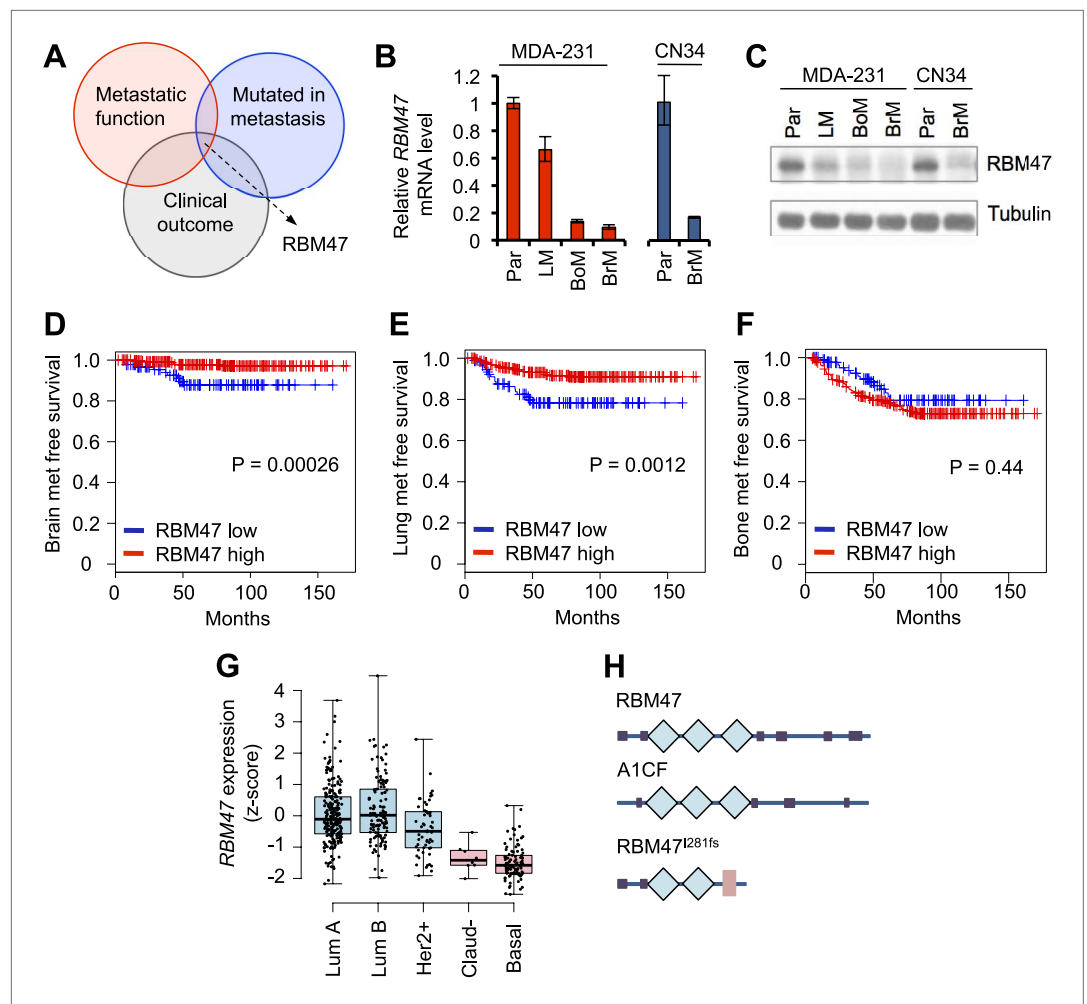


---

## Figures and figure supplements

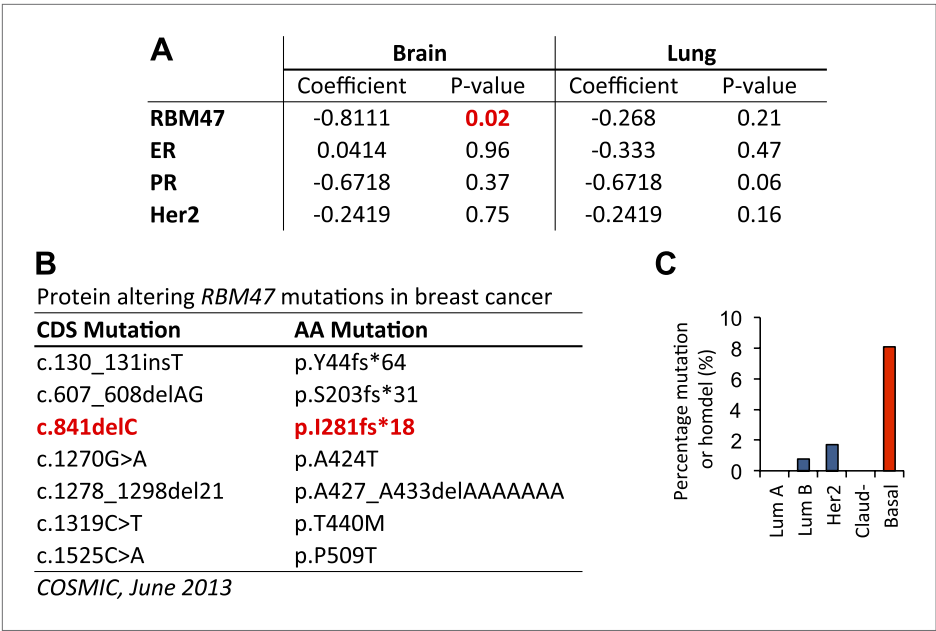
Loss of the multifunctional RNA-binding protein RBM47 as a source of selectable metastatic traits in breast cancer

**Sakari Vanharanta, et al.**

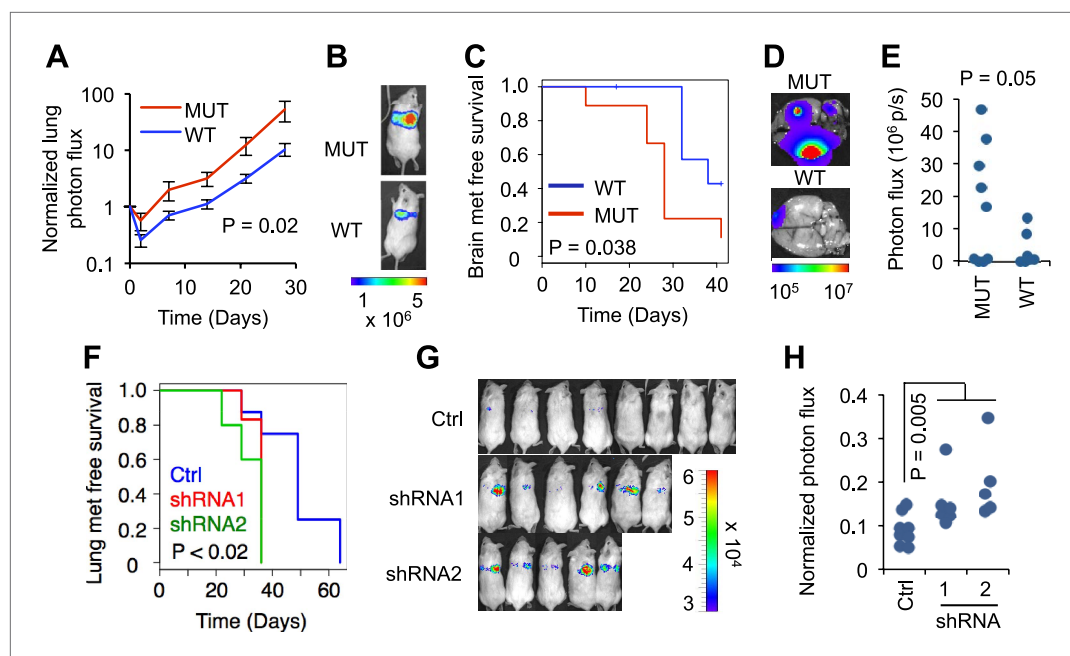


**Figure 1.** RBM47 expression associated with breast cancer progression. **(A)** A schematic of the analytical approach. Genes identified as mutated in a breast cancer brain metastasis by *Ding et al. (2010)* where compared to metastasis-associated gene expression traits in both clinical data sets and experimental model systems. This identified RBM47 as a putative breast cancer suppressor gene. **(B)** RBM47 mRNA expression measured by quantitative real-time RT-PCR in two cell line systems of breast cancer metastasis. In both panels, the data are normalized to the parental cell line (Par). Error bars represent 95% confidence intervals obtained from multiple PCR reactions. LM, lung metastatic derivative; BoM, bone metastatic derivative; BrM, brain metastatic derivative. **(C)** RBM47 protein expression measured by Western blotting. Samples as in **(B)**. Tubulin used as a loading control. **(D)** Brain metastasis free survival in a cohort of 368 untreated breast cancer patients. Cases classified based on RBM47 mRNA expression, bottom 1/3 in blue, top 2/3 in red. p-value derived from a Cox proportional hazards model using RBM47 expression as a continuous variable. **(E)** Lung metastasis free survival in a cohort of 368 untreated breast cancer patients. Cases classified based on RBM47 mRNA expression, bottom 1/3 in blue, top 2/3 in red. p-value derived from a Cox proportional hazards model using RBM47 expression as a continuous variable. **(F)** Bone metastasis free survival in a cohort of 368 untreated breast cancer patients. Cases classified based on RBM47 mRNA expression, bottom 1/3 in blue, top 2/3 in red. p-value derived from a Cox proportional hazards model using RBM47 expression as a continuous variable. **(G)** RBM47 expression as measured by RNA-seq in the TCGA data set of 748 patients. Samples grouped based on breast cancer molecular subtype: luminal A, luminal B, Her2 positive, claudin low and basal. RBM47 expression is lower in claudin low and basal subtypes, both of which are associated with poor patient outcome. **(H)** A schematic showing the predicted protein structure of RBM47, its closest homologue A1CF, a known RNA-binding protein, and the predicted structure of RBM47<sup>I281fs</sup> mutant. Blue diamonds represent RRM motifs, pink rectangle represents a truncated RRM motif.

DOI: [10.7554/eLife.02734.003](https://doi.org/10.7554/eLife.02734.003)

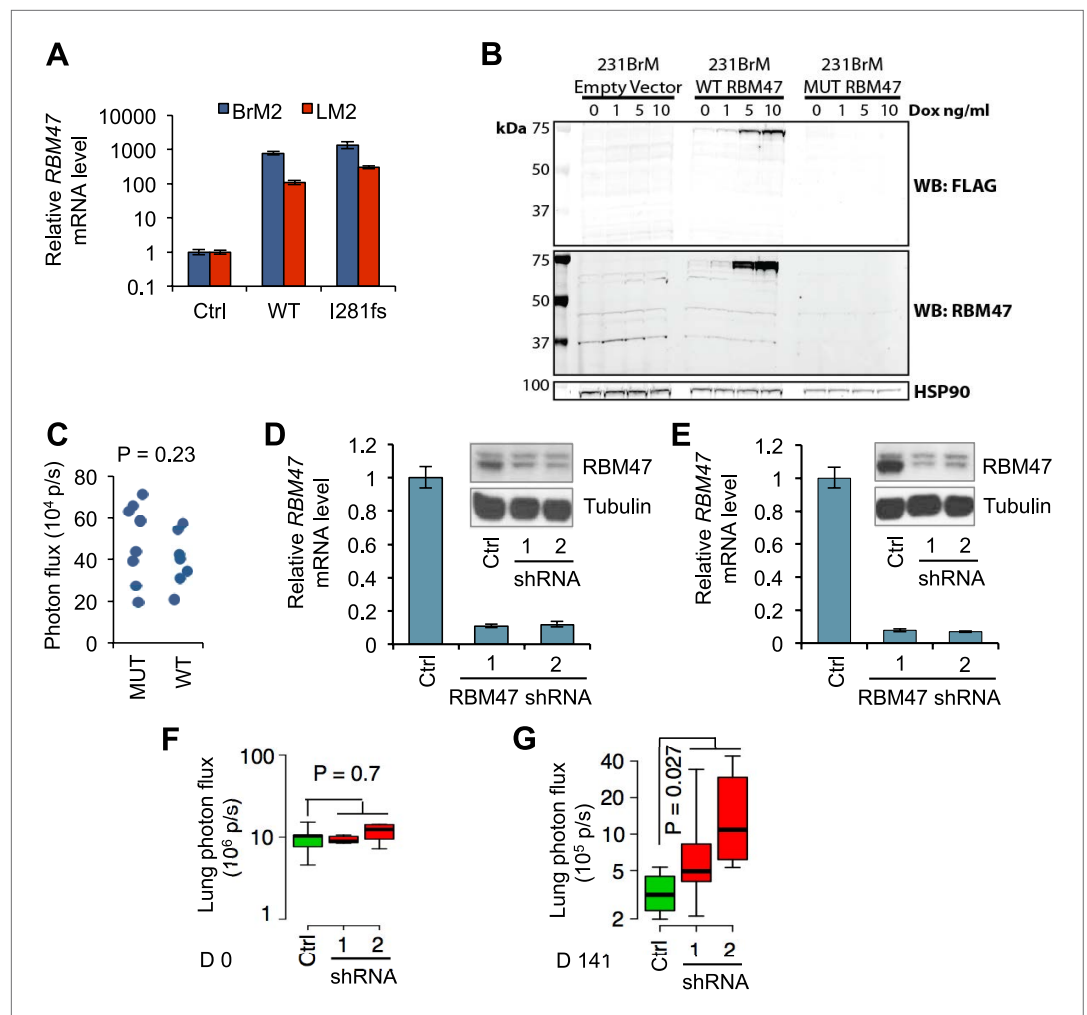


**Figure 1—figure supplement 1.** *RBM47* expression and genetic alterations in human breast cancer.  
DOI: 10.7554/eLife.02734.004



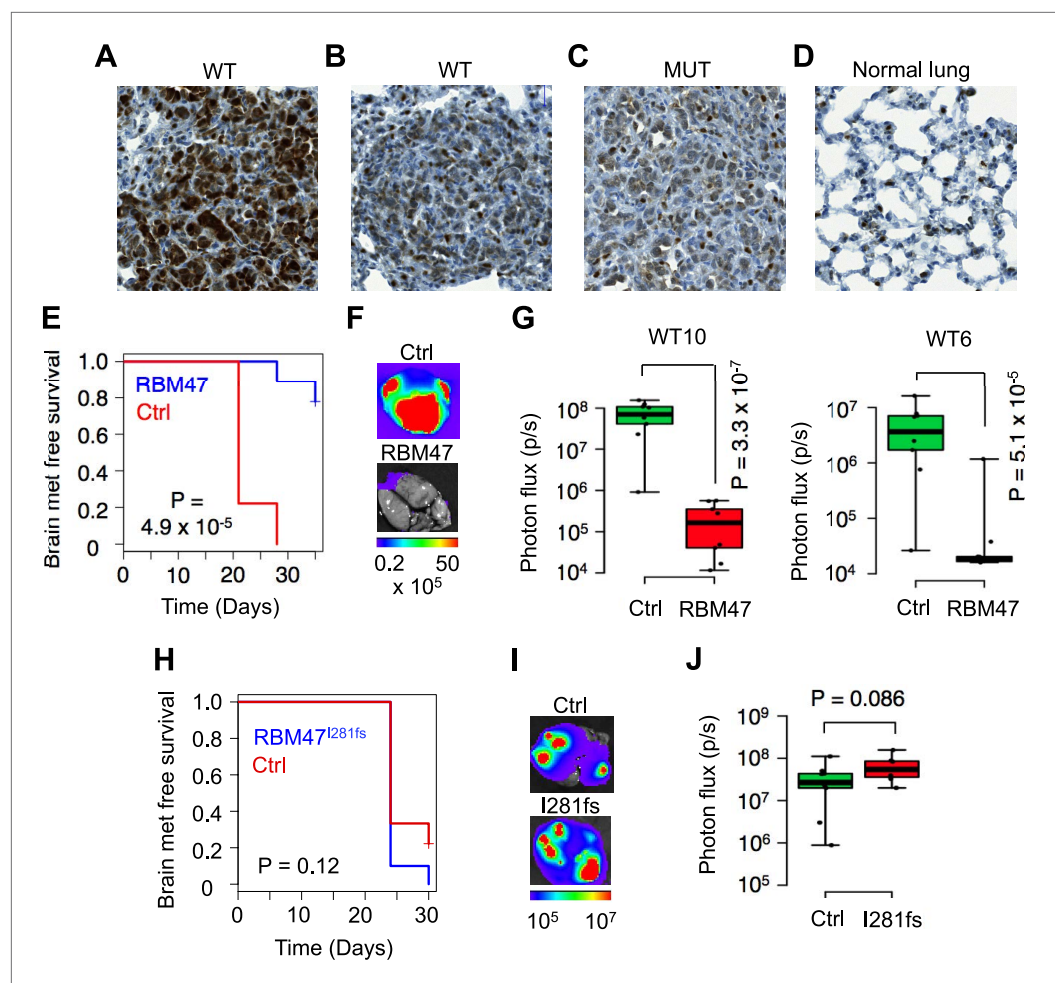
**Figure 2.** RBM47 suppresses metastatic breast cancer progression. **(A)** Normalized lung photon flux in mice after tail vein inoculation of MDA231-LM2 cells expressing either wild type RBM47 or RBM47<sup>1281fs</sup>. p-value calculated utilizing repeated measures two-way ANOVA. N = 6 for RBM47<sup>1281fs</sup>, N = 8 for RBM47. **(B)** Representative bioluminescence images from the experiment shown in **(A)**. The color scale shows bioluminescence (photons/second). **(C)** Brain metastasis free survival as determined by in vivo bioluminescence imaging in mice after intracardiac inoculation of MDA231-BrM2 cells expressing either wild type RBM47 or RBM47<sup>1281fs</sup>. p-value calculated using the log-rank test. N = 9 for RBM47<sup>1281fs</sup>, N = 7 for RBM47. **(D)** Representative bioluminescence images from the experiment shown in **(C)**. The color scale shows bioluminescence (photons/second). **(E)** Ex vivo quantification of bioluminescence from brain metastases on day 42 of the experiment shown in panel **(C)**. P-value calculated by two-tailed Student's t test. **(F)** Lung metastasis free survival as determined by in vivo bioluminescence imaging in mice after tail vein inoculation of parental MDA231 cells expressing either control vector (pGIPZ) or hairpins against RBM47 (shRNA1 and shRNA2). p-values calculated using the log-rank test. N = 8 for Ctrl group, N = 6 for shRNA1, N = 5 for shRNA2. **(G)** In vivo bioluminescence imaging on day 36 of the experiment shown in **(F)** demonstrates earlier emergence of detectable lung metastasis for the RBM47 knockdown groups when compared to the control animals. The color scale shows bioluminescence (photons/second). **(H)** Quantification of bioluminescence for the time point shown in **(G)**. Data normalized to signal on day 0 for each animal. p-value calculated using the Wilcoxon rank-sum test.

DOI: [10.7554/eLife.02734.005](https://doi.org/10.7554/eLife.02734.005)



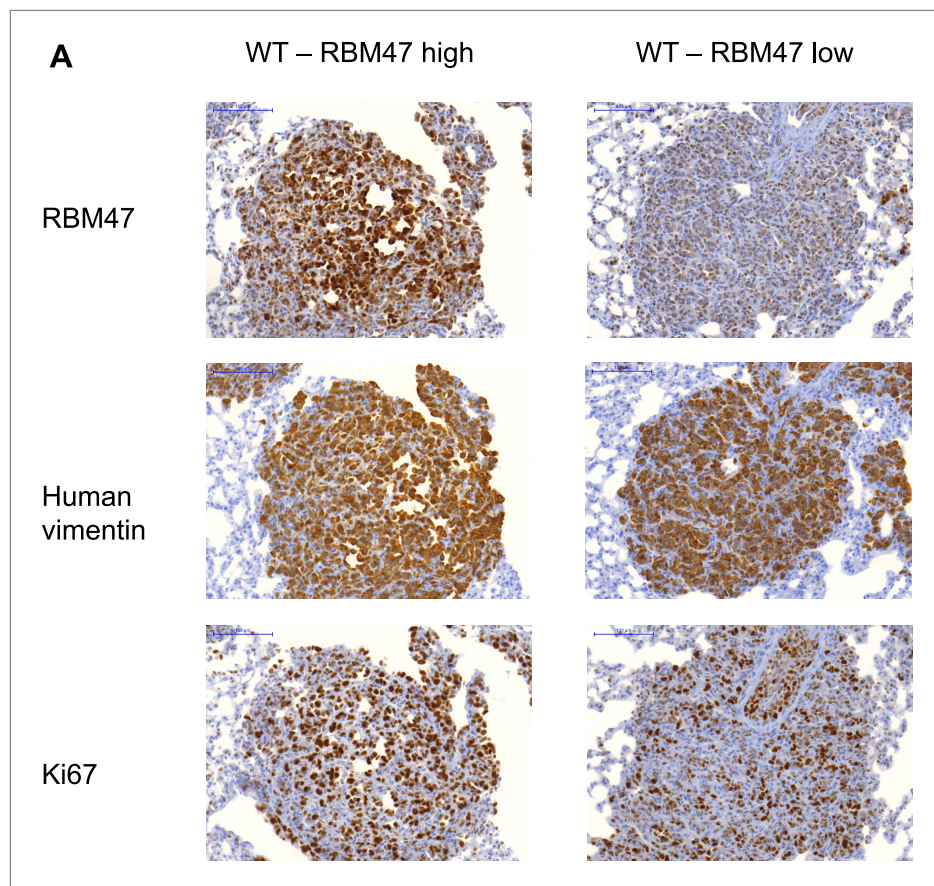
**Figure 2—figure supplement 1.** RBM47 suppresses breast cancer progression.

DOI: [10.7554/eLife.02734.006](https://doi.org/10.7554/eLife.02734.006)



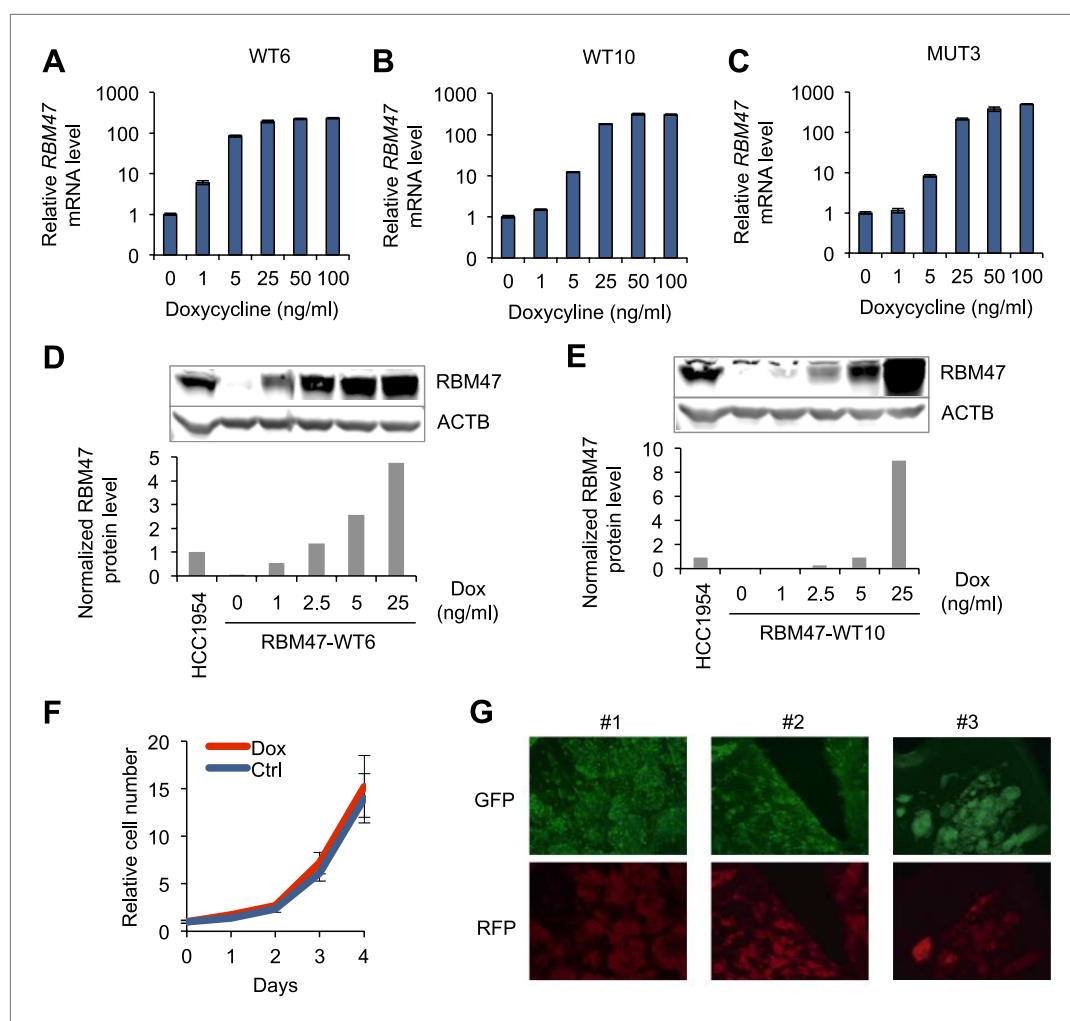
**Figure 3.** Clonal heterogeneity in RBM47 sensitivity. (A) A lung metastatic nodule with strong RBM47 expression in a mouse inoculated with RBM47-transduced MDA231-LM2 cells. RBM47 protein expression detected by immunohistochemistry using an antibody against RBM47. (B) A lung metastatic nodule with weak RBM47 expression in a mouse inoculated with RBM47-transduced MDA231-LM2 cells. RBM47 protein expression detected by immunohistochemistry. Note the clearly reduced staining when compared to panel (A). (C) A lung metastatic nodule with weak RBM47 expression in a mouse inoculated with RBM47<sup>I281fs</sup>-transduced MDA231-LM2 cells. RBM47 protein expression detected by immunohistochemistry. Staining intensity similar to that seen in panel (B). (D) RBM47 expression in normal mouse lung. RBM47 protein expression detected by immunohistochemistry. Note small cells with strong RBM47 expression, similar to those seen in panels (B) and (C). (E) Brain metastasis free survival as determined by in vivo bioluminescence imaging in mice after intracardiac inoculation of WT10 cells. The RBM47 group received doxycycline in diet. p-value calculated using the log-rank test. N = 9 for both groups. (F) Representative ex vivo bioluminescence images from brain metastases of the experiment shown in (E). The color scale shows bioluminescence (photons/second). (G) Ex vivo quantification of bioluminescence from brain metastases with and without RBM47 reintroduction. WT10 data from the experiment shown in (E). WT6 data from a similar experimental setup. p-value calculated by two-tailed Student's t test. N = 9 for Ctrl group, N = 10 for RBM47 group. (H) Brain metastasis free survival as determined by in vivo bioluminescence imaging in mice after intracardiac inoculation of MUT3 cells. The I281fs group received doxycycline in diet. p-value calculated using the log-rank test. N = 9 for Ctrl group, N = 7 for RBM47<sup>I281fs</sup> group. (I) Representative ex vivo bioluminescence images from brain metastases of the experiment shown in (H). The color scale shows bioluminescence (photons/second). (J) Ex vivo quantification of bioluminescence from brain metastases of the experiment shown in (H). p-value calculated by two-tailed Student's t test.

DOI: [10.7554/eLife.02734.007](https://doi.org/10.7554/eLife.02734.007)



**Figure 3—figure supplement 1.** RBM47 expression in experimental metastases.

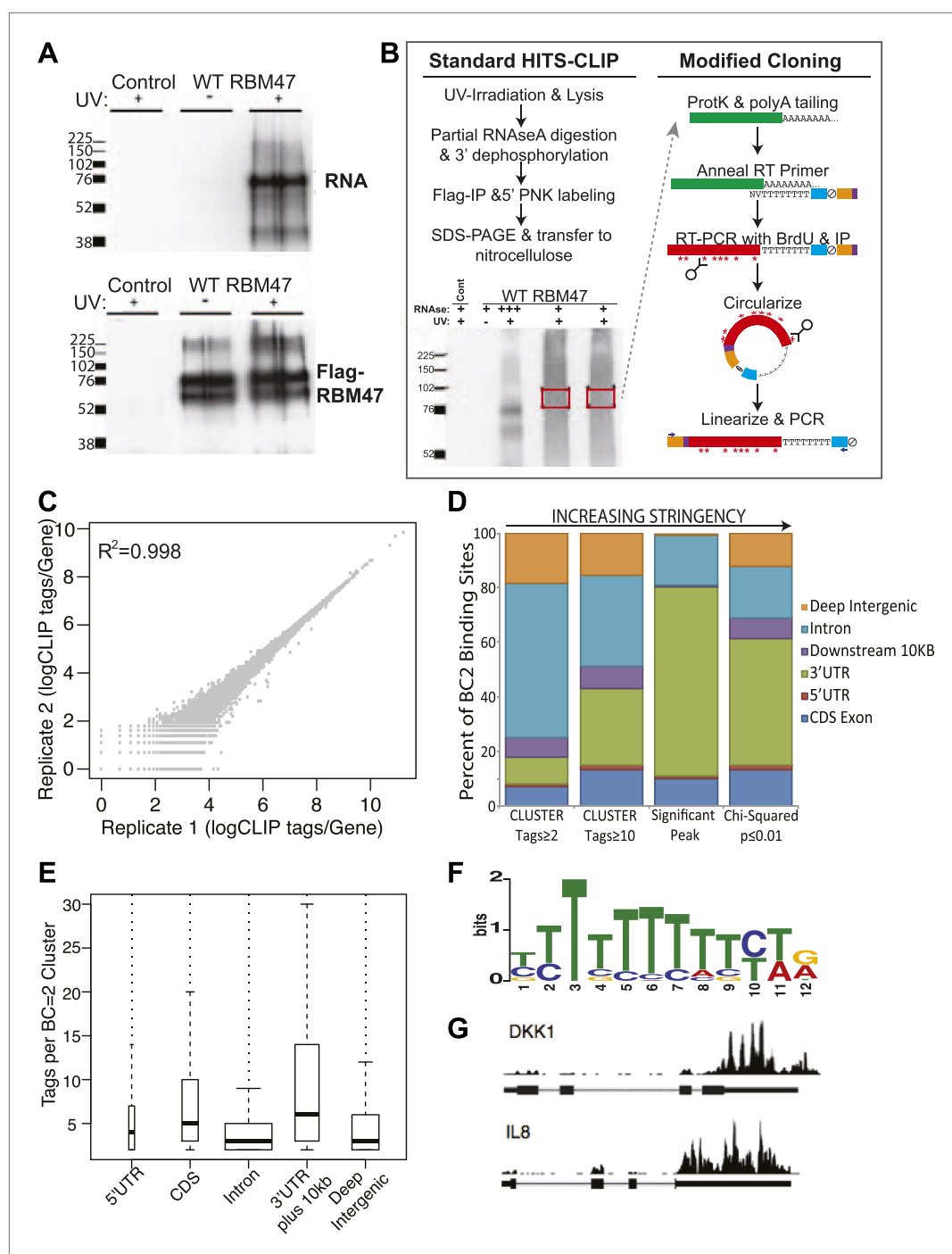
DOI: [10.7554/eLife.02734.008](https://doi.org/10.7554/eLife.02734.008)



**Figure 3—figure supplement 2.** Model systems with inducible RBM47 expression.

DOI: [10.7554/eLife.02734.009](https://doi.org/10.7554/eLife.02734.009)





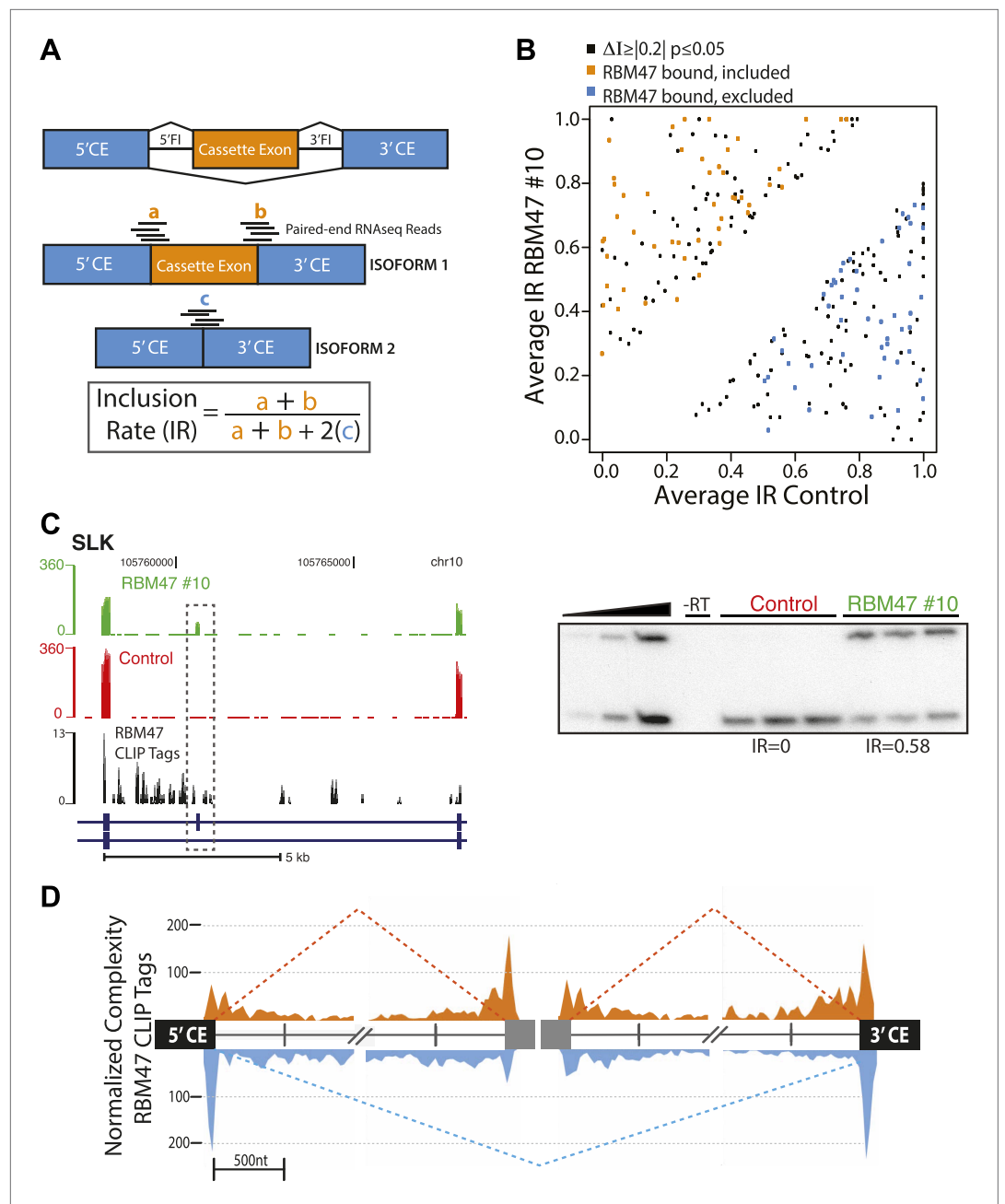
**Figure 4.** HITS-CLIP identifies genome-wide RBM47 binding maps. **(A)** Radiolabelled RNA is detectable in RBM47-expressing 231-BrM2 metastatic cells that have been UV-irradiated indicating in vivo RNA binding ability. No RNA is detected in non-crosslinked cells despite the presence of ample immunoprecipitated Flag-RBM47 protein. No RNA or protein is detected in control 231-BrM2 transduced with empty vector. Samples run in duplicate. **(B)** Schematic of the modified HITS-CLIP protocol showing autoradiogram of duplicate Flag-RBM47 samples used. Purified RBM47-bound RNA fragments (green) were polyA tailed and reverse transcribed in the presence of Brd-U using a polydT-NV primer encoding the full sense sequence of the Illumina reverse sequencing primer (blue), an abasic furan that serves as an Apel cut site ( $\xi$ ), a partial reverse complement to the Illumina forward sequencing primer (orange), and two hexamer sequences (purple): a known-sequence index for multiplexing and a degenerate barcode used to distinguish unique cDNA clones from PCR duplicates. cDNA were

Figure 4. Continued on next page

*Figure 4. Continued*

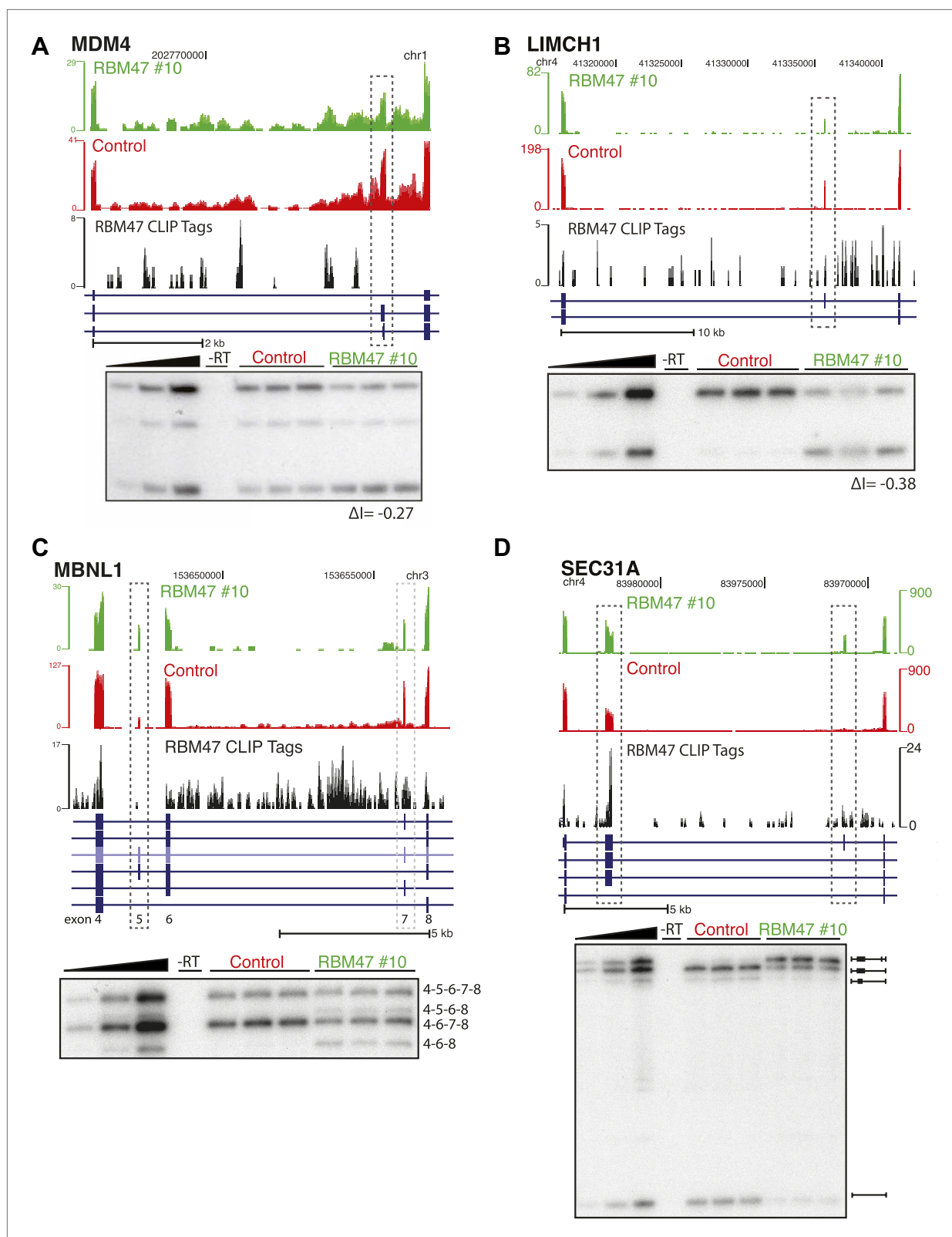
stringently purified, circularized and linearized using *Apel* to bring the Illumina sequence into correct orientation with respect to the cloned fragment, and the samples PCR amplified and deep sequenced. **(C)** RBM47 HITS-CLIP is highly reproducible between replicate experiments at the level of unique CLIP tags per transcript. **(D)** Increasing the stringency of biologically reproducible RBM47 binding site definition reveals predominant binding in 3'UTRs and intronic regions of target transcripts, with the most robust binding (tags per binding site) evident in 3'UTRs. **(E)** Distribution of tags number per biologically reproducible cluster in coding and non-coding regions of RBM47-targeted transcripts reveals a bimodal binding pattern between 3'UTRs and introns, with the latter having large numbers of reproducible yet less robust binding. **(F)** MEME analysis reveals an enrichment for polyU sequences (50 sites,  $p=2.4e^{-16}$ ) in the  $\pm 10$  nt foot print region surrounding reproducible RBM47 deletion CIMS (357 sites with  $\geq 5$  mutations,  $FDR \leq 0.01$ , **Zhang and Darnell, 2011**). **(G)** Widespread RBM47 binding is seen in target transcripts, as exemplified by binding patterns seen in the 3'UTRs of DKK1 and IL8.

DOI: [10.7554/eLife.02734.010](https://doi.org/10.7554/eLife.02734.010)



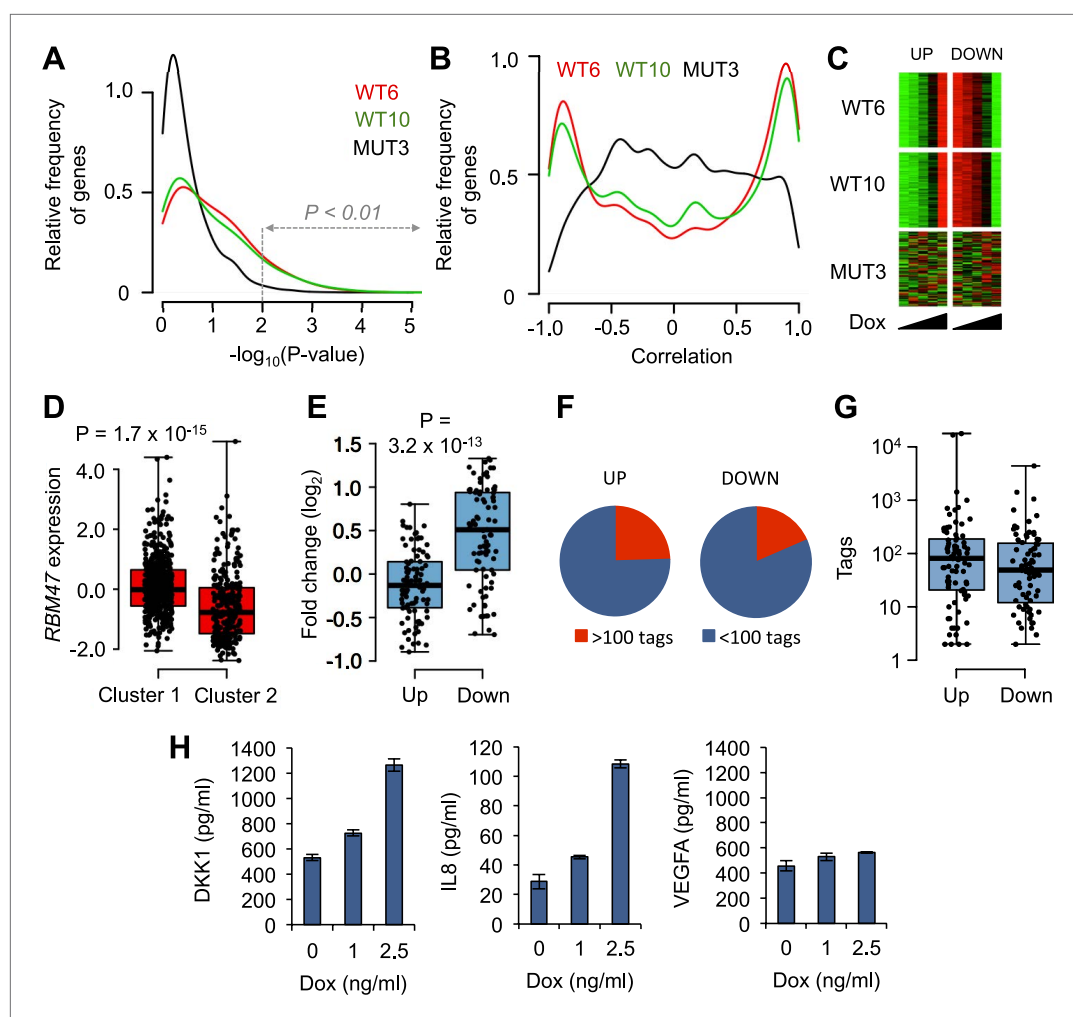
**Figure 5.** RBM47 regulates alternative splicing. **(A)** Schematic showing the method used to calculate alternative exon inclusion rates from paired-end RNA-seq data. 5'CE–5' flanking constitutive exon, 3'CE–3' flanking constitutive exon, 5'FI–5' flanking intron, 3'FI–3' flanking intron. **(B)** Scatter plot of all expressed alternatively spliced CA exons showing RBM47-dependent change in inclusion (black,  $\geq 10$  RNA-seq reads spanning exon–exon junctions,  $\Delta I \geq |0.2|$ ,  $p \leq 0.05$ ) with orange points indicating RBM47-bound and included CA exons, and blue points indicating bound and excluded exons, respectively. CA exons were considered bound given a total of tags  $\geq 10$  in BC2 tags  $\geq 5$  clusters mapping to the region spanning the start of the 5'CE to the end of the 3'CE. p-values calculated by Fisher's exact test using total isoform 1 and total isoform 2 RNA-seq reads in each condition. **(C)** Left panel shows a section of the SLK transcript (blue) that includes a CA exon (grey box). The top two panels show RNA-seq data from WT10 (green) and control cells (red), with RBM47 HITS-CLIP tags mapping to the region shown beneath in black. Increased RNA-seq signal corresponding to the CA exon is seen in the presence of RBM47 expression, while robust binding is evident in the 5'FI. Independent RT-PCR validation of this splice is shown in the right panel, with IR calculated using ImageJ analysis of autoradiograms. **(D)** Normalized complexity map of RBM47-dependent alternative splicing of CA exons. Orange and blue peaks represent binding associated with RBM47-dependent exon inclusion and exclusion, respectively.

DOI: [10.7554/eLife.02734.011](https://doi.org/10.7554/eLife.02734.011)



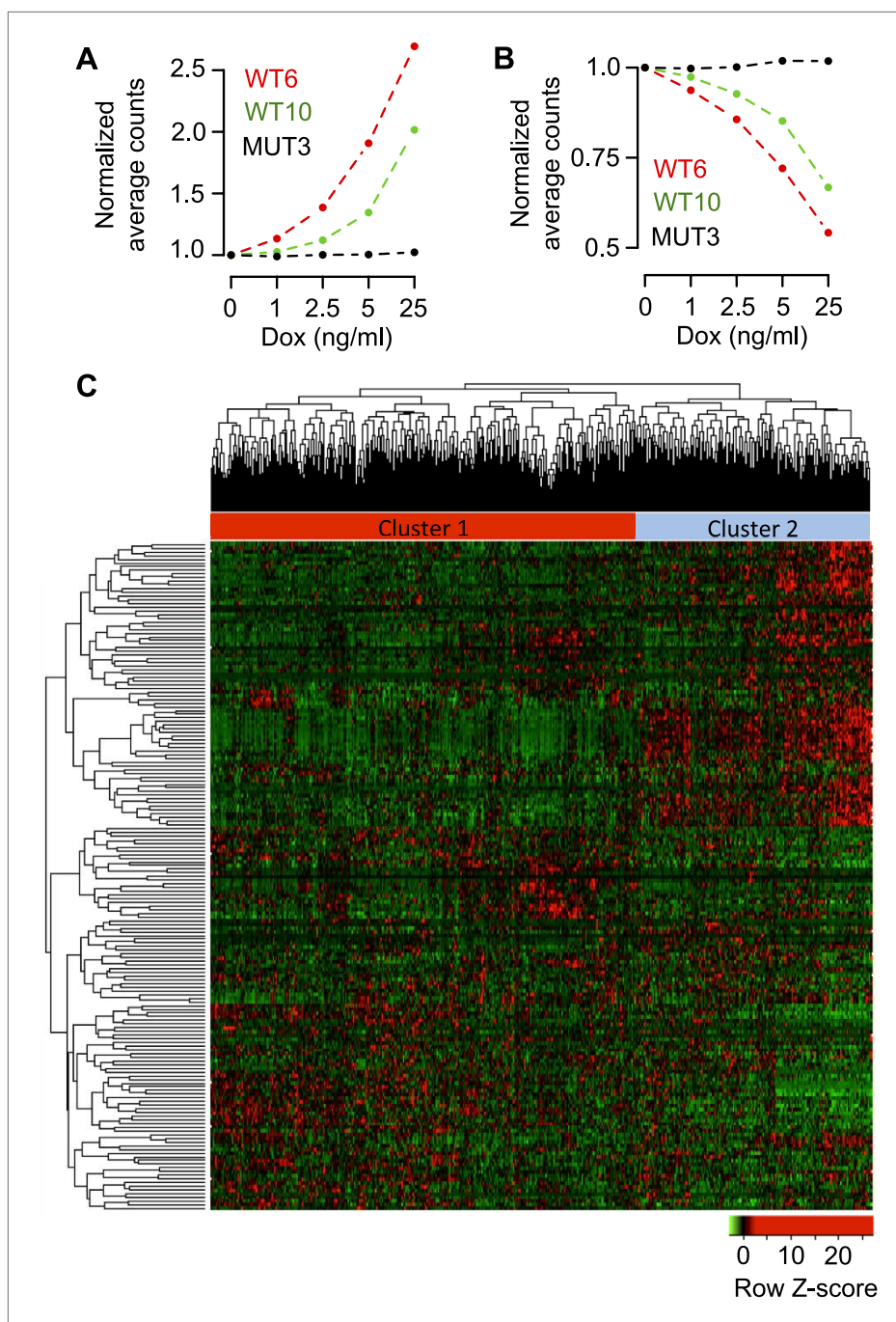
**Figure 5—figure supplement 1.** RBM47-dependent splicing events.

DOI: [10.7554/eLife.02734.012](https://doi.org/10.7554/eLife.02734.012)



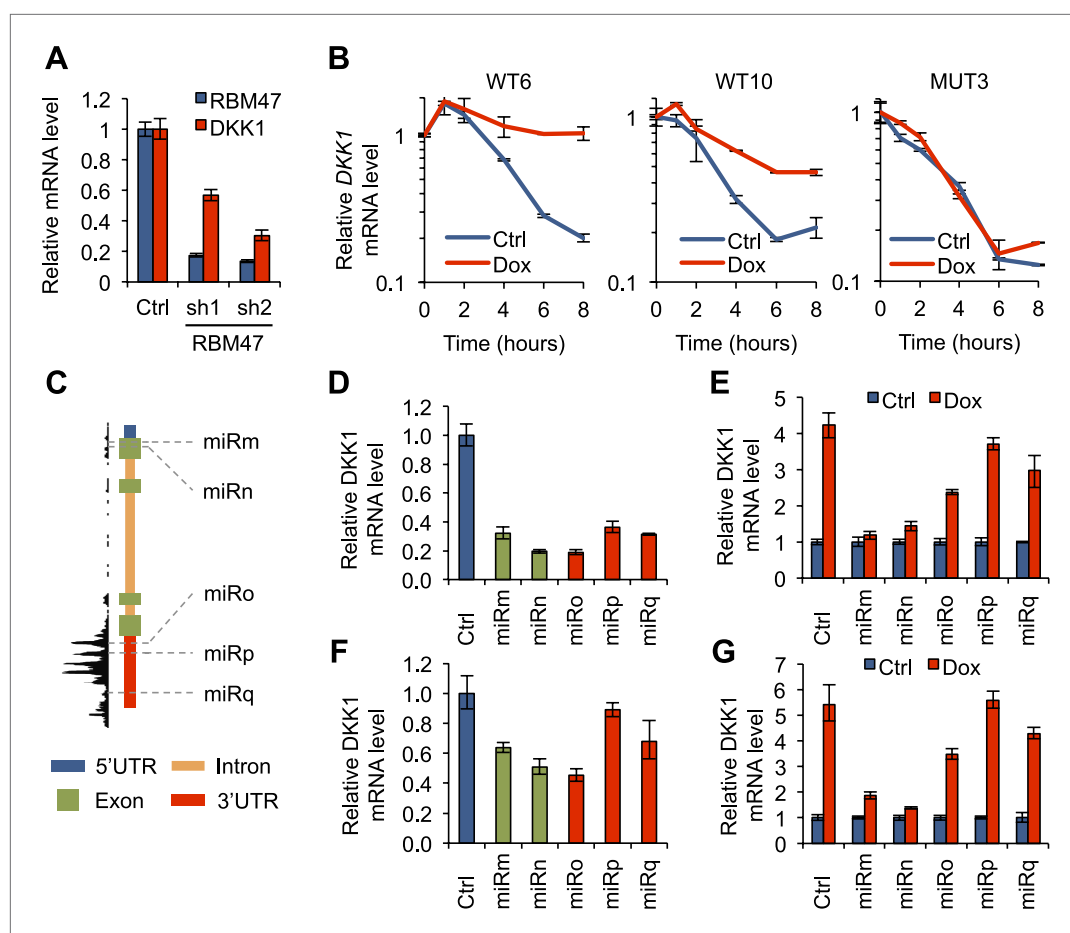
**Figure 6.** RBM47-induced changes in mRNA levels. **(A)** Distribution of p-values from correlation analysis of doxycycline concentration and gene expression for all genes in WT6, WT10 and MUT3 cell lines, respectively. Global gene expression determined by RNA-seq. **(B)** Distribution of correlation coefficients between doxycycline concentration and gene expression in WT6, WT10 and MUT3 cell lines, respectively. **(C)** Heat maps showing the top 102 positively (UP) correlated and 92 negatively (DOWN) correlated genes with RBM47 expression in WT6 and WT10 cells. The expression of these genes does not correlate with RBM47<sup>71280fs</sup> expression in the MUT3 cells. **(D)** RBM47 mRNA expression in the TCGA cohort of breast cancer samples classified by the clusters shown in **Figure 6—figure supplement 1C**. p-value determined by two-tailed Student's *t* test. **(E)** Fold change between Cluster 1 and Cluster 2 shown for the 102 positively and 92 negatively correlated RBM47 target genes, respectively. Positive fold change shows higher expression in Cluster 2, which has lower expression of RBM47. The genes that are induced upon RBM47 reintroduction tend to have lower expression in Cluster 2, and the genes that show lower expression upon RBM47 reintroduction tend to have higher expression in Cluster 2. This is in line with the experimental results shown in panel **(C)**. P-value determined by two-tailed Student's *t* test. **(F)** Pie charts showing the fraction of target genes with more than 100 RBM47 tags for both the 102 positively and 92 negatively correlated RBM47 target genes. **(G)** Tags per transcript plotted for both the positively and negatively correlated RBM47 target genes that showed more than 1 tag. The only two binding partners with  $>10^4$  tags represent DKK1 and IL8, respectively. **(H)** Secreted DKK1 and IL8 protein levels determined by ELISA in WT6 cells treated with the indicated doxycycline concentrations. VEGFA used as a control.

DOI: [10.7554/eLife.02734.013](https://doi.org/10.7554/eLife.02734.013)



**Figure 6—figure supplement 1.** Transcriptomic signature of RBM47 reintroduction.

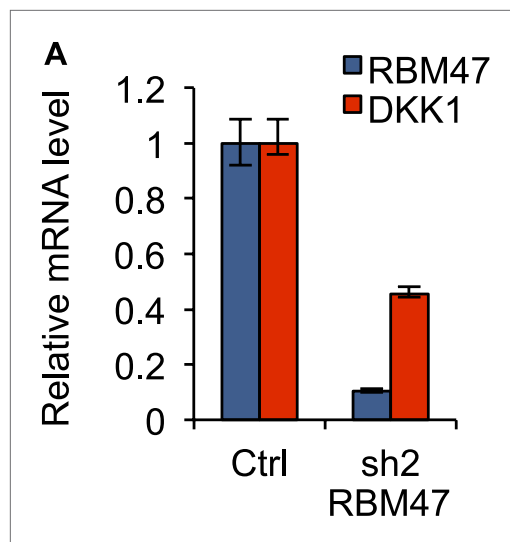
DOI: [10.7554/eLife.02734.014](https://doi.org/10.7554/eLife.02734.014)



**Figure 7.** RBM47 modulates *DKK1* mRNA stability. **(A)** *RBM47* and *DKK1* mRNA expression measured by quantitative real-time RT-PCR in SKBR3 cells expressing either control vector (pGIPZ) or hairpins against *RBM47* (shRNA1 and shRNA2). Error bars represent 95% confidence intervals obtained from multiple PCR reactions. **(B)** *DKK1* mRNA stability determined by measuring mRNA levels by quantitative real-time RT-PCR in WT6, WT10 and MUT3 cells, treated with or without doxycycline, after inhibition of transcription with actinomycin D. Data normalized to time point 0. Error bars represent 95% confidence intervals obtained from multiple PCR reactions. WT6:  $T_{1/2}$  Ctrl = 2.3 hr,  $T_{1/2}$  Dox = 9.8 hr; WT10:  $T_{1/2}$  Ctrl = 2.4 hr,  $T_{1/2}$  Dox = 5.4 hr; MUT3:  $T_{1/2}$  Ctrl = 2.1 hr,  $T_{1/2}$  Dox = 2.2 hr. **(C)** Schematic showing the locations of different *DKK1* miR-constructs in relation to *DKK1* genomic locus and RBM47 binding patterns. miRm and miRn target exon 1 that is not bound by RBM47. miRo, miRp and miRq target *DKK1* 3'UTR that is strongly bound by RBM47. **(D)** *DKK1* mRNA expression measured by quantitative real-time RT-PCR in WT6 cells expressing either control vector (pGIPZ) or the five *DKK1*-targeting miR constructs shown in panel (C). Error bars represent 95% confidence intervals obtained from multiple PCR reactions. **(E)** *DKK1* mRNA expression measured by quantitative real-time RT-PCR in WT6 cells expressing either control vector (pGIPZ) or the five *DKK1*-targeting miR constructs, with or without doxycycline treatment. Data normalized to the non-treated control for each cell line separately. Error bars represent 95% confidence intervals obtained from multiple PCR reactions. **(F)** *DKK1* mRNA expression measured by quantitative real-time RT-PCR in WT10 cells expressing either control vector (pGIPZ) or the five *DKK1*-targeting miR constructs shown in panel (C). Error bars represent 95% confidence intervals obtained from multiple PCR reactions. **(G)** *DKK1* mRNA expression measured by quantitative real-time RT-PCR in WT10 cells expressing either control vector (pGIPZ) or the five *DKK1*-targeting miR constructs, with or without doxycycline treatment. Data normalized to the non-treated control for each cell line separately. Error bars represent 95% confidence intervals obtained from multiple PCR reactions.

DOI: 10.7554/eLife.02734.015

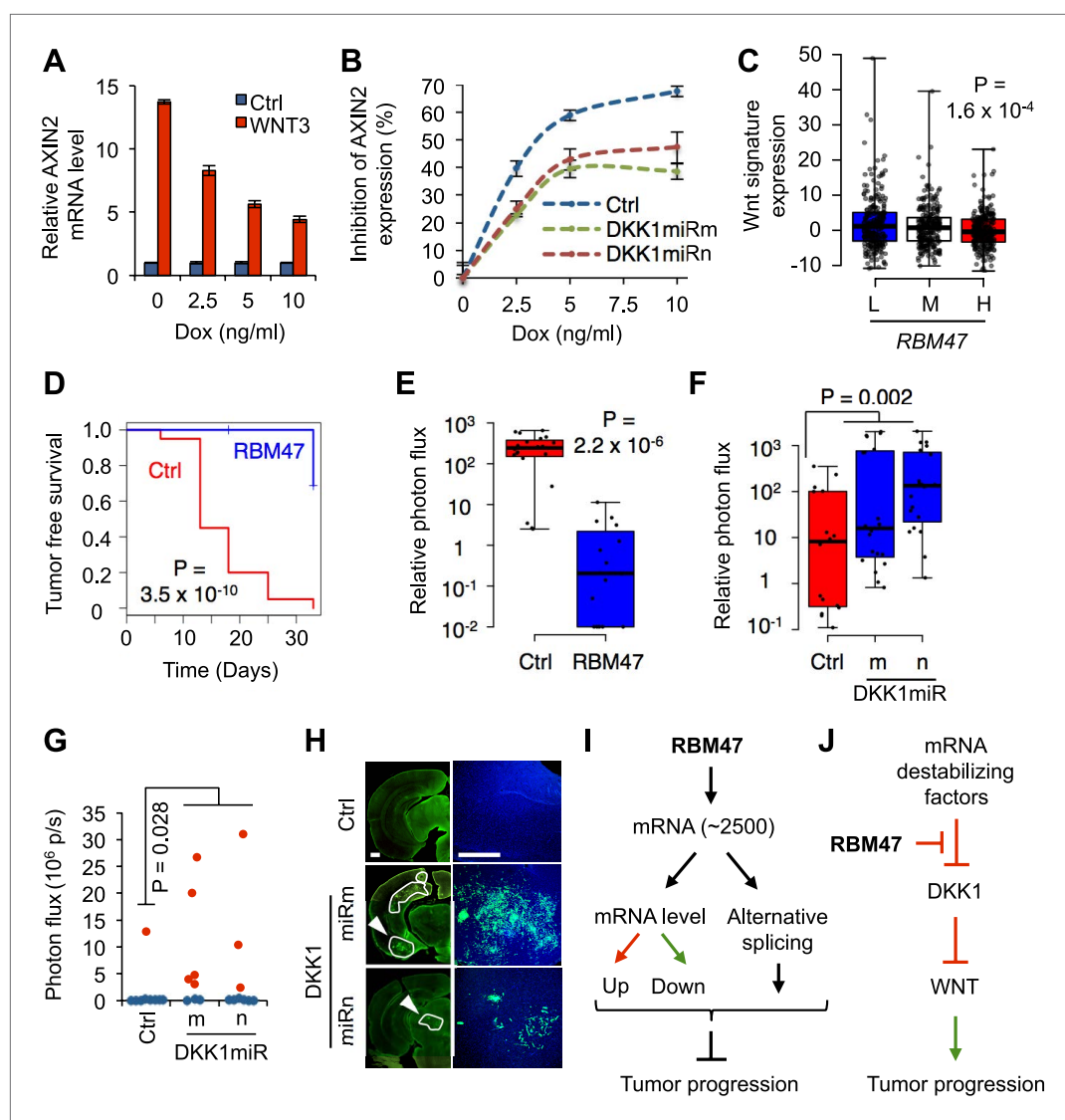




**Figure 7—figure supplement 1.** Effects of RBM47 knockdown on DKK1 mRNA.

DOI: [10.7554/eLife.02734.016](https://doi.org/10.7554/eLife.02734.016)



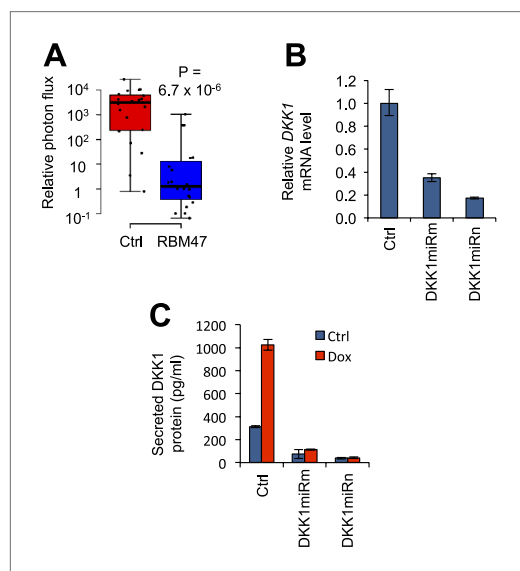


**Figure 8.** RBM47 suppresses tumor progression via Wnt inhibition. (A) *AXIN2* mRNA levels determined by quantitative real-time RT-PCR in WT6 cells treated with recombinant WNT3A in the presence of increasing concentrations of doxycycline. Error bars represent 95% confidence intervals obtained from multiple PCR reactions. (B) Normalized level of *AXIN2* mRNA inhibition as determined by quantitative real-time RT-PCR in WT6 cells transduced with either control vector or shRNA miR constructs targeting the first exon of *DKK1*. The cells were treated with recombinant WNT3A in the presence of increasing concentrations of doxycycline. Error bars represent 95% confidence intervals obtained from multiple PCR reactions. (C) Wnt pathway activity assessed in the TCGA cohort of primary breast tumors, grouped by *RBM47* expression tertiles (L, low; M, medium; H, high). Wnt signature value calculated as sum of z-scores for a curated set of 16 Wnt target genes in breast cancer. p-value determined by linear regression analysis. (D) Mammary tumor re-initiation assay. 5,000 WT6 cells implanted orthotopically in mice. RBM47 induced by doxycycline feed. Tumor growth detected by bioluminescence. p-value determined by the log-rank test. N = 20 tumors for each group. (E) Quantification of mammary tumor burden by in vivo bioluminescence imaging on day 33 of the experiment shown in (D). Data normalized to day 0 for each tumor. p-value calculated by the Wilcoxon rank-sum test. (F) Quantification of mammary tumor burden by in vivo bioluminescence imaging in mice inoculated with 231-Brm2 cells transduced with either control (pGIPZ) or *DKK1*-targeting shRNA miR constructs. Data normalized to day 0 for each tumor. p-value calculated by one-tailed Wilcoxon rank-sum test. (G) Quantification of ex vivo brain bioluminescence shown for mice inoculated intracardially with WT6 cells transduced with either control (pGIPZ) or *DKK1*-targeting shRNA miR constructs in the presence of RBM47, that is doxycycline in diet. One out of 9 (11%) control mice developed robust brain metastasis whereas 8/17 (47%) mice in the *DKK1* knockdown groups showed metastasis. p-value calculated by one-tailed Student's t test. (H) Representative Figure 8. Continued on next page

Figure 8. Continued

images of coronal brain sections analyzed for GFP immunofluorescence from the experiment shown in panel (G). Lesion contours are marked in white. Arrowheads indicate the lesions shown in higher magnification on the right; a similar brain area is shown for the control group. Scale bar 500  $\mu\text{m}$ . (I) At the global level, RBM47 binds to ~2500 target mRNAs. However, the abundance or alternative splicing of only a fraction of these change depending on RBM47 status. The target genes represent molecules from various signaling pathways. The net effect of growth promoting and inhibiting alterations determine whether RBM47 loss is beneficial for a particular cancer clone. (J) At the target mRNA level, the effects of RBM47 are dependent on the presence of other factors that modulate mRNA processing. Hence, the phenotype of RBM47 loss depends on the intracellular molecular milieu on a per transcript basis. This is exemplified by the interaction of RBM47 with *DKK1* mRNA.

DOI: [10.7554/eLife.02734.017](https://doi.org/10.7554/eLife.02734.017)



**Figure 8—figure supplement 1.** DKK1 as a mediator of RBM47-dependent tumor suppression.

DOI: [10.7554/eLife.02734.018](https://doi.org/10.7554/eLife.02734.018)