# Genome wide tracking of RNA polymerase II and associated factors during transcript elongation in *Arabidopsis thaliana*

DISSERTATION

ZUR ERLANGUNG DES DOKTORGRADES

DER NATURWISSENSCHAFTEN (DR. RER. NAT.)

DER FAKULTÄT BIOLOGIE UND VORKLINISCHE MEDIZIN

DER UNIVERSITÄT REGENSBURG

vorgelegt von

**Simon Obermeyer**

aus Painten

im Jahr

2023

Das Promotionsgesuch wurde eingereicht am: **03.05.2023**

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

Unterschrift: .....

**Simon Obermeyer**

This thesis is composed of the following manuscripts:

[1] **Obermeyer, S.**, Kapoor, H., Markusch, H. and Grasser, K.D. (2023), Transcript elongation by RNA polymerase II in plants: factors, regulation and impact on gene expression. Plant J. https://doi.org/10.1111/tpj.16115

[2] **Obermeyer S**, Stöckl R, Schnekenburger T, Kapoor H, Stempfl T, Schwartz U, Grasser KD. TFIIS is crucial during early transcript elongation for transcriptional reprogramming in response to heat stress. J Mol Biol. 2023 Jan 30;435(2):167917. doi: 10.1016/j.jmb.2022.167917. Epub 2022 Dec 9. PMID: 36502880.

[3] **Obermeyer S**, Stöckl R, Schnekenburger T, Moehle C, Schwartz U, Grasser KD. Distinct role of subunits of the Arabidopsis RNA polymerase II elongation factor PAF1C in transcriptional reprogramming. Front Plant Sci. 2022 Sep 29;13:974625. doi: 10.3389/fpls.2022.974625. PMID: 36247629; PMCID: PMC9558118.

[4] Michl-Holzinger P, **Obermeyer S**, Markusch H, Pfab A, Ettner A, Bruckmann A, Babl S, Längst G, Schwartz U, Tvardovskiy A, Jensen ON, Osakabe A, Berger F, Grasser KD. Phosphorylation of the FACT histone chaperone subunit SPT16 affects chromatin at RNA polymerase II transcriptional start sites in Arabidopsis. Nucleic Acids Res. 2022 May 20;50(9):5014-5028. doi: 10.1093/nar/gkac293. PMID: 35489065; PMCID: PMC9122599.

[5] Markusch, H., Michl-Holzinger, P., **Obermeyer, S.**, Thorbecke, C., Bruckmann, A., Babl, S., Längst, G., Osakabe, A., Berger, F. and Grasser, K.D. (2023), Elongation factor 1 is a component of the Arabidopsis RNA polymerase II elongation complex and associates with a subset of transcribed genes. New Phytol. New Phytol. 2023 Apr;238(1):113-124. doi: 10.1111/nph.18724. Epub 2023 Feb 1. PMID: 36627730.

[6] **Obermeyer, S.**, Schrettenbrunner, L., Stoeckl, R., Schwartz, U. and Grasser, K.D. Different elongation factors distinctly modulate RNA polymerase II transcription in *Arabidopsis*.(submitted)

The personal contribution from Simon Obermeyer **(SO)** in the manuscripts has been adapted from the author contributions section in the manuscripts:

[1] KDG wrote the manuscript. **SO** performed the structural modeling for Figure 2. **SO**, HM and HK contributed to the article and all authors approved the submitted version.

[2] **SO**, RS, and TS performed the experimental procedures; **SO**, RS and US analysed the next-generation sequencing data; **SO**, TS, and KDG designed the research; KDG wrote the manuscript and all authors approved the submitted version. All authors contributed to the article and approved the submitted version.

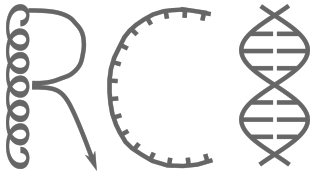
[3] **SO**, RS, and TS performed the experimental procedures; **SO**, RS, US, and CM analysed the next-generation sequencing data; **SO**, CM, and KDG designed the research; KDG wrote the manuscript and all authors approved the submitted version. All authors contributed to the article and approved the submitted version.

[4] GL, FB and KDG designed the research and supervised the project. PM-H, **SO**, HM, AP, SB and AO performed experiments. HM, PM-H, **SO**, AE, AP, AO, AT and AB analysed data. KDG wrote the manuscript. All authors discussed the results, commented the manuscript and approved the final version.

[5] GL, FB and KDG designed the research and supervised the project. HM, PM-H, **SO**, CT, SB and AO performed experiments. HM, PM-H, **SO**, SB, AB and AO analysed data. KDG wrote the manuscript. All authors discussed the results, commented the manuscript and approved the final version.

[6] KDG and **SO** designed and supervised the study. **SO**, LS and RS performed the experiments. **SO**, RS and US analysed the data. KDG and **SO** wrote the manuscript. All authors discussed the results and approved the final manuscript.

Die vorliegende Arbeit wurde im Zeitraum von April 2018 bis April 2023 am Lehrstuhl für Zellbiologie und Pflanzenbiochemie des Institutes für Pflanzenwissenschaften der Fakultät für Biologie und Vorklinische Medizin der Universität Regensburg unter Anleitung von Prof. Dr. Klaus D. Grasser angefertigt.

# Contents

# Chapter 1

# General Introduction

## 1.1 From DNA to RNA - the RNA Polymerase II

At the core of the central dogma of molecular biology, RNA polymerases enable most known organisms to transcribe genetic information into RNA. While RNAs can have regulatory, enzymatic or structuring functions, messenger RNA (mRNA) is the key molecule to deliver information to the ribosome. Transcription of genes encoding for mRNA is carried out by the RNA polymerase II (RNAPII). This molecular machinery consists of 12 subunits (Bushnell and Kornberg, 2003) that can carry out the basic role of reading the sequential information stored in DNA and adding ribonucleotides to a growing RNA chain. However, compared to RNAPI and RNAPIII, RNAPII has by far the most diverse set of targets. Eukaryotic RNAPI only has very specific targets, including the transcription of 25S, 18S, and 5.8S rRNA. Targets of RNAPIII include 5S rRNA, tRNA, 7SL RNA, U6 snRNA (Dieci et al., 2007). RNAPII targets are ranging from the permanent transcription of housekeeping genes to highly up- or downregulated stress response genes. Additionally RNAPII transcribes miRNAs, snRNAs, snoRNAs and other non-coding RNAs (Hsin and Manley, 2012). To enable specificity and accuracy for such a wide range of targets, a substantial number of additional factors are involved in fine tuning RNAPII dependent transcription in higher eukaryotes.

### 1.1.1 Initiation of transcription

The most obvious step of transcriptional regulation is the rate at which RNAPII is recruited to the promoter of genes. This is a well-studied process which, in short, consists of the interplay of conserved sequence elements and regulation of the general accessibility of the region. Different sequence elements directly on the DNA are the basic way to regulate the transcription of a gene and are present in all branches of life (Werner and Grohmann, 2011). In a well-coordinated sequence of events, several highly conserved steps regulate transcription initiation. Consequently, this mechanism is very defined and as that probably the best studied one during the transcription cycle, reviewed for example by Farnung and Vos, 2021. For RNAPII mediated transcription this step begins with the assembly of the preinitiation complex, composed of the core RNAPII and general transcription factors. This complex can unwind the DNA and form the "transcription bubble". In many cases this complex is pre-loaded on the promoter of genes and the transcription is paused until further signaling occurs. Depending on the interplay of several DNA elements and general transcription factors, this is already sufficient to achieve a good range of control.

### 1.1.2 Transcript elongation

The main focus of this work is based on the regulatory steps of transcript elongation. Compared to initiation, this step during transcript synthesis is more complex and depends on the recruitment, release and/or modification of a variety of so called transcript elongation factors (TEFs). These factors are needed to overcome obstacles during elongation. Obstacles during transcription can range from challenging sequences like repetitive regions or regions with differences in GC-composition to co-transcriptional processes like splicing to transcription over nucleosomes, including a vast amount of possible histone modifications (Chen et al., 2018). Despite the complexity of the regulation of transcript elongation, it could be the first layer of control of transcription that developed during the course of evolution (Werner and Grohmann, 2011). The following introduction of TEFs shows just a few that are well-described in several model organisms and include those this study is focused on.

#### 1.1.2.1 TFIIS

TFIIS was described in *Arabidopsis thaliana* by Grasser et al., 2009 as a nuclear protein involved in gene expression. It is a three-domain protein with an N-terminal domain, a middle domain acting as a linker and a C-terminal domain reaching into the active center of the RNAPII, where it is essential to stimulate the weak intrinsic RNA cleavage activity of the RNAPII (Kettenberger et al., 2003). Later it was shown that the RNA cleavage activity of TFIIS is crucial for rescuing backtracked and/or arrested RNAPII elongation complexes in *Arabidopsis* (Antosz et al., 2020). TFIIS deficient mice proved to have a severely impacted development (Ito et al., 2006). In yeast (Koyama et al., 2003) and in *Arabidopsis* (Grasser et al., 2009), reduced levels of TFIIS have shown only minor effects on the phenotype of the respective organisms under perfect growth conditions.

#### 1.1.2.2 The PAF1 complex

The multisubunit polymerase-associated factor 1 complex (PAF1C) consists of six subunits in plants, named ELF7 (homologue of PAF1), ELF8 (homologue of CTR9), VIP4 (homologue of LEO1), VIP5 (homologue of RTF1), VIP3 (homologue of WDR61/SKI8) and CDC73 (homologue of CDC73) in *Arabidopsis thaliana*. PAF1C has been shown to play a stabilising role in elongation beyond the pause-release step. This role correlates with several post-translational histone modifications like decreased H3K36me3 and H2b ubiquitination at target genes (Hou et al., 2019). It was shown that knockdown of human PAF1 (Arabidopsis: ELF7) and CTR9 (Arabidopsis: ELF8) displayed similar severe phenotypes (Kim et al., 2010). In plants, a functional role of PAF1C was described as a protein complex that regulates the transition from vegetative to reproductive development (van Lijsebettens and Grasser, 2014). *Arabidopsis* CDC73 was shown to play a role for a very small subset of genes including the *FLOWERING LOCUS C* (Yu and Michaels, 2010).

#### 1.1.2.3 The histone chaperone FACT

FACT (facilitates chromatin transcription) was identified to be essential to enable RNAPII for the transcription over nucleosomal templates *in vitro* (Orphanides et al., 1998). It consists of two subunits named SPT16 (suppressor of Ty 16) and SSRP1 (structure-specific recognition protein 1) (Orphanides et al., 1999). The subunits are highly conserved in many species, including in *Arabidopsis thaliana*. One prominent exception to the otherwise rather consistent naming of the FACT subunits within eukaryotic model organisms is yeast, where SSRP1 is named POB3. This FACT subunit has high sequence similarity to the homologues in other species but lacks an HMG-box domain. Since the first identification of the FACT complex, extensive research was conducted on this complex, gaining insights into the various roles of FACT in transcriptional control by establishing a specific chromatin context. This not only includes removal of nucleosomes but also reassembly and general accessibility of chromatin. Besides its role in RNAPII mediated transcription, FACT is also involved in chromatin repair and DNA replication (Formosa and Winston, 2020).

#### 1.1.2.4 SPT4-SPT5 heterodimer

As SPT5 is the only universally conserved transcription factor present in all branches of life, it acts as an indicator that the regulation of transcript elongation was the first to develop (Werner and Grohmann, 2011). Several functions were assigned to the SPT4-SPT5 dimer. These include the regulation of pausing, promoting productive elongation and the coordination of co-transcriptional mRNA processing (Song and Chen, 2022). In *Arabidopsis*, previous studies described two genes encoding SPT4 and SPT5, each. Mutant lines with reduced levels of the constitutively expressed SPT5-2 were not viable. For mutant lines with reduced levels of expression of SPT4, several severe growth defects were described and linked to defective transcript elongation (Dürr et al., 2014).

#### 1.1.2.5 ELF1

Elongation factor 1 (ELF1) was first described in yeast (Prather et al.,2005). It is a small zinc finger protein which interacts with RNAPII at the DNA entry tunnel (Ehara et al., 2017). Later it was shown to facilitate transcription through nucleosomes by adjusting the position of histones (Ehara et al., 2019). A similar function of ELF1 was also described in archaea, where ELF1 in conjunction with SPT4-SPT5 enables productive elongation of RNAPII (Blombach et al., 2021).

### 1.1.3 Termination

A last step of direct transcriptional control is the termination of the polymerisation process. As a variety of transcription units are target of RNAPII mediated transcription, there are also differences in the termination mechanism. For protein coding genes, however, this process usually involves recognition of a polyadenylation signal by the cleavage and polyadenylation complex, composed of the cleavage and polyadenylation specificity factor (CPSF), the cleavage stimulatory factor (CstF), and the cleavage factors (CF) I and II (Eaton and West, 2020). This is accompanied by a slowdown of the RNAPII. The exonuclease Xrn2 co-transcriptionally degrades RNA past the polyadenylisation signal and finally displaces the RNAPII (Cortazar et al., 2019).

## 1.2 Differences in the genome architecture of higher eukaryotes

Many plants are highly complex eukaryotes and in many ways their genome architecture, organization, and complexity is comparable to that of other complex eukaryotes like mammals. On the other hand, plants are sessile organisms and cannot easily evade extreme environmental changes. Thus, extreme and quick transcriptional flexibility is of special importance for those organisms to react to environmental cues. It has long been established that simply counting the number of genes is not a good predictor of the complexity of a species. The genome size itself is also not the single most important predictor. The best example for this is the genome size of the two widely studied plant model organisms *Arabidopsis thaliana* and *Zea mays*. While the genome size of *Arabidopsis thaliana* is approximately 135 megabase pairs (according to "The Arabidopsis Information Resource" (TAIR)) the maize genome consists of 2.4 gigabases (according to the NAM Sequencing Consortium).
This is within the range of the latest predicted human genome size of approximately 3.1 gigabases (Nurk et al., 2022) and ∼18x larger than the *Arabidopsis thaliana* genome. The length of typical transcription units of protein coding genes is approximately 2.2 kb in *Arabidopsis* compared to approximately 66.6 kb in humans. However, the approximate number of nuclear, protein-coding genes is comparable, or rather even higher in *Arabidopsis* (∼28000 in *Arabidopis*; ∼20000 in humans) (Derelle et al., 2006; Cheng et al., 2017; Piovesan et al., 2016; Nurk et al., 2022). In this context, it is of special interest to investigate the role of many conserved transcript elongation factors and how their role may differ between such highly different transcriptional challenges.

## 1.3 Additional layers of control in transcription

Arguably, multi-cellular organisms and especially higher eukaryotes are more complex than prokaryotes and therefore need a wider range and higher precision in the control of gene expression. To achieve this, additional layers of control have evolved. One of the most striking differences between prokaryotes and eukaryotes is the evolution of histones and the controlled compaction of DNA that those proteins bring with it. Bacterial gene promoter are generally open and often show a base level of initiation. On the other hand, eukaryotic promoter can be shut down completely due to the compaction mediated by the interaction of DNA with histones. Additional elements as enhancers, silencers, and insulators contribute to an increased control over gene expression. A key feature of eukaryotic cells is the packaging and compaction of DNA into chromatin. This is done by wrapping the DNA around histone octamers, in general consisting of two of the subunits H2A, H2B, H3 and H4, each. One Nucleosome comprises approximately 147 bp of DNA and one histone octamer (Luger et al., 1997). While this compaction of DNA is advantageous to organise large genomes, it also allows for a very strict control of transcription by restricting access to genes, promoter regions and upstream elements. The positioning, modifications and strength of interaction of histones with DNA directly influence transcription (Orphanides and Reinberg, 2000).

## 1.4 Transcriptional plasticity in the context of environmental stress

As sessile organisms, plants depend on quick and robust adaptability of gene expression to cope with challenging environmental conditions. Those conditions include abiotic stresses like temperature stress, salinity stress, light stress and drought stress as well as biotic stresses like transcriptional response to bacteria. Extensive data was generated and collected in case of *Arabidopsis thaliana* to determine expression patterns (Kilian et al., 2007; Winter et al.; 2007). This has shown that the expression pattern of very specific subsets of genes is quickly induced or repressed. Especially in case of heat stress, the transcriptional response is well-known (Guo et al., 2016) and inducing the heat shock response can be used to get insight into specific transcription related processes (Kumar and Wigge, 2010; Guo et al., 2016; Cortijo et al., 2017).

## 1.5 Scope of the thesis

Starting with well-established mutant lines with reduced expression of certain transcript elongation factors, the goal of this study was to get detailed knowledge on the genome wide distribution and the impact on transcription of the respective factors. Several deep sequencing approaches were combined to this end: nuclear RNA-Sequencing was used to determine differences between TEF mutant lines and wild type plants under steady state conditions as well as acute stress conditions. Chromatin immunoprecipitation (ChIP) was performed with antibodies directed against the native TEFs as well as different phosphorylation states of the carboxy-terminal domain of the largest RNAPII subunit (RNAPII-CTD) to determine the genome wide distribution of the factors in relation to the polymerase. Micrococcal nuclease (MNase) digestion of chromatin was followed up by immunoprecipitation with antibodies directed against histone H3 and deep sequencing in order to determine exact nucleosome positions in the plant genome. Finally, several published MNase-Sequencing datasets with and without immunopreciptation steps, ChIP-Sequencing datasets with immunoprecipitation directed against histone modifications and ChIP-Sequencing as well as NET-Sequencing (native elongating transcript sequencing) datasets with immunoprecipitation steps directed against different phosphorylation states of the RNAPII-CTD were reanalyzed. While well-tested antibodies against the CTD of RNAPII are commercially available, a major part of the work included creating and testing in-house made antibodies directed against transcript elongation factors and adapting the ChIP protocol to enable specific binding of the antibodies on the one hand and suitability for deep sequencing on the other hand. While many of those antibodies failed or did not show a specific signal, immunoprecipitation with several was successful. Novel insights into the function and distribution of SPT4-SPT5 (chapter 2 and chapter 7), subunits of the PAF1C (ELF7 and CDC73) (chapter 2, chapter 4 and chapter 7), subunits of FACT (SSRP1 and SPT16) (chapter 2, chapter 5 and chapter 7) and ELF1 (chapter 2 and chapter 6) in *Arabidopsis thaliana* could be obtained.

## 1.6 Publications

The following chapters comprise six articles. Five articles (chapter 2, chapter 3, chapter 4, chapter 5 and chapter 6) have been published in peer-reviewed journals and one article (chapter 7) has been submitted to a peer-reviewed journal. The PhD candidate, Simon Obermeyer, has authored four articles as first author and two as co-author. In chapter two to seven, the published and submitted versions of the manuscripts are printed, including main figures and tables. Supplementary figures, tables and data can be found online as indicated in the respective chapter. Supplementary figures, tables and data for the submitted manuscript are appended in the respective chapter (chapter 7). All articles, supplementary figures, tables and data are also included in the digital appendix of this thesis.

# Chapter 2

# Transcript elongation by RNA polymerase II in plants: factors, regulation and impact on gene expression

This peer-reviewed article was published in *The Plant Journal* in 2023

the plant journal

*The Plant Journal* (2023) doi: 10.1111/tpj.16115

SPECIAL ISSUE REVIEW

# Transcript elongation by RNA polymerase II in plants: factors, regulation and impact on gene expression

**Simon Obermeyer, Henna Kapoor, Hanna Markusch and Klaus D. Grasser***

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

Received 19 December 2022; revised 12 January 2023; accepted 17 January 2023.
*For correspondence (e-mail klaus.grasser@ur.de).

## SUMMARY

**Transcriptional elongation by RNA polymerase II (RNAPII) through chromatin is a dynamic and highly regulated step of eukaryotic gene expression. A combination of transcript elongation factors (TEFs) including modulators of RNAPII activity and histone chaperones facilitate efficient transcription on nucleosomal templates. Biochemical and genetic analyses, primarily performed in *Arabidopsis*, provided insight into the contribution of TEFs to establish gene expression patterns during plant growth and development. In addition to summarising the role of TEFs in plant gene expression, we emphasise in our review recent advances in the field. Thus, mechanisms are presented how aberrant intragenic transcript initiation is suppressed by repressing transcriptional start sites within coding sequences. We also discuss how transcriptional interference of ongoing transcription with neighbouring genes is prevented. Moreover, it appears that plants make no use of promoter-proximal RNAPII pausing in the way mammals do, but there are nucleosome-defined mechanism(s) that determine the efficiency of mRNA synthesis by RNAPII. Accordingly, a still growing number of processes related to plant growth, development and responses to changing environmental conditions prove to be regulated at the level of transcriptional elongation.**

**Keywords: RNA polymerase II, Chromatin, nucleosome, Arabidopsis, PAF1C, TFIIS, SPT4-SPT5, SPT6, FACT, PELF1.**

## INTRODUCTION

In eukaryotes, all protein-coding genes are transcribed by the 12-subunit RNA polymerase II (RNAPII) enzyme. Hence, RNAPII transcription plays a key role in differential gene expression to ensure that appropriate amounts of mRNAs are produced in a spatially and temporally coordinated manner. This is instrumental for growth, development and response to environmental conditions. The basis is established by controlling transcriptional initiation through a variety of transcription factors that bind to specific DNA elements of target genes. Diverse sets of transcription factors (and cofactors) bound to these *cis*-regulatory regions cooperate to determine in a combinatorial way the efficiency of RNAPII transcriptional initiation (Brkljacic & Grotewold, 2017; Reiter et al., 2017). For years, the following transcript synthesis was regarded as a simple polymerisation reaction, elongating the growing mRNA molecule. However, it became apparent that transcript elongation by RNAPII is a rather dynamic and discontinuous phase of the transcription cycle, which is highly regulated. Accordingly, a variety of functionally distinct transcript elongation factors (TEFs) modulate different aspects of RNAPII progression on chromatin templates (Chen et al., 2018; Kwak & Lis, 2013; Osman & Cramer, 2020; Sims et al., 2004). TEFs can be divided into various groups, such as modulators of RNAPII activity, facilitators of chromatin transcription (i.e. histone chaperones, ATP-dependent chromatin-remodelling complexes) and enzymes writing/erasing covalent histone modifications within transcribed regions (Kwak & Lis, 2013; Sims et al., 2004; van Lijsebettens & Grasser, 2014). In our review, we focus on the mechanism of chromatin transcription by RNAPII, but we largely omit the wide field of transcription-related post-translational histone modifications (i.e. acetylation, methylation, mono-ubiquitination), as that has been covered by a range of comprehensive review articles (Espinosa-Cores et al., 2020; Feng & Shen, 2014; Grasser et al., 2021; Jarosz et al., 2020; Leng, Thomas, et al., 2020; Xiao et al., 2016). Likewise, regarding the emerging research area on co-transcriptional mRNA processing in plants, we refer to

excellent recent reviews (Godoy Herz & Kornblihtt, 2019; Marquardt et al., 2023; Qin et al., 2022). In the course of our overviews of mRNA production by RNAPII, we present information derived from various organisms, most importantly plants (which means almost exclusively *Arabidopsis*), mammals and yeast. We would like to emphasise that one has to be aware that there are enormous differences between these organisms regarding important genomic parameters that undoubtedly influence the process of transcript elongation. To name but a few, the length of transcribed regions varies from typically a few thousand base pairs in yeast and *Arabidopsis* to tens or even hundreds of thousands in mammals, or the number/size of introns that increases substantially from yeast to *Arabidopsis* to mammals. In line with that, it is especially advantageous to comparatively consider the findings from different organisms, as despite these differences they still share many highly conserved TEFs. At the same time, it remains to be seen to which extent findings in *Arabidopsis* can be transferred to other plants such as maize that seem to share some genomic features (i.e. genome size) rather with mammalian systems.

## RNAPII AND THE TRANSCRIPT ELONGATION COMPLEX (TEC)

From the largest subunit of RNAPII, termed NRPB1, extends the long, repetitive and basically unstructured carboxy-terminal domain (CTD). The CTD consists of tandem heptapeptide repeats with the consensus sequence $Y_1S_2P_3T_4S_5P_6S_7$. The number of repeats varies, for instance, 26 in *Saccharomyces cerevisiae* to 52 in *Homo sapiens* (Harlen & Churchman, 2017; Jeronimo et al., 2016), while the CTD of *Arabidopsis thaliana* RNAPII contains 15 consensus and 19 divergent repeats (Hajheidari et al., 2013). RNAPII is recruited to promoters with an unphosphorylated CTD to form the pre-initiation complex, but during the advancing transcription cycle the RNAPII-CTD is dynamically phosphorylated (and modified by other post-translational modifications), which influences the interaction with factors modulating the transcription process as well as co-transcriptional events (Harlen & Churchman, 2017; Jeronimo et al., 2016). Particularly well studied is the phosphorylation of residues S2 and S5 within the CTD repeats (referred to as S2P and S5P) during transcript elongation, which thus serve as marks for elongating RNAPII. Still the exact distribution of these marks over RNAPII transcribed regions differs somewhat between organisms such as yeast and human (Harlen & Churchman, 2017; Jeronimo et al., 2016). The distribution of RNAPII-S2P and -S5P in *Arabidopsis* is again different. The typical accumulation of S5P at the transcriptional start site (TSS) seen in yeast and human, depending on the detection method, is at least less pronounced in *Arabidopsis* (Antosz et al., 2020; Zhu et al., 2018) and this mark is rather distributed over the transcribed region (Figure 1). RNAPII-S2P coverage increases towards the transcriptional end site (TES) and exhibits a prominent peak downstream of the TES. Changes in the CTD phosphorylation state during the transcription cycle play important roles, determining the transitions between initiation, elongation and termination (Harlen & Churchman, 2017; Jeronimo et al., 2016). The CTD phosphorylation, for instance, contributes to the recruitment of certain TEFs to form the RNAPII elongation complex.

Initially, the composition of the RNAPII-TEC was elucidated in yeast by biochemical approaches that were supplemented by the analysis of genetic interactions between genes encoding various TEFs (Krogan et al., 2002; Lindstrom et al., 2003; Squazzo et al., 2002). Isolation of the TEC from *Arabidopsis* cells using reciprocal tagging of different TEFs in combination with affinity purification and mass spectrometry (AP-MS) essentially confirmed the findings regarding the yeast TEC, but identified also additional

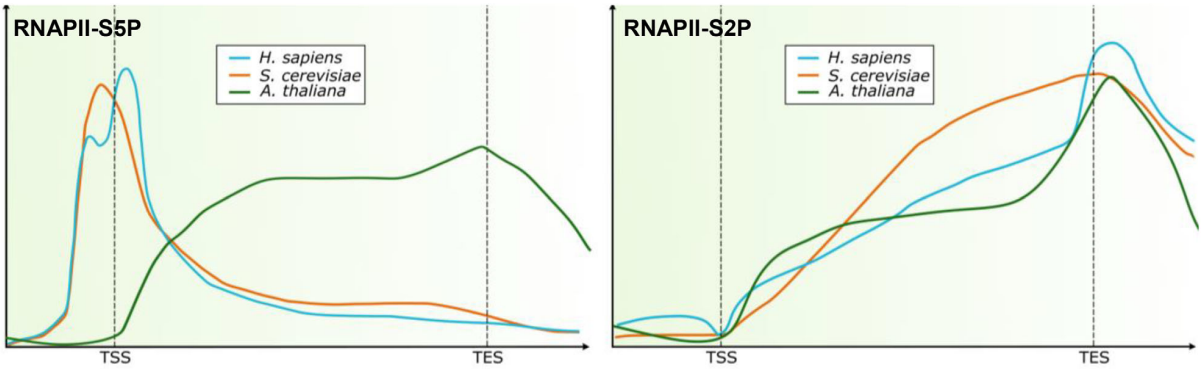



**Figure 1.** Chromatin immunoprecipitation-sequencing (ChIP-seq) profiles of phosphorylated residues of the RNAPII-CTD (S2P, S5P) across protein-coding genes in humans and budding yeast along with *Arabidopsis*.
Profiles for yeast and humans are depicted schematically as presented in Harlen & Churchman (2017). For *Arabidopsis* profiles, data plotted over highly expressed protein-coding genes (Antosz et al., 2020) have been superimposed. The second peak observed for RNAPII-S5P upstream of the transcriptional start site (TSS) in humans relates to divergent transcription, which seems to be a comparatively rare event in *Arabidopsis* (Hetzel et al., 2016; Kindgren et al., 2020; Thieffry et al., 2020; Zhu et al., 2018).

interactors (Antosz et al., 2017). RNAPII co-purified with the TEFs PAF1C, TFIIS, SPT4-SPT5, SPT6L and FACT, while P-TEFb appeared not to stably associate with the TEC in *Arabidopsis*. In addition, chromatin remodelling complexes, NAP1 histone chaperones and several histone-modifying enzymes including Elongator were found to associate with the TEC (Antosz et al., 2017). Recently, the TEF ELF1 was identified as another component of the *Arabidopsis* TEC, although it appears to associate only with a subpopulation of elongating RNAPII molecules (Markusch et al., 2023). During the past few years, advances in cryo-electron microscopy in combination with X-ray crystallography enabled enormous progress resolving the three-dimensional structure of the RNAPII-TEC of yeast and mammals. Thus, many details about the association of various TEFs with RNAPII along with the position of the DNA template and the nascent mRNA have been uncovered (Ehara et al., 2017; Vos et al., 2018, 2020). These studies demonstrated that the TEFs cover a major portion of the RNAPII surface, and that some TEFs are (tightly) connected with each other, such as SPT4-SPT5, SPT6 and PAF1C or SPT4-SPT5 and ELF1 (Figure 2). Some TEFs are placed at strategic positions, for instance, ELF1 at the 'DNA entry tunnel' for passage of the downstream DNA, and SPT4-SPT5 at the 'DNA exit tunnel'. Moreover, it became clear that many components of the TEC (i.e. TEFs) are mutually exclusive with components of the initiation complex (i.e. general transcription factors), indicating that the exchange of initiation factors with elongation factors is crucial for establishing a functional TEC and to block reassociation of initiation factors (Ehara et al., 2017; Vos et al., 2018, 2020). Further structural analyses provided also insight into the mechanism of how RNAPII with the assistance of the histone chaperone FACT and other TEFs can pass nucleosomes during transcript elongation (Ehara et al., 2019, 2022; Farnung et al., 2021, 2022; Liu et al., 2020). Initially, RNAPII (with SPT6, IWS1, PAF1C and ELF1 forming the downstream edge of the TEC) approaches the downstream nucleosome (Figure 2), starting to unwrap DNA from proximal histones and exposing the proximal H2A-H2B. FACT is successively engaged in histone and DNA contacts, eventually releasing the FACT-histone intermediate, which is then transferred presumably to the upstream edge of the TEC (formed by SPT4-SPT5, SPT6 and PAF1C). Finally, the nucleosome reassembles, explaining how nucleosomal histones and histone modifications are preserved in the wake of RNAPII passage (Ehara et al., 2019, 2022; Farnung et al., 2021; Filipovski et al., 2022). Multiple acidic regions occurring in SPT4-SPT5, SPT6, ELF1, PAF1C and FACT may contribute to binding basic nucleosomal regions exposed during transcript elongation, thereby facilitating histone transfer and nucleosome reassembly (Kujirai & Kurumizaka, 2020). Because nucleosomes in the path of RNAPII can cause backtracking of the TEC, TFIIS-stimulated cleavage of the backtracked RNA can promote nucleosome passage (Farnung et al., 2022; Kireeva et al., 2005; Nock et al., 2012). Hence, various TEFs assist RNAPII in transcribing nucleosomal templates that proceeds with an amazing rate of 1–3 kb $min^{-1}$ *in vivo* (Muniz et al., 2021). Currently, there is no structural information regarding the TEC from plants but, in view of the evolutionary conservation of RNAPII and most of the TEFs, a marked resemblance to the findings on the yeast/mammalian TEC is expected.

## FUNCTION OF RNAPII-TEFS

### Polymerase-associated factor 1 complex (PAF1C)

The multifunctional PAF1C originally was discovered and characterised in *S. cerevisiae*, where it consists of five subunits, while in metazoa it is composed of six subunits

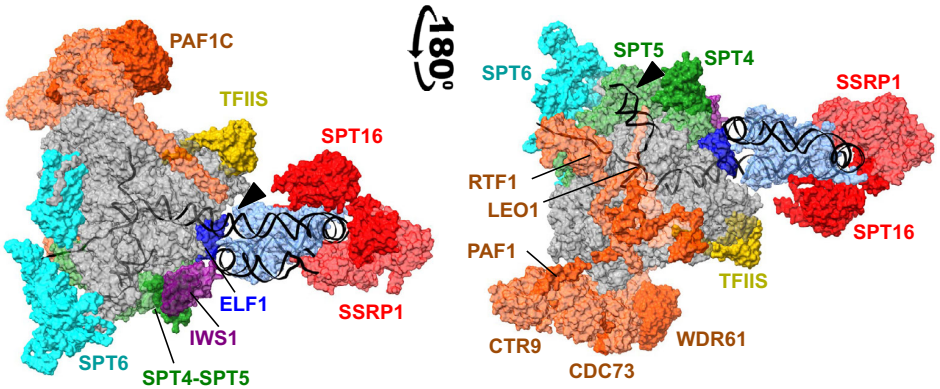



**Figure 2.** Structure of the RNAPII transcript elongation complex (TEC) engaged in nucleosome transcription.
The TEC consists of the 12-subunit RNAPII associated with several transcript elongation factors (TEFs). The shown structure is composed of an alignment of the PDB entries 7UNC (RNAPII core, PAF1C, TFIIS, SPT6, SPT4-SPT5, histones, DNA, RNA) (Filipovski et al., 2022) with 7NKY (SSRP1, SPT16) (Farnung et al., 2021) and 7XSE (ELF1, IWS1) (Ehara et al., 2022). The DNA entry- and exit-tunnels are indicated by arrowheads in the left and right views, respectively. RNAPII is indicated in grey, nucleosomal histones in light blue, DNA and RNA in black, and the colour of the respective TEFs is indicated.

(PAF1, CTR9, LEO1, RTF1, CDC73 and WDR61/SKI8). PAF1C associates with elongating RNAPII stabilising the TEC and stimulating transcription. In addition, PAF1C links transcript elongation with post-translational histone modifications over transcribed regions, including various histone methylations (i.e. H3K4me2/3, H3K36me3, K3K79me2/3) as well as H2B mono-ubiquitination (Francette et al., 2021; Jaehning, 2010). Plant PAF1C was initially recognised because of its role in the transition from vegetative to reproductive development, and consequently the subunits were named after the *Arabidopsis* mutants (Table 1) – early flowering (*elf*) and/or vernalisation independence (*vip*): PAF1 (At-ELF7), CTR9 (At-ELF8, At-VIP6), LEO1 (At-VIP4), RTF1 (At-VIP5), WDR61/SKI8 (At-VIP3) and CDC73 (At-CDC73, At-PHP) (He et al., 2004; Oh et al., 2004; Park et al., 2010; Yu & Michaels, 2010; Zhang et al., 2003; Zhang & van Nocker, 2002). Co-immunoprecipitation experiments and AP-MS analyses demonstrated that as in metazoa, *Arabidopsis* PAF1C is composed of six subunits (Antosz et al., 2017; Oh et al., 2004). Mutant *Arabidopsis* plants deficient in PAF1C subunits share early flowering phenotypes that are associated with reduced expression of the floral repressor *FLC* (and paralogs) (He et al., 2004; Oh et al., 2004; Park et al., 2010; Yu & Michaels, 2010; Zhang et al., 2003; Zhang & van Nocker, 2002). Moreover, also the temperature-responsive transition to flowering depends on PAF1C (Nasim et al., 2022). Besides the regulation of flowering, the phyllotactic regularity of spatial auxin-dependent patterning at meristems requires the action of PAF1C (Fal et al., 2017). In addition, PAF1C proved to be involved in plant responses to abiotic stress conditions. Mutants deficient in PAF1C subunits did not react as wild-type plants to mechanical stimulation inflicted by repeated touch and failed to induce touch-responsive transcripts (Jensen et al., 2017). When exposed to elevated NaCl concentrations, the growth of *elf7* and *elf8* mutants was clearly reduced when compared with wild-type plants, whereas *cdc73* plants exhibited rather enhanced tolerance. In line with their susceptibility to salt stress, the transcriptional response of *elf7* plants upon exposure to salt was severely reduced, while the response of *cdc73* plants was hardly affected (Obermeyer et al., 2022; Zhang et al., 2022).

**SPT4-SPT5 heterodimer**

SPT5 is the only elongation factor that is conserved across all domains of life from bacteria (there termed NusG) to mammals. In eukaryotes and archaea, SPT5 associates with SPT4. While yeast *spt4Δ* strains are viable, SPT5 is essential for life in various eukaryotes. SPT4-SPT5 influences pleiotropic functions during RNAPII transcription, including the regulation of pausing, promoting productive elongation and coordination of co-transcriptional mRNA processing (Decker, 2021; Hartzog & Fu, 2013; Song & Chen, 2022). In *Arabidopsis*, both SPT4 and SPT5 are encoded by two genes each. While *SPT5-1* seems to be expressed in pollen, *SPT5-2* is expressed ubiquitously throughout the plant. Homozygous individuals of three independent *spt5-2* mutant lines proved not viable, and SPT4-RNAi lines (depending on the degree of downregulation of the *SPT4*-mRNAs, which is paralleled by reduced amounts of the SPT5-2 protein) exhibited severe growth and developmental defects (Dürr et al., 2014). In line with the decreased expression of auxin-related genes (particularly *AUX/IAA* genes) in the SPT4-RNAi plants, various defects caused by impaired auxin signalling were observed, including decreased lateral root density and reduced leaf venation. In the SPT4-RNAi plants, elevated levels of RNAPII were detected over the transcribed regions (including those of downregulated genes) suggesting transcript elongation defects (Dürr et al., 2014). In *Arabidopsis*, in addition to the generally conserved RNAPII-associated SPT4-SPT5 heterodimer, SPT4 was found to directly interact with the plant-specific SPT5L/KTF1 protein that in collaboration with RNAPV modulates transcriptional silencing by RNA-directed DNA methylation (Dürr et al., 2014; He et al., 2009; Köllen et al., 2015).

**RNA cleavage-stimulating factor TFIIS**

TFIIS is a three-domain protein with the N-terminal domain I, the middle domain II and the C-terminal domain III. Domain II and the linker between domains II and III are required for RNAPII interaction. Domain III including the highly conserved acidic β-hairpin reaches to the enzyme active site, inducing extensive conformational changes to the TEC and stimulating the weak intrinsic RNA cleavage activity of RNAPII. Transcript cleavage is required to rescue backtracked/arrested RNAPII elongation complexes, thereby facilitating transcription through blocks to elongation including nucleosomes (Fish & Kane, 2002; Kettenberger et al., 2003; Noe Gonzalez et al., 2021). In accordance with the normal growth of yeast *tfIIsΔ* cells, the vegetative development of *Arabidopsis* plants lacking TFIIS is comparable to wild-type, but mutant seeds exhibit impaired dormancy (Grasser et al., 2009). The seed dormancy phenotype is caused by reduced expression of the *DOG1* gene in *tfIIs* mutant seeds (Liu et al., 2011; Mortensen & Grasser, 2014). Constitutive expression of a mutant TFIIS variant (termed TFIISmut) that efficiently inhibits the RNA cleavage activity of RNAPII proved lethal in *tfIIs* plants (Antosz et al., 2020; Dolata et al., 2015). Induced, transient expression of TFIISmut in *tfIIs* resulted in severe growth defects and transcriptomic changes. In addition, transcription-related redistribution of elongating RNAPII towards the TSSs was observed, predominantly to the position of the +1 nucleosome (first nucleosome following the TSS). Therefore, RNAPII backtracking/arrest appears to occur frequently *in planta* and TFIIS-mediated RNAPII-reactivation is essential for efficient transcription (Antosz et al., 2020). Despite their wild-

**Table 1** TEFs characterised in *Arabidopsis*

| TEF | Synonyms | Complex | Molecular function | Mutant phenotype | Reference |
|---|---|---|---|---|---|
| TFIIS | Dst1 | TFIIS | Modulates RNAPII properties<br>Stimulates the intrinsic transcript cleavage activity of RNAPII | Seed dormancy<br>Heat sensitivity | Grasser et al. (2009)<br>Antosz et al. (2020)<br>Szádeczky-Kardoss et al. (2022)<br>Obermeyer et al. (2023) |
| ELF1 | ELOF1 | ELF1 | Modulates RNAPII properties | Synergistic phenotypes with mutants lacking other TEFs | Markusch et al. (2023) |
| SPT4 | DSIF | SPT4-SPT5 | Modulates RNAPII properties | Growth and development, auxin response | Dürr et al. (2014) |
| SPT5 | DSIF | SPT4-SPT5 | Modulates RNAPII properties<br>Phosphorylated by CDKD; 2 | Growth and development, auxin response | Dürr et al. (2014) |
| IWS1 | SPN1 | IWS1 | Modulates histone modifications | Brassinosteroid-/nitrogen-dependent gene expression<br>ROS homeostasis | Li et al. (2010)<br>Widiez et al. (2011)<br>Bellegarde et al. (2019) |
| SPT6L | SPT6 | SPT6 | Histone chaperone<br>Recruitment of chromatin remodellers | Embryo basal–apical polarity | Gu et al. (2012)<br>Chen et al. (2019)<br>Shu et al. (2022) |
| VIP3 | SKI8/WDR61 | PAF1C | Modulates histone modification | Reproductive development<br>Touch response<br>Phyllotaxis | Zhang et al. (2003)<br>Fal et al. (2017)<br>Jensen et al. (2017)<br>Nasim et al. (2022) |
| VIP4 | LEO1 | PAF1C | Modulates histone modification | Reproductive development<br>Reduced seed dormancy | Zhang & van Nocker (2002)<br>Liu et al. (2011)<br>Nasim et al. (2022) |
| VIP5 | RTF1 | PAF1C | Modulates histone modifications | Reproductive development<br>Reduced seed dormancy | Oh et al. (2004)<br>Liu et al. (2011)<br>Nasim et al. (2022) |
| ELF7 | PAF1 | PAF1C | Modulates histone modifications | Reproductive development<br>Reduced salt tolerance | He et al. (2004)<br>Liu et al. (2011)<br>Obermeyer et al. (2022)<br>Zhang et al. (2022)<br>Nasim et al. (2022) |
| ELF8/ VIP6 | CTR9 | PAF1C | Modulates histone modifications | Reproductive development<br>Reduced seed dormancy<br>Reduced salt tolerance | He et al. (2004)<br>Oh et al. (2004)<br>Liu et al. (2011)<br>Fal et al. (2017)<br>Obermeyer et al. (2022)<br>Zhang et al. (2022)<br>Nasim et al. (2022) |
| CDC73/ PHP | | PAF1C | Modulates histone modifications | Reproductive development<br>Enhanced salt tolerance | Park et al. (2010)<br>Yu and Michaels (2010)<br>Obermeyer et al. (2022)<br>Nasim et al. (2022) |
| SSRP1 | Pob3 | FACT | Histone chaperone<br>Suppression of cryptic transcription | Vegetative, reproductive development<br>Reduced seed dormancy<br>Anthocyanin synthesis upon light-induction | Lolas et al. (2010)<br>Nielsen et al. (2019)<br>Michl-Holzinger et al. (2019)<br>Pfab et al. (2018)<br>Michl-Holzinger et al. (2022) |
| SPT16 | | FACT | Histone chaperone<br>Suppression of cryptic transcription<br>Phosphorylation by CK2 modulates nucleosome occupancy at TSSs | Vegetative, reproductive development<br>Anthocyanin synthesis upon light-induction | Duroux et al. (2004)<br>Lolas et al. (2010)<br>Pfab et al. (2018)<br>Nielsen et al. (2019)<br>Michl-Holzinger et al. (2022) |

ROS, reactive oxygen species; TEF, transcript elongation factor; TSS, transcriptional start site.


*The Plant Journal*, (2023), doi: 10.1111/tpj.16115

type-like growth under standard conditions, *tflls* mutants were recently found to be highly sensitive to elevated temperatures. While the Col-0 wild-type and *TFIIS* overexpressing plants survived exposure at 37°C for 2 days, it proved lethal for *tflls* mutants (Obermeyer et al., 2023; Szádeczky-Kardoss et al., 2022). Further analyses demonstrated that particularly early heat stress response is dramatically impaired in *tflls* mutants along with altered alternative splicing pattern of hundreds of transcripts under heat stress conditions (Szádeczky-Kardoss et al., 2022). In the absence of TFIIS particularly heat-stress-induced genes are expressed at lower levels than in wild-type. At genes upregulated upon heat stress in *tflls* plants promoter-proximal accumulation of RNAPII occurred at +1 nucleosomes from which the histone variant H2A.Z was evicted in a temperature-dependent manner. Thus, the heat-stress-induced promoter-proximal accumulation of RNAPII in *tflls* conforms to that seen upon TFIISmut expression (Obermeyer et al., 2023). Therefore, these studies demonstrated that TFIIS is required for efficient reprogramming of gene expression to establish plant thermotolerance (Obermeyer et al., 2023; Szádeczky-Kardoss et al., 2022).

**Zinc-finger protein ELF1**

ELF1 is a small zinc-finger protein that is conserved in eukaryotes and some archaea. Originally, it was identified as a TEF in yeast by virtue of the synthetic lethality of the *elf1Δ* mutant in combination with mutations in genes encoding other known TEFs (Prather et al., 2005). ELF1 (mammalian orthologue termed ELOF1) localises to regions actively transcribed by RNAPII (Mayer et al., 2010; Prather et al., 2005; Rossi et al., 2021) and, because of its steady association with the yeast RNAPII elongation complex during *in vitro* transcription, it was designated core elongation factor (Joo et al., 2019). A putative orthologue of yeast ELF1 is encoded in plant genomes and, recently, *Arabidopsis* ELF1 was experimentally examined. Recombinant ELF1 interacted *in vitro* with DNA, histones and nucleosomes. ELF1 is a nuclear protein, and AP-MS analyses demonstrated that it co-purified with RNAPII and various TEFs including SPT4-SPT5, SPT6L, IWS1, PAF1C and FACT (Markusch et al., 2023). Plants lacking ELF1 basically have wild-type appearance, which is in agreement with the yeast *elf1Δ* mutation that caused no significant growth defect (Prather et al., 2005). Analyses of *Arabidopsis* double-mutants revealed distinct genetic interactions between *elf1* and mutants deficient in other elongation factors. Characteristic of TEFs, ELF1 associated with genomic regions occupied by elongating RNAPII, as evident from the coverage of the RNAPII-S2P chromatin immunoprecipitation signal. However, ELF1 occupied only one-third of the RNAPII transcribed loci with preference for inducible rather than constitutively expressed genes (Markusch et al., 2023). Taken together, these results indicate that *Arabidopsis* ELF1 represents a functional plant orthologue of ELF1/ELOF1 of other organisms.

**Histone chaperone SPT6L and IWS1**

The RNAPII-associated histone chaperone SPT6 is involved in various steps of gene expression, including transcription and histone post-translational modifications. Initially, SPT6 was reported to bind the phosphorylated CTD of RNAPII, but more recently interactions with the RNAPII stalk region and the phosphorylated linker region were recognised (Sdano et al., 2017; Vos et al., 2018). SPT6 can facilitate RNAPII transcriptional elongation and suppresses aberrant intragenic transcriptional initiation by maintaining chromatin structure during ongoing transcription. In addition, by interaction with histone modifiers it influences post-translational histone modifications over transcribed chromatin (Duina, 2011; Kato et al., 2013). In the *Arabidopsis* genome, two potential orthologues are encoded, SPT6L (At1g65440) and SPT6 (At1g63210) (Gu et al., 2012), of which *SPT6L* appears to be widely expressed throughout the plant, whereas the *SPT6* transcript is barely detectable (or not expressed at all) in most tissues (Sullivan et al., 2019). The phenotype of *spt6* mutants was indistinguishable from wild-type, but *spt6l* mutants were characterised by embryos with defective apical-basal polarity, leading to embryo lethality (Gu et al., 2012). More recently, colleagues succeeded in growing *spt6l* mutant seedlings *in vitro* for molecular analyses, which revealed intriguing results (Chen et al., 2019). The genome-wide association of SPT6L correlated with that of RNAPII and, in *spt6l* plants, RNAPII occupancy was clearly reduced. Expression of a SPT6L-variant (SPT6LΔtSH2) that is defective in RNAPII interaction could partially rescue the *spt6l* phenotype and the SPT6LΔtSH2 protein accumulated at the TSS, suggesting an RNAPII-independent recruitment mechanism and a role of SPT6L during early elongation. Moreover, upon exposure to heat stress, SPT6L is rapidly recruited particularly to the TSS of heat-responsive genes, while RNAPII coverage increased over the transcribed region (Chen et al., 2019). In another study, it was uncovered that SPT6L can interact with SWI2/SNF2-type ATP-dependent chromatin remodellers SYD/BRM, linking them to the RNAPII transcription machinery (Shu et al., 2022). In line with that, SYD/BRM colocalise with SPT6L to TSS regions of many genes, and the chromatin-association of SYD/BRM is severely reduced in *spt6l* plants. Finally, SPT6L and SYD/BRM are involved in the regulation of nucleosome and RNAPII occupancy at TSS, thus contributing to early transcript elongation by RNAPII (Shu et al., 2022). The elongation factor IWS1 is a direct interactor of SPT6 that has been also characterised in plants. *Arabidopsis* IWS1 was found to interact with brassinosteroid-regulated transcription factor BES1. It has been proposed that BES1 recruits IWS1 to target genes to promote transcriptional elongation (Li et al., 2010). In this scenario, IWS1 may assist recruitment of the histone methyltransferase SDG8 to brassinosteroid-

regulated genes mediating gene expression (Wang et al., 2014). Another line of research connected IWS1 with the regulation of nitrate uptake. Under conditions of high nitrogen supply, IWS1 (there termed HNI9) mediated increased levels of H3K27me3 at the *NRT2.1* locus repressing its expression (Widiez et al., 2011). Moreover, under high nitrogen supply, IWS1 is required for the expression of detoxification genes to maintain reactive oxygen species (ROS) homeostasis (Bellegarde et al., 2019).

### Histone chaperone FACT

The conserved heterodimeric histone chaperone FACT consists of the SPT16 and SSRP1 subunits, and is involved in various DNA-dependent processes in chromatin including transcription (Formosa & Winston, 2020; Grasser, 2020; Gurova et al., 2018). Details of FACT-nucleosome interactions have been further elucidated in a recent study, illustrating that the structure of FACT resembles a unicycle, consisting of a saddle and fork that is engaged in extensive interactions of SSRP1 and SPT16 with nucleosomal DNA and all histones (Liu et al., 2020). Both SPT16 and SSRP1 (Pob3 in yeast) co-purify efficiently with components of PAF1C and SPT4-SPT5, both from yeast and *Arabidopsis* cells (Antosz et al., 2017; Krogan et al., 2002; Lindstrom et al., 2003; Squazzo et al., 2002), but recent structural studies indicate that FACT interacts with the TEC more loosely/dynamically than the above-mentioned TEFs (Ehara et al., 2022; Farnung et al., 2021). *Arabidopsis* FACT is widely expressed throughout the plant and localises to the transcriptionally active euchromatin (Duroux et al., 2004; Ikeda et al., 2011; Pfab et al., 2018). It associates with transcribed regions of active, protein-coding genes in a transcription-dependent manner (Antosz et al., 2017; Duroux et al., 2004; Perales & Más, 2007). FACT is essential for viability in *Arabidopsis* (Frost et al., 2018; Lolas et al., 2010), and decreased *SSRP1*/*SPT16* expression levels cause a variety of vegetative and reproductive defects including increased number of leaves, early bolting, impaired circadian rhythm and reduced seed dormancy (Lolas et al., 2010; Ma et al., 2018; Michl-Holzinger et al., 2019). Upon exposure to high-light stress, several anthocyanin biosynthetic genes were induced in *ssrp1*/*spt16* mutants to a lesser extent than in the wild-type and, accordingly, the mutant plants depleted in FACT accumulated lower amounts of anthocyanin pigments (Pfab et al., 2018). Recently, detailed mass spectrometric analyses revealed that both FACT subunits isolated from *Arabidopsis* cells were post-translationally modified. Four acetylation sites were mapped in the basic C-terminal region of SSRP1 and phosphorylation of three Ser/Thr residues (catalysed by protein kinase CK2) were identified in the acidic C-terminal region of SPT16 (Michl-Holzinger et al., 2022). Mutational analysis revealed only mild effects for the SSRP1 acetylation sites, while a non-phosphorylatable version of SPT16 displayed reduced histone interaction and failed to complement growth and developmental phenotypes of *spt16* mutant plants. At a subset of genes, expression of the non-phosphorylatable SPT16 version resulted in enrichment of histone H3 upstream of TSSs in a region that usually is nucleosome-depleted. Therefore, SPT16 phosphorylation might be required to establish correct nucleosome occupancy at the TSS of active genes (Michl-Holzinger et al., 2022). NAP1 histone chaperones co-purified with the *Arabidopsis* RNAPII-TEC (Antosz et al., 2017), and the histone chaperones ASF1 and HIRA share some properties with FACT and SPT6L (Layat et al., 2021; Zhong et al., 2022), suggesting that beyond FACT and SPT6L additional histone chaperones may be involved in RNAPII chromatin transcription.

## RESTRICTING TRANSCRIPTION

A fundamental challenge to gene regulation is preventing that transcription of a gene inappropriately influences expression of neighbouring genes. In metazoa, for instance, insulators ensure that transcriptional interference of that kind is suppressed (Chen & Lei, 2019; Schoborg & Labrador, 2014), while in plants such mechanism(s) are rather unclear. Interestingly, in *Arabidopsis* recently three BORDER proteins (BDR1-3) were identified that are enriched in intergenic regions and were reported to prevent interference between closely spaced genes (Yu et al., 2019). The BDRs contain a SPOC domain (found also in the SPEN family of transcriptional repressors) and a TFIIS-like domain. Single-mutants did not show clear phenotypes, whereas *bdr* triple-mutant plants exhibit severely reduced growth. BDRs apparently cause 3′ accumulation of RNAPII at the upstream gene and, in the absence of BDRs, RNAPII accumulation is reduced at this position and seems to be shifted into the promoter region of the downstream gene. The authors conclude that BDRs acting as negative elongation factors inhibit transcriptional interference by preventing RNAPII from intruding into the promoters of closely spaced tandem downstream genes (Yu et al., 2019). In view of the relatively small size and close spacing of *Arabidopsis* genes (mentioned above) in combination with the limited resolution of chromatin immunoprecipitation-sequencing (ChIP-seq) experiments, the functional principle of the BDRs remains to be clarified. Further analyses revealed that *bdr* triple-mutant plants are late flowering and fail to suppress the expression of the floral repressor *FLC*. Consistent with that compared with wild-type plants, in the *bdr* triple-mutants reduced levels of repressive histone H3K27me3 were detected over the *FLC* locus, while activating H3K4me3 was increased in the 5′ region (Yu et al., 2021). Over BDR-repressed genes high levels of RNAPII coverage were observed, albeit low levels of the corresponding mRNAs were produced. These findings imply that BDRs promote accumulation of paused or slowly elongating RNAPII over the body of BDR-repressed genes (Yu et al., 2021).

Another fundamental principle of eukaryotic transcription control is implemented by a general shutdown of transcription except that brought about by specific, positive regulatory mechanisms. This requires that TSSs of inactive genes as well as the vast number of intra- and intergenic TSSs are repressed, which is accomplished by wrapping DNA in nucleosomes (Kornberg & Lorch, 2020). As mentioned above, passage of RNAPII during transcript elongation displaces or disrupts nucleosomes along transcribed regions, and certain histone chaperones are involved in maintaining nucleosomal integrity over transcribed regions (Venkatesh & Workman, 2015). In this scenario, an important function of FACT in *Arabidopsis* chromatin was discovered by mapping thousands of intragenic TSS positions in *spt16* and *ssrp1* mutants that were not detected in wild-type plants (Nielsen et al., 2019). Hence, FACT is required for repression of aberrant intragenic transcript initiation at positions that in part are characterised by elevated histone H3K4me1 levels (Nielsen et al., 2019). Because a similar function of FACT was observed in other organisms (Formosa & Winston, 2020; Gurova et al., 2018), safeguarding transcriptional fidelity by restricting transcriptional initiation by RNAPII to promoters may be a key role of FACT (and possibly some other histone chaperones).

## REGULATION OF EARLY ELONGATION

The elongation rates of RNAPII transcription vary between genes, but also within transcribed regions, and elongation is particularly slow during early elongation (Core & Adelman, 2019; Jonkers & Lis, 2015). RNAPII elongation rates influence also the mRNA synthesis in *Arabidopsis* under different experimental conditions through transcriptional as well as co-transcriptional mechanisms (Godoy Herz et al., 2019; Leng, Ivanov, et al., 2020; Wu et al., 2016). In metazoa, promoter-proximal pausing of engaged RNAPII plays an important role in shaping the transcriptome. RNAPII typically pauses within a narrow region of 25–50 nt downstream of the TSS and is released in a controlled manner into productive elongation. The interplay of SPT4-SPT5 and the four-subunit NELF complex is a critical determinant of establishment and release of the paused state (Core & Adelman, 2019; Dollinger & Gilmour, 2021). NELF orthologues are not encoded in yeast, nematode and plant genomes. The fact that the organisms that exhibit stable, regulated RNAPII pausing are those that encode NELF implies a functional connection (Core & Adelman, 2019). In line with that, various genome-wide studies in *Arabidopsis* failed to detect a (closely) related RNAPII pausing mechanism. However, recent GRO-seq, NET-seq and RNAPII ChIP-seq analyses are compatible with the option of a nucleosome-defined mechanism for promoter-proximal RNAPII stalling that may contribute to regulate RNAPII transcript elongation (Antosz et al., 2020; Hetzel et al., 2016; Kindgren et al., 2020; Liu et al., 2021; Zhu et al., 2018). Particularly, the position of the +1 nucleosome represents a critical site for RNAPII stalling in *Arabidopsis* upon exposure to low or high temperatures (Kindgren et al., 2020; Obermeyer et al., 2023), or upon inhibition of the RNAPII transcript cleavage activity (Antosz et al., 2020). Nucleosomes generally represent obstacles to RNAPII transcription, and their dynamics/stability is determined by a complex interplay of nucleosomal histone composition, histone post-translational modifications, nucleosome occupancy and positioning. The +1 nucleosome is a major barrier to RNAPII transcription that in a way controls the transition from initiation to productive elongation (Kujirai & Kurumizaka, 2020; Lai & Pugh, 2017). Correspondingly, +1 nucleosomes exhibit special features including that they

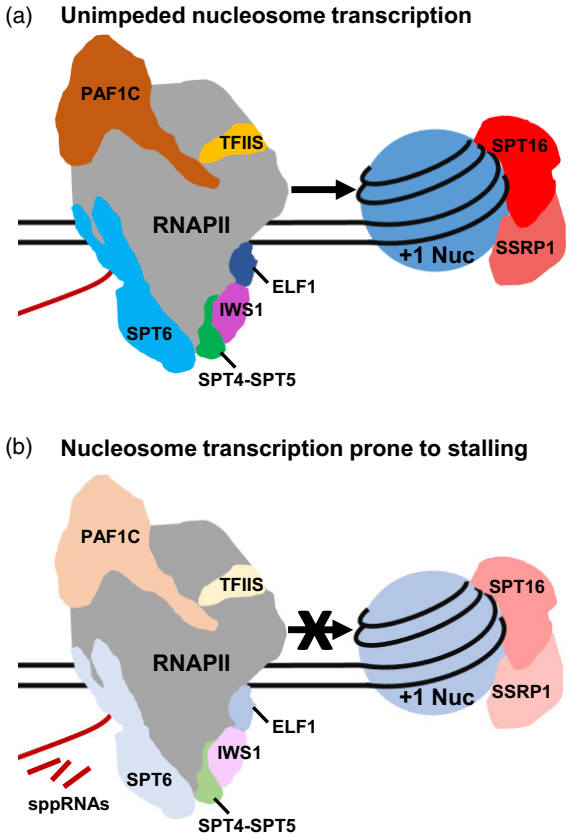



**Figure 3.** RNAPII-transcript elongation complex (TEC) approaching a +1 nucleosome under different conditions.
(a) RNAPII associated with all required transcript elongation factors (TEFs) during early elongation encounters a +1 nucleosome, which assisted by the TEFs is successfully passed and mRNA is synthesised.
(b) RNAPII associated with an incomplete set of TEFs (indicated by faded colours) encounters an altered version of the +1 nucleosome (e.g. differing in the presence/absence of a histone variant or modification, indicated by faded colour). The absence of one or several TEFs and/or the altered +1 nucleosome causes stalling of the elongating RNAPII and incomplete nucleosome passage. A reduced amount of mRNA (or none at all) is synthesised, instead sppRNAs could be produced, a process that may influence the synthesis of the corresponding mRNA.

are well-positioned and enriched in the histone variant H2A.Z (Lai & Pugh, 2017; Talbert & Henikoff, 2017). While H2A.Z incorporation into gene body nucleosomes is associated with gene repression, the correlation of H2A.Z incorporation at TSSs and plant gene expression is rather unclear and may depend on additional chromatin features (Jarillo & Piñeiro, 2015; Lei & Berger, 2020; Probst et al., 2020). In case of *Arabidopsis* genes that are upregulated upon heat stress, assistance of TFIIS is required for RNAPII to transcribe through +1 nucleosomes that are depleted in H2A.Z, suggesting that eviction of H2A.Z in this scenario results in a higher barrier for RNAPII passage (Obermeyer et al., 2023), which is in line with the lower *in vitro* stability of *Arabidopsis* H2A.Z containing nucleosomes (Osakabe et al., 2018). Intriguingly, in another study using GRO-seq, engaged RNAPII was found to accumulate downstream of TSSs upon heat stress during the immediate transcriptional response in *Arabidopsis* (Liu et al., 2021). Various studies provided evidence that additional TEFs including SPT4-SPT5, SPT6, ELF1 and FACT as well as ATP-dependent chromatin remodellers facilitate nucleosome transcription in yeast and metazoa (Gamarra & Narlikar, 2021; Kujirai & Kurumizaka, 2020). Consistently, SWI/SNF-type chromatin remodelling complexes were found recently to be targeted to nucleosomes at TSSs of *Arabidopsis* genes, likely contributing to tune early RNAPII transcript elongation (Diego-Martin et al., 2022; Shu et al., 2022). The identification of short promoter-proximal RNAs (sppRNAs) at ~14% of transcribed *Arabidopsis* genes, whose length is essentially defined by the +1 nucleosome, suggests that RNAPII stalling at this position can result in transcriptional termination (Thomas et al., 2020). Transcription events initiating at the same TSS can form full-length mRNAs or alternatively sppRNAs. At some genes increased mRNA production at the expense of sppRNAs was observed, but with most genes mRNA synthesis appears to be independent of sppRNA production (Thomas et al., 2020). In summary, it can be stated that early RNAPII transcript elongation in plants is a subject of extensive regulation (Figure 3), albeit likely by mechanisms (partially) distinct from those implemented in mammals. Depending on conditions, different molecular mechanisms including TEFs, chromatin remodellers, histone modifications and variants, and possibly sppRNAs presumably act in a concerted manner at TSS/+1 nucleosome to determine the efficiency of mRNA synthesis by RNAPII.

## ACKNOWLEDGEMENTS

Our research is supported by the German Research Foundation (DFG) through grants Gr1159/16-1 and SFB960/A6 to K.D.G. Open Access funding enabled and organized by Projekt DEAL.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## REFERENCES

**Antosz, W., Deforges, J., Begcy, K., Bruckmann, A., Poirier, Y., Dresselhaus, T. *et al.*** (2020) Critical role of transcript cleavage in Arabidopsis RNA polymerase II transcriptional elongation. *Plant Cell*, **32**, 1449–1463.

**Antosz, W., Pfab, A., Ehrnsberger, H.F., Holzinger, P., Köllen, K., Mortensen, S.A. *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.

**Bellegarde, F., Maghiaoui, A., Boucherez, J., Krouk, G., Lejay, L., Bach, L. *et al.*** (2019) The chromatin factor HNI9 and ELONGATED HYPOCOTYL5 maintain ROS homeostasis under high nitrogen provision. *Plant Physiology*, **180**, 582–592.

**Brkljacic, J. & Grotewold, E.** (2017) Combinatorial control of plant gene expression. *Biochimica et Biophysica Acta*, **1860**, 31–40.

**Chen, C., Shu, J., Li, C., Thapa, R.K., Nguyen, V., Yu, K. *et al.*** (2019) RNA polymerase II-independent recruitment of SPT6L at transcription start sites in Arabidopsis. *Nucleic Acids Research*, **47**, 6714–6725.

**Chen, D. & Lei, E.P.** (2019) Function and regulation of chromatin insulators in dynamic genome organization. *Current Opinion in Cell Biology*, **58**, 61–68.

**Chen, F.X., Smith, E.R. & Shilatifard, A.** (2018) Born to run: control of transcription elongation by RNA polymerase II. *Nature Reviews. Molecular Cell Biology*, **19**, 464–478.

**Core, L. & Adelman, K.** (2019) Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes & Development*, **33**, 960–982.

**Decker, T.-M.** (2021) Mechanisms of transcription elongation factor DSIF (Spt4-Spt5). *Journal of Molecular Biology*, **433**, 166657.

**Diego-Martin, B., Pérez-Alemany, J., Candela-Ferre, J., Corbalán-Acedo, A., Pereyra, J., Alabadí, D. *et al.*** (2022) The TRIPLE PHD FINGERS proteins are required for SWI/SNF complex-mediated +1 nucleosome positioning and transcription start site determination in Arabidopsis. *Nucleic Acids Research*, **50**, 10399–10417.

**Dolata, J., Guo, Y., Kolowerzo, A., Smolinski, D., Brzyzek, G., Jarmolowski, A. *et al.*** (2015) NTR1 is required for transcription elongation checkpoints at alternative exons in *Arabidopsis*. *The EMBO Journal*, **34**, 544–558.

**Dollinger, R. & Gilmour, D.S.** (2021) Regulation of promoter proximal pausing of RNA polymerase II in metazoans. *Journal of Molecular Biology*, **433**, 166897.

**Duina, A.A.** (2011) Histone chaperones Spt6 and FACT: similarities and differences in modes of action at transcribed genes. *Genetics Research International*, **2011**, 625210.

**Duroux, M., Houben, A., Růzicka, K., Friml, J. & Grasser, K.D.** (2004) The chromatin remodelling complex FACT associates with actively transcribed regions of the Arabidopsis genome. *The Plant Journal*, **40**, 660–671.

**Dürr, J., Lolas, I.B., Sørensen, B.B., Schubert, V., Houben, A., Melzer, M. *et al.*** (2014) The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in Arabidopsis. *Nucleic Acids Research*, **42**, 4332–4347.

**Ehara, H., Kujirai, T., Fujino, Y., Shirouzu, M., Kurumizaka, H. & Sekine, S.-I.** (2019) Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. *Science*, **363**, 744–747.

**Ehara, H., Kujirai, T., Shirouzu, M., Kurumizaka, H. & Sekine, S.-I.** (2022) Structural basis of nucleosome disassembly and reassembly by RNAPII elongation complex with FACT. *Science*, **377**, eabp9466. Available from: https://doi.org/10.1126/science.abp9466

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

**Espinosa-Cores, L., Bouza-Morcillo, L., Barrero-Gil, J., Jiménez-Suárez, V., Lázaro, A., Piqueras, R. *et al.*** (2020) Insights into the function of the NuA4 complex in plants. *Frontiers in Plant Science*, **11**, 125.

**Fal, K., Liu, M., Duisembekova, A., Refahi, Y., Haswell, E.S. & Hamant, O.** (2017) Phyllotactic regularity requires the Paf1 complex in Arabidopsis. *Development*, **144**, 4428–4436.

**Farnung, L., Ochmann, M., Engeholm, M. & Cramer, P.** (2021) Structural basis of nucleosome transcription mediated by Chd1 and FACT. *Nature Structural & Molecular Biology*, **28**, 382–387.

**Farnung, L., Ochmann, M., Garg, G., Vos, S.M. & Cramer, P.** (2022) Structure of a backtracked hexasomal intermediate of nucleosome transcription. *Molecular Cell*, **82**, 3126–3134.e7.

**Feng, J. & Shen, W.H.** (2014) Dynamic regulation and function of histone monoubiquitination in plants. *Frontiers in Plant Science*, **5**, 83.

**Filipovski, M., Soffers, J.H.M., Vos, S.M. & Farnung, L.** (2022) Structural basis of nucleosome retention during transcription elongation. *Science*, **376**, 1313–1316.

**Fish, R.N. & Kane, C.M.** (2002) Promoting elongation with transcript cleavage stimulatory factors. *Biochimica et Biophysica Acta*, **1577**, 287–307.

**Formosa, T. & Winston, F.** (2020) The role of FACT in managing chromatin: disruption, assembly, or repair? *Nucleic Acids Research*, **48**, 11929–11941.

**Francette, A.M., Tripplehorn, S.A. & Arndt, K.M.** (2021) The Paf1 complex: a keystone of nuclear regulation operating at the interface of transcription and chromatin. *Journal of Molecular Biology*, **433**, 166979.

**Frost, J.M., Kim, M.Y., Park, G.T., Hsieh, P.-H., Nakamura, M., Lin, S.J.H. *et al.*** (2018) FACT complex is required for DNA demethylation at heterochromatin during reproduction in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, **115**, E4720–E4729.

**Gamarra, N. & Narlikar, G.J.** (2021) Collaboration through chromatin: motors of transcription and chromatin structure. *Journal of Molecular Biology*, **433**, 166876.

**Godoy Herz, M.A. & Kornblihtt, A.R.** (2019) Alternative splicing and transcription elongation in plants. *Frontiers in Plant Science*, **10**, 309.

**Godoy Herz, M.A., Kubaczka, M.G., Brzyżek, G., Servi, L., Krzyszton, M., Simpson, C. *et al.*** (2019) Light regulates plant alternative splicing through the control of transcriptional elongation. *Molecular Cell*, **73**, 1066–1074.e3.

**Grasser, K.D.** (2020) The FACT histone chaperone: tuning gene transcription in the chromatin context to modulate plant growth and development. *Frontiers in Plant Science*, **11**, 85.

**Grasser, K.D., Rubio, V. & Barneche, F.** (2021) Multifaceted activities of the plant SAGA complex. *Biochimica et Biophysica Acta*, **1864**, 194613.

**Grasser, M., Kane, C.M., Merkle, T., Melzer, M., Emmersen, J. & Grasser, K.D.** (2009) Transcript elongation factor TFIIS is involved in *Arabidopsis* seed dormancy. *Journal of Molecular Biology*, **386**, 598–611.

**Gu, X.L., Wang, H., Huang, H. & Cui, X.F.** (2012) SPT6L encoding a putative WG/GW-repeat protein regulates apical-basal polarity of embryo in *Arabidopsis*. *Molecular Plant*, **5**, 249–259.

**Gurova, K., Chang, H.-W., Valieva, M.E., Sandlesh, P. & Studitsky, V.M.** (2018) Structure and function of the histone chaperone FACT - resolving FACTual issues. *Biochimica et Biophysica Acta*, **1861**, 892–904.

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

**Harlen, K.M. & Churchman, L.S.** (2017) The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nature Reviews. Molecular Cell Biology*, **18**, 263–273.

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

**He, X.J., Hsu, Y.F., Zhu, S., Wierzbicki, A.T., Pontes, O., Pikaard, C.S. *et al.*** (2009) An effector of RNA-directed DNA methylation in *Arabidopsis* is an ARGONAUTE 4- and RNA-binding protein. *Cell*, **137**, 498–508.

**He, Y., Doyle, M.R. & 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 & Development*, **18**, 2774–2784.

**Hetzel, J., Duttke, S.H., Benner, C. & Chory, J.** (2016) Nascent RNA sequencing reveals distinct features in plant transcription. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 12316–12321.

**Ikeda, Y., Kinoshita, Y., Susaki, D., Ikeda, Y., Iwano, M., Takayama, S. *et al.*** (2011) HMG domain containing SSRP1 is required for DNA demethylation and genomic imprinting in Arabidopsis. *Developmental Cell*, **21**, 589–596.

**Jaehning, J.A.** (2010) The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochimica et Biophysica Acta*, **1799**, 279–388.

**Jarillo, J.A. & Piñeiro, M.** (2015) H2A.Z mediates different aspects of chromatin function and modulates flowering responses in *Arabidopsis*. *The Plant Journal*, **83**, 96–109.

**Jarosz, M., van Lijsebettens, M. & Woloszynska, M.** (2020) Plant elongator-protein complex of diverse activities regulates growth, development, and immune responses. *International Journal of Molecular Sciences*, **21**, 6912.

**Jensen, G.S., Fal, K., Hamant, O. & Haswell, E.S.** (2017) The RNA polymerase-associated factor 1 complex is required for plant touch responses. *Journal of Experimental Botany*, **68**, 499–511.

**Jeronimo, C., Collin, P. & Robert, F.** (2016) The RNA polymerase II CTD: the increasing complexity of a low-complexity protein domain. *Journal of Molecular Biology*, **428**, 2607–2622.

**Jonkers, I. & Lis, J.T.** (2015) Getting up to speed with transcription elongation by RNA polymerase II. *Nature Reviews. Molecular Cell Biology*, **16**, 167–177.

**Joo, Y.J., Ficarro, S.B., Chun, Y., Marto, J.A. & Buratowski, S.** (2019) In vitro analysis of RNA polymerase II elongation complex dynamics. *Genes & Development*, **33**, 578–589.

**Kato, H., Okazaki, K. & Urano, T.** (2013) Spt6: two fundamentally distinct functions in the regulation of histone modification. *Epigenetics*, **8**, 1249–1253.

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

**Kindgren, P., Ivanov, M. & Marquardt, S.** (2020) Native elongation transcript sequencing reveals temperature dependent dynamics of nascent RNAPII transcription in Arabidopsis. *Nucleic Acids Research*, **48**, 2332–2347.

**Kireeva, M.L., Hancock, B., Cremona, G.H., Walter, W., Studitsky, V.M. & Kashlev, M.** (2005) Nature of nucleosomal barrier to RNA polymerase II. *Molecular Cell*, **18**, 108.

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

**Kornberg, R.D. & Lorch, Y.** (2020) Primary role of the nucleosome. *Molecular Cell*, **79**, 371–375.

**Krogan, N.J., Kim, M., Ahn, S.H., Zhong, G., Kobor, M.S., Cagney, G. *et al.*** (2002) RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a targeted proteomics approach. *Molecular Cell Biology*, **22**, 6979–6992.

**Kujirai, T. & Kurumizaka, H.** (2020) Transcription through the nucleosome. *Current Opinion in Structural Biology*, **61**, 42–49.

**Kwak, H. & Lis, J.T.** (2013) Control of transcriptional elongation. *Annual Reviews of Genetics*, **47**, 483–508.

**Lai, W.K.M. & Pugh, B.F.** (2017) Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nature Reviews. Molecular Cell Biology*, **18**, 548–562.

**Layat, E., Bourcy, M., Cotterell, S., Zdzieszyńska, J., Desset, S., Duc, C. *et al.*** (2021) The histone chaperone HIRA is a positive regulator of seed germination. *International Journal of Molecular Sciences*, **22**, 4031.

**Lei, B. & Berger, F.** (2020) H2A variants in Arabidopsis: versatile regulators of genome activity. *Plant Communications*, **1**, 100015.

**Leng, X., Ivanov, M., Kindgren, P., Malik, I., Thieffry, A., Brodersen, P. *et al.*** (2020) Organismal benefits of transcription speed control at gene boundaries. *EMBO Reports*, **21**, e49315.

**Leng, X., Thomas, Q., Rasmussen, S.H. & Marquardt, S.** (2020) A G(enomic) P(ositioning)S(ystem) for plant RNAPII transcription. *Trends in Plant Science*, **25**, 744–764.

**Li, L., Ye, H., Guo, H. & Yin, Y.** (2010) *Arabidopsis* IWS1 interacts with transcription factor BES1 and is involved in plant steroid hormone brassinosteroid regulated gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 3918–3923.

**Lindstrom, D.L., Squazzo, S.L., Muster, N., Burckin, T.A., Wachter, K.C., Emigh, C.A. *et al.*** (2003) Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. *Molecular Cell Biology*, **23**, 1368–1378.

**Liu, M., Zhu, J. & Dong, Z.** (2021) Immediate transcriptional responses of Arabidopsis leaves to heat shock. *Journal of Integrative Plant Biology*, **63**, 468–483.

**Liu, Y., Geyer, R., van Zanten, M., Carles, A., Li, Y., Hörold, A. *et al.*** (2011) Identification of the Arabidopsis REDUCED DORMANCY 2 gene uncovers a role for the polymerase associated factor 1 complex in seed dormancy. *PLoS One*, **6**, e22241.

**Liu, Y., Zhou, K., Zhang, N., Wei, H., Tan, Y.Z., Zhang, Z. *et al.*** (2020) FACT caught in the act of manipulating the nucleosome. *Nature*, **577**, 426–431.

**Lolas, I.B., Himanen, K., Grønlund, J.T., Lynggaard, C., Houben, A., Melzer, M. *et al.*** (2010) The transcript elongation factor FACT affects Arabidopsis vegetative and reproductive development and genetically interacts with HUB1/2. *The Plant Journal*, **61**, 686–697.

**Ma, Y., Gil, S., Grasser, K.D. & Mas, P.** (2018) Targeted recruitment of the basal transcriptional machinery by LNK clock components controls the circadian rhythms of nascent RNAs in Arabidopsis. *Plant Cell*, **30**, 907–924.

**Markusch, H., Michl-Holzinger, P., Obermeyer, S., Thorbecke, C., Bruckmann, A., Babl, S. *et al.*** (2023) ELF1 is a component of the Arabidopsis RNA polymerase II elongation complex and associates with a subset of transcribed genes. *The New Phytologist*. Available from: https://doi.org/10.1111/NPH.18724

**Marquardt, S., Petrillo, E. & Manavella, P.A.** (2023) Cotranscriptional RNA processing and modification in plants. *Plant Cell*. Available from: https://doi.org/10.1093/plcell/koac309

**Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Söding, J. & Cramer, P.** (2010) Uniform transitions of the general RNA polymerase II transcription complex. *Nature Structural & Molecular Biology*, **17**, 1272–1278.

**Michl-Holzinger, P., Mortensen, S.A. & Grasser, K.D.** (2019) The SSRP1 subunit of the histone chaperone FACT is required for seed dormancy in *Arabidopsis*. *Journal of Plant Physiology*, **236**, 105–108.

**Michl-Holzinger, P., Obermeyer, S., Markusch, H., Pfab, A., Ettner, A., Bruckmann, A. *et al.*** (2022) Phosphorylation of the FACT histone chaperone subunit SPT16 affects chromatin at RNA polymerase II transcriptional start sites in *Arabidopsis*. *Nucleic Acids Research*, **50**, 5014–5028.

**Mortensen, S.A. & Grasser, K.D.** (2014) The seed dormancy defect of *Arabidopsis* mutants lacking the transcript elongation factor TFIIS is caused by reduced expression of the *DOG1* gene. *FEBS Letters*, **588**, 47–51.

**Muniz, L., Nicolas, E. & Trouche, D.** (2021) RNA polymerase II speed: a key player in controlling and adapting transcriptome composition. *The EMBO Journal*, **40**, e105740.

**Nasim, Z., Susila, H., Jin, S., Youn, G. & Ahn, J.H.** (2022) Polymerase II-associated factor 1 complex-regulated FLOWERING LOCUS C-clade genes repress flowering in response to chilling. *Frontiers in Plant Science*, **13**, 817356.

**Nielsen, M., Ard, R., Leng, X., Ivanov, M., Kindgren, P., Pelechano, V. *et al.*** (2019) Transcription-driven chromatin repression of intragenic transcription start sites. *PLoS Genetics*, **15**, e1007969.

**Nock, A., Ascano, J.M., Barrero, M.J. & Malik, S.** (2012) Mediator-regulated transcription through the +1 nucleosome. *Molecular Cell*, **48**, 837–848.

**Noe Gonzalez, M., Blears, D. & Svejstrup, J.Q.** (2021) Causes and consequences of RNA polymerase II stalling during transcript elongation. *Nature Reviews. Molecular Cell Biology*, **22**, 3–21.

**Obermeyer, S., Stöckl, R., Schnekenburger, T., Kapoor, H., Stempfl, T., Schwartz, U. *et al.*** (2023) TFIIS is crucial during early transcript elongation for transcriptional reprogramming in response to heat stress. *Journal of Molecular Biology*, **435**, 167917.

**Obermeyer, S., Stöckl, R., Schnekenburger, T., Moehle, C., Schwartz, U. & Grasser, K.D.** (2022) Distinct role of subunits of the Arabidopsis RNA polymerase II elongation factor PAF1C in transcriptional reprogramming. *Frontiers in Plant Science*, **13**, 974625.

**Oh, S., Zhang, H., Ludwig, P. & van Nocker, S.** (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.

**Osakabe, A., Lorkovic, Z.J., Kobayashi, W., Tachiwana, H., Yelagandula, R., Kurumizaka, H. *et al.*** (2018) Histone H2A variants confer specific properties to nucleosomes and impact on chromatin accessibility. *Nucleic Acids Research*, **46**, 7675–7685.

**Osman, S. & Cramer, P.** (2020) Structural biology of RNA polymerase II transcription: 20 years on. *Annual Review of Cell and Development Biology*, **36**, 1–34.

**Park, S., Oh, S., Ek-Ramos, J. & van Nocker, S.** (2010) PLANT HOMOLOGOUS TO PARAFIBROMIN is a component of the PAF1 complex and assists in regulating expression of genes within H3K27ME3-enriched chromatin. *Plant Physiology*, **153**, 821–831.

**Perales, M. & Más, P.** (2007) A functional link between rhythmic changes in chromatin structure and the Arabidopsis biological clock. *Plant Cell*, **19**, 2111–2123.

**Pfab, A., Breindl, M. & Grasser, K.D.** (2018) The Arabidopsis histone chaperone FACT is required for stress-induced expression of anthocyanin biosynthetic genes. *Plant Molecular Biology*, **96**, 367–374.

**Prather, D., Krogan, N.J., Emili, A., Greenblatt, J.F. & Winston, F.** (2005) Identification and characterization of Elf1, a conserved transcription elongation factor in *Saccharomyces cerevisiae*. *Molecular Cell Biology*, **25**, 10122–10135.

**Probst, A.V., Desvoyes, B. & Gutierrez, C.** (2020) Similar yet critically different: the distribution, dynamics and function of histone variants. *Journal of Experimental Botany*, **71**, 5191–5204.

**Qin, Y., Long, Y. & Zhai, J.** (2022) Genome-wide characterization of nascent RNA processing in plants. *Current Opinion in Plant Biology*, **69**, 102294.

**Reiter, F., Wienerroither, S. & Stark, A.** (2017) Combinatorial function of transcription factors and cofactors. *Current Opinion in Genetics & Development*, **43**, 73–81.

**Rossi, M.J., Kuntala, P.K., Lai, W.K.M., Yamada, N., Badjatia, N., Mittal, C. *et al.*** (2021) A high-resolution protein architecture of the budding yeast genome. *Nature*, **592**, 309–314.

**Schoborg, T. & Labrador, M.** (2014) Expanding the roles of chromatin insulators in nuclear architecture, chromatin organization and genome function. *Cellular and Molecular Life Sciences*, **71**, 4089–4113.

**Sdano, M.A., Fulcher, J.M., Palani, S., Chandrasekharan, M.B., Parnell, T.J., Whitby, F.G. *et al.*** (2017) A novel SH2 recognition mechanism recruits Spt6 to the doubly phosphorylated RNA polymerase II linker at sites of transcription. *eLife*, **6**, e28723.

**Shu, J., Ding, N., Liu, J., Cui, Y. & Chen, C.** (2022) Transcription elongator SPT6L regulates the occupancies of the SWI2/SNF2 chromatin remodelers SYD/BRM and nucleosomes at transcription start sites in Arabidopsis. *Nucleic Acids Research*, **50**, 12754–12767.

**Sims, R.J., Belotserkovskaya, R. & Reinberg, D.** (2004) Elongation by RNA polymerase II: the short and long of it. *Genes & Development*, **18**, 2437–2468.

**Song, A. & Chen, F.X.** (2022) The pleiotropic roles of SPT5 in transcription. *Transcription*, **13**, 53–69.

**Squazzo, S.L., Costa, P.J., Lindstrom, D.L., Kumer, K.E., Simic, R., Jennings, J.L. *et al.*** (2002) The Paf complex physically and functionally associates with transcription elongation factors *in vivo*. *The EMBO Journal*, **21**, 1764–1174.

**Sullivan, A., Purohit, P.K., Freese, N.H., Pasha, A., Esteban, E., Waese, J. *et al.*** (2019) An 'eFP-Seq Browser' for visualizing and exploring RNA sequencing data. *The Plant Journal*, **100**, 641–654.

**Szádeczky-Kardoss, I., Szaker, H.M., Verma, R., Darkó, É., Pettkó-Szandtner, A., Silhavy, D. *et al.*** (2022) Elongation factor TFIIS is essential for heat stress adaptation in plants. *Nucleic Acids Research*, **50**, 1927–1950.

**Talbert, P.B. & Henikoff, S.** (2017) Histone variants on the move: substrates for chromatin dynamics. *Nature Reviews. Molecular Cell Biology*, **18**, 115–126.

**Thieffry, A., Vigh, M.L., Bornholdt, J., Ivanov, M., Brodersen, P. & Sandelin, A.** (2020) Characterization of *Arabidopsis thaliana* promoter Bidirectionality and antisense RNAs by inactivation of nuclear RNA decay pathways. *Plant Cell*, **32**, 1845–1867.

**Thomas, Q.A., Ard, R., Liu, J., Li, B., Wang, J., Pelechano, V. *et al.*** (2020) Transcript isoform sequencing reveals widespread promoter-proximal transcriptional termination in Arabidopsis. *Nature Communications*, **11**, 2589.

**van Lijsebettens, M. & Grasser, K.D.** (2014) Transcript elongation factors: shaping transcriptomes after transcript initiation. *Trends in Plant Science*, **19**, 717–726.

**Venkatesh, S. & Workman, J.L.** (2015) Histone exchange, chromatin structure and the regulation of transcription. *Nature Reviews. Molecular Cell Biology*, **16**, 178–189.

**Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H. *et al.*** (2018) Structure of activated transcription complex pol II-DSIF-PAF-SPT6. *Nature*, **560**, 607–612.

**Vos, S.M., Farnung, L., Linden, A., Urlaub, H. & Cramer, P.** (2020) Structure of complete pol II-DSIF-PAF-SPT6 transcription complex reveals RTF1 allosteric activation. *Nature Structural & Molecular Biology*, **27**, 668–677.

**Wang, X., Chen, J., Xie, Z., Liu, S., Nolan, T., Ye, H. *et al.*** (2014) Histone lysine methyltransferase SDG8 is involved in brassinosteroid-regulated gene expression in *Arabidopsis thaliana*. *Molecular Plant*, **7**, 1303–1315.

**Widiez, T., El Kafafi, S., Girin, T., Berr, A., Ruffel, S., Krouk, G. *et al.*** (2011) High nitrogen insensitive 9 (HNI9)-mediated systemic repression of root NO3- uptake is associated with changes in histone methylation.

*Proceedings of the National Academy of Sciences of the United States of America*, **108**, 13329–13334.

**Wu, Z., Ietswaart, R., Liu, F., Yang, H., Howard, M. & Dean, C.** (2016) Quantitative regulation of *FLC* via coordinated transcriptional initiation and elongation. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 218–223.

**Xiao, J., Lee, U.S. & Wagner, D.** (2016) Tug of war: adding and removing histone lysine methylation in *Arabidopsis*. *Current Opinion in Plant Biology*, **34**, 41–53.

**Yu, X., Martin, P.G.P. & Michaels, S.D.** (2019) BORDER proteins protect expression of neighboring genes by promoting 3′ pol II pausing in plants. *Nature Communications*, **10**, 4359.

**Yu, X., Martin, P.G.P., Zhang, Y., Trinidad, J.C., Xu, F., Huang, J. *et al.*** (2021) The BORDER family of negative transcription elongation factors regulates flowering time in Arabidopsis. *Current Biology*, **31**, 5377–5384.e5.

**Yu, X. & Michaels, S.D.** (2010) The *Arabidopsis* Paf1c complex component CDC73 participates in the modification of FLOWERING LOCUS C chromatin. *Plant Physiology*, **153**, 1074–1084.

**Zhang, H., Li, X., Song, R., Zhan, Z., Zhao, F., Li, Z. *et al.*** (2022) Cap-binding complex assists RNA polymerase II transcription in plant salt stress response. *Plant, Cell & Environment*, **45**, 2780–2793.

**Zhang, H., Ransom, C., Ludwig, P. & van Nocker, S.** (2003) Genetic analysis of early flowering mutants in Arabidopsis defines a class of pleiotropic developmental regulator required for expression of the flowering-time switch flowering locus C. *Genetics*, **164**, 347–358.

**Zhang, H. & van Nocker, S.** (2002) The VERNALIZATION INDEPENDENCE 4 gene encodes a novel regulator of FLOWERING LOCUS C. *The Plant Journal*, **31**, 663–673.

**Zhong, Z., Wang, Y., Wang, M., Yang, F., Thomas, Q.A., Xue, Y. *et al.*** (2022) Histone chaperone ASF1 mediates H3.3-H4 deposition in Arabidopsis. *Nature Communications*, **13**, 6970.

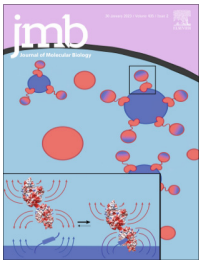
**Zhu, J., Liu, M., Liu, X. & Dong, Z.** (2018) RNA polymerase II activity revealed by GRO-seq and pNET-seq in *Arabidopsis*. *Nature Plants*, **4**, 1112–1123.

# Chapter 3

# TFIIS Is Crucial During Early Transcript Elongation for Transcriptional Reprogramming in Response to Heat Stress

This peer-reviewed article was published in the *Journal of Molecular Biology* in 2023

Communication



# TFIIS Is Crucial During Early Transcript Elongation for Transcriptional Reprogramming in Response to Heat Stress

**Simon Obermeyer [1†], Richard Stöckl [1†], Tobias Schnekenburger [1], Henna Kapoor [1], Thomas Stempfl [2], Uwe Schwartz [3] and Klaus D. Grasser [1*]**

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

*2 - **Center of Excellence for Fluorescent Bioanalytics (KFB),** University of Regensburg, Am Biopark 9, D-93053 Regensburg, Germany*

*3 - **NGS Analysis Centre,** Biology and Pre-Clinical Medicine, University of Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany*

***Correspondence to Klaus D. Grasser:** Klaus.Grasser@ur.de (K.D. Grasser)*

https://doi.org/10.1016/j.jmb.2022.167917

***Edited by Sepideh Khorasanizadeh***

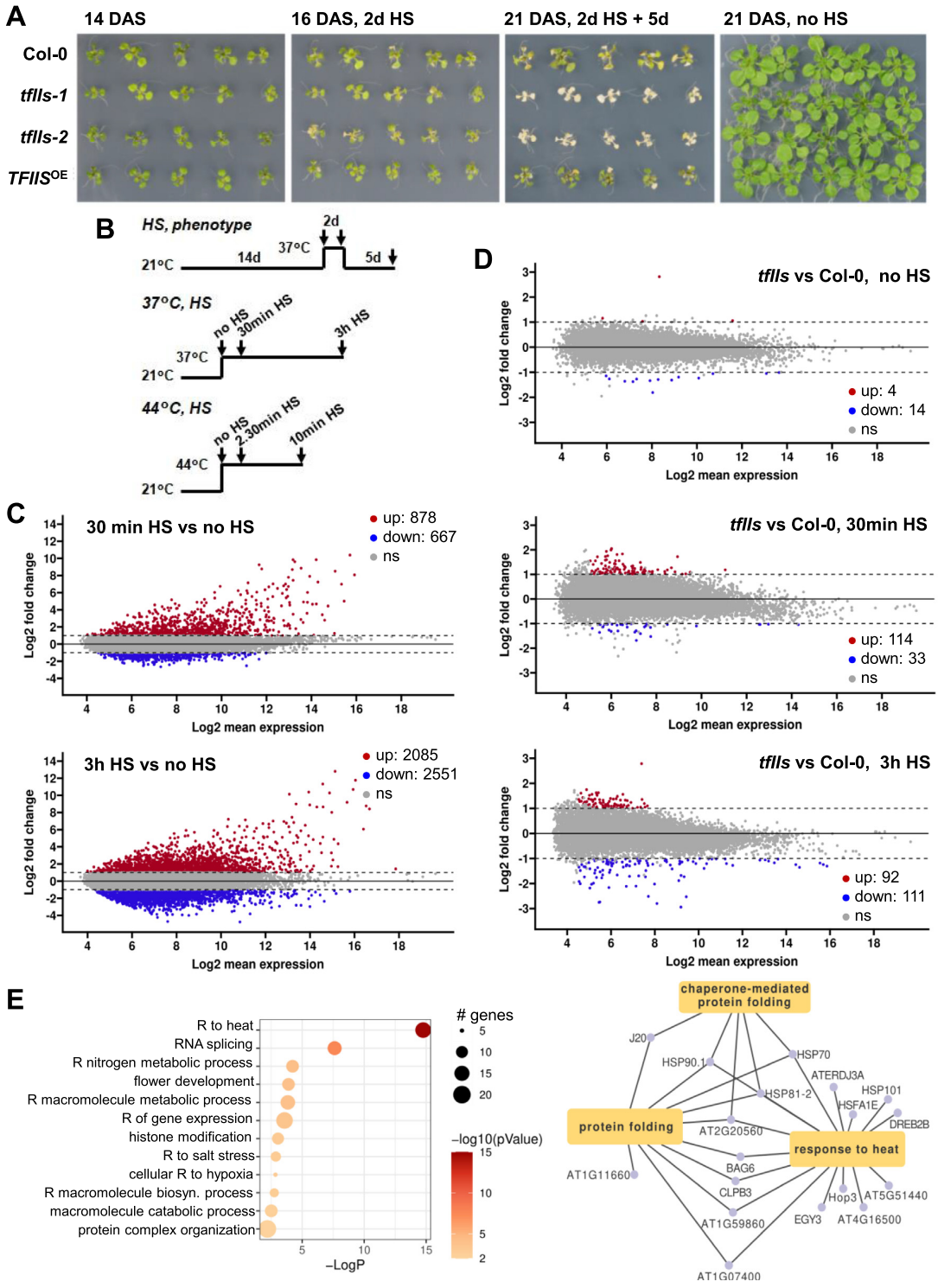
## Abstract

In addition to the stage of transcriptional initiation, the production of mRNAs is regulated during elongation. Accordingly, the synthesis of mRNAs by RNA polymerase II (RNAPII) in the chromatin context is modulated by various transcript elongation factors. TFIIS is an elongation factor that stimulates the transcript cleavage activity of RNAPII to reactivate stalled elongation complexes at barriers to transcription including nucleosomes. Since *Arabidopsis tfIIs* mutants grow normally under standard conditions, we have exposed them to heat stress (HS), revealing that *tfIIs* plants are highly sensitive to elevated temperatures. Transcriptomic analyses demonstrate that particularly HS-induced genes are expressed at lower levels in *tfIIs* than in wildtype. Mapping the distribution of elongating RNAPII uncovered that in *tfIIs* plants RNAPII accumulates at the +1 nucleosome of genes that are upregulated upon HS. The promoter-proximal RNAPII accumulation in *tfIIs* under HS conditions conforms to that observed upon inhibition of the RNAPII transcript cleavage activity. Further analysis of the RNAPII accumulation downstream of transcriptional start sites illustrated that RNAPII stalling occurs at +1 nucleosomes that are depleted in the histone variant H2A.Z upon HS. Therefore, assistance of early transcript elongation by TFIIS is required for reprogramming gene expression to establish plant thermotolerance.



## Introduction

The synthesis of eukaryotic pre-mRNAs by RNAPII is precisely controlled as a prerequisite for proper development and appropriate response to changing conditions. Transcript abundance can be adjusted by controlling transcriptional initiation by sophisticated mechanisms involving a multitude of transcription factors that functionally interact with specific DNA sequence elements. Beyond that a variety of functionally distinct transcript elongation factors associate with RNAPII to regulate the efficiency of mRNA synthesis during transcript elongation. They comprise for instance modulators of RNAPII activity and histone chaperones, and are particularly important for transcript elongation in the chromatin context.[1–3] Among the modulators of RNAPII activity is TFIIS that can facilitate transcription through blocks to elongation.[4–5] The conserved acidic β-hairpin of TFIIS reaches to the


A
14 DAS
16 DAS, 2d HS
21 DAS, 2d HS + 5d
21 DAS, no HS
Col-0
tflls-1
tflls-2
TFIIS^OE
B
HS, phenotype
21°C
14d
37°C
2d
5d
37°C, HS
no HS
30min HS
3h HS
44°C, HS
2.30min HS
10min HS
C
30 min HS vs no HS
up: 878
down: 667
ns
3h HS vs no HS
up: 2085
down: 2551
D
tflls vs Col-0, no HS
up: 4
down: 14
tflls vs Col-0, 30min HS
up: 114
down: 33
tflls vs Col-0, 3h HS
up: 92
down: 111
Log2 fold change
Log2 mean expression
E
R to heat
RNA splicing
R nitrogen metabolic process
flower development
R macromolecule metabolic process
R of gene expression
histone modification
R to salt stress
cellular R to hypoxia
R macromolecule biosyn. process
macromolecule catabolic process
protein complex organization
–LogP
# genes
–log10(pValue)
chaperone-mediated protein folding
protein folding
response to heat
J20
HSP90.1
HSP70
ATERDJ3A
HSP81-2
HSP101
HSFA1E
DREB2B
AT2G20560
BAG6
AT1G11660
CLPB3
Hop3
AT5G51440
EGY3
AT4G16500
AT1G59860
AT1G07400

enzyme active site, inducing extensive conformational changes to the elongation complex and stimulating the weak intrinsic RNA cleavage activity of RNAPII.[6] Transcript cleavage is required to rescue backtracked/arrested RNAPII elongation complexes.[7–8] Nucleosomes represent obstacles to transcript elongation and TFIIS can promote efficient nucleosome passage by RNAPII *in vitro*[9–11] and *in vivo*.[12] In view of the biochemical importance of TFIIS for transcript elongation it was surprising that yeast *tfIIs* cells grow normally.[13] Likewise, the vegetative development of *Arabidopsis tfIIs* mutants resembles that of wildtype, albeit mutant seeds show impaired dormancy.[14] At the same time reactivation of stalled/backtracked RNAPII by TFIIS-stimulated transcript cleavage contributes to correct transcriptional output in yeast and *Arabidopsis*.[15–16]

## Results and Discussion

In view of the apparent dispensability of TFIIS for vegetative growth under standard conditions, we explored the performance of *Arabidopsis tfIIs* plants upon exposure to abiotic stress conditions that provoke transcriptional reprogramming. Consistent with a very recent report,[17] it became apparent in our analyses that *tfIIs* plants are highly sensitive to elevated ambient temperatures. Exposure to 37 °C for 48 h proved lethal for plants of two different *tfIIs* alleles, whereas Col-0 wildtype and TFIIS overexpressing plants similarly survived the HS (Figure 1(A,B)). Detection of *HSP70/101* transcripts and of the HSP70 protein served to evaluate induction kinetics of gene expression under our experimental conditions (Supplemental Figure 1). To examine genome-wide transcriptional changes of Col-0 and *tfIIs* upon exposure to heat, nuclear RNA was isolated from untreated plants and plants exposed to 37 °C for 30 min or 3 h. We intended to generate information on transcriptional output (freshly synthesised unspliced/spliced mRNAs including nascent transcripts) rather than steady-state mRNA levels obtained with total poly (A) mRNA.[18–19] Therefore, we (i) isolated nuclear RNA (rather than total RNA), (ii) used rRNA depletion (rather than poly(A) enrichment) and (iii) made use of low RNA size cut-off (≥25 nt). Analysis by high throughput sequencing revealed that compared to untreated plants, after 30 min HS, 1545 differentially expressed genes (DEGs) were detected with Col-0 and *tfIIs*, while after 3 h HS 4636 DEGs were detected (Figure 1(C)). Consistent with their comparable growth and development under control conditions[14] only 18 DEGs between *tfIIs* and Col-0 were observed, whereas 147 DEGs and 203 DEGs were detected after 30 min and 3 h HS, respectively (Figure 1(D)). Interestingly, also *TFIIS* is 2.7-fold upregulated after 3 h HS in Col-0. In total, 356 DEGs were identified that differ between Col-0 and *tfIIs* at the three conditions (no HS, 30 min HS, 3 h HS). Using unsupervised k-means clustering, these DEGs can be divided into five distinct clusters according to their change of expression over time (Supplemental Figure 2). GO term analysis of the DEGs misregulated in *tfIIs* revealed no specific enrichment of HS-related genes in clusters A, B and E. However, a clear enrichment of genes responsive to heat (encoding e.g. heat shock transcription factors, heat shock proteins, factors involved in protein folding) was detected in clusters C and D, which include genes generally upregulated upon HS at different kinetics, but to a significantly lower extent in *tfIIs* (Figure 1(E)). Improper expression of these heat-responsive genes likely contributes to the impaired heat tolerance of *tfIIs* plants.

Our transcriptomic analysis of *tfIIs* relative to wildtype upon HS differs in several respects from that performed by Szádeczky-Kardoss et al.[17] Thus, we have analysed nuclear RNA (compared with total RNA) and we have measured transcript levels before HS and after 30 min or 3 h HS at 37 °C (compared with before HS and 1 h or 24 h HS at 37 °C). Finally, the bioinformatic analyses of the obtained data were performed in a different way. Still, comparable principal conclusions can

**Figure 1.** Plants lacking TFIIS are highly sensitive to heat stress (HS) and TFIIS is required for the transcriptional response to HS. (A) Exposure to 37 °C for 2d is lethal for *tfIIs*, but not for wildtype plants. Col-0 wildtype, two *tfIIs* mutant lines (*tfIIs-1, tfIIs-2*) and TFIIS overexpressing plants ($TFIIS^{OE}$) grown at 21 °C were documented 14 days after stratification (DAS), then 2d after HS at 37 °C (16 DAS), after 5d recovery at 21 °C (21 DAS). For comparison plants (21 DAS) grown without HS are shown. (B) Temperature scheme used for phenotyping (in A), for molecular analyses upon 37 °C HS and 44 °C HS. (C) Exposure to 37 °C induces transcriptional reprogramming. HS at 37 °C for 30 min or 3 h results in major transcriptomic changes in Col-0 and *tfIIs* when compared to plants without HS. (D) The lack of TFIIS affects the transcriptional response to HS. In absence of HS only minor transcriptomic changes are observed between Col-0 and *tfIIs* that are clearly enhanced upon 3 h HS. (E) Transcription of genes responsive to HS is particularly affected in *tfIIs* plants. GO term analysis of the DEGs of gene clusters C,D (cf. Supplemental Figure 2) that upon HS are upregulated in *tfIIs* to a lesser extent than in Col-0 (left) with "R" indicating "response". Gene network plot of enriched processes in cluster C (right). Differentially expressed genes (small grey nodes) are connected with their associated biological processes (big yellow nodes). Shown are processes significantly enriched and related to heat response.

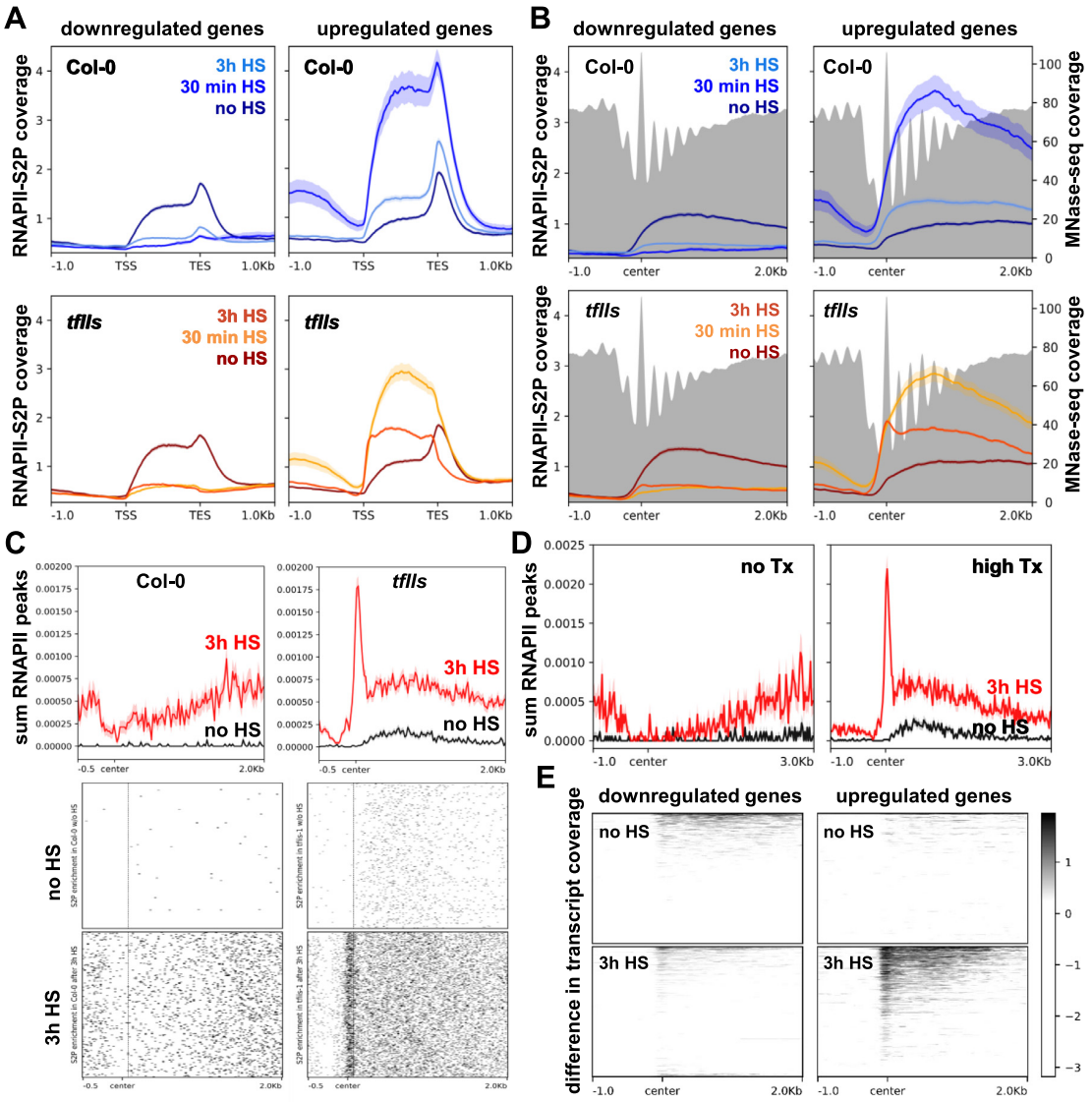



**Figure 2.** Upon 3 h HS elongating RNAPII accumulates at the +1 nucleosome of upregulated genes. (A) In absence of TFIIS, the distribution of RNAPII is shifted towards the transcriptional start site of HS-induced genes. Metagene profiles of RNAPII distribution along genes up- and downregulated (n = 1792 and n = 2222, respectively) upon HS in Col-0 and *tflls*. RNAPII-S2P coverage over scaled transcribed regions from transcriptional start site (TSS) to transcriptional end site (TES) determined by ChIP-seq in Col-0 (top) and *tflls* (bottom) after 30 min/3h HS or without . Lines represent the accumulation of RNAPII-S2P, while the shaded area indicates the standard error of the mean. (B) In *tflls* plants, RNAPII accumulates at the position of +1 nucleosomes of 3 h HS-induced genes. Metagene profiles of RNAPII distribution (after 30 min/3h HS or without) around TSS of up- and downregulated genes centred to the position of the +1 nucleosome. RNAPII coverage in Col-0 (top) and *tflls* (bottom) is plotted over nucleosomal positions (grey) determined by MNase-seq. (C) Distribution of RNAPII over all protein-coding genes illustrates HS-dependent accumulation of RNAPII at the +1 nucleosome in *tflls*. Sum of RNAPII-S2P ChIP-seq peaks (top) and heatmap (bottom) over all non-overlapping protein-coding genes (n = 26293) after 3 h HS or without, centred to the +1 nucleosome. (D) The accumulation of RNAPII in *tflls* at the +1 nucleosome is dependent on the transcriptional activity (Tx). Sum of RNAPII-S2P ChIP-seq peaks over genes with different transcriptional activity centred to the +1 nucleosome. Two groups of non-overlapping protein coding TAIR10 annotated genes were defined according to transcripts per million in 14-DAS Col-0 seedlings (left, <1 TPM, n = 6527; right, >10 TPM, n = 5539). Upon 3 h HS, RNAPII accumulates at the +1 nucleosome of highly transcribed genes. (E) RNA-seq counts are enriched in *tflls* at the position of the +1 nucleosome. Median difference in read coverage between *tfiis* and Col-0 centred over +1 nucleosomes of up- and downregulated genes without HS (top) or after 3 h HS (bottom).

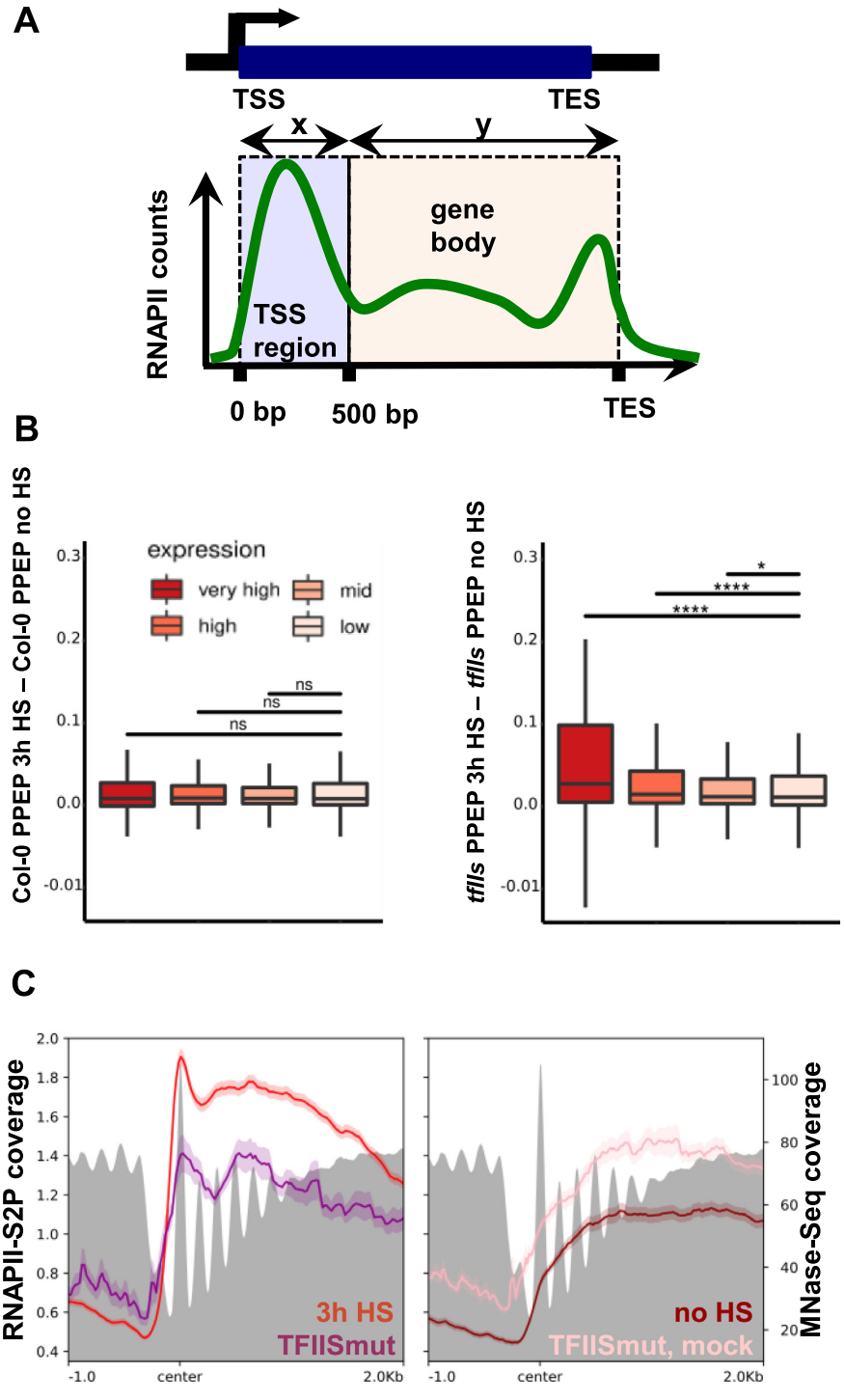
be drawn from both studies. We have compared the specific misregulation of a subset of genes in *tflls* with the results of Szádeczky-Kardoss et al.[17] While many genes were generally upregulated upon HS in both datasets (1551 in Szádeczky-Kardoss et al.; 2085 in our dataset), only a relatively small subset of HS-upregulated genes are misregulated in *tflls* in both datasets (14.4% in Szádeczky-Kardoss et al.; 9.7% in our dataset). Gene Set Enrichment Analysis[20] of the *tflls* misregulated genes of the Szádeczky-Kardoss et al. data (*tflls* vs Col-0, 1 h HS) with our dataset (*tflls* vs Col-0, 3 h HS) demonstrated that generally genes upregulated in *tflls* in Szádeczky-Kardoss et al. are also upregulated in our data and vice versa (Supplemental Figure 3). Finally, in line with our findings GO term analysis of *tflls*-dependent DEGs in Szádeczky-Kardoss et al.[17] revealed processes including response to heat stress and protein folding (cf. Figure 1(E)).

To elucidate how the absence of TFIIS impairs transcription upon HS, we mapped the distribution of elongating RNAPII genome-wide. To this end chromatin immunoprecipitation combined with high throughput sequencing (ChIP-seq) was performed using an antibody specific for RNAPII-S2P comparing Col-0 and *tflls* plants with or without exposure to 30 min/3h HS. Analysis of the data demonstrated that expectedly in both genotypes RNAPII coverage was increased upon HS along genes that are upregulated by HS, while RNAPII coverage was decreased upon HS along genes that are downregulated by HS (Figure 2 (A)). The RNAPII distribution pattern along down-regulated genes differed only slightly between *tflls* and Col-0. However, with genes upregulated upon HS, after 30 min/3h HS in *tflls* a clear shift of RNAPII distribution towards the transcriptional start site (TSS) was detected, in combination with a striking decline of signal at the transcriptional end site (TES) (Figure 2(A)), suggesting reduced RNAPII processivity. To examine possible correlation with chromatin features, the RNAPII ChIP-seq data was superimposed with nucleosomal coverage tracks derived from genome-wide micrococcal nuclease (MNase) cleavage analysis followed by histone H3 immunoprecipitation (MNase-seq). Centring the data to the +1 nucleosomal position revealed that in *tflls* at upregulated genes a prominent peak of RNAPII accumulation is visible at the position of the +1 nucleosome upon 3 h HS. This peak does not occur in Col-0, upon the shorter 30 min HS or without HS (Figure 2(B)). Therefore, we focussed on the data obtained upon 3 h HS for further analysis. Even when analysing the distribution of called RNAPII-peaks in *tflls* compared to Col-0 over all protein-coding genes, a remarkable enrichment of RNAPII in *tflls* upon 3 h HS is apparent (Figure 2(C)). The RNAPII accumulation depends on transcriptional activity of the affected genes (Figure 2(D)). In accord with the RNAPII ChIP-seq data, centring the above-mentioned RNA-seq data to the +1 nucleosome demonstrated accumulation of transcript read counts at the position of the +1 nucleosome in *tflls* relative to Col-0 upon exposure to 3 h HS (Figure 2(E). Col-0 and *tflls* were exposed to a rapid HS of 2.30 and 10 min at 44 °C (using a water bath to achieve faster temperature change,[21]) to compare short-term responses with the moderately quick responses described above. Relative to control conditions, the rapid HS resulted in detection of 1181 DEGs, of which 485 overlap with those of 30 min/3 h HS at 37 °C and of the 831 DEGs detected after 10 min HS at 44 °C, 239 and 124 DEGs were also down- and upregulated, respectively, upon 3 h HS at 37 °C (Supplemental Figure 4(A,B)). The genes upregulated under both conditions can be divided into two clusters: genes that are comparably upregulated in both genotypes after 10 min (rapid response, comprising predominantly the GO term “ribosome”) and genes that in both genotypes are upregulated already after 2.30 min, but in *tflls* do not reach wildtype-levels after 10 min (very rapid response, comprising predominantly the GO term “response to heat”) (Supplemental Figure 4(C,D)). ChIP-seq using an antibody directed against RNAPII-S2P after 10 min at 44 °C demonstrated that RNAPII accumulated in *tflls* at the +1 nucleosome of the very rapidly responding genes, while no accumulation was detected without HS and at the cluster of less rapidly responding genes and in Col-0 (Supplemental Figure 4(E)).

Using the 3 h-HS RNAPII ChIP-seq data we calculated the promoter proximal enrichment of RNAPII-S2P (PPEP) within the TSS region relative to the gene body.[16] In line with the accumulation at the +1 nucleosome, in *tflls* PPEP increased significantly upon HS depending on the transcriptional activity of the respective genes, whereas no change in PPEP was observed in Col-0 (Figure 3 (A,B)). A TFIIS mutant variant (termed TFIISmut) efficiently inhibits the transcript cleavage activity of RNAPII. Expression of TFIISmut in the *Arabidopsis tflls* mutant background resulted in RNAPII accumulation at a position overlapping with the +1 nucleosome.[16] Therefore, we reanalysed this data and superimposed the RNAPII coverage tracks with the RNAPII ChIP-seq tracks of HS-induced genes in *tflls* with and without HS. This comparison revealed that HS-induction and induced expression of TFIISmut similarly resulted in RNAPII accumulation at the +1 nucleosome (Figure 3(C)), suggesting that the TFIIS-stimulated transcript cleavage by RNAPII is required for passage of RNAPII through the +1 nucleosome.

In various organisms including plants, it is well established that the histone variant H2A.Z is enriched at the TSS, particularly in the +1 nucleosome, but still the functional relation between H2A.Z and RNAPII transcription remains

enigmatic.[22–24] Moreover, H2A.Z is closely linked to the thermosensory response in *Arabidopsis*. The high H2A.Z occupancy at the *HSP70* gene +1 nucleosome decreased drastically at elevated temperatures.[25] At non-inducible temperatures, H2A.Z nucleosomes are enriched at temperature-responsive genes and upon exposure to high temperatures H2A.Z is rapidly evicted, most likely by specific exchange for H2A. This results in a transcriptional cascade, triggering plant responses to increased temperatures.[26] In view of the enrichment of H2A.Z at the +1 nucleosome and its involvement in the response to elevated temperatures, we examined a possible connection with TFIIS in this scenario. The RNAPII-S2P distribution in *tfIIs* and Col-0 with and without exposure to HS



was centred to the +1 nucleosome of genes with low or high H2A.Z coverage.[27] In *tfIIs* exposed to HS, a prominent peak of RNAPII accumulation is apparent at the +1 nucleosome of genes with low H2A.Z coverage that is not seen in Col-0 or in the absence of HS, and that is substantially smaller at genes enriched in H2A.Z (Figure 4(A,B)). Moreover, dividing genes that are up- or downregulated in *tfIIs* upon 3 h HS (cf. Figure 1(C)) into categories with low or high H2A.Z coverage,[27] revealed that RNAPII-S2P accumulates selectively at +1 nucleosomes of upregulated genes with low H2A.Z coverage (Supplemental Figure 5(A)). Comparing groups of genes that are up- or downregulated in *tfIIs* relative to Col-0 upon 3 h HS (cf. Figure 1(D)) confirmed that at the +1 nucleosomal position H2A.Z is distinctly more depleted at the downregulated genes (Supplemental Figure 5(B)). We further analysed the interplay of TFIIS with H2A.Z at the +1 nucleosome by mapping RNAPII-S2P distribution in *tfIIs* and Col-0 over genes that display different transcriptional response to elevated temperature as described by Cortijo et al.[26] We compared predominantly repressed genes that are involved in metabolic processes with genes rapidly and markedly upregulated upon HS at 37 °C that are enriched in the GO terms of heat and light responses. At repressed genes with and without HS the H2A.Z coverage remains unchanged, whereas a significant depletion of H2A.Z occurs at the upregulated genes, particularly at the +1 nucleosome.[26] Reanalysis of this data along with our RNAPII-S2P ChIP-seq data demonstrates that in absence of HS, RNAPII coverage in *tfIIs* and Col-0 differs only slightly (Figure 4(C)). However, upon HS there is a very distinct peak of RNAPII accumulation in *tfIIs* at the +1 nucleosome of the rapidly upregulated genes, while this effect does not notably occur in Col-0 and at repressed genes. This effect is seen similarly with clusters C and D of the five gene clusters defined above (cf. Supplemental Figure 2) that are upregulated upon HS in *tfIIs* to a lesser extent than in Col-0. In absence of HS, for genes of both clusters a prominent H2A.Z peak is observed at the +1 nucleosome that is substantially decreased upon HS at 37 °C, while the H3 coverage is rather unchanged (Supplemental Figure 6). The decrease in H2A.Z is paralleled by accumulation of RNAPII at the +1 nucleosome, which is most prominent with genes of cluster C and becomes apparent already after 30 min of HS. The correlation of H2A.Z depletion and RNAPII accumulation at the +1 nucleosome is also evident from genome browser views of individual genes (Supplemental Figure 7). Taken together these analyses reveal that TFIIS is required under acute HS conditions for RNAPII to transcribe efficiently through the +1 nucleosome of induced genes depleted in H2A.Z, whereas under normal growth conditions (and higher H2A.Z occupancy) early transcript elongation appears to proceed rather unrestrained in absence of TFIIS.

Nucleosomes represent general translocation barriers for transcript elongation, resulting in backtracking of RNAPII. TFIIS associates with elongating RNAPII along transcribed regions of active genes.[15,17,28–29] There TFIIS stimulates the RNA cleavage activity of RNAPII and thus can promote reactivation of backtracked RNAPII and passage through nucleosomes, which is particularly relevant at the +1 nucleosome that contributes to control the transition from initiation to productive transcript elongation.[5,30] In plants, the promoter-proximal region including the +1 nucleosome is an object of various modulating mechanisms mediated by histone modifications/variants, histone chaperones, chromatin remodelling complexes and short promoter-proximal RNAs that adjust properties in this critical chromatin region to transcriptional requirements.[31–37] Of particular importance regarding our study is that HS results in rapid depletion of H2A.Z from the +1 nucleosome of temperature-responsive *Arabidopsis* genes.[25–26] In absence of

**Figure 3.** Promoter-proximal enrichment of RNAPII (PPEP) at the +1 nucleosome of HS-induced genes upon TFIISmut expression and depending on transcriptional activity. (A) Schematic representation of the RNAPII transcription unit including promoter-proximal TSS region and rest of the transcribed region. (B) PPEP increases significantly in *tfIIs* related to transcriptional activity, whereas no change is observed in Col-0. PPEP was calculated [PPEP = RNAPII counts (TSS region/gene body) * (x/y)] as described by Antosz et al.[16] using the RNAPII-S2P ChIP-seq data for Col-0 and *tfIIs*. Genes were divided into expression bins according to transcripts per million (TPM) in Col-0 3 h after heat treatment (very high expression > 300 TPM (n = 130); high expression = 30–300 TPM (n = 1648); medium expression = 10–30 TPM (n = 4166) and low expression = 1–10 TPM (n = 9467). Genes with gene body < 100 bp or less than 10 reads in the promoter proximal region or gene body in any sample were omitted. Asterisks indicate significant differences within groups according to Wilcoxon signed-rank test (ns: $P > 0.05$, *$P < 0.05$, ****$P < 0.0001$). (C) PPEP at the +1 nucleosome at HS-induced genes and upon induced expression of TFIISmut in *Arabidopsis tfIIs* plants that inhibits the RNA cleavage activity of RNAPII, resulting in RNAPII accumulation at +1 nucleosomes.[16] RNAPII-S2P coverage based on ChIP-seq data after 3 h HS or untreated (this work) or after induced expression of TFIISmut or mock-induced.[16] RNAPII coverage upon 3 h HS (left panel, red) resembles that upon induced TFIISmut expression (left panel, purple), while that of mock TFIISmut induction (right panel, pink) reflects the situation in unstressed plants (right panel, dark red). The nucleosomal positions at 21 °C based on MNase-seq data are indicated in grey.

TFIIS, RNAPII accumulates predominantly at the +1 nucleosome of HS-induced genes that are depleted in H2A.Z (Figures 2–4), suggesting that TFIIS is required for efficient passage of RNAPII through H2A.Z-depleted +1 nucleosomes, which may be particularly important under conditions of massive transcriptional reprogramming (Figure 1). Despite considerable effort, the consequences of H2A.Z depletion from the +1 nucleosome for RNAPII transcription have not been unequivocally clarified in different eukaryotic systems including plants,[22–24] although some studies report a higher barrier for RNAPII passage,[38–39] thus possibly requiring TFIIS action. Still the functional relationship between H2A.Z and RNAPII transcription may depend on the specific chromatin context including exact nucleosome position, underlying DNA sequence and histone modifications.[40–41] The fact that HS-induced promoter-proximal stalling of RNAPII reflects the accumulation of RNAPII upon expression of TFIISmut,[16] argues for a role of TFIIS stimulated transcript cleavage in passage of the +1 nucleosome at elevated temperatures in plants. This is in agreement with various *in vitro* studies, demonstrating that TFIIS promotes transcription through nucleosomes primarily by reactivating stalled RNAPII rather than preventing pausing/arrest in a cleavage-independent manner.[9–11,42]

In mammals, prevalent RNAPII promoter-proximal pausing occurs in a narrow region between +25 and +50 nt downstream of the TSS.[43] Various genome-wide studies in *Arabidopsis* failed to detect a (closely) related RNAPII pausing mechanism,[16,44–46] which may be related to the absence of the elongation factor NELF in plants[3,47] that contributes to control pausing in mammals.[43] However, recent GRO-seq, NET-seq and RNAPII ChIP-seq analyses are compatible with the option of a nucleosome-defined mechanism for

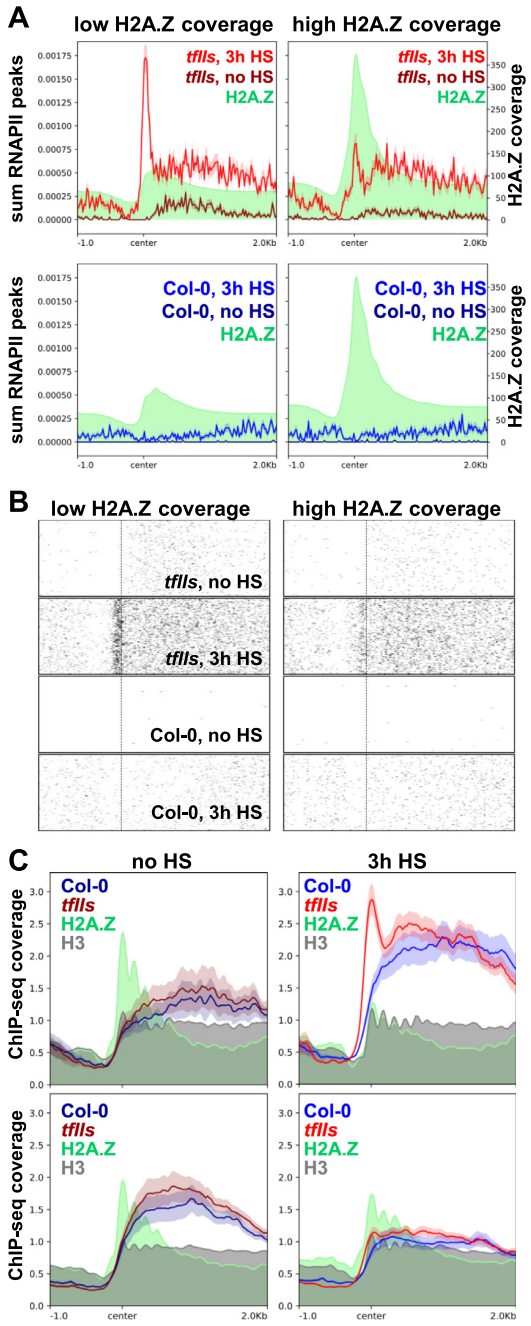



**Figure 4.** Under HS conditions TFIIS is required for RNAPII to transcribe efficiently through the +1 nucleosome of induced genes depleted in the histone variant H2A.Z. (A) RNAPII accumulates at +1 nucleosomes of genes depleted in H2A.Z. Sum of RNAPII-S2P ChIP-seq peaks over genes (n = 10000, each) with low (left) or high (right) H2A.Z ChIP-seq signals. RNAPII distribution is shown in *tfIIs* upon 3 h HS or without (top) and in Col-0 (bottom) along with the H2A.Z coverage according to Wollmann et al.[27] in green. (B) The accumulation of RNAPII in *tfIIs* upon HS at genes with low H2A.Z ChIP-seq signals visualised as a heatmap. (C) In absence of TFIIS, HS-induced H2A.Z depletion is associated with RNAPII accumulation at the +1 nucleosome of genes upregulated upon HS. RNAPII coverage upon 3 h HS at 37 °C or without based on the RNAPII-S2P ChIP-seq data centred to the +1 nucleosome is plotted along with the H2A.Z (green) and H3 coverage (grey) obtained by shifting plants from 17 to 37 °C according to Cortijo et al..[26] Two gene clusters were compared that are either rapidly upregulated (top, n = 92) or repressed (bottom, n = 321) upon HS, respectively. At the rapidly upregulated genes, the decrease of H2A.Z coverage in combination with the rather unchanged H3 coverage after 4 h at 37 °C suggests a specific exchange of H2A.Z for H2A rather than nucleosome eviction.[26] In *tfIIs*, the HS-induced H2A.Z depletion at the rapidly inducible genes is accompanied by RNAPII accumulation at the +1 nucleosome, whereas this does not occur in Col-0 and without HS or at the repressed genes.

promoter-proximal RNAPII stalling that may play a part in regulating transcriptional dynamics in plants.[16,45–46] The +1 nucleosome represents the most notable position of promoter-proximal RNAPII stalling in *Arabidopsis*[16,45] and it remains to be seen whether/how this is mechanistically linked to the production of the recently identified class of short promoter-proximal RNAs.[36] Coincident with our results under HS conditions, NET-seq analysis revealed that exposure of *Arabidopsis* plants to low temperature of 4 °C affected RNAPII stalling at the +1 nucleosome, which may facilitate adapting gene expression in response to temperature changes.[45]

Our data indicates that the RNAPII stalling at the +1 nucleosome in *tfIIs* plants leads to reduced transcriptional processivity that ultimately results in decreased upregulation of various HS-induced genes. In conclusion, our results demonstrate that *Arabidopsis* TFIIS is required for transcriptional reprogramming upon HS by promoting early RNAPII transcript elongation, albeit effects on alternative splicing of heat-regulated transcripts[17] likely contribute to the function of TFIIS in plant heat stress adaption. Future experimentation will clarify, whether TFIIS is involved in the response to other types of environmental cues that are also linked to H2A.Z +1 nucleosomes such as osmotic stress[48] or thermomorphogenesis that at mildly elevated temperatures induces various morphological adaptations to withstand suboptimal temperatures.[49–50]

## Materials and Methods

### Plant cultivation and heat stress (HS) treatment

Seeds of *Arabidopsis thaliana* ecotype Col-0 were sown on solid 0.5x MS medium and after stratification for 48 h at 4 °C in the dark, the plates were transferred to a plant incubator (PolyKlima) with long day settings (16 h light, 110 $\mu mol \cdot m^{-2} \cdot s^{-1}$) at 21 °C and 8 h darkness at 18 °C. T-DNA insertion mutants *tfIIs-1* and *tfIIs-2* were previously described[14] as well as *TFIIS* overexpressing plants.[51] For all molecular analyses the *tfIIs-1* allele was used. For HS plants were transferred to 37 °C for 30 min up to 48 h followed by recovery for 5d under standard conditions. For molecular analyses following HS, plants were harvested without recovery period. For rapid HS, sealed plates were transferred to a water bath at 44 °C for 2.30 or 10 min.[21]

### Antibodies and immunoblotting

For immunoblotting, the following primary and secondary antibodies were used: anti-HSP70 (cat. no. AS08 371, Agrisera) and anti-UAP56[52]; secondary HRP-coupled α-rabbit antibody (cat. no. 12–348, Sigma Aldrich). Total protein extracts[53] were separated by SDS-PAGE and electro-transferred onto a PVDF membrane. For immunoblot analysis[33] SuperSignal R West Pico Chemiluminescent substrate (Thermo Fisher Scientific) was used and chemiluminescence was detected using the ChemiDoc MP system (BioRad).

### Isolation of RNA and cDNA synthesis

Snap-frozen 14-DAS plants were homogenised using a TissueLyserII (Qiagen) and after resuspension of the material in Nuclear Extraction Buffer (10 mM HEPES-NaOH, pH 8; 0.4 M sucrose; 5 mM 2-mercapto-ethanol supplemented with proteinase inhibitor cocktail), nuclei were isolated by centrifugation as previously described.[33,54] RNA was extracted from isolated nuclei using the TRIzol method (Invitrogen) and further purified using the Monarch RNA Cleanup Kit (New England Biolabs). After DNase treatment reverse transcription was performed using 1.5 μg of RNA, random hexameric primers and 200 U Reverse Transcriptase (Thermo Fisher Scientific) as previously described.[55]

### RT-qPCR analyses

Amplification with cDNA as a template was performed using the Kapa SYBR FAST system (ThermoFisher Scientific) and gene-specific primers (Table S1) with a Mastercycler ep realplex 2 (Eppendorf) as previously described.[53–54] The qPCR measurements were analysed using the ΔΔCt method.

### Transcript profiling by RNA-seq

Profiling of nuclear RNAs was performed as previously described.[55] RNA isolated using the TRIzol method (Invitrogen) was further purified and DNase-treated using the Monarch RNA Cleanup Kit (New England Biolabs). Library preparation and RNA-seq were performed at the Genomics Core Facility (University of Regensburg, www.kfb-regensburg.de), employing the following modules: NuGEN Universal Plus RNA-Seq with NuQuant User Guide v3 (Tecan Genomics) in combination with Arabidopsis rRNA AnyDeplete module, and the KAPA Library Quantification Kit-Illumina/ABI Prism (Roche Sequencing Solutions). To judge final library complexities (vs PCR duplicates) unique molecular tags were used.[56] Equimolar amounts of each library (three biological replicates each per genotype and condition) were sequenced on an Illumina NextSeq 2000 instrument controlled by Control Software (NCS, v1.4.0.39521), using a 50 cycle P3 Flow Cell with dual index, paired-end run parameters. Image analysis and base calling were done by the Real Time Analysis Software (RTA, v3.9.2). The resulting '.cbcl' files were converted into '.fastq' files with the bcl2fastq (v2.20) software.

### RNA-seq data analysis

Quality control was performed using FastQC (v0.11.9) and MultiQC (v1.11).[57] After the initial quality assessment, the molecular tag data was recorded in the header for each read via umi_tools (v1.1).[56] Reads with low base quality and adapter contaminations were removed using trimmomatic (v0.39, 'ILLUMINACLIP:NGS_contaminants.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36').[58] The remaining reads were mapped to the TAIR10 genome (https://www.arabidopsis.org/)[59] using STAR (v2.7.9a, '--outFilterType BySJout --outFilterMultimapNmax 20 --alignSJoverhangMin -- alignSJDBoverhangMin 1 --outFilterMismatchNmax 999 --alignIntronMin 10 --alignIntronMax 1000000 --outFilterMismatchNoverReadLmax 0.04 --outSAMmultNmax 1 –outMultimapperOrder Random'). The resulting '.bam' files were filtered to only include alignments with MAPQ score $\geq$10, sorted, and indexed using samtools (v1.3),[60] resulting in 23.0–32.4 million and 17.7–35.9 million uniquely mapped reads for the 3 h- and 10 min-HS, respectively. Finally, the molecular tag data was used in conjunction with the alignments to remove technical duplicates using umi_tools (v1.1). Reproducibility of the biological replicates was examined by principle component analysis (PCA) and the pairwise correlation (Pearson) was calculated (Supplemental Figures 8 and 9). For the differential gene expression analysis, the resulting files from the pipeline outlined above were used to create a count table using the featureCounts function of the rsubread package (v2.4.3),[61] which was then analysed using DESeq2 (v1.30.1)[62] and the tidybulk package (v1.2.1).[63] For differential expression analysis between Col-0 and *tfiis* over all time points the DESeq2 (v1.30.1) likelihood ratio test (LRT) function was used and significantly differential expressed genes (FDR < 0.01) were clustered using unsupervised k-means clustering. GO-term enrichments were performed using Metascape[64] using multiple gene lists with all previously defined cluster. The gene network analysis was performed using cytoscape.[65]

### ChIP sequencing (ChIP-seq)

Chromatin immunoprecipitation (ChIP) was essentially performed as previously described.[16,33] Plants (14 DAS) were crosslinked with formaldehyde and used for isolation of nuclei, before chromatin was sheared using a Bioruptor pico device (Diagenode). Immunoprecipitation was performed using antibodies directed against RNAPII-S2P (abcam, ab5095). For ChIP-seq, libraries were generated using NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs) and the final libraries (three biological replicates each per genotype and condition) were sequenced by the Genomics Core Facility at the University of Regensburg using NextSeq 2000 (Illumina). Reads were aligned to the TAIR10 genome (https://www.arabidopsis.org/) using Bowtie2[66] and coverage tracks were calculated with deeptools "bamCoverage". Downstream analysis was mainly performed using the deepTools2 suite (version 3.5.0)[67] and quality control was performed at several steps using FastQC.[57] 12.3–19.0 million reads per sample mapped against the TAIR10 genome and regions with aberrant coverage or low sequence complexity were removed.[68] After confirming high pairwise correlations, the biological replicates were merged and CPM normalised (Supplemental Figures 10 and 11). Macs3 (version 3.0.0a7) "callpeak" function was used to call peaks in *tfIIs* over Col-0 as control. The summits of the resulting peak files were converted to .bw tracks using deepTools bedGraphToBigWig. Expression levels of genes (divided into quartiles according to transcripts per million) was deduced from the RNA-sequencing data of Col-0 under control conditions.

### Calculation of promoter proximal enrichment of RNAPII (PPEP)

PPEP was calculated as previously described[16] using the following function: PPEP = RNAPII counts (TSS region/ gene body) * (x/y) with the TSS region defined as the region from the TSS to 500 bp downstream and the gene body defined as the region 500 bp downstream the TSS to the TES. X is defined as the length of the TSS region and y being the length of the gene body. The change in promoter proximal pausing after heat shock was calculated as: $\Delta$PPEP = PPEP (3 h HS) - PPEP (no HS).

### MNase-seq

Plants (14 DAS, at 21 °C) were fixed using formaldehyde followed by isolation of nuclei as described.[16,33] Isolated nuclei were washed twice in MNase Buffer (20 mM Tris/HCl pH 7.5, 300 mM sucrose, 3 mM $CaCl_2$, supplemented with proteinase inhibitor mix) and resuspended in MNase Buffer. Nuclei were split into 3 samples and MNase digestion was performed using MNase (Sigma) at 37 °C for different times followed by immunoprecipitation with an anti-H3 antibody (cat. no. ab1791, Abcam) as described.[16,33]

Libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs) using 7 PCR cycles for library amplification. Sequencing was performed at the Genomics Core Facility at the University of Regensburg. In brief, the libraries were bead purified with Agencourt AMPure XP magnetic beads (Beckman Coulter) and quantified using the KAPA Library Quantification Kit (Roche). Equimolar amounts of each library (three replicates) were sequenced using an Illumina NextSeq 2000 System controlled by the Control Software (NCS, v1.4.0.39521), using one 100 cycles P3 Flow Cell with the single index, paired-

end run parameters. Image analysis and base calling were done by the Real Time Analysis Software (RTA, v3.9.2). The resulting '.cbcl' files were converted into '.fastq' files with the bcl2fastq (v2.20) software.

### MNase-seq data analysis

The quality of the raw reads and the following steps was monitored using FastQC (v0.11.8) and MultiQC (v1.9), as well as statistics created with samtools view (v1.10) and picard CollectInsertSizeMetrics (v2.25.6). Reads with low base quality and adapter contaminations were removed using trimmomatic in paired-end mode (v0.38, 'ILLUMINACLIP:NGS_contaminants.fa:2:30:10:2:keepBothReads LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36′). Remaining paired reads were then mapped to the TAIR10 genome using bowtie2 in paired-end mode (v2.4.2, '--local -t --very-sensitive-local --no-discordant'). The resulting '.sam' files were converted to '.bam' files, sorted, and indexed using samtools (v1.10). The alignments were then filtered according to insert size using deeptools alignmentSieve (v3.1.3). Alignments with a mapping quality $\geq$30, a minimum insert length of 140, and a maximum insert length of 200 were defined as of mono-nucleosomal origin, resulting in 105.8–141.7 million uniquely mapped reads per replicate. Alignments that mapped to regions blacklisted as described,[68] as well as alignments that mapped to the mitochondrial and plastid sequences were discarded. Three replicates were merged and mono-nucleosomal positions were called using DANPOS (v3.0) using the parameters 'dpos --extend 70 -c 119481543 -u 0 -z 1 -a 1 -e 1′. +1 nucleosomal positions were extracted as the first nucleosomal position after TAIR10 annotated transcription start sites with a minimal overlap of 140 bp.

## CRediT authorship contribution statement

**Simon Obermeyer:** Investigation, Formal analysis, Validation, Visualization, Writing – review & editing. **Richard Stöckl:** Investigation, Formal analysis, Validation, Visualization. **Tobias Schnekenburger:** Investigation. **Henna Kapoor:** Investigation. **Thomas Stempfl:** Investigation, Data curation. **Uwe Schwartz:** Investigation, Data curation. **Klaus D. Grasser:** Conceptualization, Funding acquisition, Supervision, Writing – original draft.

## DATA AVAILABILITY

The RNA-seq, ChIP-seq and MNase-seq data of this study has been deposited in the NCBI Sequence Read Archive and are accessible through the accession code PRJNA877815.

## Acknowledgement

We thank Isabel Bäurle and coworkers for valuable advice regarding the establishment of rapid heat stress treatments. This work was presented at the European Workshop on Plant Chromatin in Prague, Czech Republic, May 2022.

## Conflict of Interest

The authors declare no conflicts of interests.

## Funding

This research was supported by the German Research Foundation (DFG) through grants Gr1159/14-2 and SFB960/A6 to KDG.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2022.167917.

*Received 31 October 2022;*
*Accepted 4 December 2022;*
*Available online 9 December 2022*

***Keywords***:
*Arabidopsis*;
Chromatin;
Histone;
H2A.Z;
Nucleosome

† These authors contributed equally: Simon Obermeyer, Richard Stöckl.

## References

1. Chen, F.X., Smith, E.R., Shilatifard, A., (2018). Born to run: control of transcription elongation by RNA polymerase II. *Nature Rev. Mol. Cell Biol.* **19**, 464–478.
2. Kwak, H., Lis, J.T., (2013). Control of transcriptional elongation. *Ann. Rev. Genet.* **47**, 483–508.
3. van Lijsebettens, M., Grasser, K.D., (2014). Transcript elongation factors: shaping transcriptomes after transcript initiation. *Trends Plant Sci.* **19**, 717–726.
4. Fish, R.N., Kane, C.M., (2002). Promoting elongation with transcript cleavage stimulatory factors. *Biochim. Biophys. Acta* **1577**, 287–307.
5. Noe Gonzalez, M., Blears, D., Svejstrup, J.Q., (2021). Causes and consequences of RNA polymerase II stalling during transcript elongation. *Nature Rev. Mol. Cell Biol.* **22**, 3–21.

6. Kettenberger, H., Armache, K.-J., Cramer, P., (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* **114**, 347–357.
7. Wang, D., Bushnell, D.A., Huang, X., Westover, K.D., Levitt, M., Kornberg, R.D., (2009). Structural basis of transcription: backtracked RNA Polymerase II at 3.4 angstrom resolution. *Science* **324**, 1203–1206.
8. Cheung, A.C.M., Cramer, P., (2011). Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature* **471**, 249–253.
9. Farnung, L., Ochmann, M., Garg, G., Vos, S.M., Cramer, P., (2022). Structure of a backtracked hexasomal intermediate of nucleosome transcription. *Mol. Cell* **82**, 3126–3134.e7.
10. Kireeva, M.L., Hancock, B., Cremona, G.H., Walter, W., Studitsky, V.M., Kashlev, M., (2005). Nature of nucleosomal barrier to RNA polymerase II. *Mol. Cell* **18**, 108.
11. Nock, A., Ascano, J.M., Barrero, M.J., Malik, S., (2012). Mediator-regulated transcription through the +1 nucleosome. *Mol. Cell* **48**, 837–848.
12. Churchman, L.S., Weissman, J.S., (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469**, 368–373.
13. Nakanishi, T., Nakano, A., Nomura, K., Sekimizu, K., Natori, S., (1992). Purification, gene cloning, and gene disruption of the transcription elongation factor S-II in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **267**, 13200–13204.
14. Grasser, M., Kane, C.M., Merkle, T., Melzer, M., Emmersen, J., Grasser, K.D., (2009). Transcript elongation factor TFIIS is involved in *Arabidopsis* seed dormancy. *J. Mol. Biol.* **386**, 598–611.
15. Sigurdsson, S., Dirac-Svejstrup, A.B., Svejstrup, J.Q., (2010). Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. *Mol. Cell* **38**, 202–210.
16. Antosz, W., Deforges, J., Begcy, K., Bruckmann, A., Poirier, Y., Dresselhaus, T., Grasser, K.D., (2020). Critical Role of Transcript Cleavage in Arabidopsis RNA Polymerase II Transcriptional Elongation. *Plant Cell* **32**, 1449–1463.
17. Szádeczky-Kardoss, I., Szaker, H.M., Verma, R., Darkó, É., Pettkó-Szandtner, A., Silhavy, D., Csorba, T., (2022). Elongation factor TFIIS is essential for heat stress adaptation in plants. *Nucleic Acids Res.* **50**, 1927–1950.
18. Dhaliwal, N.K., Mitchell, J.A., (2016). Nuclear RNA Isolation and Sequencing. *Methods Mol. Biol.* **1402**, 63–71.
19. Zaghlool, A., Ameur, A., Nyberg, L., Halvardson, J., Grabherr, M., Cavelier, L., Feuk, L., (2013). Efficient cellular fractionation improves RNA sequencing analysis of mature and nascent transcripts from human tissues. *BMC Biotechnol.* **13**, 99.
20. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., et al., (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2**, 100141
21. Stief, A., Altmann, S., Hoffmann, K., Pant, B.D., Scheible, W.-R., Bäurle, I., (2014). Arabidopsis miR156 Regulates Tolerance to Recurring Environmental Stress through SPL Transcription Factors. *Plant Cell* **26**, 1792–1807.
22. Colino-Sanguino, Y., Clark, S.J., Valdes-Mora, F., (2022). The H2A.Z-nuclesome code in mammals: emerging functions. *Trends Genet.* **38**, 273–289.
23. Jarillo, J.A., Piñeiro, M., (2015). H2A.Z mediates different aspects of chromatin function and modulates flowering responses in *Arabidopsis*. *Plant J.* **83**, 96–109.
24. Lei, B., Berger, F., (2020). H2A Variants in Arabidopsis: Versatile Regulators of Genome Activity. *Plant communications* **1**, 100015
25. Kumar, S.V., Wigge, P.A., (2010). H2A.Z-containing nucleosomes mediate the thermosensory response in *Arabidopsis*. *Cell* **140**, 136–147.
26. Cortijo, S., Charoensawan, V., Brestovitsky, A., Buning, R., Ravarani, C., Rhodes, D., van Noort, J., Jaeger, K.E., et al., (2017). Transcriptional Regulation of the Ambient Temperature Response by H2A.Z Nucleosomes and HSF1 Transcription Factors in Arabidopsis. *Mol. Plant* **10**, 1258–1273.
27. Wollmann, H., Stroud, H., Yelagandula, R., Tarutani, Y., Jiang, D., Jing, L., Jamge, B., Takeuchi, H., et al., (2017). The histone H3 variant H3.3 regulates gene body DNA methylation in Arabidopsis thaliana. *Genome Biol.* **18**, 94.
28. Adelman, K., Marr, M.T., Werner, J., Saunders, A., Ni, Z., Andrulis, E.D., Lis, J.T., (2005). Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS. *Mol. Cell* **17**, 103–112.
29. Pokholok, D.K., Hannett, N.M., Young, R.A., (2002). Exchange of RNA Polymerase II Initiation and Elongation Factors during Gene Expression In Vivo. *Mol. Cell* **9**, 799–809.
30. Kujirai, T., Kurumizaka, H., (2020). Transcription through the nucleosome. *Curr. Opin. Struct. Biol.* **61**, 42–49.
31. Bieluszewski, T., Sura, W., Dziegielewski, W., Bieluszewska, A., Lachance, C., Kabza, M., Szymanska-Lejman, M., Abram, M., et al., (2022). NuA4 and H2A.Z control environmental responses and autotrophic growth in Arabidopsis. *Nature Commun.* **13**, 277.
32. Diego-Martin, B., Pérez-Alemany, J., Candela-Ferre, J., Corbalán-Acedo, A., Pereyra, J., Alabadí, D., Jami-Alahmadi, Y., Wohlschlegel, J., et al., (2022). The TRIPLE PHD FINGERS proteins are required for SWI/SNF complex-mediated +1 nucleosome positioning and transcription start site determination in Arabidopsis. *Nucleic Acids Res.* **50**, 10399–10417.
33. Michl-Holzinger, P., Obermeyer, S., Markusch, H., Pfab, A., Ettner, A., Bruckmann, A., Babl, S., Längst, G., et al., (2022). Phosphorylation of the FACT histone chaperone subunit SPT16 affects chromatin at RNA polymerase II transcriptional start sites in *Arabidopsis*. *Nucleic Acids Res.* **50**, 5014–5028.
34. Potok, M.E., Wang, Y., Xu, L., Zhong, Z., Liu, W., Feng, S., Naranbaatar, B., Rayatpisheh, S., et al., (2019). Arabidopsis SWR1-associated protein methyl-CpG-binding domain 9 is required for histone H2A.Z deposition. *Nature Commun.* **10**, 3352.
35. Torres, E.S., Deal, R.B., (2019). The histone variant H2A.Z and chromatin remodeler BRAHMA act coordinately and antagonistically to regulate transcription and nucleosome dynamics in Arabidopsis. *Plant J.* **99**, 144–162.
36. Thomas, Q.A., Ard, R., Liu, J., Li, B., Wang, J., Pelechano, V., Marquardt, S., (2020). Transcript isoform sequencing reveals widespread promoter-proximal transcriptional termination in Arabidopsis. *Nature Commun.* **11**, 2589.
37. Wang, Y., Zhong, Z., Zhang, Y., Xu, L., Feng, S., Rayatpisheh, S., Wohlschlegel, J.A., Wang, Z., et al., (2020). NAP1-RELATED PROTEIN1 and 2 negatively regulate H2A.Z abundance in chromatin in Arabidopsis. *Nature Commun.* **11**, 2887.

38. Rudnizky, S., Bavly, A., Malik, O., Pnueli, L., Melamed, P., Kaplan, A., (2016). H2A.Z controls the stability and mobility of nucleosomes to regulate expression of the LH genes. *Nature Commun.* **7**, 12958.
39. Weber, C.M., Ramachandran, S., Henikoff, S., (2014). Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol. Cell* **53**, 819–830.
40. Cole, L., Kurscheid, S., Nekrasov, M., Domaschenz, R., Vera, D.L., Dennis, J.H., Tremethick, D.J., (2021). Multiple roles of H2A.Z in regulating promoter chromatin architecture in human cells. *Nature Commun.* **12**, 2524.
41. Chen, Z., Gabizon, R., Brown, A.I., Lee, A., Song, A., Díaz-Celis, C., Kaplan, C.D., Koslover, E.F., et al., (2019). High-resolution and high-accuracy topographic and transcriptional maps of the nucleosome barrier. *Elife* **8**
42. Ishibashi, T., Dangkulwanich, M., Coello, Y., Lionberger, T. A., Lubkowska, L., Ponticelli, A.S., Kashlev, M., Bustamante, C., (2014). Transcription factors IIS and IIF enhance transcription efficiency by differentially modifying RNA polymerase pausing dynamics. *Proc. Natl. Acad. Sci. USA* **111**, 3419–3424.
43. Core, L., Adelman, K., (2019). Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev.* **33**, 960–982.
44. Hetzel, J., Duttke, S.H., Benner, C., Chory, J., (2016). Nascent RNA sequencing reveals distinct features in plant transcription. *Proc. Natl. Acad. Sci. USA* **113**, 12316–12321.
45. Kindgren, P., Ivanov, M., Marquardt, S., (2020). Native elongation transcript sequencing reveals temperature dependent dynamics of nascent RNAPII transcription in Arabidopsis. *Nucleic Acids Res.* **48**, 2332–2347.
46. Zhu, J., Liu, M., Liu, X., Dong, Z., (2018). RNA polymerase II activity revealed by GRO-seq and pNET-seq in *Arabidopsis*. *Nature Plants* **4**, 1112–1123.
47. Hartzog, G.A., Fu, J., (2013). The Spt4-Spt5 complex: a multi-faceted regulator of transcription elongation. *Biochim. Biophys. Acta* **1829**, 105–115.
48. Sura, W., Kabza, M., Karlowski, W.M., Bieluszewski, T., Kus-Slowinska, M., Pawełoszek, Ł., Sadowski, J., Ziolkowski, P.A., (2017). Dual Role of the Histone Variant H2A.Z in Transcriptional Regulation of Stress-Response Genes. *Plant Cell* **29**, 791–807.
49. van der Woude, L.C., Perrella, G., Snoek, B.L., van Hoogdalem, M., Novák, O., van Verk, M.C., van Kooten, H.N., Zorn, L.E., et al., (2019). HISTONE DEACETYLASE 9 stimulates auxin-dependent thermomorphogenesis in Arabidopsis thaliana by mediating H2A.Z depletion. *Proc. Natl. Acad. Sci. USA* **116**, 25343–25354.
50. Xue, M., Zhang, H., Zhao, F., Zhao, T., Li, H., Jiang, D., (2021). The INO80 chromatin remodeling complex promotes thermomorphogenesis by connecting H2A.Z eviction and active transcription in Arabidopsis. *Mol. Plant* **14**, 1799–1813.
51. Mortensen, S.A., Grasser, K.D., (2014). The seed dormancy defect of *Arabidopsis* mutants lacking the transcript elongation factor TFIIS is caused by reduced expression of the *DOG1* gene. *FEBS Letter* **588**, 47–51.
52. Kammel, C., Thomaier, M., Sørensen, B.B., Schubert, T., Längst, G., Grasser, M., Grasser, K.D., (2013). *Arabidopsis* DEAD-box RNA helicase UAP56 interacts with both RNA and DNA as well as with mRNA export factors. *PLoS One* **8**, e60644.
53. Antosz, W., Pfab, A., Ehrnsberger, H.F., Holzinger, P., Köllen, K., Mortensen, S.A., Bruckmann, A., Schubert, T., 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.
54. Pfab, A., Bruckmann, A., Nazet, J., Merkl, R., Grasser, K. D., (2018). 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.
55. Obermeyer, S., Stöckl, R., Schnekenburger, T., Moehle, C., Schwartz, U., Grasser, K.D., (2022). Distinct role of subunits of the Arabidopsis RNA polymerase II elongation factor PAF1C in transcriptional reprogramming. *Front. Plant Sci.* **13**, 974625
56. Smith, T., Heger, A., Sudbery, I., (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499.
57. Ewels, P., Magnusson, M., Lundin, S., Käller, M., (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048.
58. Bolger, A.M., Lohse, M., Usadel, B., (2014). Trimmomatic:a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.
59. Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., et al., (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* **40**, D1202–D1210.
60. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., et al., (2021). Twelve years of SAMtools and BCFtools. *GigaScience* **10**
61. Liao, Y., Smyth, G.K., Shi, W., (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* **47**, e47.
62. Love, M.L., Huber, W., Anders, S., (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.
63. Mangiola, S., Molania, R., Dong, R., Doyle, M.A., Papenfuss, A.T., (2021). tidybulk: an R tidy framework for modular transcriptomic data analysis. *Genome Biol.* **22**, 42.
64. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., Chanda, S.K., (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Commun.* **10**, 1523.
65. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., et al., (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504.
66. Langmead, B., Salzberg, S.L., (2012). Fast gapped-read alignment with Bowtie 2. *Nature Meth.* **9**, 357–359.
67. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., et al., (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165.
68. Quadrana, L., Bortolini Silveira, A., Mayhew, G.F., LeBlanc, C., Martienssen, R.A., Jeddeloh, J.A., Colot, V., (2016). The Arabidopsis thaliana mobilome and its impact at the species level. *Elife* **5**, e15716.

# Chapter 4

# Distinct role of subunits of the Arabidopsis RNA polymerase II elongation factor PAF1C in transcriptional reprogramming

This peer-reviewed article was published in the journal *Frontiers in Plant Science* in 2022

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 29 September 2022
DOI 10.3389/fpls.2022.974625



OPEN ACCESS

EDITED BY
Sara Farrona,
National University of Ireland
Galway, Ireland

REVIEWED BY
Rafal Archacki,
University of Warsaw, Poland
Inna Lermontova,
Leibniz Institute of Plant Genetics and
Crop Plant Research (IPK), Germany

*CORRESPONDENCE
Klaus D. Grasser
Klaus.Grasser@ur.de

SPECIALTY SECTION
This article was submitted to
Plant Genetics, Epigenetics and
Chromosome Biology,
a section of the journal
Frontiers in Plant Science

RECEIVED 21 June 2022
ACCEPTED 05 September 2022
PUBLISHED 29 September 2022

CITATION
Obermeyer S, Stöckl R,
Schnekenburger T, Moehle C,
Schwartz U and Grasser KD (2022)
Distinct role of subunits of the
*Arabidopsis* RNA polymerase II
elongation factor PAF1C in
transcriptional reprogramming.
*Front. Plant Sci.* 13:974625.
doi: 10.3389/fpls.2022.974625



# Distinct role of subunits of the *Arabidopsis* RNA polymerase II elongation factor PAF1C in transcriptional reprogramming

Simon Obermeyer[1], Richard Stöckl[1], Tobias Schnekenburger[1], Christoph Moehle[2], Uwe Schwartz[3] and Klaus D. Grasser[1]*

[1]Cell Biology & Plant Biochemistry, Biochemistry Centre, University of Regensburg, Regensburg, Germany, [2]Center of Excellence for Fluorescent Bioanalytics (KFB), University of Regensburg, Regensburg, Germany, [3]NGS Analysis Centre, Biology and Pre-Clinical Medicine, University of Regensburg, Regensburg, Germany

Transcript elongation by RNA polymerase II (RNAPII) is dynamic and highly regulated, thereby contributing to the implementation of gene expression programs during plant development or in response to environmental cues. The heterohexameric polymerase-associated factor 1 complex (PAF1C) stabilizes the RNAPII elongation complex promoting efficient transcript synthesis. In addition, PAF1C links transcriptional elongation with various post-translational histone modifications at transcribed loci. We have exposed *Arabidopsis* mutants deficient in the PAF1C subunits ELF7 or CDC73 to elevated NaCl concentrations to provoke a transcriptional response. The growth of *elf7* plants was reduced relative to that of wildtype under these challenging conditions, whereas *cdc73* plants exhibited rather enhanced tolerance. Profiling of the transcriptional changes upon NaCl exposure revealed that *cdc73* responded similar to wildtype. Relative to wildtype and *cdc73*, the transcriptional response of *elf7* plants was severely reduced in accord with their greater susceptibility to NaCl. The data also imply that CDC73 is more relevant for the transcription of longer genes. Despite the fact that both ELF7 and CDC73 are part of PAF1C the strikingly different transcriptional response of the mutants upon NaCl exposure suggests that the subunits have (partially) specific functions.

KEYWORDS

***Arabidopsis thaliana*, chromatin, histone modifications, PAF1C, RNA polymerase II, transcript elongation**

## Introduction

In eukaryotic cells, hundreds of proteins regulate the transcription of genes by RNA polymerase II (RNAPII). The majority of these proteins control transcriptional initiation *via* interaction with promoter regions. Functionally distinct so-called transcript elongation factors (TEFs) modulate the efficiency of mRNA synthesis by different mechanisms after the initiation stage. Some TEFs act as histone modifiers, histone chaperones or chromatin remodelers, manipulating properties of the nucleosomal template, whereas others adjust the catalytic activity of RNAPII directly (Sims et al., 2004; Kwak and Lis, 2013; Chen et al., 2018). A remarkable example of a regulator of transcriptional elongation by RNAPII is the multifunctional polymerase-associated factor 1 complex (PAF1C) that originally was discovered and characterized in *Saccharomyces cerevisiae* (Wade et al., 1996; Mueller and Jaehning, 2002). Yeast PAF1C is composed of five subunits Paf1, Ctr9, Leo1, Rtf1 and Cdc73 and it interacts directly with elongating RNAPII (Jaehning, 2010). In mammals, PAF1C contains an additional subunit WDR61 (Ski8), while RTF1 appears to be less stably associated (Francette et al., 2021). Recent cryo-electron microscopy studies revealed detailed structural insight into the architecture and molecular interactions of elongation competent RNAPII associated with PAF1C and other TEFs (Vos et al., 2018; Vos et al., 2020). The structural information suggests that association of PAF1C with RNAPII allosterically contributes to the stabilization of the elongation complex, which is in line with its stimulation of productive transcript synthesis (Rondón et al., 2004; Hou et al., 2019). Moreover, PAF1C links transcript elongation with various post-translational histone modifications over transcribed regions, including H2B mono-ubiquitination (Wood et al., 2003; Xiao et al., 2005), H3K4me2/3 and H3K79me2/3 (Krogan et al., 2003; Ng et al., 2003; Wood et al., 2003), as well as H3K36me3 (Chu et al., 2007). Thus, PAF1C acts at the interface of transcription and chromatin, modulating the progression of RNAPII during elongation.

In plants, PAF1C is conserved and was initially recognized because of its role in the transition from vegetative to reproductive development (van Lijsebettens and Grasser, 2014). The PAF1C subunits were named after the *Arabidopsis* mutants, whose analysis for their developmental phenotype led to identification of the genes encoding the respective subunits (He et al., 2004; Oh et al., 2004). Accordingly, most of the subunits based on the mutants were termed early flowering (*elf*) and/or vernalization independence (*vip*): PAF1 (At-ELF7), CTR9 (At-ELF8, At-VIP6), LEO1 (At-VIP4), RTF1 (At-VIP5), WDR61/SKI8 (At-VIP3) and CDC73 (At-CDC73, At-PHP) (He et al., 2004; Oh et al., 2004; Park et al., 2010). Co-immunoprecipitation and affinity-purification in combination with mass spectrometry analyses revealed that in common with mammals, *Arabidopsis* PAF1C consists of these six subunits (Oh et al., 2004; Antosz et al., 2017). *Arabidopsis* mutants deficient in PAF1C subunits VIP3, VIP4, VIP5, VIP6/ELF8 and ELF7 exhibit marked early flowering phenotypes that are commonly associated by reduced expression of the floral repressor *FLC* (and paralogs) (Zhang and van Nocker, 2002; Zhang et al., 2003; He et al., 2004; Oh et al., 2004). Moreover, these mutants are significantly smaller than wild type and often exhibit defects in inflorescence and flower morphology (Zhang and van Nocker, 2002; Zhang et al., 2003; He et al., 2004; Fal et al., 2017). Mutants defective in the CDC73/PHP subunit share the early flowering phenotype (and reduced *FLC* expression), but are of wild type size and display normal flower development (Park et al., 2010; Yu and Michaels, 2010). The decreased expression of *FLC* in mutants lacking PAF1C subunits is likely mediated by altered levels and distribution of histone H3 methylation marks at the *FLC* locus (e.g. H3K4me3, H3K27me3, H3K36me2) (He et al., 2004; Oh et al., 2008; Xu et al., 2008; Park et al., 2010; Yu and Michaels, 2010). PAF1C also modulates the temperature-responsive transition to flowering (Nasim et al., 2022). Beyond that PAF1C is required for plant response to repeated touch stimuli, as *Arabidopsis* mutants deficient in VIP3, VIP5 and VIP6 did not react to the mechanical stimulation as the wild type. VIP3 proved necessary for the touch-induced upregulation of the *TCH3* and *TCH4* mRNAs, a characteristic feature of the response to mechanical stimulation (Jensen et al., 2017).

The diversity of phenotypes associated with deficiency of PAF1C subunits in various organisms can be attributed to the misexpression of target genes and downstream processes. In yeast and mammals, there is accumulating evidence that inactivation of different PAF1C subunit genes to some extent causes distinct phenotypes, arguing for subunit specificity (Francette et al., 2021). This aspect has not been addressed in plants under transcriptional challenging conditions. Since chromatin-based reprogramming in reaction to environmental cues is crucial for plant growth (Kim, 2021), we studied here the transcriptional response of *Arabidopsis* PAF1C subunit mutants upon exposure to elevated NaCl concentrations.

## Materials and methods

### Plant cultivation and documentation

Seeds of *Arabidopsis thaliana* Col-0 were sown on solid 0.5x MS medium (Murashige and Skoog, 1962) and after stratification for 48h at 4°C in the dark, the plates were transferred to a plant incubator (PolyKlima) with long day settings (16h light (110 $\mu mol \cdot m^{-2} \cdot s^{-1}$) at 21°C and 8h darkness at 18°C. In some experiments the 0.5x MS medium was supplemented with 50 or 100 mM NaCl and for root growth assays plants were grown on vertically oriented 0.5x MS plates. Plant phenotypes including root growth were documented as previously described (Dürr et al., 2014; Antosz et al., 2020). Seeds

of T-DNA insertion lines were obtained from the Arabidopsis stock centre (http://arabidopsis.info/): *elf7-2*, *elf7-3*, *elf8-1* (He et al., 2004); *cdc73-1*, *cdc73-2* (Park et al., 2010; Yu and Michaels, 2010) and GK_270G12 harboring a T-DNA insertion in exon 14 of *ELF8* (At2g06210) (Kleinboelting et al., 2012) termed *elf8-4*. Plant genotypes were examined by PCR analysis of genomic DNA isolated from leaves as previously described (Dürr et al., 2014; Antosz et al., 2020) using gene- and insertion-specific primers (Table S1).

## Isolation of RNA and cDNA synthesis

Nuclei were isolated from aerial parts of 7 days after stratification (DAS) plants as previously described (Pfab et al., 2017) and analyzed by fluorescence microscopy and immunoblotting as previously described (Antosz et al., 2017; Pfab et al., 2018). RNA was extracted from frozen nuclei using the TRIzol method (Invitrogen). After DNase treatment reverse transcription was performed using 1.5 µg of RNA, random hexameric primers and 200 U Reverse Transcriptase (Thermo Fisher Scientific) as previously described (Pfaff et al., 2018).

## RT-qPCR analyses

Amplification with cDNA as a template was performed using the Kapa SYBR FAST system (ThermoFisher Scientific) and gene-specific primers (Table S1) with a Mastercycler ep realplex 2 (Eppendorf) as previously described (Antosz et al., 2017). The qPCR measurements were analyzed using the ΔΔCt method implemented with the 'pcr' package (Ahmed and Kim, 2018).

## Transcript profiling by RNA-seq

Nuclear RNA isolated using the TRIzol method (Invitrogen) was further purified and DNase-treated using the Monarch RNA Cleanup Kit (New England Biolabs). Library preparation and RNA-seq were performed at the Genomics Core Facility (University of Regensburg, www.kfb-regensburg.de), employing the following modules: NuGEN Universal Plus RNA-Seq with NuQuant User Guide v3 (Tecan Genomics) in combination with Arabidopsis rRNA AnyDeplete module, the Illumina NextSeq 2000 System (Illumina), and the KAPA Library Quantification Kit-Illumina/ABI Prism (Roche Sequencing Solutions). To judge final library complexities (*vs.* PCR duplicates) unique molecular tags were used (Smith et al., 2017). Equimolar amounts of each library were sequenced on an Illumina NextSeq 2000 instrument controlled by the NextSeq 2000 Control Software (NCS, v1.4.0.39521), using 50 cycles P3 Flow Cell with the dual index, paired-end run parameters. Image analysis and base calling were done by the Real Time Analysis Software (RTA, v3.9.2). The resulting '.cbcl' files were converted into '.fastq' files with the bcl2fastq (v2.20) software.

## RNA-seq data analysis

Quality control was performed using fastQC (v0.11.9) and multiQC (v1.11) (Ewels et al., 2016). After the initial quality assessment, the molecular tag data was recorded in the header for each read *via* umi_tools (v1.1) (Smith et al., 2017). Reads with low base quality and adapter contaminations were removed using trimmomatic (v0.39, 'ILLUMINACLIP : NGS_contaminants. fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36') (Bolger et al., 2014). The remaining reads were mapped to the TAIR10 genome (Lamesch et al., 2012) using STAR (v2.7.9a, '–outFilterType BySJout –outFilterMultimapNmax 20 –alignSJoverhangMin 8 –alignSJDBoverhangMin 1

–outFilterMismatchNmax 999 –alignIntronMin 10 –alignIntronMax 1000000

–outFilterMismatchNoverReadLmax 0.04 –outSAMmultNmax 1 –outMultimapperOrder Random'). The resulting '.bam' files were filtered to only include alignments with MAPQ score ≥ 10, sorted, and indexed using samtools (v1.3) (Danecek et al., 2021). Finally, the molecular tag data was used in conjunction with the alignments to remove technical duplicates using umi_tools (v1.1).

For the differential gene expression analysis, the resulting files from the pipeline outlined above were used to create a count table using the featureCounts function of the rsubread package (v2.4.3) (Liao et al., 2019), which was then analyzed using DESeq2 (v1.30.1) (Love et al., 2014) and the tidybulk package (v1.2.1) (Mangiola et al., 2021).

Differential splicing analysis was performed using DEXSeq v1.40 (Anders et al., 2012; Reyes et al., 2013). First, the exon annotation was prepared using the provided python script dexseq_prepare_annotation.py with the option '-r no'. Next, the read coverage of each exon was calculated using the dexseq_count.py script. Finally, salt stress-induced alternative splicing of either Col-0 or *cdc73-2* samples was tested using default settings and an FDR of 0.05.

## ChIP sequencing

Chromatin immunoprecipitation (ChIP) was essentially performed as previously described (Antosz et al., 2020; Michl-Holzinger et al., 2022). Plants (14-DAS *in vitro* grown) were crosslinked with formaldehyde and used for isolation of nuclei, before chromatin was sheared using a Bioruptor pico device (Diagenode). Immunoprecipitation was performed using antibodies directed against RNAPII-S2P (abcam, ab5095). For ChIP-seq, libraries were generated using NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs) and the

final libraries (3 replicates each) were sequenced by the Genomics Core Facility at the University of Regensburg using NextSeq 2000 (Illumina). Reads were aligned to the TAIR10 genome (https://www.arabidopsis.org/) using Bowtie2 (Langmead and Salzberg, 2012) and coverage tracks were calculated with deeptools "bamCoverage". Downstream analysis was mainly performed using the deepTools2 suite (version 3.5.0) (Ramírez et al., 2016) and quality control was performed at several steps using FastQC (Ewels et al., 2016). Reads were mapped against the TAIR10 genome and regions with aberrant coverage or low sequence complexity were removed (Quadrana et al., 2016). After confirming high pairwise correlations, the biological replicates were merged and CPM normalized.

# Results

## Various PAF1C subunit mutants respond differently when exposed to elevated NaCl concentrations

Based on preliminary tests with different environmental stress conditions we focused on elevated NaCl concentrations to provoke transcriptional reprogramming in PAF1C subunit mutants. Initially, seeds of the T-DNA insertion lines *elf7-2*, *elf7-3*, *elf8-1, elf8-4, cdc73-1*, *cdc73-2* and of the wild type of Col-0 were sown on MS medium containing 100 mM NaCl and on MS control medium. In presence of 100 mM NaCl after 14d all genotypes showed reduced growth, when compared with plants grown on control medium (Figure 1A). Compared to Col-0, the *elf7* and *elf8* mutants were affected to a greater extent, while the *cdc73* mutants grew rather better than Col-0. These observations were reassessed with plants grown on control medium for seven days, before transfer onto control medium or onto medium containing 100 mM NaCl. The growth of the transferred plants confirmed the previous finding that *elf7-2/3* and *elf8-1/4* plants were more strongly affected by NaCl than Col-0, while *cdc73-1/2* were affected less than Col-0 (Figure 1B). The aerial parts of the different genotypes were quantified after growth in presence or absence of 100 mM NaCl. Only minor differences between the genotypes were observed on control medium regarding rosette diameter and fresh weight (Figures 1C, D). However, when grown in presence of 100 mM NaCl, rosette diameter and fresh weight of *elf7* and *elf8* plants were significantly reduced, while *cdc73* were somewhat bigger than Col-0. Furthermore, the relative growth of the primary root of the different genotypes was monitored on vertically oriented plates, comparing root elongation in presence of 100 mM NaCl with that of the same genotype on control medium (Supplementary Figure S1). Quantification of root length revealed that the roots of Col-0 and *cdc73* grew comparably, whereas the relative root growth of *elf7* and *elf8* plants was clearly decreased (Figure 1E). Thus, compared to the Col-0 wild type, exposure to 100 mM NaCl has a profound negative effect on the growth of *elf7* and *elf8*, while the lack of CDC73 apparently has no adverse effect on the tolerance to NaCl.

To figure out the time it takes in our experimental setup until transcriptional induction is traceable, we determined the levels of the *LTP4* mRNA that is strongly up-regulated in *Arabidopsis* plants in response to NaCl (Winter et al., 2007). Using RT-qPCR, the *LTP4* transcript was quantified (relative to transcripts of three housekeeping genes) at different times after transfer of Col-0 plants onto medium containing 100 mM NaCl or control medium. Consistent with the induction kinetics of other studies (Kilian et al., 2007), the *LTP4* transcript, but not the two controls, were clearly up-regulated in our setup after 3h and 9h of exposure to 100 mM NaCl, while after 1h no induction was detected (Figure 2). For the following analyses addressing transcriptomic changes upon NaCl exposure, we focused on the three genotypes Col-0, *elf7-3* and *cdc73-2*.

## ELF7 is required for efficient transcriptional response upon NaCl exposure

To learn whether the differential response to NaCl exposure of *elf7-3* relative to Col-0 and *cdc73-2* is paralleled by distinct transcriptional changes, we performed genome-wide transcript profiling of these three genotypes using RNA-seq. We intended to generate information on transcriptional output (freshly synthesized unspliced/spliced mRNAs including nascent transcripts) rather than steady-state mRNA levels obtained with total poly(A) mRNA (Zaghlool et al., 2013; Dhaliwal and Mitchell, 2016). Therefore, we (i) isolated nuclear RNA (rather than total RNA) (Supplementary Figure S2), (ii) used rRNA depletion (rather than poly(A) enrichment) and (iii) made use of low RNA size cut-off (≥25 nt). Four biological replicates of RNA of each genotype with or without 3h exposure to 100 mM NaCl were used for preparation of sequencing libraries. Analysis of the RNA-seq data revealed >35 million unique reads per genotype (Table S2) and that the biological replicates of the analyzed genotypes/conditions yielded robust results (Supplementary Figure S3), except for one of the *elf7-3*/NaCl samples that was removed as outlier. Analysis of genes differentially expressed in *elf7-3* or *cdc73-2* compared to Col-0 under control conditions identified 12 and 722 differentially expressed genes (DEGs) in *cdc73-2* and *elf7-3*, respectively (Supplementary Figure S4). Gene ontology (GO) term analysis of the DEGs in *elf7-3* identified genes responsive to salt stress as a prominent term (Supplementary Figure S5), while due to the small number of DEGs in *cdc73-2* the analysis showed no significant enrichment. The differential expression of only subsets of genes is in general agreement with *Arabidopsis* mutants deficient in other TEFs (van Lijsebettens and Grasser, 2014), although the

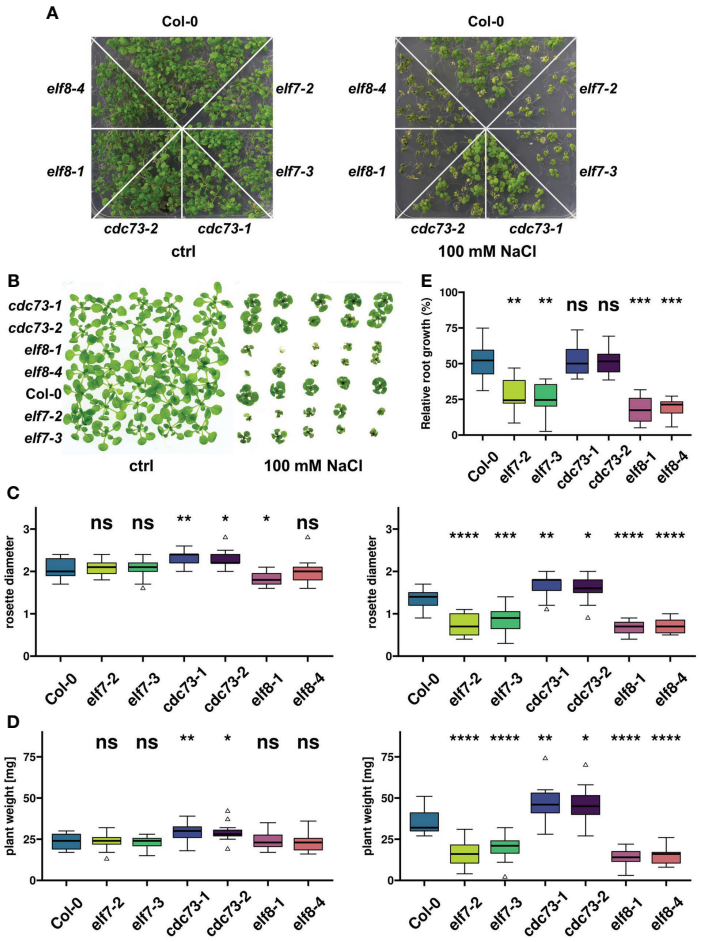



FIGURE 1
Growth of PAF1C subunit mutants is differentially affected by exposure to elevated NaCl concentrations. **(A)** Seeds of the indicated genotypes were sown on solid MS medium containing 100 mM NaCl or control medium (ctrl), and the resulting plants were documented at 14 days after stratification (14 DAS). **(B)** After growth on solid MS medium for 7 DAS, plants of the indicated genotypes were transferred to control medium or to medium containing 100 mM NaCl. Plants were documented after additional 18d (ctrl) or 21d (100 mM NaCl) of growth. **(C, D)** 7 DAS plants were grown in absence or presence of 100 mM NaCl for another 11d (ctrl, left) or 14d (100 mM NaCl, right), before rosette diameter [cm] **(C)** or fresh weight [mg] **(D)** was determined (n = 15). **(E)** Plants were grown on vertical MS plates for 7 DAS, before they were transferred to control medium or medium containing 100 mM NaCl. Length of the primary roots was recorded directly after transfer (0h) and 96h after transfer (n = 8). Growth after 96h in presence of NaCl is depicted as percentage of growth of the same genotype on control medium. The relative growth of each genotype was compared to Col-0 using Wilcoxon signed-rank test (ns, not significant; *, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$; ****, $p \leq 0.0001$).

transcriptomic difference to wild type is remarkably low in case of *cdc73-2*.

Analysis of differential gene expression with or without NaCl treatment for the genotypes separately revealed numerous significantly up- or down-regulated genes in Col-0 and *cdc73-2* (Figures 3A, B). In contrast, a clearly lower number of DEGs with mostly smaller expression changes was detected in *elf7-3* (Figure 3C). Principal component analysis (PCA) revealed a strict clustering of the samples according to NaCl treatment and control treatment (Supplementary Figure S6) that together with the observed Pearson correlation (Supplementary Figure S3) indicates the reproducibility of the obtained data. The transcriptomic analysis illustrates that 926 and 831 nuclear-encoded genes are differentially expressed upon NaCl treatment in Col-0 and *cdc73-2*, respectively (Figure 3D). A major part (568 genes) of the DEGs appears to be equally regulated in both genotypes. Moreover, a greater number of genes is up-regulated in both genotypes (675 and 507 for Col-0 and *cdc73-2*, respectively) when compared to the down-regulated genes (251 and 324 for Col-0 and *cdc73-2*, respectively) (Figures 3D, E). In *elf7-3*, 338 DEGs are detected upon NaCl treatment, of which 225 and 113 are up- and downregulated, respectively.

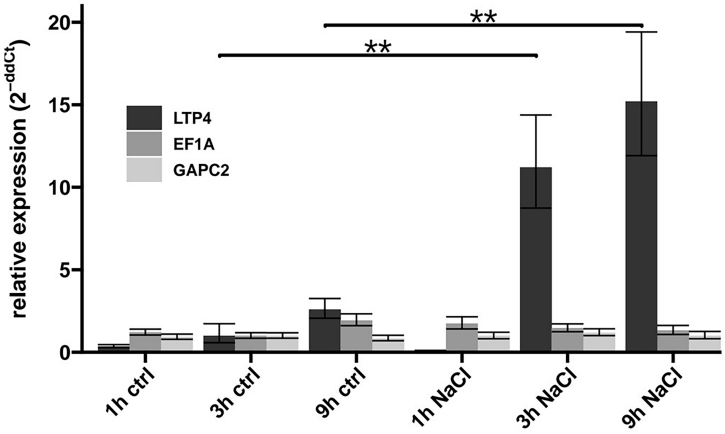



FIGURE 2
Changes in transcript levels upon exposure to NaCl. After growth on solid MS medium for 7 DAS, Col-0 plants were transferred to control medium (ctrl) or to medium containing 100 mM NaCl (NaCl). After different periods of the indicated treatment, RNA was isolated and the transcript levels of *LTP4* and reference genes (*EF1A*, *GAPC2*, *ACT2*) quantified by RT-qPCR. The depicted relative expression was normalised to the level of *ACT2* transcript in Col-0. The bars represent the mean relative expression of three technical replicates and error bars represent SD. The relative expression of *LTP4* was compared between control and NaCl treated condition of the same timepoints using Wilcoxon´s signed-rank test indicated by brackets (**, $p \leq 0.01$).

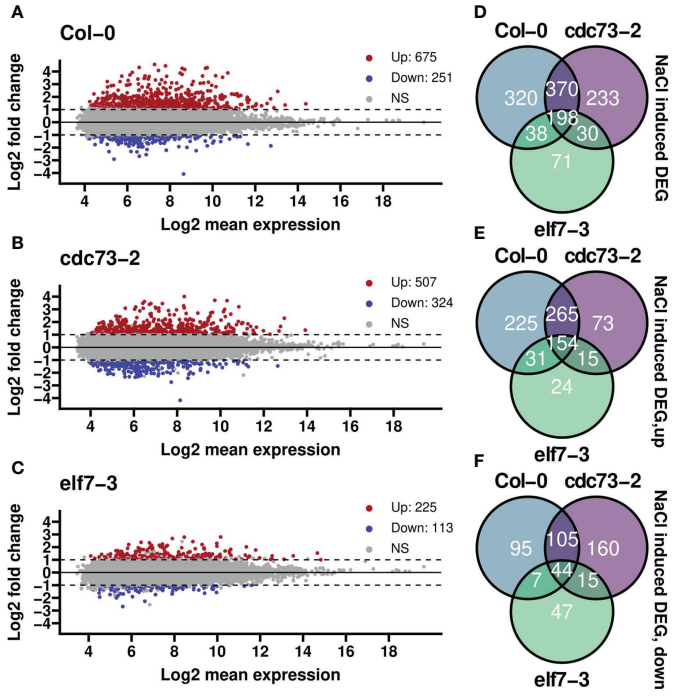



FIGURE 3
Differential gene expression analysis of *elf7-3* and *cdc73-2* relative to Col-0 upon NaCl treatment. **(A–C)** differential gene expression between treatment and control groups for each genotype separately. Highlighted in red are genes that are significantly upregulated after salt stress (log-fold change, LFC≥ 1 and *padjusted* ≤ 0.05), while in blue are genes that are significantly downregulated upon NaCl exposure (LFC≤ –1 and *padjusted* ≤ 0.05). NS = not significant. Numbers in the legends represent the gene counts in that group. **(D, E)** Venn diagrams summarising the number of differentially expressed genes in the different genotypes (as in **A–C**) upon NaCl treatment with total number of differentially expressed genes **(D)**, upregulated **(E)** or downregulated **(F)** genes. ns, not significant.

Hierarchical clustering analysis of the 500 genes that showed most variable transcript levels upon exposure to NaCl was visualized as a heatmap. Here, in agreement with the PCA (Supplementary Figure S6), predominantly two clusters are apparent, representing the two different conditions "control" and "NaCl" (Figure 4). As shown above, compared to Col-0 and *cdc73-2*, *elf7-3* exhibit decreased changes in transcript levels upon exposure to NaCl. Further analysis revealed that both the potential of *elf7-3* plants to induce and to repress transcription is decreased (Supplementary Figure S7). The NaCl-induced transcriptomic changes, occurring in Col-0 upon exposure to NaCl (926 DEGs) were further examined by GO term analysis. As expected (Golldack et al., 2014; Yang and Guo, 2018), this analysis demonstrated predominant enrichment of the GO term "response to salt stress" in all three genotypes (Supplementary Figure S8).

The differential transcriptional response observed with the different genotypes is also evident at the level of individual genes. As exemplified by the NaCl-inducible wax synthase/acyl-CoA: diacylglycerol acyltransferase (*WSD1*) gene, which plays a critical role in wax ester synthesis (Abdullah et al., 2021). Upon NaCl exposure substantially increased transcript levels of *WSD1* are detected in Col-0 and *cdc73-2*, while expression of the gene is hardly altered in *elf7-3* (Figure 5A). Another situation is observed for NaCl-repressed *DISEASE RELATED NONSPECIFIC LIPID 26 TRANSFER PROTEIN 1* (*DRN1*) gene, required for defense against pathogens as well as for normal seedling growth under salinity stress (Dhar et al., 2020). The *DRN1* transcript levels are reduced in presence of NaCl in Col-0 and *cdc73-2*, but not in *elf7-3* (Figure 5B). In case of the *ICL* gene encoding isocitrate lyase that plays a role in plant salt tolerance through the glyoxylate cycle (Yuenyong et al., 2019), no transcriptional change occurs in response to NaCl. However, relative to Col-0 and *elf7-3* distinctly elevated *ICL* transcript levels are detected in *cdc73-2* (Figure 5C). As discussed below, the differential transcriptional response might provide an explanation for the major difference of the three genotypes regarding their tolerance towards NaCl.

Further evaluation of genes specifically up- or down-regulated upon NaCl treatment in Col-0 or *cdc73-2* revealed that in Col-0 the average length of up-regulated transcripts is greater than that of the down-regulated transcripts, whereas the inverse is seen in *cdc73-2* (Figure 6A). The number of exons in these genes is higher in genes down-regulated in *cdc73-2* relative to the up-regulated genes, while regarding number of exons there is no significant difference among the DEGs specifically regulated in Col-0 (Figure 6B). Together these findings suggest that in the absence of CDC73 the processivity of RNAPII is decreased, which is particularly relevant for the transcription of longer genes. Analysis of differential splicing events induced upon NaCl exposure in either *cdc73-2* or Col-0 revealed only very few events (Supplementary Figure 9). This suggests that the greater length of down-regulated genes does not correlate with differential splicing events, but rather with a reduced transcriptional processivity, albeit the sequencing depth and read length of the experimental setup was not sufficient for a comprehensive spicing analysis.

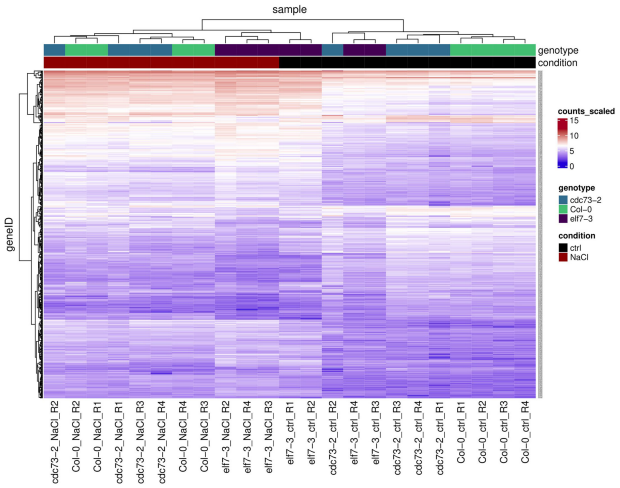



FIGURE 4
Heatmap and hierarchical clustering of the 500 most variable genes identified by transcript profiling of *elf7-3* and *cdc73-2* relative to Col-0 upon NaCl treatment. These genes were selected from the scaled counts of the count table using the subread package "featureCounts" and the count table includes the counts for all replicates, conditions and genotypes. Relative transcript abundance is shown as heatmap using the indicated scale (LFC of scaled counts). Genotypes and condition (ctrl: control; NaCl: 3h 100 mM NaCl) are colour-coded on top.

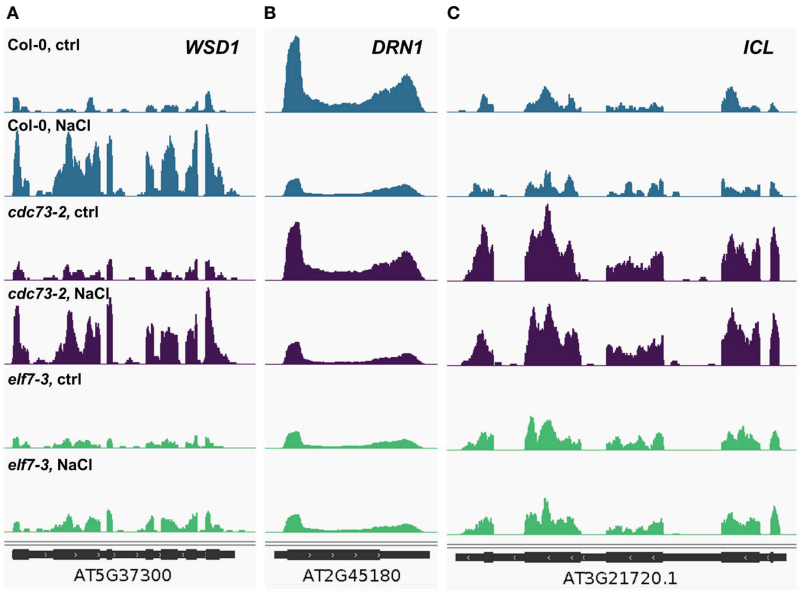



FIGURE 5
Changes in transcript levels in the different genotypes at individual loci dependent on the presence of NaCl. Coverage of merged RNA-seq samples was normalised to the effective *Arabidopsis* genome size and visualised using the Integrative Genomics Viewer. **(A)** coverage of *WSD1* (At5g37300), a gene induced by elevated salt concentrations. **(B)** coverage of *DRN1* (At2g45180), a gene repressed by elevated salt concentrations. **(C)** coverage of *ICL* (At3g21720), a gene, whose expression enhances salt tolerance.

Chromatin immunoprecipitation (ChIP) in combination with high throughput sequencing (ChIP-seq) was employed to examine the distribution of RNAPII over genes, whose expression is dependent on the presence/absence of ELF7. Chromatin of Col-0 and *elf7-3* was analyzed with an antibody specific for elongating RNAPII (RNAPII-S2P). Analyses of the ChIP-seq data revealed 12.3-18.8 million high quality reads per genotype and the biological replicates show robust correlation (Supplementary Figure S10). RNAPII coverage was compared for genes up- or downregulated in *elf7-3 vs*. Col-0 under standard conditions or upon exposure to NaCl. For upregulated genes RNAPII coverage is mildly increased, whereas for downregulated genes RNAPII coverage is clearly reduced in *elf7-3* (Figure 7). The genes with highly reduced RNAPII coverage include, for instance, also the above-mentioned *DRN1* gene (cf. Figure 5), whose expression is required for salt tolerance (Dhar et al., 2020). Therefore, ELF7 is particularly necessary for transcription of genes mis-regulated in *elf7-3* and transcriptional upregulation appears to depend more strongly on ELF7 (cf. Supplementary Figure S4).

## Discussion

PAF1C travels with RNAPII to regulate transcript elongation and to modulate chromatin structure of transcribed regions (Jaehning, 2010; Francette et al., 2021). In plants, so far mainly the role of PAF1C in developmental processes was studied (He et al., 2004; Oh et al., 2004; Park et al., 2010; Yu and Michaels, 2010; Fal et al., 2017). In few cases the observed mutant phenotype (s) could be correlated with misexpressed target genes, such as the above-mentioned link between *FLC* expression and time of flowering (Zhang et al., 2003; He et al., 2004; Oh et al., 2004; Xu et al., 2008; Park et al., 2010). Generally, *Arabidopsis* mutants deficient in various TEFs (e.g. Elongator, SPT4-SPT5, TFIIS) exhibit differential expression of only subsets of genes (Nelissen et al., 2005; Grasser et al., 2009; Dürr et al., 2014). Likewise, relative to wild type under normal growth conditions, based on microarray hybridization experiments (Park et al., 2010) or RNA-seq of total mRNA (Nasim et al., 2022), several hundreds of DEGs were observed for mutants deficient in PAF1C subunits, except for *cdc73*, whose transcriptome was similar to that of wildtype. Consistently, using high-throughput sequencing of nuclear mRNAs, we identified 12 and 722 DEGs in *cdc73-2* and *elf7-3*, respectively. Although it is difficult to directly compare transcriptomics datasets obtained with different techniques and analyzed using distinct statistical methods, the results in general support the emerging view that in plants TEFs are required for the correct transcription of only subsets of genes (van Lijsebettens and Grasser, 2014).

Recently, PAF1C (i.e. VIP3, VIP5, VIP6) was also identified as a critical factor in the plant response to repeated mechanic

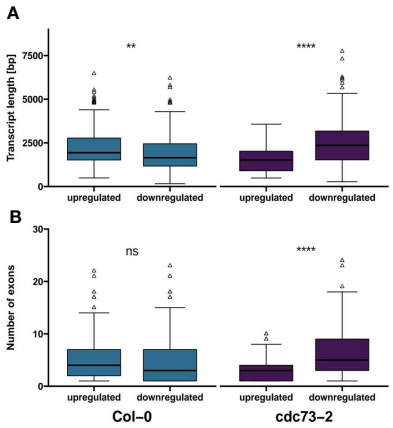



FIGURE 6
Transcript length and number of exons of genes uniquely differentially expressed in Col-0 or *cdc73-2*. **(A)** average transcript length of genes differentially expressed either in Col-0 or *cdc73-2*. **(B)** number of exons of genes differentially expressed either in Col-0 or *cdc73-2*. Data was evaluated using Wilcoxon signed-rank test (**, $p \leq 0.01$; ****, $p \leq 0.0001$; ns, not significant).

stimulation (Jensen et al., 2017) and in temperature-responsive flowering (Nasim et al., 2022). We have comparatively exposed different PAF1C subunit mutants to elevated concentrations of NaCl to provoke transcriptional response. In presence of NaCl, *elf7* and *elf8* displayed more severe decrease in growth of the aerial parts than Col-0, whereas the rosette size of *cdc73* rather exceeded that of Col-0. Growth of the primary root upon NaCl exposure was similarly decreased in Col-0 and *cdc73-1/cdc73-2*, whereas root growth of *elf7-2/elf7-3* and *elf8-1/elf8-4* was considerably more strongly reduced. Recently, it was reported – based on a similar, albeit not identical experimental setup – that the root growth of *elf7-3* in presence of 100 mM NaCl was comparable to that of Col-0 (Li et al., 2019). In view of that we accurately measured the relative root growth of both *elf7* alleles in presence of NaCl relative to their growth on control medium, taking into consideration the distinct growth of the mutants compared to Col-0 in absence of NaCl. Thereby, compared to the other genotypes we determined significantly decreased root growth of *elf7-2* and *elf7-3* (and *elf8-1/elf8-4*) in presence of NaCl.

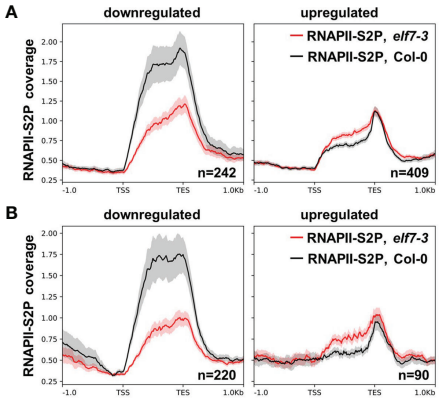



FIGURE 7
RNAPII coverage over genes up- and downregulated in *elf7-3*. Metagene plots of ChIP-seq analysis of Col-0 and *elf7-3* using an antibody directed against RNAPII-S2P over DEGs elf7-3/Col-0 under standard condtions **(A)** or upon exposure to NaCl **(B)**. Mean signals of the three biological replicates were averaged (line) and the shaded area represents the SEM for the three replicates at each position. The tracks were scaled over transcribed regions from transcriptional start site (TSS) to transcriptional end site (TES).

To record rather early transcriptional responses, we profiled the transcriptomes of *elf7-3* and *cdc73-2* in comparison to Col-0 after 3h of NaCl exposure, analyzing nuclear mRNAs including nascent transcripts. Col-0 and to a slightly lesser extent *cdc73-2* exhibited clear transcriptomic changes upon NaCl exposure, whereas *elf7-3* showed a comparatively clearly decreased response to this stress condition (Figure 3) that is also reflected by an altered RNAPII coverage in *elf7-3* (Figure 7). A major part of the genes differentially expressed in Col-0 upon NaCl exposure were also differentially expressed in *cdc73-2*, illustrating that both genotypes respond similarly to the presence of NaCl. With *elf7-3* plants, the presence of NaCl resulted in distinctly lower transcriptional response, indicating that ELF7 plays an important role in the response to elevated NaCl concentrations. In agreement with our findings, a recent study concluded that Col-0 and *cdc73 Arabidopsis* plants grown at 23°C have a similar transcriptome, and that *elf7* and *vip3/4/5/6* share a large number of DEGs (Nasim et al., 2022). In yeast as well as in mammals, PAF1C subunit mutants are also associated with a range of phenotypes suggesting (partial) subunit specificity (Betz et al., 2002; Akanuma et al., 2007; Chu et al., 2007; Cheung et al., 2008; Wang et al., 2008). Furthermore, yeast PAF1C subunit mutants exhibit various defects in the tolerance to environmental stress conditions (Shi et al., 1996; Betz et al., 2002; García et al., 2016). Particularly, inactivation of PAF1 and CTR9 result in severe mutant phenotypes in accord with their central structural role within PAF1C (Kim et al., 2010; Chu et al., 2013; Vos et al., 2020). Moreover, the different PAF1C subunits are involved to a variable extent in modulating the genomic distribution of a range of histone marks including H2B mono-ubiquitination and various transcription-related H3 methylations (Jaehning, 2010; Francette et al., 2021).

Analysis of the genes that are differentially expressed upon exposure of *Arabidopsis* plants to NaCl demonstrated that a greater part of the DEGs is comparably regulated in Col-0 and *cdc73-2*, illustrating that a similar transcriptional response occurs in both genotypes. In contrast, a comparatively lower transcriptional response can be observed with *elf7-3*, suggesting that ELF7 is crucial for this type of stress response, and

accordingly *elf7-3* (and *elf7-2*) plants are rather susceptible to NaCl. At the same time the above-mentioned finding that the expression of salt-responsive genes is altered in *elf7-3* plants under control conditions, may influence the adaptability of these plants upon exposure to NaCl, although the mechanism, how the lack of ELF7 influences gene expression under these conditions remains unknown. Interestingly, the tolerance to NaCl of *cdc73-2* plants is even more pronounced than that of Col-0. Elevated expression of gene(s) that enhance the tolerance to NaCl such as the *ICL* gene (Yuenyong et al., 2019) could contribute to the improved performance of *cdc73* plants under conditions of increased salt concentrations. The *ICL* transcript is not regulated by the presence of NaCl, but there are significantly elevated levels of the transcript in *cdc73-2* relative to Col-0 and *elf7-3* (Figure 5C) that may augment the resistance of *cdc73-2* to NaCl. However, several non-overlapping DEGs in *cdc73-2* represent salt-responsive genes (Supplementary Figure S11) that may also contribute to the salt tolerance of *cdc73-2*. Examination of DEGs that are specifically regulated in Col-0 or *cdc73-2* revealed a greater average length of the transcripts down-regulated in *cdc73-2* compared to Col-0. In line with that the genes down-regulated in *cdc73-2* contain a greater number of exons. Our comparative analysis of alternative splicing events in Col-0 and *cdc73-2* revealed only few events, but we consider this outcome rather inconclusive. Together these findings suggest that in the absence of CDC73 the processivity of RNAPII is decreased, which is particularly relevant for the transcription of longer genes. In line with that, depletion of PAF1C in mouse cells results in decreased RNAPII processivity and reduced elongation rate (Hou et al., 2019). In conclusion, our study demonstrates that the *Arabidopsis* PAF1C subunit mutants deficient in ELF7 and CDC73 respond very distinctly to NaCl, which is reflected by the different transcriptional response. Plants lacking CDC73 transcriptionally respond to NaCl similar to Col-0, but still – likely because of altered expression of certain gene(s) – exhibit increased tolerance to NaCl. In contrast, the transcriptional response of plants lacking ELF7 is decreased and consequently the plants are markedly sensitive to NaCl exposure. Therefore, our analyses provide evidence for PAF1C subunit specificity in plant response to environmental conditions.

## Data availability statement

The data presented in the study are deposited in the Sequence Read Archive (SRA) repository, accession number PRJNA816434 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA8164).

## Author contributions

SO, RS, and TS performed the experimental procedures; SO, RS, US, and CM analyzed the next-generation sequencing data; SO, CM, and KDG designed the research; KDG wrote the manuscript and all authors approved the submitted version. All authors contributed to the article and approved the submitted version.

## Funding

This research was supported by the German Research Foundation (DFG) through grants Gr1159/14-2 and SFB960/A6 to KDG.

## Acknowledgments

We thank Mathias Gradl for contributions to the project and the Nottingham Arabidopsis Stock Centre (NASC) for providing *Arabidopsis* T-DNA insertion lines.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer IL declared a past co-authorship with one of the authors KDG to the handling editor.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.974625/full#supplementary-material

# References

Abdullah, H. M., Rodriguez, J., Salacup, J. M., Castañeda, I. S., Schnell, D. J., Pareek, A., et al. (2021). Increased cuticle waxes by overexpression of WSD1 improves osmotic stress tolerance in arabidopsis thaliana and camelina sativa. *Int.J.Mol.Sci.* 22, 5173. doi: 10.3390/ijms22105173

Ahmed, M., and Kim, D. R. (2018). Pcr: an r package for quality assessment, analysis and testing of qPCR data. *PeerJ.* 6, e4473. doi: 10.7717/peerj.4473

Akanuma, T., Koshida, S., Kawamura, A., Kishimoto, Y., and Takada, S. (2007). Paf1 complex homologues are required for notch-regulated transcription during somite segmentation. *EMBO Rep.* 8, 858–863. doi: 10.1038/sj.embor.7401045

Anders, S., Reyes, A., and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. *Genome Res.* 22, 2008–2017. doi: 10.1101/gr.133744.111

Antosz, W., Deforges, J., Begcy, K., Bruckmann, A., Poirier, Y., Dresselhaus, T., et al. (2020). Critical role of transcript cleavage in arabidopsis RNA polymerase II transcriptional elongation. *Plant Cell* 32, 1449–1463. doi: 10.1105/tpc.19.00891

Antosz, W., Pfab, A., Ehrnsberger, H. F., Holzinger, P., Köllen, K., Mortensen, S. A., 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. doi: 10.1105/tpc.16.00735

Betz, J. L., Chang, M., Washburn, T. M., Porter, S. E., Mueller, C. L., and Jaehning, J. A. (2002). Phenotypic analysis of Paf1/RNA polymerase II complex mutations reveals connections to cell cycle regulation, protein synthesis, and lipid and nucleic acid metabolism. *Mol. Genet. Genomics* 268, 272–285. doi: 10.1007/s00438-002-0752-8

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic:a flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Chen, F. X., Smith, E. R., and Shilatifard, A. (2018). Born to run: control of transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* 19, 464–478. doi: 10.1038/s41580-018-0010-5

Cheung, V., Chua, G., Batada, N. N., Landry, C. R., Michnick, S. W., Hughes, T. R., et al. (2008). Chromatin- and transcription-related factors repress transcription from within coding regions throughout the saccharomyces cerevisiae genome. *PLoS Biol.* 6, e277. doi: 10.1371/journal.pbio.0060277

Chu, X., Qin, X., Xu, H., Li, L., Wang, Z., Li, F., et al. (2013). Structural insights into Paf1 complex assembly and histone binding. *Nucleic Acids Res.* 41, 10619–10629. doi: 10.1093/nar/gkt819

Chu, Y., Simic, R., Warner, M. H., Arndt, K. M., and Prelich, G. (2007). Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes. *EMBO J.* 26, 4646–4656. doi: 10.1038/sj.emboj.7601887

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10, giab008. doi: 10.1093/gigascience/giab008

Dhaliwal, N. K., and Mitchell, J. A. (2016). Nuclear RNA isolation and sequencing. *Methods Mol. Biol.* 1402, 63–71. doi: 10.1007/978-1-4939-3378-5_7

Dhar, N., Caruana, J., Erdem, I., and Raina, R. (2020). An arabidopsis DISEASE RELATED NONSPECIFIC LIPID TRANSFER PROTEIN 1 is required for resistance against various phytopathogens and tolerance to salt stress. *Gene* 753, 144802. doi: 10.1016/j.gene.2020.144802

Dürr, J., Lolas, I. B., Sørensen, B. B., Schubert, V., Houben, A., Melzer, M., et al. (2014). The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in arabidopsis. *Nucleic Acids Res.* 42, 4332–4347. doi: 10.1093/nar/gku096

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi: 10.1093/bioinformatics/btw354

Fal, K., Liu, M., Duisembekova, A., Refahi, Y., Haswell, E. S., and Hamant, O. (2017). Phyllotactic regularity requires the Paf1 complex in arabidopsis. *Development* 144, 4428–4436. doi: 10.1242/dev.154369

Francette, A. M., Tripplehorn, S. A., and Arndt, K. M. (2021). The Paf1 complex: A keystone of nuclear regulation operating at the interface of transcription and chromatin. *J. Mol. Biol.* 433, 166979. doi: 10.1016/j.jmb.2021.166979

García, P., Del Encinar Dedo, J., Ayté, J., and Hidalgo, E. (2016). Genome-wide screening of regulators of catalase expression: Role of a transcription complex and histone and tRNA modification complexes on adaptation to stress. *J. Biol. Chem.* 291, 790–799. doi: 10.1074/jbc.M115.696658

Golldack, D., Li, C., Mohan, H., and Probst, N. (2014). Tolerance to drought and salt stress in plants: Unraveling the signaling networks. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00151

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. doi: 10.1016/j.jmb.2008.12.066

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. doi: 10.1101/gad.1244504

Hou, L., Wang, Y., Liu, Y., Zhang, N., Shamovsky, I., Nudler, E., et al. (2019). Paf1C regulates RNA polymerase II progression by modulating elongation rate. *Proc. Natl. Acad. Sci. U.S.A.* 116, 14583–14592. doi: 10.1073/pnas.1904324116

Jaehning, J. A. (2010). The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochim. Biophys. Acta* 1799, 279–388. doi: 10.1016/j.bbagrm.2010.01.001

Jensen, G. S., Fal, K., Hamant, O., and Haswell, E. S. (2017). The RNA polymerase-associated factor 1 complex is required for plant touch responses. *J. Exp. Bot.* 68, 499–511. doi: 10.1093/jxb/erw439

Kilian, J., Whitehead, D., Horak, J., Wanke, D., Weinl, S., Batistic, O., et al. (2007). The AtGenExpress global stress expression data set: Protocols, evaluation and model data analysis of UV-b light, drought and cold stress responses. *Plant J.* 50, 347–363. doi: 10.1111/j.1365-313X.2007.03052.x

Kim, J.-H. (2021). Multifaceted chromatin structure and transcription changes in plant stress response. *Int.J.Mol.Sci.* 22, 2013. doi: 10.3390/ijms22042013

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. doi: 10.1016/j.cell.2009.12.050

Kleinboelting, N., Huep, G., Kloetgen, A., Viehoever, P., and Weisshaar, B. (2012). GABI-kat SimpleSearch: New features of the arabidopsis thaliana T-DNA mutant database. *Nucleic Acids Res.* 40, D1211–D1215. doi: 10.1093/nar/gkr1047

Krogan, N. J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M. A., 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. doi: 10.1016/s1097-2765(03)00091-1

Kwak, H., and Lis, J. T. (2013). Control of transcriptional elongation. *Ann. Rev. Genet.* 47, 483–508. doi: 10.1146/annurev-genet-110711-155440

Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., et al. (2012). The arabidopsis information resource (TAIR): Improved gene annotation and new tools. *Nucleic Acids Res.* 40, D1202–D1210. doi: 10.1093/nar/gkr1090

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nat. Meth.* 9, 357–359. doi: 10.1038/nmeth.1923

Liao, Y., Smyth, G. K., and Shi, W. (2019). The r package rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* 47, e47. doi: 10.1093/nar/gkz114

Li, Y., Yang, J., Shang, X., Lv, W., Xia, C., Wang, C., et al. (2019). SKIP regulates environmental fitness and floral transition by forming two distinct complexes in arabidopsis. *New Phytol.* 224, 321–335. doi: 10.1111/nph.15990

Love, M. L., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8

Mangiola, S., Molania, R., Dong, R., Doyle, M. A., and Papenfuss, A. T. (2021). Tidybulk: an r tidy framework for modular transcriptomic data analysis. *Genome Biol.* 22, 42. doi: 10.1186/s13059-020-02233-7

Michl-Holzinger, P., Obermeyer, S., Markusch, H., Pfab, A., Ettner, A., Bruckmann, A., et al. (2022). Phosphorylation of the FACT histone chaperone subunit SPT16 affects chromatin at RNA polymerase II transcriptional start sites in arabidopsis. *Nucleic Acids Res.* 50, 5014–5028. doi: 10.1093/nar/gkac293

Mueller, C. L., and Jaehning, J. A. (2002). Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex. *Mol.Cell.Biol.* 22, 1971–1980. doi: 10.1128/MCB.22.7.1971-1980.2002

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassay with tobacco tissue cultures. *Physiol.Plant.* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x

Nasim, Z., Susila, H., Jin, S., Youn, G., and Ahn, J. H. (2022). Polymerase II-associated factor 1 complex-regulated FLOWERING LOCUS c-clade genes repress flowering in response to chilling. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.817356

Nelissen, H., Fleury, D., Bruno, L., Robles, P., de Veylder, L., Traas, J., et al. (2005). The elongata mutants identify a functional elongator complex in plants with a role in cell proliferation during organ growth. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7754–7759. doi: 10.1073/pnas.0502600102

Ng, H. H., Robert, F., Young, R. A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* 11, 709–719. doi: 10.1016/s1097-2765(03)00092-3

Oh, S., Park, S., and van Nocker, S. (2008). Genic and global functions for Paf1C in chromatin modification and gene expression in arabidopsis. *PLoS Genet.* 4, e1000077. doi: 10.1371/journal.pgen.1000077

Oh, S., Zhang, H., Ludwig, P., and van Nocker, S. (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. doi: 10.1105/tpc.104.026062

Park, S., Oh, S., Ek-Ramos, J., and van Nocker, S. (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. doi: 10.1104/pp.110.155838

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.* 1629, 105–122. doi: 10.1007/978-1-4939-7125-1_8

Pfab, A., Grønlund, J. T., Holzinger, P., Längst, G., and Grasser, K. D. (2018). The arabidopsis histone chaperone FACT: Role of the HMG-box domain of SSRP1. *J. Mol. Biol.* 430, 2747–2759. doi: 10.1016/j.jmb.2018.06.046

Pfaff, C., Ehrnsberger, H. F., Flores-Tornero, M., Sørensen, B. B., Schubert, T., Längst, G., et al. (2018). ALY RNA-binding proteins are required for nucleocytosolic mRNA transport and modulate plant growth and development. *Plant Physiol.* 177, 226–240. doi: 10.1104/pp.18.00173

Quadrana, L., Bortolini Silveira, A., Mayhew, G. F., LeBlanc, C., Martienssen, R. A., Jeddeloh, J. A., et al. (2016). The arabidopsis thaliana mobilome and its impact at the species level. *Elife* 5, e15716. doi: 10.7554/eLife.15716

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., et al. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160–W165. doi: 10.1093/nar/gkw257

Reyes, A., Anders, S., Weatheritt, R. J., Gibson, T. J., Steinmetz, L. M., and Huber, W. (2013). Drift and conservation of differential exon usage across tissues in primate species. *Proc. Natl. Acad. Sci. U.S.A.* 110, 15377–15382. doi: 10.1073/pnas.1307202110

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. doi: 10.1038/sj.embor.7400045

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. doi: 10.1128/MCB.16.2.669

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. doi: 10.1101/gad.1235904

Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res.* 27, 491–499. doi: 10.1101/gr.209601.116

van Lijsebettens, M., and Grasser, K. D. (2014). Transcript elongation factors: shaping transcriptomes after transcript initiation. *Trends Plant Sci.* 19, 717–726. doi: 10.1016/j.tplants.2014.07.002

Vos, S. M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H., et al. (2018). Structure of activated transcription complex pol II-DSIF-PAF-SPT6. *Nature* 560, 607–612. doi: 10.1038/s41586-018-0440-4

Vos, S. M., Farnung, L., Linden, A., Urlaub, H., and Cramer, P. (2020). Structure of complete pol II-DSIF-PAF-SPT6 transcription complex reveals RTF1 allosteric activation. *Nat. Struct. Mol. Biol.* 27, 668–677. doi: 10.1038/s41594-020-0437-1

Wade, P. A., Werel, W., Fentzke, R. C., Thompson, N. E., Leykam, J. F., Burgess, R. R., et al. (1996). A novel collection of accessory factors associated with yeast RNA polymerase II. *Protein Expr. Purif.* 8, 85–90. doi: 10.1006/prep.1996.0077

Wang, P., Bowl, M. R., Bender, S., Peng, J., Farber, L., Chen, J., et al. (2008). Parafibromin, a component of the human PAF complex, regulates growth factors and is required for embryonic development and survival in adult mice. *Mol.Cell.Biol.* 28, 2930–2940. doi: 10.1128/MCB.00654-07

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An "Electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2, e718. doi: 10.1371/journal.pone.0000718

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. doi: 10.1074/jbc.C300269200

Xiao, T., Kao, C.-F., Krogan, N. J., Sun, Z.-W., Greenblatt, J. F., Osley, M. A., et al. (2005). Histone H2B ubiquitylation is associated with elongating RNA polymerase II. *Mol.Cell.Biol.* 25, 637–651. doi: 10.1128/MCB.25.2.637-651.2005

Xu, L., Zhao, Z., Dong, A., Soubigou-Taconnat, L., Renou, J. P., Steinmetz, A., et al. (2008). Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in arabidopsis thaliana. *Mol.Cell.Biol.* 28, 1348–1360. doi: 10.1128/MCB.01607-07

Yang, Y., and Guo, Y. (2018). Unraveling salt stress signaling in plants. *J.Integr.Plant Biol.* 60, 796–804. doi: 10.1111/jipb.12689

Yuenyong, W., Sirikantaramas, S., Qu, L.-J., and Buaboocha, T. (2019). Isocitrate lyase plays important roles in plant salt tolerance. *BMC Plant Biol.* 19, 472. doi: 10.1186/s12870-019-2086-2

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. doi: 10.1104/pp.110.158386

Zaghlool, A., Ameur, A., Nyberg, L., Halvardson, J., Grabherr, M., Cavelier, L., et al. (2013). Efficient cellular fractionation improves RNA sequencing analysis of mature and nascent transcripts from human tissues. *BMC Biotechnol.* 13, 99. doi: 10.1186/1472-6750-13-99

Zhang, H., Ransom, C., Ludwig, P., and van Nocker, S. (2003). Genetic analysis of early flowering mutants in arabidopsis defines a class of pleiotropic developmental regulator required for expression of the flowering-time switch flowering locus c. *Genetics* 164, 347–358. doi: 10.1093/genetics/164.1.347

Zhang, H., and van Nocker, S. (2002). The VERNALIZATION INDEPENDENCE 4 gene encodes a novel regulator of FLOWERING LOCUS c. *Plant J.* 31, 663–673. doi: 10.1046/j.1365-313x.2002.01380.x

# Chapter 5

# Phosphorylation of the FACT histone chaperone subunit SPT16 affects chromatin at RNA polymerase II transcriptional start sites in *Arabidopsis*

This peer-reviewed article was published in the journal *Nucleic Acids Research* in 2022

*5014–5028* *Nucleic Acids Research, 2022, Vol. 50, No. 9* *Published online 30 April 2022*
*https://doi.org/10.1093/nar/gkac293*

# Phosphorylation of the FACT histone chaperone subunit SPT16 affects chromatin at RNA polymerase II transcriptional start sites in *Arabidopsis*

**Philipp Michl-Holzinger[1], Simon Obermeyer[1], Hanna Markusch[1], Alexander Pfab[1], Andreas Ettner[1], Astrid Bruckmann[2], Sabrina Babl[3], Gernot Längst[3], Uwe Schwartz[4], Andrey Tvardovskiy[5], Ole N. Jensen[5], Akihisa Osakabe[6], Frédéric Berger[6] and Klaus D. Grasser[1,*]**

[1]Department of Cell Biology & Plant Biochemistry, Centre for Biochemistry, University of Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany, [2]Institute for Biochemistry I, Centre for Biochemistry, University of Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany, [3]Institute for Biochemistry III, Centre for Biochemistry, University of Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany, [4]NGS Analysis Centre, Biology and Pre-Clinical Medicine, University of Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany, [5]Department of Biochemistry and Molecular Biology, VILLUM Center for Bioanalytical Sciences, University of Southern Denmark, Odense, Denmark and [6]Gregor Mendel Institute (GMI), Austrian Academy of Sciences, Vienna BioCenter (VBC), Dr. Bohr-Gasse 3, 1030 Vienna, Austria

Received September 15, 2021; Revised April 12, 2022; Editorial Decision April 13, 2022; Accepted April 19, 2022

## ABSTRACT

**The heterodimeric histone chaperone FACT, consisting of SSRP1 and SPT16, contributes to dynamic nucleosome rearrangements during various DNA-dependent processes including transcription. In search of post-translational modifications that may regulate the activity of FACT, SSRP1 and SPT16 were isolated from *Arabidopsis* cells and analysed by mass spectrometry. Four acetylated lysine residues could be mapped within the basic C-terminal region of SSRP1, while three phosphorylated serine/threonine residues were identified in the acidic C-terminal region of SPT16. Mutational analysis of the SSRP1 acetylation sites revealed only mild effects. However, phosphorylation of SPT16 that is catalysed by protein kinase CK2, modulates histone interactions. A non-phosphorylatable version of SPT16 displayed reduced histone binding and proved inactive in complementing the growth and developmental phenotypes of *spt16* mutant plants. In plants expressing the non-phosphorylatable SPT16 version we detected at a subset of genes enrichment of histone H3 directly upstream of RNA polymerase II transcriptional start sites (TSSs) in a region that usually is nucleosome-depleted. This suggests that some genes require phosphorylation of the SPT16 acidic region for establishing the correct nucleosome occupancy at the TSS of active genes.**

## INTRODUCTION

Packaging nuclear DNA into nucleosomes constitutes the basic structural unit of eukaryotic chromatin. Due to their stability, nucleosomes serve as general repressors of transcription and other DNA-dependent processes (1,2). Accordingly, various mechanisms exist that facilitate the transcription of chromatin templates by destabilization/disassembly of nucleosomes (3). So-called histone chaperones play a fundamental role in regulating nucleosomal dynamics. They represent a diverse group of proteins that functionally interact with core histones to assemble/disassemble nucleosome particles without consuming energy in form of ATP. In this way, histone chaperones are key factors that determine nucleosome properties, for instance, by modulating the distribution of histone variants or epigenetic information, thereby defining functionally distinct chromatin landscapes (4–6).

FACT is a heterodimeric histone chaperone consisting of SSRP1 (Pob3 in yeast) and SPT16 that originally was identified in yeast and mammalian cells (7–10). Its established name traces back to the finding that FACT facilitates chromatin transcription by promoting *in vitro* transcription by RNA polymerase II (RNAPII) from reconstituted chromatin templates. Meanwhile it became apparent that

*To whom correspondence should be addressed. Tel: +49 941 9433032; Fax: +49 941 9433353; Email: klaus.grasser@biologie.uni-regensburg.de
Present address: Andrey Tvardovskiy, Institute of Functional Epigenetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany.

FACT contributes to other chromatin-dependent processes including recombination, replication and repair (11–14), and hence the established abbreviation may well stand more broadly for facilitates chromatin transactions. Regarding transcription by RNAPII, FACT has been implicated in initiation (15–18) as well as in the elongation stage (8,9,19,20). Importantly, FACT can disassemble and reassemble nucleosomes; thereby it is involved in both overcoming and maintaining the nucleosomal barrier to DNA-dependent processes in the chromatin context. To accomplish that, FACT establishes contacts with different nucleosomal targets including H2A–H2B dimers, H3–H4 tetramers and DNA (21–26). A recent structural analysis of FACT in complex with nucleosomal particles revealed that the overall shape of FACT resembles a unicycle, consisting of a saddle and fork that is engaged in multiple interactions with histones and DNA (27). Reversible nucleosome reorganization and uncoiling of nucleosomal DNA from the histone core, increasing DNA accessibility, are brought about by FACT interfering with DNA-histone contacts (24,25,28,29). Moreover, FACT plays a critical role in nucleosome reassembly, for instance, following RNAPII passage to maintain chromatin signature and to repress transcription initiation from cryptic sites within coding regions (25,30–32).

In plants, there is also a variety of histone chaperones (33,34) and their loss affects regulation of growth and development or the response to environmental stress conditions (35–37). Likewise, FACT consisting of SSRP1 and SPT16 is conserved in a wide variety of plant species (38,39). Both subunits are essential for viability in *Arabidopsis* (40,41) and are widely expressed in almost all cell types (42–44). Plants deficient in SSRP1 or SPT16 display various defects in vegetative and reproductive development including increased number of leaves and inflorescences, early flowering, reduced seed production and impairment of proper expression of parentally regulated genes during seed development (40,41,43). FACT copurified with the RNAPII transcript elongation complex from *Arabidopsis* cells and associated with active protein-coding genes in a transcription dependent manner (41,42,45–47), suggesting a role in transcriptional elongation. At the same time *Arabidopsis* FACT efficiently represses intragenic transcriptional initiation (48).

Currently, there is only limited information about the potential regulation of FACT function via post-translational modifications (12). Protein kinase CK2 *in vitro* phosphorylates human and maize SSRP1 altering DNA interactions of the protein (49,50). *Drosophila* SSRP1 is phosphorylated at CK2 consensus sites in insect cells and the phosphorylation has subtle effects on *in vitro* nucleosome interactions of SSRP1 (51,52). Human SPT16 is phosphorylated within its acidic domain in insect cells and the phosphorylation influences the interaction with a 112-bp octasome (23). Since plant FACT has not been systematically analysed for post-translational modifications and their function, we isolated both subunits from *Arabidopsis* cells and using mass spectrometry four acetylation sites were detected in SSRP1, while three phosphorylation sites were mapped in SPT16. Mutation of the SSRP1 acetylation sites revealed only mild effects, whereas mutation of the SPT16 phosphorylation sites affected its interaction with histones and the ability of the protein to complement efficiently the mutant phenotype of *spt16* plants. In addition, the expression of a non-phosphorylatable SPT16 variant resulted in a distinct histone H3 enrichment upstream of RNAPII transcriptional start sites (TSSs) of a subset of genes.

## MATERIALS AND METHODS

### Plasmid constructions

The required gene or cDNA sequences were amplified by PCR with KAPA DNA polymerase (PeqLab) using an *Arabidopsis thaliana* genomic DNA or cDNA as template and the primers (providing also the required restriction enzyme cleavage sites) listed in Supplementary Table S1. Mutations were introduced by overlap extension PCR. The PCR fragments were inserted into suitable plasmids using standard methods. All plasmid constructions were checked by DNA sequencing, and details of the plasmids generated in this work are summarized in Supplementary Table S2.

### Plant material

Seeds of *Arabidopsis thaliana* (Col-0, L*er*) were stratified in darkness for 48 h at 4°C and plants were grown at 21°C on soil in a phytochamber or on MS medium in plant incubators under long-day conditions (45,53). The T-DNA insertion lines *ssrp1-1* and *spt16-1* used for the presented experiments were previously reported (41). *Agrobacterium*-mediated plant transformation, characterization of transgenic lines and PCR-based genotyping of plants using primers specific for DNA insertions and target genes (Supplementary Table S1) was performed as previously described (53,54). Primary transformants were selected and analysed by PCR-based genotyping. Plants harbouring the transgene were screened by immunoblotting for uniform expression and were examined for consistent phenotype. Based on these analyses three independent plant lines per transgene were typically selected for further experimentation, of which one line was analysed in detail. Plant phenotypes were observed and documented as previously described (41,53), while some analyses were performed using the software Leaf-GP (55).

### Affinity purification and characterization of GS-tagged proteins from *Arabidopsis* cells

*Arabidopsis* suspension cultured PSB-D cells were maintained and transformed as previously described (56). Proteins of 15 g cultured cells were isolated and GS-tagged proteins (2× protein G domains and streptavidin-binding peptide) were purified by IgG affinity chromatography as previously described (45,53). 1.5 μg of in solution trypsin digested sample were analysed by LC–MS/MS using a Proxeon Easy nLC system (Thermo Scientific) essentially as previously described (57). Samples were loaded through a two-column system (analytical- (15 cm) and trap-column (3 cm) with an internal diameter of 75 μm filled C18-AQ 3 μm (150 × 4 mm) ReproSil-Pur beads. Separations were performed using gradient elution at RT with a flow rate of 300 nl/min. The mobile phase consisted of a linear gradient containing 0.1% formic acid (v/v) (eluent A) and 0.1%

formic acid in 95% acetonitrile (eluent B). The LC system was coupled to a Q Exactive hybrid quadrupole-Orbitrap mass spectrometer (Thermo Scientific) via an electrospray ionization source operated in positive ion mode. Mass spectra were acquired using a top 12 high collision dissociation (HCD) data-dependent acquisition (DDA) method. The obtained peptide mass fingerprints were analysed using the proteome discoverer 2.0 software, searching the UniProt database. Enzyme specificity was set to trypsin allowing up to two missed cleavages. The search included cysteine carbamidomethylation as a fixed modification, protein N-terminal acetylation, acetylation of lysine and arginine, mono-, di- and tri- methylation of lysine, mono- and dimethylation of arginine, oxidation of methionine and phosphorylation of serine, threonine, and tyrosine as variable modifications. Peptide identification was based on a search with a mass deviation of the precursor ion up to 2.5 ppm after recalibration, and the allowed fragment mass deviation was set to 0.01 Da. Analysis of *in vitro* phosphorylated proteins and subtraction of the experimental background of contaminating proteins were performed as previously described (45).

### Antibodies and immunoblotting

For immunoblotting, the following primary and secondary antibodies were used: anti-H2B (cat. no. ab1790, Abcam) and anti-H3 (cat. no. ab1791, Abcam); anti-SSRP1 and anti-SPT16 (42); secondary HRP-coupled α-rabbit antibody (Sigma Aldrich). Total protein extracts from PSB-D cultures were prepared as described before (45) and nuclear proteins were isolated as for ChIP (53,58). Following SDS-PAGE proteins were electro-transferred onto a PVDF membrane. For immunoblot analysis SuperSignal R West Pico Chemiluminescent substrate (Thermo Fisher Scientific) was used and chemiluminescence was detected using a Multiimage FluorChem FC2 instrument (Alpha Innotech) or the ChemiDoc MP system (Bio-Rad).

### Recombinant proteins

Using the respective expression vectors (Supplementary Table S2) truncated/mutated SSRP1 proteins were expressed in *E. coli* as 6xHis-tagged fusion proteins that were purified by two-step chromatography (Ni-NTA–agarose, phenyl-sepharose) as previously described (59). Using the respective expression vectors (Supplementary Table S2) truncated/mutated SPT16 proteins were expressed in *Escherichia coli* as 6xHis-tagged or glutathione S transferase (GST) tagged fusion proteins that were purified by Ni-NTA–agarose or glutathione-sepharose chromatographies, respectively, as previously described (53,59). Recombinant *Arabidopsis* and human histones H2A–H2B were prepared as previously described (60,61), while recombinant maize CK2α was purified by three-step chromatography as previously described (62).

### Circular dichroism (CD)

Proteins (10 μM) were analysed in 50 mM $KH_2PO_4$ using a Jasco J-815 CD spectropolarimeter using a 0.02 cm cell. The buffer signal was subtracted and the measured ellipticity was converted into mean residue ellipticity as described before (63).

### Electrophoretic mobility shift assay (EMSA)

DNA and different concentrations of protein were incubated in 1x EMSA buffer (10 mM Tris–HCl pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, 5% (v/v) glycerol, 0.01 mg/ml BSA) for 10 min and analysed in 6% polyacrylamide TBE gels. DNA was visualized using ethidium bromide staining.

### GST pull-down assays

Recombinant SPT16 proteins fused to glutathione S transferase (GST) were mixed with equimolar amounts of recombinant *Arabidopsis* H2A–H2B or with bovine cytochrome *c* (Sigma Aldrich) in GST buffer (0.2–0.35 M NaCl, 25 mM HEPES pH 7.6, 0.05% (v/v) NP40, 5 mM DTT, 10% (v/v) glycerol, 2 mM $MgCl_2$). For some assays GST-AD/WT was phosphorylated *in vitro* by CK2α (as described below) prior to the binding reaction. Following incubation (30 min at 30°C) glutathione sepharose beads were added and the samples were incubated for 3 h on a rotating wheel at 4°C. Beads were washed three times in GST buffer before bound proteins were eluted by boiling in protein loading buffer and analysed by SDS-PAGE.

### MicroScale thermophoresis (MST)

For MST, GST-AD proteins were labelled with a 647-NHS-ester (Dyomics). MST experiments were performed in MST buffer (200 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4$, 1.8 mM $KH_2PO_4$, pH 7.4) with constant target concentration of 10 nM and a 1:2 ligand titration starting at 36 μM H2A–H2B. Measurements were performed in premium capillaries at 7% LED, 60% MST at 25°C on a Monolith NT.$115_{Pico}$ device (NanoTemper Technologies). The data was analyzed using the MO.affinity analysis software (NanoTemper Technologies) and binding affinity was determined at 2.5s laser On time. Data was fitted using the $K_D$ model and fraction bound was calculated by dividing the $\Delta F_{norm}$ value of each point by the amplitude of the fitted curve, resulting in values from 0 to 1 (0 = unbound, 1 = bound).

### Protein kinase CK2 *in vitro* phosphorylation assays

Radioactive and non-radioactive CK2 phosphorylation assays were performed as previously described (62). For radioactive assays, recombinant SPT16 proteins were reacted in the presence of [γ-$^{32}$P]ATP with recombinant maize CK2α. Phosphate incorporation was monitored by SDS-PAGE and scanning of the gels using a Cyclone Storage phosphorimager (Canberra Packard). For non-radioactive assays, recombinant SPT16 proteins were phosphorylated by recombinant purified maize CK2α in the presence of unlabelled ATP and phosphate incorporation was monitored by acetic acid urea PAGE (AU-PAGE) (64) and by mass spectrometry.

### Chromatin immunoprecipitation (ChIP)

ChIP was essentially performed as previously described (53,58) using 1 g of 14-days after stratification (14-DAS) *in vitro* grown plants that were crosslinked with 1% (v/v) formaldehyde. Chromatin was sonicated using a Bioruptor Pico (Diagenode) and immunoprecipitation with antibodies specific for the C-terminal part of histone H3 (cat. no. ab1791, Abcam) was performed using magnetic Dynabeads Protein A (Thermo Fisher Scientific). To generate libraries for high-throughput sequencing the NEBNext Ultra II DNA Library Prep Kit with index primers NEBNext multiplex oligos for Illumina (New England BioLabs) were used. Sequencing of the libraries (75-nt, single-end) was performed by the Genomics Core Facility at the University of Regensburg (http://www.kfb-regensburg.de/) using an Illumina NextSeq 500 instrument. Reads were aligned to the TAIR10 genome (https://www.arabidopsis.org/) using Bowtie2 (65) and coverage tracks were calculated with DANPOS2 (66). Downstream analysis was mainly performed using the deepTools2 suite (version 3.5.0) (67) and quality control was performed at several steps using FastQC (68). Three replicates were obtained for each genotype and 12.6–16.2 million reads per sample mapped uniquely to the *Arabidopsis* genome (average Phred quality score >34 for each library). Expression levels of genes (divided into quartiles according to logTPM) was deduced from RNA sequencing data of *in vitro* grown 6-DAS Col plants (NCBI sequence read archive, PRJNA758800). MNase sequencing data (69) were reanalysed. After preprocessing reads were aligned to the *Arabidopsis* TAIR10 genome using Bowtie2 (65). Only reads with mononucleosomal length (140–200 bp) were used for further analysis and genome-wide nucleosome occupancy maps were generated using the dpos option of the DANPOS2 (66).

## RESULTS

### Mapping post-translational modifications of the FACT subunits

To identify post-translational modifications of the FACT subunits, SSRP1 and SPT16 were isolated from *Arabidopsis* suspension cultured cells. We used PSB-D cell lines expressing SSRP1 or SPT16 fused to a GS affinity tag as well as cells expressing the unfused GS tag that served as a control (45). GS fusion proteins (and the unfused GS tag) were isolated from total protein extracts by IgG affinity purification (45,53) and analysed by SDS-PAGE. Several prominent protein bands co-eluted with SSRP1-GS and SPT16-GS (Figure 1A). Notably, SSRP1 and SPT16 co-purified with each other, as expected (42), and histones were found to be enriched in the eluates. Subsequently, the eluted proteins were digested with trypsin and analysed by liquid chromatography coupled with mass spectrometry (LC–MS). Modified peptides detected with high confidence (FDR < 0.01) were considered. The mass spectrometric analyses revealed four acetylation sites mapping to lysine residues (Supplementary Figure S1A) in the HMG-box domain of SSRP1 (K594, K599) and to the flanking basic region (K539, K549) (Figure 1B, C). SSRP1 proteins of various plant species contain lysine residues at respective positions. Moreover, two high-confidence phosphorylation sites map to serine residues (Supplementary Figure S1B) within the intrinsically disordered acidic domain of SPT16 (S1033, S1035) (Figure 1B,C). An additional phosphorylation site (T1023) adjacent to the phospho-serine was detected with lower confidence, but it is indicated in our data, since it was also identified in our *in vitro* phosphorylation experiments (see below).

### Mild effects caused by acetylation of SSRP1

Since protein acetylation can alter functional properties of diverse HMG-box proteins including their DNA interactions (70,71), we produced different recombinant SSRP1 proteins comprising the C-terminal region including the HMG-box DNA-binding domain (DBD). In addition to the wild-type variant (DBD/WT), we generated proteins in which the four acetylated lysines (Supplementary Figure S2A, B) were changed to non-acetylatable alanine or to acetylation-mimicking glutamine residues (DBD/4xA or DBD/4xQ, respectively). Proteins purified by two-step chromatography (Supplementary Figure S2C) were analysed for possible structural differences using CD spectroscopy and for their DNA-binding capacities using electrophoretic mobility shift assays (EMSAs). Far-UV CD spectra of the three proteins (Supplementary Figure S2D) reflect the largely α-helical structure of HMG-box domains (72,73). Relative to DBD/WT the spectra of the mutated proteins indicate some loss in the amount of α-helix (mild decrease of the negative peak around 222 nm) and a gain in random coil (minimum around 200 nm). HMG-box proteins often bind selectively to certain DNA structures including DNA minicircles or four-way junction (4wj) DNA (72,73). Therefore, we reacted increasing concentrations of the recombinant proteins with 4wj DNA. As evident from the disappearance of the unbound DNA probe and the formation of protein complexes with lower electrophoretic mobility, the three variants displayed no clear differences in their affinity for the 4wj DNA (Supplementary Figure S2E). Relative to the WT protein, we observed with the mutant variants a reduced tendency of forming distinctly recognisable protein/DNA complexes in the EMSA analysis. Similarly, the affinity of the mutant variants for linear DNA differed only slightly from that of the WT protein (Supplementary Figure S2F).

Acetylation of HMG-box proteins can influence their subcellular localization, as exemplified by mammalian HMGB1 (74). Moreover, two of the identified SSRP1 acetylation sites are situated within a 20-aa basic region adjacent to the HMG-box domain (cf. Supplementary Figure S2A), which serves as nuclear localization sequence in maize SSRP1 (75). Therefore, SSRP1/WT and the mutant version SSRP1/4xQ both fused to eGFP were expressed in Col-0 plants. Analysis of leaf cells of the transgenic plants by confocal laser scanning microscopy (CLSM) demonstrated that both eGFP fusion proteins localize to the nucleus (Supplementary Figure S3A), suggesting that the acetylatable lysine residues do not interfere with nuclear localization.

To explore the potential role of the identified acetylatable residues of SSRP1 in plant growth and development, we used a complementation approach. The *Arabidopsis* trans-

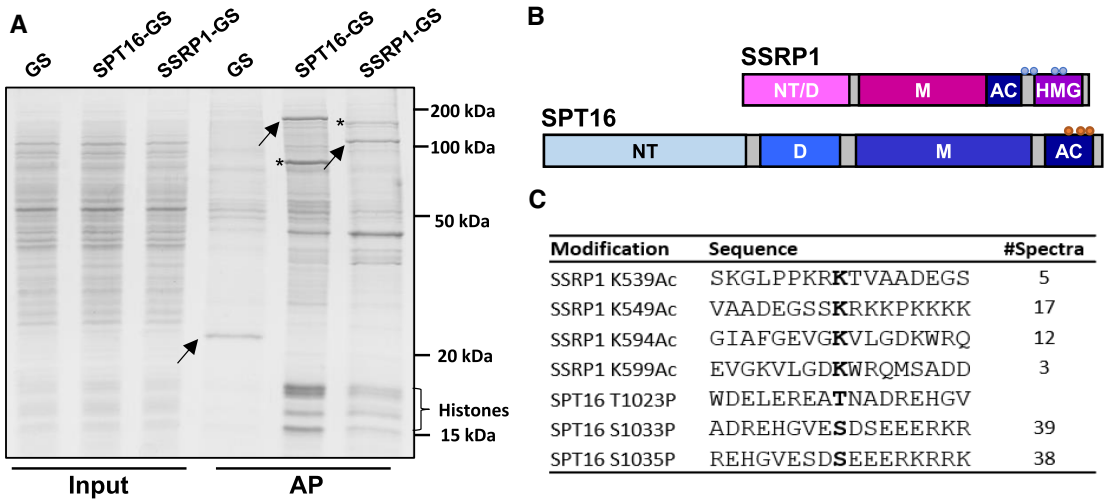



**Figure 1.** Mapping of acetylation sites in SSRP1 and phosphorylation sites in SPT16. (**A**) Affinity purification of FACT subunits from Arabidopsis cells. The indicated GS-fusion proteins were purified from PSB-D cells by IgG affinity chromatography from total protein extracts. Isolated proteins were separated by SDS-PAGE and stained with Coomassie brilliant blue. The unfused GS-tag (GS) served as control in these experiments. Arrows indicate the bands corresponding to the bait proteins fused to the GS-tag. Asterisks indicate the respective complex partner. (**B**) Schematic representation of the SSRP1 and SPT16. SSRP1 is composed of an N-terminal domain (that is required for heterodimerization (NT/D) with SPT16), middle domain (M), acidic region (AC) and HMG-box domain (HMG), while SPT16 consists of N-terminal (NT), dimerization (D), middle (M), and acidic (AC) domains. (**C**) MS analysis lead to the detection of acetylation and phosphorylation sites in SSRP1 and SPT16, respectively. The list summarizes the modified sites (bold) that were identified in the MS analyses. Shown are modifications, which were found in more than one high confidence spectrum, except T1023P which was detected with low confidence and included here because of additional results (see below). The identified modified residues are indicated in the scheme in (B) with acetylation sites in SSRP1 (lysine, blue circles) and phosphorylation sites in SPT16 (serine/threonine, orange circles).

poson insertion mutant *ssrp1-1* is homozygous lethal and exhibits wild-type phenotype in the heterozygous state (41). Constructs driving the expression of wild-type SSRP1/WT or of the mutated SSRP1/4xR and SSRP1/4xQ variants were introduced into heterozygous *ssrp1-1* plants. The presence of the three complementation constructs allowed the isolation of plants homozygous for the *ssrp1-1* mutation (Supplementary Figure S3B), demonstrating that the expression of SSRP1 and the mutant variants restored growth of the otherwise lethal *ssrp1-1* mutant. The SSRP1 proteins were expressed at similar levels in the different plant lines (Supplementary Figure S3C). Comparative phenotypic analysis of the L*er* wild-type control along with the complementation lines illustrated that all plants basically had wild-type appearance (Figure 2A). No striking differences were observed between the plant lines regarding bolting time of the plants or various growth determinants including leaf area and height of the plants (Figure 2B)—parameters that are clearly affected with homozygous viable plants of the *ssrp1-2* allele (41). We concluded that although acetylation of SSRP1 affects to some extent protein structure and its interaction with DNA, our analyses did not reveal distinct impact of the identified acetylatable residues in SSRP1 on plant growth and development under standard growth conditions.

### SPT16 is phosphorylated by protein kinase CK2

Inspection of the amino acid sequence around the serine residues in SPT16 identified as phospho-sites, revealed that S1033/S1035 are situated within consensus phosphorylation sites of protein kinase CK2 (Figure 3A). Target sites of CK2 often reside in acidic regions of substrate proteins within a minimal consensus sequence (i.e. S/T-X-X-D/E/pS) (76). To examine the *in vitro* capacity of CK2 to phosphorylate SPT16, the acidic domain (AD, D955-R1075) was generated as 6xHis-tagged protein. The purified protein was incubated with the recombinant catalytic subunit (CK2α) of maize CK2 in the presence of radiolabelled ATP. Separation of the phosphorylation reactions by SDS-PAGE and visualization of the $^{32}$P-incorporation into the substrate protein demonstrated that CK2α efficiently phosphorylated the acidic domain of SPT16 (Figure 3B). Phosphorylation of the acidic domain by CK2α in the presence of unlabelled ATP did not alter the electrophoretic mobility of the protein compared to the non-phosphorylated protein in SDS-PAGE analysis, whereas phosphorylation by CK2α resulted in a clear mobility shift of SPT16-AD in acetic acid urea PAGE analysis (Figure 3C). To identify sites phosphorylated by CK2α *in vitro*, the phosphorylated acidic domain (as in Figure 3C) was proteolytically cleaved and analysed by LC–MS. Mass spectrometric analyses demonstrated that in addition to serine residues S1033/S1035, threonine T1023 was distinctly phosphorylated in our *in vitro* phosphorylation assay (Figure 3D). Since the same phospho-sites were detected on SPT16 isolated from *Arabidopsis* cells, it was likely that protein kinase CK2 phosphorylates SPT16 *in planta*. In support of this idea, several CK2 variants robustly co-purified with SPT16-GS and SSRP1-GS in our AP-MS analyses of *Arabidopsis* cells (Supplementary Table S3). Since the acidic domain of SPT16 contains several additional CK2 consensus sites (Figure 3A), we wondered whether (some of) these sites may be also phosphorylated by CK2, but perhaps escape

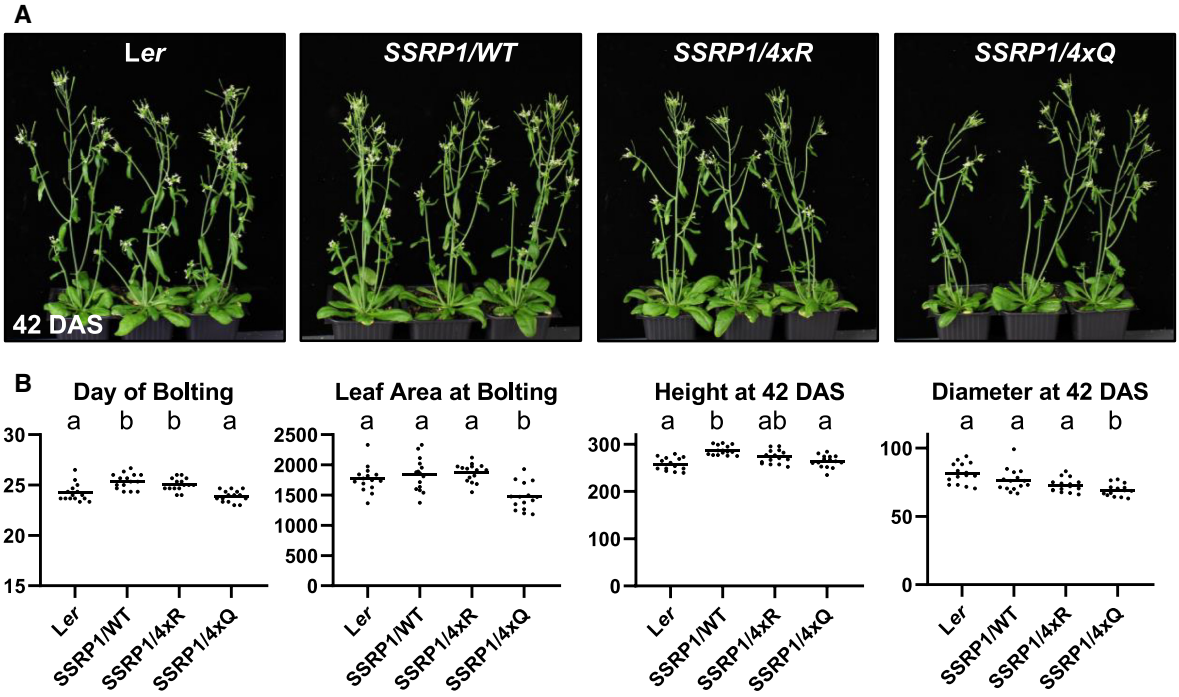


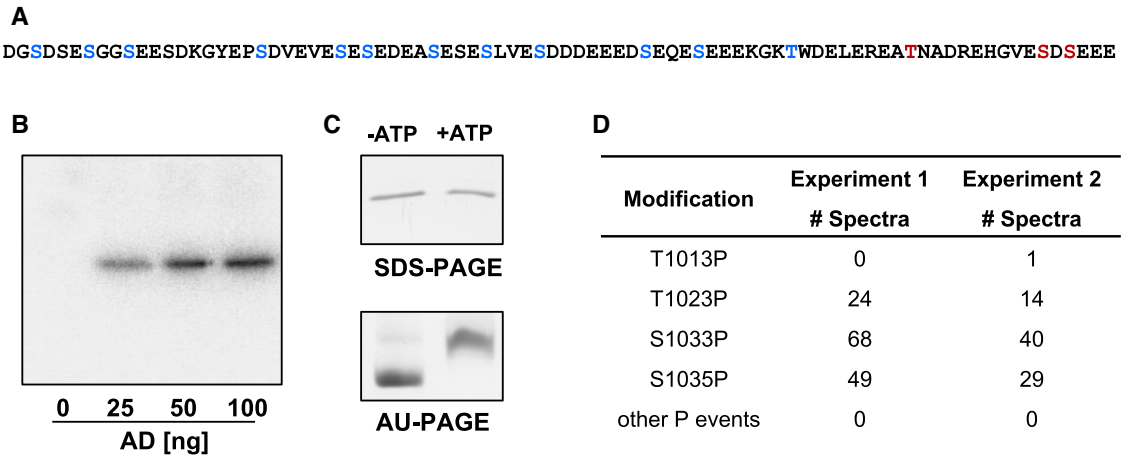

**Figure 2.** Plants expressing SSRP1 acetylation site mutants are mildly affected. Plants of the *ssrp1-1* mutant line were transformed with constructs driving the expression of wild-type SSRP1 (SSRP1/WT), the four acetylated lysine residues changed to arginine (K539R/K549R/K594R/K599R, termed SSRP1/4xR) or the four acetylated lysine residues changed to glutamine (K539Q/K549Q/K594Q/K599Q, termed SSRP1/4xQ) and were phenotypically analysed relative to the wild-type (L*er*). (**A**) Representative individuals of the different transgenic lines relative to L*er* (42 DAS) are shown. (**B**) Quantification of phenotypic parameters. Day of bolting, leaf area at bolting [mm2], height at 42 DAS [mm], rosette diameter at 42 DAS [mm], were determined. Data comprised the measurements of 15 individual plants of each genotype in three biological replicates and was analysed by one-way ANOVA. The line in the scatterplot marks the average value and letters above the scatter indicate the outcome of the Tukey´s test ($P$-value < 0.05).

mass spectrometric detection. Therefore, a mutated version of the acidic domain (with non-phosphorylatable changes in T1023V, S1033A, S1035A termed AD/VAA) was generated (Supplementary Figure S4A). The mutated version and the wild-type acidic domain were comparatively phosphorylated using the *in vitro* $^{32}$P-CK2α assay (Supplementary Figure S4B). Relative to the wild-type protein the mutated version was phosphorylated to a lesser extent. Quantification of the $^{32}$P-incorporation demonstrated that phosphorylation is decreased by ∼40% (Supplementary Figure S4C). This illustrates that the three identified residues are target sites of CK2, but additional residues within the acidic domain can be phosphorylated by this protein kinase, although the corresponding phospho-peptides were not detected in our LC–MS analyses of SPT16 isolated from plant cells. This could be related to low stoichiometry of the *in vivo* phosphorylations, or technical issues including ion suppression and the fact that the peptides are highly negatively charged.

### Influence of phosphorylation on SPT16 interaction with histones

The acidic domain of SPT16 substantially contributes to FACT-histone interactions (22,24,26). To examine whether the phosphorylation of the acidic domain alters histone binding, we produced a non-phosphorylatable and a phosphomimic variant of the acidic domain (AD) by replacing the phosphorylated S/T residues as above (T1023V, S1033A, S1035A termed AD/VAA) or with acidic residues (T1023E, S1033D, S1035D, termed AD/EDD), respectively. Recombinant AD/VAA, AD/EDD and wild-type versions of the acidic domain fused to GST were expressed in *E. coli* and purified. The three GST-AD fusion proteins along with unfused GST and potential binding partners, heterodimers of recombinant *Arabidopsis* histone H2A and H2B and cytochrome C (CytC) were analysed by SDS-PAGE (Figure 4A), illustrating their purity. The GST-AD proteins were used for GST pull-down assays with histone H2A–H2B. Histones were incubated with GST-AD/WT, GST-AD/VAA and GST-AD/EDD at different NaCl concentrations. Following immobilization on glutathione sepharose beads and washing of the beads, bound proteins were eluted and the GST fusion proteins were separated from co-eluting histones by SDS-PAGE (Figure 4B). Quantification of experimental triplicates revealed that GST-AD/VAA interacted ∼41% less efficiently with the histones than GST-AD/WT, while the interaction of GST-AD/EDD was ∼46% more efficiently than that of GST-AD/WT (Figure 4C). To directly assess the influence of CK2-mediated phosphorylation on the GST-AD interaction with histone H2A–H2B, GST-AD/WT was phospho-

*5020 Nucleic Acids Research, 2022, Vol. 50, No. 9*




| Modification | Experiment 1 # Spectra | Experiment 2 # Spectra |
|---|---|---|
| T1013P | 0 | 1 |
| T1023P | 24 | 14 |
| S1033P | 68 | 40 |
| S1035P | 49 | 29 |
| other P events | 0 | 0 |

**Figure 3.** Protein kinase CK2 phosphorylates SPT16 within the acidic domain (AD). (**A**) Amino acid sequence of the acidic domain (AD, D955-R1075) of SPT16. Detected phosphosites (S1033, S1035 and T1023) in SPT16 isolated from *Arabidopsis* cells are indicated in red, while additional minimal CK2 consensus sites are indicated in blue. (**B**) Different amounts of SPT16-AD phosphorylated *in vitro* by recombinant maize CK2α in the presence of [γ-$^{32}$P]ATP after analysis by SDS-PAGE and phosphoimaging. (**C**) Reaction of SPT16-AD with CK2α in absence or presence of unlabelled ATP after separation by SDS-PAGE (top) or AU-PAGE (bottom) visualised by Coomassie staining. (**D**) SPT16-AD samples phosphorylated *in vitro* (as in C) were digested with trypsin and analysed by LC–MS/MS. Two replicates were analysed revealing phosphorylation events comparable to those detected *in vivo*.

rylated by CK2 *in vitro* prior to the binding assay, or mock-phosphorylated (no ATP added) (Figure 4D). Quantification of experimental triplicates demonstrated that CK2-phosphorylated GST-AD/WT interacted ∼44% more efficiently with the histones than the mock-phosphorylated protein (Figure 4E). The specificity of the AD-histone interaction is also evident from the lack of binding of H2A–H2B to unfused GST and the lack of binding of the small basic protein CytC to GST-AD/EDD in the GST pull-down assay (Figure 4F). Additionally, the histone interaction of the GST-AD variants was analysed in solution using microscale thermophoresis. The measurements confirmed the higher affinity for histones of GST-AD/EDD ($K_d$ = 1.73 μM), while GST-AD/WT and GST-AD/VAA ($K_d$ = 3.64 and 4.18 μM, respectively) bound similarly to the histones (Supplementary Figure S5).

To further evaluate the role of the AD of SPT16 in histone interactions, the wild-type version SPT16/WT-GS and for comparison mutated variants of SPT16 were expressed in *Arabidopsis* suspension cultured cells. In the mutated variants the three phosphorylatable S/T residues of the AD were replaced as above by non-phosphorylatable residues (termed SPT16/VAA) or by phosphomimicking residues (termed SPT16/EDD). IgG affinity purification from the different cell lines resulted in co-purification of SSRP1 with the three SPT16-GS versions (Figure 4G). Moreover, histones were clearly enriched in the eluates of the SPT16/WT-GS, SPT16/VAA-GS and SPT16/EDD-GS proteins, but not with the unfused GS control. The co-eluted histone bands were noticeably more prominent with SPT16/WT-GS and SPT16/EDD-GS when compared with SPT16/VAA-GS. This observation was confirmed by immunoblot analyses using antibodies directed against histones H2B and H3 (Figure 4H). Together these experiments suggest that the phosphorylation of SPT16 increases interaction with nucleosomal histones.

### Role of SPT16 phosphorylation in plant growth and development

To study the possible involvement of the three identified SPT16 phosphorylation sites in plant development, we expressed wild type and mutant versions of SPT16 fused with the red fluorescent protein (TagRFP) in *spt16* T-DNA insertion plants under control of the *SPT16* promoter. Coding sequences were fused to the sequence encoding TagRFP to enable the detection of protein expression and localization. The expression constructs were introduced into *spt16-1* plants (Supplementary Figure S6A). This mutant allele expresses reduced levels of SPT16 (41) and immunoblot analyses of nuclear protein extracts using a SPT16-specific antibody revealed that the different SPT16-TagRFP fusion proteins are expressed in the different plant lines along with the endogenous SPT16 (Supplementary Figure S6B, C). CLSM confirmed the nuclear localization of SPT16-TagRFP and of the corresponding phospho-site variants in root tips (Supplementary Figure S6D). Phenotypic analyses of the *spt16* plants expressing the different SPT16 variants relative to the Col-0 wild-type control demonstrated that expression of SPT16 and the phosphomimic variant partially complemented the mutant phenotype, whereas the non-phosphorylatable variant hardly imparted complementation (Figure 5A). We analysed in detail some of the phenotypic characteristics that are clearly affected in *spt16* mutant plants (41). Thus, the early bolting phenotype of *spt16* is partially complemented by the expression of SPT16/WT and SPT16/EDD (resembling Col-0), whereas the bolting of the plants expressing SPT16/VAA is comparable to that of *spt16* (Figure 5B). The reduced rosette diameter of *spt16* is also more efficiently complemented by the expression of SPT16/WT and SPT16/EDD. Likewise, the increased number of primary inflorescences of *spt16* is reduced to a greater extent by the expression of SPT16/WT and SPT16/EDD relative to SPT16/VAA. The ability to complement mu-

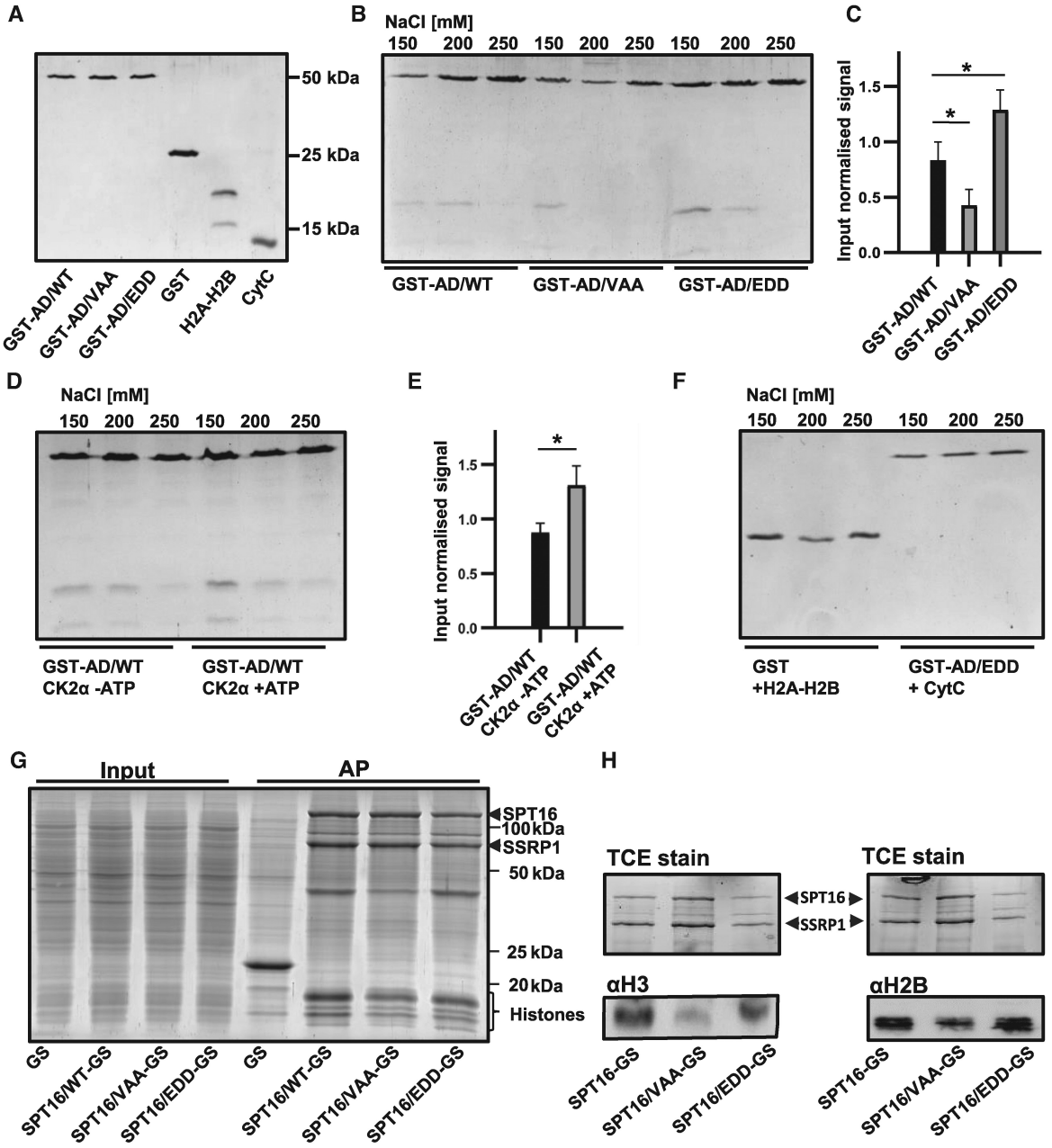



**Figure 4.** Phosphorylation of the SPT16 acidic domain modulates interactions with histones. (**A**) SDS-PAGE analysis of GST-AD/WT, GST-AD/VAA (with T1023V/S1033A/S1035A), GST-AD/EDD (with T1023E, S1033D, S1035D), unfused GST, recombinant *Arabidopsis* H2A–H2B and bovine Cytochrome C (CytC) visualised by Coomassie staining. (**B**) Comparative GST pull-down assays performed with GST-AD/WT, GST-AD/VAA and GST-AD/EDD, and with H2A–H2B in the presence of the indicated concentrations of NaCl. Eluted proteins were analysed by SDS-PAGE and Coomassie staining. (**C**) Quantification of H2B normalised to the GST-bait protein in presence of 150 mM NaCl. Statistical analysis of three replicates was performed using an unpaired t-test (p-value $< 0.05$). (**D**) Comparative GST pull-down assays performed with GST-AD/WT after phosphorylation with CK2α in absence/presence of ATP, respectively. (**E**) Quantification of H2B normalised to the GST-bait protein in presence of 150 mM NaCl. Statistical analysis of three replicates was performed using an unpaired t-test (*P*-value $< 0.05$). (**F**) Control assays showing either unfused GST in pull-down experiments with *Arabidopsis* H2A–H2B (left), or GST-AD/EDD with CytC (right) in presence of the indicated NaCl concentrations, reveal no interaction. (**G**) Input samples (Input) and eluates of IgG affinity purifications (AP) of *Arabidopsis* PSB-D cells expressing the unfused GS-tag, SPT16/WT-GS or the phosphositemutants SPT16/VAA-GS and SPT16/EDD-GS. Proteins were analysed by SDS-PAGE and Coomassie staining. (**H**) Immunoblot analyses (bottom) of eluates from the IgG affinity purification (as in **G**) probed with antibodies specific for histone H3 (left) and H2B (right), along with TCE stains (top) of the same SDS-PAGE gel that was used for the respective immunoblot analysis, serving as loading control. SPT16 and SSRP1 are highlighted by triangles. TCE: 2,2,2-trichlorethanol.

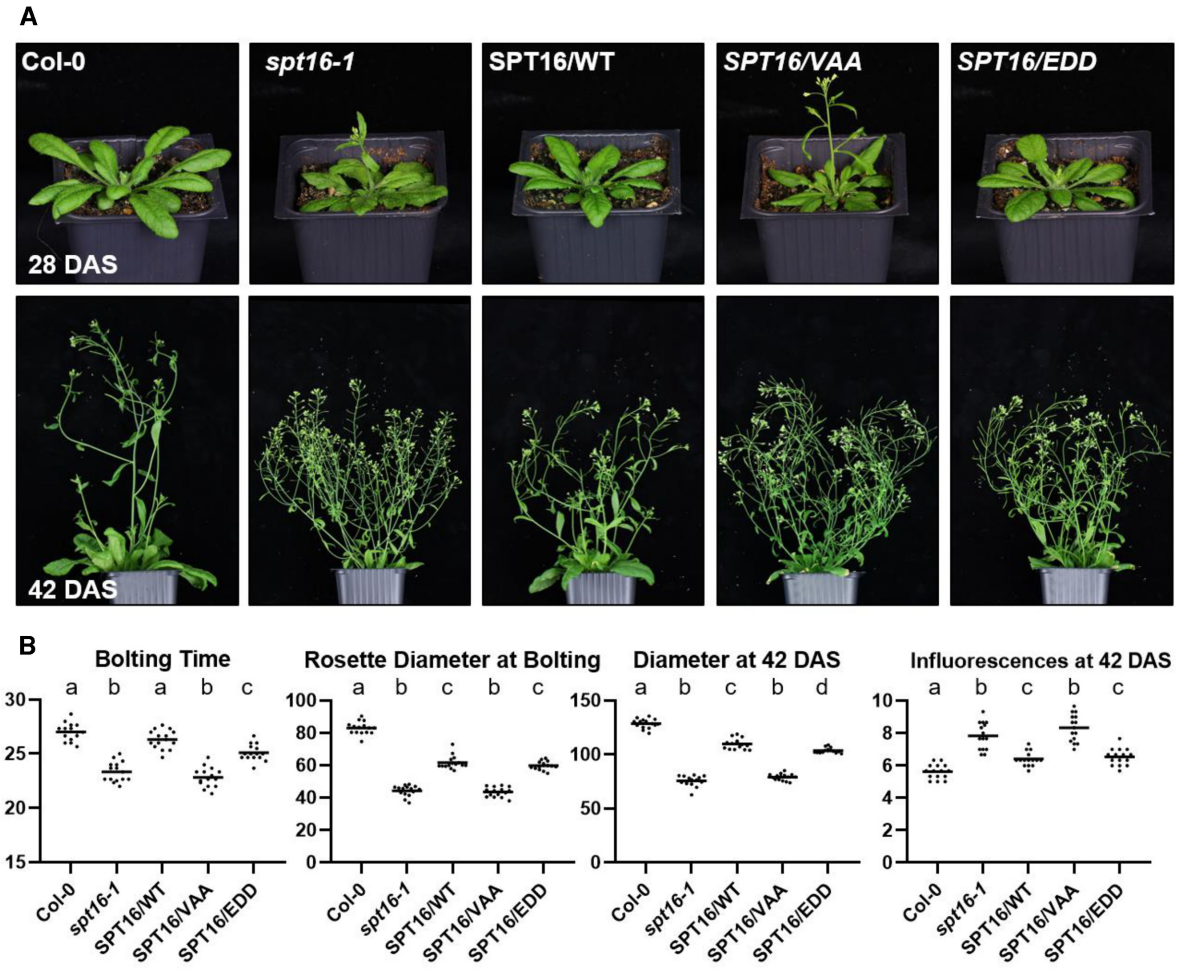



**Figure 5.** Phosphorylation of SPT16 acidic domain influences plant growth and development. Plants of the *spt16-1* mutant line were transformed with constructs driving the expression of wild-type SPT16 (termed SPT16/WT), the non-phosphorylatable (T1023V/S1033A/S1035A, termed SPT16/VAA) and phosphomimic (T1023E/S1033D/S1035D, termed SPT16/EDD) variants, and were phenotypically analysed relative to the wild-type (Col-0) and the *spt16-1* line. (**A**) Representative individuals of the different transgenic lines relative to Col-0 and *spt16-1* are shown at 28 DAS (top) and 42 DAS (bottom). (**B**) Quantification of phenotypic parameters. Day of bolting, rosette diameter at bolting [mm], rosette diameter at 42 DAS [mm], and number of primary inflorescences at 42 DAS were determined. Data comprised the measurements of 15 individual plants of each genotype in three biological replicates and was analysed by one-way ANOVA. The line in the scatterplot marks the average value and letters above the scatter indicate the outcome of the Tukey´s test ($P$-value $< 0.05$).

tant phenotypes also illustrated the functionality of the expressed SPT16 fusion proteins, which is in agreement with findings regarding SPT16 fusion proteins in other organisms (77,78). Thus, taken together the complementation analyses revealed that the expression of SPT16/VAA in *spt16* background resembles dominant-negative features and that phosphorylation of SPT16 is important for the impact of SPT16 on growth and development.

### Effects on chromatin at transcriptional start sites

In view of the function of FACT as a histone chaperone and the distinct abilities of SPT16 and the corresponding phospho-site variants to complement the defects of *spt16* mutant plants, we examined possible alterations in the chromatin of the different genotypes. The genome-wide distribution of H3 was used as a proxi for nucleosome occupancy and compared using ChIPseq between the following genetic backgrounds, Col-0 wild-type, *spt16*, *spt16* expressing SPT16/WT, the non-phosphorylatable SPT16/VAA or the phosphomimic SPT16/EDD variants. Analysis of the ChIPseq data revealed that biological replicates of the analysed genotypes yielded robust results (Supplementary Figure S7). Consistent with previous studies on the genomic distribution of H3 in *Arabidopsis* (79,80), we observed in all genotypes an H3 enrichment over transposable elements (Supplementary Figure S8A) and a characteristically lower occupancy upstream of RNA polymerase II transcription start sites (TSSs) particularly of highly transcribed genes (Supplementary Figure S8B). The profiles of Col-0, *spt16* and *spt16* expressing the phosphomimic variant SPT16/EDD were comparable, whereas *spt16* expressing

the non-phosphorylatable variant SPT16/VAA showed a prominent H3 occupancy peak upstream of the TSS (peak centre ∼120 bp upstream of TSS), whose extent correlated with the transcript levels of the respective genes (Figure 6A). The H3 enrichment upstream of the TSS observed specifically for the *spt16* plants expressing the SPT16/VAA variant (Supplementary Figure S8C) is accompanied by lower H3 occupancy in the gene body (Figure 6B). The peak of H3 enrichment upstream of the TSS coincides with the position of nucleosome depleted regions and DNase hypersensitive sites reported in several studies analysing *Arabidopsis* chromatin (79–82). It is conceivable that nucleosome depletion upstream of the TSS is restricted in the presence of the non-phosphorylatable SPT16/VAA variant resulting in H3 accumulation. Therefore, H3 distribution was comparatively analysed for genes that showed a prominent H3 enrichment at the TSS (≥50% higher coverage in SPT16/VAA as compared to SPT16/WT, $P < 0.05$) in the presence of SPT16/VAA (384 genes) relative to genes, whose H3 occupancy was unchanged (3284 genes). While with the latter group of genes an essentially uniform H3 distribution was observed in all genotypes, the former group showed H3 enrichment upstream of the TSS that was strikingly enhanced in presence of the SPT16/VAA variant (Supplementary Figure S9). ChIP-qPCR analysis of the association of the SPT16 variants with the region upstream of the TSS of three loci that exhibit H3 enrichment revealed that SPT16 is detected, however, at levels clearly lower than within the transcribed region. No significant differences between the variants were observed (Supplementary Figure S10). Both in maize and *Arabidopsis*, nucleosomes were identified directly upstream of the TSS that proved particularly nuclease sensitive (69,83). The differential MNase digestion data obtained for *Arabidopsis* chromatin (69) were analysed comparatively for the two above-mentioned groups of genes. For both groups of genes, within the transcribed region high and low MNase digestion similarly revealed the characteristic nucleosomal pattern, albeit a more intense peak of the +1 nucleosome was observed for the genes with H3 enrichment upstream of the TSS upon expression of the SPT16/VAA variant (Supplementary Figure S11). However, MNase digestion of these genes varied markedly upstream of the TSS (left panel), whereas a much smaller difference between high and low digestion occurred with the group of genes with unchanged H3 distribution (right panel). There is a prominent peak at the position of the -1 nucleosome upon low MNase digestion that is substantially weaker upon high MNase digestion (indicative of an unstable/fragile nucleosome) specifically with the group of genes showing H3 accumulation upstream of the TSS in the presence of the non-phosphorylatable SPT16/VAA. Centring H3 ChIPseq tracks and MNase-seq tracks over the TSSs of all protein-coding genes clarified the correlation of H3 enrichment observed with SPT16/VAA and of the region of most pronounced variability in MNase accessibility (Supplementary Figure S12). The H3 occupancy based on the ChIPseq analysis plotted in comparison to the differential MNase digestion (69) at individual loci illustrates these effects (Figure 6C). Here it is evident that the H3 enrichment upstream of the TSS observed in the presence of the non-phosphorylatable SPT16/VAA variant (but not with SPT16/WT or SPT16/EDD) coincides with the peak of an unstable/fragile nucleosome that is preferentially detected after low rather than high MNase digestion. In conclusion, these experiments suggest that at a subgroup of RNAPII-transcribed genes phosphorylation of SPT16 is required for establishing the correct nucleosome occupancy at the TSS.

## DISCUSSION

Histone chaperones assist transcription of nucleosomal templates and concurrently maintain proper chromatin signature over transcribed genes. A prominent representative is the FACT heterodimer consisting of SSRP1 and SPT16 that promotes RNAPII progression through chromatin and assists other DNA-dependent processes. We identified post-translational modifications that may modulate the activity of FACT in *Arabidopsis* cells. Our analyses did not reveal a major role for the four acetylation sites that were mapped within the C-terminal region of SSRP1 comprising the HMG-box domain and the adjacent basic region. However, our study showed that three *in vivo* phosphorylation sites identified in the acidic domain of SPT16 impact deeply the function of FACT and its role in growth and development of *Arabidopsis*. The same three sites (and additional sites) were phosphorylated by protein kinase CK2 *in vitro*, suggesting that CK2 is the enzyme catalysing the phosphorylation *in planta*. In agreement with that several α and β subunits of the heterodimeric protein kinase CK2 family co-purified with the *Arabidopsis* RNAPII transcript elongation complex (45), similar to the situation in yeast (84). Comparing the amino acid sequences of the *Arabidopsis* SPT16 acidic domain with those of SPT16 proteins of other organisms revealed multiple CK2 consensus phosphorylation sites in all SPT16 proteins (Supplementary Figure S13). Phosphorylation of several residues within the acidic domain was also observed, when human SPT16 was isolated from insect cells (23). Therefore, CK2-mediated phosphorylation of SPT16 might be a common mechanism, as it has been described for SSRP1 (49–52). Since members of the plant CK2 family have been implicated in multiple developmental and stress-responsive pathways (85,86), it will be attractive to investigate whether phosphorylation of FACT plays a role in these processes.

In comparison to the wild-type acidic domain, the phosphomimic variant of the acidic domain (AD/EDD) bound *in vitro* more tightly to H2A–H2B heterodimers. Likewise, SPT16/WT and SPT16/EDD co-purified more efficiently with histones from *Arabidopsis* cells, when compared with the non-phosphorylatable SPT16/VAA variant. Therefore, the phosphorylatable SPT16/WT and phosphomimic SPT16/EDD variants interact efficiently with histones, whereas the histone-interactions of the non-phosphorylatable SPT16/VAA variant are reduced. The acidic region of SPT16 (and SSRP1) substantially contribute to the interaction with histones and are required for the nucleosome reorganizing function of FACT (22,24,26,27). CK2-mediated phosphorylation of the acidic region of SPT16 (and SSRP1) could fine-tune the interaction of FACT with nucleosomes. In line with that the phosphorylation of the acidic region of human SPT16 is cru-

*5024 Nucleic Acids Research, 2022, Vol. 50, No. 9*

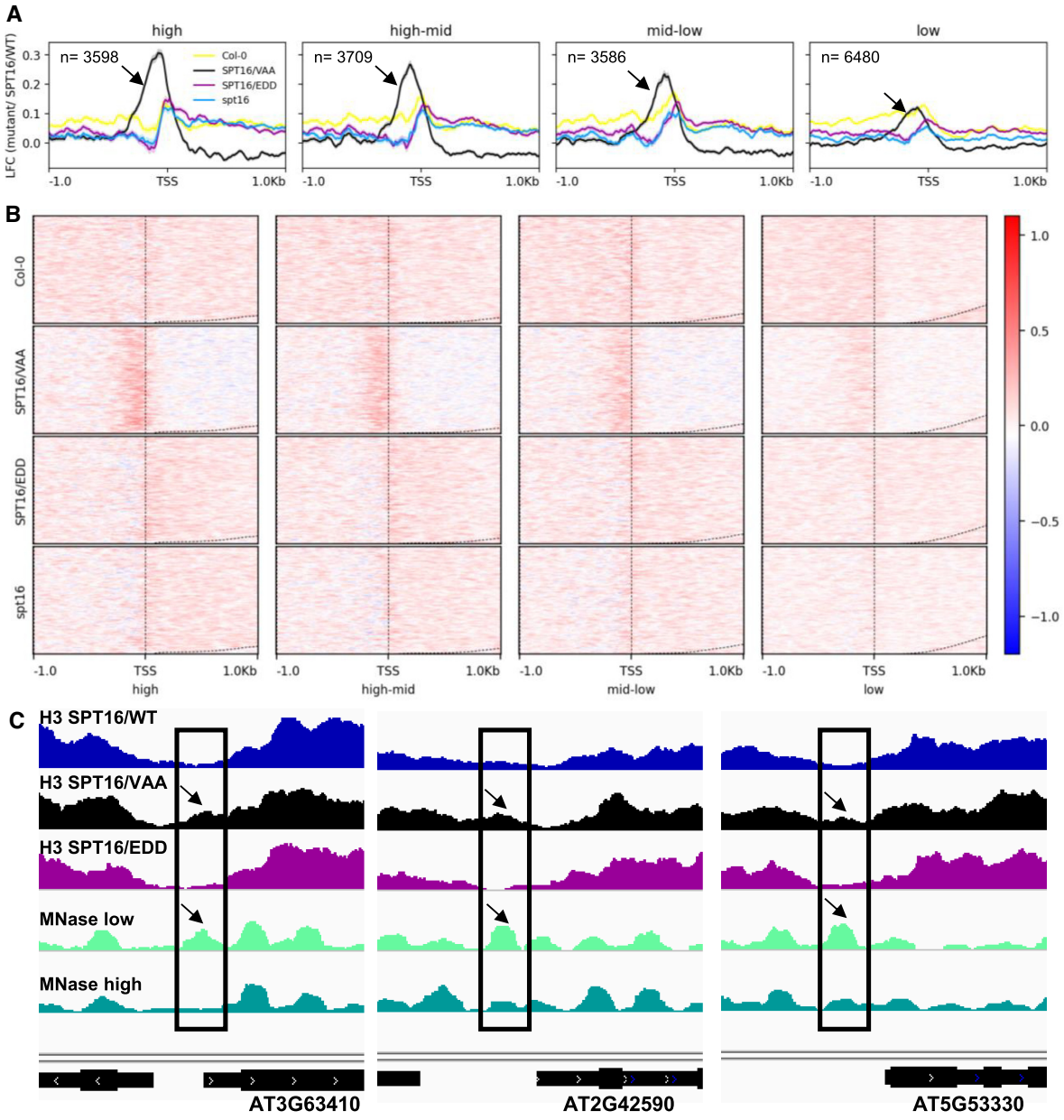



**Figure 6.** Enrichment of histone H3 in the presence of the non-phosphorylatable SPT16/VAA upstream of the TSS dependent on transcription. (**A**) Metagene plots (LFC, log-fold change) of H3 ChIPseq data (10-bp bins) in Col-0, *spt16* and *spt16* expressing SPT16/VAA or SPT16/EDD normalised to those expressing SPT16/WT. Mean signals of the biological replicates were averaged (line) and the tracks represent the SEM for the three replicates at each position (shaded area). The tracks are centred around the TSS and the panels represent non-overlapping protein-coding genes grouped according to transcript levels (high, high-mid, mid-low, low) based on RNAseq data obtained with 6 DAS Col-0 plants. Arrows indicate the transcription dependent peak of H3 enrichment in SPT16/VAA expressing plants. (**B**) Heatmaps illustrating the distribution of H3 in the different plants lines. H3 occupancy is normalised to the plants expressing SPT16/WT and genes in the different panels are grouped related to transcript levels as above. The relative H3 occupancy is colour coded as indicated. (**C**) H3 enrichment upstream of the TSS in presence of SPT16/VAA coincides with differential MNase peaks. Representative examples of H3 ChIPseq profiles representing a 1-kb region around the TSS were aligned with differential MNase (high/low MNase) sequencing data (69). Black boxes indicate a 150-bp window, while the arrows indicate an H3 peak occurring in *spt16* plants expressing the SPT16/VAA variant and a nucleosomal peak that preferentially is observed upon low MNase digestion. Gene models are shown at the bottom.

cial for the interaction with histones of a 112-bp octasome (23). Expression of the varying SPT16 variants in *spt16* plants complemented mutant phenotypes to different extents. SPT16/WT and SPT16/EDD complemented rather efficiently, for instance, the early bolting or increased number of primary inflorescences of *spt16* plants. In contrast, the SPT16/VAA variant proved inactive in reversing the mutant phenotype of *spt16* plants. This demonstrates that the phosphorylation of the SPT16 acidic domain is important for the *in vivo* functionality of FACT.

Analysis of the genomic distribution of histone H3 revealed that mutation of phosphorylation sites in the FACT subunit SPT16 impacted on nucleosome occupancy upstream of the TSS. In presence of the non-phosphorylatable SPT16/VAA variant a peak of H3 enrichment was detected particularly with highly transcribed genes. In contrast to that H3 distribution of plants expressing the phosphorylatable SPT16/WT and the phosphomimic SPT16/EDD variants resembled that of wild-type plants. As expected the association of the SPT16 variants with the regions upstream of the TSSs did not differ significantly. We rather think that SPT16 is phosphorylated by CK2 after FACT recruitment to chromatin. Although both FACT and CK2 associate with the RNAPII elongation complex, their occupancy in chromatin could differ and co-localization of FACT and CK2 may define SPT16 phosphorylation-regulated genomic target sites. The genomic region of the SPT16/VAA-specific H3 enrichment corresponds with a position in *Arabidopsis* chromatin that is characterised as nucleosome depleted region and DNase hypersensitive site (79–82). Increased histone occupancy upstream of the TSS occurred also in yeast cells upon SPT16 depletion (87). Interestingly, the site of H3 enrichment in presence of the non-phosphorylatable SPT16/VAA variant is in accordance with a notable differentially MNase sensitive peak that most likely represents unstable/fragile -1 nucleosomes (69) occurring also in yeast and metazoan chromatin (88–90). In our analyses, this feature correlated with high transcript levels and also in yeast unstable/fragile -1 nucleosomes are linked to transcriptional activity (88). One should expect that H3 enrichment related to persistence of an (unstable/fragile) nucleosome upstream of the TSS interferes with binding of transcription factors and/or RNAPII recruitment, and hence leads to mis-regulation and/or reduced transcriptional output. The role of unstable/fragile nucleosomes currently has not been clarified in plant transcription. In yeast, these unstable nucleosomes are proposed to contribute to poising genes for transcription (91), which is also conceivable in plants (69). Although our ChIPseq analysis convincingly demonstrated distinct H3 accumulation upstream of the TSS at 384 genes, we still expect that the magnitude of the effect is underestimated from our experiments primarily for two reasons: (i) we have mutated three amino acid residues within the acidic domain of SPT16 that were confirmed phosphorylation sites, but as detailed above there are several additional CK2 consensus sites that quite possibly are phosphorylated *in planta* and likely could increase phosphorylation-dependent impact. (ii) The mutated and wild-type SPT16 expression constructs were introduced into the *spt16-1* mutant line, in which SPT16 levels are reduced to <50% (41), but the mutant plants still contain wild-type SPT16 that may mask the effects of the introduced SPT16 variants.

In yeast, the abundant histone chaperones FACT and SPT6 collaborate in managing nucleosomes (87,92) and the activity of SPT6 is also under control of CK2-catalysed phosphorylation (93). The *Arabidopsis* orthologue SPT6L can be recruited to the TSS independent of RNAPII and plays a role during early transcription (94). Therefore, it is conceivable that FACT and SPT6L jointly regulate early transcriptional elongation in plants, since both histone chaperones and the protein kinase associate with RNAPII (45,84). Our identification of CK2-mediated phosphorylation of SPT16 and its implication on nucleosome occupancy at TSSs as well as its influence on plant development paved the way for further studies addressing the interplay of histone chaperones and CK2 in the regulation of chromatin dynamics in the *Arabidopsis* model.

## DATA AVAILABILITY

The ChIPseq data have been deposited to the NCBI sequence read archive with the accession code PRJNA748845, while mass spectrometry data have been deposited to the PRIDE database with the accession code PXD028534.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank Ina Weig-Meckl for performing qPCR analyses, Klaus Tiefenbach for assistance in recording CD spectra, Tina Ravnsborg for help analysing MS data, and the Nottingham Arabidopsis Stock Centre (NASC) for providing *Arabidopsis* T-DNA insertion lines.

## FUNDING

German Research Foundation (DFG) [Gr1159/14-2 and SFB960/A6 to K.D.G.]; proteomics and mass spectrometry research at SDU is supported by generous grants to the VILLUM Center for Bioanalytical Sciences (VILLUM Foundation) [7292 to O.N.J., PRO-MS]; Danish National Mass Spectrometry Platform for Functional Proteomics [5072-00007B to O.N.J.]; Danish National Research Foundation Center for Epigenetics [DNRF 82]. Funding for open access charge: DFG.
*Conflict of interest statement.* None declared.

## REFERENCES

1. Kornberg,R.D. and Lorch,Y. (2020) Primary role of the nucleosome. *Mol. Cell*, **79**, 371–375.
2. Li,B., Carey,M. and Workman,J.L. (2007) The role of chromatin during transcription. *Cell*, **128**, 707–719.
3. Kujirai,T. and Kurumizaka,H. (2020) Transcription through the nucleosome. *Curr. Opin. Struct. Biol.*, **61**, 42–49.
4. Das,C., Tyler,J.K. and Churchill,M.E.A. (2010) The histone shuffle: histone chaperones in an energetic dance. *Trends Biochem. Sci.*, **35**, 476–489.

5. Gurard-Levin,Z.A., Quivy,J.-P. and Almouzni,G. (2014) Histone chaperones: assisting histone traffic and nucleosome dynamics. *Ann. Rev. Biochem.*, **83**, 487–517.
6. Hammond,C.M., Strømme,C.B., Huang,H., Patel,D.J. and Groth,A. (2017) Histone chaperone networks shaping chromatin function. *Nat. Rev. Mol. Cell Biol.*, **18**, 141–158.
7. Brewster,N.K., Johnston,G.C. and Singer,R.A. (1998) Characterization of the CP complex, an abundant dimer of Cdc68 and Pob3 proteins that regulates yeast transcriptional activation and chromatin repression. *J. Biol. Chem.*, **273**, 21972–21979.
8. Orphanides,G., LeRoy,G., Chang,C.-H., Luse,D.S. and Reinberg,D. (1998) FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell*, **92**, 105–116.
9. Orphanides,G., Wu,W.H., Lane,W.S., Hampsey,M. and Reinberg,D. (1999) The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature*, **400**, 284–288.
10. Wittmeyer,J., Joss,L. and Formosa,T. (1999) Spt16 and Pob3 of *Saccharomyces cerevisiae* form an essential, abundant heterodimer that is nuclear, chromatin-associated, and copurifies with DNA polymerase alpha. *Biochemistry*, **38**, 8961–8971.
11. Formosa,T. and Winston,F. (2020) The role of FACT in managing chromatin: disruption, assembly, or repair?*Nucleic Acids Res.*, **48**, 11929–11941.
12. Gurova,K., Chang,H.-W., Valieva,M.E., Sandlesh,P. and Studitsky,V.M. (2018) Structure and function of the histone chaperone FACT - Resolving FACTual issues. *Biochim. Biophys. Acta*, **1861**, 892–904.
13. Belotserkovskaya,R. and Reinberg,D. (2004) Facts about FACT and transcript elongation through chromatin. *Curr. Opin. Genet. Dev.*, **14**, 139–146.
14. Zhou,K., Liu,Y. and Luger,K. (2020) Histone chaperone FACT FAcilitates Chromatin Transcription: mechanistic and structural insights. *Curr. Opin. Struct. Biol.*, **65**, 26–32.
15. Malone,E.A., Clark,C.D., Chiang,A. and Winston,F. (1991) Mutations in SPT16/CDC68 suppress cis- and trans-acting mutations that affect promoter function in Saccharomyces cerevisiae. *Mol. Cell. Biol.*, **11**, 5710–5717.
16. Ransom,M., Williams,S.K., Dechassa,M.L., Das,C., Linger,J., Adkins,M., Liu,C., Bartholomew,B. and Tyler,J.K. (2009) FACT and the proteasome promote promoter chromatin disassembly and transcriptional initiation. *J. Biol. Chem.*, **284**, 23461–23471.
17. Petrenko,N., Jin,Y., Dong,L., Wong,K.H. and Struhl,K. (2019) Requirements for RNA polymerase II preinitiation complex formation in vivo. *Elife*, **8**, e43654.
18. Yu,Y., Yarrington,R.M. and Stillman,D.J. (2020) FACT and Ash1 promote long-range and bidirectional nucleosome eviction at the HO promoter. *Nucleic. Acids. Res.*, **48**, 10877–10889.
19. Belotserkovskaya,R., Oh,S., Bondarenko,V.A., Orphanides,G., Studitsky,V.M. and Reinberg,D. (2003) FACT facilitates transcription-dependent nucleosome alteration. *Science*, **301**, 1090–1093.
20. Hsieh,F.-K., Kulaeva,O.I., Patel,S.S., Dyer,P.N., Luger,K., Reinberg,D. and Studitsky,V.M. (2013) Histone chaperone FACT action during transcription through chromatin by RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 7654–7659.
21. Hondele,M., Stuwe,T., Hassler,M., Halbach,F., Bowman,A., Zhang,E.T., Nijmeijer,B., Kotthoff,C., Rybin,V., Amlacher,S. *et al.* (2013) Structural basis of histone H2A–H2B recognition by the essential chaperone FACT. *Nature*, **499**, 111–114.
22. Kemble,D.J., McCullough,L.L., Whitby,F.G., Formosa,T. and Hill,C.P. (2015) FACT disrupts nucleosome structure by binding H2A–H2B with conserved peptide motifs. *Mol. Cell*, **60**, 294–306.
23. Mayanagi,K., Saikusa,K., Miyazaki,N., Akashi,S., Iwasaki,K., Nishimura,Y., Morikawa,K. and Tsunaka,Y. (2019) Structural visualization of key steps in nucleosome reorganization by human FACT. *Sci. Rep.*, **9**, 10183.
24. Tsunaka,Y., Fujiwara,Y., Oyama,T., Hirose,S. and Morikawa,K. (2016) Integrated molecular mechanism directing nucleosome reorganization by human FACT. *Genes Dev.*, **30**, 673–686.
25. Wang,T., Liu,Y., Edwards,G., Krzizike,D., Scherman,H. and Luger,K. (2018) The histone chaperone FACT modulates nucleosome structure by tethering its components. *Life Sci. Alliance*, **1**, e201800107.
26. Winkler,D.D., Muthurajan,U.M., Hieb,A.R. and Luger,K. (2011) Histone chaperone FACT coordinates nucleosome interaction through multiple synergistic binding events. *J. Biol. Chem.*, **286**, 41883–41892.
27. Liu,Y., Zhou,K., Zhang,N., Wei,H., Tan,Y.Z., Zhang,Z., Carragher,B., Potter,C.S., D'Arcy,S. and Luger,K. (2020) FACT caught in the act of manipulating the nucleosome. *Nature*, **577**, 426–431.
28. Xin,H., Takahata,S., Blanksma,M., McCullough,L., Stillman,D.J. and Formosa,T. (2009) yFACT induces global accessibility of nucleosomal DNA without H2A–H2B displacement. *Mol. Cell*, **35**, 365–376.
29. Valieva,M.E., Armeev,G.A., Kudryashova,K.S., Gerasimova,N.S., Shaytan,A.K., Kulaeva,O.I., McCullough,L.L., Formosa,T., Georgiev,P.G., Kirpichnikov,M.P. *et al.* (2016) Large-scale ATP-independent nucleosome unfolding by a histone chaperone. *Nat. Struct. Mol. Biol.*, **23**, 1111–1116.
30. Kaplan,C.D., Laprade,L. and Winston,F. (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science*, **301**, 1096–1099.
31. Mason,P.B. and Struhl,K. (2003) The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo. *Mol. Cell. Biol.*, **23**, 8323–8333.
32. Cheung,V., Chua,G., Batada,N.N., Landry,C.R., Michnick,S.W., Hughes,T.R. and Winston,F. (2008) Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. *PLoS Biol.*, **6**, e277.
33. Kumar,A. and Vasudevan,D. (2020) Structure-function relationship of H2A–H2B specific plant histone chaperones. *Cell Stress Chaperones*, **25**, 1–17.
34. Tripathi,A.K., Singh,K., Pareek,A. and Singla-Pareek,S.L. (2015) Histone chaperones in *Arabidopsis* and rice: genome-wide identification, phylogeny, architecture and transcriptional regulation. *BMC Plant Biol.*, **15**, 42.
35. Zhu,Y., Dong,A. and Shen,W.H. (2013) Histone variants and chromatin assembly in plant abiotic stress responses. *Biochim. Biophys. Acta*, **1819**, 343–348.
36. Probst,A.V. and Mittelsten Scheid,O. (2015) Stress-induced structural changes in plant chromatin. *Curr. Opin. Plant Biol.*, **27**, 8–16.
37. Otero,S., Desvoyes,B. and Gutierrez,C. (2014) Histone H3 dynamics in plant cell cycle and development. *Cytogenet. Genome Res.*, **143**, 114–124.
38. Grasser,K.D. (2020) The FACT histone chaperone: tuning gene transcription in the chromatin context to modulate plant growth and development. *Front. Plant Sci.*, **11**, 85.
39. Zhou,W., Zhu,Y., Dong,A. and Shen,W.-H. (2015) Histone H2A/H2B chaperones: from molecules to chromatin-based functions in plant growth and development. *Plant J.*, **83**, 78–95.
40. Frost,J.M., Kim,M.Y., Park,G.T., Hsieh,P.H., Nakamura,M., Lin,S.J.H., Yoo,H., Choi,J., Ikeda,Y., Kinoshita,T. *et al.* (2018) *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, E4720–E4729.
41. Lolas,I.B., Himanen,K., Grønlund,J.T., Lynggaard,C., Houben,A., Melzer,M., van Lijsebettens,M. and Grasser,K.D. (2010) The transcript elongation factor FACT affects *Arabidopsis* vegetative and reproductive development and genetically interacts with *HUB1/2*. *Plant J.*, **61**, 686–697.
42. Duroux,M., Houben,A., Ruzicka,K., Friml,J. and Grasser,K.D. (2004) The chromatin remodelling complex FACT associates with actively transcribed regions of the *Arabidopsis* genome. *Plant J.*, **40**, 660–671.
43. Ikeda,Y., Kinoshita,Y., Susaki,D., Iwano,M., Takayama,S., Higashiyama,T., Kakutani,T. and Kinoshita,T. (2011) HMG domain containing SSRP1 is required for DNA demethylation and genomic imprinting in *Arabidopsis*. *Dev. Cell*, **21**, 589–596.
44. Pfab,A., Breindl,M. and Grasser,K.D. (2018) The *Arabidopsis* histone chaperone FACT is required for stress-induced expression of anthocyanin biosynthetic genes. *Plant Mol. Biol.*, **96**, 367–374.
45. 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) Composition of the *Arabidopsis* RNA polymerase II transcript elongation complex reveals the interplay between elongation and mRNA processing factors. *Plant Cell*, **29**, 854–870.

46. Ma,Y., Gil,S., Grasser,K.D. and Mas,P. (2018) Targeted recruitment of the basal transcriptional machinery by LNK clock components controls the circadian rhythms of nascent RNAs in *Arabidopsis*. *Plant Cell*, **30**, 907–924.
47. Perales,M. and Más,P. (2007) A functional link between rhythmic changes in chromatin structure and the *Arabidopsis* biological clock. *Plant Cell*, **19**, 2111–2123.
48. Nielsen,M., Ard,R., Leng,X., Ivanov,M., Kindgren,P., Pelechano,V. and Marquardt,S. (2019) Transcription-driven chromatin repression of Intragenic transcription start sites. *PLoS Genet.*, **15**, e1007969.
49. Krohn,N.M., Stemmer,C., Fojan,P., Grimm,R. and Grasser,K.D. (2003) Protein kinase CK2 phosphorylates the high mobility group domain protein SSRP1, inducing the recognition of UV-damaged DNA. *J. Biol. Chem.*, **278**, 12710–12715.
50. Li,Y., Keller,D.M., Scott,J.D. and Lu,H. (2005) CK2 phosphorylates SSRP1 and inhibits its DNA-binding activity. *J. Biol. Chem.*, **280**, 11869–11875.
51. Aoki,D., Awazu,A., Fujii,M., Uewaki,J.I., Hashimoto,M., Tochio,N., Umehara,T. and Tate,S.I. (2020) Ultrasensitive change in nucleosome binding by multiple phosphorylations to the intrinsically disordered region of the histone chaperone FACT. *J. Mol. Biol.*, **432**, 4637–4657.
52. Tsunaka,Y., Toga,J., Yamaguchi,H., Tate,S., Hirose,S. and Morikawa,K. (2009) Phosphorylated intrinsically disordered region of FACT masks its nucleosomal DNA binding elements. *J. Biol. Chem.*, **284**, 24610–24621.
53. 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.
54. Pfab,A., Grønlund,J.T., Holzinger,P., Längst,G. and Grasser,K.D. (2018) The *Arabidopsis* histone chaperone FACT: role of the HMG-box domain of SSRP1. *J. Mol. Biol.*, **430**, 2747–2759.
55. Zhou,J., Applegate,C., Alonso,A.D., Reynolds,D., Orford,S., Mackiewicz,M., Griffiths,S., Penfield,S. and Pullen,N. (2017) Leaf-GP: an open and automated software application for measuring growth phenotypes for Arabidopsis and wheat. *Plant Meth.*, **13**, 117.
56. van Leene,J., Eeckhout,D., Cannoot,B., Winne,N., Persiau,G., van de Slijke,E., Vercruysse,L., Dedecker,M., Vandepoele,K., Martens,L. *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.
57. Kovalchuk,S.I., Jensen,O.N. and Rogowska-Wrzesinska,A. (2019) FlashPack: fast and simple preparation of ultrahigh-performance capillary columns for LC-MS. *Mol. Cell. Proteomics*, **18**, 383–390.
58. Antosz,W., Deforges,J., Begcy,K., Bruckmann,A., Poirier,Y., Dresselhaus,T. and Grasser,K.D. (2020) Critical role of transcript cleavage in Arabidopsis RNA polymerase II transcriptional elongation. *Plant Cell*, **32**, 1449–1463.
59. Ritt,C., Grimm,R., Fernandez,S., Alonso,J.C. and Grasser,K.D. (1998) Basic and acidic regions flanking the HMG domain of maize HMGa modulate the interactions with DNA and the self-association of the protein. *Biochemistry*, **37**, 2673–2681.
60. Osakabe,A., Lorkovic,Z.J., Kobayashi,W., Tachiwana,H., Yelagandula,R., Kurumizaka,H. and Berger,F. (2018) Histone H2A variants confer specific properties to nucleosomes and impact on chromatin accessibility. *Nucleic. Acids. Res.*, **46**, 7675–7685.
61. Luger,K., Rechsteiner,T.J. and Richmond,T.J. (1999) Expression and purification of recombinant histones and nucleosome reconstitution. *Methods Mol. Biol.*, **119**, 1–16.
62. Stemmer,C., Schwander,A., Bauw,G., Fojan,P. and Grasser,K.D. (2002) Protein kinase CK2 differentially phosphorylates maize chromosomal high mobility group B (HMGB) proteins modulating their stability and DNA interactions. *J. Biol. Chem.*, **277**, 1092–1098.
63. Myers,J.K., Pace,C.N. and Scholtz,J.M. (1997) Helix propensities are identical in proteins and peptides. *Biochemistry*, **36**, 10923–10929.
64. Shechter,D., Dormann,H.L., Allis,C.D. and Hake,S.B. (2007) Extraction, purification and analysis of histones. *Nat. Protoc.*, **2**, 1445–1457.
65. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Meth.*, **9**, 357–359.
66. Chen,K., Xi,Y., Pan,X., Li,Z., Kaestner,K., Tyler,J., Dent,S., He,X. and Li,W. (2013) DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res.*, **23**, 341–351.
67. Ramírez,F., Ryan,D.P., Grüning,B., Bhardwaj,V., Kilpert,F., Richter,A.S., Heyne,S., Dündar,F. and Manke,T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.*, **44**, W160–W165.
68. Ewels,P., Magnusson,M., Lundin,S. and Käller,M. (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, **32**, 3047–3048.
69. Pass,D.A., Sornay,E., Marchbank,A., Crawford,M.R., Paszkiewicz,K., Kent,N.A. and Murray,J.A.H. (2017) Genome-wide chromatin mapping with size resolution reveals a dynamic sub-nucleosomal landscape in Arabidopsis. *PLoS Genet.*, **13**, e1006988.
70. Assenberg,R., Webb,M., Connolly,E., Stott,K., Watson,M., Hobbs,J. and Thomas,J.O. (2008) A critical role in structure-specific DNA binding for the acetylatable lysine residues in HMGB1. *Biochem. J.*, **411**, 553–561.
71. Pelletier,G., Stefanovsky,V.Y., Faubladier,M., Hirschler-Lazkiewicz,I., Savard,J., Rothblum,L.I., Côté,J. and Moss,T. (2000) Competitive recruitment of CBP and Rb-HDAC regulates UBF acetylation and ribosomal transcription. *Mol. Cell*, **6**, 1059–1066.
72. Malarkey,C.S. and Churchill,M.E. (2012) The high mobility group box: the ultimate utility player of a cell. *Trends Biochem. Sci.*, **37**, 553–562.
73. Antosch,M., Mortensen,S.A. and Grasser,K.D. (2012) Plant proteins containing high mobility group box DNA-binding domains modulate different nuclear processes. *Plant Physiol.*, **159**, 875–883.
74. Bonaldi,T., Talamo,F., Scaffidi,P., Ferrera,D., Porto,A., Bachi,A., Rubartelli,A., Agresti,A. and Bianchi,M.E. (2003) Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion. *EMBO J.*, **22**, 5551–5560.
75. Röttgers,K., Krohn,N.M., Lichota,J., Stemmer,C., Merkle,T. and Grasser,K.D. (2000) DNA-interactions and nuclear localisation of the chromosomal HMG domain protein SSRP1 from maize. *Plant J.*, **23**, 395–405.
76. Litchfield,D.W. (2003) Protein kinase CK2: structure, regulation and role in cellular decisions of life and death. *Biochem. J.*, **369**, 1–15.
77. Wienholz,F., Zhou,Di, Turkyilmaz,Y., Schwertman,P., Tresini,M., Pines,A., van Toorn,M., Bezstarosti,K., Demmers,J.A.A. and Marteijn,J.A. (2019) FACT subunit Spt16 controls UVSSA recruitment to lesion-stalled RNA Pol II and stimulates TC-NER. *Nucleic Acids Res.*, **47**, 4011–4025.
78. Vanssay,A., Touzeau,A., Arnaiz,O., Frapporti,A., Phipps,J. and Duharcourt,S. (2020) The Paramecium histone chaperone Spt16-1 is required for Pgm endonuclease function in programmed genome rearrangements. *PLoS Genet.*, **16**, e1008949.
79. Stroud,H., Otero,S., Desvoyes,B., Ramirez-Parra,E., Jacobsen,S.E. and Gutierrez,C. (2012) Genome-wide analysis of histone H3.1 and H3.3 variants in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 5370–5375.
80. Zhang,T., Zhang,W. and Jiang,J. (2015) Genome-wide nucleosome occupancy and positioning and their impact on gene expression and evolution in plants. *Plant Physiol.*, **168**, 1406–1416.
81. 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.
82. Zhang,W., Zhang,T., Wu,Y. and Jiang,J. (2012) Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in Arabidopsis. *Plant Cell*, **24**, 2719–2731.
83. Vera,D.L., Madzima,T.F., Labonne,J.D., Alam,M.P., Hoffman,G.G., Girimurugan,S.B., Zhang,J., McGinnis,K.M., Dennis,J.H. and Bass,H.W. (2014) Differential nuclease sensitivity profiling of chromatin reveals biochemical footprints coupled to gene expression and functional DNA elements in maize. *Plant Cell*, **26**, 3883–3893.
84. Krogan,N.J., Kim,M., Ahn,S.H., Zhong,G., Kobor,M.S., Cagney,G., Emili,A., Shilatifard,A., Buratowski,S. and Greenblatt,J.F. (2002) RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a targeted proteomics approach. *Mol.Cell.Biol.*, **22**, 6979–6992.
85. Vilela,B., Pagès,M. and Riera,M. (2015) Emerging roles of protein kinase CK2 in abscisic acid signaling. *Front. Plant Sci.*, **6**, 966.

86. Mulekar,J.J. and Huq,E. (2014) Expanding roles of protein kinase CK2 in regulating plant growth and development. *J. Exp. Bot.*, **65**, 2883–2893.
87. Jeronimo,C., Poitras,C. and Robert,F. (2019) Histone recycling by FACT and Spt6 during transcription prevents the scrambling of histone modifications. *Cell Rep.*, **28**, 1206–1218.
88. Kubik,S., Bruzzone,M.J. and Shore,D. (2017) Establishing nucleosome architecture and stability at promoters: roles of pioneer transcription factors and the RSC chromatin remodeler. *Bioessays*, **39**, 1600237.
89. Mieczkowski,J., Cook,A., Bowman,S.K., Mueller,B., Alver,B.H., Kundu,S., Deaton,A.M., Urban,J.A., Larschan,E., Park,P.J. *et al.* (2016) MNase titration reveals differences between nucleosome occupancy and chromatin accessibility. *Nat. Commun.*, **7**, 11485.
90. Schwartz,U., Németh,A., Diermeier,S., Exler,J.H., Hansch,S., Maldonado,R., Heizinger,L., Merkl,R. and Längst,G. (2019) Characterizing the nuclease accessibility of DNA in human cells to map higher order structures of chromatin. *Nucleic Acids Res.*, **47**, 1239–1254.
91. Xi,Y., Yao,J., Chen,R., Li,W. and He,X. (2011) Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation. *Genome Res.*, **21**, 718–724.
92. McCullough,L., Connell,Z., Petersen,C. and Formosa,T. (2015) The abundant histone chaperones Spt6 and FACT collaborate to assemble, inspect, and maintain chromatin structure in *Saccharomyces cerevisiae*. *Genetics*, **201**, 1031–1045.
93. Gouot,E., Bhat,W., Rufiange,A., Fournier,E., Paquet,E. and Nourani,A. (2018) Casein kinase 2 mediated phosphorylation of Spt6 modulates histone dynamics and regulates spurious transcription. *Nucleic Acids Res.*, **46**, 7612–7630.
94. Chen,C., Shu,J., Li,C., Thapa,R.K., Nguyen,V., Yu,K., Yuan,Z.-C., Kohalmi,S.E., Liu,J., Marsolais,F. *et al.* (2019) RNA polymerase II-independent recruitment of SPT6L at transcription start sites in Arabidopsis. *Nucleic Acids Res.*, **47**, 6714–6725.

# Chapter 6

# Elongation factor 1 is a component of the *Arabidopsis* RNA polymerase II elongation complex and associates with a subset of transcribed genes

This peer-reviewed article was published in the journal *New Phytologist* in 2023

New Phytologist

Research

# Elongation factor 1 is a component of the *Arabidopsis* RNA polymerase II elongation complex and associates with a subset of transcribed genes

**Hanna Markusch[1], Philipp Michl-Holzinger[1], Simon Obermeyer[1], Claudia Thorbecke[1], Astrid Bruckmann[2], Sabrina Babl[3], Gernot Längst[3], Akihisa Osakabe[4], Frédéric Berger[4] and Klaus D. Grasser[1]**

[1]Cell Biology & Plant Biochemistry, Centre for Biochemistry, University of Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany; [2]Institute for Biochemistry I, Centre for Biochemistry, University of Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany; [3]Institute for Biochemistry III, Centre for Biochemistry, University of Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany; [4]Gregor Mendel Institute (GMI), Austrian Academy of Sciences, Vienna BioCenter (VBC), Dr. Bohr-Gasse 3, 1030 Vienna, Austria

Author for correspondence:
*Klaus D. Grasser*
*Email:* klaus.grasser@ur.de

Received: *20 July 2022*
Accepted: *24 December 2022*

*New Phytologist* (2023) **238:** 113–124
**doi**: 10.1111/nph.18724

**Key words:** *Arabidopsis thaliana*, chromatin, gene activity, mRNA synthesis, RNA polymerase II, transcript elongation.

## Summary

- Elongation factors modulate the efficiency of mRNA synthesis by RNA polymerase II (RNAPII) in the context of chromatin, thus contributing to implement proper gene expression programmes. The zinc-finger protein elongation factor 1 (ELF1) is a conserved transcript elongation factor (TEF), whose molecular function so far has not been studied in plants.
- Using biochemical approaches, we examined the interaction of *Arabidopsis* ELF1 with DNA and histones *in vitro* and with the RNAPII elongation complex *in vivo*. In addition, cytological assays demonstrated the nuclear localisation of the protein, and by means of double-mutant analyses, interplay with genes encoding other elongation factors was explored. The genome-wide distribution of ELF1 was addressed by chromatin immunoprecipitation.
- ELF1 isolated from *Arabidopsis* cells robustly copurified with RNAPII and various other elongation factors including SPT4-SPT5, SPT6, IWS1, FACT and PAF1C. Analysis of a CRISPR-Cas9-mediated gene editing mutant of *ELF1* revealed distinct genetic interactions with mutants deficient in other elongation factors. Moreover, ELF1 associated with genomic regions actively transcribed by RNAPII. However, ELF1 occupied only *c.* 33% of the RNAPII transcribed loci with preference for inducible rather than constitutively expressed genes.
- Collectively, these results establish that *Arabidopsis* ELF1 shares several characteristic attributes with RNAPII TEFs.

## Introduction

Production of mRNAs by RNA polymerase II (RNAPII) is characterised by several consecutive steps including promoter recruitment, initiation, elongation and termination. Over the last years, it became clear that in addition to initiation, the elongation stage is dynamic and heavily regulated. Consequently, a range of transcript elongation factors (TEFs) was identified that facilitate efficient mRNA synthesis from nucleosomal templates (Sims *et al.*, 2004; Kwak & Lis, 2013; Chen *et al.*, 2018; Osman & Cramer, 2020). During ongoing transcription, TEFs serve diverse functions assisting the progression of RNAPII through repressive chromatin. Accordingly, some TEFs modulate the catalytic properties and processivity of RNAPII, while others act as histone chaperones or chromatin remodelling factors assisting progression of RNAPII through nucleosomes. Another group of TEFs modifies transcribed chromatin through deposition/erasure of transcription-related histone marks such as mono-ubiquitination or a variety of methylations and acetylations (Sims *et al.*, 2004; Kwak & Lis, 2013; Chen *et al.*, 2018; Osman & Cramer, 2020).

A specific example of the heterogenous family of TEFs is elongation factor 1 (ELF1) that originally was identified in a yeast genetic screen by virtue of the synthetic lethality of the *elf1Δ* mutant in combination with mutations in genes encoding other known TEFs (Prather *et al.*, 2005). ELF1 is a small zinc-finger protein that is conserved in eukaryotes and some archaea. It was found to interact functionally with various other TEFs including SPT4-SPT5, TFIIS, FACT and PAF1C, and it preferentially localises to genomic regions that are actively transcribed by RNAPII (Prather *et al.*, 2005; Mayer *et al.*, 2010; Rossi *et al.*, 2021). Based on its steady association with the yeast RNAPII elongation complex during *in vitro* transcription, ELF1 was designated as core elongation factor (Joo *et al.*, 2019). Using a combination of cryo-EM and X-ray crystallography, it recently could be clarified that ELF1 directly interacts with RNAPII at the DNA entry tunnel of downstream DNA (Ehara *et al.*, 2017). Further *in vitro* studies demonstrated that ELF1 (particularly in cooperation with

SPT4-SPT5) promoted progression of RNAPII through the nucleosomal barrier during transcriptional elongation (Ehara *et al.*, 2019). The human orthologue ELOF1 can facilitate RNAPII ubiquitination and is also involved in transcription-coupled DNA repair (van der Weegen *et al.*, 2021). In archaea, ELF1 was identified as part of the elongation complex, promoting productive transcript elongation (Blombach *et al.*, 2021).

In plants, various TEFs have been identified and studies primarily in *Arabidopsis* have demonstrated that by establishing proper gene expression programmes they modulate growth and development (van Lijsebettens & Grasser, 2014). A number of TEFs associate with elongating *Arabidopsis* RNAPII to form the active elongation complex that also integrates the cooperation with cotranscriptional processes (Antosz *et al.*, 2017). *Arabidopsis* factors modulating RNAPII properties during elongation (e.g. TFIIS and SPT4-SPT5) influence hormone signalling, stress response and germination (Grasser *et al.*, 2009; Dürr *et al.*, 2014; Antosz *et al.*, 2020; Szádeczky-Kardoss *et al.*, 2022), while transcription-related histone chaperones (e.g. SPT6L, FACT) are required for normal embryo development, seedling establishment and developmental transitions (Lolas *et al.*, 2010; Gu *et al.*, 2012; Chen *et al.*, 2019). Various factors that affect the deposition of histone marks over transcribed regions (e.g. PAF1-C, HUB1/2 and SDG proteins) have profound effects on many aspects of plant growth and development (He *et al.*, 2004; Oh *et al.*, 2004; Fleury *et al.*, 2007; Xu *et al.*, 2008; Berr *et al.*, 2010; Fiorucci *et al.*, 2019).

A putative orthologue of yeast ELF1 (and mammalian ELOF1) is encoded in *Arabidopsis* (and other plant genomes), but so far, its molecular function has not been studied. Therefore, we have characterised ELF1 from *Arabidopsis* that proved to be a nuclear protein associated with elongating RNAPII. ELF1 interacts *in vitro* with DNA and histones, and it localises to genomic regions that are actively transcribed by RNAPII. Examination of respective mutant plants demonstrated genetic interactions between *ELF1* and genes encoding other known TEFs. Taken together, these findings illustrate that *Arabidopsis* ELF1 is the counterpart of yeast ELF1 (and mammalian ELOF1) acting in RNAPII transcriptional elongation.

## Materials and Methods

### Recombinant protein production

Using pET24 expression vectors (Supporting Information Table S1), ELF1 proteins fused to a GB1/6xHis-tag were expressed in *Escherichia coli* and bound to Ni-NTA-agarose (Qiagen). Full-length and truncated ELF1 proteins were eluted by cleavage with the HRV 3C protease (Sigma), removing the GB1/6xHis-tag. Subsequently, recombinant proteins were purified by ion-exchange FPLC using a Resource S Column (GE Healthcare, Freiburg, Germany) and proteins were eluted with a linear gradient 0–1 M NaCl in buffer D (10 mM sodium phosphate, pH 7.0, 1 mM EDTA, 1 mM dithiothreitol, 0.5 mM PMSF) as described previously (Grasser *et al.*, 1996). Recombinant *Arabidopsis* histones H2A-H2B were prepared as described previously (Osakabe *et al.*, 2018). Purified proteins were characterised by SDS-PAGE and mass spectrometry.

### Electrophoretic mobility shift assays

DNA-binding of recombinant ELF1 proteins was examined by electrophoretic mobility shift assays (EMSAs) with Cy5-labelled DNA oligonucleotides (Table S2) that were either annealed to linear double strands or to four-way junctions (4wjs). Binding reactions were analysed in 6% polyacrylamide 1× TBE gels and DNA was visualised using a ChemiDoc MP Imaging System (Bio-Rad). Interactions with nucleosomes and the corresponding nucleosomal DNA were examined as described previously (Pfab *et al.*, 2018).

### Intramolecular crosslinking with EDC

Recombinant ELF1 proteins were incubated for different periods in buffer D (10 mM sodium phosphate, pH 7.0, 1 mM EDTA, 1 mM dithiothreitol, 0.5 mM PMSF) with a final concentration of 20 mM 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC; Sigma) from a freshly prepared stock as described previously (Thomsen *et al.*, 2004). Subsequently, samples were analysed by SDS-PAGE and Coomassie staining.

### GST pull-down assays

Protein interaction assays were essentially performed as described previously (Michl-Holzinger *et al.*, 2022), with recombinant ELF1 proteins fused to glutathione S transferase (GST) mixed with equimolar amounts of recombinant *Arabidopsis* H2A-H2B or with bovine cytochrome C (CytC; Sigma Aldrich) in GST buffer (0.2–0.35 M NaCl, 25 mM HEPES pH 7.6, 0.05% (v/v) NP40, 5 mM DTT, 10% (v/v) glycerol, 2 mM $MgCl_2$). Following incubation (30 min at 30°C), glutathione-sepharose beads were added and the samples were incubated for 3 h on a rotating wheel at 4°C. Beads were washed three times in GST buffer before bound proteins were eluted by boiling in protein loading buffer and analysed by SDS-PAGE.

### Affinity purification and characterisation of GS-tagged ELF1 from *Arabidopsis* cells

*Arabidopsis* suspension-cultured PSB-D cells were maintained and transformed as described previously (van Leene *et al.*, 2015). Protein isolation, purification of GS-tagged (protein G and streptavidin-binding peptide-tagged) ELF1 using IgG-coupled magnetic beads, mass spectrometry and data analyses including removal of experimental background were performed as described previously (Dürr *et al.*, 2014; Antosz *et al.*, 2017).

### Plant material

Seeds of *Arabidopsis thaliana* (L.) Heynh. (ecotype Col-0) were stratified in darkness for 48 h at 4°C, and plants were grown at 21°C on soil in a phytochamber or on MS medium in plant

incubators (PolyKlima, Freising, Germany) under long-day conditions (Antosz *et al.*, 2017; Michl-Holzinger *et al.*, 2022). Col-0 plants expressing eGFP-NLS were described previously (Pfab *et al.*, 2018).

### CRISPR-Cas9-mediated gene editing

CRISPR-Cas9 gene-edited plant lines were generated in the Col-0 background utilising egg cell-specific promoter-controlled Cas9 vector systems (Wang *et al.*, 2015). A specific sgRNA sequence was selected using the tool CRISPR-P 2.0 (Liu *et al.*, 2017). For the *ELF1* sgRNA, complementary oligonucleotides (Table S2) were annealed and inserted into pHEE401E (Wang *et al.*, 2015) via Golden Gate assembly using BsaI restriction sites, generating the transformation vector (Table S1) that was introduced into *Arabidopsis* Col-0 by *Agrobacterium*-mediated transformation as described previously (Lolas *et al.*, 2010; Dürr *et al.*, 2014). For genotyping of CRISPR-Cas9 mutants, genomic DNA was isolated from leaves and used as template for PCR with primers specified in Table S2, before the amplified fragments were analysed by DNA sequencing.

### Generation and analysis of *Arabidopsis* double mutants

*Arabidopsis* double-mutant plants (all ecotype Col-0) were generated by genetic crossing, as described previously (Lolas *et al.*, 2010), of gene-edited *elf1-1* and *tfIIs-1* (SALK_056755; Grasser *et al.*, 2009), *ssrp1-2* (SALK_001283; Lolas *et al.*, 2010) or *iws1-1* (SALK_056238; Li *et al.*, 2010). Complementation analyses were performed with double-mutant plants crossed with *elf1-1*$^{C1}$ (to be described later). Obtained plants were verified by PCR-based genotyping (using primers listed in Table S2) and DNA sequencing. Plant phenotypes were generally analysed and documented as described previously (Lolas *et al.*, 2010; Dürr *et al.*, 2014), while leaf serration was determined according to Bilsborough *et al.* (2011) and leaf angles according to Hopkins *et al.* (2008).

### Confocal laser scanning microscopy

Confocal laser scanning microscopy (CLSM) was performed using a Zeiss LSM 980 Airyscan2, equipped with a 10× NA 0.3, a 20× NA 0.3, a 40× Oil 1.3 or 63× Oil NA 1.3 objective. eGFP and propidium iodide (PI) were excited using a VIS Laser at 488 and 561 nm, respectively. The emission of eGFP was detected at 500–550 nm, while the emission of PI was detected at 570–620 nm.

### ChIP sequencing

Chromatin immunoprecipitation (ChIP) was essentially performed as described previously (Antosz *et al.*, 2020; Michl-Holzinger *et al.*, 2022). *Arabidopsis* plants (14 d after stratification, 14-DAS *in vitro* grown) were crosslinked with formaldehyde and used for isolation of nuclei, before chromatin was sheared using a Bioruptor Pico device (Diagenode, Seraing, Belgium). Immunoprecipitation was performed using antibodies directed against GFP (ab290; Abcam, Berlin, Germany) and S2P-modified (Ser2-phosphorylated heptapeptides of the carboxy-terminal domain of NRPB1) RNAPII (ab5095; Abcam). For ChIP sequencing (ChIP-Seq), libraries were generated using NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, Frankfurt, Germany) and the final libraries (four and three replicates each for ELF1-eGFP and S2P-RNAPII, respectively) were sequenced by the Genomics Core Facility at the University of Regensburg (http://www.kfb-regensburg.de/) using NextSeq 2000 (Illumina, San Diego, CA, USA). Reads were aligned to the TAIR10 genome (https://www.arabidopsis.org/) using Bowtie2 (Langmead & Salzberg, 2012), and coverage tracks were calculated with deepTools 'bamCoverage' (Table S3). Downstream analysis was mainly performed using the deepTools2 suite (v.3.5.0; Ramírez *et al.*, 2016), and quality control was performed at several steps using FastQC (Ewels *et al.*, 2016). In case of RNAPII-S2P, 12.3–18.8 M reads and for ELF1-eGFP 12.1–23.1 M reads mapped against the TAIR10 genome. Genomic regions with aberrant coverage or low sequence complexity were filtered out, as described previously (Quadrana *et al.*, 2016). After confirming high pairwise correlations, the biological replicates were merged and CPM normalised. Macs3 callpeak was used to call peaks over input as control. Gene Ontology (GO)-term enrichments have been performed with Shiny GO (Ge *et al.*, 2020). Expression levels of genes (divided into genes > 28 and < 0.01 TPMs) were deduced from RNA sequencing data (Michl-Holzinger *et al.*, 2022) of *in vitro* grown 6-DAS Col-0 plants under standard conditions, which was performed in three replicates.

## Results

### ELF1 interacts with DNA, histones and nucleosomes

To identify putative *Arabidopsis* orthologues of yeast ELF1 and human ELOF1, we searched the *Arabidopsis* database (https://www.arabidopsis.org/) with the Blastp program using these amino acid sequences as a query. The search revealed a single clear hit (AGI locus At5g46030) encoding a 120-aa protein (13.9 kDa) rich in charged amino acid residues, in the following termed *Arabidopsis* ELF1. Alignments revealed that *Arabidopsis* ELF1 shares 32.9%, 33.1%, 59.2% and 81.0% amino acid sequence identity with its orthologues from *Saccharomyces cerevisiae*, *Homo sapiens*, *Oryza sativa* and *Brassica napus*, respectively (Fig. S1). It is comprised of a basic N-terminal region, the central Zn-finger region, the RNAPII-binding region and an acidic C-terminal region. A major part of the ELF1 sequences is considerably conserved, but the acidic region is lacking in metazoan sequences and is quite heterogeneous in plant sequences (Fig. S1). For some reason, the acidic region appears to be more extended in ELF1 sequences of *Brassicacea*.

Recombinant ELF1 proteins were expressed in *E. coli* and purified to examine their molecular interactions. In addition to full-length ELF1, truncated versions lacking the basic N-terminal region (ELF1ΔN) or the acidic C-terminal region (ELF1ΔC) were produced (Fig. 1a,b). DNA binding of the ELF1 proteins was analysed using EMSAs with four-way junction (4wj) DNA

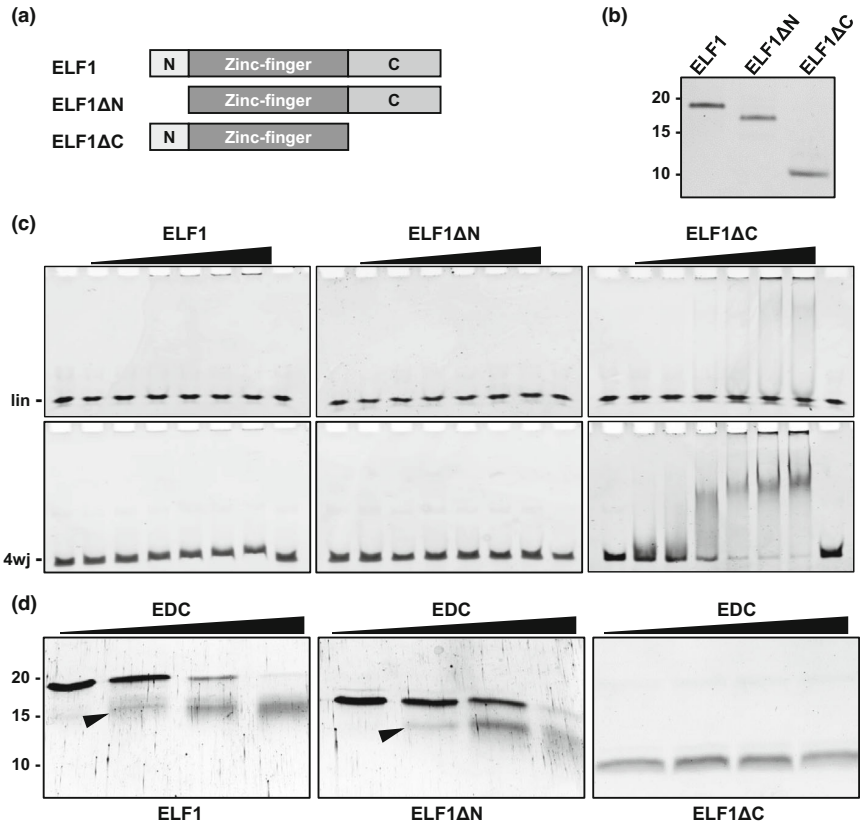



**Fig. 1** DNA interactions of ELF1 are inhibited by the acidic C-terminal region. (a) Schematic illustration of ELF1 and truncated versions. N, basic N-terminal region; zinc-finger region; C, acidic C-terminal region. (b) Purified recombinant ELF1 proteins analysed by SDS-PAGE and stained with Coomassie. (c) Electrophoretic mobility shift assays of increasing concentrations of the indicated ELF1 proteins (0, 1, 2.5, 5, 7.5, 10, 15, 0 μM, from left to right) and linear (lin, upper) or four-way junction DNA (lower). (d) The indicated ELF1 proteins were chemically crosslinked for increasing periods (0, 5, 10 and 20 min) with EDC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide), before proteins were separated on tricine polyacrylamide gels and stained with Coomassie. Intramolecular crosslinking is evident with ELF1 and ELF1ΔN from the reaction time-dependent appearance of an additional, faster migrating protein band (arrowhead).

or the corresponding linear double-stranded DNA. Four-way junction DNA was chosen, since it resembles some features of a nucleosome, as two juxtaposed DNA duplexes of the junction in a way mimic the entry/exit point of the nucleosomal DNA (Zlatanova & van Holde, 1998).

Increasing concentrations of the ELF1 proteins were incubated with the DNA, and the formation of protein/DNA complexes was examined by electrophoresis. Binding of ELF1 was detectable only with the 4wj DNA, resulting (relative to the DNA in absence of protein) in slightly reduced migration of the DNA band with increasing protein input (Fig. 1c, left). No DNA interaction was traceable with ELF1ΔN (Fig. 1c, middle), whereas marked binding of ELF1ΔC to both types of DNA was observed (Fig. 1c, right). As evident from the disappearance of the free DNA, ELF1ΔC interacted with higher affinity with the 4wj DNA than with the linear DNA, although in both cases the relatively fuzzy migration of the protein/DNA complexes suggests rather nonspecific interaction. The remarkably increased DNA interaction of ELF1ΔC likely is caused by the absence of the acidic C-terminal region that might interact intramolecularly with basic parts of the protein. Chemical crosslinking using the zero-length agent EDC (which cross-links carboxyl and amino groups) was used to address possible intramolecular interactions of the acidic C-terminal region in ELF1. Full-length and truncated versions of ELF1 were reacted with EDC for various periods and subsequently analysed by SDS-PAGE. Intramolecular crosslinking is evident with ELF1 and ELF1ΔN from the reaction time-dependent appearance of an additional, faster migrating protein band (Fig. 1d). By contrast, ELF1ΔC lacking the acidic region could not be crosslinked, suggesting that in ELF1 and ELF1ΔN, the acidic region interacts with basic parts of the protein (basic N-terminal region and/or central region), thereby reducing the affinity for DNA.

The interaction of ELF1 with histones was studied using GST pull-down assays. ELF1 and ELF1ΔC were produced as GST fusion proteins and immobilised on gluthatione-sepharose beads. Added H2A-H2B bound to GST-ELF1, but not to GST-ELF1ΔC (Fig. 2a), suggesting that the acidic C-terminal region is required for histone interaction. Similar to the histone interaction of the acidic region of SPT16 (Michl-Holzinger *et al.*, 2022), the specificity of the ELF1–histone interaction is evident from the lack of binding of H2A-H2B to unfused GST and the lack of binding of the small basic protein CytC to GST-ELF1 in the GST pull-down assay.

Using EMSAs, the interaction of ELF1 proteins with reconstituted nucleosomes (and the corresponding 147-bp DNA fragment) was examined. No interaction of ELF1 and ELF1ΔN with DNA or nucleosomes could be detected, whereas ELF1ΔC bound both to DNA and nucleosome particles (Fig. 2b). With increasing protein concentration, the migration of the nucleosome band was gradually more retarded. In conclusion, the *in vitro* interaction studies reveal that the acidic C-terminal region of ELF1 is required for the binding to histones, but it inhibits the binding of ELF1 to DNA and nucleosomes.

New Phytologist





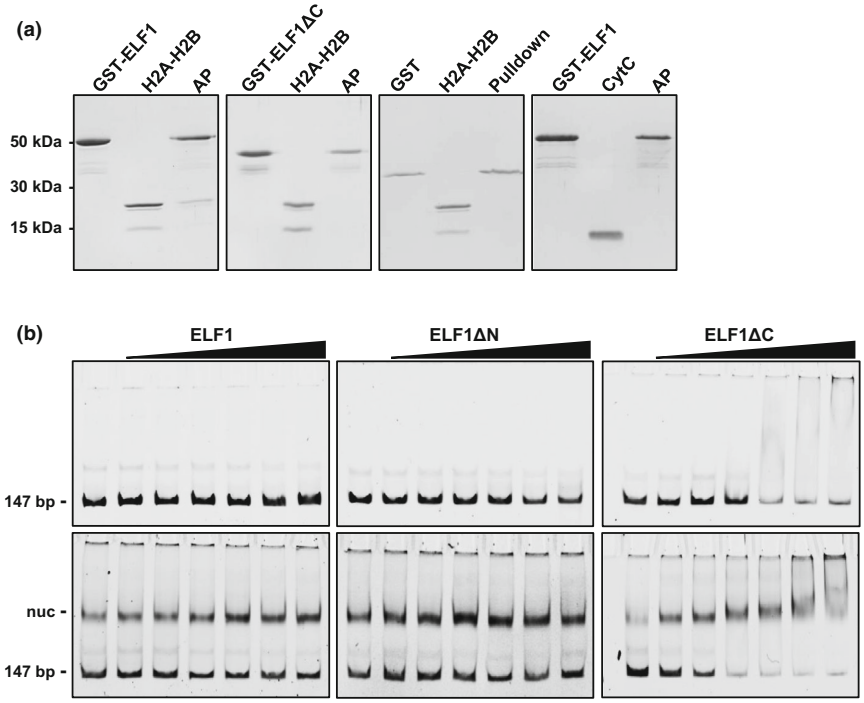



**Fig. 2** ELF1 C-terminal region inhibits interactions with DNA and nucleosomes, but is required for histone binding. (a) Interaction of ELF1 with histones analysed by glutathione S transferase GST pull-down assays. The indicated GST-ELF1 fusion proteins (or unfused GST) were incubated with *Arabidopsis* H2A-H2B histones (or with cytochrome C, CytC). Binding reactions were bound to glutathione sepharose and the eluates of the affinity purification (AP) were analysed along with the input samples by SDS–PAGE and Coomassie staining. Fifty per cent of each input was loaded on an 18% SDS gel. (b) Interaction of full-length and truncated ELF1 proteins with DNA and human nucleosomes analysed by electrophoretic mobility shift assays. Increasing concentrations of the indicated ELF1 proteins (0, 1, 2.5, 5, 7.5, 10 and 15 μM, from left to right) were incubated with 147-bp DNA (upper) or a mixture of 147-bp DNA and reconstituted human mono-nucleosomes containing the 147-bp DNA (lower).

## ELF1 is a component of the RNAPII transcript elongation complex

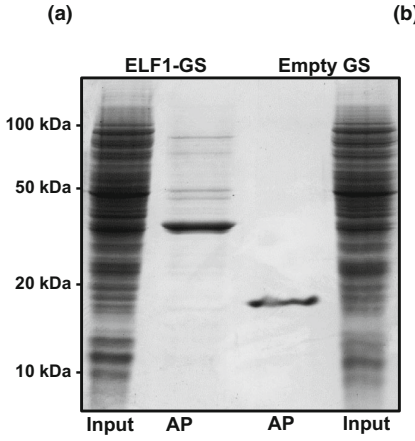
To investigate the possible association with RNAPII, ELF1 was expressed as GS-tagged fusion protein in *Arabidopsis* PSB-D suspension-cultured cells, an approach that has been used to study other RNAPII-associated proteins (Dürr *et al.*, 2014; Antosz *et al.*, 2017). ELF1-GS and interacting proteins were isolated from cell extracts by IgG affinity purification using mild conditions to preserve protein interactions. SDS–PAGE analysis revealed the enrichment of the ELF1-GS bait protein along with a number of copurifying protein bands that were not detected with the unfused GS-tag (Fig. 3a). Proteins in the eluates were identified after tryptic digestion by mass spectrometry. Strikingly, a number of RNAPII subunits were found to robustly copurify with ELF1 (Fig. 3b). In addition, several TEFs reproducibly co-eluted efficiently with ELF1 and RNAPII such as FACT (SSRP1 and SPT16), SPT4-SPT5, SPT6L, IWS1 and PAF1C (ELF7, VIP3, VIP4 and CDC73). These interactors were not identified in the control experiment with the unfused GS-tag or in comparable IgG affinity purifications of unrelated nuclear proteins (Municio *et al.*, 2021; Cheng *et al.*, 2022). This result suggests that *Arabidopsis* ELF1 along with other TEFs associates with RNAPII to form the transcript elongation complex in accord with the situation in yeast (Ehara *et al.*, 2017, 2019).

## Functional inactivation of *ELF1* causes mild effects and ELF1 localises to cell nuclei

No useful T-DNA insertion mutant could be identified to study the *in planta* function of ELF1. Therefore, we generated CRISPR-Cas9-mediated gene-edited lines in the Col-0 background for that purpose. The gene editing resulted in insertion of an additional base pair within the coding sequence of exon 1 leading to a frameshift and a premature translational stop codon (Fig. S2a), hereafter referred to as *elf1-1*. The mutant plants are basically indistinguishable from the Col-0 wild-type, except that the height of mature plants is slightly increased (Fig. S2b). Likewise, the *elf1Δ* mutation caused no significant growth defect in yeast (Prather *et al.*, 2005). Following transformation of *elf1-1* plants with a complementation construct, the expression of an ELF1-eGFP fusion protein under control of the *ELF1* promoter (*c.* 4 kb upstream of transcriptional start site (TSS)) in three independent plant lines (termed *elf1-1*$^{C1}$, *elf1-1*$^{C5}$ and *elf1-1*$^{C6}$) did not affect plant phenotype, besides reducing the increased height of *elf1-1* (Fig. S2c–e). Analysis of these plants using CLSM demonstrated that ELF1-eGFP localises to the cell nucleus (Fig. 4), as reported for human ELOF1 (van der Weegen *et al.*, 2021). In addition, the microscopic survey detected ELF1-eGFP in all analysed *Arabidopsis* leaf and root cells (Fig. 4), suggesting ubiquitous expression, consistent with publicly available *Arabidopsis* mRNA expression data regarding *ELF1* (Winter *et al.*, 2007).

## Genetic interaction of *ELF1* with genes coding for other transcript elongation factors

To further elucidate the involvement of ELF1 in elongation by RNAPII, we generated double-mutant deficient in ELF1 and different previously characterised TEFs. Thus, *elf1-1* was crossed with *tfIIs-1* (Grasser *et al.*, 2009), *ssrp1-2* (Lolas *et al.*, 2010) and *iws1-1* (Li *et al.*, 2010). The obtained double-mutant plants were




(b)

**Transcript elongation related proteins copurifying with ELF1-GS**

| Detected | Interactor | AGI | Complex |
|---|---|---|---|
| 2597.5/3 | NRPB1 | AT4G35800 | RNAPII |
| 2616.6/3 | NRPB2 | AT4G21710 | RNAPII |
| 1296.3/3 | NRPB3 | AT2G15430 | RNAPII |
| 421.8/3 | NRPB5A | AT3G22320 | RNAPII |
| 340.2/2 | NRPB7 | AT5G59180 | RNAPII |
| 134.1/2 | NRPB8A | AT1G54250 | RNAPII |
| 181.3/2 | NRPB8B | AT3G59600 | RNAPII |
| 131.9/2 | NRPB9A | AT3G16980 | RNAPII |
| 132.9/2 | NRPB11 | AT3G52090 | RNAPII |
| 1355.9/3 | SPT16 | AT4G10710 | FACT |
| 816.3/3 | SSRP1 | AT3G28730 | FACT |
| 87.2/2 | SPT4-1 | AT5G08565 | SPT4-SPT5 |
| 1133.2/3 | SPT5-2 | AT4G08350 | SPT4-SPT5 |
| 791.4/3 | ELF1 | AT5G46030 | ELF1 |
| 1040.3/2 | SPT6L | AT1G65440 | SPT6 |
| 359.7/3 | IWS1 | AT1G32130 | IWS1 |
| 230.5/3 | ELF7 | AT1G79730 | PAF1C |
| 343.9/3 | VIP4 | AT5G61150 | PAF1C |
| 92.3/3 | VIP5 | AT1G61040 | PAF1C |
| 332.7/3 | CDC73 | AT3G22590 | PAF1C |

**Fig. 3** Affinity purification of ELF1 from *Arabidopsis* cells. (a) Total protein extract of PSB-D cells expressing ELF1-GS or the unfused GS-tag (input) was analysed by SDS-PAGE and Coomassie staining, along with the eluate of the corresponding IgG affinity purifications (AP). (b) Transcript elongation-related proteins (i.e. transcript elongation factors, RNAPII subunits) that copurified with ELF1-GS, as identified by mass spectrometry. Numbers in the left column indicate the respective average Mascot scores and the number of times the interactor was detected in three independent affinity purifications; only proteins are listed that were detected at least twice in three experiments with significant Mascot scores in the ELF1-GS samples, but not in controls.

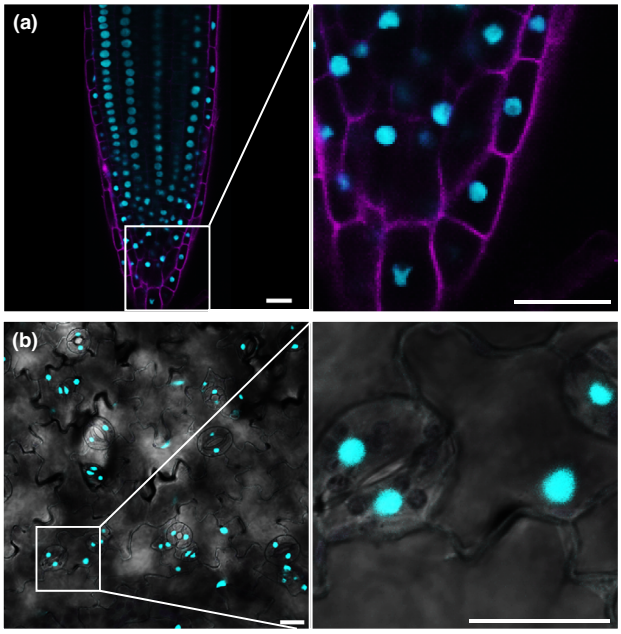



**Fig. 4** ELF1 localises to nuclei of root and leaf cells. Root tips and leaves of *elf1-1* plants expressing ELF1-eGFP fusion proteins under control of the *ELF1* promoter were analysed using CLSM. (a) In root tips of 7-d after stratification (DAS) plants eGFP fluorescence is visible in cyan, while propidium iodide (PI) staining is in magenta. (b) eGFP fluorescence in the abaxial surface of leaves from 10-DAS plants (epidermal cells, guard cells). Bar, 20 μm.

phenotypically analysed to detect possible genetic interactions between *ELF1* and the genes encoding other TEFs. Regarding the first type of double mutant, both the *elf1-1* and *tfIIs-1* parental plants exhibit essentially wild-type appearance, whereas the *elf1-1 tfIIs-1* double mutant is substantially smaller (Fig. 5a–d). The reduced growth of the double mutant is efficiently complemented in *elf1-1 tfIIs1*$^{C}$ plants by expression of ELF1-eGFP under control of the *ELF1* promoter (Fig. 5e), which is also evident from the measured fresh weight (Fig. 5v) and the rosette diameter (Fig. S3a,b) of the plant lines. Other determined phenotypic parameters differ only slightly between the different genotypes (Fig. S3a,b). Concerning the second type of double mutant, the *elf1-1 ssrp1-2* plants show striking leaf serration (leaf margin protrusions; Bilsborough *et al.*, 2011) that is greatly enhanced compared with the *ssrp1-2* parental plants and complemented in *elf1-1 ssrp1-2*$^{C}$ plants (Fig. 5f–m,w). Rosette diameter and bolting time are minimally changed in *elf1-1 ssrp1-2* plants compared with the *ssrp1-2* single mutant (Fig. S4a,b). Regarding the third type of double mutant, when compared to the parental lines the *elf1-1 iws1-1* double mutant is very strikingly late bolting (Fig. S5a,b) and it proved sterile. In addition, the leaf angle/inclination (angle a leaf deviates from horizontal; Hopkins *et al.*, 2008) of *elf1-1 iws1-1* plants is notably more erect than that of the parental lines and this feature is complemented in *elf1-1 iws1-1*$^{C}$ plants (Fig. 5n–u,x). Compared with the parental lines, other phenotypic parameters are only mildly affected in the double mutant (Fig. S5a,b). Together, the genetic interactions of *ELF1* with genes encoding the TEFs TFIIS, SSRP1 and IWS1 highlight a role of ELF1 in RNAPII transcript elongation.

## ELF1 associates with a subset of regions actively transcribed by RNAPII

Using the above-mentioned *elf1-1* plants expressing ELF1-eGFP, ChIP-Seq analyses were performed to examine the genome-wide distribution of ELF1. To that end, ChIP was carried out with antibodies directed against eGFP, while as ChIP-Seq controls served comparable ChIP analyses with plants expressing GFP-NLS (GFP fused to a nuclear localisation signal) and ChIP input samples. Analysis of the ChIP-Seq data revealed that biological replicates yielded robust results that group together and are

New Phytologist





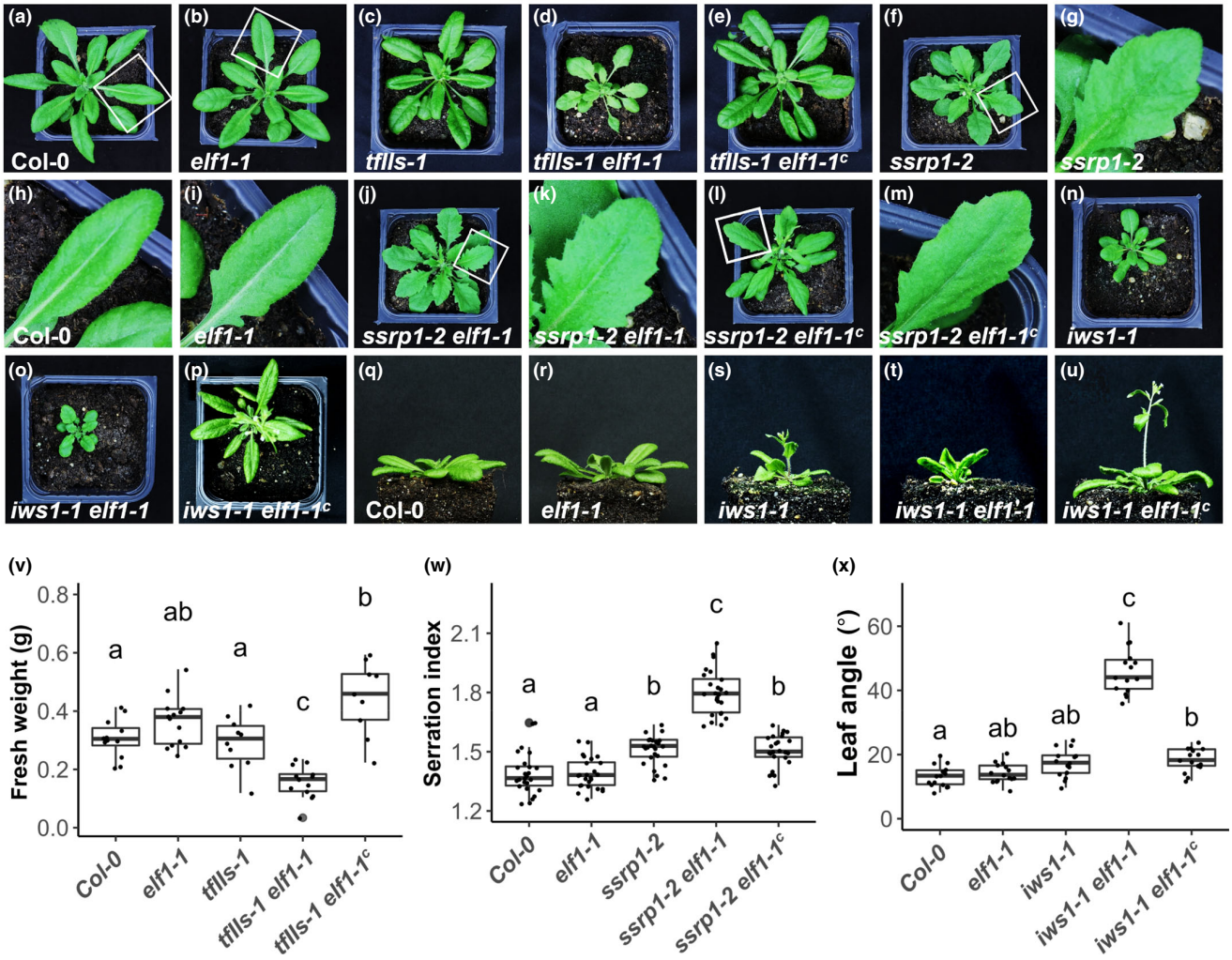



**Fig. 5** In combination with other transcript elongation factor mutations, *elf1-1* leads to specific phenotypes that can be complemented by *pELF1::ELF1-eGFP*. (a–u) documentation of representative individuals illustrating overall and specific phenotypes of the indicated genotypes (Col-0, single- and double mutants, as well as complemented double mutants indicated by (c)). Compared with the parental lines, *tflls-1 elf1-1* exhibits a striking growth defect (d) and (v), while *ssrp1-2 elf1-1* is characterised by a prominent serrated leaf phenotype (j) and (w). The leaf phenotype is illustrated by the overall view of the rosette in (a, b, f, j, l) and the boxed areas are magnified in (g–i, k, m). Relative to the parental lines, the *iws1-1 elf1-1* double mutant exhibits remarkably more erect leaves (lateral view) due to a greater leaf angle (t, x). (v–x) Phenotypical analyses are visualised by plotting the median in boxplots with hinges corresponding to the first and third quartile, whisker range is 1.5 inter-quartile range of the respective hinge, and outliers are represented by dots. The outcome of statistical analysis of the indicated phenotypes using one-way ANOVA followed by Tukey's multiple comparison test is indicated by letters. All pictures have been taken, and statistical analysis has been performed with 28-d after stratification (DAS) plants.

clearly distinct from the controls (Fig. S6). While the controls did not yield distinguishable enrichment in genomic coverage, ChIP with ELF1-eGFP exhibited markedly increased coverage over RNAPII transcribed regions, beginning at TSSs with a distinctive maximum around transcriptional end sites (TESs; Fig. 6a). The distribution pattern over the transcribed regions resembles the profile of elongating RNAPII (Hetzel *et al.*, 2016; Zhu *et al.*, 2018; Yu *et al.*, 2019; Antosz *et al.*, 2020), and therefore, the ChIP coverage of ELF1-eGFP was compared with that determined for RNAPII-S2P (Fig. S7a). Both ELF1-eGFP and RNAPII-S2P are enriched over the transcribed regions of highly transcribed genes (based on RNA-Seq data) with a distinctive peak around the TES, but only background levels are detected over nontranscribed genes (Fig. 6b). Dividing the genes into four groups depending on transcript levels demonstrated that RNAPII-S2P and ELF1-eGFP coverage increases with higher transcript levels (Fig. S7b), illustrating the correlation of ELF1 occupancy with ongoing RNAPII transcription that is typical of elongation factors. Interestingly, of the protein-coding genes with a peak for RNAPII-S2P enrichment ($n$ = 9969) only *c.* 33% ($n$ = 3253) of these genes show also a peak for ELF1-eGFP enrichment (Fig. 6c).

This is also apparent at the level of individual genes, when comparing the distribution of the ChIP signals of ELF1-eGFP and RNAPII-S2P along with the corresponding transcript levels and controls (Fig. 6d). Here, next to transcribed loci with clear

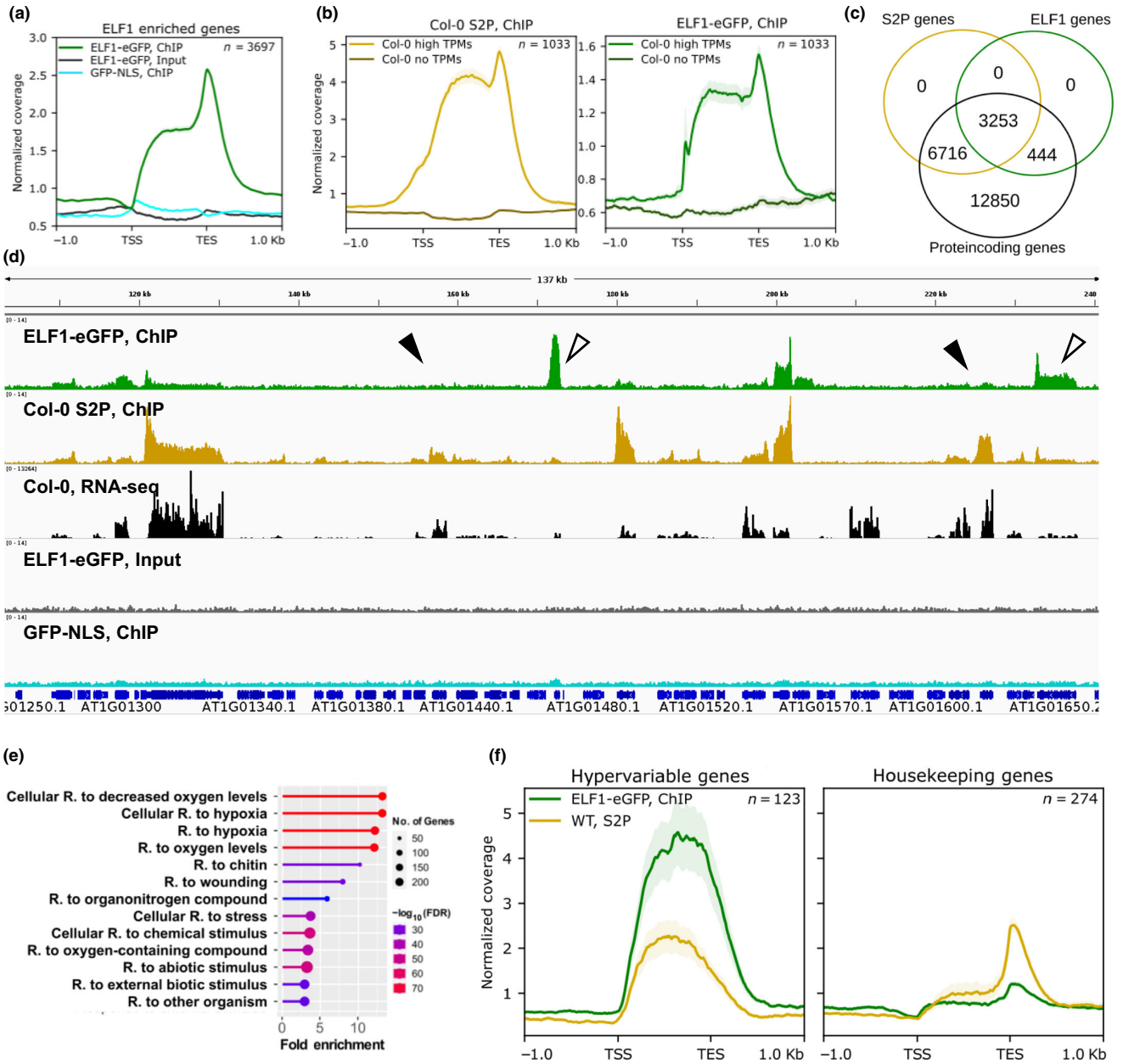



**Fig. 6** ELF1 is enriched over transcribed regions of a subset of RNAPII transcribed genes. (a) Metagene analysis of ChIP sequencing (ChIP-Seq) data obtained using α-GFP-specific antibodies with *elf1-1*$^{C}$ plants (expressing ELF1-eGFP) and control plants (expressing GFP-NLS) as well as *elf1-1*$^{C}$ input samples. (b) ELF1-GFP and RNAPII-S2P chromatin immunoprecipitation (ChIP) signals were plotted over highly transcribed genes (> 28 TPM, based on RNA-Seq analysis) and nontranscribed genes (< 0.01 TPM). Mean signals of the biological replicates were averaged (line) and the tracks represent the SEM for the replicates at each position (shaded area). (c) Number of genes with ELF1-GFP occupancy relative to the number of genes with RNAPII-S2P coverage and the total number of protein-coding genes. (d) ChIP-Seq profiles of individual genes aligned with RNA-Seq data and controls, illustrating that some transcribed loci exhibit RNAPII-S2P coverage, but no ELF1-GFP enrichment (indicated by closed arrowheads), whereas at other loci, there is a particularly strong ELF1-GFP signal relative to the RNAPII-S2P coverage (indicated by open arrowheads). Gene models are shown at the bottom. (e) Gene Ontology (GO)-term analysis of genes showing the highest ELF1-GFP coverage relative to RNAPII-S2P coverage ($n$ = 1000). Gene Ontology terms were sorted as indicated according to fold enrichment and number of genes per class with 'R.' representing 'response'. (f) ELF1-GFP and RNAPII-S2P ChIP signals were plotted over hypervariable and housekeeping genes (according to Zilberman *et al*., 2007; Aceituno *et al*., 2008).

RNAPII-S2P and ELF1-eGFP coverage, there are transcribed loci with RNAPII-S2P coverage, but without detectable ELF1-eGFP association (indicated by closed arrowheads). In addition, at few other loci there are – relative to the RNAPII-S2P coverage – remarkably prominent ELF1-GFP signals (indicated by open arrowheads). Loci with high ELF1-eGFP coverage are also apparent from unsupervised k-means clustering analysis (Fig. S8a). Gene Ontology analysis of genes showing the highest ELF1-

eGFP coverage demonstrated a clear enrichment for gene classes that are responsive to various stimuli (Fig. 6e). We further explored that by comparatively analysing the association of ELF1-eGFP with hypervariable relative to housekeeping genes (Zilberman *et al.*, 2007; Aceituno *et al.*, 2008). While 34 out of 123 hypervariable genes are highly enriched in ELF1-eGFP, only five out of 274 housekeeping genes exhibit prominent ELF1-eGFP coverage (Fig. S8b). This difference in ELF1-eGFP coverage is also evident from plotting the ChIP signals over these two gene classes (Fig. 6f). We extended this analysis by extracting genes of the GO terms 'responsive to stimulus' from TAIR10 (https://www.arabidopsis.org/). The ChIP signals of RNAPII-S2P and ELF1-eGFP colocalise over those genes as well as genes of the subgroups that are responsive to abiotic and biotic stimuli (Fig. S9). Interestingly, relative to the RNAPII-S2P signal, the ELF1-eGFP signal is particularly enriched on genes responsive to biotic stimuli. Sequence motif analysis using HOMER (Heinz *et al.*, 2010) of the 797 genes with high ELF1-eGFP coverage (cf. Fig. S8) revealed consistent results, as motifs including those of WRKY and AP2/ERF factors indicative of biotic and abiotic stress-responsive genes (Phukan *et al.*, 2016; Xie *et al.*, 2019) were clearly enriched (Fig. S10). Taken together, ELF1 associates over regions actively transcribed by RNAPII, which represents a characteristic feature of TEFs (Sims *et al.*, 2004; Kwak & Lis, 2013; van Lijsebettens & Grasser, 2014). Beyond that, ELF1 occupancy is detected over a subset of transcribed loci with a preference for inducible rather than for constitutively expressed genes.

## Discussion

Besides SPT4-SPT5 and TFIIS, ELF1 is considered a 'basal' TEF that is conserved not only in eukaryotes, but also in archaea (Ehara & Sekine, 2018). Sequences with similarity to yeast ELF1 (and mammalian ELOF1) are encoded in plant genomes and in this work, we provide evidence that *Arabidopsis* ELF1 plays a role in transcript elongation by RNAPII. Apart from its amino acid sequence conservation compared with the orthologues from yeast and metazoa, ELF1 localises to nuclei in *Arabidopsis* cells. It prominently copurifies with RNAPII and various TEFs including SPT4-SPT5, SPT6L, IWS1, PAF1C and FACT, in line with results obtained with ELF1 proteins of other organisms (Ehara *et al.*, 2017, 2019; van der Weegen *et al.*, 2021). Moreover, characteristic for elongation factors ChIP experiments demonstrated that ELF1 specifically associates with the transcribed region of genes actively transcribed by RNAPII, resembling the distribution of *Arabidopsis* RNAPII-S2P and that of ELF1 in yeast (Mayer *et al.*, 2010; Rossi *et al.*, 2021). Finally, *ELF1* exhibits genetic interactions with genes encoding other *Arabidopsis* elongation factors such as TFIIS, SSRP1 and IWS1. Yeast *ELF1* genetically interacted with *TFIIS* and genes encoding additional TEFs such as SPT4-SPT5, SPT6, SPT16 and PAF1C (Prather *et al.*, 2005), while human *ELOF1* showed interaction with *SPT4-SPT5, SPT6, CTR9* and *LEO1* (van der Weegen *et al.*, 2021).

In agreement with the yeast *elf1Δ* mutation that caused no significant growth defect (Prather *et al.*, 2005), *Arabidopsis elf1-1* exhibits basically wild-type appearance. A rice mutant defective in a possible *ELF1* orthologue termed *OsTEF1* showed reduced tillering capacity and retarded growth of seminal roots (Paul *et al.*, 2012), while inactivation of a putative wheat orthologue termed *TaTEF-7A* modulates grain number per spike (Zheng *et al.*, 2014). The observation that the *elf1-1* mutation does not cause significant growth defects, resembles the situation with *tfIIs* mutant plants that also grow similar to wild-type under standard conditions (Grasser *et al.*, 2009), but exhibit a striking sensitivity when exposed to elevated temperatures (Szádeczky-Kardoss *et al.*, 2022). Moreover, *Arabidopsis* plants defective in PAF1C subunits display reduced tolerance to increased salt concentrations and are affected in the response to mechanical stimulation (Jensen *et al.*, 2017; Obermeyer *et al.*, 2022; Zhang *et al.*, 2022). Likewise, expression of ELF1 may prove relevant under certain environmental conditions, since ELF1 associates preferentially with genes that are responsive to biotic and abiotic stress responses. Combination of the *elf1-1* mutation with genes encoding other *Arabidopsis* TEFs resulted in distinct phenotypic alterations. Thus, despite the wild-type appearance of the parental lines the *elf1-1 tfIIs-1* double mutant exhibits markedly reduced growth. However, the *elf1-1 ssrp1-2* plants are characterised by prominent leaf serration, while *elf1-1 iws1-1* plants are sterile and display strikingly altered leaf inclination resulting in more vertically oriented leaves. The phenotypes observed with the various double-mutant combinations suggest that the interaction of ELF1 with other TEFs may result in (partially) different transcriptional output in *Arabidopsis*.

Because of its constant association with the elongation complex during transcription, yeast ELF1 is designated as core elongation factor (Joo *et al.*, 2019), and likewise, archaeal ELF1 is a general part of the elongation complex (Blombach *et al.*, 2021). Based on our ChIP-Seq analyses, *Arabidopsis* ELF1 associates only with *c.* 33% of the RNAPII transcribed loci. Interestingly, ELF1 is detected preferentially along inducible genes rather than at constitutively expressed genes. Generally, the relative magnitude of the ELF1 and RNAPII-S2P ChIP-Seq signals appears to vary considerably, which may suggest that some genes require a greater amount of ELF1 for proper transcription, while other genes can be efficiently transcribed in the absence of ELF1. In addition to increased sensitivity of current mass spectrometry and the small size of the ELF1 protein, its association with only a subpopulation of RNAPII elongation complexes could account for the fact that ELF1 was not detected in previous mass spectrometric analyses of the *Arabidopsis* RNAPII elongation complex (Antosz *et al.*, 2017).

Recent *in vitro* transcription experiments on reconstituted nucleosomal templates and structural cryo-EM analyses indicated that *Komagataella pastoris* ELF1 promotes nucleosome transcription by RNAPII. Particularly, in combination with SPT4-SPT5, ELF1 exhibited a strong synergistic effect on RNAPII progression on the nucleosome (Ehara *et al.*, 2019). ELF1 and SPT4-SPT5 together reshape the downstream edge of the RNAPII elongation complex. In this scenario, the conserved basic N-terminal region of ELF1 could interact with the nucleosomal DNA, assisting dissociation of histone–DNA contacts, which is favourable for

RNAPII progression through nucleosomes (Ehara *et al.*, 2019). Our *in vitro* interaction studies demonstrated that *Arabidopsis* ELF1 binds to histones via its C-terminal acidic region, while the N-terminal basic region proved relevant for DNA interactions. Individual ELF1 appears to exist in an autoinhibited conformation, in which the acidic C-terminal region can be crosslinked intramolecularly with basic parts of the protein. This reminds of chromosomal HMGB proteins, whose acidic C-terminal region interacts with the basic DNA-binding parts of the protein to modulate DNA interactions (Thomsen *et al.*, 2004; Stott *et al.*, 2014). In case of *Arabidopsis* ELF1, during chromatin transcription by RNAPII, the acidic C-terminal region could interact with nucleosomal histones, at the same time releasing the basic N-terminal region from autoinhibition to interact with the DNA of the approached nucleosome, facilitating separation of histone–DNA contacts. According to a current model (Kujirai & Kurumizaka, 2020), ELF1 may contribute to the multiple acidic regions occurring in various TEFs including ELF1, SPT4-SPT5, PAF1C, FACT and SPT6 that in the RNAPII elongation complex may bind basic regions of the nucleosome exposed by the remodelling process during ongoing transcription.

Our study has revealed that ELF1 is a component of the RNAPII elongation complex in plants. Its role in transcriptional elongation is also evident from genetic interactions of *ELF1* with genes encoding other TEFs such as TFIIS, SSRP1 and IWS1, as well as its association with the transcribed region of active genes. ELF1 is enriched particularly along inducible RNAPII transcribed genes, basically sharing the genomic distribution with elongating RNAPII. The molecular analyses presented here provide a valuable resource and may inspire future studies addressing the contribution of transcriptional elongation to adjusting plant gene expression programmes under diverse conditions.

## Acknowledgements

We thank Serena Herzinger and Emilia Polz for contributions to the project, Eduard Hochmuth for recording mass spectra, the Genomics Core Facility at the University of Regensburg (http://www.kfb-regensburg.de/) for high-throughput sequencing and the Nottingham Arabidopsis Stock Centre (NASC) for providing *Arabidopsis* T-DNA insertion lines. This research was supported by the German Research Foundation (DFG) through grants Gr1159/14-2 and SFB960/A6 to KDG. Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

None declared.

## Author contributions

GL, FB and KDG designed the research and supervised the project. HM, PM-H, SO, CT, SB and AO performed experiments. HM, PM-H, SO, SB, AB and AO analysed data. KDG wrote the manuscript. All authors discussed the results, commented the manuscript and approved the final version.

## ORCID

Klaus D. Grasser https://orcid.org/0000-0002-7080-5520

## Data availability

Sequencing data have been deposited to the NCBI sequence read archive with the accession nos.: ChIP-Seq data of ELF1-eGFP and GFP-NLS (accession no. PRJNA823592), ChIP-Seq data of RNAPII-S2P (accession no. PRJNA826267) and RNA-Seq 6-DAS Col-0 (accession no. PRJNA758800).

## References

**Aceituno FF, Moseyko N, Rhee SY, Gutiérrez RA. 2008.** The rules of gene expression in plants: organ identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*. *BMC Genomics* **9**: 438.

**Antosz W, Deforges J, Begcy K, Bruckmann A, Poirier Y, Dresselhaus T, Grasser KD. 2020.** Critical role of transcript cleavage in Arabidopsis RNA polymerase II transcriptional elongation. *Plant Cell* **32**: 1449–1463.

**Antosz W, Pfab A, Ehrnsberger HF, Holzinger P, Köllen K, Mortensen SA, 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.

**Berr A, McCallum EJ, Ménard R, Meyer D, Fuchs J, Dong A, Shen W-H. 2010.** Arabidopsis SET DOMAIN GROUP2 is required for H3K4 trimethylation and is crucial for both sporophyte and gametophyte development. *Plant Cell* **22**: 3232–3248.

**Bilsborough GD, Runions A, Barkoulas M, Jenkins HW, Hasson A, Galinha C, Laufs P, Hay A, Prusinkiewicz P, Tsiantis M. 2011.** Model for the regulation of *Arabidopsis thaliana* leaf margin development. *Proceedings of the National Academy of Sciences, USA* **108**: 3424–3429.

**Blombach F, Fouqueau T, Matelska D, Smollett K, Werner F. 2021.** Promoter-proximal elongation regulates transcription in archaea. *Nature Communications* **12**: 5524.

**Chen C, Shu J, Li C, Thapa RK, Nguyen V, Yu K, Yuan Z-C, Kohalmi SE, Liu J, Marsolais F *et al.* 2019.** RNA polymerase II-independent recruitment of SPT6L at transcription start sites in Arabidopsis. *Nucleic Acids Research* **47**: 6714–6725.

**Chen FX, Smith ER, Shilatifard A. 2018.** Born to run: control of transcription elongation by RNA polymerase II. *Nature Reviews. Molecular Cell Biology* **19**: 464–478.

**Cheng J, Xu L, Bergér V, Bruckmann A, Yang C, Schubert V, Grasser KD, Schnittger A, Zheng B, Jiang H. 2022.** H3K9 demethylases IBM1 and JMJ27 are required for male meiosis in *Arabidopsis thaliana*. *New Phytologist* **235**: 2252–2269.

**Dürr J, Lolas IB, Sørensen BB, Schubert V, Houben A, Melzer M, Deutzmann R, Grasser M, Grasser KD. 2014.** The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in Arabidopsis. *Nucleic Acids Research* **42**: 4332–4347.

**Ehara H, Kujirai T, Fujino Y, Shirouzu M, Kurumizaka H, Sekine S-I. 2019.** Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. *Science* **363**: 744–747.

**Ehara H, Sekine S-I. 2018.** Architecture of the RNA polymerase II elongation complex: new insights into Spt4/5 and Elf1. *Transcription* **9**: 286–291.

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

**Ewels P, Magnusson M, Lundin S, Käller M. 2016.** MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**: 3047–3048.

**Fiorucci A-S, Bourbousse C, Concia L, Rougée M, Deton-Cabanillas A-F, Zabulon G, Layat E, Latrasse D, Kim SK, Chaumont N *et al.* 2019.**

Arabidopsis S2Lb links AtCOMPASS-like and SDG2 activity in H3K4me3 independently from histone H2B monoubiquitination. *Genome Biology* **20**: 100.

**Fleury D, Himanen K, Cnops G, Nelissen H, Boccardi TM, Maere S, Beemster GTS, Neyt P, Anami S, Robles P *et al.* 2007.** The *Arabidopsis thaliana* homolog of yeast BRE1 has a function in cell cycle regulation during early leaf and root growth. *Plant Cell* **19**: 417–432.

**Ge SX, Jung D, Yao R. 2020.** ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**: 2628–2629.

**Grasser KD, Grimm R, Ritt C. 1996.** Maize chromosomal HMGc: two closely related structure-specific DNA-binding proteins specify a second type of plant HMG-box protein. *The Journal of Biological Chemistry* **271**: 32900–32906.

**Grasser M, Kane CM, Merkle T, Melzer M, Emmersen J, Grasser KD. 2009.** Transcript elongation factor TFIIS is involved in *Arabidopsis* seed dormancy. *Journal of Molecular Biology* **386**: 598–611.

**Gu XL, Wang H, Huang H, Cui XF. 2012.** SPT6L encoding a putative WG/GW-repeat protein regulates apical-basal polarity of embryo in *Arabidopsis*. *Molecular Plant* **5**: 249–259.

**He Y, Doyle MR, Amasino RM. 2004.** PAF1-complex-mediated histone methylation of *FLOWERING LOCUS C* chromatin is required for the vernalization-responsive, winter-annual habit in *Arabidopsis*. *Genes & Development* **18**: 2774–2784.

**Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010.** Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular Cell* **38**: 576–589.

**Hetzel J, Duttke SH, Benner C, Chory J. 2016.** Nascent RNA sequencing reveals distinct features in plant transcription. *Proceedings of the National Academy of Sciences, USA* **113**: 12316–12321.

**Hopkins R, Schmitt J, Stinchcombe JR. 2008.** A latitudinal cline and response to vernalization in leaf angle and morphology in *Arabidopsis thaliana* (Brassicaceae). *New Phytologist* **179**: 155–164.

**Jensen GS, Fal K, Hamant O, Haswell ES. 2017.** The RNA polymerase-associated factor 1 complex is required for plant touch responses. *Journal of Experimental Botany* **68**: 499–511.

**Joo YJ, Ficarro SB, Chun Y, Marto JA, Buratowski S. 2019.** *In vitro* analysis of RNA polymerase II elongation complex dynamics. *Genes & Development* **33**: 578–589.

**Kujirai T, Kurumizaka H. 2020.** Transcription through the nucleosome. *Current Opinion in Structural Biology* **61**: 42–49.

**Kwak H, Lis JT. 2013.** Control of transcriptional elongation. *Annual Review of Genetics* **47**: 483–508.

**Langmead B, Salzberg SL. 2012.** Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

**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. *Nature protocols* **10**: 169–187.

**Li L, Ye H, Guo H, Yin Y. 2010.** *Arabidopsis* IWS1 interacts with transcription factor BES1 and is involved in plant steroid hormone brassinosteroid regulated gene expression. *Proceedings of the National Academy of Sciences, USA* **107**: 3918–3923.

**van Lijsebettens M, Grasser KD. 2014.** Transcript elongation factors: shaping transcriptomes after transcript initiation. *Trends in Plant Science* **19**: 717–726.

**Liu H, Ding Y, Zhou Y, Jin W, Xie K, Chen L-L. 2017.** CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. *Molecular Plant* **10**: 530–532.

**Lolas IB, Himanen K, Grønlund JT, Lynggaard C, Houben A, Melzer M, van Lijsebettens M, Grasser KD. 2010.** The transcript elongation factor FACT affects Arabidopsis vegetative and reproductive development and genetically interacts with HUB1/2. *The Plant Journal* **61**: 686–697.

**Mayer A, Lidschreiber M, Siebert M, Leike K, Söding J, Cramer P. 2010.** Uniform transitions of the general RNA polymerase II transcription complex. *Nature Structural & Molecular Biology* **17**: 1272–1278.

**Michl-Holzinger P, Obermeyer S, Markusch H, Pfab A, Ettner A, Bruckmann A, Babl S, Längst G, Schwartz U, Tvardovskiy A *et al.* 2022.** Phosphorylation of the FACT histone chaperone subunit SPT16 affects chromatin at RNA polymerase II transcriptional start sites in *Arabidopsis*. *Nucleic Acids Research* **50**: 5014–5028.

**Municio C, Antosz W, Grasser KD, Kornobis E, van Bel M, Eguinoa I, Coppens F, Bräutigam A, Lermontova I, Bruckmann A *et al.* 2021.** The Arabidopsis condensin CAP-D subunits arrange interphase chromatin. *New Phytologist* **230**: 972–987.

**Obermeyer S, Stöckl R, Schnekenburger T, Moehle C, Schwartz U, Grasser KD. 2022.** Distinct role of subunits of the Arabidopsis RNA polymerase II elongation factor PAF1C in transcriptional reprogramming. *Frontiers in Plant Science* **13**: 974625.

**Oh S, Zhang H, Ludwig P, van Nocker S. 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.

**Osakabe A, Lorkovic ZJ, Kobayashi W, Tachiwana H, Yelagandula R, Kurumizaka H, Berger F. 2018.** Histone H2A variants confer specific properties to nucleosomes and impact on chromatin accessibility. *Nucleic Acids Research* **46**: 7675–7685.

**Osman S, Cramer P. 2020.** Structural biology of RNA polymerase II transcription: 20 years on. *Annual Review of Cell and Developmental Biology* **36**: 1–34.

**Paul P, Awasthi A, Rai AK, Gupta SK, Prasad R, Sharma TR, Dhaliwal HS. 2012.** Reduced tillering in Basmati rice T-DNA insertional mutant OsTEF1 associates with differential expression of stress related genes and transcription factors. *Functional & Integrative Genomics* **12**: 291–304.

**Pfab A, Grønlund JT, Holzinger P, Längst G, Grasser KD. 2018.** The Arabidopsis histone chaperone FACT: role of the HMG-box domain of SSRP1. *Journal of Molecular Biology* **430**: 2747–2759.

**Phukan UJ, Jeena GS, Shukla RK. 2016.** WRKY transcription factors: molecular regulation and stress responses in plants. *Frontiers in Plant Science* **7**: 760.

**Prather D, Krogan NJ, Emili A, Greenblatt JF, Winston F. 2005.** Identification and characterization of Elf1, a conserved transcription elongation factor in *Saccharomyces cerevisiae*. *Journal of Molecular Cell Biology* **25**: 10122–10135.

**Quadrana L, Bortolini Silveira A, Mayhew GF, LeBlanc C, Martienssen RA, Jeddeloh JA, Colot V. 2016.** The *Arabidopsis thaliana* mobilome and its impact at the species level. *eLife* **5**: e15716.

**Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016.** deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research* **44**: W160–W165.

**Rossi MJ, Kuntala PK, Lai WKM, Yamada N, Badjatia N, Mittal C, Kuzu G, Bocklund K, Farrell NP, Blanda TR *et al.* 2021.** A high-resolution protein architecture of the budding yeast genome. *Nature* **592**: 309–314.

**Sims RJ, Belotserkovskaya R, Reinberg D. 2004.** Elongation by RNA polymerase II: the short and long of it. *Genes & Development* **18**: 2437–2468.

**Stott K, Watson M, Bostock MJ, Mortensen SA, Travers A, Grasser KD, Thomas JO. 2014.** Structural insights into the mechanism of negative regulation of single-box high mobility group proteins by the acidic tail domain. *The Journal of Biological Chemistry* **289**: 29817–29826.

**Szádeczky-Kardoss I, Szaker HM, Verma R, Darkó É, Pettkó-Szandtner A, Silhavy D, Csorba T. 2022.** Elongation factor TFIIS is essential for heat stress adaptation in plants. *Nucleic Acids Research* **50**: 1927–1950.

**Thomsen MS, Franssen L, Launholt D, Fojan P, Grasser KD. 2004.** Interactions of the basic N-terminal and the acidic C-terminal domains of the maize chromosomal HMGB1 protein. *Biochemistry* **43**: 8029–8037.

**Wang Z-P, Xing H-L, Dong L, Zhang H-Y, Han C-Y, Wang X-C, Chen Q-J. 2015.** Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. *Genome Biology* **16**: 144.

**van der Weegen Y, de Lint K, van den Heuvel D, Nakazawa Y, Mevissen TET, van Schie JJM, San Martin Alonso M, Boer DEC, González-Prieto R, Narayanan IV *et al.* 2021.** ELOF1 is a transcription-coupled DNA repair factor that directs RNA polymerase II ubiquitylation. *Nature Cell Biology* **23**: 595–607.

**Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. 2007.** An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* **2**: e718.

**Xie Z, Nolan TM, Jiang H, Yin Y. 2019.** AP2/ERF transcription factor regulatory networks in hormone and abiotic stress responses in Arabidopsis. *Frontiers in Plant Science* **10**: 228.

**Xu L, Zhao Z, Dong A, Soubigou-Taconnat L, Renou JP, Steinmetz A, Shen W-H. 2008.** Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in *Arabidopsis thaliana*. *Journal of Molecular Cell Biology* **28**: 1348–1360.

**Yu X, Martin PGP, Michaels SD. 2019.** BORDER proteins protect expression of neighboring genes by promoting 3′ Pol II pausing in plants. *Nature Communications* **10**: 4359.

**Zhang H, Li X, Song R, Zhan Z, Zhao F, Li Z, Jiang D. 2022.** Cap-binding complex assists RNA polymerase II transcription in plant salt stress response. *Plant, Cell & Environment* **45**: 2780–2793.

**Zheng J, Liu H, Wang Y, Wang L, Chang X, Jing R, Hao C, Zhang X. 2014.** TEF-7A, a transcript elongation factor gene, influences yield-related traits in bread wheat (*Triticum aestivum* L.). *Journal of Experimental Botany* **65**: 5351–5365.

**Zhu J, Liu M, Liu X, Dong Z. 2018.** RNA polymerase II activity revealed by GRO-seq and pNET-seq in *Arabidopsis*. *Nature Plants* **4**: 1112–1123.

**Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. 2007.** Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genetics* **39**: 61–69.

**Zlatanova J, van Holde K. 1998.** Binding to four-way junction DNA: a common property of architectural proteins? *The FASEB Journal* **12**: 421–431.

## Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

**Fig. S1** Amino acid sequence alignment of ELF1 of different species.

**Fig. S2** Generation of the *elf1-1* gene editing mutant and *elf1-1 pELF1::ELF1-eGFP* complementation line.

**Fig. S3** Phenotype of *tfIIs elf1* double-mutant plants.

**Fig. S4** Phenotype of *ssrp1 elf1* double-mutant plants.

**Fig. S5** Phenotype of *iws1 elf1* double-mutant plants.

**Fig. S6** Chromatin immunoprecipitation sequencing replicates show high pairwise correlation.

**Fig. S7** Comparative coverage profiles of ELF1-eGFP and RNA polymerase II-S2P.

**Fig. S8** Differential coverage of genes with ELF1.

**Fig. S9** ELF1-eGFP and RNAPII-S2P coverage over genes responsive to stimuli.

**Fig. S10** Motif analysis of ELF1-enriched genes.

**Table S1** Plasmids used in this study.

**Table S2** Oligonucleotides used in this study.

**Table S3** Summary of chromatin immunoprecipitation sequencing statistics for this project.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

# Chapter 7

# Different elongation factors distinctly modulate RNA polymerase II transcription in *Arabidopsis*

This article was submitted to a peer-reviewed journal in April 2023

# Different elongation factors distinctly modulate RNA polymerase II transcription in *Arabidopsis*

Simon Obermeyer[1], Lukas Schrettenbrunner[1], Richard Stöckl[1], Uwe Schwartz[2] and Klaus D. Grasser[1,*]

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

[2]NGS Analysis Centre, Biology and Pre-Clinical Medicine, University of Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany

[*]To whom correspondence should be addressed:
Tel: +49-941-9433032; Fax: +49-941-9433352; Email: Klaus.Grasser@ur.de
ORCID: 0000-0002-7080-5520

## ABSTRACT

Various transcript elongation factors (TEFs) including modulators of RNA polymerase II (RNAPII) activity and histone chaperones tune the efficiency of transcription in the chromatin context. TEFs are involved in establishing gene expression patterns during growth and development in *Arabidopsis*, while little is known about the genomic distribution of the TEFs and the way they facilitate transcription. We have mapped the genome-wide occupancy of the *Arabidopsis* elongation factors SPT4-SPT5, PAF1C and FACT, relative to that of elongating RNAPII phosphorylated at residues S2/S5 within the carboxyterminal domain. The distribution of SPT4-SPT5 along transcribed regions closely resembles that of RNAPII-S2P, while the occupancy of FACT and PAF1C is rather related to that of RNAPII-S5P. Under transcriptionally challenging conditions, mutant plants lacking the corresponding TEFs are differentially impaired in transcript synthesis. Strikingly, in plants deficient in PAF1C, defects in transcription across intron-exon borders are observed that are cumulative along transcribed regions. Instead, FACT is involved in adjusting nucleosomal occupancy upstream of transcriptional start sites. Moreover, under stress conditions FACT is particularly required for transcriptional upregulation and to promote RNAPII transcription through +1 nucleosomes. Thus, *Arabidopsis* TEFs are differently distributed along transcribed regions, and are distinctly required during transcript elongation especially upon transcriptional reprogramming.


## INTRODUCTION

Protein-coding genes are transcribed by RNA polymerase II (RNAPII), which plays a fundamental role in differential gene expression to ensure that appropriate amounts of mRNAs are produced in a spatially and temporally coordinated manner. Proper mRNA synthesis is an essential prerequisite for growth, development and response to environmental conditions. RNAPII transcription is characterised by a cycle of events including initiation, elongation and termination, and in addition to controlling the initiation step, the efficiency of mRNA production is regulated during elongation (1,2). Throughout the transcription cycle, the carboxyterminal domain (CTD) of the largest subunit of RNAPII is dynamically phosphorylated within its heptapeptide repeats including S2P and S5P, which are characteristic of the elongating polymerase (3,4). In the cell nucleus, the DNA template for transcription is wrapped around histone octamers, forming nucleosomes that represent obstacles for RNAPII progression. Consequently, the elongation phase of transcription on chromatin templates is a dynamic and discontinuous process (5,6). Studies primarily in yeast

and mammals demonstrated that transcriptional elongation (and its coordination with co-transcriptional events) is promoted by a variety of transcript elongation factors (TEFs) to assist mRNA synthesis by RNAPII. TEFs serve diverse functions including modulating the catalytic properties and processivity of RNAPII and relieving progression of the enzyme through repressive chromatin (2,7–9). During the past few years advances in cryo-electron microscopy in combination with x-ray crystallography allowed elucidation of many details regarding the three-dimensional organisation of elongating RNAPII associated with various TEFs (9–11).

The heterodimeric TEF, SPT4-SPT5, is conserved in eukaryotes and archaea and associates with RNAPII in a transcription-dependent manner (12). SPT4-SPT5 is involved in various aspects of transcript elongation including stabilisation of the RNAPII elongation complex, transcription through nucleosomes, regulation of RNAPII pausing, as well as transcriptional termination and the coordination with co-transcriptional mRNA processing (13,14). The multifunctional TEF, PAF1C, in higher eukaryotes consists of six subunits and associates with RNAPII to stabilise the elongation complex and to promote transcription (15,16). Furthermore, PAF1C links transcript elongation with post-translational histone modifications along actively transcribed regions, including H2B mono-ubiquitination and various histone methylations (i.e. H3K4me2/3, H3K36me3, K3K79me2/3). Therefore, as a global regulatory factor PAF1C functions at the interface between chromatin and transcription (15,16). Another factor that is considered to be a TEF is the FACT histone chaperone consisting of the SSRP1 and SPT16 subunits. FACT can facilitate disassembly and reassembly of nucleosomes, thus it is effective in various DNA-dependent processes such as replication, repair and transcription (17,18). The reversible nucleosome reorganisation is accomplished by multiple contacts of the FACT subunits with histones and the nucleosomal DNA. During elongation FACT promotes transcription of RNAPII through nucleosomes, but importantly it also assists reassembly of nucleosomes after RNAPII passage (17–19).

A number of studies, particularly in the *Arabidopsis thaliana* model, have revealed that various TEFs play vital roles in governing plant growth and development (20,21). In *Arabidopsis*, each subunit of SPT4-SPT5 is encoded by two genes and the ubiquitously expressed *SPT5-2* proved essential. SPT4-RNAi plants expressing reduced amounts of *SPT4* (and SPT5-2) exhibit growth and developmental defects of which some may be caused by inadequate auxin signalling (22). Moreover, SPT4-SPT5 modulates plant thermomorphogenesis and salt tolerance (23,24). Along with other TEFs, SPT4-SPT5 associate with elongating RNAPII, forming the active elongation complex (22,25). *Arabidopsis* PAF1C consisting (as in mammals) of six subunits associates with RNAPII (25) and was initially recognised because of its role in the transition from vegetative to

reproductive development. The early flowering phenotype of plants deficient in PAF1C subunits is associated with reduced expression of the floral repressor *FLC* (26–30). In addition, PAF1C proved to be involved in plant response to mechanical stimulation inflicted by repeated touch (31) as well as in the response to elevated salt concentration (32,33). The FACT histone chaperone composed of SSRP1 and SPT16 associates with transcribed regions of *Arabidopsis* genes in a transcription-dependent manner (25,34,35). FACT proved essential for viability (36,37) and decreased *SSRP1/SPT16* expression levels resulted in a variety of vegetative and reproductive defects including increased number of leaves, early bolting, reduced seed set and impaired circadian rhythm (37,38). Moreover, FACT is critically involved in the repression of aberrant transcript initiation from within coding regions of RNAPII-transcribed genes (39).

While substantial information has been gathered about the role of TEFs in plant development, knowledge regarding genome-wide impact of TEF function on RNAPII transcription as well as their contribution to (rapid) transcriptional reprogramming is rather incomplete. Therefore, we have comparatively analysed three types of *Arabidopsis* TEF (SPT4-SPT5, PAF1C and FACT) regarding their genomic distribution relative to elongating RNAPII. In addition, their involvement in the transcriptional response to challenging conditions was examined by analysing corresponding mutant plants relative to wildtype. The comparative analysis uncovered that the three types of TEFs distinctly influence transcript elongation in the context of plant chromatin.

## MATERIALS AND METHODS

### Plant material

Seeds of *Arabidopsis thaliana* Col-0 were sown on solid 0.5x MS medium (40) and after stratification for 48h at 4 °C in the dark, the plates were transferred to a plant incubator (PolyKlima) with long day settings with 16h light (110 $\mu mol \cdot m^{-2} \cdot s^{-1}$) at 21°C and 8h darkness at 18°C. For heat stress treatment the plates were transferred to a water bath set to 44°C for 10 min and snap frozen in liquid nitrogen, as previously described (41). Plant lines used in this study were described before: *ssrp1-2, spt16-1* (37), *elf7-3* (26) and SPT4-RNAi line 3 (termed *SPT4-R3*) (22).

### Chromatin Immunoprecipitation

Chromatin immunoprecipitation (ChIP) was essentially performed as described previously (42,43). *Arabidopsis* plants grown *in vitro* (14 d after stratification, 14-DAS) were crosslinked with formaldehyde and used for isolation of nuclei, before chromatin was sheared using a Bioruptor Pico device (Diagenode, Seraing, Belgium). Immunoprecipitation was carried out

using magnetic Dynabeads Protein A (Thermo Fisher Scientific). Elongation factor ChIPs were performed with previously described antibodies directed against purified, recombinant domains of elongation factors SPT5-2 (aaT792-P1041) (22), ELF7 (aaD401-E589) (25), SSRP1 (aaV453-N646) and SPT16 (aaE416-M554) (34). RNAPII ChIPs were performed with antibodies against RNAPII-S2P (ab5095; Abcam) and RNAPII-S5P (ab5131; Abcam).

**Chromatin immunoprecipitation sequencing (ChIP-seq)**

Libraries were prepared using the NEBnext UltraII Library preparation kit with 6bp NEBNext Multiplex Oligonucleotides for Illumina according to the manufacturer with minor modifications. ChIP-seq was carried out as described in the Illumina NextSeq 500 System Guide (Illumina), and the KAPA Library Quantification Kit - Illumina/ABI Prism (Roche Sequencing Solutions). Equimolar amounts of each library were sequenced on a NextSeq 500 instrument controlled by the NextSeq Control Software (NCS) v2.2.0, using the 75 Cycles High Output Kit with single index, single-read (SR) run parameters. Image analysis and base calling were done by the Real Time Analysis Software (RTA) v2.4.11. The resulting .bcl files were converted into .fastq files with the bcl2fastq v2.18 software. ChIPseq was performed at the Genomics Core Facility “KFB - Center of Excellence for Fluorescent Bioanalytics” (University of Regensburg, Regensburg, Germany; www.kfb-regensburg.de). Sequencing resulted in an average of 17.7M reads per library (min 12.8M, max 23.9M) with an average Phred score >34 (Table S1).

**ChIP-seq data analysis**

Reads were trimmed using Trimmomatic (v0.39) (44) followed by alignment to the TAIR10 genome (https://www.arabidopsis.org/) using Bowtie2 (45) using the preset mode *--local --very-sensitive-local*. Alignment files were converted to .bam files containing only alignments with MAPQ score >10, sorted and indexed using the samtools suite (v1.16.1) (46). Coverage tracks were calculated using deepTools ‘bamCoverage’. Downstream analysis was mainly performed using the deepTools2 suite (v.3.5.0) (47). Quality control was performed at several steps using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Genomic regions with aberrant coverage or low sequence complexity were filtered out, as described previously (48). After confirming high pairwise correlations and performing hierarchical clustering analysis (Supplementary Figure S1, S2), the biological replicates (four replicates for SPT5, ELF7, SSRP1 and SPT16 and two replicates for RNAPII-S2P and -S5P) were merged and CPM normalised. Peaks were called using the MACS3 (49) *callpeak* function with RNAPII-S2P and RNAPII-S5P as control, respectively, with the following arguments *-g 119481543 -B -q 0.01 --nomodel*.

**MNase-seq data analysis**

Analyses were performed as previously described (41) and +1 nucleosomal positions were extracted as the first nucleosomal position after TAIR10 annotated transcription start sites with a minimal overlap of 140 bp. The following nucleosomal positions were extracted following the definition of the +1 nucleosome for each gene.

**Transcript profiling by RNA-seq**

Nuclear RNA isolated from 6-DAS plants using the TRIzol method (Invitrogen) was further purified and DNase-treated using the Monarch RNA Cleanup Kit (New England Biolabs) as previously described (41). Library preparation and RNAseq were carried out as described in the NuGEN Universal RNA-Seq with NuQuant User Guide v2.1 (Tecan Genomics), the Illumina NextSeq 500 System Guide (Illumina), and the KAPA Library Quantification Kit - Illumina/ABI Prism User Guide (Roche Sequencing Solutions).

In brief, 25 ng of total RNA isolated from *Arabidopsis* nuclei were reverse transcribed into first strand cDNA using a mixture of random and poly-dT primers with an integrated DNase treatment step. Second strand synthesis, using a nucleotide analogue enabling strand retention, generated double stranded cDNA (ds cDNA). The ds cDNA was fragmented to a median size of 200-500 base pairs with a S2 Ultrasonication System (Covaris) using the following settings: Intensity 5; Duty Cycle 10%; Cycles per Burst 200; Treatment Time 180 s; micro Tube AFA Fiber Snap-Cap. Next, end repair was performed to generate blunt-ended ds cDNA, followed by the ligation of the indexing adapters, a strand selection via nucleotide analogue-targeted degradation and a reduction of the *Arabidopsis* rRNA content by an AnyDeplete-mediated adaptor cleavage. Finally, cDNA libraries were created by 18 cycles PCR enrichment and purified by a magnetic bead clean-up using 0.8 volumes of the bead suspension. The libraries were quantified using the KAPA Library Quantification Kit. Equimolar amounts of each library were sequenced on a NextSeq 500 instrument controlled by the NextSeq Control Software (NCS) v2.2.0 using one 75 Cycles High Output Kit with the single index, single-read (SR) run parameters. Image analysis and base calling resulted in .bcl files, which were converted into .fastq files with the bcl2fastq v2.18 software.

Library preparation and RNAseq were performed at the Genomics Core Facility “KFB - Center of Excellence for Fluorescent Bioanalytics” (University of Regensburg, Regensburg, Germany; www.kfb-regensburg.de). Sequencing of three biological replicates per genotype and condition resulted in an average of 8.4M reads per library (min 6.9M, max 14.4M) with an average Phred score >34 (Table S2).

**RNA-seq data analysis**

Quality control of raw RNAseq reads was performed with FastQC v0.11.8

(https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and MultiQC (50). After the initial quality assessment, reads with low base quality and adapter contaminations were removed using Trimmomatic (v0.39) (44). The remaining reads were aligned to the TAIR10 genome assembly with STAR (v2.7.3a) (51). The resulting alignment maps were converted to .bam files with Samtools (v1.9) (46). Quality control of the alignment was performed with FastQC, MultiQC and Qualimap v2.2.1 (52). Read normalisation was performed with Samtools and Deeptools v3.3.0 (47). Count table was made with subread package v1.6.3 (53). Consistency of the biological replicates was evaluated by calculation of pairwise correlations and principle component analyses (Supplementary Figure S3, S4). Differential expression analysis was performed with DeSeq2 v1.24.0 (54).
Unnormalised .bam alignment files were merged and then indexed with Samtools v1.9. log2ratios of the HS libraries were calculated using the merged no HS control for each genotype as control with parallel sequencing depth normalisation with bamCompare (Deeptools v3.3.0). Internal exons and intron-exon borders were extracted using the R package TxDb.Athaliana.BioMart.plantsmart28 (version 3.2.2) with R (version 3.6.1).

## RESULTS

### Differential distribution of TEFs over transcribed regions

Chromatin immunoprecipitation combined with high throughput sequencing (ChIP-seq) was used to map the distribution of TEFs and RNAPII genome-wide. To detect native SPT4-SPT5, PAF1C and FACT, antibodies directed against *Arabidopsis* SPT5-2, ELF7 (orthologue of PAF1) and SSRP1/SPT16 (22,25,34) were used. To detect elongating RNAPII, commercial antibodies were used that are directed against the phosphorylated residues S2P and S5P of the heptapeptide repeats occurring in the CTD of the NRPB1 subunit (55). Analysis of the ChIP-seq data obtained with 14 days after stratification (DAS) Col-0 plants demonstrated that expectedly elongating RNAPII and TEFs are detected over RNAPII-transcribed regions (Fig. 1A-C). RNAPII-S2P and S5P show a divergent pattern with the S2P signal gradually increasing from the transcriptional start site (TSS) towards the transcriptional end site (TES) and exhibiting a prominent peak just downstream of the TES, while S5P is more evenly distributed throughout the gene body. The distribution of RNAPII-S2P and S5P over transcribed regions (Supplementary Figure S5) is essentially in agreement with earlier studies in *Arabidopsis* (42,56). Moreover, as recently discussed (21), the result confirms the different genomic distribution of S5P in *Arabidopsis* in comparison to that observed in yeast and mammals, which exhibits maximum enrichment around the TSS with markedly declining coverage over the transcribed region towards the TES. The profile of SPT5 closely resembles that of S2P (Fig. 1A), also evident from the high Pearson correlation of 0.95

(Supplementary Figure S6). ELF7 is spread over the transcribed region and no clear enrichment is seen at the TES (Fig. 1B). The FACT subunits, SSRP1 and SPT16, share a highly similar distribution (Fig. 1C) with a Pearson correlation of 0.99 (Supplementary Figure S6), which also illustrates the robustness of the analysis. The FACT occupancy over the transcribed region is shifted towards the TSS (relative to that of RNAPII) and more closely resembling the distribution of S5P rather than S2P. In line with their concerted involvement in transcriptional elongation the analysed TEFs and RNAPII-S2P/S5P share relatively high pairwise correlations (Supplementary Figure S6). Further analysis of the distribution patterns by hierarchical clustering confirmed the more pronounced conformity of SPT5 with S2P, and the higher conformity of FACT and ELF7 with S5P (Fig. 1D). Next the analysed genes were divided into quartiles depending on the corresponding transcript levels. SPT5 shows a very clear linear correlation with S2P, while ELF7 shows a marked correlation with S5P and it is somewhat overrepresented in the gene body of very highly transcribed genes (Fig. 2). Both SPT16 and SSRP1 exhibit striking similarity with each other and are (compared to S2P and S5P) overrepresented in the gene body of genes dependent of transcriptional activity (Fig. 2). To identify regions of TEF enrichment relative to RNAPII, the ChIP-seq signals were analysed by differential peak calling of SPT5/ELF7/FACT relative to RNAPII-S2P/S5P. This analysis revealed that in line with the above results SPT5 largely shared the distribution of RNAPII-S2P, but (other than ELF7 and FACT) was enriched upstream of the TSS relative to RNAPII-S5P (Fig. 3A). Consistently, relative to S5P, SPT5 is enriched mostly outside transcribed regions (Fig. 3D,E). The enrichment of ELF7 is essentially restricted to the gene body (Fig. 3B), where it co-localises with S5P rather than S2P (Fig. 3D,E). For FACT a characteristic accumulation is detected downstream of the TSS (Fig. 3C). This prominent promoter-proximal accumulation is seen both relative to S2P and S5P (Fig. 3D,E). Taken together, these analyses illustrate that over transcribed *Arabidopsis* genes SPT4-SPT5, PAF1C and FACT are differently distributed relative to each other and relative to RNAPII.

**Transcriptional reprogramming is distinctly affected in plants lacking different TEFs**

To further elaborate the function of the TEFs in transcriptional elongation, *in vitro* grown plants deficient in TEFs were exposed to heat stress (HS) to prompt transcriptional reprogramming. SPT4-RNAi plants (termed *SPT4-R3*, also depleted in SPT5, (22)), *elf7-3* (no full-length *ELF7* mRNA is detectable (26)), *ssrp1-2* (substantially decreased expression of SSRP1 (37)) and for comparison Col-0 wildtype plants were exposed to a rapid HS of 10min at 44°C (using a water bath to achieve fast temperature change). Sublethal HS treatment of that kind evoked altered expression of hundreds of genes (41,57) and we presumed that this scenario may assist identifying whether/which transcriptional impairments occur in the different mutants. To examine genome-wide transcriptional changes of the

plants deficient in TEFs relative to Col-0 upon exposure to heat, nuclear RNA was isolated from untreated plants (no HS) and plants exposed to HS. We intended to generate information on transcriptional output (freshly synthesised unspliced/spliced mRNAs including nascent transcripts) rather than steady-state mRNA levels obtained with total poly(A) mRNA enrichment (58,59). Therefore, we (1) isolated nuclear RNA (rather than total RNA), (2) used rRNA depletion (rather than poly(A) enrichment) and (3) made use of low RNA size cut-off (≥25 nt). Isolated RNAs were assayed by high throughput sequencing and the reads were mapped to the *Arabidopsis* genome. Under standard conditions relatively mild changes in the transcript patterns are observed for the mutants deficient in the three TEFs relative to Col-0 plants and the defects are increased with higher transcript levels (Fig. 4A-C, top panels). Upon HS, however, pronounced and distinct defects are detected for the three genotypes. A common feature evident from the metagene profiles is a transcript level-dependent shift of the mapped RNA-seq reads towards the TSS (Fig. 4A-C, bottom panels). In parallel, the transcript reads decrease along the transcribed region towards the TES, indicating decreased RNAPII processivity. This effect is most pronounced with *elf7*, while with *ssrp1* a prominent peak just downstream of the TSS becomes apparent upon HS. In case of *SPT4-R3* the drop in transcript reads over the transcribed region is less pronounced than with *elf7*. The relative shift of the RNA-seq reads towards the TSS and the decrease over the transcribed regions upon HS among the expressed gene groups is obvious over a large set of genes (Fig. 4D-F). Thus, the transcript profiling analysis highlights that the depletion of TEFs distinctly impairs transcript synthesis at different genomic regions, particularly under conditions of massive transcriptional reprogramming.

**In addition to transcribed regions SPT5 localises to a region upstream of TSSs**

SPT5 mainly localises to the transcribed region, closely resembling the distribution of RNAPII-S2P (Fig. 1). Differential ChIP-seq peak calling revealed that relative to S5P, SPT5 is also enriched upstream of the TSS (Fig. 3). When summarising the differentially called ChIP-seq peaks over closely spaced genes (<500 bp distance) and more distant genes (>500 bp distance), it became clear that a notable SPT5 peak upstream of TSSs is detected only at closely spaced genes (Fig. 5A). Unlike SPT5, independent from the gene distance, FACT and ELF7 exhibit no considerable enrichment in this region flanking the TSS (Supplementary Figure S7A-C). The SPT5 peak adjacent to transcribed regions reminded of the genomic distribution of BORDER proteins (BDR1-3) that were reported to prevent transcriptional interference of closely spaced downstream genes in *Arabidopsis*. BDR1/2 occupied regions just upstream of the TSS and/or downstream of the TES, while BDR3 showed a strong preference for the TES (60). Reanalysis of the BDR1 ChIP-seq data demonstrated a more pronounced enrichment at closely spaced genes (<500 bp distance)

when compared to more distant genes (>500 bp distance) (Supplementary Figure S7D). Because of the prevalent presence of BDR1 at TSSs, peaks of SPT5 enrichment relative to RNAPII-S5P were centred over BDR1 ChIP-seq summits (60). This analysis demonstrated that SPT5 co-localises with BDR1 upstream of TSSs, whereas SSRP1 and ELF7 do not share this genomic distribution (Fig. 5B). The heatmap shows a relative accumulation of SPT5 over the centre of BDR1 peaks while ELF7 and FACT are not enriched (Fig. 5C). Therefore, unlike PAF1C and FACT, SPT5 associates with BDR1 target sites upstream of TSSs of some closely spaced genes.

**ELF7 is required for RNAPII to transcribe efficiently over splice sites**

ELF7 was detected over transcribed regions and its distribution showed a lower correlation with RNAPII than that of SPT5 (Fig. 1). Furthermore, ELF7 exhibits differences in the distribution compared to S2P but almost no difference to S5P (Fig. 3B), and compared to Col-0 in *elf7* mutant plants, transcript reads decreased along transcribed regions, particularly upon HS-induced transcriptional reprogramming (Fig. 4). In search of gene characteristics that may be related to ELF7-dependent transcriptional defects, genes up- and downregulated in *elf7* relative to Col-0 upon HS were analysed for the length of their transcribed regions. For downregulated genes a median length of 2532 bp was determined, whereas upregulated genes were shorter (median: 1787 bp) (Supplementary Figure S8A), suggesting that PAF1C is particularly required for transcription of longer genes. Moreover, downregulated genes contained ~8.8 exons and upregulated genes contained fewer exons (~6.2 exons), while exon length differed only slightly (160 vs. 185 bp) between the two groups of genes (Supplementary Figure S8B,C). Genes downregulated in *elf7* show an increased transcript coverage throughout the gene body in Col-0, while this upregulation cannot be seen in *elf7*. However, this difference between Col-0 and *elf7* is only apparent downstream of the promoter-proximal region. Therefore, genes downregulated in *elf7* are downregulated due to defective elongation beyond the promoter-proximal region (Fig. 6A). For genes upregulated in *elf7* only minor differences are seen over the transcribed region (Fig. 6B). This indicates a distinctive reduction of transcription along transcribed regions in ELF7-dependent genes. As genes downregulated in *elf7* contained a higher number of introns (Supplementary Figure S8C), RNA-seq data of intron-less genes were compared with those of intron-containing genes. Intron-less genes showed a similar pattern in *elf7* and Col-0, whereas for intron-containing genes transcript reads were reduced in *elf7* along the transcribed region (Fig. 6C,D). Plotting the relative transcript coverage over intron-exon borders along protein coding genes showed only minor differences under standard conditions (without HS) (Fig 6E). In *elf7* and *SPT4-R3* a slight reduction of mapped reads can be seen across intron-exon borders. However, upon HS a pronounced drop of transcripts is detected in *elf7* over intron-exon

borders that is cumulative over successive intron-exon borders along genes (Fig. 6F). These effects are clearly weaker for *SPT4-R3* and not discernible for *ssrp1* plants. Plotting the distribution of RNA-seq reads over ELF7 ChIP-seq peaks shows that this transcriptional defect is also present over ELF7 enriched regions (Supplementary Figure S8D, top). At the same time these ELF7 enriched sites exhibit accumulation of S5P (Supplementary Figure S8D, bottom). This suggests that ELF7 enriched regions in Col-0 coincide with sites of transcription defects in *elf7* plants upon HS. Genes downregulated in *elf7* were additionally searched for the occurrence of prominent sequence motifs. The two most strikingly identified motifs represent recognition sites for the spliceosomal factors SRSF2 and SNRNP (Supplementary Figure S8E), while no significantly enriched motifs were detected with genes upregulated in *elf7* or with the other analysed mutants (*ssrp1, SPT4-R3*). Hence, our analyses highlight that the transcriptional elongation defects observed with *elf7* plants correlate with the exon-intron structure of affected genes.

**FACT localises to nucleosomes, modulates the access to NDRs and is required to promote transcription over +1 nucleosomes upon HS**

The distribution of FACT over transcribed regions also showed a lower correlation with RNAPII than that of SPT5 (Fig.1) with a prominent promoter-proximal enrichment relative to RNAPII downstream of TSSs (Fig. 3). Upon HS, transcript coverage accumulates in *ssrp1* plants just downstream of the annotated TSS (Fig. 4). In view of the function of FACT as a histone chaperone, its genomic distribution was superimposed with nucleosome positions derived from MNase-seq data (41). This dataset was obtained by H3 immunoprecipitation after MNase digestion, ensuring that only DNA protected by histones is sequenced (61). Both the SSRP1 and SPT16 ChIP-seq coverage is enriched over nucleosomal positions, while the nucleosomal pattern is markedly less pronounced with ELF7 and SPT5 (Fig. 7A) and RNAPII (Supplementary Figure S9A). Comparison of the FACT enrichment relative to RNAPII-S2P revealed that in general, the SSRP1 coverage correlates positively with the transcript level (Fig. 7B). The accumulation of FACT at highly transcribed genes is in line with the analyses above (Fig. 2). FACT (both SSRP1 and SPT16) and RNAPII are particularly enriched at genes that are downregulated in *ssrp1* plants relative to Col-0 upon HS, while the coverage is clearly lower at upregulated genes (Supplementary Figure S9B). With genes that are differentially expressed dependent on ELF7, the differences in ELF7 and RNAPII occupancy between up- and downregulated genes are comparatively clearly lower (Supplementary Figure S9C). This suggests that compared to PAF1C, FACT is particularly required for transcriptional upregulation in response to HS. Genes downregulated in *ssrp1* (relative to Col-0) in fact fail to be upregulated upon HS, while the genes differentially expressed in *elf7* show comparable S2P/S5P coverage in Col-0 (cf. Supplementary Figure S9B,C). Consistent

with earlier studies (62,63), along transcribed regions of genes with high or medium expression, a notable, phased nucleosome pattern is evident that is not well discernible with low- or non-expressed genes (Fig. 7B). Strikingly, high FACT occupancy is associated with a clearly more pronounced nucleosome-depleted region (NDR) upstream of the TSS. Using unsupervised k-means clustering, two gene clusters were identified according to high and low FACT occupancy relative to RNAPII-S2P based on the ChIP-seq data. Genes of both clusters were intersected with nucleosomal positions, and centring SSRP1/SPT16 ChIP-seq peaks to the +1 nucleosomal position of those genes demonstrated a distinctive accumulation of SSRP1 and SPT16 just downstream of the +1 nucleosome of FACT enriched genes (Fig. 7C). Centring MNase histone H3 ChIP-seq data (41) over the +1 nucleosomal position of the FACT enriched gene cluster revealed a striking depletion of nucleosomes upstream of the TSS, while there is no marked change in the nucleosomal pattern downstream of the TSS (Fig. 7D). The depletion of the NDR is considerably less pronounced with the FACT non-enriched gene cluster. The mean transcript levels of FACT enriched genes are higher than those of ELF7- and SPT5-enriched genes, suggesting that FACT is especially associated with high expression levels (Supplementary Figure S9D). In addition, the transcript levels of FACT non-enriched genes are lower, but still those genes are transcribed, although there is a relatively high nucleosome occupancy at the NDR. Centring the relative fold-change in transcript coverage (cf. Fig. 4) over the two gene clusters highlighted a clear accumulation of transcripts just upstream of the +1 nucleosome in the FACT enriched gene cluster after HS (Fig. 7E). The accumulation of transcript reads is distinctly weaker for the FACT non-enriched gene cluster or without HS treatment. The maximum accumulation of transcripts maps ~40 bp upstream of the centre of the +1 nucleosome, while the peak centre of FACT enrichment maps ~110 bp downstream of the +1 nucleosome (Fig. 7F-H). Centring the FACT ChIP-seq coverage and the RNA-seq reads to the +1 to +5 nucleosomal positions illustrates that the distinctive FACT occupancy as well as the accumulating RNA-seq reads upon HS occur exclusively at the +1 nucleosome (Supplementary Figure S10A,B). Furthermore, we plotted FACT ChIP-seq peaks centred to the +1 nucleosome over FACT enriched/non-enriched, ELF7 enriched/non-enriched and SPT5 enriched/non-enriched gene clusters. The specific enrichment of FACT downstream of the +1 nucleosome is only detected in the FACT enriched cluster, but not in the other gene clusters (Supplementary Figure S11A-C). The corresponding MNase H3 ChIP-seq patterns reveal that the distinctive cleared NDR upstream of the TSS is obvious in the FACT enriched cluster, but clearly less marked in the other gene clusters (Supplementary Figure S11D-G). Finally, by plotting the MNase H3 ChIP-seq signal centred to the +1 nucleosome of genes up- or downregulated in *ssrp1*, the cleared NDR is quite evident for the downregulated genes, but markedly less discernible for the upregulated genes (Supplementary Figure

S11H). Thus, the genes that exhibit particularly increased FACT coverage (Supplementary Figure S9B) are characterised by a prominent NDR upstream of the TSS (Supplementary Figure S11H) and require FACT for transcriptional upregulation. Apart from its influence on the promoter-proximal NDR our results imply that FACT is required under HS conditions for RNAPII to efficiently transcribe through +1 nucleosomes of a number of genes. Recently, the elongation factor TFIIS was found to be critically involved in plant HS response, in part also by promoting transcription through the +1 nucleosome (41,57). Therefore, we examined whether the expression of a common set of genes is affected in *tfIIs* and *ssrp1* upon HS. However, comparison of the genes downregulated upon HS revealed only a minor overlap between the genes affected in *tfIIs* and *ssrp1* (Supplementary Figure S12A). TFIIS-dependent stalling of RNAPII at the +1 nucleosome under HS conditions is associated with eviction of the histone variant H2A.Z from the +1 nucleosome (41). Hence, we examined H2A.Z occupancy in absence/presence of HS, derived from H2A.Z ChIP-seq data (64) over genes that were up- or downregulated in *ssrp1* relative to Col-0. No change in H2A.Z coverage at the +1 nucleosome was apparent in the analysis (Supplementary Figure S12B). Consequently, both FACT and TFIIS influence RNAPII transcription at the +1 nucleosome upon HS, but in different ways.

**DISCUSSION**

Eukaryotic organisms vary substantially regarding genomic features that most likely influence the process of transcript elongation by RNAPII. Besides distinctions in chromatin characteristics this includes genome size and in part related to that enormous differences in the common length of transcribed regions, on the one hand in mammals (e.g. ~66.6 kb in human, (65)) and on the other hand in *Arabidopsis* or yeast (e.g. ~2.2 kb in *Arabidopsis*, (66)). Still, the majority of TEFs are conserved among eukaryotes indicating that they may be adapted to the respective genomic context and consequently to the specific requirements for efficient transcription. Here we have examined the genomic distribution of three different types of TEFs in *Arabidopsis* and investigated their role in transcriptional reprogramming by analysing the respective mutant plants under stress conditions.

Our ChIP-seq analysis revealed that SPT5 associates with transcribed regions and its occupancy correlates with that of RNAPII, closely resembling the distribution of S2P. The enrichment of SPT5 at the TES (along with S2P) suggests that it remains associated with RNAPII until termination. Moreover, the SPT5 profile resembles that of the elongation factor ELF1 in *Arabidopsis* (67), which is in line with the synergistic action of yeast SPT4-SPT5 and ELF1 on *in vitro* transcript elongation on nucleosomal templates (68). In mammals, there is a prominent promoter-proximal SPT5 peak close to the TSS at the same position as that of another TEF termed NELF (69), consistent with the role of SPT4-SPT5 and NELF in

establishment/release of RNAPII promoter-proximal pausing immediately down-stream of TSSs (13,14). Yeast and plants do not encode NELF and supposedly do not make use of stable, regulated promoter-proximal pausing of RNAPII the way mammals do (21,70) and in agreement with that no promoter-proximal SPT5 peak is observed in yeast (71,72) and *Arabidopsis* (Figs. 1,2). In mammals and yeast, SPT4-SPT5 stabilises the elongation complex and increases RNAPII processivity (73,74) consistent with its strategic position on the elongating polymerase (75). In line with that we observed upon exposure to HS decreased mapping of transcript reads along the transcribed region towards TES in *SPT4-R3* plants. Apart from its predominant association with transcribed regions, a peak of SPT5 was detected upstream of TSSs of some closely spaced genes. There it co-localises with BDR1 that has been reported to prevent transcriptional interference with closely spaced downstream genes (60). While the SPT5 peaks co-localise with BDR1 summits, it is uncertain whether the accumulation close to TSSs might be an interference with peaks at TESs of neighbouring genes. The functional implication of SPT5 localising upstream of TSSs is unclear, but it may be related to a possible involvement in transcriptional initiation as proposed for mammalian SPT5 (76).

The *Arabidopsis* PAF1C subunit ELF7 is almost exclusively distributed over the transcribed region of protein-coding genes and shows a higher correlation with S5P rather than S2P. Thus, the genomic distribution largely resembles that of the yeast orthologues, which are detected along the gene body with declining occupancy at the TES (71,72). In mammals, PAF1C exhibits a notably different distribution, as in addition to its occupancy over the transcribed region, remarkable enrichment of PAF1C is detected at TSSs and downstream of TESs (77). We noticed in *elf7* plants a striking decline of transcript reads along the transcribed region towards TES after exposure to HS that was more pronounced with longer genes, suggesting reduced processivity. PAF1C has the capacity to stabilise the elongation complex, which facilitates transcript elongation on chromatin templates resulting in enhanced processivity (78–81). Our analyses of the transcriptional defects observed in *elf7* plants revealed a link with the occurrence of introns. Metagene profiles over genes with introns show decreased transcript coverage in *elf7* downstream of the promoter-proximal region and the effect was cumulative over consecutive intron-exon borders. The sites of impaired transcript mapping are also characterised by elevated ELF7 occupancy and are enriched in RNAPII-S5P, which is linked to co-transcriptional splicing (82,83). Ongoing transcription and co-transcriptional intron splicing are closely coupled, which is manifested by physical interactions between the RNAPII elongation complex and the splicing machinery that has been also demonstrated in *Arabidopsis* (25). In various organisms including plants, transcriptional elongation and splicing were found to influence each other in complex and not completely clarified manners (82–85). In this scenario, PAF1C could facilitate recruitment of

splicing factors such as NTC, facilitating intron splicing (86,87) and/or SRSF2/SNRNP, whose recognition motifs were markedly enriched in genes downregulated in *elf7*. As a consequence of splicing defects transcriptional elongation may be indirectly impaired. Just as well the above-mentioned stabilisation of the elongation complex by PAF1C (78–81) could promote RNAPII transcription across exon-intron borders, resulting in enhanced processivity.

The FACT subunits, SSRP1 and SPT16, co-localise over transcribed regions and compared to RNAPII-S2P/S5P are shifted towards the TSS particularly at highly transcribed genes. The distribution of *Arabidopsis* FACT is similar to that observed in yeast, while in metazoa FACT is even more enriched downstream of the TSS, which may be related to the regulated RNAPII promoter-proximal pausing (71,88). Beyond that FACT displayed a nucleosomal periodicity over transcribed genes, corresponding to that of yeast FACT. There, the periodicity was interpreted as preferential binding of FACT to nucleosomes partially unwrapped by transcribing RNAPII (89–91). This is in agreement with recent structural studies, demonstrating that upon unwrapping DNA from the TSS-proximal side of the nucleosome by the approaching RNAPII elongation complex and exposure of the proximal H2A-H2B dimer, FACT is recruited to assist further nucleosome transcription (92,93). Upon HS in *ssrp1* plants (in which SSRP1 is downregulated to <50%, (37)) transcript reads accumulated just upstream of the +1 nucleosome, close to the site of FACT enrichment (Fig. 7). This suggests that under HS conditions FACT is required for early transcript elongation, particularly for efficient transcription through the +1 nucleosome. Since there is only minor overlap between differentially expressed genes in *tfIIs* (41) and *ssrp1* relative to Col-0 upon exposure to HS, it is likely that TFIIS and FACT distinctly affect transcription through the +1 nucleosomes. Consistently, the maximum accumulation of transcripts upon HS map ~40 and ~20 bp upstream of the +1 nucleosome centre in *ssrp1* and *tfIIs*, respectively. Similar to yeast FACT (91), *Arabidopsis* FACT accumulates downstream of the +1 nucleosome. Early transcript elongation appears to be remarkably responsive to altered temperatures in *Arabidopsis*, as promoter-proximal accumulation of RNAPII downstream of TSS occurs both upon exposure to cold and heat (41,94,95). Another striking feature related to FACT was the occurrence of a prominent NDR upstream of the TSS at FACT enriched genes that is not seen with FACT non-enriched genes or with genes enriched in ELF7 or SPT5. Depletion of FACT from mouse cells leads to a loss of promoter-proximal nucleosomal occupancy resulting in a NDR upstream of TSSs that is associated with increased transcription (96). As in our experimental setup, the MNase cleavage was followed by H3-ChIP before sequencing, we conclude that the highly MNase sensitive region upstream of TSSs of FACT enriched genes is depleted in nucleosomes. This is consistent with earlier experiments that revealed that at a subset of genes phosphorylation of *Arabidopsis* SPT16 is required to establish or retain a NDR upstream of the TSS (43). NDRs play important roles in regulating RNAPII

transcription as they are sites of transcription factor binding, preinitiation complex assembly and RNAPII loading (97,98). Thus, in *Arabidopsis* and mouse the histone chaperone FACT modulates nucleosome occupancy upstream of TSSs and in *Arabidopsis* FACT is particularly required upon HS for the efficient upregulation of genes. The applied rapid HS results in a massive transcriptional reprogramming of hundreds of genes (41,57), hence apparently necessitating FACT activity to adapt to the altered conditions.

Our study analysing the genomic occupancy of three types of RNAPII elongation factors in plants revealed that SPT4-SPT5, PAF1C and FACT predominantly localise to transcribed regions, albeit with distinct distribution patterns. Examination of corresponding mutant plants under acute HS conditions highlighted that different TEFs are diversely required to elicit an appropriate transcriptional response. These insights into post-initiation regulation of gene activity in *Arabidopsis* will serve as a foundation for future studies addressing transcript elongation also in other plant species. Moreover, it represents a starting point for comparative studies to unveil the similarities and dissimilarities of RNAPII transcriptional regulation relative to yeast and metazoan models.

## DATA AVAILABILITY

ChIP-Seq data and RNA-Seq data was deposited at the sequence read archive under the BioProject PRJNA931822 and PRJNA756828, respectively. MNase-Seq and NET-Seq data is available under the BioProject PRJNA877815 and the GEO accession GSE117014, respectively.

## SUPPLEMENTARY DATA

Supplementary data are available at NAR Online.

## ACKNOWLEDGEMENTS

We would like to thank Christoph Moehle und Thomas Stempfl for continuous support of this project regarding NGS.

## FUNDING

Our research is supported by the German Research Foundation (DFG) through grants Gr1159/16-1 and SFB960/A6 to KDG.

**REFERENCES**


1. Svejstrup, J.Q. (2004) The RNA polymerase II transcription cycle: cycling through chromatin, *Biochim. Biophys. Acta,* **1677,** 64–73.
2. 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.
3. Jeronimo, C., Collin, P. and Robert, F. (2016) The RNA polymerase II CTD: the increasing complexity of a low-complexity protein domain, *J. Mol. Biol.,* **428,** 2607–2622.
4. Harlen, K.M. and Churchman, L.S. (2017) The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain, *Nat. Rev. Mol. Cell Biol.,* **18,** 263–273.
5. 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.
6. Noe Gonzalez, M., Blears, D. and Svejstrup, J.Q. (2021) Causes and consequences of RNA polymerase II stalling during transcript elongation, *Nat. Rev. Mol. Cell Biol.,* **22,** 3–21.
7. Chen, F.X., Smith, E.R. and Shilatifard, A. (2018) Born to run: control of transcription elongation by RNA polymerase II, *Nat. Rev. Mol. Cell Biol.,* **19,** 464–478.
8. Kwak, H. and Lis, J.T. (2013) Control of transcriptional elongation, *Ann. Rev. Genet.,* **47,** 483–508.
9. Osman, S. and Cramer, P. (2020) Structural Biology of RNA Polymerase II Transcription: 20 Years On, *Ann. Rev. Cell Dev. Biol.,* **36,** 1–34.
10. Ehara, H. and Sekine, S.-I. (2018) Architecture of the RNA polymerase II elongation complex: new insights into Spt4/5 and Elf1, *Transcription,* **9,** 286–291.
11. Schier, A.C. and Taatjes, D.J. (2020) Structure and mechanism of the RNA polymerase II transcription machinery, *Genes Dev.,* **34,** 465–488.
12. Hartzog, G.A. and Fu, J. (2013) The Spt4-Spt5 complex: a multi-faceted regulator of transcription elongation, *Biochim. Biophys. Acta,* **1829,** 105–115.
13. Decker, T.-M. (2021) Mechanisms of Transcription Elongation Factor DSIF (Spt4-Spt5), *J. Mol. Biol.,* **433,** 166657.
14. Song, A. and Chen, F.X. (2022) The pleiotropic roles of SPT5 in transcription, *Transcription,* **13,** 53–69.
15. Francette, A.M., Tripplehorn, S.A. and Arndt, K.M. (2021) The Paf1 Complex: A Keystone of Nuclear Regulation Operating at the Interface of Transcription and Chromatin, *J. Mol. Biol.,* **433,** 166979.
16. Jaehning, J.A. (2010) The Paf1 complex: platform or player in RNA polymerase II transcription?, *Biochim. Biophys. Acta,* **1799,** 279–388.
17. Formosa, T. and Winston, F. (2020) The role of FACT in managing chromatin: disruption, assembly, or repair?, *Nucleic Acids Res.,* **48,** 11929–11941.
18. Gurova, K., Chang, H.-W., Valieva, M.E., Sandlesh, P. and Studitsky, V.M. (2018) Structure and function of the histone chaperone FACT - Resolving FACTual issues, *Biochim. Biophys. Acta,* **1861,** 892–904.
19. Jeronimo, C. and Robert, F. (2022) The histone chaperone FACT: a guardian of chromatin structure integrity, *Transcription,* **13,** 16–38.
20. van Lijsebettens, M. and Grasser, K.D. (2014) Transcript elongation factors: shaping transcriptomes after transcript initiation, *Trends Plant Sci.,* **19,** 717–726.

21. Obermeyer, S., Kapoor, H., Markusch, H. and Grasser, K.D. (2023) Transcript elongation by RNA polymerase II in plants: factors, regulation and impact on gene expression, *Plant J.,* doi: 10.1111/tpj.16115.
22. 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.
23. Xue, M., Zhang, H., Zhao, F., Zhao, T., Li, H. and Jiang, D. (2021) The INO80 chromatin remodeling complex promotes thermomorphogenesis by connecting H2A.Z eviction and active transcription in Arabidopsis, *Mol. Plant,* **14,** 1799–1813.
24. Liaqat, A., Alfatih, A., Jan, S.U., Sun, L., Zhao, P. and Xiang, C. (2023) Transcription elongation factor AtSPT4-2 positively modulates salt tolerance in Arabidopsis thaliana, *BMC Plant Biol.,* **23,** 49.
25. Antosz, W., Pfab, A., Ehrnsberger, H.F., Holzinger, P., Köllen, K., Mortensen, S.A., Bruckmann, A., Schubert, T., Längst, G., Griesenbeck, J., Schubert, V., Grasser, M. and Grasser, K.D. (2017) 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.
26. 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.
27. Oh, S., Zhang, H., Ludwig, P. and van Nocker, S. (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.
28. Park, S., Oh, S., Ek-Ramos, J. and van Nocker, S. (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.
29. 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.
30. Nasim, Z., Susila, H., Jin, S., Youn, G. and Ahn, J.H. (2022) Polymerase II-Associated Factor 1 Complex-Regulated FLOWERING LOCUS C-Clade Genes Repress Flowering in Response to Chilling, *Front. Plant Sci.,* **13,** 817356.
31. Jensen, G.S., Fal, K., Hamant, O. and Haswell, E.S. (2017) The RNA Polymerase-Associated Factor 1 Complex Is Required for Plant Touch Responses, *J. Exp. Bot.,* **68,** 499–511.
32. Obermeyer, S., Stöckl, R., Schnekenburger, T., Moehle, C., Schwartz, U. and Grasser, K.D. (2022) Distinct role of subunits of the Arabidopsis RNA polymerase II elongation factor PAF1C in transcriptional reprogramming, *Front. Plant Sci.,* **13,** 974625.
33. Zhang, H., Li, X., Song, R., Zhan, Z., Zhao, F., Li, Z. and Jiang, D. (2022) Cap-binding complex assists RNA polymerase II transcription in plant salt stress response, *Plant Cell Environ.,* **45,** 2780–2793.
34. Duroux, M., Houben, A., Růzicka, K., Friml, J. and Grasser, K.D. (2004) The chromatin remodelling complex FACT associates with actively transcribed regions of the Arabidopsis genome, *Plant J.,* **40,** 660–671.
35. Perales, M. and Más, P. (2007) A functional link between rhythmic changes in chromatin structure and the Arabidopsis biological clock, *Plant Cell,* **19,** 2111–2123.

36. Frost, J.M., Kim, M.Y., Park, G.T., Hsieh, P.-H., Nakamura, M., Lin, S.J.H., Yoo, H., Choi, J., Ikeda, Y., Kinoshita, T., Choi, Y., Zilberman, D. and Fischer, R.L. (2018) FACT complex is required for DNA demethylation at heterochromatin during reproduction in Arabidopsis, *Proc. Natl. Acad. Sci. USA,* **115,** E4720-E4729.
37. Lolas, I.B., Himanen, K., Grønlund, J.T., Lynggaard, C., Houben, A., Melzer, M., van Lijsebettens, M. and Grasser, K.D. (2010) The transcript elongation factor FACT affects Arabidopsis vegetative and reproductive development and genetically interacts with HUB1/2, *Plant J.,* **61,** 686–697.
38. Ma, Y., Gil, S., Grasser, K.D. and Mas, P. (2018) Targeted Recruitment of the Basal Transcriptional Machinery by LNK Clock Components Controls the Circadian Rhythms of Nascent RNAs in Arabidopsis, *Plant Cell,* **30,** 907–924.
39. Nielsen, M., Ard, R., Leng, X., Ivanov, M., Kindgren, P., Pelechano, V. and Marquardt, S. (2019) Transcription-driven chromatin repression of Intragenic transcription start sites, *PLoS Genet.,* **15,** e1007969.
40. Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassay with tobacco tissue cultures, *Physiol.Plant.,* **15,** 473–497.
41. Obermeyer, S., Stöckl, R., Schnekenburger, T., Kapoor, H., Stempfl, T., Schwartz, U. and Grasser, K.D. (2023) TFIIS is crucial during early transcript elongation for transcriptional reprogramming in response to heat stress, *J. Mol. Biol.,* **435,** 167917.
42. Antosz, W., Deforges, J., Begcy, K., Bruckmann, A., Poirier, Y., Dresselhaus, T. and Grasser, K.D. (2020) Critical Role of Transcript Cleavage in Arabidopsis RNA Polymerase II Transcriptional Elongation, *Plant Cell,* **32,** 1449–1463.
43. Michl-Holzinger, P., Obermeyer, S., Markusch, H., Pfab, A., Ettner, A., Bruckmann, A., Babl, S., Längst, G., Schwartz, U., Tvardovskiy, A., Jensen, O.N., Osakabe, A., Berger, F. and Grasser, K.D. (2022) Phosphorylation of the FACT histone chaperone subunit SPT16 affects chromatin at RNA polymerase II transcriptional start sites in *Arabidopsis*, *Nucleic Acids Res.,* **50,** 5014–5028.
44. Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic:a flexible trimmer for Illumina sequence data, *Bioinformatics,* **30,** 2114–2120.
45. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2, *Nat. Meth.,* **9,** 357–359.
46. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M. and Li, H. (2021) Twelve years of SAMtools and BCFtools, *GigaScience,* **10**.
47. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F. and Manke, T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis, *Nucleic Acids Res.,* **44,** W160-5.
48. Quadrana, L., Bortolini Silveira, A., Mayhew, G.F., LeBlanc, C., Martienssen, R.A., Jeddeloh, J.A. and Colot, V. (2016) The Arabidopsis thaliana mobilome and its impact at the species level, *Elife,* **5,** e15716.
49. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W. and Liu, X.S. (2008) Model-based analysis of ChIP-Seq (MACS), *Genome Biol.,* **9,** R137.
50. Ewels, P., Magnusson, M., Lundin, S. and Käller, M. (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report, *Bioinformatics,* **32,** 3047–3048.
51. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner, *Bioinformatics,* **29,** 15–21.

52. Okonechnikov, K., Conesa, A. and García-Alcalde, F. (2016) Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data, *Bioinformatics,* **32,** 292–294.
53. Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features, *Bioinformatics,* **30,** 923–930.
54. Love, M.L., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.,* **15,** 550.
55. Haag, J.R. and Pikaard, C.S. (2011) Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing, *Nat. Rev. Mol. Cell Biol.,* **12,** 483–492.
56. Zhu, J., Liu, M., Liu, X. and Dong, Z. (2018) RNA polymerase II activity revealed by GRO-seq and pNET-seq in *Arabidopsis*, *Nat. Plants,* **4,** 1112–1123.
57. Szádeczky-Kardoss, I., Szaker, H.M., Verma, R., Darkó, É., Pettkó-Szandtner, A., Silhavy, D. and Csorba, T. (2022) Elongation factor TFIIS is essential for heat stress adaptation in plants, *Nucleic Acids Res.,* **50,** 1927–1950.
58. Dhaliwal, N.K. and Mitchell, J.A. (2016) Nuclear RNA Isolation and Sequencing, *Methods Mol. Biol.,* **1402,** 63–71.
59. Zaghlool, A., Ameur, A., Nyberg, L., Halvardson, J., Grabherr, M., Cavelier, L. and Feuk, L. (2013) Efficient cellular fractionation improves RNA sequencing analysis of mature and nascent transcripts from human tissues, *BMC Biotechnol.,* **13,** 99.
60. Yu, X., Martin, P.G.P. and Michaels, S.D. (2019) BORDER proteins protect expression of neighboring genes by promoting 3' Pol II pausing in plants, *Nat. Commun.,* **10,** 4359.
61. Mieczkowski, J., Cook, A., Bowman, S.K., Mueller, B., Alver, B.H., Kundu, S., Deaton, A.M., Urban, J.A., Larschan, E., Park, P.J., Kingston, R.E. and Tolstorukov, M.Y. (2016) MNase titration reveals differences between nucleosome occupancy and chromatin accessibility, *Nat. Commun.,* **7,** 11485.
62. Zhang, T., Zhang, W. and Jiang, J. (2015) Genome-Wide Nucleosome Occupancy and Positioning and Their Impact on Gene Expression and Evolution in Plants, *Plant Physiol,* **168,** 1406–1416.
63. 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.,* **78,** 706–714.
64. Cortijo, S., Charoensawan, V., Brestovitsky, A., Buning, R., Ravarani, C., Rhodes, D., van Noort, J., Jaeger, K.E. and Wigge, P.A. (2017) Transcriptional Regulation of the Ambient Temperature Response by H2A.Z Nucleosomes and HSF1 Transcription Factors in Arabidopsis, *Mol. Plant,* **10,** 1258–1273.
65. Piovesan, A., Caracausi, M., Antonaros, F., Pelleri, M.C. and Vitale, L. (2016) GeneBase 1.1: a tool to summarize data from NCBI gene datasets and its application to an update of human gene statistics, *Database,* **2016**.
66. Derelle, E., Ferraz, C., Rombauts, S., Rouzé, P., Worden, A.Z., Robbens, S., Partensky, F., Degroeve, S., Echeynié, S., Cooke, R., Saeys, Y., Wuyts, J., Jabbari, K., Bowler, C., Panaud, O., Piégu, B., Ball, S.G., Ral, J.-P., Bouget, F.-Y., Piganeau, G., Baets, B. de, Picard, A., Delseny, M., Demaille, J., van de Peer, Y. and Moreau, H. (2006) Genome analysis of the smallest free-living eukaryote Ostreococcus tauri unveils many unique features, *Proc. Natl. Acad. Sci. USA,* **103,** 11647–11652.
67. Markusch, H., Michl-Holzinger, P., Obermeyer, S., Thorbecke, C., Bruckmann, A., Babl, S., Längst, G., Osakabe, A., Berger, F. and Grasser, K.D. (2023) ELF1 is a component of the Arabidopsis RNA polymerase II elongation complex and associates with a subset of transcribed genes, *New Phytol.,* **238,** 113-124.

68. Ehara, H., Kujirai, T., Fujino, Y., Shirouzu, M., Kurumizaka, H. and Sekine, S.-I. (2019) Structural insight into nucleosome transcription by RNA polymerase II with elongation factors, *Science,* **363,** 744–747.
69. Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A. and Young, R.A. (2010) c-Myc regulates transcriptional pause release, *Cell,* **141,** 432–445.
70. Core, L. and Adelman, K. (2019) Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation, *Genes Dev.,* **33,** 960–982.
71. Rossi, M.J., Kuntala, P.K., Lai, W.K.M., Yamada, N., Badjatia, N., Mittal, C., Kuzu, G., Bocklund, K., Farrell, N.P., Blanda, T.R., Mairose, J.D., Basting, A.V., Mistretta, K.S., Rocco, D.J., Perkinson, E.S., Kellogg, G.D., Mahony, S. and Pugh, B.F. (2021) A high-resolution protein architecture of the budding yeast genome, *Nature,* **592,** 309–314.
72. 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.
73. Hu, S., Peng, L., Xu, C., Wang, Z., Song, A. and Chen, F.X. (2021) SPT5 stabilizes RNA polymerase II, orchestrates transcription cycles, and maintains the enhancer landscape, *Mol. Cell,* **81,** 4425-4439.
74. Wada, T., Takagi, G., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F., Buratowski, S. and Handa, H. (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.
75. 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.
76. Diamant, G., Bahat, A. and Dikstein, R. (2016) The elongation factor Spt5 facilitates transcription initiation for rapid induction of inflammatory-response genes, *Nat. Commun.,* **7,** 11547.
77. 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.
78. 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.
79. Hou, L., Wang, Y., Liu, Y., Zhang, N., Shamovsky, I., Nudler, E., Tian, B. and Dynlacht, B.D. (2019) Paf1C regulates RNA polymerase II progression by modulating elongation rate, *Proc. Natl. Acad. Sci. USA,* **116,** 14583–14592.
80. van den Heuvel, D., Spruijt, C.G., González-Prieto, R., Kragten, A., Paulsen, M.T., Di Zhou, Wu, H., Apelt, K., van der Weegen, Y., Yang, K., Dijk, M., Daxinger, L., Marteijn, J.A., Vertegaal, A.C.O., Ljungman, M., Vermeulen, M. and Luijsterburg, M.S. (2021) A CSB-PAF1C axis restores processive transcription elongation after DNA damage repair, *Nat. Commun.,* **12,** 1342.
81. Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H. and Cramer, P. (2018) Structure of activated transcription complex Pol II-DSIF-PAF-SPT6, *Nature,* **560,** 607–612.
82. Tellier, M., Maudlin, I. and Murphy, S. (2020) Transcription and splicing: A two-way street, *WIREs RNA,* **11,** e1593.
83. Giono, L.E. and Kornblihtt, A.R. (2020) Linking transcription, RNA polymerase II elongation and alternative splicing, *Biochem. J.,* **477,** 3091–3104.

84. Marquardt, S., Petrillo, E. and Manavella, P.A. (2023) Cotranscriptional RNA processing and modification in plants, *Plant Cell,* doi 10.1093/plcell/koac309.
85. Godoy Herz, M.A. and Kornblihtt, A.R. (2019) Alternative Splicing and Transcription Elongation in Plants, *Front. Plant Sci.,* **10,** 309.
86. Koncz, C., Dejong, F., Villacorta, N., Szakonyi, D. and Koncz, Z. (2012) The spliceosome-activating complex: molecular mechanisms underlying the function of a pleiotropic regulator, *Front. Plant Sci.,* **3,** 9.
87. Li, Y., Yang, J., Shang, X., Lv, W., Xia, C., Wang, C., Feng, J., Cao, Y., He, H., Li, L. and Ma, L. (2019) SKIP regulates environmental fitness and floral transition by forming two distinct complexes in Arabidopsis, *New Phytol.,* **224,** 321–335.
88. Tettey, T.T., Gao, X., Shao, W., Li, H., Story, B.A., Chitsazan, A.D., Glaser, R.L., Goode, Z.H., Seidel, C.W., Conaway, R.C., Zeitlinger, J., Blanchette, M. and Conaway, J.W. (2019) A Role for FACT in RNA Polymerase II Promoter-Proximal Pausing, *Cell Rep.,* **27,** 3770-3779.e7.
89. Martin, B.J.E., Chruscicki, A.T. and Howe, L.J. (2018) Transcription Promotes the Interaction of the FAcilitates Chromatin Transactions (FACT) Complex with Nucleosomes in Saccharomyces cerevisiae, *Genetics,* **210,** 869–881.
90. Jeronimo, C., Angel, A., Nguyen, V.Q., Kim, J.M., Poitras, C., Lambert, E., Collin, P., Mellor, J., Wu, C. and Robert, F. (2021) FACT is recruited to the +1 nucleosome of transcribed genes and spreads in a Chd1-dependent manner, *Mol. Cell,* **81,** 3542-3559.e11.
91. 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.,* **28,** 357–366.
92. Ehara, H., Kujirai, T., Shirouzu, M., Kurumizaka, H. and Sekine, S.-I. (2022) Structural basis of nucleosome disassembly and reassembly by RNAPII elongation complex with FACT, *Science,* **377,** doi 10.1126/science.abp9466.
93. Farnung, L., Ochmann, M., Engeholm, M. and Cramer, P. (2021) Structural basis of nucleosome transcription mediated by Chd1 and FACT, *Nat. Struct. Mol. Biol.,* **28,** 382–387.
94. Liu, M., Zhu, J. and Dong, Z. (2021) Immediate transcriptional responses of Arabidopsis leaves to heat shock, *J.Integr.Plant Biol.,* **63,** 468–483.
95. Kindgren, P., Ivanov, M. and Marquardt, S. (2020) Native elongation transcript sequencing reveals temperature dependent dynamics of nascent RNAPII transcription in Arabidopsis, *Nucleic Acids Res.,* **48,** 2332–2347.
96. Mylonas, C. and Tessarz, P. (2018) Transcriptional repression by FACT is linked to regulation of chromatin accessibility at the promoter of ES cells, *Life Sci. Alliance,* **1,** e201800085.
97. Brahma, S. and Henikoff, S. (2020) Epigenome Regulation by Dynamic Nucleosome Unwrapping, *Trends Biochem. Sci.,* **45,** 13–26.
98. Lai, W.K.M. and Pugh, B.F. (2017) Understanding nucleosome dynamics and their links to gene expression and DNA replication, *Nat. Rev. Mol. Cell Biol.,* **18,** 548–562.

**FIGURE LEGENDS**

**Figure 1.** Elongation factors exhibit distinct distribution over transcribed regions relative to each other and relative to RNAPII. **A-C**, Metagene profiles of TEFs and elongating RNAPII (S2P and S5P) distribution based on ChIP-seq data (based on four replicates for SPT5, ELF7, SSRP1, SPT16 and two replicates for S2P, S5P). TEF and RNAPII ChIP-seq coverage (normalised to counts per million reads) was plotted over TAIR10 annotated, non-overlapping, protein coding genes (n= 24474). Lines are colour-coded as indicated and represent the accumulation of RNAPII and elongation factors, while the shaded area indicates the standard error of the mean. For comparison input samples of the respective TEF ChIPs are shown. Genes were scaled over the annotated region and 500bp upstream and downstream of TSS and TES are shown unscaled. **D**, Pearsson correlation together with hierarchical clustering illustrates the differential distribution of TEFs and RNAPII. Pairwise correlation was calculated over all TAIR10 annotated, non-overlapping, protein coding genes (n= 24474) divided into 10 bins per gene.

**Figure 2.** RNAPII and elongation factor occupancy increases with transcript levels. **A**, Enrichment of elongating RNAPII (S2P, S5P) and elongation factors (SPT5, ELF7, SSRP1, SPT16) along transcribed regions depicted as metagene plots based on ChIP-seq data. Genes were divided into four groups according to transcript levels (logTPM, transcripts per million) based on RNA-seq data of Col-0 plants (n=6177 in each group with quartile “low Tx”: logTPM 0.0 to -1.9; quartile “mid Tx”: logTPM 0.6 to 0.0; quartile “high Tx”: logTPM 0.9 to 0.6; quartile “very high Tx”: logTPM 3.4 to 0.9). Lines are colour-coded as indicated and represent the normalised coverage of RNAPII and elongation factors, while the shaded area indicates the standard error of the mean. **B**, Differential occupancy of RNAPII and elongation factors along transcribed regions dependent on transcript levels shown as heatmaps. Genes were are divided in four groups depending on transcript levels (as in A; n=6177) and they are sorted from high expression (top) to low expression (bottom) as indicated on the left as color-coded logTPMs (transcripts per million) of genes in Col-0 (no HS).

**Figure 3.** Genomic distribution of the TEFs relative to RNAPII-S2P and -S5P. **A-C**, Metagene profiles of differential calling of ChIP-seq peaks obtained for SPT5, ELF7 and FACT (SSRP1, SPT16) relative to those obtained for RNAPII-S2P (TEF/S2P) and S5P (TEF/S5P). Lines are colour-coded as indicated and represent the accumulation of elongation factors relative to RNAPII, while the shaded area indicates the standard error of

the mean. Please note the different scale of the y-axes. **D**, Number of detected, differential peaks relative to S2P and S5P. **E**, Genomic distribution of the detected differential peaks. Schematic diagram on top depicting the proximal region (0-300 bp downstream of TSS) and gene body (transcribed region without 0-300 bp). Bottom scheme depicts the distribution of the differential peaks relative to S2P and S5P in the promoter proximal region, gene body and outside of transcribed regions. For comparison, the distribution of randomly called peaks within the genome is shown (random).

**Figure 4.** Mutants deficient in elongation factors exhibit distinct transcriptional defects. RNA-seq data (based on three replicates per genotype and condition) of mutants deficient in elongation factors (**A**, *SPT4-R3*; **B**, *elf7*; **C**, *ssrp1*) relative to Col-0 wildtype plants after 10 min HS (HS) or without HS (no HS). TSS-centred metagene profiles of the sequencing data split according to expression level of genes after HS into four groups as indicated (low Tx, mid Tx, high Tx and very high Tx with n=6177 genes for each group). Replicates were merged and the log-fold changes (LFC) between mutant and Col-0 of no HS vs. HS were calculated with TPM normalisation. **E-F**, Heatmaps illustrate the shift of mapped RNA-seq reads of the transcribed gene groups (n=6177) towards the TSS upon exposure to HS.

**Figure 5.** Beyond its prevalent coverage over transcribed regions, ChIP-seq peaks of SPT5 enriched relative to RNAPII-S5P are detected upstream of TSSs of closely spaced genes and correlate with BDR1 peaks. **A**, TAIR10 annotated, non-overlapping, protein coding genes were divided into two classes according to their distance to the next annotated gene. Genes with less than 500bp distance to the TSS or TES of the next annotated gene were classified as closely spaced (<500bp; n=9544), while genes with more than 500bp distance were classified as not closely spaced (>500bp; n=7668). SPT5 ChIP-seq peaks were centred to the TSS. **B,C**, ChIP-seq peaks of SPT5, ELF7 and FACT were centred over ChIP-seq summits (n=21334) of BDR1 (60). Lines (in A,B) are colour-coded as indicated and represent the accumulation of elongation factors, while the shaded area indicates the standard error of the mean.

**Figure 6.** The transcription defects observed in *elf7* plants over transcribed regions are correlated with intron/exon borders. **A**, RNA-seq data of HS relative to no HS (LFC) for *elf7* and Col-0 plants plotted over genes downregulated in *elf7* compared to Col-0 (n=290), centred to the TSS (position of TSS is indicated by a dotted line). **B**, RNA-seq data of *elf7*

and Col-0 plants of HS relative to no HS (LFC) plotted over genes upregulated in *elf7* compared to Col-0 (n=129), centred to the TSS. **C**, RNA-seq data of *elf7* and Col-0 plants of HS relative to no HS (LFC) plotted over genes highly expressed upon HS containing ≥10 introns (n=290), centred to the TSS. **D**, RNA-seq data of *elf7* and Col-0 plants of HS relative to no HS (LFC) plotted over genes highly expressed upon HS containing no introns (n=355), centred to the TSS. **E,F**, RNA-seq data of the respective mutant relative to Col-0 with no HS (E) and upon HS (F) over TAIR10 annotated intron-exon borders (I│E) of non-overlapping, protein-coding genes (n=24474) starting from left to right consecutively with I│E2 to I│E10. Dotted line indicates intron-exon borders.

**Figure 7.** FACT is enriched over nucleosomal positions, modulates the accessability of the promoter-proximal NDR and upon HS is required for efficient transcription through +1 nucleosomes. **A,** SSRP1 and SPT16 exhibit a more pronounced nucleosomal pattern than ELF7 and SPT5. The respective TEF ChIP-seq coverage was centred over nucleosome positions (+1 to +5) from MNase H3 ChIP-seq data (n=26295 for +1 nucleosome) (41). **B,** Transcriptional activity (Tx, orange), SSRP1 relative to RNAPII-S2P enrichment (red) and mononucleosomal coverage (grey) over all TAIR10 annotated, non-overlapping, protein-coding genes sorted by mean SSRP1 signal intensity (n=24474 ). **C-H** Peaks called from FACT (SSRP1 and SPT16) relative to RNAPII-S2P ChIP-seq data were divided into two clusters according to high FACT occupancy (FACT enriched; n=10360) and low FACT occupancy (FACT non-enriched; n=13901) using unsupervised k-means clustering. **C**, Genes from both clusters were intersected with +1 nucleosomal positions. Called peaks from SSRP1 and SPT16 ChIP-seq data were centred to the +1 nucleosomal position in the respective cluster. **D,** MNase H3 ChIP-seq data was centered to the +1 nucleosomal position in the respective cluster. **E,** Based on the RNA-seq data, LFC of *ssrp1*/Col-0 without HS and after HS was centred over the +1 nucleosomal position in the respective cluster. **F,** Mean density of FACT/S2P ChIP-seq peaks, **G,** mean MNase H3 ChIP-seq coverage (NDR seen with the FACT enriched cluster indicated by an arrow) and **H,** LFC of *ssrp1*/Col-0 based on RNA-seq after HS centered to the +1 nucleosomal position in the two above-mentioned gene clusters.

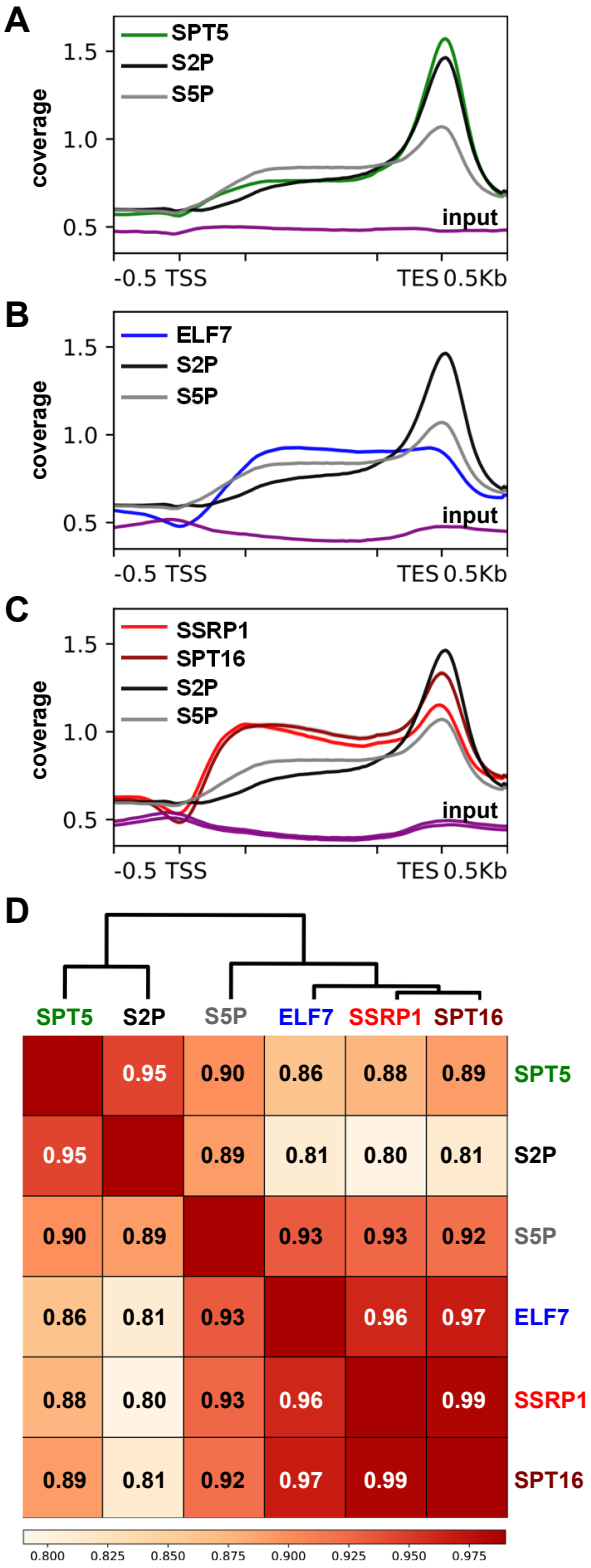



**Figure 1.** Elongation factors exhibit distinct distribution over transcribed regions relative to each other and relative to RNAPII. **A-C**, Metagene profiles of TEFs and elongating RNAPII (S2P and S5P) distribution based on ChIP-seq data (based on four replicates for SPT5, ELF7, SSRP1, SPT16 and two replicates for S2P, S5P). TEF and RNAPII ChIP-seq coverage (normalised to counts per million reads) was plotted over TAIR10 annotated, non-overlapping, protein coding genes (n= 24474). Lines are colour-coded as indicated and represent the accumulation of RNAPII and elongation factors, while the shaded area indicates the standard error of the mean. For comparison input samples of the respective TEF ChIPs are shown. Genes were scaled over the annotated region and 500bp upstream and downstream of TSS and TES are shown unscaled. **D**, Pearsson correlation together with hierarchical clustering illustrates the differential distribution of TEFs and RNAPII. Pairwise correlation was calculated over all TAIR10 annotated, non-overlapping, protein coding genes (n= 24474) divided into 10 bins per gene.

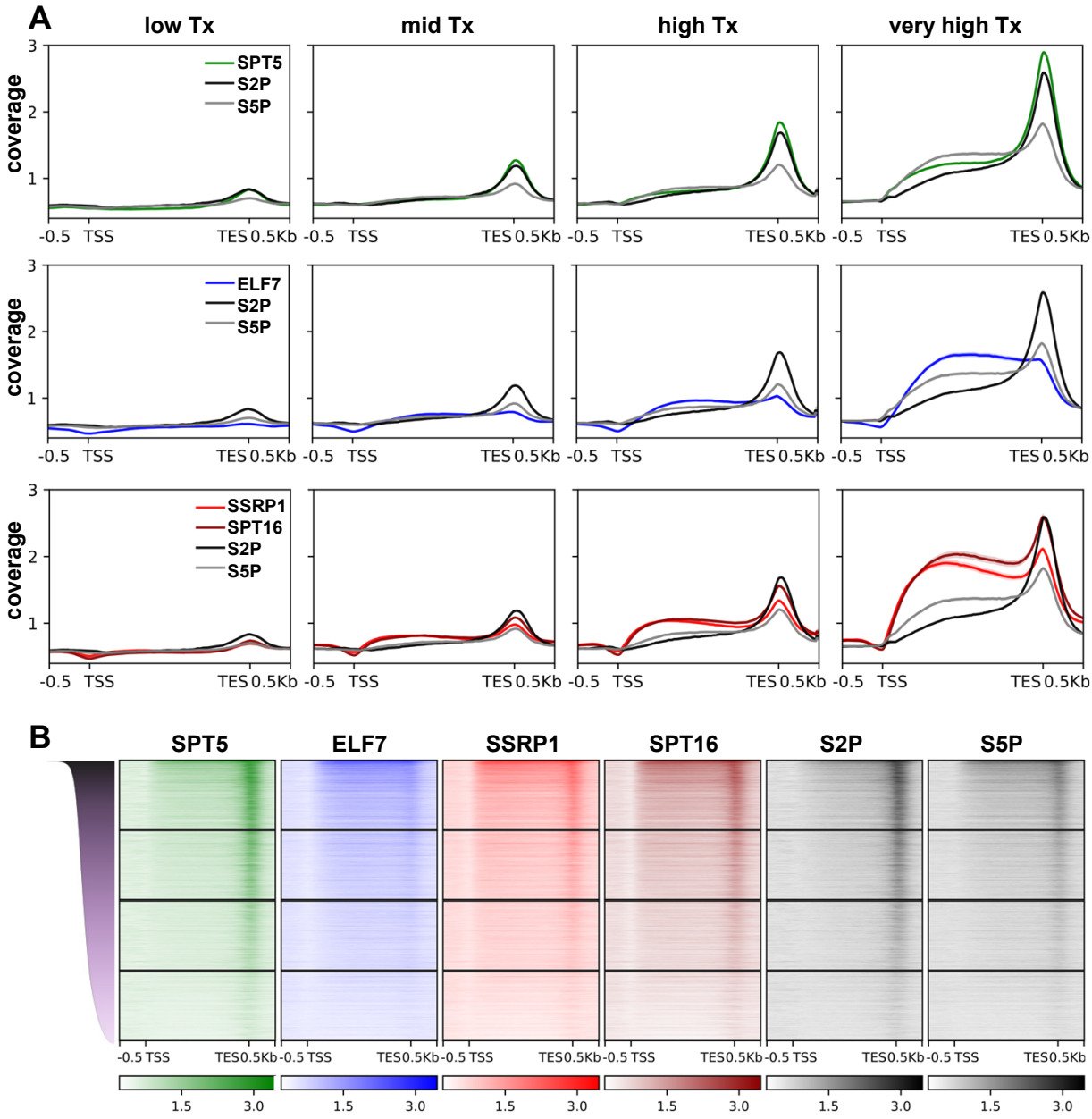



**Figure 2.** RNAPII and elongation factor occupancy increases with transcript levels. **A**, Enrichment of elongating RNAPII (S2P, S5P) and elongation factors (SPT5, ELF7, SSRP1, SPT16) along transcribed regions depicted as metagene plots based on ChIP-seq data. Genes were divided into four groups according to transcript levels (logTPM, transcripts per million) based on RNA-seq data of Col-0 plants (n=6177 in each group with quartile “low Tx”: logTPM 0.0 to -1.9; quartile “mid Tx”: logTPM 0.6 to 0.0; quartile “high Tx”: logTPM 0.9 to 0.6; quartile “very high Tx”: logTPM 3.4 to 0.9). Lines are colour-coded as indicated and represent the normalised coverage of RNAPII and elongation factors, while the shaded area indicates the standard error of the mean. **B**, Differential occupancy of RNAPII and elongation factors along transcribed regions dependent on transcript levels shown as heatmaps. Genes were are divided in four groups depending on transcript levels (as in A; n=6177) and they are sorted from high expression (top) to low expression (bottom) as indicated on the left as color-coded logTPMs (transcripts per million) of genes in Col-0 (no HS).

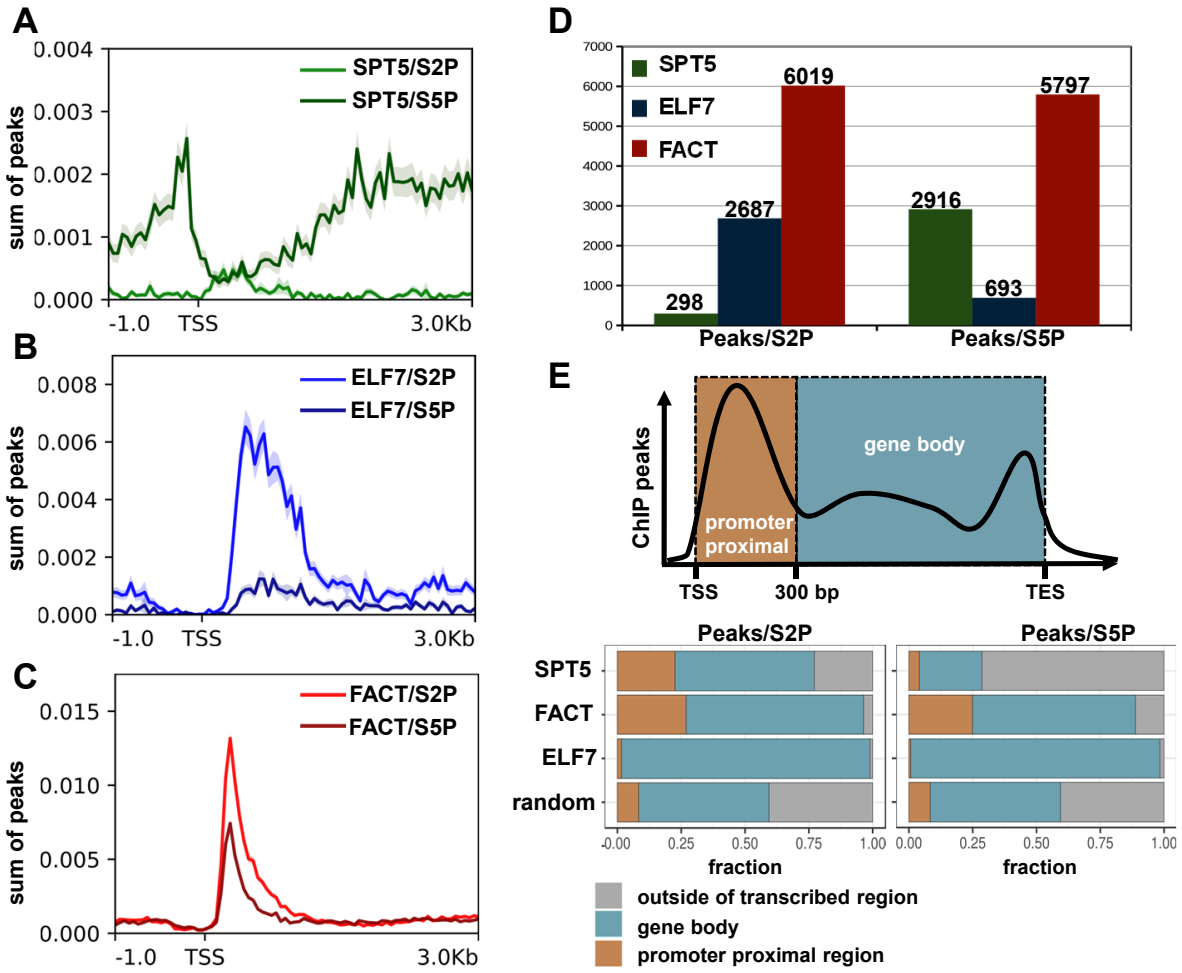



**Figure 3.** Genomic distribution of the TEFs relative to RNAPII-S2P and -S5P. **A-C**, Metagene profiles of differential calling of ChIP-seq peaks obtained for SPT5, ELF7 and FACT (SSRP1, SPT16) relative to those obtained for RNAPII-S2P (TEF/S2P) and S5P (TEF/S5P). Lines are colour-coded as indicated and represent the accumulation of elongation factors relative to RNAPII, while the shaded area indicates the standard error of the mean. Please note the different scale of the y-axes. **D**, Number of detected, differential peaks relative to S2P and S5P. **E**, Genomic distribution of the detected differential peaks. Schematic diagram on top depicting the proximal region (0-300 bp downstream of TSS) and gene body (transcribed region without 0-300 bp). Bottom scheme depicts the distribution of the differential peaks relative to S2P and S5P in the promoter proximal region, gene body and outside of transcribed regions. For comparison, the distribution of randomly called peaks within the genome is shown (random).

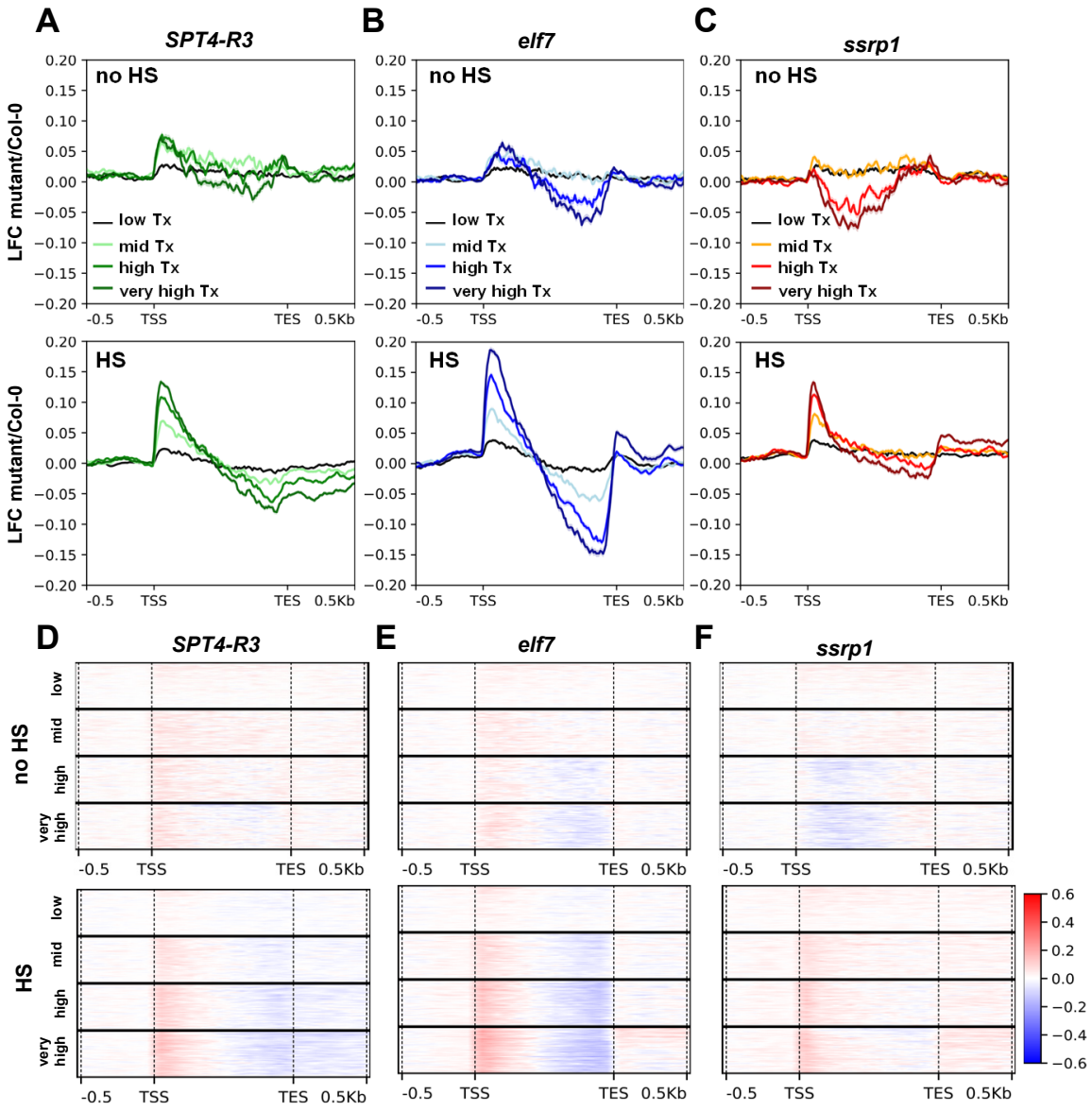



**Figure 4.** Mutants deficient in elongation factors exhibit distinct transcriptional defects. RNA-seq data (based on three replicates per genotype and condition) of mutants deficient in elongation factors (**A**, *SPT4-R3*; **B**, *elf7*; **C**, *ssrp1*) relative to Col-0 wildtype plants after 10 min HS (HS) or without HS (no HS). TSS-centred metagene profiles of the sequencing data split according to expression level of genes after HS into four groups as indicated (low Tx, mid Tx, high Tx and very high Tx with n=6177 genes for each group). Replicates were merged and the log-fold changes (LFC) between mutant and Col-0 of no HS vs. HS were calculated with TPM normalisation. **E-F**, Heatmaps illustrate the shift of mapped RNA-seq reads of the transcribed gene groups (n=6177) towards the TSS upon exposure to HS.

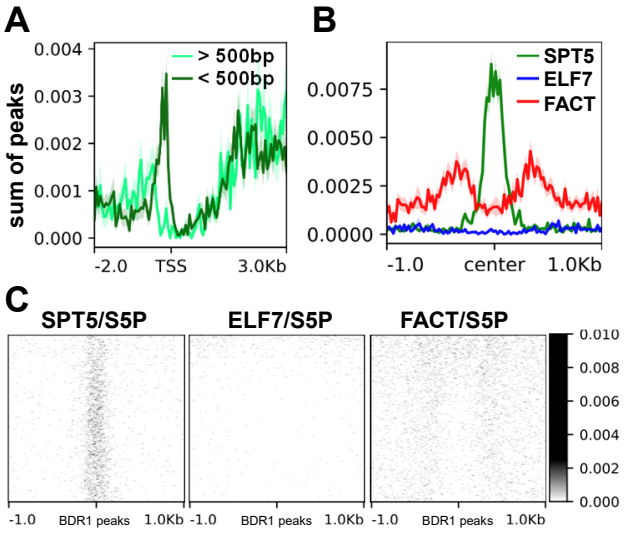



**Figure 5.** Beyond its prevalent coverage over transcribed regions, ChIP-seq peaks of SPT5 enriched relative to RNAPII-S5P are detected upstream of TSSs of closely spaced genes and correlate with BDR1 peaks. **A**, TAIR10 annotated, non-overlapping, protein coding genes were divided into two classes according to their distance to the next annotated gene. Genes with less than 500bp distance to the TSS or TES of the next annotated gene were classified as closely spaced (<500bp; n=9544), while genes with more than 500bp distance were classified as not closely spaced (>500bp; n=7668). SPT5 ChIP-seq peaks were centred to the TSS. **B,C**, ChIP-seq peaks of SPT5, ELF7 and FACT were centred over ChIP-seq summits (n=21334) of BDR1 (60). Lines (in A,B) are colour-coded as indicated and represent the accumulation of elongation factors, while the shaded area indicates the standard error of the mean.

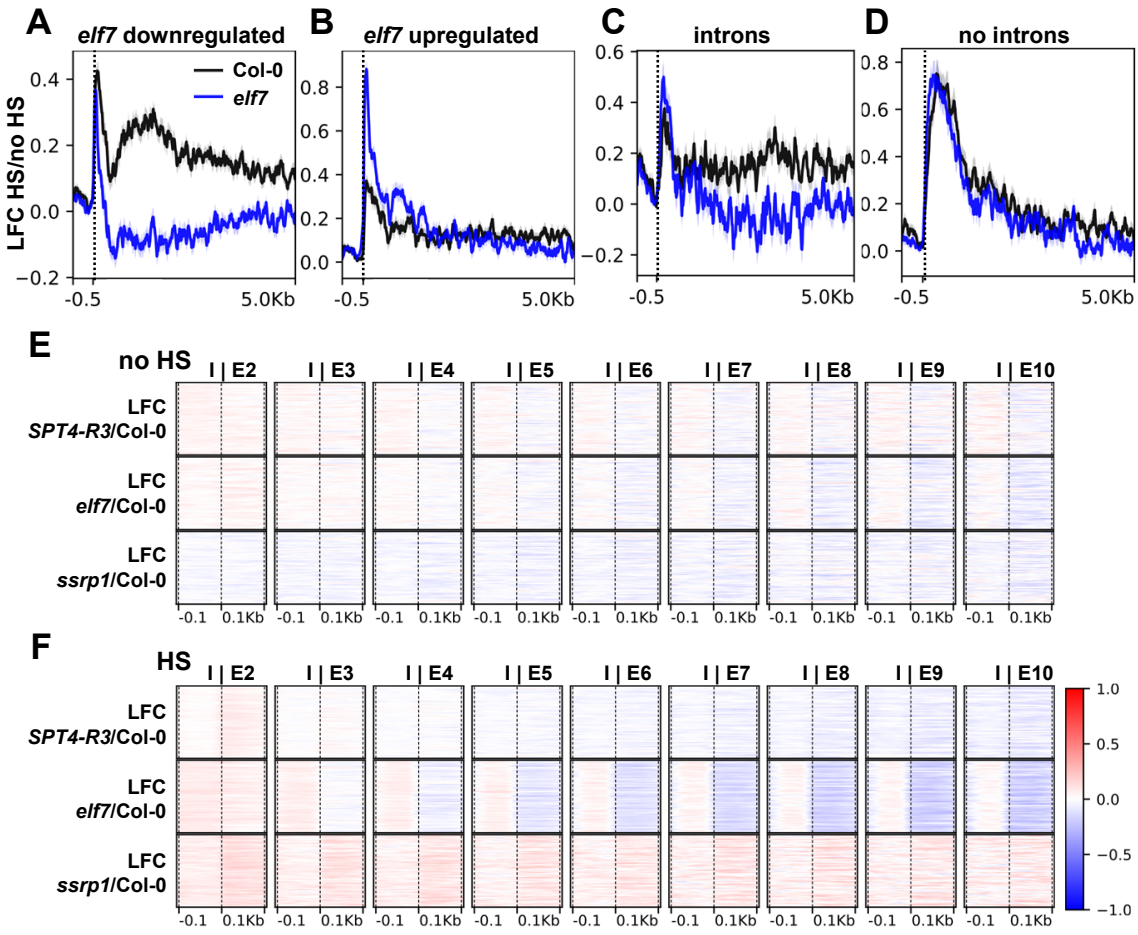



**Figure 6.** The transcription defects observed in *elf7* plants over transcribed regions are correlated with intron/exon borders. **A**, RNA-seq data of HS relative to no HS (LFC) for *elf7* and Col-0 plants plotted over genes downregulated in *elf7* compared to Col-0 (n=290), centred to the TSS (position of TSS is indicated by a dotted line). **B**, RNA-seq data of *elf7* and Col-0 plants of HS relative to no HS (LFC) plotted over genes upregulated in *elf7* compared to Col-0 (n=129), centred to the TSS. **C**, RNA-seq data of *elf7* and Col-0 plants of HS relative to no HS (LFC) plotted over genes highly expressed upon HS containing ≥10 introns (n=290), centred to the TSS. **D**, RNA-seq data of *elf7* and Col-0 plants of HS relative to no HS (LFC) plotted over genes highly expressed upon HS containing no introns (n=355), centred to the TSS. **E,F**, RNA-seq data of the respective mutant relative to Col-0 with no HS (E) and upon HS (F) over TAIR10 annotated intron-exon borders (I | E) of non-overlapping, protein-coding genes (n=24474) starting from left to right consecutively with I | E2 to I | E10. Dotted line indicates intron-exon borders.

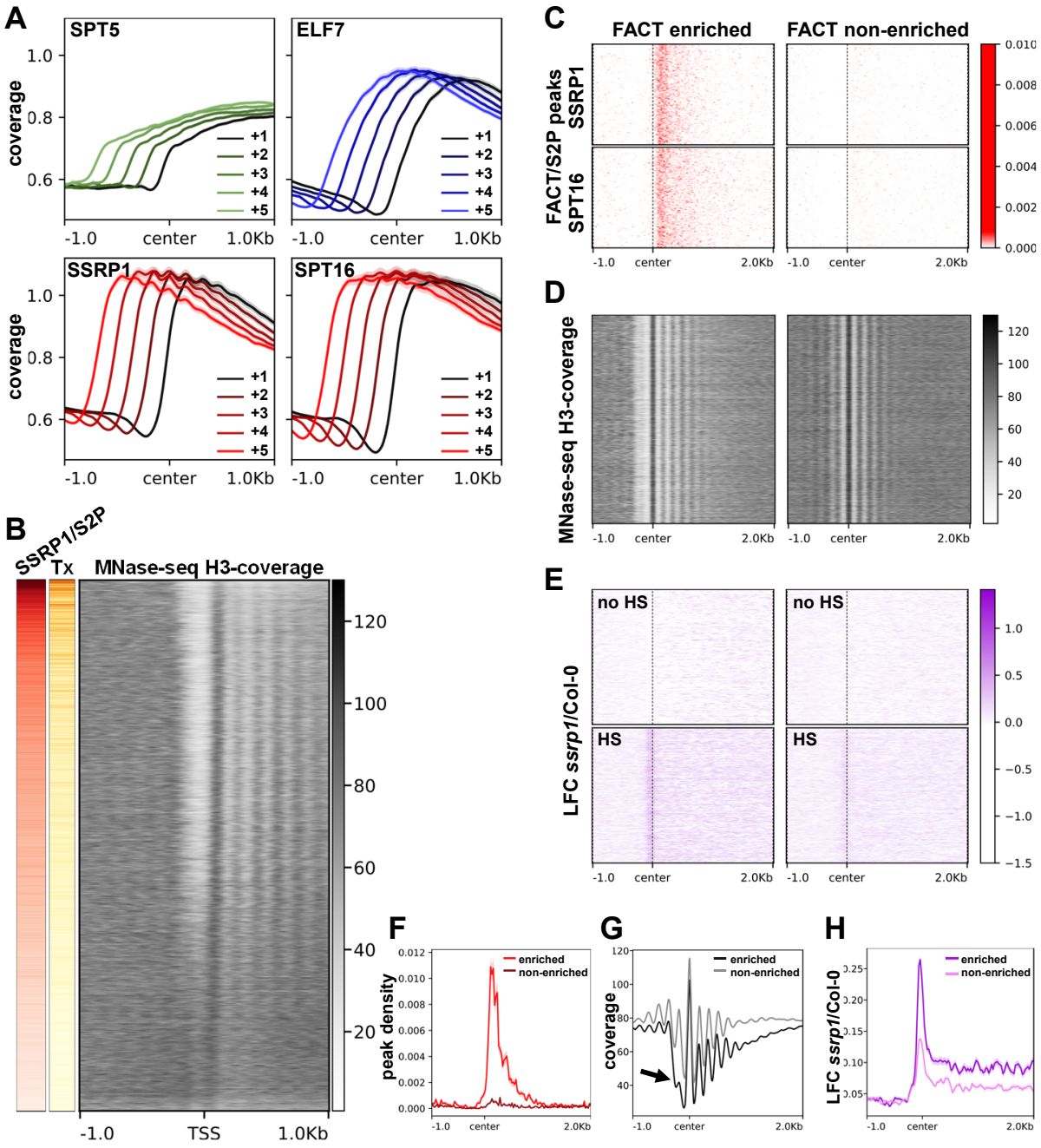
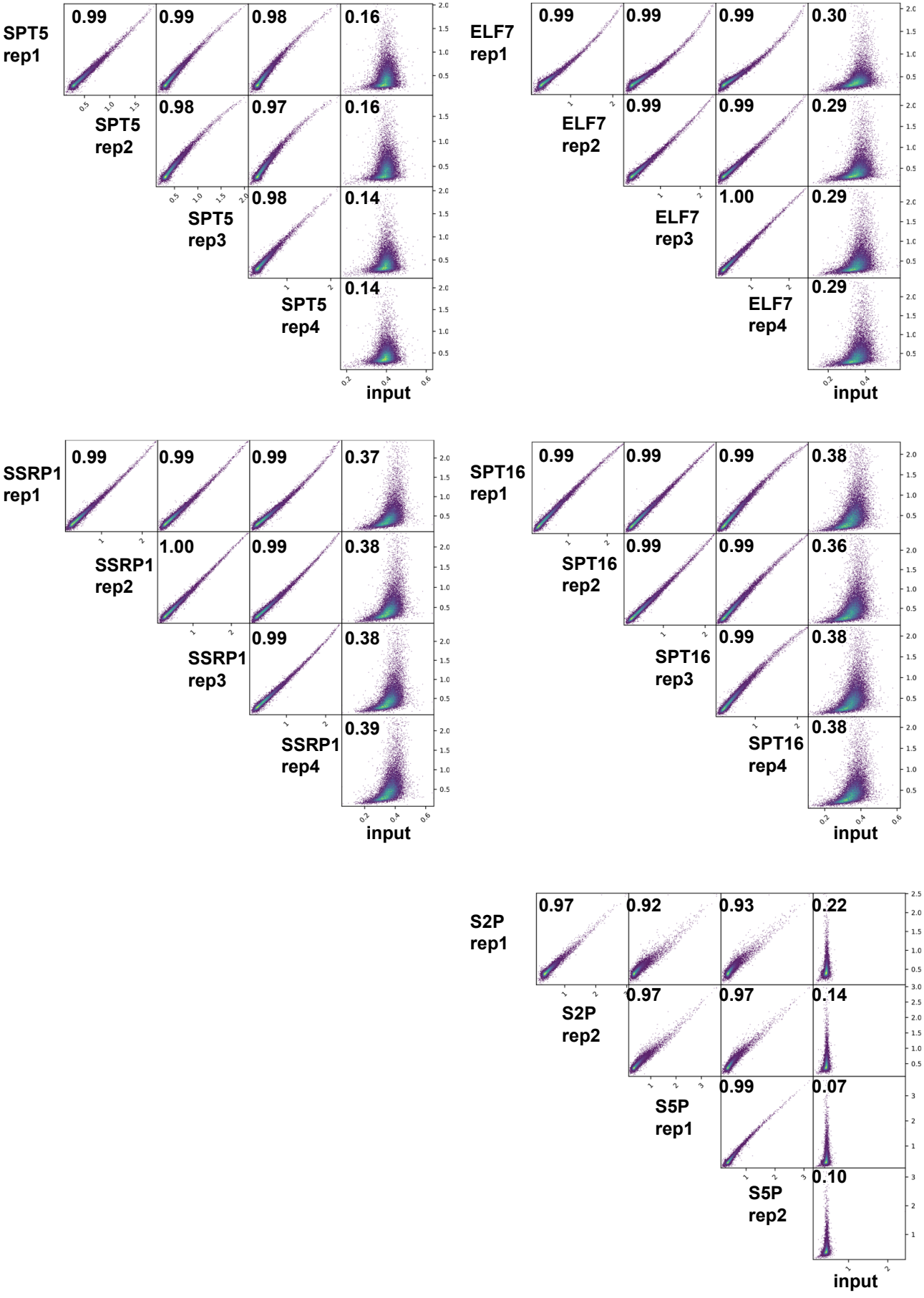
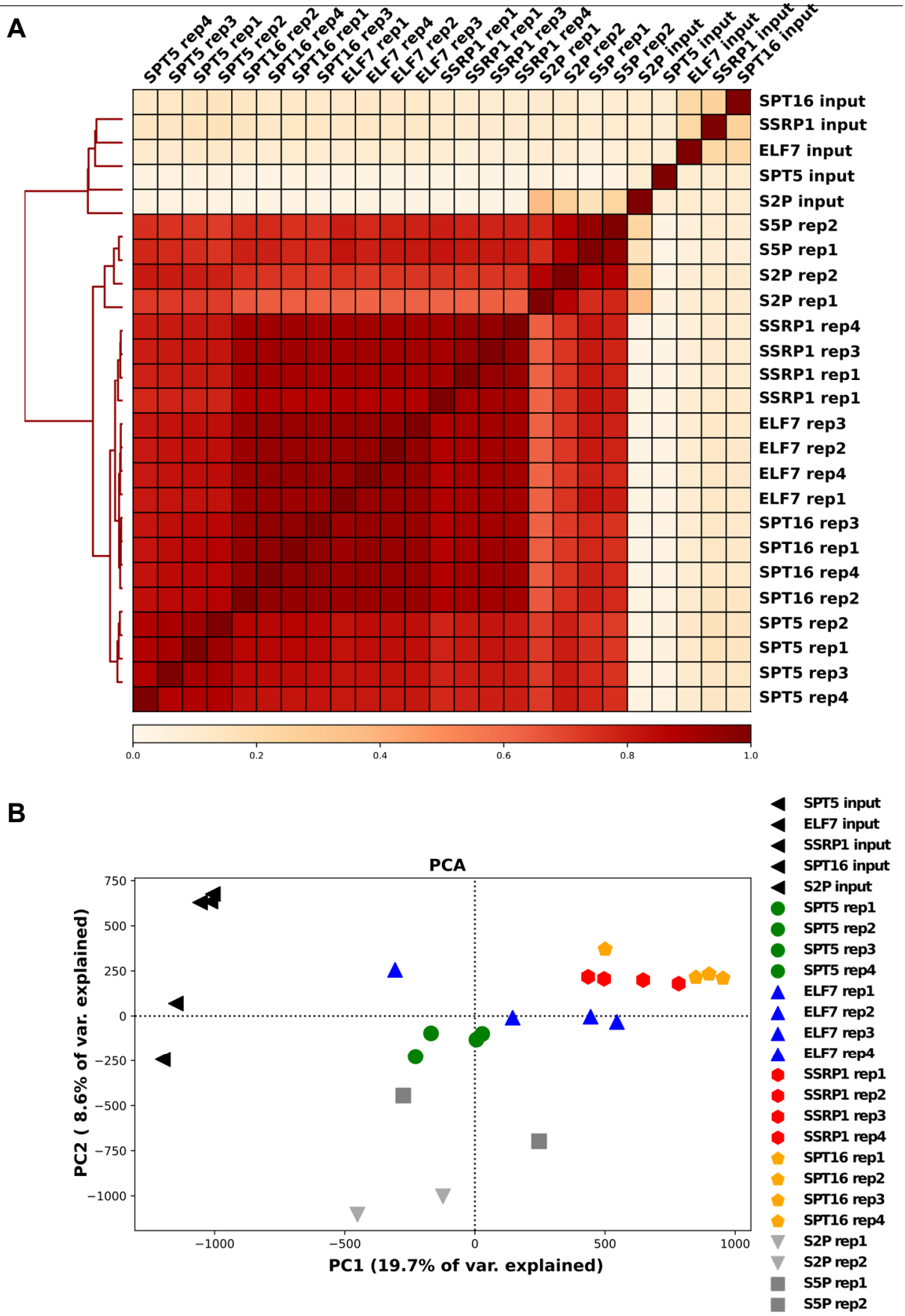


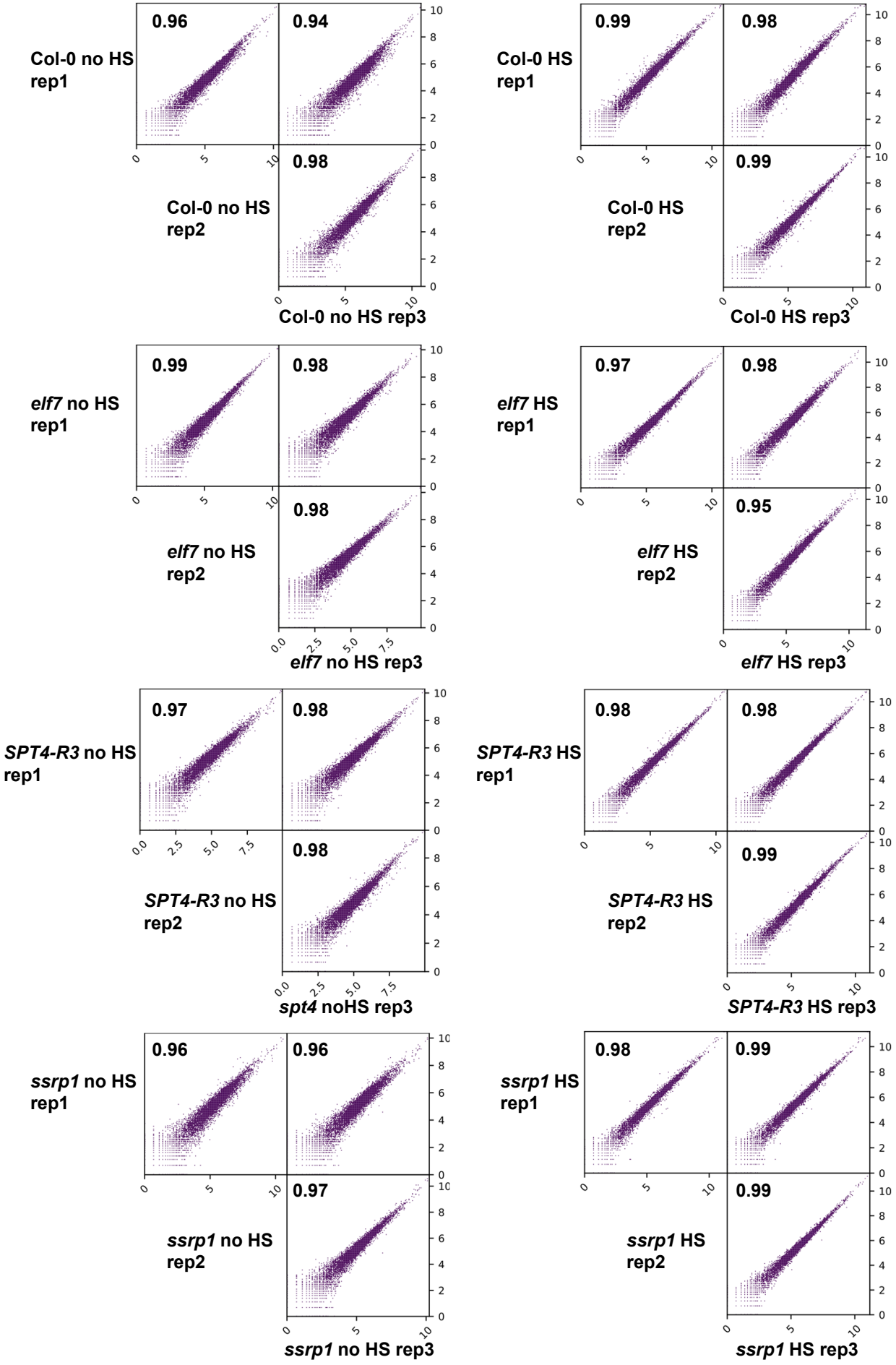

**Figure 7.** FACT is enriched over nucleosomal positions, modulates the accessability of the promoter-proximal NDR and upon HS is required for efficient transcription through +1 nucleosomes. **A,** SSRP1 and SPT16 exhibit a more pronounced nucleosomal pattern than ELF7 and SPT5. The respective TEF ChIP-seq coverage was centred over nucleosome positions (+1 to +5) from MNase H3 ChIP-seq data (n=26295 for +1 nucleosome) (41). **B,** Transcriptional activity (Tx, orange), SSRP1 relative to RNAPII-S2P enrichment (red) and mononucleosomal coverage (grey) over all TAIR10 annotated, non-overlapping, protein-coding genes sorted by mean SSRP1 signal intensity (n=24474 ). **C-H** Peaks called from FACT (SSRP1 and SPT16) relative to RNAPII-S2P ChIP-seq data were divided into two clusters according to high FACT occupancy (FACT enriched; n=10360) and low FACT occupancy (FACT non-enriched; n=13901) using unsupervised k-means clustering. **C**, Genes from both clusters were intersected with +1 nucleosomal positions. Called peaks from SSRP1 and SPT16 ChIP-seq data were centred to the +1 nucleosomal position in the respective cluster. **D,** MNase H3 ChIP-seq data was centered to the +1 nucleosomal position in the respective cluster. **E,** Based on the RNA-seq data, LFC of *ssrp1*/Col-0 without HS and after HS was centred over the +1 nucleosomal position in the respective cluster. **F,** Mean density of FACT/S2P ChIP-seq peaks, **G,** mean MNase H3 ChIP-seq coverage (NDR seen with the FACT enriched cluster indicated by an arrow) and **H,** LFC of *ssrp1*/Col-0 based on RNA-seq after HS centered to the +1 nucleosomal position in the two above-mentioned gene clusters.



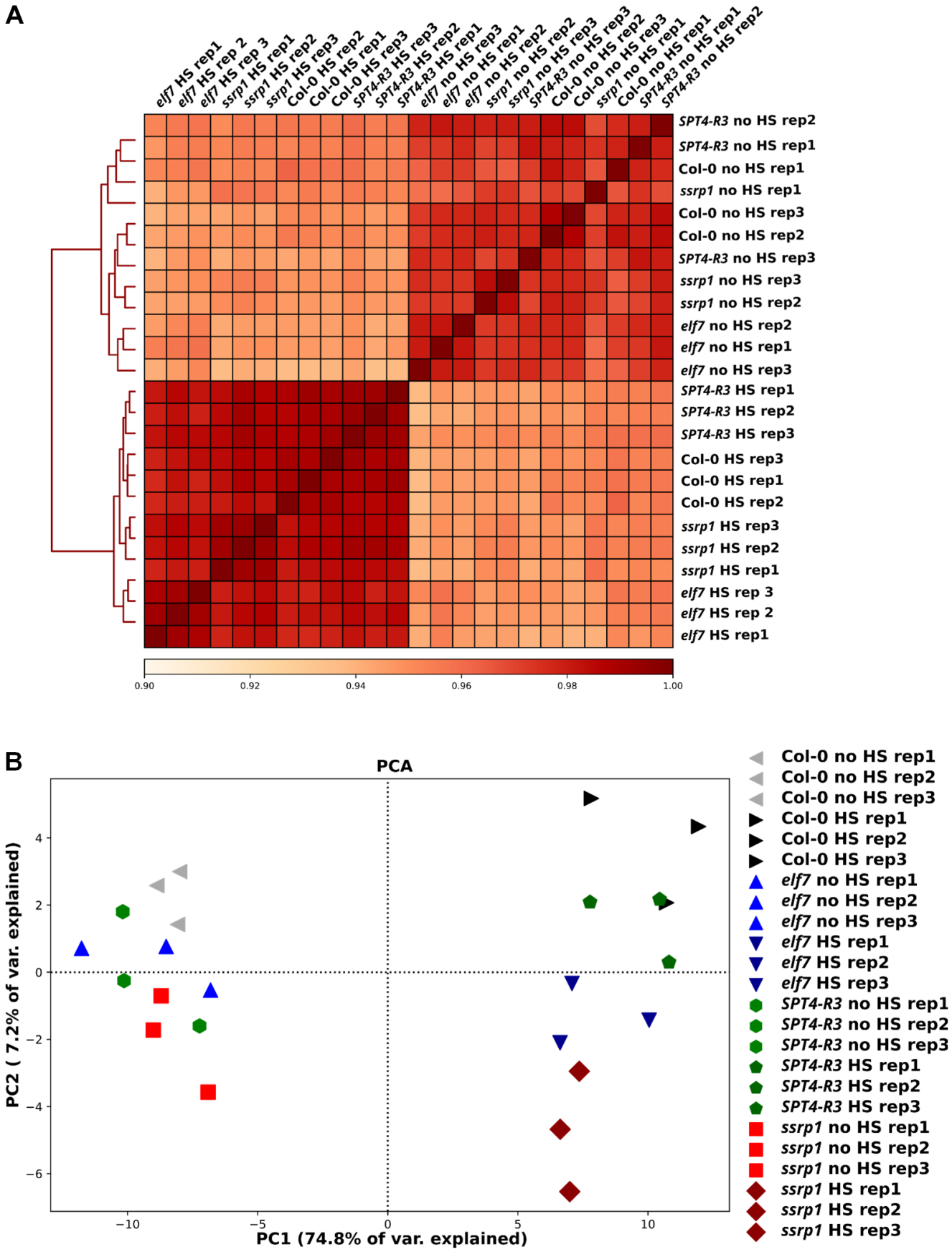

**Supplementary Figure S1.** Pairwise Pearson correlation of TEF ChIP-seq replicates (SPT5, ELF7, SSRP1, SPT16) and RNAPII ChIP-seq replicates (S2P, S5P). In addition, the respective input samples are included over all TAIR10 annotated non-overlapping protein-coding genes (n=24474).



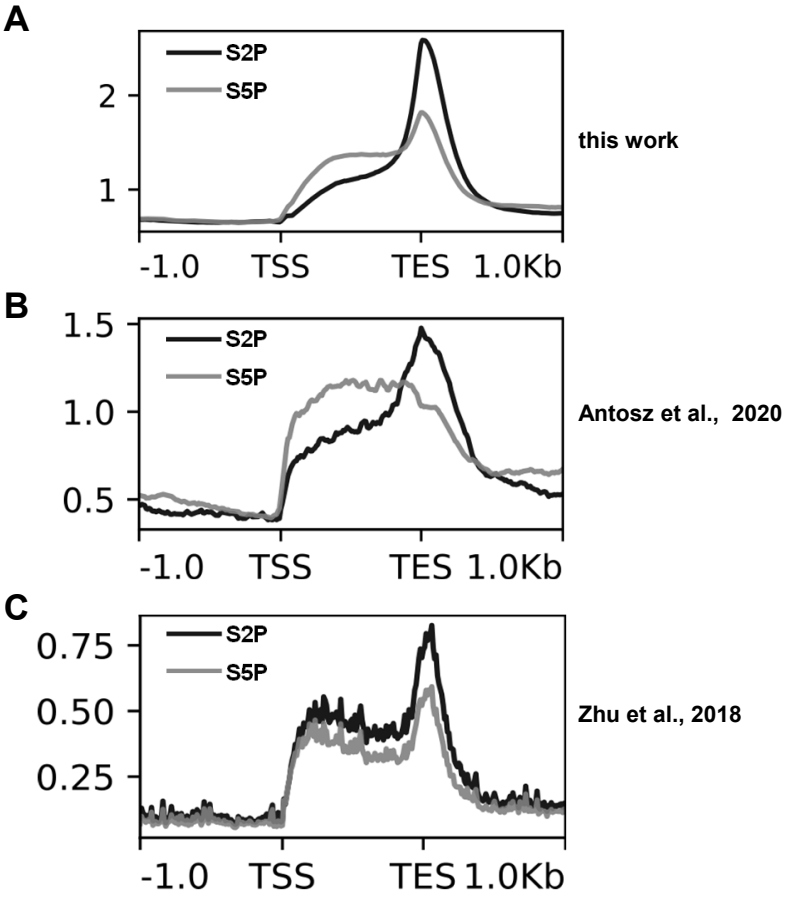

**Supplementary Figure S2.** Hierarchical clustering of pairwise correlation and principal component analysis of the TEF and RNAPII ChIP-seq data as well as input samples. **A**, Heatmap depiction of hierarchical clustering of the pairwise correlations obtained for all the ChIP-seq replicates and input samples. **B**, Principal component analysis of all ChIP-seq replicates and the respective input samples over all TAIR10 annotated non-overlapping protein-coding genes (n=24474).



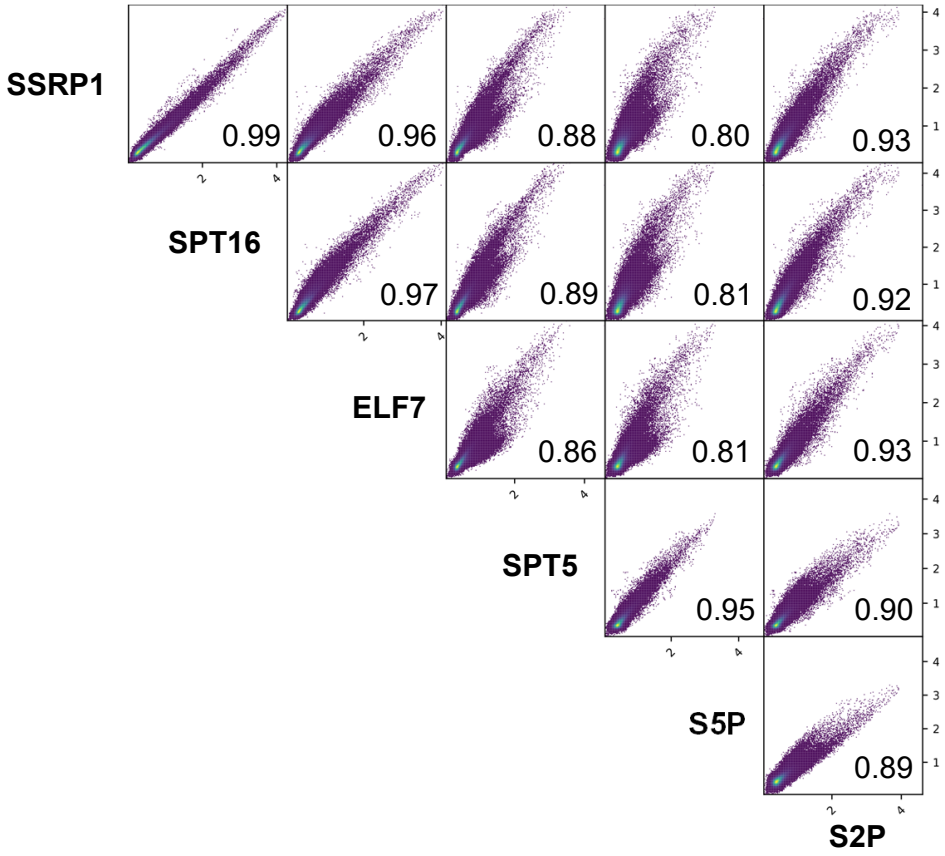

**Supplementary Figure S3.** Pairwise Pearson correlation of RNA-Seq replicates over all TAIR10 annotated non-overlapping protein-coding genes (n=24474).



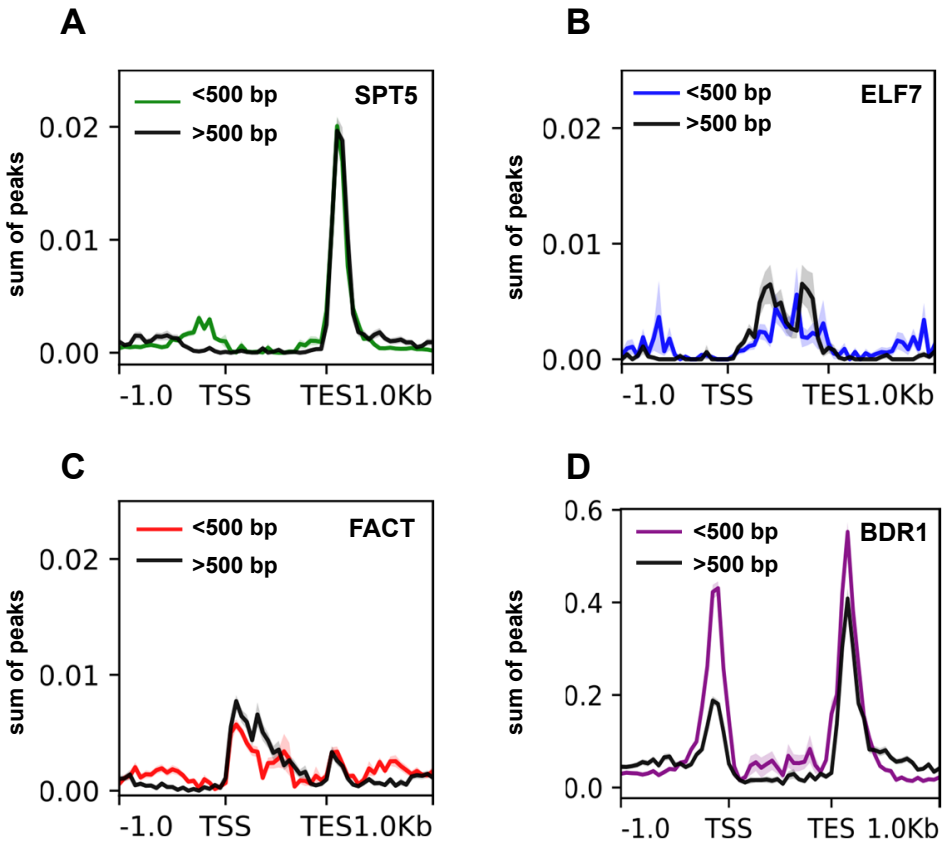

**Supplementary Figure S4.** Hierarchical clustering of pairwise correlation and principal component analysis of all RNA-seq data. **A**, Heatmap depiction of hierarchical clustering of the pairwise correlations obtained for all the RNA-seq replicates. **B**, Principal component analysis of all RNA-seq replicates over all TAIR10 annotated non-overlapping protein-coding genes (n=24474).



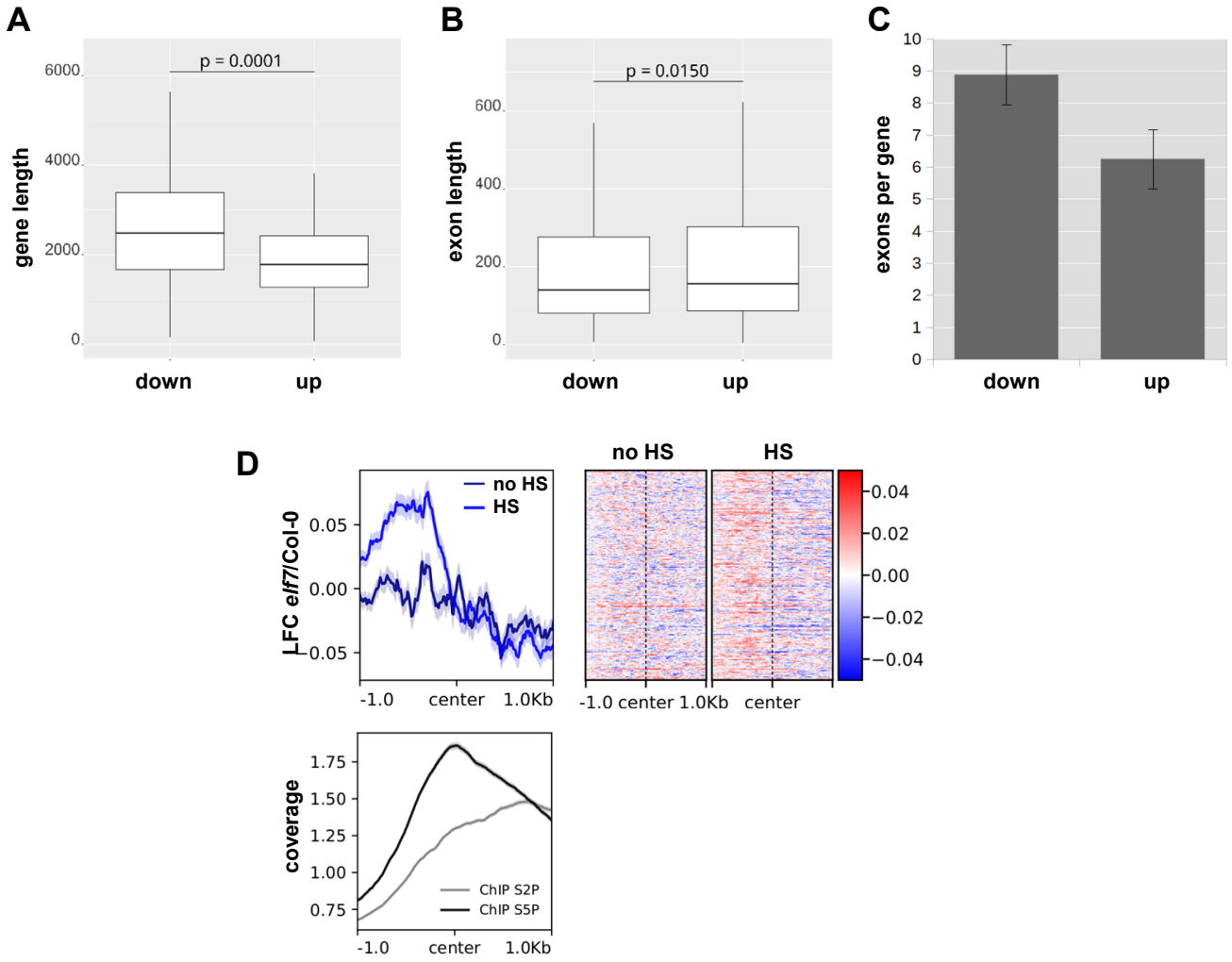

**Supplementary Figure S5.** Comparison of RNAPII-S2P and S5P distribution datasets from different studies. Metagene profiles of RNAPII (S2P and S5P) distribution based on ChIP-seq data (A,B) and NET-Seq data (C). **A**, RNAPII ChIP-Seq coverage (this work) was plotted over highly expressed non-overlapping, protein coding genes (n=6177). **B**, ChIP-seq data (SRA-accession PRJNA527114) from (42), and **C**, NET-Seq data (SRA-acccession PRJNA480966) from (56) over the same set of genes. Lines show the accumulation of RNAPII, while the shaded area indicates the standard error of the mean. Genes were scaled over the annotated region.




**Supplementary Figure S6.** Scatterplots and Pearson correlation of ChIP-seq data of TEFs and RNAPII. Pairwise correlation was calculated over all TAIR10 annotated non-overlapping protein-coding genes (n=24474). Genes were split into 10 bins per gene using the bedtools „makewindows“ function.



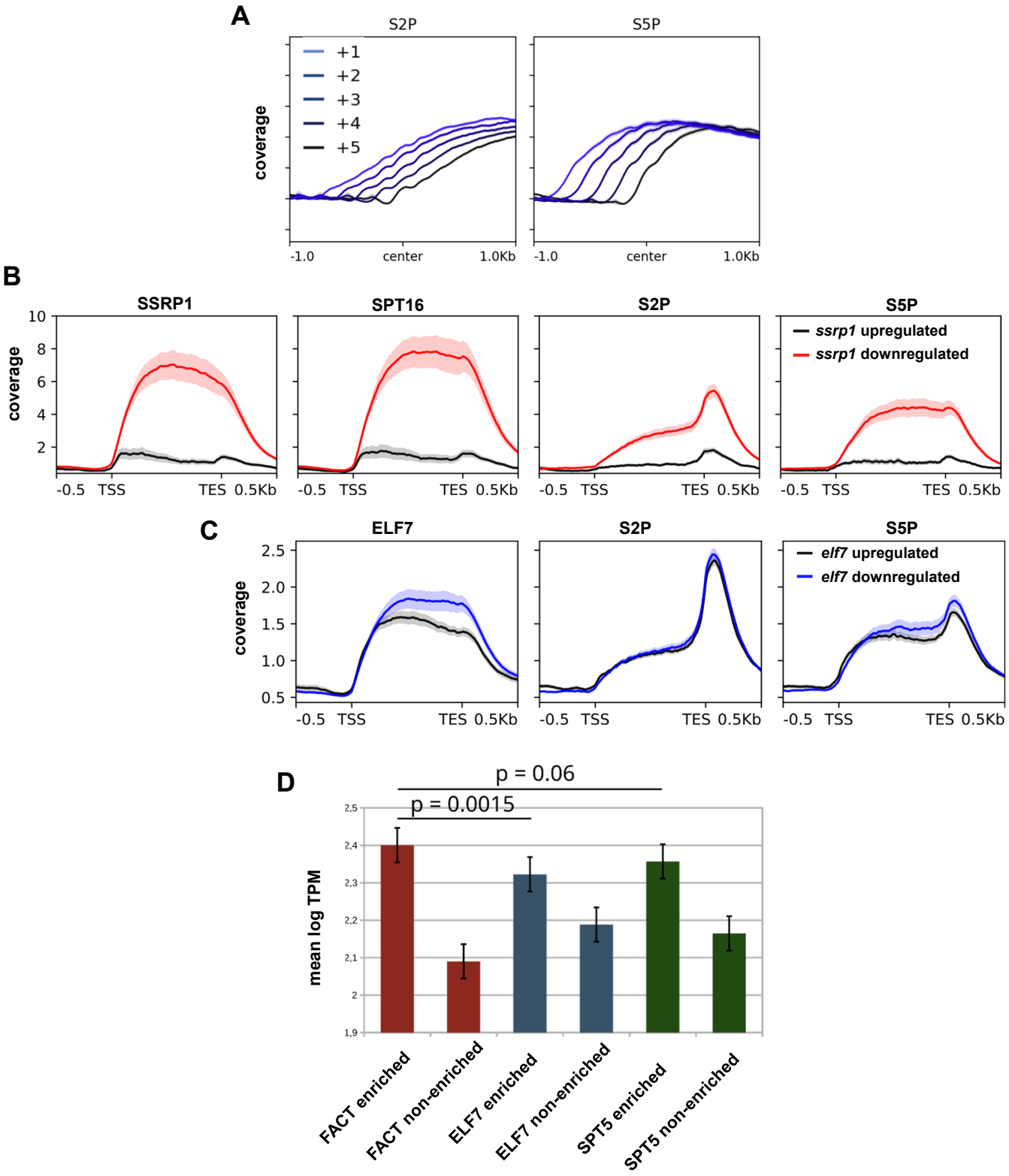

**Supplementary Figure S7.** Genomic distribution of TEFs and BDR1 depending on the distance between genes. ChIP-seq peaks of SPT5 (**A**), ELF7 (**B**), FACT (**C**) and of BDR1 (Yu et al., 2019) (**D**) were plotted over genes divided according to the distance to the next gene. Distribution of peaks over TAIR10 annotated, non-overlapping, protein coding genes with more than 500 bp distance (n=7668) from the TSS to the next gene and more than 500 bp from the TES to the next gene are indicated in black, while distribution of peaks over genes with less than 500 bp distance (n=9544) to the next gene is color coded as indicated.



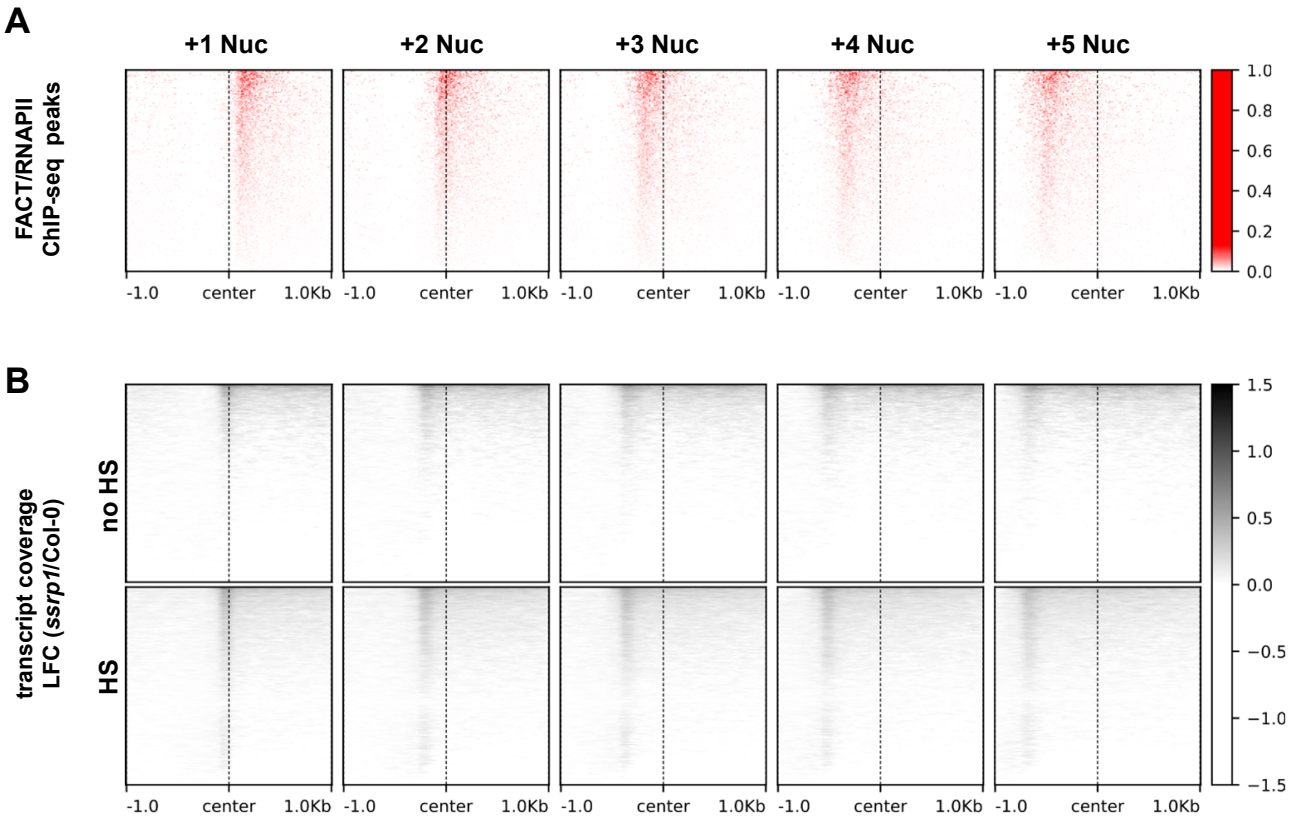

E

| Rank | Motif | P-value | log P-pvalue | % of Targets | % of Background |
|---|---|---|---|---|---|
| 1 |  | 1e-16 | -3.691e+01 | 51.90% | 28.53% |
| 2 |  | 1e-14 | -3.332e+01 | 29.76% | 12.25% |
| 3 |  | 1e-13 | -3.019e+01 | 53.63% | 32.27% |

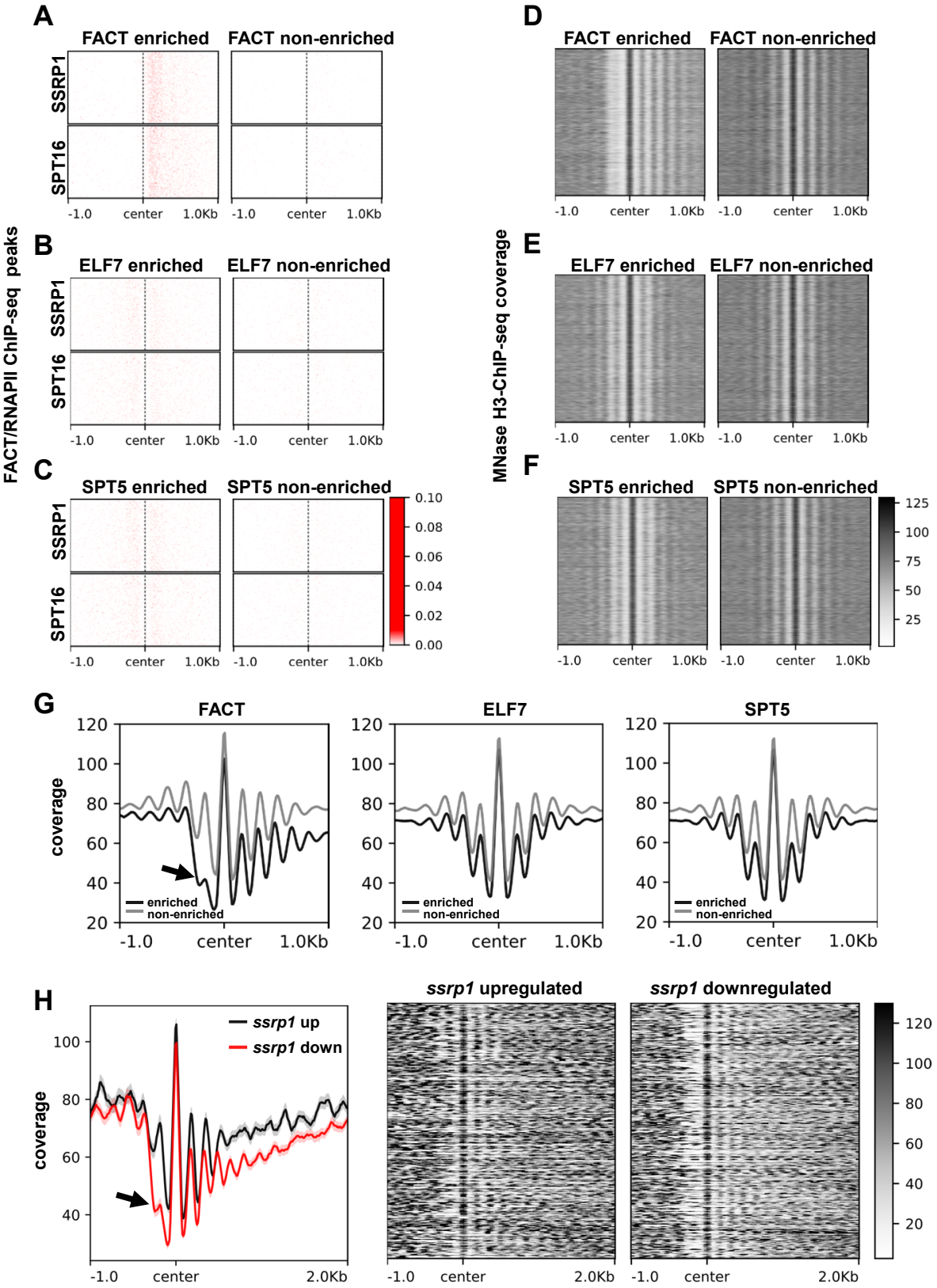
**Supplementary Figure S8.** Genes downregulated in *elf7* are longer and contain a higher number of introns than upregulated genes. **A**, Genes downregulated (n=290) in *elf7* relative to Col-0 upon HS are longer than upregulated (n=129) genes. Differentially expressed genes for *elf7* vs Col-0 upon HS were identified using Deseq2 (log-fold change cutoff of 2 and adjusted p-value < 0.05). The median gene length (TSS to TES) was determined for up- and downregulated genes and statistical significance was calculated using unpaired t-test resulting in a two-tailed P value of 0.0001. **B**, The median exon length in up and downregulated genes was determined and statistical significance was calculated using an unpaired t-test resulting in a two-tailed P value of 0.0150. **C**, The number of exons in downregulated genes is greater than in upregulated genes. The plot shows the mean number of exons per gene and the standard error of the mean. **D**, Upon HS, in *elf7* transcript reads accumulate upstream of peaks of ELF7 enrichment relative to S2P. LFC of transcripts in *elf7* relative to Col-0 (with/without HS), plotted over the center of ELF7/S2P peaks in non-overlapping protein-coding genes (n=2687; top). ChIP-seq coverage of S2P and S5P over the same areas (bottom). **E**, Motif search in genes downregulated in *elf7* (n=290) using the HOMER software revealed three significantly enriched motifs: recognition sites for the spliceosomal factors SRSF2 (1) and SNRNP (2), and an unknown *Arabidopsis* promoter (3).



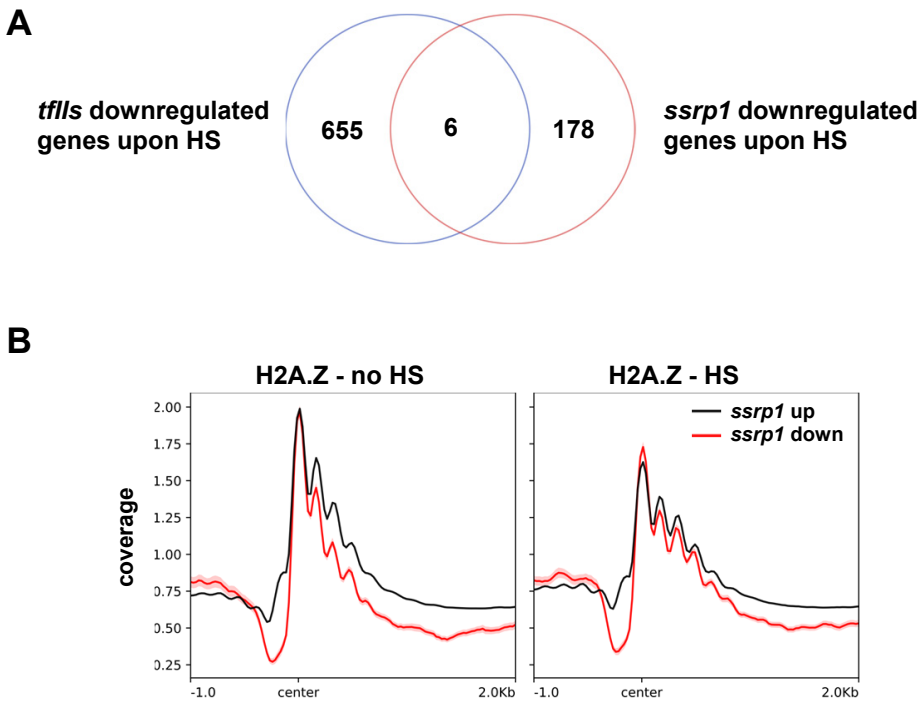

**Supplementary Figure S9.** FACT is markedly enriched at genes that are downregulated in *ssrp1*. **A**, S2P and S5P coverage centered over nucleosome positions (+1 to +5; n=26295 for +1 nucleosome) from MNase-seq data (cf. Fig 7A). **B**, SSRP1, SPT16 and RNAPII-S2P/S5P are enriched at genes downregulated in *ssrp1*. ChIP-seq coverage was plotted over genes upregulated (n= 306) or downregulated (n=367) in *ssrp1* relative to Col-0 after HS. **C**, Comparatively smaller fluctuation of ELF7 and RNAPII at genes differentially expressed in *elf7*. ChIP-seq coverage of ELF7 and RNAPII-S2P/S5P was plotted over genes upregulated (n= 750) or downregulated (n=981) in *elf7* relative to Col-0 after HS. The SEM is indicated as shaded area. **D**, Transcript levels of genes enriched or non-enriched in the indicated TEFs (according to ChIP-seq data) based on RNA-seq data, depicted as mean log TPM with p-values determined by unpaired t-test.




**Supplementary Figure S10.** FACT peaks and enrichment of RNA-seq reads at the position of the +1 nucleosome. **A**, FACT peaks (SSRP1 and SPT16, relative to RNAPII) over nucleosomal positions (+1 to +5) defined by MNase H3-ChIP-seq of all protein-coding genes (cf. Fig. 7). **B**, Relative enrichment of transcript reads *in ssrp1* relative to Col-0 plotted over nucleosomal positions (+1 to +5) of all protein coding genes defined by MNase H3-ChIP-seq. Replicates were merged and the log-fold change (LFC) between *ssrp1* and Col-0 was calculated with TPM normalisation.




**Supplementary Figure S11.** FACT enriched genes are charcterised by a remarkable NDR upstream of TSSs. Unsupervised k-means clustering was used to identify gene clusters enriched or non-enriched in FACT, ELF7 or SPT5 relative to RNAPII-S2P and data are centred to the position of the +1 nucleosome (A-G). **A**, Distribution of SSRP1 and SPT16 ChIP-seq peaks over FACT enriched (n=10360) and non-enriched (n=13901) genes. This data is almost identical to that shown in Fig. 7C, and is shown here for comparison with the ELF7/SPT5 datasets. **B**, Distribution of SSRP1 and SPT16 ChIP-seq peaks over ELF7 enriched (n=12710) and non-enriched (n=11551) genes. **C**, Distribution of SSRP1 and SPT16 ChIP-seq peaks over SPT5 enriched (n=10713)) and non-enriched (n=13548) genes. **D-F**, MNase H3 ChIP-seq coverage plotted over the same gene clusters analysed in A-C. **G**, Metagene profiles of MNase H3 ChIP-seq coverage over the gene clusters analysed in A-C with an arrow indicating the NDR apparent with FACT enriched genes. **H**, MNase H3 ChIP-seq tracks were plotted over +1 mononucleosomal positions of genes upregulated (n=306) or downregulated (n=367) in *ssrp1* relative to Col-0.




**Supplementary Figure S12.** Comparison of genes downregulated upon HS in *ssrp1* and *tflls*. **A**, Number of genes downregulated after 10 min HS at 44°C (padj < 0.05) in *tflls* (Obermeyer et al., 2023) and *ssrp1* (this work). **B**, H2A.Z ChIP-Seq coverage (Cortijo et al., 2017) without and with 37°C HS centred to the +1 nucleosomal position of *ssrp1* up- and downregulated genes. Lines represent the mean coverage and the standard error of the mean is indicated as a shaded area.

**Supplementary Table S1.** Sequencing statistics of ChIP-seq analysis

| sample_name | total reads | unique reads | duplicate reads | remaining after QC |
|---|---|---|---|---|
| **RNAPII_INPUT** | 19142506 | 18501842 | 640664 | 19080686 |
| **S2P_rep1** | 15644840 | 14653870 | 990970 | 12783373 |
| **S5P_rep1** | 15480110 | 14572446 | 907664 | 13221695 |
| **S2P_rep2** | 18310469 | 16947896 | 1362573 | 15646322 |
| **S5P_rep2** | 17327828 | 16533110 | 794718 | 15887063 |
| **SPT5_INPUT** | 20872903 | 15513775 | 5359128 | 20767919 |
| **SPT5_rep1** | 20286505 | 16109828 | 4176677 | 19575088 |
| **SPT5_rep2** | 23915710 | 18641063 | 5274647 | 23327232 |
| **SPT5_rep3** | 16718785 | 13036488 | 3682297 | 16179688 |
| **SPT5_rep4** | 18795229 | 16110548 | 2684681 | 12433360 |
| **ELF7_INPUT** | 18125511 | 12218169 | 5907342 | 18082698 |
| **ELF7_rep1** | 17457860 | 13271762 | 4186098 | 16930063 |
| **ELF7_rep2** | 17076862 | 13736967 | 3339895 | 16277265 |
| **ELF7_rep3** | 15536223 | 12666837 | 2869386 | 14818365 |
| **ELF7_rep4** | 17848064 | 15145309 | 2702755 | 12998707 |
| **SSRP1_INPUT** | 19690725 | 13340212 | 6350513 | 19642872 |
| **SSRP1_rep1** | 12988748 | 10932962 | 2055786 | 12393081 |
| **SSRP1_rep2** | 14936663 | 12425218 | 2511445 | 14449437 |
| **SSRP1_rep3** | 13840059 | 11719581 | 2120478 | 13127933 |
| **SSRP1_rep4** | 15925350 | 13397869 | 2527481 | 14981822 |
| **SPT16_INPUT** | 20686963 | 13426258 | 7260705 | 20634971 |
| **SPT16_rep1** | 17659216 | 14473482 | 3185734 | 16954657 |
| **SPT16_rep2** | 14830451 | 12474221 | 2356230 | 14054897 |
| **SPT16_rep3** | 18586399 | 15304790 | 3281609 | 17728575 |
| **SPT16_rep4** | 16539079 | 13279976 | 3259103 | 16045037 |

**Supplementary Table S2.** Sequencing statistics of RNA-seq analysis

| sample_name | total reads | uniquely mapped | multimapped | remaining after QC |
|---|---|---|---|---|
| **Col0_10min-HS_rep1** | 8350443 | 7163100 | 1114696 | 3534298 |
| **Col0_10min-HS_rep2** | 7769978 | 6618164 | 1010969 | 3107837 |
| **Col0_10min-HS_rep3** | 8160252 | 7021234 | 978289 | 3551508 |
| **Col0_no-HS_rep1** | 8305868 | 5145756 | 1271404 | 3161915 |
| **Col0_no-HS_rep2** | 8015992 | 6239835 | 1635909 | 2404824 |
| **Col0_no-HS_rep3** | 7766473 | 6139680 | 1466453 | 2247738 |
| **elf7-3_no-HS_rep1** | 7886048 | 6016090 | 1578012 | 1971501 |
| **elf7-3_no-HS_rep2** | 8313749 | 6771157 | 1400874 | 2173056 |
| **elf7-3_no-HS_rep3** | 8304892 | 4411155 | 1390833 | 2884158 |
| **elf7-3-10min-HS_rep1** | 8202507 | 6896661 | 1227444 | 3111895 |
| **elf7-3-10min-HS_rep2** | 8614288 | 7329671 | 1179023 | 3425400 |
| **elf7-3-10min-HS_rep3** | 8195811 | 6775868 | 1223958 | 3617307 |
| **spt4-R3_10min-HS_rep1** | 8411616 | 6924934 | 976045 | 3817226 |
| **spt4-R3_10min-HS_rep2** | 7696034 | 6485708 | 1065880 | 3172680 |
| **spt4-R3_10min-HS_rep3** | 8702015 | 7338323 | 1235536 | 3478978 |
| **spt4-R3_no-HS_rep1** | 8121972 | 6050334 | 1554804 | 2136885 |
| **spt4-R3_no-HS_rep2** | 6882077 | 4977204 | 1494305 | 1975339 |
| **spt4-R3_no-HS_rep3** | 7237295 | 5328737 | 1520498 | 1939883 |
| **ssrp1-2_10min-HS_rep1** | 8668808 | 7193684 | 1275091 | 3859923 |
| **ssrp1-2_10min-HS_rep2** | 8599635 | 6994247 | 1294709 | 3734123 |
| **ssrp1-2_10min-HS_rep3** | 8311354 | 6706343 | 1425706 | 3252333 |
| **ssrp1-2_no-HS_rep1** | 7057409 | 5017764 | 1272026 | 1836949 |
| **ssrp1-2_no-HS_rep2** | 11879224 | 8595952 | 2115469 | 2703888 |
| **ssrp1-2_no-HS_rep3** | 7935765 | 6079895 | 1666628 | 2027584 |

# Chapter 8

# General Discussion

## 8.1 Comparative co-localisation of transcript elongation factors

Chromatin immunoprecipitation followed by deep sequencing has been shown to be a powerful tool to map the distribution of transcript elongation factors in a genome wide manner. Comprehensive studies in model organisms are still rare, though. In 2021 such a study was published for budding yeast (Rossi et al., 2021). In Arabidopsis, such comprehensive data were not available. This thesis is a first step to compare the genome wide distribution of several transcript elongation factors in *A.thaliana* as a model organism for higher eukaryotes. To this end, robust data could be obtained for the factors ELF1, ELF7, SSRP1, SPT16 and SPT5. While this is far from the amount of factors mapped in yeast, obtaining high quality antibodies is still a rate limiting step in plants. Whenever possible, antibodies directly raised against epitopes of the respective transcript elongation factors were used rather than tagging of proteins. Therefore the risk of altering the function or distribution of the factors due to the addition of tags is reduced. The data shows specific differences for three types of transcript elongation factors. The histone chaperone FACT shows an accumulation close to transcriptional start sites and, in line with its role in remodelling histones, is enriched over nucleosomal positions. This is also highly reflective of the distribution of FACT in yeast (Rossi et al., 2021).
ELF7, as part of the PAF1 complex, is almost exclusively present in the gene body and is very highly correlated with the positioning of S5 phosphorylated RNAPII, again resembling the positioning in yeast. In combination with the observed changes in transcription this shows the role of the PAF1 complex to stabilise the transcript elongation complex after the initial phase.
Mapping of SPT5 to the genome showed a very high correlation with the distribution of S2 phosphorylated RNAPII, indicating a very strong connection with elongating RNAPII until transcription termination. Again, this is more reflective of the situation in yeast, as a notable accumulation of SPT5 as well as RNAPII is absent shortly after transcriptional start sites.
Finally, it was possible to identify ELF1 as a component of the *Arabidopsis* transcript elongation complex and its distribution on the genome could be mapped in high resolution. While its overall distribution is highly reflective of the distribution of S2P, it is notably absent over a subset of genes linked to several stress responses. If and how ELF1 could have an additional function for the transcription of stress response genes remains to be further studied.

A drawback of ChIP-sequencing is the limited resolution, which depends on the fragment size of the DNA during the ChIP protocol, especially comparing the average length of transcripts in *A.thaliana* (∼2.2 kb/transcript) with that of other higher eukaryotes (∼66.6 kb/transcript in *H.sapiens*). One way to increase the resolution to nucleotide level could be the use of native elongating transcript sequencing (NET-Seq). While this method is not well established in plants, it could provide a highly detailed map for the distribution of the RNAPII depending on different phosphorylation states within its C-terminal domain, allowing even more detailed insights into the interplay between the polymerase and additional factors contributing to the normal function of transcript elongation.

## 8.2 TFIIS is crucial for transcriptional reprogramming to enable thermotolerance

TFIIS is an RNAPII associated factor which has been shown to stimulate transcript cleavage activity when the polymerase is stalled at certain transcription barriers. In plants, a previous study has shown how expression of a mutated version of TFIIS leads to promoter proximal stalling of RNAPII. This mutation inhibits the cleavage activity of RNAPII (Antosz et al., 2020). In this study, it was possible to show that the *Arabidopsis tfIIs-1* mutant line exhibits a comparable molecular behaviour upon heat stress. While the lack of TFIIS does not have a strong phenotypic or molecular defect compared to wild type plants at perfect growth conditions, after acute stress TFIIS is necessary for proper transcription of genes needed under these conditions. RNAPII in *tfIIs-1* mutants could not transcribe through heat shock induced genes and this defect was linked to an increased accumulation of the polymerase and transcripts at the +1 nucleosomal position. Finally, it was also possible to connect the stalling of the polymerase to a depletion of H2A.Z at the +1 nucleosome. The exact consequences of including or removing H2A.Z at this site to regulate transcription is still under debate (Probst et al., 2020). However, the conclusion that this barrier is a potential stalling point for the RNAPII, requiring TFIIS to overcome the barrier, can be drawn. In conjunction with the comparable molecular phenotype described by Antosz et al., 2020, this might offer an intriguing model how TFIIS is required to cleave RNA in order to enable backtracked RNAPII to continue transcribing over nucleosomal barriers.

## 8.3 PAF1C enables proper transcript elongation

PAF1C is a protein complex that regulates transcript elongation and modifies the chromatin structure of transcribed regions by traveling with RNAPII. In plants, the role of PAF1C in developmental processes has been studied, and mutant phenotypes have been linked to misexpressed target genes, such as the expression of the *FLOWERING LOCUS C* (Yu and Michaels, 2010). In yeast and mammals, PAF1C subunit mutants exhibit various defects in the

tolerance to environmental stress conditions and inactivation of PAF1 and CTR9 resulted in severe mutant phenotypes (Francette et al., 2021). In chapter four of this thesis it is shown that *Arabidopsis* mutant lines deficient in PAF1C subunits ELF7 and CDC73 show differential gene expression, and several hundreds of differentially expressed genes were observed in these mutant lines after stress induction. PAF1C subunit mutants were exposed to elevated concentrations of NaCl to provoke a transcriptional response, and ELF7 and ELF8 deficient mutant lines displayed a more severe decrease in growth of the aerial parts than Col-0. Root growth of ELF7 and ELF8 deficient lines was also considerably more strongly reduced upon NaCl exposure. Transcriptomic analysis of *elf7-3* and *cdc73-2* in comparison to Col-0 after 3h of NaCl exposure showed a decreased response to NaCl in *elf7-3*, indicating that ELF7 plays an important role in the response to elevated NaCl concentrations.
In chapter seven the global distribution of ELF7 within the genome was mapped using antibodies raised against this PAF1C subunit. As expected, it was possible to show that ELF7 is highly associated with RNAPII. However, the association of ELF7 is significantly higher with S5 phosphorylated RNAPII compared to S2 phosphorylated RNAPII. By comparative transcriptional analysis it was shown that transcript elongation in the *elf7-3* mutant line exhibits defects located at intron-exon borders. A conclusion about the exact mechanism how the PAF1 complex is required for transcription over these barriers would be speculative. However, a direct or indirect role of the PAF1 complex is supported by the fact that *elf7-3* downregulated genes have more intron-exon borders and are enriched in binding sites for spliceosomal factors.

## 8.4 FACT modulates transcriptional start sites

Chapter five describes in detail how phosphorylation of the SPT16 subunit of FACT impacts transcriptional start sites of RNAPII transcribed genes. Phosphorylation of SPT16 is required for efficient binding of histones. In agreement with this, a non-phosphorylatable mutant line showed changes in the occupancy of histone H3 directly upstream of RNAPII transcriptional start sites. This region, often termed NDR (nucleosome depleted region) usually has to be depleted of histones to enable binding of transcription initiation components. A comparable retention of histones in this region after altering SPT16 function has been described in yeast (Jeronimo et al., 2019). Although the binding of the FACT complex is downstream of this region, the observed pattern is highly reflective of the function of FACT in yeast (Vinayachandran et al., 2018). In this thesis it was shown that FACT is partially recruited to nucleosomes and is required in particular for very highly transcribed genes. It was also possible to show that expression of a non-phosphorylatable version of SPT16 results in a lack of depletion of the NDR. If this is a direct effect remains speculative. A possible explanation could be a defect in the recruitment of CK2, as discussed in detail by Michl-Holzinger, 2021. To draw further conclusions, differential MNAse-sequencing would be a highly promising method in order to connect the function of FACT to nucleosome stability. This method, which uses different time points or protein concentrations to stop digestion of chromatin with micrococcal nuclease

(Wernig-Zorc et al., 2022) could deliver in-detail information in changes in nucleosome stability using our established FACT mutant lines. Preliminary results, described by Stoeckl, 2021, yielded promising results but are not sufficient to draw further conclusions.

## 8.5 ELF1 is a transcript elongation factor associated with a subset of RNAPII transcribed genes

ELF1 comprises a basic N-terminus, a defined Zn-Ribbon, and an acidic C-terminus only present in higher eukaryotes (Prather et al., 2005). In chapter six the first study characterising the function of ELF1 in *A.thaliana* is described. It was possible to show that ELF1 co-purifies with components of the transcript elongation complex and genetic interaction with genes encoding other transcript elongation factors was proven. Purified domains of ELF1 interacted with histones, specifically via its C-terminal acidic region. Finally, it was possible to prove nuclear localisation and, on a genome wide scale, co-localisation with S2 phosphorylated RNAPII at transcribed regions. The distribution of ELF1 in *Arabidopsis* is similar as reported in yeast (Rossi et al., 2021). Interestingly, compared to the distribution of other TEFs, ELF1 is overly enriched at a subset of genes. As these genes are almost exclusively linked to stress response reactions, this might indicate a more selective role of ELF1 for certain targets or certain conditions. In agreement with this, there were no significant changes in growth or notable phenotypic changes under standard conditions. Therefore it is possible that ELF1 acts in conjunction with other parts of the transcript elongation complex to regulate proper transcription, especially under certain challenging conditions.

# Chapter 9

# Summary

Transcript elongation in *Arabidopsis thaliana* is regulated by a variety of factors interacting with RNA polymerase II (RNAPII) during transcription. Recruitment or activation of those factors at barriers during transcription enables the polymerase to properly transcribe through genes in the context of chromatin *in vivo*. In this thesis the genome wide distribution of members of the transcript elongation complex in *Arabidopsis thaliana* was elucidated. Using chromatin immunoprecipitation in conjunction with deep sequencing, the distribution of ELF1, subunits of the FACT complex (SSRP1, SPT16), a subunit of the SPT4-SPT5 dimer (SPT5), a subunit of the PAF1 complex (ELF7), RNAPII S2P and RNAPII S5P was mapped in wild type plants. In addition, functional implications on gene expression were investigated using respective mutant lines. Finally, mutant lines were exposed to different stress conditions to change the expression of a defined subset of genes. Specific phenotypes and/or transcriptional defects could be detected for lines lacking TFIIS, ELF1, subunits of the PAF1 complex, SPT4-SPT5 and subunits of FACT. In case of TFIIS, we could connect the defect to a point mutation in the acidic loop of the C- terminal domain, previously described by Antosz et al., 2020. For FACT we could show that expression of a non-phosphorylatable variant of SPT16 changes the nucleosome occupancy at transcriptional start sites. Finally, we could determine ELF1 as a newly described transcript elongation factor in *Arabidopsis thaliana* with histone binding capability and possible implications in the correct regulation of expression of a subset of genes.

# Chapter 10

# Acknowledgements

I want to thank all people whose continuous help, support, knowledge and motivation made this thesis possible:

Klaus, Marion, Hanna, Henna, Amelie, Valentin, Philipp, Silvia, Wojciech, Sophie, Claudia, Lukas, Schneki, Rich, Matthias, Ina, Schmitzi, Kevin, Uwe, Gernot, Sara, Thomas, Christoph, Michael, Margit, Johanna, Thorsten, Mathe, Ralf, Anna.

# Chapter 11

# Bibliography

Antosz W, Deforges J, Begcy K, Bruckmann A, Poirier Y, Dresselhaus T, Grasser KD. Critical Role of Transcript Cleavage in Arabidopsis RNA Polymerase II Transcriptional Elongation. Plant Cell. 2020 May;32(5):1449-1463. doi: 10.1105/tpc.19.00891. Epub 2020 Mar 9. PMID: 32152189; PMCID: PMC7203918.

Blombach F, Fouqueau T, Matelska D, Smollett K, Werner F. Promoter-proximal elongation regulates transcription in archaea. Nat Commun. 2021 Sep 17;12(1):5524. doi: 10.1038/s41467-021-25669-2. PMID: 34535658; PMCID: PMC8448881.

Bushnell DA, Kornberg RD. Complete, 12-subunit RNA polymerase II at 4.1-A resolution: implications for the initiation of transcription. Proc Natl Acad Sci U S A. 2003 Jun 10;100(12):6969-73. doi: 10.1073/pnas.1130601100. Epub 2003 May 13. PMID: 12746498; PMCID: PMC165814.

Chen, F.X., Smith, E.R. and Shilatifard, A. Born to run: control of transcription elongation by RNA polymerase II. Nat Rev Mol Cell Biol 19, 464–478 (2018). https://doi.org/10.1038/s41580-018-0010-5

Cheng CY, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD. Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. Plant J. 2017 Feb;89(4):789-804. doi: 10.1111/tpj.13415. Epub 2017 Feb 10. PMID: 27862469.

Cortazar MA, Sheridan RM, Erickson B, Fong N, Glover-Cutter K, Brannan K, Bentley DL. Control of RNA Pol II Speed by PNUTS-PP1 and Spt5 Dephosphorylation Facilitates Termination by a "Sitting Duck Torpedo" Mechanism. Mol Cell. 2019 Dec 19;76(6):896-908.e4. doi: 10.1016/j.molcel.2019.09.031. Epub 2019 Oct 30. PMID: 31677974; PMCID: PMC6927536.

Cortijo S, Charoensawan V, Brestovitsky A, Buning R, Ravarani C, Rhodes D, van Noort J, Jaeger KE, Wigge PA. Transcriptional Regulation of the Ambient Temperature Response by H2A.Z Nucleosomes and HSF1 Transcription Factors in Arabidopsis. Mol Plant. 2017 Oct 9;10(10):1258-1273. doi: 10.1016/j .molp.2017.08.014. Epub 2017 Sep 8. PMID: 28893714; PMCID: PMC6175055.

Derelle E, Ferraz C, Rombauts S, Rouz´e P, Worden AZ, Robbens S, Partensky F, Degroeve S, Echeyni´e S, Cooke R, Saeys Y, Wuyts J, Jabbari K, Bowler C, Panaud O, Pi´egu B, Ball SG, Ral JP, Bouget FY, Piganeau G, De Baets B, Picard A, Delseny M, Demaille J, Van de Peer Y, Moreau H. Genome analysis of the smallest free-living eukaryote Ostreococcus tauri unveils many unique features. Proc Natl Acad Sci U S A. 2006 Aug 1;103(31):11647-52. doi: 10.1073/pnas.0604795103 . Epub 2006 Jul 25. PMID: 16868079; PMCID: PMC15 44224.

Dieci G, Fiorino G, Castelnuovo M, Teichmann M, Pagano A. The expanding RNA polymerase III transcriptome. Trends Genet. 2007 Dec;23(12):614-22. doi: 10.1016/j.tig.2007.09.001. Epub 2007 Oct 30. PMID: 17977614.

Dürr J, Lolas IB, Sørensen BB, Schubert V, Houben A, Melzer M, Deutzmann R, Grasser M, Grasser KD. The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in Arabidopsis. Nucleic Acids Res. 2014 Apr;42(7):4332-47. doi: 10.1093/nar/gku096. Epub 2014 Feb 4. PMID: 24497194; PMCID: PMC3985667.

Eaton JD, West S. Termination of Transcription by RNA Polymerase II: BOOM!. Trends in Genetics, Volume 36, Issue 9, 2020, Pages 664-675, ISSN 0168-9525, https://doi.org/10.1016/j.tig.2020.05.008.

Ehara H, Kujirai T, Fujino Y, Shirouzu M, Kurumizaka H, Sekine SI. Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. Science. 2019 Feb 15;363(6428):744-747. doi: 10.1126/science.aav8912. Epub 2019 Feb 7. PMID: 30733384.

Ehara H, Yokoyama T, Shigematsu H, Yokoyama S, Shirouzu M, Sekine SI. Structure of the complete elongation complex of RNA polymerase II with basal factors. Science. 2017 Sep 1;357(6354):921-924. doi: 10.1126/science.aan8552. Epub 2017 Aug 3. PMID: 28775211.

Farnung L, Vos SM. Assembly of RNA polymerase II transcription initiation complexes. Curr Opin Struct Biol. 2022 Apr;73:102335. doi: 10.1016/j.sbi.2022.102335. Epub 2022 Feb 17. PMID: 35183822; PMCID: PMC9339144.

Francette AM, Tripplehorn SA, and Arndt KM (2021). The Paf1 complex: A keystone of nuclear regulation operating at the interface of transcription and chromatin. J. Mol. Biol. 433, 166979. doi: 10.1016/j.jmb.2021.166979

Guo M, Liu JH, Ma X, Luo DX, Gong ZH, Lu MH. The Plant Heat Stress Transcription Factors (HSFs): Structure, Regulation, and Function in Response to Abiotic Stresses. Front Plant Sci. 2016 Feb 9;7:114. doi: 10.3389/fpls.2016.00114. PMID: 26904076; PMCID: PMC4746267.

Hsin JP, Manley JL. The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. 2012 Oct 1;26(19):2119-37. doi: 10.1101/gad.200303.112. PMID: 23028141; PMCID: PMC3465734.

Jeronimo C, Poitras C, Robert F. Histone Recycling by FACT and Spt6 during Transcription Prevents the Scrambling of Histone Modifications. Cell Rep. 2019 Jul 30;28(5):1206-1218.e8. doi: 10.1016/j .celrep.2019.06.097. PMID: 31365865.

Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. Plant J. 2007 Apr;50(2):347-63. doi: 10.1111/j.1365-313X.2007.03052.x. Epub 2007 Mar 21. PMID: 17376166.

Kumar SV, Wigge PA. H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. Cell. 2010 Jan 8;140(1):136-47. doi: 10.1016/j.cell.2009.11.006. PMID: 20079334.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997 Sep 18;389(6648):251-60. doi: 10.1038/38444. PMID: 9305837.

Markusch, H., Michl-Holzinger, P., Obermeyer, S., Thorbecke, C., Bruckmann, A., Babl, S., Längst, G., Osakabe, A., Berger, F. and Grasser, K.D. (2023), Elongation factor 1 is a component of the Arabidopsis RNA polymerase II elongation complex and associates with a subset of transcribed genes. New Phytol. New Phytol. 2023 Apr;238(1):113-124. doi: 10.1111/nph.18724. Epub 2023 Feb 1. PMID: 36627730.

Michl-Holzinger P. The Arabidopsis RNA Polymerase II Transcript Elongation Complex and the role of post-translational modifications on the histone chaperone FACT. 2021

Michl-Holzinger P, Obermeyer S, Markusch H, Pfab A, Ettner A, Bruckmann A, Babl S, Längst G, Schwartz U, Tvardovskiy A, Jensen ON, Osakabe A, Berger F, Grasser KD. Phosphorylation of the FACT histone chaperone subunit SPT16 affects chromatin at RNA polymerase II transcriptional start sites in Arabidopsis. Nucleic Acids Res. 2022 May 20;50(9):5014-5028. doi: 10.1093/nar/gkac293. PMID: 35489065; PMCID: PMC9122599.

Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, Vollger MR, Altemose N, Uralsky L, Gershman A, Aganezov S, Hoyt SJ, Diekhans M, Logsdon GA, Alonge M, Antonarakis SE, Borchers M, Bouffard GG, Brooks SY,... Phillippy AM. The complete sequence of a human genome. Science. 2022 Apr;376(6588):44-53. doi: 10.1126/science.abj6987. Epub 2022 Mar 31. PMID: 35357919; PMCID: PMC9186530.

Obermeyer, S., Kapoor, H., Markusch, H. and Grasser, K.D. (2023), Transcript elongation by RNA polymerase II in plants: factors, regulation and impact on gene expression. Plant J. https://doi.org/10.1111/tpj.16115

Obermeyer S, Stöckl R, Schnekenburger T, Kapoor H, Stempfl T, Schwartz U, Grasser KD. TFIIS is crucial during early transcript elongation for transcriptional reprogramming in response to heat stress. J Mol Biol. 2023 Jan 30;435(2):167917.
doi: 10.1016/j.jmb.2022.167917 . Epub 2022 Dec 9. PMID: 36502880.

Obermeyer S, Stöckl R, Schnekenburger T, Moehle C, Schwartz U, Grasser KD. Distinct role of subunits of the Arabidopsis RNA polymerase II elongation factor PAF1C in transcriptional reprogramming. Front Plant Sci. 2022 Sep 29;13:974625.
doi: 10.3389/fpls.2022.974625. PMID: 36247629; PMCID: PMC9558118.

Orphanides G, Reinberg D. RNA polymerase II elongation through chromatin. Nature. 2000 Sep 28;407(6803):471-5. doi: 10.1038/35035000 . PMID: 11028991.

Piovesan A, Caracausi M, Antonaros F, Pelleri MC, Vitale L. GeneBase 1.1: a tool to summarize data from NCBI gene datasets and its application to an update of human gene statistics. Database (Oxford). 2016 Dec 26;2016:baw153. doi: 10.1093/database/baw153. PMID: 28025344; PMCID: PMC5199132.

Prather D, Krogan NJ, Emili A, Greenblatt JF, Winston F. Identification and characterization of Elf1, a conserved transcription elongation factor in Saccharomyces cerevisiae. Mol Cell Biol. 2005 Nov;25(22):10122-35. doi: 10.1128/MCB.25.22.10122-10135.2005. PMID: 16260625; PMCID: PMC1280281.

Probst AV, Desvoyes B, Gutierrez C. Similar yet critically different: the distribution, dynamics and function of histone variants. J Exp Bot. 2020 Aug 17;71(17):5191-5204. doi: 10.1093/jxb/eraa230. PMID: 32392582.

Rossi MJ, Kuntala PK, Lai WKM, Yamada N, Badjatia N, Mittal C, Kuzu G, Bocklund K, Farrell NP, Blanda TR, Mairose JD, Basting AV, Mistretta KS, Rocco DJ, Perkinson ES, Kellogg GD, Mahony S, Pugh BF. A high-resolution protein architecture of the budding yeast genome. Nature. 2021 Apr;592(7853):309-314. doi: 10.1038/s41586-021-03314-8. Epub 2021 Mar 10. PMID: 33692541; PMCID: PMC8035251.

Stoeckl R. Transcriptomic changes in PAF1c subunit mutant lines at salt stress conditions and mapping nucleosome stability in Arabidopsis thaliana. 2021.

Vinayachandran V, Reja R, Rossi MJ, Park B, Rieber L, Mittal C, Mahony S, Pugh BF. Widespread and precise reprogramming of yeast protein-genome interactions in response to heat shock. Genome Res. 2018 Feb 14;28(3):357–66. doi: 10.1101/gr.226761.117. Epub ahead of print. PMID: 29444801; PMCID: PMC5848614.

Werner F, Grohmann D. Evolution of multisubunit RNA polymerases in the three domains of life. Nat Rev Microbiol. 2011 Feb;9(2):85-98. doi: 10.1038/nrmicro2507. PMID: 21233849.

Wernig-Zorc S, Schwartz U, Längst G. Optimized MNase-seq pipeline reveals a novel gene regulation via differential nucleosome stability. 2022.
doi: https://doi.org/10.1101/2022.12.29.521985

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS One. 2007 Aug 8;2(8):e718. doi: 10.1371/journal.pone.0000718. PMID: 17684564; PMCID: PMC1934936.