

Supplemental information

**Efficient RNA polymerase II pause release
requires U2 snRNP function**

Livia Caizzi, Sara Monteiro-Martins, Björn Schwalb, Kseniia Lysakovskaia, Jana Schmitzova, Anna Sawicka, Ying Chen, Michael Lidschreiber, and Patrick Cramer

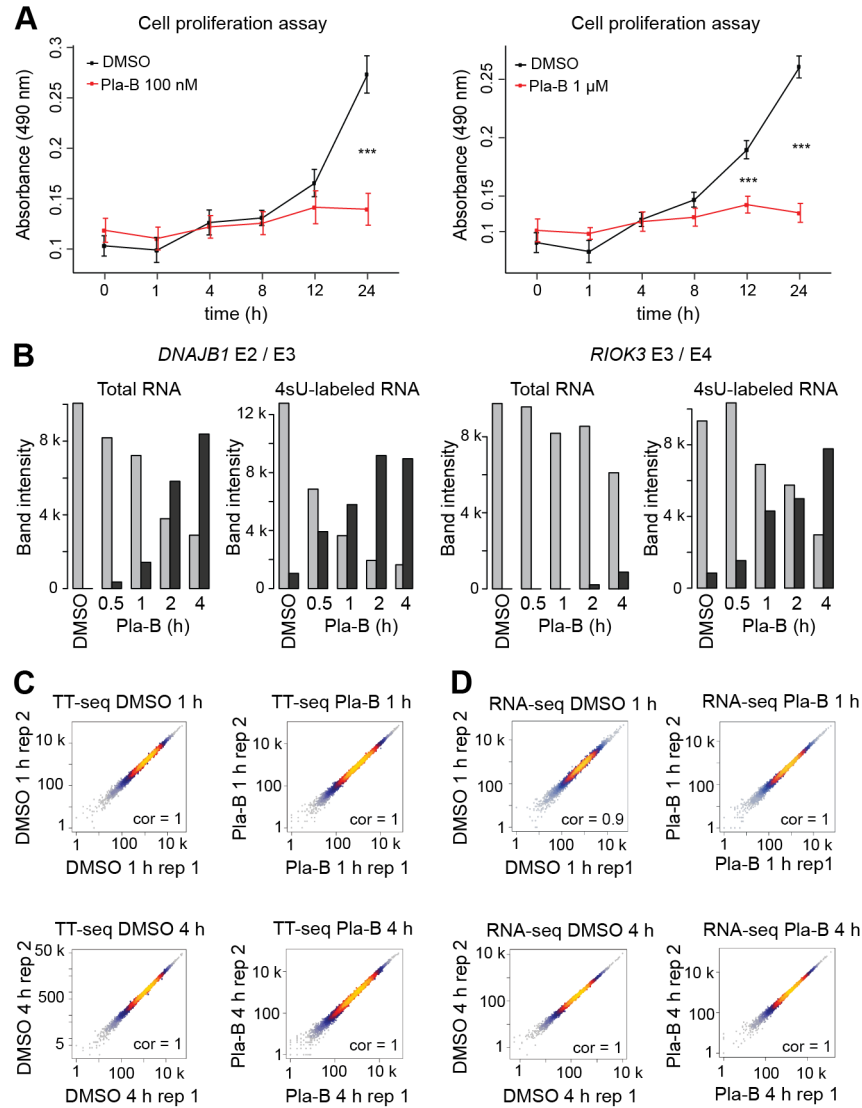


Figure S1. TT-seq and RNA-seq data are highly reproducible, related to Figure 1. (A) Proliferation assay comparing effect on cell growth of 100 nM and 1 μM Pla-B over a time course of 24 h. Geometric line represents the mean of four biological replicates and error bars represent the standard deviation. (***) p-value < 0.001 by two-way ANOVA test. **(B)** Agarose gel band intensity (ImageJ) for spliced and unspliced RT-PCR products for total and 4sU-labeled RNAs. See also **Figure 1D** and **Table S1**. **(C and D)** Scatter plots comparing TT-seq **(C)** and RNA-seq **(D)** replicates using antisense bias corrected counts for major isoforms for samples treated with DMSO or 1 μM Pla-B for 1 h (top) and 4 h (bottom). Spearman correlation of 0.9 and 1.

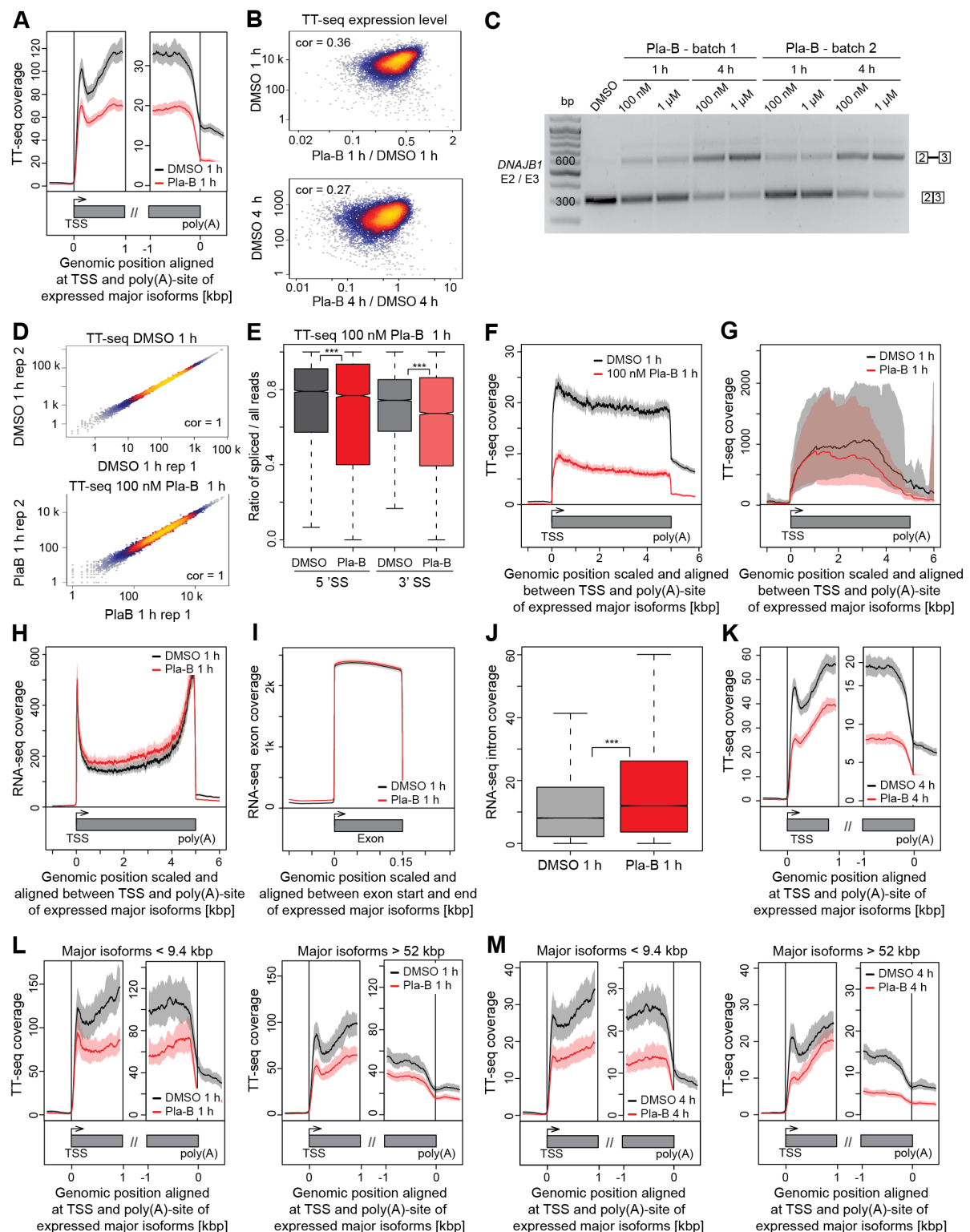


Figure S2. TT-seq monitors decrease of RNA synthesis activity upon Pla-B treatment, related to Figure 2. (A and K) Metagenesis analysis comparing DMSO and 1 μ M Pla-B treated cells for 1 h (A) and 4 h (K). TT-seq coverage was aligned at TSS for 5,465 major isoforms that exceed 1 kbp in length. Solid lines represent the averaged signal and the shaded area represents 95 % confidence interval of the mean (bootstrap). (B) Scatter plot comparing DMSO TT-seq RPK and ratio of Pla-B to DMSO TT-seq RPK upon 1 h (top) and 4 h (bottom) of

DMSO or 1 μ M Pla-B treatments. Spearman correlation of 0.36 for 1 h and 0.27 for 4 h. (C) Agarose gel showing spliced and unspliced RT-PCR products for region spanning exons 2 and 3 of *DNAJB1* of total RNA (unspliced: 597 bp, spliced: 302 bp). Pla-B concentration and batch are indicated for each time point (batch 1: G0516, batch 2: B0420). See also **Table S1**. (D) Scatter plots comparing TT-seq replicates using antisense bias corrected counts for major isoforms for samples treated with DMSO or 100 nM Pla-B (batch 2) for 1 h. Spearman correlation of 1. (E) Ratio of spliced reads over total unspliced and spliced reads upon 1 h DMSO or 100 nM Pla-B (batch 2). (***) p-value $< 2.2 \times 10^{-16}$ by Wilcoxon signed rank test. (F) Metagene analysis comparing TT-seq signal between cells treated with DMSO or 100 nM Pla-B (batch 2). TT-seq coverage was averaged for 5,535 major isoforms scaled between TSS and poly(A)-site. (G) Intronless genes meta-analysis comparing TT-seq signal between cells treated with DMSO and 1 μ M Pla-B for 1 h. TT-seq coverage was averaged for 51 intronless genes scaled between TSS and poly(A)-site. (H) Metagene analysis comparing RNA-seq signal between cells treated with DMSO and 1 μ M Pla-B for 1 h. RNA-seq coverage was averaged for 5,535 major isoforms scaled between TSS and poly(A)-site. (I) Meta-analysis comparing RNA-seq signal over exons between cells treated with DMSO and 1 μ M Pla-B for 1 h. RNA-seq coverage was averaged for 46,243 exons scaled between TSS and poly(A)-site. (J) Boxplot comparison of antisense bias corrected counts on 52,593 introns between RNA-seq samples treated with DMSO and 1 μ M Pla-B for 1 h. Black bars represent the median values for each group. Lower and upper boxes are the first and third quartiles, respectively. The ends of the whiskers extend the box by 1.5 times the interquartile range. Outliers are not drawn. (***) p-value $< 2.2 \times 10^{-16}$ by Wilcoxon signed rank test. (L and M) Metagene analysis for 1 h (L) and 4 h (M) DMSO or 1 μ M Pla-B treated cells comparing major isoforms belonging to the lower (left) and upper (right) quartiles of length. TT-seq coverage aligned at TSS and poly(A)-site for 1,314 (lower) and 1,384 (upper) major isoforms that exceed 1 kbp in length.

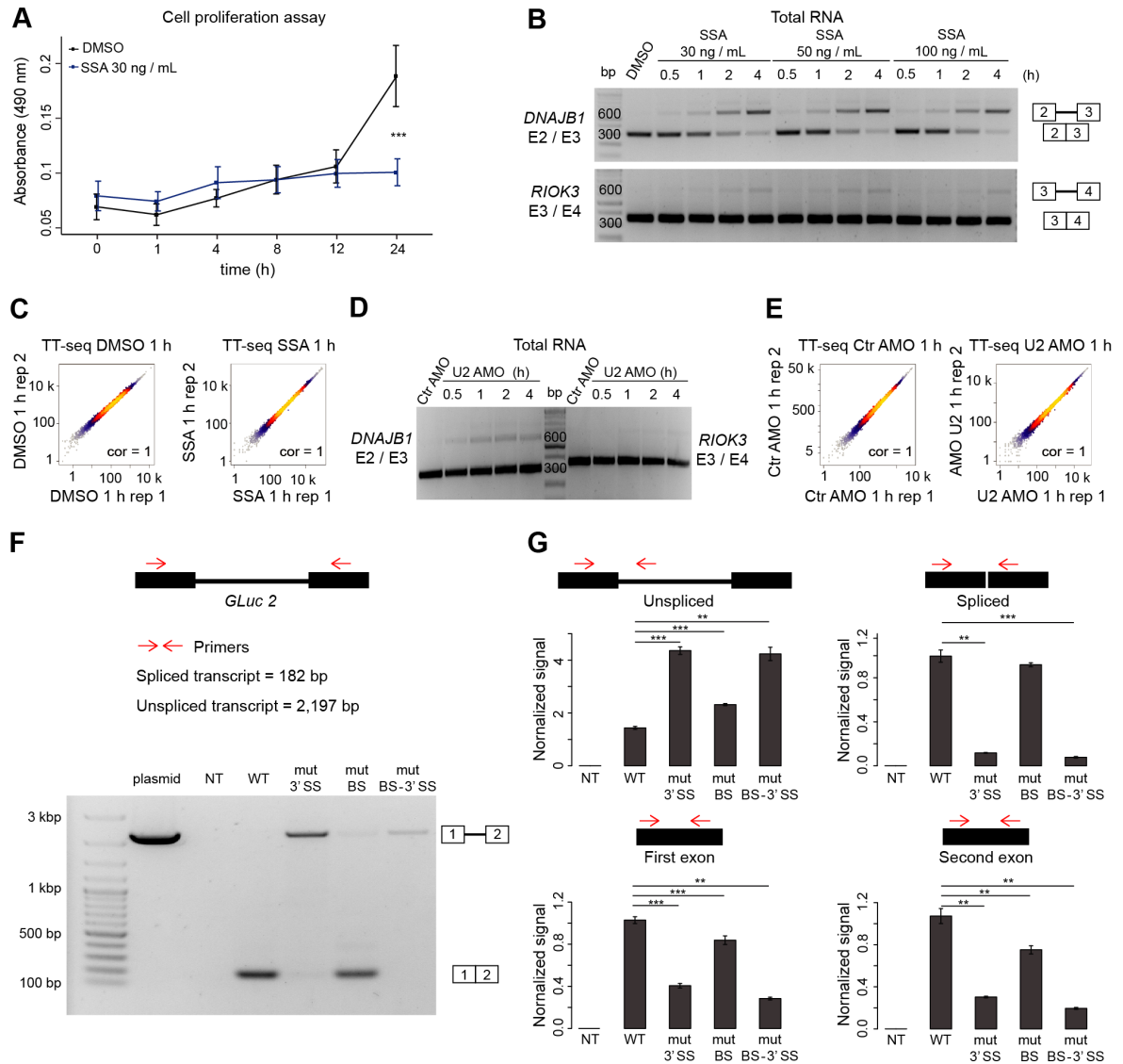


Figure S3. TT-seq monitors decrease of RNA synthesis activity upon SSA and U2 AMO treatments, related to Figure 3. (A) Proliferation assay comparing effect on cell growth of 30 ng / mL SSA over a time course of 24 h. Geometric line represents the mean of four biological replicates and error bars represent the standard deviation. (***) p-value < 0.001 by two-way ANOVA test. **(B)** Agarose gel showing spliced and unspliced RT-PCR products for total RNA upon 30 min, 1 h, 2 h and 4 h of 30 ng / mL SSA treatment and 4 h DMSO for regions spanning the exons 2 and 3 of *DNAJB1* (unspliced: 597 bp, spliced: 302 bp) and exons 3 and 4 of *RIOK3* (unspliced: 650 bp, spliced: 352 bp). See also **Table S1**. **(C and E)** Scatter plots comparing TT-seq replicates using antisense bias corrected counts for major isoforms for samples treated with DMSO or 30 ng / mL SSA **(C)** and 75 μ M Ctr AMO or 75 μ M U2 AMO **(E)** for 1 h. Spearman correlation of 1. **(D)** Agarose gel showing spliced and unspliced RT-PCR products for total RNA upon 30 min, 1 h, 2 h and 4 h of 75 μ M U2 AMO treatment and 4 h 75 μ M Ctr AMO, as in **(B)**. **(F)** pCMV-GLuc 2 plasmid containing an intron of 1,982 bp originating from the human genome position chr3: 89596735-89598673 (top, Methods). Agarose gel showing spliced and unspliced RT-PCR products for plasmid, non-transfected control (NT), wild-type consensus sequences for branch site (BS) and 3' SS (WT), mutated sequence of 3' SS (mut 3' SS), mutated sequence of BS (mut BS), and mutated both sequence of BS and 3' SS (mut BS - 3' SS)

(bottom). Image used for this figure was taken from gel of replicate 1 (see original images in Mendeley repository data). **(G)** Transcription level for unspliced and spliced transcripts (top) and for first and second exons (bottom) in NT, WT, mut 3' SS, mut BS, mut BS - 3' SS conditions. Geometric bar represents the mean of three biological replicates and error bars represent the standard deviation. (***) p-value < 0.001 by Welch's *t*-test. See also **Table S3**.

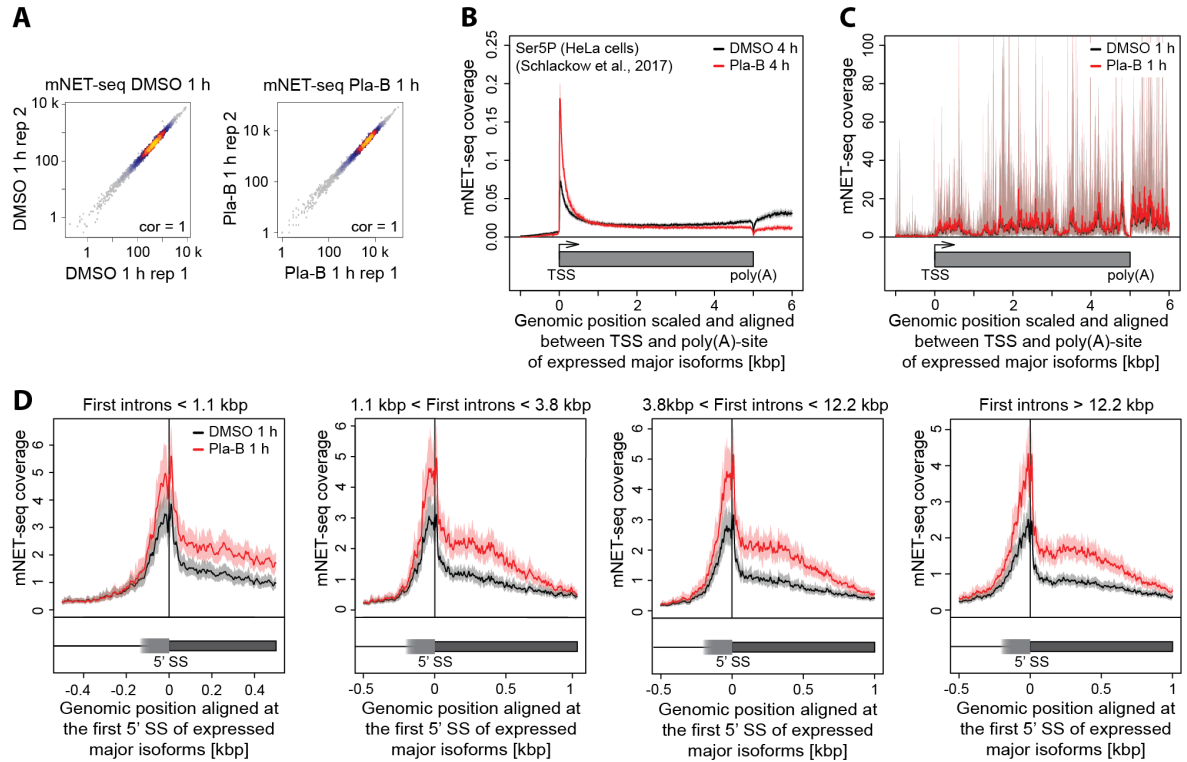


Figure S4. mNET-seq data are highly reproducible and show accumulation of Pol II beyond the 5' SS after Pla-B treatment, related to Figure 4. (A) Scatter plots comparing mNET-seq replicates using antisense bias corrected counts for major isoforms for samples treated with DMSO and 1 μ M Pla-B for 1 h. Spearman correlation of 1. (B) Metagene analysis comparing the averaged mNET-seq signal for CTD ser5P between HeLa cells treated with DMSO and 1 μ M Pla-B for 4 h. mNET-seq coverage was averaged for 5,535 major isoforms scaled between TSS and poly(A)-site. Solid lines represent the averaged signal and the shaded area represents 95 % confidence interval of the mean (bootstrap). Datasets for mNET-seq against CTD ser5P in HeLa cells upon 4 h DMSO and 1 μ M Pla-B treatments were obtained from (Schlackow et al., 2017). (C) Intronless genes meta-analysis comparing mNET-seq signal between cells treated with DMSO control and 1 μ M Pla-B for 1 h. TT-seq coverage was averaged for 51 intronless genes scaled between TSS and poly(A)-site. (D) First intron meta-analysis comparing the averaged mNET-seq signal between cells treated with DMSO and 1 μ M Pla-B for 1 h. First introns were divided into 4 quartiles based on first intron length.

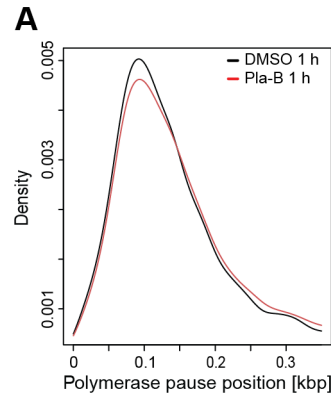


Figure S5. Polymerase pause positions are largely unaltered by Pla-B treatment, related to Figure 5. (A) Polymerase pause position calculated for 3,626 intron-containing major isoforms that exceed 10 kbp in length and had an identified pause position (Methods) for both DMSO and 1 μ M Pla-B treated samples.

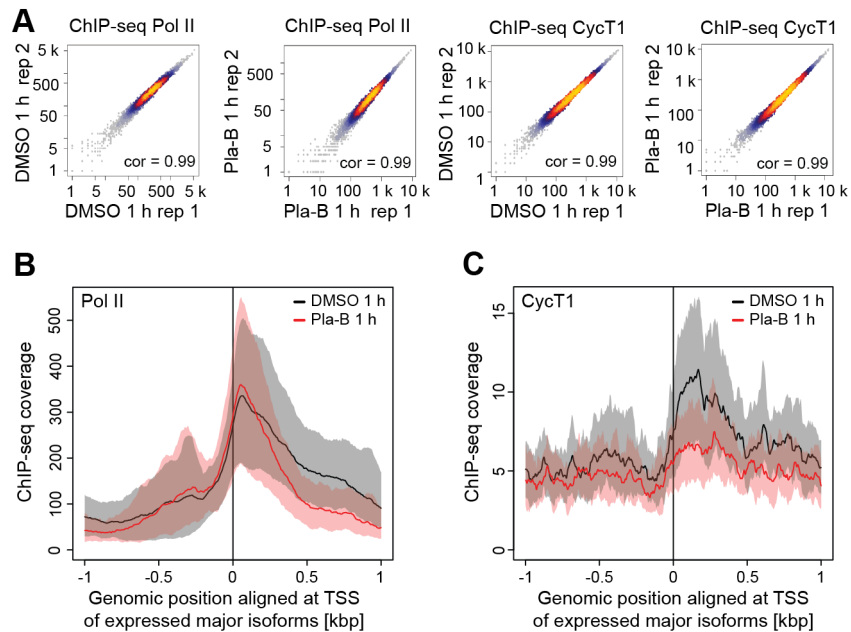


Figure S6. ChIP-seq data are highly reproducible and show a mild effect of Pla-B treatment on intronless genes, related to Figure 6. (A) Scatter plots comparing ChIP-seq replicates using counts for major isoforms for samples treated with DMSO and 1 μ M Pla-B. Spearman correlation of 0.99. (B and C) Intronless genes meta-analysis comparing ChIP-seq signal between cells treated with DMSO and 1 μ M Pla-B for 1 h. ChIP-seq coverage for Pol II (B) and CycT1 (C) was averaged for a total of 51 intronless genes that exceed 1 kbp in length. Solid lines represent the averaged signal and the shaded area represents 95 % confidence interval of the mean (bootstrap).

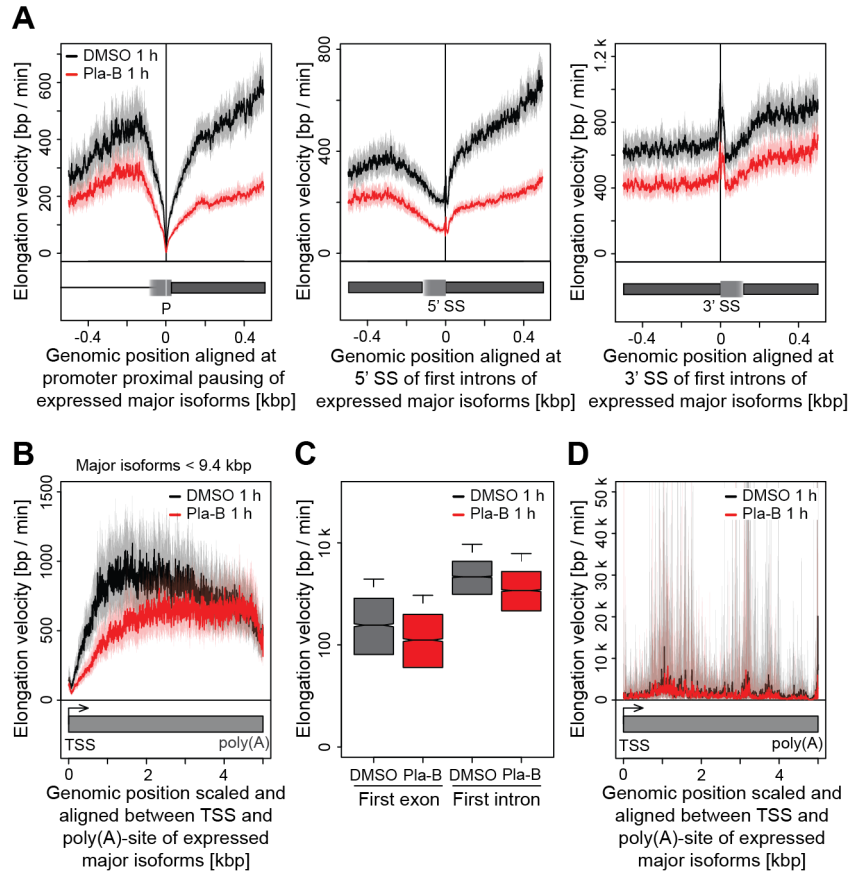


Figure S7. Pla-B treatment strongly decreases Pol II elongation velocity in the promoter-proximal region and in short transcripts, related to Figure 7. (A) Metagenome analysis comparing the elongation velocity in cells treated for 1 h with DMSO or 1 μ M Pla-B. Elongation velocity was averaged for 3,626 first exons aligned at the pause site defined on the DMSO treated samples (left). Elongation velocity was averaged for 5,350 first introns and aligned at the 5' SS (middle) or at the 3' SS (right). Solid lines represent the averaged signal and the shaded area represents 95 % confidence interval of the mean (bootstrap). **(B)** Metagenome analysis comparing the elongation velocities between cells treated with DMSO or 1 μ M Pla-B for 1 h for major isoforms belonging to the lowest quartile of major isoforms length (< 9.4 kbp). Elongation velocity was averaged for 1,384 major isoforms scaled between TSS (left) and poly(A)-site (right). **(C)** Elongation velocities for the first exon and first intron in DMSO and 1 μ M Pla-B treated cells for 1 h. Elongation velocity distribution was calculated for 5,350 first exons and first introns (unique exons were excluded). **(D)** Intronless genes meta-analysis comparing the elongation velocity in cells treated with 1 h DMSO or 1 μ M Pla-B. TT-seq coverage was averaged for 51 intronless genes scaled between TSS and poly(A)-site.

Table S1. Primers used for RT-PCR. Related to Figure 1 and Figures S1-3

Primer sequences used for total and 4sU-labelled RNA semiquantitative-PCRs.

Gene	Primer sequence
<i>DNAJB1</i> _RT-PCR_agarose gel_F	GAACCAAAATCACTTTCCCCAAGGAAGG
<i>DNAJB1</i> _RT-PCR_agarose gel_R	AATGAGGTCCCCACGTTTCTCGGGTGT
<i>RIOK3</i> _RT-PCR_agarose gel_F	GCTGAAGGACCATTATTAATTACTGGAG
<i>RIOK3</i> _RT-PCR_agarose gel_R	TTCTTGCTGTGTTCTTTCTCCCACA

Table S2. Oligonucleotides used for RNA extension assay. Related to Figure 2

Oligonucleotides sequences used for *JUNB* scaffold template DNA, 5'-biotin-labeled non-template DNA and 5'-FAM labeled RNA primer.

Name	Oligos sequence
<i>JUNB</i> -A-less-T-DNA	GAAACCCCCAGCACCCAGCACCCAGCAGGCACCGAGGCTGG CCTGGCCGCTCTCAAGGTCCCA
<i>JUNB</i> -A-less-NT-DNA	TTTTTTGGGACCTTGAGAGCGGCCAGGCCAGCCTCGGTGCCT GCTGGGTGCTGGGTGCTGGGGGTTTC
<i>JUNB</i> -U-RNA	UUUUUUUCAGGCCAGCC

Table S3. Primers used for reporter gene splicing assay. Related to Figure S3

Primer sequences used for gene reporter splicing assay real time PCRs and semiquantitative-PCRs.

Name	Primers sequence
<i>GAPDH</i> _RT-qPCR_F	TCACCAGGGCTGCTTTTA
<i>GAPDH</i> _RT-qPCR_R	TTCACACCCATGACGAACA
<i>CLuc</i> _RT-qPCR_F	gagtagaggccgcaggatggttag
<i>CLuc</i> _RT-qPCR_R	cttcagtcctgtcttggtcagcacag
<i>GLuc</i> 2_exon 1_RT-qPCR_F	caaagtctgtttgccctgatctgc
<i>GLuc</i> 2_exon 1_RT-qPCR_R	cttcccgcggtcagcatc
<i>GLuc</i> 2_exon 2_RT-qPCR_F	gcttgccaacgtgcagtgttc
<i>GLuc</i> 2_exon 2_RT-qPCR_R	cttgatctgtccacctggccc
<i>GLuc</i> 2_unspliced_RT-qPCR_F	gcgagGTAAGAGTgaatttgccaC
<i>GLuc</i> 2_unspliced_RT-qPCR_R	gaacctggattattcactggaaccaatactc
<i>GLuc</i> 2_spliced_RT-qPCR_F	gcataggcgaggcgatcgctc
<i>GLuc</i> 2_spliced_RT-qPCR_R	gtcagaacactgcacgttggc
<i>GLuc</i> 2_exon1/2_RT-PCR_agarose gel_F	gatctgcctgtccacatcaagtg
<i>GLuc</i> 2_exon1/2_RT-PCR_agarose gel_R	cgacctgtgcgatgaactgctc