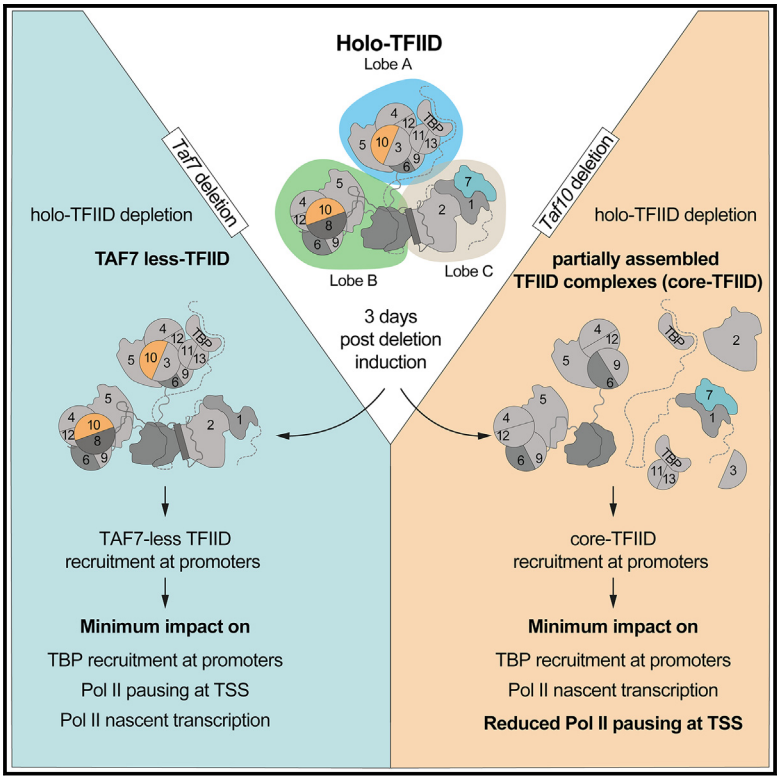
Cell Reports

Article

RNA polymerase II transcription initiation in holo-TFIID-depleted mouse embryonic stem cells

Graphical abstract



Highlights

- TAF7 or TAF10 depletion leads to the formation of different partial TFIID complexes
- Nascent Pol II transcription is not strongly affected in holo-TFIID-depleted cells
- TBP is recruited at the promoter in holo-TFIID-depleted cells
- Pol II pausing is affected in the absence of TAF10

Authors

Vincent Hisler, Paul Bardot, Dylane Detilleux, ..., Ferenc Müller, László Tora, Stéphane D. Vincent

Correspondence

vincent@igbmc.fr

In brief

Hisler et al. reveal that TAF7 and TAF10 depletion affects holo-TFIID sequential assembly, leading to the formation of partial TFIID complexes. Transcription is active and TBP is recruited at promoters when holo-TFIID is depleted. Partial TFIID complexes may sustain active transcription but cannot replace holo-TFIID over several cell divisions and/or development.

Hisler et al., 2024, Cell Reports *43*, 114791
October 22, 2024
https://doi.org/10.1016/j.celrep.2024.114791

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

RNA polymerase II transcription initiation in holo-TFIID-depleted mouse embryonic stem cells

Vincent Hisler,[1,2,3,4] Paul Bardot,[1,2,3,4,8] Dylane Detilleux,[1,2,3,4,8] Andrea Bernardini,[1,2,3,4] Matthieu Stierle,[1,2,3,4] Emmanuel Garcia Sanchez,[1,2,3,4] Claire Richard,[1,2,3,4] Lynda Hadj Arab,[1,2,3,4] Cynthia Ehrhard,[1,2,3,4] Bastien Morlet,[1,2,3,4,5] Yavor Hadzhiev,[7] Matthieu Jung,[1,2,3,4,6] Stéphanie Le Gras,[1,2,3,4,6] Luc Négroni,[1,2,3,4,5] Ferenc Müller,[7] László Tora,[1,2,3,4] and Stéphane D. Vincent[1,2,3,4,9,*]

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), 67400 Illkirch, France
[2]CNRS, UMR7104, 67400 Illkirch, France
[3]INSERM, U1258, 67400 Illkirch, France
[4]Université de Strasbourg, 67400 Illkirch, France
[5]Proteomics Platform (IGBMC), 67400 Illkirch, France
[6]GenomEast (IGBMC), 67400 Illkirch, France
[7]Institute of Cancer and Genomic Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham B15 2TT, UK
[8]These authors contributed equally
[9]Lead contact
*Correspondence: vincent@igbmc.fr
https://doi.org/10.1016/j.celrep.2024.114791

SUMMARY

The recognition of core promoter sequences by TFIID is the first step in RNA polymerase II (Pol II) transcription initiation. Metazoan holo-TFIID is a trilobular complex, composed of the TATA binding protein (TBP) and 13 TBP-associated factors (TAFs). Why and how TAFs are necessary for the formation of TFIID domains and how they contribute to transcription initiation remain unclear. Inducible TAF7 or TAF10 depletion, followed by comprehensive analysis of TFIID subcomplex formation, chromatin binding, and nascent transcription in mouse embryonic stem cells, result in the formation of a TAF7-lacking TFIID or a minimal core-TFIID complex, respectively. These partial complexes support TBP recruitment at promoters and nascent Pol II transcription at most genes early after depletion, but importantly, TAF10 is necessary for efficient Pol II pausing. We show that partially assembled TFIID complexes can sustain Pol II transcription initiation but cannot replace holo-TFIID over several cell divisions and/or development.

INTRODUCTION

RNA polymerase II (Pol II) is responsible for the transcription of all protein-coding genes and is regulated, allowing gene-by-gene variable expression levels depending on the cellular context (reviewed in Roeder[1]). Consequently, Pol II transcription dysfunction is associated with pathologies such as developmental defects, cancer, and metabolic or neural diseases. Pol II transcription is first regulated by the binding of specific transcription factors to enhancers, which recruit different classes of transcriptional co-activators, creating a favorable context for transcription initiation. Pol II recruitment via the formation of the pre-initiation complex (PIC) on active promoters is the obligatory step for transcription initiation.

The PIC is composed of six general transcription factors (GTFs), TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH, and Pol II (reviewed in Roeder[1]). TFIID is the first GTF to bind to the promoter, initiating the nucleation of the PIC. In metazoans, holo-TFIID is composed of the TATA binding protein (TBP) and 13 TBP-associated factors (TAFs)[2,3] and is recruited to the promoter by multiple mechanisms. TFIID interacts with histone marks associated with open chromatin via TAF1 and TAF3.[4,5] Moreover, TFIID interacts with DNA motifs such as the TATA box via TBP,[6] but also with downstream elements (Motif Ten Element [MTE] and Downstream Promoter Element [DPE])[7] which are crucial for the recognition of the large majority of mammalian promoters that lack a TATA box.

Holo-TFIID is composed of three lobes (Figure 1A).[3,8] Lobes A and B contain TAFs sharing a conserved histone fold domain (HFD) allowing specific heterodimerization. The HEAT domains of the two copies of TAF6, with TAF1 and TAF8, connect lobes A and B and, with TAF2, constitute lobe C.[3,8] Co-translational assembly of TFIID[9,10] occurs in multiple steps and leads to the buildup of subcomplexes. The core-TFIID (TAF5, TAF4/TAF12, and TAF6/TAF9 in two copies) is the basic structure of lobes A and B.[2,3,8] Addition of the 3-TAF submodule (TAF2, TAF8, and TAF10) to the core-TFIID results in the formation of the 8-TAF complex, which contains a full B lobe.[11] Then, the pre-assembled TFIID building blocks engage co-translationally with the nascent TAF1 polypeptide.[10] TAF7 and TAF11/TAF13 interact in the cytoplasm,[12] and it was proposed that the biochemically reconstituted TAF1/TAF7/TAF11/TAF13/TBP subcomplex (S-TAF) could integrate with the 8-TAF complex to form the holo-TFIID.[13]

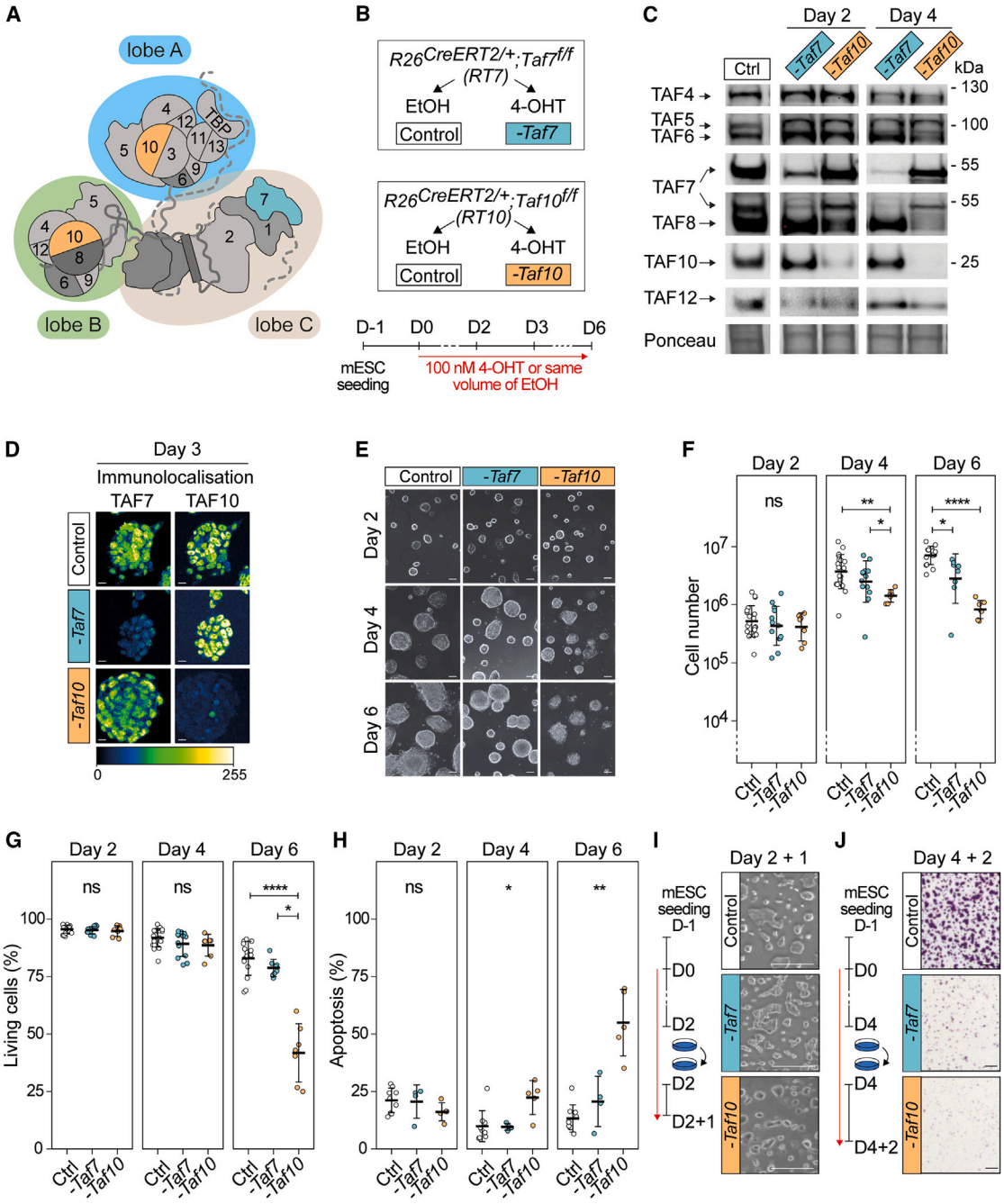


Figure 1. Phenotypic analysis of the conditional depletion of TAF7 or TAF10 in mESCs

(A) Trilobular structure of TFIID.

(B) Deletion of *Taf7* (−*Taf7*) in *R26*$^{CreERT2/+}$;*Taf7*$^{f/f}$ (*RT7*) and *Taf10* (−*Taf10*) in *R26*$^{CreERT2/+}$;*Taf10*$^{f/f}$ (*RT10*) mESCs. Control cells were treated with EtOH.

(C) Western blot analyses of TAF4, TAF5, TAF6, TAF7, TAF8, TAF10, and TAF12 protein expression after *Taf7* or *Taf10* deletion at day (D) 2 and D4. As a control (Ctrl), *RT7* cells were treated for 2 days. The Ponceau staining is displayed at the bottom.

(D) Immunolocalization of TAF7 and TAF10 in *RT7* and *RT10* cells at D3. As a control, *RT10* cells were treated for 3 days. Color scale: Green Fire Blue LUT.

(E) Colony growth at D2, D4, and D6.

(F–H) Total number of cells (F), percentage of living cells (G), and percentage of apoptotic cells (H) at D2, D4, and D6. (E) Ctrl, $n = 5$; −*Taf7*, $n = 2$; D4: −*Taf10*, $n = 3$ biological replicates for each day. (F and G) Ctrl: D2, $n = 20$; D4, $n = 20$; D6, $n = 15$. −*Taf7*: D2, $n = 13$; D4, $n = 13$; D6, $n = 8$. −*Taf10*: D2, $n = 7$; D4, $n = 7$; D6, $n = 7$

(legend continued on next page)

How the three lobes formed by the TAF subcomplexes are required for transcription initiation has not been investigated. The variable composition of TFIID and variable phenotypes suggest both essential and redundant roles for TAFs forming TFIID variant complexes (reviewed in Levine et al.[14]). First, non-canonical TFIID complexes have been described in yeast, *Drosophila*, and human cells.[15–18] Second, some *TAF*s, such as *TAF4* and *TAF9*, have paralogs[14] and are co-expressed with their relative paralog.[19–21] As TAF4B and TAF9B are part of TFIID[22,23] in two copies, this suggests the possible combination of these homologs within TFIID.

TFIID is important *in vivo*, as mutations in several *Taf* genes lead to peri-implantation lethality in the mouse.[24–27] Interestingly, *Taf2*$^{-/-}$, *Taf5*$^{-/-}$, *Taf8*$^{-/-}$, and *Taf10*$^{-/-}$ inner cell masses (ICMs) fail at outgrowth *in vitro*, while the trophectoderm is not affected,[24,25,27] suggesting that these TAFs are not required in all cell types. Moreover, while TAF7 is important for early thymocyte proliferation and differentiation, it is not required for thymocyte final differentiation.[26] Similarly, conditional deletion of *Taf10* in various cellular contexts has different effects depending on the developmental stage and the cell type.[23,28,29] Interestingly, TAF10 depletion severely affects TFIID assembly,[23,25,29] whereas TAF7 depletion does not.[26] Overall, TFIID composition and requirements are variable depending on the cellular context; however, as the data have been gathered from different systems, it is not possible to draw comparative conclusions about the function of partial and/or holo-TFIID complexes in cellular homeostasis.

In this study, we investigated the biochemical and genetic functions of *Taf7* and *Taf10*, coding for subunits of distinct TFIID domains (TAF7 lobe C and TAF10 lobes A and B) and integrating the complex at different steps (TAF7 later, TAF10 earlier) in comparable conditions, in pluripotent mouse embryonic stem cells (mESCs) and in the embryonic mesoderm. We monitored the consequences of TAF7 or TAF10 depletion on holo-TFIID assembly and chromatin distribution, as well as on Pol II transcription, to test whether partial TFIID complexes can sustain Pol II transcription. Holo-TFIID is drastically depleted upon *Taf7* or *Taf10* deletion. TAF7 depletion resulted in the assembly of a TAF7-less TFIID, whereas TAF10 depletion led to the formation of a core-TFIID complex. Notably, TBP was still present on most promoters upon either TAF7 or TAF10 depletion, coupled with mostly intact Pol II transcription, 3 days after induction of the deletion. However, a small number of genes were affected in both depletions, with more severity in TAF10-depleted cells. Consistently, conditional deletion in the mesoderm led to more severe phenotype upon *Taf10* than *Taf7* loss. Altogether, our data suggest that in the absence of holo-TFIID, partial TFIID modules broadly support nascent transcription with distinct requirements on a subset of genes in mESCs.

RESULTS

Depletion of distinct TFIID subunits causes different phenotypic severities in mESCs

To get insight into the molecular consequences of deletion of *Taf7* or *Taf10* during stem cell maintenance, mESCs were derived from blastocysts carrying the inducible ubiquitously expressed *R26*^{CreERT2} allele[30] associated with *Taf7*^{f} or *Taf10*^{f} (*R26*$^{CreERT2/+}$;*Taf7*$^{f/f}$ and *R26*$^{CreERT2/+}$;*Taf10*$^{f/f}$, hereafter called *RT7* and *RT10*, respectively). The deletion of *Taf7* (−*Taf7*) or *Taf10* (−*Taf10*) was induced by the addition of 4-hydroxytamoxifen (4-OHT) at day 0 (Figure 1B). 4-OHT treatment did not induce secondary defects (Figures S1A and S1B). Deletion efficiency was monitored by western blot (WB) analyses, and depletion of TAF7 protein in *RT7* mESCs was observed as early as day 2 and almost complete at day 4 (Figures 1C, S1C, and S1D). Similarly, TAF10 depletion in *RT10* mESCs was observed as early as day 2 and no longer detectable at day 4 (Figures 1C, S1C, and S1D). In these conditions, the expression of TAF4, TAF5, TAF6, and TAF12 TFIID subunits was not affected, and in agreement with our earlier studies, depletion of TAF10 resulted in the destabilization of TAF8[9,23] (Figures 1C and S1C). The depletion of TAF7 and TAF10 proteins is nearly homogeneous within the cellular population, as shown by immunolocalization (Figures 1D and S1E). Altogether, we have established an efficient cellular model to study the effects of the depletion of TAF7 or TAF10 proteins.

Next, we investigated the effect of the depletion on cell proliferation and viability. *RT7* and *RT10* mutant cells form smaller colonies compared to control cells (Figures 1E and S1F), indicating that both TAF7 and TAF10 are required for mESC maintenance, in agreement with the incapacity of *Taf7*$^{-/-}$ and *Taf10*$^{-/-}$ ICMs to grow *in vitro*.[25,26]

We then evaluated cell growth in control and mutant conditions. There were no detectable differences in cell numbers between the control either mutant cell line at day 2 (Figure 1F). Depletion of TAF7 resulted in a minor decrease in cell number only at day 6. In contrast, TAF10 depletion resulted in a significant reduction in cell number already at day 4 (Figure 1F). To assess whether this effect is due to a cell proliferation decrease, we quantified cells in S phase. We observed a reduction tendency in a percentage of S-phase cells in both mutants from day 3 (Figure S1G), suggesting a minor impact of TAF7 and TAF10 depletion on cell-cycle progression. As decrease in cell growth might be due to reduced viability, we investigated cell viability (Figure 1G). No differences between control and *RT7* or *RT10* mutant cells were observed at days 2 and 4. However, at day 6, while depletion of TAF7 did not have any significant effect, depletion of TAF10 reduced viability 2-fold. Analysis of apoptosis showed that the percentage of apoptotic cells ranged from 10% to 25% from day 2 to day 6 in control and mutant *RT7*

biological replicates. (H) Ctrl, $n = 9$; *Taf7*, $n = 4$; *Taf10*, $n = 5$ biological replicates for each day). Two independent experiments were conducted. The bars correspond to the mean ± SD. Kruskal-Wallis test followed by Dunn *post hoc* test: ns, not significant; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$.
(I) Cell density after passage at D2 and 1 day of extra culture.
(J) Cell density evaluated by crystal violet staining after passage at D4 and 2 days of culture. The control conditions correspond to *RT7* cells (E, I, and J). Scale bars, 15 μm (D), 50 μm (E), and 150 μm (I and J).

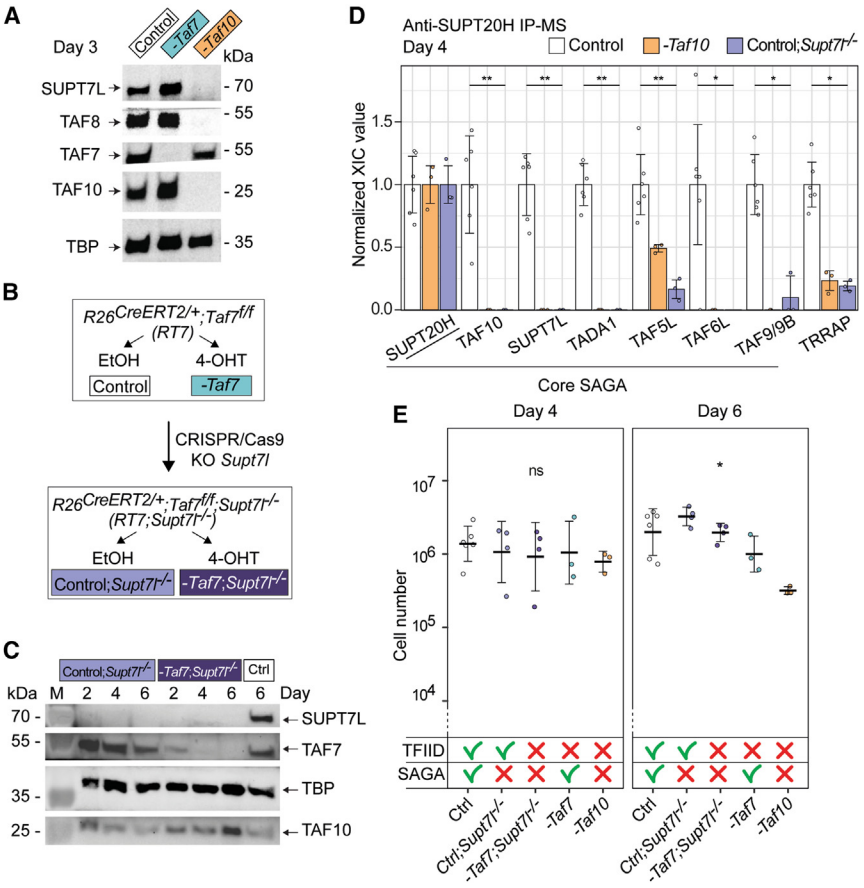


Figure 2. The more severe phenotype in TAF10-depleted mESCs is not due to SAGA assembly defect

(A) Western blot analyses of SUPT7L, TAF7, TAF8, TAF10, and TBP expression in TAF7- or TAF10-depleted cells after 3 days of treatment. Control *RT7* cells were treated with EtOH.

(B) Generation of the $R26^{CreERT2/+};Taf7^{f/f};Supt7l^{-/-}$ (*RT7;Supt7l*$^{-/-}$) mESCs.

(C) Western blot analyses of SUPT7L, TAF7, TBP, and TAF10 expression in *RT7;Supt7l*$^{-/-}$ cells after 2, 4, and 6 days of treatment with EtOH (Control;*Supt7l*$^{-/-}$) or 4-OHT (−*Taf7;Supt7l*$^{-/-}$). Control *RT7* cells were treated 6 days. M, molecular weight marker.

(D) Anti-SUPT20H IP-MS analyses on nuclear-enriched lysates from *RT10* and *RT7* cells treated 4 days with EtOH (Control, *RT7*, and *RT10* data merged), *RT10* mESCs with 4-OHT (−*Taf10*), and *RT7;Supt7l*$^{-/-}$ cells with EtOH (Control;*Supt7l*$^{-/-}$). For each protein, the XIC (extracted ion chromatograms) values of Control;*Supt7l*$^{-/-}$ and −*Taf10* lysates were normalized to those of *RT7* and *RT10* control cells treated with EtOH, respectively. Control, $n = 2$ biological replicates × 3 technical replicates; −*Taf10*, $n = 1 \times 3$; Control;*Supt7l*$^{-/-}$, $n = 1 \times 3$.

(E) Total number of cells at day (D) 4 and D6. *RT7* and *RT10* cells treated with EtOH were merged as control (Ctrl). The impact on TFIID and SAGA assembly is indicated at the bottom. D4 and D6: Ctrl, $n = 6$; Ctrl;*Supt7l*$^{-/-}$, $n = 4$; −*Taf7;Supt7l*$^{-/-}$, $n = 4$; −*Taf7*, $n = 3$; −*Taf10*, $n = 3$ biological replicates. Kruskal-Wallis test: ns, not significant; $^{*}p < 0.05$ and $^{**}p < 0.01$ (D and E). Means ± SD are shown (D and E).

cells (Figure 1H), indicating that TAF7 depletion does not induce apoptosis. In contrast, while no significant increase in apoptosis could be observed in TAF10-depleted cells at days 2 and 4, at day 6, the apoptosis rate increased 2-fold (Figure 1H). Finally, we evaluated the ability of the cells to form colonies after replating. When the mutant cells were split on day 2 and analyzed 1 day later (D2 + 1), no significant difference was observed in colony size or number of living cells for the two mutants (Figures 1I and S1H). In contrast, when the mutant cells were split at day 4 and analyzed 2 days after (D4 + 2), their capacity to form colonies was severely impaired (Figures 1J and S1I), indicating that both TAF7 and TAF10 are required for mESC maintenance from day 4 onward.

Altogether, these data indicate that TAF10 is required for the growth and survival of mESCs, whereas TAF7 depletion only weakly impacts mESC growth and survival. Importantly, this difference is not the consequence of the expression of TAF7 paralogs, as they are not expressed in controls nor in the mutant cells (Figures S1J–S1L).

The severity of the TAF10-depletion phenotype is not caused by defects in SAGA assembly

While TAF7 is a TFIID-specific subunit, TAF10 is also shared with the Supt-Ada-Gcn5 acetyltransferase (SAGA) co-activator complex.[31] SAGA assembly is defective in *Taf10* conditional mutant embryos.[23] Therefore, the difference in phenotype severity between mutant *RT7* and *RT10* mESCs could be associated with an impairment in both TFIID and SAGA function in *RT10* cells. To test this hypothesis, we focused on SUPT7L, the HFD partner of TAF10 in SAGA.[32] Importantly, *Supt7l* deletion in mESCs disrupts SAGA assembly.[33] WB analysis of SUPT7L expression in the *RT7* line indicated that TAF7 depletion does not affect SUPT7L levels as expected, but surprisingly, TAF10 depletion resulted in the loss of SUPT7L expression (Figure 2A), indicating that TAF10 is required for the stability of its two HFD partners, TAF8 and SUPT7L. These results indicate that loss of *Supt7l* or *Taf10* has similar effects on SAGA assembly. To deplete TAF7 in a context where SAGA assembly is impaired, we deleted *Supt7l* in the *RT7* line (hereafter called *RT7;Supt7l*$^{-/-}$) (Figure 2B). Co-depletion of SUPT7L and TAF7 in mutant *RT7;Supt7l*$^{-/-}$ cells was validated by WB (Figure 2C). Immunoprecipitation coupled with mass spectrometry (IP-MS) using antibodies against SUPT20H, a core SAGA subunit, confirmed that SAGA is not assembled in the absence of TAF10 and/or SUPT7L[23,33] (Figure 2D).

Supt7l loss of function results only in a slight growth defect in similar culture conditions,[33] and our data indicated that the growth and survival of 4-OHT-treated *RT7;Supt7l*$^{-/-}$ mESCs are not affected (Figures 2E, S2A, and S2B). Depletion of TAF7 in the *Supt7l*$^{-/-}$ background did not change significantly the cell

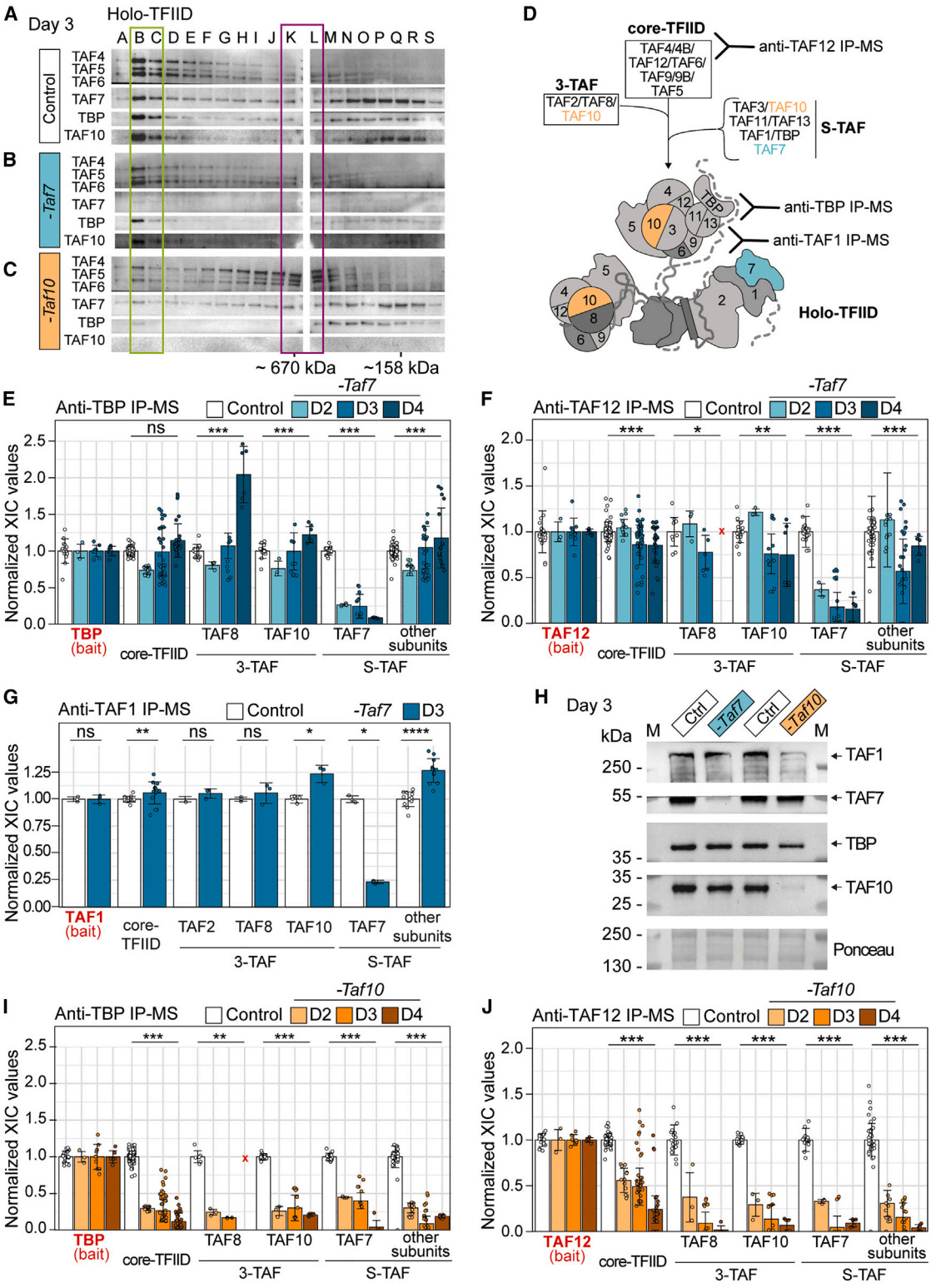


Figure 3. Depletion of TAF7 or TAF10 differentially affects TFIID assembly

(A–C) Gel filtration coupled to western blot analysis of TAF4, TAF5, TAF6, TAF7, TAF10, and TBP expression in *RT10* mESCs treated with EtOH (Control) (A) and TAF7- and TAF10-depleted mESCs treated with 4-OHT (B and C) for 3 days ($n = 1$). Letters on the top correspond to the fractions. Positions of the complexes are indicated by colored boxes.

(legend continued on next page)

number, cell viability, or cell death compared to TAF7 depletion alone (Figures 2E, S2A, and S2B). Therefore, SAGA disruption does not aggravate the phenotype of TAF7 loss, strongly suggesting that the difference in the observed severity between mutant *RT7* and mutant *RT10* mESCs is not caused by the disruption of SAGA in mutant *RT10* cells but by the different molecular consequences associated with the loss of TAF7 or TAF10 in TFIID.

Depletion of TAF7 or TAF10 leads to the formation of distinct partial TFIID complexes

To analyze the molecular changes occurring in TFIID assembly following depletion of TAF7 or TAF10, we first performed gel filtration on *RT7* or *RT10* mESC nuclear-enriched whole-cell extracts (NWCEs) collected at day 3, in control and mutant conditions. In control cells, fractions B and C contained the TFIID complex, as suggested by the molecular weight and the co-localization of TBP and all the tested TAFs (Figure 3A, green box). In mutant *RT7* extracts, the profile is very similar to that of the control lysate, except for the absence of TAF7, indicating that a TFIID complex without TAF7 (hereafter called TAF7-less TFIID) is present in these cells (Figure 3B, green box). In mutant *RT10* cells, however, very little high-molecular-weight TAF complex could be detected. Consequently, TAF4, TAF5, and TAF6 were relocated to lower-molecular-weight complexes, peaking at about 670 kDa fractions (Figure 3C, purple box), indicating that the holo-TFIID complex assembly was perturbed after TAF10 depletion.

To explore the composition of the partial complexes after TAF7 or TAF10 depletion, we performed IP-MS using anti-TAF12 and anti-TBP antibodies on NWCEs after 2, 3, and 4 days of treatment. As TAF12 is part of the core-TFIID and TBP is only part of the holo-TFIID complex, anti-TAF12 IP detects all the intermediate partial TFIID complexes assembling on core-TFIID, whereas anti-TBP IP should immunoprecipitate only holo-TFIID[10] (Figure 3D). As TBP is also part of the SL1 and TFIIIB complexes, which are involved in Pol I and Pol III transcription initiation, respectively,[34] we tested whether TAF7 or TAF10 depletion affects the interaction between TBP and SL1 subunits or the BRF1 subunit of TFIIIB (Figures S3A and S3B). Remarkably, as in the *TCre;Taf10* mutant embryos,[23] TAF10 depletion resulted in increased interaction between TBP and its non-TFIID partners, particularly BRF1 (Figure S3B). These data indicate that some TBP is relocated to these complexes when TAF10 is depleted, suggesting an impairment in TBP incorporation into TFIID. This relocation was not observed when TAF7 was depleted (Figure S3A), suggesting that TBP incorporation into TFIID is not affected under these conditions.

TAF7 depletion resulted in the loss of TAF7 in the anti-TBP or -TAF12 IP-ed complexes, without causing significant changes in the IP-ed subunits of TFIID (Figures 3E, 3F, S3C, and S3D). In both IP-MSs, the signal detected for each of the TFIID subunits remained relatively stable on each day of the treatment analyzed. In the anti-TBP IP at day 4, we observed an increase in the co-immunoprecipitation of TAF2 and TAF8, which do not interact directly with TAF7 but are localized in lobe C of holo-TFIID. This indicates that TAF7 depletion has a limited effect on TFIID assembly, resulting in a TAF7-less TFIID complex, as suggested by the gel-filtration analysis (Figures 3A and 3B) and confirmed by the anti-TAF1 IP-MS in TAF7-depleted cells (Figures 3G and S3G). In contrast, TAF10 depletion caused a severe impairment of TFIID assembly, as TAF1 levels are strongly decreased in TAF10-depleted cell lysates (Figure 3H). This is also supported by the changes in the TFIID subunits detected in anti-TBP (Figures 3I, 3E, and S3E) or anti-TAF12 IPs (Figures 3J and S3F). First, the interaction between TBP and all the TAFs strongly decreased (about 4-fold already at day 2) in mutant *RT10* extracts (Figures 3I and S3E). Second, anti-TAF12 IP-MS showed that when TAF10 is depleted, TAF12 maintains a stronger interaction with TAF5, TAF6, TAF4/4B, and TAF9/9B core-TFIID subunits, but interactions with the other TAFs or TBP are severely decreased (Figures 3J and S3F). In particular, the signals of the 3-TAF and S-TAF complex subunits decreased by about 4-fold at day 2 in the *RT10* mutant extracts compared to the control condition (Figure 3J). Signals of TAF5, TAF6, and TAF9/9B decreased by only 2-fold (Figure S3F) and gradually decreased on the following days, while TAF4/4B-TAF12 interaction was maintained up to day 3 (Figure S3F). Altogether, these data indicate that the core-TFIID is the main remaining partial TFIID complex present in the mutant *RT10* mESCs and confirm that TAF10 presence is important for the recruitment of other subunits on the core-TFIID to assemble the holo-TFIID.

Despite the efficient depletion of TAF7 and TAF10 at day 3 (Figure S1C), TAF7 and TAF10 could still be detected in the IP-MS performed on mutant *RT7* and mutant *R10* cells, respectively, suggesting the presence of remaining holo-TFIID. To estimate the potential fraction of stable holo-TFIID, we reanalyzed the IP-MS and the gel-filtration data at day 3 (Figure 3, see STAR Methods). Altogether, we estimated that 31% and 24% of holo-TFIID is still present in TAF7- and TAF10-depleted cells in the IP-MS experiments (Figures S3H and S3I). Using the gel-filtration data, we estimated that 18% and 28.5% of holo-TFIID is detected in TAF7- and TAF10-depleted cells, respectively (Figures S3J and S3K). Therefore, we estimated that a maximum

(D) Detection of holo-TFIID and TFIID submodules by the different antibodies.
(E and F) Relative normalized XIC values of TFIID subunits from anti-TBP (E) and anti-TAF12 (F) IP-MS from *RT7* mESCs at day (D) 2, D3, and D4 after EtOH (Control) or 4-OHT (−*Taf7*) treatment.
(G) Normalized relative XIC values for TFIID subunits from anti-TAF1 IP-MS in control and TAF7-depleted cells at D3.
(H) Western blot analyses of TAF1, TAF7, TAF10, and TBP expression after TAF7 (−*Taf7*) or TAF10 (−*Taf10*) depletion at D3. M, molecular weight marker.
(I and J) Relative normalized XIC values of TFIID subunits from anti-TBP (I) and anti-TAF12 (J) IP-MS from *RT10* mESCs at D2, D3, and D4 after EtOH (Control) or 4-OHT (−*Taf10*) treatment. In (E)–(G), (I), and (J), subunits of the same submodules (see D) were merged, except for the bait proteins and TAF2, TAF7, TAF8, and TAF10. TAF2 data were not taken into account because TAF2 was poorly detected in the IP-MS from controls. Core-TFIID: TAF4/4B, TAF5, TAF6, TAF9/9B, and TAF12. S-TAF: TAF1, TAF3, TAF11, TAF13, and TBP. D2, $n = 1$ biological replicate × 3 technical replicates; D3, $n = 3 \times 3$; D4, $n = 2 \times 3$ (E, F, I, and J); $n = 3$ technical replicates (G). Each dot corresponds to one measure of one subunit. Red crosses indicate proteins not detected in the control condition. Means ± SD are shown, Kruskal-Wallis test: ns, not significant; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$.

of 24.5% and 26% of remaining holo-TFIID persists in the TAF7- and TAF10-depleted cells, respectively, at day 3.

Altogether, these data demonstrate that TFIID assembly is differentially affected by TAF7 or TAF10 depletion. A TAF7-less TFIID persists in mutant *RT7* cells; however, in mutant *RT10* cells, the holo-TFIID complex is not assembled, and a partial core-TFIID-like complex containing TAF4/4B, TAF5, TAF6, TAF9/9B, and TAF12, but lacking TBP, remains.

TBP is recruited to the chromatin in TAF7- and TAF10-depleted mESCs

As at day 3 holo-TFIID depletion is well established without a major phenotype in both TAF7- and TAF10-depleted mESCs, we decided to focus our analyses at that stage to avoid potential secondary effects. As TAF7 or TAF10 depletion results in distinct partially assembled TAFs and incomplete TFIID complexes, either containing or lacking TBP, respectively, we investigated whether TBP would be recruited on promoters under these conditions. To this end, first we analyzed whole-cell extracts (WCEs), cytoplasmic-enriched extracts (CEs), nuclear-enriched extracts (NEs), and chromatin extracts (ChrEs) from control mESCs at day 3.[35] Enrichment of H3 in ChrE and tubulin in CE confirmed the fractionation (Figure S4A). TAF4, TAF5, TAF6, and TBP were more intense in the control ChrE, indicating a tight association of these subunits with the chromatin. While TAF7 was strongly depleted in all the fractions, there was no obvious difference in the levels of the tested TAF proteins between control and 4-OHT-treated RT7 cells at day 3 (Figure S4B), further reinforcing the presence of a TAF7-less TFIID complex in the TAF7-depleted cells. Interestingly, TBP levels appeared only mildly decreased in the TAF7-depleted condition, suggesting that the loading of TBP on the chromatin is only slightly affected in TAF7-depleted cells. Remarkably, when TAF10 was depleted, TAF8 was not detected as expected[9,23] in all the fractions; however, no obvious differences could be observed in TAF4, TAF5, TAF6, and TAF7, as well as TBP, levels (Figure S4C), suggesting that TBP recruitment at the promoters may not be significantly affected in TAF10-depleted cells.

TAF7-less TFIID and partial core TAF complexes are recruited at the promoters in TAF7- and TAF10-depleted mESCs, respectively

To evaluate the distribution of TBP on promoters in control and TAF7- or TAF10-depleted mESCs, we conducted anti-TBP CUT&RUN analyses at day 3. Annotation analyses indicated that TBP peaks were primarily located in promoter regions, as expected (Figures S5A–S5D). Comparison of the accumulation of TBP at the positions of the TBP-bound regions in the control conditions showed that TBP is still present on most promoters in TAF7- (Figures 4A and S5E) and in TAF10-depleted cells (Figures 4B and S5F). Nevertheless, the mean profile of TBP indicated that in both conditions, there is a reduction in TBP levels present at all the TBP-bound regions (Figures 4C, 4G, S5G, and S5K); however, no significant differential peak could be detected in TAF7-depleted cells, and only 16 differential TBP peaks could be detected in TAF10-depleted cells (Figures 4K, 4L, S5O, and S5P).

We carried out anti-TAF7, anti-TAF10, and anti-TAF12 CUT&RUN to measure the impact of TAF7 or TAF10 depletion on the accumulation of TAF7, TAF10, and TAF12 at the positions of TBP-bound promoters in the controls. As expected, we observed a strong depletion of TAF7 in TAF7-depleted mESCs (Figures 4D and S5H), but no decrease of TAF10 or TAF12 was detected at these sites (Figures 4E, 4F, S5I, and S5J), further indicating that under TAF7 depletion, a TAF7-less TFIID is binding to these promoters. Moreover, we observed a decrease of TAF10 in TAF10-depleted mESCs at the TBP-bound sites together with a strong decrease of TAF7 at these promoters (Figures 4I, 4H, S5L, and S5M). Importantly, TAF12 accumulation was not affected in the TAF10-depleted cells (Figures 4J and S5N), suggesting that a core-TFIID-like complex containing TAF12 may bind to these sites, possibly together with TBP.

Global Pol II activity is maintained in TAF7- and TAF10-depleted mESCs

After TAF7 or TAF10 depletion, the amount of holo-TFIID drastically decreased in mESCs. We therefore investigated the impact of the remaining partial TAF-containing complexes on nascent Pol II transcription. No differences in the incorporation of 5-ethynyl uridine (EU) after 2 to 5 days of treatment could be observed between control and TAF7- or TAF10-depleted cells (Figure S6A). To assess active Pol II transcription, we investigated the phosphorylation status of the C-terminal domain of the Pol II subunit RPB1. WB analysis of RPB1 at day 3 indicated that the hypo- (Pol II-A) and hyper- (Pol II-O) phosphorylated RPB1 levels were comparable in both mutant and control conditions (Figure S6B). We then specifically analyzed the phosphorylation of the serine 5 ($RPB1^{pSer5}$, pSer5) and serine 2 ($RPB1^{pSer2}$, pSer2) associated with the initiation and elongation status of Pol II, respectively (reviewed in Harlen and Churchman[36]) by immunofluorescence (IF) and WB. Similarly, no obvious difference in $RPB1^{pSer2}$ or $RPB1^{pSer5}$ phosphorylation signal could be observed between control and mutant *RT7* and *RT10* cells (Figures 5A–5D), indicating that Pol II transcription initiation and elongation are active and are not obviously affected after TAF7 or TAF10 depletion.

To further understand the effect of TAF7 or TAF10 depletion on Pol II transcription, we carried out anti-Pol II chromatin immunoprecipitation sequencing (ChIP-seq) experiments at day 3. No significant differences could be observed in the TAF7-depleted cells at the transcription start site (TSS), in the gene body region, or at the transcription end site (TES) (Figures 5E and 5F). However, surprisingly, a significant strong decrease in Pol II accumulation could be observed in the TAF10-depleted cells at the TSS (Figures 5G and 5H), while on the metagene profiles of Pol II, detection in the gene body and TES region was less impacted (Figure 5H). Furthermore, analyses of the pausing ratios indicated a reduction in Pol II pausing at the TSS region compared to the gene body in both mutant cells; however, the pausing effect was more pronounced in TAF10-depleted cells (Figures 5I, 5J, S6C, and S6D).

Altogether, these data suggest that despite a dramatic decrease in the amount of holo-TFIID and a decreased enrichment of Pol II in TAF10-depleted cells at TSSs, the global activity of the Pol II is maintained in TAF7- or TAF10-depleted mESCs.

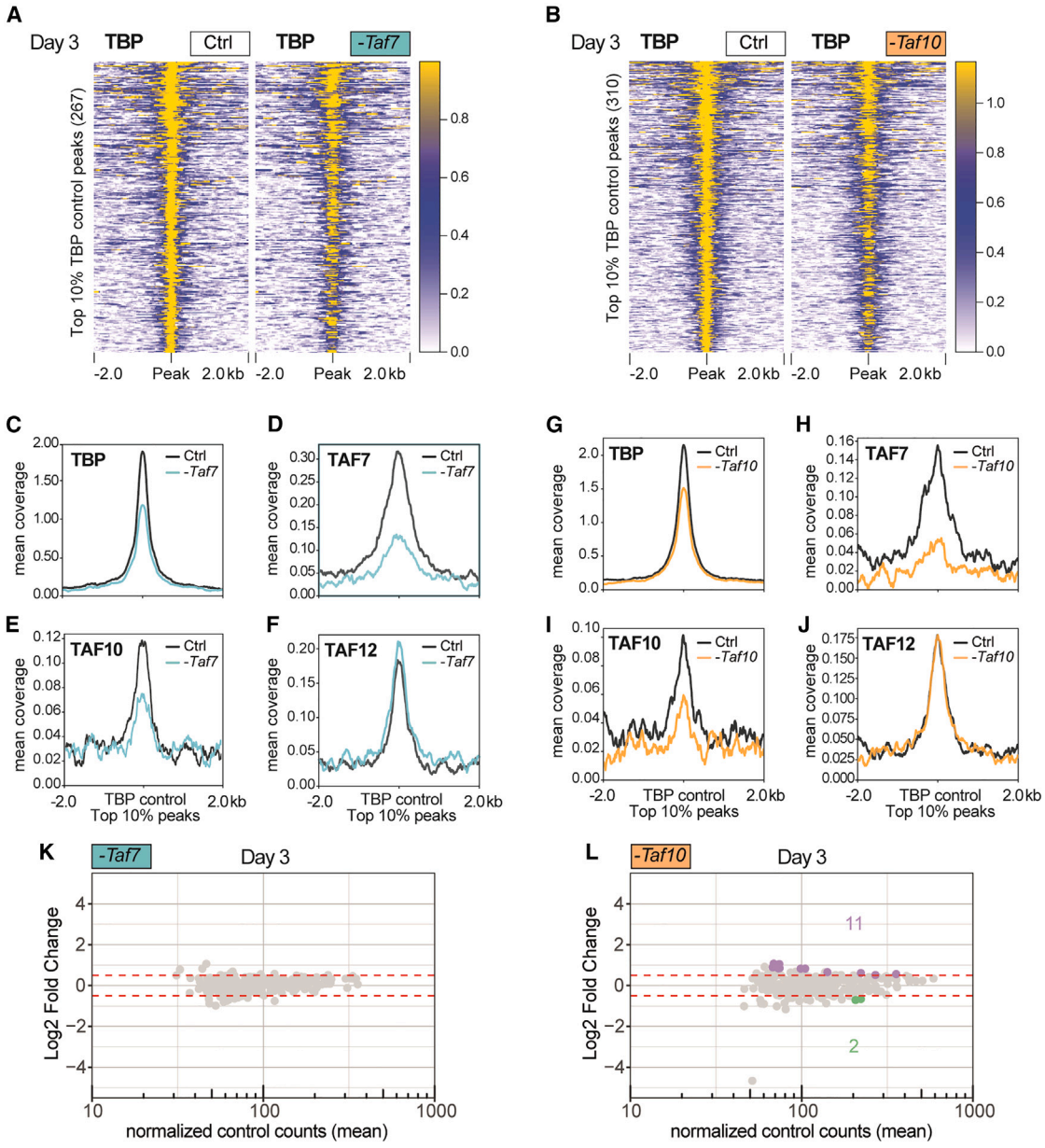


Figure 4. TBP is recruited at promoters in TAF7- or TAF10-depleted mESCs

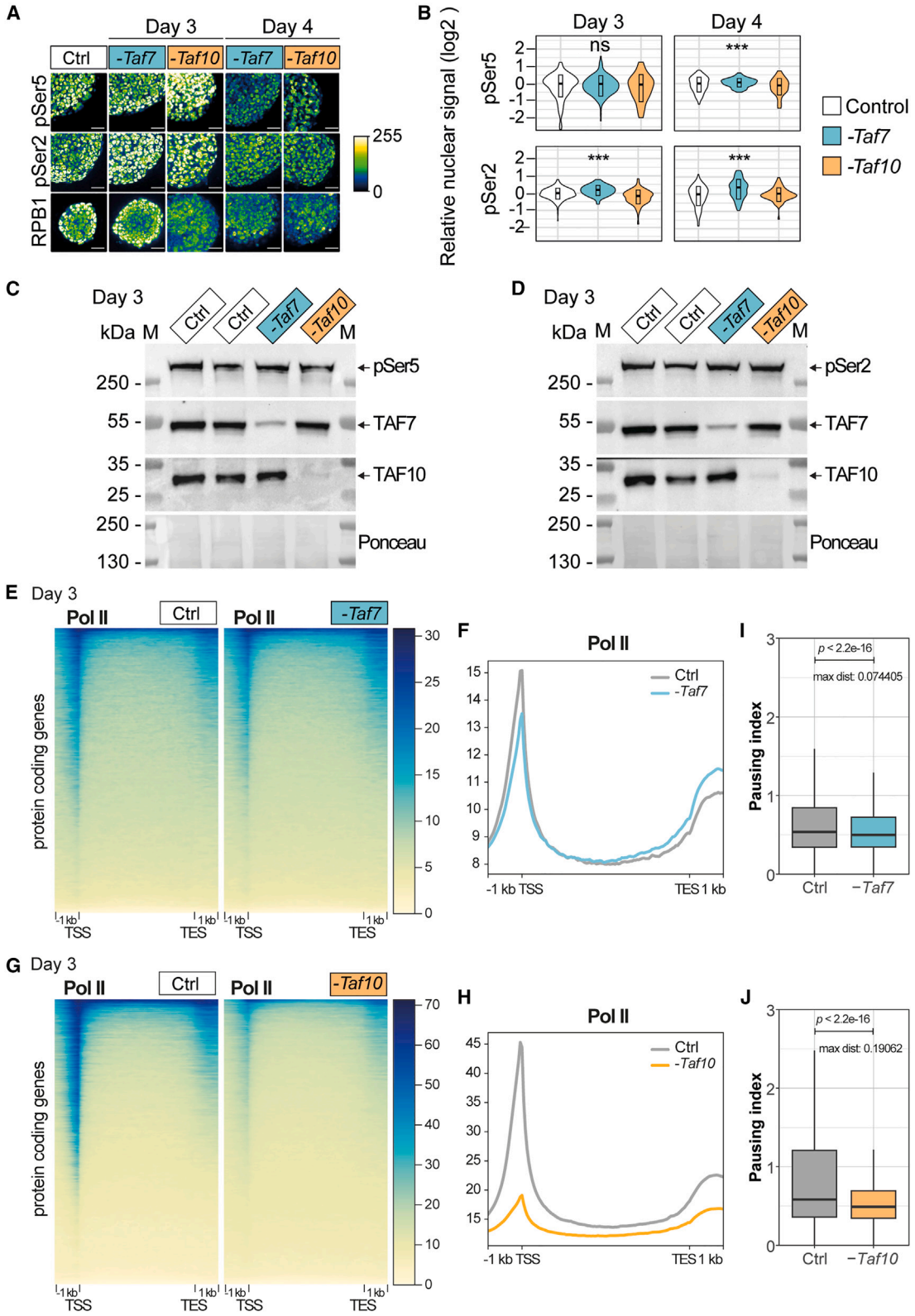
(A and B) TBP distribution at the position of the top 10% (fold enrichment) anti-TBP CUT&RUN peaks (±2 kb) detected in the control condition and in *RT7* control and TAF7-depleted (A) and *RT10* control and TAF10-depleted (B) mESCs at D3.

(C–J) Mean profiles of TBP (C and G), TAF7 (D and H), TAF10 (E and I), and TAF12 (F and J) protein accumulation at the TBP control peaks in control (black line) and TAF7-depleted (blue line) (C–F) and in control (black line) and TAF10-depleted (orange line) (G–J) mESCs.

(K and L) MA plots of differential peak analyses in TAF7-depleted (K) and TAF10-depleted (L) mESCs at D3, with the number of significantly down- (green) or upregulated (purple) TBP peaks (adjusted $p \leq 0.05$, $|\log2 \text{ fold change}| \geq 0.5$). No significant peaks were detected in the TAF7-depleted cells.

TAF7 or TAF10 depletion has a limited impact on nascent Pol II transcription

As global Pol II transcription is maintained in the TAF7- and TAF10-depleted mESCs, we investigated whether TAF7 or TAF10 depletion would specifically impact Pol II active transcription in mESCs by analysis of nascent RNA using 4-Thiouridine (4-sU) at day 3 (Figure 6A). As spikes, we added unlabeled yeast as well as *Drosophila* S2 4-sU-labeled RNA. Only reads from 4-sU-labeling experiments could be mapped on the *Drosophila* and mouse genomes, while no reads could be mapped on the yeast genome, confirming the enrichment of newly synthetized RNAs (Figure 6B). Principal-component (PCA) analyses indicated that in both mESC lines, depleted and control condition samples were well separated (Figures 6C and 6D).



(legend on next page)

Differential expression analyses confirmed that TAF7 and TAF10 depletion is not accompanied by a global reduction in Pol II transcription, as only a small fraction of the nascent transcripts is affected (Figures 6E and 6F). In TAF7-depleted mESCs, 0.40% of transcripts were significantly downregulated versus 0.14% upregulated (Figure 6E), while in TAF10-depleted mESCs, 1.34% of transcripts were significantly downregulated versus 0.08% upregulated (Figure 6F) in the 4-OHT-treated compared to the control conditions, with a threshold of −0.5 log2 fold change and an adjusted *p* value of 0.05. While there is no obvious difference in the ratio of down- versus upregulated transcripts in TAF7-depleted mESCs, there are more downregulated transcripts in TAF10-depleted cells (Figures 6E and 6F), suggesting a more pronounced decrease in global transcription in these cells (Figure 6G), which could be explained by the more affected Pol II chromatin distribution observed in TAF10-depleted cells (Figures 5G and 5H). Comparison of the fold changes (Figure 6H) indicates that there is a correlation between the *RT7* and the *RT10* data (Pearson correlation coefficient: 0.519). However, there is only a partial overlap of commonly differentially regulated transcripts (Figure S6E), and overall, only GO categories related to translation were significantly enriched in mutant *RT10* cells (Figures S6F–S6H).

The intersection of our genomic data analyses suggested that when nascent transcription is not significantly affected by TAF7 or TAF10 depletion, the TBP peaks were not affected, as seen at different loci on UCSC genome browser snapshots (Figures 6I, 6J, S6I, and S6J). However, while Pol II chromatin distribution is not obviously affected in TAF7-depleted cells, it is severely affected in TAF10-depleted cells, with an important decrease at the TSS and only a slight reduction in the gene body (Figures 6I, S6I, and S6J). In contrast, in the case of the *Hes1* gene, which is significantly downregulated in TAF10-depleted mESCs (Figure 6J), Pol II was barely detectable, and TBP was evicted from the *Hes1* promoter, indicating that genes that are downregulated following TAF10 depletion seem to have lost TBP as well as Pol II recruitment and elongation. As the *Hes1* promoter contains a TATA box, we investigated whether TATA-like motifs were enriched in the promoters of downregulated transcripts in TAF10-depleted cells but could not detect a significant difference with non-affected promoters (Figure S6K).

In conclusion, our data indicate that Pol II global transcription is not seriously affected in conditions where holo-TFIID is depleted in either TAF7- or TAF10-depleted mESCs. As in our biochemical analyses of TFIID and TFIID-like complexes, TAF10 depletion has a stronger effect than TAF7 depletion on the transcription of a subset of genes. This suggests that the TAF7-less TFIID complex allows efficient recruitment of Pol II on promoters and that the partial TAF assemblies detected in the TAF10-depleted cells, while also being globally sufficient to sustain Pol II transcription, but not pausing, fail to do so for a subset of genes.

Embryonic deletion of *Taf7* or *Taf10* induces different embryonic phenotypes with distinct severities

Our experiments indicate that Pol II transcription is not majorly affected at day 3 in steady-state conditions (Figures 5 and 6). However, at day 6, the mutant mESCs are affected, especially the TAF10-depleted cells (Figure 1), suggesting that, while no major defects are detected initially, Pol II transcription is actually affected on a longer term. To analyze the effect of the TAF7 and TAF10 depletion in a more dynamic and *in vivo* transcriptional context, we compared the conditional deletion of *Taf10* or *Taf7* using the *T-Cre* transgenic line, which is active in the early posterior mesoderm.[37] We had previously shown that depletion of TAF10 in the early mesoderm results in a growth arrest, while cyclic transcription in the pre-somitic mesoderm associated with somitogenesis is not initially affected.[23] The efficiency of TAF7 depletion was assessed by IF (Figures S7A and S7B). Tg(T-Cre/+);*Taf7*$^{f/f}$ and Tg(T-Cre/+);*Taf10*$^{f/f}$ mutant embryos (hereafter called *TCre;Taf7* and *TCre;Taf10*, respectively) are similar at E9.5 and display a similar reduction in the somite index (Figures 7A–7C and S7C). While no limb buds are present in *TCre;Taf10* mutant embryos,[23] *TCre;Taf7* embryos displayed forelimb buds, although smaller compared to the control (Figures 7A and 7B). At E10.5, both *TCre;Taf10* and *TCre;Taf7* mutant embryos display a growth arrest in the trunk and tail regions (Figures 7D–7F). After E10.5, more differences could be observed between the two mutants. First, enlarged pericardia are observed in all the *TCre;Taf7* mutant embryos (Figure 7E) suggesting that *Taf7* deletion leads to cardiovascular defects not observed in *TCre;Taf10* mutant or control embryos (Figures 7D and 7F). Second, while *TCre;Taf10* mutants could not be retrieved after E11.5 due to placenta and allantois degeneration,[23] *TCre;Taf7* mutant embryos are still detected at E12.5 (Figure 7H). Third, the presence of blood cells is obvious in the vasculature of controls and *TCre;Taf7* mutant embryos (Figures 7J and 7K), but not of *TCre;Taf10* mutant embryos (Figure 7L). Altogether, these data indicate that *Taf7* or *Taf10* conditional deletion in the same genetic context leads to similar growth arrest phenotypes, but different outcomes. *Taf10*

Figure 5. Depletion of holo-TFIID in the absence of TAF7 or TAF10 has only limited defects on Pol II global transcription

(A) Representative views of immunofluorescence using anti-RPB1, anti-RPB1^{pSer2}, and anti-RPB1^{pSer5} antibodies on *RT7* and *RT10* mESCs treated with 4-OHT (−*Taf7* or −*Taf10*) and with EtOH (Ctrl) at day (D) 3 and D4. Color scale: Green Fire Blue LUT. Scale bars: 50 μm.

(B) Quantification of RPB1^{pSer5} and RPB1^{pSer2} nuclear signal represented as violin plots. D3 and D4, *n* = 1 biological replicate. Kruskal-Wallis test: ns, not significant; **p* < 0.05, ***p* < 0.01, and ****p* < 0.001.

(C and D) Western blot analyses of TAF7, TAF10, and RPB1^{pSer5} (C) and RPB1^{pSer2} (D) expression in *RT7* and *RT10* mESCs treated with EtOH (Ctrl) and TAF7- (−*Taf7*) and TAF10- (−*Taf10*) depleted mESCs at D3. The Ponceau staining is displayed at the bottom. M, molecular weight marker.

(E–H) Heatmaps (E and G) and mean profiles (F and H) of Pol II distribution on protein-coding genes (from −1 kb upstream of the TSS to +1 kb downstream of the TES) in *RT7* control (Ctrl) and TAF7-depleted (−*Taf7*) (E and F) and *RT10* control (Ctrl) and TAF10-depleted (−*Taf10*) (G and H) mESCs at D3. TSS, transcription start site; TES, transcription end site.

(I and J) Pausing index calculated in *RT7* control (Ctrl) and TAF7-depleted (−*Taf7*) mESCs (I) and in *RT0* control (Ctrl) and TAF10-depleted (−*Taf10*) mESCs (J) at D3. Kolmogorov-Smirnov test. The *p* values and maximal distances (max dist) are indicated.

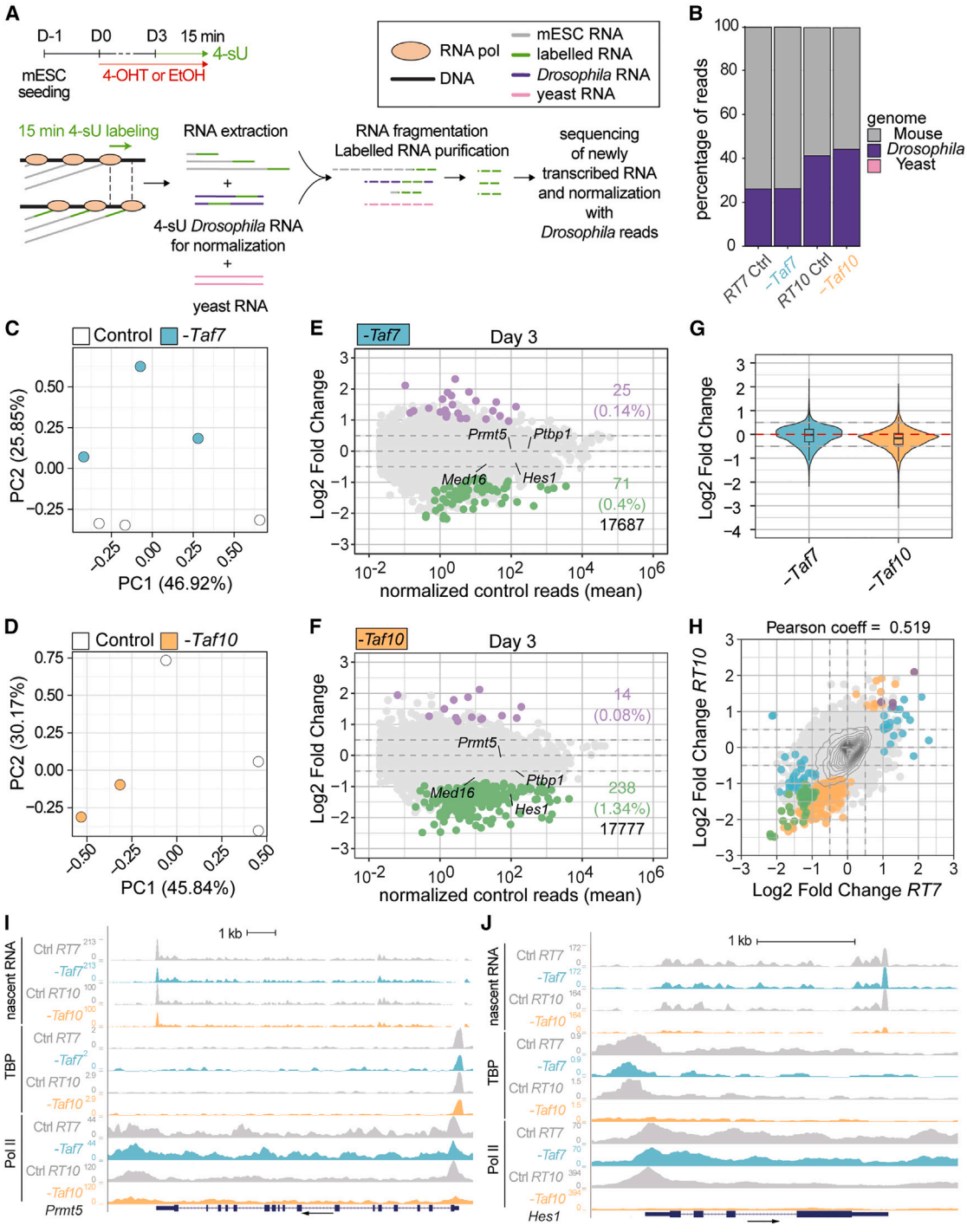


Figure 6. TAF7 loss has a milder impact on RNA Pol II gene transcription than TAF10 loss after 3 days of treatment

(A) Experimental strategy.

(B) Percentage of mouse (gray), *Drosophila* (purple), and yeast (pink) reads.

(C and D) Principal-component analysis (PCA) of *RT7* (C) and *RT10* (D) control (white dot) and mutant (colored dot) mESC nascent transcriptomes.

(E and F) MA plots of mutant versus control *RT7* (E) and *RT10* (F) mESCs, with the numbers and percentages of significantly down- (green) or upregulated (purple) protein-coding transcripts displayed on the right and total transcripts shown at the bottom (20 normalized reads per gene length in kilobases threshold, adjusted $p \leq 0.05$ (FDR-corrected Wald test) and $|\log2 \text{ fold change}| \geq 1$).

(G) Global log2 fold change comparisons in TAF7- and TAF10-depleted mESCs.

(legend continued on next page)

conditional deletion leads to a more severe phenotype associated with lack of forelimbs and red blood cells, while *Taf7* conditional deletion leads to a milder phenotype and is associated with cardiovascular defects. These data confirm the difference in phenotype severity observed in TAF7- versus TAF10-depleted mESCs and further support that TAF10 depletion has a stronger impact on TFIID assembly compared to TAF7-depletion.

In conclusion, our *in vivo* results demonstrate that embryonic deletion of either *Taf7* or *Taf10* induces different phenotypes with distinct severities. Thus, while our *in vitro* data demonstrated that partial TFIID modules are able to sustain nascent Pol II transcription in mESCs, these partial TFIID assemblies cause embryonic lethality over time, as they do not support all active gene regulatory changes during development.

DISCUSSION

Different requirements for TAF7 and TAF10 during development

We observed a difference in phenotype severity between TAF7 and TAF10 depletion. TAF10 is important for mESC survival, as observed in *Taf10* mutant F9 cells and in the ICM of *Taf10*$^{-/-}$ mutant blastocysts.[25,38] In contrast, TAF7 is not essential for mESC survival, in agreement with the survival of *Taf7*$^{-/-}$ blastocysts until hatching at E4.5.[26] Importantly, as TAF10 is also part of the transcriptional co-activator complex SAGA, we confirmed that the difference in phenotype in the *RT10* mESCs cells was not a consequence of a defect in SAGA assembly, and remarkably, we observed that TAF10 was required for the stability of SUPT7L, as it is for TAF8.[9,23] It is tempting to speculate that TAF10 is also required for the stability of its other HFD partner in lobe A, TAF3. As a consequence, destabilization of TAF8 would lead to the loss of TAF2 from lobe B, and loss of TAF3 is predicted to impair integration of TAF11/TAF13 in lobe A, thus preventing TBP integration. The severity of this domino effect in the absence of TAF10 potentially explains the stronger phenotype observed in TAF10-depleted mESCs and *TCre;Taf10* embryos.

We interpret the increased severity in TAF10-depleted cells compared to TAF7-depleted cells to be from different impacts on holo-TFIID assembly (see below). However, in the *Taf7* and *Taf10* conditional mutant embryos, we observed two striking phenotypes specific to each TAF: enlarged pericardium in the *TCre;Taf7* embryos and a lack of blood cells in the *TCre;Taf10* embryos. This suggests that TAF7 and TAF10 are associated with specific regulatory functions despite the fact that TAF7 and TAF10 are both subunits of holo-TFIID. This might also suggest that TAF7-less TFIID cannot regulate genes necessary for cardiovascular development, while red blood cell differentiation is more dependent on TAF10-regulated pathways, as supported by the direct TAF10/GATA1 interaction during erythropoiesis.[39] Importantly, the lack of TAF10 seriously affects both holo-TFIID and holo-SAGA assemblies, but does not affect many gene regulatory pathways important for other developmental processes. This observation is in line with our finding that partial TFIID complexes can still regulate Pol II gene expression from many developmental gene promoters.

Sequential assembly of holo-TFIID

Assembly of holo-TFIID is the result of the sequential incorporation of different TFIID submodules on the TAF1 nascent scaffold.[2,10–12] Our data support this model, as holo-TFIID formation is arrested at different steps in our TAF7- or TAF10-depleted cells. Indeed, our IP-MS analyses indicate that in the absence of TAF7, a TAF7-less TFIID complex containing TBP is formed, in agreement with TAF7 being the last subunit to be recruited on TAF1 during TFIID assembly.[10] In contrast, in TAF10-depleted cells, we could detect only the core-TFIID complex. In normal conditions, the next step after core-TFIID assembly is the recruitment of the 3-TAF module.[2,11] In the absence of TAF10, as it is required for the stability of its HFD partners (this study, Kamenova et al.,[9] and Bardot et al.[23]), there is no 3-TAF complex formed. Altogether, our data confirm that *in vivo*, the holo-TFIID complex is sequentially assembled from different building blocks.

Active Pol II transcription in TAF7- and TAF10-depleted mESCs

A striking observation from our study is that despite an obvious impact on TFIID assembly, no major impact on nascent Pol II transcription could be observed. Conditional deletion of *Taf10* during somitogenesis did not impact the active transcription of cyclic genes associated with the formation of the somites.[23] In this precedent study, we showed a limited effect on the steady-state transcriptome, but we could not exclude a compensation of transcription initiation defects by a buffering of RNA decay.[40,41] In the present study, we showed a lack of obvious impact on active transcription. Only a very limited proportion of expressed transcripts were significantly and differentially affected in TAF7- or TAF10-depleted cells after 3 days of treatment, when our IP-MS data showed that the assembly of TFIID is impaired. Surprisingly, while no obvious difference could be observed in Pol II chromatin distribution in TAF7-depleted cells, the lack of TAF10 resulted in a massive decrease in Pol II recruitment at TSSs, with only a minor decrease in gene bodies. These data suggest that despite the limited effect of TAF10 depletion on active Pol II transcription in steady-state conditions, the Pol II transcription initiation process is nevertheless affected on all genes, as indicated by the slight global downregulation shift of log2 fold change distribution when TAF10 is depleted, correlating with the more severe phenotype observed in TAF10-depleted cells. We also showed that, while TAF7 or TAF10 depletion did not affect cell viability when passaged at day 2 of treatment, clonal growth was affected when the depleted cells were passaged at day 4 of treatment (Figures 1I and 1J). These data indicate that TAF7 and TAF10 depletion is

(H) Correlation plot of the log2 fold changes in TAF7- and TAF10-depleted mESCs. Significantly differentially regulated transcripts are highlighted as follows: only in TAF7-depleted cells (blue), only in TAF10-depleted cells (orange), and downregulated (green) or upregulated (purple) in both depleted mESCs.
(I and J) UCSC genome browser views of nascent RNA, TBP distribution, and Pol II at *Prmt5* (I) and *Hes1* (J) loci between control (Ctrl *RT7* and Ctrl *RT10*), TAF7-depleted (−*Taf7*), and TAF10-depleted (−*Taf10*) mESCs. The y axes indicate the genomic coverage, and arrows indicate direction of transcription.

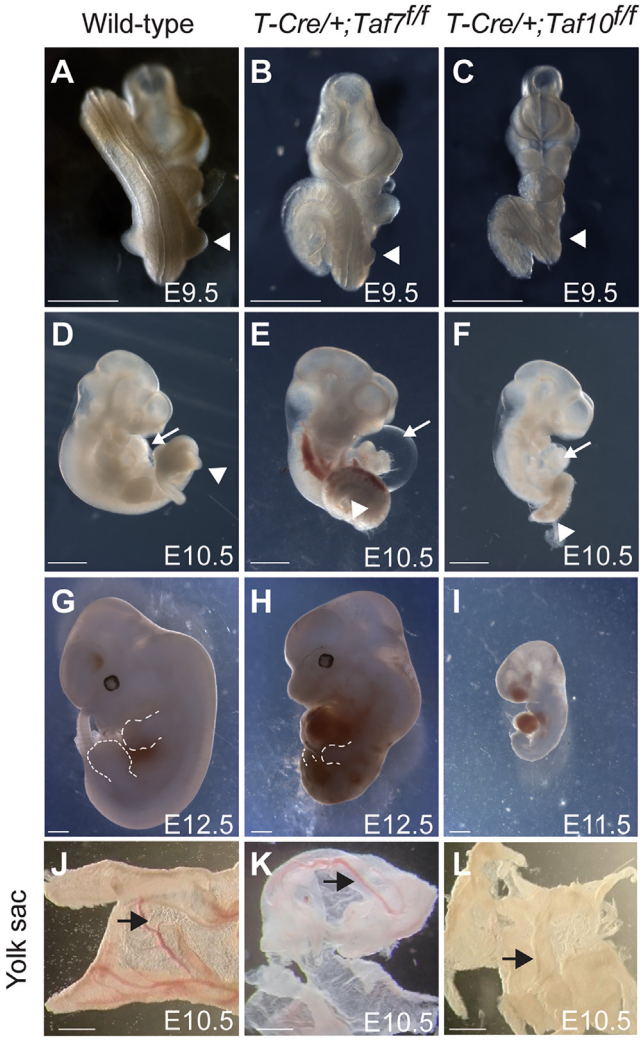


Figure 7. Conditional deletion of *Taf7* or *Taf10* in the early mesoderm results in similar yet different phenotypes

(A–I) Whole-mount views of wild-type (A, D, and G) and Tg(T-Cre/+);*Taf7*$^{f/f}$ (B, E, and H) and Tg(T-Cre/+);*Taf10*$^{f/f}$ (C, F, and I) mutant embryos at E9.5 (A–C, $n > 4$), E10.5 (D–F, $n > 4$), and E12.5 (G and H, $n > 2$). As no Tg(T-Cre/+);*Taf10*$^{f/f}$ mutant embryos could be recovered at E12.5, an E11.5 Tg(T-Cre/+);*Taf10*$^{f/f}$ mutant embryo is shown (I). White arrowheads, forelimbs (A–C) and hindlimbs (D–F); white arrows, heart (D–F); dashed lines, limb buds (G and H).

(J–L) Whole-mount views of wild-type (J), Tg(T-Cre/+);*Taf7*$^{f/f}$ (K), and Tg(T-Cre/+);*Taf10*$^{f/f}$ (L) embryo yolk sacs ($n > 3$). Black arrows, blood vessels. Scale bars, 1 mm.

actually damaging for the mESCs but that during a window of time, the cells are able to compensate.

These surprising observations can be explained by several mutually non-exclusive hypotheses. A first hypothesis is that residual holo-TFIID complexes in TAF7- or TAF10-depleted cells are responsible for the maintenance of active transcription. Using our gel-filtration and IP-MS data, we estimated that at day 3, the holo-TFIID fraction remaining is 24.5% and 26% in TAF7- and TAF10-depleted cells, respectively. These numbers are actually overestimations, as they are indirect measurements; however, they reflect the reality of the persistence of holo-TFIID already observed in E9.5 mouse *TCre;Taf10* embryos.[23] It is conceivable that chromatin-bound holo-TFIID has a longer half-life and is very difficult to destroy in the time frame of our studies. This would suggest that only a very minimal amount of holo-TFIID may be required to maintain active transcription in a steady-state situation. However, TBP has a short residence time at promoters of seconds,[42] and as the mESCs are still dividing, especially in the TAF7-depleted cells, which are almost indistinguishable from the control *RT7* cells up to 6 days of treatment, remaining chromatin-bound TFIID complexes or TBP would be dislodged during DNA synthesis and diluted out in the cell population. Thus, this hypothesis cannot solely explain our observations.

A second hypothesis is that mESCs are not using a holo-TFIID complex. It has been proposed that human ESCs (hESCs) have a non-canonical TBP-containing TAF complex, as expression of several TAF proteins, including TAF7, but not TAF10, was not detected.[43] However, as shown in the same study and in our present data, this is not the case in mESCs. Differential requirements for TAF4, TAF8, and TAF12 have been reported, as *Taf4* deletion does not impact mESC growth, in contrast to *Taf8* conditional deletion and TAF12 depletion.[19,44,45] Interestingly, depletion of TAF8 and TAF12 leads to a strong defect in holo-TFIID assembly.[44,45] Our data are in favor of a TAF7-less TFIID complex almost as active as the holo-TFIID, as we did not observe major difference in the proportion in TAF proteins, except TAF7, in our IP-MS or in our CUT&RUN analyses. These data are in agreement with the observation that TAF7 is not required for mature T cell survival and differentiation and that in these conditions, an apparent TAF7-less TFIID was observed, although depletion efficiency of TAF7 was only about 50%.[26]

A third hypothesis is that holo-TFIID is not required to maintain active transcription reinitiation, as proposed in Tatarakis et al.[29] During the very first round of Pol II transcription initiation, holo-TFIID would be required to create a scaffold, a PIC, that would allow some GTFs to remain at the promoter facilitating reinitiation and the recruitment of a new Pol II.[46,47] Genes that have been already transcribed once would not necessarily require holo-TFIID, maybe only TBP and/or some partial TAF complexes, for the reinitiation of transcription. To further test this hypothesis, it would be interesting to differentiate our *RT7* and *RT10* mESCs to evaluate the impact of the depletion of TAF7 or TAF10 on the activation of a specific new gene expression program. It would be important to test differentiation in various cell types, as the loss of TAF10 has a differential effect depending on the cellular context.[23,25,28,29]

Transcription initiation by partial TFIID complexes

Recent structural studies have proposed that holo-TFIID plays a major role in PIC assembly and deposition of TBP on core promoters.[3,7,8] In this model, TBP acts as an anchor on the DNA, allowing the positioning of the PIC independent of the sequence of the core promoter.[48] Therefore, holo-TFIID and TBP appear central to the process of Pol II transcription initiation. However, evidence suggests that Pol II transcription initiation does not always require TBP or holo-TFIID. *In vitro* transcription initiation can be mediated by the TBP-free TAF complex (TFTC), a mixture of

SAGA and other TAF10-containing TAF complexes.[49,50] The snRNA activating protein complex (SNAPC) recruits TBP on Pol II small nuclear (snRNA) coding genes.[51] In *Drosophila*, a non-TFIID complex containing the TBP paralog TBPL1 controls *H1* and ribosomal protein-coding gene transcription initiation,[52–55] and different PIC complexes not containing TBP and associated with different classes of promoters have also been described recently.[56] In the mouse, a complex containing the TBP paralog TBPL2, lacking TBP and TAF proteins, mediates transcription during oocyte growth.[57] More strikingly, transcription initiation independent of TBP and its paralogs has been described for a subset of genes during *Xenopus* development,[58] and two recent studies have shown that acute depletion of TBP has no major impact on Pol II transcription without compensation by a TBP paralog.[59,60]

Our data are in favor of transcription initiation mediated by partially assembled TFIID complexes in the mESCs in steady-state conditions. On one hand, our data support the idea that the TAF7-less TFIID complex is able to mediate transcription initiation, at least up to a certain degree, as the TAF7-depleted mESCs are sensitive to late replating and the *TCre;Taf7* embryos display a phenotype. On the other hand, a partially assembled TFIID-like complex is present in TAF10-depleted cells, and while Pol II enrichment at the TSS is severely affected, the transcription activity of many genes is not majorly affected. It is not clear how the core-TFIID could mediate transcription initiation. It has been recently shown that the *in vitro*-assembled S-TAF complex promotes transcription initiation *in vitro*.[13] However, the interaction of TBP with all the TAFs decreases in our anti-TBP IP-MS, suggesting that the S-TAF complex does not exist *in vivo* in TAF10-depleted cells. These data, in addition to TBP relocation to Pol III and Pol I machineries (Figures S3A and S3B), confirm the structural data indicating that the core-TFIID does not interact strongly with TBP.[3,7,8,12,61,62] Moreover, it is not clear how the core-TFIID could be imported in the nucleus, as the nuclear localization signal (NLS) of TAF8, incorporated in the 8-TAF complex, is important for this import.[2,11,32] Alternatively, the TAF1 nascent protein, which already interacts with TBP, could recruit the core-TFIID and be imported into the nucleus. As TAF1 interacts structurally with TAF6,[8,10] such a complex could interact with DNA via TAF4 and TAF1, as well as TBP, as observed in the holo-TFIID/PIC structure.[8] The apparent size of 670 kDa of the partial TFIID complex in our gel-filtration experiment supports the existence of such a complex. However, this interaction might be weak, as it could not be detected in our IP-MS experiments. This could be explained by the absence of TAF8, but also by the destabilization of TAF1 in TAF10-depleted cells (Figure 3H), suggesting that incomplete holo-TFIID assembly on TAF1 nascent protein leads to a destabilization of the latter.

In conclusion, our study supports the vision of a much more dynamic and flexible Pol II transcription initiation machinery. Interestingly, this flexibility has some impact on human physiopathology, as a growing number of mutations have been identified in *TAF1*, *TAF2*, *TAF4*, *TAF6*, *TAF8*, and *TAF13* genes in patients associated with neurodevelopmental defects and intellectual disabilities.[44,63,64] Surprisingly, these mutations mainly affect the brain development and function but without any major effect on the other organs, clearly suggesting that the transcription initiation machinery has a certain degree of adaptability but that the cellular context is a main constraint.

Limitations of the study

In our study, we tested the functional importance of TAF7 and TAF10 using an inducible CRE/*LoxP* approach. While the genomic deletion is very efficient and fast, this approach does not target the proteins that have already been translated. However, we evaluated the potential persistence of some TAF7 or TAF10 protein at different time points, by IF, WB, and IP-MS. Moreover, we analyzed the status of the TFIID complex to confirm the impact of the depletion of TAF7 or TAF10.

As our study is more long term compared to an acute depletion approach, we decided to perform our transcriptomic and genomic analyses at day 3, which corresponds to an efficient depletion of the targeted proteins and to a stage where no important cellular phenotype could be observed. However, as our cells were depleted of TAF7 or TAF10 for days, we cannot rule out any secondary effects, although, as we show, there were no major defects in transcription. This does not rule out the total absence of phenotype, as cells became affected in the longer term, but our study does not allow the analysis of the longer-term defects.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources and reagents should be directed to and will be fulfilled by the lead contact, Stéphane D. Vincent (vincent@igbmc.fr).

Materials availability

Cell lines generated in this study are available upon request with MTA completion.

Data and code availability

- Nascent RNA sequencing, CUT&RUN, and ChIP-seq data have been deposited at GEO[65] GEO: GSE245196. The mass spectrometry proteomics data have been deposited with the ProteomeXchange Consortium via the PRIDE[66] partner repository PRIDE: PXD046459. Raw data for the WB and cell counting experiments were deposited on Mendeley, Mendeley Data: https://doi.org/10.17632/m66wgvw7sm.2. All datasets are publicly available as of the date of publication and accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this study is available from the lead contact (Stéphane D. Vincent, vincent@igbmc.fr) upon request.

ACKNOWLEDGMENTS

We thank all members of the Tora lab for protocols, thoughtful discussions, and suggestions; C. Ebel and M. Philipps for help with FACS sorting; S. Falcone, M. Gendron, and the IGBMC animal facility for animal caretaking; and the PluriCell East platform for cell culture services. We thank E. Compe for the anti-PSer2 and anti-PSer5 antibodies. This work was supported by funds from the Agence Nationale de la Recherche (ANR-19-CE11-0003-02, ANR-19-CE12-0029-01, ANR-20-CE12-0017-03, and ANR-22-CE11-0013-01 to L.T.; ANR-23-CE12-0034 to S.D.V.), NIH MIRA (R35GM139564) and NSF (award no. 1933344) to L.T., and Fondation pour la Recherche Médicale (EQU-2021-03012631) to S.D.V. and L.T. This work, as part of the ITI 2021–2028 program of the University of Strasbourg, was also supported by IdEx Unistra (ANR-10IDEX-0002), SFRI-STRAT'US project (ANR 20-SFRI-0012), and EUR

IMCBio (ANR-17-EURE-0023) under the framework of the French Investments for the Future Program. Sequencing was performed by the GenomEast Platform, a member of the France Génomique consortium (ANR-10-INBS-0009). V.H. was a recipient of a fellowship from the University of Strasbourg Doctoral School ED414.

AUTHOR CONTRIBUTIONS

V.H., P.B., D.D., A.B., E.G.S., C.R., M.S., Y.H., C.E., and L.H.A. performed experiments. B.M. and L.N. performed the mass spectrometry analyses; M.J., S.L.G., V.H., Y.H., and S.D.V. performed the bioinformatic analyses; and V.H., L.T., and S.D.V. designed the experiments with the help of F.M. S.D.V. conceived and supervised the project and wrote the paper with support from F.M., V.H., and L.T.

DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Mice
  - Generation and maintenance of mESCs
- METHOD DETAILS
  - Collection of mouse embryos
  - Calculation of the somite index
  - Embryo sectioning and immunolocalization
  - Genotyping and screening
  - 4-Hydroxytamoxifen (4-OHT) treatment
  - mESC phenotypic analysis
  - Depletion analysis
  - TFIID complex composition analysis
  - Transcription analysis
  - Cleavage under targets and release using nuclease (CUT&RUN) and library preparation
  - Pol II ChIP-seq
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Immunofluorescence and EdU signal quantification
  - IP-MS data analysis
  - Nascent 4-sU RNA sequencing data analysis
  - CUT&RUN data analysis
  - Pol II Chip-seq data analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.114791.

Received: January 9, 2024
Revised: July 9, 2024
Accepted: September 7, 2024
Published: September 30, 2024

REFERENCES

1. Roeder, R.G. (2019). 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. Nat. Struct. Mol. Biol. *26*, 783–791. https://doi.org/10.1038/s41594-019-0287-x.
2. Bieniossek, C., Papai, G., Schaffitzel, C., Garzoni, F., Chaillet, M., Scheer, E., Papadopoulos, P., Tora, L., Schultz, P., and Berger, I. (2013). The architecture of human general transcription factor TFIID core complex. Nature *493*, 699–702. https://doi.org/10.1038/nature11791.
3. Patel, A.B., Louder, R.K., Greber, B.J., Grünberg, S., Luo, J., Fang, J., Liu, Y., Ranish, J., Hahn, S., and Nogales, E. (2018). Structure of human TFIID and mechanism of TBP loading onto promoter DNA. Science *362*, eaau8872. https://doi.org/10.1126/science.aau8872.
4. Jacobson, R.H., Ladurner, A.G., King, D.S., and Tjian, R. (2000). Structure and Function of a Human TAFII250 Double Bromodomain Module. Science *288*, 1422–1425. https://doi.org/10.1126/science.288.5470.1422.
5. Vermeulen, M., Mulder, K.W., Denissov, S., Pijnappel, W.W.M.P., van Schaik, F.M.A., Varier, R.A., Baltissen, M.P.A., Stunnenberg, H.G., Mann, M., and Timmers, H.T.M. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell *131*, 58–69. https://doi.org/10.1016/j.cell.2007.08.016.
6. Buratowski, S., Hahn, S., Sharp, P.A., and Guarente, L. (1988). Function of a yeast TATA element-binding protein in a mammalian transcription system. Nature *334*, 37–42. https://doi.org/10.1038/334037a0.
7. Louder, R.K., He, Y., López-Blanco, J.R., Fang, J., Chacón, P., and Nogales, E. (2016). Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. Nature *531*, 604–609. https://doi.org/10.1038/nature17394.
8. Chen, X., Qi, Y., Wu, Z., Wang, X., Li, J., Zhao, D., Hou, H., Li, Y., Yu, Z., Liu, W., et al. (2021). Structural insights into preinitiation complex assembly on core promoters. Science *372*, eaba8490. https://doi.org/10.1126/science.aba8490.
9. Kamenova, I., Mukherjee, P., Conic, S., Mueller, F., El-Saafin, F., Bardot, P., Garnier, J.-M., Dembélé, D., Capponi, S., Timmers, H.T.M., et al. (2019). Co-translational assembly of mammalian nuclear multisubunit complexes. Nat. Commun. *10*, 1740. https://doi.org/10.1038/s41467-019-09749-y.
10. Bernardini, A., Mukherjee, P., Scheer, E., Kamenova, I., Antonova, S., Mendoza Sanchez, P.K., Yayli, G., Morlet, B., Timmers, H.T.M., and Tora, L. (2023). Hierarchical TAF1-dependent co-translational assembly of the basal transcription factor TFIID. Nat. Struct. Mol. Biol. *30*, 1141–1152. https://doi.org/10.1038/s41594-023-01026-3.
11. Trowitzsch, S., Viola, C., Scheer, E., Conic, S., Chavant, V., Fournier, M., Papai, G., Ebong, I.-O., Schaffitzel, C., Zou, J., et al. (2015). Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules. Nat. Commun. *6*, 6011. https://doi.org/10.1038/ncomms7011.
12. Gupta, K., Watson, A.A., Baptista, T., Scheer, E., Chambers, A.L., Koehler, C., Zou, J., Obong-Ebong, I., Kandiah, E., Temblador, A., et al. (2017). Architecture of TAF11/TAF13/TBP complex suggests novel regulation properties of general transcription factor TFIID. Elife *6*, e30395. https://doi.org/10.7554/elife.30395.
13. Fant, C.B., Levandowski, C.B., Gupta, K., Maas, Z.L., Moir, J., Rubin, J.D., Sawyer, A., Esbin, M.N., Rimel, J.K., Luyties, O., et al. (2020). TFIID Enables RNA Polymerase II Promoter-Proximal Pausing. Mol. Cell *78*, 785–793.e8. https://doi.org/10.1016/j.molcel.2020.03.008.
14. Levine, M., Cattoglio, C., and Tjian, R. (2014). Looping back to leap forward: transcription enters a new era. Cell *157*, 13–25. https://doi.org/10.1016/j.cell.2014.02.009.
15. Jacq, X., Brou, C., Lutz, Y., Davidson, I., Chambon, P., and Tora, L. (1994). Human TAFII30 is present in a distinct TFIID complex and is required for transcriptional activation by the estrogen receptor. Cell *79*, 107–117.
16. Kaufmann, J., Verrijzer, C.P., Shao, J., and Smale, S.T. (1996). CIF, an essential cofactor for TFIID-dependent initiator function. Gene Dev. *10*, 873–886. https://doi.org/10.1101/gad.10.7.873.
17. Martin, J., Halenbeck, R., and Kaufmann, J. (1999). Human Transcription Factor hTAFII150 (CIF150) Is Involved in Transcriptional Regulation of Cell Cycle Progression. Mol. Cell Biol. *19*, 5548–5556. https://doi.org/10.1128/mcb.19.8.5548.

18. Papai, G., Tripathi, M.K., Ruhlmann, C., Werten, S., Crucifix, C., Weil, P.A., and Schultz, P. (2009). Mapping the initiator binding Taf2 subunit in the structure of hydrated yeast TFIID. Structure (London, England : 1993) *17*, 363–373. https://doi.org/10.1016/j.str.2009.01.006.

19. Langer, D., Martianov, I., Alpern, D., Rhinn, M., Keime, C., Dollé, P., Mengus, G., and Davidson, I. (2016). Essential role of the TFIID subunit TAF4 in murine embryogenesis and embryonic stem cell differentiation. Nat. Commun. *7*, 11063–11116. https://doi.org/10.1038/ncomms11063.

20. Gura, M.A., Mikedis, M.M., Seymour, K.A., de Rooij, D.G., Page, D.C., and Freiman, R.N. (2020). Dynamic and regulated TAF gene expression during mouse embryonic germ cell development. PLoS Genet. *16*, e1008515. https://doi.org/10.1371/journal.pgen.1008515.

21. Herrera, F.J., Yamaguchi, T., Roelink, H., and Tjian, R. (2014). Core promoter factor TAF9B regulates neuronal gene expression. Elife *3*, e02559. https://doi.org/10.7554/elife.02559.

22. Liu, W.-L., Coleman, R.A., Grob, P., King, D.S., Florens, L., Washburn, M.P., Geles, K.G., Yang, J.L., Ramey, V., Nogales, E., and Tjian, R. (2008). Structural Changes in TAF4b-TFIID Correlate with Promoter Selectivity. Mol. Cell *29*, 81–91. https://doi.org/10.1016/j.molcel.2007.11.003.

23. Bardot, P., Vincent, S.D., Fournier, M., Hubaud, A., Joint, M., Tora, L., and Pourquié, O. (2017). The TAF10-containing TFIID and SAGA transcriptional complexes are dispensable for early somitogenesis in the mouse embryo. Development *144*, 3808–3818. https://doi.org/10.1242/dev.146902.

24. Voss, A.K., Thomas, T., Petrou, P., Anastassiadis, K., Schöler, H., and Gruss, P. (2000). Taube nuss is a novel gene essential for the survival of pluripotent cells of early mouse embryos. Development *127*, 5449–5461. https://doi.org/10.1242/dev.127.24.5449.

25. Mohan, W.S., Scheer, E., Wendling, O., Metzger, D., and Tora, L. (2003). TAF10 (TAFII30) Is Necessary for TFIID Stability and Early Embryogenesis in Mice. Mol. Cell Biol. *23*, 4307–4318. https://doi.org/10.1128/mcb.23.12.4307-4318.2003.

26. Gegonne, A., Tai, X., Zhang, J., Wu, G., Zhu, J., Yoshimoto, A., Hanson, J., Cultraro, C., Chen, Q.-R., Guinter, T., et al. (2012). The General Transcription Factor TAF7 Is Essential for Embryonic Development but Not Essential for the Survival or Differentiation of Mature T Cells. Mol. Cell Biol. *32*, 1984–1997. https://doi.org/10.1128/mcb.06305-11.

27. He, X.D., Phillips, S., Hioki, K., Majhi, P.D., Babbitt, C., Tremblay, K.D., Pobezinsky, L.A., and Mager, J. (2024). TATA-binding associated factors have distinct roles during early mammalian development. Dev. Biol. *511*, 53–62. https://doi.org/10.1016/j.ydbio.2024.04.002.

28. Indra, A.K., Mohan, W.S., Frontini, M., Scheer, E., Messaddeq, N., Metzger, D., and Tora, L. (2005). TAF10 is required for the establishment of skin barrier function in foetal, but not in adult mouse epidermis. Dev. Biol. *285*, 28–37. https://doi.org/10.1016/j.ydbio.2005.05.043.

29. Tatarakis, A., Margaritis, T., Martinez-Jimenez, C.P., Kouskouti, A., Mohan, W.S., Haroniti, A., Kafetzopoulos, D., Tora, L., and Talianidis, I. (2008). Dominant and redundant functions of TFIID involved in the regulation of hepatic genes. Mol. Cell *31*, 531–543. https://doi.org/10.1016/j.molcel.2008.07.013.

30. Ventura, A., Kirsch, D.G., McLaughlin, M.E., Tuveson, D.A., Grimm, J., Lintault, L., Newman, J., Reczek, E.E., Weissleder, R., and Jacks, T. (2007). Restoration of p53 function leads to tumour regression in vivo. Nature *445*, 661–665. https://doi.org/10.1038/nature05541.

31. Helmlinger, D., and Tora, L. (2017). Sharing the SAGA. Trends Biochem. Sci. *42*, 850–861. https://doi.org/10.1016/j.tibs.2017.09.001.

32. Soutoglou, E., Demény, M.A., Scheer, E., Fienga, G., Sassone-Corsi, P., and Tora, L. (2005). The nuclear import of TAF10 is regulated by one of its three histone fold domain-containing interaction partners. Mol. Cell Biol. *25*, 4092–4104. https://doi.org/10.1128/mcb.25.10.4092-4104.2005.

33. Fischer, V., Plassard, D., Ye, T., Reina-San-Martin, B., Stierle, M., Tora, L., and Devys, D. (2021). The related coactivator complexes SAGA and ATAC control embryonic stem cell self-renewal through acetyltransferase-independent mechanisms. Cell Rep. *36*, 109598. https://doi.org/10.1016/j.celrep.2021.109598.

34. Vannini, A., and Cramer, P. (2012). Conservation between the RNA polymerase I, II, and III transcription initiation machineries. Mol. Cell *45*, 439–446. https://doi.org/10.1016/j.molcel.2012.01.023.

35. Gillotin, S. (2018). Isolation of Chromatin-bound Proteins from Subcellular Fractions for Biochemical Analysis. Bio. Protoc. *8*, e3035. https://doi.org/10.21769/bioprotoc.3035.

36. 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. https://doi.org/10.1038/nrm.2017.10.

37. Perantoni, A.O., Timofeeva, O., Naillat, F., Richman, C., Pajni-Underwood, S., Wilson, C., Vainio, S., Dove, L.F., and Lewandoski, M. (2005). Inactivation of FGF8 in early mesoderm reveals an essential role in kidney development. Development *132*, 3859–3871. https://doi.org/10.1242/dev.01945.

38. Metzger, D., Scheer, E., Soldatov, A., and Tora, L. (1999). Mammalian TAF(II)30 is required for cell cycle progression and specific cellular differentiation programmes. EMBO J. *18*, 4823–4834. https://doi.org/10.1093/emboj/18.17.4823.

39. Papadopoulos, P., Gutiérrez, L., Demmers, J., Scheer, E., Pourfarzad, F., Papageorgiou, D.N., Karkoulia, E., Strouboulis, J., van de Werken, H.J.G., van der Linden, R., et al. (2015). TAF10 Interacts with the GATA1 Transcription Factor and Controls Mouse Erythropoiesis. Mol. Cell Biol. *35*, 2103–2118. https://doi.org/10.1128/mcb.01370-14.

40. Sun, M., Schwalb, B., Schulz, D., Pirkl, N., Etzold, S., Larivière, L., Maier, K.C., Seizl, M., Tresch, A., and Cramer, P. (2012). Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. Genome Res. *22*, 1350–1359. https://doi.org/10.1101/gr.130161.111.

41. Timmers, H.T.M., and Tora, L. (2018). Transcript Buffering: A Balancing Act between mRNA Synthesis and mRNA Degradation. Mol. Cell *72*, 10–17. https://doi.org/10.1016/j.molcel.2018.08.023.

42. Patange, S., Ball, D.A., Wan, Y., Karpova, T.S., Girvan, M., Levens, D., and Larson, D.R. (2022). MYC amplifies gene expression through global changes in transcription factor dynamics. Cell Rep. *38*, 110292. https://doi.org/10.1016/j.celrep.2021.110292.

43. Maston, G.A., Zhu, L.J., Chamberlain, L., Lin, L., Fang, M., and Green, M.R. (2012). Non-canonical TAF complexes regulate active promoters in human embryonic stem cells. Elife *1*, e00068. https://doi.org/10.7554/elife.00068.

44. El-Saafin, F., Curry, C., Ye, T., Garnier, J.-M., Kolb-Cheynel, I., Stierle, M., Downer, N.L., Dixon, M.P., Negroni, L., Berger, I., et al. (2018). Homozygous TAF8 mutation in a patient with intellectual disability results in undetectable TAF8 protein, but preserved RNA polymerase II transcription. Hum. Mol. Genet. *27*, 2171–2186. https://doi.org/10.1093/hmg/ddy126.

45. Sun, F., Sun, T., Kronenberg, M., Tan, X., Huang, C., and Carey, M.F. (2021). The Pol II preinitiation complex (PIC) influences Mediator binding but not promoter–enhancer looping. Gene Dev. *35*, 1175–1189. https://doi.org/10.1101/gad.348471.121.

46. Hawley, D.K., and Roeder, R.G. (1987). Functional steps in transcription initiation and reinitiation from the major late promoter in a HeLa nuclear extract. J. Biol. Chem. *262*, 3452–3461. https://doi.org/10.1016/s0021-9258(18)61372-9.

47. Zawel, L., Kumar, K.P., and Reinberg, D. (1995). Recycling of the general transcription factors during RNA polymerase II transcription. Gene Dev. *9*, 1479–1490. https://doi.org/10.1101/gad.9.12.1479.

48. Bernardini, A., Hollinger, C., Willgenss, D., Müller, F., Devys, D., and Tora, L. (2023). Transcription factor IID parks and drives preinitiation complexes at sharp or broad promoters. Trends Biochem. Sci. *48*, 839–848. https://doi.org/10.1016/j.tibs.2023.07.009.

49. Wieczorek, E., Brand, M., Jacq, X., and Tora, L. (1998). Function of TAFII-containing complex without TBP in transcription by RNA polymerase II. Nature *393*, 187–191. https://doi.org/10.1038/30283.

50. Demény, M.A., Soutoglou, E., Nagy, Z., Scheer, E., Jànoshàzi, A., Richardot, M., Argentini, M., Kessler, P., and Tora, L. (2007). Identification of a Small TAF Complex and Its Role in the Assembly of TAF-Containing Complexes. PLoS One *2*, e316. https://doi.org/10.1371/journal.pone.0000316.

51. Rengachari, S., Schilbach, S., Kaliyappan, T., Gouge, J., Zumer, K., Schwarz, J., Urlaub, H., Dienemann, C., Vannini, A., and Cramer, P. (2022). Structural basis of SNAPc-dependent snRNA transcription initiation by RNA polymerase II. Nat. Struct. Mol. Biol. *29*, 1159–1169. https://doi.org/10.1038/s41594-022-00857-w.

52. Isogai, Y., Keleş, S., Prestel, M., Hochheimer, A., and Tjian, R. (2007). Transcription of histone gene cluster by differential core-promoter factors. Genes Dev. *21*, 2936–2949. https://doi.org/10.1101/gad.1608807.

53. Parry, T.J., Theisen, J.W.M., Hsu, J.-Y., Wang, Y.-L., Corcoran, D.L., Eustice, M., Ohler, U., and Kadonaga, J.T. (2010). The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. Genes Dev. *24*, 2013–2018. https://doi.org/10.1101/gad.1951110.

54. Wang, Y.L., Duttke, S.H.C., Chen, K., Johnston, J., Kassavetis, G.A., Zeitlinger, J., and Kadonaga, J.T. (2014). TRF2, but not TBP, mediates the transcription of ribosomal protein genes. Genes Dev. *28*, 1550–1555. https://doi.org/10.1101/gad.245662.114.

55. Baumann, D.G., and Gilmour, D.S. (2017). A sequence-specific core promoter-binding transcription factor recruits TRF2 to coordinately transcribe ribosomal protein genes. Nucleic Acids Res. *45*, 10481–10491. https://doi.org/10.1093/nar/gkx676.

56. Serebreni, L., Pleyer, L.M., Haberle, V., Hendy, O., Vlasova, A., Loubiere, V., Nemčko, F., Bergauer, K., Roitinger, E., Mechtler, K., and Stark, A. (2023). Functionally distinct promoter classes initiate transcription via different mechanisms reflected in focused versus dispersed initiation patterns. EMBO J *42*, e113519. https://doi.org/10.15252/embj.2023113519.

57. Yu, C., Cvetesic, N., Hisler, V., Gupta, K., Ye, T., Gazdag, E., Negroni, L., Hajkova, P., Berger, I., Lenhard, B., et al. (2020). TBPL2/TFIIA complex establishes the maternal transcriptome through oocyte-specific promoter usage. Nat. Commun. *11*, 6439. https://doi.org/10.1038/s41467-020-20239-4.

58. Gazdag, E., Jacobi, U.G., van Kruijsbergen, I., Weeks, D.L., and Veenstra, G.J.C. (2016). Activation of a T-box-Otx2-Gsc gene network independent of TBP and TBP-related factors. Development (Cambridge, England) *143*, 1340–1350. https://doi.org/10.1242/dev.127936.

59. Santana, J.F., Collins, G.S., Parida, M., Luse, D.S., and Price, D.H. (2022). Differential dependencies of human RNA polymerase II promoters on TBP, TAF1, TFIIB and XPB. Nucleic Acids Res. *50*, 9127–9148. https://doi.org/10.1093/nar/gkac678.

60. Kwan, J.Z.J., Nguyen, T.F., Uzozie, A.C., Budzynski, M.A., Cui, J., Lee, J.M.C., Van Petegem, F., Lange, P.F., and Teves, S.S. (2023). RNA Polymerase II transcription independent of TBP in murine embryonic stem cells. Elife *12*, e83810. https://doi.org/10.7554/elife.83810.

61. Kokubo, T., Yamashita, S., Horikoshi, M., Roeder, R.G., and Nakatani, Y. (1994). Interaction between the N-terminal domain of the 230-kDa subunit and the TATA box-binding subunit of TFIID negatively regulates TATA-box binding. Proc National Acad Sci *91*, 3520–3524. https://doi.org/10.1073/pnas.91.9.3520.

62. Anandapadamanaban, M., Andresen, C., Helander, S., Ohyama, Y., Siponen, M.I., Lundström, P., Kokubo, T., Ikura, M., Moche, M., and Sunnerhagen, M. (2013). High-resolution structure of TBP with TAF1 reveals anchoring patterns in transcriptional regulation. Nat. Struct. Mol. Biol. *20*, 1008–1014. https://doi.org/10.1038/nsmb.2611.

63. Wong, K.M., Jepsen, W.M., Efthymiou, S., Salpietro, V., Sanchez-Castillo, M., Yip, J., Kriouile, Y., Diegmann, S., Dreha-Kulaczewski, S., Altmüller, J., et al. (2022). Mutations in TAF8 cause a neurodegenerative disorder. Brain *145*, 3022–3034. https://doi.org/10.1093/brain/awac154.

64. Janssen, B.D.E., van den Boogaard, M.J.H., Lichtenbelt, K., Seaby, E.G., Stals, K., Ellard, S., Newbury-Ecob, R., Dixit, A., Roht, L., Pajusalu, S., et al. (2022). De novo putative loss-of-function variants in TAF4 are associated with a neuro-developmental disorder. Hum. Mutat. *43*, 1844–1851. https://doi.org/10.1002/humu.24444.

65. Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. *30*, 207–210. https://doi.org/10.1093/nar/30.1.207.

66. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. *47*, D442–D450. https://doi.org/10.1093/nar/gky1106.

67. Martianov, I., Velt, A., Davidson, G., Choukrallah, M.-A., and Davidson, I. (2016). TRF2 is recruited to the pre-initiation complex as a testis-specific subunit of TFIIA/ALF to promote haploid cell gene expression. Sci. Rep. *6*, 32069. https://doi.org/10.1038/srep32069.

68. Brou, C., Chaudhary, S., Davidson, I., Lutz, Y., Wu, J., Egly, J.M., Tora, L., and Chambon, P. (1993). Distinct TFIID complexes mediate the effect of different transcriptional activators. EMBO J. *12*, 489–499. https://doi.org/10.1002/j.1460-2075.1993.tb05681.x.

69. Dubois, M.-F., Vincent, M., Vigneron, M., Adamczewski, J., Egly, J.-M., and Bensaude, O. (1997). Heat-shock inactivation of the TFIIH-associated kinase and change in the phosphorylation sites on the C-terminal domain of RNA polymerase II. Nucleic Acids Res. *25*, 694–700. https://doi.org/10.1093/nar/25.4.694.

70. Gyenis, Á., Umlauf, D., Újfaludi, Z., Boros, I., Ye, T., and Tora, L. (2014). UVB Induces a Genome-Wide Acting Negative Regulatory Mechanism That Operates at the Level of Transcription Initiation in Human Cells. PLoS Genet. *10*, e1004483. https://doi.org/10.1371/journal.pgen.1004483.

71. Chapman, R.D., Heidemann, M., Albert, T.K., Mailhammer, R., Flatley, A., Meisterernst, M., Kremmer, E., and Eick, D. (2007). Transcribing RNA Polymerase II Is Phosphorylated at CTD Residue Serine-7. Science *318*, 1780–1782. https://doi.org/10.1126/science.1145977.

72. Nagy, Z., Riss, A., Fujiyama, S., Krebs, A., Orpinell, M., Jansen, P., Cohen, A., Stunnenberg, H.G., Kato, S., and Tora, L. (2010). The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes. Cell. Mol. Life Sci. *67*, 611–628. https://doi.org/10.1007/s00018-009-0199-8.

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

74. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetJ *17*, 10–12. https://doi.org/10.14806/ej.17.1.200.

75. 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. https://doi.org/10.1093/nar/gkw257.

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

77. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675. https://doi.org/10.1038/nmeth.2089.

78. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589. https://doi.org/10.1016/j.molcel.2010.05.004.

79. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169. https://doi.org/10.1093/bioinformatics/btu638.

80. 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. https://doi.org/10.1186/gb-2008-9-9-r137.

81. 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. https://doi.org/10.1093/bioinformatics/bts635.

82. Meers, M.P., Bryson, T., and Henikoff, S. (2019). A streamlined protocol and analysis pipeline for CUT&RUN chromatin profiling. Preprint at bioRxiv. https://doi.org/10.1101/569129.

83. Smith, R., and Tostengard, A.R. (2020). Quantitative Evaluation of Ion Chromatogram Extraction Algorithms. J. Proteome Res. *19*, 1953–1964. https://doi.org/10.1021/acs.jproteome.9b00768.

84. Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106. https://doi.org/10.1186/gb-2010-11-10-r106.

85. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

86. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

87. Amemiya, H.M., Kundaje, A., and Boyle, A.P. (2019). The ENCODE Blacklist: Identification of Problematic Regions of the Genome. Sci. Rep. *9*, 9354. https://doi.org/10.1038/s41598-019-45839-z.

88. Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics *21*, 3439–3440. https://doi.org/10.1093/bioinformatics/bti525.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
rabbit polyclonal anti-TAF1 (WB, IP)	Abcam	Cat# ab264327
mouse monoclonal anti-TAF4 (WB)	in house	32TA-2B9; Mohan et al.[25]
mouse monoclonal anti-TAF5 (WB)	in house	1TA-1C2; Jacq et al.[15]
mouse monoclonal anti-TAF6 (WB)	in house	25TA-2G7; Langer et al.[19]
abbit polyclonal anti-TAF7 (WB, IF)	in house	19TA-2C7; Langer et al.[19]
rabbit polyclonal anti-TAF7 (WB, IF, CUT&RUN)	in house	3407; Bardot &Vincent et al. [23]
mouse monoclonal anti-TAF7L (WB)	in house	46TA-2D5; Martianov et al.[67]
mouse monoclonal anti-TAF8 (WB)	in house	3478; Bardot &Vincent et al. [23]
mouse monoclonal anti-TAF10 (WB, IF, CUT&RUN)	in house	6TA-2B11; Mohan et al.[25]
mouse monoclonal anti-TAF12 (WB, IP, CUT&RUN)	in house	22TA-2A1; Langer et al.[19]
mouse monoclonal anti-TBP (WB, IP)	in house	3TF1-3G3; Brou et al.[68]
rabbit polyclonal anti-TBP (CUT&RUN)	Abcam	Cat# ab2817
mouse monoclonal anti-SUPT7L (WB)	Bethyl laboratories	Cat# A302-803A
mouse monoclonal anti-RPB1(CTD) (WB)	in house	1PB-7C2; Dubois et al.[69]
mouse monoclonal anti-RPB1(CTD) (ChIP-seq)	in house	CTD-7G5; Gyenis et al.[70]
rat monoclonal anti-RPB1pSer5 (IF, WB)	GmbH antibody service	CTD4-3E8; Chapman et al.[71]
rat monoclonal anti-RPB1pSer2 (IF, WB)	GmbH antibody service	CTD4-3E10; Chapman et al.[71]
mouse monoclonal anti-GST (IP)	in house	15TF2-1D10; Nagy et al.[72]
Alexa Fluor® 488-labelled goat anti-rabbit IgG	Life Technologies	Cat# A-11008; RRID: AB_143165
Alexa Fluor® 546-labelled goat anti-mouse IgG	Life Technologies	Cat# A-11003; RRID: AB_2534071
Alexa Fluor® 488-labelled goat anti-rat IgG	Life Technologies	Cat# A-11006; RRID: AB_2534074
Peroxydase AffiniPure™ goat anti-rabbit IgG(H+L)	Jackson ImmunoResearch	Cat# 111-035- 144; RRID: AB_2307391
Peroxydase AffiniPure™ $F(ab')_2$ Fragment goat anti-mouse IgG, Fc_γ fragment specific	Jackson ImmunoResearch	Cat# 111-036-071; RRID: AB_2338524
Chemicals, peptides, and recombinant proteins		
4′,6-diamidino-2-phenylindole dihydrochloride (DAPI)	Molecular Probes	Cat# D1306
4-Hydroxytamoxifen (4-OHT)	Sigma Aldrich	Cat# H7904
4-Thiouridine (4-sU)	Glentham Life Science	Cat# GN6085
5-Ethynyl Uridine (5-EU)	Thermo Fisher Scientific	Cat# E13045
BlueTrypan Staining 0.4%	Invitrogen T10288	Cat# T10288
CHIR99021	Axon Medchem	Cat# 1386
Crystal Violet	Sigma Aldrich	Cat# C3886-25G
Complete Protease Inhibitor Cocktail (cOmplete), EDTA free	Roche	Cat# 11873580001
EZ-link™ HPDP-Biotin	Thermo Fisher Scientific	Cat# 21341
Leukemia inhibitory factor (LIF)	in house	N/A
Protein G Sepharose	GE healthcare	Cat# 17-0618-05
Dynabeads™ Protein G	Invitrogen	Cat# 10003D
Random hexamer primer	Thermo Fisher Scientific	Cat# SO142
RNasin	Promega	Cat# N2111
RiboPure – Yeast Kit	Invitrogen	Cat# AM1926
TRI® Reagent (Trizol)	Molecular Research Center Inc.	Cat# TR188
VECTASHIELD® Mounting Media without DAPI	VectorLabs	Cat# H-1000

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
AMPure XP beads	Beckman-Coulter	Cat# A63882
Click-it RNA Imaging Kits	Invitrogen	Cat# C10329
EdU Staining Proliferation kit	Abcam	Cat# ac222421
LightCycler® 480 SYBR® Green 2x PCR Master Mix I	Roche	Cat# 4887352001
MicroPlex library preparation kit v3	Diagenode	Cat# C05010001
NucleoSpin RNA XS, RNA extraction kit	Machery-Nagel	Cat# 740902.50
Pierce™ ECL Western Blotting Substrate	Thermo Fisher Scientific	Cat# 32109
QuantiTect Reverse Transcription Kit	Qiagen	Cat# 205311
Protein Assay Dye Reagent Concentrate	Bio-Rad	Cat# 5000006
RiboPure – Yeast Kit	Invitrogen	Cat# AM1926
SPRIselect beads	Beckman-Coulter	Cat# B23319
SuperScript IV Reverse Transcriptase	Invitrogen	Cat# 18090050
TruSeq Stranded Total RNA LT Sample Prep Kit with Ribo-Zero Gold	Illumina	Cat# RS-122-2301
μMACS Streptavidin Kit	Miltenyi Biotec	Cat# 130-074-101
Deposited data		
Mass-spectrometry proteomics TFIID complex anti-TAF12 IPs from control and TAF7 or TAF10 depleted mESCs	This study	PRIDE: PXD046459
Mass-spectrometry proteomics TFIID complex anti-TBP IPs from control and TAF7 or TAF10 depleted mESCs	This study	PRIDE: PXD046459
Nascent transcriptomic data from control and TAF7 or TAF10 depleted mESCs	This study	GEO: GSE245196
CUT&RUN from control and TAF7 or TAF10 depleted mESCs	This study	GEO: GSE245196
Anti-Pol II ChIP-seq from control and TAF7 or TAF10 depleted mESCs	This study	GEO: GSE245196
Western blots and cell counting raw data	This study	Mendeley Data: https://doi.org/10.17632/m66wgvw7sm.2
Experimental models: Cell lines		
R26$^{CreERT2/+}$;*Taf7*$^{f/f}$mouse ES cells	This study	N/A
R26$^{CreERT2/+}$;*Taf10*$^{f/f}$mouse ES cells	Bardot & Vincent et al.[23]	N/A
R26$^{CreERT2/+}$;*Taf7*$^{f/f}$;*Sup7l*$^{-/-}$mouse ES cells	This study	N/A
E14tg2a.4 mouse ES cells	BayGenomics	N/A
CD1 WT mouse fibroblasts	IGBMC PluriCell East platform	N/A
S2 Drosophila cells	IGBMC PluriCell East platform	N/A
Experimental models: Organisms/strains		
Tg(*T-Cre*)	Perantoni et al.[37]	N/A
R26^{CreERT2}	Ventura et al.[30]	MGI: 3699244
Taf10^{f}	Mohan et al.[25]	MGI: 3606185
Taf7^{f}	Gegone et al.[26]	MGI: 5430373
Tg(*T-Cre*/+);*Taf7*$^{f/f}$	This study	N/A
Tg(*T-Cre*/+);*Taf10*$^{f/f}$	Bardot & Vincent et al.[23]	N/A
Oligonucleotides		
For sgRNA sequences, see Table S1	This study	N/A
For primer sequences, see Table S2	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Adobe Illustrator 28.5	Adobe	https://www.adobe.com
Bowtie2	Langemead et al.[73]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Cutadapt	Martin et al.[74]	https://cutadapt.readthedocs.io/en/v1.10/
deepTools	Ramirez et al.[75]	https://github.com/deeptools/deepTools
DESeq2	Love et al.[76]	https://bioconductor.org/packages/
ImageJ	Schneider et al.[77]	https://imagej.net/software/imagej/
Image Lab Software	Bio Rad	https://www.bio-rad.com/fr-fr/product/image-lab-software?ID=KRE6P5E8Z
Homer	Heinz et al.[78]	http://homer.ucsd.edu/homer/
htseq-count	Anders et al.[79]	https://htseq.readthedocs.io/en/master/
MACS2	Zhang et al.[80]	https://hbctraining.github.io/
Proteome Discoverer 2.2	Thermo Fisher Scientific	https://www.thermofisher.com/fr/fr/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-software/multi-omics-data-analysis/proteome-discoverer-software.html
R	R-project	https://CRAN.R-project.org/
RStudio	RStudio	https://www.rstudio.com/categories/rstudio-ide/
STAR	Dobin et al.[81]	https://github.com/alexdobin/STAR

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mice

Animal experimentation was carried out according to animal welfare regulation and guidelines of the French Ministry of Agriculture and Ministry of Higher Education, Research, and Innovation. The original mouse lines (*Taf7*^{f}, *Taf10*^{f}, *R26*^{CreERT2}, Tg(*T-Cre*)) have already been described[25,26,30,37]. Adult mice (8-weeks old) were bred to produce E9.5 to E12.5 embryos. The sex of the embryos was not determined.

Generation and maintenance of mESCs

E3.5 blastocysts were collected from *R26*$^{CreERT2/+}$;*Taf7*$^{f/f}$ X *Taf7*$^{f/f}$ or *R26*$^{CreERT2/+}$;*Taf10*$^{f/f}$ X *Taf10*$^{f/f}$ mating. Uteruses were collected, embryos were flushed out with M2 medium (37°C) and placed in 96 well plates coated with mouse embryonic fibroblast (MEF) feeders in 2i+LIF medium (DMEM medium supplemented with 15% fetal calf serum ES-tested, 2 mM L-glutamine, 0.1% ß-mercaptoethanol, 100 UI/ ml penicillin and 100 mg/ml streptomycin, 0.1 mM non-essential amino acids, 100 μL/50 mL of leukemia inhibitory factor (LIF), 3 μM CHIR99021 and 1 mM PD0325901) at 37°C under 5% CO_2. For initial amplification, mESCs were maintained on feeders until frozen in DMEM medium supplemented with 30% fetal calf serum and 20% DMSO. For experiment, mESCs were grow on gelatin. All the clones established were mycoplasma-free and were used for experiments before passage 35.

R26$^{CreERT2/+}$;*Taf7*$^{f/f}$;*Supt7l*$^{-/-}$ mESCs were generated from the clone *RT7#13*. These cells were transfected at 70% confluence with plasmid constructs containing *Cas9-EGFP* and gRNA (Table S1) using Lipofectamine 2000 kit. Two days after, single EGFP^{+} cells were isolated in 96 well plates using the BD FACS Aria TM II (BD Biosciences), amplified and frozen.

Mouse ES E14 cells (BayGenomics) were cultured on plates coated with 0.1% gelatine solution in 1x PBS (Dutcher, Cat# P06-20410) using DMEM medium supplemented with 15% foetal calf serum ES-tested (Thermo Fisher Scientific, Cat# 10270-106), 2 mM L-glutamine (Thermo Fisher Scientific, Cat# 25030-024), 0.1% β-mercaptoethanol (Thermo Fisher Scientific, Cat#31350-010), 100 UI/mL penicillin and 100 μg/mL streptomycin (Thermo Fisher Scientific, Cat# 15140-122), 0.1 mM non-essential amino acids (Thermo Fisher Scientific, Cat# 11140-035) and 1,500 U/mL home made LIF, 3 μM CHIR99021 (axon medchem, Cat# 1386) and 1 μM PD0325901 (axon medchem, Cat#1408). Cells were grown at 37°C with 5% CO_2 levels. Cells were passaged every second day.

Drosophila melanogaster Schneider S2 cells were grown in Schneider's Drosophila Medium (Invitrogen Cat#21720-24) supplemented with 10% fetal calf serum (Sigma Amdrich Cat# F7524), 50 UI/mL penicillin and 50 μg/ml streptomycin (Thermo Fisher Scientific, Cat# 15140-122) at 22°C in normal air conditions.

METHOD DETAILS

Collection of mouse embryos

Embryos were collected in PBS, fixed in 4% PFA/PBS for 1 hour at 4°C under agitation and rinsed three times in PBS. The embryos were imaged using a Leica MZ16 macroscope coupled to a CoolSnap-Pro color camera (RS Photometrics).

Calculation of the somite index

Somite index was calculated by subtracting the mean of the somite number of a litter from the somite number of each littermate.

Embryo sectioning and immunolocalization

Fixed embryos were equilibrated in 30% sucrose/PBS (3 h, 4°C), embedded in Cryomatrix (Thermo Fischer) and stored at -80°C. Fifteen micrometers-sections were cut on a Leica cryostat.

Sections were rehydrated in PBS, permeabilized in 0.5% Triton X-100/PBS (Sigma-Aldrich) for 40 min at RT (room temperature), blocked in AB buffer (3% Bovine Serum Albumin (BSA), 1% goat serum, 0.1% Tween 20 in PBS) for 40 min at RT and rewashed in 0.1% Triton X-100/PBS. Primary antibodies listed in key resources table were diluted 1/1000 in AB buffer and incubated overnight at 4°C. Sections were then washed three times in 0.1% Triton X-100/PBS for 40 min each. Secondary antibody (key resources table) was diluted 1/1000 in 1 μg/mL DAPI (4′,6-diamidino-2-phenylindole dihydrochloride)/AB buffer and incubated for 1 h at RT. The sections were then washed in 0.1%Triton X-100/PBS several times. The slides were mounted in Vectashield® and imaged with a TCS SP5 (Leica) laser-scanning microscope (20x Plan APO objective). The pictures are shown with the LUT "Green Fire Blue" scale.

Genotyping and screening

Mouse tail tips were digested in 300 μL of 200 μg/mL of proteinase K in tail digestion buffer (10 mM Tris pH 7, 200 mM NaCl, 5 μM EDTA, 0.2% SDS). Embryonic yolk sacs and mESCs were digested in 100 μL of 200 μg/mL of proteinase K in 1X PCR buffer. 0.6 μL of lysate was used in 25 μL of PCR reaction with Taq DNA polymerase kit. Primers are listed in Table S2.

4-Hydroxytamoxifen (4-OHT) treatment

$5x10^4$ or $8x10^5$ cells were seeded in 6 well plates or in one P100 petri dish, respectively, at day -1 (D-1). At D0, cells were treated with 100 nM of 4-OHT in a final volume of 2 mL or 12 mL, respectively. Control cells were treated with ethanol (0.1% EtOH final) in the same final volume. The experiments were performed at D2, D4 or D6. The 6 well plates were used for phenotypic analyses, and the P100 petri dishes for proteomic and transcriptomic analyses.

mESC phenotypic analysis

Cell counting

Trypsinized cells were resuspended in a dozen to hundreds of μL of PBS according to the size of the cell pellet. Cells were stained with Trypan Blue and counted on a Countess II Automated Cell Counter (Thermo Fischer Scientific). Different clones were used: *RT7*: 5 clones (#5, #7, #8, #13, #15), *RT10*: 5 clones (#3, #6, #9, #15, #41), *RT7; Supt7l*$^{-/-}$: 3 clones (#45, #71, #72).

Apoptosis assay

Dead floating and attached cells were collected and stained with the FITC Annexin V Apoptosis Detection Kit and counterstained with propidium iodide PI (Thermo Fischer Scientific) and analyzed using BD FACS Celesta (BD Bioscience). A minimum of 10,000 events were recorded. Different independent clones were used: *RT7*: 4 clones (#5, #7, #8, #13), *RT10*: 6 clones (#3, #6, #9, #15, #19, #41), *RT7; Supt7l*$^{-/-}$: 3 clones (#3, #71, #72).

Measuring of surface area of colonies

Cells were rinsed with PBS, fixed with 4% PFA for 30 min at 4°C, washed with PBS, stained for 30 min with 0.1% crystal violet dye and rewashed with PBS. Pictures were taken with a macroscope M420 (Leica) coupled with a CoolSNAP camera (RS Photometrics). Different independent clones were used: *RT7*: 2 clones (#5, #13), *RT10*: 3 clones (#3, #6, #9).

Cell cycle analysis by EdU

Cells were plated on gelatinized round glass slides. They were incubated with 10 μM of 5-ethynyl-2'-deoxyuridine (EdU, Thermo Fischer Scientific) for 3 h and then fixed with 4% PFA for 15 min at RT. They were then stained with the EdU Staining Proliferation Kit (Abcam, Cat# ab222421). Slides were mounted, cells were imaged with a TCS SP5 (Leica) laser-scanning microscope (20x Plan APO objective).

Depletion analysis

Whole cell extract

Cells were pelleted, rinsed, resuspended in one equivalent volume of WCE buffer (50 mM Tris HCl pH7.9, 25% Glycerol, 0.2mM EDTA, 5 mM $MgCl_2$, 600 mM KCl, 0.5% NP40, 1 mM DTT, 1X cOmplete) and incubated for 30 min on ice. Then, 3 volumes of IP0 (25mM Tris HCl pH7.9, 5% Glycerol, 0.1% NP40, 1 mM DTT, 1X PIC) were added and incubated for 30 min on ice. Proteins presented in the supernatant were recovered after high-speed centrifugation.

Cytoplasmic, nuclear, chromatin extract

This method was previously published in [35]. One volume of cell pellet is lysed in 2 volumes of ice-cold E1 buffer (50 mM HEPES-KOH, pH7.5, 140 mM NaCl, 1 mM EDTA, pH8, 10% glycerol, 0.5% NP40, 0.25% Triton-X-100, 1 mM DTT, 1X cOmplete protease inhibitor (PIC)) by up-and-down pipetting. The suspension is centrifugated at 1100 g at 4°C for 2min. The supernatant is collected as the cytoplasmic extract (CE). The pellet is resuspended and incubated 10 min on 5 volumes of ice-cold E1 buffer. After 2 min of centrifugation at 1100 g, the pellet is resuspended in 1 volume of ice-cold E2 buffer (10 mM Tris-HCl pH8.0, 200 mM NaCl, 1 mM EDTA pH8.0, 1 mM EGTA pH8.0, 1xPIC). The suspension is centrifugated at 1100 g at 4°C for 2 min. The supernatant is collected as the nuclear extract (NE). The pellet is resuspended and incubated 10 min on 2 volumes of ice-cold E2 buffer. After 2 min of centrifugation at 1100 g, the pellet is resuspended in 1 volume of ice-cold E3 buffer (50 mM Tris-HCl pH6.8, 20 mM NaCl, 1 mM MgCl2, 1% NP-40, 1xPIC) and transferred in a new clean tube. 1/1000 of benzonase (Sigma, Cat# E1014) is added and chromatin digestion is achieved during 30 min at RT under agitation. The suspension is centrifugated at 16000 g at 4°C for 10 min. The supernatant is collected as the chromatin extract (ChrE).

Western blot (WB)

Protein concentrations were measured by Bradford method. Twenty μg of WCE or 15 μg of CE, NE or ChrE were boiled for 5 min in 100 mM Tris HCl pH 6.8, 30% glycerol, 4% SDS, 0.2% Bromophenol Blue, 100 mM DTT, resolved in 10% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane (Protran, Amersham). After blocking in 3% milk/PBS, primary antibodies diluted 1/1000 in 0.3% milk/PBS (key resources table) were incubated overnight at 4°C. After 3 washes in 0.05% Tween 20/PBS, HRP-coupled secondary antibodies (key resources table) diluted 1/10000 in 0.3% milk/PBS were incubated 2 h at RT followed by ECL detection (Thermo Fisher Scientific) in a ChemiDoc Touch Imaging System (Bio Ras). Western blot images were processed on Bio Rad Image Lab Software (version 5.2.1).

Immunofluorescence (IF) on mESC

Cells plated on gelatinized round glass slides were rinsed, fixed with 4% PFA/PBS for 10 min at RT, rinsed again and then permeabilized in 0.1% Triton X-100/PBS for 20 min at RT. After PBS washes, primary antibodies (key resources table) diluted in 10% fetal calf serum (FCS)/PBS were incubated for 1 h at RT. Cells were rinsed twice with 0.02% Triton X-100/PBS for 5 min. Secondary antibodies (key resources table) were diluted 1/1000 in 1 μg/mL DAPI/10% FCS/PBS and incubated for 1 h at RT. Cells were then washed twice for 5 min in 0.02% Triton X-100/PBS. Slides were mounted, cells were imaged with a TCS SP5 (Leica) laser-scanning microscope (20x Plan APO objective).

RT-qPCR

RNAs were extracted using TRI® Reagent (Molecular Research Center Inc), precipitated in isopropanol, washed with 75% EtOH, resuspended in RNase-free water and quantified with a Nanodrop (Thermo Fisher Scientific). For RNA extraction from testis, tissues were stocked with a B pestle in a glass dounce grinder (Kimble) followed by high-speed centrifugation.

Reverse Transcription (RT) was performed using with 1 μg of total RNA using QuantiTect Reverse Transcription Kit (Qiagen) in T100 Bio-Rad machine. For qPCR, cDNAs were diluted 5 times and amplified using LightCycler 480 SYBR Green 2x PCR Master Mix I with 0.6 mM of forward and reverse primers (Sigma Aldrich, Table S2) in 8 μL of reaction volume. qPCR reaction was realized using a LightCycler 480 machine (Roche). Normalized values correspond to $(1+\text{efficiency})^{-\Delta\Delta Ct}$.

TFIID complex composition analysis

Nuclear enriched whole cell extracts

Cells (RT7#13 or RT10#41) were pelleted, rinsed, and resuspended in one volume of ice-cold Hypotonic Buffer (1 mL for 1 g of cells) (10 mM Tris pH8, 1.5 mM $MgCl_2$, 10 mM KCl supplemented with cOmplete protease inhibitor mix 1x (Roche). Cells were then lysed by 10 gentle strokes with a B pestle in a glass Dounce grinder (Kimble). After 10 min of centrifugation at 9000 g, the pellet, which contains nuclei, was resuspended in one volume of High Salt Buffer (20 mM Tris pH8, 1.5 mM $MgCl_2$, 450 mM NaCl, 0.2 mM EDTA, 25% glycerol, 0.5% NP40, supplemented with cOmplete protease inhibitor mix 1x). Nuclei were lysed by 10 gentle strokes and incubated on ice for 30 min. After 10 min of centrifugation at 9000 g, the supernatant was recovered as nuclear enriched whole cell extract (NWCE).

IP-MS analysis

Immunoprecipitation (IP). One mg of NWCE was first incubated for 1 h with 120 μL of Protein-G Sepharose beads (GE healthcare) in 1 mL of IP100 buffer (25 mM Tris HCl pH7.9, 10% Glycerol, 0.1% NP40, 5 mM $MgCl_2$, 100 mM KCl, 1X cOmplete) at 4°C under gentle agitation. NWCE was isolated and incubated with 10 to 30 μL of antibodies (key resources table) during 2 h at 4°C under gentle agitation and then incubated with fresh 120 μL of Protein-G Sepharose beads overnight. Beads were then washed at 4°C twice with IP500 buffer (25 mM Tris HCl pH 7.9, 10% Glycerol, 0.1% NP40, 5 mM $MgCl_2$, 500 mM KCl, 1X cOmplete) under gentle agitation and then three times with IP100 buffer, each time for 10 min. Immunoprecipitated proteins were eluted with 50 μL of acid Glycine buffer (0.1 M glycine pH2.8) directly buffered with 1 μL of 10 mM Tris HCl pH8. Each immunoprecipitation was verified by western blot by loading 20 μL of input and of supernatant and 15 μL of eluted proteins.

Liquid digestion. Eluted proteins were TCA-precipitated overnight at 4°C. Samples were then centrifuged at 16000g for 30 min at 4°C. Pellets were washed twice with 500 μL cold acetone, centrifuged at 16000g for 10 min at 4°C, denatured with 8 M urea in Tris-HCl 0.1 mM, reduced with 5 mM TCEP for 30 min, then alkylated with 10 mM iodoacetamide for 30 minutes in the dark. Both reduction and alkylation were performed at room temperature and under agitation (100g). Double digestion was performed with endoproteinase Lys-C (Wako) at a ratio of 1/100 (enzyme/proteins) in 8 M urea for 4 h, followed by an overnight modified trypsin

digestion (Promega) at a ratio of 1/100 (enzyme/proteins) in 2 M urea. Both LysC and Trypsin digestions were performed at 37°C. Peptide mixtures were then desalted on C18 spin-columns and dried on Speed-Vacuum before LC-MS/MS analysis.

LC-MS/MS analysis. Each sample was analyzed in triplicate (experimental triplicate) using an Ultimate 3000 nano-RSLC (Thermo Scientific, San Jose California) coupled in line with a LTQ-Orbitrap ELITE mass spectrometer via a nano-electrospray ionization source (Thermo Scientific, San Jose California).

Peptide mixtures were loaded on a C18 Acclaim PepMap100 trap-column (75 μm ID x 2 cm, 3 μm, 100Å, Thermo Fisher Scientific) for 3.5 minutes at 5 μL/min with 2% ACN, 0.1% FA in H_2O and then separated on a C18 Accucore nano-column (75 μm ID x 50 cm, 2.6 μm, 150Å, Thermo Fisher Scientific) with a 100 min linear gradient from 5% to 50% buffer B (A: 0.1% FA in H_2O / B: 99% ACN, 0.1% FA in H_2O), then a 20 min linear gradient from 50% to 70% buffer B, followed with 10 min at 99% B and 10 min of regeneration at 5% B. The total duration was set to 140 min at a flow rate of 200nL/min. The oven temperature was kept constant at 40°C.

The mass spectrometer was operated in positive ionization mode, in data-dependent mode with survey scans from m/z 300-1600 acquired in the Orbitrap at a resolution of 240,000 at m/z 400. The 20 most intense peaks (TOP20) from survey scans were selected for further fragmentation in the Linear Ion Trap with an isolation window of 2.0 Da and were fragmented by CID with normalized collision energy of 35%. Unassigned and single charged states were rejected.

The Ion Target Value for the survey scans (in the Orbitrap) and the MS2 mode (in the Linear Ion Trap) were set to 1E6 and 5E3 respectively and the maximum injection time was set to 100 ms for both scan modes. Dynamic exclusion was used. Exclusion duration was set to 30 s, repeat count was set to 1 and exclusion mass width was ± 10 ppm.

Gel filtration (GF)

One mg of NWCE was diluted twice in the GF buffer (20 mM Tris pH8, 1.5 mM $MgCl_2$, 450 mM NaCl, 0.2 mM EDTA,10% glycerol, 0.5% NP40), centrifuged for 10 min at 16000 rpm then passed at 0.4 mL/min through a Superose 6 GL 10/300 column (Sigma Aldrich) previously equilibrated with GF buffer. About 60 fractions of 250 μL were collected and 27 μL of each fraction were analyzed by western blot.

Estimation of remaining holo-TFIID in the IP-MS and gel filtration

IP-MS. In the TAF7 depleted cells, the detection of TAF7 in the 2 IPs is a direct indication of the remaining holo-TFIID as TAF7 does not interact directly with TAF12 or TBP. To estimate the fraction of remaining holo-TFIID in the TAF7-depleted cells, we used the normalized XIC values of TAF7.

TAF10 is present in different intermediates of the TFIID assembly but is also part of the SAGA complex. As the anti-TAF12 IP-MS detect all the intermediates from the core-TFIID to the holo-TFIID, the remaining fraction of holo-TFIID in the TAF10-depleted cells in this experiment was assessed by averaging the normalized XIC values of the S-TAF subunits (TAF1, TAF7, TAF11, TAF13 and TBP), as the S-TAF does not directly interact with TAF12. To estimate the remaining fraction of holo-TFIID in the anti-TBP IP-MS, we used the average of the normalized XIC values of the core-TFIID complex subunits, as it does not directly interact with TBP. These estimations were performed from day 0 to day 4 and plotted as a regression curve using ggplot2 version 3.4.0.

Gel filtration. As the intensity on the blots were not identical between the different conditions, we normalized the data. For TAF10-depleted cells, for each protein, we normalized the signal intensity per fraction to the total signal intensity over all the fractions. We used the added normalized signal intensities of TAF4, TAF5 and TAF6 in control and TAF10-depleted fractions to calculate the ratio of the signal in depleted over control conditions in each fraction. As holo-TFIID is contained in the high molecular weight fractions, the values of fractions B to D were used as a proxy for the quantification of holo-TFIID.

For TAF7 depleted cells, as there is no major difference between the control and the *-Taf7* condition, except for the depletion of TAF7, we only focused on fractions B to D where the majority of the signal is present. Similarly, we normalized the signal intensity of fraction B, C or D to the total signal intensity over all fractions B to D. We then used the added normalized signal intensities of TAF4, TAF5 and TAF6 in control and TAF7 depleted fractions to normalize the signal intensity detected for TAF7, in each fraction and each condition. We used the normalized intensity of TAF7 in fractions B to D as a proxy of the remaining holo-TFIID in depleted cells.

Transcription analysis

Analysis of newly synthesized RNA by EU labelling

Cells (RT7#13 or RT10#41) plated on gelatinized round glass slides were incubated with 1 mM of 5-ethynyl-uridine (EU, Thermo Fisher Scientific, E10345) for 1 h, fixed with 4% PFA for 15 min at RT and rinsed. They were then stained using the Click-it RNA Imaging Kits (Invitrogen, C10329) according to the manufacturer's guidelines. Cells were imaged with a TCS SP5 (Leica) laser-scanning microscope (20x Plan APO objective).

Newly synthesized 4-sU RNA sequencing

4-sU RNA labeling and purification. The protocol for newly synthesized RNA sequencing is based on published protocols [33]. Briefly, mESCs were seeded at D-1, treated with either EtOH (control) or 4-OHT (mutant) from D0 to D3 with a reseeding/amplification step at D2. On D3, cells were labelled with 500 μM 4-sU for 15 min. In parallel, drosophila S2 cells were also labelled with 500 μM 4-sU for 15 min. RNAs were extracted with Trizol, precipitated with isopropanol, washed with 75% ethanol, and resuspended in DEPC-treated water. DNAse treatment was performed using the TURBO DNA-free kit. Non labelled *S.cerevisiae* RNAs were isolated using the RiboPure - Yeast kit (Invitrogen). RNAs were measured with the Qbit machine using the Quant-it RNA Broad Range kit. 200 μg of mESC RNAs were mixed with 25 μg of S2 RNAs and 25 μg of yeast RNAs. The mixture was precipitated again, resuspended in 130 μL of DEPC-treated water and then fragmented with Covaris E220 sonicator. One μL of RNA were collected before and after

sonication to check fragmentation on a 1% agarose gel and on bioanalyzer. The fragmented RNA is expected to have an average size of 1.5 kb. To isolate newly synthesized 4-sU labelled RNA fragments, RNAs were first biotinylated with Biotin-HPDP molecules, then combined with streptavidin magnetic beads and finally isolated on column (μMACS streptavidin beads and kit, Miltenyi). RNAs from flowthrough and elution are then precipitated in absolute EtOH with 0.1 mg/mL glycogen and 300 mM NaOAc (pH 5.2) overnight at -20°C, washed with 75% EtOH and resuspended in respectively 150 μL and 15 μL of DEPC-treated water.

The quality of 4-sU labelled RNA fragment purification was checked by RT-qPCR prior sequencing. Two μL of eluted RNAs and 7.5 μL of RNAs from the flowthrough were collected into respectively 12 μL and 2.5 μL of DEPC-treated water. RT was carried out using Superscript IV kit and qPCR by 480 SYBR green I Master kit. Primers are listed in Table S2. Good purification is characterized by a lower amount of yeast RNA and a higher amount of mouse intronic sequence in the purified sample than in the flowthrough sample after normalization to Drosophila RNAs (ΔCt) and RNA from EtOH flowthrough sample (-ΔΔCt).

Library preparation. Fifteen to 50 ng of eluted RNA was used for the library preparation using TruSeq Stranded Total RNA LT Sample Prep Gold Kit (Illumina, RS-122-230) according to the Illumina protocol with the following modifications. Four-thiouridine-labelled RNA was cleaned up using 1.8× RNA Clean AMPure XP beads (A63882, Beckman-Coulter) and fragmented using divalent cations at 94°C for 1 min without depletion of rRNA. While double stranded cDNA synthesis and adapter ligation were performed according to manufacturer instructions, the number of PCR cycles for library amplification was reduced to 10 cycles. After purification using SPRIselect beads (B23319, Beckman-Coulter), libraries were sequenced with 1×50 bp on a HiSeq4000 System (Illumina).

Cleavage under targets and release using nuclease (CUT&RUN) and library preparation

CUT&RUN experiments were performed in biological duplicate as described in [82]. Briefly, mESC (RT7#13 or RT10#41) were treated with EtOH or 4-OHT for 3 days and 250,000 cells were used per CUT&RUN sample. Cells were washed and resuspended in wash buffer and incubated for 10 min at room temperature with 10 μL of concanavalin A-coated beads (Bangs Laboratories, BP531). Cells bound to beads were permeabilized using 0.05% digitonin and incubated with the appropriate antibody at 4°C overnight. Protein A-MNAse (pA-MN) was added to a final concentration of 700 ng/mL and incubated at 4°C for 1 h on a tube rotator. After washing, pA-MN was activated with 2 μL of 100 mM CaCl2 and digestion was performed for 30 min at 0°C. The reaction was stopped with 100 μL of stop buffer containing 2 pg/mL of heterologous spike-in DNA from yeast. Release of the DNA fragments was achieved by incubating samples at 37°C during 15 min. DNA was extracted using NucleoSpin columns (Macherey-Nagel) and eluted in 30 μL of NE buffer.

CUT&RUN-seq libraries were generated using the MicroPlex library preparation kit v3 (Diagenode), following the manufacturer's instructions, except that the stage 4 of the library amplification PCR was performed with a combined annealing-extension step for 10 s at 60°C, and stage 5 with an extension step for 10 s at 60°C for 7 cycles. Yield and size distribution were quantified on a 2100 Bioanalyzer instrument (Agilent). Sequencing was performed by the GenomEast platform (IGBMC) on an Illumina NextSeq 2000 (PE-50, 20 million reads).

Pol II ChIP-seq

Duplicates of RT10 mESCs were treated with EtOH or 4-OHT for 3 days, then cells were cross-linked with 1% formaldehyde for 10 min and quenched with 125 mM glycine for 5 min at room temperature with gentle shaking. Cells were quickly rinsed in cold PBS twice then scraped on ice and centrifuged at 4°C at 1250 g for 3 min. After a rinse in cold PBS, cells were centrifuged at 4°C at 1250 g for 3 min and snap-frozen in liquid nitrogen. Cell pellets were thawed on ice and resuspended in 1 mL lysis buffer 1 (50 mM HEPES–KOH pH 7.5 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP40, 0.25% Triton X-100) supplemented with protease inhibitors and incubated at 4°C on a rocker for 10 min. Lysates were centrifuged at 100 g at 4°C for 5 min. Pellets were resuspended with 1 mL lysis buffer 2 (10 mM Tris pH 8.0 1 mM EDTA 0.5 mM EGTA 200 mM NaCl) supplemented with protease inhibitors and incubated at 4°C on a rocker for 10 min, then centrifuged at 100 g at 4°C for 5 min. Pellets were resuspended in 1 mL shearing buffer (0.1% SDS, 1 mM EDTA, 10 mM Tris HCl pH 8.0) supplemented with protease inhibitors, then centrifuged at 100 g at 4°C for 5 min. Pellets were resuspended in 500 μL shearing buffer, transferred in a 1 mL covaris milliTUBE and sonicated with a Covaris E220 sonicator for 8 min with 5% duty, 140 peak incident power and 200 cycles per burst. The sonicated lysates were centrifuged at 16000 g for 15 min at 4°C to pellet cellular debris. Sonicated extracts were pre-cleared at 4°C with 25 μL Dynabeads Protein G, in the meantime 25 μL Dynabeads Protein G were pre-bound with 3 μg of anti-Pol II antibody (7G5) at 4°C for 6 h. Pre-cleared extracts were incubated with antibody/beads mix overnight at 4°C on a rotating wheel. Beads were washed for 10 min at 4°C on a rotating wheel with, successively, low-salt buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), high-salt buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), LiCl buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA, 1% Nonidet-P40, 1% Na-deoxycholate), and twice with TE buffer (10 mM Tris-HCl pH8.0, 1 mM EDTA). DNA was recovered using 2x100 μL of elution buffer (0.1 M NaHCO3, 1% SDS). Crosslinks were reversed by adding 8 μL of 5 M NaCl, incubating samples overnight at 65°C and treating them with RNAse A and proteinase K for 2 h at 55°C.

Five ng of eluted DNA were used for the library preparation using the MicroPlex library preparation kit v3 (Diagenode) following the manufacturer's instructions. Sequencing was performed by the GenomEast platform (IGBMC) on an Illumina NextSeq 2000 (PE-50).

QUANTIFICATION AND STATISTICAL ANALYSIS

Details for individual experiments including number of biological and technical replicates and statistical tests performed can be found in the figure legends. For unique comparison (n = 2), non-parametric Mann & Whitney tests were performed. For multiple comparisons (n >2), non-parametric Kruskal-Wallis tests were performed, followed by a Dunn post hoc test when possible. For distribution comparisons, Kolmogorov & Smirnov test were performed. Comparisons were considered statistically significant with a *p*.value below 0.05.

Analyses were performed using custom scripts available on request (ImageJ version 1.53q, R software version 4.0.2, [77]). R analyses were achieved using ggplot2 version 3.4.0, tibble 3.1.8, tidyr 1.2.1, readr 2.1.3, purr 0.3.5, dplyr 1.1.10, stringr 1.4.1, forcats 0.5.2) and rstatix (version 4.2.2, https://rpkgs.datanovia.com/rstatix/) for statistical analysis.

Immunofluorescence and EdU signal quantification

Hundreds of measurements for individual nuclei were performed for the different quantifications. ImageJ was used to create a mask for the nuclei via the DAPI signal using the 'Substract background' and 'watershed' tools. This mask was then used on the other channels to measure signal intensity. The ImageJ macro is available upon request. For the statistical analyses, non parametric Kruskal-Wallis tests were performed.

IP-MS data analysis

Proteins were identified by database searching using SequestHT (Thermo Fisher Scientific) with Proteome Discoverer 2.2 software (Thermo Fisher Scientific) on *Mus musculus* database (Swissprot, non-reviewed, release 2019_08_07, 55121 entries). Precursor and fragment mass tolerances were set at 7 ppm and 0.5 Da respectively, and up to 2 missed cleavages were allowed. Oxidation (M) was set as variable modification, and Carbamidomethylation (C) as fixed modification. Peptides were filtered with a false discovery rate (FDR) at 1%, rank 1 and proteins were identified with 1 unique peptide.

Each experiment was analyzed separately using Extracted Ion Chromatogram (XIC) values [83]. First, only peptides whose mean XIC values from triplicate measurements of the control condition (EtOH) were greater than the mean XIC values from triplicate measurements of the mock IP were kept for further analysis. Note that TAF4 and TAF4B were analyzed together in a virtual TAF4.4B protein and TAF9 and TAF9B in a TAF9.9B protein. Second, protein XIC values (PXV) were calculated by averaging XIC values of peptides kept belonging to the same proteins (1). Third, PXV were normalized by the mean of PXV from triplicate measurements of the mock IP (ΔPXV_x) (2) and then by the normalized mean of PXV from triplicate measurements of the bait ($\Delta PXV_{bait(x)}$) (3). Last, fold change (FC) was calculated from one condition (EtOH or 4-OHT) and the average values of the control condition (EtOH) (4).

$$PXV_x = \frac{\sum_{i=1}^{n} XIC_{peptide_{(x)}}}{n} \quad \text{(Equation 1)}$$

$$\Delta PXV_x = PXV_{x(IP)} - \frac{\sum_{j=1}^{n} PXV_{x_{(IPmock)}}}{n} \quad \text{(Equation 2)}$$

$$\Delta PXV_{bait(x)} = \Delta PXV_x \Big/ \frac{\sum_{j=1}^{n} \Delta PXV_{bait}}{n} \quad \text{(Equation 3)}$$

$$FC = \Delta PXV_{bait(x)} \Big/ \frac{\sum_{j=1}^{n} \Delta PXV_{EtOH(bait(x))}}{n} \quad \text{(Equation 4)}$$

where x is the protein of interest, i is the number of peptides belonging to the same protein and j is the number of measurements performed on the same IP under the same conditions.

In the graph, each point corresponds to one FC value and the bar plot is the average of the FC values from different IP experiments, corresponding to the same condition, for a protein or group of proteins belonging to the same sub-complex. For the statistical analyses, non parametric Kruskal-Wallis tests were performed for multiple comparisons.

Nascent 4-sU RNA sequencing data analysis

Experiments were carried out with 3 biological replicates, 1 dataset (TAF10-depleted condition) was omitted because of technical issue. Reads were preprocessed to remove adapter, polyA, low-quality sequences (Phred quality score below 20) and reads shorter

than 40 bases. These preprocessing steps were performed using cutadapt (version 1.10, [74]). Reads were mapped to rRNA sequences using bowtie (version 2.2.8, [73]) and reads mapping to rRNA sequences were then removed for further analysis. Evaluation of the percentage of reads mapping to *Mus musculus*, *Drosophila melanogaster* and *Saccharomyces cerevisiae* was carried out using Fastq-screen (version 0.11.3). Reads were mapped onto the mm10 and BDGP6 assembly of *Mus musculus* and *Drosophila melanogaster* genome using STAR (version 2.5.3a, [81]). Gene expression quantification was performed from uniquely aligned reads using htseq-count (version 0.6.1p1, [79]), with annotations from Ensembl version 93 and ''union'' mode. Only non-ambiguously assigned reads to a gene have been retained for further analyses. Data were normalized using size factors computed with the median-of-ratios method, proposed in [84], on *Drosophila melanogaster* counts. Principal Component Analysis was computed on variance stabilizing transformed data calculated with the method proposed in [76], using size factors computed from *Drosophila melanogaster* counts. Comparisons of interest were performed using Wald statistic test for differential expression and implemented in the Bioconductor package DESeq2 (version 1.16.1, [76]). Eulerr plots were generated using eulerr package (version 4.2.2). In genome browser view, normalized mESCs read counts were used.

CUT&RUN data analysis

Experiments were carried out with 2 biological replicates. Data were preprocessed with cutadapt v4.0 [74] to trim adapter sequences (Nextera Transposase Sequence) from 3' end of reads. Cutadapt was used with the following parameters '-a AGATCGGAAGAG -A AGATCGGAAGAG -m 25:25'. Reads were mapped to Mus musculus genome (assembly mm10) using Bowtie2 v2.4.4 [73] with default parameters except for ''–end-to-end –very-sensitive –no-mixed –no-discordant -I 10 -X 700''. BigWig files were generated using deeptools bamCoverage v3.5 [75] with the following parameters ''-bs 10 –normalizeUsing CPM –effectiveGenomeSize 2652783500 –skipNonCoveredRegions –extendReads''. Bigwig files of mean signal per condition were generated using deeptools bamCompare v3.5 [75] with the following parameters ''-of bigwig –operation mean –effectiveGenomeSize 2652783500 –normalizeUsing CPM –scaleFactorsMethod None –extendReads''.

The peak calling was done with Macs2 v2.2.7.1 [80] with default parameters except ''-f BAMPE -q 0.1''. Peaks were annotated relative to genomic features using Homer v4.11 [78] (annotations got extracted from gtf file downloaded from Ensembl 102).

Heatmap and mean profile

The tool deeptools computeMatrix v3.55 [75] was used to generate a count matrix at the positions of interest (union of peaks of all datasets to compare) and finally the tool deeptools plotProfile v3.5 was used to generate mean profile plots and deeptools plotHeatmap v3.5 was used to generate heatmaps. Top 10% peaks were filtered using an in-house R script, using the fold enrichment score as input.

Differential peak analysis

Detected peaks (all peaks) were combined to get the union of all peaks (per condition RT7 or RT10) using the tool Bedtools merge v2.30.0 [85]. Then, the number of reads per merged peaks were computed using deeptools multiBamSummary BED-file v3.5.0 [75]. Data were normalized using the method proposed by Anders and Huber [79] using read counts per peaks. Comparisons of interest were performed using the method proposed by [76] implemented in the DESeq2 Bioconductor library (DESeq2 v1.42.1). Resulting *P*.values were adjusted for multiple testing using the Benjamini and Hochberg method.

Pol II Chip-seq data analysis

Experiments were conducted in biological duplicates. Reads were mapped to mouse genome (mm10) using Bowtie2 v2.5.0 [73] with default parameters except for ''–mm''. Reads which mapping quality is below 10 were removed using samtools v1.15.1 [86] with the command line "samtools view -b -q 10". Then, reads falling into Encode blacklisted regions v2 [87] were removed using BEDtools intersect v2.30.0 [85]. BigWig files were generated using Deeptools bamCoverage [85] v3.5.4 with the following parameters ''-bs 10 -p 10–skipNonCoveredRegions –extendReads''. Deeptools bamCoverage was used with ''—scaleFactor'' and scale factors determined based on the background signals calculated on the 30% least enriched regions.

Heatmap and mean profile

The tool Deeptools computeMatrix v3.5.0 was used to generate a count matrix at the positions of interest and Deeptools plotHeatmap v3.5.0 was used to generate heatmaps and plotProfile v3.5.0 was used to create mean profiles [75]. Regions of interest were either all TSS of the Mouse Genome mm10 with annotation from Ensembl v102 or TBP peaks top10 or all peaks.

Pausing index

Pausing index was described as such: Pausing index = 5' read density / gene body density. Gene positions were extracted from Ensembl v102 using R scripts and the R/Bioconductor package biomaRt v2.50.0 [88]. The number of reads per gene bodies and 5' regions was counted using BEDtools intersect v2.30.0 [85]. The 5′ region of genes was defined as the regions comprised between TSS-30nt and TSS + 300nt. Gene body region was defined as the region comprised between TSS + 301nt and TTS-300nt. Gene bodies were all normalized to be 1000 nt long. Moreover, 1 was added to the gene body and 5' read counts so that illegal division does not occur. Statistical comparison of the distributions was performed using a Kolmogorov & Smirnov test.