Supplemental Material

DDX54 regulates transcriptome dynamics during DNA damage response

Miha Milek[1], Koshi Imami[1], Neelanjan Mukherjee[1], Francesca De Bortoli[2], Ulrike Zinnall[1], Orsalia Hazapis[1], Christian Trahan[3,5], Marlene Oeffinger[3,4,5], Florian Heyd[2], Uwe Ohler[1,6,7], Matthias Selbach[1,8] and Markus Landthaler[1,9,*]

[1]**Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin Institute for Medical Systems Biology, Germany.**

[2]**Department of Biology, Chemistry, Pharmacy, Freie Universität Berlin, Institute of Chemistry and Biochemistry, Laboratory of RNA Biochemistry, Berlin, Germany.**

[3]**Institut de Recherches Cliniques de Montréal, Montréal, Canada.**

[4]**Faculty of Medicine, Division of Experimental Medicine, McGill University, Montréal, Canada.**

[5]**Département de Biochimie, Faculté de Médecine, Université de Montréal, Montréal, Canada.**

[6]**Department of Biology, Humboldt-Universität zu Berlin, Germany.**

[7]**Department of Computer Science, Humboldt-Universität zu Berlin, Germany.**

[8]**Charite-Universitätsmedizin Berlin, Germany.**

[9]**IRI Life Sciences, Institute of Biology, Humboldt-Universität zu Berlin, Germany**

***Corresponding author:** Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin Institute for Medical Systems Biology, Robert-Rössle-Str. 10, 13125 Berlin, Germany. Phone: +49 30 9406-3026. Fax: +49 30 9406-49160.

Email: markus.landthaler@mdc-berlin.de

TABLE OF CONTENTS

SUPPLEMENTAL METHODS

Cell culture, plasmids and treatments

HEK293 Flp-In T-REx cell line was from purchased from Thermo Fisher Scientific and MCF-7 from ATCC (HTB-22). MCF-7 Flp-In cell line (a kind gift from R. Agami's laboratory, (Slobodin et al. 2017)) was generated by transfection of MCF-7 cells with a *Sca*I-linearized pFRT/lacZeo plasmid (Thermo Fisher), followed by zeocin selection (500 µg/ml) and monoclonal expansion. Unless stated otherwise, all cell lines were cultured in high glucose Dulbeco's modified Eagle's medium (DMEM, Thermo Fisher Scientific), supplemented with 10% (v/v) fetal bovine serum, 2 mM L-glutamine, 100 U/ml penicillin/streptomycin (all from Thermo Fisher Scientific).

Vector pDONR223 carrying the DDX54 coding sequence corresponding to isoform 2 (NM_024072.3; NP_076977.3) was obtained from the hORFeome V5.1 collection. DDX54 coding sequence was recombined into pFRT/TO/FLAG/HA-DEST (Addgene ID: 26360), pFRT/TO/RFP and pDEST5-BirA-FLAG-N-term-pcDNA5-FRT-TO (Couzens et al. 2013) using LR Clonase II (Thermo Fischer Scientific). Vector encoding C-terminal FLAG/HA-tagged DDX54 was obtained by sub-cloning via *Kpn*I and *Bam*HI sites (see Oligonucleotides) into pcDNA5/FRT/TO (Thermo Fisher Scientific). All plasmids were submitted to Addgene.

To generate stable MCF-7 Flp-In cell lines overexpressing FLAG/HA-DDX54 or RFP-DDX54; and HEK293 Flp-In T-REx overexpressing BirA-FLAG/DDX54, cells were co-transfected with the vectors in combination with the pOG44 vector (Thermo Fisher). Briefly, MCF-7 Flp-In cells were co-transfected in a 6-well format by mixing 150 µl of Opti-MEM with 3 µg plasmid DNA (5:1 ratio of pOG44 to destination vector) and 9 µl Fugene HD (Promega). After a 10-min incubation, the transfection mixture was added to the cells. Cells were re-plated into 10 cm dishes after 48 hours and allowed to attach overnight. Hygromycin (50 µg/ml, InvivoGen) was added the next day and cells were selected for 2-3 weeks resulting in expansion of monoclonal colonies. Generation of HEK293 Flp-In T-REx stable cell lines was performed as previously described (Gregersen et al. 2014).

Isolation of poly(A)$^+$ RNA bound proteins

For these experiments, we used both label-free MCF-7 cells growing in the standard DMEM formulation, as well as SILAC “heavy” labeled HEK293 cells. For SILAC "heavy" labeling, HEK293 Flp-In T-REx cells were grown in high glucose SILAC DMEM (PAA, E15-086) supplemented with 10% (v/v) dialyzed fetal bovine serum (Sigma Aldrich), 2 mM L-glutamine, 100 U/ml penicillin/streptomycin (both from Thermo Fischer Scientific), 0.398 mM $^{13}C_6$,$^{15}N_2$ L-arginine and 0.798 mM $^{13}C_6$,$^{15}N_2$ L-lysine (Cambridge Isotope Laboratories) for at least 7 passages. We then treated the cells (40 15-cm dishes) with 200 µM 4SU (ChemGenes) for 16 h and crosslinked them by exposing them to 0.2 J/cm^2 UV light at 365nm (Stratalinker). Cells were immediately washed in PBS, scraped-off, pelleted, flash frozen in liquid nitrogen and stored at -80 ºC until further processing.

For large scale oligo(dT) affinity purifications we used 60 15-cm culture dishes per one replicate of label-free MCF-7 cells. Cells were then treated with 200 µM 4SU for 16h, IR-exposed (10 Gy by a cesium-137 γ-ray source) or left untreated. One hour after IR exposure, all cell culture plates were irradiated with UV light at 365 nm (0.2 J/cm^2) on ice and harvested immediately. To ensure similar harvesting times after exposure to ionizing radiation (IR), not more than 20 15-cm dishes were harvested at a time.

We next lysed the cell pellets obtained from 60 plates in 60 ml of lysis/binding buffer (1% lithium dodecylsulphate, 100 mM Tris[pH 7.5], 10 mM EDTA, 500 mM LiCl, 5 mM DTT, Complete Mini EDTA-free protease inhibitor (Roche)), and passed them 4 times through a 23G needle, followed by 2 passes through a 26G needle. SILAC "heavy" cells were lysed in the same ratio (25 plates of cells in 25 ml of lysis/binding buffer) and processed as above. An equal volume (12 ml, ratio 1:5) of the SILAC "heavy" lysate was mixed with both control and IR-exposed label-free lysates. Oligo(dT) affinity purifications were then performed with 200 µl bed volume (equal to 15 ml original bead volume) of magnetic oligo(dT) Dynabeads (Ambion, Thermo Fischer Scientific) for 1 h at room temperature. Beads were then captured on an in-house-produced 50-ml magnets and supernatants were saved for 2 additional rounds of oligo(dT) affinity purification. Beads were washed twice with 1 lysate volume of lysis/binding buffer, followed by 3 washes with wash buffer

(50 mM Tris[pH 7.5], 140 mM LiCl, 2 mM EDTA, 0.5 % IGEPAL CA-630, 0.5 mM DTT) and 1 wash in low salt elution buffer (10 mM Tris[pH 7.5]). Protein-RNA complexes were then heat-eluted (2 min, 80 ºC) in 2 ml of elution buffer (twice for depletion 1 and once for depletions 2 and 3 to give a total volume of 8 ml per sample). For mass spectrometric analysis RNA was digested by incubation with RNAse I (10 U/ml, Thermo Fischer Scientific) and benzonase (125 U/ml, Merck) for 1 h at 37ºC in elution buffer containing 1mM $MgCl_2$. Eluates were next concentrated using Amicon Ultra-15 centrifugal filter units (Millipore UFC901024) according to the manufacturer's instruction.

Sample preparation for proteome analysis

Proteins were precipitated with methanol-chloroform extraction (Wessel and Flugge 1984) and resuspended in 50 µL of 8 M urea and 0.1 M Tris[pH 8]. Proteins were reduced with 10 mM DTT at room temperature for 30 min and alkylated with 50 mM iodoacetamide at room temperature for 30 min in the dark. Proteins were first digested by lysyl endopeptidase (LysC) (Wako) at a LysC-to-protein ratio of 100:1 (w/w) at room temperature for 3 h. Afterwards, the sample solution was diluted to final concentration of 2 M urea with 50 mM ammonium bicarbonate. Trypsin (Promega) digestion was performed at a trypsin-to-protein ratio of 100:1 (w/w) under constant agitation at room temperature for 16 h. Peptides were desalted with C18 Stage Tips (Rappsilber et al. 2003) prior to LC-MS/MS analysis.

Mass spectrometry

Reversed-phase liquid chromatography was performed by employing an EASY nLC II (Thermo Fisher) using self-made fritless C18 microcolumns (Ishihama et al., 2002, PMID: 12498253) (75 µm ID packed with ReproSil-Pur C18-AQ 3-µm resin, Dr. Maisch GmbH) connected on-line to the electrospray ion source (Proxeon) of a Q Exactive mass spectrometer or Q Exactive plus (Thermo Fisher). Peptide samples were eluted at a flow rate of 200 nl/min with a 5 to 50 % acetonitrile gradient in 0.1% formic acid over 2 h. Settings for MS analysis is as follows: one full scan (resolution 70,000; m/z 300–1,700; target value 3e6; maximum injection 20 ms) followed by top 10 MS/MS scans (resolution 17,500; target value 1e6; maximum injection 60 ms) using higher-energy

collisional dissociation (HCD) (isolation width, 2; normalized collision energy, 26). The Q Exactive instrument was operated in the data dependent mode with a full scan in the Orbitrap followed by up to 10 consecutive MS/MS scans. Ions with an unassigned charge state and singly charged ions were rejected. Former target ions selected for MS/MS were dynamically excluded for 30 s.

Processing and analysis of mass spectrometry data

All raw data were analyzed and processed by MaxQuant (v1.3.0.5 or 1.5.1.2) (Cox and Mann 2008). Default settings were kept except that 'match between runs' was turned on. Search parameters included two missed cleavage sites, cysteine carbamidymethyl fixed modification, and variable modifications including methionine oxidation and protein N-terminal acetylation. The peptide mass tolerance was 6 ppm and the MS/MS tolerance was 20 ppm. Database search was performed with Andromeda (Cox et al. 2011) against UniProt/Swiss-Prot human database (downloaded on 2012-06) with common serum contaminants and enzyme sequences. False discovery rate (FDR) was set to 1% at peptide and at protein level.

Immunostaining, microscopy and Western analysis

Cells were seeded onto sterile coverslips, allowed to attach overnight and exposed to actinomycin D, UVC or IR at indicated conditions. Next, cells were fixed in 1 % (v/v) formaldehyde/PBS for 15 min at room temperature in the dark, followed by a 10-minute permeabilization step in 0.2 % (v/v) Triton-X/PBS. Cells were then washed in PBS and blocked in 20 % (v/v) FBS in 0.05 % (v/v) Tween 20/PBS (PBST) for 1h at room temperature. Primary antibody incubations (see Supplemental Table S9) were performed in PBST for 3h or overnight at 4 ºC. After washing, incubation with secondary anti-mouse IgG, conjugated to Alexa-488 (Thermo Fisher Scientific) was performed. Nuclei were labeled with 10 µg/ml Hoechst 33342 and images were acquired on a Zeiss Axio Observer Z1 or a Keyence BZ-X710 microscope using 63x magnification.

For Western blotting analysis of poly(A)$^+$ RNA-bound proteins, MCF-7 cells were grown in 15-cm dishes (2 per sample) until 80 % confluence, incubated with 200 µM 4SU overnight, followed by exposure to IR (10 Gy). After 1h of incubation, cells were UV crosslinked (365 nm, 0.2 J/cm^2),

followed by harvesting and lysis in lysis/binding buffer. After oligo(dT) affinity purifications, Western analysis of inputs, supernatants and eluates was performed using the antibodies listed in Supplemental Table S9.

PAR-CLIP experiments

MCF-7 Flp-In cells expressing FLAG/HA-tagged DDX54 were grown in standard DMEM to 80% confluence. For a typical experiment, fifteen 15-cm dishes were used. They were incubated with 200 µM 4SU (ChemGenes) for 16 h, IR-exposed (10 Gy), UV-crosslinked at 365 nm (0.2 J/cm^2), harvested and stored at -80 °C. Cell pellets were lysed in approximately 3 cell pellet volumes (5 ml) of high salt lysis buffer (50 mM Tris[pH 7.5], 500 mM NaCl, 1% (v/v) IGEPAL-CA630, 1 mM DTT, complete EDTA-free protease inhibitor cocktail (Roche)), incubated for 30 minutes on ice followed by 10-second sonication at 80% amplitude. Immunoprecipitation, small RNA cloning and cDNA library generation were performed as previously described (Hafner et al. 2010; Hafner et al. 2012) with the following modifications. The first and second RNase T1 treatments were performed at 1 U/µl for 10 min at 22 °C and at 2 U/µl in 80 µl of IP wash buffer (50 mM HEPES[pH 7.5], 300 mM KCl, 0.05% (v/v) IGEPAL CA-630, 0.5 mM DTT) for 7 min at 22 °C, respectively. Crosslinked protein-RNA complexes were resolved by SDS-PAGE using Novex 3-8% gradient Tris-acetate gels (Thermo Fischer Scientific). Proteins and protein-RNA complexes were transferred to a nitrocellulose membrane (Whatman) and exposed to phosphorimager screen. ^{32}P-RNA-FLAG/HA-DDX54 complex migrating at approximately 115 kDa was excised and digested by proteinase K (Roche, 4 mg/ml for 1 h at 55 °C). After phenol-chloroform extraction and precipitation, RNA was ligated to 3' adaptor (see Oligonucleotides) and gel-purified. Two RNA size ranges per each sample (corresponding to 22-35 and 50-100 nucleotides of unligated fragments) were excised, 5'adaptor ligated, reverse transcribed and PCR-amplified to obtain cDNA libraries. cDNA libraries were amplified using Phusion High-Fidelity DNA polymerase (Thermo Fischer Scientific) and utilizing differing barcode- containing reverse PCR primers (see Oligonucleotides). This was followed by excision of approximately 140-180 bp sized fragments from 2.5 % (w/v) agarose gels and subsequent purification by QIAquick gel extraction kit (Qiagen). Massively parallel sequencing

was carried out on HiSeq2000 Illumina instrument using 1x51+7 cycle and 1x101+7 cycle runs. Clustering and amplification of cDNA libraries was performed according to manufacturer's instruction.

PAR-CLIP analysis

We used the previously published computational PAR-CLIP pipeline (v0.97a) (Lebedeva et al. 2011; Jens 2016). Briefly, the reads were demultiplexed and adapter-removed (Flexbar v2.5), collapsed (Lebedeva et al. 2011; Jens 2016) and trimmed (FASTX Toolkit, v0.0.14). Next, the reads were aligned to hg19 and a set of RefSeq pre-mRNA and mRNA sequences using BWA 0.5.8c. Unique alignments were converted to a pileup and read clusters scored for characteristic conversions and read variability. After stringent-false positive filtering (using antisense clusters as a decoy database and a false discovery rate of 0.05) remaining clusters were written as bed files. See (Lebedeva et al. 2011; Jens 2016) for details. Conservative sets of DDX54 binding sites were defined by only retaining overlapping reads with T-C transitions in both biological replicates. Consensus sets of DDX54 binding sites were defined as a pooled average of the two biological replicates, retaining overlapping reads with T-C transitions in at least one of the two biological replicates.

For T-C transition event counts, we sequentially aligned reads to reference references in the following order starting with human pre-rRNA (U13369.1) followed by rRNA (NR_023363.1, NR_003285.2, NR_003287.2, NR_003286.2), other ncRNA (Ensembl), tRNA (GtRNADb), mtDNA (AF347015.1) and finally the human genome (hg19, iGenomes). T-C transitions were extracted from BAM files using SAMtools (v.0.1.19) and a perl script (Schueler et al. 2014). Next, using intersectBed (BEDtools), the number of T-C transitions in different transcript regions and biotypes as defined in GENCODE v19 annotation were counted, while excluding the annotations that overlapped the used reference mapping sequences (rRNA, snRNA, snoRNA, 7SK, 7SL, TERC, tRNA). Final categories given in Fig. 2B were defined as follows:

- (pre-)rRNA: all T-C transitions mapping to pre-rRNA and rRNA sequences; genomic alignments to rRNA loci were discarded.

- intronic: all T-C transitions not mapping to exons but mapping to introns of protein-coding genes;

- CDS: all T-C transitions mapping to CDS exons

- 3'UTR: all T-C transitions mapping to 3'UTR exons

- 5'UTR: all T-C transitions mapping to 5'UTR exons

- lncRNA: all T-C transitions mapping to exons of the following Gencode v19 categories; processed_transcript, lincRNA, antisense, sense_overlapping, sense_3prime_overlapping_ncrna.

- other ncRNA: all T-C transitions mapping to exons of Gencode v19 miRNA category, all T-C transitions mapping to snRNA, snoRNA, 7SK, 7SL, TERC and tRNA sequences. Genomic alignments to snRNA, snoRNA, 7SK, 7SL, TERC and tRNA loci were discarded.

To detect spliced reads in PAR-CLIP sequencing libraries (Supplemental Table S2), reads were first aligned to ncRNA sequences as described above, followed by mapping of the remaining reads to the genome, where STAR RNA-seq aligner (ver 2.4.2a) was used with the following parameters: --outFilterType BySJout --outFilterMultimapNmax 20 --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --outFilterMismatchNoverLmax 0.05 --outFilterMatchNmin 16 --outFilterScoreMinOverLread 0 --outFilterMatchNminOverLread 0 --alignIntronMin 20 --alignIntronMax 1000000.

For 7-mer enrichment analysis we determined nucleotide frequencies based on Markov models (Eddy 2004) in the sequence of 20 nucleotides (nt) upstream and downstream from the most frequent T-C transition position within a consensus DDX54 binding site. Background regions of the same length were selected either upstream or downstream from DDX54 binding sites with no overlap and were required to be located in the same transcript region as the DDX54 binding site (exons, introns or UTRs). To determine the occurrence of exonic spliced enhancers (ESE) relative to the most frequent DDX54 T-C transition in a window of 40 nt within DDX54 binding sites we searched for ESE consensus motifs obtained from (Cartegni et al. 2003).

For secondary structure predictions around 3' splice sites we compared DDX54-bound and unbound introns. Regions starting 60 nt upstream and ending 50 nt downstream of 3' splice sites were folded using the Vienna RNA package (Lorenz et al. 2011) using RNAplfold program (-W 80 -L 40 –u 1). Base pair probabilities were averaged per nucleotide position over all sequences and

differences between both DDX54 intron classes were assessed with the Kolmogorov-Smirnov test. For secondary structure predictions around DDX54 crosslinking sites, we folded regions 50 nt upstream and 50 nt downstream from the most frequent T-C transition within top 2000 FDR-ranked conservative DDX54 binding sites and performed the same analysis as described above.

VARNA (v3.92) was used for visualization of secondary snRNA structures.

Additional 4SU-seq analysis

To obtain percent intron retention scores, reads from each sample were first separately mapped to the human genome (hg19) and *D. melanogaster* genome (BDGP6) using TopHat2 (v2.1.0). To manipulate the alignment BAM/SAM files we used SAMtools (v0.1.19) (Li et al. 2009). We intersected human and *D. melanogaster* alignment BAM files and removed the reads that aligned to both genomes. The reads that uniquely mapped to the human genome were extracted from the remaining reads by filtering for the NH:i:1 attribute in the BAM/SAM file.

Next, we performed an intron-centric analysis by using junction and spliced read counts in all non-overlapping Refseq introns and their upstream or downstream exons as described previously (Braunschweig et al. 2014). Specifically, we used SAMtools and BEDtools (v2.25.0) (Quinlan and Hall 2010) to count unspliced reads in introns (I), 5' and 3' exon-intron junctions (E1I and IE2, respectively) and spliced reads in exon-exon junctions (E1E2). Briefly, we demanded a minimum overlap of eight nucleotides over any given junction for a read to be counted. Read counts were then normalized using DESeq2 size factors obtained from *D. melanogaster* exon read counts from HTSeq-count outputs. Per-intron PIR scores were computed as follows: PIR=100x average(E1I, IE2)/(E1E2 + average(E1I, IE2)), where E1I,IE2 and E1E2 were normalized read counts for the junctions described above. We only considered PIR scores for introns with sufficient read evidence by demanding an empirical criterion of median (EI1, IE2, I) + E1E2 to be higher than 10. In addition, a between-replicate mean PIR score was only computed for introns where the computation of PIR score was possible for both replicates.

To infer synthesis, degradation and processing rates we used INSPEcT v26 Bioconductor package (de Pretis et al. 2015). Instead of intronic and exonic FPKMs, which are typically used in INSPEcT,

we provided primary and mature FPKM values from RSEM outputs. This would prevent the potential overestimation of mature sequences due to considering all exonic sequences as mature. The remaining INSPEcT parameters were set to default.

To quantify and define differentially spliced events we ran SUPPA (Alamancos et al. 2015) using default parameters. We provided TPM values from RSEM outputs.

BioID proximity ligation assays and co-immunoprecipitation

To perform BirA-DDX54 proximity ligation assays (BioID) we used published protocols (Couzens et al. 2013). Briefly, four 15-cm dishes of doxycycline-inducible stably overexpressing BirA-FLAG/DDX54 HEK293 Flp-In T-REx cells at 90% confluence were used per sample (approximately $1x10^9$ cells). To detect differences in BirA-DDX54 interacting and proximal proteins after IR exposure, we first induced cells with 1 µg/ml doxycycline (Sigma) for 24 h, followed by IR exposure (10 Gy) and addition of 250 µM biotin. Biotin-treated non-doxycycline-induced cells were used as negative control. Three and a half hours after the addition of biotin, cells were harvested and BioID was performed as described below.

Briefly, cells were lysed in modified RIPA buffer (50 mM Tris[pH 7.5], 150 mM NaCl, 1% (v/v) IGEPAL CA-630, 1 mM EDTA, 1 mM EGTA, 0.1 % (m/v) SDS, 0.5 % (m/v) sodium deoxycholate, complete EDTA-free protease inhibitor cocktail (Roche)) for 15 min on ice, passed 8 times through 21 G needle and sonicated (six 5-second pulses, 30% amplitude). Next, 250 U benzonase were added and lysates were incubated horizontally on ice for 1 h with agitation. Lysates were cleared (15.000g, 15 min, 4 °C), filtered through a 5µm Supor membrane and incubated with 40 µl (bed volume) of RIPA buffer-washed streptavidin sepharose (17-5113-01, GE Healthcare) for 3 h at 4 °C with rotation. Sepharose was then washed twice in RIPA buffer (1 ml), twice in wash buffer (50 mM Hepes[pH 8.0], 100 mM KCl, 10% (v/v) glycerol, 2 mM EDTA, 0.1% (v/v) IGEPAL-CA630) and 3 times in 50 mM ammonium bicarbonate[pH 8.0] (ABC). For mass spectrometry analysis, beads were further processed for alkylation and tryptic digestion in ABC. For Western analysis, an aliquot of washed beads (10%) was incubated at 90 °C for 5 minutes in 95% formamide containing 10mM

EDTA and 1 mM biotin, followed by addition of Laemmli buffer, reboiling (2 min) and loading on SDS-PAGE gel.

For anti-FLAG co-immunprecipitation (co-IP) experiments, one hour after IR exposure (10 Gy), two 15-cm dishes of FLAG-HA/DDX54 MCF-7 Flp-In cells at 80 % confluence were washed with ice-cold PBS. Cells were directly lysed in 4 ml of ice-cold lysis buffer (1 % (v/v) IGEPAL-CA630, 50 mM Tris[pH 7.5], 300 mM NaCl, complete EDTA-free protease inhibitor cocktail (Roche)), followed by sonication (3 times with a 5-second pulse at 30 % amplitude) and incubation for 30 min on ice. Lysates were cleared by centrifugation (15.000g, 15 min, 4 °C), treated with RNAse T1 (Thermo Fisher) or left untreated. Low (+), intermediate (++) and high (+++) RNAse T1 treatments corresponded to 2.5 U/µl for 5 min at 22 °C, 5 U/µl for 5 min at 37 °C and 50 U/µl for 5 min at 37 °C, respectively. Aliquots (5 %) of the input lysate were removed and subsequently used for RNA extraction or Western analysis. Anti-FLAG M2 monoclonal antibody (Sigma Aldrich: F3165) or IgG1 isotype control (Sigma Aldrich, M5284) -conjugated magnetic Dynabeads Protein G (Thermo Fisher) (40 µl per sample) were prepared as described (Spitzer et al. 2014). Next, beads were mixed with cell lysates and incubated for 1 h at 4 °C with rotation. Afterwards, supernatants were removed and beads were washed four times in 0.5 % (v/v) IGEPAL-CA630, 50 mM Tris[pH 7.5], 300 mM KCl, complete EDTA-free protease inhibitor cocktail (Roche). Beads were re-suspended in sample loading buffer and analyzed along with input lysate samples by Western analysis using antibodies listed in Supplemental Table S9.

Purification of DDX54-FLAG/HA

DDX54-FLAG/HA was transiently overexpressed in Flp-In T-REx HEK293 cells and purified by anti-FLAG immunoprecipitation. Briefly, 7 plates of Flp-In T-REx HEK293 cells were grown to 50-60 % confluence and transfected with pcDNA5/FRT/TO-DDX54-FLAG/HA (20 µg DNA per plate) using the calcium phosphate method (Graham and van der Eb 1973; Jordan et al. 1996). After 24 h, media was exchanged for fresh culture media containing 4 µg/ml doxycycline. Cells were harvested 72 hours post transfection in ice-cold PBS, washed and lysed in 10 ml of ice-cold high salt lysis buffer (1 % (v/v) IGEPAL CA-630, 50 mM Tris[pH 7.5], 500 mM NaCl, 0.5 mM DTT,

complete EDTA-free protease inhibitor cocktail (Roche)) and sonicated (three 5-second pulses at 30% amplitude). Next, 250 U of benzonase were added and lysates were incubated horizontally on ice for 30 min with agitation. Lysates were cleared (15.000g, 15 min, 4 °C), filtered through a 5 µm Supor membrane and incubated with 100 µl (bed volume) of pre-washed anti-FLAG M2 agarose gel (Sigma) for 2 h at 4 °C. Next, the beads were washed 4 times in 5 ml of ice-cold wash buffer (0.5 % (v/v) IGEPAL CA-630, 50 mM Tris[pH 7.5], 400 mM KCl, 0.5 mM DTT, complete EDTA-free protease inhibitor cocktail (Roche)) and 1 time in 1 ml of FLAG elution buffer (20 mM Tris[pH 7.5], 100 mM NaCl, 10 % (v/v) glycerol). Next, the elution was performed at room temperature for 45 minutes with rotation in 200 µl of FLAG elution buffer containing 0.5 µg/µl FLAG peptide. Afterwards, samples were centrifuged and the supernatants were stored at -80 °C until further use. Lysates from untransfected cells were prepared in high salt lysis buffer containing 0.1 µg/µl FLAG peptide and processed exactly as described above.

TCGA data analysis

We integrated our PAR-CLIP data with TCGA expression z-scores and PSI values for all alternative events defined in (Sebestyen et al. 2016). PAR-CLIP T-C transitions were considered if they were present in the following alternatively spliced events according to (Alamancos et al. 2015):

RI: intronic regions including 50 nucleotides up- and downstream of 5’ (e1) and 3’ (s2) splice sites, respectively.

SE: alternative exon including 20 nucleotides up- and downstream of 3’ (s2) and 5’ (e2) splice sites, respectively.

MX: (main) alternative exon including 20 nucleotides up- and downstream of 3’ (s2) and 5’ (e2) splice sites, respectively.

A3: region between the 5’ splice site (e1) and the alternative 3’ splice site (s2) including 20 nucleotides up- or downstream.

A5: region between the alternative 5’ splice site (e2) and the 3’ splice site (s3) including 20 nucleotides up- or downstream, respectively.

T-C transitions obtained from PAR-CLIP experiments were intersected with the above regions using intersectBed function (BEDtools). Next, the total number of T-C transitions per alternative splicing event was computed for all four PAR-CLIP libraries (2 control and 2 IR samples). A cutoff of 25 T-C transitions per splicing event was used to define DDX54-bound events. Events with not more than one total T-C transition were considered unbound.

We used two approaches to evaluate the effect of DDX54 binding and expression on tumor-specific alternatively spliced events. In the first approach, for each event, we first computed Spearman correlation coefficients between DDX54 expression and PSI values for all normal and tumor samples. Next, we compared the distributions of Spearman correlation coefficients for DDX54-bound and non-bound events and evaluated significant differences in coefficients by Kolmogorov-Smirnov tests. For retained intron events, the result is shown in Fig. 6E. In the second approach, we first classified the normal and tumor samples into terciles based on differences in DDX54 expression. Briefly, paired normal and tumor samples were used to determine a per-sample difference in z-score. Next, we defined terciles of delta(z) values using R's quantile function and assigned samples into three groups based on DDX54 expression (q1 – low, q2 – medium, q3 – high). Next, we computed differences between tumor and normal median PSIs across all samples in a given tercile. Differences in ΔPSI values between highly DDX54-bound and non-bound alternative events were assessed using Kolmogorov-Smirnov's test. The results of this approach are shown in Supplemental Fig. S10C, D.

Oligonucleotides

DDX54 C-terminal FLAG/HA sub-cloning primers

DDX54 *Kpn*I forward ACGCATGGTACCGATGGCGGCCGACAAGG

DDX54 *Bam*HI reverse AGTTTAGGATCCTCAAGCGTAATCGGGCACGTCATAAGGGTACTTGTC ATCGTCGTCCTTGTAGTCAAGCATCCTCTTCCGCATCTTGC

siRNAs

DDX54_s1 rGrArArGrArArGrUrCrUrGrGrArGrGrCrUrUrC[dT][dT]

DDX54_as1 rGrArArGrCrCrUrCrCrArGrArCrUrUrCrUrUrC[dT][dT]

DDX54_s3 rUrCrArCrUrArArGrGrArGrCrUrArGrGrCrArA[dT][dT]

DDX54_as3 rUrUrGrCrCrUrArGrCrUrCrCrUrUrArGrUrGrA[dT][dT]

DDX54_s4 rCrCrArArArCrUrGrCrUrGrGrUrGrGrArArUrU[dT][dT]

DDX54_as4 rArArUrUrCrCrArCrCrArGrCrArGrUrUrUrGrG[dT][dT]

DDX54_s5 rGrArArGrArCrArArGrArArGrArArGrArUrUrA[dT][dT]

DDX54_as5 rUrArArUrCrUrUrCrUrUrCrUrUrGrUrCrUrUrC[dT][dT]

Small RNA cloning adaptors

5'adaptor oR5-NN rGrUrUrCrArGrArGrUrUrCrUrArCrArGrUrCrCrGrArCrGrArUrCrNrN

3' adaptor NN-RA3 /5Phos/NNT GGA ATT CTC GGG TGC CAA GG/3InvdT/

RT primer GCCTTGGCACCCGAGAATTCCA

PAR-CLIP PCR amplification

Forward primer

RP1 AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGA

Reverse Index Primers

RNA PCR Primer, Index 1 (RPI1)

CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCCTTGGCACCCGAGAATTCC

A

RNA PCR Primer, Index 2 (RPI2)

CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCCTTGGCACCCGAGAATTCC

A

RNA PCR Primer, Index 3 (RPI3)

CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCCTTGGCACCCGAGAATTCC

A

RNA PCR Primer, Index 4 (RPI4)

CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCCTTGGCACCCGAGAATTCC

A

RNA PCR Primer, Index 5 (RPI5)

CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCCTTGGCACCCGAGAATTCC

A

RNA PCR Primer, Index 6 (RPI6)

CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCCTTGGCACCCGAGAATTCC

A

RNA PCR Primer, Index 7 (RPI7)

CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCCTTGGCACCCGAGAATTCC

A

RNA PCR Primer, Index 8 (RPI8)

CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCCTTGGCACCCGAGAATTCC

A

PCR primers (*in vitro* splicing assay):

MINX exon1 forward GAATACACGGAATTCGAGCTC

MINX exon2 reverse GATCCCCACTGGAAAGACC

RT-PCR primers (coordinates of retained intron events are given for each gene):

geneName;geneId:intronNumber; chromosome-exon1Start:exon1End_exon2Start:exon2End:

strand

PPP3CA:NM_001130691:12; chr4-101944587:101947218_101953424:101953521:-

*PPP3CA*_exon_forward AGCGGAGTACTTTCTGGAGG

*PPP3CA*_intron_forward AGGTCCCGTCACTTTGTTCT

*PPP3CA*_exon_reverse AGGCGGCATCCTCTCATTAA

MDM2:NM_002392:9;chr12-69229609:69229764_69230452:69230529:+

*MDM2*_exon_forward GAAGAAGGACAAGAACTCTCAGA

*MDM2*_intron_forward TGTGGGTAAGGATTTCTCTCTCC

*MDM2*_exon_reverse AATGAATCTGTATCACTCTCCCC

XRCC6:NM_001469:1;chr22-42017295:42017349_42017994:42018090:+

*XRCC6*_exon_forward CATGCGTGGATTGTCGTCTT

*XRCC6*_intron_forward CCTGAAACGTGAGGGATAGC

*XRCC6*_exon_reverse ATGACTCCCACCCTGACATG

WDR55:NM_017706:2;chr5-140047819:140047919_140048000:140048087:+

*WDR55*_exon-exon_forward AAACCAAGGAGCTCTGGTCA

*WDR55*_exon-exon_reverse ATGAGCCTTGGAAACACGTC

DNAJC2:NM_014377:8;chr7-102962958:102963079_102963150:102963241:-

*DNAJC2*_exon-exon_forward TCGTGATGAGAGGAGATGGA

*DNAJC2*_exon-exon_reverse TGCTCCTTCCGTTTAGCTTC

CCND1:NM_053056:3;chr11-69458600:69458759_69462762:69462910:+

*CCND1*_exon_forward AAACAGATCATCCGCAAACA

*CCND1*_intron_forward CTCAGAGCCAGCACACAGG

*CCND1*_exon_reverse GGCGGATTGGAAATGAACTT

ERCC3:NM_000122:5;chr2-128046913:128047095_128047265:128047400:-

*ERCC3*_exon-exon_forward CCGAGAATGCCGCTTAAGAA

*ERCC3*_exon-exon_reverse AGTCAAACAGGTCCATGGGG

RAD23A;NM_005053:5;chr19-13059500:13059627_13059895:13059973:+

*RAD23A*_exon-exon_forward GAGCCGTGGAGTATCTGCT

*RAD23A*_exon-exon_reverse CTCCTGGACAGAACCGTGTT

*alphaTub84B*_forward TGTCGCGTGTGAAACACTTC

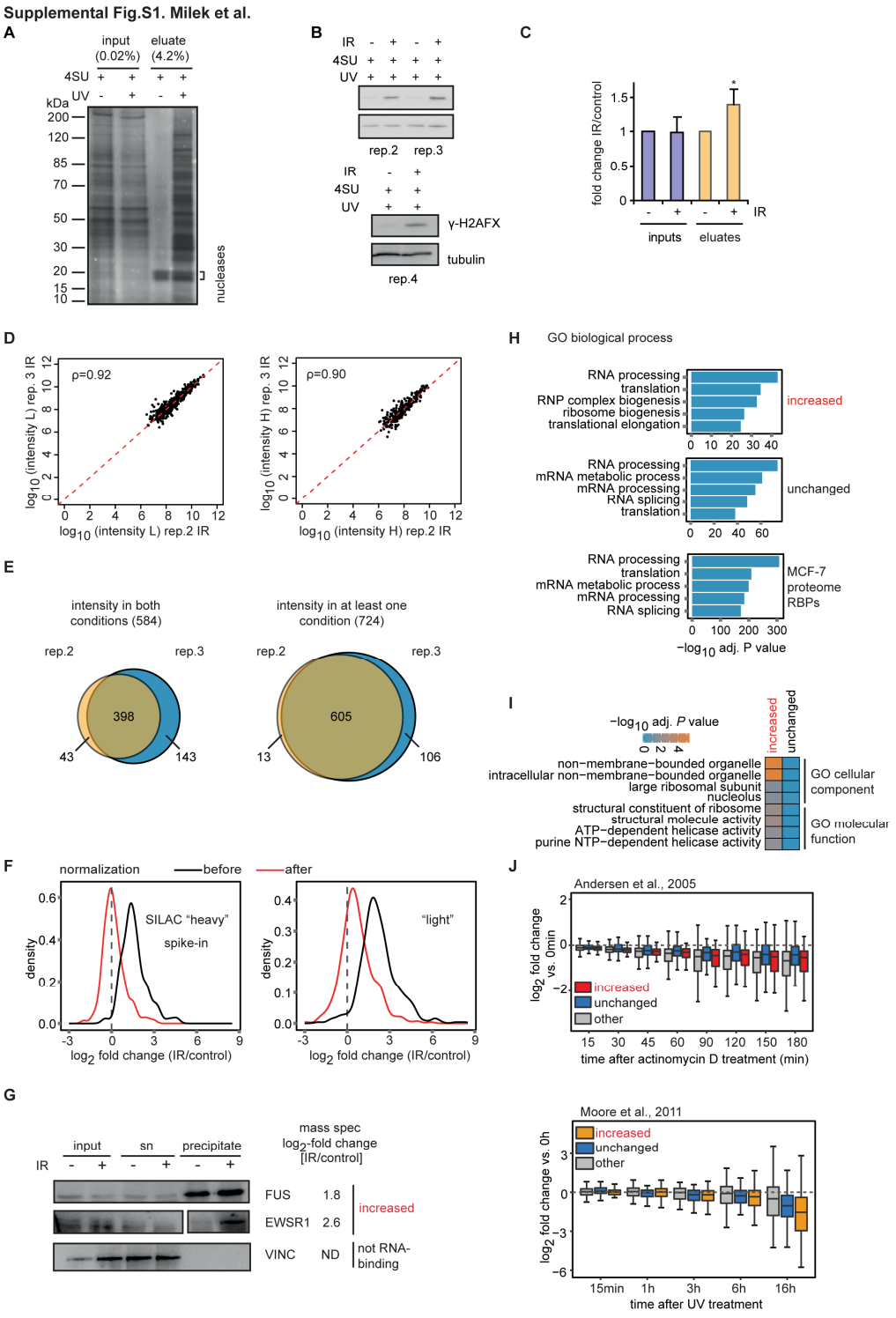
*alphaTub84B*_reverse AGCAGGCGTTTCCAATCTG

SUPPLEMENTAL FIGURES

Supplemental Fig. S1.

Differential mRNA interactome capture shows increased binding of proteins to poly(A)$^+$ RNA upon exposure to ionizing radiation. (A) MCF-7 cells were treated with 4SU (200 µM) for 16 h and exposed to UV 365nm (0.2 J/cm^2). After oligo(dT) purification protein-RNA complexes were heat-eluted, nuclease treated and resolved with SDS-PAGE. Proteins were visualized using silver staining. (B) Western analysis of phosphorylated H2AFX upon 4SU treatment, UV and/or IR exposure in total cell lysates for replicates 2-4 (replicate 1 is shown in Fig. 1). (C) Comparison of protein amounts between IR-exposed and control samples. Total band intensity from silver stained SDS-PAGE gels (n=3) was quantified by ImageJ. Average fold changes between IR-exposed and control samples from three independent experiments are shown for input and eluate lanes (*p<0.05, Student's *t* test). (D) Correlation between "light" and "heavy" summed peptide intensities for replicates 2 and 3 (IR condition). Each dot denotes a protein group with $\log_{10}$-transformed intensity. Spearman's rho (ρ) is given as a measure of correlation between replicates. (E) Per-replicate overlap between identified proteins with either intensity in both control and IR-exposed conditions (left) or intensity in at least one condition (right). Absolute numbers are given. (F) Effect of SILAC "heavy" intensity-derived normalization factor on the distribution of $\log_2$-transformed fold changes between IR-exposed and control cells for "heavy" (left) and "light" intensity data (right). (G) Validation of MS results by Western analysis. Shown are representative Western blots to detect proteins belonging to increased and non-RNA binders as negative control. (H) GO term enrichment analysis (DAVID) of proteins with increased, unchanged binding to mRNA and MCF-7 proteome RBPs. Negative $\log_{10}$-transformed adjusted *P* values (Benjamini-Hochberg) are plotted for enriched biological process GO terms. No background gene set was used. (I) GO enrichment analysis (DAVID) of proteins with increased and unchanged mRNA binding. RBPs detected in MCF-7 whole proteome analysis were used as the background gene set. Negative $\log_{10}$ adjusted P values (Benjamini-Hochberg) are shown in the heatmap. (J) Proteins with unchanged and increased mRNA binding were intersected with two nucleolar proteome dynamics datasets

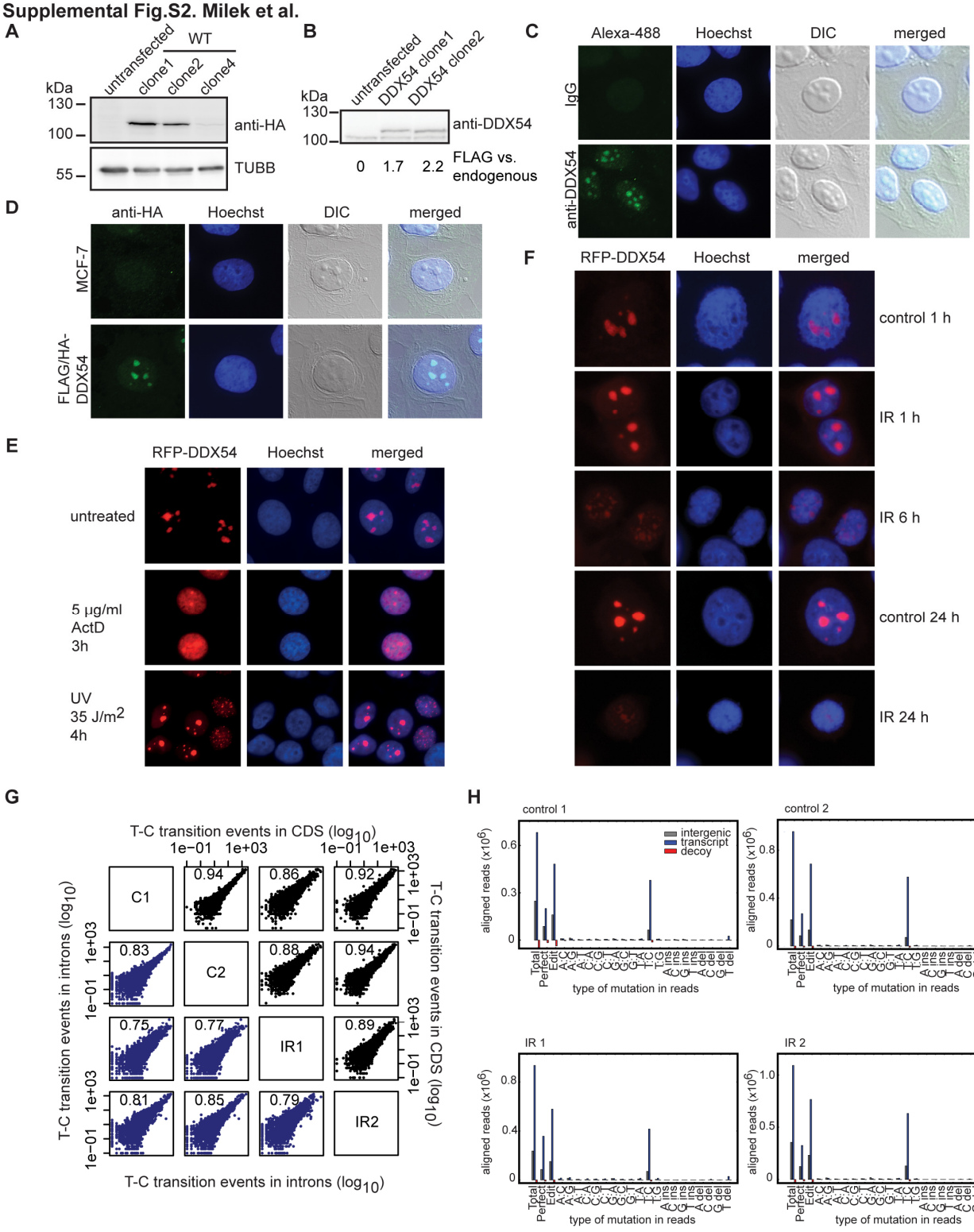
(Andersen et al. 2005; Moore et al. 2011). Non-overlapping nucleolar proteins were classified as "other". Normalized log_2-transformed fold changes in nucleolar protein abundance after different time periods post actinomycin D treatment (1 µg/ml, top) or UV exposure (35 J/m^2, bottom) are shown.



Supplemental Fig. S2.

Localization of DDX54 and PAR-CLIP diagnostics. (A) Western analysis of three clones of cell lines stably overexpressing FLAG/HA-DDX54. Clone 1 was used for further experiments. (B) Anti-DDX54 antibodies were used to detect both recombinant and endogenous DDX54 in FLAG/HA-DDX54 stable cell lysates using Western analysis. (C) Localization of endogenous DDX54 in MCF-7 Flp-In cells. Immunofluorescence of fixed and permeabilized MCF-7 cells stained with IgG isotype control (top row) and anti-DDX54 antibody (bottom row). (D) Localization of recombinant FLAG/HA-DDX54. MCF-7 (top) and FLAG/HA-DDX54 stably overexpressing cells (bottom) were stained with anti-HA antibody and imaged. (E) Re-localization of DDX54 detected by fluorescence microscopy. MCF-7 cells stably expressing RFP-DDX54 were treated with actinomycin D (5 µg/ml) or exposed to UVC (35 J/m^2). After indicated time periods, they were fixed, Hoechst-stained and imaged. (F) Re-localization of DDX54 detected by fluorescence microscopy. MCF-7 cells stably expressing RFP-DDX54 were exposed to IR (10 Gy) and fixed after indicated time periods, Hoechst-stained and imaged. (G) Correlation between number of PAR-CLIP T-C transitions in coding sequence (CDS) exons (black) or introns (violet) in RefSeq transcripts. Log_{10}-transformed number of T-C transitions per transcript compartment is shown. Spearman's rho is indicated for each scatter plot. (H) DDX54 PAR-CLIP alignment statistics. Y-axes show number of uniquely mapped reads to hg19 per sample. Possible mutations in sequencing reads are shown (substitutions, deletions and insertions as well as total number of aligned reads, reads with perfect match (perfect) and reads containing mutations (edit)).

Supplemental Fig.S2. Milek et al.



Supplemental Fig. S3.

Impact of DDX54 on pre-rRNA processing. (A) DDX54 PAR-CLIP T-C transition profile over the transcribed region of human rDNA locus for control (black) and actinomycin D (red) treated cells (5 µg/ml, 3h). Normalized absolute number of T-C transition events per position is shown. Mature rRNA sequences and transcribed spacers are depicted as blue boxes and black lines, respectively. Pre-rRNA cleavage sites are denoted with orange arrows. ETS, external transcribed spacers; ITS, internal transcribed spacers. (B) Total RNA was extracted from mock- and siDDX54-transfected MCF-7 cells using different siRNAs (1, 4 and 5). Northern blotting using probes hybridizing to either 5'ETS, ITS1 or ITS2 regions of the pre-rRNA to detect processing intermediates was performed. RNA loading was monitored by methylene blue staining.

47S, full-length transcribed pre-rRNA.

45S, produced by endonucleolytic cleavage of 47S at sites 01 and 02.

43S, produced by endonucleolytic cleavage of 45S at site A0.

41S, produced by endonucleolytic cleavage of 43S at site 1.

30S, produced by endonucleolytic cleavage of 45S at site 2.

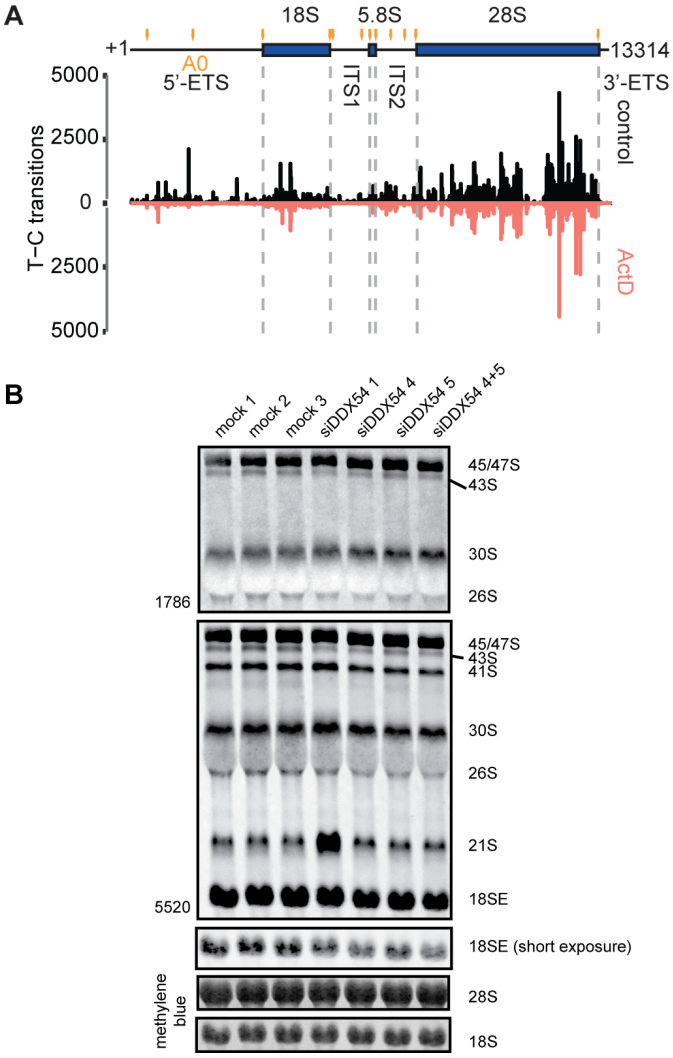
26S, produced by endonucleolytic cleavage of 30S at site A0.

21S, produced by endonucleolytic cleavage of 26S (site 1) or, alternatively, 41S (site 2).

21SC, produced by 3'exonucleolytic cleavage of 21S.

18SE, produced by 3'exonucleolytic cleavage and endonucleolytic cleavage of 21SC at site E.

Supplemental Fig. S3. Milek et al.

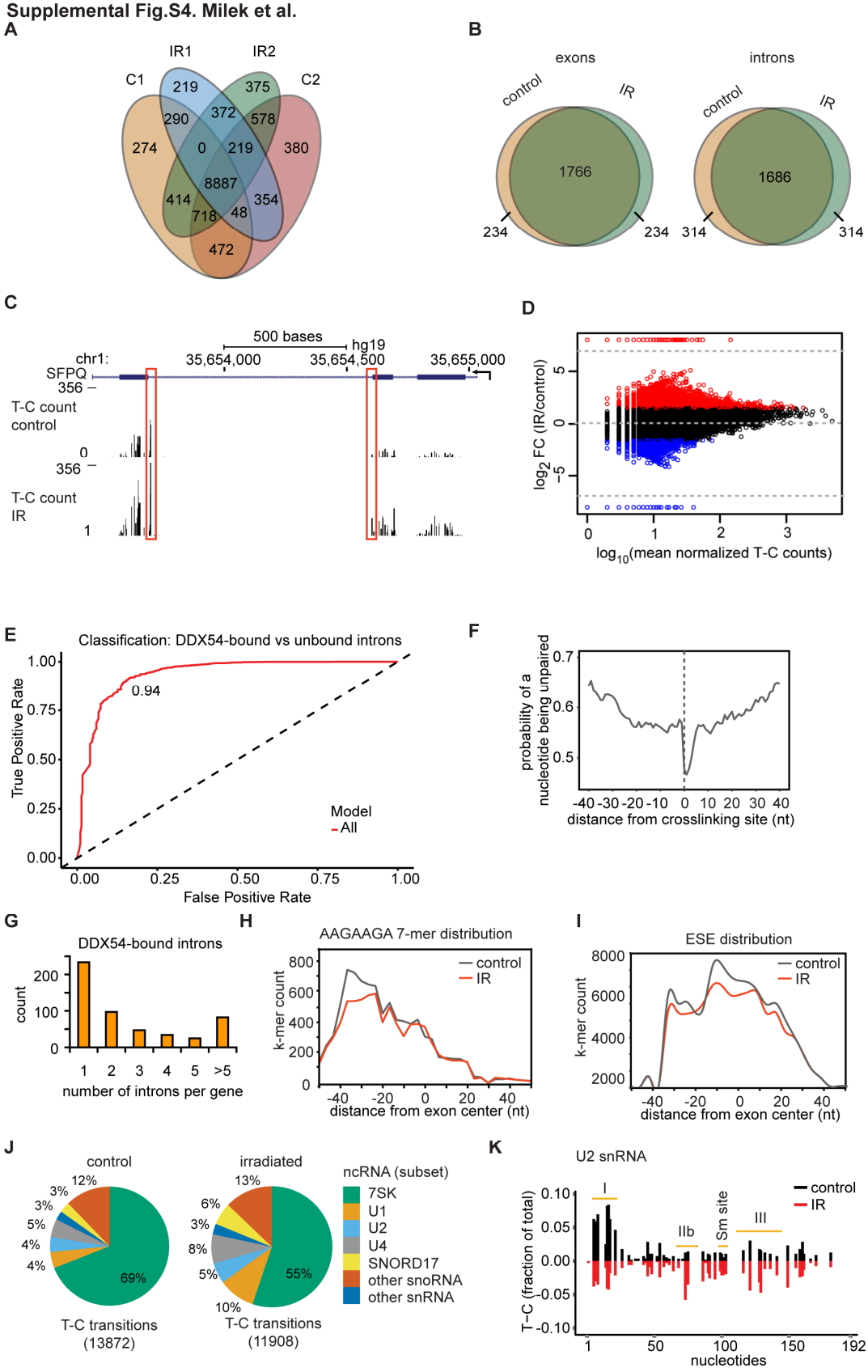


Supplemental Fig. S4.

PAR-CLIP reveals increased binding of DDX54 to acceptor splice sites upon exposure to IR.

(A) Overlap in the number of PAR-CLIP target transcripts determined in different replicates is shown by Venn diagrams. (B) Overlap of top 2000 DDX54 target transcripts between control and IR samples ranked by number of T-C transitions in exons (left) or introns (right). (C) UCSC browser snapshot in a portion of *SFPQ* locus depicting T-C transitions (absolute numbers) in control and IR (1h) DDX54 PAR-CLIP samples. Red frames show regions around donor and acceptor splice sites of exons 7 and 8, respectively. (D) MA plot of $\log_2$- transformed fold changes in DDX54 binding occupancy (IR vs. control). For each merged DDX54 binding site between control and IR-exposed conditions, the number of T-C transitions was counted, normalized between conditions, total T-C count per transcript and the expression level obtained from RNA-seq datasets (1h post IR). Binding sites with higher occupancy (log2-fold change >1.5) are shown in red, while those with decreased occupancy are denoted in blue (log2-fold change <-1.5) (E) The area under receiver operating characteristic (ROC) curve for predictions of DDX54-bound vs. unbound introns using 5-fold cross-validation. (F) Average probabilities of nucleotides being unpaired in the regions of 40 nt upstream and downstream from crosslinking sites in top 2000 DDX54 binding sites selected from the conservative PAR-CLIP set based on ranking by FDR values. Light grey shaded region shows significantly different base pairing probabilities between control and IR exposed samples ($P<10^{-4}$, K-S test). (G) Absolute counts of DDX54-bound introns (n=1705) per gene (n=518). (H) Distributions of DDX54 PAR-CLIP derived AAGAAGA motifs and (I) all exon splice enhancer (ESE) elements overlapping DDX54 binding sites in regions relative to exon centers for control and IR exposed (10 Gy, 1h) conditions. (J) Percentages of DDX54 PAR-CLIP T-C transition events for a subset of nuclear ncRNAs in control and IR-exposed (10Gy, 1h) cells. The number of considered T-C transition events is given in parentheses. (K) DDX54 PAR-CLIP T-C transition profile over human U2 snRNA for control (black) and IR exposed (10Gy, 1h) cells (red). Fraction of total U2 snRNA T-C transition events is shown per position. Positions of the Sm binding site and stem loops (I, IIb and III) are denoted by yellow lines.

Supplemental Fig.S4. Milek et al.

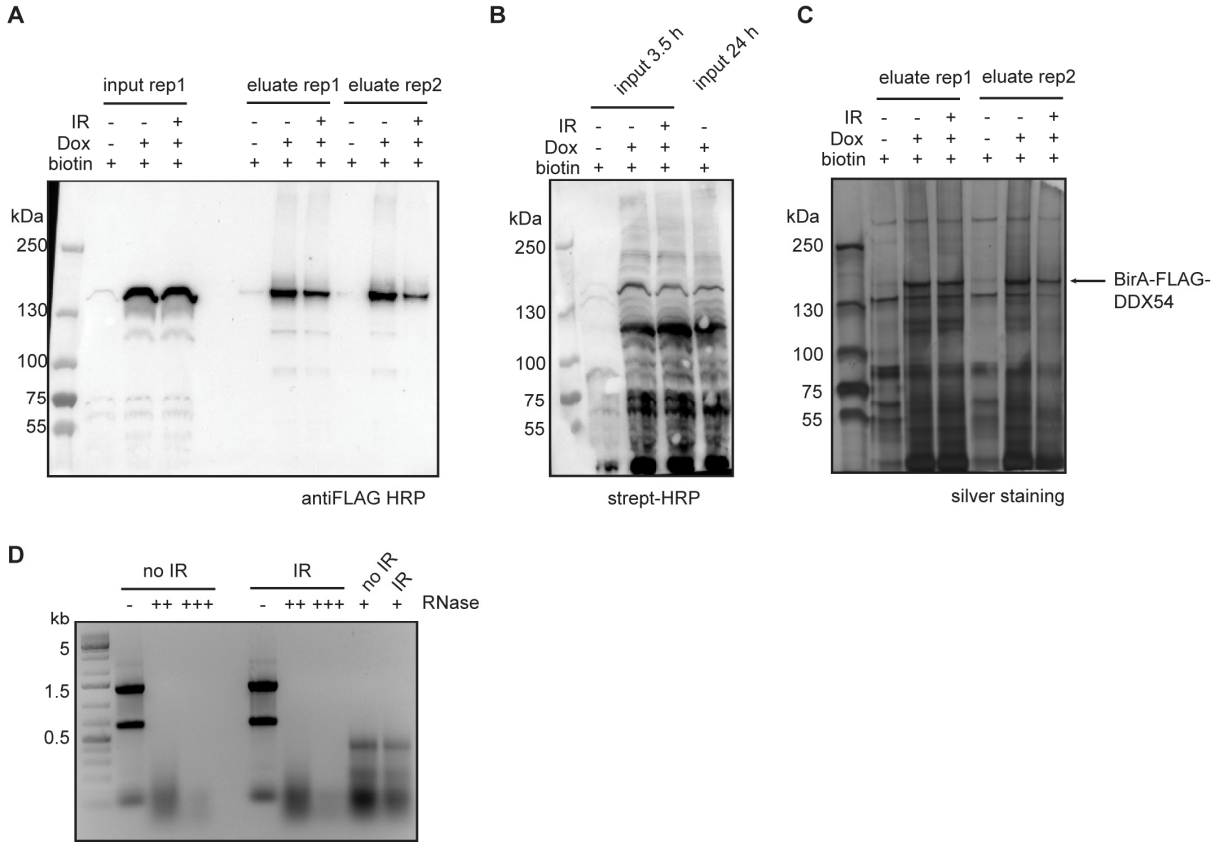


Supplemental Fig. S5.

Proximity-dependent biotinylation assay of BirA-FLAG-DDX54.

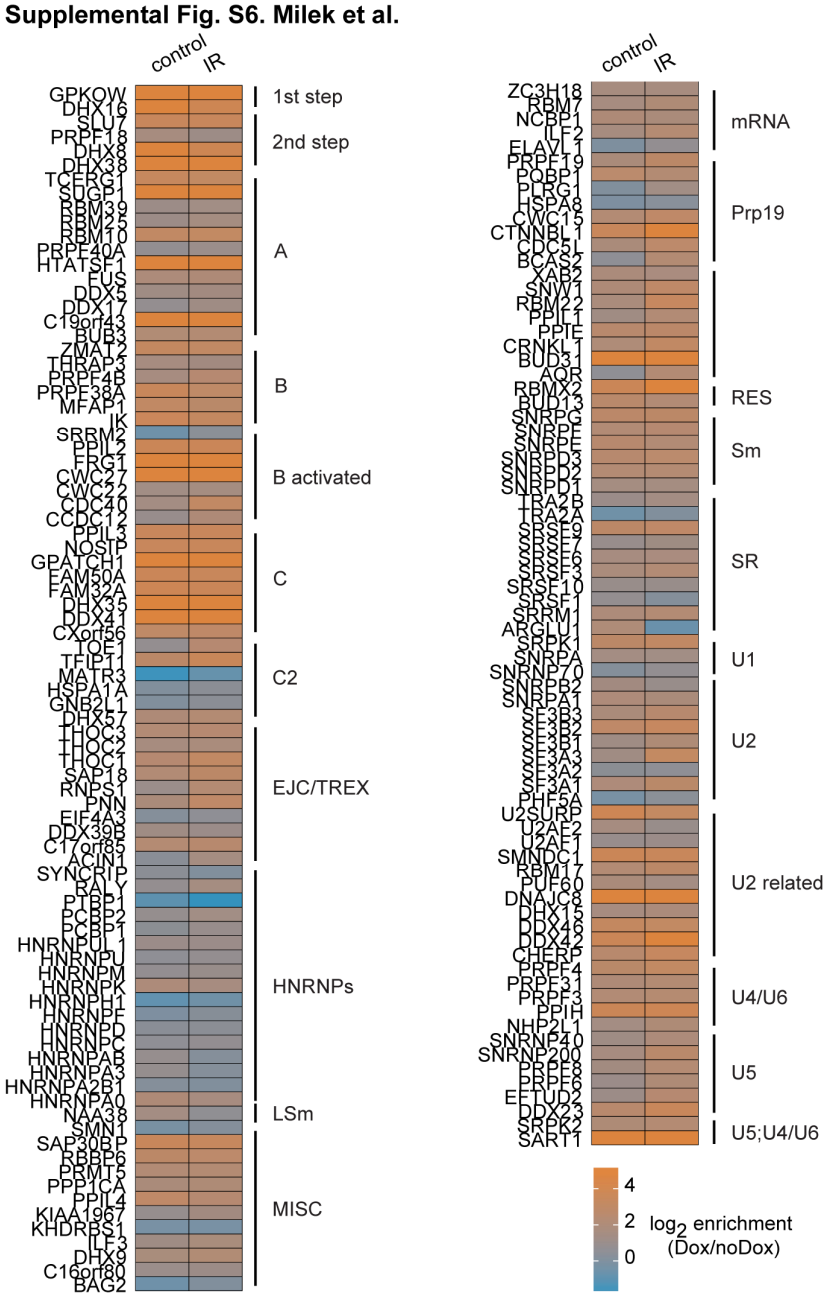
(A) Western analysis of BioID input lysates and eluates after streptavidin affinity purification. Cells were treated with doxycycline (Dox, 1 µg/ml, 16 h) or left untreated (noDox), followed by IR exposure (10 Gy, 1h) and incubated with 250 µM biotin for 3.5 hours. After cell lysis and strepatavidin affinity purifications BirA-FLAG-DDX54 was detected by anti-FLAG antibody conjugated to HRP. (B) Analysis of biotin labeling of cellular proteins was performed by probing input lysates with streptavidin-HRP. A 3.5-hour labeling time was used for the comparative control/IR experiment and compared to the 24-hour time point with 50 µM biotin incubation. (C) Total proteins detected by silver staining in eluates after streptavidin affinity purifications. Treatments were performed as described in (A). (D) Gel electrophoresis of total RNA fragments obtained from lysates of IR-exposed or untreated MCF-7 cells. Lysates were incubated in the absence (-) or presence of low (+), intermediate (++) or high (+++) RNase T1 concentration at 2.5 U/µl for 5 min at 22 °C, 5 U/µl for 5 min at 37 °C and 50 U/µl of RNAse T1 for 5 min at 37 °C, respectively. RNA was then extracted and analyzed on a 1% agarose gel.

Supplemental Fig.S5. Milek et al.



Supplemental Fig. S6.

Heatmap of log_2-transformed enrichment values in control and IR exposed (10 Gy) cells for a subset of BioID-detected proteins categorized by spliceosomal subcomplex according to (Hegele et al. 2012).

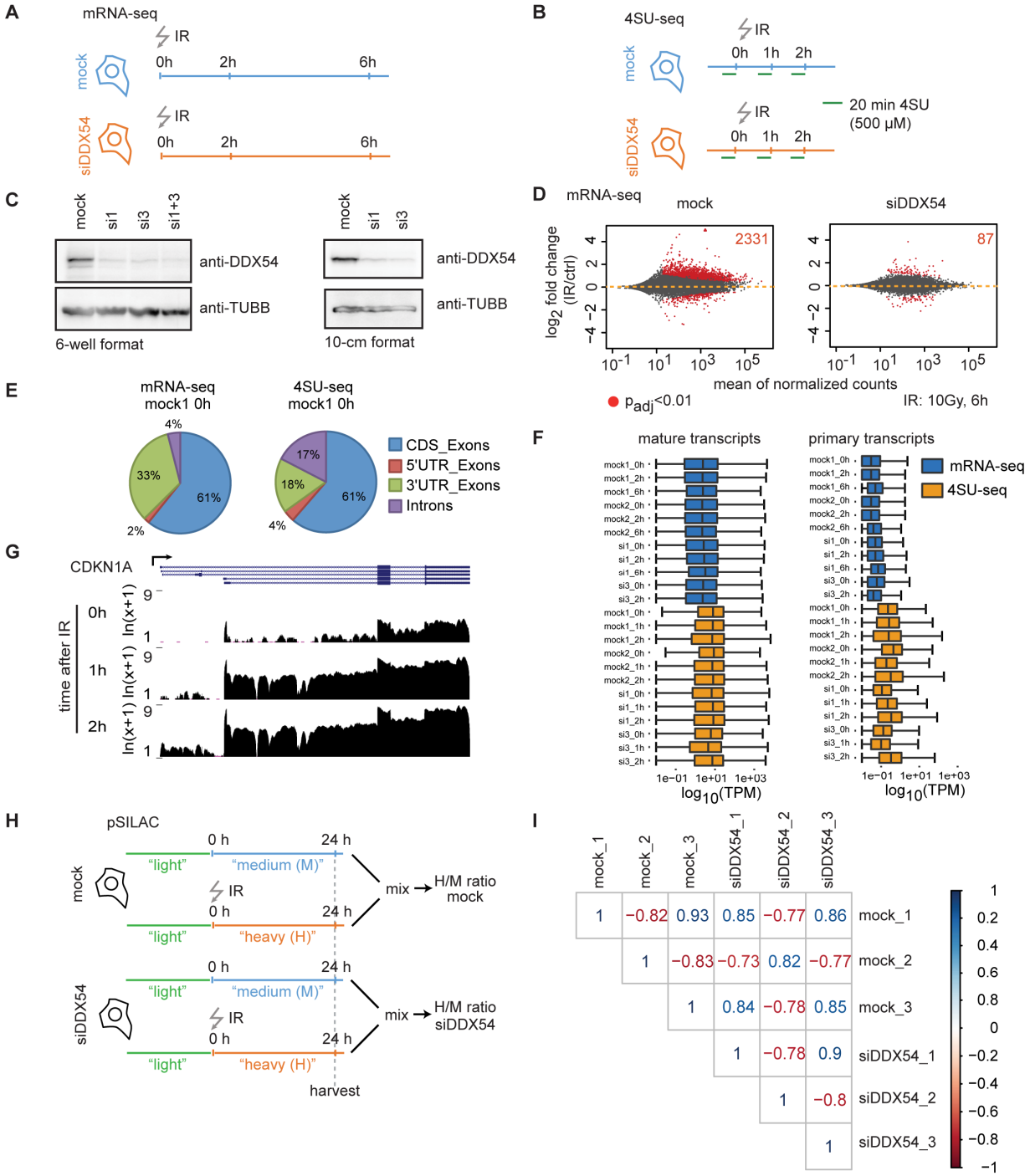


Supplemental Fig. S7.

DDX54 is required for the expression of DNA damage induced mRNAs via post-transcriptional mechanisms.

Experimental design for mRNA-seq (A) and 4SU-seq (B). DDX54 knockdowns were performed for 48 h followed by IR exposure (10 Gy) and harvesting after the indicated time periods. For 4SU-seq, cells were pulsed with 4SU (500µM) 20 min before harvesting. (C) Validation of DDX54 knockdown. MCF-7 Flp-In cells were reverse transfected as described in Supplemental Methods in a 6-well (left) or 10-cm dish format (right). Western analysis using anti-DDX54 antibody was performed. (D) Differential expression analysis (DESeq2) of mRNA-seq data between IR exposed (10 Gy, 6h) and untreated MCF-7 cells that were either mock- or siDDX54-transfected. Number of upregulated DE transcripts is indicated in upper right corners of the scatter plots (red). (E) Pie charts depicting the percentage of reads (mock 0 h, replicate 1) mapping to RefSeq transcript regions (considered regions CDS, 5'UTR 3'UTR and introns) in mRNA-seq (left) and 4SU-seq (right). (F) Quantification of primary and mature transcript abundance by mRNA-seq (blue) and 4SU-seq (orange). Box plots of all $\log_{10}$-transformed TPM values corresponding to mature (mRNA) and primary (pre-mRNA) transcripts obtained from both sequencing protocols are shown. (G) UCSC browser snapshot of *CDKN1A* locus showing the 4SU-seq read coverage upon different time periods after IR exposure. (H) Experimental design for pulsed SILAC (pSILAC) experiment. DDX54 knockdowns were performed for 48 h in the presence of SILAC "light" media. Immediately after IR exposure (10 Gy) cell cultures were switched to media containing SILAC "heavy" (IR) or "medium-heavy" (no IR) amino-acids and cells were incubated for another 24 h. MS analysis was performed to obtain normalized heavy-to-medium (H/M) ratios that reflect changes in protein synthesis upon IR vs. control. (I) Pairwise comparisons between pSILAC replicates. Pearson correlation coefficients between $\log_2$-transformed H/M ratios (3 or more ratio counts) are given. IR exposed and control samples were labeled "heavy" and "medium-heavy" for replicates 1 and 3. Replicate 2 is a label-swap experiment and anti-correlates with the other samples.

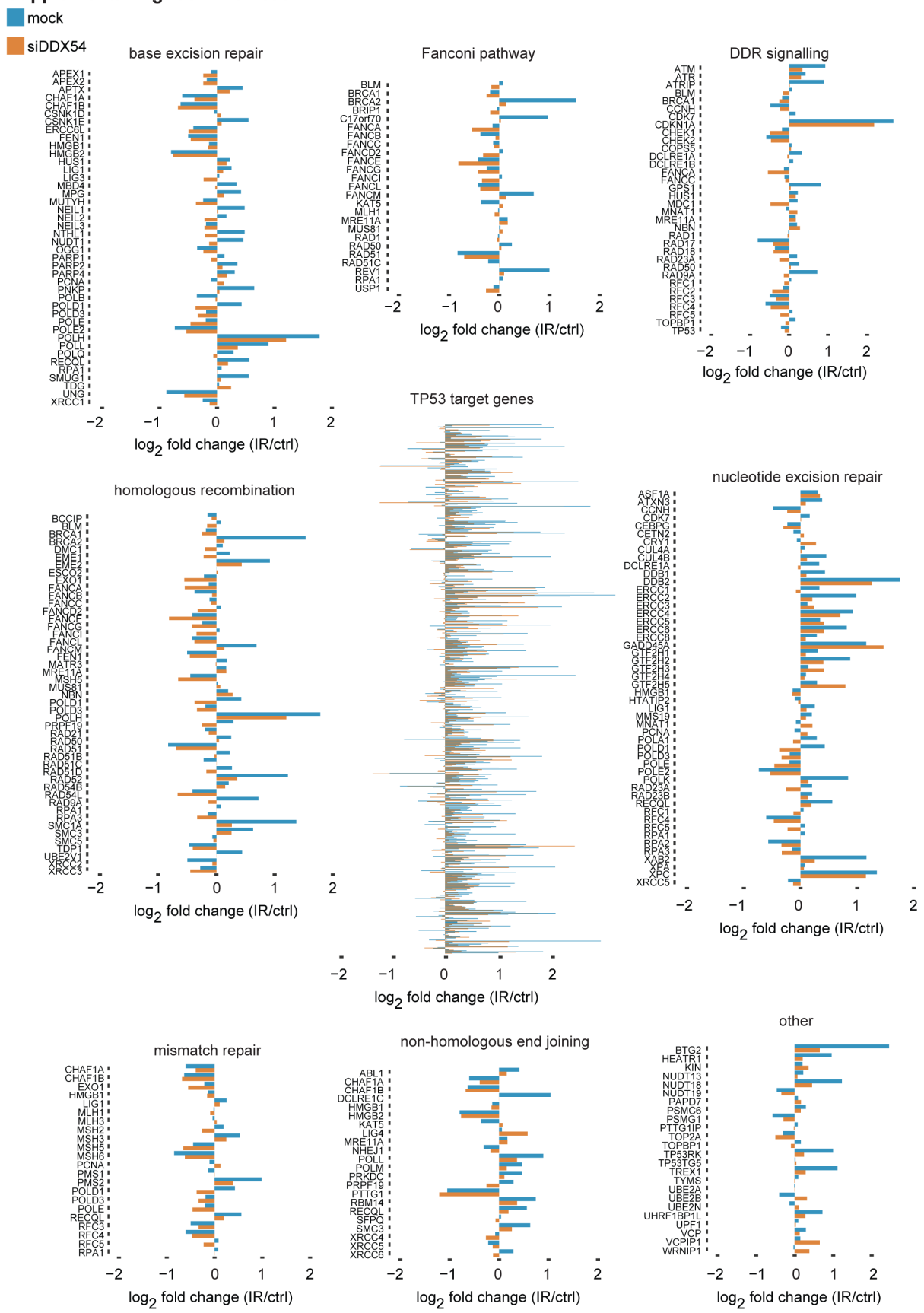
Supplemental Fig.S7. Milek et al.



Supplemental Fig. S8.

Changes in mRNA abundance for DNA damage related genes (Wood et al. 2001; Milanowska et al. 2011; Nikulenkov et al. 2012; Menendez et al. 2013; Mjelle et al. 2015) 6 h after IR exposure. DESeq2-generated $\log_2$-transformed fold changes (IR vs. control) are plotted for mock (blue) and siDDX54 (orange) conditions. A full set of TP53 target genes is shown.

Supplemental Fig. S8. Milek et al.



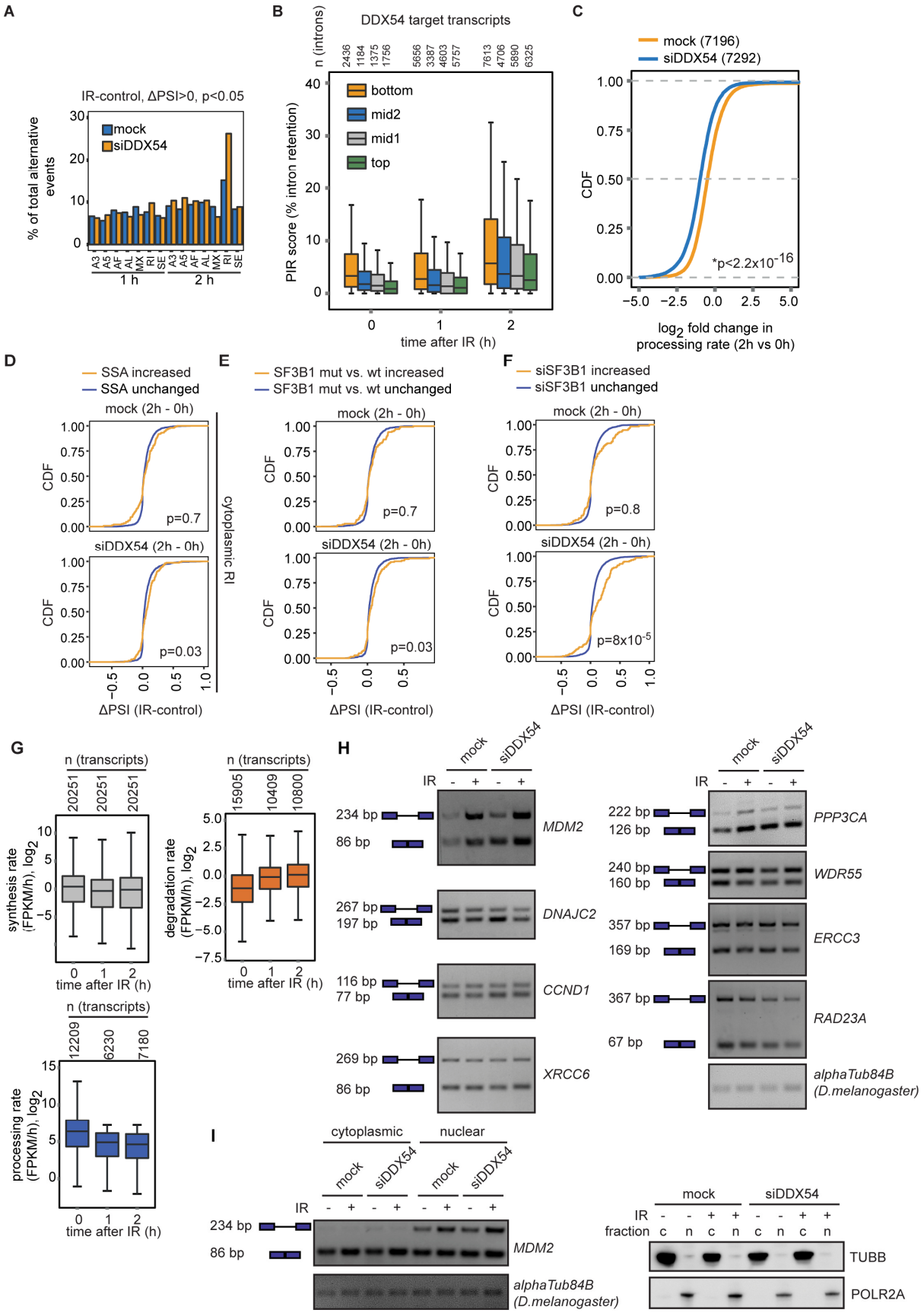
Supplemental Fig. S9.

DDX54 binding decreases IR-induced intron retention by increasing pre-mRNA processing rates of its target transcripts.

(A) Different classes of alternative events (A3, 3' alternative splice site; A5, 5' alternative splice site: AF, alternative first exons; AL, alternative last exons; MX, mutually exclusive exons; RI, retained intons; SE, skipping exons) were quantified by SUPPA (Alamancos et al. 2015) and compared between conditions (ΔPSI, IR- control). Shown are percentages of dfferential events that increase upon IR (ΔPSI>0, p<0.05) over all quantified events in a given class. (B) Box-plot of PIR scores for transcripts categorized according to expression-normalized consensus DDX54 PAR-CLIP T-C transition events. Numbers of introns considered in each group are given above the plots. (C) Cumulative density plots of log_2-transformed fold changes in processing rate 2 h vs. 0 h after IR exposure (10 Gy) are shown for mock- or siDDX54 treated cells. K-S test between both conditions was performed. (D) Comparison between spliceostatin A (SSA) and DDX54-regulated retained intron events. Introns with increased retention (adj. P <0.05, ΔPSI >0.05) or unchanged (adj. P >0.05, absolute ΔPSI <0.05) upon SSA treatment detected in the cytoplasm (Yoshimoto et al. 2017) were identified by SUPPA. Distributions of ΔPSI values (IR – control) between indicated time periods were compared for both mock and siDDX54 conditions. To assess the significant differences between RI groups Wilcoxon rank sum tests were used. (E) Retained (adj. P <0.05, ΔPSI >0.05) and unchanged (adj. P >0.05, absolute dPSI <0.05) RI events were determined in SF3B1 K700E mutant vs. wild-type background (Kesarwani et al. 2016). Distributions of ΔPSI values (IR – control) between indicated time periods were compared for both mock and siDDX54 conditions. To assess the significant differences between RI groups Wilcoxon rank sum tests were used. (F) Retained (ΔPSI>0.1) and unchanged (absolute ΔPSI<0.1) RI events were determined upon SF3B1 knockdown (Kfir et al. 2015). Distributions of ΔPSI values (IR – control) between indicated time periods were compared for both mock and siDDX54 conditions. To assess the significant differences between RI groups Wilcoxon rank sum tests were used. (G) Box-plot of log_2-transformed synthesis (red), degradation (orange) and processing (blue) rates as determined from RNA-seq read counts for total and 4SU-incorporated RNA fractions obtained after different time

periods after IR exposure (10 Gy) using INSPEcT (de Pretis et al. 2015). Numbers of considered transcripts for each time point are given above the plot. (H) RT-PCR analysis of retained intron events upon DDX54 knockdown (48 h) and IR exposure (2 h). Representative images of 2.5 % ethidium-bromide stained agarose gels are shown. Expected lengths of unspliced and spliced products are given. (I) *(Left)* Subcellular fractionation was performed according to (Nojima et al. 2015). RT-PCR analysis of unspliced and spliced MDM2 isoforms in cytoplasmic and nuclear fractions of MCF-7 cells that were either mock- or siDDX54- transfected (72 h) and/or IR exposed (10 Gy). Cells were harvested 2h post IR exposure. *(Right)* Western analysis of cytoplasmic and nuclear fractions, detecting specific cytoplasmic (TUBB, beta-tubulin) and nuclear markers (POLR2A, RNA Pol II).

Supplemental Fig.S9. Milek et al.

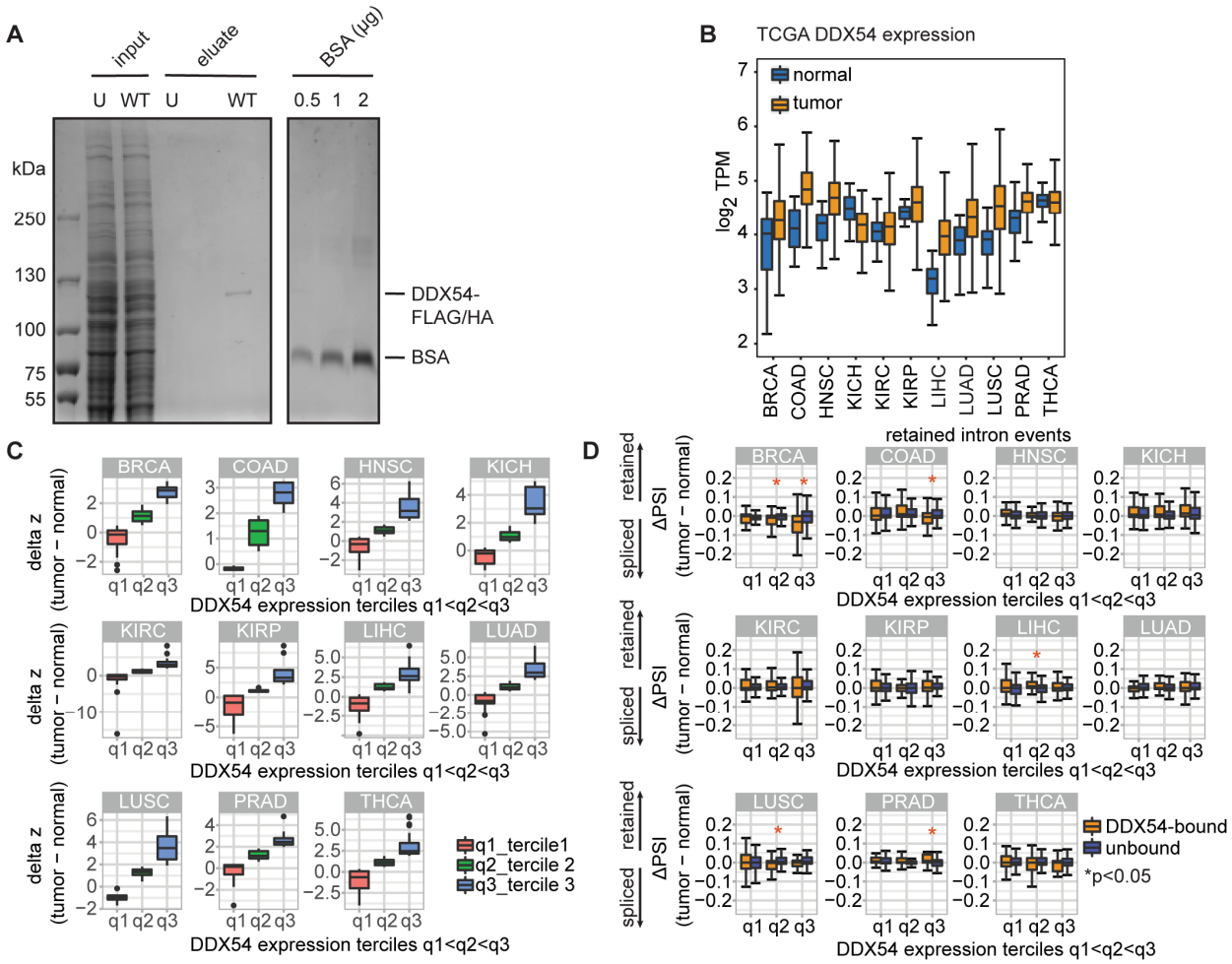


Supplemental Fig. S10.

DDX54 affects *in vitro* splicing efficacy of weak splice site-containing transcripts and promotes cell survival. Impact of DDX54 expression on splicing of introns in breast and colon tumor tissues.

(A) To purify DDX54-FLAG/HA HEK293 Flp-In T-Rex cells were transiently transfected with pcDNA5/FRT/TO-DDX54-FLAG/HA and anti-FLAG immunoprecipitation was carried out, followed by FLAG peptide elution. Coomassie-staining of SDS-PAGE gel of FLAG pepetide eluate is shown. (B) Boxplot of DDX54 expression values (TPM) obtained from RNA-seq datasets from 11 normal and tumor TCGA samples. See Fig. 6E for description of tumor types. (C) Boxplots of differences in DDX54 expression z scores between matched tumor and normal samples classified into expression terciles (q1 - lowest, q2 - intermediate, q3 - high). (D) Differences between PSI values of retained intron events in DDX54-bound vs. unbound introns (based on PAR-CLIP classification) in 3 different classes of DDX54 tumor expression (based on TCGA RNA-seq expression values; q1 – lowest, q2 – intermediate, q3 –high).

Supplemental Fig. S10. Milek et al.



LIST OF SUPPLEMENTAL TABLES

Supplemental Table S1. List of poly(A)$^+$ RNA interacting proteins with increased or unchanged binding.

Supplemental Table S2. DDX54 PAR-CLIP alignment statistics.

Supplemental Table S3. Exonic and intronic features used for random forest classifier.

Supplemental Table S4. List of BioID-detected proteins.

Supplemental Table S5. Primary and mature transcript TPMs, read counts and FPKMs.

Supplemental Table S6. List of pSILAC quantified proteins.

Supplemental Table S7. Percent intron retention (PIR) scores determined from 4SU-seq.

Supplemental Table S8. Synthesis, processing and degradation rates.

Supplemental Table S9. List of antibodies.

REFERENCES

Alamancos GP, Pages A, Trincado JL, Bellora N, Eyras E. 2015. Leveraging transcript quantification for fast computation of alternative splicing profiles. *Rna* **21**: 1521-1531.

Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M. 2005. Nucleolar proteome dynamics. *Nature* **433**: 77-83.

Braunschweig U, Barbosa-Morais NL, Pan Q, Nachman EN, Alipanahi B, Gonatopoulos-Pournatzis T, Frey B, Irimia M, Blencowe BJ. 2014. Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res* **24**: 1774-1786.

Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. 2003. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* **31**: 3568-3571.

Couzens AL, Knight JD, Kean MJ, Teo G, Weiss A, Dunham WH, Lin ZY, Bagshaw RD, Sicheri F, Pawson T et al. 2013. Protein interaction network of the mammalian Hippo pathway reveals mechanisms of kinase-phosphatase interactions. *Sci Signal* **6**: rs15.

Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**: 1367-1372.

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **10**: 1794-1805.

de Pretis S, Kress T, Morelli MJ, Melloni GE, Riva L, Amati B, Pelizzola M. 2015. INSPEcT: a computational tool to infer mRNA synthesis, processing and degradation dynamics from RNA- and 4sU-seq time course experiments. *Bioinformatics* **31**: 2829-2835.

Eddy SR. 2004. What is a hidden Markov model? *Nat Biotechnol* **22**: 1315-1316.

Graham FL, van der Eb AJ. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**: 456-467.

Gregersen LH, Schueler M, Munschauer M, Mastrobuoni G, Chen W, Kempa S, Dieterich C, Landthaler M. 2014. MOV10 Is a 5' to 3' RNA helicase contributing to UPF1 mRNA target degradation by translocation along 3' UTRs. *Mol Cell* **54**: 573-585.

Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jr., Jungkamp AC, Munschauer M et al. 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**: 129-141.

Hafner M, Renwick N, Farazi TA, Mihailovic A, Pena JT, Tuschl T. 2012. Barcoded cDNA library preparation for small RNA profiling by next-generation sequencing. *Methods* **58**: 164-170.

Hegele A, Kamburov A, Grossmann A, Sourlis C, Wowro S, Weimann M, Will CL, Pena V, Luhrmann R, Stelzl U. 2012. Dynamic protein-protein interaction wiring of the human spliceosome. *Mol Cell* **45**: 567-580.

Jens M. 2016. A Pipeline for PAR-CLIP Data Analysis. *Methods Mol Biol* **1358**: 197-207.

Jordan M, Schallhorn A, Wurm FM. 1996. Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucleic Acids Res* **24**: 596-601.

Kesarwani AK, Ramirez O, Gupta AK, Yang X, Murthy T, Minella AC, Pillai MM. 2016. Cancer-associated SF3B1 mutants recognize otherwise inaccessible cryptic 3' splice sites within RNA secondary structures. *Oncogene*.

Kfir N, Lev-Maor G, Glaich O, Alajem A, Datta A, Sze SK, Meshorer E, Ast G. 2015. SF3B1 association with chromatin determines splicing outcomes. *Cell Rep* **11**: 618-629.

Lebedeva S, Jens M, Theil K, Schwanhausser B, Selbach M, Landthaler M, Rajewsky N. 2011. Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. *Mol Cell* **43**: 340-352.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078-2079.

Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**: 26.

Menendez D, Nguyen TA, Freudenberg JM, Mathew VJ, Anderson CW, Jothi R, Resnick MA. 2013. Diverse stresses dramatically alter genome-wide p53 binding and transactivation landscape in human cancer cells. *Nucleic Acids Res* **41**: 7286-7301.

Milanowska K, Krwawicz J, Papaj G, Kosinski J, Poleszak K, Lesiak J, Osinska E, Rother K, Bujnicki JM. 2011. REPAIRtoire--a database of DNA repair pathways. *Nucleic Acids Res* **39**: D788-792.

Mjelle R, Hegre SA, Aas PA, Slupphaug G, Drablos F, Saetrom P, Krokan HE. 2015. Cell cycle regulation of human DNA repair and chromatin remodeling genes. *DNA Repair (Amst)* **30**: 53-67.

Moore HM, Bai B, Boisvert FM, Latonen L, Rantanen V, Simpson JC, Pepperkok R, Lamond AI, Laiho M. 2011. Quantitative proteomics and dynamic imaging of the nucleolus reveal distinct responses to UV and ionizing radiation. *Mol Cell Proteomics* **10**: M111 009241.

Nikulenkov F, Spinnler C, Li H, Tonelli C, Shi Y, Turunen M, Kivioja T, Ignatiev I, Kel A, Taipale J et al. 2012. Insights into p53 transcriptional function via genome-wide chromatin occupancy and gene expression analysis. *Cell Death Differ* **19**: 1992-2002.

Nojima T, Gomes T, Grosso AR, Kimura H, Dye MJ, Dhir S, Carmo-Fonseca M, Proudfoot NJ. 2015. Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell* **161**: 526-540.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841-842.

Rappsilber J, Ishihama Y, Mann M. 2003. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* **75**: 663-670.

Schueler M, Munschauer M, Gregersen LH, Finzel A, Loewer A, Chen W, Landthaler M, Dieterich C. 2014. Differential protein occupancy profiling of the mRNA transcriptome. *Genome Biol* **15**: R15.

Sebestyen E, Singh B, Minana B, Pages A, Mateo F, Pujana MA, Valcarcel J, Eyras E. 2016. Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks. *Genome Res* **26**: 732-744.

Slobodin B, Han R, Calderone V, Vrielink JA, Loayza-Puch F, Elkon R, Agami R. 2017. Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-adenosine Methylation. *Cell* **169**: 326-337 e312.

Spitzer J, Hafner M, Landthaler M, Ascano M, Farazi T, Wardle G, Nusbaum J, Khorshid M, Burger L, Zavolan M et al. 2014. PAR-CLIP (Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation): a step-by-step protocol to the transcriptome-wide identification of binding sites of RNA-binding proteins. *Methods Enzymol* **539**: 113-161.

Wessel D, Flugge UI. 1984. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* **138**: 141-143.

Wood RD, Mitchell M, Sgouros J, Lindahl T. 2001. Human DNA repair genes. *Science* **291**: 1284-1289.

Yoshimoto R, Kaida D, Furuno M, Burroughs AM, Noma S, Suzuki H, Kawamura Y, Hayashizaki Y, Mayeda A, Yoshida M. 2017. Global analysis of pre-mRNA subcellular localization following splicing inhibition by spliceostatin A. *Rna* **23**: 47-57.