Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/5/3/eaat9820/DC1

Supplementary Materials for

The endogenous retrovirus-derived long noncoding RNA TROJAN promotes triple-negative breast cancer progression via ZMYND8 degradation

Xi Jin, Xiao-En Xu*, Yi-Zhou Jiang, Yi-Rong Liu, Wei Sun, Ya-Jie Guo, Yi-Xing Ren, Wen-Jia Zuo, Xin Hu, Sheng-Lin Huang, Hong-Jie Shen, Fei Lan, Yun-Fei He, Guo-Hong Hu, Gen-Hong Di, Xiang-Huo He, Da-Qiang Li, Suling Liu*, Ke-Da Yu*, Zhi-Ming Shao*

*Corresponding author. Email: zhimingshao@yahoo.com (Z.-M.S.); xiaoenxu@163.com (X.-E.X.); yukeda@163.com (K.-D.Y.); suling@fudan.edu.cn (S.L.)

Published 6 March 2019, *Sci. Adv.* **5**, eaat9820 (2019)
DOI: 10.1126/sciadv.aat9820

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

(available at advances.sciencemag.org/cgi/content/full/5/3/eaat9820/DC1)

SUPPLEMENTARY FIGURE LEGENDS

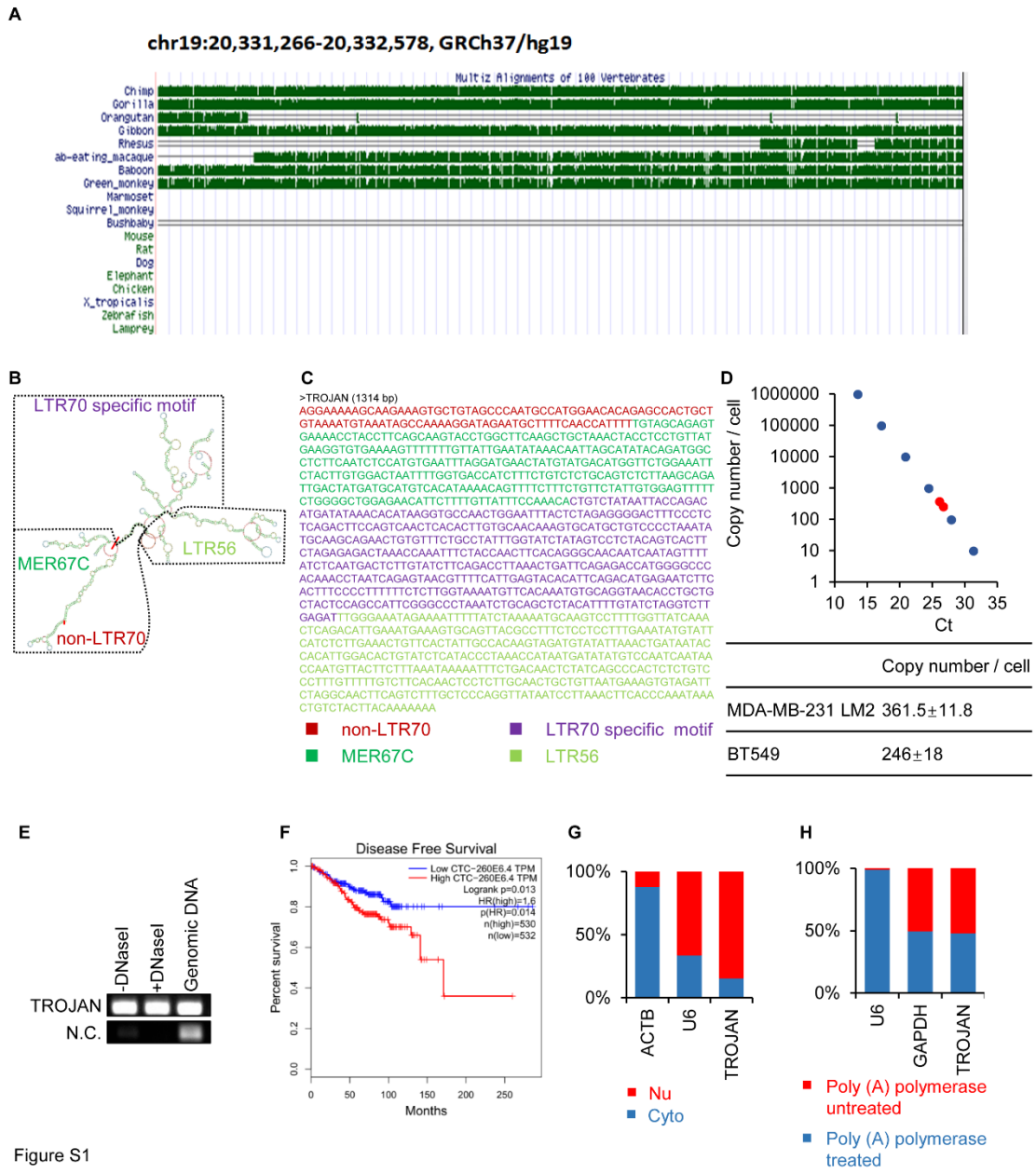


Figure S1

Fig. S1. TROJAN properties. (**A**) Sequence homology of TROJAN among different species (UCSC genome browser database). (**B**) TROJAN secondary structure predicted by RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). (**C**) The full-length sequence of TROJAN identified by RACE assay. (**D**) The estimated copy number of TROJAN per cell. The red dots indicate molecules per cell. Total RNA was collected from 10^5 MDA-MB-231 LM2 and BT549 cells and analyzed by quantitative reverse transcription (qRT)-PCR. We first calculated that the

total RNA per cell was approximately 125 pg. The weight of 1 mol TROJAN (6.02×10^{23} copies) was 292544.69 g. A standard curve (blue dots) was generated by serial dilution of *in vitro* transcribed TROJAN RNA. (**E**) Validation of TROJAN expression using a pair of primers. Prior to reverse transcription, total MDA-MB-231 LM2 RNA was extracted and treated with or without deoxyribonuclease I (DNase I). N.C.: a negative control primer targeting a nontranscribed location in the genome of breast cancer cell lines (LINC01614). (**F**) Kaplan-Meier analysis of the disease-free survival based on TROJAN (CTC-260E6.4) in the TCGA data set. A log-rank test was used to determine statistical significance between the high expression group (n = 530) and the low expression group (n = 532). (**G**) TROJAN RNA expression in different subcellular fractions of MDA-MB-231 LM2 cells. Red indicates the nuclear fraction (Nu), and blue indicates the cytoplasmic (Cyto) fraction. The percentages were calculated by qRT-PCR. (**H**) Relative TROJAN expression in total RNA extracted and treated with or without poly(A) polymerase. The percentages were calculated by qRT-PCR.

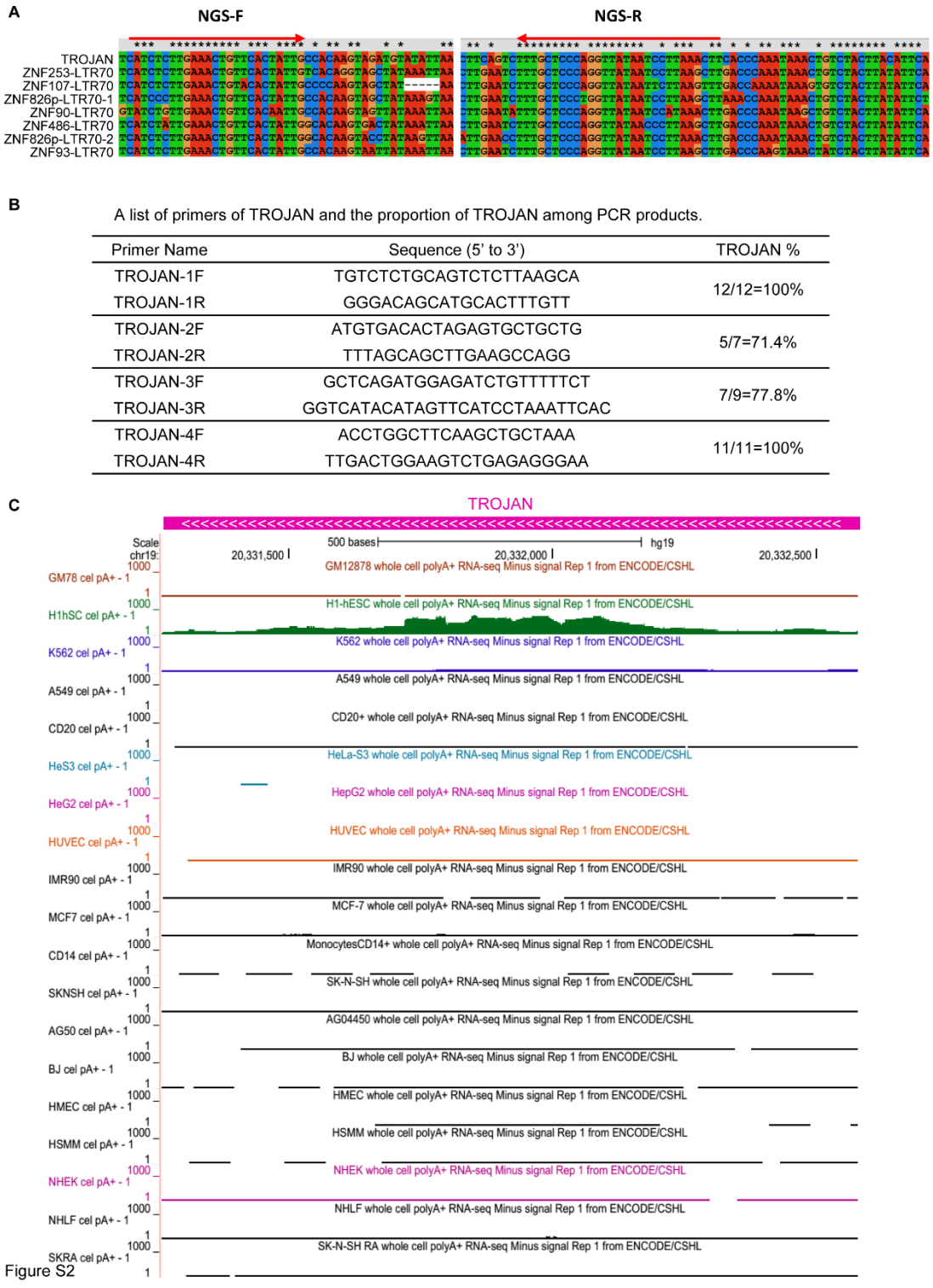


Fig. S2. Expression pattern of TROJAN. (**A**) The compatible pair of primers covering eight expressed LTR70s used for RNA-Seq to analyze the constituent ratio of LTR70 in Fig. 1. (**B**) A list of primers for TROJAN and the proportion of TROJAN in the PCR products. (**C**) The RNA-Seq data were acquired from the UCSC genome browser database (Long RNA-seq from ENCODE/Cold Spring Harbor Lab).

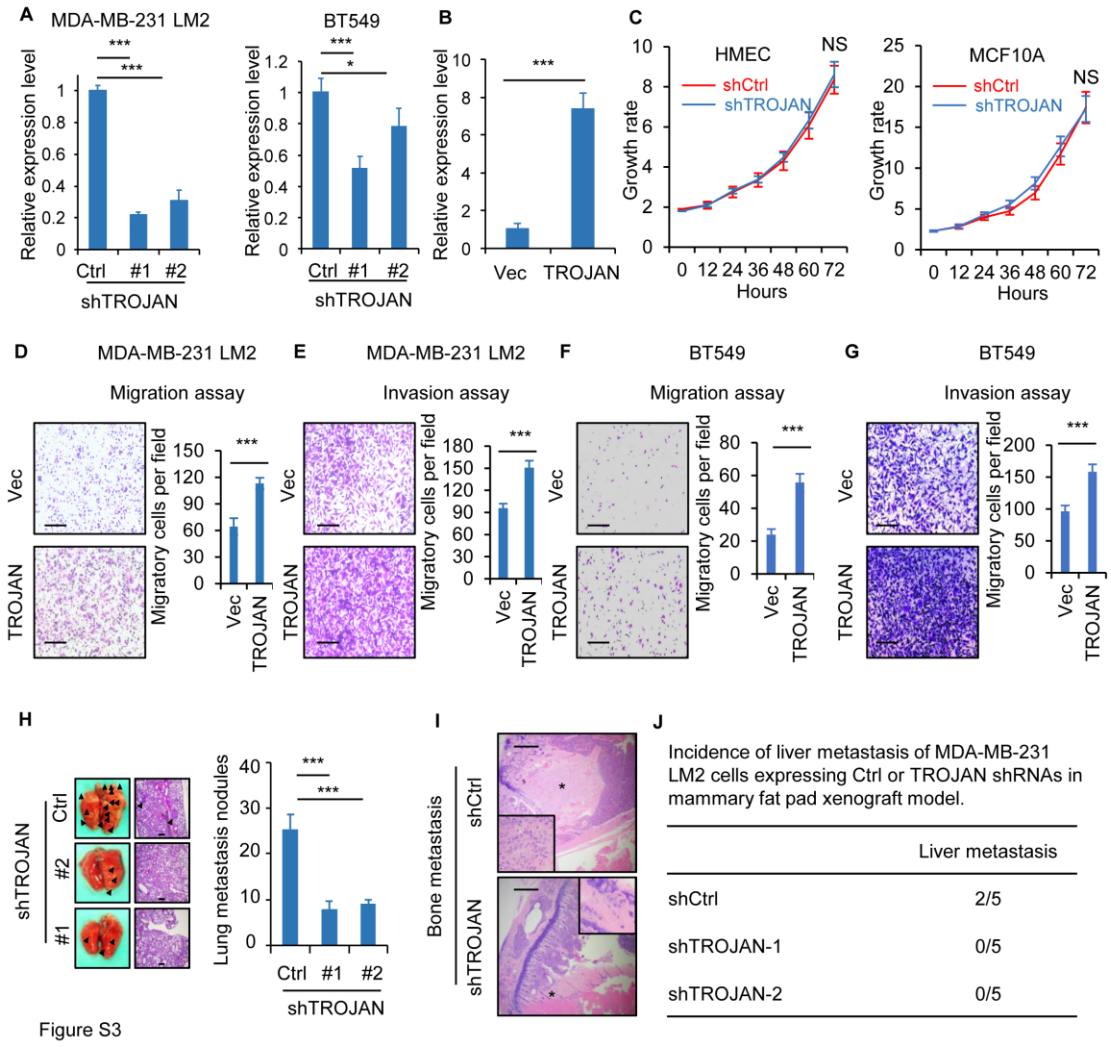


Fig. S3. TROJAN phenotype. (**A**) Quantitative reverse transcription (qRT)-PCR analysis of relative TROJAN transcription in MDA-MB-231 LM2 and BT549 cells expressing control (Ctrl) or TROJAN short hairpin RNA (shRNA). The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. (**B**) qRT-PCR detection of relative TROJAN transcription in MDA-MB-231 LM2 cells overexpressing a control vector (Vec) or TROJAN. The data are presented as the mean ± the s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. (**C**) *In vitro* growth curves of HMEC and MCF-10A cells expressing Ctrl or TROJAN shRNA. The data are presented as the mean ± s.d.; n = 3 independent experiments. (**D**) *In vitro* transwell migration assay of MDA-MB-231 LM2 cells expressing Vec or TROJAN. The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. Scale bar: 200 μm.

(**E**) *In vitro* transwell invasion assay of MDA-MB-231 LM2 cells expressing Vec or TROJAN. The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. Scale bar: 200 μm. (**F**) *In vitro* transwell migration assay of BT549 cells expressing Vec or TROJAN. The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. Scale bar: 200 μm. (**G**) *In vitro* transwell invasion assay of BT549 cells expressing Vec or TROJAN. The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. Scale bar: 200 μm. (**H**) The numbers of lung metastases and metastatic nodules are shown. The data are presented as the mean ± s.d.; two-tailed unpaired Student's *t*-test. Scale bar: 200 μm. (**I**) H&E staining of bone metastasis of SCP2 cells expressing Ctrl or TROJAN shRNA. Scale bar: 50 μm. The asterisks indicate bone metastasis. (**J**) Incidence of liver metastasis of MDA-MB-231 LM2 cells expressing Ctrl or TROJAN shRNA in the mammary fat pad xenograft model. *p < 0.05, ***p < 0.001. NS: not significant.

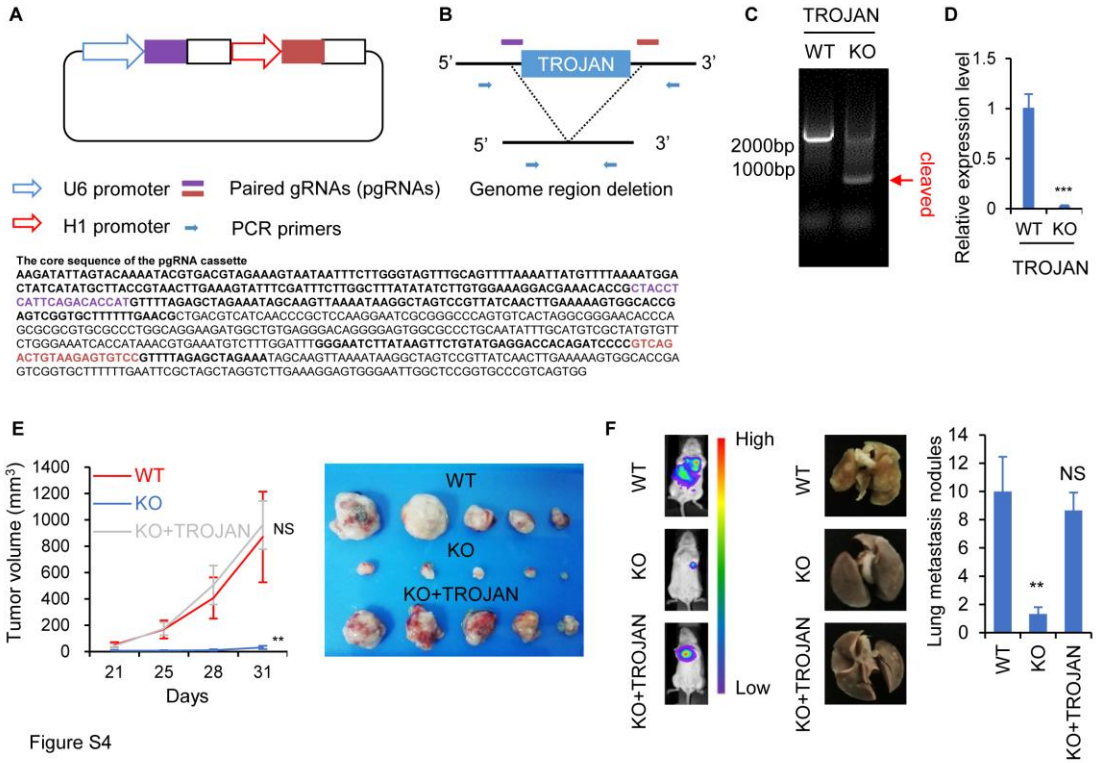


Fig. S4. Construction of the TROJAN knockout cell line. (**A**) Schematic diagram of the modification of the plasmid LentiCRISPRv2 and the construction of a plasmid

expressing paired guide RNAs (pgRNAs) flanking TROJAN. (**B**) Schematic diagram of the deletion of TROJAN. (**C**) PCR products generated from the genomic sequence around TROJAN showing the efficiency of targeted deletion in MDA-MB-231 LM2 cells expressing pgRNAs. TROJAN wild type (WT) or TROJAN knockout (KO). (**D**) Quantitative reverse transcription (qRT)-PCR analysis of relative TROJAN transcription levels in MDA-MB-231 LM2 cells with TROJAN WT or TROJAN KO. The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's t-test. (**E**) *In vivo* growth of MDA-MB-231 LM2 cells (31 days; n = 5) with TROJAN WT, TROJAN KO or TROJAN KO plus ectopically expressed TROJAN (KO + TROJAN). Tumor volume quantification (left) and representative tumor images (right) are shown. The data are presented as the mean ± s.d.; two-tailed unpaired Student's t-test. (**F**) *In vivo* intravenous xenograft mouse model of MDA-MB-231 LM2 cells (42 days; n = 3) with TROJAN WT, TROJAN KO or TROJAN KO plus exogenous TROJAN expression (KO + TROJAN). The relative bioluminescence intensity (BLI) indicates lung metastasis. The numbers of lung metastases and metastatic nodules are shown. The data are presented as the mean ± s.d.; two-tailed unpaired Student's t-test. $^{**}p < 0.01$, $^{***}p < 0.001$.

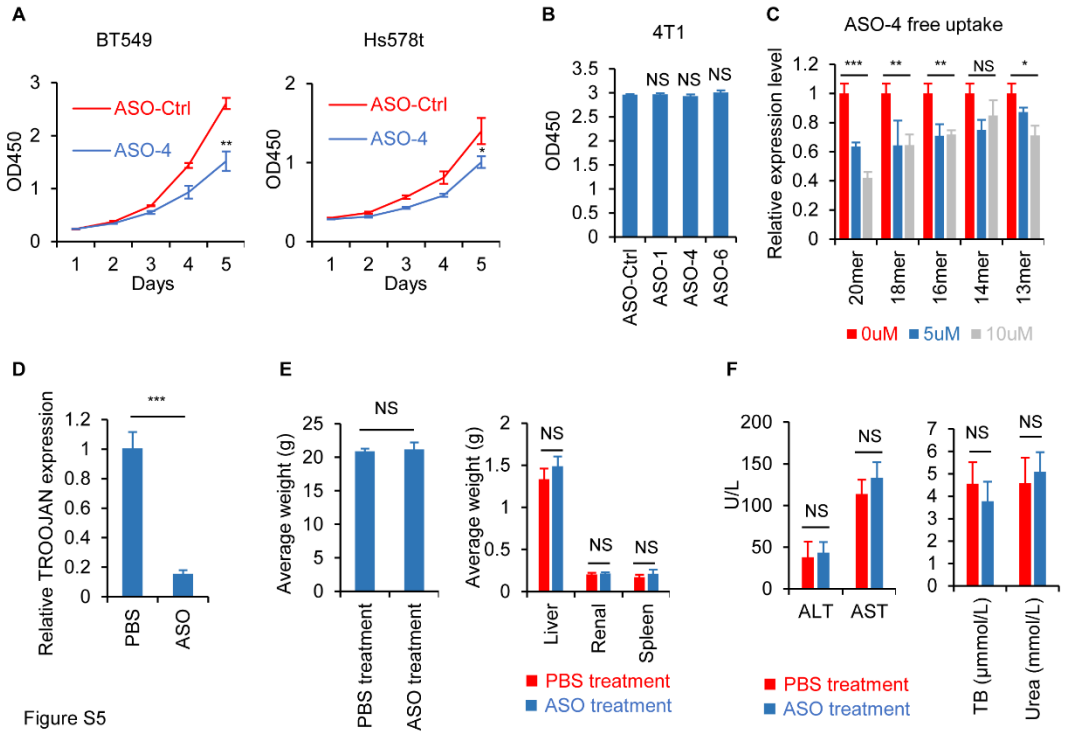
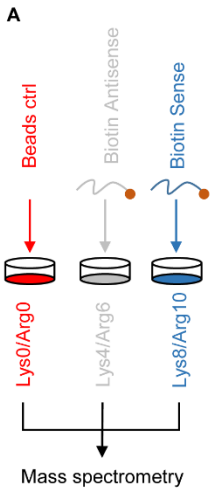


Fig. S5. Potential therapeutic role of TROJAN in breast cancer progression.

(**A**) *In vitro* growth assay of BT549 and Hs578t cells transfected with ASO-4 targeting TROJAN or Ctrl. The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. (**B**) *In vitro* growth assay of murine 4T1 breast cancer cells transfected with ASOs targeting TROJAN or Ctrl. CCK8 reagent was added after 72 hours, and the OD450 was determined. The data are presented as the mean ± s.d.; n = 3 independent experiments. (**C**) *In vitro* free uptake assay in MDA-MB-231 LM2 cells. Six different concentrations of ASO were assessed (0, 0.5, 1, 2.5, 5 and 10 μM). The 0 μM group was set as the control. Quantitative reverse transcription (qRT)-PCR detection of relative TROJAN expression. The data are presented as the mean ± s.d.; n = 3 independent experiments. (**D**) qRT-PCR analysis of relative TROJAN expression in the metastatic nodules of mice treated with an ASO targeting TROJAN or PBS. The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. ***p < 0.001. (**E**) Weights of the mice, as well as those of their livers, kidneys and spleens, in the PBS and ASO treatment groups. The data are presented as the mean ± s.d.; two-tailed unpaired Student's *t*-test. (**F**) Serum chemistry markers of liver function (ALT, AST

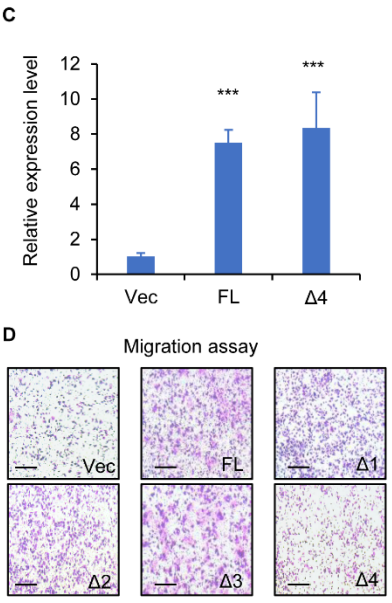
and TB) and renal function (urea) in the PBS and ASO treatment groups. ALT: alanine aminotransferase; AST: aspartate aminotransferase; and TB: total bilirubin. The data are presented as the mean ± s.d.; two-tailed unpaired Student's *t*-test. *p < 0.05, **p < 0.01 and ***p < 0.001. NS: not significant.

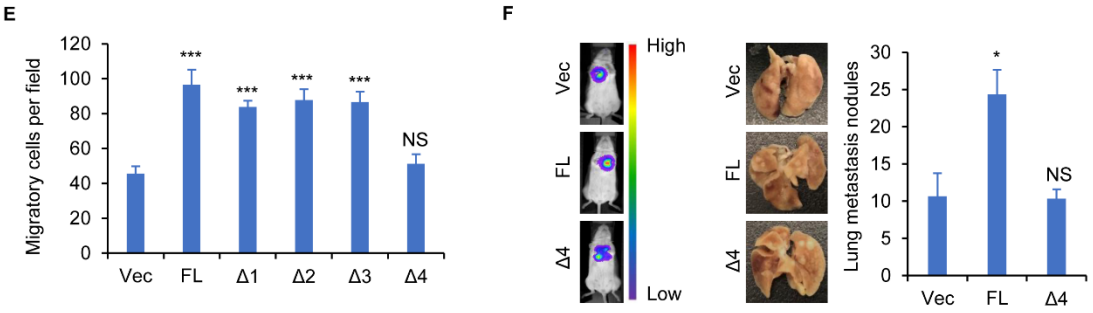


B

Top 10 potential TROJAN specific interaction proteins identified by mass spectrometry

Gene Names	Intensity		
	Ctrl	Antisense	Sense
ZMYND8	0	0	589390
OPHN1	0	0	340820
SEPT14;SEPT7	0	0	307700
CAPN1	0	0	190620
TUBA4A	0	0	154190
TMEM177	0	0	129670
FAM65A	0	0	126670
STMN1;STMN2	0	0	109330
SUB1	0	0	105010
DSTN	0	0	85993





G Highly homologous LTR56s

	LTR56 genome location
LTR56-1	chr17:56475049-56475486
LTR56-2	chr20:50392171-50392498
LTR56-3	chr8:99846205-99846626
LTR56-4	chr7:48150141-48150487
LTR56-5	chr8:6557630-6557994

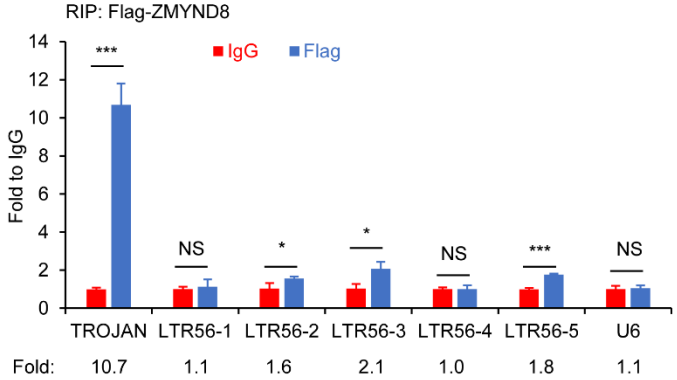


Figure S6

Fig. S6. TROJAN associates with the ZMYND8 protein. (**A**) Schematic diagram of the RNA pull-down assay combined with stable isotope labeling with amino acids (SILAC)-based quantitative proteomics and the top 10 potential TROJAN interaction proteins identified by mass spectrometry according to intensity. (**B**) Top 10 potential TROJAN-specific interacting proteins identified by mass spectrometry. (**C**) Quantitative reverse transcription (qRT)-PCR analysis of relative transcription levels from control vector (Vec), TROJAN full-length (FL) and mutant TROJAN (Δ4). The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. (**D**) *In vitro* transwell migration assay of MDA-MB-231 LM2 cells expressing control empty vector (Vec), full-length TROJAN (FL) or mutant TROJAN (Δ1-4). Scale bar: 200 μm. (**E**) Quantification of the transwell migration assay. The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. (**F**) *In vivo* intravenous xenograft mouse model of MDA-MB-231 LM2 cells (42 days; n = 3) expressing control empty vector (Vec), full-length TROJAN (FL) or mutant TROJAN (Δ4). The numbers of lung metastases and metastatic nodules are shown. The data are presented as the mean ± s.d.; two-tailed unpaired Student's *t*-test. (**G**) Representative LTR56 genome locations (left) and RNA immunoprecipitation (RIP) assays (right) with subsequent qRT-PCR assays in HEK293T cells ectopically expressing full-length Flag-tagged ZMYND8. The relative quantification of TROJAN and LTR56 expression in RNA-protein complexes immunoprecipitated with IgG or Flag antibodies is shown. The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. *$p < 0.05$ and ***$p < 0.001$. NS: not significant.

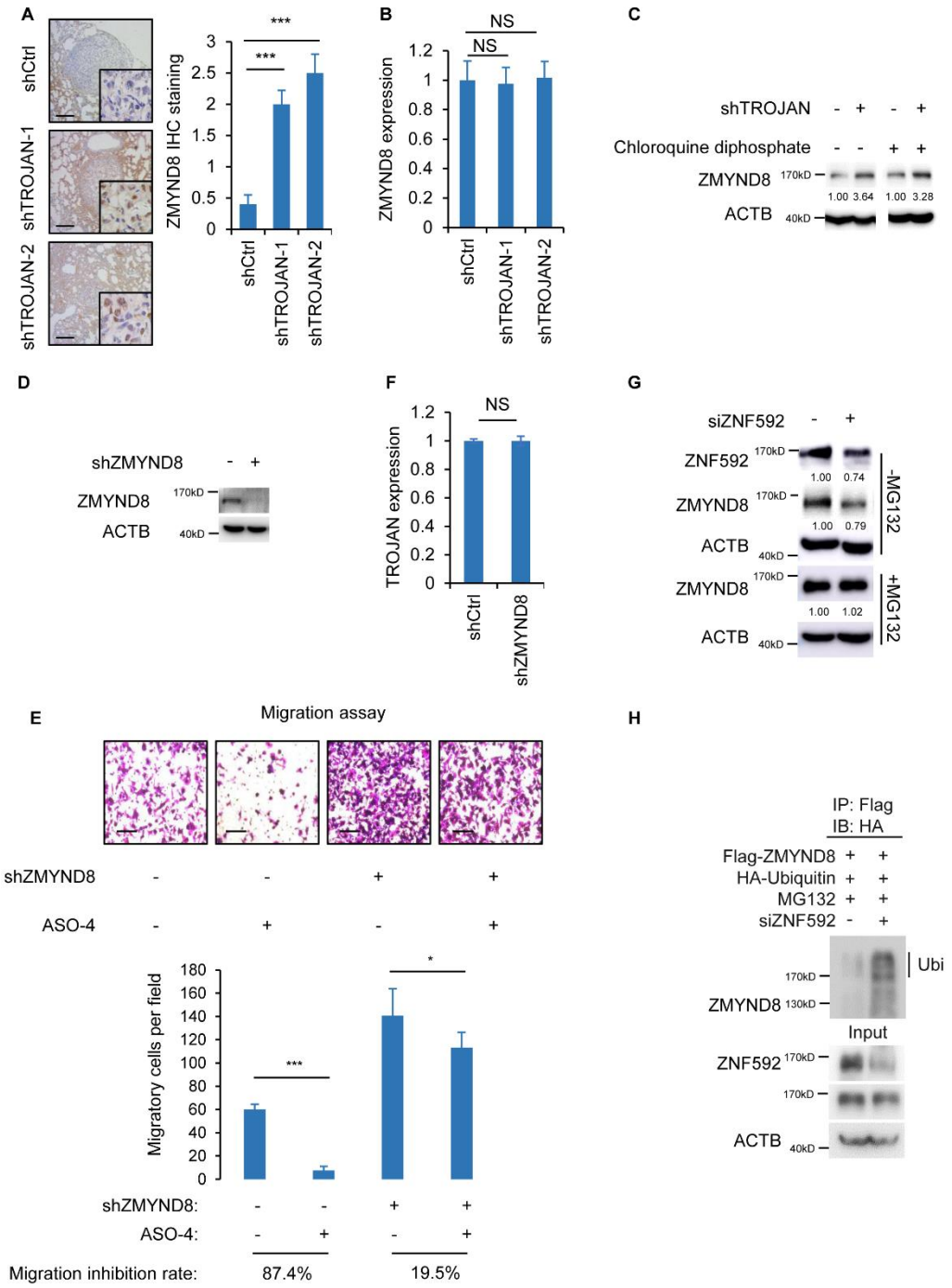


Figure S7

Fig. S7. TROJAN degrades the ZMYND8 protein. (**A**) Immunohistochemical staining of ZMYND8 in lungs isolated from mice administered MDA-MB-231 LM2 cells expressing short hairpin RNAs (shRNAs) targeting TROJAN or control (Ctrl) by tail vein injection. Scale bar: 50 μm. The quantitative data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. (**B**)

Quantitative reverse transcription (qRT)-PCR analysis of relative ZMYND8 transcription after TROJAN knockdown using shRNAs. The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. NS: not significant. (**C**) Western blot images showing ZMYND8 expression in MDA-MB-231 LM2 cells expressing control or TROJAN shRNA treated with DMSO (control) or chloroquine diphosphate. (**D**) Western blot images showing ZMYND8 expression in MDA-MB-231 LM2 cells expressing Ctrl or ZMYND8 shRNA. (**E**) Transwell migration assay of MDA-MB-231 LM2 cells treated with ASO-4 and/or ZMYND8 shRNA. The data are presented as the mean ± s.d. ***$p < 0.001$, n = 3 independent experiments; two-tailed unpaired Student's *t*-test. Scale bar: 200 μm. (**F**) qRT-PCR analysis of relative TROJAN transcription after ZMYND8 knockdown using shRNAs. The data are presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. NS: not significant. (**G**) Western blot images showing ZMYND8 expression in MDA-MB-231 LM2 cells expressing control or ZNF592 small interfering RNA (siRNA) treated with DMSO (control) or MG132. (**H**) Western blot images showing ZMYND8-associated ubiquitination in control and ZNF592-downregulated HEK293T cells ectopically expressing full-length Flag-tagged ZMYND8 and treated with MG132. *$p < 0.05$ and ***$p < 0.001$. NS: not significant.

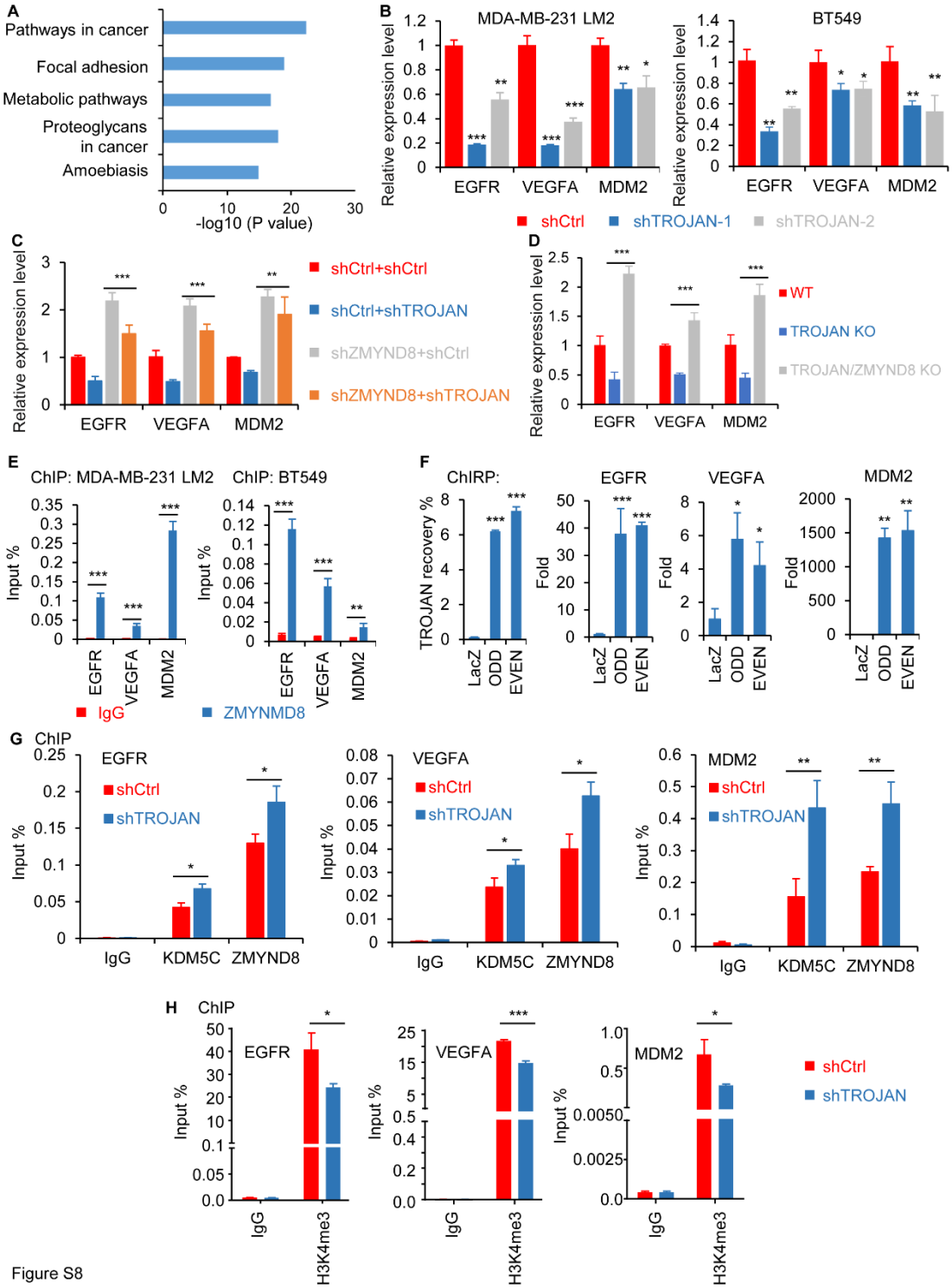


Figure S8

Fig. S8. TROJAN and ZMYND8 regulate the transcription of diverse target genes. (**A**) Pathway analysis of potential TROJAN target genes identified by microarray. (**B**) Quantitative reverse transcription (qRT)-PCR analysis of the relative transcription levels of the potential target genes in MDA-MB-231 LM2 and BT549 cells expressing control or TROJAN short hairpins RNAs (shRNAs). The data are

presented as the mean ± s.d.; n = 3 independent experiments; two-tailed unpaired Student's *t*-test. (**C**) qRT-PCR detection of the target genes in MDA-MB-231 LM2 cells expressing TROJAN and/or ZMYND8 shRNAs. The data are presented as the mean ± s.d.; n = 3 independent experiments. (**D**) qRT-PCR detection of the target genes in MDA-MB-231 LM2 cells with TROJAN knockout (KO) and/or ZMYND8 KO. The data are presented as the mean ± s.d.; n = 3 independent experiments. (**E**) qRT-PCR analysis of ZMYND8 occupying the EGFR, VEGFA and MDM2 promoters after chromatin immunoprecipitation (ChIP) assay. (**F**) qRT-PCR analysis of TROJAN RNA recovery and enrichment of the promoter regions of potential target genes after chromatin isolation by RNA purification (ChIRP) assay. LacZ was set as the negative control. The data are presented as the mean ± s.d.; two-tailed unpaired Student's *t*-test. (**G**) qRT-PCR analysis of KDM5C and ZMYND8 at the EGFR, VEGFA and MDM2 promoters after ChIP assay in MDA-MB-231 LM2 cells expressing control or TROJAN shRNA. (**H**) qRT-PCR analysis of H3K4me3 at the EGFR, VEGFA and MDM2 promoters after ChIP assay in MDA-MB-231 LM2 cells expressing control or TROJAN shRNA. The data are presented as the mean ± s.d.; two-tailed unpaired Student's *t*-test. *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$. NS: not significant.

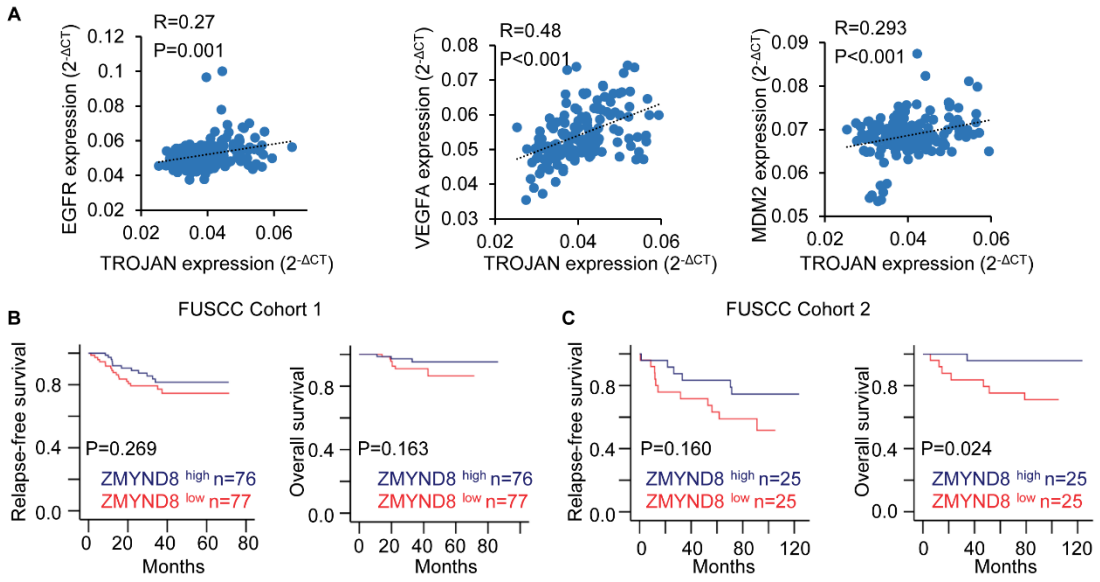


Figure S9

Fig. S9. TROJAN regulates breast cancer progression via ZMYND8. (**A**) Correlations of TROJAN with EGFR, VEGFA and MDM2 expression as determined by Pearson correlation analyses. (**B, C**) Kaplan-Meier analyses of the association of ZMYND8 with relapse-free survival and overall survival in FUSCC cohorts 1 (**B**) and 2 (**C**).

SUPPLEMENTARY TABLE LEGENDS AND TABLES

Table S1. TNBC-related HERVs, basic patient information, primers, and MS data for proteins pulled down by TROJAN and SYSL1.