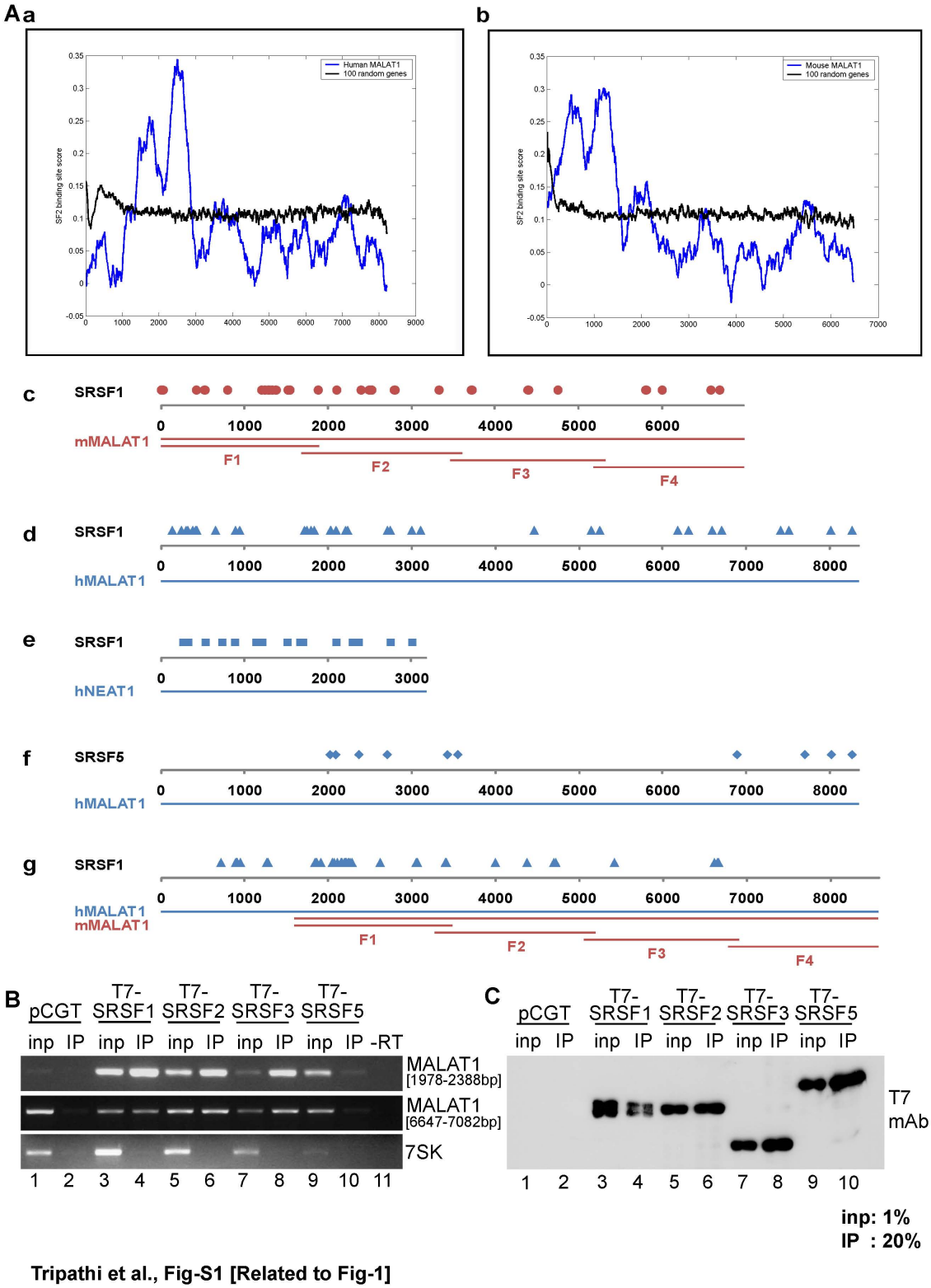
Molecular Cell, Volume *39*

Supplemental Information

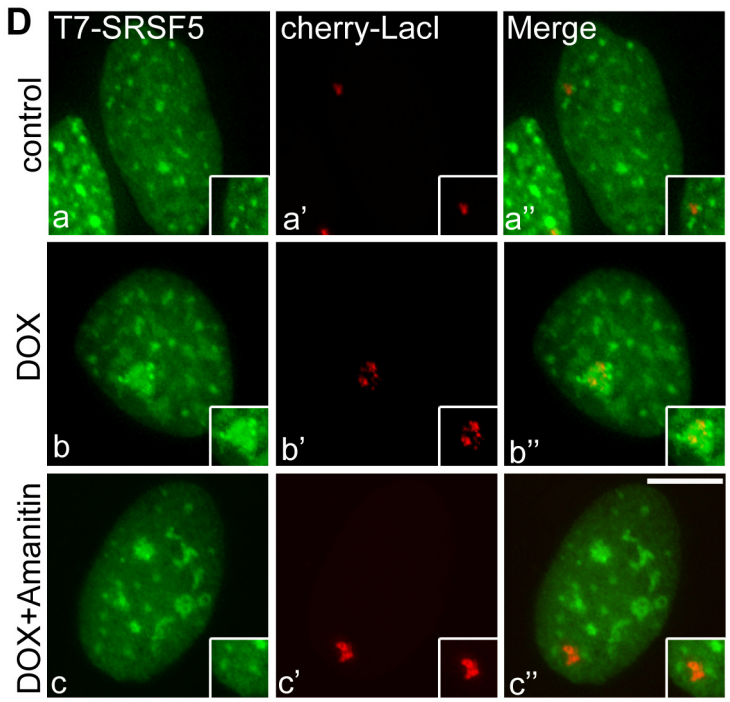
The Nuclear-Retained Noncoding RNA MALAT1 Regulates Alternative Splicing by Modulating SR Splicing Factor Phosphorylation

Vidisha Tripathi, Jonathan D. Ellis, Zhen Shen, David Y. Song, Qun Pan, Andrew T. Watt, Susan M. Freier, C. Frank Bennett, Alok Sharma, Paula A. Bubulya, Benjamin J. Blencowe, Supriya G. Prasanth, and Kannanganattu V. Prasanth

Figure S1. MALAT1 interacts with a specific set of SR splicing factors
(Aa-b) Detection of RNAcompete SRSF1 binding sites in MALAT1 RNA versus randomly selected mRNAs. The SRSF1 binding sites (from 7-mers derived from RNAcompete analysis) is compared between human (a) or mouse (b) MALAT1 (blue) and 100 randomly selected mRNA sequences (black) within 500-nucleotide sliding windows. The statistical significance of the difference in overall scored SRSF1 binding sites between human or mouse MALAT1 and randomly selected mRNAs was calculated using Wilcoxon Signed Rank Test (p-value = 1.0674e-217 [human] and p = 6.6882e-062 [mouse]). (Ac-e) Diagram showing the potential binding sites of SRSF1 in mouse MALAT1 (Ac; ~28 binding sites) and human MALAT1 (Ad; ~34 sites), NEAT1 (Ae; ~14 sites) and SRSF5 in hMALAT1 (Af; ~10 sites), analyzed using ESE finder 3 program (Smith et al., 2006). The stringency threshold values used for SRSF1 and SRSF5 were 3.9 and 5.34 respectively, whereas the suggested threshold values suggested by ESE finder 3 were 1.95 (SRSF1) and 2.67 (SRSF5). (Ag) Diagram showing SRSF1 binding sites in hMALAT1, obtained by SRSF1 CLIP-Seq analysis (~49 sites) (Sanford et al., 2009) (data kindly provided by Dr. J. Sanford, UCSC, USA). Note that the 3'end region (from ~6.7kb to 8.5kb) of human MALAT1 contains minimum SRSF1 binding sites and this region shows high sequence similarity with F4 region of mouse MALAT1 (~80%).
(B) RT-PCR from T7-antibody immunoprecipitated samples (lanes 2, 4, 6, 8 & 10) from cells expressing T7-pCGT vector alone (lanes 1 & 2), T7-SRSF1 (lanes 3 & 4), T7-SRSF2 (lanes 5 & 6), T7-SRSF3 (lanes 7 & 8) and T7-SRSF5 (lanes 9 & 10) using 2 sets of primers against MALAT1 revealed interaction of MALAT1 with SRSF1 (lane 4), SRSF2 (lane 6) and SRSF3 (lane 8) but not with SRSF5 (lane 10). RT-PCR using 7SK primer set was used as negative control.
(C) Immunoblot analysis using T7-mAb from T7-mAb immunoprecipitated samples of cells expressing T7-tag alone (pCGT), T7-SRSF1, T7-SRSF2, T7-SRSF3 and T7-SRSF5. 'Inp' denotes input (1%) and 'IP' denotes immunoprecipitated sample (20%).
(D) Immunolocalization using T7 antibody (a,b,c) in control (a-a"), DOX treated (b-b") and DOX + α-amanitin treated (c-c") U2OS 2-6-3 CLTon cells showing the distribution of transiently expressed T7-SRSF5 (a-c) at the reporter gene locus (stained by Cherry-LacI; red; a',b'c') only when the locus was transcriptionally active (b-b"). Note that the T7-SRSF5 localized to nuclear speckles in all the cells. The inset shows the magnified locus region.
(E) RNA-IP using T7 antibody followed by qPCR in DOX-treated U2OS 2-6-3 CLTon cell extracts transiently expressing T7-SRSF5 showing the interaction of T7-SRSF5 with reporter RNA (CFP). NEAT1 RNA was used as negative control for qPCR analysis and T7-PSP1 IP was used as negative control to estimate the level of non-specific interaction of a T7-tagged protein with reporter CFP RNA. The error bar represents mean ± SD of three independent experiments.



Tripathi et al., Fig-S1 [Related to Fig-1]


D
T7-SRSF5
cherry-LacI
Merge
control
DOX
DOX+Amanitin
a
a'
a''
b
b'
b''
c
c'
c''

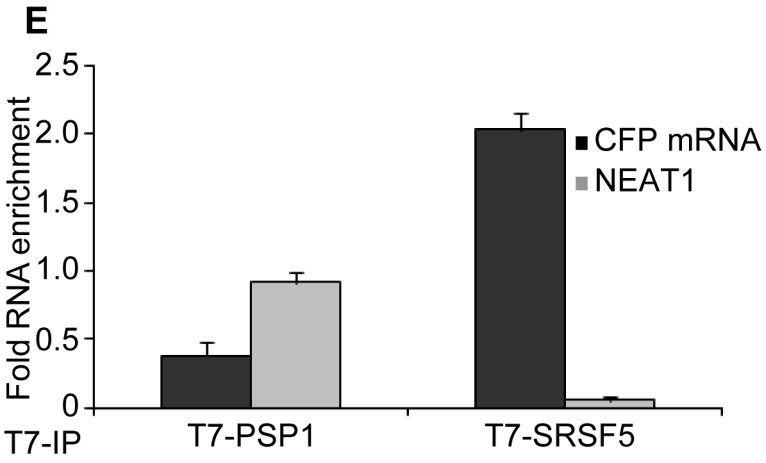

E
Fold RNA enrichment
2.5
2.0
1.5
1.0
0.5
0
CFP mRNA
NEAT1
T7-IP
T7-PSP1
T7-SRSF5

Figure S2. Nuclear distribution of exogenously expressed mMALAT1 in HeLa cells is not influenced by endogenous hMALAT1

(A) RNA-IP using SRSF1 antibody followed by qPCR in HeLa cell extracts expressing mMALAT1 mutants (F1, F2, F3 & F4) revealed specific interaction between SRSF1 and RNAs from F1, F2 and F3 region. SRSF1 showed weak interaction with NEAT1 RNA. The error bar represents mean ± SD.

(B) HeLa cells were transiently transfected with mMALAT1 full-length cDNA (a-a" & b-b") and mutant F1 construct (c-c" & d-d"). Co-RNA FISH using mouse (a-d) and human (a'-d') specific MALAT1 probes revealed that mMALAT1 full length (a & b) and F1-mutant RNA (c & d) showed similar localization in control (a-a" & c-c") and hMALAT1-depleted (b-b" & d-d") cells (compare b with a & d with c). DNA is counterstained with DAPI (blue). The bar represents 10 µm.

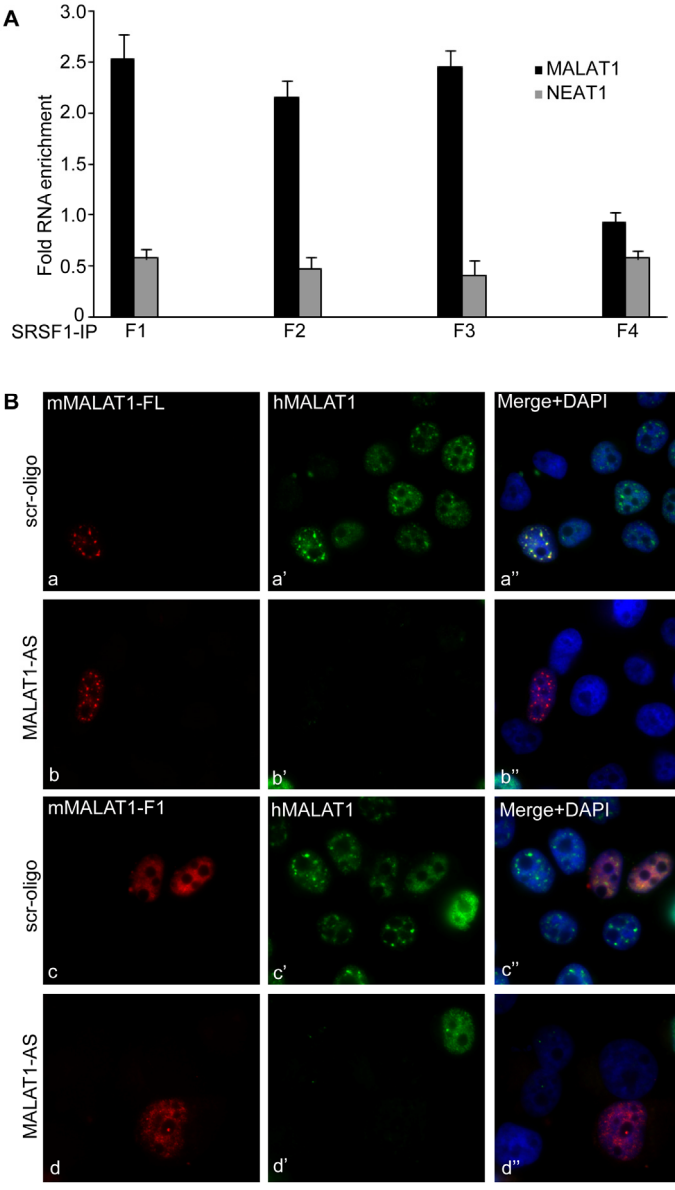


Figure S3. PRP6 and SON splicing factors influence the distribution of MALAT1 to nuclear speckles

(A) Immunoblot of whole-cell extract from HeLa cells transfected (two times at an interval of 24 hrs) with SRSF1 siRNA duplexes or control luciferase siRNA (GL3si) using SRSF1 antibody revealed ~90% knock down of SRSF1 in SRSF1 siRNA treated cells. MEK immunoblot was used as a loading control.

(B) Immunoblot of whole-cell extract from cells transfected (two times at an interval of 24 hrs) with two independent PRP6 siRNA duplexes (PRP6si1 & PRP6si2) or control luciferase siRNA (GL3si) and harvested after 24 hrs post siRNA treatment. Efficacy of RNAi was assessed by immunoblotting to detect endogenous PRP6 (Makarov et al., 2000).

(C) MALAT1 RNA-FISH (a-c) in control (a-a") and two independent PRP6 siRNA-treated (b-b" & c-c") YFP-SRSF1 expressing HeLa cells (a', b' & c') revealed homogeneous distribution of MALAT1 upon PRP6 depletion (compare b & c with a). The bar represents 10 μm.

(D) RNA-IP using T7 antibody followed by qPCR in HeLa cell extracts expressing either T7-PRP6 alone or co-expressing T7-PRP6 with mMALAT1 mutants (F1, F2, F3 & F4) revealed increased levels of interaction of T7-PRP6 with endogenous hMALAT1 or with F1 and F3 mutant RNAs. T7-PRP6 also interacted F2 and F4 RNA, but to a lesser extent than F1 and F3. The error bar represents mean ± SD of three independent experiments.

(E) YFP-SRSF1 (green; a') stably expressing HeLa cells were transfected with an independent siRNA duplex against SON (SONsi1). MALAT1 showed homogeneous nuclear distribution (red; a) in SON-depleted HeLa cells. DNA is counterstained with DAPI (blue; a"). The bar represents 10 μm.

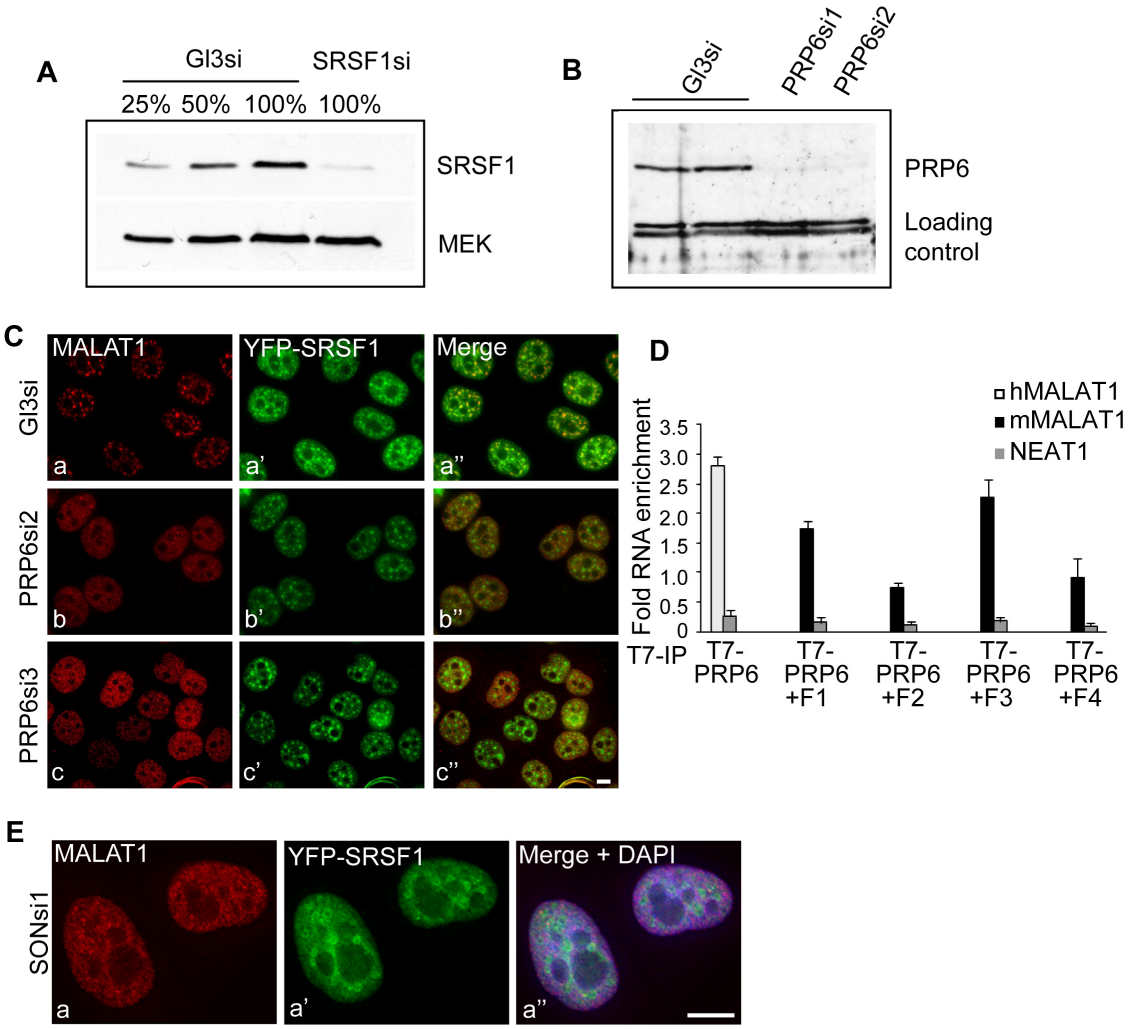

A
Gl3si
SRSF1si
25% 50% 100% 100%
SRSF1
MEK
B
Gl3si
PRP6si1
PRP6si2
PRP6
Loading control
C
MALAT1
YFP-SRSF1
Merge
Gl3si
PRP6si2
PRP6si3
a
a'
a''
b
b'
b''
c
c'
c''
D
hMALAT1
mMALAT1
NEAT1
Fold RNA enrichment
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
T7-IP
T7-PRP6
T7-PRP6 +F1
T7-PRP6 +F2
T7-PRP6 +F3
T7-PRP6 +F4
E
MALAT1
YFP-SRSF1
Merge + DAPI
SONsi1
a
a'
a''

Figure S4. MALAT1 RNA influences the localization of splicing factors in nuclear speckles

(A) qPCR using MALAT1 specific primers in scr-oligo or MALAT1 specific antisense oligo-transfected (MALAT1-AS1 and AS2) HeLa cells showing significant levels of MALAT1 knock down in MALAT1 oligo-treated cells. The error bars represent mean and SD of three independent experiments.

(B) U2AF-65 immunostaining in scr-oligo (a-a') or MALAT1 antisense oligo-treated (b-b') HeLa cells revealed reduced association of endogenous U2AF-65 to nuclear speckles in MALAT1-depleted cells. Bar represents 10 µm. DNA is counterstained with DAPI.

(C) B"-U2 snRNP immunostaining in MALAT1 antisense oligo-treated HeLa cells expressing YFP-SRSF1 revealed reduced association of B"-U2 snRNP to nuclear speckles. In 40% of MALAT1-depleted cells, SRSF1 continue to localize to nuclear speckles (a) (n = 500). Bar represents 10 µm.

(D) MALAT1 RNA-FISH (a-b) in scr-oligo (a-a') or MALAT1 antisense oligo-treated (b-b') HeLa cells expressing GFP-UAP56 (a'-b') revealed localization of GFP-UAP56 to nuclear speckles in MALAT1-depleted cells. Bar represents 10 µm.

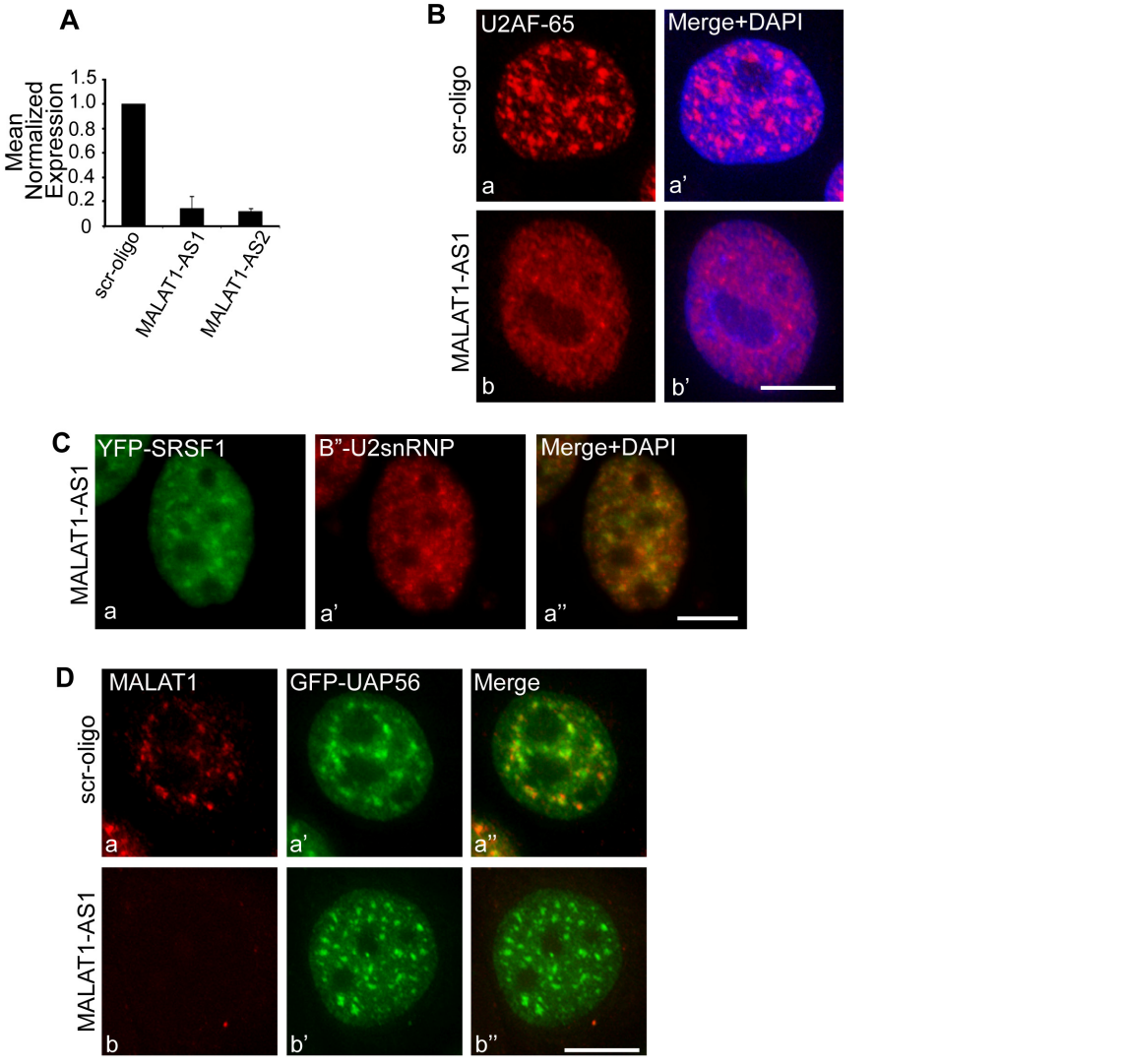


Figure S5. MALAT1 depletion results in mitotic delay followed by nuclear break down

Time-lapse microscopy was performed in control (scr-oligo) and MALAT1-AS2 antisense oligonucleotides-treated HeLa cells that stably express YFP-H2B. Live cell imaging was initiated 5 hrs after the antisense oligo transfection. The control oligo-treated cells completed division within 90 mins. The MALAT1-AS2-treated cells were stuck at prometaphase for prolonged duration (more than 180 min) and then divided to produce a multinucleated cell (see also movies 1 [scr-oligo] & 2 [MALAT1-AS2]).

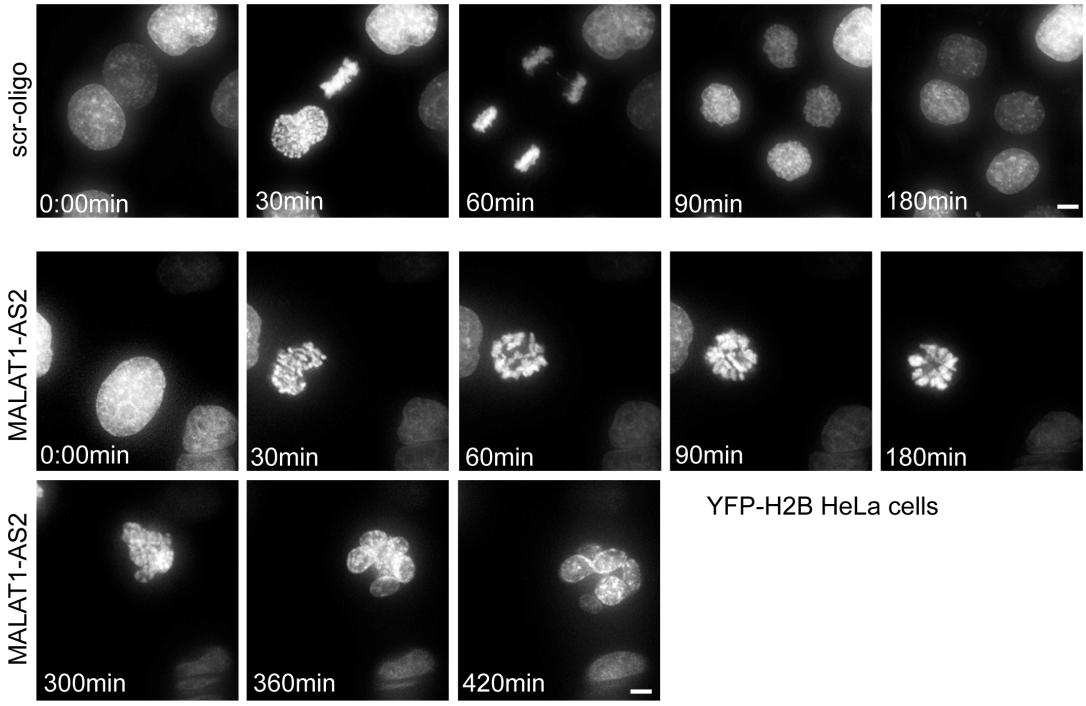


Figure S6A. MALAT1-depletion affects cellular levels of phosphorylated SRSF1

Control (scr-oligo; lanes 1-4) and MALAT1 antisense oligonucleotides (MALAT1-AS2; lanes 5-8) transfected HeLa cells were extracted using 300 mM NaCl salt into soluble supernatant (s; lanes 1, 3, 5 & 7) and insoluble chromatin pellet (p; lanes 2, 4, 6 & 8) fractions. Further, the supernatant and pellet fractions were incubated in presence (s; lanes 3 & 7, p; 4 & 8) or absence (s; 1 & 5, p; 2 & 6) of Antarctic phosphatase (AP). Immunoblot analysis using SRSF1 antibody revealed that the fast migrating form of SRSF1 present in MALAT1 depleted pellet fraction (low mol. wt. band; lane 6) showed similar mobility in presence or absence of AP treatment (compare lane 6 with 8) and the mobility was comparable to phosphatase treated SRSF1 from total extracts (lane 9). Orc2 was used as a loading control.

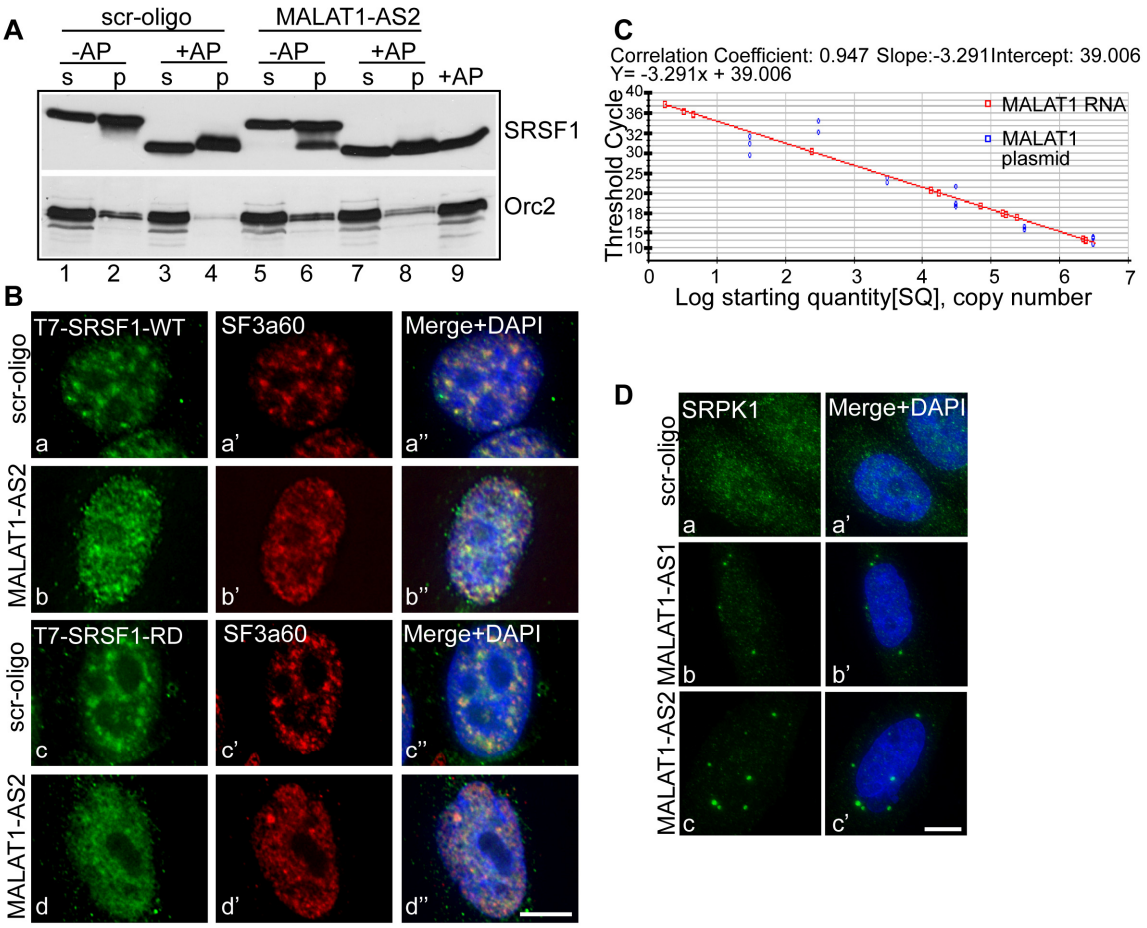
Figure S6B. MALAT1 depletion affects the nuclear speckle localization of WT and phosphomimetic mutant of SRSF1

Immunolocalization using T7 (a,b,c,d) and SF3a60 antibodies (a',b',c',d') in CSK pre-extracted (required for efficient SF3a60 antibody staining) and formaldehyde (2%) fixed control (scr-oligo; a-a" & c-c") and MALAT1-depleted (MALAT1-AS; b-b" & d-d") HeLa cells expressing either T7-SRSF1-WT (a-b) or T7-SRSF1-RD phosphomimetic

mutant (c-d) revealed that a large population of SRSF1-WT (compare b with a; 62% [31 out 50 T7 +ve cells]) and SRSF1-RD (compare d with c; 68% [34 out of 50 T7 +ve cells]) was mislocalized from nuclear speckles and showed more homogenous nuclear distribution upon MALAT1 depletion. Note that a population of SRSF1-RD expressing cells (30% [15 out of 50 T7 +ve cells]) showed homogenous nuclear distribution even in scr-oligo treated cells. DNA is counterstained with DAPI. The bar represents 10 μm.

Figure S6C MALAT1 is an abundant nuclear ncRNA
The RNA copy number (number of RNA molecules/cell) of MALAT1 in HeLa cells was estimated using qPCR analysis. Standard curve representing the amplification efficiencies of the ten-fold serial dilutions of the pGEM-Teasy hMALAT1 template. Real-time PCR amplifications were performed using the MALAT1 specific primers. The calculated Ct values were plotted against the log of the initial amount of pGEM-Teasy hMALAT1 molecules (10^{10}-10^{16}) to generate the standard curve. Squares and circles represent MALAT1 standards (pGEM-Teasy-hMALAT1) and MALAT1 RNA from cells respectively.

Figure S6D. MALAT1-depleted cells show aberrant localization of SR protein kinase, SRPK1
Immunolocalization using SRPK1 antibody in scr-oligo transfected HeLa cells showed homogenous cytoplasmic and nuclear staining of SRPK1 (a-a'). However, MALAT1 depletion using two independent antisense oligonucleotides (b & c) resulted in the redistribution of SRPK1 into large cytoplasmic aggregates. DNA is counterstained using DAPI. The bar represents 10 μm.


A
scr-oligo
MALAT1-AS2
-AP
+AP
-AP
+AP
s p s p s p s p +AP
SRSF1
Orc2
1 2 3 4 5 6 7 8 9
B
T7-SRSF1-WT
SF3a60
Merge+DAPI
scr-oligo
a
a'
a''
MALAT1-AS2
b
b'
b''
T7-SRSF1-RD
SF3a60
Merge+DAPI
scr-oligo
c
c'
c''
MALAT1-AS2
d
d'
d''
C
Correlation Coefficient: 0.947 Slope:-3.291Intercept: 39.006
Y= -3.291x + 39.006
Threshold Cycle
MALAT1 RNA
MALAT1 plasmid
Log starting quantity[SQ], copy number
D
SRPK1
Merge+DAPI
scr-oligo
a
a'
MALAT1-AS1
b
b'
MALAT1-AS2
c
c'

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cell Culture

HeLa, U2OS and EpH4 cells were grown in DMEM containing high glucose, supplemented with penicillin-streptomycin and 10% fetal bovine serum (FBS) (Hyclone, Logan, UT).

Plasmid constructs and DNA transfection

Mouse MALAT1 cDNA (6982bp) was PCR amplified and cloned into pSV40-EYFP-N1 vector (BD Biosciences, Clonetech, CA) by replacing the YFP cDNA using Apa1-Not1 enzymes: pSV40-mMALAT1. Using primers from overlapping regions of the cDNA, four mutant constructs of mMALAT1 were further cloned in the same vector: pSV40-mMalat1-F1 (1-1892bp), F2 (1677-3600bp), F3 (3476-5331bp) and F4 (5185-6982bp). The T7-tagged plasmids used in the RNA-IP experiments include pCGT vector, pCGT-SRSF1, pCGT-SRSF1ΔRRM1, pCGT-SRSF1ΔRRM2, pCGT-SRSF1ΔRS, pCGT-SRSF2, pCGT-SRSF3, pCGT-SRSF5 (a kind gift from Dr. A. Krainer, CSHL, USA), pCGT-PSP1 and pCGT-PRP6. pCGT-SRSF1-FF-DD was constructed by PCR amplifying the SRSF1-FF-DD from pET19b-SRSF1-FF-DD (kind gifts from Drs. A. Krainer, CSHL USA and J. Caceres, MRC Human Genetics Unit, Edinburgh) and cloned into pCGT vector using Xba1-BamH1 sites. pCGT-SRSF1-RD was provided by (Drs. J. Sanford, University of California at SC, USA and J. Caceres, MRC Human Genetics Unit, Edinburgh. Fluorescent protein-tagged plasmid DNAs used in the present study include pEGFP-SF1, U2AF65, U2AF35 and UAP56 (gift from Dr. M. Carmo-Fonseca, University of Lisbon, Portugal); pECFP-SRSF3 (gift from Dr. J. Caceres, MRC Human

Genetics Unit, Edinburgh); pEYFP-siR-SON, pEYFP-siR-SON-1-2008 (Sharma et al., 2010), pEYFP-SRSF1 and SRSF2. Plasmid DNAs were individually transfected (2µg) into cells using lipofectamine 2000 (Invitrogen, Carlsbad, CA). After transfection, cells were seeded onto coverslips and processed for antisense treatment, immunofluorescence localization or RNA-FISH 24 hrs post transfection.

Antibodies

Antibodies used in the present study include, B" (U2-snRNP, mIgG, IF, 1: 100), SF3a60 (rIgG, IF, 1: 200, WB 1:2000) (Kramer et al., 1994), SRSF1 (mIgG, mAb103, IF, 1: 300) (Hanamura et al., 1998), SRSF1 (mIgG, mAb96, WB, 1: 1000) (Hanamura et al., 1998), SRSF2 (mIgG, WB, 1:20) (Cavaloc et al., 1999), 3C5 (mIgM, IF, 1: 100; WB, 1:100) (Turner and Franchi, 1987), mAB104 (mIgM, WB, 1:4) (Roth et al., 1990), UAP56 (rIgG, IF, 1:50) (Kota et al., 2008), U2AF-65 (rIgG, IF, 1:100, WB, 1:1000), PRP6 (rIgG, WB, 1: 1000) (Makarov et al., 2000), Orc2 (rIgG, WB, 1: 2000), MEK (rIgG, WB, 1:2000), T7 (mIgG, IF, 1:1000, WB, 1:5000) (Novus Biologicals), SON (rIgG, IF, 1:5000) (Sharma et al., 2010), SRPK1 (rIgG, IF, 1:200).

RNA Fluorescence in situ hybridization (RNA-FISH) and immunofluorescence staining

To detect MALAT1 RNA, cells were rinsed briefly in PBS and then fixed in 4% formaldehyde in PBS (pH 7.4) for 15 min at RT. Further, the cells were permeabilized in PBS containing 0.5% Triton X-100 and 5mM VRC (New England Biolabs Inc., USA) on ice for 10 min; washed with PBS 3x 10min and rinsed once in 2XSSC prior to

hybridization. Hybridization was carried out using nick-translated cDNA probe (Abbott Molecular, USA) in a moist chamber at 37°C for 12–16 hr as described earlier (Prasanth et al., 2005). For colocalization studies, after RNA-FISH, cells were again fixed for 5 min in 2% formaldehyde and immunofluorescence staining were performed as described earlier (Prasanth et al., 2005). *In vivo* Br-UTP incorporation analysis, U2 snRNA and poly A^+ RNA FISH were performed as described (Huang et al., 1994; Sacco-Bubulya and Spector, 2002). Fluorescence mages were acquired using either a DeltaVision RT (Olympus, 60X, 1.42 NA; Applied precision, USA) or Axioimager Z1 (Zeiss, 63X 1.4 NA) microscopes and Z-stacks spanning the entire nuclei were acquired (Delta Vision RT) and the images were processed using either Axiovision (Axioimager) or SoftWoRx (Delta Vision RT) software.

Live cell imaging

HeLa cells stably expressing YFP-H2B were seeded onto glass-bottom containing 35 mm Petri dishes (Mactech, USA) 24 hrs prior to antisense oligo treatment. Cells were transfected two or three times (gap of 24hrs between each transfection) with control or MALAT1 AS oligonucleotides using Lipofectamine RNAi Max reagent. 5 hrs after the final transfection, cells were incubated with L-15 medium (without phenol red) containing 30% FBS and were placed into a 37^{o}C environmental chamber on the stage of a DeltaVision RT imaging system. The nuclei were imaged acquiring Z-stacks (20 stacks, 1 μm thickness) for every 30 min for 12-24 hrs using 60X, 1.42 NA objective lense. The images were processed using SoftWoRx (Delta Vision RT) software.

RT-PCR Assays

Total cellular RNA was extracted from the control, MALAT1 depleted cells and T7-SRSF1 overexpressing cells using Trizol (Invitrogen) and polyA$^+$ RNA was isolated using oligo-dT columns (NucleoTrap Midi kit, Clontech, USA) as per the manufacturer's instructions. One-step RT-PCR was performed using 15ng of poly A$^+$ RNA as template with gene specific primers (5μM) (Sigma-Genosys, USA). The reaction mix was supplemented with α-^{32}P-dCTP (Perkin Elmer, USA). The total reaction product was run onto a 7.5% polyacrylamide gel, dried and exposed to a phosphoimager screen (Storm, Amersham Biosciences). The quantification of the isoform abundance was performed using image analyzer (ImageQuant) and the percentage inclusion of the exon was derived. The percentage exon inclusion represents the difference between the intensities of the two isoforms. The primer sequences are provided at the end of supplementary experimental procedures.

Quantitative real-time PCR (qPCR)

Total cellular RNA was isolated using Trizol (Invitrogen) according to the manufacturer's instructions and reverse transcribed into cDNA using Multscribe Reverse transcriptase and Random Hexamers (Applied Biosystems). qPCRs were performed using the iCycler iQ system (Biorad, CA, USA). Transcript levels were quantitated against a standard curve by Real-Time RT-PCR using the SYBR® Green I fluorogenic dye and data analyzed using the iCycler iQ system software.

Primer sets showing comparably high efficiencies were used for the analyses.

For the RNA-IP experiment, the qPCR results were analyzed using the comparative Ct method (Pfaffl, 2001). The qPCR levels of the IP samples were normalized against the input of the same samples and calibrated to mIgG or pCGT vector control. Fold RNA enrichment represents log of RQ value. Histograms show mean ± SD obtained for three independent experiments. The knock down analyses of MALAT1 RNA was calculated using Q-gene, a Microsoft Excel script package (Muller et al., 2002).

The MALAT1 qPCR values were first normalized against the GAPDH levels of the same samples and then adjusted such that the MALAT1 levels of the control samples (scr-oligo treated) had a value of 1. Histograms represent mean normalized expression (MNE) ± SD.

Cellular fractionation and RNA slot blot

Nuclear and Cytoplasmic fractionation and RNA isolation were described previously (Prasanth et al., 2005). RNA slot blot was conducted using a slot blot apparatus (Minifold apparatus, S & S Inc.) and hybridization using oligonucleotide probes (^{32}P-γ-ATP labeled; Oligo-dT, U6 snRNA and Lysine-4 tRNA probes) were performed as per manufacturers instructions (Ambion Inc. Austin, Texas). The quantification of radioactive signal was performed using Image J software.

***In vivo* reporter cell line system**

The U2OS 2-6-3 cells (Janicki et al., 2004) were transfected with a plasmid containing pVITRO2.Cherry Lac Repressor (LacI) and Tetracycline activator (rTa), and selected with Hygromycin and G418 to generate 2-6-3 CLTon cells (Zhen and Prasanth,

unpublished data). U2OS-2-6-3 CLTon cells were grown in DMEM with 10% Tet system approved FBS (Clontech). Cells grown on coverslips to 50-60% confluency were transfected with 1µg pCGT-SRSF5 using Lipofectamine 2000 reagent. 24 hrs post transfection cells were fixed in 2% (w/v) formaldehyde and used for immunoflurescence analysis (NO DOX). For DOX treatment, 24 hrs post DNA transfection, cells were incubated with DOX (1µg/ml) for 3hrs at 37^{o}C/5% CO_2 incubator. For DOX followed by α-amanitin treatment, α-amanitin was added to the cells that were already treated with DOX and incubated for another 5 hrs at 37^{o}C/5% CO_2 in presence of both the drugs.

Salt fractionation and Immunoblotting

Control and MALAT1 antisense oligo-transfected HeLa cells were scraped in the medium and washed with ice-cold PBS. Extraction was performed in lysis buffer (50 mM Tris; pH 7.2, 0.5% NP-40, 10% glycerol, 1mM DTT) containing 150 mM or 300 mM salt (NaCl) for 30 min at 4°C and the cells were further centrifuged at 10,000g for 10min at 4°C. Loading dye was directly added to the supernatant and pellet and boiled for 5min and finally loaded onto a 15% polyacrylamide gel. All the buffers contained protease and phosphatase inhibitors unless otherwise mentioned.

For the dephosphorylation of extracts, the supernatant and pellet after the salt extraction were directly incubated at 37°C for 30min in presence of 5U of Antarctic Phosphatase (New England Biolabs, USA) along with 1X phosphatase buffer. Further, the samples were boiled after the addition of loading dye and loaded onto gel. Immunoblotting was performed as described (Prasanth et al., 2003).

SRSF1 binding site motif analysis

To determine whether SRSF1 binding sites are more enriched in MALAT1 RNA, we performed an SRSF1 binding site motif analysis using RNAcompete derived binding preferences for this factor (Ray et al., 2009), RNAcompete is a microarray based method for the systematic analysis of RNA binding specificities to determine linear or structured binding preferences of RNA binding proteins. Human and mouse MALAT1 RNA sequences were scored for RNAcompete 7-mer binding site sequences for SRSF1 within a sliding window of 500 nucleotides. The same analysis was performed on 100 randomly selected mRNAs (of at least 1 kb) from both human and mouse. The numbers of sliding windows sampled across the randomly selected mRNAs was normalized to generate equal numbers as MALAT1, and SRSF1 binding site scores were determined as averages across the 100 mRNAs.

MALAT1 RNA copy number analysis

Total RNA was isolated from 4.2 x 10^6 HeLa cells using Trizol (Invitrogen) and the concentration was determined spectrophotometrically using different dilutions of the RNA. cDNA was synthesized using 1µg of the total RNA in a 50 µl reaction.

Serial ten-fold dilutions (10^{10}-10^{16} molecules) of pGEM-Teasy hMALAT1 plasmid were used as a reference molecule for the standard curve calculation. All the qPCR reactions were performed in a 20 µl reaction mix containing 1/5-1/100 volumes of the cDNA preparation. Real-time quantitations were performed using the Bio-rad iCycler iQ system (BioRad, Hercules, CA, USA). The fluorescence threshold value was calculated using the iCycler iQ system software.

Primer sequences

MGEA6 For	CTGAAACAGAGCTTAAATTTGAAC
MGEA6 Rev	CTGGCGGAGGAAACATCATCC
B-MYB For	GGAAGTCTTCTGACCAACTGGC
B-MYB Rev	GCAGCATGTTTCTGGTGCAGGGG
CDK7 For	CCACCGTTTACAAGGCCAGAG
CDK7 Rev	CCAATGTTGATGTAAATATTCTAATCCTTG
HMG2L1 For	CACAAGAAGAAGAGGAAGCACTC
HMG2L1 Rev	ACTGACTTTCTTCCTGTGGTGCT
CAMK2B For	ACAGTGCAGCCGCCACCAG
CAMK2B Rev	TTAGCGTCTTCATCCTCTATGGTG
TLK2 For	GGAAAGGCACTCCTAGGGGA
TLK2 Rev	AGGCGGAGGTCGATAAGCTG
ARHGEF1 For	AACAGCCAGTTCCAGAGCCTG
ARHGEF1 Rev	AAAGGCGACGTTGGGAGGGA
PAX2 For	CCCAATGGAGATTCCCAGAGTG
PAX2 Rev	CCATTCCTGCCAGGGTGGAG
SAT1 For	TGCCTGGTTGCAGAAGTGCC
SAT1 Rev	GGTCATAGGTAAAATAGTACATGGC
MALAT1-Apa1-F1 For	GACTGAGGGCCCCAGGCATTCAGGCAGCGA
MALAT1-Not1-F1 Rev	GACTGAGCGGCCGCTCTGGGCCTTGGCAGTCA
MALAT1- Apa1-F2 For	GACTGAGGGCCCCTTGAGGTCGGCAACATGG
MALAT1- Not1-F2 Rev	GACTGAGCGGCCGCATCAATTCTTTACTTTTG
MALAT1- Apa1-F3 For	GACTGAGGGCCCAAAATAATGTCATGTCTC
MALAT1- Not1-F3 Rev	GACTGAGCGGCCGCCAAGAATGTTGCTTGTCT
MALAT1- Apa1-F4 For	GACTGAGGGCCCTAACAGGCTGAGTGTTGA
MALAT1- Not1-F4 Rev	GACTGAGCGGCCGCGATATTTAGTTTTTATTTC
MALAT1- 1978-For	GAAGGAGCGCTAACGATTTG
MALAT1-2388-Rev	TCTCCTGGACTTGGCAGTCT
MALAT1- 6647-For	GGCAGGAGAGACAACAAAGC
MALAT1-7082-Rev	CCTCGACACCATCGTTACCT
7SK For	GACATCTGTCACCCCATTGA
7SK rev	GCCTCATTTGGATGTGTCTG
GAPDH For	TCACCAGGGCTGCTTTTAAC
GAPDH rev	TTCTAGACGGCAGGTCAGGT
hNEAT1 For	TCGGGTATGCTGTTGTGAAA
hNEAT1 Rev	ACCTTCATTGGCTTGAGTGG
qPCR Primers sequences:	
hMALAT1 For	GACGGAGGTTGAGATGAAGC
hMALAT1 Rev	ATTCGGGGCTCTGTAGTCCT
mMalat1F1R1 For	CGTTTGAAGGCATGAGTTGG
mMala1F1R1 Rev	TGCCTCCCAAGTGCTAGGAT
mMalat1F2R2 For	AGCTTTTGAGGGCTGACTGC

mMalat1F2R2 Rev	CCATTCATTCCCCTCTGAGC
mMalat1F3R3 For	CTTTTCCCCCACATTTCCAA
mMalat1F3R3 Rev	CTCGTGGCTCAAGTGAGGTG
mMalat1F4R4 For	GGTTACCAGCCCAAACCTCA
mMalat1F4R4 Rev	TTCTGGAAAAGCTGGGGAAA
hNEAT1 For	TCGGGTATGCTGTTGTGAAA
hNEAT1 Rev	TGACGTAACAGAATTAGTTCTTACCA
ACTIN-B For	TGCGTTACACCCTTTCTTGA
ACTIN-B Rev	AAAGCCATGCCAATCTCATC
GAPDH For	GATCATCAGCAATGCCTCCT
GAPDH Rev	TGTGGTCATGAGTCCTTCCA
CFP For	CGACAACCACTACCTGAGCA
CFP Rev	GAACTCCAGCAGGACCATGT

SUPPLEMENTAL REFERENCES

Cavaloc, Y., Bourgeois, C.F., Kister, L., and Stevenin, J. (1999). The splicing factors 9G8 and SRSF3 transactivate splicing through different and specific enhancers. RNA (New York, N.Y *5*, 468-483.

Hanamura, A., Caceres, J.F., Mayeda, A., Franza, B.R., Jr., and Krainer, A.R. (1998). Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors. RNA (New York, N.Y *4*, 430-444.

Huang, S., Deerinck, T.J., Ellisman, M.H., and Spector, D.L. (1994). In vivo analysis of the stability and transport of nuclear poly(A)+ RNA. The Journal of cell biology *126*, 877-899.

Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R.H., and Spector, D.L. (2004). From silencing to gene expression: real-time analysis in single cells. Cell *116*, 683-698.

Kota, K.P., Wagner, S.R., Huerta, E., Underwood, J.M., and Nickerson, J.A. (2008). Binding of ATP to UAP56 is necessary for mRNA export. Journal of cell science *121*, 1526-1537.

Kramer, A., Legrain, P., Mulhauser, F., Groning, K., Brosi, R., and Bilbe, G. (1994). Splicing factor SF3a60 is the mammalian homologue of PRP9 of S.cerevisiae: the conserved zinc finger-like motif is functionally exchangeable in vivo. Nucleic acids research *22*, 5223-5228.

Makarov, E.M., Makarova, O.V., Achsel, T., and Luhrmann, R. (2000). The human homologue of the yeast splicing factor prp6p contains multiple TPR elements and is stably associated with the U5 snRNP via protein-protein interactions. J Mol Biol *298*, 567-575.

Muller, P.Y., Janovjak, H., Miserez, A.R., and Dobbie, Z. (2002). Processing of gene expression data generated by quantitative real-time RT-PCR. BioTechniques *32*, 1372-1374, 1376, 1378-1379.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic acids research *29*, e45.
Prasanth, K.V., Prasanth, S.G., Xuan, Z., Hearn, S., Freier, S.M., Bennett, C.F., Zhang, M.Q., and Spector, D.L. (2005). Regulating gene expression through RNA nuclear retention. Cell *123*, 249-263.
Prasanth, K.V., Sacco-Bubulya, P.A., Prasanth, S.G., and Spector, D.L. (2003). Sequential entry of components of the gene expression machinery into daughter nuclei. Molecular biology of the cell *14*, 1043-1057.
Ray, D., Kazan, H., Chan, E.T., Pena Castillo, L., Chaudhry, S., Talukder, S., Blencowe, B.J., Morris, Q., and Hughes, T.R. (2009). Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. Nature biotechnology *27*, 667-670.
Roth, M.B., Murphy, C., and Gall, J.G. (1990). A monoclonal antibody that recognizes a phosphorylated epitope stains lampbrush chromosome loops and small granules in the amphibian germinal vesicle. The Journal of cell biology *111*, 2217-2223.
Sacco-Bubulya, P., and Spector, D.L. (2002). Disassembly of interchromatin granule clusters alters the coordination of transcription and pre-mRNA splicing. The Journal of cell biology *156*, 425-436.
Sanford, J.R., Wang, X., Mort, M., Vanduyn, N., Cooper, D.N., Mooney, S.D., Edenberg, H.J., and Liu, Y. (2009). Splicing factor SFRS1 recognizes a functionally diverse landscape of RNA transcripts. Genome research *19*, 381-394.
Sharma, A., Takata, H., Shibahara, K.I., Bubulya, A., and Bubulya, P.A. (2010). Son Is Essential for Nuclear Speckle Organization and Cell Cycle Progression. Molecular biology of the cell.
Smith, P.J., Zhang, C., Wang, J., Chew, S.L., Zhang, M.Q., and Krainer, A.R. (2006). An increased specificity score matrix for the prediction of SRSF1-specific exonic splicing enhancers. Human molecular genetics *15*, 2490-2508.
Turner, B.M., and Franchi, L. (1987). Identification of protein antigens associated with the nuclear matrix and with clusters of interchromatin granules in both interphase and mitotic cells. Journal of cell science *87 (Pt 2)*, 269-282.