls homologous recombination in Arabidopsis thaliana. Mol. Cell *16*, 479–485.

Fuchs, G., Hollander, D., Voichek, Y., Ast, G., and Oren, M. (2014). Cotranscriptional histone H2B monoubiquitylation is tightly coupled with RNA polymerase II elongation rate. Genome Res. *24*, 1572–1583.

Fukaya, T., Lim, B., and Levine, M. (2017). Rapid Rates of Pol II Elongation in the Drosophila Embryo. Curr. Biol. *27*, 1387–1391.

Gaertner, B., and Zeitlinger, J. (2014). RNA polymerase II pausing during development. Dev. Camb. Engl. *141*, 1179–1183.

Galburt, E.A., Grill, S.W., Wiedmann, A., Lubkowska, L., Choy, J., Nogales, E., Kashlev, M., and Bustamante, C. (2007). Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. Nature *446*, 820–823.

Gamba, P., and Zenkin, N. (2018). Transcription fidelity and its roles in the cell. Curr. Opin. Microbiol. *42*, 13–18.

Gamba, P., James, K., and Zenkin, N. (2017). A link between transcription fidelity and pausing in vivo. Transcription *8*, 99–105.

García-Muse, T., and Aguilera, A. (2016). Transcription-replication conflicts: how they occur and how they are resolved. Nat. Rev. Mol. Cell Biol. *17*, 553–563.

Gaudinier, A., Tang, M., and Kliebenstein, D.J. (2015). Transcriptional networks governing plant metabolism. Curr. Plant Biol. *3–4*, 56–64.

Gaykalova, D.A., Kulaeva, O.I., Volokh, O., Shaytan, A.K., Hsieh, F.-K., Kirpichnikov, M.P., Sokolova, O.S., and Studitsky, V.M. (2015). Structural analysis of nucleosomal barrier to transcription. Proc. Natl. Acad. Sci. *112*, E5787–E5795.

Gelot, C., Magdalou, I., and Lopez, B.S. (2015). Replication Stress in Mammalian Cells and Its Consequences for Mitosis. Genes *6*, 267–298.

Ghavi-Helm, Y., Michaut, M., Acker, J., Aude, J.-C., Thuriaux, P., Werner, M., and Soutourina, J. (2008). Genome-wide location analysis reveals a role of TFIIS in RNA polymerase III transcription. Genes Dev. *22*, 1934–1947.

Gilchrist, D.A., Dos Santos, G., Fargo, D.C., Xie, B., Gao, Y., Li, L., and Adelman, K. (2010). Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell *143*, 540–551.

Gilmour, D.S., and Lis, J.T. (1986). RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells. Mol. Cell. Biol. *6*, 3984–3989.

Gómez-Herreros, F., de Miguel-Jiménez, L., Millán-Zambrano, G., Peñate, X., Delgado-Ramos, L., Muñoz-Centeno, M.C., and Chávez, S. (2012). One step back before moving forward: Regulation of transcription elongation by arrest and backtracking. FEBS Lett. *586*, 2820–2825.

Gout, J.-F., Thomas, W.K., Smith, Z., Okamoto, K., and Lynch, M. (2013). Large-scale detection of in vivo transcription errors. Proc. Natl. Acad. Sci. *110*, 18584–18589.

Gout, J.-F., Li, W., Fritsch, C., Li, A., Haroon, S., Singh, L., Hua, D., Fazelinia, H., Smith, Z., Seeholzer, S., et al. (2017). The landscape of transcription errors in eukaryotic cells. Sci. Adv. *3*, e1701484.

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.

Grefen, C., Donald, N., Hashimoto, K., Kudla, J., Schumacher, K., and Blatt, M.R. (2010). A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. Plant J. *64*, 355–365.

Guglielmi, B., Soutourina, J., Esnault, C., and Werner, M. (2007). TFIIS elongation factor and Mediator act in conjunction during transcription initiation in vivo. Proc. Natl. Acad. Sci. *104*, 16062–16067.

Gutiérrez, G., Millán-Zambrano, G., Medina, D.A., Jordán-Pla, A., Pérez-Ortín, J.E., Peñate, X., and Chávez, S. (2017). Subtracting the sequence bias from partially digested MNase-seq data reveals a general contribution of TFIIS to nucleosome positioning. Epigenetics Chromatin *10*, 58.

Hajheidari, M., Farrona, S., Huettel, B., Koncz, Z., and Koncz, C. (2012). CDKF;1 and CDKD Protein Kinases Regulate Phosphorylation of Serine Residues in the C-Terminal Domain of Arabidopsis RNA Polymerase II. Plant Cell *24*, 1626–1642.

Hajheidari, M., Koncz, C., and Eick, D. (2013). Emerging roles for RNA polymerase II CTD in Arabidopsis. Trends Plant Sci. *18*, 633–643.

Harlen, K.M., and Churchman, L.S. (2017). Subgenic Pol II interactomes identify region-specific transcription elongation regulators. Mol. Syst. Biol. *13*, 900.

Harlen, K.M., Trotta, K.L., Smith, E.E., Mosaheb, M.M., Fuchs, S.M., and Churchman, L.S. (2016). Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. Cell Rep. *15*, 2147–2158.

Hartzog, G.A., and Fu, J. (2013). The Spt4-Spt5 complex: a multi-faceted regulator of transcription elongation. Biochim. Biophys. Acta *1829*, 105–115.

Hartzog, G.A., Wada, T., Handa, H., and Winston, F. (1998). Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae. Genes Dev. *12*, 357–369.

He, Y., Doyle, M.R., and Amasino, R.M. (2004). PAF1-complex-mediated histone methylation of FLOWERING LOCUS C chromatin is required for the vernalization-responsive, winter-annual habit in Arabidopsis. Genes Dev. *18*, 2774–2784.

Heidemann, M., Hintermair, C., Voß, K., and Eick, D. (2013). Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. Biochim. Biophys. Acta *1829*, 55–62.

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.

Hetzel, J., Duttke, S.H., Benner, C., and Chory, J. (2016). Nascent RNA sequencing reveals distinct features in plant transcription. Proc. Natl. Acad. Sci. U. S. A. *113*, 12316–12321.

Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M., and Bustamante, C. (2009). Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. Science *325*, 626–628.

Holstege, F.C., Fiedler, U., and Timmers, H.T. (1997). Three transitions in the RNA polymerase II transcription complex during initiation. EMBO J. *16*, 7468–7480.

Howe, K.J., Kane, C.M., and Ares, M. (2003). Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae. RNA N. Y. N *9*, 993–1006.

Imashimizu, M., Kireeva, M.L., Lubkowska, L., Gotte, D., Parks, A.R., Strathern, J.N., and Kashlev, M. (2013). Intrinsic translocation barrier as an initial step in pausing by RNA polymerase II. J. Mol. Biol. *425*, 697–712.

Imashimizu, M., Takahashi, H., Oshima, T., McIntosh, C., Bubunenko, M., Court, D.L., and Kashlev, M. (2015). Visualizing translocation dynamics and nascent transcript errors in paused RNA polymerases in vivo. Genome Biol. *16*, 98.

Irvin, J.D., Kireeva, M.L., Gotte, D.R., Shafer, B.K., Huang, I., Kashlev, M., and Strathern, J.N. (2014). A Genetic Assay for Transcription Errors Reveals Multilayer Control of RNA Polymerase II Fidelity. PLOS Genet. *10*, e1004532.

Ishibashi, T., Dangkulwanich, M., Coello, Y., Lionberger, T.A., Lubkowska, L., Ponticelli, A.S., Kashlev, M., and Bustamante, C. (2014). Transcription factors IIS and IIF enhance transcription efficiency by differentially modifying RNA polymerase pausing dynamics. Proc. Natl. Acad. Sci. U. S. A. *111*, 3419–3424.

Ito, T., Arimitsu, N., Takeuchi, M., Kawamura, N., Nagata, M., Saso, K., Akimitsu, N., Hamamoto, H., Natori, S., Miyajima, A., et al. (2006). Transcription elongation factor S-II is required for definitive hematopoiesis. Mol. Cell. Biol. *26*, 3194–3203.

Iwabuchi, K., Li, B., Bartel, P., and Fields, S. (1993). Use of the two-hybrid system to identify the domain of p53 involved in oligomerization. Oncogene *8*, 1693–1696.

Izban, M.G., and Luse, D.S. (1992). The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'----5' direction in the presence of elongation factor SII. Genes Dev. *6*, 1342–1356.

Jaehning, J.A. (2010). The Paf1 complex: platform or player in RNA polymerase II transcription? Biochim. Biophys. Acta *1799*, 379–388.

James, K., Gamba, P., Cockell, S.J., and Zenkin, N. (2017). Misincorporation by RNA polymerase is a major source of transcription pausing in vivo. Nucleic Acids Res. *45*, 1105–1113.

Jeon, C., Yoon, H., and Agarwal, K. (1994). The transcription factor TFIIS zinc ribbon dipeptide Asp-Glu is critical for stimulation of elongation and RNA cleavage by RNA polymerase II. Proc. Natl. Acad. Sci. U. S. A. *91*, 9106–9110.

Jiang, C., and Pugh, B.F. (2009). Nucleosome positioning and gene regulation: advances through genomics. Nat. Rev. Genet. *10*, 161–172.

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.

Joubès, J., and Chevalier, C. (2000). Endoreduplication in higher plants. Plant Mol. Biol. *43*, 735–745.

Juntawong, P., Girke, T., Bazin, J., and Bailey-Serres, J. (2014). Translational dynamics revealed by genome-wide profiling of ribosome footprints in Arabidopsis. Proc. Natl. Acad. Sci. *111*, E203–E212.

Karakasili, E., Burkert-Kautzsch, C., Kieser, A., and Sträßer, K. (2014). Degradation of DNA damage-independently stalled RNA polymerase II is independent of the E3 ligase Elc1. Nucleic Acids Res. *42*, 10503–10515.

Kettenberger, H., Armache, K.-J., and Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. Cell *114*, 347–357.

Kettenberger, H., Armache, K.-J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Mol. Cell *16*, 955–965.

Khraiwesh, B., Zhu, J.-K., and Zhu, J. (2012). Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. Biochim. Biophys. Acta *1819*, 137–148.

Kim, B., Nesvizhskii, A.I., Rani, P.G., Hahn, S., Aebersold, R., and Ranish, J.A. (2007). The transcription elongation factor TFIIS is a component of RNA polymerase II preinitiation complexes. Proc. Natl. Acad. Sci. U. S. A. *104*, 16068–16073.

Kim, J., Guermah, M., McGinty, R.K., Lee, J.-S., Tang, Z., Milne, T.A., Shilatifard, A., Muir, T.W., and Roeder, R.G. (2009). RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. Cell *137*, 459–471.

Kim, J., Guermah, M., and Roeder, R.G. (2010). The Human PAF1 Complex Acts in Chromatin Transcription Elongation Both Independently and Cooperatively with SII/TFIIS. Cell *140*, 491–503.

Kireeva, M.L., Hancock, B., Cremona, G.H., Walter, W., Studitsky, V.M., and Kashlev, M. (2005). Nature of the Nucleosomal Barrier to RNA Polymerase II. Mol. Cell *18*, 97–108.

Klopf, E., Moes, M., Amman, F., Zimmermann, B., Pelchrzim, F. von, Wagner, C., and Schroeder, R. (2018). Nascent RNA signaling to yeast RNA Pol II during transcription elongation. PLOS ONE *13*, e0194438.

Köllen, K., Dietz, L., Bies-Etheve, N., Lagrange, T., Grasser, M., and Grasser, K.D. (2015). The zinc-finger protein SPT4 interacts with SPT5L/KTF1 and modulates transcriptional silencing in Arabidopsis. FEBS Lett. *589*, 3254–3257.

Komarnitsky, P., Cho, E.-J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev. *14*, 2452–2460.

Kowalik, K.M., Shimada, Y., Flury, V., Stadler, M.B., Batki, J., and Bühler, M. (2015). The Paf1 complex represses small-RNA-mediated epigenetic gene silencing. Nature *520*, 248–252.

Koyama, H., Ito, T., Nakanishi, T., and Sekimizu, K. (2007). Stimulation of RNA polymerase II transcript cleavage activity contributes to maintain transcriptional fidelity in yeast. Genes Cells Devoted Mol. Cell. Mech. *12*, 547–559.

Krajewski, W.A., Li, J., and Dou, Y. (2018). Effects of histone H2B ubiquitylation on the nucleosome structure and dynamics. Nucleic Acids Res. *46*, 7631–7642.

Krogan, N.J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Ryan, O.W., Golshani, A., Johnston, M., et al. (2003). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol. Cell *11*, 721–729.

Kudo, T., Sasaki, Y., Terashima, S., Matsuda-Imai, N., Takano, T., Saito, M., Kanno, M., Ozaki, S., Suwabe, K., Suzuki, G., et al. (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.

Kurepa, J., Karangwa, C., Duke, L.S., and Smalle, J.A. (2010). Arabidopsis sensitivity to protein synthesis inhibitors depends on 26S proteasome activity. Plant Cell Rep. *29*, 249–259.

Kwak, H., and Lis, J.T. (2013). Control of transcriptional elongation. Annu. Rev. Genet. *47*, 483–508.

Lademann, C.A., Renkawitz, J., Pfander, B., and Jentsch, S. (2017). The INO80 Complex Removes H2A.Z to Promote Presynaptic Filament Formation during Homologous Recombination. Cell Rep. *19*, 1294–1303.

Landick, R. (2009). Functional divergence in the growing family of RNA polymerases. Struct. Lond. Engl. 1993 *17*, 323–325.

Lange, U., and Hausner, W. (2004). Transcriptional fidelity and proofreading in Archaea and implications for the mechanism of TFS-induced RNA cleavage. Mol. Microbiol. *52*, 1133–1143.

Laptenko, O., Kim, S.-S., Lee, J., Starodubtseva, M., Cava, F., Berenguer, J., Kong, X.-P., and Borukhov, S. (2006). pH-dependent conformational switch activates the inhibitor of transcription elongation. EMBO J. *25*, 2131–2141.

Larschan, E., Bishop, E.P., Kharchenko, P.V., Core, L.J., Lis, J.T., Park, P.J., and Kuroda, M.I. (2011). X chromosome dosage compensation via enhanced transcriptional elongation in *Drosophila*. Nature *471*, 115–118.

Lavigne, M.D., Konstantopoulos, D., Ntakou-Zamplara, K.Z., Liakos, A., and Fousteri, M. (2017). Global unleashing of transcription elongation waves in response to genotoxic stress restricts somatic mutation rate. Nat. Commun. *8*, 2076.

Lee, C., Li, X., Hechmer, A., Eisen, M., Biggin, M.D., Venters, B.J., Jiang, C., Li, J., Pugh, B.F., and Gilmour, D.S. (2008). NELF and GAGA Factor Are Linked to Promoter-Proximal Pausing at Many Genes in Drosophila. Mol. Cell. Biol. *28*, 3290–3300.

Lee, Y., Kim, M., Han, J., Yeom, K.-H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. EMBO J. *23*, 4051–4060.

Lermontova, I., Schubert, V., Fuchs, J., Klatte, S., Macas, J., and Schubert, I. (2006). Loading of Arabidopsis Centromeric Histone CENH3 Occurs Mainly during G2 and Requires the Presence of the Histone Fold Domain. Plant Cell *18*, 2443–2451.

Li, G., Liu, S., Wang, J., He, J., Huang, H., Zhang, Y., and Xu, L. (2014). ISWI proteins participate in the genome-wide nucleosome distribution in Arabidopsis. Plant J. Cell Mol. Biol. *78*, 706–714.

Li, H., Zhang, Z., Wang, B., Zhang, J., Zhao, Y., and Jin, Y. (2007). Wwp2-mediated ubiquitination of the RNA polymerase II large subunit in mouse embryonic pluripotent stem cells. Mol. Cell. Biol. *27*, 5296–5305.

Lidschreiber, M., Leike, K., and Cramer, P. (2013). Cap Completion and C-Terminal Repeat Domain Kinase Recruitment Underlie the Initiation-Elongation Transition of RNA Polymerase II. Mol. Cell. Biol. *33*, 3805–3816.

Lindstrom, D.L., and Hartzog, G.A. (2001). Genetic interactions of Spt4-Spt5 and TFIIS with the RNA polymerase II CTD and CTD modifying enzymes in Saccharomyces cerevisiae. Genetics *159*, 487–497.

Ling, Y., Smith, A.J., and Morgan, G.T. (2006). A sequence motif conserved in diverse nuclear proteins identifies a protein interaction domain utilised for nuclear targeting by human TFIIS. Nucleic Acids Res. *34*, 2219–2229.

Lisica, A., Engel, C., Jahnel, M., Roldán, É., Galburt, E.A., Cramer, P., and Grill, S.W. (2016). Mechanisms of backtrack recovery by RNA polymerases I and II. Proc. Natl. Acad. Sci. U. S. A. *113*, 2946–2951.

Liu, M.-J., Seddon, A.E., Tsai, Z.T.-Y., Major, I.T., Floer, M., Howe, G.A., and Shiu, S.-H. (2015). Determinants of nucleosome positioning and their influence on plant gene expression. Genome Res. *25*, 1182–1195.

Liu, P., Kenney, J.M., Stiller, J.W., and Greenleaf, A.L. (2010). Genetic organization, length conservation, and evolution of RNA polymerase II carboxyl-terminal domain. Mol. Biol. Evol. *27*, 2628–2641.

Liu, Y., Kung, C., Fishburn, J., Ansari, A.Z., Shokat, K.M., and Hahn, S. (2004). Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. Mol. Cell. Biol. *24*, 1721–1735.

Liu, Y., Warfield, L., Zhang, C., Luo, J., Allen, J., Lang, W.H., Ranish, J., Shokat, K.M., and Hahn, S. (2009). Phosphorylation of the Transcription Elongation Factor Spt5 by Yeast Bur1 Kinase Stimulates Recruitment of the PAF Complex. Mol. Cell. Biol. *29*, 4852–4863.

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. Cell Mol. Biol. *61*, 686–697.

Lozano, R., Booth, G.T., Omar, B.Y., Li, B., Buckler, E.S., Lis, J.T., Jannink, J.-L., and Carpio, D.P.D. (2018). RNA polymerase mapping in plants identifies enhancers enriched in causal variants. BioRxiv 376640.

Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991). The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. Proc. Natl. Acad. Sci. U. S. A. *88*, 10004–10008.

Lu, X., Zhu, X., Li, Y., Liu, M., Yu, B., Wang, Y., Rao, M., Yang, H., Zhou, K., Wang, Y., et al. (2016). Multiple P-TEFbs cooperatively regulate the release of promoter-proximally paused RNA polymerase II. Nucleic Acids Res. *44*, 6853–6867.

Luse, D.S., and Studitsky, V.M. (2011). The mechanism of nucleosome traversal by RNA polymerase II. RNA Biol. *8*, 581–585.

Luse, D.S., Spangler, L.C., and Újvári, A. (2011). Efficient and rapid nucleosome traversal by RNA polymerase II depends on a combination of transcript elongation factors. J. Biol. Chem. *286*, 6040–6048.

Mao, P., Meas, R., Dorgan, K.M., and Smerdon, M.J. (2014). UV damage-induced RNA polymerase II stalling stimulates H2B deubiquitylation. Proc. Natl. Acad. Sci. *111*, 12811–12816.

Marazzi, I., Ho, J.S.Y., Kim, J., Manicassamy, B., Dewell, S., Albrecht, R.A., Seibert, C.W., Schaefer, U., Jeffrey, K.L., Prinjha, R.K., et al. (2012). Suppression of the antiviral response by an influenza histone mimic. Nature *483*, 428–433.

Martinez-Rucobo, F.W., and Cramer, P. (2013). Structural basis of transcription elongation. Biochim. Biophys. Acta *1829*, 9–19.

Marton, H.A., and Desiderio, S. (2008). The Paf1 complex promotes displacement of histones upon rapid induction of transcription by RNA polymerase II. BMC Mol. Biol. *9*, 4.

Mavrich, T.N., Jiang, C., Ioshikhes, I.P., Li, X., Venters, B.J., Zanton, S.J., Tomsho, L.P., Qi, J., Glaser, R.L., Schuster, S.C., et al. (2008). Nucleosome organization in the Drosophila genome. Nature *453*, 358–362.

Mayekar, M.K., Gardner, R.G., and Arndt, K.M. (2013). The recruitment of the Saccharomyces cerevisiae Paf1 complex to active genes requires a domain of Rtf1 that directly interacts with the Spt4-Spt5 complex. Mol. Cell. Biol. *33*, 3259–3273.

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.

Mayer, A., Landry, H.M., and Churchman, L.S. (2017). Pause & Go: from the discovery of RNA polymerase pausing to its functional implications. Curr. Opin. Cell Biol. *46*, 72–80.

McDonald, S.M., Close, D., Xin, H., Formosa, T., and Hill, C.P. (2010). Structure and biological importance of the Spn1-Spt6 interaction, and its regulatory role in nucleosome binding. Mol. Cell *40*, 725–735.

McIvor, E.I., Polak, U., and Napierala, M. (2010). New insights into repeat instability. RNA Biol. *7*, 551–558.

McKay, B.C., and Cabrita, M.A. (2013). Arresting transcription and sentencing the cell: the consequences of blocked transcription. Mech. Ageing Dev. *134*, 243–252.

Mejia, Y.X., Nudler, E., and Bustamante, C. (2015). Trigger loop folding determines transcription rate of Escherichia coli's RNA polymerase. Proc. Natl. Acad. Sci. *112*, 743–748.

Min, I.M., Waterfall, J.J., Core, L.J., Munroe, R.J., Schimenti, J., and Lis, J.T. (2011). Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. Genes Dev. *25*, 742–754.

Miropolskaya, N., Esyunina, D., and Kulbachinskiy, A. (2017). Conserved functions of the trigger loop and Gre factors in RNA cleavage by bacterial RNA polymerases. J. Biol. Chem. *292*, 6744–6752.

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

Morris, D.P., and Greenleaf, A.L. (2000). The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. J. Biol. Chem. *275*, 39935–39943.

Mortensen, S.A. (2012). "Molecular analysis of the transcript elongation factor TFIIS and the histone chaperone FACT in Arabidopsis thaliana". PhD thesis. University of Regensburg.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. Nat. Genet. *39*, 1507–1511.

Narita, T., Yamaguchi, Y., Yano, K., Sugimoto, S., Chanarat, S., Wada, T., Kim, D., Hasegawa, J., Omori, M., Inukai, N., et al. (2003). Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex. Mol. Cell. Biol. *23*, 1863–1873.

Narsai, R., Howell, K.A., Millar, A.H., O'Toole, N., Small, I., and Whelan, J. (2007). Genome-wide analysis of mRNA decay rates and their determinants in Arabidopsis thaliana. Plant Cell *19*, 3418–3436.

Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y., and Adelman, K. (2010). Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. Science *327*, 335–338.

Nelissen, H., De Groeve, S., Fleury, D., Neyt, P., Bruno, L., Bitonti, M.B., Vandenbussche, F., Van der Straeten, D., Yamaguchi, T., Tsukaya, H., et al. (2010). Plant Elongator regulates auxin-related genes during RNA polymerase II transcription elongation. Proc. Natl. Acad. Sci. U. S. A. *107*, 1678–1683.

Ng, H.H., Dole, S., and Struhl, K. (2003). The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. J. Biol. Chem. *278*, 33625–33628.

Nguyen, V.T., Giannoni, F., Dubois, M.F., Seo, S.J., Vigneron, M., Kédinger, C., and Bensaude, O. (1996). In vivo degradation of RNA polymerase II largest subunit triggered by alpha-amanitin. Nucleic Acids Res. *24*, 2924–2929.

Nock, A., Ascano, J.M., Barrero, M.J., and Malik, S. (2012). Mediator-regulated transcription through the +1 nucleosome. Mol. Cell *48*, 837–848.

Nojima, T., Gomes, T., Grosso, A.R.F., Kimura, H., Dye, M.J., Dhir, S., Carmo-Fonseca, M., and Proudfoot, N.J. (2015). Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. Cell *161*, 526–540.

Nudler, E. (2009). RNA polymerase active center: the molecular engine of transcription. Annu. Rev. Biochem. *78*, 335–361.

Nudler, E. (2012). RNA Polymerase Backtracking in Gene Regulation and Genome Instability. Cell *149*.

Oh, S., Zhang, H., Ludwig, P., and Nocker, S. van (2004). A Mechanism Related to the Yeast Transcriptional Regulator Paf1c Is Required for Expression of the Arabidopsis FLC/MAF MADS Box Gene Family. Plant Cell *16*, 2940–2953.

Oh, S., Park, S., and Nocker, S. van (2008). Genic and Global Functions for Paf1C in Chromatin Modification and Gene Expression in Arabidopsis. PLOS Genet. *4*, e1000077.

Olmsted, V.K., Awrey, D.E., Koth, C., Shan, X., Morin, P.E., Kazanis, S., Edwards, A.M., and Arrowsmith, C.H. (1998). Yeast Transcript Elongation Factor (TFIIS), Structure and Function I: NMR STRUCTURAL ANALYSIS OF THE MINIMAL TRANSCRIPTIONALLY ACTIVE REGION. J. Biol. Chem. *273*, 22589–22594.

Opalka, N., Chlenov, M., Chacon, P., Rice, W.J., Wriggers, W., and Darst, S.A. (2003). Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. Cell *114*, 335–345.

Orel, N., Kyryk, A., and Puchta, H. (2003). Different pathways of homologous recombination are used for the repair of double-strand breaks within tandemly arranged sequences in the plant genome. Plant J. Cell Mol. Biol. *35*, 604–612.

Otero, G., Fellows, J., Li, Y., de Bizemont, T., Dirac, A.M., Gustafsson, C.M., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J.Q. (1999). Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation. Mol. Cell *3*, 109–118.

Pal, M., and Luse, D.S. (2003). The initiation-elongation transition: lateral mobility of RNA in RNA polymerase II complexes is greatly reduced at +8/+9 and absent by +23. Proc. Natl. Acad. Sci. U. S. A. *100*, 5700–5705.

Pal, M., Ponticelli, A.S., and Luse, D.S. (2005). The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. Mol. Cell *19*, 101–110.

Palangat, M., Renner, D.B., Price, D.H., and Landick, R. (2005). A negative elongation factor for human RNA polymerase II inhibits the anti-arrest transcript-cleavage factor TFIIS. Proc. Natl. Acad. Sci. U. S. A. *102*, 15036–15041.

Park, S., Oh, S., Ek-Ramos, J., and Nocker, S. van (2010). PLANT HOMOLOGOUS TO PARAFIBROMIN Is a Component of the PAF1 Complex and Assists in Regulating Expression of Genes within H3K27ME3-Enriched Chromatin. Plant Physiol. *153*, 821–831.

Parsa, J.-Y., Boudoukha, S., Burke, J., Homer, C., and Madhani, H.D. (2018). Polymerase pausing induced by sequence-specific RNA-binding protein drives heterochromatin assembly. Genes Dev. *32*, 953–964.

Paule, M.R., and White, R.J. (2000). Transcription by RNA polymerases I and III. Nucleic Acids Res. *28*, 1283–1298.

Perales, R., and Bentley, D. (2009). "Cotranscriptionality": The Transcription Elongation Complex as a Nexus for Nuclear Transactions. Mol. Cell *36*, 178–191.

Peterlin, B.M., and Price, D.H. (2006). Controlling the elongation phase of transcription with P-TEFb. Mol. Cell *23*, 297–305.

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

Pfab, A., Bruckmann, A., Nazet, J., Merkl, R., and Grasser, K.D. (2018a). 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. J. Mol. Biol. *430*, 1479–1494.

Pfab, A., Grønlund, J.T., Holzinger, P., Längst, G., and Grasser, K.D. (2018b). The Arabidopsis Histone Chaperone FACT: Role of the HMG-Box Domain of SSRP1. J. Mol. Biol. *430*, 2747–2759.

Pinskaya, M., Ghavi-Helm, Y., Mariotte-Labarre, S., Morillon, A., Soutourina, J., and Werner, M. (2014). PHD and TFIIS-Like Domains of the Bye1 Transcription Factor Determine Its Multivalent Genomic Distribution. PLoS ONE *9*.

Pokholok, D.K., Hannett, N.M., and Young, R.A. (2002). Exchange of RNA Polymerase II Initiation and Elongation Factors during Gene Expression In Vivo. Mol. Cell *9*, 799–809.

Poli, J., Gerhold, C.-B., Tosi, A., Hustedt, N., Seeber, A., Sack, R., Herzog, F., Pasero, P., Shimada, K., Hopfner, K.-P., et al. (2016). Mec1, INO80, and the PAF1 complex cooperate to limit transcription-replication conflicts through RNAPII removal during replication stress. Genes Dev. *30*, 337–354.

Pontier, D., Yahubyan, G., Vega, D., Bulski, A., Saez-Vasquez, J., Hakimi, M.-A., Lerbs-Mache, S., Colot, V., and Lagrange, T. (2005). Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. Genes Dev. *19*, 2030–2040.

Prather, D.M., Larschan, E., and Winston, F. (2005). Evidence that the elongation factor TFIIS plays a role in transcription initiation at GAL1 in Saccharomyces cerevisiae. Mol. Cell. Biol. *25*, 2650–2659.

Pruneski, J.A., Hainer, S.J., Petrov, K.O., and Martens, J.A. (2011). The Paf1 complex represses SER3 transcription in Saccharomyces cerevisiae by facilitating intergenic transcription-dependent nucleosome occupancy of the SER3 promoter. Eukaryot. Cell *10*, 1283–1294.

Qian, X., Gozani, S.N., Yoon, H., Jeon, C., Agarwal, K., and Weiss, M.A. (1993). Novel zinc finger motif in the basal transcriptional machinery: Three-dimensional NMR studies of the nucleic acid binding domain of transcriptional elongation factor TFIIS. Biochemistry *32*, 9944–9959.

Qiu, H., Hu, C., Gaur, N.A., and Hinnebusch, A.G. (2012). Pol II CTD kinases Bur1 and Kin28 promote Spt5 CTR-independent recruitment of Paf1 complex. EMBO J. *31*, 3494–3505.

Ramanathan, Y., Rajpara, S.M., Reza, S.M., Lees, E., Shuman, S., Mathews, M.B., and Pe'ery, T. (2001). Three RNA polymerase II carboxyl-terminal domain kinases display distinct substrate preferences. J. Biol. Chem. *276*, 10913–10920.

Rapsomaniki, M.A., Kotsantis, P., Symeonidou, I.-E., Giakoumakis, N.-N., Taraviras, S., and Lygerou, Z. (2012). easyFRAP: an interactive, easy-to-use tool for qualitative and quantitative analysis of FRAP data. Bioinforma. Oxf. Engl. *28*, 1800–1801.

Ratner, J.N., Balasubramanian, B., Corden, J., Warren, S.L., and Bregman, D.B. (1998). Ultraviolet Radiation-induced Ubiquitination and Proteasomal Degradation of the Large Subunit of RNA Polymerase II IMPLICATIONS FOR TRANSCRIPTION-COUPLED DNA REPAIR. J. Biol. Chem. *273*, 5184–5189.

Reines, D., Chamberlin, M.J., and Kane, C.M. (1989). Transcription elongation factor SII (TFIIS) enables RNA polymerase II to elongate through a block to transcription in a human gene in vitro. J. Biol. Chem. *264*, 10799–10809.

Reines, D., Conaway, R.C., and Conaway, J.W. (1999). Mechanism and regulation of transcriptional elongation by RNA polymerase II. Curr. Opin. Cell Biol. *11*, 342–346.

Rhee, H.S., and Pugh, B.F. (2011). Comprehensive Genome-wide Protein-DNA Interactions Detected at Single-Nucleotide Resolution. Cell *147*, 1408–1419.

Roghanian, M., Yuzenkova, Y., and Zenkin, N. (2011). Controlled interplay between trigger loop and Gre factor in the RNA polymerase active centre. Nucleic Acids Res. *39*, 4352–4359.

Rondón, A.G., Gallardo, M., García-Rubio, M., and Aguilera, A. (2004). Molecular evidence indicating that the yeast PAF complex is required for transcription elongation. EMBO Rep. *5*, 47–53.

Ruan, W., Lehmann, E., Thomm, M., Kostrewa, D., and Cramer, P. (2011). Evolution of Two Modes of Intrinsic RNA Polymerase Transcript Cleavage. J. Biol. Chem. *286*, 18701–18707.

Sadeghi, L., Prasad, P., Ekwall, K., Cohen, A., and Svensson, J.P. (2015). The Paf1 complex factors Leo1 and Paf1 promote local histone turnover to modulate chromatin states in fission yeast. EMBO Rep. *16*, 1673–1687.

Santisteban, M.S., Hang, M., and Smith, M.M. (2011). Histone variant H2A.Z and RNA polymerase II transcription elongation. Mol. Cell. Biol. *31*, 1848–1860.

Saunders, A., Core, L.J., and Lis, J.T. (2006). Breaking barriers to transcription elongation. Nat. Rev. Mol. Cell Biol. *7*, 557–567.

Scheibe, R., Backhausen, J.E., Emmerlich, V., and Holtgrefe, S. (2005). Strategies to maintain redox homeostasis during photosynthesis under changing conditions. J. Exp. Bot. *56*, 1481–1489.

Schlücking, K., Edel, K.H., Köster, P., Drerup, M.M., Eckert, C., Steinhorst, L., Waadt, R., Batistic, O., and Kudla, J. (2013). A new β-estradiol-inducible vector set that facilitates easy construction and efficient expression of transgenes reveals CBL3-dependent cytoplasm to tonoplast translocation of CIPK5. Mol. Plant *6*, 1814–1829.

Schuermann, D., Fritsch, O., Lucht, J.M., and Hohn, B. (2009). Replication Stress Leads to Genome Instabilities in Arabidopsis DNA Polymerase δ Mutants. Plant Cell *21*, 2700–2714.

Schweikhard, V., Meng, C., Murakami, K., Kaplan, C.D., Kornberg, R.D., and Block, S.M. (2014). Transcription factors TFIIF and TFIIS promote transcript elongation by RNA polymerase II by synergistic and independent mechanisms. Proc. Natl. Acad. Sci. U. S. A. *111*, 6642–6647.

Selth, L.A., Sigurdsson, S., and Svejstrup, J.Q. (2010). Transcript Elongation by RNA Polymerase II. Annu. Rev. Biochem. *79*, 271–293.

Serizawa, H., Conaway, J.W., and Conaway, R.C. (1993). Phosphorylation of C-terminal domain of RNA polymerase II is not required in basal transcription. Nature *363*, 371–374.

Shandilya, J., and Roberts, S.G.E. (2012). The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. Biochim. Biophys. Acta *1819*, 391–400.

Sheldon, K.E., Mauger, D.M., and Arndt, K.M. (2005). A Requirement for the Saccharomyces cerevisiae Paf1 Complex in snoRNA 3'End Formation. Mol. Cell *20*, 225–236.

Sheridan, R.M., Fong, N., D'Alessandro, A., and Bentley, D.L. (2019). Widespread Backtracking by RNA Pol II Is a Major Effector of Gene Activation, 5' Pause Release, Termination, and Transcription Elongation Rate. Mol. Cell *73*, 107-118.e4.

Shi, X., Finkelstein, A., Wolf, A.J., Wade, P.A., Burton, Z.F., and Jaehning, J.A. (1996). Paf1p, an RNA polymerase II-associated factor in Saccharomyces cerevisiae, may have both positive and negative roles in transcription. Mol. Cell. Biol. *16*, 669–676.

Sigurdsson, S., Dirac-Svejstrup, A.B., and Svejstrup, J.Q. (2010). Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. Mol. Cell *38*, 202–210.

Simic, R., Lindstrom, D.L., Tran, H.G., Roinick, K.L., Costa, P.J., Johnson, A.D., Hartzog, G.A., and Arndt, K.M. (2003). Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. EMBO J. *22*, 1846–1856.

Sims, R.J., Belotserkovskaya, R., and Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. Genes Dev. *18*, 2437–2468.

Sims, R.J., Rojas, L.A., Beck, D.B., Bonasio, R., Schüller, R., Drury, W.J., Eick, D., and Reinberg, D. (2011). The C-terminal domain of RNA polymerase II is modified by site-specific methylation. Science *332*, 99–103.

Snapp, E. (2005). Design and Use of Fluorescent Fusion Proteins in Cell Biology. Curr. Protoc. Cell Biol. Editor. Board Juan Bonifacino Al *CHAPTER*, Unit-21.4.

Somesh, B.P., Sigurdsson, S., Saeki, H., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J.Q. (2007). Communication between Distant Sites in RNA Polymerase II through Ubiquitylation Factors and the Polymerase CTD. Cell *129*, 57–68.

Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003). Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. EMBO J. *22*, 2234–2244.

Steurer, B., and Marteijn, J.A. (2017). Traveling Rocky Roads: The Consequences of Transcription-Blocking DNA Lesions on RNA Polymerase II. J. Mol. Biol. *429*, 3146–3155.

Steurer, B., Janssens, R.C., Geverts, B., Geijer, M.E., Wienholz, F., Theil, A.F., Chang, J., Dealy, S., Pothof, J., van Cappellen, W.A., et al. (2018). Live-cell analysis of endogenous GFP-RPB1 uncovers rapid turnover of initiating and promoter-paused RNA Polymerase II. Proc. Natl. Acad. Sci. U. S. A. *115*, E4368–E4376.

Surpin, M., Larkin, R.M., and Chory, J. (2002). Signal transduction between the chloroplast and the nucleus. Plant Cell *14 Suppl*, S327-338.

Svejstrup, J.Q. (2003). Rescue of arrested RNA polymerase II complexes. J. Cell Sci. *116*, 447–451.

Svejstrup, J.Q. (2007). Contending with transcriptional arrest during RNAPII transcript elongation. Trends Biochem. Sci. *32*, 165–171.

Svetlov, V., and Nudler, E. (2013). Basic mechanism of transcription by RNA polymerase II. Biochim. Biophys. Acta *1829*, 20–28.

Swanson, M.S., and Winston, F. (1992). SPT4, SPT5 and SPT6 interactions: effects on transcription and viability in Saccharomyces cerevisiae. Genetics *132*, 325–336.

Sydow, J.F., Brueckner, F., Cheung, A.C.M., Damsma, G.E., Dengl, S., Lehmann, E., Vassylyev, D., and Cramer, P. (2009). Structural Basis of Transcription: Mismatch-Specific Fidelity Mechanisms and Paused RNA Polymerase II with Frayed RNA. Mol. Cell *34*, 710–721.

Szeberenyi, J. (2006). The effect of α-amanitin on RNA polymerase II ubiquitination*. Biochem. Mol. Biol. Educ. *34*, 386–389.

Tehranchi, A.K., Blankschien, M.D., Zhang, Y., Halliday, J.A., Srivatsan, A., Peng, J., Herman, C., and Wang, J.D. (2010). The transcription factor DksA prevents conflicts between DNA replication and transcription machinery. Cell *141*, 595–605.

Thomas, M.J., Platas, A.A., and Hawley, D.K. (1998). Transcriptional fidelity and proofreading by RNA polymerase II. Cell *93*, 627–637.

Timm, S., Florian, A., Wittmiß, M., Jahnke, K., Hagemann, M., Fernie, A.R., and Bauwe, H. (2013). Serine acts as a metabolic signal for the transcriptional control of photorespiration-related genes in Arabidopsis. Plant Physiol. *162*, 379–389.

Tomson, B.N., and Arndt, K.M. (2013). The many roles of the conserved eukaryotic Paf1 complex in regulating transcription, histone modifications, and disease states. Biochim. Biophys. Acta *1829*, 116–126.

Tramier, M., Zahid, M., Mevel, J.-C., Masse, M.-J., and Coppey-Moisan, M. (2006). Sensitivity of CFP/YFP and GFP/mCherry pairs to donor photobleaching on FRET determination by fluorescence lifetime imaging microscopy in living cells. Microsc. Res. Tech. *69*, 933–939.

Traverse, C.C., and Ochman, H. (2016). Conserved rates and patterns of transcription errors across bacterial growth states and lifestyles. Proc. Natl. Acad. Sci. U. S. A. *113*, 3311–3316.

Tsugama, D., Liu, S., and Takano, T. (2011). A rapid chemical method for lysing Arabidopsis cells for protein analysis. Plant Methods *7*, 22.

Ubeda-Tomás, S., Federici, F., Casimiro, I., Beemster, G.T.S., Bhalerao, R., Swarup, R., Doerner, P., Haseloff, J., and Bennett, M.J. (2009). Gibberellin signaling in the endodermis controls Arabidopsis root meristem size. Curr. Biol. CB *19*, 1194–1199.

Ullah, Z., Lee, C.Y., Lilly, M.A., and DePamphilis, M.L. (2009). Developmentally programmed endoreduplication in animals. Cell Cycle Georget. Tex *8*, 1501–1509.

Uptain, S.M., Kane, C.M., and Chamberlin, M.J. (1997). Basic mechanisms of transcript elongation and its regulation. Annu. Rev. Biochem. *66*, 117–172.

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

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

Van Lijsebettens, M., and Grasser, K.D. (2014). Transcript elongation factors: shaping transcriptomes after transcript initiation. Trends Plant Sci. *19*, 717–726.

Van Oss, S.B., Cucinotta, C.E., and Arndt, K.M. (2017). Emerging Insights into the Roles of the Paf1 Complex in Gene Regulation. Trends Biochem. Sci. *42*, 788–798.

Vannini, A., and Cramer, P. (2012). Conservation between the RNA polymerase I, II, and III transcription initiation machineries. Mol. Cell *45*, 439–446.

Vergara, Z., Sequeira-Mendes, J., Morata, J., Peiró, R., Hénaff, E., Costas, C., Casacuberta, J.M., and Gutierrez, C. (2017). Retrotransposons are specified as DNA replication origins in the gene-poor regions of Arabidopsis heterochromatin. Nucleic Acids Res. *45*, 8358–8368.

Vermulst, M., Denney, A.S., Lang, M.J., Hung, C.-W., Moore, S., Mosely, M.A., Thompson, J.W., Madden, V., Gauer, J., Wolfe, K.J., et al. (2015). Transcription Errors Induce Proteotoxic Stress and Shorten Cellular Lifespan. Nat. Commun. *6*, 8065.

Verrier, L., Taglini, F., Barrales, R.R., Webb, S., Urano, T., Braun, S., and Bayne, E.H. (2015). Global regulation of heterochromatin spreading by Leo1. Open Biol. *5*, 150045.

Vinayachandran, V., Reja, R., Rossi, M.J., Park, B., Rieber, L., Mittal, C., Mahony, S., and Pugh, B.F. (2018). Widespread and precise reprogramming of yeast protein-genome interactions in response to heat shock. Genome Res.

Vos, S.M., Farnung, L., Urlaub, H., and Cramer, P. (2018a). Structure of paused transcription complex Pol II-DSIF-NELF. Nature *560*, 601–606.

Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H., and Cramer, P. (2018b). Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature *560*, 607–612.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F., et al. (1998). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev. *12*, 343–356.

Wade, P.A., Werel, W., Fentzke, R.C., Thompson, N.E., Leykam, J.F., Burgess, R.R., Jaehning, J.A., and Burton, Z.F. (1996). A Novel Collection of Accessory Factors Associated with Yeast RNA Polymerase II. Protein Expr. Purif. *8*, 85–90.

Walmacq, C., Kireeva, M.L., Irvin, J., Nedialkov, Y., Lubkowska, L., Malagon, F., Strathern, J.N., and Kashlev, M. (2009). Rpb9 subunit controls transcription fidelity by delaying NTP sequestration in RNA polymerase II. J. Biol. Chem. *284*, 19601–19612.

Wang, X. (2014). Identification of Marker Genes for Cancer Based on Microarrays Using a Computational Biology Approach. Curr. Bioinforma. *9*, 140–146.

Wang, D., Bushnell, D.A., Huang, X., Westover, K.D., Levitt, M., and Kornberg, R.D. (2009). Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. Science *324*, 1203–1206.

Wang, P., Hendron, R.-W., and Kelly, S. (2017). Transcriptional control of photosynthetic capacity: conservation and divergence from Arabidopsis to rice. New Phytol. *216*, 32–45.

Weber, C.M., Ramachandran, S., and Henikoff, S. (2014). Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. Mol. Cell *53*, 819–830.

Weidtkamp-Peters, S., and Stahl, Y. (2017). The Use of FRET/FLIM to Study Proteins Interacting with Plant Receptor Kinases. Methods Mol. Biol. Clifton NJ *1621*, 163–175.

Weilbaecher, R.G., Awrey, D.E., Edwards, A.M., and Kane, C.M. (2003). Intrinsic transcript cleavage in yeast RNA polymerase II elongation complexes. J. Biol. Chem. *278*, 24189–24199.

Weiner, A., Hughes, A., Yassour, M., Rando, O.J., and Friedman, N. (2010). High-resolution nucleosome mapping reveals transcription-dependent promoter packaging. Genome Res. *20*, 90–100.

Werner, F., and Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. Nat. Rev. Microbiol. *9*, 85–98.

Wery, M., Shematorova, E., Van Driessche, B., Vandenhaute, J., Thuriaux, P., and Van Mullem, V. (2004). Members of the SAGA and Mediator complexes are partners of the transcription elongation factor TFIIS. EMBO J. *23*, 4232–4242.

West, M.L., and Corden, J.L. (1995). Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. Genetics *140*, 1223–1233.

Williams, L.A., and Kane, C.M. (1996). Isolation and characterization of the Schizosaccharomyces pombe gene encoding transcript elongation factor TFIIS. Yeast Chichester Engl. *12*, 227–236.

Wind, M., and Reines, D. (2000). Transcription elongation factor SII. BioEssays News Rev. Mol. Cell. Dev. Biol. *22*, 327–336.

Wood, A., Schneider, J., Dover, J., Johnston, M., and Shilatifard, A. (2003). The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. J. Biol. Chem. *278*, 34739–34742.

Wu, J., and Xu, W. (2012). Histone H3R17me2a mark recruits human RNA polymerase-associated factor 1 complex to activate transcription. Proc. Natl. Acad. Sci. U. S. A. *109*, 5675–5680.

Wu, S., Shi, Y., Mulligan, P., Gay, F., Landry, J., Liu, H., Lu, J., Qi, H.H., Wang, W., Nickoloff, J.A., et al. (2007). A YY1–INO80 complex regulates genomic stability through homologous recombination–based repair. Nat. Struct. Mol. Biol. *14*, 1165–1172.

Xiao, T., Kao, C.-F., Krogan, N.J., Sun, Z.-W., Greenblatt, J.F., Osley, M.A., and Strahl, B.D. (2005). Histone H2B ubiquitylation is associated with elongating RNA polymerase II. Mol. Cell. Biol. *25*, 637–651.

Xu, Y., Bernecky, C., Lee, C.-T., Maier, K.C., Schwalb, B., Tegunov, D., Plitzko, J.M., Urlaub, H., and Cramer, P. (2017). Architecture of the RNA polymerase II-Paf1C-TFIIS transcription elongation complex. Nat. Commun. *8*, 15741.

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. (2006). P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. Mol. Cell *21*, 227–237.

Yang, Y., Li, W., Hoque, M., Hou, L., Shen, S., Tian, B., and Dynlacht, B.D. (2016). PAF Complex Plays Novel Subunit-Specific Roles in Alternative Cleavage and Polyadenylation. PLOS Genet. *12*, e1005794.

Yin, K., Ueda, M., Takagi, H., Kajihara, T., Sugamata Aki, S., Nobusawa, T., Umeda-Hara, C., and Umeda, M. (2014). A dual-color marker system for in vivo visualization of cell cycle progression in Arabidopsis. Plant J. Cell Mol. Biol. *80*, 541–552.

Yu, X., and Michaels, S.D. (2010). The Arabidopsis Paf1c Complex Component CDC73 Participates in the Modification of FLOWERING LOCUS C Chromatin. Plant Physiol. *153*, 1074–1084.

Yu, M., Yang, W., Ni, T., Tang, Z., Nakadai, T., Zhu, J., and Roeder, R.G. (2015). RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. Science *350*, 1383–1386.

Yuzenkova, Y., and Zenkin, N. (2010). Central role of the RNA polymerase trigger loop in intrinsic RNA hydrolysis. Proc. Natl. Acad. Sci. U. S. A. *107*, 10878–10883.

Yuzenkova, Y., Bochkareva, A., Tadigotla, V.R., Roghanian, M., Zorov, S., Severinov, K., and Zenkin, N. (2010). Stepwise mechanism for transcription fidelity. BMC Biol. *8*, 54.

Zamudio, J.R., Kelly, T.J., and Sharp, P.A. (2014). Argonaute-bound small RNAs from promoter-proximal RNA Polymerase II. Cell *156*, 920–934.

Zatreanu, D., Mitter, R., Tumini, E., Williams, H., Gregersen, L., Roma, S., Stewart, A., Aguilera, A., and Svejstrup, J. (2019). Elongation Factor TFIIS Prevents Transcription Stress and R-Loop Accumulation to Maintain Genome Stability (Rochester, NY: Social Science Research Network).

Zenkin, N., and Yuzenkova, Y. (2015). New Insights into the Functions of Transcription Factors that Bind the RNA Polymerase Secondary Channel. Biomolecules *5*, 1195–1209.

Zhang, C., Cao, L., Rong, L., An, Z., Zhou, W., Ma, J., Shen, W.-H., Zhu, Y., and Dong, A. (2015a). The chromatin-remodeling factor AtINO80 plays crucial roles in genome stability maintenance and in plant development. Plant J. Cell Mol. Biol. *82*, 655–668.

Zhang, J., Palangat, M., and Landick, R. (2010). Role of the RNA polymerase trigger loop in catalysis and pausing. Nat. Struct. Mol. Biol. *17*, 99–104.

Zhang, L., Silva, D.-A., Pardo-Avila, F., Wang, D., and Huang, X. (2015b). Structural Model of RNA Polymerase II Elongation Complex with Complete Transcription Bubble Reveals NTP Entry Routes. PLoS Comput. Biol. *11*, e1004354.

Zhang, L., Pardo-Avila, F., Unarta, I.C., Cheung, P.P.-H., Wang, G., Wang, D., and Huang, X. (2016). Elucidation of the Dynamics of Transcription Elongation by RNA Polymerase II using Kinetic Network Models. Acc. Chem. Res. *49*, 687–694.

Zhang, T., Zhang, W., and Jiang, J. (2015c). Genome Wide Nucleosome Occupancy and Positioning and Their Impact on Gene Expression and Evolution in Plants. Plant Physiol. *168*, 1406–1416.

Zheng, Q., and Wang, X.-J. (2008). GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. Nucleic Acids Res. *36*, W358–W363.

Zuo, J., Niu, Q.W., and Chua, N.H. (2000). Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. Plant J. Cell Mol. Biol. *24*, 265–273.

Acknowledgements

I would like to thank all the people that have supported me or contributed to my work during my PhD.

Foremost, I would like to thank my PhD supervisor Prof. Dr. Klaus Grasser for the opportunity to conduct my research in his group. I am thankful for all the guidance and for the thorough evaluation of my work.

I would like to thank Prof. Dr. Thomas Dresselhaus for creating friendly and stimulating environment in the department and for leading my examination committee as a chairperson.

I would like to especially thank my mentoring team: Dr. Joachim Griesenbeck and Dr. Wen-Hui Shen for their encouragement and insightful comments widening my research from various perspectives. In addition, I would like to thank Dr. Joachim Griesenbeck for correcting this thesis and Dr. Wen-Hui Shen for hosting me in his research group as an intern. Besides my advisors, I would like to thank Prof. Dr. Herbert Tschochner for the evaluation of my work as a 3rd examiner.

In addition, I would like to thank all members of Chromatin in Plants – European Training and Mobility: Prof. Dr. Mieke Van Lijsebettens, Prof. Dr. Andreas Houben, Prof. Dr. Paloma Mas, Dr. Marc De Block, Dr. Wen-Hui Shen and every ChIP-ET fellow for creating an outstanding scientific network which stimulated my research and development.

Also I thank Dr. Astrid Bruckmann, Dr. Rasha ElBashir, Dr. Thomas Strempfl, Dr. Alexandre Berr, Dr. Jörg Fuchs and Dr. Jan Medenbach for providing their services and help with different experiments.

I am profoundly grateful to Dr. Kevin Begcy and Dr. Jules Deforges for the bioinformatical analysis of sequencing data, all valuable suggestions and patience for my remarks.

My sincere thanks go to my fellow labmates: Alex, Martin, Lena, Julius, Brian, Silvia, Hans, Philipp, Tina, Irene, Amelie, Antje, Marion, Michael and all students that I encountered. Especially I would like to thank Alex and Martin for showing me German lifestyle, Silvia for every ChIP-on-ChiP we did together, Hans for being a lab hero, Philipp for calling my work “a beast”, Kevin for his hidden joy after every RM victory and all of you just for being good friends.

Finally, my heartfelt thanks go to my beloved wife Marianna. Without your unconditional support, endless patience and countless sacrifices I couldn’t have come this far.