Supplemental Figures to:

Systematic Analysis of A-to-I RNA Editing Upon Release of ADAR from the Nucleolus

Ruben Lattuca[1,2], Rümeyza Bascetin[1], Vincent Detours[2,#], and Denis L.J. Lafontaine[1,#]

[1]RNA Molecular Biology, Fonds de la Recherche Scientifique (F.R.S./FNRS), Université libre de Bruxelles, Biopark campus, B-6041, Belgium
[2]IRIBHM Jacques E. Dumont, Université libre de Bruxelles, CP602 808 route de Lennik, B-1070 Brussels, Belgium
[#] equal contribution, co-lead

CORRESPONDENCE
vincent.detours@ulb.be
denis.lafontaine@ulb.be

ORCID ID
Denis Lafontaine 0000-0001-7295-6288

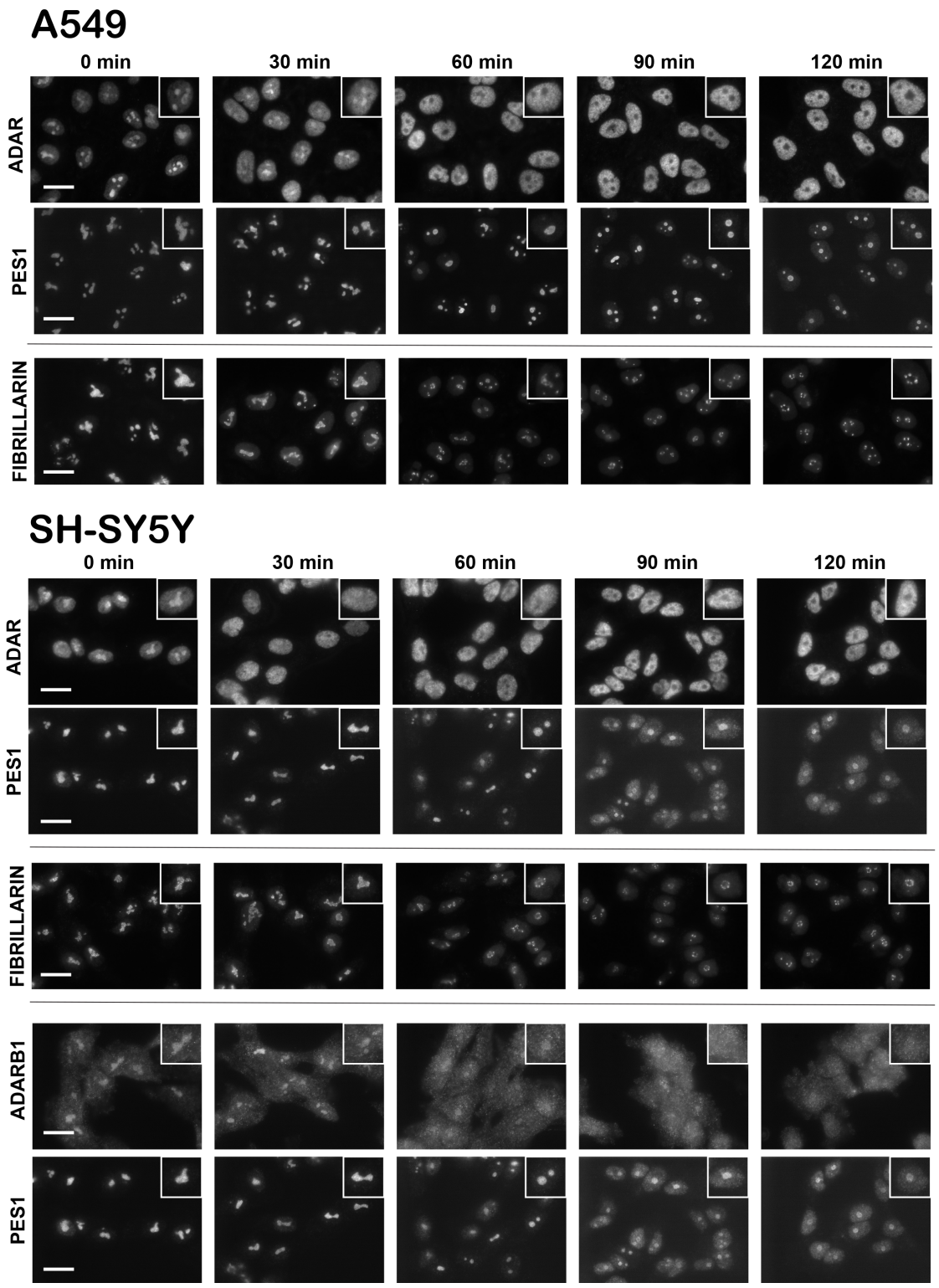

Figure S1: Extended version of rationale with additional colocalization studies

A549 or SH-SY5Y cells treated with Act D for the indicated times were immuno-stained with antibodies specific to ADAR, ADARB1, fibrillarin, or PES1 and imaged by fluorescence microscopy. Scale bar, 5 µm.

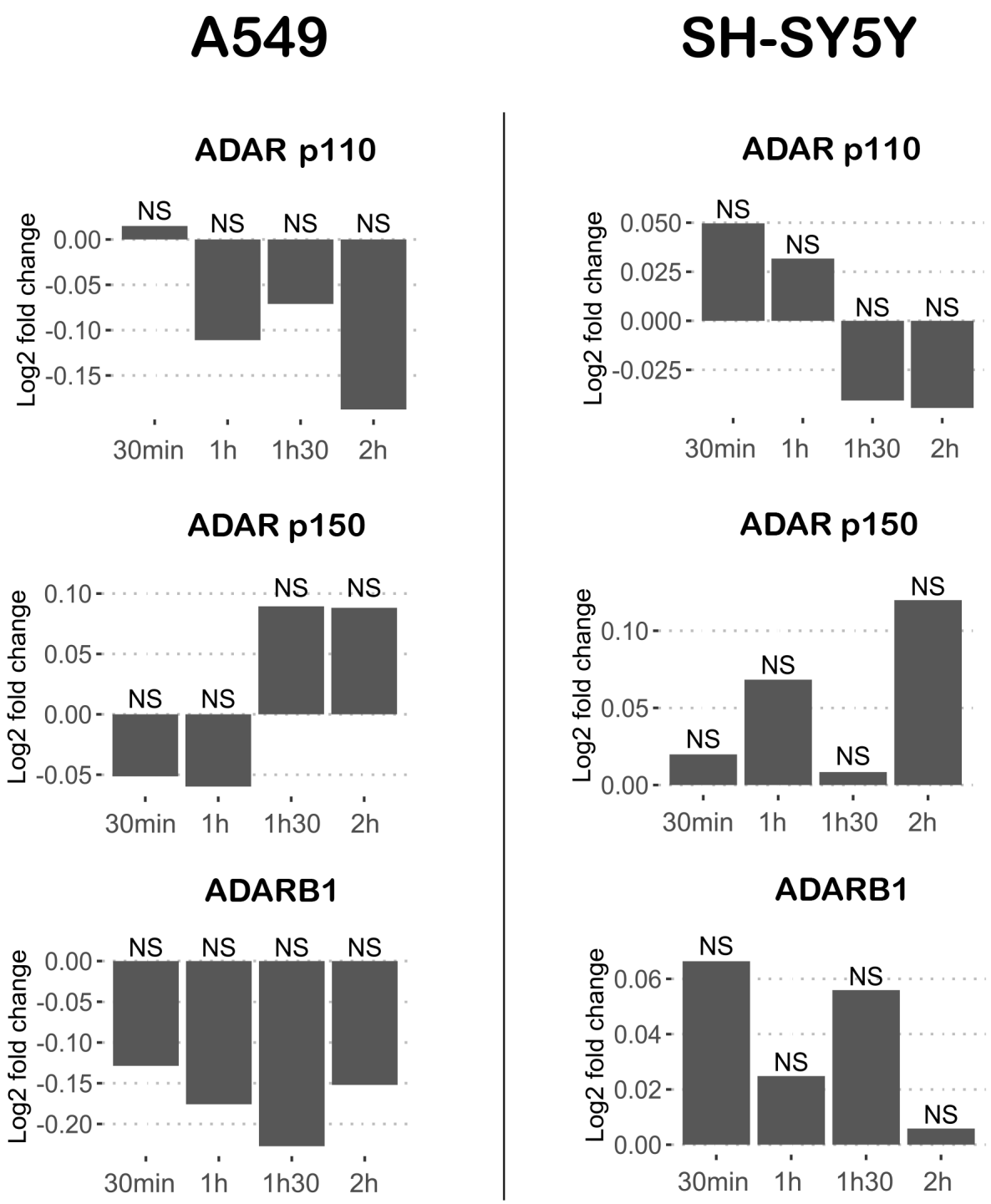

Figure S2: ADAR isoform and ADARB1 expression do not change during actinomycin D treatment

Differential expression of ADAR p110, ADAR p150 and ADARB1 between control and treated cells. None of the expression variations of ADAR isoforms or ADARB1 are significant (all *q*-values > 0.1). NS, non-significant.

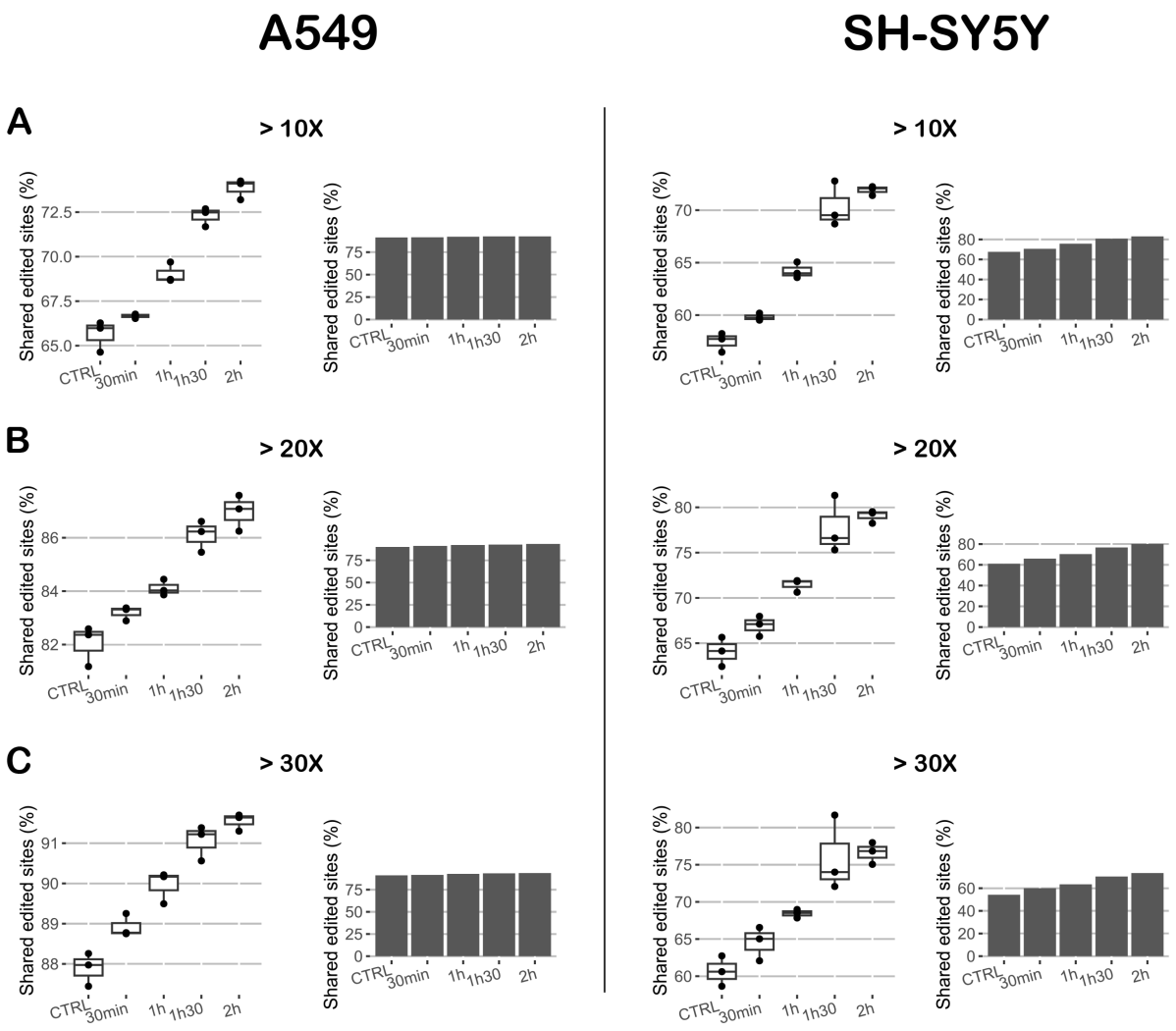

Figure S3: Percentage of shared edited sites in our dataset and the REDIportal editing event database

The REDIportal database contains a total of 15,683,856 editing events. Boxplots show the percentage of common edited sites between our dataset and the REDIportal database, as well as across replicates in A549 and SH-SY5Y cells, at coverage thresholds of > 10x (**A**), > 20x (**B**) and > 30x (**C**).

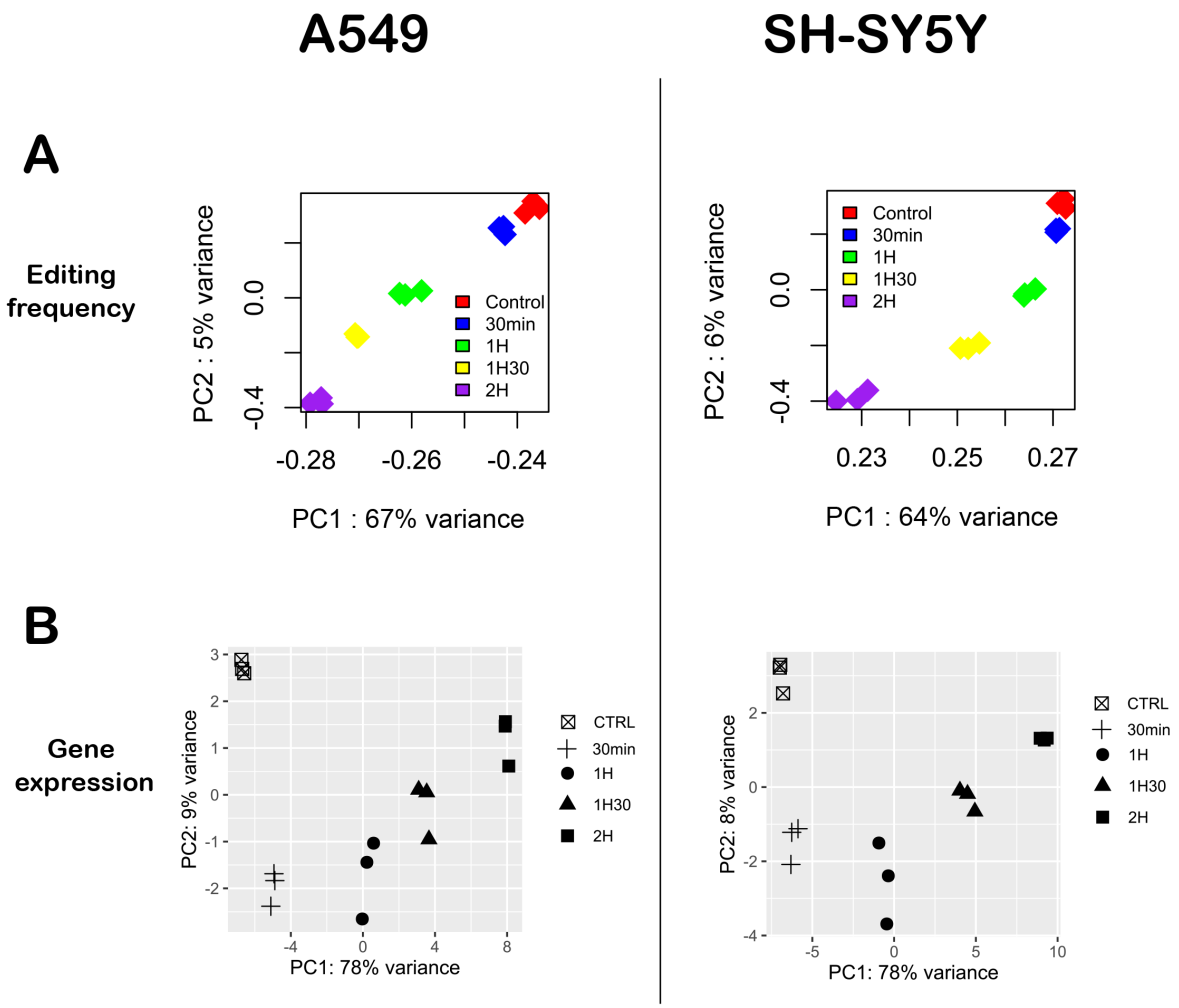

Figure S4: Consistency of gene expression and editing frequency changes between replicates

Principal component analysis of editing frequency (**A**) and gene expression (**B**) in the A549 and SH-SY5Y cell lines.

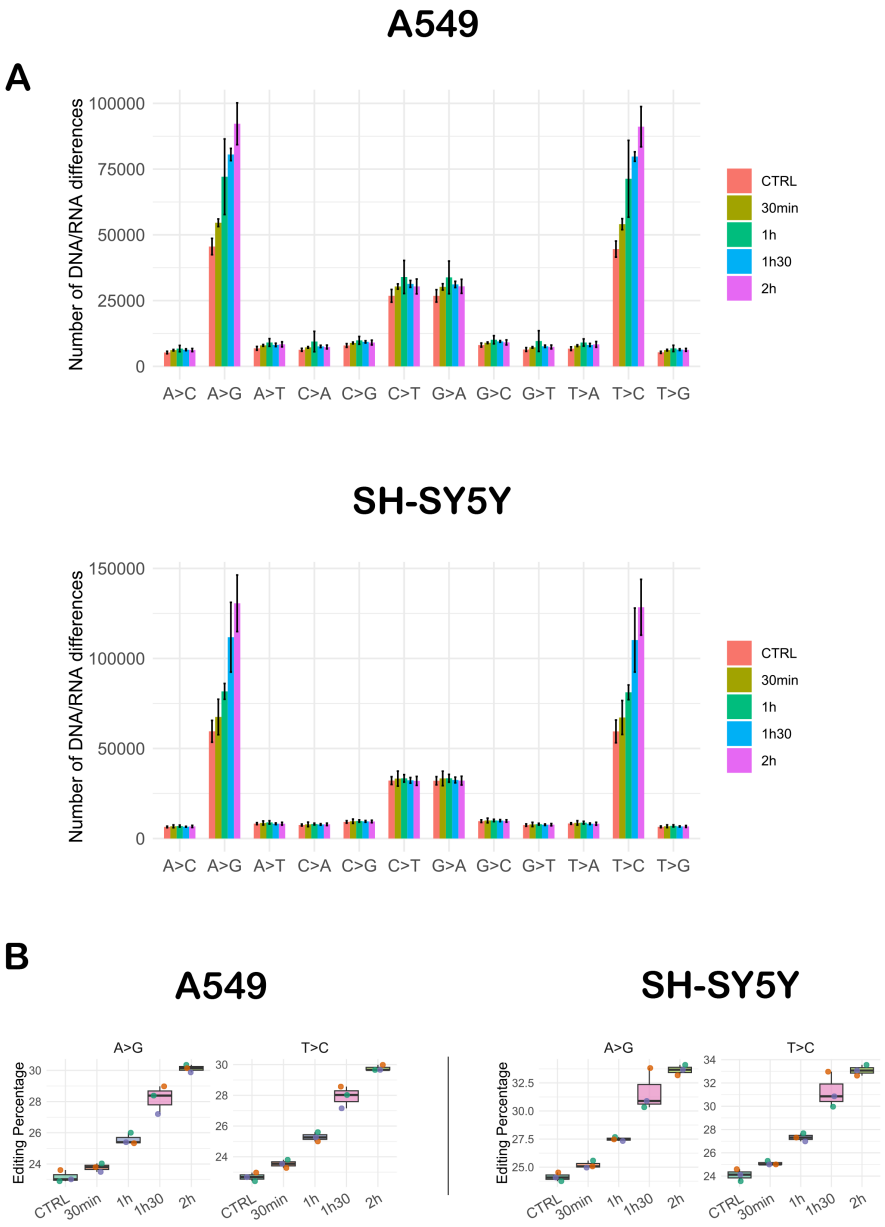

Figure S5: Distribution of all 12 DNA/RNA differences over actinomycin treatment time

A, Distribution of the 12 DNA/RNA differences computed for each time-point in A549 (left) and SH-SY5Y (right). DNA/RNA differences were selected with a ≥ 0.1 frequency and a ≥ 10x coverage threshold in each replicate.

B, Proportion of A>G and T>C DNA/RNA differences in A549 (left) and SH-SY5Y (right) for each time point. In A549 cells, A>G and T>C proportions significantly increase from 23.17% to 30.12% and from 22.68% to 29.76% after two hours of treatment (adjusted *p*-values = 5.34×10^{-23} and 7.68×10^{-23}). In SH-SY5Y cells, A>G and T>C proportions significantly increase from 24.11% to 33.65% and from 24.1% to 33.1% after two hours of treatment (adjusted *p*-values = 1.43×10^{-73} and 6.53×10^{-48}).

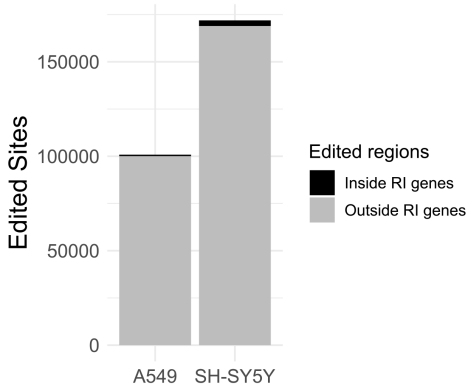

Figure S6: Edited sites are mainly located in region without differential intron retention events

Number of edited sites detected in genes with differential retained intron (RI) events (*q*-values < 0.05) after 2 hours of treatment. Edited sites were selected with an ≥ 0.1 editing frequency, a ≥ 10x coverage and detected in at least two replicates.

In A549 and SH-SY5Y, 749 and 2953 edited sites were detected in genes with retained intron events, accounting for only 0.75% and 1.75% (black portions of the histograms) of the total edited sites at 2 hours.

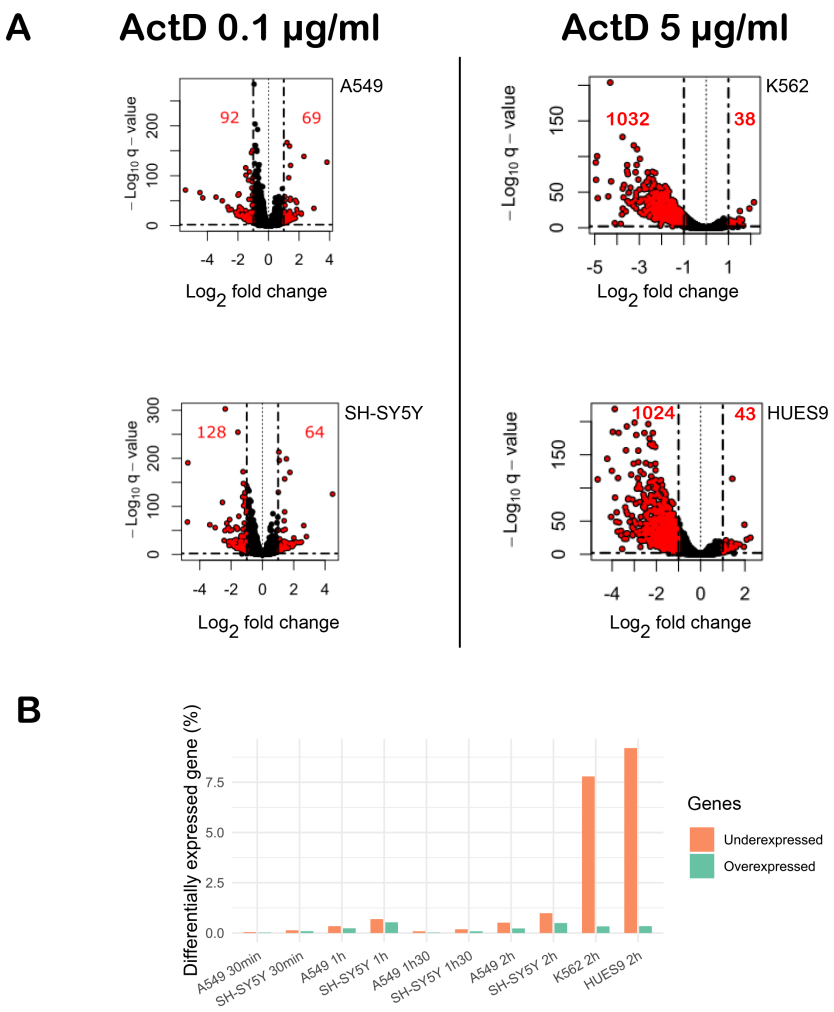

Figure S7: Impact of different doses of actinomycin D on gene expression

A, Differentially expressed genes after 2 h of treatment with either a low (0.1 µg/ml, left; this work) or high (5 µg/ml, right; Ref. [76]) concentration of Act D ($|\log_2 FC| \geq 1$, *q*-value 0.01).

At 0.1 µg/ml Act D, only 92 and 128 genes were downregulated in A549 and SH-SY5Y cells, respectively. In contrast, at 5 µg/ml Act D, 1,032 and 1,024 genes were downregulated in K562 and HUES9 cells, respectively.

B, Comparison of the proportion of downregulated and upregulated genes among all expressed genes (read count ≥ 10 per gene per sample) at each time point following treatment with 0.1 µg/ml Act D (in A549 and SH-SHY5Y cells) and after 2 h of treatment with 5 µg/ml Act D (in K562 and HUES9 cells).

After 2 h of low-dose treatment, only 0.51% and 0.98% of expressed genes were downregulated in A549 and SH-SY5Y, respectively. In contrast, 7.79% and 9.2% of expressed genes were downregulated in K562 and HUES9, respectively, following high-dose treatment.

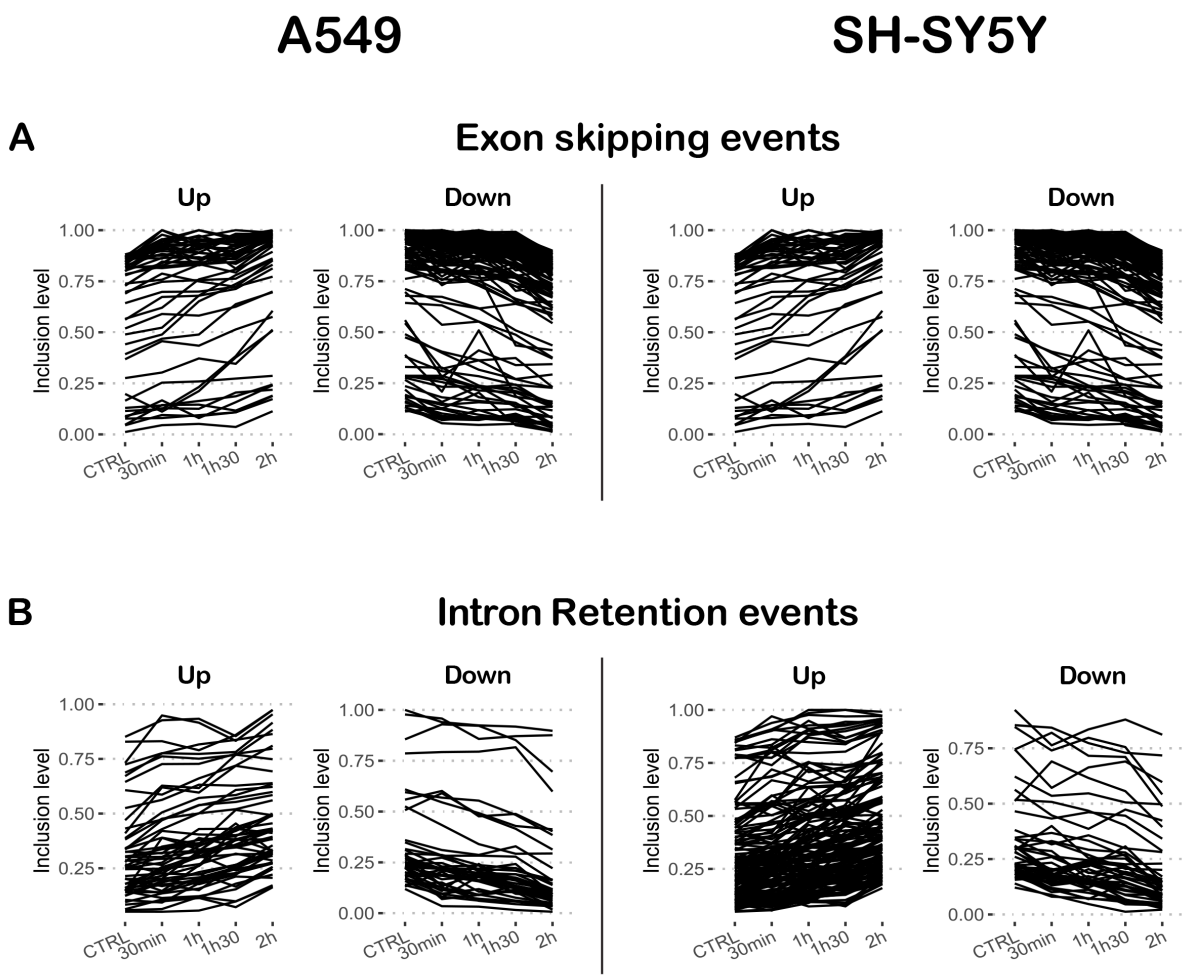

Figure S8: Aberrant splicing events occur continuously across treatment

Aberrant splicing events, either exon skipping or intron retention, are plotted as "inclusion levels" across treatment. The data illustrate, at the level of individual genes, that these effects are continuous.

A, Monotonic increase (Up) or decrease (Down) in inclusion levels for genes with exon skipping. *Up* indicates greater inclusion and less exon skipping; *Down* indicates reduced inclusion and more exon skipping.

B, Monotonic increase (Up) or decrease (Down) in inclusion levels for genes with intron retention. *Up* indicates greater inclusion and increased intron retention; *Down* indicates reduced inclusion and less intron retention.

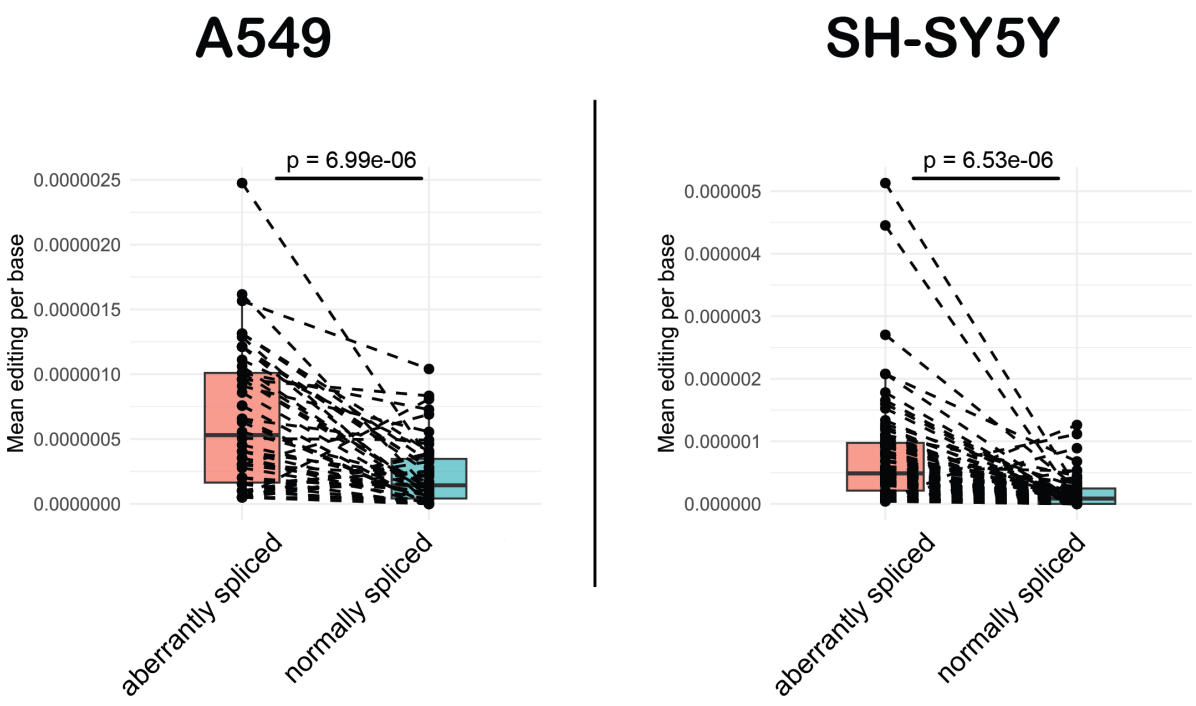

Figure S9: Increasing number of edited sites in aberrantly spliced regions within genes

Each dot represents a gene with the mean editing of its transcript in intronic regions flanking aberrantly spliced versus normally spliced exons. Transcripts of individual genes showed a significant increase in editing per base in regions flanking an aberrant exon inclusion/exclusion event (*p*-values = 6.99×10^{-6}, 6.53×10^{-6}). The number of edited sites is normalized to the summed length of both intronic regions flanking the spliced/unspliced exon.

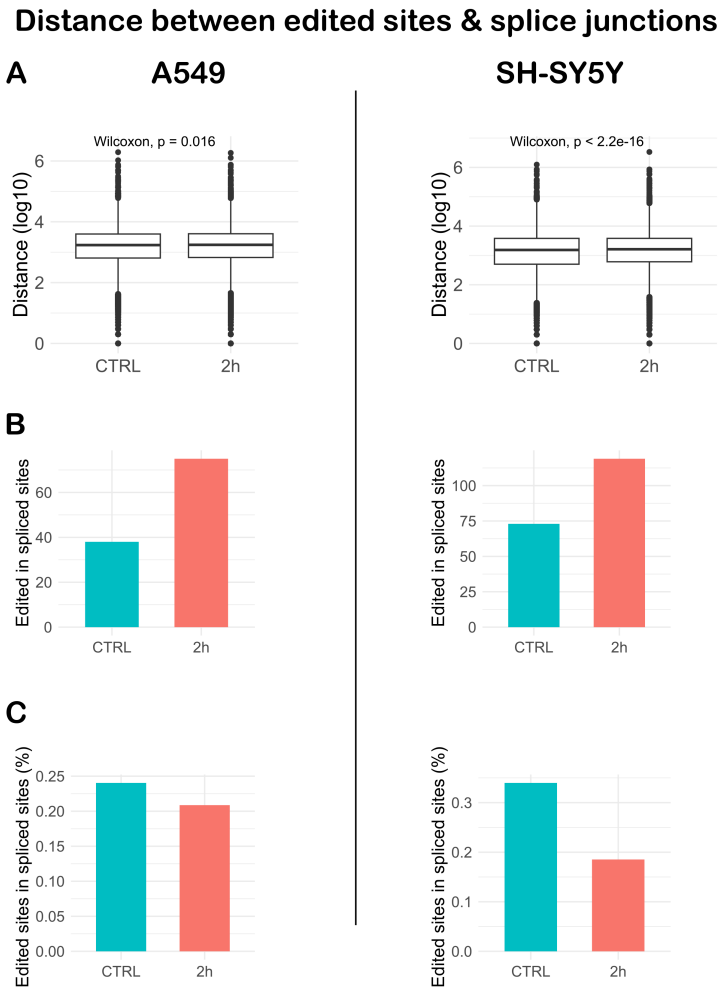

Figure S10: Assessment of the distance between edited sites and splice junctions

A, Distance between edited sites and the nearest splice junction in A549 (left) and SH-SY5Y (right). For each edited site detected in control and 2 h-treated cells (editing frequency ≥ 0.1, coverage ≥ 10x, detected in at least two replicates), we calculated the distance to the closest splice junction.

In both cell lines, the median distance to the nearest splice junction differed significantly between control and 2 h-treated cells (Wilcoxon test: A549 —CTRL median distance = 1,717 bp, 2 h = 1,749, *p*-value = 0.016; SH-SY5Y —CTRL median = 1,535 bp, 2 h = 1,624, *p*-value = 2.2×10^{-16}. However, the magnitude of the difference in median distances was small and did not indicate a shift toward the splice junction.

B-**C**, Number (**B**) and percentage (**C**) of edited sites at the two splice sites flanking the junction in A549 (left) and SH-SY5Y (right).

We observed increased editing at the splicing sites flanking the junction (**B**), but these changes appeared to reflect a global increase in editing rather than a specific alteration at the junction itself (**C**).

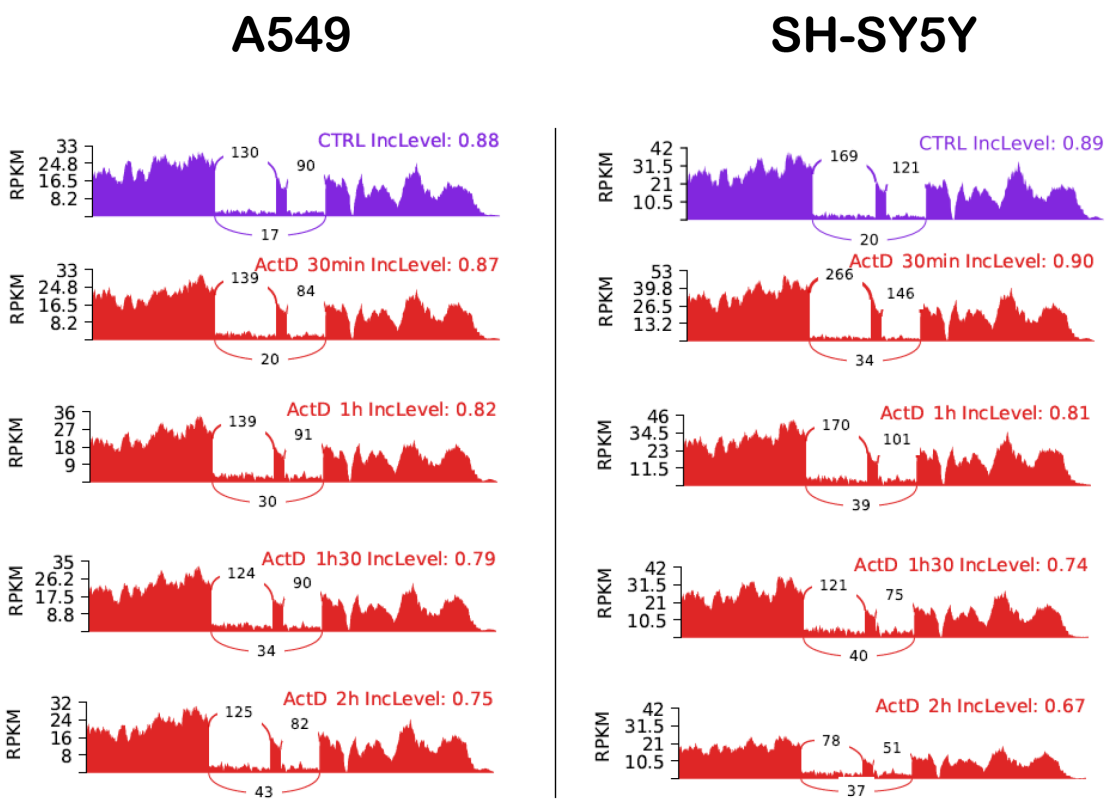

Figure S11: Gradually decreasing exon inclusion level in ZNF217

Continuous decrease in ZNF217 exon 4 inclusion in A549 and SH-SY5Y cells with treatment time. Analysis performed with rMATS.

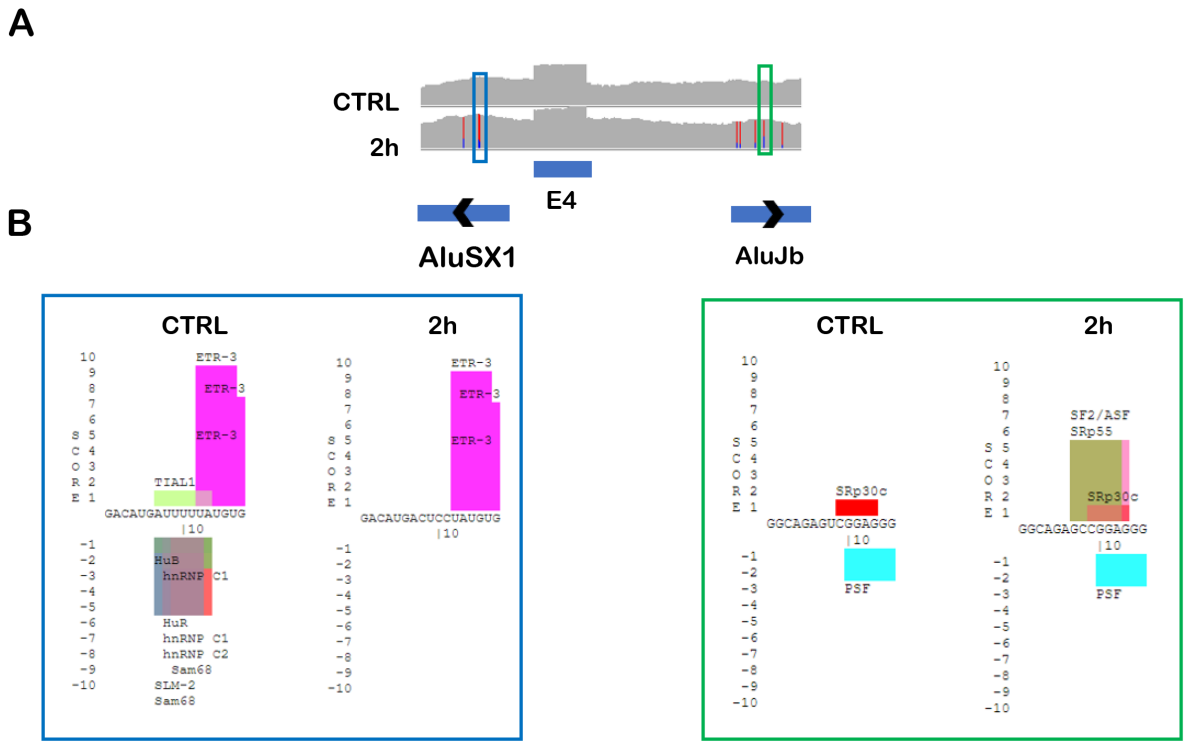

Figure S12: *De novo* editing in Alu elements flanking differentially spliced Exon 4 of ZNF217 may generate splicing via regulatory elements (SRE)

A, Genome track of alternatively spliced ZNF217 exon 4 for control and 2-h-treated samples (SH-SY5Y cells). Read coverage is displayed on a log scale. ZNF217 exon 4 is flanked by two Alu sequences (AluSX1 and AluJb), which are edited in 2-h-treated cells (blue/red sticks). Blue and green boxes depict regions where editing may alter sequence recognition (see panel **B**).

B, Analysis of the effect of Alu editing on splicing regulatory elements.

SpliceAid was used to identify splicing regulatory elements (SREs) in the Alu sequences flanking ZNF217 exon 4. The algorithm looks for motifs in a database of strictly experimentally determined target RNA sequences. SpliceAid returns positive and negative scores corresponding, respectively, to sequences promoting exon exclusion or inclusion. Editing in AluSX1 may decrease exon inclusion by repressing Sam68, hnRNP C1/C2, and HuB binding. Editing in AluJb may promote exon exclusion by inducing SF2 and SRp55 binding.

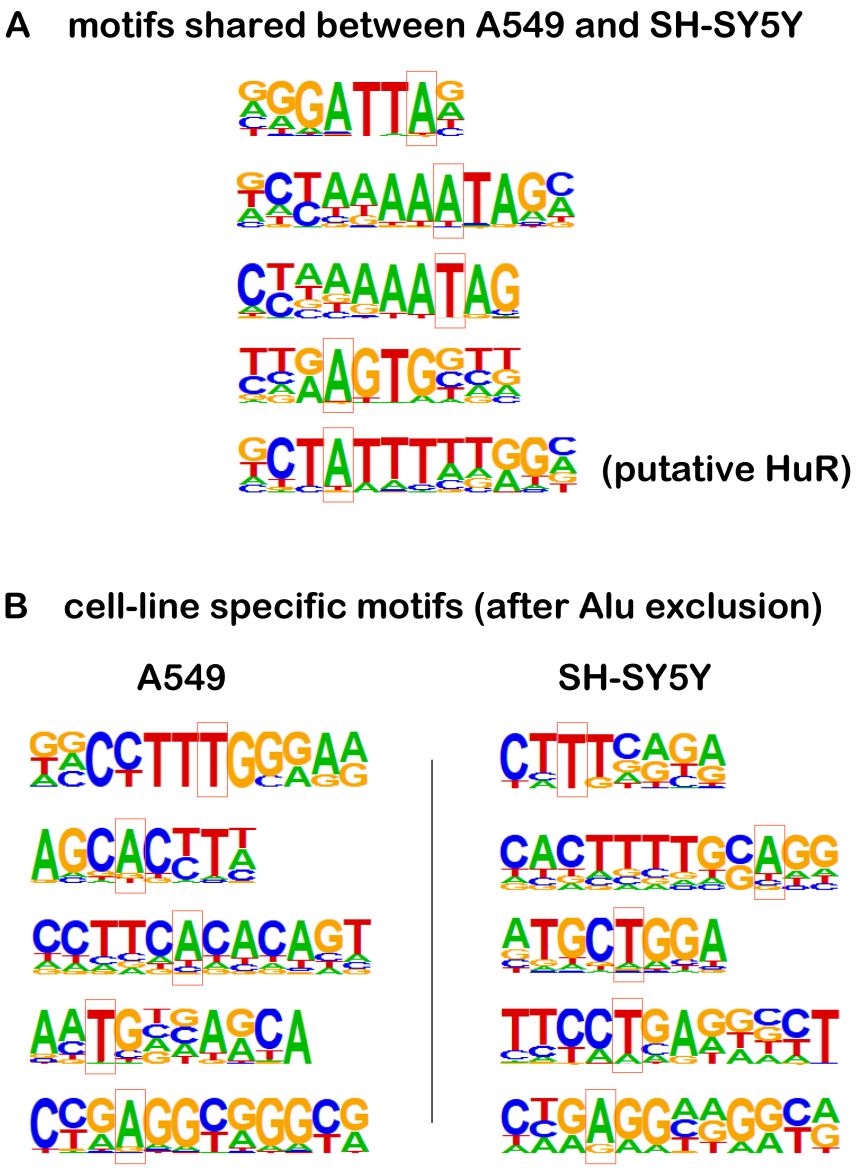

Figure S13: Detection of ADAR motif enrichment

A, Top enriched motifs in 2 h-treated A549 and SH-SY5Y cells compared to untreated cells. Enriched motifs were detected with HOMER (hypergeometric test, adjusted *p*-value < $2x10^{-16}$). Both cell lines share the same top enriched motifs enriched in Alu sequences. The edited A residue is boxed in red; when the conserved motif was on the complementary strand, a T is boxed instead.

B, Same as (**A**) but using Alu sequences as background. Top enriched motifs are not shared anymore between A549 and SH-SY5Y cell lines. The significance of these enriched motifs remains to be established.

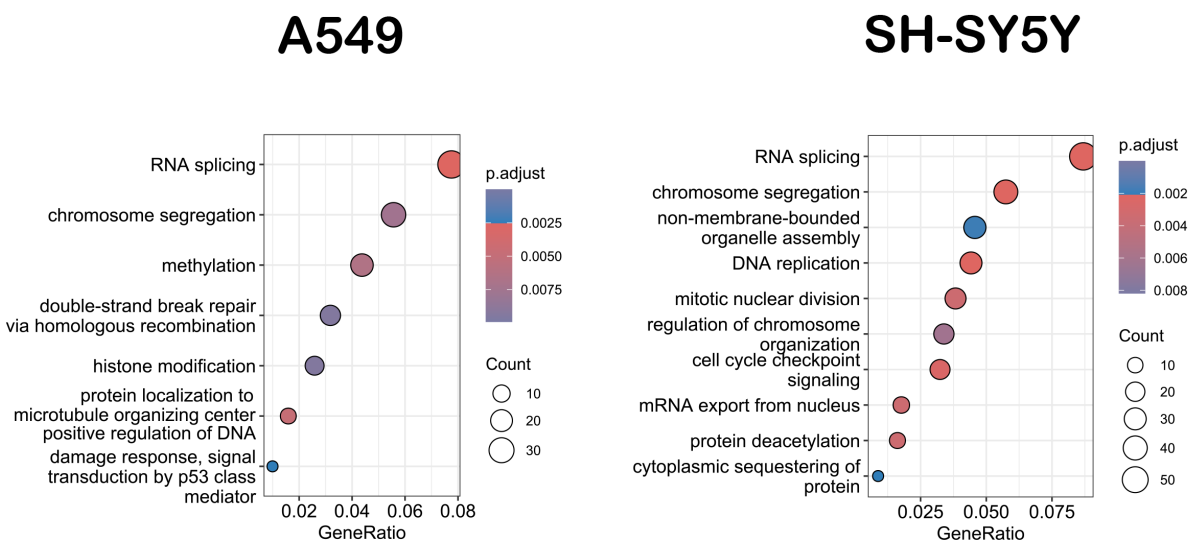

Figure S14: Pathways enrichment analysis of genes alternatively spliced upon release of ADAR from the nucleolus

ClusterProfiler was used to perform pathway enrichment analysis of genes whose transcript undergo alternative splicing events in A549 and SH-SY5Y cells upon release of ADAR from the nucleolus. The ClusterProfiler package computes a GeneRatio, corresponding to the ratio of input genes that are annotated with a GO term. A hypergeometric test was performed to assess gene set enrichment (adjusted *p*-value < 0.01). Overall, splicing functions are overrepresented among genes whose transcripts undergo ADAR nucleolar release-dependent skipping or inclusion events.

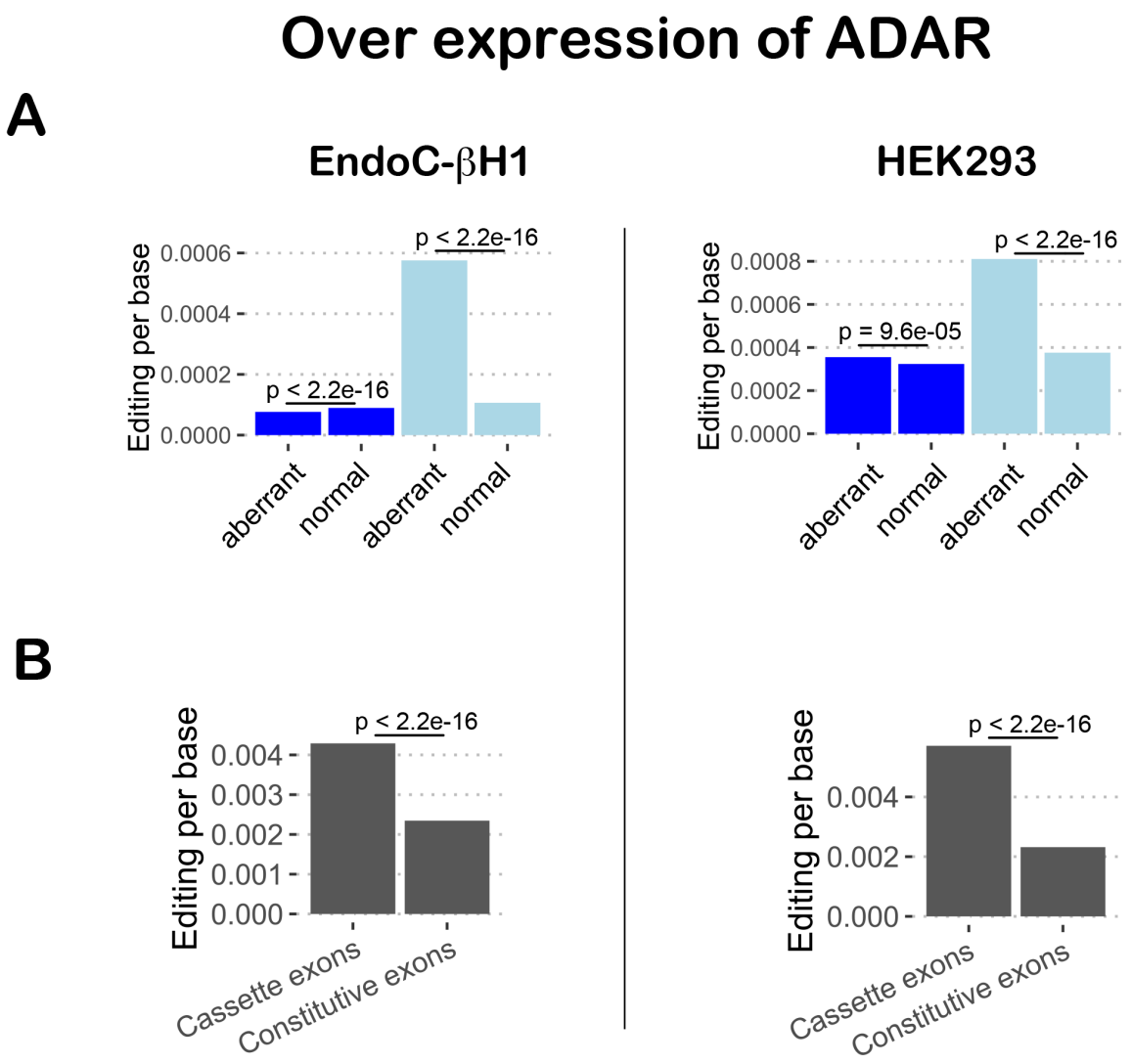

Figure S15: Upon overexpression of ADAR, editing levels increase in intronic segments flanking aberrantly spliced exons and in cassette exons

A, Same as **Fig 5B** for EndoC-βH1 and HEK293 cells overexpressing ADAR (ADAR OE).

B, Same as **Fig 5C** for EndoC-βH1 and HEK293 cells overexpressing ADAR (ADAR OE).

SUPPLEMENTAL TABLE LEGENDS

Table S1 (related to Fig 2B): edited sites dataset

All edited sites detected in 2 cell lines (A549 and SH-SY5Y), 5 time points (0, 30, 60, 90, and 120 min.), and triplicate samples (3 independent cultures)

Table S2 (related to Fig 2A): differentially edited sites dataset

2 cell lines, 4 time points (30, 60, 90, and 120 min)

Table S3 (related to Fig 4A): differentially expressed genes dataset

2 cell lines, 1 time point (2 h)

Table S4 (related to Fig 4A): differentially expressed isoforms dataset

2 cell lines, 1 time point (2 h)

Table S5 (related to Fig 5): alternative splicing events dataset

2 cell lines, 1 time point (2 h)