

Supplementary Information

The MTR4/hnRNP complex surveils aberrant polyadenylated RNAs with multiple exons

Kenzui Taniue, Anzu Sugawara, Chao Zeng, Han Han, Xinyue Gao, Yuki Shimoura, Atsuko Nakanishi Ozeki, Rena Onoguchi-Mizutani, Masahide Seki, Yutaka Suzuki, Michiaki Hamada, Nobuyoshi Akimitsu

Supplementary Figures

Supplementary Fig. 1 MTR4 destabilizes 3XTs.

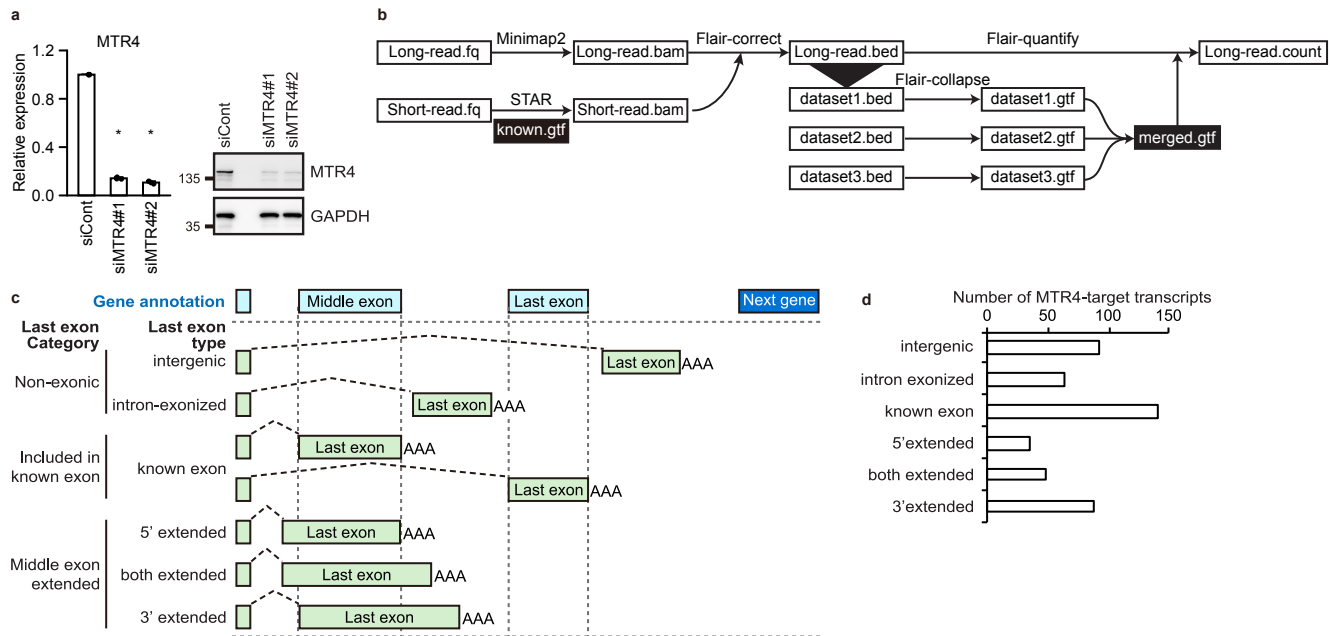
Supplementary Fig. 2 The PAXT, but not the NEXT and TRAMP complexes, regulates the expression of 3XTs.

Supplementary Fig. 3 hnRNP complex regulates the expression of multi- but not mono-exon 3XTs.

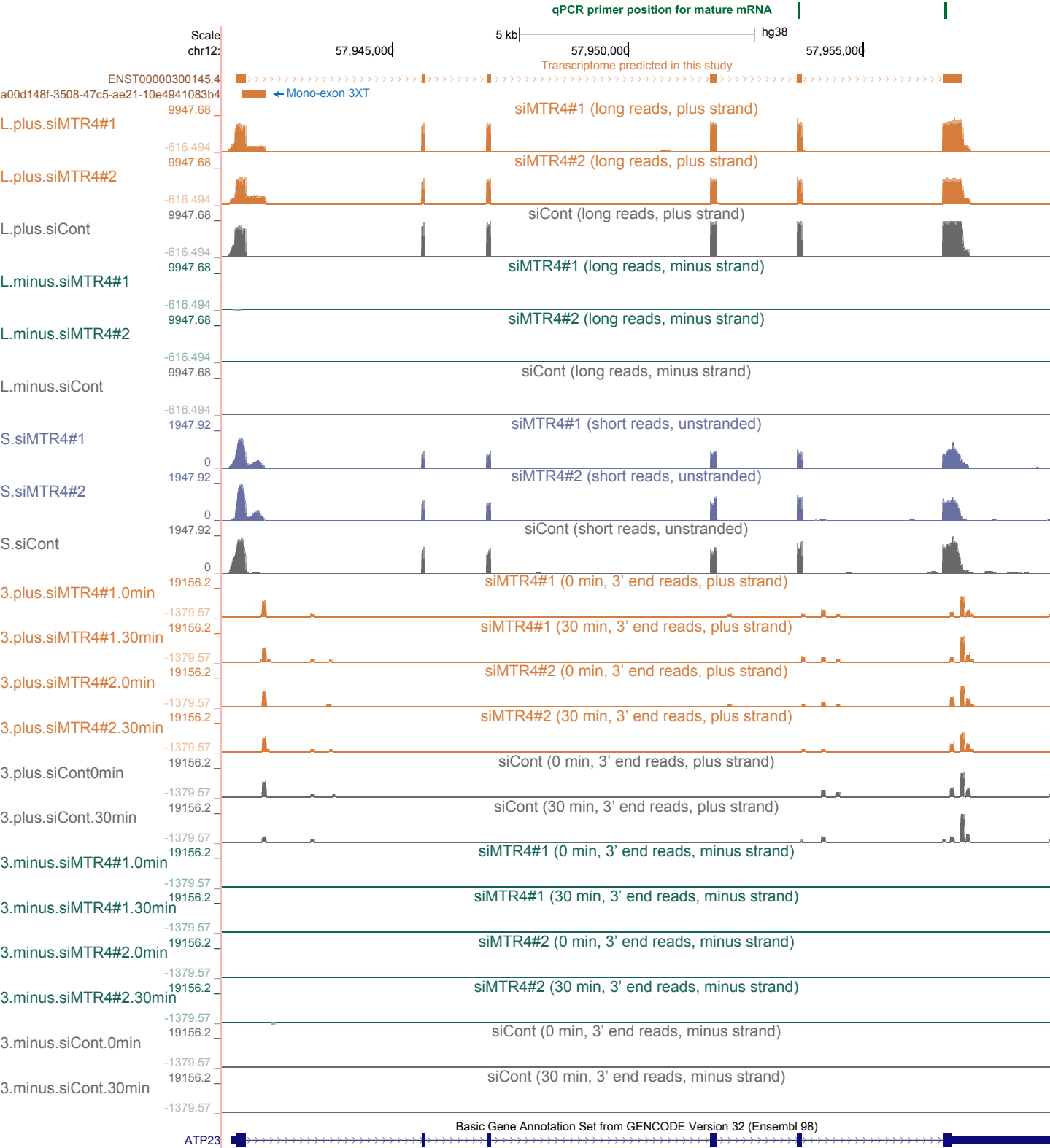
Supplementary Fig. 4 hnRNP complex recognizes multi-exon 3XTs to promote the degradation of host 3XTs.

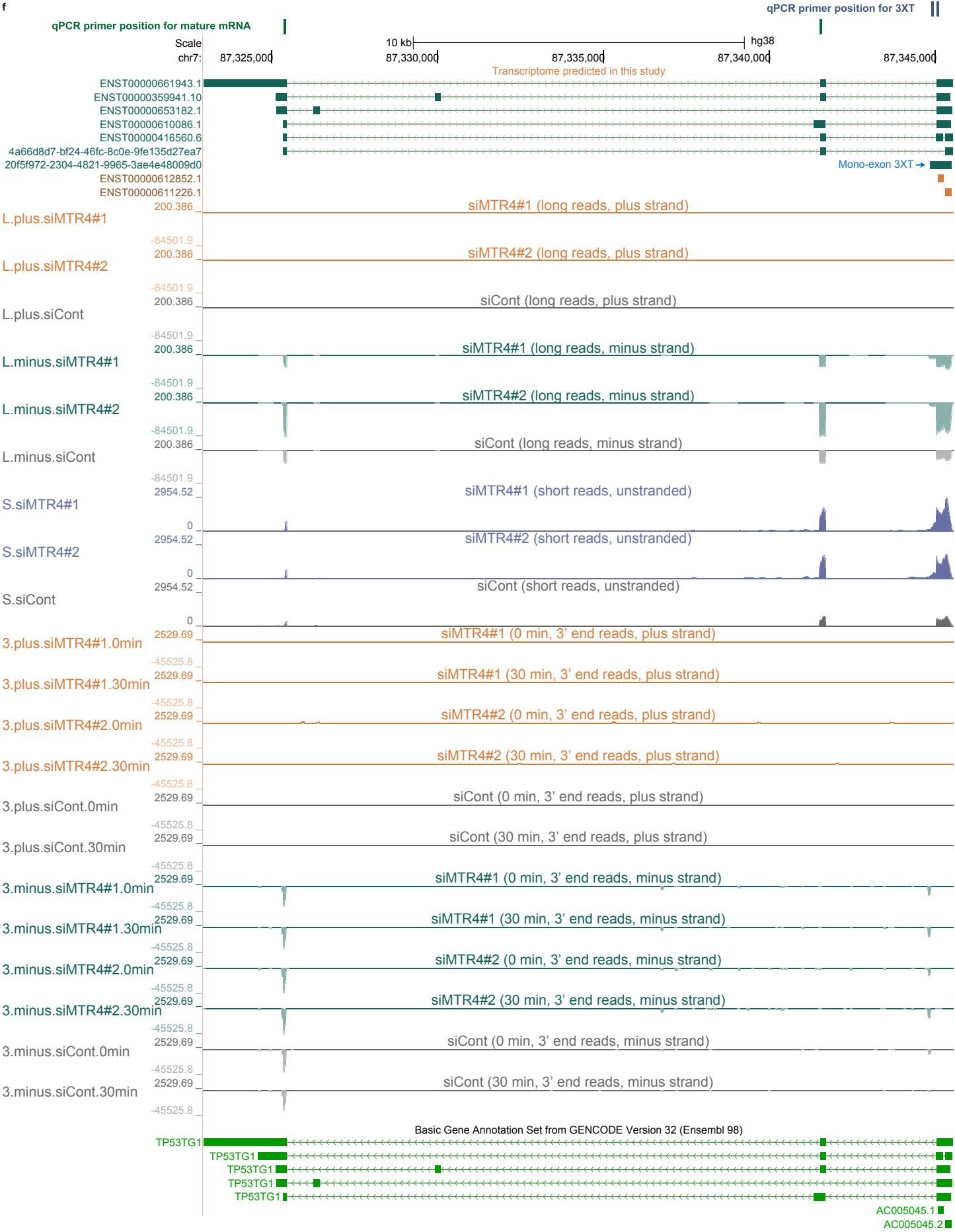
Supplementary Fig. 5 KCTD13 3XT protein forms a liquid-like assembly.

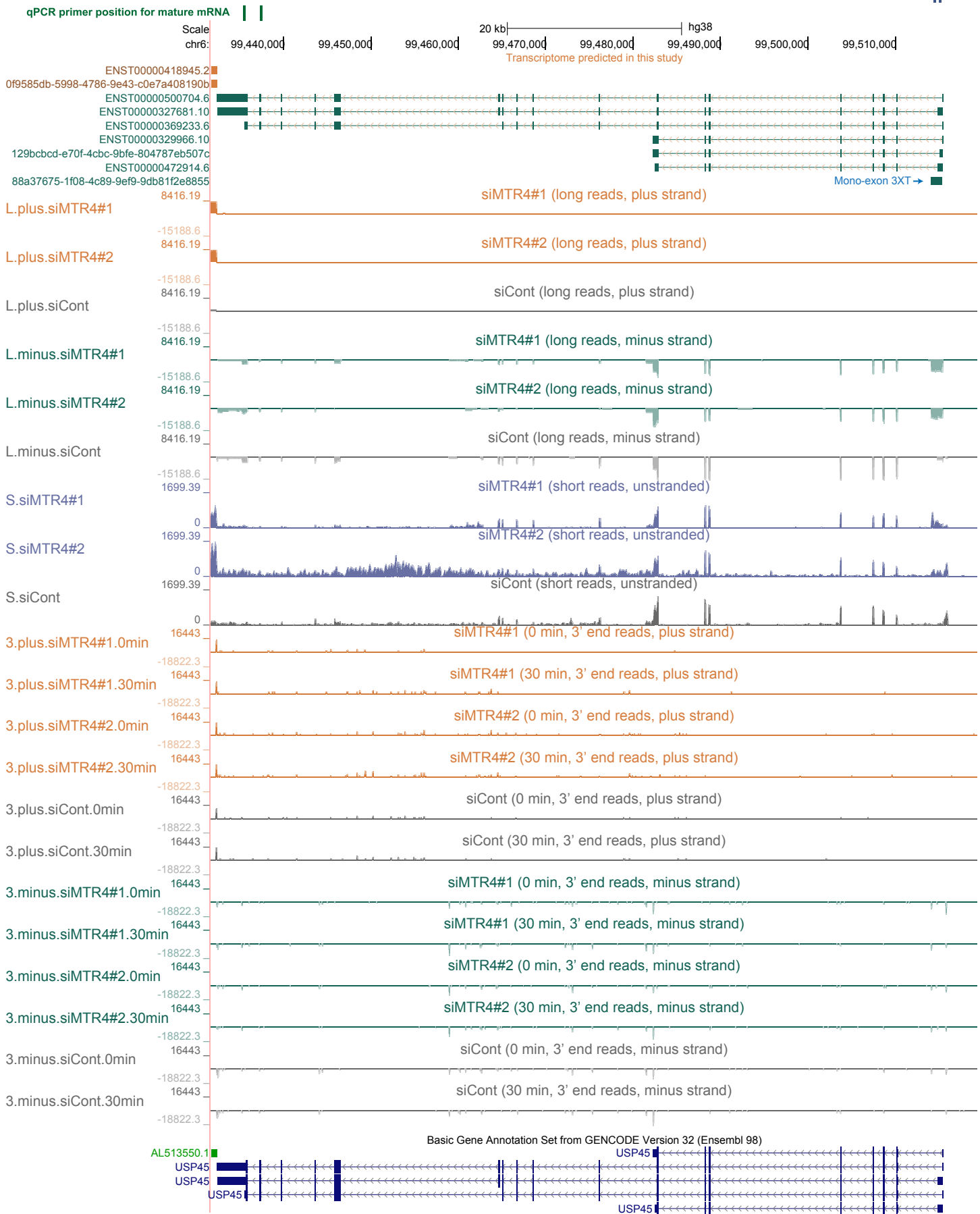
Supplementary References



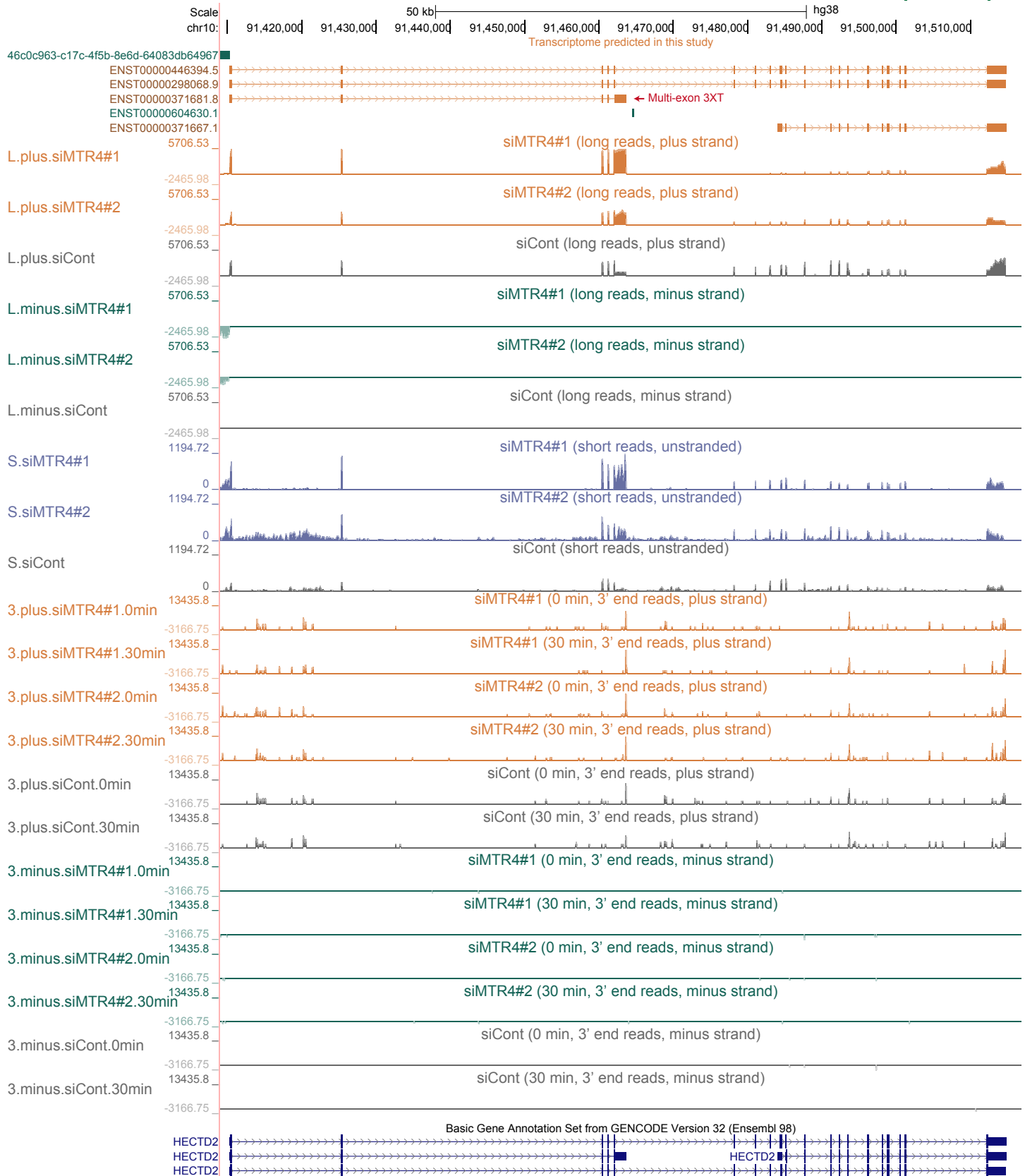
e qPCR primer position for 3XT

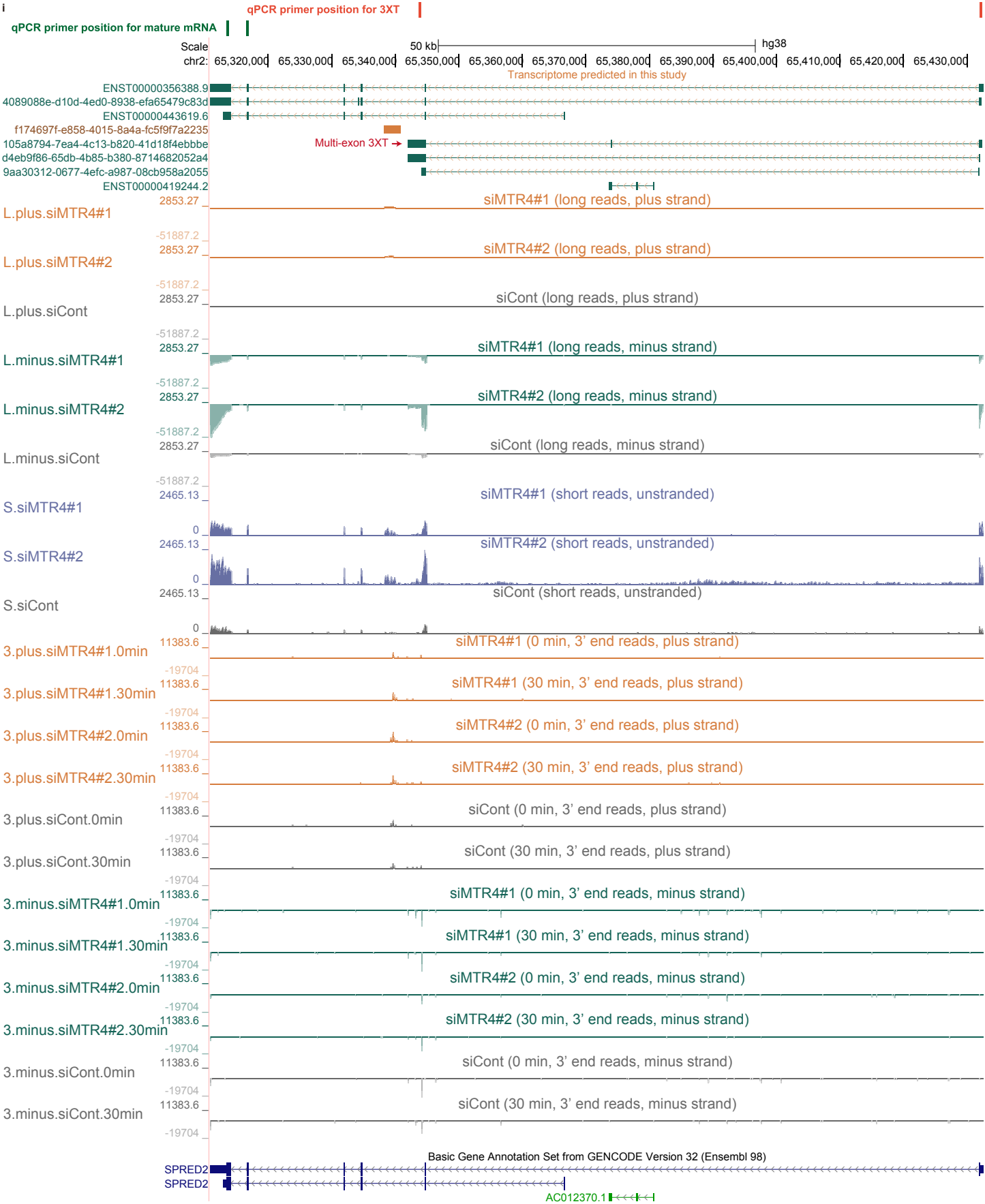


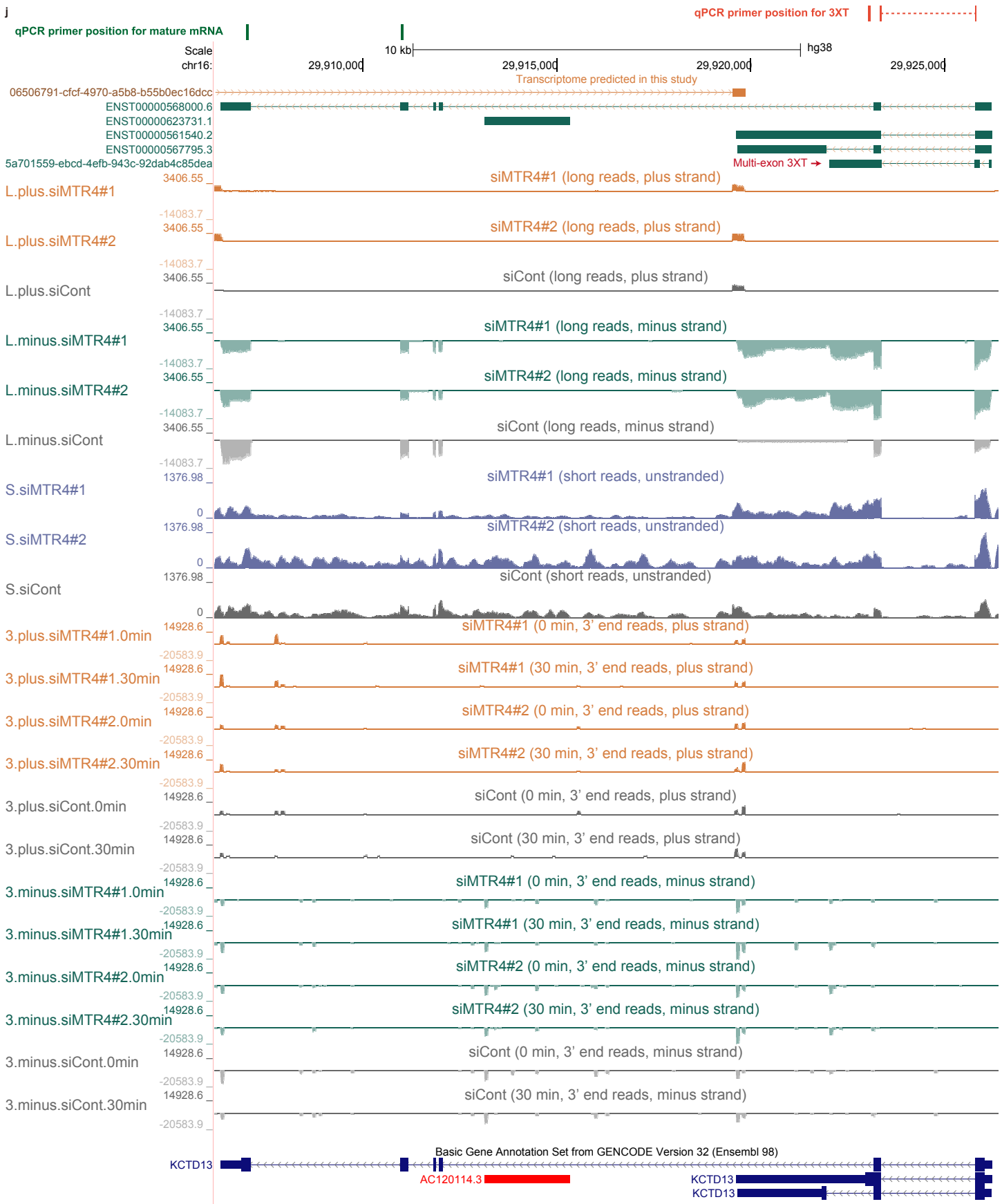


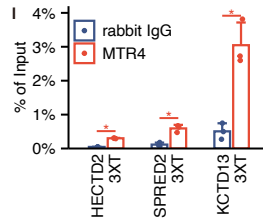
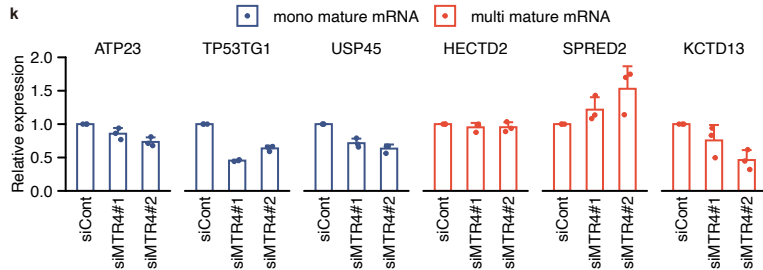


qPCR primer position for mature mRNA



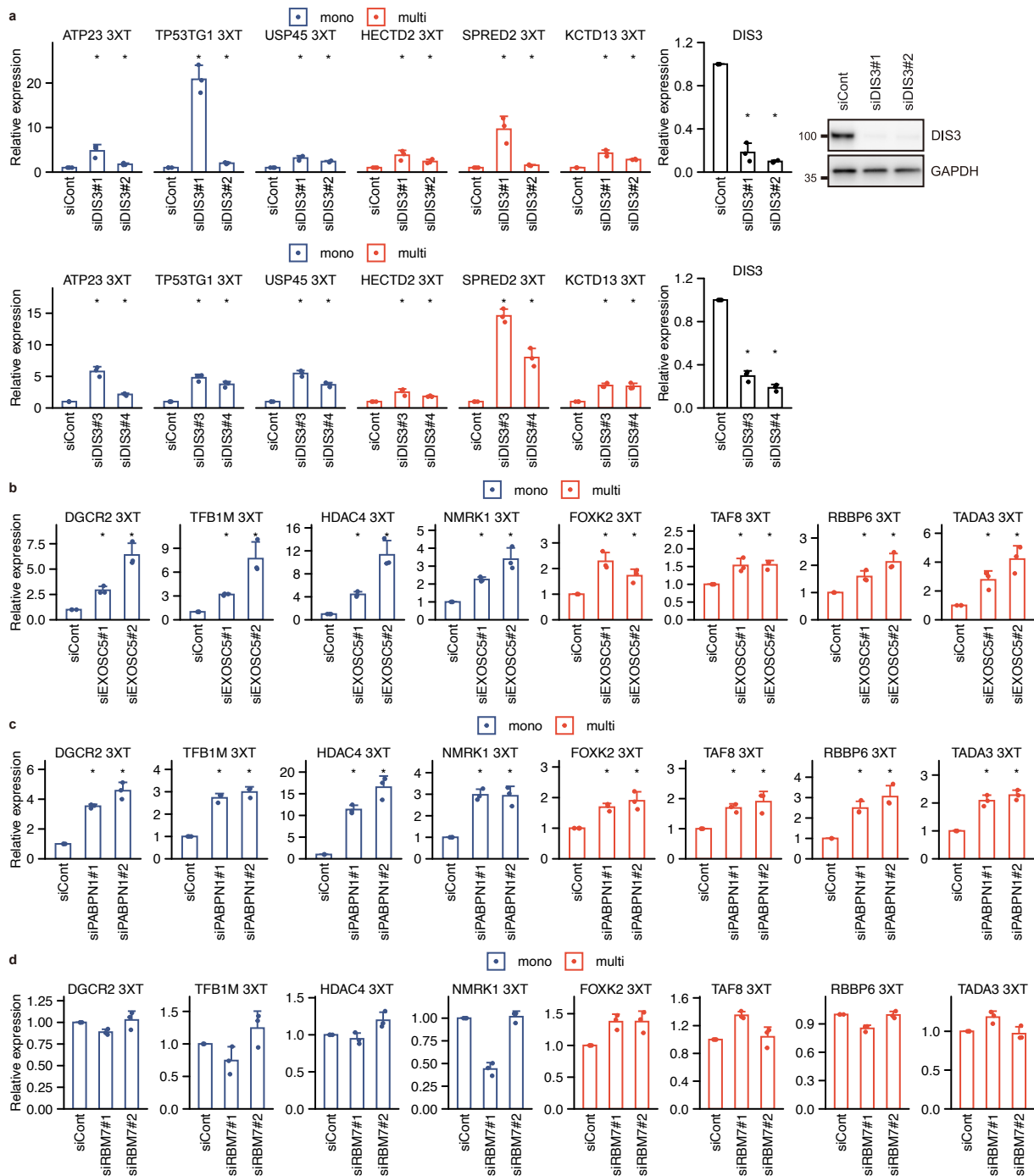


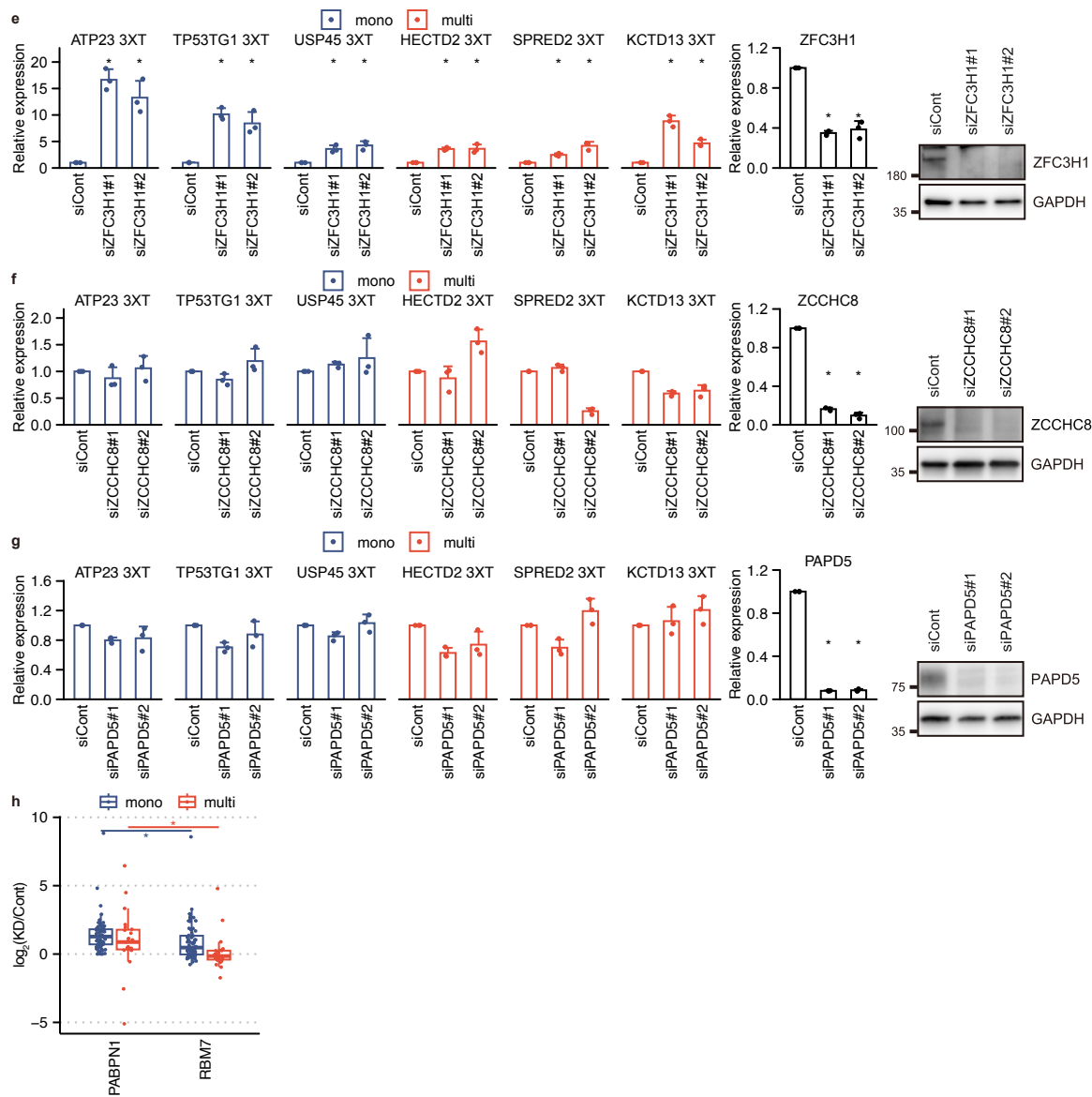




Supplementary Fig. 1 | MTR4 destabilizes 3XTs. a (left) qRT-PCR analysis of *MTR4* in HeLa cells transfected with siMTR4. Results are expressed as the mean $\pm$ s.d. ($n = 3$ biologically independent replicates). Paired two-sided Student's t-test; $*p < 0.05$. (right) HeLa cells transfected with siMTR4 subjected to immunoblotting analysis with anti-MTR4 or anti-GAPDH antibodies. GAPDH was used as a loading control. **b** Schematic representation of FLAIR analysis. Reads obtained by long- and short-read sequencing are aligned to the genome using Minimap2 or STAR, respectively. Flair-correct was used to correct the misaligned splice sites in long-read data using available genome annotations and short-read splice junctions. Flair-collapse defined high-confidence transcripts from corrected reads. Transcripts illustrated in all datasets were merged as a new annotation. The final quantification was determined by Flair-quantify using the new annotation. **c** The classification of APA features according to the status of the last exon. "Intergenic"; the last exon is located between the last exon of the originally annotated mature mRNA and the downstream gene. "Intron-exonized"; the last exon is located within the intron. "Known exon"; the last exon is located at the originally annotated exon. "5' Extended", "3' extended", "both extended"; the last exon has been extended into a primarily intronic region at the 5', 3', or both ends. **d** The number of APA features regulated by MTR4. **e–g** Screenshots from the UCSC genome browser displaying the *ATP23* (**e**), *TP53TG1* (**f**) and *USP45* (**g**) loci, which are genes with mono-exon 3XTs, with long-read (L), short-read (S), and 3' sequence (3) data from MTR4-depleted HeLa cells. **h–j** Screenshots from the UCSC genome browser displaying *HECTD2* (**h**), *SPRED2* (**i**) and *KCTD13* (**j**) loci, which are genes with multi-exon 3XTs, with long-read (L), short-read (S), and 3' sequence (3) data from MTR4-depleted HeLa cells. The qRT-PCR primers to detect mono 3XTs (blue), multi 3XTs (orange), and their corresponding originally annotated mature mRNAs (green) are indicated on the top of the screenshots. **k** qRT-PCR analysis of

original mRNAs of 3XTs in HeLa cells transfected with siMTR4. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). **I** Lysates from HeLa cells were subjected to immunoprecipitation with an anti-MTR4 antibody followed by qRT-PCR analysis to detect the indicated mRNAs. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; * $p < 0.05$. The exact p -values are **a** *MTR4*: $p = 0.00003$ (siCont vs siMTR4#1), $p = 0.00005$ (siCont vs siMTR4#2); **I** *HECTD2* 3XT: $p = 0.00073$ (rabbit IgG vs MTR4 IP), *SPRED2* 3XT: $p = 0.01043$ (rabbit IgG vs MTR4 IP), *KCTD13* 3XT: $p = 0.01074$ (rabbit IgG vs MTR4 IP). Source data are provided as a Source Data file.

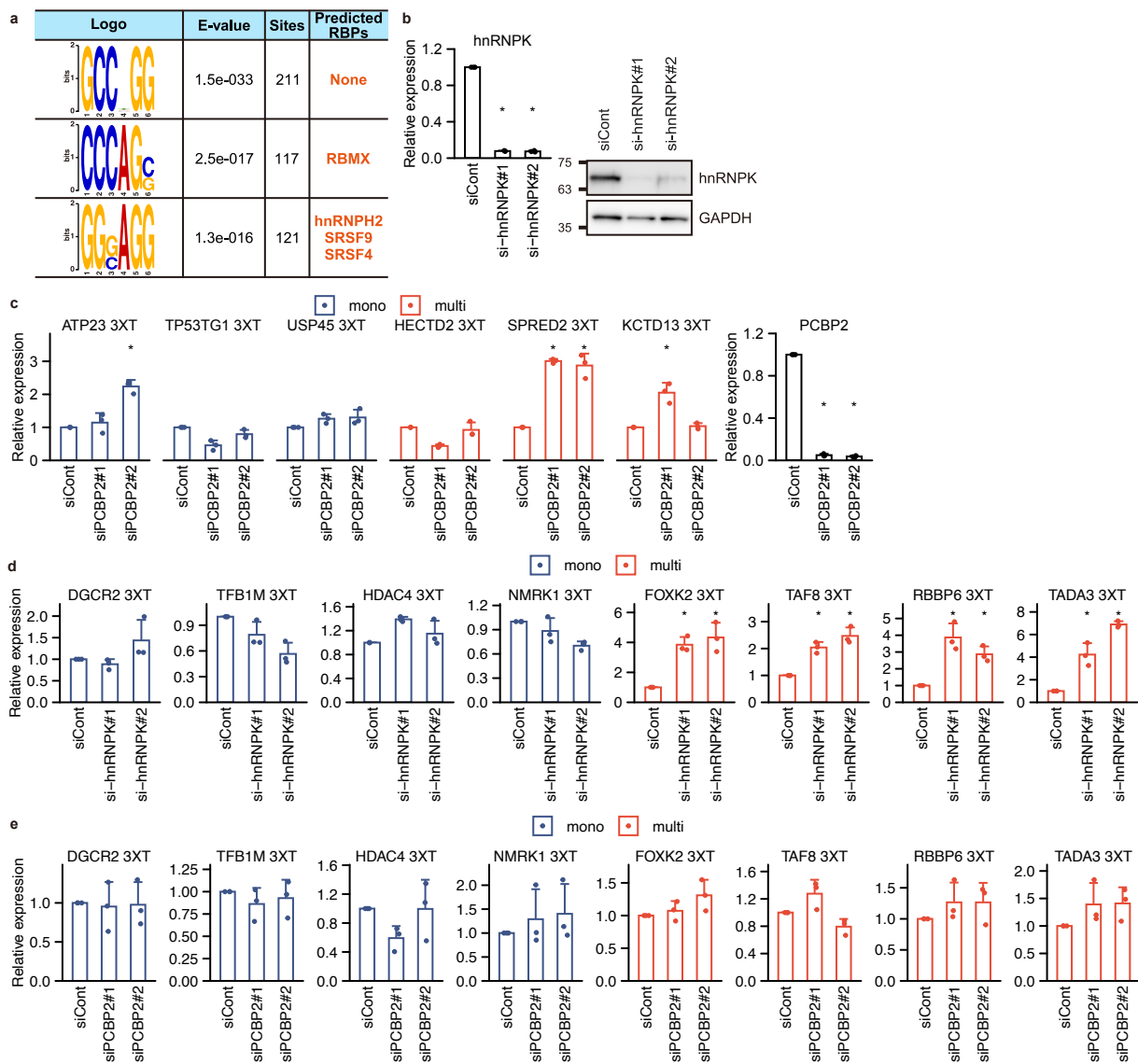


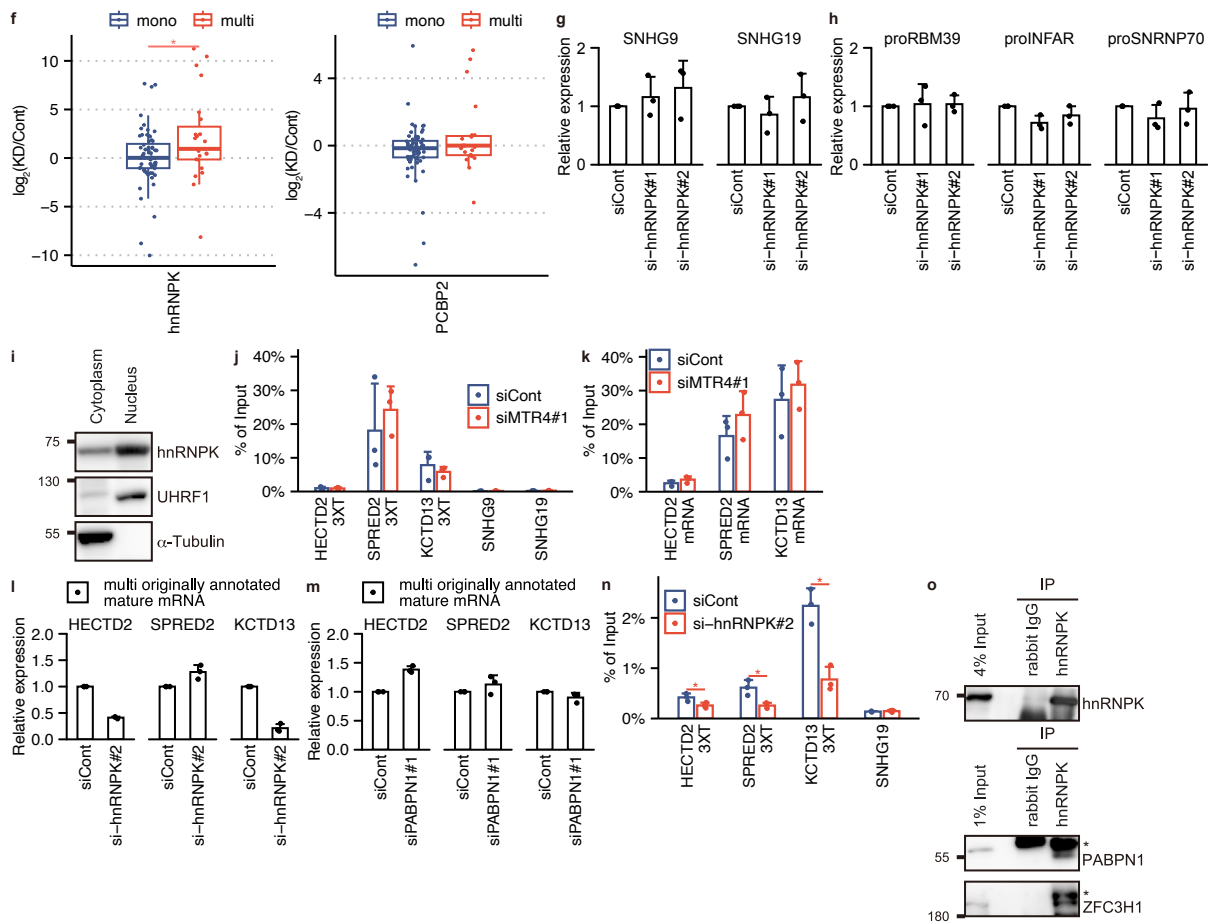


Supplementary Fig. 2 | The PAXT, but not the NEXT and TRAMP complexes, regulates the expression of 3XTs. a (left) qRT-PCR analysis of mono- (blue) and multi-exon (red) 3XT expression in HeLa cells transfected with siRNA targeting *DIS3*. (middle) qRT-PCR analysis of *DIS3* expression in HeLa cells transfected with siRNA targeting *DIS3*. qRT-PCR results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; $*p < 0.05$. (right) Cell lysates from HeLa cells transfected with siRNA targeting *DIS3* were subjected to immunoblotting analysis with anti-*DIS3* or anti-GAPDH antibodies. GAPDH was used as a loading control. **b–d** qRT-PCR analysis of additional mono- (blue, upper) and multi-exon (red, lower) 3XT expression in HeLa cells transfected with siRNA targeting *EXOSC5* (**b**), *PABPN1* (**c**) or *RBM7* (**d**). Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; $*p < 0.05$. **e–g**, (left) qRT-PCR analysis of mono- (blue) and multi-exon (red) 3XT expression in HeLa cells transfected with siRNA targeting *ZFC3H1* (**e**), *ZCCHC8* (**f**) or *PAPD5* (**g**). (middle) qRT-PCR analysis of *ZFC3H1*, *ZCCHC8* or *PAPD5* expression in HeLa cells transfected with siRNA targeting *ZFC3H1* (**e**), *ZCCHC8* (**f**) or *PAPD5* (**g**). qRT-PCR results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; $*p < 0.05$. (right) Cell lysates from HeLa cells transfected with siRNA targeting indicated genes were subjected to immunoblotting analysis with anti-*ZFC3H1* (**e**), anti-*ZCCHC8* (**f**), anti-*PAPD5* (**g**) or anti-GAPDH antibodies. GAPDH was used as a loading control. **h** MTR4-target 3XT expression levels in *PABPN1*- and *RBM7*-knockdown HeLa cells¹. Brunner-Munzel test; $*p < 0.05$. The exact *p*-values are **a** *ATP23* 3XT: $p = 0.04350$ (siCont vs siDIS3#1), $p = 0.03105$ (siCont vs siDIS3#2), *TP53TG1* 3XT: $p = 0.00825$ (siCont vs siDIS3#1), $p = 0.01099$ (siCont vs siDIS3#2), *USP45* 3XT: $p = 0.01798$ (siCont vs siDIS3#1), $p = 0.00536$ (siCont vs siDIS3#2), *HECTD2* 3XT: $p = 0.04683$

(siCont vs siDIS3#1), $p = 0.03610$ (siCont vs siDIS3#2), *SPRED2* 3XT: $p = 0.03562$
 (siCont vs siDIS3#1), $p = 0.02311$ (siCont vs siDIS3#2), *KCTD13* 3XT: $p = 0.01527$
 (siCont vs siDIS3#1), $p = 0.00132$ (siCont vs siDIS3#2), *DIS3*: $p = 0.00367$ (siCont vs
 siDIS3#1), $p = 0.00004$ (siCont vs siDIS3#2), *ATP23* 3XT: $p = 0.00821$ (siCont vs
 siDIS3#3), $p = 0.00836$ (siCont vs siDIS3#4), *TP53TG1* 3XT: $p = 0.00659$ (siCont vs
 siDIS3#3), $p = 0.00837$ (siCont vs siDIS3#4), *USP45* 3XT: $p = 0.00345$ (siCont vs
 siDIS3#3), $p = 0.00543$ (siCont vs siDIS3#4), *HECTD2* 3XT: $p = 0.03619$ (siCont vs
 siDIS3#3), $p = 0.00448$ (siCont vs siDIS3#4), *SPRED2* 3XT: $p = 0.00208$ (siCont vs
 siDIS3#3), $p = 0.01384$ (siCont vs siDIS3#4), *KCTD13* 3XT: $p = 0.00570$ (siCont vs
 siDIS3#3), $p = 0.01260$ (siCont vs siDIS3#4), *DIS3* 3XT: $p = 0.00153$ (siCont vs
 siDIS3#3), $p = 0.00048$ (siCont vs siDIS3#4); **b** *DGCR2* 3XT: $p = 0.01287$ (siCont vs
 siEXOSC5#1), $p = 0.01504$ (siCont vs siEXOSC5#2), *TFB1M* 3XT: $p = 0.00107$ (siCont
 vs siEXOSC5#1), $p = 0.03129$ (siCont vs siEXOSC5#2), *HDAC4* 3XT: $p = 0.00610$
 (siCont vs siEXOSC5#1), $p = 0.01869$ (siCont vs siEXOSC5#2), *NMRK1* 3XT: $p =$
 0.00374 (siCont vs siEXOSC5#1), $p = 0.02234$ (siCont vs siEXOSC5#2), *FORK2* 3XT:
 $p = 0.02302$ (siCont vs siEXOSC5#1), $p = 0.03897$ (siCont vs siEXOSC5#2), *TAF8* 3XT:
 $p = 0.04316$ (siCont vs siEXOSC5#1), $p = 0.01584$ (siCont vs siEXOSC5#2), *RBBP6*
 3XT: $p = 0.03968$ (siCont vs siEXOSC5#1), $p = 0.02425$ (siCont vs siEXOSC5#2),
TADA3 3XT: $p = 0.03806$ (siCont vs siEXOSC5#1), $p = 0.02626$ (siCont vs
 siEXOSC5#2); **c** *DGCR2* 3XT: $p = 0.00083$ (siCont vs siPABPN1#1), $p = 0.00805$
 (siCont vs siPABPN1#2), *TFB1M* 3XT: $p = 0.00410$ (siCont vs siPABPN1#1), $p =$
 0.00447 (siCont vs siPABPN1#2), *HDAC4* 3XT: $p = 0.00305$ (siCont vs siPABPN1#1),
 $p = 0.00904$ (siCont vs siPABPN1#2), *NMRK1* 3XT: $p = 0.00577$ (siCont vs
 siPABPN1#1), $p = 0.01690$ (siCont vs siPABPN1#2), *FORK2* 3XT: $p = 0.01008$ (siCont
 vs siPABPN1#1), $p = 0.03397$ (siCont vs siPABPN1#2), *TAF8* 3XT: $p = 0.01267$ (siCont

vs siPABPN1#1), $p = 0.04387$ (siCont vs siPABPN1#2), *RBBP6* 3XT: $p = 0.01599$ (siCont vs siPABPN1#1), $p = 0.02196$ (siCont vs siPABPN1#2), *TADA3* 3XT: $p = 0.01099$ (siCont vs siPABPN1#1), $p = 0.00637$ (siCont vs siPABPN1#2); **e** *ATP23* 3XT: $p = 0.00544$ (siCont vs siZFC3H1#1), $p = 0.02164$ (siCont vs siZFC3H1#2), *TP53TG1* 3XT: $p = 0.00529$ (siCont vs siZFC3H1#1), $p = 0.02686$ (siCont vs siZFC3H1#2), *USP45* 3XT: $p = 0.02324$ (siCont vs siZFC3H1#1), $p = 0.01957$ (siCont vs siZFC3H1#2), *HECTD2* 3XT: $p = 0.00383$ (siCont vs siZFC3H1#1), $p = 0.03038$ (siCont vs siZFC3H1#2), *SPRED2* 3XT: $p = 0.01090$ (siCont vs siZFC3H1#1), $p = 0.01775$ (siCont vs siZFC3H1#2), *KCTD13* 3XT: $p = 0.00611$ (siCont vs siZFC3H1#1), $p = 0.01223$ (siCont vs siZFC3H1#2), *ZFC3H1*: $p = 0.00038$ (siCont vs siZFC3H1#1), $p = 0.00618$ (siCont vs siZFC3H1#2); **f** *ZCCHC8*: $p = 0.00011$ (siCont vs siZCCHC8#1), $p = 0.00035$ (siCont vs siZCCHC8#2); **g** *PAPD5*: $p = 2.93524\text{E-}06$ (siCont vs siPAPD5#1), $p = 0.00005$ (siCont vs siPAPD5#2); **h** mono: $p = 0.00032$ (PABPN1 vs RBM7), multi: $p = 0.00085$ (PABPN1 vs RBM7). Source data are provided as a Source Data file.

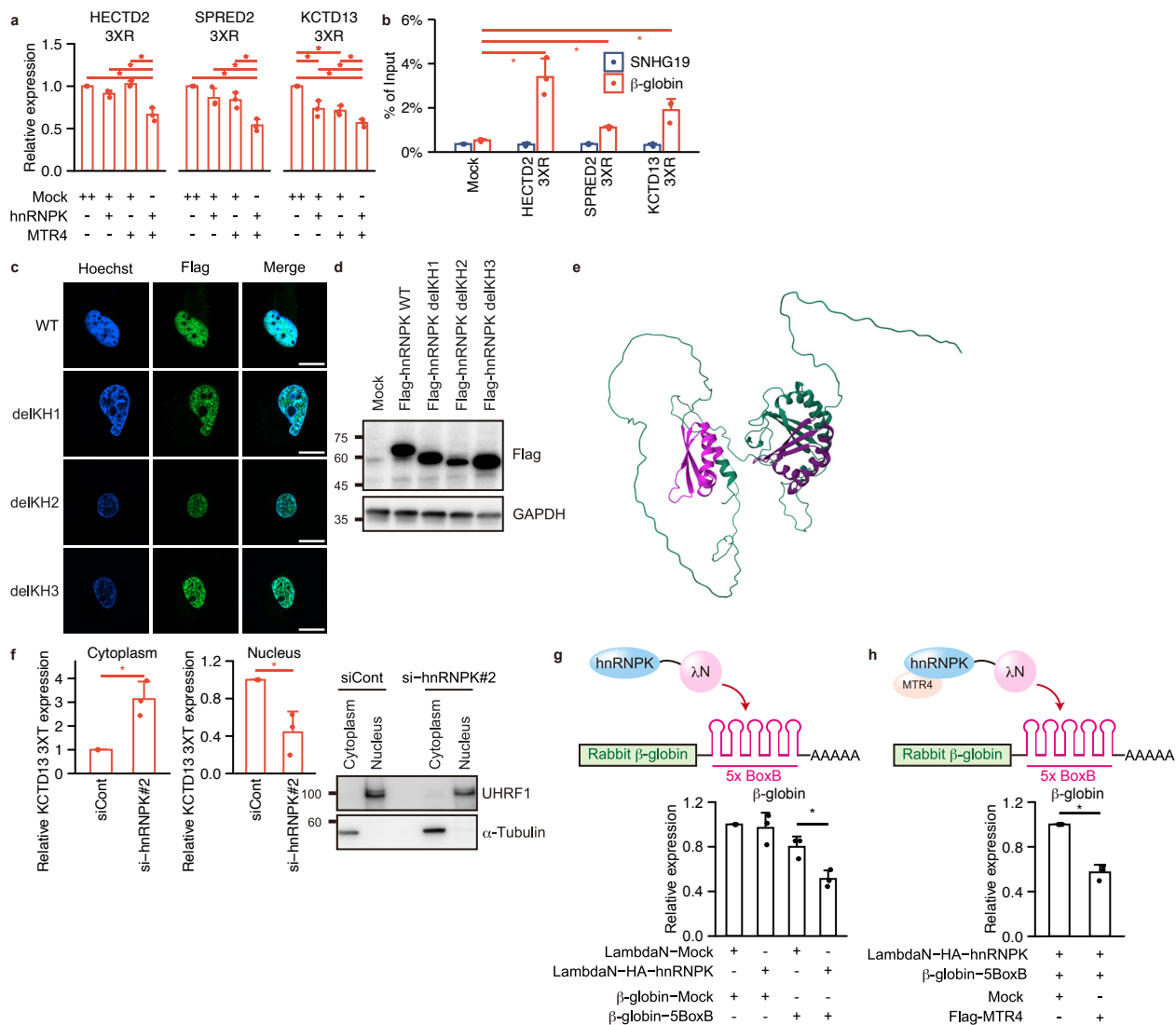




Supplementary Fig. 3 | hnRNPK regulates the expression of multi- but not mono-exon 3XTs. **a** The table shows the results of MEME motif enrichment analysis using mono-exon 3XR. E-value; the expected number of motifs with the given log likelihood ratio (or higher), and with the same width and site count, that one would find in a similarly sized set of random sequences. Sites; the number of sites contributing to the construction of the motif. Predicted RBPs (RNA-binding proteins); each motif-matched RBP speculated by Tomtom using different distance metrics. **b** (left) qRT-PCR analysis of *hnRNPK* in HeLa cells transfected with si-hnRNPK. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; $*p < 0.05$. (right) HeLa cells transfected with si-hnRNPK subjected to immunoblotting analysis with anti-hnRNPK or anti-GAPDH antibodies. GAPDH was used as a loading control. **c** (left) qRT-PCR analysis of mono- (blue) and multi-exon (red) 3XT expression in HeLa cells transfected with siPCBP2. (right) qRT-PCR analysis of *PCBP2* expression in HeLa cells transfected with siPCBP2. Results are expressed as the mean $\pm$ s.d. (n = 3). $*p < 0.05$. **d, e** qRT-PCR analysis of additional mono- (blue, upper) and multi-exon (red, lower) 3XT expression in HeLa cells transfected with siRNA targeting *hnRNPK* (**d**) and *PCBP2* (**e**). Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; $*p < 0.05$. **f** MTR4-target 3XT expression levels in cells in which *hnRNPK*² (left) and *PCBP2*³ (right) have been knocked down. Unpaired two-sided Student's t-test; $*p < 0.05$. **g** qRT-PCR analysis of *SNHG9* and *SNHG19* RNAs, which are target RNAs of the PAXT complex, in HeLa cells transfected with si-hnRNPK. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). **h** qRT-PCR analysis of *proRBM39*, *proIFNAR* and *proSNRNP70* RNAs, which are target RNAs of the NEXT complex, in HeLa cells transfected with si-hnRNPK. Results are expressed as the mean $\pm$ s.d. (n = 3

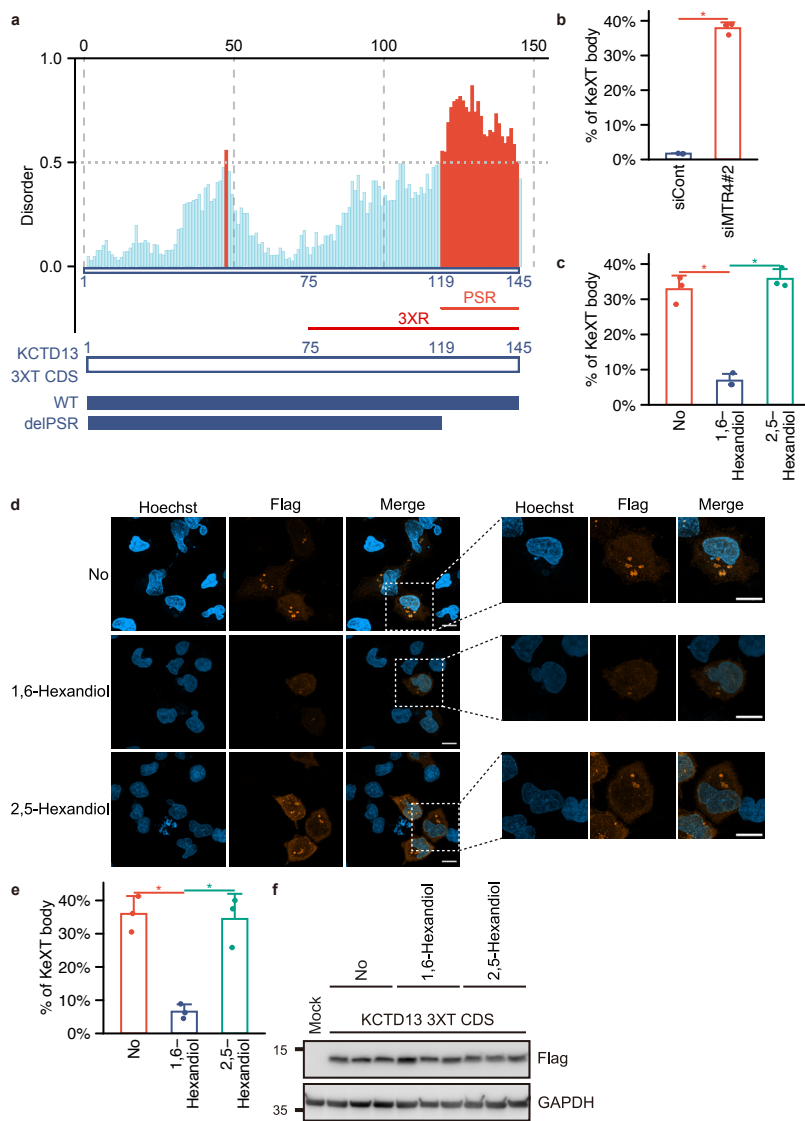
biologically independent replicates). **i** hnRNPk mainly localized in the nucleus. HeLa cells were subjected to subcellular fractionation, and the amounts of hnRNPk in each fraction were evaluated by immunoblotting analysis with the indicated antibodies. α -Tubulin was used as a cytoplasm-specific marker. UHRF1 was used as a nuclear marker. This experiment was repeated independently two times with similar results. **j** Lysates from HeLa cells transfected with siRNA targeting *MTR4* were subjected to immunoprecipitation with an anti-hnRNPk antibody followed by qRT-PCR analysis to detect the indicated RNAs. *SNHG9* and *SNHG19* RNAs were used as negative controls. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). **k** Lysates from HeLa cells transfected with siRNA targeting *MTR4* were subjected to immunoprecipitation with an anti-hnRNPk antibody followed by qRT-PCR analysis to detect the indicated mRNAs. *SNHG9* and *SNHG19* RNAs were used as negative controls. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). **l, m** qRT-PCR analysis of *HECTD2*, *SPRED2*, and *KCTD13* mRNAs, which are mature mRNAs of multi-exon 3XT, in HeLa cells transfected with si-hnRNPk#2 (**l**) and siPABPN1#1 (**m**). Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). **n** Lysates from HeLa cells transfected with siRNA targeting *hnRNPk* were subjected to immunoprecipitation with an anti-MTR4 antibody followed by qRT-PCR analysis to detect the indicated mRNAs. *SNHG9* RNA was used as a negative control. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; * $p < 0.05$. **o** hnRNPk is associated with PAXT components *in vivo*. Lysates from HeLa cells were subjected to immunoprecipitation with an anti-hnRNPk antibody or rabbit IgG followed by immunoblotting analysis with anti-ZFC3H1, anti-PABPN1 or anti-hnRNPk antibodies. *, IgG heavy chain. This experiment was repeated independently two times with similar

results. The exact p -values are **b** *hnRNPK*: $p = 0.00001$ (siCont vs si-hnRNPK#1), $p = 0.00003$ (siCont vs si-hnRNPK#2); **c** *ATP23* 3XT: $p = 0.02395$ (siCont vs siPCBP2#2), *SPRED2* 3XT: $p = 0.00033$ (siCont vs siPCBP2#1), $p = 0.01202$ (siCont vs siPCBP2#2), *KCTD13* 3XT: $p = 0.02602$ (siCont vs siPCBP2#1), *PCBP2*: $p = 0.00006$ (siCont vs siPCBP2#1), $p = 0.00002$ (siCont vs siPCBP2#2); **d** *FORK2* 3XT: $p = 0.01074$ (siCont vs si-hnRNPK#1), $p = 0.02789$ (siCont vs si-hnRNPK#2), *TAF8* 3XT: $p = 0.01259$ (siCont vs si-hnRNPK#1), $p = 0.01493$ (siCont vs si-hnRNPK#2), *RBBP6* 3XT: $p = 0.02731$ (siCont vs si-hnRNPK#1), $p = 0.02036$ (siCont vs si-hnRNPK#2), *TADA3* 3XT: $p = 0.03124$ (siCont vs si-hnRNPK#1), $p = 0.00077$ (siCont vs si-hnRNPK#2); **f** *hnRNPK*: $p = 0.02650$ (mono vs multi); **n** *HECTD2* 3XT: $p = 0.00883$ (siCont vs si-hnRNPK#2), *SPRED2* 3XT: $p = 0.04309$ (siCont vs si-hnRNPK#2), *KCTD13* 3XT: $p = 0.00916$ (siCont vs si-hnRNPK#2). Source data are provided as a Source Data file.



Supplementary Fig. 4 | hnRNP-K recognizes multi-exon 3XRs to promote the degradation of host 3XTs. **a** HeLa cells were transfected with mock, hnRNP-K and/or MTR4 expression vectors along with a construct containing a 3XR-fused rabbit β -globin gene and then subjected to qRT-PCR analysis. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; * p < 0.05. **b** Lysates from HeLa cells transfected with 3XR-fused rabbit β -globin gene were subjected to immunoprecipitation with an anti-hnRNP-K antibody followed by qRT-PCR analysis to detect the indicated RNAs. *SNHG19* RNA was used as a negative control. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; * p < 0.05. **c** Detection of hnRNP-K proteins in HeLa cells transfected with wild-type (WT) and deletion mutant (delKH1~3) Flag-tagged hnRNP-K. KH; K homology. Scale bar, 10 μ m. This experiment was repeated independently two times with similar results. **d** The lysates from HeLa cells transfected with wild-type (WT) and deletion mutant (delKH1~3) Flag-tagged hnRNP-K were subjected to immunoblotting analysis with anti-Flag or anti-GAPDH antibodies. GAPDH was used as a loading control. This experiment was repeated independently two times with similar results. **e** The hnRNP-K structure predicted by AlphaFold. KH1 (pink) and KH3 (purple) domains are in the proximity of each other. **f** hnRNP-K knockdown caused the cytoplasmic accumulation of *KCTD13* 3XTs. (left, middle) HeLa cells transfected with siCont or si-hnRNP-K#2 were subjected to subcellular fractionation, and multi-exon 3XT abundance in the cytoplasm (left) and in the nucleus (middle) was evaluated by qRT-PCR. The levels of multi-3XTs were normalized to *GAPDH* mRNA. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; * p < 0.05. (right) The lysates were subjected to immunoblotting analysis with the indicated antibodies. α -Tubulin was

used as a cytoplasmic marker. UHRF1 was used as a nuclear marker. Cytoplasm; cytoplasmic fraction. Nucleus; nuclear fraction. **g** (upper) Schematic representation of the tethering assay. (lower) HeLa cells were transfected with λ N-mock or λ N-HA-hnRNPK, β -globin-5BoxB reporter or the β -globin control, and EGFP. β -globin mRNA levels were normalized to EGFP levels and measured at 24 hours after transfection. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; * $p < 0.05$. **h** (upper) Schematic representation of the tethering assay. (lower) HeLa cells were transfected with mock or Flag-tagged MTR4, λ N-HA-hnRNPK, β -globin-5BoxB reporter, and EGFP. β -globin mRNA levels were normalized to EGFP levels and measured at 24 hours after transfection. Results are expressed as the mean $\pm$ s.d. (n = 3 biologically independent replicates). Paired two-sided Student's t-test; * $p < 0.05$. The exact p -values are **a** *HECTD2* 3XR: $p = 0.01791$ (Mock vs MTR4 + hnRNPK), $p = 0.04428$ (hnRNPK vs MTR4 + hnRNPK), $p = 0.00676$ (MTR4 vs MTR4 + hnRNPK), *SPRED2* 3XR: $p = 0.00788$ (Mock vs MTR4 + hnRNPK), $p = 0.00779$ (hnRNPK vs MTR4 + hnRNPK), $p = 0.02116$ (MTR4 vs MTR4 + hnRNPK), *KCTD13* 3XR: $p = 0.00369$ (Mock vs MTR4 + hnRNPK), $p = 0.02647$ (hnRNPK vs MTR4 + hnRNPK), $p = 0.00495$ (MTR4 vs MTR4 + hnRNPK), $p = 0.03896$ (Mock vs hnRNPK), $p = 0.01278$ (Mock vs MTR4); **b** β -globin: $p = 0.02652$ (Mock vs *HECTD2* 3XR), $p = 0.00811$ (Mock vs *SPRED2* 3XR), $p = 0.04544$ (Mock vs *KCTD13* 3XR); **f** Cytoplasm: $p = 0.03761$ (siCont vs si-hnRNPK#2); Nucleus: $p = 0.04838$ (siCont vs si-hnRNPK#2); **g** β -globin: $p = 0.04483$ (β -globin-5BoxB + LambdaN-mock vs β -globin-5BoxB + LambdaN-HA-hnRNPK); **h** β -globin: $p = 0.00814$ (Mock vs Flag-MTR4). Source data are provided as a Source Data file.



Supplementary Fig. 5 | KCTD13 3XT protein forms a liquid-like assembly. a

(upper) The dSCOPE score for each residue. The phase separation region (PSR) predicted by dSCOPE is marked in orange. (lower) Schematic representation of the wild-type (WT) and deletion mutant (delPSR) proteins translated from *KCTD13* 3XT coding sequences (CDS). **b** The graph shows the mean proportion of cells containing endogenous KeXT bodies. Results are expressed as the mean $\pm$ s.d. (n = 3). siCont; 61, 56, and 60 cells. siMTR4#2; 59, 64, and 62 cells. $*p < 0.05$. **c** The graph shows the mean proportion of cells containing endogenous KeXT bodies treated with or without 1,6-hexandiol or 2,5-hexandiol under MTR4 depletion. Results are expressed as the mean $\pm$ s.d. (n = 3). No; 56, 63, and 61 cells. 1,6-hexanediol; 52, 52, and 66 cells. 2,5-hexanediol; 56, 59, and 58 cells. $*p < 0.05$. **d** KeXT bodies in HeLa cells in the presence of 1,6-hexandiol or 2,5-hexandiol. Representative maximum projection immunofluorescence image of HeLa cells transfected with Flag-tagged KCTD13 3XT protein, treated with or without 1,6-hexandiol or 2,5-hexandiol, and immunostained with anti-Flag antibody (orange). Nuclei were stained with Hoechst33258 (blue). Scale bar, 10 μ m. **e** The graph shows the mean proportion of cells containing KeXT bodies which consisted of Flag-tagged KCTD13 3XT protein. Results are expressed as the mean $\pm$ s.d. (n = 3). No; 63, 97, and 131 cells. 1,6-hexanediol; 176, 177, and 135 cells. 2,5-hexanediol; 56, 70, and 89 cells. $*p < 0.05$. **f** Cell lysates from HeLa cells transfected with Flag-tagged KCTD13 3XT protein and treated with or without 1,6-hexanediol or 2,5-hexanediol were subjected to immunoblotting analysis with anti-Flag or anti-GAPDH antibodies. GAPDH was used as a loading control. The exact p -values are **b** $p = 0.00079$ (siCont vs siMTR4#2); **c** $p = 0.00389$ (1,6-hexandiol vs No), $p = 0.00621$ (1,6-hexandiol vs 2,5-hexandiol); **e** $p = 0.00093$ (1,6-hexandiol vs No), $p =$

0.00357 (1,6-hexandiol vs 2,5-hexandiol). Source data are provided as a Source Data file.

Supplementary References

1. Meola, N. *et al.* Identification of a Nuclear Exosome Decay Pathway for Processed Transcripts. *Mol Cell* **64**, 520–533 (2016).
2. Lubelsky, Y. & Ulitsky, I. Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature* **555**, 107–111 (2018).
3. Wang, X. *et al.* PCBP2 Posttranscriptional Modifications Induce Breast Cancer Progression via Upregulation of UFD1 and NT5E. *Mol Cancer Res* **19**, 86–98 (2021).