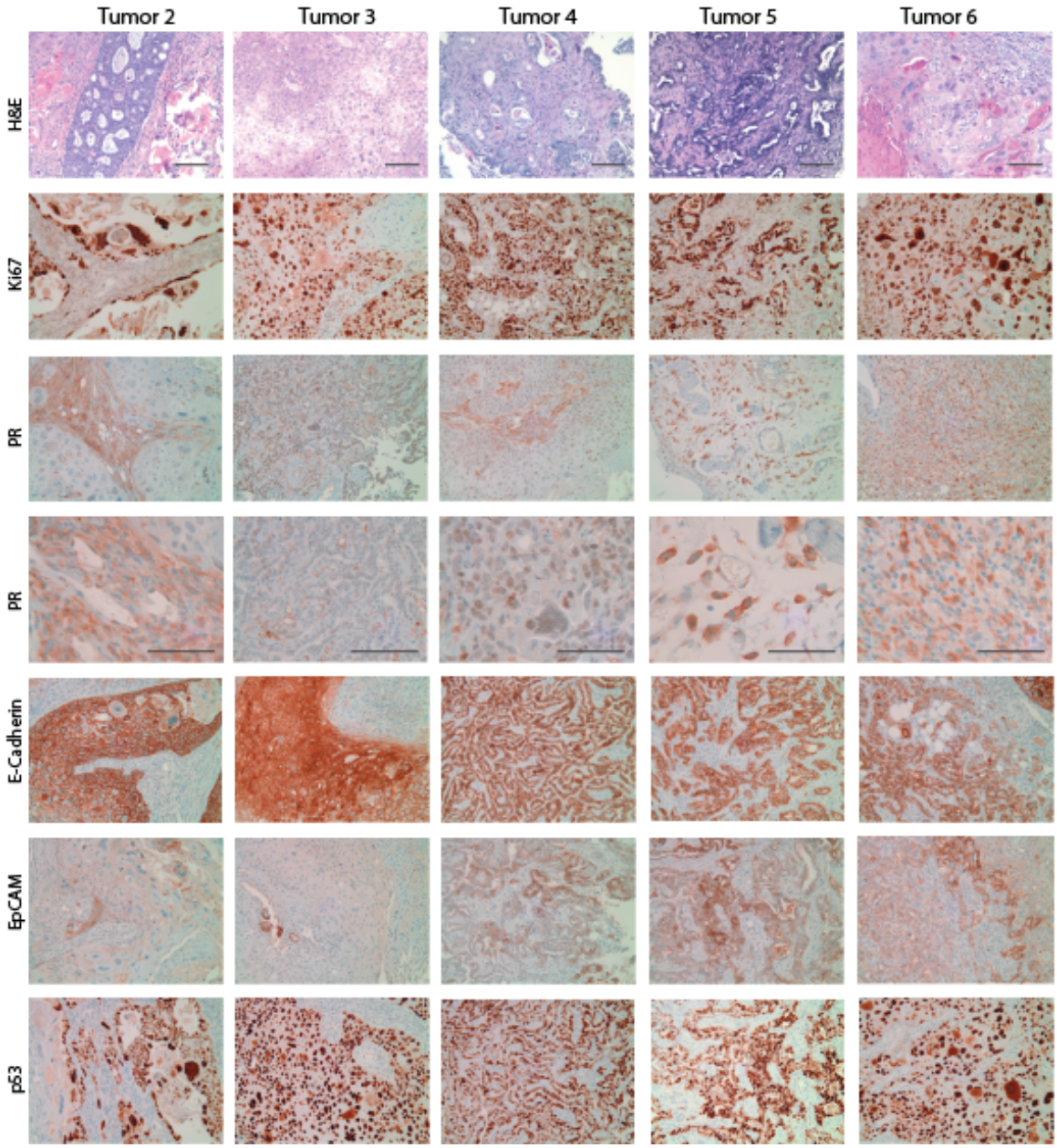
Cell Reports, Volume *28*

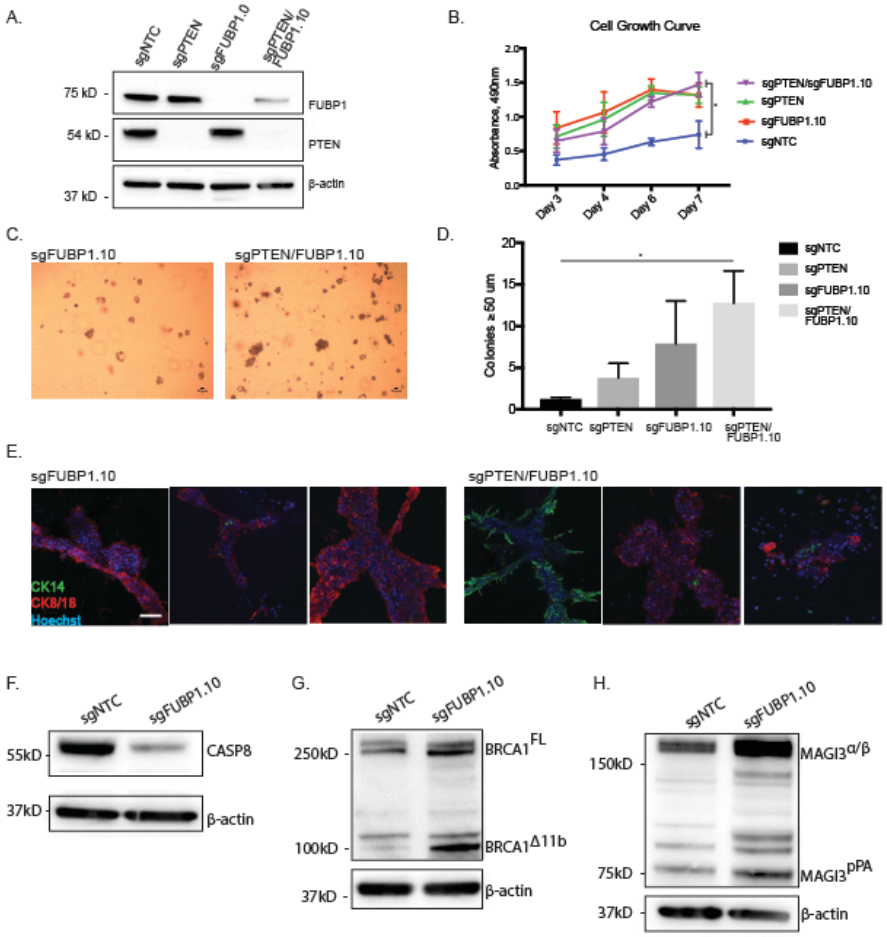
Supplemental Information

Identification of *FUBP1* as a Long Tail Cancer Driver and Widespread Regulator of Tumor Suppressor and Oncogene Alternative Splicing

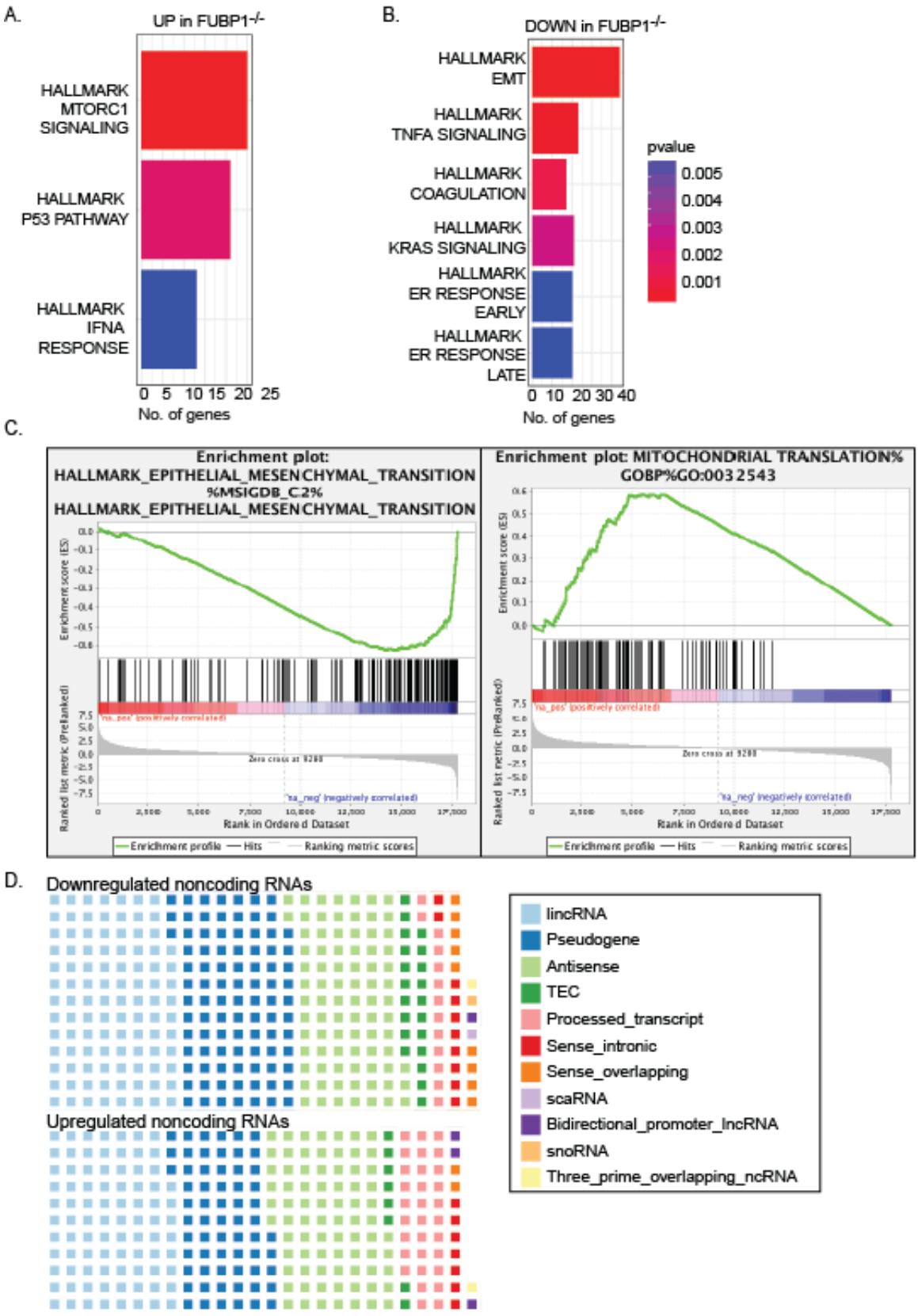
Jessica S. Elman, Thomas K. Ni, Kristen E. Mengwasser, Dexter Jin, Ania Wronski, Stephen J. Elledge, and Charlotte Kuperwasser



Supplementary Figure S1. Additional histology of tumors shows variable phenotypes and heterogeneous staining for epithelial, proliferation, and hormone markers, Related to Figure 1. Scale bars = 100 μm.

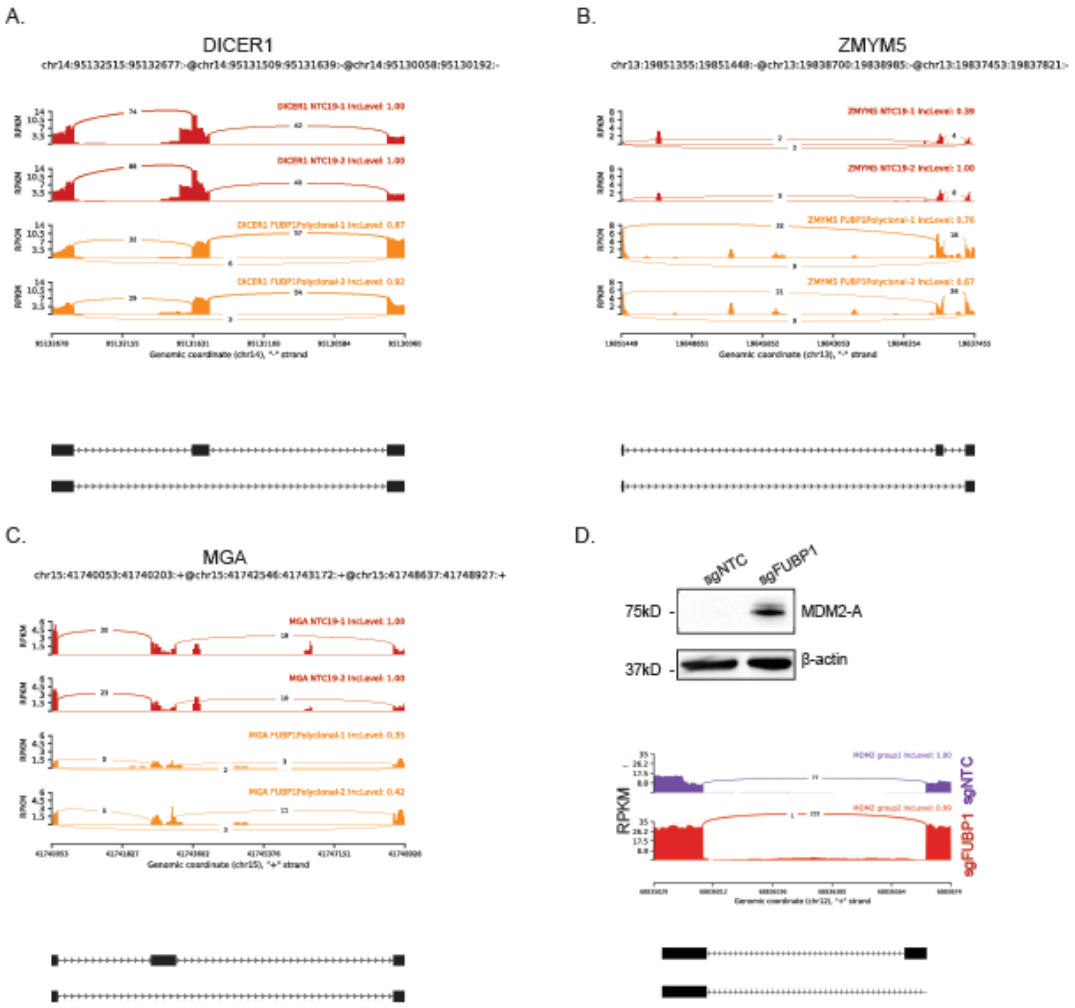


Supplementary Figure S2. Transformed phenotype in MCF10F cells generated with an alternative FUBP1-targeting sgRNA, Related to Figures 2 and 3. (A) Western blot of lysates from MCF10F cells transduced with NTC, PTEN, FUBP1, or PTEN + FUBP1 CRISPR-Cas9 sgRNA to show knockout of corresponding genes. (B) Proliferation of the indicated cell lines over 7 days, measured by MTS assay, analyzed with an ANOVA with a multiple comparisons test. (C) Soft agar growth assays for the indicated cell lines. Images show representative soft agar fields for the indicated cell lines after two weeks. Scale bar represents 50 μm. Analyzed with one-way ANOVA with a multiple comparisons test against sgNTC. (D) Quantification of soft agar colonies in the indicated cell lines after two weeks. (E) Representative IF images of the indicated cell lines after 10 days in hydrogels. Green = $CK14^{+}$, Red = $CK8/18^{+}$. Nuclei stained with Hoechst (blue). Scale bars represent 100 μm. (F-H) Validation of alternative splice events in FUBP1-null MCF10F cells generated with an alternative FUBP1-targeting sgRNA. Data are presented as mean ± SEM, n = 3 biological replicates per cell line. *$p < 0.05$ (two-tailed Student's t-tests).

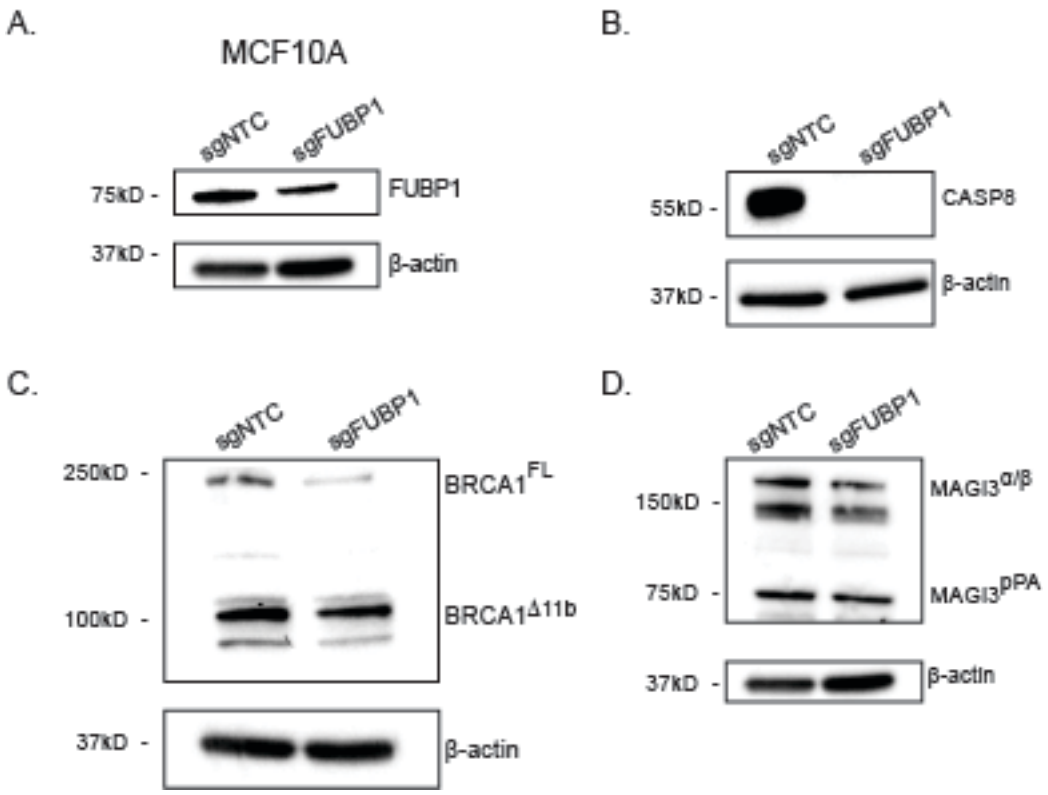


Supplementary Figure S3. FUBP1 loss results in stress-related signaling, Related to

Figure 3. (A) Universal gene enrichment analysis was performed and visualized with the R package clusterProfiler using the MSigDB collection of annotated gene sets for up-regulated and (B) down-regulated genes in *FUBP1*-null cells. The most significantly upregulated genes in FUBP1-null cells were enriched for signaling related to mTORC, p53, and IFN-α pathways. The most significantly downregulated genes were enriched for pathways related to epithelial to mesenchymal transition (EMT), TNF-α, KRAS, and estrogen response signaling. (C) GSEA analysis of up- and down-regulated genes in FUBP1-null cells revealed strong negative enrichment for EMT-related genes (*TGFB1*, *SNAI2*, *MFAP5*) (D) Waffle plot illustrating distribution of significantly non-coding RNAs among up and down-regulated genes in FUBP1-null cells (FDR ≤ 0.05, log2 fold change ≥ |1.5|).

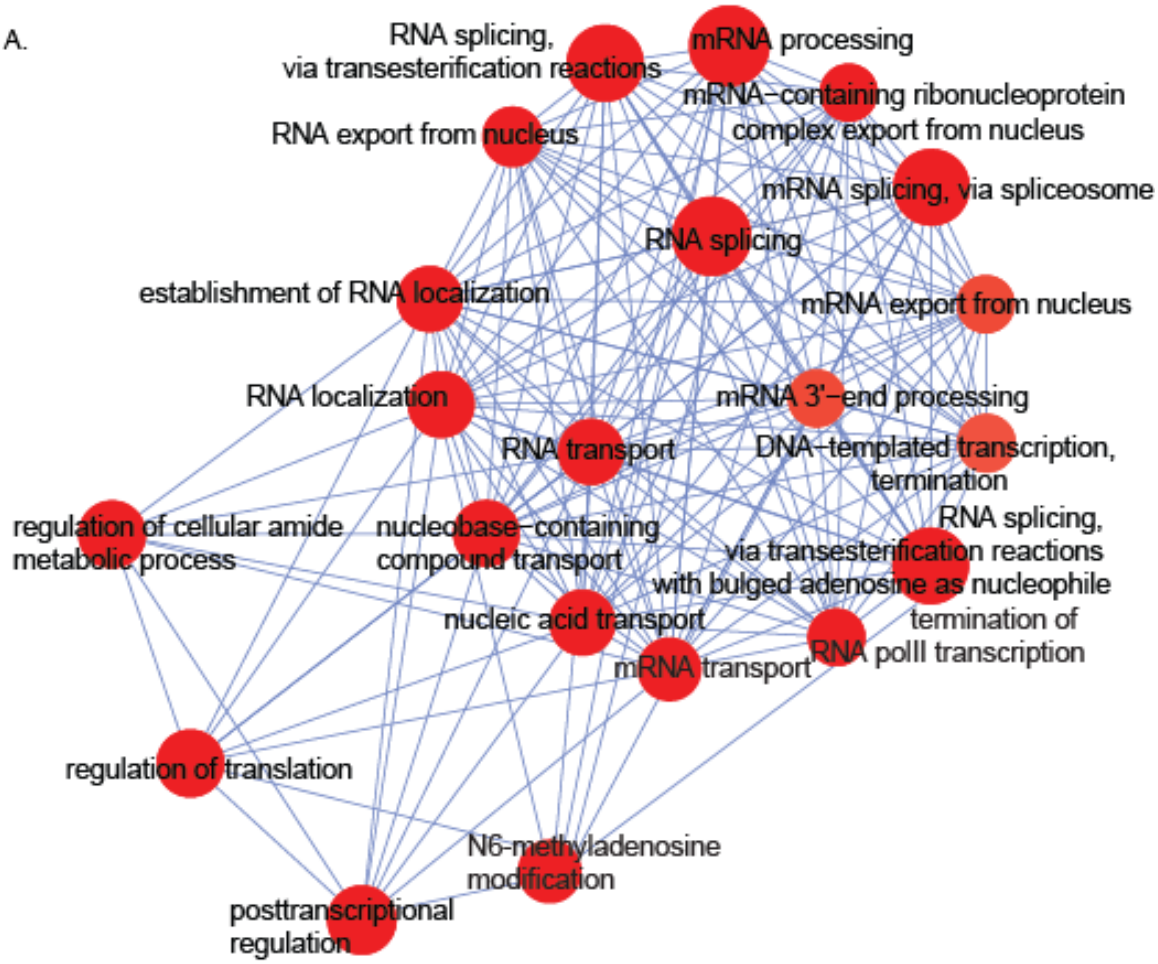


Supplementary Figure S4. Additional alternative splicing events in *FUBP1*-null and NTC cells, Related to Figure 3. (A-C) Sashimi plots for alternative splicing of *DICER1, ZMYM5*, and *MGA* in NTC and *FUBP1*-null cells. Per-base expression is plotted on y-axis, genomic coordinates on x-axis, and mRNA isoforms are shown on bottom (exons in black, introns as lines with arrow-heads). (D) Western blots of NTC and *FUBP1*-null cell lysates for MDM2 and corresponding sashimi plot of alternative splicing. Y-axis represents a modified reads per kilobase of transcript (RPKM), per a million mapped reads. Peaks indicate number of junction reads. Below, annotation of alternative isoforms.



Supplementary Figure S5. Assessment of alternative splice events in FUBP1-null MCF10A cells, Related to Figure 3. (A) Western blot of lysates from MCF10A cells transduced with NTC or FUBP1 CRISPR-Cas9 sgRNA to show knockout of FUBP1. (B-D) Western blot of lysates from NTC and FUBP1-null MCF10A cells to show alternative splicing of corresponding proteins.

A.



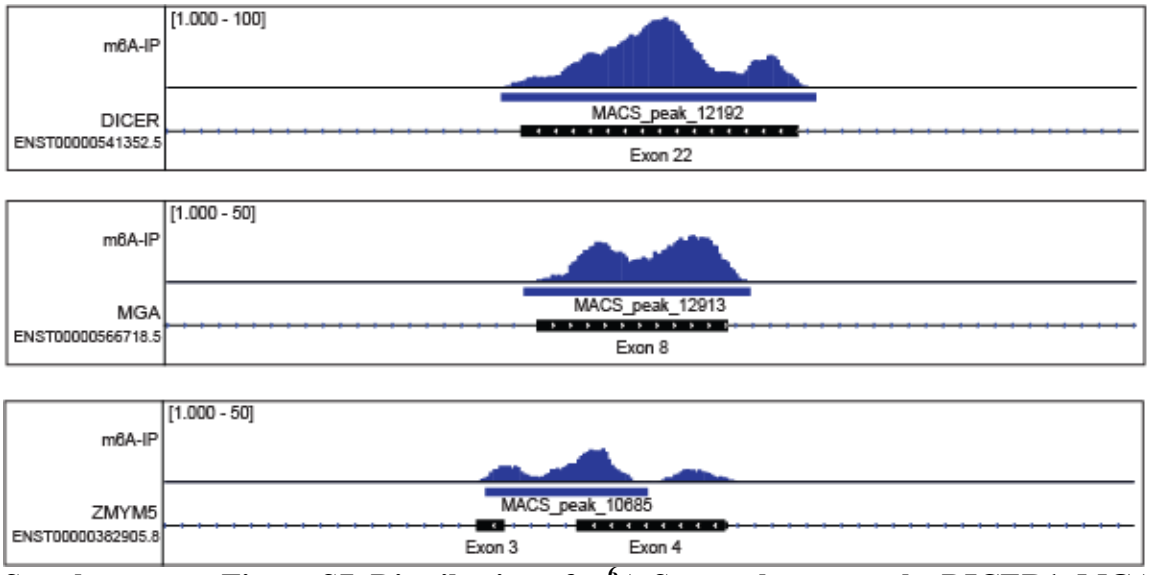
B.

Protein name	# unique peptides	Out of 3 experiments
RBM15*	15	3
ELAVL1*	11	3
SYNCRIP	20	3
DHX36	16	3
MATR3*	22	2
HNRNPA2B1*	22	2
IGF2BP1*	17	2
HNRNPR	18	2
VIRMA*	18	2

*Validated with IP/IB

Supplementary Figure S6. Functional clustering of proteins that co-immunoprecipitated with FUBP1, detected by IP/MS, Related to Figure 4. (A) The R package clusterProfiler was used to create a network that groups proteins based on biological process using the "enrichGO" function with a Benjamini & Hochberg adjustment method and a p value cutoff of 0.01. The network was generated with the

“enrichMap” command in clusterProfiler to show the top 20 enriched biological processes among proteins that uniquely co-immunoprecipitated with FUBP1. (B) Table summarizing proteins related to m^6A modification that co-immunoprecipitated with FUBP1-V5, detected by IP/MS in 3/3 (orange) or 2/3 (yellow) experiments.



Supplementary Figure S7. Distribution of m^{6}A-Seq peaks across the DICER1, MGA, and ZMYM5 loci, **Related to Figure 4.** Based on analysis of previously published m^{6}A-Seq data in HepG2 cells. The locations of the putative m^{6}A sites are indicated within exons directly upstream of splice sites yielding AS transcripts found in FUBP1-null cells