Online Supporting Information for

"Roles of Small GTPases in Acquired Tamoxifen Resistance in MCF-7 Cells Revealed by Targeted, Quantitative Proteomic Analysis"

Ming Huang[a] and Yinsheng Wang[a,b,*]

[a]Environmental Toxicology Graduate Program and [b]Department of Chemistry, University of California, Riverside, Riverside, CA 92521-0403, USA.

[*]To whom correspondence should be addressed. Email: Yinsheng.Wang@ucr.edu

Table of Contents:

The Skyline files for the MRM spectral library containing the MS/MS spectra and iRT information for all targeted small GTPase peptides and 10 standard BSA peptides are available at PeptideAtlas with the identifier number of PASS01267.

I. Supplementary Experimental Section

MRM spectral library. On the basis of data acquired from shotgun proteomic analyses of the tryptic digestion mixtures of low-molecular weight proteins (15–37 kDa) from the protein lysates of 9 cell lines derived from different human tissue origins, a Skyline[1] MRM spectral library for small GTPases was constructed recently.[2] Tandem mass spectra and chromatographic retention time of unique peptides derived from the targeted small GTPases were incorporated into the library. The transitions corresponding to the formation of the three most abundant y-ions based on the MS/MS acquired from shotgun proteomic analyses were selected [3]. The MRM library encompassed 432 distinct peptides representing 113 non-redundant small GTPases encoded by unique genes.

Data source for bioinformatic analyses. Patient RNAseq data were obtained from The Cancer Genome Atlas (TCGA) via cBioPortal (http://www.cbioportal.org/).[4] The complete clinical data files were downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). The Cancer Cell Line Encyclopedia (CCLE) (http://www.broadinstitute.org/ccle/home) were interrogated for the comprehensive evaluation of mRNA expression for candidate genes among more than 100 breast cancer cell lines.[5] Publicly available transcriptomic profiles with accession numbers GSE3494, GSE4922, GSE6434, GSE24460, GSE26495 and GSE42568 were downloaded from the GEO database and analyzed using R (version 3.4.3).

Patient survival analysis. Kaplan–Meier survival curves were generated using an online database Kaplan-Meier plotter (kmplotter.com) for breast cancer.[6] Data were analyzed using the JetSet best probe set to analyze gene expression and relapse-free survival (RFS). Briefly, gene names were entered into the database to obtain Kaplan–Meier survival plot where the hazard ratio (HR), 95% confidence intervals (95% CI) and logrank *p* values were calculated and displayed.

Immunoblotting. Total protein lysate was prepared from cell pellet using ice-cold CelLytic M cell lysis reagent (Sigma-Aldrich, MO) containing protease inhibitor cocktail (1:100). After cell lysis, the protein concentration was determined using Quick Start™ Bradford Protein Assay (Bio-Rad). Approximately 10–20 µg whole-cell protein lysates, mixed with 4×Laemmli SDS loading buffer, were loaded onto 10% polyacrylamide gels and, after electrophoresis, the proteins were transferred onto nitrocellulose membranes. After blocking with 5% non-fat milk in PBS with 0.1% Tween-20 (PBST) for 1 h at 25 °C, the membranes were incubated with primary antibodies against human RAB27A (Abcam; rabbit polyclonal, 1:5,000), RAB27B (Proteintech; rabbit polyclonal, 1:5,000), RAB31 (4D12. Santa Cruz; rabbit polyclonal, 1:2,000), or β-actin (Thermo Fisher; rabbit polyclonal, 1:10,000). After overnight incubation with primary antibodies at 4°C with 5% bovine serum albumin (BSA) in PBST, the membranes were then incubated with peroxidase-labeled donkey anti-rabbit secondary antibody (Thermo Fisher; 1:10,000) or mouse m-IgGκ BP-HRP (Santa Cruz; 1:10,000) for 1 h at 25 °C. Amersham ECL Prime Western Blot Detecting Reagent (GE Healthcare) was used for visualization of protein bands.

Plasmid construction. pLKO.1-shRAB31 and pLKO.1-scramble plasmids were constructed by inserting a short hairpin double-stranded oligonucleotide targeting RAB31 or a scrambled sequence into the pLKO.1 lentiviral vector. The sequences of shRAB31 oligonucleotides are listed in **Table S2**. Human RAB31 cDNA was amplified from an in-house cDNA library and subcloned into the p3×Flag-CMV10 vector.

Generation of stable knockdown cell lines. The lentiviral vectors pLKO.1-shRAB31 and pLKO.1-scramble were generated as described above. Recombinant lentiviruses were produced by co-transfection of HEK293T cells with the pLKO.1-scramble or pLKO.1-shRAB31 shRNA plasmids, envelope plasmid pLTR-G (Addgene #17532) and packaging plasmid pCMV-dR8.2 dvpr (Addgene #8455). Lentivirus-containing supernatant was harvested and filtered through 0.45-μm pore size filters at 48 h post-transfection. Infection of MCF-7 or T47D cells with recombinant lentivirus was conducted in the presence of 5 μg/ml polybrene. After removal of virus, the cells were selected in 1 μg/mL puromycin-containing medium for 3 days to eliminate uninfected cells. After selection, the cells were maintained in medium containing 1 μg/mL puromycin, and were used for subsequent experiments.

References:

(1) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. *Bioinformatics* **2010**, *26*, 966-968.
(2) Huang, M.; Qi, T. F.; Li, L.; Zhang, G.; Wang, Y. *Cancer Res* **2018**, *78*, 5431-5445.
(3) Liebler, D. C.; Zimmerman, L. J. *Biochemistry* **2013**, *52*, 3797-3806.
(4) Cerami, E.; Gao, J. J.; Dogrusoz, U.; Gross, B. E.; Sumer, S. O.; Aksoy, B. A.; Jacobsen, A.; Byrne, C. J.; Heuer, M. L.; Larsson, E.; Antipin, Y.; Reva, B.; Goldberg, A. P.; Sander, C.; Schultz, N. *Cancer Discovery* **2012**, *2*, 401-404.
(5) Barretina, J.; Caponigro, G.; Stransky, N.; Venkatesan, K.; Margolin, A. A.; Kim, S.; Wilson, C. J.; Lehar, J.; Kryukov, G. V.; Sonkin, D.; Reddy, A.; Liu, M.; Murray, L.; Berger, M. F.; Monahan, J. E.; Morais, P.; Meltzer, J.; Korejwa, A.; Jane-Valbuena, J.; Mapa, F. A., et al. *Nature* **2012**, *483*, 603-607.
(6) Gyorffy, B.; Lanczky, A.; Eklund, A. C.; Denkert, C.; Budczies, J.; Li, Q.; Szallasi, Z. *Breast Cancer Res Treat* **2010**, *123*, 725-731.

Table S1A. A complete list of all small GTPases quantified in the scheduled LC-MRM analyses from two sets of forward and one set of reverse SILAC experiments, classified by different sub-families.

Table S1B. A complete list of all small GTPase peptides quantified in the scheduled LC-MRM analysis from two sets of forward- and one set of reverse-SILAC experiments.

Table S2. A table showing oligo sequences used for the construction of shRNA plasmids using the pLKO.1 lentiviral vector.

shRAB31-1	CCGGTGCTAAGGAATACGCTGAATC**CTCGAG**GATTCAGCGTATTCCTTAGCATTTTTG
shRAB31-2	CCGGCAAAGTTGAGAAGCCAACCAT**CTCGAG**ATGGTTGGCTTCTCAACTTTGTTTTTG
shScramble	CCGGTCCTAAGGTTAAGTCGCCCTCG**CTCGAG**CGAGGGCGACTTAACCTTAGGTTTTTG

Note: CTCGAG denotes the hairpin region.

III. Supplementary Figures

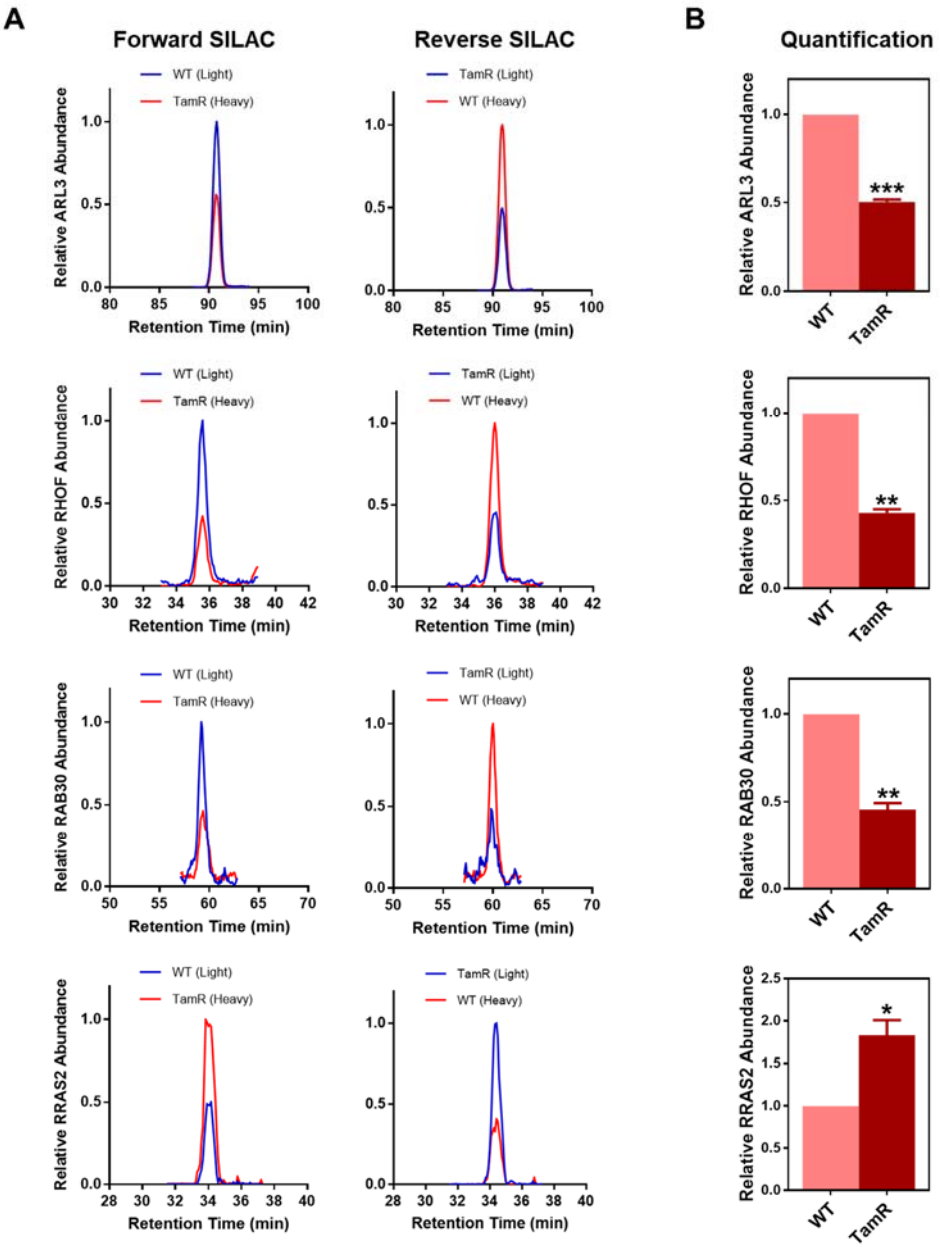

Figure S1. LC-MRM for the quantification of the relative levels of expressions of ARL3, RHOF, RAB30, and RRAS2 proteins in the paired MCF-7/WT and MCF-7/TamR cells. (A) Representative extracted-ion chromatograms (XICs) for the quantification of the ARL3, RHOF, RAB30, and RRAS2 proteins in one forward and one reverse SILAC labeling experiments; **(B)** Quantification results of the LC-MRM analyses (n = 3). The *p* values were calculated by using a paired two-tailed Student's *t* test (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). Error bars represent standard deviations.

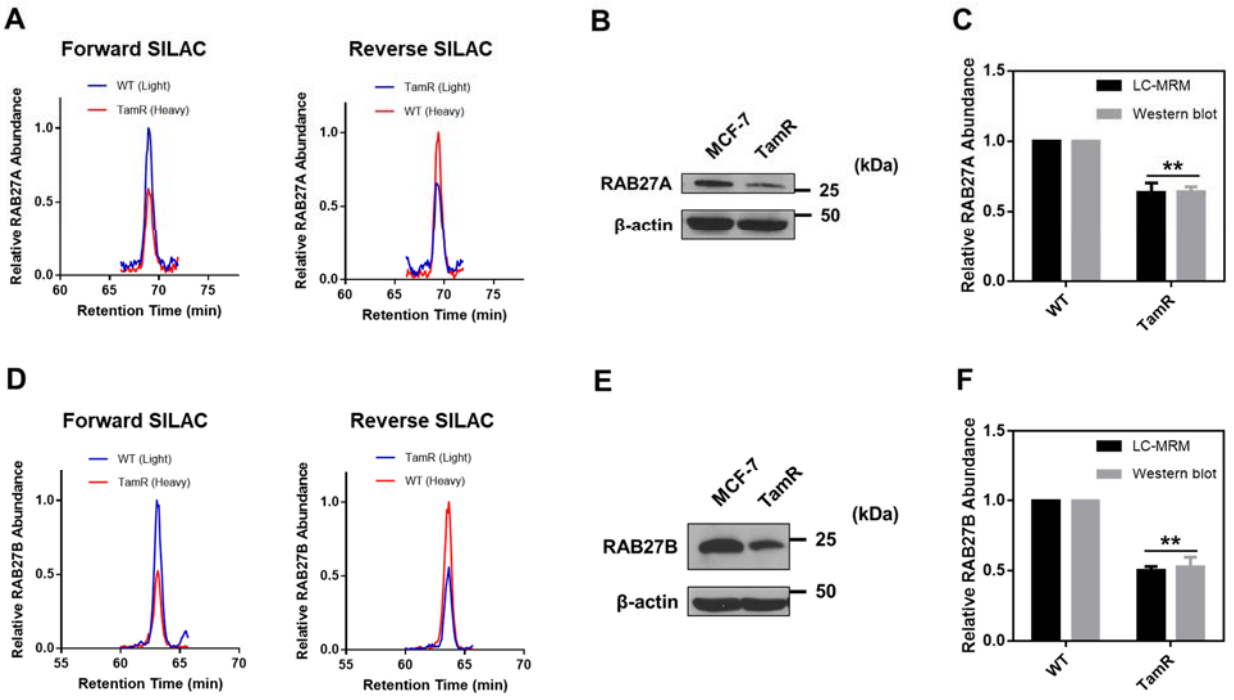

Figure S2. LC-MRM and Western-blot analyses for the quantification of the relative levels of expressions of RAB27A and RAB27B proteins in the paired MCF-7/WT and MCF-7/TamR cells. (A) Representative XICs for the quantification of the tryptic peptide FLALGDSGVGK derived from RAB27A in one forward and one reverse SILAC labeling experiments; **(B)** Validation of the differential expression of RAB27A in MCF-7/WT and MCF-7/TamR cells by Western-blot analysis; **(C)** Comparison of quantification results obtained from LC-MRM and Western-blot analyses (n = 3) in panels (A) and (B); **(D)** Representative XICs for the quantification of the tryptic peptide LLALGDSGVGK derived from RAB27B in one forward and one reverse SILAC labeling experiments; **(E)** Validation of the differential expression of RAB27B in MCF-7/WT and MCF-7/TamR cells by Western-blot analysis; **(F)** Comparison of quantification results obtained from LC-MRM and Western-blot analyses (n = 3) in panels (D) and (E). The *p* values were calculated by using a paired two-tailed Student's *t* test (** $p < 0.01$). Error bars represent standard deviations.

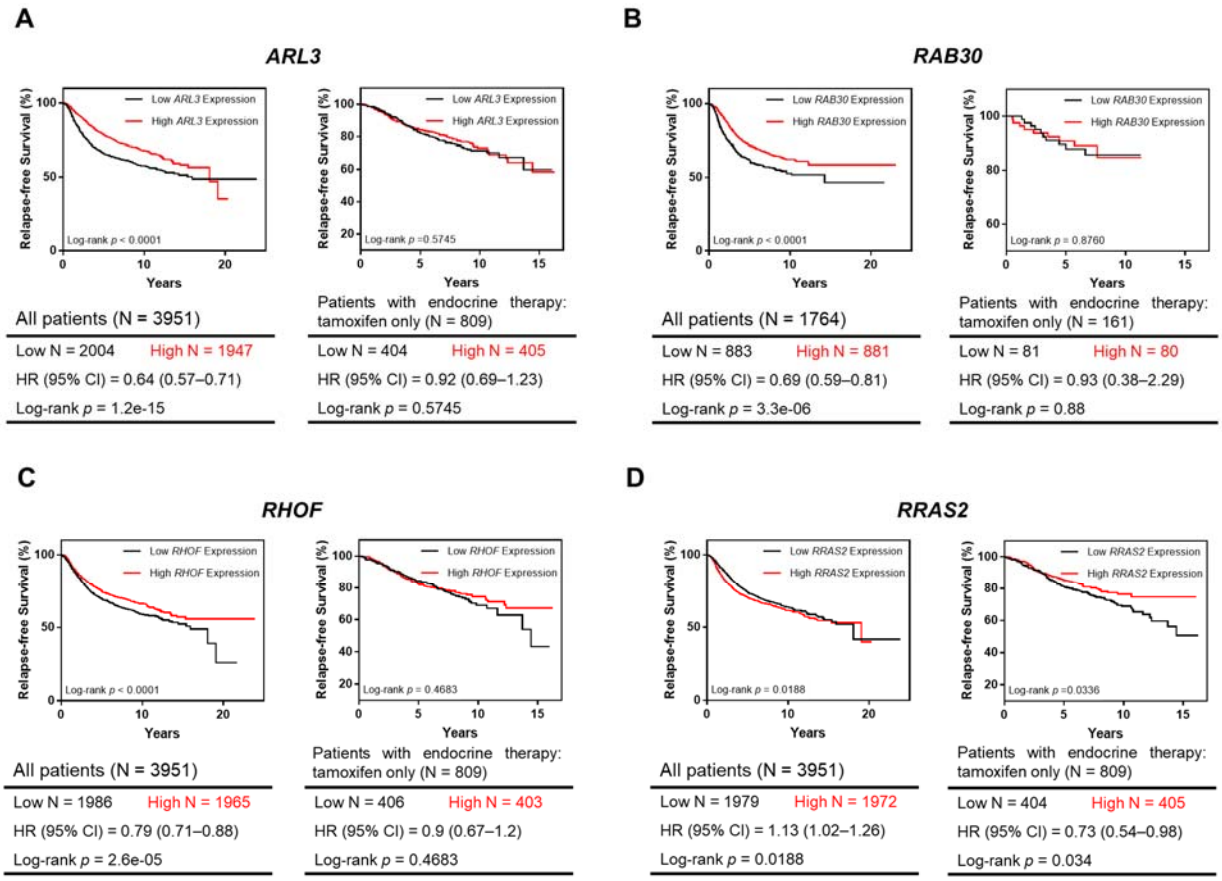

Figure S3. Representative Kaplan–Meier survival curves for the implications of *ARL3*, *RHOF*, *RAB30*, and *RRAS2* mRNA expressions in survival of breast cancer patients. Kaplan–Meier survival analyses for relapse-free survival (RFS) in all breast cancer patients (left panel) and in ER-positive breast cancer patients receiving only treatment with tamoxifen but not chemotherapy (right panel). The patient population was stratified by median mRNA expression levels of the **(A)** *ARL3*, **(B)** *RAB30*, **(C)** *RHOF*, and **(D)** *RRAS2* genes, respectively. The *p* values were calculated by using a log-rank test. Number of patients, hazard ratios (HR) and 95% confidence intervals (95% CI) are indicated.

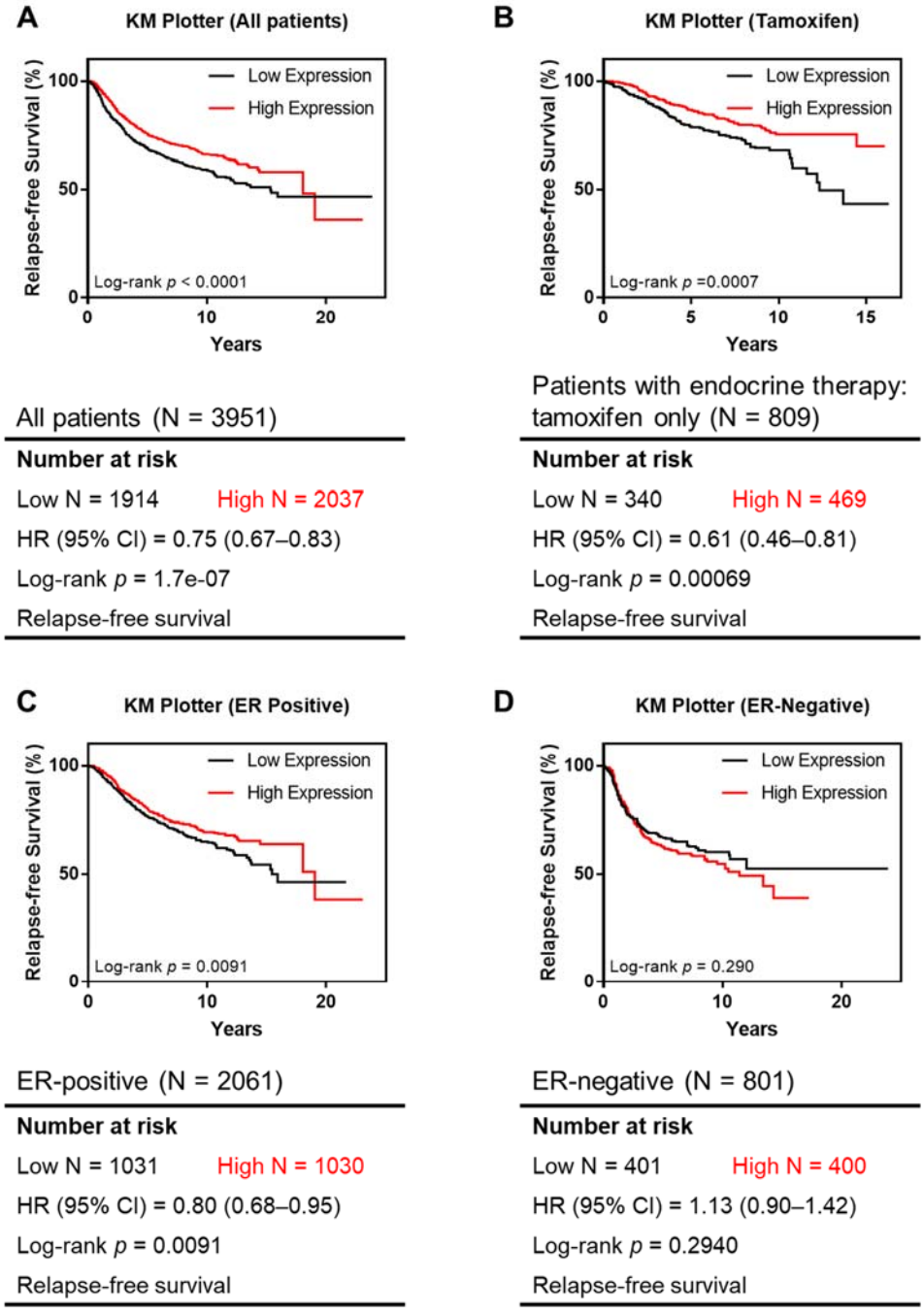

Figure S4. *RAB31* expression predicts breast cancer patient outcome. Kaplan–Meier survival analyses of the implications of mRNA expression of the *RAB31* gene for relapse-free survival (RFS) in breast cancer patients: **(A)** all breast cancer patients (n = 3951); **(B)** ER-positive breast cancer patients receiving tamoxifen but without chemotherapy (n = 809); **(C)** ER-positive breast cancer patients (n = 2061); **(D)** ER-negative breast cancer patients (n = 801). The patient population was stratified by median *RAB31* mRNA expression levels. The *p* value was calculated by using a log-rank test. Analysis was performed using KM-plotter (kmplot.com/analysis).

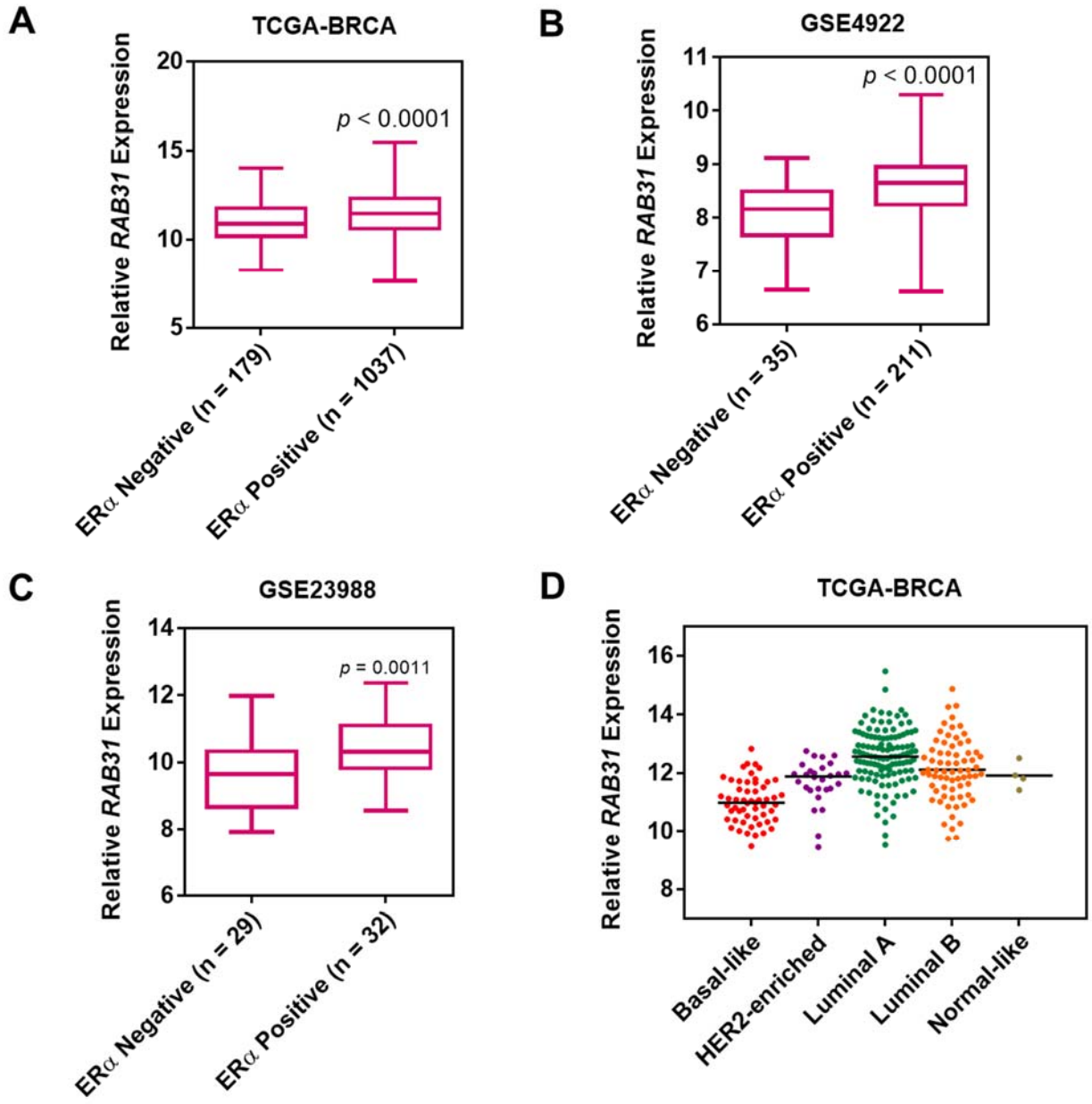

Figure S5. *RAB31* expression is correlated with ER status and breast cancer subtypes. Box-whisker plot showing correlated expressions of *RAB31* and *ESR1*, the latter of which encodes estrogen receptor alpha (ERα), in the: **(A)** TCGA-BRCA cohort; **(B)** GSE4922 dataset; **(C)** GSE23988 dataset. Whisker shows 10th and 90th percentile; box boundaries show 25th and 75th percentile and the lines represent median values. **(D)** A scatter plot showing *RAB31* expression across the five molecular subtypes of breast cancer (n = 831) in the TCGA-BRCA cohort, and the lines represent median value. The *p* values for the box-whisker plots were calculated by using an unpaired two-tailed Student's *t* test.

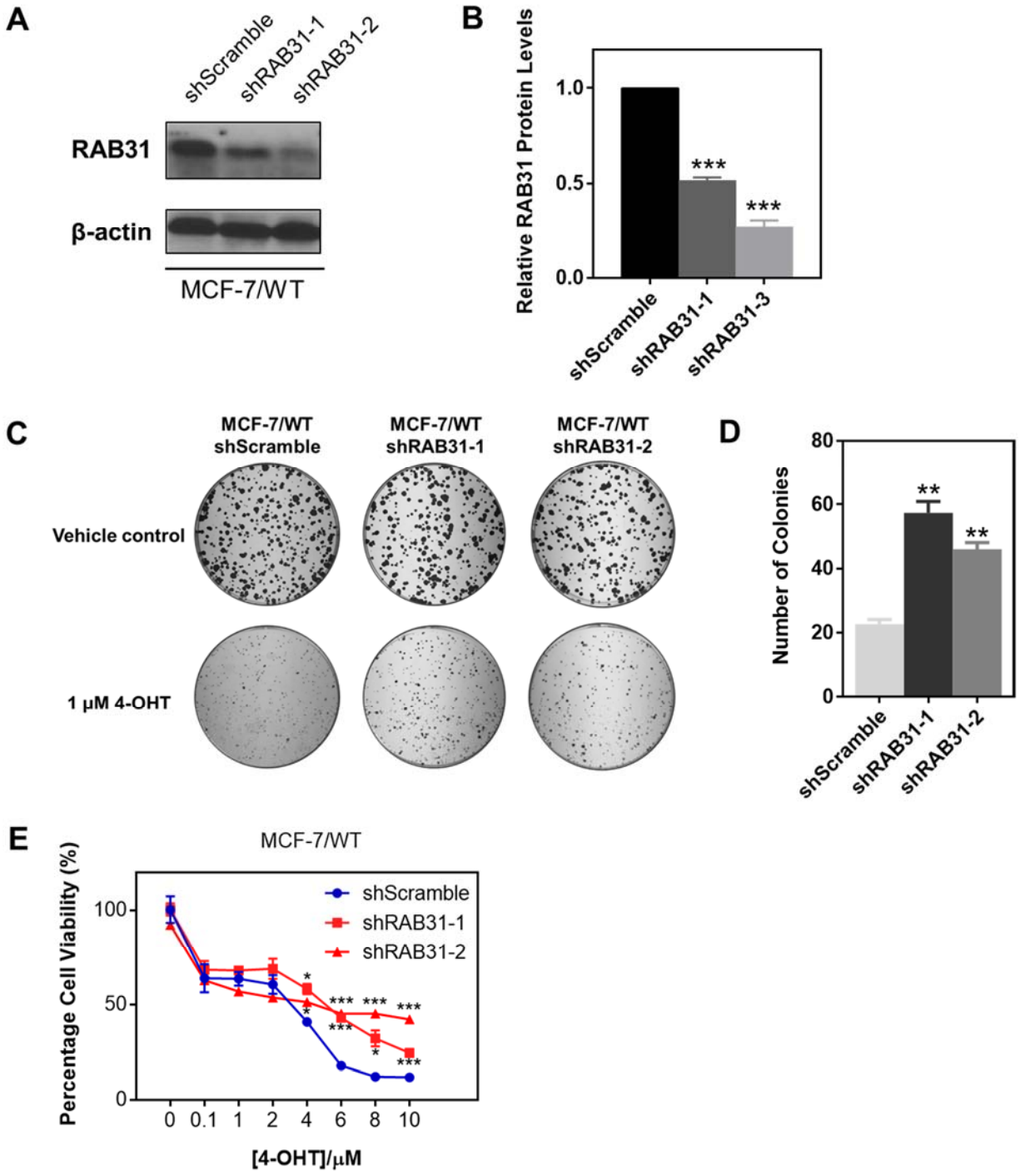

Figure S6. RAB31 knockdown confers increased tamoxifen resistance. (A) Validation of stable knockdown of RAB31 in MCF-7 cells by Western-blot analysis; **(B)** Quantification results for Western-blot validation of stable shRAB31 knockdown MCF-7 cell lines; **(C)** Colony formation assay for stable shRAB31/shScramble MCF-7 cells; **(D)** Quantification results for colony formation assay; **(E)** Cell proliferation assay for stable shRAB31/shScramble MCF-7 cells upon treatment with different doses of 4-OHT. The *p* values were calculated by using a paired two-tailed Student's *t* test (** $p < 0.01$). Error bars represent standard deviations.

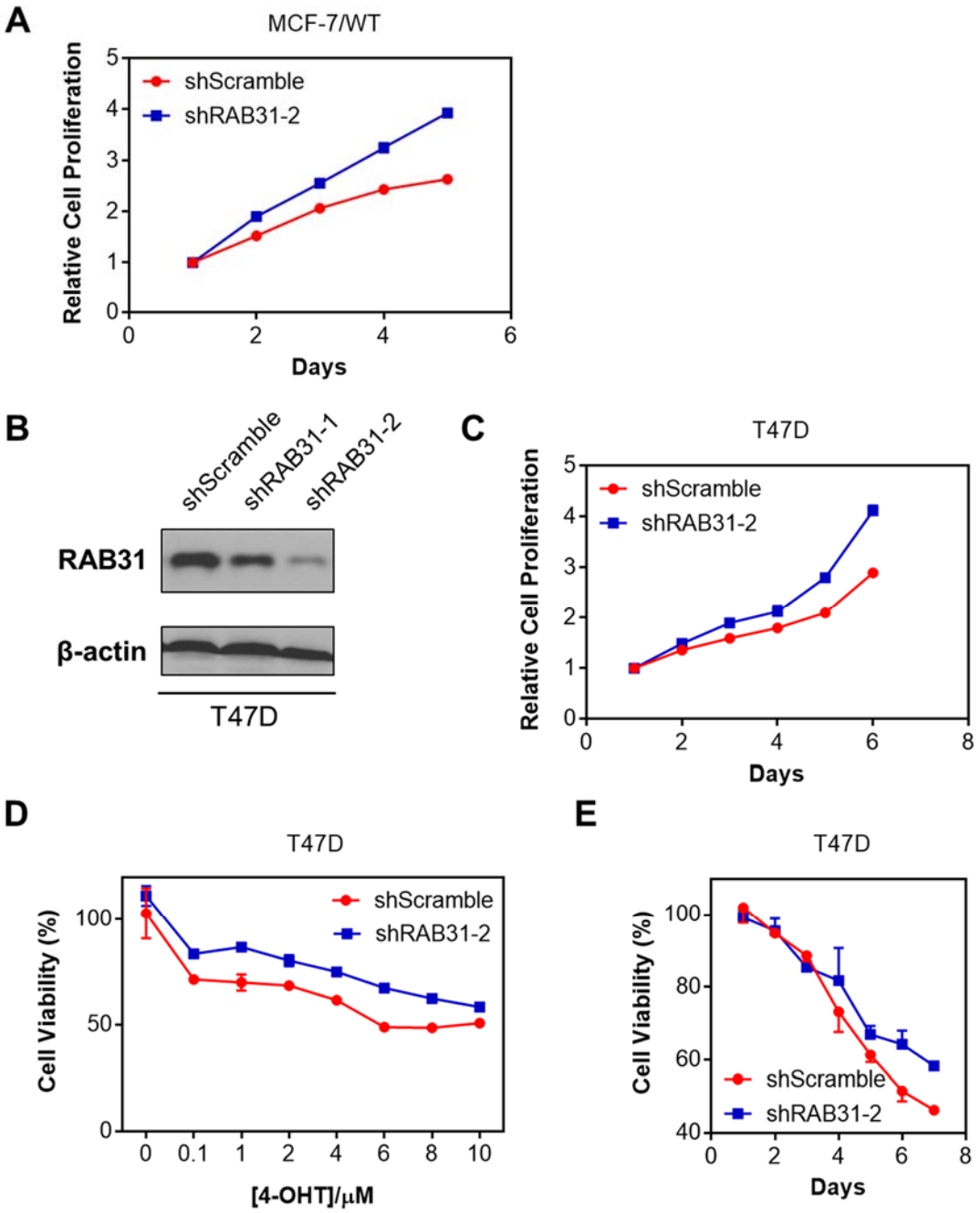

Figure S7. RAB31 knockdown renders elevated tamoxifen resistance and proliferation rates. **(A)** Cell proliferation assay for stable shRAB31-2 and shScramble MCF-7 cells; **(B)** Validation of stable knockdown of RAB31 in T47D cell lines by Western-blot analysis; **(C)** Cell proliferation assay for stable shRAB31-2 and shScramble T47D cells; **(D)** Cell proliferation assay for stable shRAB31-2 and shScramble T47D cells under dose-dependent 4-OHT treatment; **(E)** Cell proliferation assay for stable shRAB31-2 and shScramble T47D cells under time-dependent 1 μM 4-OHT treatment. The data represent the means±standard deviations of results from three parallel experiments. Some error bars appear to be smaller than the symbols.