Differential effects of pre-mRNA splicing inhibitors on RNA polymerase II transcription

Michael Tellier[1,2*], Gilbert Ansa[1] and Shona Murphy[1*].

[1] Sir William Dunn School of Pathology, University of Oxford, Oxford, OX1 3RE, United Kingdom

[2] Current address: Department of Molecular and Cell Biology, University of Leicester, Leicester, LE1 7RH, United Kingdom

* To whom correspondence should be addressed. Tel: (+44) 1162297011. Email: mt477@leicester.ac.uk

Correspondence may also be addressed to Shona Murphy. Tel: (+44) 1865275616 Email: shona.murphy@path.ox.ac.uk

Summary

The production of eukaryotic mRNA requires transcription by RNA polymerase (pol) II and co-transcriptional processing, including capping, splicing, and cleavage and polyadenylation (CPA). Pol II can positively affect co-transcriptional processing through interaction of factors with its carboxyl terminal domain (CTD), comprising 52 repeats of the heptapeptide Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7. Small molecule inhibitors of the splicing factor SF3B1 cause loss of the transcription elongation factor P-TEFb from protein-coding gene templates and major transcription defects, indicating that splicing can, in turn, positively affect transcription. To understand better the relationship between pre-mRNA splicing and pol II transcription, we have investigated the effect of two other splicing inhibitors, Madrasin and Isoginkgetin, on transcription. We found that Madrasin rapidly inhibits pre-mRNA splicing, whereas Isoginkgetin affects transcription before any detectable effect on pre-mRNA splicing. Interestingly, we found that both of these small molecules promote general downregulation of transcription and global transcriptional readthrough, including on intronless and histone genes. Both small molecules affect the association of the mRNA CPA complex with chromatin, likely explaining the transcription termination defect. However, splicing inhibition is not necessarily associated with transcriptional readthrough as small molecule inhibitors of SF3B1 or knockdown of splicing factors do not cause a global transcription termination defect.

Introduction

The expression of eukaryotic protein-coding genes requires both transcription by RNA polymerase (pol) II and co-transcriptional processes, which can include RNA capping, splicing, and cleavage and polyadenylation (1). Recent research has highlighted how transcription and pre-mRNA splicing can influence each other (2-7). Coupling between transcription and pre-mRNA splicing is thought to occur, at least partially, via the carboxyl-terminal domain (CTD) of the large subunit of pol II, which is composed in humans of 52 repeats of the heptapeptide Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7. The pol II CTD can be modified by several post-translational modifications, including protein phosphorylation on Tyr1, Ser2, Thr4, Ser5, and Ser7 (8). While Ser2P is associated with transcription elongation and Ser5P with transcription initiation and pol II pausing (9), pol II enriched for Ser5P has also been found to specifically interact with splicing factors (10,11), providing a physical link between the pol II complex and the spliceosome. A direct interaction between the pol II complex and the U1 snRNP was also observed by cryo-EM (12). In addition, the pol II elongation rate, which can be modified by cellular stresses or by mutating pol II, can affect the inclusion or the exclusion of some exons from the final mRNA in yeasts, mammals, and plants (2-4,6,13-17). It also works the other way round as inhibition of pre-mRNA splicing by Spliceostatin A (SSA) or Pladienolide B (PlaB), two small molecule inhibitors of the U2 snRNP protein SF3B1, affect pol II transcription (5,7,18,19), pol II CTD phosphorylation (5), and recruitment of the key transcription elongation kinase, P-TEFb (7), which is composed of cyclin-dependent kinase (CDK)9 and cyclin T1. Interestingly, P-TEFb has also been found to be recruited to chromatin by the splicing factors SFPQ (20) and SRSF2 (21,22).

Coupling between pre-mRNA splicing and cleavage and polyadenylation (CPA) is also required to define the last exon of protein-coding genes and to promote mRNA CPA and transcription termination

(1). This coupling involves interactions between splicing factors, including U2AF65 and the SF3B proteins, and CPA factors, including CPSF proteins and the poly(A) polymerase (23-26). A consequence of this coupling is that failure to splice pre-mRNA is associated with a transcription termination defect, as the poly(A) site is no longer recognised by the mRNA CPA complex (18,27,28). Transcription of intronless and histone genes is instead not sensitive to short inhibition of pre-mRNA splicing by SF3B1 inhibitors (7).

To further investigate the effects of pre-mRNA splicing inhibition on transcription elongation and transcription termination, we used two small molecule inhibitors of the early spliceosome, Madrasin (29) and Isoginkgetin (30). We have also re-analysed previously published RNA-seq datasets of different SF3B1 inhibitors, including SSA and Sudemycin C1, D1, and K (31,32). We found that while a short treatment of Madrasin decreases pre-mRNA splicing, Isoginkgetin is a poor splicing inhibitor. Inhibition of pre-mRNA splicing by Madrasin downregulates transcription and promotes a global readthrough on protein-coding genes. Surprisingly, these inhibitors also cause a transcription termination defect on intronless and histone genes. Re-analysis of ChIP-seq of some splicing factors shows that these RNA-binding proteins are also found on intronless and histone genes, indicating that the presence of introns is not required for their binding. Inhibition of pre-mRNA splicing with Madrasin affects the chromatin association of the elongation factors SPT5 and CDC73 and of the mRNA CPA factor, CPSF73, on intron-containing, intronless, and histone genes, potentially explaining the transcriptional readthrough induced by this splicing inhibitor. While Isoginkgetin does not inhibit pre-mRNA splicing, we found that this small molecule generally downregulates transcription and also promotes a global transcriptional readthrough on intron-containing, intronless, and histone genes. Interestingly, Isoginkgetin also affects the chromatin association of mRNA CPA factors. In contrast, re-analysis of RNA-seq after treatment of cells with different SF3B1 inhibitors show a limited transcription termination defect, detected only on a few intron-containing genes, while re-analysis of GRO-seq following knockdown of splicing factors do not show any transcriptional readthrough. These results indicate that splicing inhibition is not necessarily associated with a transcription termination defect.

MATERIAL AND METHODS

Cell culture

HeLa cells were obtained from ATCC (ATCC® CCL-2™) and grown in DMEM medium supplemented with 10% foetal calf serum, 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mM L-glutamine at 37°C and 5% Co2. HeLa cells were treated with 30-90 µM of Madrasin (Sigma-Aldrich) or 30 µM of Isoginkgetin (Merck) for the time indicated in the figures. As a negative control, HeLa cells were treated with DMSO. Cells were routinely checked for mycoplasma contamination using Plasmo Test Mycoplasma Detection Kit (InvivoGen, rep-pt1).

Protein extraction and Western blot

Western blot analysis was performed on chromatin extracts as previously described in the mNET-seq procedure (33). A ~ 80% confluent 15 cm dish was washed twice with ice-cold PBS and scrapped in 5 ml of ice-cold PBS. The cells were pelleted at 420 g for 5 min at 4°C. After discarding the supernatant, the cells were resuspended in 4 ml of ice-cold HLB + N buffer (10 mM Tris–HCl (pH 7.5), 10 mM NaCl,

2.5 mM MgCl2 and 0.5% (vol/vol) NP-40) and incubated on ice for 5 min. The cell pellets were then underlayed with 1 ml of ice-cold HLB + NS buffer (10 mM Tris–HCl (pH 7.5), 10 mM NaCl, 2.5 mM MgCl2, 0.5% (vol/vol) NP-40 and 10% (wt/vol) sucrose). Following centrifugation at 420 g for 5 min at 4°C, the nuclear pellets were resuspended by pipetting up and down in 125 µl of NUN1 buffer (20 mM Tris–HCl (pH 7.9), 75 mM NaCl, 0.5 mM EDTA and 50% (vol/vol) glycerol) and moved to a new 1.5 ml ice-cold microcentrifuge tube. Following the addition of 1.2 ml of ice-cold NUN2 buffer (20 mM HEPES-KOH (pH 7.6), 300 mM NaCl, 0.2 mM EDTA, 7.5 mM MgCl2, 1% (vol/vol) NP-40 and 1 M urea), the tubes were vortexed at maximum speed for 10 s and incubated on ice for 15 min with a vortexing step of 10 s every 3 min. The samples were centrifuged at 16,000 g for 10 min at 4°C and the supernatant discarded while the chromatin pellets were washed with 500 µl of ice-cold PBS and then with 100 µl of ice-cold water. The chromatin pellet was then digested in 100 µl of water supplemented with 1 µl of Benzonase (25–29 units, Merck Millipore) for 15 min at 37°C in a thermomixer at 1,400 rpm. For loading, 10 µg of proteins were boiled in 1× LDS plus 100 mM DTT.

For whole-cell extract, cells were washed in ice-cold PBS twice, collected in ice-cold PBS with an 800 g centrifugation for 5 min at 4°C. The pellets were resuspended in RIPA buffer supplemented with protease inhibitor cocktail and PhosSTOP, kept on ice for 30 min with a vortexing step every 10 min. After centrifugation at 14,000 g for 15 min at 4°C, the supernatants were kept and quantified with the Bradford method. For loading, 20 µg of proteins were boiled in 1× LDS plus 100 mM DTT.

Western blots were performed with NuPAGE Novex 4–12% Bis–Tris Protein Gels (Life Technologies) with the following primary antibodies: Rpb1 NTD (D8L4Y) Rabbit mAb (14958S, Cell Signaling Technology), Phospho-Rpb1 CTD (Ser2) (E1Z3G) Rabbit mAb (13499S, Cell Signaling Technology), Phospho-Rpb1 CTD (Ser5) (D9N5I) Rabbit mAb (13523S, Cell Signaling Technology), Rabbit anti-CPSF73 Antibody, Affinity Purified (A301-091A, Bethyl Laboratories), Rabbit anti-CPSF100 Antibody, Affinity Purified (A301-581A, Bethyl Laboratories), SPT6 (D6J9H) Rabbit mAb (15616S, Cell Signaling Technology), Proteintech SUMO1 Monoclonal antibody (67559-1-IG-20UL, Proteintech Europe), Proteintech SUMO2/3 Polyclonal antibody (11251-1-AP-20UL, Proteintech Europe), and Histone H3 (ab1791, Abcam). Secondary antibodies were purchased from Merck Millipore (Goat Anti-Rabbit IgG Antibody, HRP-conjugate, 12–348, and Goat Anti-Mouse IgG Antibody, HRP conjugate, 12–349), the chemiluminescent substrate (SuperSignal West Pico PLUS) from Thermo Fisher, and the membranes visualized on an iBright FL1000 Imaging System (Thermo Fisher). Ponceau S – Proteomics Grade staining (VWR International) was performed according to the manufacturer's instructions.

Quantifications were performed with Image Studio Lite version 5.2.5. For each antibody, the quantification has been calculated as: (IP treated signal / histone H3 treated signal) / (IP DMSO signal / histone H3 DMSO signal).

Chromatin immunoprecipitation followed by qPCR (ChIP-qPCR)

ChIP was performed as previously described (34). HeLa cells were grown in 100- or 150-mm dishes until they reached ~80% confluence. The cells were fixed with 1% formaldehyde at room temperature for 10 minutes with shaking. Quenching was performed with 125 mM Glycine for 5 minutes at room temperature with shaking. Cells were washed twice with ice-cold PBS and then scrapped in ice-cold PBS and pelleted for 10 minutes at 1,500 rpm for 10 min at 4°C. Lysis was performed on ice for 10

minutes with ChIP lysis buffer (10 mM Tris–HCl ph8.0, 0.25% Triton X-100, 10 mM EDTA, and protease inhibitor cocktail) before being centrifuged at 1,500 g for 5 min at 4°C. Pellets were washed once with ChIP Wash buffer (10 mM Tris–HCl pH8.0, 200 mM NaCl, 1 mM EDTA, and protease inhibitor cocktail). Pellets were then resuspended in ChIP Sonication buffer (10 mM Tris–HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, protease inhibitor cocktail, and phosphatase inhibitor), incubated 10 minutes on ice, and sonicated for 30 cycles, 30 seconds on/30 seconds off using a Bioruptor Pico (Diagenode). The samples were centrifuged at 13,000 rpm for 15 minutes at 4°C. Supernatants were transferred to new Eppendorf tubes and pre-cleared with 10 µl of Protein G Dynabeads for 30 minutes on a rotating wheel at 4°C. For each IP, 60 µg of chromatin was incubated overnight on a rotating wheel at 4°C with the following antibodies: normal rabbit IgG (2729S, Cell Signaling), RNA polymerase II antibody (NBP2-32080, Novus Biologicals), CDC73 (A300-170A, Bethyl Laboratories), SUPT5H antibody (A300-868A, Bethyl Laboratories), CPSF73 (A301-091A, Bethyl Laboratories), and CPSF100 (A301-581A, Bethyl Laboratories). Protein G Dynabeads were saturated overnight with RIPA buffer (10 mM Tris–HCl pH8.0, 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% Triton X-100, 0.1% sodium deoxycholate) supplemented with 10 mg/ml of bovine serum albumin on a rotating wheel at 4°C.

The chromatin-antibody mix was incubated for 1h on a rotating wheel at 4°C with pre-blocked Protein G Dynabeads. IgG supernatant was kept as input. Beads were then washed three times with ice-cold RIPA buffer, three times with ice-cold High Salt buffer (10 mM Tris–HCl pH8.0, 500 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% Triton X-100, 0.1% sodium deoxycholate), twice with LiCl Wash buffer (10 mM Tris–HCl pH8.0, 250 mM LiCl, 1 mM EDTA, 1% NP-40, 1% sodium deoxycholate), and twice with TE buffer (10 mM Tris–HCl pH 7.5, 1 mM EDTA). The samples were eluted twice with Elution buffer (100 mM NaHCO3, 1% SDS, 10 mM DTT) from the beads on a thermomixer for 15 minutes at 25°C at 1,400 rpm. Each input were diluted one in ten in Elution buffer. The samples were treated with RNase A for 30 minutes at 37°C before reversal of the crosslink with 200 mM NaCl for 5h at 65°C. DNA-protein mix was precipitated with 2.5x volume of 100% ethanol overnight at -20°C. The ethanol was removed after centrifugation for 15 minutes at 13,000 rpm 4°C. The pellets were resuspended in TE, 5x Proteinase K buffer (50 mM Tris–HCl pH 7.5, 25 mM EDTA, 1.25% SDS) and Proteinase K (20 mg/ml), and incubated for 2h at 45°C. DNA was purified with a QIAGEN PCR purification kit and kept at -20°C.

ChIP samples were analysed by real-time qPCR with a QuantiTect SYBR Green PCR kit (QIAGEN) and a Rotor-Gene RG-3000 (Corbett Research). Each reaction was performed with: 1 µl of template, 1 µl of primer pair mix (10 µM), 3 µl of water and 5 µl of SYBR Green Mix (2×). The sequences of primers used in this study can be found in Table 1. The thermo-cycling parameters were: 95°C for 15 min followed by 40 cycles of 94°C for 15 s, 57°C for 20 s and 72°C for 25 s. The Roto-Gene Q Series Software was used to calculate the threshold cycle (Ct) value. Signals are presented as a percentage of Input DNA after removal of the IgG background signal. Each ChIP sample was measured in triplicate by qPCR. The number of biological replicates is indicated on each figure.

RNA preparation and PCR/qPCR

Total RNA was extracted from ~70% confluent HeLa cells grown in 6-wells plate using a RNeasy Micro Kit (Qiagen) with a DNase step according to the manufacturer's instructions. Reverse-transcription (RT) was performed with 500 ng of RNA using random hexamers or oligo(dT) with the SuperScript III kit

(Invitrogen) according to the manufacturer's instructions. PCR reactions were performed with the Phusion High-Fidelity DNA polymerase (NEB) according to the manufacturer's instructions. The sequence of primers used for PCR is given in Table 1. Quantifications were performed with Image Studio Lite version 5.2.5. The quantifications are shown as % of unspliced reads and calculated as follow: (unspliced RNA signal) / (spliced RNA signal + unspliced RNA signal) * 100.

cDNA was amplified by qPCR with a QuantiTect SYBR Green PCR kit (QIAGEN) and a Rotor-Gene RG-3000 (Corbett Research). The sequence of primers used for qPCR is given in Table 1. Values are normalized to the 7SK non-coding RNA (random primers), or the GAPDH mRNA (oligo d(T)), that have been used as control. Experiments were replicated at least three times to ensure reproducibility, and each RNA sample was measured in triplicate by qPCR.

5'EU

HeLa cells were seeded at 40-50% confluency on coverslips in a 6-well plate, then treated the following day with Madrasin or Isoginkgetin for the time indicated on the figure before being incubated with 1 mM 5-ethylnyl uridine (5-EU) for 1 h to label newly synthesized RNA. Cells were then fixed by transferring coverslips to 3.7 % formaldehyde diluted in PBS for 15 min, washed with PBS, and then permeabilised in 0.5 % Triton X-100 diluted in PBS for 15 min at room temperature. Global nascent RNA transcription was detected using Click-iT® RNA Alexa Fluor 488 Imaging kit (ThermoFisher, C10329) according to the manufacturer's instructions. Following Click-iT reaction, cellular DNA was immunostained with Hoechst 33342 diluted 1:1000 in PBS for 15 min at room temperature, protected from light. Coverslips were then washed twice with PBS, then mounted onto slides using Fluoromount-G (SouthernBiotech, 0100-01) containing DAPI. Slides were visualised with an inverted multi-channel fluorescence microscope (Evos M7000; ThermoFisher). Fluorescent intensity was analysed and quantified with ImageJ (version 1.53q) by first using the DAPI channel (λEx/λEm (with DNA) = 350/461 nm) image to define the nucleus, then GFP channel (λEx/λEm = 495/519 nm) to measure the fluorescent intensity within the nucleus. Data were plotted by box-and-whisker plot with GraphPad Prism 9.4.1 software with the following settings: boxes: 25-75 percentile range; whiskers: min-max values; horizontal bars: median.

mNET-seq

mNET-seq was carried out as previously described (33). In brief, the chromatin fraction was isolated from 4×10^7 HeLa cells treated with DMSO or Madrasin for 30 minutes. Chromatin was digested in 100 µl of MNase (40 units/µl) reaction buffer for 2 min at 37°C in a thermomixer at 1,400 rpm. MNase was inactivated with the addition of 10 µl EGTA (25mM) and soluble digested chromatin was collected after centrifugation at 13 000 rpm for 5 minutes at 4°C. The supernatant was diluted with 400 µl of NET-2 buffer and antibody-conjugated beads were added. Antibodies used: Pol II (MABI0601, MBL International) and Ser5P (MABI0603, MBL International). Immunoprecipitation was performed at 4°C on a rotating wheel at 16 rpm for 1h. The beads were washed six times with 1 ml of ice-cold NET-2 buffer and once with 100 µl of 1x PNKT (1x PNK buffer and 0.05% Triton X-100) buffer. Washed beads were incubated in 200 µl PNK reaction mix for six minutes in a Thermomixer at 1,400 rpm at 37°C. The beads were then washed once with 1 ml of NET-2 buffer and RNA was extracted with Trizol reagent. RNA was suspended in urea Dye and resolved on a 6% TBU gel (Invitrogen) at 200 V for 5 minutes. In

order to size select 35–100 nt RNAs, a gel fragment was cut between BPB and XC dye markers. For each sample, several small holes were made with a 25G needle at the bottom of a 0.5 ml tube, which is then placed in a 1.5 ml tube. Gel fragments were placed in the layered tube and broken down by centrifugation at 12,000 rpm for one minute. The small RNAs were eluted from the gel using RNA elution buffer (1 M NaOAc and 1 mM EDTA) for 1 hour on a rotating wheel at 16 rpm at room temperature. Eluted RNAs were purified with SpinX column (Coster) and two glass filters (Millipore) and the resulting flow-through RNA was precipitated overnight with ethanol. RNA libraries were prepared according to manual of Truseq small RNA library prep kit (Illumina). Deep sequencing (Hiseq4000, Illumina) was conducted by the high throughput genomics team of the Wellcome Trust Centre for Human Genetics (WTCHG), Oxford.

Bioinformatics analyses

Transcription units' annotation

Gencode V35 annotation, based on the GRCh38 version of the human genome, was used to obtain the list of protein-coding genes. To obtain the list of expressed mRNA isoforms, Salmon version 0.14.1 (35) was used on two HeLa nucleoplasm RNA-seq (GSE110028 (36)) and the most expressed mRNA isoform of each gene was kept for further analysis. HeLa chromatin RNA-seq signal (GSE110028 (36)) was then quantified with bedtools mutlicov to keep only the most expressed mRNA isoform with more than 10 reads in two biological replicates.

For the intron-containing, intronless, and histone genes, the mRNAs isoforms were separated in 193 intronless genes, 40 histone genes, and 2,236 intron-containing genes with a maximal gene size corresponding to the size of the longest intronless gene. The list of used exons was extracted from Gencode V35 by using the annotation of each most expressed mRNA isoform. The first and final exons of each mRNA isoform were then removed to keep only internal exons.

mNET-seq analysis

Adapters were trimmed with Cutadapt version 1.18 (37) in paired-end mode with the following options: --minimum-length 10 -q 15,10 -j 16 – A GATCGTCGGACTGTAGAACTCTGAAC – a AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC. Trimmed reads were mapped to the human GRCh38.p13 reference sequence with STAR version 2.7.3a (38) and the parameters: --runThreadN 16 --readFilesCommand gunzip -c -k --limitBAMsortRAM 20000000000 --outSAMtype BAM SortedByCoordinate. SAMtools version 1.9 (39) was used to retain the properly paired and mapped reads (-f 3). A custom python script (33) was used to obtain the 3' nucleotide of the second read and the strandedness of the first read. Strand-specific bam files were generated with SAMtools. FPKM-normalized bigwig files were created with deepTools2 version 3.4.2 (40) bamCoverage tool with the parameters -bs 1 -p max --normalizeUsing RPKM.

ChIP-seq analysis

ChIP-seq of Input, SFPQ, SRSF1, SRSF3, SRSF4, SRSF7, SRSF9, U2AF1, U2AF2, hnRNPC, hnRNPH1, hnRNPK, hnRNPL, hnRNPLL, hnRNPUL1, and snRNP70 in K562 and HepG2 cells are from GSE120104 (41). Adapters were trimmed with Cutadapt with the following options: --minimum-length 10 -q 15,10 -j 16 -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA. Trimmed reads were mapped to the human GRCh38.p13 reference genome with STAR and the parameters: --runThreadN 16 --

readFilesCommand gunzip -c -k –alignIntronMax 1 --limitBAMsortRAM 20000000000 --outSAMtype BAM SortedByCoordinate. SAMtools was used to retain the properly mapped reads and to remove PCR duplicates. Reads mapping to the DAC Exclusion List Regions (accession: ENCSR636HFF) were removed with BEDtools version 2.29.2 (42). FPKM-normalized bigwig files were created with deepTools2 bamCoverage tool with the parameters -bs 10 -p max –e 200 --normalizeUsing RPKM.

RNA-seq analysis

Chromatin and nucleoplasm RNA-seq in HeLa cells were obtained from GSE110028 (36). RNA-seq in HeLa cells for DMSO, Methanol, and eight hours treatment with Isoginkgetin or spliceostatin A (SSA) were taken from GSE72055 (31). RNA-seq in HeLa cells for DMSO and 3h treatment with SSA, Sudemycin (Sud) C1, or SudK were taken from E-MTAB-6060 (32). RNA-seq in Rh18 cells for DMSO and Sudemycin D1treated for 8h were obtained from GSE102539 (43).

Adapters were trimmed with Cutadapt in paired-end mode with the following options: --minimum-length 10 -q 15,10 -j 16 -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA -A AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT. Trimmed reads were mapped to the human GRCh38.p13 reference sequence with STAR and the and the ENCODE parameters: --runThreadN 16 --limitBAMsortRAM 20000000000 --outSAMtype BAM SortedByCoordinate --quantMode GeneCounts --outFilterMultimapNmax 20 --outFilterType BySJout --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --outFilterMismatchNmax 999 -- alignIntronMin 20 --alignIntronMax 1000000 --alignMatesGapMax 1000000. SAMtools was used to retain the properly paired and mapped reads (-f 3) and to produce strand-specific bam files. HTSeq version 1.99.2 (44) count tool and the Gencode v35 annotation has been used to obtain the number of counts per gene. DESeq2 version 1.30.1 (45) was then used to perform the differential expression analysis. Any gene with a two-fold fold change and an adjusted p-value < 0.05 was considered as significant. FPKM-normalized bigwig files were created with deepTools bamCoverage tool with the parameters -bs 10 -p max --normalizeUsing RPKM.

GRO-seq analysis

GRO-seq of Control, SRSF4 KD, U2AF1 KD, U2AF2 KD, hnRNPH1 KD, hnRNPK KD, hnRNPL KD, and hnRNPLL KD in HepG2 cells are from GSE120104 (41) and were analysed as in (46). Adapters were trimmed with Cutadapt with the following options: --minimum-length 10 -q 15,10 -j 16 -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA. The remaining rRNA reads were removed by mapping the trimmed reads to the rRNA genes defined in the Human ribosomal DNA complete repeating unit (GenBank: U13369.1) with STAR and the parameters --runThreadN 16 --readFilesCommand gunzip -c –k --outReadsUnmapped Fastx --limitBAMsortRAM 20000000000 --outSAMtype BAM SortedByCoordinate.

The unmapped reads were mapped to the human GRCh38.p13 reference genome with STAR and the parameters: --runThreadN 16 --limitBAMsortRAM 20000000000 --outSAMtype BAM SortedByCoordinate. SAMtools was used to retain the properly mapped reads and merged biological replicates. Reads mapping to the DAC Exclusion List Regions (accession: ENCSR636HFF) were removed with BEDtools. FPKM-normalized bigwig files were created with deepTools2 bamCoverage tool with the parameters -bs 10 -p max –e 200 --normalizeUsing RPKM.

Splicing efficiency

RNA-seq splicing efficiency was calculated by first parsing each bam file to obtain the list of spliced and unspliced reads with the awk command (awk '/^@/|| $6 ~ /N/' for spliced reads and awk '/^@/|| $6 !~ /N/' for unspliced reads). The splicing efficiency was then calculated as the number of spliced reads over total reads with BEDtools multicov –s –split.

Quantification

mNET-seq signal was quantified with bedtools multicov tool and normalized to 100 million reads and to the length of the quantified region (RP100M per bp). The regions used for quantification were defined as the 5' splice site as defined in the transcription units' annotation part. Boxplots were produced in GraphPad Prism 9.5.

Metagene profiles

Metagene profiles of genes scaled to the same length were generated with Deeptools2 computeMatrix tool with a bin size of 10 bp and the plotting data obtained with plotProfile –outFileNameData tool. Graphs representing the (IP / Input) signal (ChIP-seq), the RNA-seq, or the mNET-seq signal were then created with GraphPad Prism 9.5. Metagene profiles are shown as the average of two biological replicates.

P values and significance tests

P-values were computed with unpaired two-tailed Student's t test, Wilcoxon rank sum test, or Wilcoxon signed-rank test, as indicated in the legends. Statistical tests were performed in GraphPad Prism 9.5.

RESULTS

Characterization of the pre-mRNA splicing inhibitors Isoginkgetin and Madrasin

We first characterised whether a short treatment with Madrasin or Isoginkgetin inhibits pre-mRNA splicing in HeLa cells. Treatment of HeLa cells with 30 μM, 60 μM, or 90 μM Madrasin for 30 minutes, followed by RT-PCR analysis, indicates concentration-dependent inhibition of splicing of the intron located between exons 4 and 5 of *BRD2* (Figure 1A and 1B). To confirm this result, we performed mNET-seq in HeLa cells following 30 minutes treatment with DMSO (Control) or 90 μM Madrasin (Figure 1C-E), using an anti-pol II Ser5P antibody, whose peaks at 5' splice sites correspond to the first cleavage step of co-transcriptional splicing (11). Treatment with Madrasin results in an approximately 50% decrease in Ser5P signal at 5' splice sites, with some 5' splice sites more affected than others (see single gene examples in Figure 1E). The effect of Madrasin on pre-mRNA splicing was also confirmed by a general decrease in the association of mature small nuclear (sn)RNAs with pol II Ser5P following treatment (Supplementary Figure 1), a phenotype also observed with the SF3B1 inhibitor PlaB (11).

In contrast, treatment of HeLa cells with 30 μM Isoginkgetin, from 1h to 6h, only starts to cause a detectable inhibition of pre-mRNA splicing from 6h as analysed by RT-PCR on *DNAJB1* and *BRD2* genes (Figure 1F and 1G). Re-analysis of previously published total RNA-seq datasets (31,47) in HeLa cells following 6h, 12h, or 18h treatment with 30 μM Isoginkgetin show that pre-mRNA splicing inhibition by Isoginkgetin also starts to be detectable at 6h but is only significantly affected at 12h and 18h timepoints (Figure 1H). These results indicate that Isoginkgetin unlikely acts on pre-mRNA splicing as

a primary effect, in agreement with a previous observation indicating that Isoginkgetin affects transcription rather than pre-mRNA splicing after 6h treatment (47).

Inhibition of pre-mRNA splicing with Madrasin affects transcription of protein-coding genes and promotes transcriptional readthrough

To investigate the effect of pre-mRNA splicing inhibition on nascent transcription, we first performed 5'EU incorporation in HeLa cells treated with DMSO, 100 μM DRB, which inhibits CDK9 and therefore transcription elongation by pol II, or 90 μM Madrasin for 1h (Figure 2A and 2B). Madrasin treatment decreases 5'EU incorporation, indicating a reduction of nascent transcription, but not as strongly as DRB. To determine whether Madrasin affects ribosomal (r)DNA and/or protein-coding genes transcription, we performed qRT-PCR with random primers following 30- or 60-minutes treatment with Madrasin (Figure 2C and D). We found that inhibition of pre-mRNA splicing by Madrasin affects the expression of protein-coding genes but not transcription of rDNA by pol I. To understand better the effect of Madrasin on nascent transcription, we performed mNET-seq in HeLa cells with an anti-total pol II antibody following 30 minutes treatment with Madrasin (Figure 2E). Madrasin treatment causes a decrease in pol II pausing and in pol II level across the gene body while the pol II level increases downstream of poly(A) site, indicating a transcription termination defect. We confirmed these results by pol II ChIP-qPCR on the protein-coding gene *KPNB1* (Figure 2F and 2G).

Transcription of intronless and histone genes is also affected by Madrasin

As inhibition of pre-mRNA splicing by Madrasin promotes a general transcription termination defect on intron-containing protein-coding genes, we have analysed the effect of this drug on intronless genes with poly(A) sites and on intronless replication-activated histone genes. As intron-containing genes are generally longer than intronless genes, we selected three groups of expressed genes in HeLa cells: histone genes (n=40), intronless genes (n=193), and intron-containing protein-coding genes with a gene size inferior or equal to the longest intronless genes (n=2,236) (Supplementary Figure 2A). The RNA level from the intron-containing and intronless genes is similar in chromatin RNA-seq, while histone genes produce more RNA (Supplementary Figure 2B). As expected, transcription termination of the intron-containing genes is affected by Madrasin (Supplementary Figure 2C). Surprisingly, Madrasin also affects nascent transcription of intronless genes with a general decrease in RNA production coupled to a transcription termination defect (Figure 3A and Supplementary Figure 2D). Interestingly, Madrasin activates the expression of some intronless genes, such as *JUN*, while also promoting transcriptional readthrough (Figure 3B). We confirmed these results with pol II ChIP-qPCR (Figure 3C). Histone genes are also affected by Madrasin, mostly promoting a transcription termination defect without affecting the level of nascent transcripts across the coding region (Figure 3D and Supplementary Figure 2E). The transcription termination defect on histone genes was confirmed by pol II ChIP-qPCR on the *HIST1H1C* gene (Figure 2E).

Splicing factors are recruited on intronless and histone genes

While the cellular target(s) of Madrasin is unknown, this small molecule inhibits both pre-mRNA splicing and transcription termination of intronless and histone genes. We have therefore investigated whether splicing factors are recruited exclusively to intron-containing genes by re-analysing ChIP-seq of multiple splicing factors and hnRNP proteins, including SRSF1, SRSF3, SRSF4, SRSF7, SRSF9, hnRNPC, hnRNPH1, hnRNPK, hnRNPLL, hnRNPL, hnRNPUL1, SFPQ, snRNP70, U2AF1, and U2AF2 performed in HepG2 and K562 cell lines (41) (Supplementary Figure 3). Surprisingly, these splicing factors are generally associated with the chromatin of intron-containing, intronless, and histone genes in both cell lines, indicating that the presence of introns is not required for their recruitment.

Inhibition of pre-mRNA splicing with Madrasin affects the chromatin association of elongation and termination factors

As inhibition of pre-mRNA splicing affects transcription elongation and termination, we investigated by ChIP-qPCR whether the recruitment of two elongation factors, SPT5 and CDC73, and of the pre-mRNA cleavage enzyme, CPSF73, are affected by Madrasin on an intron-containing gene, *KPNB1*, an intronless gene, *JUN*, and a histone gene, *HIST1H1C*, (Figure 4 and Supplementary Figure 4). When factor occupancy is ratioed to pol II, Madrasin treatment causes an increase in SPT5, CDC73, and CPSF73 levels on *KPNB1* across the gene body and close to the poly(A) site (up to the primer pA + 2.6) while further downstream of the poly(A) site (primers pA + 4.1 and TD), the recruitment of these factors is either unchanged or decreased compared to the control. For *JUN*, while transcription is stimulated by Madrasin, there a decrease in recruitment of SPT5, CDC73, and CPSF73 to the chromatin when ratioed to total pol II. For *HIST1H1C*, a slight decrease in SPT5 and CDC73 is observed after Madrasin treatment when ratioed to pol II while for CPSF73, the decrease is specific to primer T1, which is located downstream of the TES where RNA cleavage occurs. In the case of CPSF73, the effect of Madrasin is explained by the change in pol II signal as the CPSF73 signal remains similar between the control and treatment conditions (Supplementary Figure 4C).

Isoginkgetin treatment affects pol I and pol II transcription

As treatment up to 6h with Isoginkgetin does not markedly affect pre-mRNA splicing (Figure 1F-H), we have investigated the effect of this drug on transcription. We first monitored 5'EU incorporation in HeLa cells following a time-course of 1h to 6h treatment with Isogkingetin (Figure 5A and 5B). It takes approximately 5h of Isoginkgetin treatment to see a marked decrease in 5'EU incorporation. We next investigated by qRT-PCR whether Isoginkgetin affects pol I and/or pol II transcription (Figure 5C and 5D). Transcription of rDNA decreases after 2h treatment with Isoginkgetin while a detectable effect on the production of poly(A)+ mRNA takes at least 3h (*NR4A3, LDLR*, and *KPNB1*). There is a general downregulation for protein-coding genes following Isoginkgetin. However, *JUN* is upregulated, although this appears not to be at the level of transcription as measured by pol II ChIP (Figure 5E). A general downregulation of expression of protein-coding genes by this drug is also apparent after re-analysis of previously published RNA-seq datasets (6h, 12h, and 18h Isoginkgetin treatment), which shows that protein-coding genes are generally downregulated while non-coding genes tend to be upregulated (Supplementary Figure 5A and 5B). *ATF3* expression was previously shown to be upregulated in HeLa

cells after a 24h treatment with Isoginkgetin (48), while we observed a decrease in ATF3 expression after 6h treatment (Figure 5D). We performed qRT-PCR following a 24h treatment with Isoginkgetin and found that both *JUN* and *ATF3* are upregulated while four other protein-coding genes are downregulated (Supplementary Figure 5C). As changes in the level of poly(A)+ mRNA take more time to appear than changes in transcription, we performed pol II ChIP-qPCR following Isoginkgetin treatment (Figure 5E-G and Supplementary Figure 5D). We did not observe any change in total pol II across *KPNB1* after treating HeLa cells with Isoginkgetin for 30 minutes (Supplementary Figure 5D). However, an Isoginkgetin time-course from 1h to 6h shows a transcription termination defect after treatment for 1h on *KPNB1* (primers pA +1.4 to TD), *JUN* (primer TD), and *HIST1H1C* (primer TD2) (Figure 5E-G). From the 2h timepoint onwards, the pol II level across these three genes decreases, indicating a loss of transcription, which would explain the downregulation in poly(A)+ mRNA production apparent from 3h treatment onwards.

Isoginkgetin treatment also promotes a general transcription termination defect

Since we observed an increase in pol II level downstream of three protein-coding genes, *KPNB1*, *JUN*, and *HIST1H1C*, following Isoginkgetin treatment, we have investigated whether the transcription termination defect is a general trend. We re-analysed the previously published RNA-seq dataset where HeLa cells were treated with DMSO, as a control, or with Isoginkgetin for 12h (31) (Supplementary Figure 6A-C). We found that Isoginkgetin promotes a transcription termination defect on intron-containing, intronless, and histone genes, indicating that similarly to Madrasin, this small molecule promotes global transcriptional readthrough by pol II.

To understand better how Isoginkgetin could promote a transcription termination defect, we purified the chromatin fraction of HeLa cells treated with DMSO or Isoginkgetin for 1h, 2h, or 4h and performed western blots for total pol II, Ser2P, Ser5P, the elongation factor SPT6, and the mRNA CPA factors CPSF73 and CPSF100 (Figure 6A). We observed a loss of total pol II and SPT6 on chromatin over time while the pol II CTD marks were not as strongly affected, apart from an increase in Ser5P at the 1h timepoint. Interestingly, we found a specific loss of CPSF73 and CPSF100 from the chromatin after 1h treatment followed by a gradual increase in chromatin association of these two proteins. To confirm these observations, we performed pol II, Ser2P, Ser5P, and CPSF100 ChIP-qPCR across the *KPNB1* gene following Isoginkgetin treatment for 1h, 1h30, and 2h (Figure 6B). The changes in pol II profile looks similar to what we previously observed after Isoginketin treatment (Figure 4E) and the changes in Ser2P profile match the changes in total pol II, indicating that this mark is largely unaffected. However, the Ser5P level increases at the promoter region after 1h treatment but returns to control levels after 1h30, in agreement with the chromatin western blot. Importantly, we observed a loss of CPSF100 across *KPNB1* following Isoginkgetin treatment for the three timepoints, indicating that Isoginkgetin affects chromatin association of the mRNA CPA complex.

Isoginkgetin is a bioflavonoid, a group of small molecules that have been shown to regulate the SUMO post-translational modification (49). The SUMO modification is present on mRNA splicing and CPA factors and regulates their activities (49-51). To determine whether 1h, 2h, or 4h treatment of Isoginkgetin changes SUMO1 or SUMO2/3 levels, we performed western blots on HeLa whole cell

extract (Figure 6C). SUMO2/3 levels increase slightly only after 4h treatment, indicating that the effects we observe on transcription and mRNA CPA is unlikely to be mediated by changes in SUMOylation.

SF3B1 inhibitors and knockdown of splicing factors do not induce a general transcription termination defect

The majority of the available splicing inhibitors target the U2 snRNP protein, SF3B1 (52). To determine whether SF3B1 inhibition also promotes a global transcription termination defect, we re-analysed several HeLa RNA-seq datasets where either a 3h treatment with SSA, Sudemycin (Sud) C1, and Sudemycin K, or an 8h treatment with SSA and Sudemycin D1 (Rh18 cells) was performed (31,32 ,43) (Figure 7A-C and Supplementary Figure 7A-D). While treatment with SSA for 3h and 8h and with Sudemycin D1 for 8h affects pre-mRNA splicing, 3h treatment with Sudemycins C1 and K do not inhibit pre-mRNA splicing. No general defect on transcription termination is apparent for any of the SF3B1 inhibitors or treatment times. Inhibition of SF3B1 promotes a down-regulation of histone mRNA levels with a bigger decrease at 8h than at 3h (Figure 7A-C), likely due to cell cycle defects (53). It has been reported (54) that inhibition of SF3B1 with PlaB causes a transcription termination defect for a class of genes known as “Downstream Of Genes” (DOGs) (55), which are a group of genes known to be sensitive to cellular stresses (55-58). We have confirmed these observations on some of the DOGs, such as *RPL9* and *VAPA*, where a transcription termination defect is observed after 3h and 8h treatment with the different SF3B1 inhibitors.

To investigate whether knockdown of splicing factors and of hnRNP proteins could promote transcriptional readthrough, we have re-analysed the HepG2 GRO-seq datasets from (41) (Supplementary Figure 8A). While knockdown of some splicing factors, such as SRSF4, U2AF1, or U2AF2 can affect transcription of intron-containing, intronless, or histone genes, we do not observe a transcription termination defect. Similarly, no transcriptional readthrough is detected after knockdown of hnRNP proteins, although nascent transcription level can be affected. Contrary to the effect of SF3B1 inhibitors, there was no transcriptional readthrough on “DOGs” following knockdown of splicing factors or hnRNP proteins (Supplementary Figure 8B).

DISCUSSION

We set out to test whether inhibition of pre-mRNA splicing affects transcription elongation and transcription termination using two small molecules, Madrasin and Isoginkgetin, that were found to inhibit pre-mRNA splicing following a 24h treatment (29,30). To limit confounding effects associated with long treatment times, we have carried out short time courses of treatment. However, we found that only Madrasin inhibits pre-mRNA splicing following a 30 minutes treatment. Interestingly, Madrasin also promotes general transcriptional downregulation and transcriptional readthrough, which is observed on intron-containing, intronless, and histone genes. Potentially explaining the Madrasin-induced transcriptional dysregulation, we found that Madrasin affects the chromatin association of elongation and mRNA CPA factors. While a short treatment of Isoginkgetin does not affect pre-mRNA splicing, we found that this small molecule rapidly promotes transcriptional downregulation and transcriptional readthrough on intron-containing, intronless, and histone genes. Like Madrasin, Isoginkgetin affects the

chromatin association of mRNA CPA factors, potentially explaining the transcriptional readthrough. We note that while both Madrasin and Isoginkgetin promote global transcriptional readthrough, this is not the case for the splicing inhibitors targeting the U2 snRNP factor SF3B1, which cause readthrough on only a limited number of genes, called “DOGs”. In addition, knockdown of splicing factors and hnRNP proteins does not markedly affect transcription termination, indicating that splicing inhibition is not necessarily associated with transcriptional readthrough.

Long-reads sequencing approaches have shown that non-spliced intron-containing RNAs are associated with transcriptional readthrough (18,28). In addition, recognition of the final 3’ splice site and mRNA CPA are closely coupled (59-61). It is therefore surprising that inhibition of pre-mRNA splicing with small molecules targeting SF3B1 or knockdown of different splicing factors does not result in a general transcription termination defect. Therefore, out of Madrasin, Isoginkgetin, SF3B1 inhibitors, and splicing factor knockdown, only Madrasin causes both pre-mRNA splicing inhibition and transcriptional readthrough. However, the cellular target(s) of Madrasin is unknown. Our results therefore raise the question of whether this small molecule is causing transcriptional readthrough because of pre-mRNA splicing loss. We cannot rule out that Madrasin is inhibiting multiple proteins, for example a splicing factor and a mRNA CPA factor. From our results, the link to transcriptional readthrough seems to be whether the recruitment of mRNA CPA factors is affected rather than an effect on pre-mRNA splicing. In support of this idea, inhibition of SF3B1 with PlaB or SSA promotes premature termination of pol II (18,19), which relies on mRNA CPA (18,19), indicating that SF3B1 inhibition does not promote a loss of mRNA CPA factors from chromatin and therefore transcriptional readthrough, contrary to Isoginkgetin and Madrasin. Another important aspect that will require further investigation is the effect of these small molecule inhibitors on pol II elongation rate, which is known to affect pre-mRNA splicing and mRNA CPA (2-4,6,15,62).

An initially surprising observation was the sensitivity of intronless and histone genes to Madrasin treatment, which also argue for a lack of specificity of Madrasin towards pre-mRNA splicing. However, re-analysis of ChIP-seq of splicing factors and hnRNP proteins has shown that these RNA-binding proteins are also present on intronless and histone genes, indicating that the presence of introns is not required for their recruitment. In addition, knockdown of some of these splicing factors and hnRNP proteins affects the expression of intronless and histone genes, supporting splicing-independent recruitment and functions of these RNA binding proteins (63-66). The functions of these slicing factors on intronless and histone genes remains however to be investigated.

An interesting observation from this study is the diversity in transcriptional readthrough. For example, Madrasin and Isoginkgetin are both promoting global readthrough, including on intron-containing, intronless, and histone genes. In comparison, SF3B1 inhibitors but also different cellular stresses, such as KCl or influenza infection, affect only DOGs (54-57). As DOGs are generated following cellular stresses, transcriptional readthrough induced by SF3B1 inhibition on DOGs could indicate that SF3B1 inhibitors are also promoting cellular stress. This could be especially the case for the Sudemycins C1 and K, which do not effectively inhibit pre-mRNA splicing but cause transcriptional readthrough at DOGs, indicating a splicing-independent effect. A link between splicing defects and cellular stress has been shown with mutations in the splicing factor U2AF1 that is associated with the development of

stress granules (67). In the case of Madrasin and Isoginkgetin, the transcription termination defect is global, which indicates another or additional mechanisms to just cellular stress. From our data on CPSF73 and CPSF100, we propose that the chromatin association of mRNA CPA factors are affected by Madrasin and Isoginkgetin. However, it remains to be determined whether these two small molecule inhibitors directly or indirectly target the mRNA CPA machinery.

In line with a previous report (47), our results clearly indicate that Isoginkgetin is not a potent pre-mRNA splicing inhibitor. Splicing defects take more than 6h of Isoginkgetin treatment to appear, strongly suggesting that this is a secondary effect. However, Isoginkgetin affects pol II transcription in ~1h and pol I transcription in ~2h. The most rapid and strongest effect we observe is the loss of the mRNA CPA factors CPSF73 and CPSF100 from chromatin, which likely explains the global transcriptional readthrough promoted by Isoginkgetin treatment. Additionally, Isoginkgetin can also affect post-transcriptional regulation, as exemplified by the *JUN* gene. While we observed a clear increase in *JUN* poly(A)+ mRNA, transcription of *JUN* by pol II was decreased by Isoginkgetin, indicating a stabilisation of the *JUN* mRNA. The re-analysis of Isoginkgetin RNA-seq datasets also shows a clear enrichment of non-coding (nc)RNAs in upregulated genes, with ncRNAs representing approximately 2/3 of these genes. This upregulation may be explained by the transcriptional readthrough from upstream protein-coding genes into ncRNA genes, transforming them into expressed genes for the differential expression analysis software.

Understanding the coupling between transcription, pre-mRNA splicing, and mRNA CPA remains technically challenging. Based on our and other groups' analyses using Madrasin and Isoginkgetin, it is clearly important to use short timepoints to test potential splicing inhibitors rather than 24h treatment in order to limit the influence of secondary/indirect effects.

DATA AVAILABILITY

mNET-seq data have been deposited with the Gene Expression Omnibus under accession number GSE221279.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENT

We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics for sequencing.

Author contributions: G.A. and M.T. carried out the experimental analyses. M.T. performed the bioinformatic analyses. M.T. and S.M. supervised the research. M.T. wrote the paper with input from all the authors.

FUNDING

This work was supported by the Wellcome Trust Investigator Awards [WT106134AIA and WT210641/Z/18/Z to S.M.]. Funding for open access charge: Wellcome Trust.

CONFLICT OF INTEREST

None declared.

Table and Figure legends

Name	Forward primer	Reverse primer
PCR		
DNAJB1 E2/3	GGCCTGATGGGTCTTATCTATGG	AATGAGGTCCCCACGTTTCTCGGGTGT
BRD2 E4/5	CAAAATTATAAAACAGCCTATGGACATG	TTTTCCAGCGTTTGTGCCATTAGGA
ChIP-qPCR		
KPNB1 TSS	TTACTTCCTCCCTCCAAATGGG	ACAGCCTCCCTTCCTTCTTTC
KPNB1 TSS+3.2	GCCCAGAGAACAAGAAATCG	GGAATGGACAAGCTGTGTTG
KPNB1 TSS+20.2	TGCAAGAGCCAGTGGGAACACTT	CCTCTACTCAGCAATGATACTTC
KPNB1 pA-4.8	CTGAGGAAACTGAAGAACCAAG	GAAGGCAGTGCTTGCCAGAAT
KPNB1 pA-2.9	GAGGAGTGTGCACGGATGCTGAA	CCAAGATGGCCGATGTTATGG
KPNB1 pA-0.4	TAGTTACCGTCTGCTTGGGAAGATG	CCTCTGACAGCAAGTCCAACATT
KPNB1 pA+1.4	GACTCATCACACCAAGGTCAC	GATAGTGCTGGGAAGGAAATGG
KPNB1 pA+2.6	GTACATCTCAGCTTTGGCATATG	GCCCAGAACATAGCAGGCATTGC
KPNB1 pA+4.1	GTTTCACCGTGTTAGCCAGGATGG	CCACAGCCATGTTCATTTCTGC
KPNB1 TD (pA +11.3)	AGGAGCATGGCTTTTCTCTG	TCATGCTGGAACTGGTTGAG
KPNB1 +12.8	TGGAAGTTCTGATGCTCCTG	TCAAGTTAATGCCCCTGCTC
JUN TSS	TTTAGGGGTTGACTGGTAGCAG	TGTCTGTCTGTCTGCCTGAC
JUN GB	TGACCGCGACTTTTCAAAGC	GCGCAGGGTTAATTAAGATGCC
JUN pA	ATCAAGTGGCATGTGCTGTG	TGCCACCAATTCCTGCTTTG
JUN TD	TCACCCCTTCATGTTTCTGC	GTGTCACTTACGCCAAGACTTG

HIST1H1C GB	ACTCTGGTGCAAACGAAAGG	TTAACCTTGGGCTTGGCTTC
HIST1H1C T1	TCCCTCTGTTTGACATCCATGG	TGAGCAGCTTATCTCCAGAAGC
HIST1H1C T2	TCACGGGTTCAAGCGATTCTC	AAAATTAGCCGGGCATGGTG
Neg	TGGTACAACCACAGCTCAGTG	AAGCTGGACATGGTTGTGTG
qRT-PCR		
ATF3	AACCACAGTCAGTGGAGAGATG	TTCTCACAGCTGCAAACACC
LDLR	TTTGACGGGACTTCAGGTTC	TCCCTTGTGACATCTTCACG
NR4A3	AAGCCACCAGCTGTTAATGG	GCAATGCTGTTAGAGGAGCAG
PHLDB2	GAAACGACTTCAGGCAAGTCTC	GCCATGTTTTAGGAAAGAGCAC
SLCO4A1	TCCTCTTCTTTGCCATAGCC	ACAAGTTTCCAGGCCATCTG
ATP5F1B	ACCCATTGAAGAAGCTGTGG	CAATCAAGGCTCTTGTGCAG
TARS1	CATGGAAAAGGAGGAACAGC	CTTTGCCAAACTCCTCCAAG
H1-2 (HIST1H1C)	ACTCTGGTGCAAACGAAAGG	TTAACCTTGGGCTTGGCTTC
TUBB	ATATGTTCCTCGTGCCATCC	TTTGGCCCAGTTGTTACCTG
GAPDH	CAACGACCACTTTGTCAAGC	TTCCTCTTGTGCTCTTGCTG
KPNB1 (poly(A)+)	GAGGAGTGTGCACGGATGCTGAA	CCAAGATGGCCGATGTTATGG
KPNB1 (Random primers)	CAGTCTGTGATGGCATTTAAG	TCTGACTTTCAAACCATTTCACCTCC
JUN	TGACCGCGACTTTTCAAAGC	GCGCAGGGTTAATTAAGATGCC
7SK	CTGATCTGGCTGGCTAGGCGGG	GAAGACCGGTCCTCCTCTATCGG
Pre-rRNA	CCTGCTGTTCTCTCGCGCGTCCGAG	AACGCCTGACACGCACGGCACGGAG

Table 1: Sequences of primers used for PCR, ChIP-qPCR, and qRT-PCR.

REFERENCES

1. Tellier, M., Maudlin, I. and Murphy, S. (2020) Transcription and splicing: A two-way street. *Wiley Interdiscip Rev RNA*, **11**, e1593.
2. Howe, K.J., Kane, C.M. and Ares, M., Jr. (2003) Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae. *RNA*, **9**, 993-1006.
3. de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D. and Kornblihtt, A.R. (2003) A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell*, **12**, 525-532.
4. Dujardin, G., Lafaille, C., de la Mata, M., Marasco, L.E., Munoz, M.J., Le Jossic-Corcos, C., Corcos, L. and Kornblihtt, A.R. (2014) How slow RNA polymerase II elongation favors alternative exon skipping. *Mol Cell*, **54**, 683-690.

5. Koga, M., Hayashi, M. and Kaida, D. (2015) Splicing inhibition decreases phosphorylation level of Ser2 in Pol II CTD. *Nucleic Acids Res*, **43**, 8258-8267.
6. Godoy Herz, M.A., Kubaczka, M.G., Brzyzek, G., Servi, L., Krzyszton, M., Simpson, C., Brown, J., Swiezewski, S., Petrillo, E. and Kornblihtt, A.R. (2019) Light Regulates Plant Alternative Splicing through the Control of Transcriptional Elongation. *Mol Cell*, **73**, 1066-1074 e1063.
7. Caizzi, L., Monteiro-Martins, S., Schwalb, B., Lysakovskaia, K., Schmitzova, J., Sawicka, A., Chen, Y., Lidschreiber, M. and Cramer, P. (2021) Efficient RNA polymerase II pause release requires U2 snRNP function. *Mol Cell*, **81**, 1920-1934 e1929.
8. Zaborowska, J., Egloff, S. and Murphy, S. (2016) The pol II CTD: new twists in the tail. *Nat Struct Mol Biol*, **23**, 771-777.
9. Komarnitsky, P., Cho, E.J. and Buratowski, S. (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev*, **14**, 2452-2460.
10. Harlen, K.M., Trotta, K.L., Smith, E.E., Mosaheb, M.M., Fuchs, S.M. and Churchman, L.S. (2016) Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. *Cell Rep*, **15**, 2147-2158.
11. Nojima, T., Rebelo, K., Gomes, T., Grosso, A.R., Proudfoot, N.J. and Carmo-Fonseca, M. (2018) RNA Polymerase II Phosphorylated on CTD Serine 5 Interacts with the Spliceosome during Co-transcriptional Splicing. *Mol Cell*, **72**, 369-379 e364.
12. Zhang, S., Aibara, S., Vos, S.M., Agafonov, D.E., Luhrmann, R. and Cramer, P. (2021) Structure of a transcribing RNA polymerase II-U1 snRNP complex. *Science*, **371**, 305-309.
13. Eperon, L.P., Graham, I.R., Griffiths, A.D. and Eperon, I.C. (1988) Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell*, **54**, 393-401.
14. Braberg, H., Jin, H., Moehle, E.A., Chan, Y.A., Wang, S., Shales, M., Benschop, J.J., Morris, J.H., Qiu, C., Hu, F. *et al.* (2013) From structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II. *Cell*, **154**, 775-788.
15. Fong, N., Kim, H., Zhou, Y., Ji, X., Qiu, J., Saldi, T., Diener, K., Jones, K., Fu, X.D. and Bentley, D.L. (2014) Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev*, **28**, 2663-2676.
16. Saldi, T., Cortazar, M.A., Sheridan, R.M. and Bentley, D.L. (2016) Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing. *J Mol Biol*, **428**, 2623-2635.
17. Aslanzadeh, V., Huang, Y., Sanguinetti, G. and Beggs, J.D. (2018) Transcription rate strongly affects splicing fidelity and cotranscriptionality in budding yeast. *Genome Res*, **28**, 203-213.
18. Sousa-Luis, R., Dujardin, G., Zukher, I., Kimura, H., Weldon, C., Carmo-Fonseca, M., Proudfoot, N.J. and Nojima, T. (2021) POINT technology illuminates the processing of polymerase-associated intact nascent transcripts. *Mol Cell*, **81**, 1935-1950 e1936.
19. Yoshimoto, R., Chhipi-Shrestha, J.K., Schneider-Poetsch, T., Furuno, M., Burroughs, A.M., Noma, S., Suzuki, H., Hayashizaki, Y., Mayeda, A., Nakagawa, S. *et al.* (2021) Spliceostatin A interaction with SF3B limits U1 snRNP availability and causes premature cleavage and polyadenylation. *Cell Chem Biol*, **28**, 1356-1365 e1354.
20. Takeuchi, A., Iida, K., Tsubota, T., Hosokawa, M., Denawa, M., Brown, J.B., Ninomiya, K., Ito, M., Kimura, H., Abe, T. *et al.* (2018) Loss of Sfpq Causes Long-Gene Transcriptopathy in the Brain. *Cell Rep*, **23**, 1326-1341.

21. Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S. and Fu, X.D. (2008) The splicing factor SC35 has an active role in transcriptional elongation. *Nat Struct Mol Biol*, **15**, 819-826.
22. Ji, X., Zhou, Y., Pandit, S., Huang, J., Li, H., Lin, C.Y., Xiao, R., Burge, C.B. and Fu, X.D. (2013) SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell*, **153**, 855-868.
23. Gunderson, S.I., Vagner, S., Polycarpou-Schwarz, M. and Mattaj, I.W. (1997) Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation. *Genes Dev*, **11**, 761-773.
24. Vagner, S., Vagner, C. and Mattaj, I.W. (2000) The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. *Genes Dev*, **14**, 403-413.
25. Kyburz, A., Friedlein, A., Langen, H. and Keller, W. (2006) Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3' end processing and splicing. *Mol Cell*, **23**, 195-205.
26. Tellier, M., Zaborowska, J., Neve, J., Nojima, T., Hester, S., Fournier, M., Furger, A. and Murphy, S. (2022) CDK9 and PP2A regulate RNA polymerase II transcription termination and coupled RNA maturation. *EMBO Rep*, **23**, e54520.
27. Zhao, N., Sebastiano, V., Moshkina, N., Mena, N., Hultquist, J., Jimenez-Morales, D., Ma, Y., Rialdi, A., Albrecht, R., Fenouil, R. *et al.* (2018) Influenza virus infection causes global RNAPII termination defects. *Nat Struct Mol Biol*, **25**, 885-893.
28. Reimer, K.A., Mimoso, C.A., Adelman, K. and Neugebauer, K.M. (2021) Co-transcriptional splicing regulates 3' end cleavage during mammalian erythropoiesis. *Mol Cell*, **81**, 998-1012 e1017.
29. Pawellek, A., McElroy, S., Samatov, T., Mitchell, L., Woodland, A., Ryder, U., Gray, D., Luhrmann, R. and Lamond, A.I. (2014) Identification of small molecule inhibitors of pre-mRNA splicing. *J Biol Chem*, **289**, 34683-34698.
30. O'Brien, K., Matlin, A.J., Lowell, A.M. and Moore, M.J. (2008) The biflavonoid isoginkgetin is a general inhibitor of Pre-mRNA splicing. *J Biol Chem*, **283**, 33147-33154.
31. Tseng, C.K., Wang, H.F., Burns, A.M., Schroeder, M.R., Gaspari, M. and Baumann, P. (2015) Human Telomerase RNA Processing and Quality Control. *Cell Rep*, **13**, 2232-2243.
32. Vigevani, L., Gohr, A., Webb, T., Irimia, M. and Valcarcel, J. (2017) Molecular basis of differential 3' splice site sensitivity to anti-tumor drugs targeting U2 snRNP. *Nat Commun*, **8**, 2100.
33. Nojima, T., Gomes, T., Grosso, A.R.F., Kimura, H., Dye, M.J., Dhir, S., Carmo-Fonseca, M. and Proudfoot, N.J. (2015) Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell*, **161**, 526-540.
34. Tellier, M., Zaborowska, J., Caizzi, L., Mohammad, E., Velychko, T., Schwalb, B., Ferrer-Vicens, I., Blears, D., Nojima, T., Cramer, P. *et al.* (2020) CDK12 globally stimulates RNA polymerase II transcription elongation and carboxyl-terminal domain phosphorylation. *Nucleic Acids Res*, **48**, 7712-7727.
35. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. and Kingsford, C. (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*, **14**, 417-419.
36. Nojima, T., Tellier, M., Foxwell, J., Ribeiro de Almeida, C., Tan-Wong, S.M., Dhir, S., Dujardin, G., Dhir, A., Murphy, S. and Proudfoot, N.J. (2018) Deregulated

Expression of Mammalian lncRNA through Loss of SPT6 Induces R-Loop Formation, Replication Stress, and Cellular Senescence. *Mol Cell*, **72**, 970-984 e977.
37. Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011*, **17**, 3.
38. 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.
39. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Genome Project Data Processing, S. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.
40. Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F. and Manke, T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res*, **44**, W160-165.
41. Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R., Blue, S.M., Chen, J.Y., Cody, N.A.L., Dominguez, D. *et al.* (2020) A large-scale binding and functional map of human RNA-binding proteins. *Nature*, **583**, 711-719.
42. Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841-842.
43. Wu, G., Fan, L., Edmonson, M.N., Shaw, T., Boggs, K., Easton, J., Rusch, M.C., Webb, T.R., Zhang, J. and Potter, P.M. (2018) Inhibition of SF3B1 by molecules targeting the spliceosome results in massive aberrant exon skipping. *RNA*, **24**, 1056-1066.
44. Anders, S., Pyl, P.T. and Huber, W. (2015) HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166-169.
45. 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.
46. Tellier, M. and Murphy, S. (2020) Incomplete removal of ribosomal RNA can affect chromatin RNA-seq data analysis. *Transcription*, **11**, 230-235.
47. Boswell, S.A., Snavely, A., Landry, H.M., Churchman, L.S., Gray, J.M. and Springer, M. (2017) Total RNA-seq to identify pharmacological effects on specific stages of mRNA synthesis. *Nat Chem Biol*, **13**, 501-507.
48. Vanzyl, E.J., Sayed, H., Blackmore, A.B., Rick, K.R.C., Fernando, P. and McKay, B.C. (2020) The spliceosome inhibitors isoginkgetin and pladienolide B induce ATF3-dependent cell death. *PLoS One*, **15**, e0224953.
49. Pawellek, A., Ryder, U., Tammsalu, T., King, L.J., Kreinin, H., Ly, T., Hay, R.T., Hartley, R.C. and Lamond, A.I. (2017) Characterisation of the biflavonoid hinokiflavone as a pre-mRNA splicing modulator that inhibits SENP. *Elife*, **6**.
50. Vethantham, V., Rao, N. and Manley, J.L. (2007) Sumoylation modulates the assembly and activity of the pre-mRNA 3' processing complex. *Mol Cell Biol*, **27**, 8848-8858.
51. Pozzi, B., Bragado, L., Will, C.L., Mammi, P., Risso, G., Urlaub, H., Luhrmann, R. and Srebrow, A. (2017) SUMO conjugation to spliceosomal proteins is required for efficient pre-mRNA splicing. *Nucleic Acids Res*, **45**, 6729-6745.
52. Stanley, R.F. and Abdel-Wahab, O. (2022) Dysregulation and therapeutic targeting of RNA splicing in cancer. *Nat Cancer*, **3**, 536-546.
53. Petasny, M., Bentata, M., Pawellek, A., Baker, M., Kay, G. and Salton, M. (2021) Splicing to Keep Cycling: The Importance of Pre-mRNA Splicing during the Cell Cycle. *Trends Genet*, **37**, 266-278.

54. Castillo-Guzman, D., Hartono, S.R., Sanz, L.A. and Chédin, F. (2020) SF3B1-targeted Splicing Inhibition Triggers Global Alterations in Transcriptional Dynamics and R-Loop Metabolism. *bioRxiv*, 2020.2006.2008.130583.
55. Vilborg, A., Passarelli, M.C., Yario, T.A., Tycowski, K.T. and Steitz, J.A. (2015) Widespread Inducible Transcription Downstream of Human Genes. *Mol Cell*, **59**, 449-461.
56. Vilborg, A., Sabath, N., Wiesel, Y., Nathans, J., Levy-Adam, F., Yario, T.A., Steitz, J.A. and Shalgi, R. (2017) Comparative analysis reveals genomic features of stress-induced transcriptional readthrough. *Proc Natl Acad Sci U S A*, **114**, E8362-E8371.
57. Bauer, D.L.V., Tellier, M., Martinez-Alonso, M., Nojima, T., Proudfoot, N.J., Murphy, S. and Fodor, E. (2018) Influenza Virus Mounts a Two-Pronged Attack on Host RNA Polymerase II Transcription. *Cell Rep*, **23**, 2119-2129 e2113.
58. Melnick, M., Gonzales, P., Cabral, J., Allen, M.A., Dowell, R.D. and Link, C.D. (2019) Heat shock in C. elegans induces downstream of gene transcription and accumulation of double-stranded RNA. *PLoS One*, **14**, e0206715.
59. Niwa, M., Rose, S.D. and Berget, S.M. (1990) In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev*, **4**, 1552-1559.
60. Dye, M.J. and Proudfoot, N.J. (1999) Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. *Mol Cell*, **3**, 371-378.
61. Custodio, N., Carmo-Fonseca, M., Geraghty, F., Pereira, H.S., Grosveld, F. and Antoniou, M. (1999) Inefficient processing impairs release of RNA from the site of transcription. *EMBO J*, **18**, 2855-2866.
62. Cortazar, M.A., Sheridan, R.M., Erickson, B., Fong, N., Glover-Cutter, K., Brannan, K. and Bentley, D.L. (2019) Control of RNA Pol II Speed by PNUTS-PP1 and Spt5 Dephosphorylation Facilitates Termination by a "Sitting Duck Torpedo" Mechanism. *Mol Cell*, **76**, 896-908 e894.
63. Patel, S.B., Novikova, N. and Bellini, M. (2007) Splicing-independent recruitment of spliceosomal small nuclear RNPs to nascent RNA polymerase II transcripts. *J Cell Biol*, **178**, 937-949.
64. Spiluttini, B., Gu, B., Belagal, P., Smirnova, A.S., Nguyen, V.T., Hebert, C., Schmidt, U., Bertrand, E., Darzacq, X. and Bensaude, O. (2010) Splicing-independent recruitment of U1 snRNP to a transcription unit in living cells. *J Cell Sci*, **123**, 2085-2093.
65. Wu, H., Sun, S., Tu, K., Gao, Y., Xie, B., Krainer, A.R. and Zhu, J. (2010) A splicing-independent function of SF2/ASF in microRNA processing. *Mol Cell*, **38**, 67-77.
66. Mohan, N., Kumar, V., Kandala, D.T., Kartha, C.C. and Laishram, R.S. (2018) A Splicing-Independent Function of RBM10 Controls Specific 3' UTR Processing to Regulate Cardiac Hypertrophy. *Cell Rep*, **24**, 3539-3553.
67. Biancon, G., Joshi, P., Zimmer, J.T., Hunck, T., Gao, Y., Lessard, M.D., Courchaine, E., Barentine, A.E.S., Machyna, M., Botti, V. *et al.* (2022) Precision analysis of mutant U2AF1 activity reveals deployment of stress granules in myeloid malignancies. *Mol Cell*, **82**, 1107-1122 e1107.

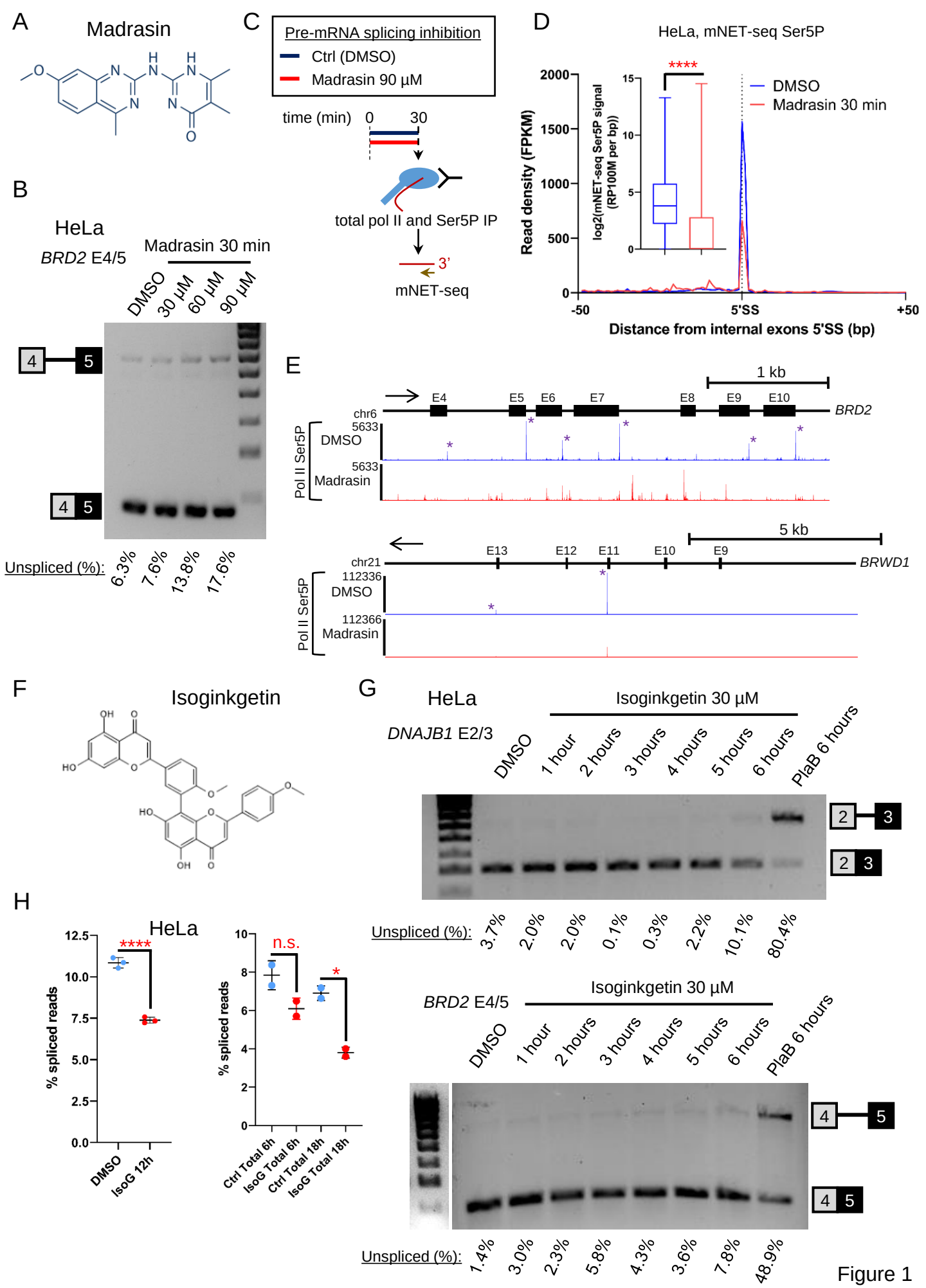
A
Madrasin
B
HeLa
BRD2 E4/5
Madrasin 30 min
DMSO
30 µM
60 µM
90 µM
4
5
Unspliced (%):
6.3%
7.6%
13.8%
17.6%
C
Pre-mRNA splicing inhibition
Ctrl (DMSO)
Madrasin 90 µM
time (min) 0 30
total pol II and Ser5P IP
3'
mNET-seq
D
HeLa, mNET-seq Ser5P
Read density (FPKM)
log2(mNET-seq Ser5P signal (RP100M per bp))
****
DMSO
Madrasin 30 min
-50
5'SS
+50
Distance from internal exons 5'SS (bp)
E
1 kb
E4
E5
E6
E7
E8
E9
E10
chr6
BRD2
Pol II Ser5P
5633
DMSO
5633
Madrasin
5 kb
E13
E12
E11
E10
E9
chr21
BRWD1
112336
DMSO
112366
Madrasin
F
Isoginkgetin
G
HeLa
DNAJB1 E2/3
Isoginkgetin 30 µM
DMSO
1 hour
2 hours
3 hours
4 hours
5 hours
6 hours
PlaB 6 hours
2
3
Unspliced (%):
3.7%
2.0%
2.0%
0.1%
0.3%
2.2%
10.1%
80.4%
BRD2 E4/5
Isoginkgetin 30 µM
DMSO
1 hour
2 hours
3 hours
4 hours
5 hours
6 hours
PlaB 6 hours
4
5
Unspliced (%):
1.4%
3.0%
2.3%
5.8%
4.3%
3.6%
7.8%
48.9%
H
HeLa
% spliced reads
****
DMSO
IsoG 12h
n.s.
*
Ctrl Total 6h
IsoG Total 6h
Ctrl Total 18h
IsoG Total 18h

Figure 1

Figure 1: Characterization of the pre-mRNA splicing inhibitors Isoginkgetin and Madrasin. **(A)** Chemical structure of Madrasin. **(B)** RT-PCR with primers amplifying the intron located between exons 4 and 5 of *BRD2*. HeLa cells were treated with DMSO or 30 μM, 60 μM, or 90 μM Madrasin for 30 minutes. The location of the spliced and unspliced RNA is shown on the left of the panel. The percentages of unspliced RNA compared to total (spliced + unspliced) are shown below. **(C)** Schematic of the mNET-seq technique. **(D)** Metagene profile of mNET-seq Ser5P treated with DMSO (blue) or 90 μM Madrasin (red) for 30 minutes around the 5' splice site (5'SS) of internal exons, i.e. neither first or last exons. Quantification of the mNET-seq Ser5P DMSO (blue) and Madrasin (red) signals are shown as boxplots, with min to max with first quartile, median, and third quartile. Statistical test: Wilcoxon signed-rank test. P-value: **** < 0.0001. **(E)** Screenshot of the genome browser mNET-seq Ser5P DMSO (blue) and Madrasin (red) tracks of the protein-coding genes *BRD2* and *BRWD1*. Purple stars indicate the location of co-transcriptional splicing events. The arrow indicates the sense of transcription. **(F)** Chemical structure of Isoginkgetin. **(G)** RT-PCR with primers amplifying the intron located between exons 2 and 3 of *DNAJB1* or the intron located between exons 4 and 5 of *BRD2*. HeLa cells were treated with DMSO, 30 μM Isoginkgetin for 1h to 6h, or 1 μM Pladienolide B for 6h. The locations of the spliced and unspliced RNA are shown on the right of the panel. The percentages of unspliced RNA compared to total (spliced + unspliced) are shown below. **(H)** Percentage of spliced reads in total RNA-seq treated with DMSO (blue) or Isoginkgetin (red) for the time indicated on the figure. Each point represents a biological replicate. Statistical test: Wilcoxon rank sum test. P-value: n.s. not significant, * < 0.05, **** < 0.0001.

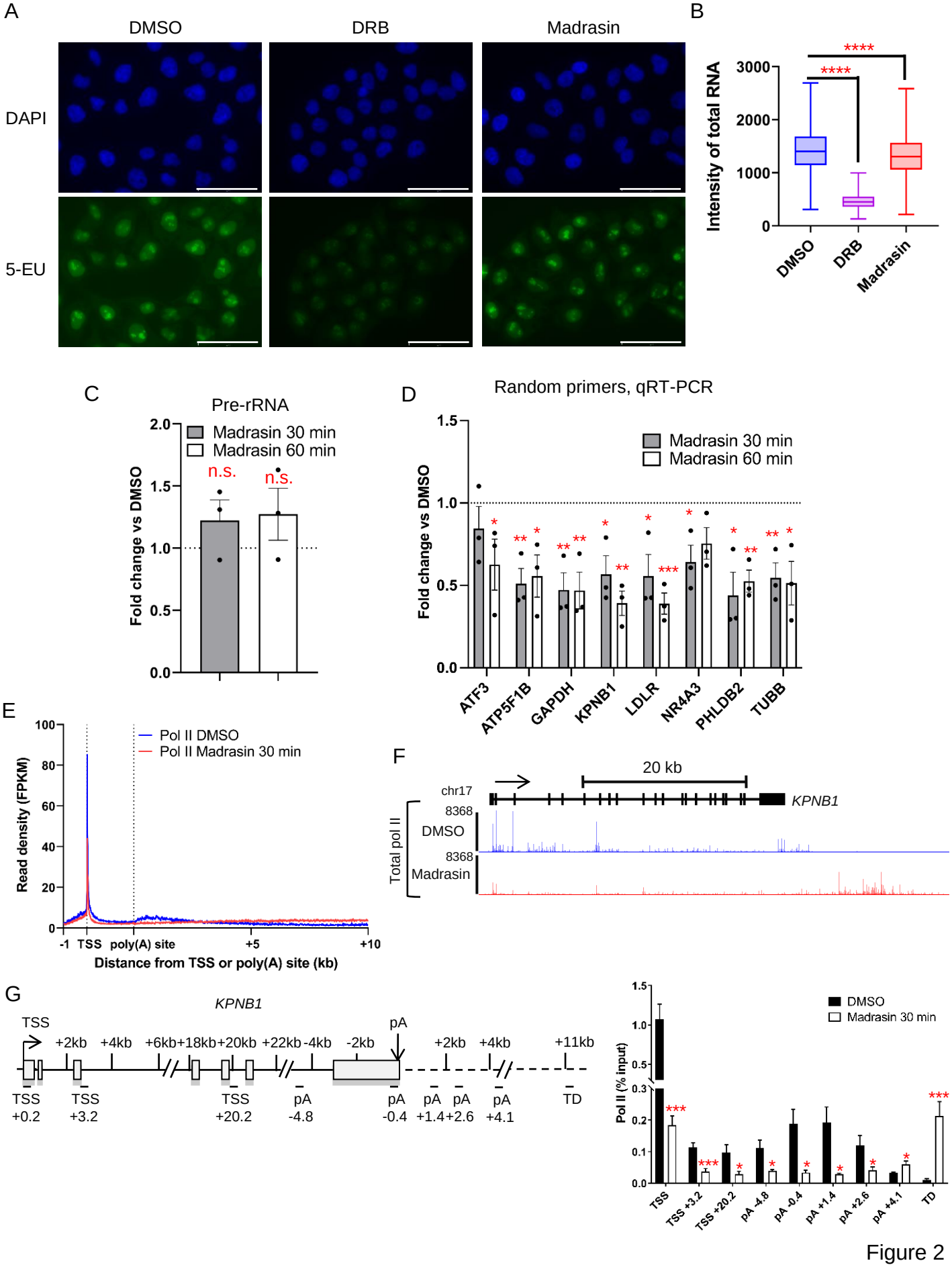
A
DMSO
DRB
Madrasin
DAPI
5-EU
B
Intensity of total RNA
3000
2000
1000
0
****
****
DMSO
DRB
Madrasin
C
Pre-rRNA
Madrasin 30 min
Madrasin 60 min
n.s.
n.s.
Fold change vs DMSO
2.0
1.5
1.0
0.5
0.0
D
Random primers, qRT-PCR
Madrasin 30 min
Madrasin 60 min
Fold change vs DMSO
1.5
1.0
0.5
0.0
ATF3
ATP5F1B
GAPDH
KPNB1
LDLR
NR4A3
PHLDB2
TUBB
E
Pol II DMSO
Pol II Madrasin 30 min
Read density (FPKM)
100
80
60
40
20
0
-1
TSS
poly(A) site
+5
+10
Distance from TSS or poly(A) site (kb)
F
20 kb
chr17
KPNB1
8368
DMSO
8368
Madrasin
Total pol II
G
KPNB1
TSS
+2kb
+4kb
+6kb
+18kb
+20kb
+22kb
-4kb
-2kb
pA
+2kb
+4kb
+11kb
TSS +0.2
TSS +3.2
TSS +20.2
pA -4.8
pA -0.4
pA +1.4
pA +2.6
pA +4.1
TD
DMSO
Madrasin 30 min
Pol II (% input)
1.5
1.0
0.5
0.3
0.2
0.1
0.0
TSS
TSS +3.2
TSS +20.2
pA -4.8
pA -0.4
pA +1.4
pA +2.6
pA +4.1
TD

Figure 2

Figure 2: Inhibition of pre-mRNA splicing with Madrasin affects transcription of protein-coding genes and promotes transcriptional readthrough

(A) Representative images of immunofluorescence analysis or 5'EU incorporation in HeLa cells treated with DMSO, 100 μM DRB, a CDK9 inhibitor, or 90 μM Madrasin for 1h. EU (green), DAPI (blue), scale bars: 50 μm. **(B)** Quantification of 5'EU intensity per nucleus for DMSO (blue), DRB (purple), and Madrasin (red). Boxplot settings are: min to max values with the box showing 25-75 percentile range. > 10,000 nuclei were quantified per condition. Statistical test: Kruskal-Wallis test. P-value: **** < 0.0001. **(C)** qRT-PCR with primers amplifying a region from the pol I transcribed pre-rRNA. HeLa cells were treated with DMSO or 90 μM Madrasin for 30 minutes (grey) or 60 minutes (white). cDNA was generated with random hexamers. Values are normalised to the 7SK snRNA and shown as relative to DMSO, mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: n.s. not significant. **(D)** qRT-PCR with primers amplifying different protein-coding genes. HeLa cells were treated with DMSO or 90 μM Madrasin for 30 minutes (grey) or 60 minutes (white). cDNA was generated with random hexamers. Values are normalised to the 7SK snRNA and shown as relative to DMSO, mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, ** < 0.01, *** < 0.001. **(E)** Metagene profile of total pol II mNET-seq performed in HeLa cells treated with DMSO (blue) or 90 μM Madrasin for 30 minutes (red) on scaled expressed protein-coding genes. **(F)** Screenshot of the genome browser total pol II mNET-seq DMSO (blue) and Madrasin (red) tracks of the protein-coding gene *KPNB1*. The arrow indicates the sense of transcription. **(G)** Total pol II ChIP-qPCR of the *KPNB1* protein-coding gene in HeLa cells treated with DMSO (black) or 90 μM Madrasin for 30 minutes (white). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, *** < 0.001.

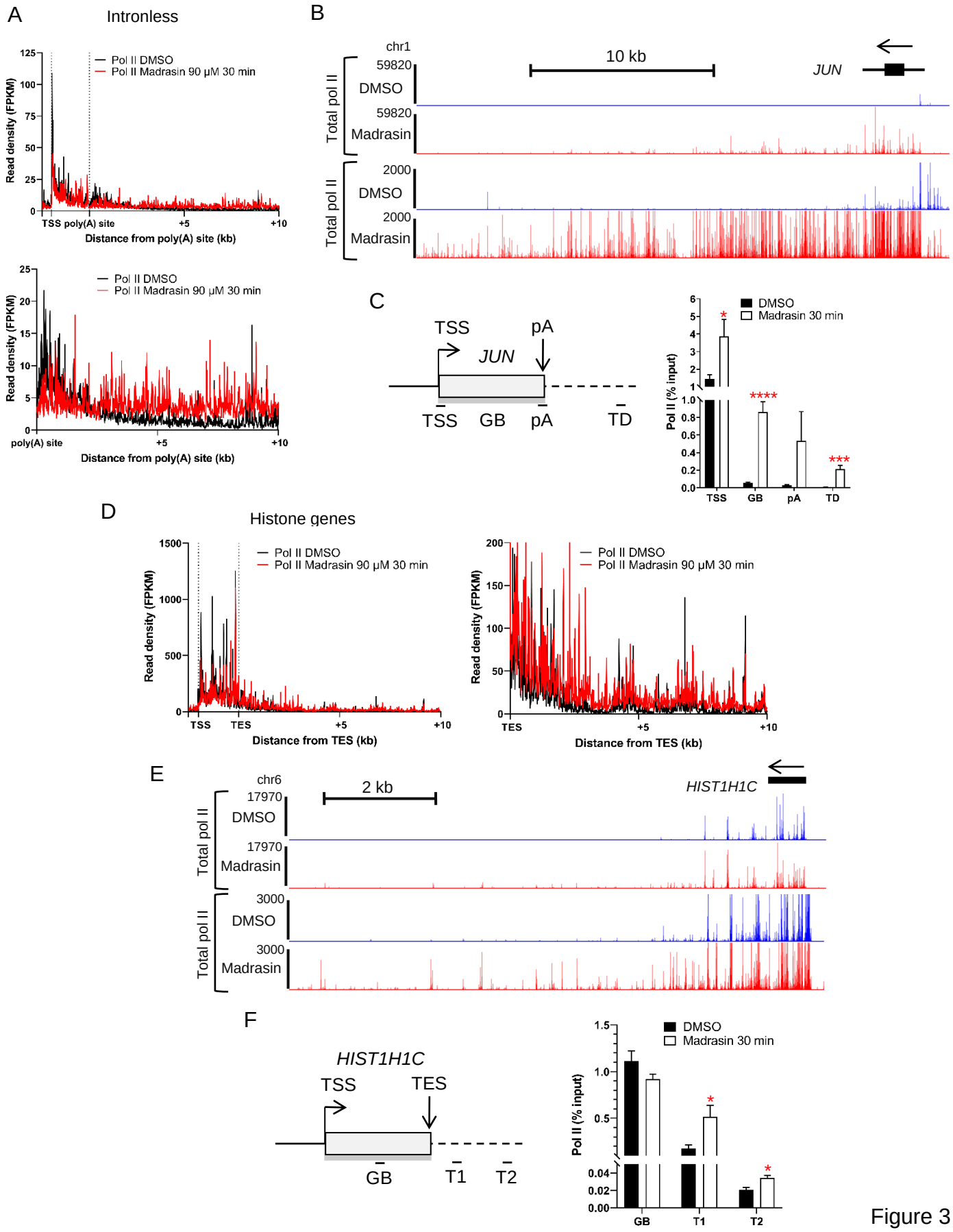
A
Intronless
Pol II DMSO
Pol II Madrasin 90 μM 30 min
Read density (FPKM)
TSS poly(A) site
+5
+10
Distance from poly(A) site (kb)
poly(A) site
B
chr1
10 kb
JUN
Total pol II
59820
DMSO
Madrasin
2000
C
TSS
pA
JUN
GB
TD
DMSO
Madrasin 30 min
Pol II (% input)
D
Histone genes
TES
Distance from TES (kb)
E
chr6
2 kb
HIST1H1C
17970
3000
F
T1
T2

Figure 3

Figure 3: Intronless and histone genes are also affected by Madrasin

(A) Metagene profiles of total pol II mNET-seq performed in HeLa cells treated with DMSO (blue) or 90 μM Madrasin for 30 minutes (red) on scaled expressed intronless genes. **(B)** Screenshot of the genome browser total pol II mNET-seq DMSO (blue) and Madrasin (red) tracks of the intronless gene *JUN*. The arrow indicates the sense of transcription. **(C)** Total pol II ChIP-qPCR of the *JUN* intronless gene in HeLa cells treated with DMSO (black) or 90 μM Madrasin for 30 minutes (white). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, *** < 0.001, **** < 0.0001. **(D)** Metagene profiles of total pol II mNET-seq performed in HeLa cells treated with DMSO (blue) or 90 μM Madrasin for 30 minutes (red) on scaled expressed histone genes. **(E)** Screenshot of the genome browser total pol II mNET-seq DMSO (blue) and Madrasin (red) tracks of the histone gene *HIST1H1C*, also known as *H1-2*. The arrow indicates the sense of transcription. **(F)** Total pol II ChIP-qPCR of the *HIST1H1C* histone gene in HeLa cells treated with DMSO (black) or 90 μM Madrasin for 30 minutes (white). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05.

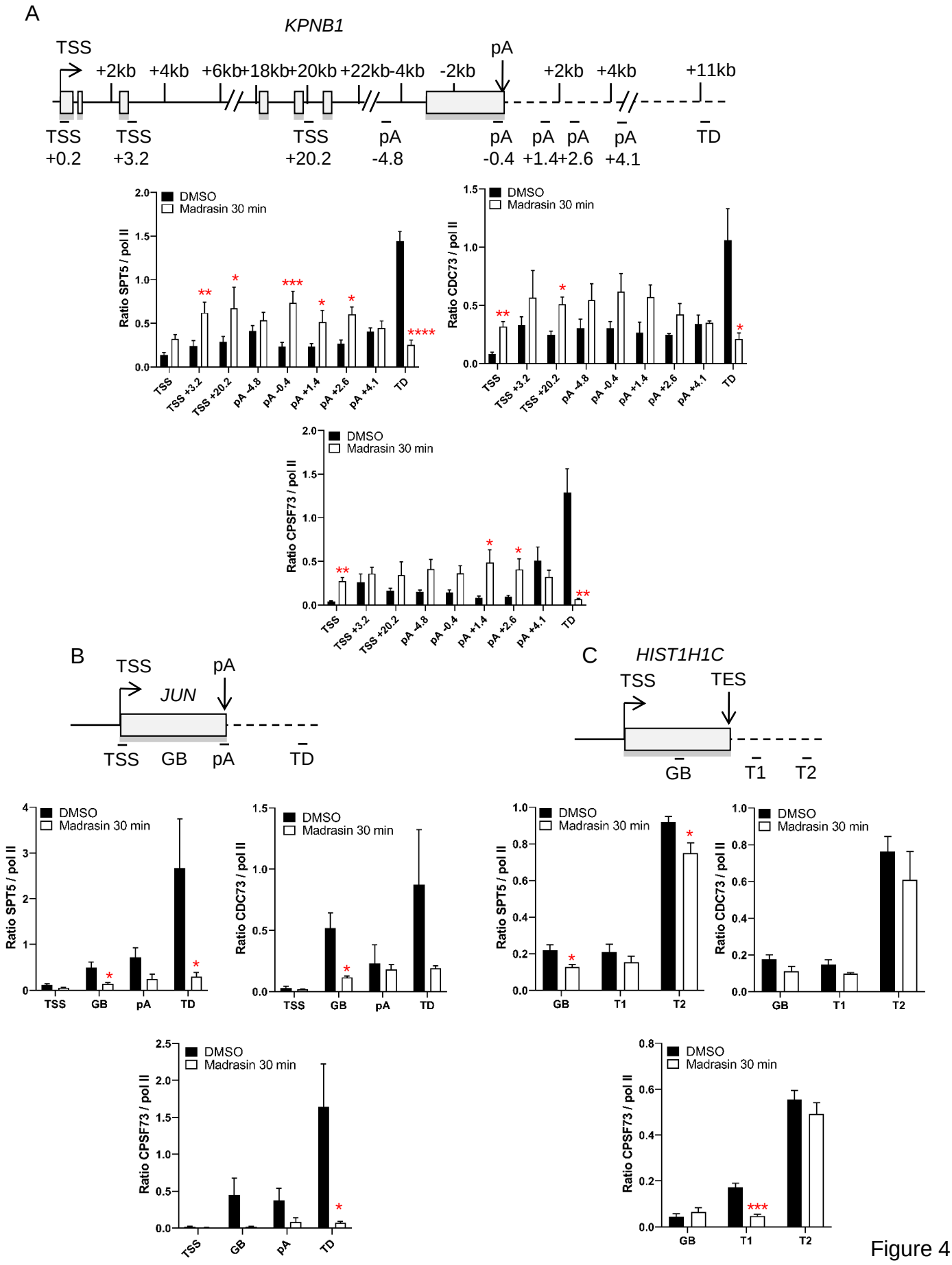
A
KPNB1
TSS
pA
+2kb +4kb +6kb +18kb +20kb +22kb -4kb -2kb +2kb +4kb +11kb
TSS +0.2
TSS +3.2
TSS +20.2
pA -4.8
pA -0.4
pA +1.4
pA +2.6
pA +4.1
TD
DMSO
Madrasin 30 min
Ratio SPT5 / pol II
Ratio CDC73 / pol II
Ratio CPSF73 / pol II
B
JUN
GB
C
HIST1H1C
TES
T1
T2

Figure 4

Figure 4: Inhibition of pre-mRNA splicing with Madrasin affects the chromatin association of elongation and termination factors

(A) Ratios of SPT5 / total pol II, CDC73 / total pol II, or CPSF73 / total pol II from ChIP-qPCR on the *KPNB1* intron-containing gene in HeLa cells treated with DMSO (black) or 90 μM Madrasin for 30 minutes (white). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001. **(B)** Ratios of SPT5 / total pol II, CDC73 / total pol II, or CPSF73 / total pol II from ChIP-qPCR on the *JUN* intronless gene in HeLa cells treated with DMSO (black) or 90 μM Madrasin for 30 minutes (white). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05. **(C)** Ratios of SPT5 / total pol II, CDC73 / total pol II, or CPSF73 / total pol II from ChIP-qPCR on the *HIST1H1C* histone gene in HeLa cells treated with DMSO (black) or 90 μM Madrasin for 30 minutes (white). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, *** < 0.001.

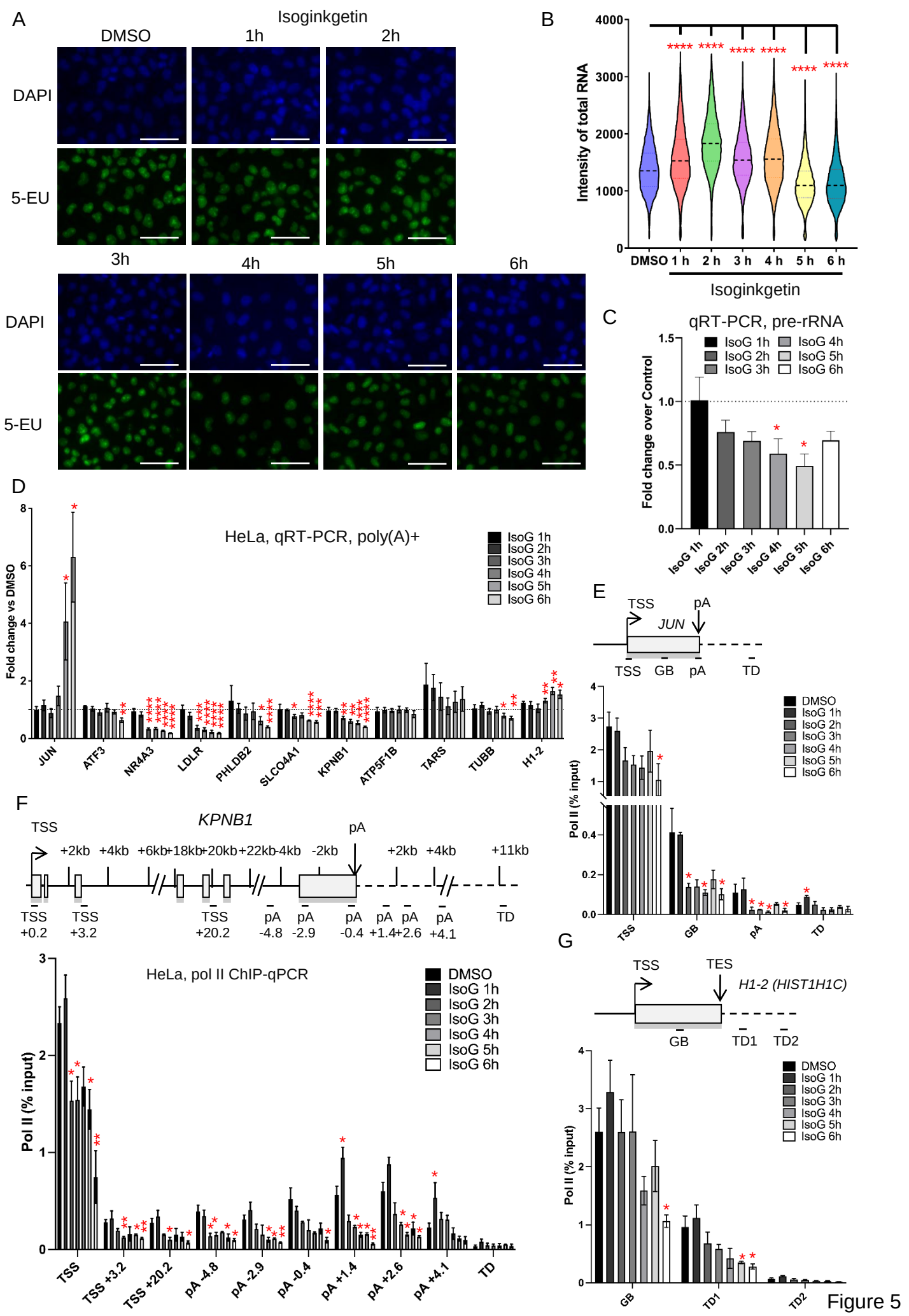
A
Isoginkgetin
DMSO
1h
2h
DAPI
5-EU
3h
4h
5h
6h
B
Intensity of total RNA
DMSO 1 h 2 h 3 h 4 h 5 h 6 h
Isoginkgetin
C
qRT-PCR, pre-rRNA
IsoG 1h IsoG 2h IsoG 3h IsoG 4h IsoG 5h IsoG 6h
Fold change over Control
D
HeLa, qRT-PCR, poly(A)+
Fold change vs DMSO
JUN ATF3 NR4A3 LDLR PHLDB2 SLCO4A1 KPNB1 ATP5F1B TARS TUBB H1-2
E
TSS pA
JUN
TSS GB pA TD
DMSO
Pol II (% input)
F
KPNB1
TSS
pA
+2kb +4kb +6kb +18kb +20kb +22kb -4kb -2kb +2kb +4kb +11kb
TSS +0.2 TSS +3.2 TSS +20.2 pA -4.8 pA -2.9 pA -0.4 pA +1.4 pA +2.6 pA +4.1 TD
HeLa, pol II ChIP-qPCR
TSS TSS +3.2 TSS +20.2 pA -4.8 pA -2.9 pA -0.4 pA +1.4 pA +2.6 pA +4.1 TD
G
TSS TES
H1-2 (HIST1H1C)
GB TD1 TD2

Figure 5

Figure 5: Isoginkgetin treatment affects pol I and pol II transcription

(A) Representative images of immunofluorescence analysis or 5'EU incorporation in HeLa cells treated with DMSO or 30 μM Isoginkgetin for 1h to 6h. EU (green), DAPI (blue), scale bars: 50 μm. **(B)** Quantification of 5'EU intensity per nucleus for DMSO (blue) and Isoginkgetin (other colours). Boxplot settings are: min to max values with the box showing 25-75 percentile range. 8,466 nuclei were quantified per condition. Statistical test: Kruskal-Wallis test. P-value: **** < 0.0001. **(C)** qRT-PCR with primers amplifying a region from the pol I transcribed pre-rRNA. HeLa cells were treated with DMSO or 30 μM Madrasin for 1h to 6h. cDNA was generated with random hexamers. Values are normalised to the *7SK* snRNA and shown as relative to DMSO, mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: n* < 0.05. **(D)** qRT-PCR with primers amplifying different protein-coding genes. HeLa cells were treated with DMSO or 30 μM Isoginkgetin for 1h to 6h. cDNA was generated with oligo(dT). Values are normalised to the *GAPDH* protein-coding gene and shown as relative to DMSO, mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001. **(E)** Total pol II ChIP-qPCR of the *KPNB1* intron-containing gene in HeLa cells treated with DMSO (black) or 30 μM Isoginkgetin for 1h to 6h. Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, ** < 0.01. **(F)** Total pol II ChIP-qPCR of the *JUN* intronless gene in HeLa cells treated with DMSO (black) or 30 μM Isoginkgetin for 1h to 6h. Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05. **(G)** Total pol II ChIP-qPCR of the *HIST1H1C* histone gene in HeLa cells treated with DMSO (black) or 30 μM Isoginkgetin for 1h to 6h. Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05.

A

HeLa, chromatin

Isoginkgetin

DMSO 1h 2h 4h

Total pol II

1 1.08 0.86 0.55

Ser2P

1 0.80 0.62 0.63

Ser5P

1 1.34 0.80 0.73

SPT6

1 0.82 0.64 0.26

CPSF73

1 0.58 1.11 1.63

CPSF100

1 0.69 1.20 2.36

Histone H3

B

KPNB1

TSS +2kb +4kb +6kb +18kb +20kb +22kb -4kb -2kb pA +2kb +4kb +11kb

TSS +0.2 TSS +3.2 TSS +20.2 pA -4.8 pA -2.9 pA -0.4 pA +1.4 pA +2.6 pA +4.1 TD

HeLa, pol II ChIP-qPCR

Pol II (% input)

DMSO
IsoG 1h
IsoG 1h30
IsoG 2h

TSS TSS +3.2 TSS +20.2 pA -4.8 pA -2.9 pA -0.4 pA +1.4 pA +2.6 pA +4.1 TerDef

HeLa, Ser2P ChIP-qPCR

Ser2P (% Input)

DMSO
IsoG 1h
IsoG 1h30
IsoG 2h

TSS TSS +3.2 TSS +20.2 pA -4.8 pA -2.9 pA -0.4 pA +1.4 pA +2.6 pA +4.1 TerDef

HeLa, Ser5P ChIP-qPCR

Ser5P (% Input)

DMSO
IsoG 1h
IsoG 1h30
IsoG 2h

TSS TSS +3.2 TSS +20.2 pA -4.8 pA -2.9 pA -0.4 pA +1.4 pA +2.6 pA +4.1 TerDef

HeLa, CPSF100 ChIP-qPCR

CPSF100 (% Input)

DMSO
IsoG 1h
IsoG 1h30
IsoG 2h

TSS TSS +3.2 TSS +20.2 pA -4.8 pA -2.9 pA -0.4 pA +1.4 pA +2.6 pA +4.1 TerDef

C

HeLa, whole-cell extract

DMSO Isoginkgetin 1h 2h 4h

SUMO1

DMSO Isoginkgetin 1h 2h 4h

Ponceau S

DMSO Isoginkgetin 1h 2h 4h

SUMO2/3

DMSO Isoginkgetin 1h 2h 4h

Ponceau S

Figure 6

Figure 6: Isoginkgetin treatment also promotes a general transcription termination defect

(A) Western blots on the chromatin fraction of HeLa cells treated with DMSO or with Isoginkgetin for 1h, 2h, or 4h. The antibodies are indicated on the right, with histone H3 used as a loading control. Quantifications of the western blots are shown below each panel. Each quantification has been normalised to histone H3 signal and then to the DMSO condition. **(B)** ChIP-qPCR of total pol II, Ser2P, Ser5P, and CPSF100 across the *KPNB1* intron-containing gene in HeLa cells treated with DMSO (black) or 30 μM Isoginkgetin for 1h, 1h30, or 2h. Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001. **(C)** Western blot on whole cell extract of HeLa cells treated with DMSO or with Isoginkgetin for 1h, 2h, or 4h. The antibodies are indicated at the bottom, with Ponceau S staining used as loading control.

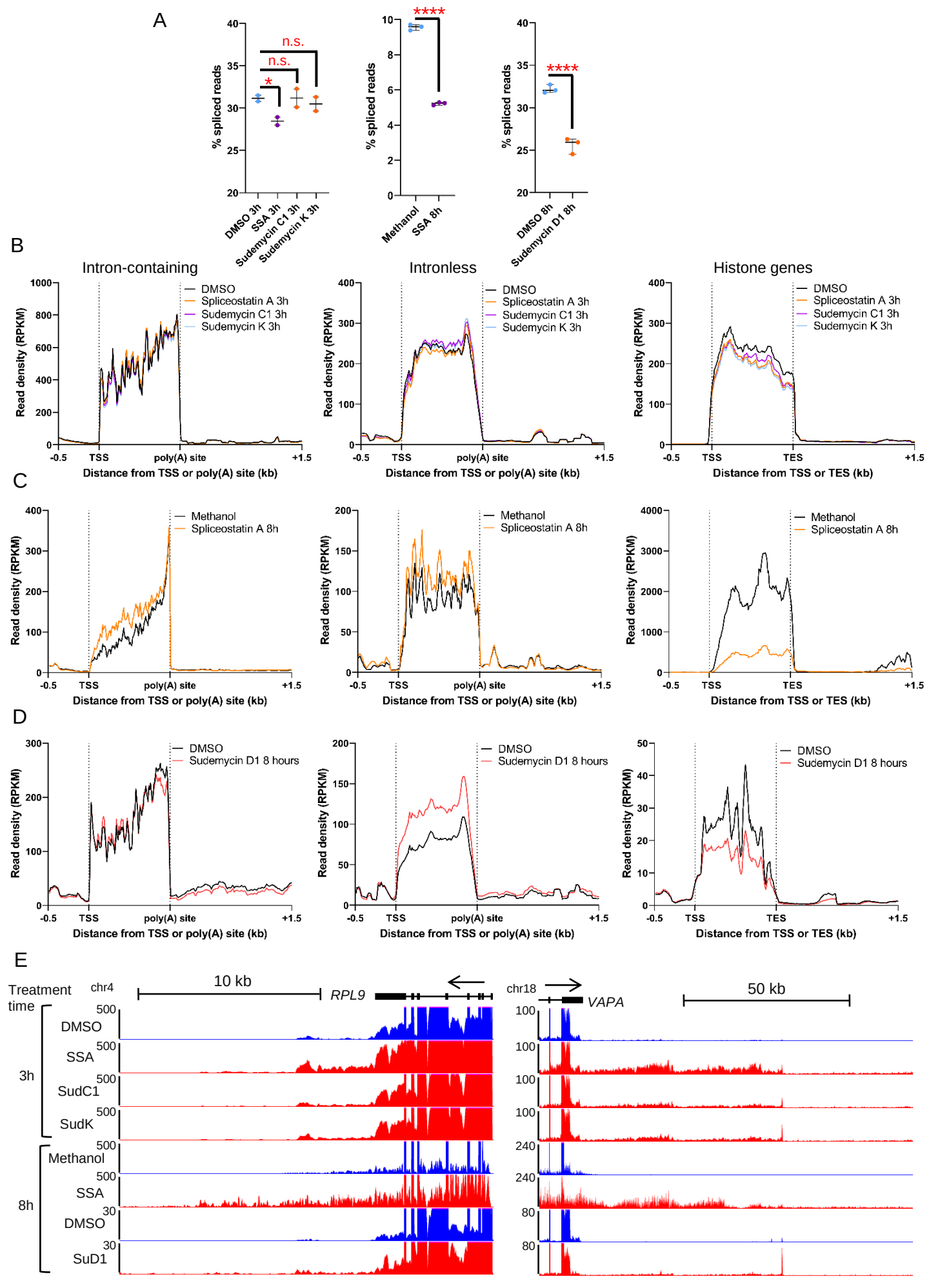
A
% spliced reads
n.s.
n.s.
*
DMSO 3h
SSA 3h
Sudemycin C1 3h
Sudemycin K 3h
****
Methanol
SSA 8h
****
DMSO 8h
Sudemycin D1 8h
B
Intron-containing
Intronless
Histone genes
Read density (RPKM)
DMSO
Spliceostatin A 3h
Sudemycin C1 3h
Sudemycin K 3h
-0.5
TSS
poly(A) site
+1.5
Distance from TSS or poly(A) site (kb)
TES
Distance from TSS or TES (kb)
C
Methanol
Spliceostatin A 8h
D
Sudemycin D1 8 hours
E
Treatment
time
chr4
10 kb
RPL9
chr18
VAPA
50 kb
3h
8h
DMSO
SSA
SudC1
SudK
Methanol
SuD1

Figure 7

Figure 7: SF3B1 inhibitors and knockdown of splicing factors do not induce a general transcription termination defect

(A) Percentage of spliced reads in total RNA-seq treated with DMSO (blue), Methanol (blue), SSA (purple), or Sudemycins C1, K, or D1 (orange), for the time indicated on the figure. Each point represents a biological replicate. Statistical test: Wilcoxon rank sum test. P-value: n.s. not significant, * < 0.05, **** < 0.0001. **(B)** Metagene profile of total RNA-seq performed in HeLa cells treated with DMSO (black), Spliceostatin A (orange), Sudemycin C1 (purple), or Sudemycin K (blue) for 3h on scaled expressed intron-containing, intronless, and histone genes. The metagene profiles are shown as the average of two biological replicates. **(C)** Metagene profile of total RNA-seq performed in HeLa cells treated with methanol (black) or Spliceostatin A (orange) for 8h on scaled expressed intron-containing, intronless, and histone genes. The metagene profiles are shown as the average of two biological replicates. **(D)** Metagene profile of total RNA-seq performed in Rh18 cells treated with DMSO (black) or Sudemycin D1 (red) for 8h on scaled expressed intron-containing, intronless, and histone genes. The metagene profiles are shown as the average of two biological replicates. **(E)** Screenshots of the genome browser total RNA-seq tracks for DMSO or methanol (blue) or following SF3B1 inhibition (red) with Spliceostatin A, Sudemycin C1, Sudemycin K, or Sudemycin D1 for 3h or 8h. The DOGs protein-coding genes *RPL9* and *VAPA* are shown. The arrow indicates the sense of transcription.