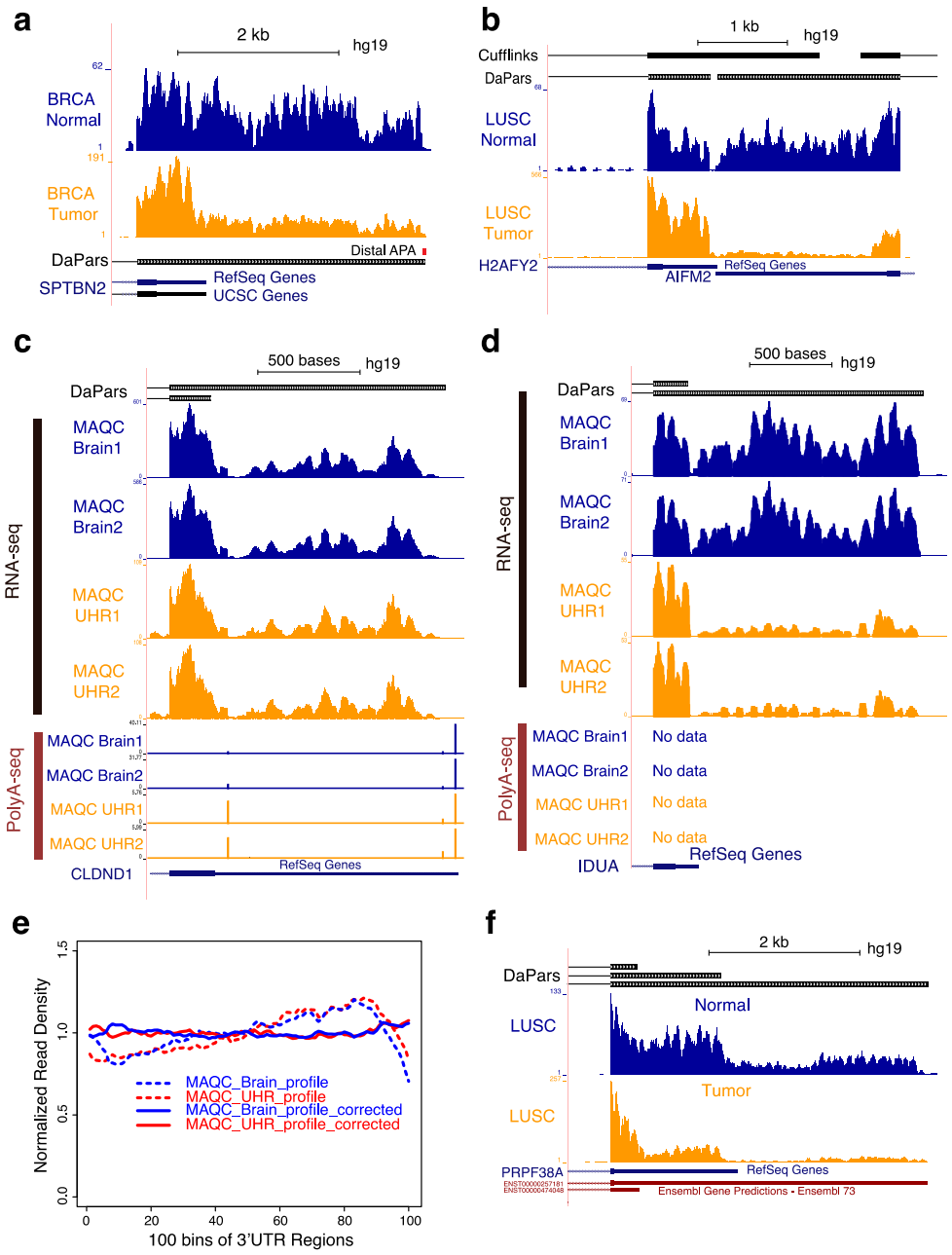
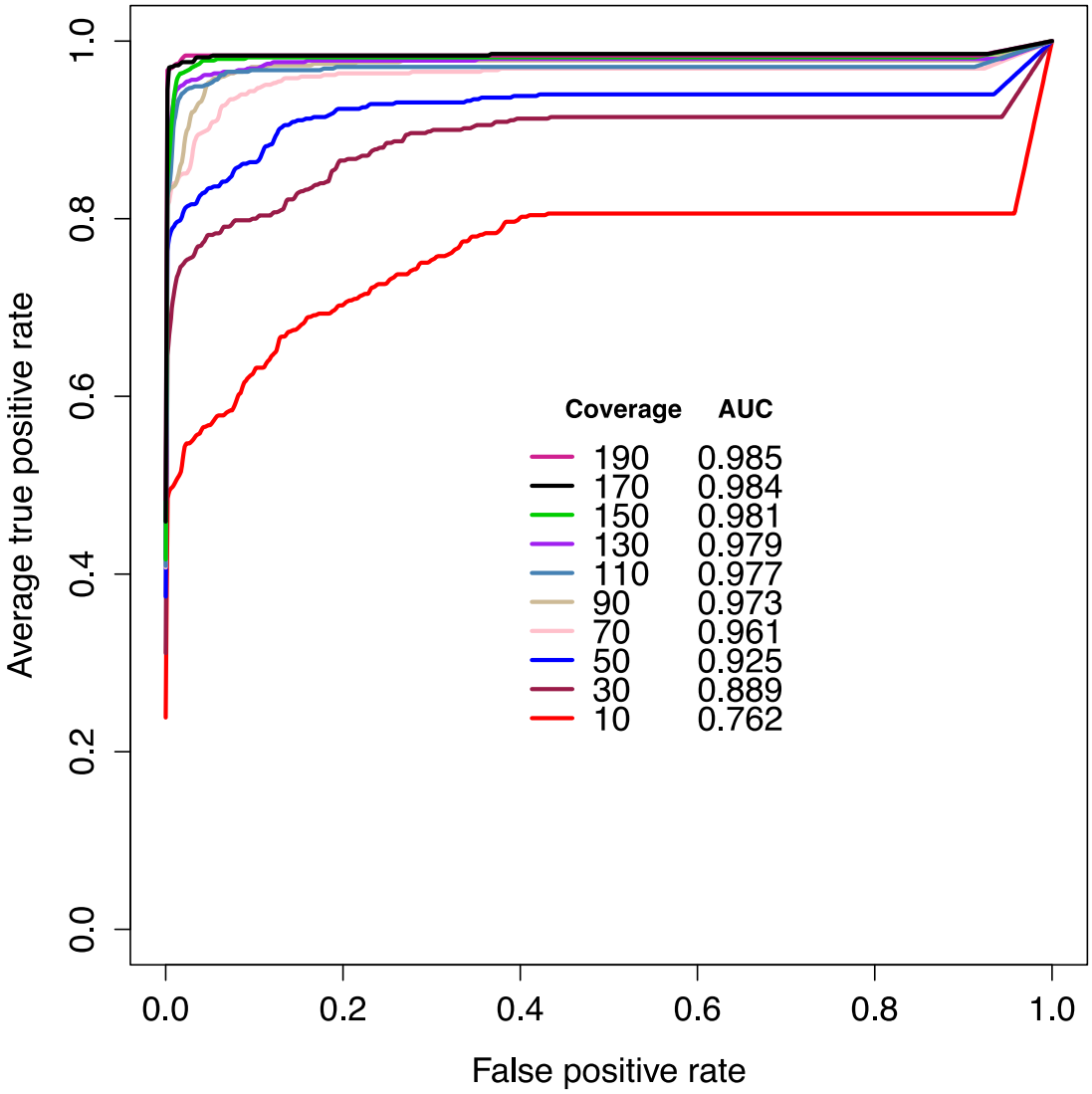
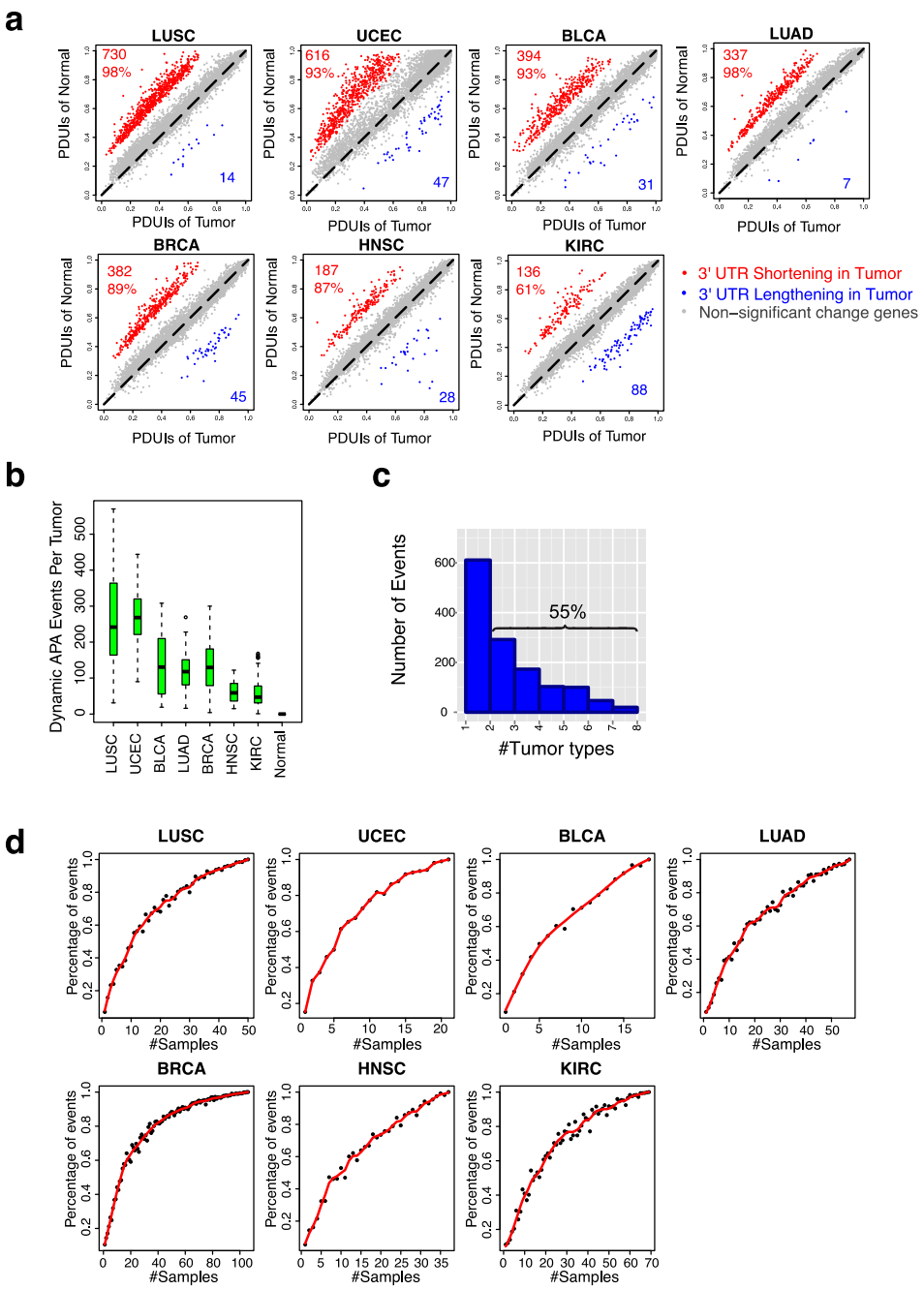


<u>Supplementary Figure 1</u>: More details about the DaPars algorithm.

(a) DaPars identifies an annotation independent distal polyA site directly from RNA-seq. (b) DaPars separates two transcripts that may overlap through automatic cutoff selection for the detection of distal polyA sites. (c) Examples of dynamic APA events detected by PolyA-seq but are not significant in DaPars analysis of RNA-seq. (d) Examples of dynamic APA events detected by DaPars analysis of RNA-seq but are absent in PolyA-seq. (e) The observed (dash lines) non-uniform and corrected (solid lines) uniform RNA-seq profiles on 3' UTR regions from MAQC UHR and Brain RNA-seq datasets. (f) DaPars is able to identify more than 2 dynamic APA events.

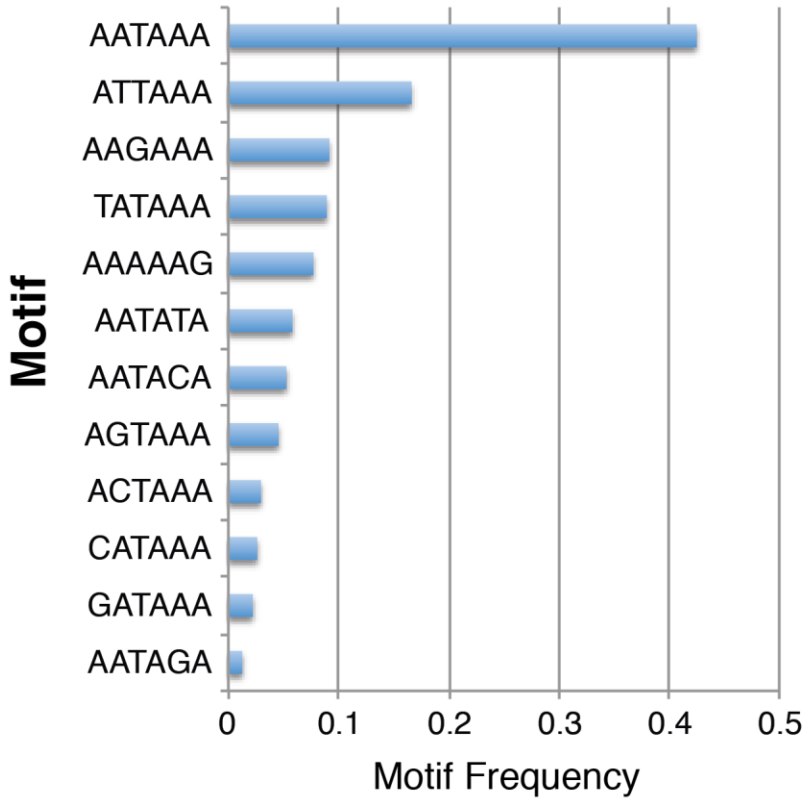


Supplementary Figure 2: The Receiver Operating Characteristic (ROC) curves of DaPars simulation at different coverage levels. AUC: Area Under the ROC Curve.

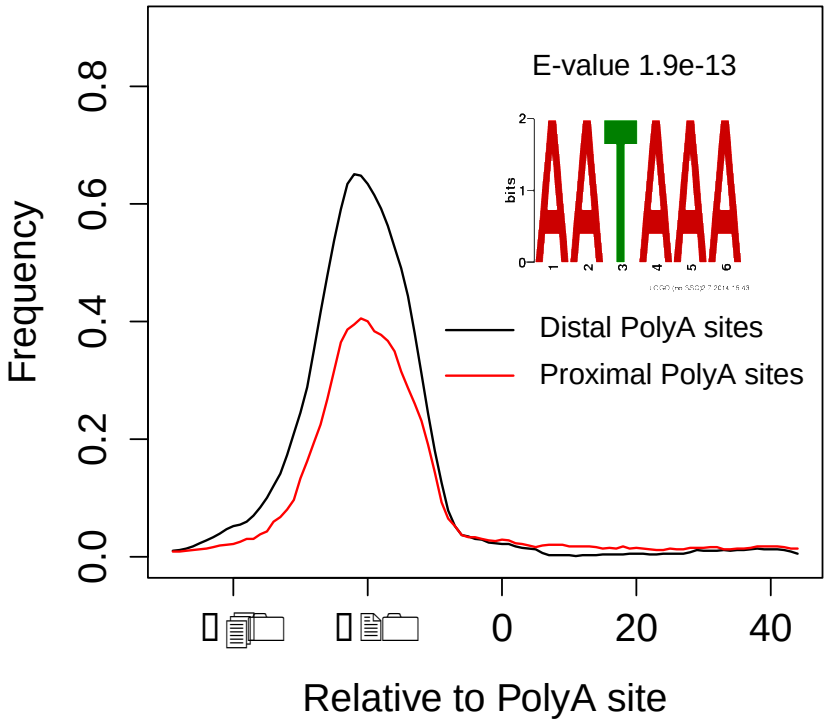


<u>Supplementary Figure 3:</u> DaPars identified dynamic APA events between tumor and matched normal.

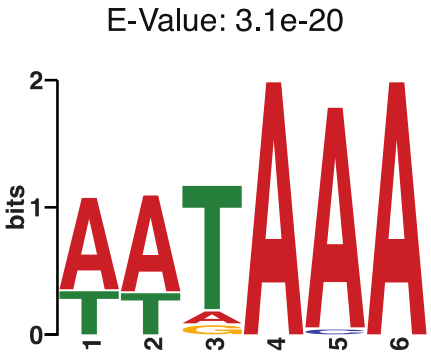
(a) Scatter plots of PDUIs of all genes in each tumor type. The integral number indicates the number of shortening (red) or lengthening (blue) events. The percentage of shortening is also listed. (b) Numbers of dynamic APA events between tumor and normal in each tumor type, and in a normal-vs-normal comparison as a negative control. (c) Dynamic APA events shared across tumor types. (d) Saturation analysis within each tumor type. Each point represents a random subset of patients with smaller sizes.



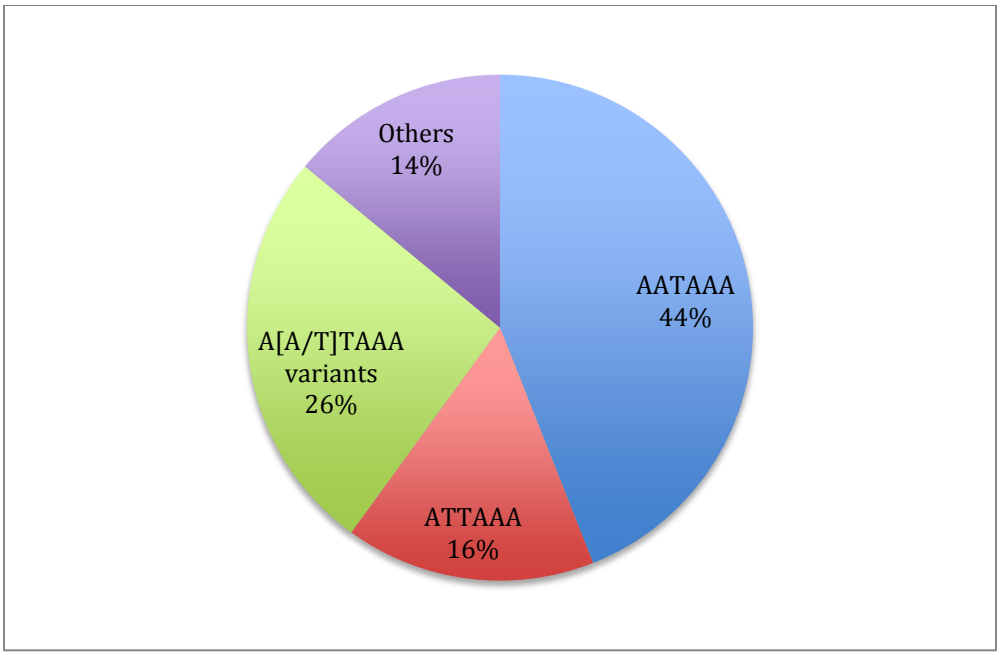
<u>Supplementary Figure 4:</u> Prevalence of Polyadenylation signal motifs in the 50 bp upstream regions of those identified dynamic proximal APA sites.



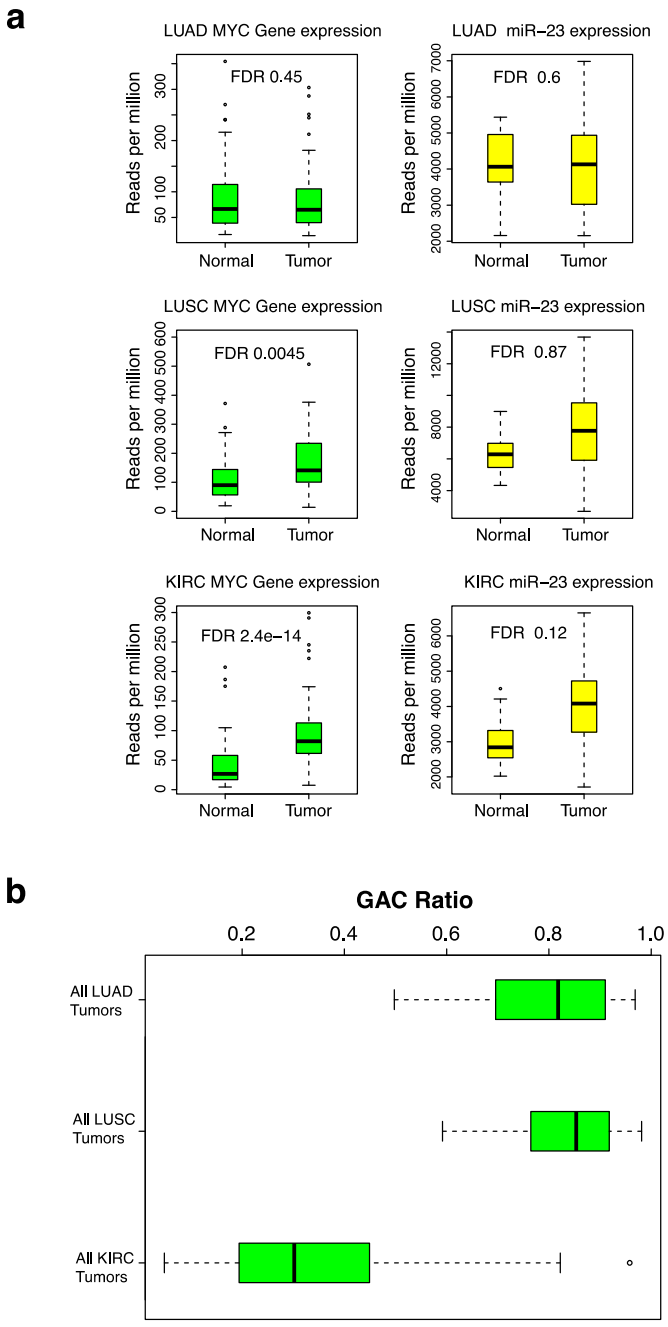
Supplementary Figure 5: Discriminative Motif identified by DREME between distal and proximal PolyA sites of recurrent 3'UTR shortening genes.



Supplementary Figure 6: Enriched polyA motif from the adjacent regions of those *de novo* APA sites that do not coincide with previously annotated polyA sites.

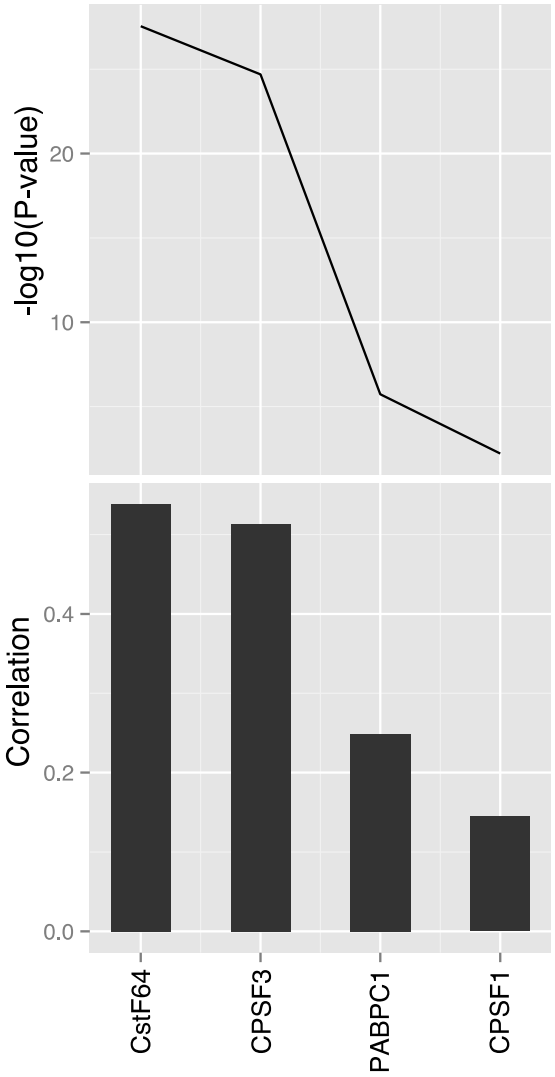


Supplementary Figure 7: Polyadenylation signal motif distribution in 50bp upstream of predicted novel PolyA sites.

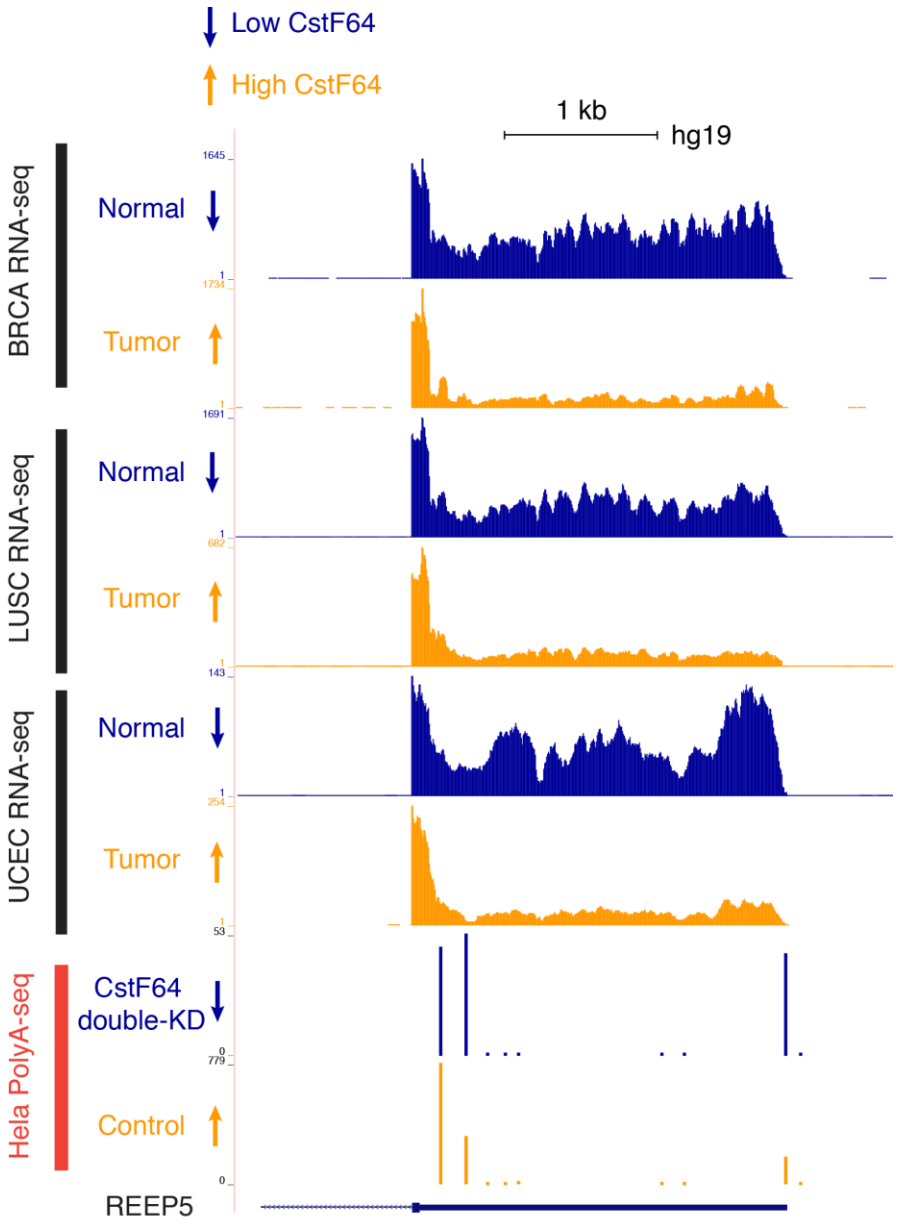


<u>Supplementary Figure 8:</u> Regulation of the *GLS* gene.

(a) Differential expression analysis of *Myc* and *miR-23*.
(b) Boxplots of *GAC* ratios of all tumors in LUSC, LUAD and KIRC.

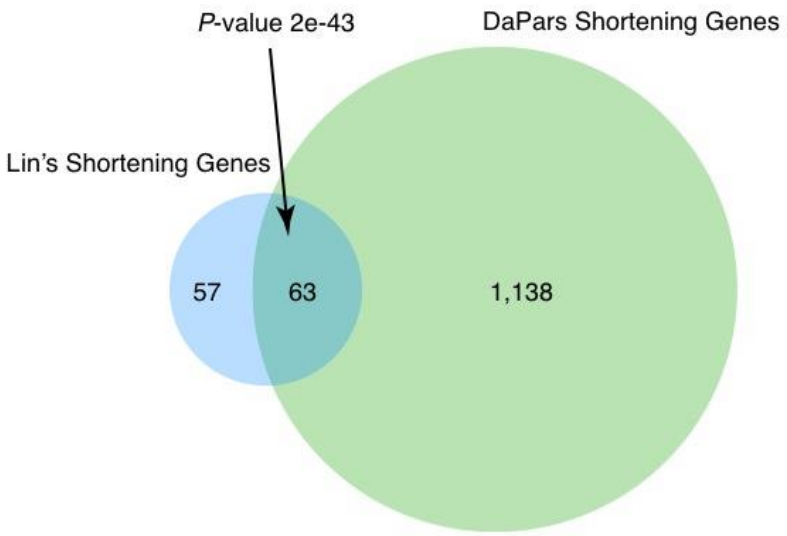


Supplementary Figure 9: Correlation between 4 PolyA factors expression fold-change and number of 3′ UTR shortening events per sample. Y-axes represent Spearman's correlations and *P*-values, respectively.

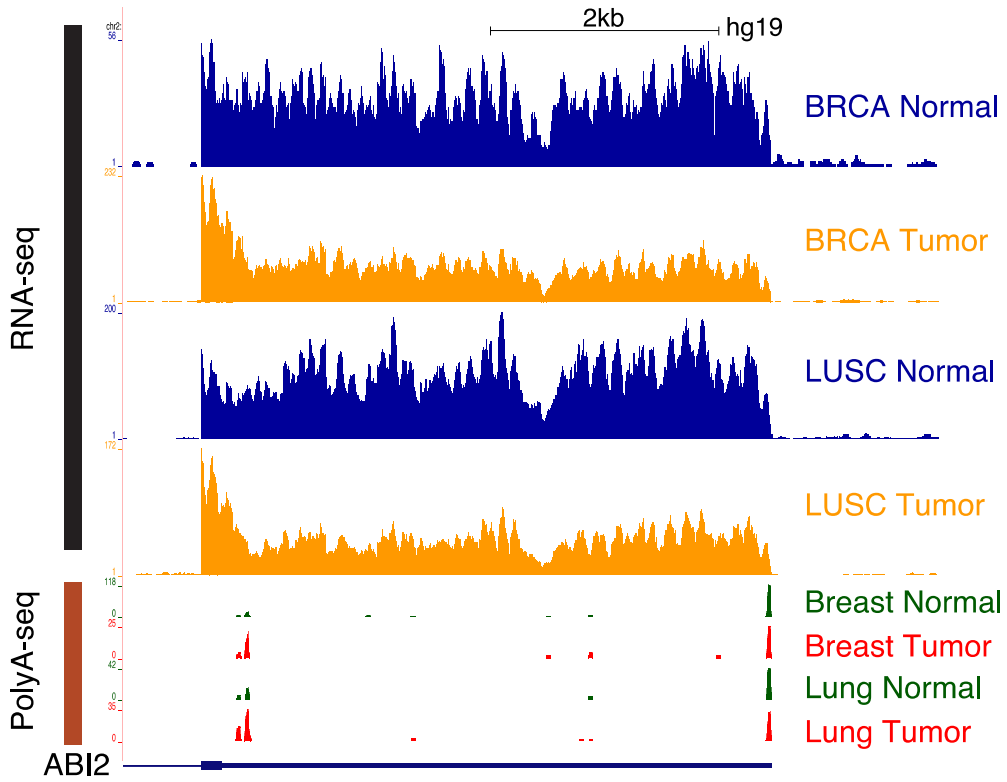


<u>Supplementary Figure 10:</u> ***CstF64*** **regulated 3′ UTR usage in Hela and tumor samples.**

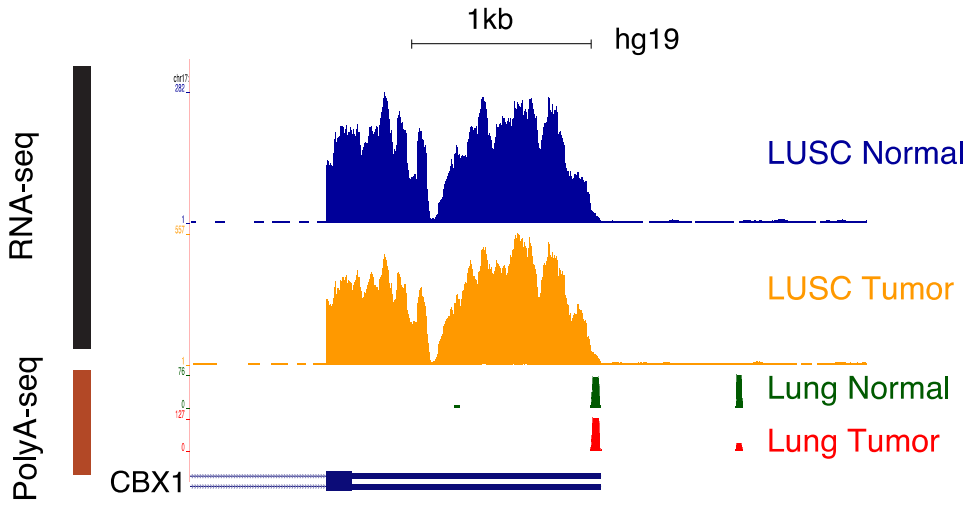
Compared with matched normal samples, tumor samples (LUSC, BRCA and UCEC) with higher *CstF64* expression show 3' UTR shortening, which resembles the PolyA-seq data in Hela cells, where Hela control with higher *CstF64* expression has 3' UTR shortening than Hela with *CstF64* double knockdown (*CstF64* double-KD).



Supplementary Figure 11: Venn plot of the comparison between Lin's PolyA-seq shortening gene list and our recurrent 3'UTR shortening gene list based on DaPars analysis. The significance of overlap was calculated by Fisher's exact test.



Supplementary Figure 12: Consistent example of Lin's PolyA-seq and TCGA RNA-seq



Supplementary Figure 13: Example of inconsistence between Lin's PolyA-seq and TCGA RNA-seq

Abbreviation	Cancer	#Tumor and matched normal pairs
BLCA	Bladder urothelial carcinoma	18
BRCA	Breast invasive carcinoma	106
LUAD	Lung adenocarcinoma	57
LUSC	Lung squamous cell carcinoma	50
HNSC	Head and Neck squamous cell carcinoma	37
KIRC	Kidney renal clear cell carcinoma	71
UCEC	Uterine corpus endometrioid carcinoma	21

Supplementary Table 1: Tumor types and number of samples in each tumor type.

Tumor type	Selected gene	ΔPDUI			Tumor mRNA expression			ΔPDUI vs exp correlation
		P-value (Cox)	HR	HR 95% CI	P-value (Cox)	HR	HR 95% CI	
BRCA	*SYNCRIP*	9.3e-05(***)	12.78	3.56-45.84	0.007(**)	1.02	1.00-1.03	-
	EPSTI1	0.24	1.60	0.73-3.51	0.09	0.98	0.95-1.00	-
	ATP5S	0.009(**)	0.29	0.11-0.73	0.06	0.90	0.80-1.00	-
LUSC	*RAB23*	0.003(**)	0.13	0.03-0.50	0.11	0.95	0.89-1.01	-
KIRC	*TMCO7*	0.006(**)	3.43	1.41-8.31	0.49	0.93	0.77-1.13	-
	PLXDC2	0.014(*)	3.48	1.28-9.43	0.07	1.02	0.99-1.05	-

Tumor type	Selected gene	ΔPDUI			mRNA log2 fold change			ΔPDUI vs log2FC correlation
		P-value (Cox)	HR	HR 95% CI	P-value (Cox)	HR	HR 95% CI	
BRCA	*SYNCRIP*	9.3e-05(***)	12.78	3.56-45.84	0.11	1.56	0.90-2.69	-
	EPSTI1	0.24	1.60	0.73-3.51	0.89	1.02	0.98-1.40	-
	ATP5S	0.009(**)	0.29	0.11-0.73	0.19	0.65	0.35-1.23	-0.22
LUSC	*RAB23*	0.003(**)	0.13	0.03-0.50	0.66	0.88	0.49-1.56	-
KIRC	*TMCO7*	0.006(**)	3.43	1.41-8.31	0.19	0.57	0.24-1.33	-
	PLXDC2	0.014(*)	3.48	1.28-9.43	0.02(*)	1.95	1.11-3.41	0.26

<u>Supplementary Table 2:</u> Multivariate Cox proportional hazards model for selected dynamic APA events of survival analysis in BRCA, LUSC and KIRC.

HR, hazard ratio; CI, confidence interval; '***', '**' and '*' correspond to P< 0.001,

P<0.01 and P<0.05, respectively. Hazard ratio exceeding 1 indicates poor prognosis for

patients possessing shorter 3′ UTR and HR less than 1 means good prognosis for patients possessing shorter 3′ UTR. If the 95% confidence interval of one covariate for HR includes 1, there may be no difference in the HR attributable to this covariate. In the last column, only the correlation with *P*-value < 0.05 is presented, otherwise indicated by '-'.