

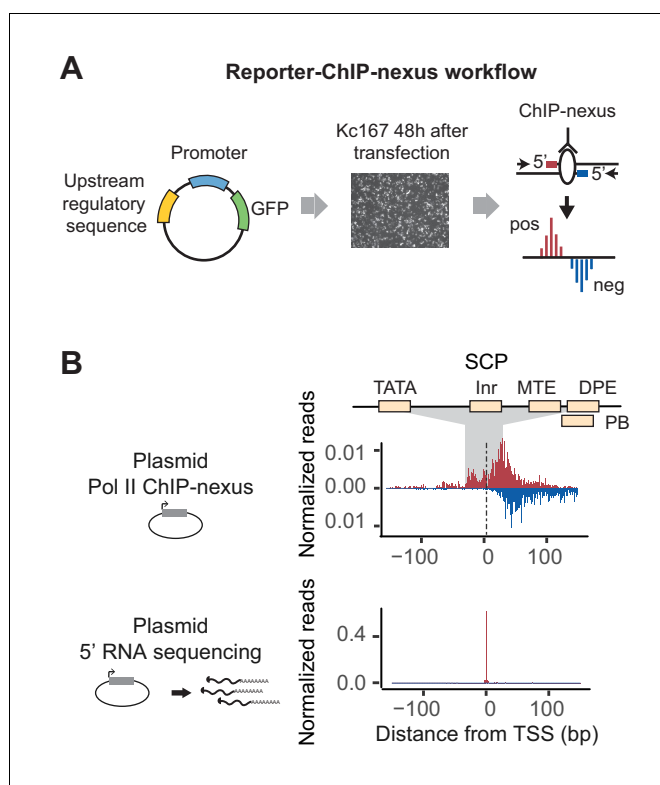


---

## Figures and figure supplements

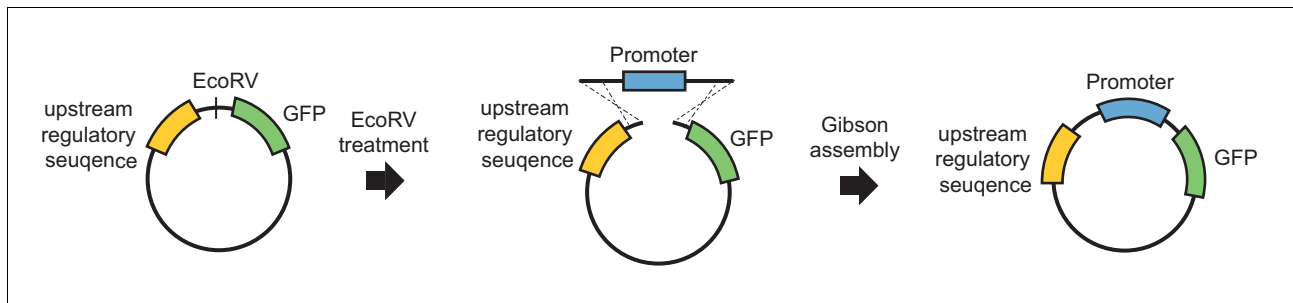
Reporter-ChIP-nexus reveals strong contribution of the *Drosophila* initiator sequence to RNA polymerase pausing

**Wanqing Shao et al**



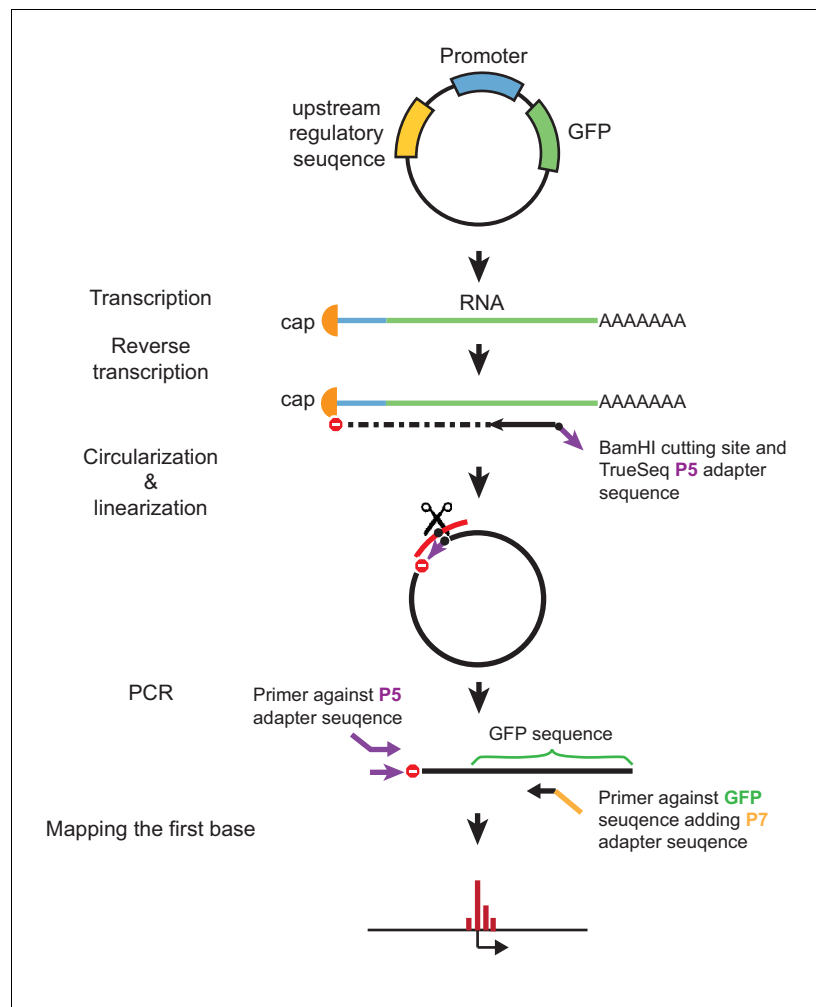
**Figure 1.** Reporter-ChIP-nexus captures paused Pol II. (A) Reporter-ChIP-nexus is performed by cloning *Drosophila pseudoobscura* promoters or synthetic promoters into a simple GFP reporter and transfecting into *D. melanogaster* Kc167 cells. The whole cell lysate from cross-linked cells is used to perform Pol II ChIP-nexus. Exonuclease stop bases are then mapped and shown on the positive strand in red above the line, while reads from the negative end are shown in blue below the line. (B) Results of reporter-ChIP-nexus reveal strong Pol II pausing at the synthetic super core promoter (SCP), which contains the core promoter elements TATA, Inr, MTE, DPE and PB (top). The position of transcriptional initiation is mapped by sequencing the 5' end of the produced RNA (bottom). The results for the SCP promoter show that the vast majority of RNAs start at the expected site of initiation.

DOI: <https://doi.org/10.7554/eLife.41461.002>



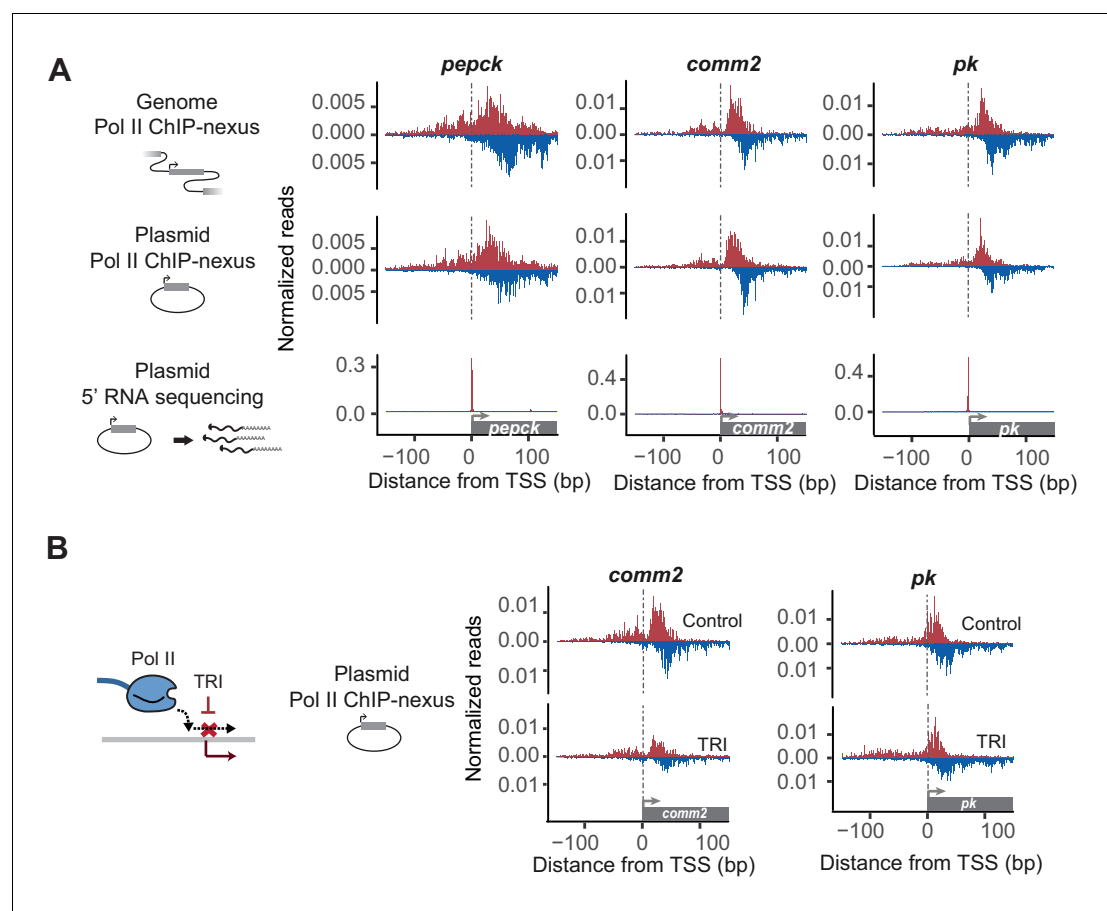
**Figure 1—figure supplement 1.** A simple GFP reporter that allows the fast insertion of any promoter sequence. To perform reporter-ChIP-nexus, we constructed a simple GFP reporter plasmid with upstream regulatory sequences. An EcoRV cutting site was used to linearize the plasmid and insert a promoter of interest using Gibson assembly (see Materials and methods for more information).

DOI: <https://doi.org/10.7554/eLife.41461.003>



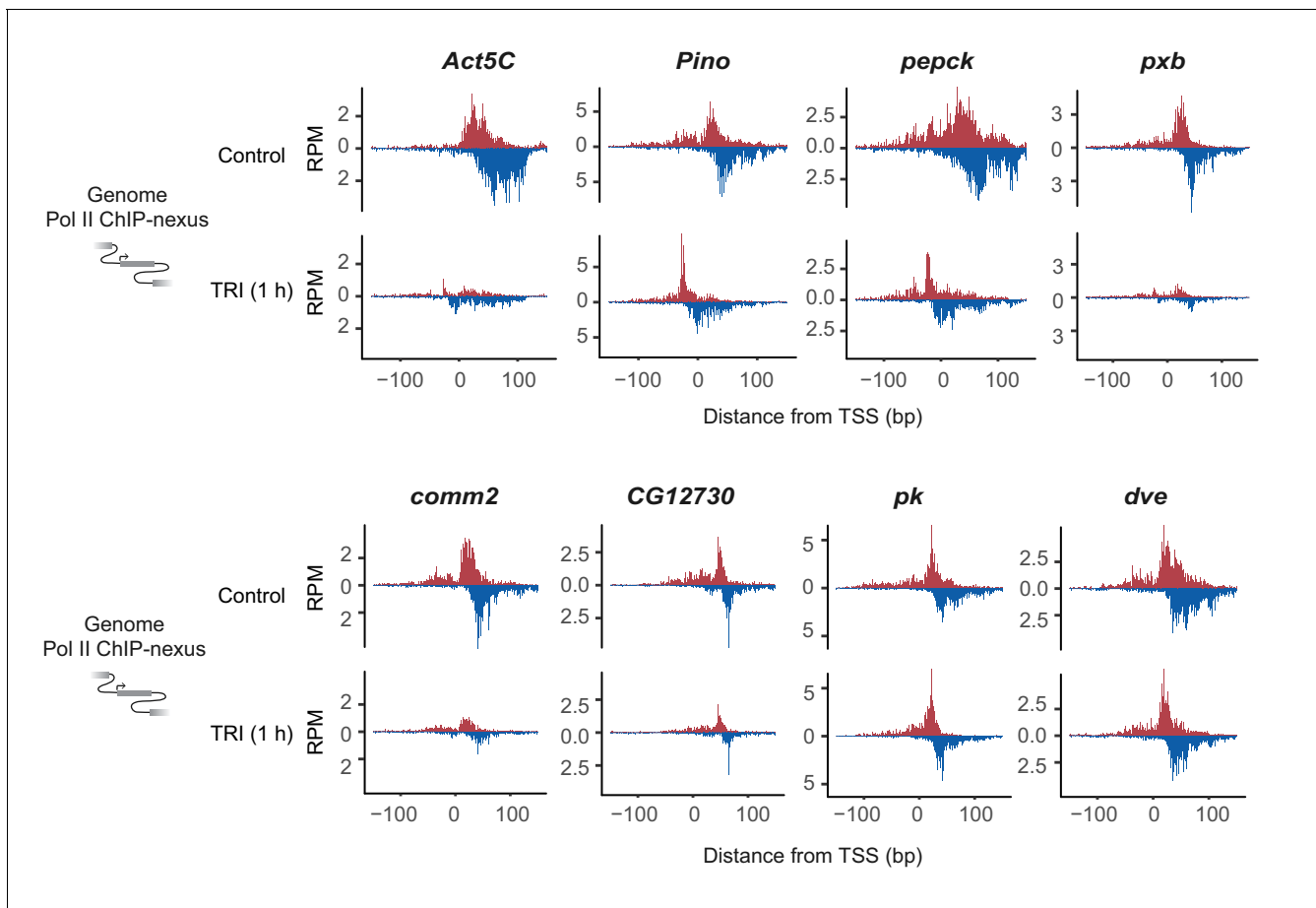
**Figure 1—figure supplement 2.** Workflow for gene-specific 5' RNA sequencing, a method similar to RNA amplification of cDNA ends (RACE). RNA transcribed from the reporter was reverse transcribed using a primer that has 3' sequences complementary to the GFP sequence. The 5' of this primer contains a BamHI restriction cutting site and a partial reverse TruSeq P5 adapter sequence (purple). When the reverse transcriptase reaction is complete (usually stopping at the cap), each generated cDNA contains the TruSeq P5 adapter sequence. After purification, the single-stranded cDNA is circularized to bring the TruSeq P5 adapter sequence to the cDNA 3' end, where the reverse transcriptase stops. The product is then linearized again by digesting with BamHI (scissors) after making the restriction enzyme cut site double-stranded by annealing it to a complementary oligos (red). The resulting DNA is then amplified in a PCR with a primer that contains the full-length P5 sequence (matching the partial P5 sequence), as well as a primer that primes on the GFP sequence (black) and has the full-length P7 adapter sequence at the 5' end (yellow). The library is then sequenced with a P5 primer and after alignment to the plasmid, the first base is recorded. These bases correspond to the site where the reverse transcriptase stopped transcribing, which should occur at the 5' end of the RNA, where the cap is found.

DOI: <https://doi.org/10.7554/eLife.41461.004>



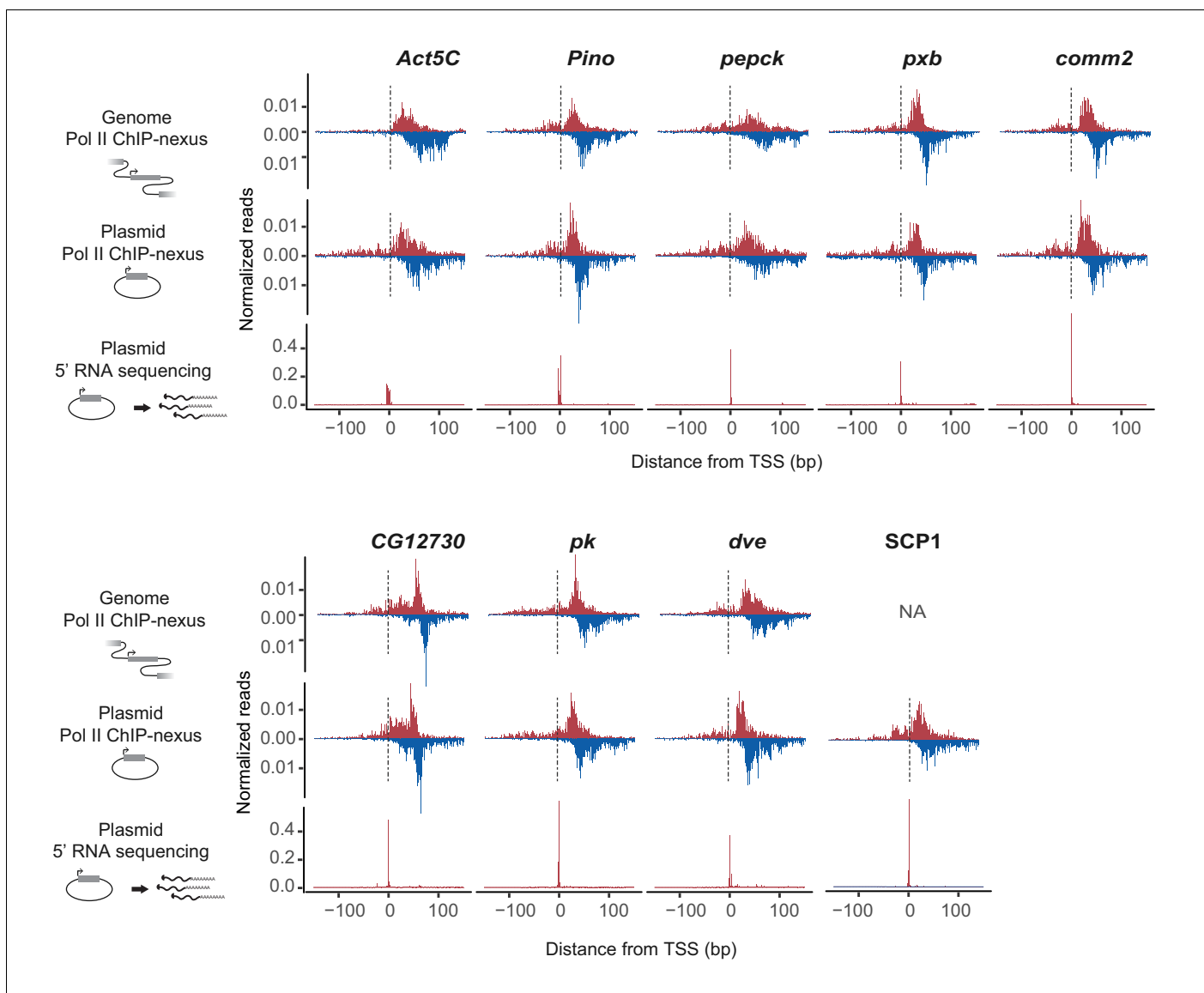
**Figure 2.** Reporter-ChIP-nexus recapitulates the endogenous Pol II pausing profile. **(A)** To obtain the endogenous Pol II pattern of *D. pseudoobscura* promoters in the genome, ChIP-nexus was performed in a *D. pseudoobscura* cell line. Results are shown for the *pepck*, *comm2* and *pk* promoters. The same promoters were then examined using reporter-ChIP-nexus in *D. melanogaster* Kc167 cells, which yielded patterns very similar to the endogenous Pol II profiles. The transcription initiation sites were confirmed by 5' RNA sequencing. **(B)** To determine the stability of paused Pol II on the plasmid, transfected cells were treated with either DMSO (Control) or triptolide (TRI) for 1 hr. The results show a relative reduction in Pol II similar to that of the endogenous loci after treating the *D. pseudoobscura* cell line with TRI (see **Figure 2—figure supplement 4**). Therefore, reporter-ChIP-nexus reveals gene-specific Pol II pausing stability on a plasmid.

DOI: <https://doi.org/10.7554/eLife.41461.005>



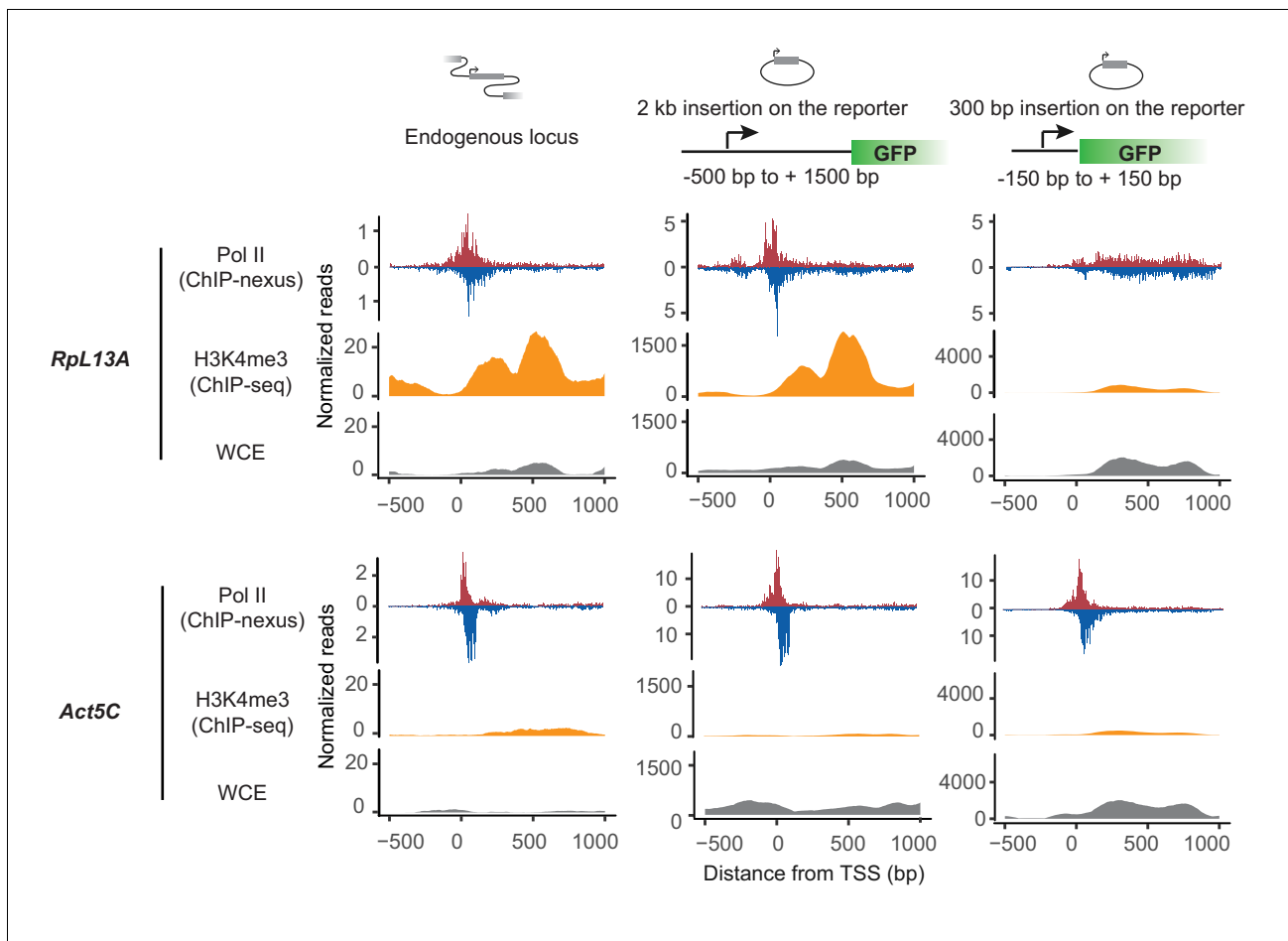
**Figure 2—figure supplement 1.** Paused Pol II stability measurements at eight *D. pseudoobscura* promoters. Pol II ChIP-nexus was performed on the *D. pseudoobscura* cell line ML83-63. The stability of paused Pol II was analyzed by comparing the Pol II profile under control and triptolide (TRI) treated conditions. The eight *D. pseudoobscura* promoters shown in this figure have high Pol II signal and various degrees of Pol II pausing. The sequences of these promoters were cloned into our reporter and used for later mutagenesis experiments.

DOI: <https://doi.org/10.7554/eLife.41461.006>



**Figure 2—figure supplement 2.** Reporter-ChIP-nexus recapitulates endogenous Pol II pausing. The ChIP-nexus profile of endogenous *D. pseudoobscura* promoters (first row) strongly resembles the profile detected by reporter-ChIP-nexus after cloning the *D. pseudoobscura* promoters into the reporter and transfection of the plasmid into Kc167 cells (second row). The artificial super core promoter (SCP) also produces a strong Pol II pausing profile. The transcription start site on the reporter was mapped for each promoter using gene-specific 5' RNA sequencing (third row) and was found to map to the expected location.

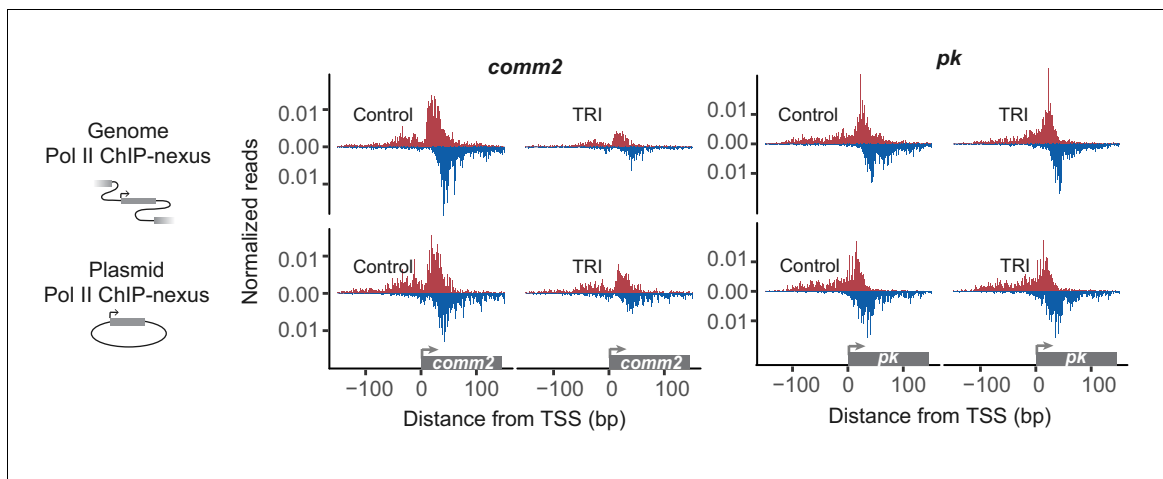
DOI: <https://doi.org/10.7554/eLife.41461.007>



**Figure 2—figure supplement 3.** Larger promoter region insertion is required for recapitulating Pol II pausing at *RpL13A* on the plasmid. Pol II ChIP-nexus and H3K4me3 ChIP-seq profiles at *D. pseudoobscura* *RpL13A* and *Act5C* endogenous loci and that on the plasmids with 300 bp or 2 kb promoter region insertion. Both *RpL13A* and *Act5C* show strong pausing profile at the endogenous loci. However, in contrast to *Act5C*, where Pol II pausing can be observed on the plasmid with both 300 bp and 2 kb promoter region insertion, *RpL13A* pausing profile can only be recapitulated on the plasmid when the larger 2 kb region was inserted. In addition, although *Act5C* show minimal enrichment of H3K4me3, high H3K4me3 signal can be observed at *RpL13A* endogenous locus, and this H3K4me3 signal was only recapitulated on the plasmid with the 2 kb promoter region insertion. This correlation between Pol II profile and H3K4me3 enrichment suggests that endogenous chromatin context is important for establishing paused Pol II at *RpL13A*.

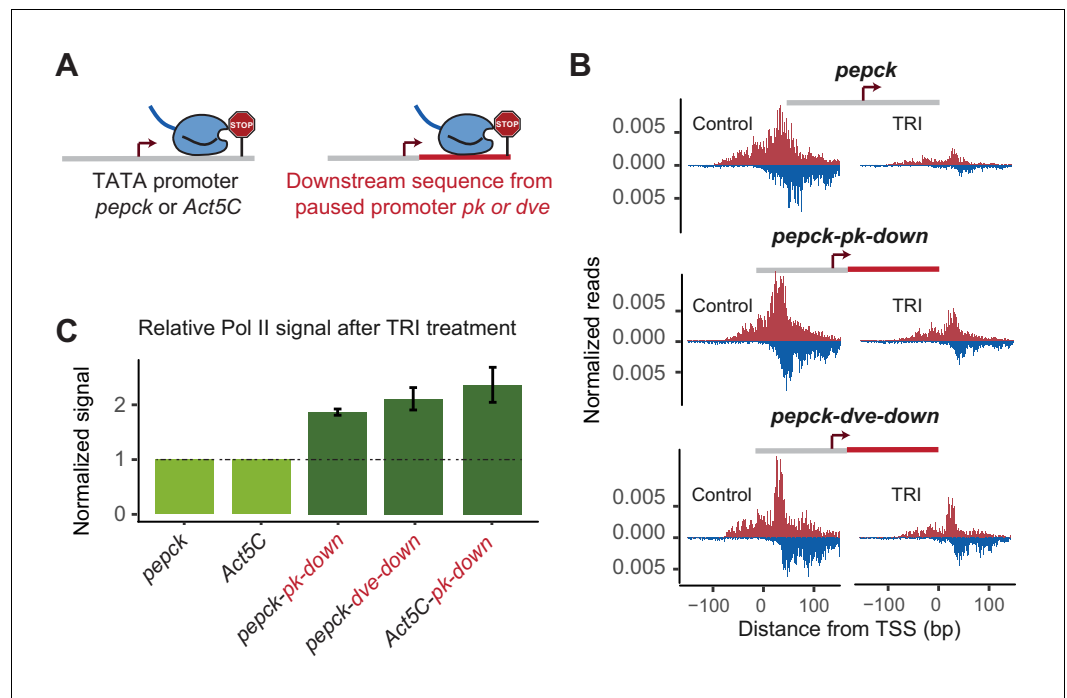
DOI: <https://doi.org/10.7554/eLife.41461.008>





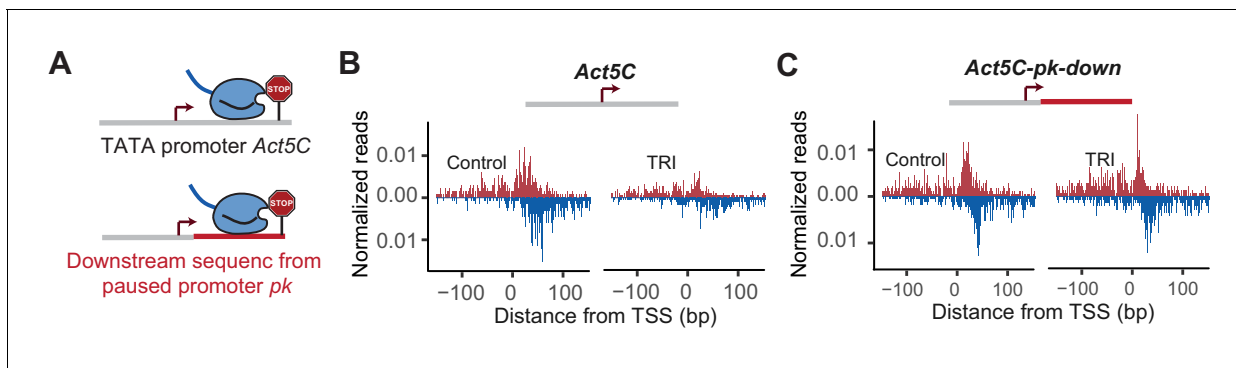
**Figure 2—figure supplement 4.** Reporter-ChIP-nexus recapitulates gene-specific Pol II pausing stability. The stability of paused Pol II was examined on the plasmids and at the endogenous loci by treatment with triptolide (TRI). After TRI treatment, paused Pol II is lost, and the degree of persistence relative to the DMSO-treated control condition is a measurement for the stability of paused Pol II. The examples of the *comm2* and *pk* promoters show that the relative loss of Pol II on the plasmid is similar to that at the endogenous loci.

DOI: <https://doi.org/10.7554/eLife.41461.009>



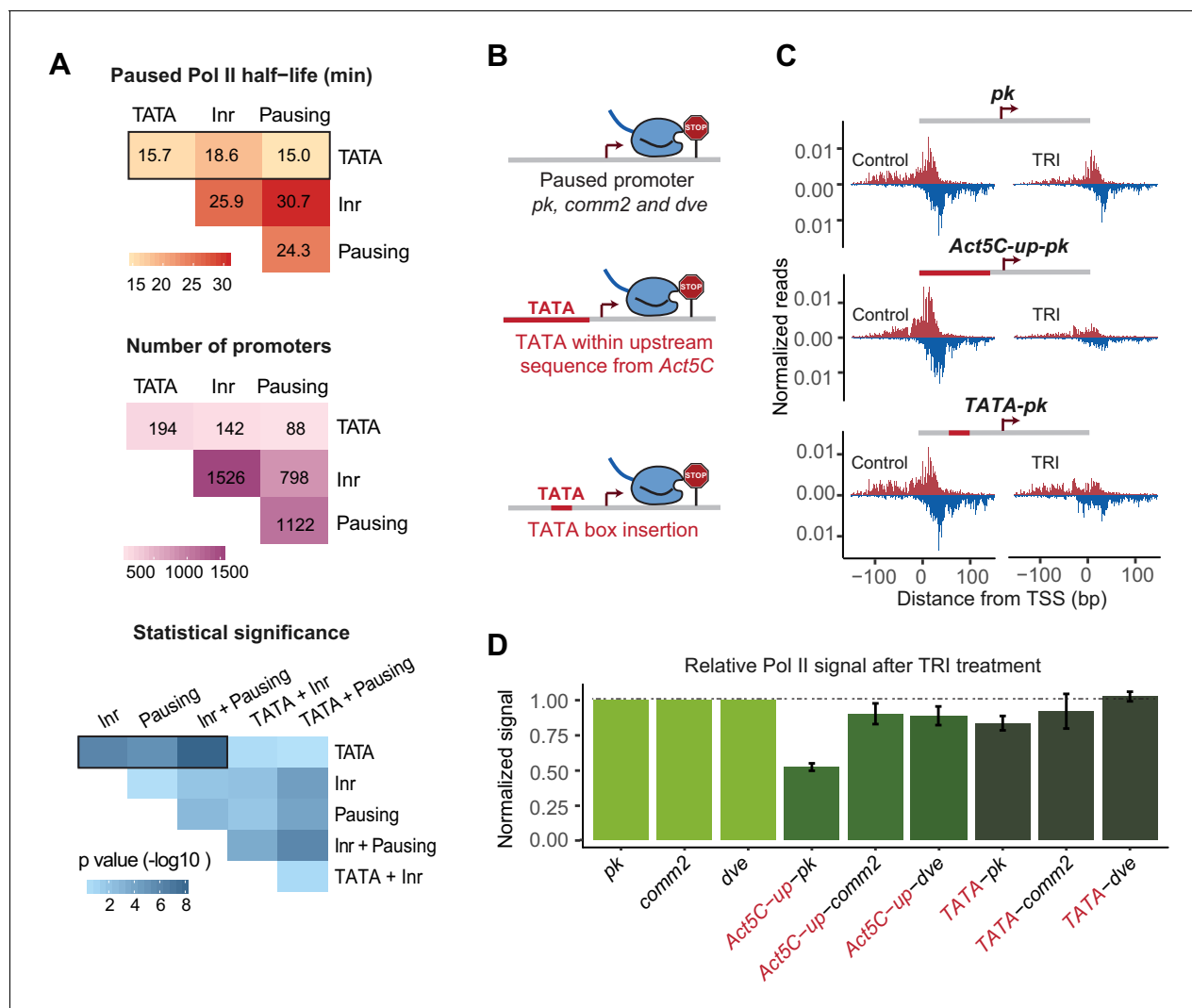
**Figure 3.** Changes in downstream promoter sequences alter paused Pol II stability. (A) The *pepck* and *Act5C* downstream sequences were replaced with that from the stably paused promoter *pk* or *dve* (fusion site: 8 bp after the TSS). (B) Pol II ChIP-nexus data on the plasmids after transfection into Kc167 cells with or without treatment with triptolide (TRI) for 5 min. The wild-type *pepck* promoter shows a strong reduction of paused Pol II. At the *pepck-pk-down* and *pepck-dve-down* fusion promoters, the same TRI treatment did not reduce paused Pol II to the same extent as at the wild-type promoter, suggesting an increase in the paused Pol II stability as a result of changing the downstream promoter sequence. (C) To quantify the difference in paused Pol II stability between different constructs, the ratio of paused Pol II before and after TRI treatment is calculated using two biological replicates. The ratio for the wild-type promoter is then normalized to 1 (light green), and the relative change in this ratio for the fusion promoter is shown on the right (dark green). Error bars refer to the standard error of the mean (TRI treatment: 5 min for all promoters).

DOI: <https://doi.org/10.7554/eLife.41461.010>



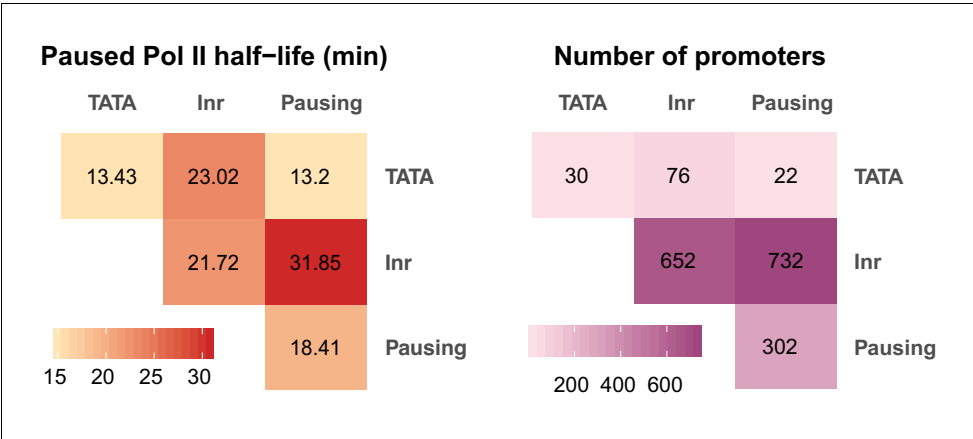
**Figure 3—figure supplement 1.** Downstream promoter sequences influences Pol II pausing at *Act5C*. Pol II ChIP-nexus profile at *Act5C* and *Act5C-pk-down* fusion promoters under control condition and after 5 min Triptolide (TRI) treatment.

DOI: <https://doi.org/10.7554/eLife.41461.011>



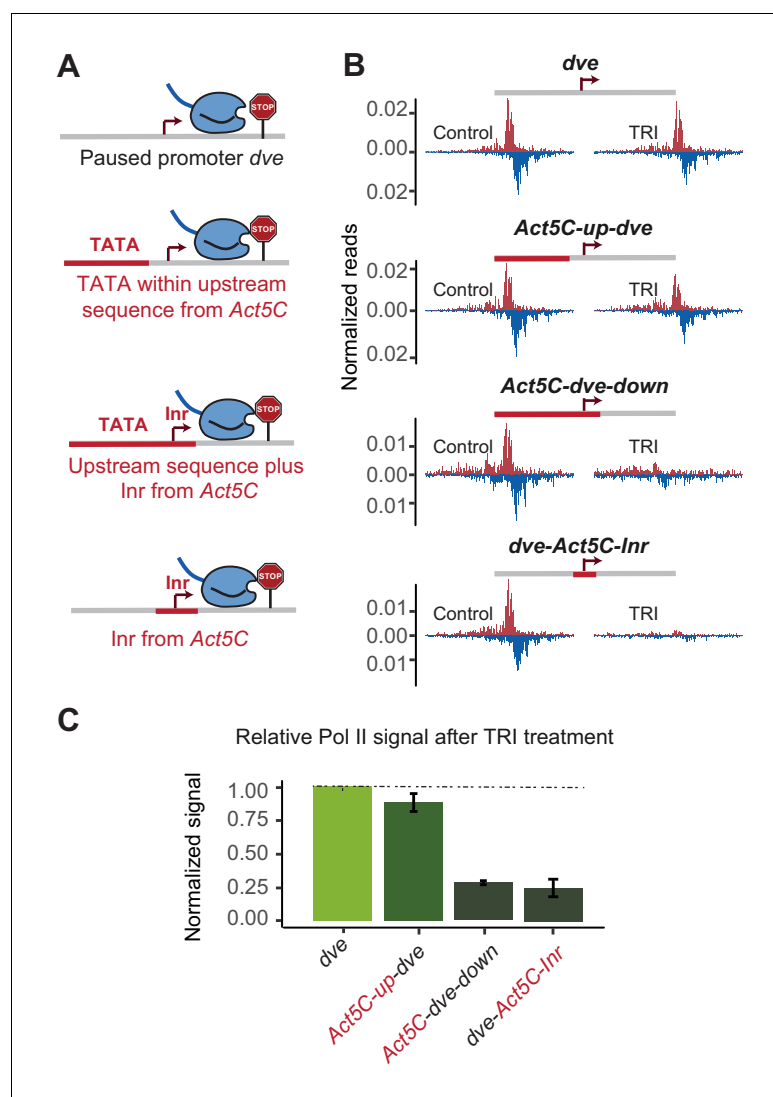
**Figure 4.** A TATA insertion may reduce Pol II pausing. (A) Analysis of paused Pol II half-lives as a function of various core promoter element combinations. Median paused Pol II half-life of promoters with different combinations of core promoter elements (top) and the number of promoters classified (middle) are shown. A Wilcoxon rank sum test was then used to test whether the various promoter combinations have significantly different Pol II half-lives from each other (bottom). Core promoter elements other than those used for classification were allowed to occur in each group (e.g. Inr and pausing elements can occur at the 194 promoters with TATA box). Similar results were obtained when no promoter element other than those used for classification were allowed (mutually exclusive model **Figure 4—figure supplement 1**). (B) Experimental strategy to functionally test the role of TATA in Pol II pausing: the upstream promoter sequences of the paused promoter *pk*, *comm2* and *dve* were replaced with that of the TATA promoter *Act5C* (fusion site: 7 bp before the TSS), or a canonical TATA box sequence (TATAAAAA) was inserted at 31 bp upstream of the TSS. (C) The Pol II ChIP-nexus profiles for wild-type *pk*, the upstream region replacement (*Act5C-up-pk*) and the TATA insertion (*TATA-pk*) under control conditions (left) and after 1 hr treatment with triptolide (TRI, right) show that Pol II pausing is reduced after the *pk* promoter acquired a TATA box. (D) Quantification of the relative changes in paused Pol II stability for all fusion promoters relative to the wild-type promoters (TRI treatment: 1 hr for *pk* and *dve* based promoters, 40 min for *comm2* based promoters). Note that only the *pk* promoter shows a clear reduction in Pol II pausing after insertion of a TATA box.

DOI: <https://doi.org/10.7554/eLife.41461.012>



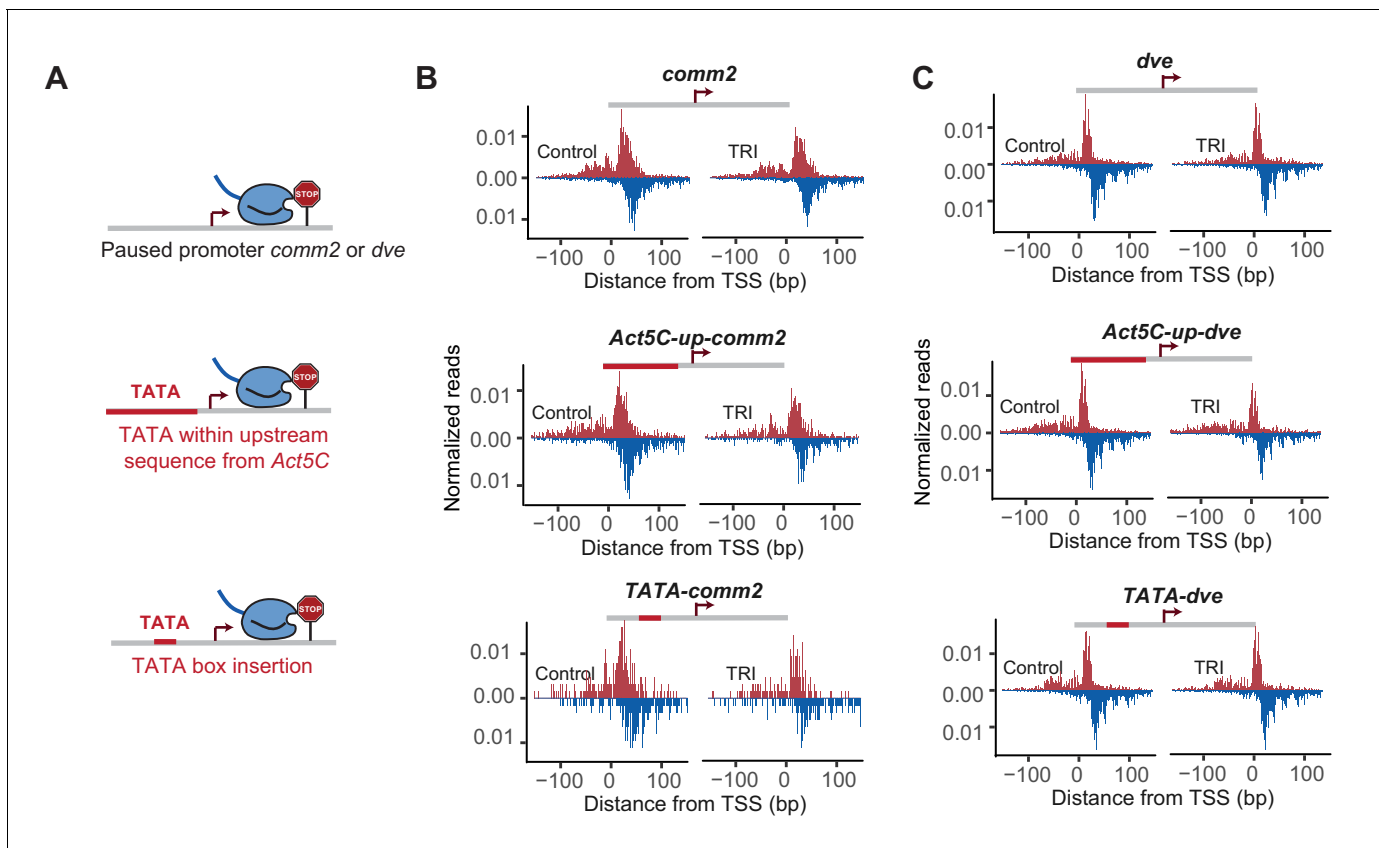
**Figure 4—figure supplement 1.** Correlation between paused Pol II half-life and core promoter elements using a mutually exclusive model. To better understand the relationship between core promoter elements and the half-life of paused Pol II genome-wide, we calculated the median half-life for each element and its possible combination with other elements with a mutually exclusive model (See Materials and method for details). Left: median paused Pol II half-life. Right: number of promoters in each category.

DOI: <https://doi.org/10.7554/eLife.41461.013>



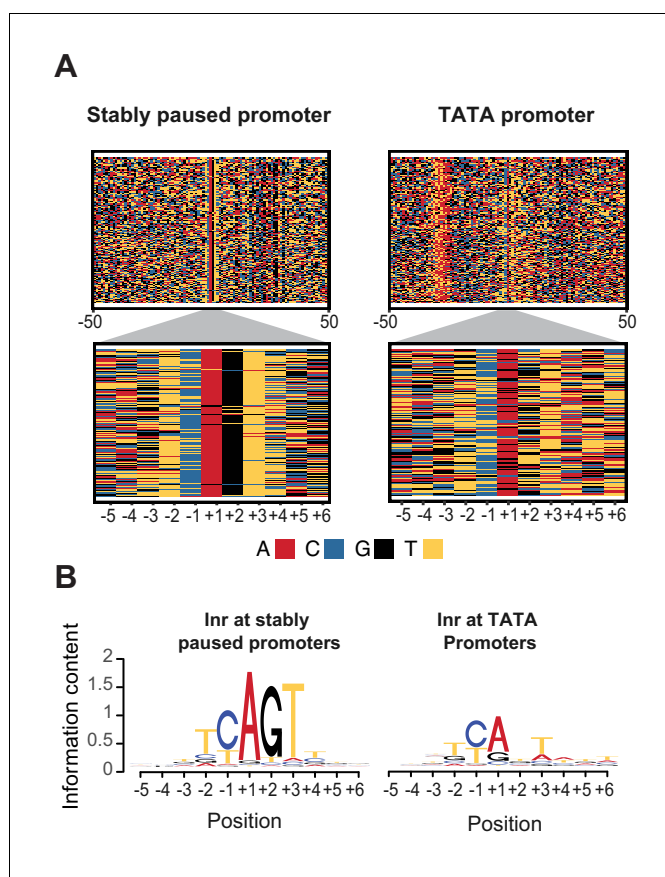
**Figure 5.** The initiator sequence plays an important role in Pol II pausing. (A) Experimental series on the *dve* promoter testing the role of the Inr in conjunction with a TATA-containing region. Either only the upstream region was replaced with that of the *Act5C* promoter (fusion site 7 bp before the TSS), the upstream region and the Inr were replaced (fusion site 8 bp after the TSS), or only the Inr was replaced (16 bp region around the TSS). (B) Pol II ChIP-nexus profiles under control conditions and after treatment with triptolide (TRI) for 1 hr show a strong reduction of paused Pol II after replacing the Inr sequence. (C) Quantification of the relative changes in paused Pol II stability for all fusion promoters relative to the wild-type promoter (TRI treatment: 1 hr for all the promoters).

DOI: <https://doi.org/10.7554/eLife.41461.014>



**Figure 5—figure supplement 1.** Effect of TATA insertion at *comm2* and *dve*. Pol II ChIP-nexus profile at *comm2*, *Act5C-up-comm2*, *TATA-comm2*, *dve*, *Act5C-up-dve* and *TATA-dve* promoters under control condition and after Triptolide (TRI) treatment (40 min for *comm2* and 1 hr for *dve*).

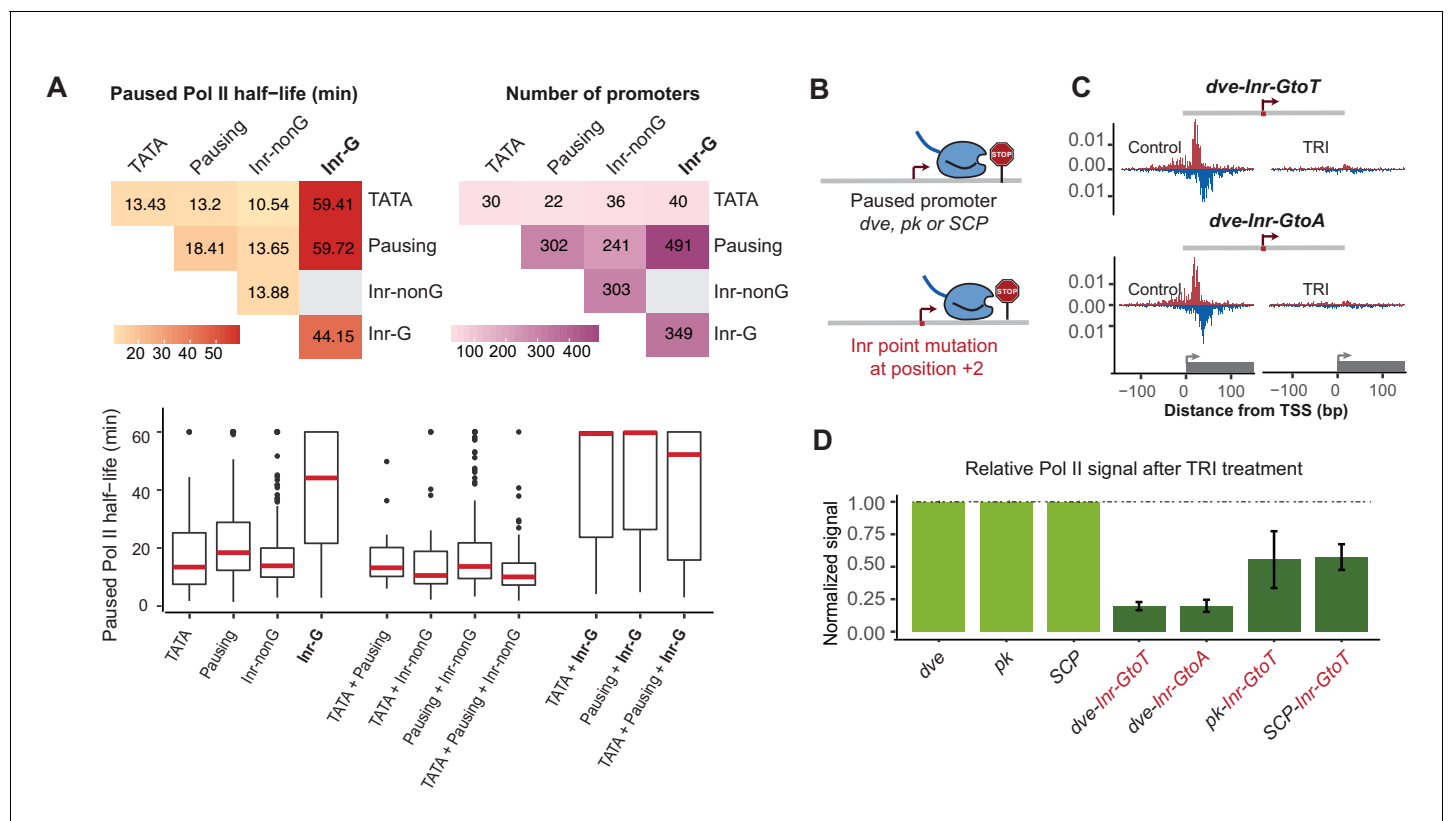
DOI: <https://doi.org/10.7554/eLife.41461.015>



**Figure 6.** Initiator sequences differ between TATA and stably paused promoters. (A) The sequence of stably paused promoters (132 randomly selected from the 490 promoters that did not have a TATA box and had a paused Pol II half-life longer than 60 min) are shown on the left as colored letters for a 100 bp window centered on the transcription start site (top). TATA promoters (132 promoters with paused Pol II half-lives shorter than 30 min) are shown on the right as comparison. A higher magnification of the initiator (Inr) sequences below shows clear differences between the two promoter types. (B) The consensus motif for the Inr sequences of stably paused promoters match more closely the canonical Inr consensus sequence (TCAKTY) and preferentially contains a G at the +2 position in most cases, whereas the Inr sequences of TATA promoters are more degenerate without enrichment of a G at the +2 position.

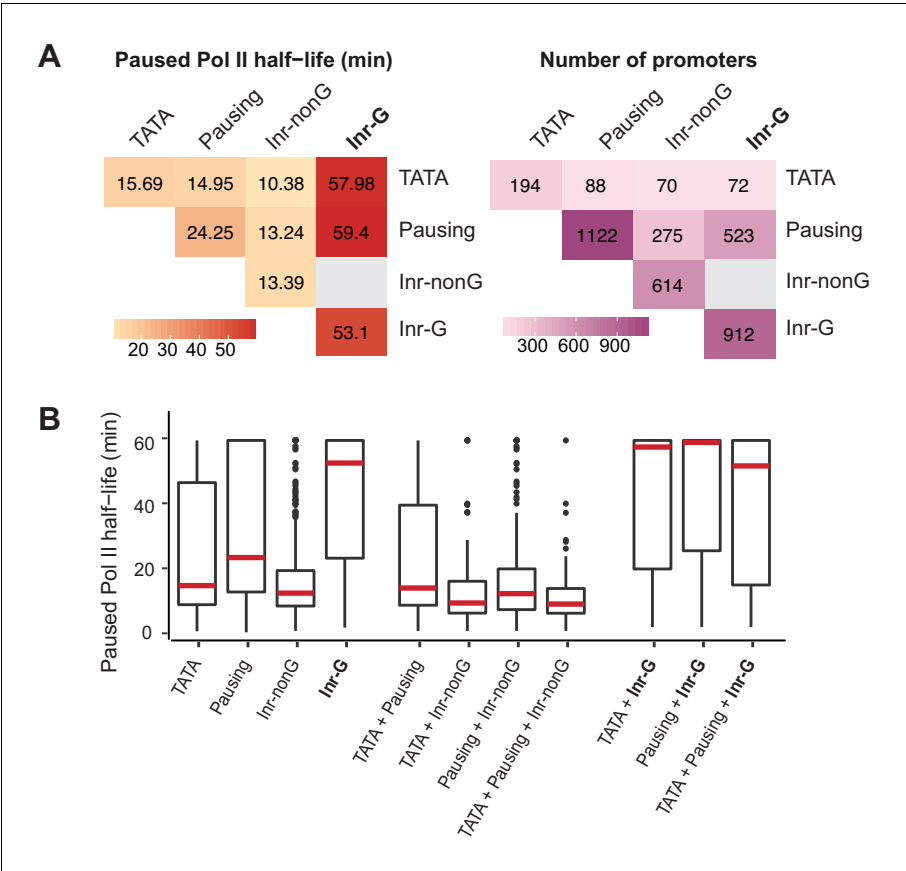
DOI: <https://doi.org/10.7554/eLife.41461.016>



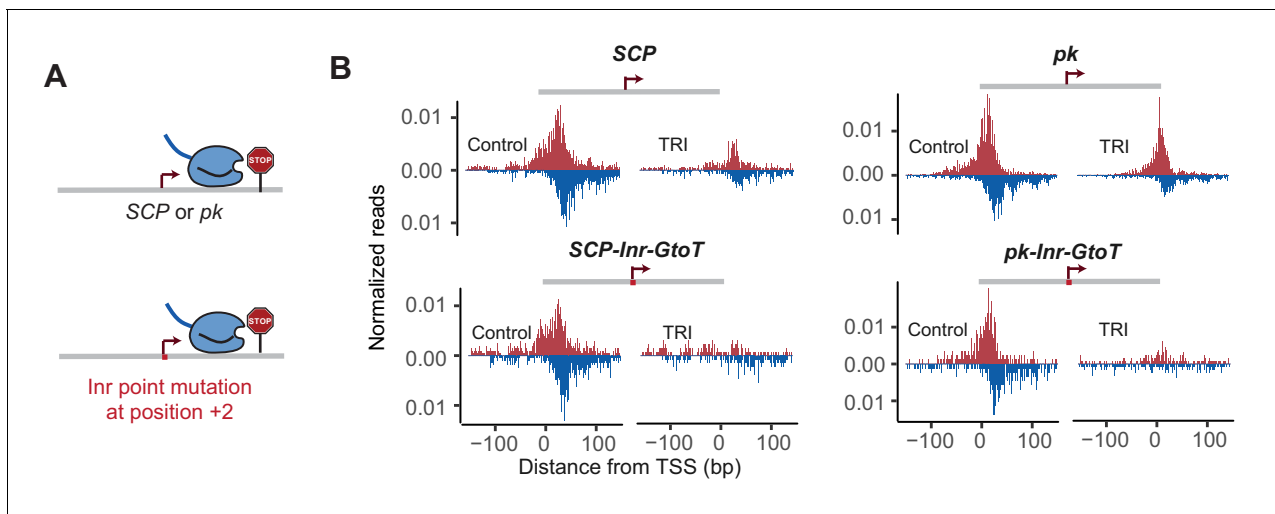


**Figure 7.** The G at Inr + 2 position is critical for stable Pol II pausing. **(A)** Analysis of paused Pol II half-lives as a function of core promoter element combinations after separating the Inr sequences into those that contain a G at position +2 (Inr-G) versus those that do not contain a G at this position (Inr-nonG). Median paused Pol II half-life (left), promoter numbers (right) and boxplot of their distribution (bottom, median in red) are shown for promoters with different combinations of core promoter elements. The promoters only contain the indicated promoter element, excluding any other promoter element (e.g. the 491 promoters with pausing elements and Inr-G variant do not contain TATA box). Similar results were obtained using a non-mutually exclusive model (**Figure 7—figure supplement 1**). Note that strong Pol II pausing was observed in all combinations that contain the Inr-G variant. Results from testing between combinations using the Wilcoxon rank test are shown in Table S4. **(B)** Experimental strategy to test the effect of mutating the G at Inr + 2 position of the Inr to T or A at the stably paused promoters *dve*, *pk* and the synthetic Super Core Promoter (SCP). **(C)** Pol II ChIP-nexus profiles after treatment with triptolide (TRI) for 1 hr show a strong reduction in Pol II pausing after the G was mutated at the *dve* promoter. **(D)** Quantification of the relative changes in paused Pol II stability for all fusion promoters relative to the corresponding wild-type promoter. Error bars are from replicate experiments. The duration of TRI treatment was 1 hr for *pk* and *dve* derived promoters, 30 min for SCP derived promoters.

DOI: <https://doi.org/10.7554/eLife.41461.017>

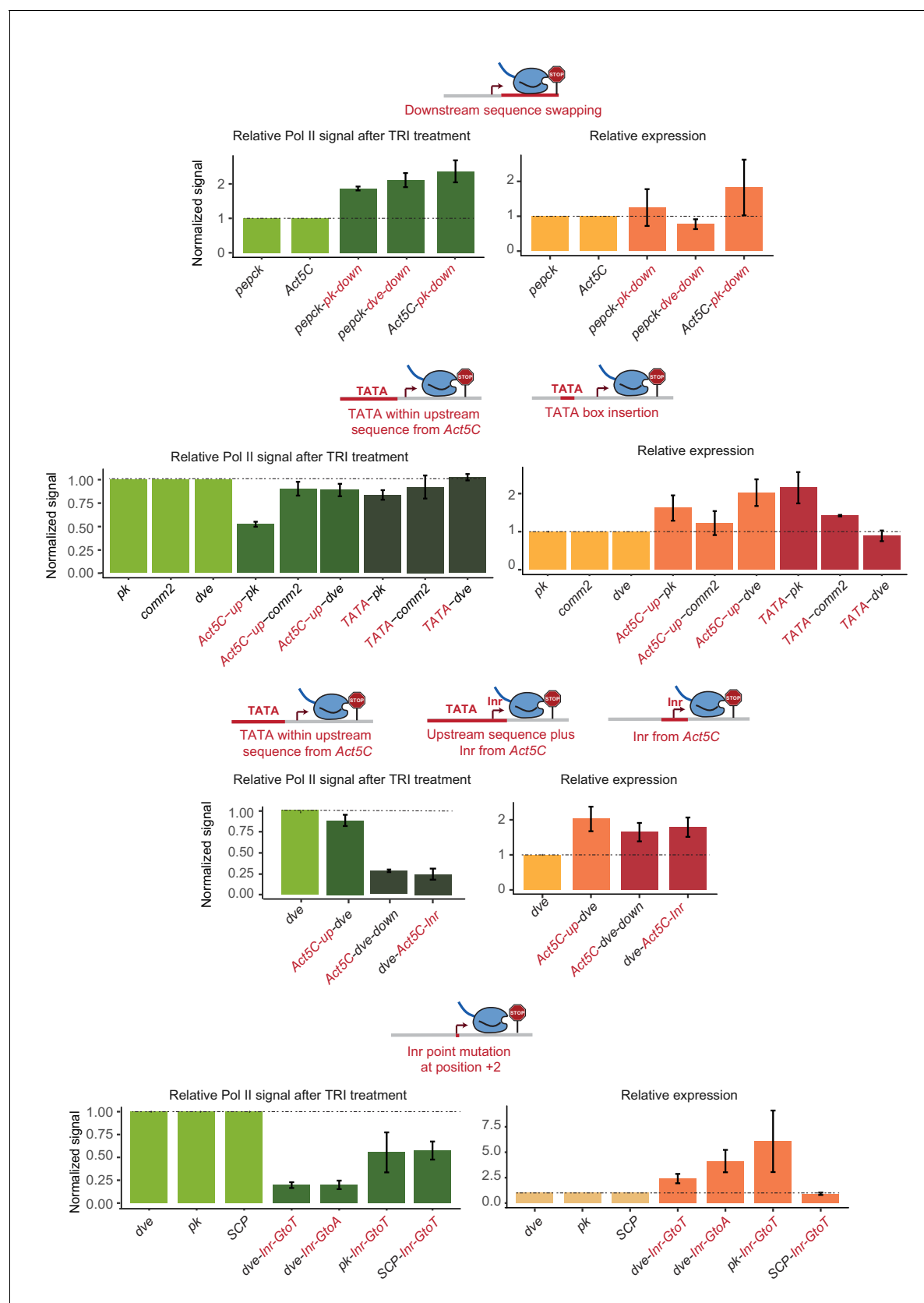


**Figure 7—figure supplement 1.** Correlation between paused Pol II half-life and Inr variants using a non-mutually exclusive model. (A) Median paused Pol II half-life (left) and promoter numbers (right) at promoters with different combinations of core promoter elements using a non-mutually exclusive model. (B) Paused Pol II half-life at promoters with different combinations of core promoter elements using a non-mutually exclusive model.  
DOI: <https://doi.org/10.7554/eLife.41461.018>



**Figure 7—figure supplement 2.** Mutating the G at Inr + 2 position reduces Pol II pausing at SCP and *pk*. Pol II ChIP-nexus profile at SCP, *SCP-Inr-GtoT*, *pk* and *pk-Inr-GtoT* fusion promoters under control condition and Triptolide (TRI) treatment (30 min for SCP, 1 hr for *pk*).

DOI: <https://doi.org/10.7554/eLife.41461.019>



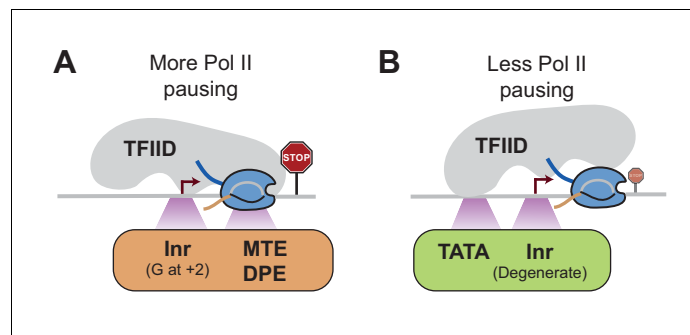
**Figure 7—figure supplement 3.** Transcript level changes after altering the promoter sequence. To study the effect of altered promoter sequences on RNA expression, we performed RT-qPCR for all experiments. The relative Pol II pausing stability (green) and relative RNA expression levels (yellow) are

Figure 7—figure supplement 3 continued on next page

*Figure 7—figure supplement 3 continued*

shown. Measurements for the wild-type promoter are normalized to 1, and the relative values for the mutated promoters are shown next to it. Errors bars refer to standard error among two to three biological replicates. Note that the error is large due the variability in transfection efficiency between independent samples. Despite the variability, the change in Pol II pausing stability was in most cases inversely correlated with that of RNA transcription. Thus, a loss in Pol II pausing was usually associated with an increase in transcription. This is consistent with Pol II pausing having an inhibitory effect on transcription. We cannot exclude the possibility that the causality is reverse though, that is that an increase in promoter strength leads to a reduction in Pol II pausing. It is unclear, however, whether promoter strength can be uncoupled from Pol II pausing, and we cannot measure them separately. In our model, promoter strength and Pol II pausing are both determined by TFIID, which binds to both the upstream and downstream core promoter elements we studied.

DOI: <https://doi.org/10.7554/eLife.41461.020>



**Figure 8.** Model of how promoter sequences may influence Pol II pausing through TFIID Since the core promoter elements that we analyzed are bound by TFIID, we hypothesize that TFIID affects Pol II pausing dependent on the position it binds to the promoter. (A) At stably paused promoters, pausing elements and consensus Inr sequences with a G at +2 position promote TFIID-downstream DNA interactions, which prolong Pol II pausing. (B) At TATA-containing promoters, the TATA box and a degenerate Inr sequence favor TFIID interactions with upstream DNA, which promote Pol II pause release.

DOI: <https://doi.org/10.7554/eLife.41461.021>