Isoginkgetin and Madrasin are poor splicing inhibitors

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. Michael Tellier: mt477@leicester.ac.uk; Shona Murphy: shona.murphy@path.ox.ac.uk

Keywords: RNA polymerase II, transcription, RNA splicing, splicing inhibitors, Isoginkgetin, Madrasin.

Abstract

The production of eukaryotic mRNAs requires transcription by RNA polymerase (pol) II and co-transcriptional processing, including capping, splicing, and cleavage and polyadenylation. 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, and pol II elongation rate can regulate splicing. Splicing, in turn, can also affect transcriptional activity and transcription elongation defects are caused by some splicing inhibitors. Multiple small molecule inhibitors of splicing are now available, some of which specifically target SF3B1, a U2 snRNP component. SF3B1 inhibition results in a general downregulation of transcription elongation, including premature termination of transcription caused by increased use of intronic poly(A) sites. Here, we have investigated the effect of Madrasin and Isoginkgetin, two non-SF3B1 splicing inhibitors, on splicing and transcription. Surprisingly, we found that both Madrasin and Isoginkgetin affect transcription before any effect on splicing, indicating that their effect on pre-mRNA splicing is likely to be indirect. Both small molecules promote a general downregulation of transcription. Based on these and other published results, we conclude that these two small molecules should not be considered as primarily pre-mRNA splicing inhibitors.

Introduction

The production of a mature mRNA from an eukaryotic protein-coding gene requires both transcription by RNA polymerase (pol) II and co-transcriptional processes, which include RNA capping, splicing, and cleavage and polyadenylation (1). Recent research has highlighted how transcription and pre-mRNA splicing can influence each other (2-8). 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 phosphorylation of Tyr1, Ser2, Thr4, Ser5, and Ser7 (9). While Ser2-P is associated with transcription elongation and Ser5P with transcription initiation and pol II pausing (10), Ser2-P and Ser5-P have also been found to specifically interact with splicing factors (11, 12), 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 (13). In addition, the pol II elongation rate, which can be

modified by cellular stresses or by mutating pol II, can affect the inclusion of exons in or the exclusion of exons from the mature mRNA in yeasts, mammals, and plants (2-4, 6, 14-18). In turn, 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, 19, 20), pol II CTD phosphorylation (5), and recruitment of the key transcription elongation kinase, P-TEFb (7, 21), which is composed of cyclin-dependent kinase (CDK)9 and Cyclin T1.

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 (22-25). A consequence of this coupling is that failure to splice the terminal intron in pre-mRNA is associated with a transcription termination defect, as the poly(A) site is no longer recognised by the mRNA CPA complex (19, 26, 27).

To further investigate the effects of pre-mRNA splicing inhibition on transcription elongation and transcription termination, we used two small molecules thought to be inhibitors of the early spliceosome, Madrasin (28) and Isoginkgetin (29). We found that in contrast to the SF3B1 inhibitors, PlaB and Herboxidiene (HB), Madrasin and Isoginkgetin are poor splicing inhibitors. However, treatment with Madrasin and Isoginkgetin globally downregulates pol II transcription, which was previously observed for Isoginkgetin (30). These two small molecules should therefore not be considered to be primarily pre-mRNA splicing inhibitors.

Materials and Methods

Cell culture

HeLa cells were obtained from ATCC (ATCC® CCL-2™) and grown in DMEM medium supplemented with 10% foetal bovine serum, 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mM L-glutamine at 37°C and 5% CO_2. HeLa cells were treated with 30-90 µM Madrasin (Sigma-Aldrich), 30 µM Isoginkgetin (Merck), 1 µM Herboxidiene (Cayman Chemical), 1 µM Pladienolide B (Cambridge Bioscience), or 100 µM 5,6-dichlorobenzimidazone-1-β-D-ribofuranoside (DRB, Sigma) 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 blotting

Western blot analysis was performed on chromatin extracts as previously described in the mNET-seq procedure (31). A ~ 80% confluent 15 cm dish was washed twice with ice-cold PBS and scrapped in 5 ml of ice-cold PBS. After pelleting the cells at 420 g for 5 min at 4°C and removing 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 transferred into 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 tube was 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 sample was centrifuged at 16,000 g for 10 min at 4°C and the supernatant discarded while the chromatin pellet was 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), 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 (32). 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 min with shaking. Quenching was performed with 125 mM Glycine for 5 min at room temperature with shaking. Cells were washed twice with ice-cold PBS and then scrapped in ice-cold PBS and pelleted for 10 min at 1,500 rpm for 10 min at 4°C. Lysis was performed on ice for 10 min 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 min 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 min at 4°C. Supernatants were transferred to new Eppendorf tubes and pre-cleared with 10 µl of Protein G Dynabeads for 30 min 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), Phospho-Rpb1 CTD (Ser2) (E1Z3G) Rabbit mAb (13499S, Cell Signaling Technology), Phospho-Rpb1 CTD (Ser5) (D9N5I) Rabbit mAb (13523S, Cell Signaling Technology), CDC73 (A300-170A, Bethyl Laboratories), SUPT5H antibody (A300-868A, Bethyl Laboratories), and CPSF73 (A301-091A, 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 min at 25°C at 1,400 rpm. Each input was diluted 1/10 in Elution buffer. The samples were treated with RNase A for 30 min 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 min 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.

Name	Forward primer	Reverse primer
PCR		
DNAJB1 E2/3	GAACCAAAATCACTTTCCCCAAGGAAGG	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
H1-2 GB	ACTCTGGTGCAAACGAAAGG	TTAACCTTGGGCTTGGCTTC
H1-2 T1	TCCCTCTGTTTGACATCCATGG	TGAGCAGCTTATCTCCAGAAGC

H1-2 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	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.

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 9.5 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 (31). In brief, the chromatin fraction was isolated from 4×10^7 HeLa cells treated with DMSO or 90 µM Madrasin for 30 min. 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 (25 mM) and soluble digested chromatin was collected after centrifugation at 13 000 rpm for 5 min 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 6 min 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 min. 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 1 min. The small RNAs were eluted from the gel using RNA elution buffer (1 M NaOAc and 1 mM EDTA) for 1h 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 as in (8). To obtain the list of expressed mRNA isoforms, Salmon version 0.14.1 (33) was used on two HeLa nucleoplasm RNA-seq (GSE110028 (34)) and the most expressed mRNA isoform of each gene was kept for further analysis. HeLa chromatin RNA-seq signal (GSE110028 (34)) was then quantified with BEDtools version 2.29.2 (35) multicov 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 (36) 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 (37) and the parameters: --runThreadN 16 --readFilesCommand gunzip -c -k --limitBAMsortRAM 20000000000 --outSAMtype BAM SortedByCoordinate. SAMtools version 1.9 (38) was used to retain the properly paired and mapped reads (-f 3). A custom python script (31) 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 (39) bamCoverage tool with the parameters -bs 1 -p max --normalizeUsing RPKM.

RNA-seq analysis

RNA-seq in HeLa cells treated for 12h with DMSO or 35 µM Isoginkgetin were taken from GSE72055 (40). Total RNA-seq in HeLa cells for 6h or 18h treatment with DMSO or 30 µM Isoginkgetin were taken from GSE86857 (30). RNA-seq were analysed as in (41). 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. 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 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 (42) count tool

and the Gencode v35 annotation has been used to obtain the number of counts per gene. DESeq2 version 1.30.1 (43) 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.

Splicing efficiency

RNA-seq splicing efficiency was calculated as previously (8) 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 for each expressed protein-coding gene with BEDtools multicov –s –split.

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 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

Isoginkgetin and Madrasin are poor splicing inhibitors

We first determined whether a short treatment with Madrasin or Isoginkgetin (Fig 1A) inhibits pre-mRNA splicing in HeLa cells as observed for PlaB and HB (7, 21). As the effect of Madrasin is less well characterised than Isoginkgetin, we first determined which concentration of Madrasin would be optimal for short-term inhibition. We treated HeLa cells with 30 µM, 60 µM, or 90 µM Madrasin for 30 min followed by RT-PCR analysis (S1A Fig). We observed a concentration-dependent inhibition of splicing of the intron located between exons 4 and 5 of *BRD2* and selected 90 µM as a working concentration. To compare the efficiency of the different splicing inhibitors, we treated HeLa cells for

1h with DMSO, 1 μM PlaB, 1 μM HB, 30 μM Isoginkgetin, or 90 μM Madrasin and performed RT-PCR for exons 2 and 3 of *DNAJB1* and exons 4 and 5 of *BRD2* (Fig 1B). While SF3B1 inhibition with PlaB and HB shows, as expected, a clear increase in intron retention for both genes, Isoginkgetin and Madrasin have little effect on pre-mRNA splicing of these two genes. As SF3B1 inhibition also affects transcription (5, 7, 19, 21), we investigated nascent transcription activity via 5'EU incorporation in HeLa cells treated for 1h with DMSO, 100 μM DRB, a CDK9 inhibitor, 1 μM PlaB, 1 μM HB, 30 μM Isoginkgetin, or 90 μM Madrasin (Fig 1C and D). As we previously observed, CDK9 and SF3B1 inhibition results in a reduction in 5'EU incorporation (21) whereas Madrasin, but not Isoginkgetin, decreases the 5'EU signal.

These results indicate that both Madrasin and Isoginkgetin are poor splicing inhibitors but that Madrasin affects transcription.

Fig 1: Isoginkgetin and Madrasin are poor splicing inhibitors.

(A) Chemical structure of Madrasin and Isoginkgetin. **(B)** RT-PCR with primers amplifying the intron located between exons 2 and 3 of *DNAJB1* or exons 4 and 5 of *BRD2*. HeLa cells were treated with DMSO, 1 μM PlaB, 1 μM HB, 30 μM Isoginkgetin, or 90 μM Madrasin for 1h. 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)** Representative images of immunofluorescence analysis or 5'EU incorporation in HeLa cells treated with DMSO, 100 μM DRB, a CDK9 inhibitor, 1 μM PlaB, 1 μM HB, 30 μM Isoginkgetin, or 90 μM Madrasin for 1h. EU (green), DAPI (blue), scale bars: 50 μm. **(D)** Quantification of 5'EU intensity per nucleus for DMSO, DRB, PlaB, HB, Isoginkgetin, and Madrasin. 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: n.s. not significant, **** < 0.0001.

Madrasin decreases transcription of protein-coding genes

As Madrasin reduces 5'EU signal, we first investigated by qRT-PCR whether rDNA transcription is decreased after 30 min or 60 min treatment with 90 μM Madrasin (S1B Fig). As the pre-rRNA level is not affected by Madrasin, we investigated pol II nascent transcription by performing total pol II and Ser5-P mNET-seq in HeLa cells following 30 min treatment with DMSO (Control) or 90 μM Madrasin

(Fig 2A and B and S1C Fig). Madrasin treatment causes a decrease in pol II pausing and in pol II levels across the gene body on protein-coding genes while the pol II level increases downstream of poly(A) sites, indicating a transcription termination defect. We confirmed the decrease in pol II transcription following Madrasin treatment in mNET-seq by qRT-PCR on eight protein-coding genes (Fig 2C) and by pol II ChIP-qPCR on our model protein-coding gene *KPNB1* (Fig 2D and E). To understand whether the effect of Madrasin on transcription could be mediated by a loss of transcription elongation and termination factors, we performed ChIP-qPCR on the *KPNB1* gene of the elongation factors SPT5 and CDC73, a subunit of the PAF1 complex, and of CPSF73, the pre-mRNA cleavage enzyme, in HeLa cells treated with DMSO or 90 µM Madrasin for 30 min (Fig 2E and F). When the results are not ratioed to pol II (Fig 2E), the effect of Madrasin on SPT5 and CDC73 mimics the effect on pol II, with a reduction across the whole gene body and an increase in the delayed transcription termination region and the CPSF73 signal is not affected by Madrasin treatment. When ratioed to pol II, there is a relative increase in SPT5, CDC73, and CPSF73 across the gene body of *KPNB1* but a reduced level of the three proteins on the delayed transcription termination region following Madrasin treatment (Fig 2F).

These results indicate that while Madrasin has a limited effect on pre-mRNA splicing, it downregulates pol II transcription of protein-coding genes and causes a transcription termination defect.

Fig 2: Madrasin decreases transcription of protein-coding genes

(A) Schematic of the mNET-seq technique. **(B)** Metagene profile of total pol II mNET-seq treated with DMSO (blue) or 90 µM Madrasin (red) for 30 min on scaled expressed protein-coding genes. **(C)** qRT-PCR with primers amplifying different protein-coding genes. HeLa cells were treated with DMSO or 90 µM Madrasin for 30 min (red) or 60 min (orange). 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. **(D)** Screenshot of the genome browser total pol II mNET-seq DMSO (blue) and Madrasin (red) tracks on the protein-coding gene *KPNB1*. The arrow indicates the sense of transcription. **(E)** Total pol II, SPT5, CDC73, and CPSF73 ChIP-qPCR across the protein-coding gene *KPNB1* in HeLa cells treated with DMSO (blue) or 90 µM Madrasin for 30 min (red). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, ** < 0.01, *** < 0.001. **(F)** Ratios of SPT5 /

total pol II, CDC73 / total pol II, or CPSF73 / total pol II from ChIP-qPCR on the intron-containing gene *KPNB1* in HeLa cells treated with DMSO (blue) or 90 µM Madrasin for 30 min (red). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001.

Transcription of intronless and histone genes is also affected by Madrasin

As Madrasin promotes a general transcriptional defect on protein-coding genes, which also include intronless and histone 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) (S2A Fig). The RNA level from the intron-containing and intronless genes is similar in chromatin RNA-seq, while histone genes produce more RNA (S2B Fig). As expected, transcription of intron-containing genes is similarly affected by Madrasin as all protein-coding genes (S2C Fig). Madrasin also decreases nascent transcription of intronless genes and causes a higher pol II signal downstream of the poly(A) site (Fig 3A and S2D Fig). Interestingly, Madrasin activates the expression of some intronless genes, such as *JUN* (Fig 3B). We confirmed the increase in pol II, SPT5, and CDC73, but not CPSF73, on *JUN* after Madrasin treatment by ChIP-qPCR (Fig 3C and S2E Fig). Histone genes are also affected by Madrasin, mostly promoting a higher pol II signal downstream of the histone genes without affecting the level of nascent transcripts across the coding region (Fig 3D and E and S2F Fig). While pol II ChIP-qPCR on the *H1-2* gene confirmed the mNET-seq profile, Madrasin treatment reduces SPT5 and CDC73, but not CPSF73, on this histone gene (Fig 2F and S2G Fig).

These results indicate that Madrasin globally affects pol II transcription on intron-containing and intronless genes, and in the case of histone genes, mostly on transcription downstream of the normal transcription termination region.

Fig 3: Transcription of intronless and histone genes is 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 min (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, SPT5, CDC73, and CPSF73 ChIP-qPCR on the intronless gene *JUN* in HeLa cells treated with DMSO (blue) or 90 µM Madrasin for 30 min (red). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, ** < 0.01, *** < 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 min (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 *H1-2*. The arrow indicates the sense of transcription. **(F)** Total pol II, SPT5, CDC73, and CPSF73 ChIP-qPCR on the histone gene *H1-2* in HeLa cells treated with DMSO (blue) or 90 µM Madrasin for 30 min (red). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05.

Isoginkgetin treatment affects pol I and pol II transcription

While a 1h treatment with 30 µM Isoginkgetin did not affect pre-mRNA splicing and 5'EU incorporation (Fig 1B-D), we investigated whether a longer treatment time would result in a change in splicing and transcription as Isoginkgetin was already shown to affect transcription after 6h of treatment (30). We treated HeLa cells with 30 µM Isoginkgetin from 1h to 6h and found virtually no effect on pre-mRNA splicing as analysed by RT-PCR on *DNAJB1* and *BRD2* genes (Fig 4A). Re-analysis of previously-published total RNA-seq datasets (30, 40) in HeLa cells following 6h, 12h, or 18h treatment with 30 µM Isoginkgetin confirms that pre-mRNA splicing is not affected by Isoginkgetin at 6h but is significantly decreased at 12h and 18h (Fig 4B). Isoginkgetin is therefore unlikely to act on pre-mRNA splicing as a primary effect, in agreement with (30). To test whether Isoginkgetin affects transcription prior to 6h treatment, we first monitored 5'EU incorporation in HeLa cells following a time-course of 1h to 6h treatment with 30 µM Isoginkgetin (Fig 4C and D). It takes approximately 5h of Isoginkgetin treatment to see a marked decrease in 5'EU incorporation while shorter treatment time points show rather an increase in 5'EU incorporation. To determine whether pre-rRNA transcription could be responsible for the change in 5'EU incorporation, we investigated whether Isoginkgetin affects pol I transcription by qRT-PCR (Fig 4E). Transcription of rDNA starts to decrease after 2h treatment with Isoginkgetin but becomes significantly reduced only at 4h and 5h.

Fig 4: Isoginkgetin is a poor splicing inhibitor

(A) 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 PlaB 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. **(B)** 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. **(C)** 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. **(D)** Quantification of 5'EU intensity per nucleus for DMSO and Isoginkgetin. 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. **(E)** qRT-PCR with primers amplifying a region from the pol I transcribed pre-rRNA. HeLa cells were treated with DMSO or 30 µM Isoginkgetin 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.

We next investigated by oligo d(T) qRT-PCR the expression of 11 protein-coding genes in HeLa cells treated with DMSO or 30 µM Isoginkgetin from 1h to 6h (Fig 5A). We observed a detectable effect on the production of poly(A)+ mRNA from 3h onwards (*NR4A3, LDLR*, and *KPNB1*), with generally a downregulation of protein-coding gene expression, with the exception of the intronless gene *JUN*, following Isoginkgetin treatment. 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 (S3A and B Fig). *ATF3* expression was previously shown to be upregulated in HeLa cells after a 24h treatment with Isoginkgetin (44), while we observed a decrease in ATF3 expression after 6h treatment (Fig 5A). We performed qRT-PCR following a 24h treatment with Isoginkgetin and confirmed that both *JUN* and *ATF3* are upregulated while four other protein-coding genes are downregulated (S3C Fig). 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 (Fig 5B-D and S3D Fig). We did not observe any change in total pol II across *KPNB1* after treating HeLa cells with Isoginkgetin for 30 min (S3D Fig). However, an Isoginkgetin time-course from 1h to 6h shows a decrease in pol II signal from 2h onwards across *KPNB1*, *JUN*, and *H1-2* (Fig 5B-D), which would explain the downregulation in poly(A)+ mRNA production apparent from 3h treatment onwards.

To determine whether Isoginkgetin could affect pol II CTD phosphorylation before transcription, we purified the chromatin fraction of HeLa cells treated with DMSO or 30 µM Isoginkgetin for 1h, 2h, or 4h and performed western blots for total pol II, Ser2-P, and Ser5-P (Fig 5E). We generally observed a loss of total pol II and of Ser2 and Ser5 phosphorylation on chromatin over time, with the exception of an increase in Ser5-P at the 1h timepoint. To confirm these observations, we performed pol II, Ser2-P, and Ser5-P ChIP-qPCR across the *KPNB1* gene following Isoginkgetin treatment for 1h, 1h30, and 2h (Fig 5F). The changes in total pol II profile looks similar to what we previously observed after Isoginkgetin treatment (Fig 5B) and the changes in Ser2-P profile match the changes in total pol II, indicating that this mark is largely unaffected. However, the Ser5-P level increases at the promoter region after 1h treatment but returns to control levels after 1h30, in agreement with the chromatin western blot.

Isoginkgetin is a bioflavonoid, a group of small molecules that have been shown to regulate the SUMO post-translational modification (45). The SUMO modification is present on mRNA splicing and CPA factors and regulates their activities (45-47). To determine whether 1h, 2h, or 4h treatment of 30 µM Isoginkgetin modifies SUMO1 or SUMO2/3 levels, we performed western blots on HeLa whole cell extract (S3E Fig). We only observed a slight increase in total SUMO2/3 level after 4h treatment.

These results indicate that Isoginkgetin affects pol I and pol II transcription and pol II CTD phosphorylation rather than pre-mRNA splicing. The effect of Isoginkgetin on transcription is also not likely mediated via a global change in SUMOylation.

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

(A) 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 protein-coding gene *GAPDH* and shown as relative to DMSO, mean ± SEM, n = 4

biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001. **(B)** Total pol II ChIP-qPCR on the intron-containing gene *KPNB1* in HeLa cells treated with DMSO 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. **(C)** Total pol II ChIP-qPCR on the intronless gene *JUN* in HeLa cells treated with DMSO 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. **(D)** Total pol II ChIP-qPCR of the histone gene *H1-2* in HeLa cells treated with DMSO 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. **(E)** Western blots of total pol II, Ser2-P, Ser5-P, and histone H3 on the chromatin fraction of HeLa cells treated with DMSO or 30 µM Isoginkgetin for 1h, 2h, or 4h. 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. **(F)** ChIP-qPCR of total pol II, Ser2-P, and Ser5-P across the intron-containing gene *KPNB1* in HeLa cells treated with DMSO 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.

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 (28, 29). To limit confounding effects associated with long treatment times, we have carried out short time courses of treatment. However, we rather found that both Madrasin and Isoginkgetin are poor splicing inhibitors and should therefore not considered as primarily inhibitors of pre-mRNA splicing.

While Madrasin has a limited effect on pre-mRNA splicing following a 30 min treatment, this small molecule promotes a general transcriptional downregulation and a delay in transcription termination, which is observed on intron-containing, intronless, and histone genes. The transcriptional dysregulation resulting from Madrasin could potentially explain the longer-term indirect effect on splicing via changes in proteins interacting with pol II, as we observed for the elongation factors SPT5 and CDC73. Another non-mutually exclusive possibility is a higher pol II elongation rate following Madrasin treatment, which could also explain the delayed transcription termination, as pol II

elongation rate is known to regulate RNA processing (2-4, 6, 16, 48). However, the cellular target(s) of Madrasin is unknown.

In line with a previous report (30), 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 and pol II CTD Ser2 and Ser5 phosphorylation in ~1h and pol I transcription in ~2h. 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 partially explained by the transcriptional readthrough previously observed for Isoginkgetin (30, 40) (S4 Fig) from upstream protein-coding genes into ncRNA genes, transforming them into expressed genes in the differential expression analysis software. However, we could not observe a transcription termination defect nor the increase in total, but not poly(A)+, RNA-seq signal in the first intron (30) following Isoginkgetin treatment by total pol II ChIP-qPCR. As Isoginkgetin mimics RNA exosome inhibition (40), it is therefore likely that the accumulation of poorly processed transcripts (increased signal in the first intron, increased signal downstream of the poly(A) site, and increased non-polyadenylated ncRNAs) are due to a loss of RNA exosome activity rather than a transcriptional defect. The cellular target(s) of Isoginkgetin is also unknown. Isoginkgetin can partially inhibit SENP1 (45), a protease deconjugating SUMO modification, *in vitro,* and in agreement with this, we start to see an increase in global SUMO2/3 levels after 4h treatment. It remains to be determined whether the effect of Isoginkgetin on pol I and pol II transcription is mediated by rapid changes (< 4h) in SUMO modifications on proteins involved in transcription.

Understanding the coupling between transcription, pre-mRNA splicing, and mRNA CPA remains technically challenging. Based on our and other groups' analyses, Madrasin and Isoginkgetin should not be considered splicing inhibitors. It is also important to use short timepoints to screen for potential splicing inhibitors rather than 24h treatment in order to limit the influence of secondary/indirect effects.

Acknowledgments

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

REFERENCES

1. Tellier M, Maudlin I, Murphy S. Transcription and splicing: A two-way street. Wiley Interdiscip Rev RNA. 2020;11(5):e1593.
2. Howe KJ, Kane CM, Ares M, Jr. Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae. RNA. 2003;9(8):993-1006.
3. de la Mata M, Alonso CR, Kadener S, Fededa JP, Blaustein M, Pelisch F, et al. A slow RNA polymerase II affects alternative splicing in vivo. Mol Cell. 2003;12(2):525-32.
4. Dujardin G, Lafaille C, de la Mata M, Marasco LE, Munoz MJ, Le Jossic-Corcos C, et al. How slow RNA polymerase II elongation favors alternative exon skipping. Mol Cell. 2014;54(4):683-90.
5. Koga M, Hayashi M, Kaida D. Splicing inhibition decreases phosphorylation level of Ser2 in Pol II CTD. Nucleic Acids Res. 2015;43(17):8258-67.
6. Godoy Herz MA, Kubaczka MG, Brzyzek G, Servi L, Krzyszton M, Simpson C, et al. Light Regulates Plant Alternative Splicing through the Control of Transcriptional Elongation. Mol Cell. 2019;73(5):1066-74 e3.
7. Caizzi L, Monteiro-Martins S, Schwalb B, Lysakovskaia K, Schmitzova J, Sawicka A, et al. Efficient RNA polymerase II pause release requires U2 snRNP function. Mol Cell. 2021;81(9):1920-34 e9.
8. Henfrey C, Murphy S, Tellier M. Regulation of mature mRNA levels by RNA processing efficiency. NAR Genom Bioinform. 2023;5(2):lqad059.
9. Zaborowska J, Egloff S, Murphy S. The pol II CTD: new twists in the tail. Nat Struct Mol Biol. 2016;23(9):771-7.
10. Komarnitsky P, Cho EJ, Buratowski S. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev. 2000;14(19):2452-60.
11. Harlen KM, Trotta KL, Smith EE, Mosaheb MM, Fuchs SM, Churchman LS. Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. Cell Rep. 2016;15(10):2147-58.
12. Nojima T, Rebelo K, Gomes T, Grosso AR, Proudfoot NJ, Carmo-Fonseca M. RNA Polymerase II Phosphorylated on CTD Serine 5 Interacts with the Spliceosome during Co-transcriptional Splicing. Mol Cell. 2018;72(2):369-79 e4.
13. Zhang S, Aibara S, Vos SM, Agafonov DE, Luhrmann R, Cramer P. Structure of a transcribing RNA polymerase II-U1 snRNP complex. Science. 2021;371(6526):305-9.
14. Eperon LP, Graham IR, Griffiths AD, Eperon IC. Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? Cell. 1988;54(3):393-401.
15. Braberg H, Jin H, Moehle EA, Chan YA, Wang S, Shales M, et al. From structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II. Cell. 2013;154(4):775-88.
16. Fong N, Kim H, Zhou Y, Ji X, Qiu J, Saldi T, et al. Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. Genes Dev. 2014;28(23):2663-76.
17. Saldi T, Cortazar MA, Sheridan RM, Bentley DL. Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing. J Mol Biol. 2016;428(12):2623-35.

18. Aslanzadeh V, Huang Y, Sanguinetti G, Beggs JD. Transcription rate strongly affects splicing fidelity and cotranscriptionality in budding yeast. Genome Res. 2018;28(2):203-13.
19. Sousa-Luis R, Dujardin G, Zukher I, Kimura H, Weldon C, Carmo-Fonseca M, et al. POINT technology illuminates the processing of polymerase-associated intact nascent transcripts. Mol Cell. 2021;81(9):1935-50 e6.
20. Yoshimoto R, Chhipi-Shrestha JK, Schneider-Poetsch T, Furuno M, Burroughs AM, Noma S, et al. Spliceostatin A interaction with SF3B limits U1 snRNP availability and causes premature cleavage and polyadenylation. Cell Chem Biol. 2021;28(9):1356-65 e4.
21. Ansa G, Murphy S, Tellier M. Inhibition of SF3B1 affects recruitment of P-TEFb to chromatin through multiple mechanisms. bioRxiv. 2024:2024.06.26.600844.
22. Gunderson SI, Vagner S, Polycarpou-Schwarz M, Mattaj IW. Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation. Genes Dev. 1997;11(6):761-73.
23. Vagner S, Vagner C, Mattaj IW. The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. Genes Dev. 2000;14(4):403-13.
24. Kyburz A, Friedlein A, Langen H, Keller W. Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3' end processing and splicing. Mol Cell. 2006;23(2):195-205.
25. Tellier M, Zaborowska J, Neve J, Nojima T, Hester S, Fournier M, et al. CDK9 and PP2A regulate RNA polymerase II transcription termination and coupled RNA maturation. EMBO Rep. 2022;23(10):e54520.
26. Zhao N, Sebastiano V, Moshkina N, Mena N, Hultquist J, Jimenez-Morales D, et al. Influenza virus infection causes global RNAPII termination defects. Nat Struct Mol Biol. 2018;25(9):885-93.
27. Reimer KA, Mimoso CA, Adelman K, Neugebauer KM. Co-transcriptional splicing regulates 3' end cleavage during mammalian erythropoiesis. Mol Cell. 2021;81(5):998-1012 e7.
28. Pawellek A, McElroy S, Samatov T, Mitchell L, Woodland A, Ryder U, et al. Identification of small molecule inhibitors of pre-mRNA splicing. J Biol Chem. 2014;289(50):34683-98.
29. O'Brien K, Matlin AJ, Lowell AM, Moore MJ. The biflavonoid isoginkgetin is a general inhibitor of Pre-mRNA splicing. J Biol Chem. 2008;283(48):33147-54.
30. Boswell SA, Snavely A, Landry HM, Churchman LS, Gray JM, Springer M. Total RNA-seq to identify pharmacological effects on specific stages of mRNA synthesis. Nat Chem Biol. 2017;13(5):501-7.
31. Nojima T, Gomes T, Grosso ARF, Kimura H, Dye MJ, Dhir S, et al. Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. Cell. 2015;161(3):526-40.
32. Tellier M, Zaborowska J, Caizzi L, Mohammad E, Velychko T, Schwalb B, et al. CDK12 globally stimulates RNA polymerase II transcription elongation and carboxyl-terminal domain phosphorylation. Nucleic Acids Res. 2020;48(14):7712-27.
33. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14(4):417-9.
34. Nojima T, Tellier M, Foxwell J, Ribeiro de Almeida C, Tan-Wong SM, Dhir S, et al. Deregulated Expression of Mammalian lncRNA through Loss of SPT6 Induces R-Loop Formation, Replication Stress, and Cellular Senescence. Mol Cell. 2018;72(6):970-84 e7.
35. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2.

36. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011. 2011;17(1):3.
37. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
38. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.
39. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44(W1):W160-5.
40. Tseng CK, Wang HF, Burns AM, Schroeder MR, Gaspari M, Baumann P. Human Telomerase RNA Processing and Quality Control. Cell Rep. 2015;13(10):2232-43.
41. Tellier M, Murphy S. Incomplete removal of ribosomal RNA can affect chromatin RNA-seq data analysis. Transcription. 2020;11(5):230-5.
42. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-9.
43. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
44. Vanzyl EJ, Sayed H, Blackmore AB, Rick KRC, Fernando P, McKay BC. The spliceosome inhibitors isoginkgetin and pladienolide B induce ATF3-dependent cell death. PLoS One. 2020;15(12):e0224953.
45. Pawellek A, Ryder U, Tammsalu T, King LJ, Kreinin H, Ly T, et al. Characterisation of the biflavonoid hinokiflavone as a pre-mRNA splicing modulator that inhibits SENP. Elife. 2017;6.
46. Vethantham V, Rao N, Manley JL. Sumoylation modulates the assembly and activity of the pre-mRNA 3' processing complex. Mol Cell Biol. 2007;27(24):8848-58.
47. Pozzi B, Bragado L, Will CL, Mammi P, Risso G, Urlaub H, et al. SUMO conjugation to spliceosomal proteins is required for efficient pre-mRNA splicing. Nucleic Acids Res. 2017;45(11):6729-45.
48. Cortazar MA, Sheridan RM, Erickson B, Fong N, Glover-Cutter K, Brannan K, et al. Control of RNA Pol II Speed by PNUTS-PP1 and Spt5 Dephosphorylation Facilitates Termination by a "Sitting Duck Torpedo" Mechanism. Mol Cell. 2019;76(6):896-908 e4.

Supporting information

S1 Fig. Madrasin decreases transcription of protein-coding genes

(**A**) 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 min. 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. **(B)** 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 min (red) or 60 min (orange). 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. **(C)** Metagene profile of pol II CTD Ser5-P

mNET-seq performed in HeLa cells treated with DMSO (blue) or 90 µM Madrasin for 30 min (red) on scaled expressed protein-coding genes.

S2 Fig. Transcription of intronless and histone genes is also affected by Madrasin

(A) Violin plots, defined as min to max values, first and third quartiles (thin dotted lines), and mean (thicker dotted line), showing the $\log_{10}$ of gene length distribution (in bp) for intron-containing (red), intronless (blue), and histone (black) genes. Statistical test: Wilcoxon rank sum test. P-value: *** < 0.001, **** < 0.0001. **(B)** Violin plots, defined as min to max values, first and third quartiles (thin dotted lines), and mean (thicker dotted line), showing the $\log_{10}$ of gene expression distribution, from two biological HeLa chromatin RNA-seq experiments, for intron-containing (red), intronless (blue), and histone (black) genes. Statistical test: Wilcoxon rank sum test. P-value: ns: not significant, **** < 0.0001. **(C)** Metagene profile of total pol II mNET-seq performed in HeLa cells treated with DMSO (blue) or 90 µM Madrasin for 30 min (red) on scaled expressed intron-containing genes. **(D)** Screenshots of the genome browser total pol II mNET-seq DMSO (blue) and Madrasin (red) tracks of the intronless genes *SOX4* and *SF3B5*. The arrow indicates the sense of transcription. **(E)** Ratios of SPT5 / total pol II, CDC73 / total pol II, or CPSF73 / total pol II from ChIP-qPCR on the intronless gene *JUN* in HeLa cells treated with DMSO (blue) or 90 µM Madrasin for 30 min (red). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05. **(F)** Screenshot of the genome browser total pol II mNET-seq DMSO (blue) and Madrasin (red) tracks of the histone gene *H1-4*. The arrow indicates the sense of transcription. **(G)** Ratios of SPT5 / total pol II, CDC73 / total pol II, or CPSF73 / total pol II from ChIP-qPCR on the histone gene *H1-2* in HeLa cells treated with DMSO (blue) or 90 µM Madrasin for 30 min (red). Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: * < 0.05, *** < 0.001.

S3 Fig. Isoginkgetin treatment affects pol I and pol II transcription

(A) MA plot of total RNA-seq in HeLa cells treated with DMSO or Isoginkgetin for 6h or 18h. Significantly upregulated genes are shown in red while significantly downregulated genes are shown in blue. Other genes are shown in grey. The numbers of differentially expressed genes are shown on the figure. **(B)** MA plot of total RNA-seq in HeLa cells treated with DMSO or Isoginkgetin for 12h. Significantly upregulated genes are shown in red while significantly downregulated genes are shown

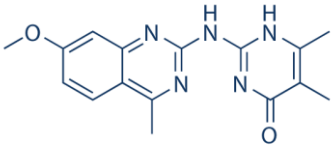
in blue. Other genes are shown in grey. The numbers of differentially expressed genes are shown on the figure. **(C)** qRT-PCR with primers amplifying different protein-coding genes. HeLa cells were treated with DMSO or 30 µM Isoginkgetin for 24h. cDNA was generated with oligo(dT). Values are normalised to the protein-coding gene *GAPDH* 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, **** < 0.0001. **(D)** Total pol II ChIP-qPCR of the intron-containing gene *KPNB1* in HeLa cells treated with DMSO (black) or 30 µM Isoginkgetin (white) for 30 min. Mean ± SEM, n = 3 biological replicates. Statistical test: two-tailed unpaired t-test. P-value: not significant. **(E)** Western blot on whole cell extract of HeLa cells treated with DMSO or 30 µM Isoginkgetin for 1h, 2h, or 4h. The antibodies are indicated at the bottom, with Ponceau S staining used as loading control.

S4 Fig. Isoginkgetin treatment affects transcription termination in RNA-seq

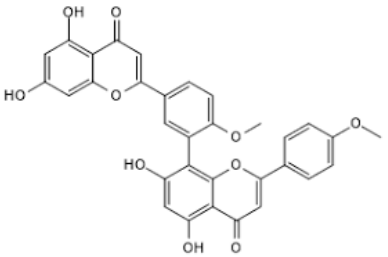
(A) Metagene profile of total RNA-seq performed in HeLa cells treated with DMSO (blue) or Isoginkgetin (red) for 12h on scaled expressed intron-containing, intronless, and histone genes. The metagene profiles are shown as the average of two biological replicates. **(B)** Screenshots of the genome browser tracks for HeLa RNA-seq treated with DMSO (blue) or Isoginkgetin (red) for 12h on the intron-containing gene *KPNB1*, intronless genes *SOX4*, *SF3B5*, and *JUN*, and histone genes *H1-2*, *H1-4*, and *H2BC5*. The arrow indicates the sense of transcription.

A

Madrasin

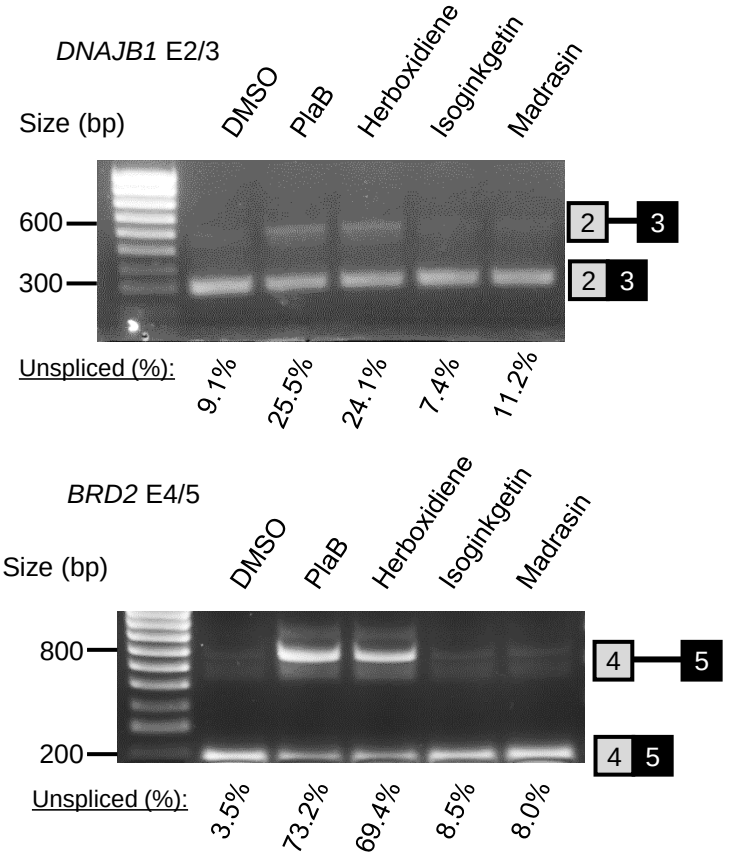


Isoginkgetin



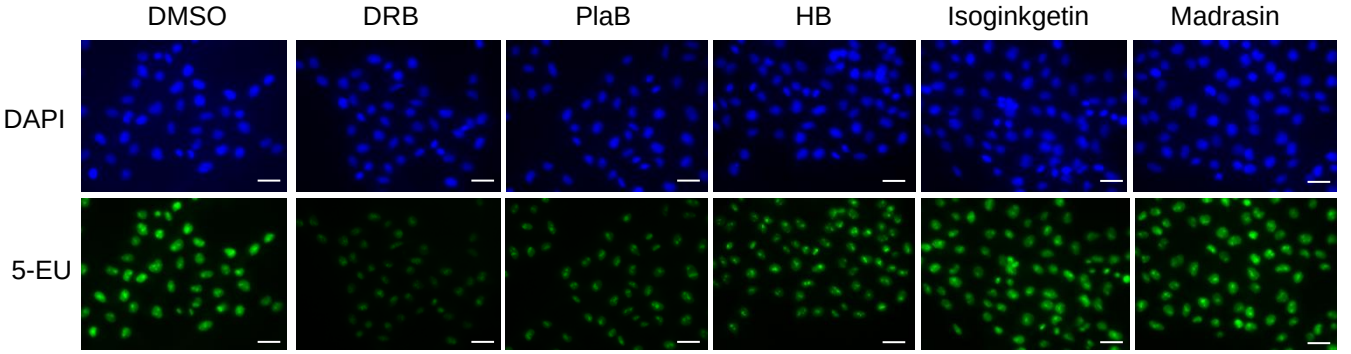
B

HeLa, 1h treatment



C

HeLa, 1h treatment



D

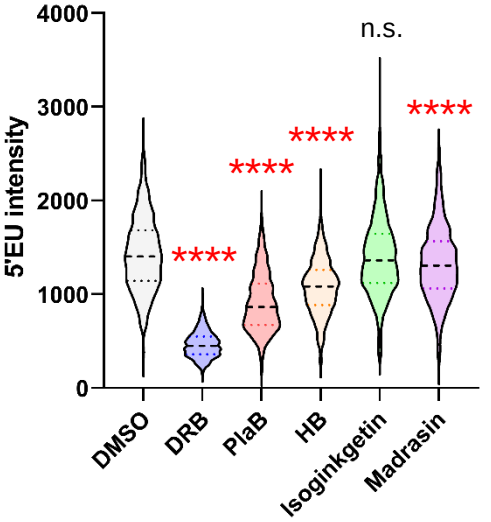


Fig 1

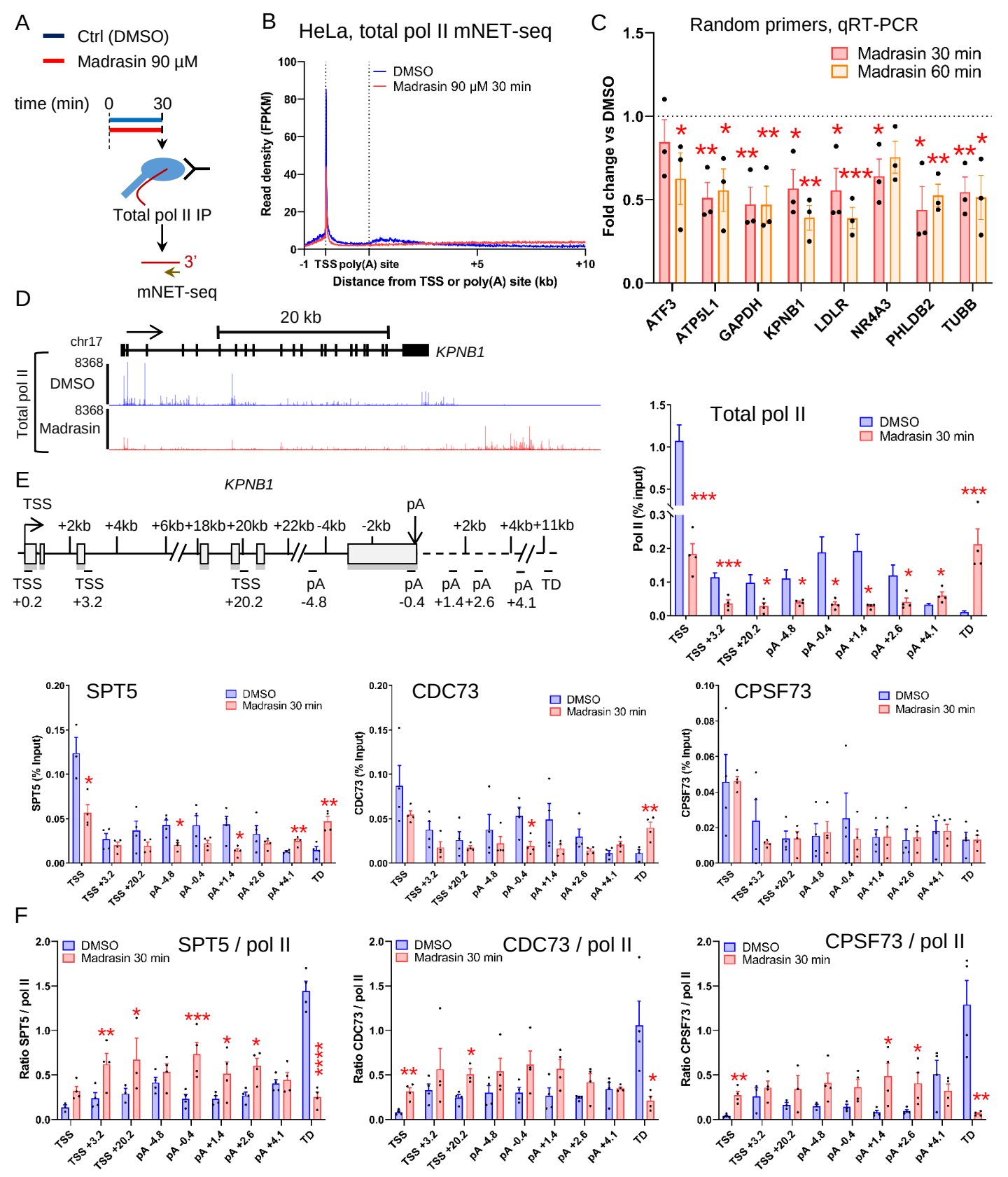

A
Ctrl (DMSO)
Madrasin 90 µM
time (min) 0 30
Total pol II IP
3'
mNET-seq
B
HeLa, total pol II mNET-seq
DMSO
Madrasin 90 µM 30 min
Read density (FPKM)
-1 TSS poly(A) site +5 +10
Distance from TSS or poly(A) site (kb)
C
Random primers, qRT-PCR
Madrasin 30 min
Madrasin 60 min
Fold change vs DMSO
ATF3 ATP5L1 GAPDH KPNB1 LDLR NR4A3 PHLDB2 TUBB
D
20 kb
chr17
KPNB1
Total pol II
DMSO
Madrasin
E
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
Total pol II
Pol II (% input)
SPT5
SPT5 (% Input)
CDC73
CDC73 (% Input)
CPSF73
CPSF73 (% Input)
F
SPT5 / pol II
Ratio SPT5 / pol II
CDC73 / pol II
Ratio CDC73 / pol II
CPSF73 / pol II
Ratio CPSF73 / pol II

Fig 2

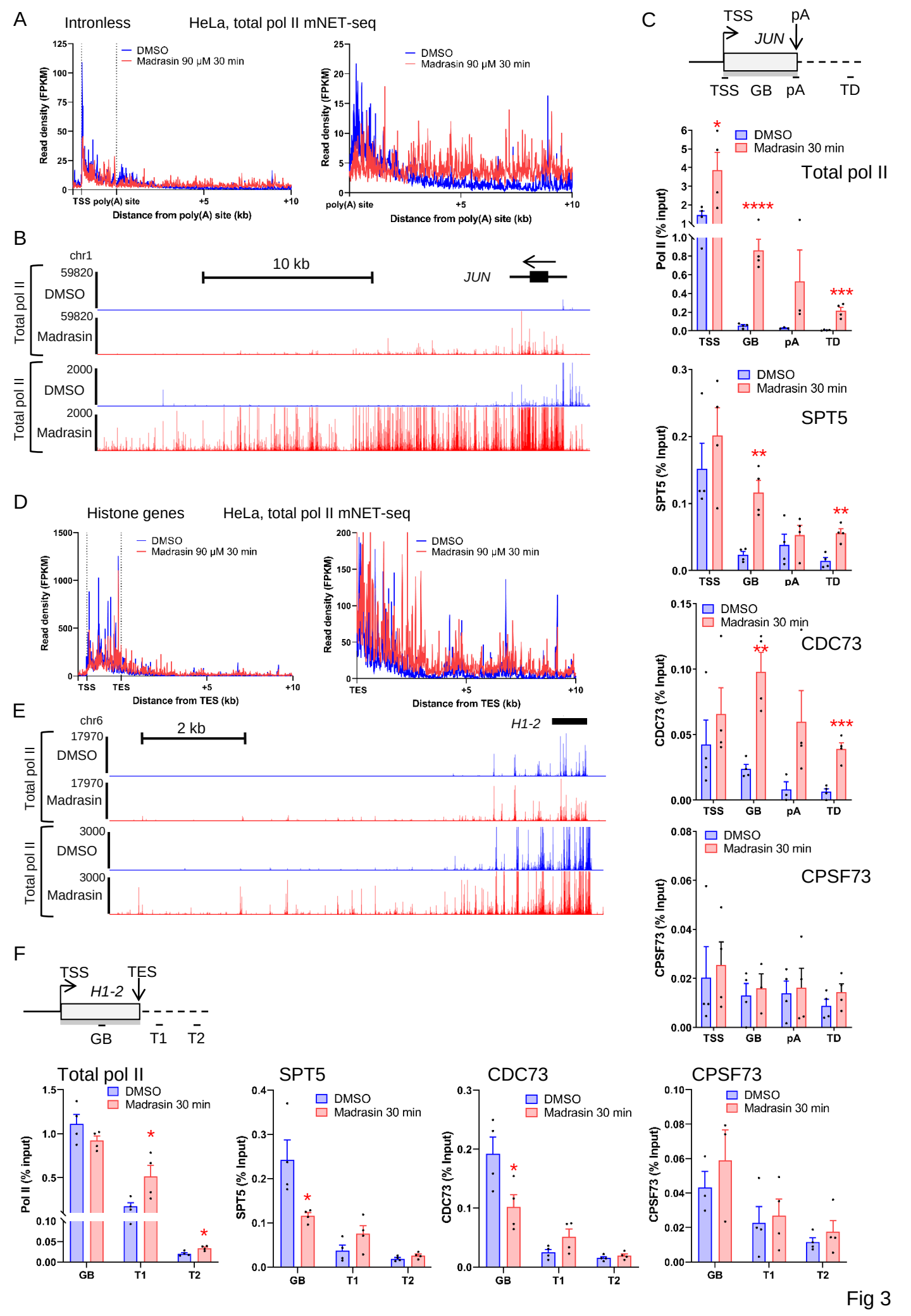

A
Intronless
HeLa, total pol II mNET-seq
DMSO
Madrasin 90 μM 30 min
Read density (FPKM)
TSS poly(A) site
+5
+10
Distance from poly(A) site (kb)
poly(A) site
B
chr1
10 kb
JUN
Total pol II
59820
DMSO
Madrasin
2000
C
TSS
pA
JUN
GB
TD
Total pol II
Madrasin 30 min
Pol II (% input)
SPT5
SPT5 (% Input)
CDC73
CDC73 (% Input)
CPSF73
CPSF73 (% Input)
D
Histone genes
Distance from TES (kb)
E
chr6
2 kb
H1-2
17970
3000
F
T1
T2

Fig 3

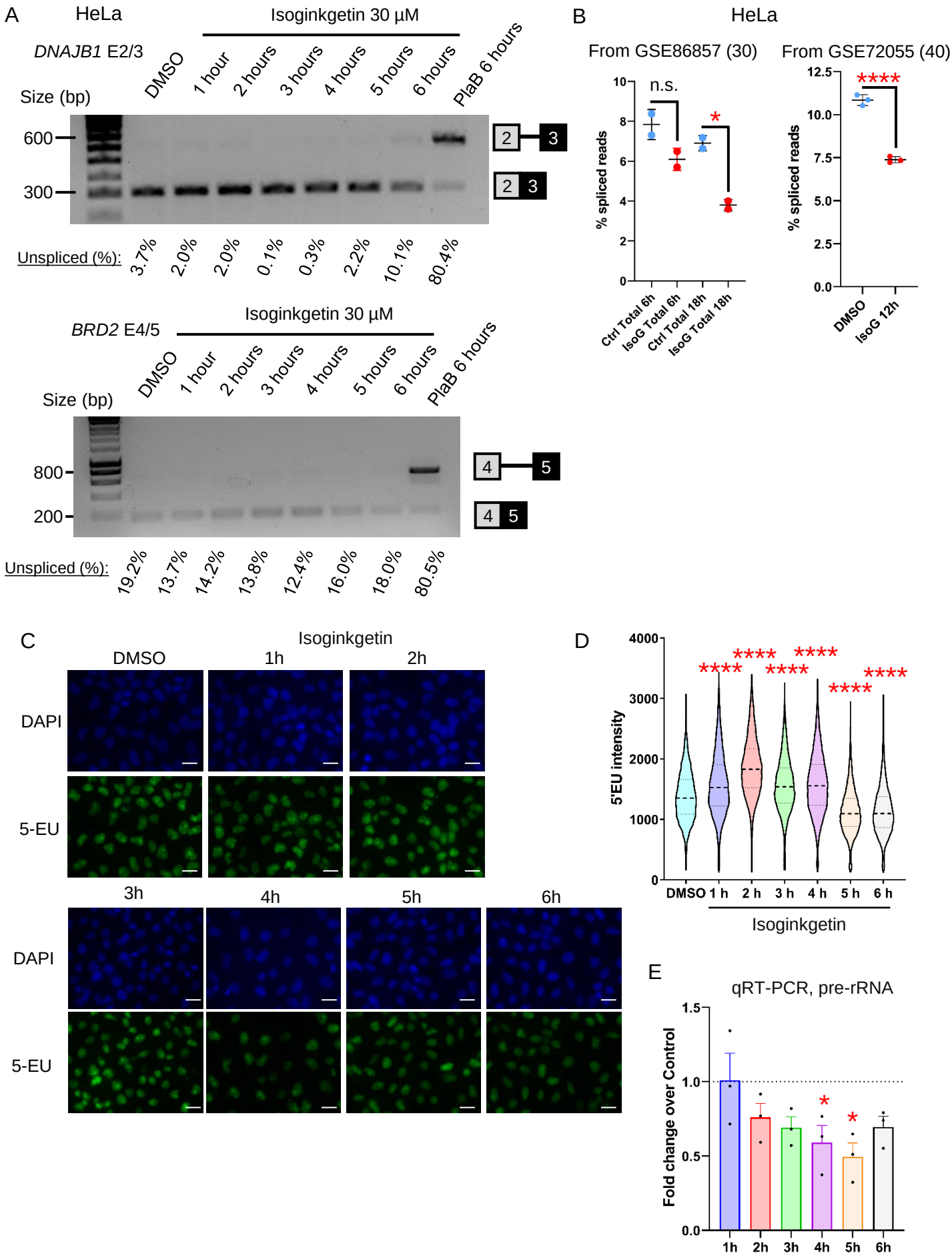

A
HeLa
Isoginkgetin 30 µM
DNAJB1 E2/3
DMSO
1 hour
2 hours
3 hours
4 hours
5 hours
6 hours
PlaB 6 hours
Size (bp)
600
300
2
3
Unspliced (%):
3.7%
2.0%
2.0%
0.1%
0.3%
2.2%
10.1%
80.4%
BRD2 E4/5
800
200
4
5
19.2%
13.7%
14.2%
13.8%
12.4%
16.0%
18.0%
80.5%
B
HeLa
From GSE86857 (30)
From GSE72055 (40)
% spliced reads
n.s.
*
****
Ctrl Total 6h
IsoG Total 6h
Ctrl Total 18h
IsoG Total 18h
DMSO
IsoG 12h
C
Isoginkgetin
DMSO
1h
2h
3h
4h
5h
6h
DAPI
5-EU
D
5'EU intensity
DMSO
1 h
2 h
3 h
4 h
5 h
6 h
Isoginkgetin
E
qRT-PCR, pre-rRNA
Fold change over Control

Fig 4

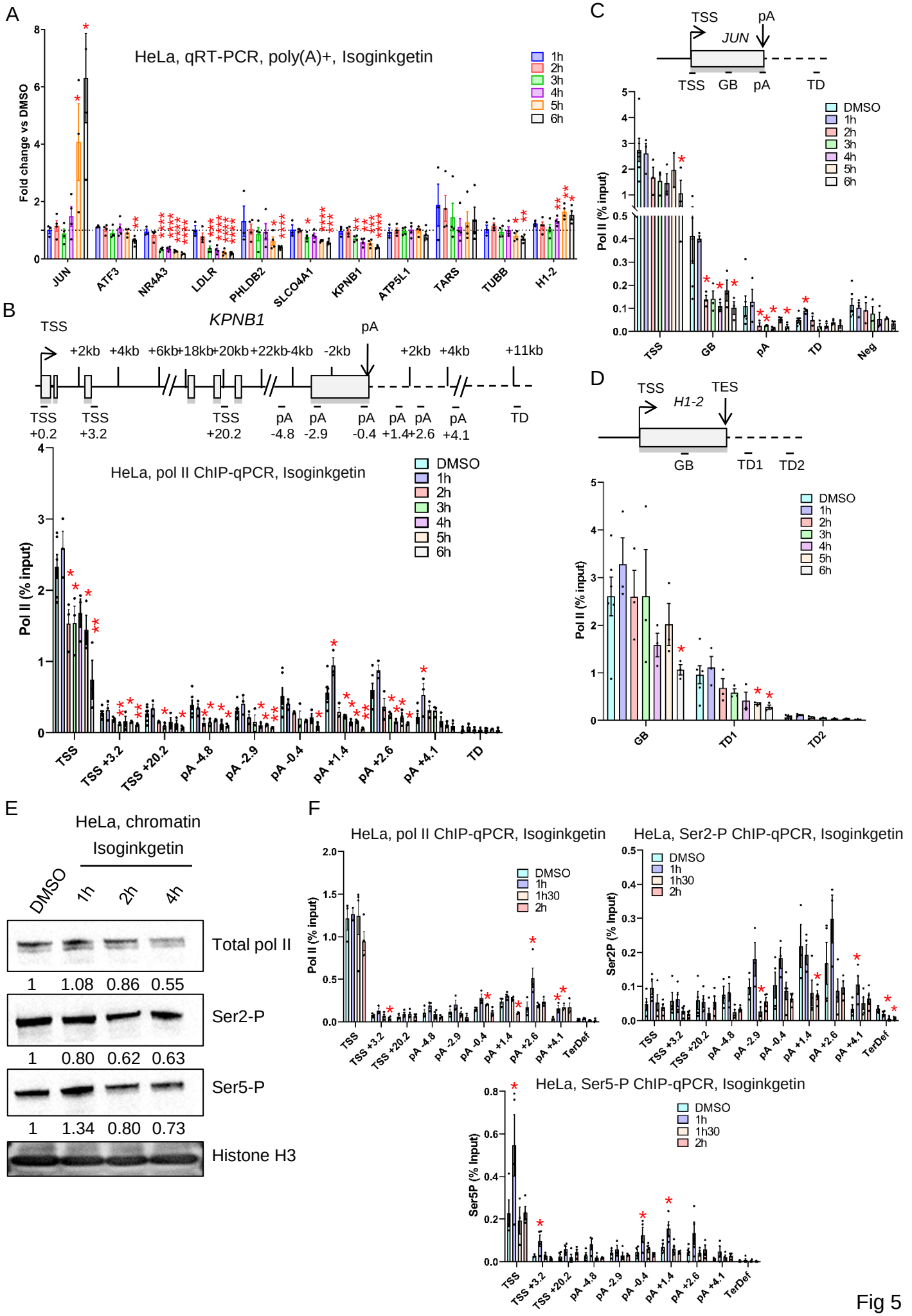

A
HeLa, qRT-PCR, poly(A)+, Isoginkgetin
Fold change vs DMSO
1h
2h
3h
4h
5h
6h
JUN
ATF3
NR4A3
LDLR
PHLDB2
SLCO4A1
KPNB1
ATP5L1
TARS
TUBB
H1-2
B
KPNB1
TSS
pA
+2kb +4kb +6kb +18kb +20kb +22kb -4kb -2kb +2kb +4kb +11kb
TSS +0.2
TSS +3.2
TSS +20.2
pA -4.8
pA -2.9
pA -0.4
pA +1.4
pA +2.6
pA +4.1
TD
HeLa, pol II ChIP-qPCR, Isoginkgetin
Pol II (% input)
DMSO
1h
2h
3h
4h
5h
6h
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
C
TSS
pA
JUN
TSS
GB
pA
TD
Pol II (% input)
DMSO
1h
2h
3h
4h
5h
6h
TSS
GB
pA
TD
Neg
D
TSS
TES
H1-2
GB
TD1
TD2
Pol II (% input)
DMSO
1h
2h
3h
4h
5h
6h
GB
TD1
TD2
E
HeLa, chromatin
Isoginkgetin
DMSO
1h
2h
4h
Total pol II
1 1.08 0.86 0.55
Ser2-P
1 0.80 0.62 0.63
Ser5-P
1 1.34 0.80 0.73
Histone H3
F
HeLa, pol II ChIP-qPCR, Isoginkgetin
Pol II (% input)
DMSO
1h
1h30
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, Ser2-P ChIP-qPCR, Isoginkgetin
Ser2P (% Input)
DMSO
1h
1h30
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, Ser5-P ChIP-qPCR, Isoginkgetin
Ser5P (% Input)
DMSO
1h
1h30
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

Fig 5