

**Online Supporting Information for**

**“A Targeted Quantitative Proteomic Approach for Probing Altered Protein  
Expression of Small GTPases Associated with Colorectal Cancer Metastasis”**

Ming Huang<sup>a</sup> and Yinsheng Wang<sup>a,b,\*</sup>

<sup>a</sup>Environmental Toxicology Graduate Program and <sup>b</sup>Department of Chemistry, University of  
California, Riverside, Riverside, CA 92521-0403, USA.

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

## Table of Contents:

### I. Supplementary Experimental Section

### II. Supplementary Tables

**Table S1.** A complete list of small GTPase proteins (Table S1A) and peptides (Table S1B) quantified in the scheduled LC-MRM analysis from two sets of forward and one set of reverse SILAC experiments (in Excel).

**Table S2.** A list of oligodeoxyribonucleotides sequences used for the construction of shRNA plasmids using the pLKO.1 lentiviral vector, and primers used for RT-qPCR.

### III. Supplementary Figures

**Figure S1.** Scheduled LC-MRM for assessing the differential expression of small GTPases in paired SW480/SW620 cells.

**Figure S2.** In-depth coverage of the small GTPase proteome facilitated by the Ge-LC-MRM-based quantification.

**Figure S3.** LC-MRM quantification of RHOB, RHOF, RHOG, and RAB6 in SW480/SW620 cells.

**Figure S4.** Validation of differential expression of RAB27A and ARF4 in SW480/SW620 cells.

**Figure S5.** Prognostic values of ARF4, RHOF and RAB6B in CRC patient cohorts.

**Figure S6.** Higher expression of *SAR1A* in CRC confers unfavorable patient prognosis.

**Figure S7.** Higher expression of *RAB31* is associated with worse CRC patient outcome and higher disease stages.

**Figure S8.** RAB31 knockdown modulates the migratory and invasive capacities and cell proliferation of SW480 cells.

### IV. Supplementary Skyline Files

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 PASS01326 (<http://www.peptideatlas.org/PASS/PASS01326>).

## I. Supplementary Experimental Section

**MRM Spectral Library.** On the basis of data acquired from shotgun proteomic analyses of the tryptic peptide mixtures of low-molecular weight proteins (15–37 kDa) resolved by SDS-PAGE fractionation of the protein lysates of 9 cell lines derived from different human tissue origins, a Skyline MRM spectral library for small GTPases was constructed previously.<sup>1</sup> The MRM library encompassed 432 distinct peptides representing 113 non-redundant small GTPases encoded by unique genes. Cysteine carbamidomethylation (+57 Da) was set as a fixed modification, whereas lysine (+8 Da) and arginine (+6 Da) mass shifts introduced by heavy isotope labeling were defined when *in silico*-derived transitions were generated by Skyline.

**LC-MS/MS Analysis.** All LC-MS/MS data acquired in the DDA mode were carried out on an LTQ-Orbitrap Velos mass spectrometer coupled with an EASY-nLC II HPLC system and a nanoelectrospray ionization source (Thermo Fisher Scientific, San Jose, CA). The peptides were automatically loaded onto a 4-cm trapping column (150  $\mu\text{m}$  i.d.) packed with ReproSil-Pur 120 C18-AQ resin (5  $\mu\text{m}$  in particle size and 120  $\text{\AA}$  in pore size, Dr. Maisch GmbH HPLC) at 3  $\mu\text{L}/\text{min}$ . The trapping column was coupled to a 20-cm fused silica analytical column (75  $\mu\text{m}$  i.d.) packed with ReproSil-Pur 120 C18-AQ resin (3  $\mu\text{m}$  in particle size and 120  $\text{\AA}$  in pore size, Dr. Maisch GmbH HPLC). The peptide mixtures were then separated using a 170-min linear gradient of 2–35% acetonitrile in 0.1% aqueous solution of formic acid and at a flow rate of 230 nL/min. Full-scan MS ( $m/z$  300-1500) were acquired after ion accumulation to a target value of  $1 \times 10^6$  in the Orbitrap at a resolution of 17,500 (at  $m/z$  400). The 20 most abundant precursor ions found in MS were selected for fragmentation by collisionally induced dissociation (CID) to obtain the MS/MS, where the automatic gain control (AGC) target value was set to  $1 \times 10^4$  and the normalized collision energy to 35%. Dynamic exclusion was applied with a time window of 30 s. MS raw data were searched against the IPI protein database (version 3.68, 87,061 protein entries) using the Andromeda search engine integrated into the MaxQuant software (version 1.3.0.5). Cysteine carbamidomethylation (+57 Da) was set as a fixed modification, whereas methionine oxidation (+16 Da) was set as a variable modification. SILAC doublets were defined by selecting isotope labels on lysine (+8 Da) or arginine (+6 Da).

**Data source for bioinformatic analyses.** Patient RNA-Seq data were obtained from The Cancer Genome Atlas (TCGA) via cBioPortal (<http://www.cbioportal.org/>). The data acquired from four different human gene expression microarrays with accession numbers of GSE14333, GSE17536, GSE21510, GSE39582, and the corresponding clinical information were downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>). The data were analyzed using R/Bioconductor (version 3.4.3).

**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  $\mu\text{g}$  whole-cell protein lysates, mixed with 4 $\times$  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) at 25 °C for 1 h, the membranes were incubated with primary antibodies against human ARF4 (Proteintech; rabbit polyclonal, 1:5,000), RAB6A (38-TB, Santa Cruz; mouse polyclonal, 1:1,000), pan-RAB6 (3G3, Santa Cruz; mouse polyclonal, 1:1,000), RAB27A (Abcam; rabbit polyclonal, 1:5,000), RAB31 (4D12, Santa Cruz; rabbit polyclonal, 1:2,000), SAR1A (K-44, Santa Cruz; mouse polyclonal, 1:1,000), SAR1B (AT1C7, Santa Cruz; mouse polyclonal,

1:1,000), Slug (PA5-11922, Thermo Fisher; rabbit polyclonal, 1:1,000), or  $\beta$ -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-IgG $\kappa$  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.

**Generation of stable knockdown cell lines and siRNA transfection.** The lentiviral vectors pLKO.1-shSAR1B and pLKO.1-scramble plasmids were constructed by inserting a short hairpin double-stranded oligonucleotide targeting human SAR1B or a scrambled sequence into the AgeI/EcoRI sites of the shRNA vector pLKO.1 (Addgene #10878; Cambridge, MA). Expression of the shRNA was driven by the human U6 promoter. The sequences of shSAR1B and shScramble oligonucleotides are listed in Table S2. Recombinant lentiviruses were produced by co-transfection of HEK293T cells with the shRNA plasmids pLKO.1-shScramble or pLKO.1-shSAR1B, 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- $\mu$ m pore size filters at 48 h post-transfection. Infection of SW480 cells with recombinant lentivirus was conducted in the presence of 5  $\mu$ g/ml polybrene. After removal of virus, the cells were selected in 5  $\mu$ g/mL puromycin-containing medium for 3 days to eliminate uninfected cells. After selection, the cells were maintained in a medium containing 2.5  $\mu$ g/mL puromycin and were used for subsequent experiments. For siRNA transfection, siGENOME non-targeting (NT) siRNA control (D-001210-02-05) and an siRNA sequence (5'-UGAUGUUGUGGUCAAAGUGAU-3') targeting RAB31 were purchased from Dharmacon (Lafayette, CO). siRNAs were transfected into SW480 cells using RNAiMAX (Invitrogen) following the manufacturer's protocol. Cells were either harvested or used for subsequent experiments at 72 h after transfection.

**Cell proliferation assay.** Cell proliferation was evaluated using a cell counting kit-8 (CCK-8; Dojindo Laboratories, Kumamoto, Japan). Briefly, SW480 cells were seeded in a 96-well flat-bottomed microplate (3000 cells/well) in complete growth medium (100  $\mu$ L/well) for 24 h. At the desired time points, 10  $\mu$ L of the CCK-8 dye was added to each well and, after incubation at 37°C for 4 h, the absorbance at 450 nm for the cells was recorded using the Synergy™ H1 Hybrid Multi-Mode Microplate Reader (BioTek Instruments).

**Quantitative reverse transcription PCR (RT-qPCR).** Total RNA was extracted from cells using The E.Z.N.A.® Total RNA Kit I (Omega Bio-Tech), and cDNA was synthesized via oligo(dT)<sub>18</sub> primed reverse transcription by employing M-MLV reverse transcriptase (Promega). After a 60-min reaction at 42°C, the reverse transcriptase was deactivated by heating at 75°C for 5 min. RT-qPCR experiments was performed using iQ SYBR Green Supermix kit (Bio-Rad) on a Bio-Rad iCycler system (Bio-Rad), and the PCR conditions were as follows: 95°C for 3 min; 45 cycles at 95°C for 15 s, 55°C for 30 s, and 72°C for 45 s. The comparative cycle threshold (Ct) method ( $\Delta\Delta$ Ct) was used for the relative quantification of gene expression.<sup>2</sup> The sequences of primers are listed in Table S2. Relative gene expression was normalized to that of the internal control (*GAPDH*).

## References:

1. Huang, M.; Qi, T. Y. F.; Li, L.; Zhang, G.; Wang, Y. S., A targeted quantitative proteomic approach assesses the reprogramming of small GTPases during melanoma metastasis. *Cancer Res.* **2018**, *78*, 5431-5445.
2. Livak, K. J.; Schmittgen, T. D., Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* **2001**, *25*, 402-408.
3. Ghosh, D.; Yu, H.; Tan, X. F.; Lim, T. K.; Zubaidah, R. M.; Tan, H. T.; Chung, M. C.; Lin, Q., Identification of key players for colorectal cancer metastasis by iTRAQ quantitative proteomics profiling of isogenic SW480 and SW620 cell lines. *J. Proteome Res.* **2011**, *10*, 4373-87.
4. Lee, J. G.; McKinney, K. Q.; Pavlopoulos, A. J.; Park, J. H.; Hwang, S., Identification of anti-metastatic drug and natural compound targets in isogenic colorectal cancer cells. *J. Proteomics* **2015**, *113*, 326-36.
5. Cai, R.; Huang, M.; Wang, Y. S., Targeted quantitative profiling of GTP-binding proteins in cancer cells using isotope-coded GTP probes. *Anal. Chem.* **2018**, *90*, 14339-14346.

## II. Supplementary Tables

**Table S1A.** A complete list of all small GTPases quantified in the scheduled LC-MRM analysis from two sets of forward and one set of reverse SILAC experiments, organized according to 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 S2A.** A complete list of the oligonucleotide sequences used for the construction of shRNA plasmids with the pLKO.1 lentiviral vector.

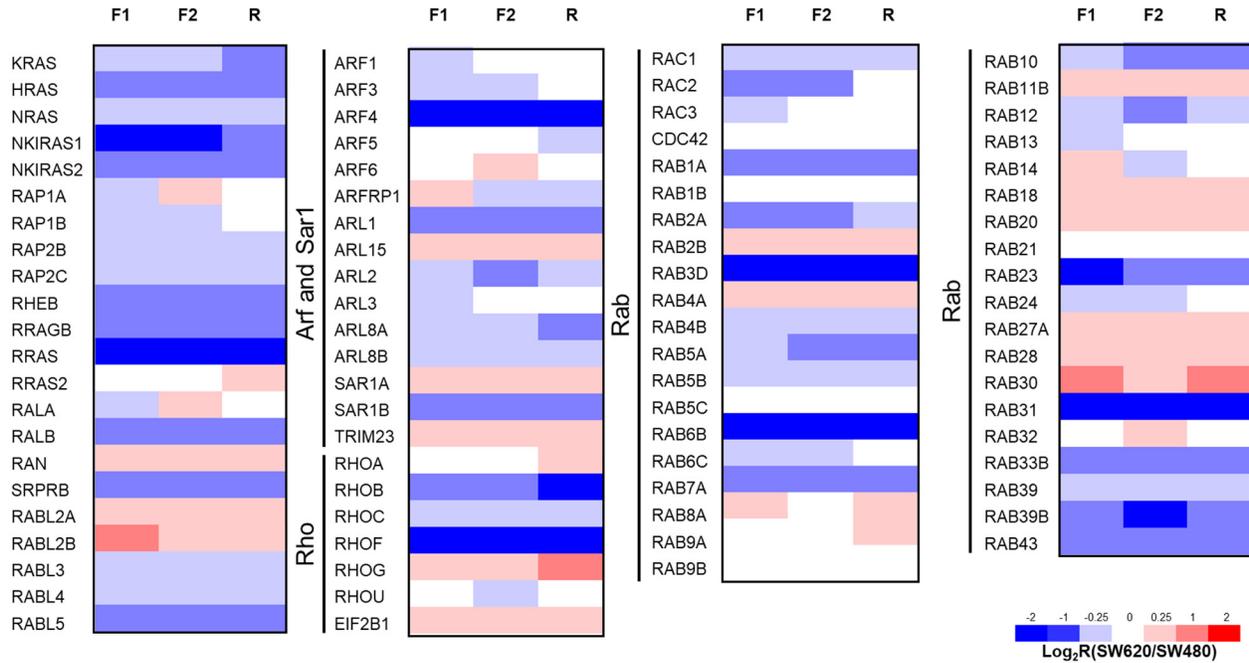
<b>shSAR1B</b>	CCGGGCGAGAGATGTTTGGTTTATACTCGAGTATAAACCAACATCTCTCGCTTTTTG
<b>shScramble</b>	CCGGTCCTAAGGTTAAGTCGCCCTCGCTCGAGCGAGGGCGACTTAACCTTAGGTTTTTG

Note: CTCGAG denotes the loop of the hairpin region.

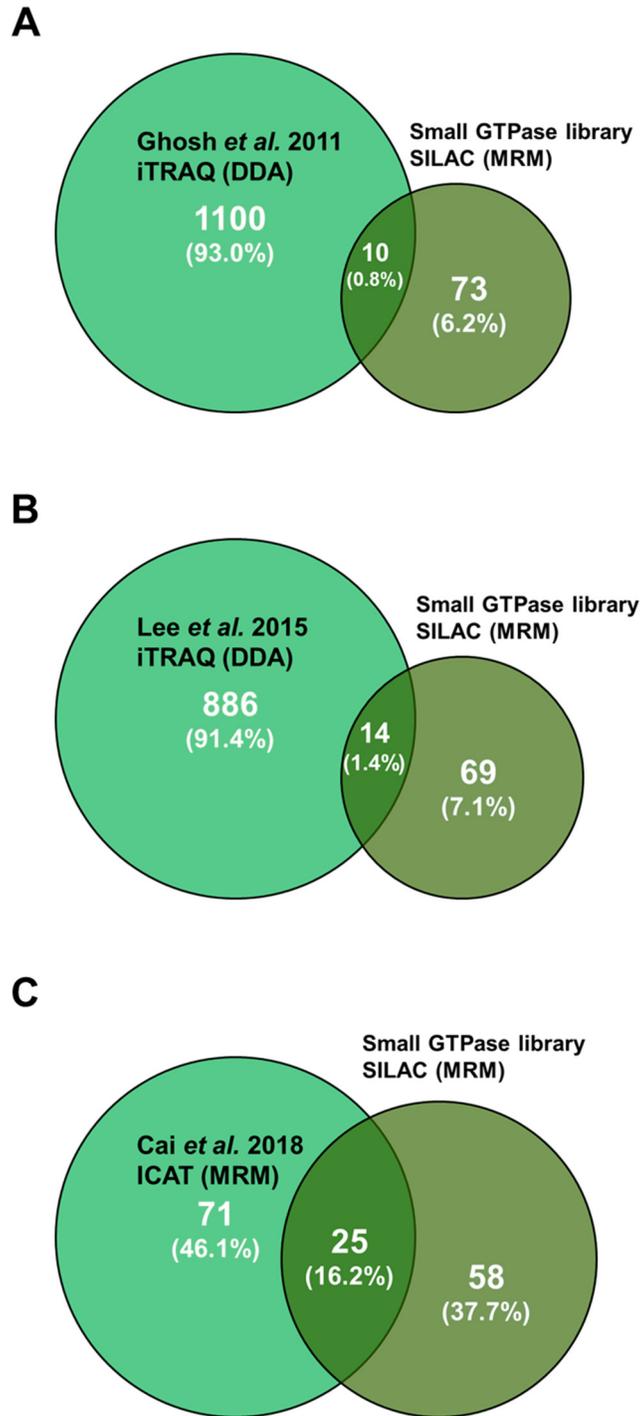
**Table S2B.** A table showing primer sequences used in the RT-qPCR experiments.

<b><i>GAPDH</i> Forward</b>	AGGGCTGCTTTTAACTCTGGT
<b><i>GAPDH</i> Reverse</b>	CCCCACTTGATTTTGGAGGGA
<b><i>SAR1B</i> Forward</b>	TCCCAACATTACATCCCCTTC
<b><i>SAR1B</i> Reverse</b>	GCCATTGATAGCAGGAAGGTAG
<b><i>CDH1</i> Forward</b>	CAACGATGGCATTTTGAAAACAG
<b><i>CDH1</i> Reverse</b>	TCACATCCAGCACATCCAC
<b><i>CDH2</i> Forward</b>	CAGAATCAGTGGCGGAGATC
<b><i>CDH2</i> Reverse</b>	CAGCAACAGTAAGGACAAACATC
<b><i>SNAI1</i> Forward</b>	ACAAGCACCAAGAGTCCG
<b><i>SNAI1</i> Reverse</b>	ATGGCAGTGAGAAGGATGTG
<b><i>SNAI2</i> Forward</b>	ACTGCTCCAAAACCTTCTCC
<b><i>SNAI2</i> Reverse</b>	TGTCATTTGGCTTCGGAGTG
<b><i>VIM</i> Forward</b>	CAACCTGGCCGAGGACAT
<b><i>VIM</i> Reverse</b>	ACGCATTGTCAACATCCTGTCT
<b><i>ZEB1</i> Forward</b>	ACCCTTGAAAGTGATCCAGC
<b><i>ZEB1</i> Reverse</b>	CATTCCATTTTCTGTCTTCCGC
<b><i>ZEB2</i> Forward</b>	TCCAGAAAAGCAGTTCCCTTC
<b><i>ZEB2</i> Reverse</b>	CACACTGATAGGGCTTCTCG

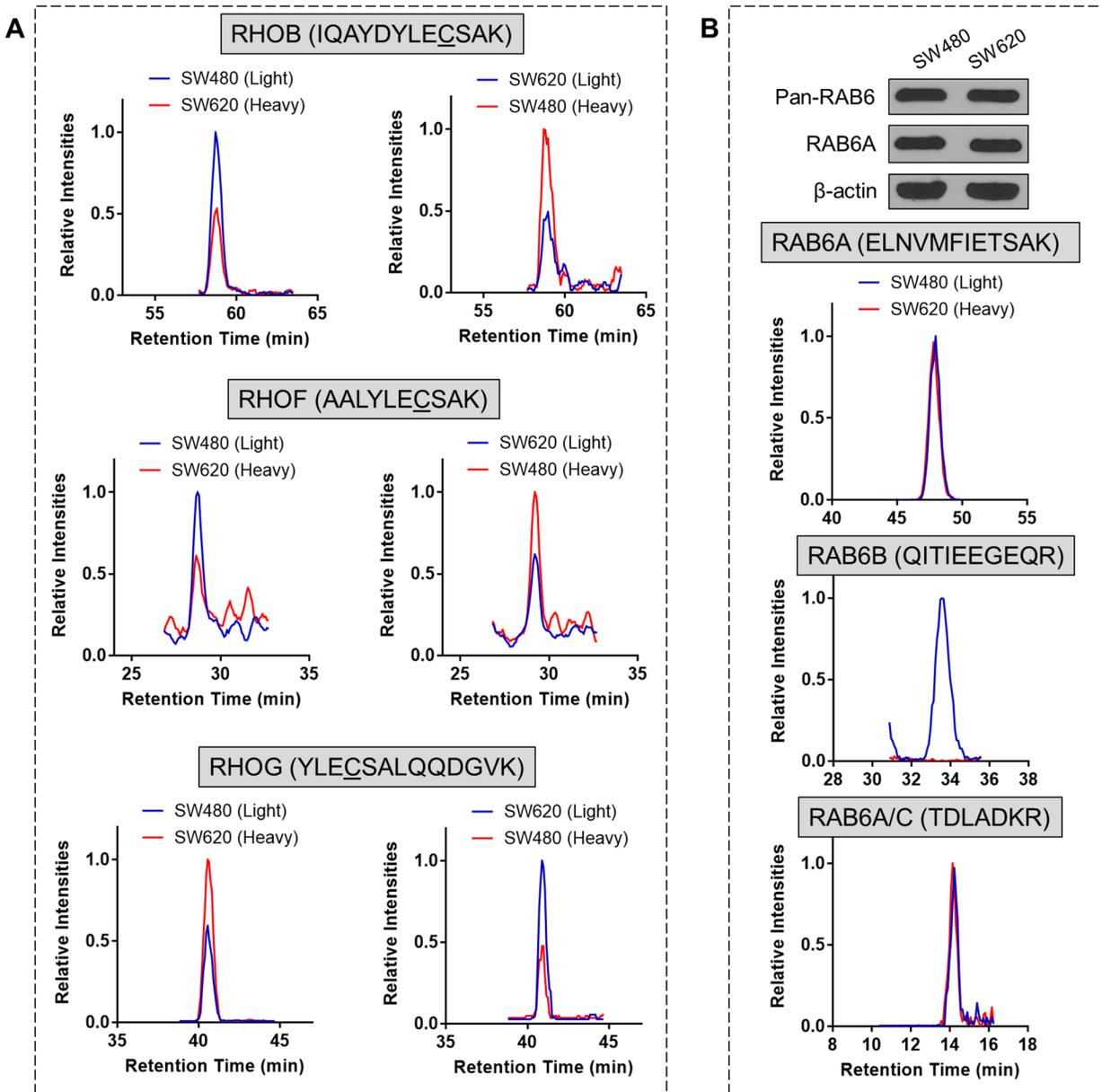
### III. Supplementary Figures



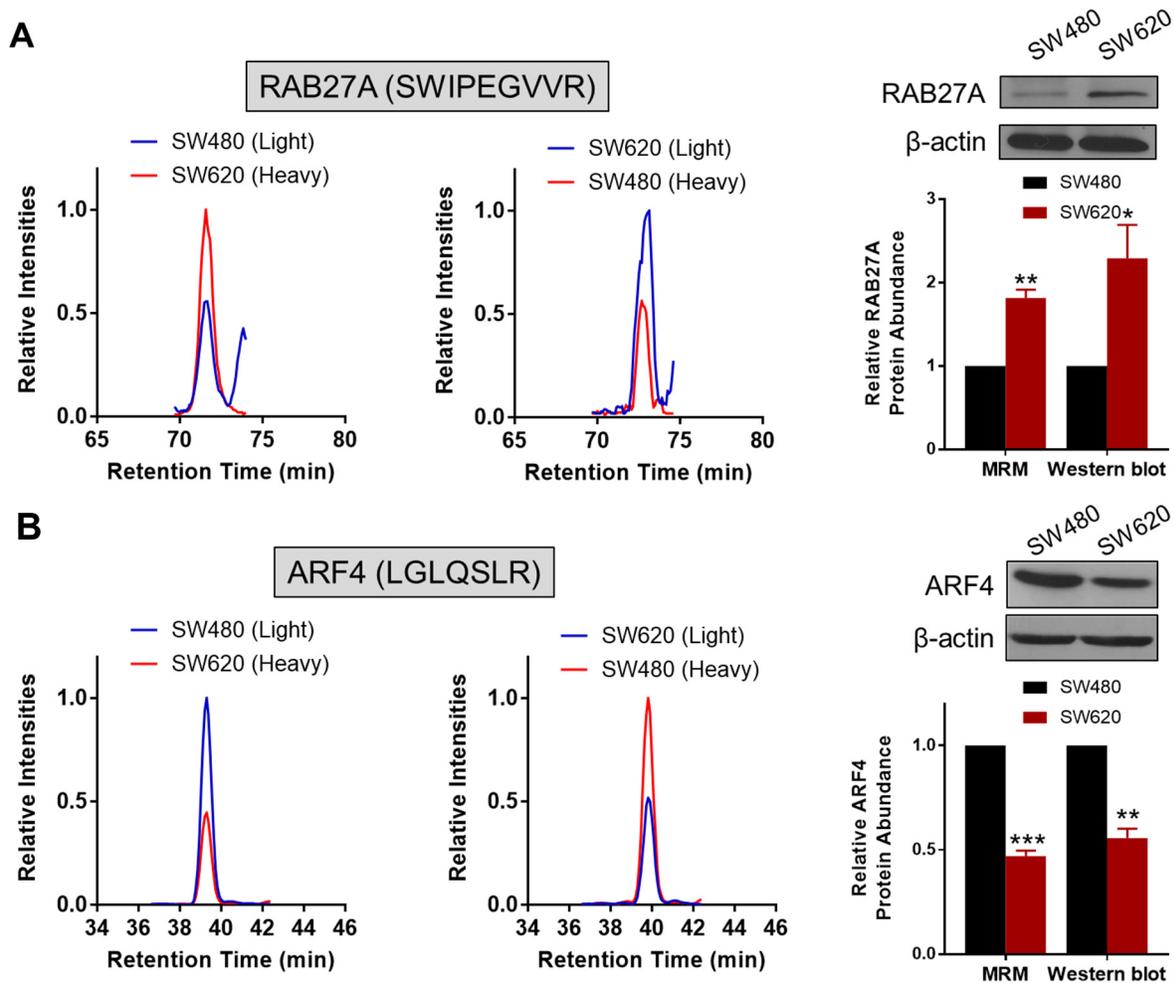
**Figure S1. Scheduled MRM analysis of differential expression of small GTPases in paired SW480/SW620 cells.** A heatmap showing the differential expression of small GTPases in paired SW480/SW620 cells. Shown are the Log<sub>2</sub>R(SW620/SW480) values obtained from two forward and one reverse SILAC labeling experiments (F1 and F2: forward experiments, R1: reverse experiment). As indicated by the scale bar, the red and blue bars designate those small GTPases that are up- and down-regulated, respectively, by at least 1.5-fold in the SW620 over SW480 cells.



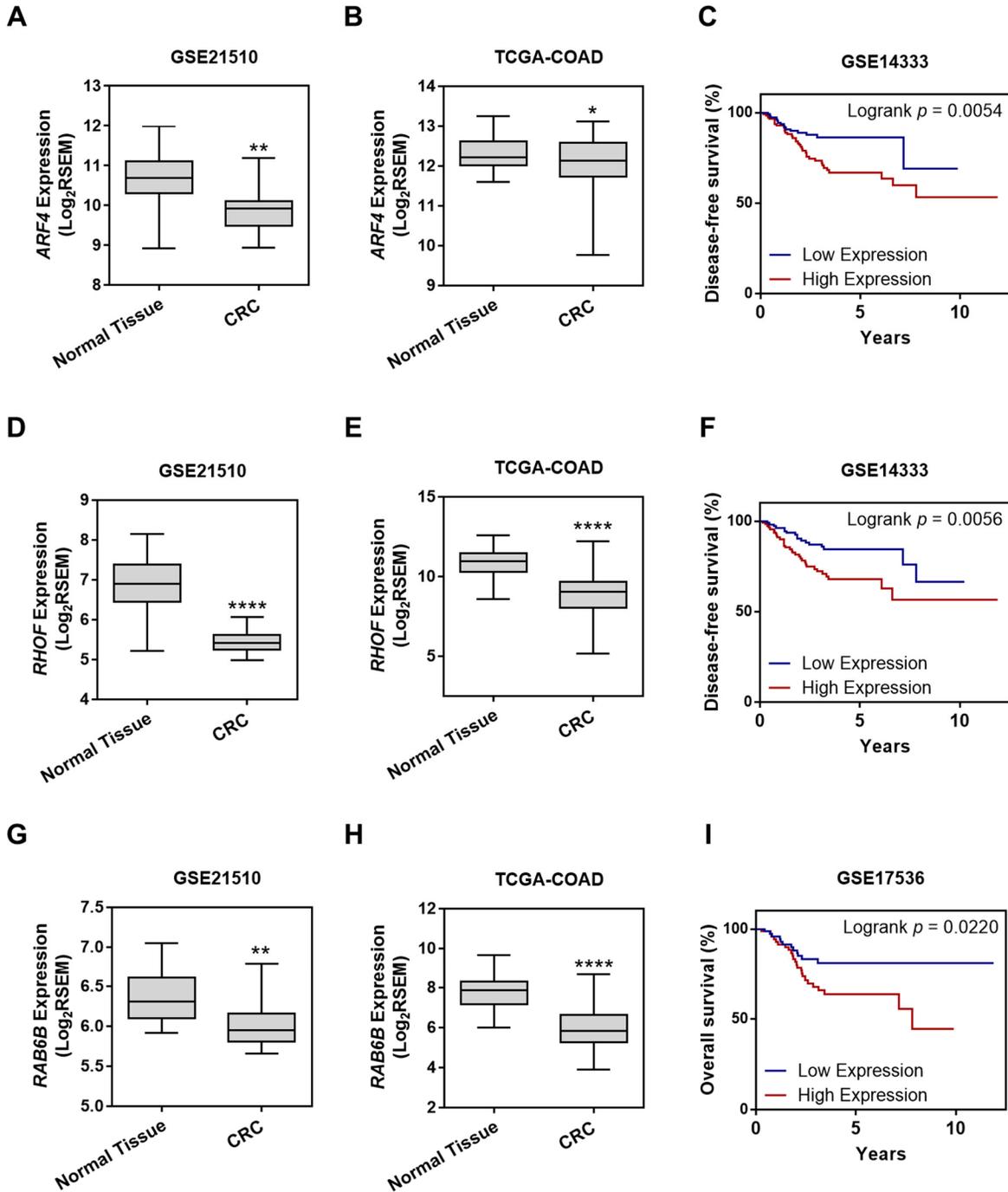
**Figure S2. In-depth coverage of the small GTPase proteome facilitated by the Ge-LC-MRM-based quantification.** Venn diagrams showing the overlapped small GTPases quantified in this study and three independent proteomic studies.<sup>3-5</sup>



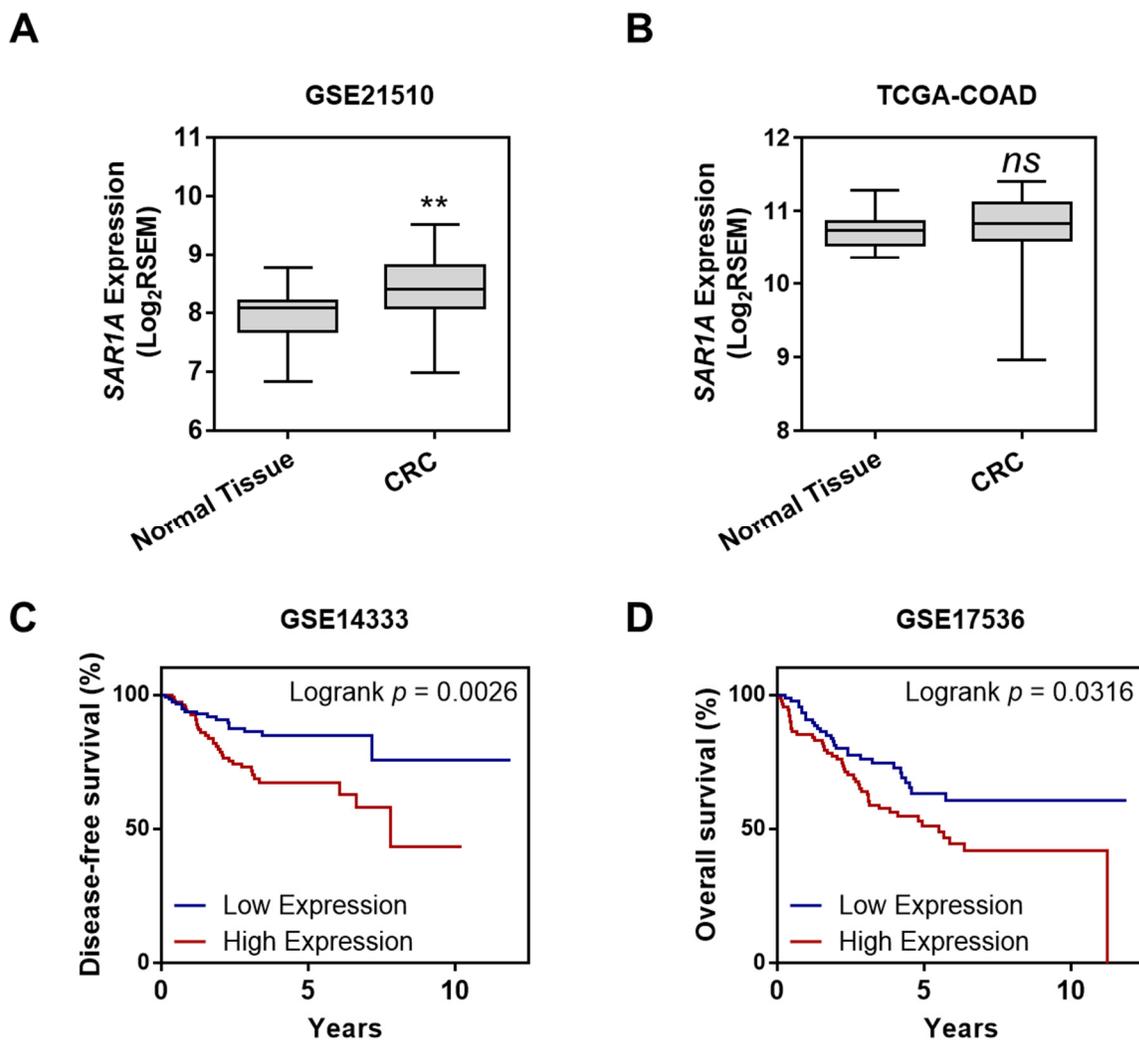
**Figure S3. LC-MRM quantification of RHOB, RHOF, RHOG, and RAB6 in SW480/SW620 cells.** (A) Selected-ion chromatograms (SICs) for the quantification of tryptic peptides IQAYDYLEC<sub>S</sub>SAK (RHOB), AALYLEC<sub>S</sub>SAK (RHOF) and YLEC<sub>S</sub>SALQQDGVK (RHOG) in one forward and one reverse SILAC labeling experiments; (B) Western blot validation of the protein abundance of pan-RAB6 (RABA/B/C) and RAB6A, and the SICs for the quantification of tryptic peptides ELNVMFIETSAK (RAB6A), QITIEEGEQR (RAB6B) and TDLADKR (RABA/C).



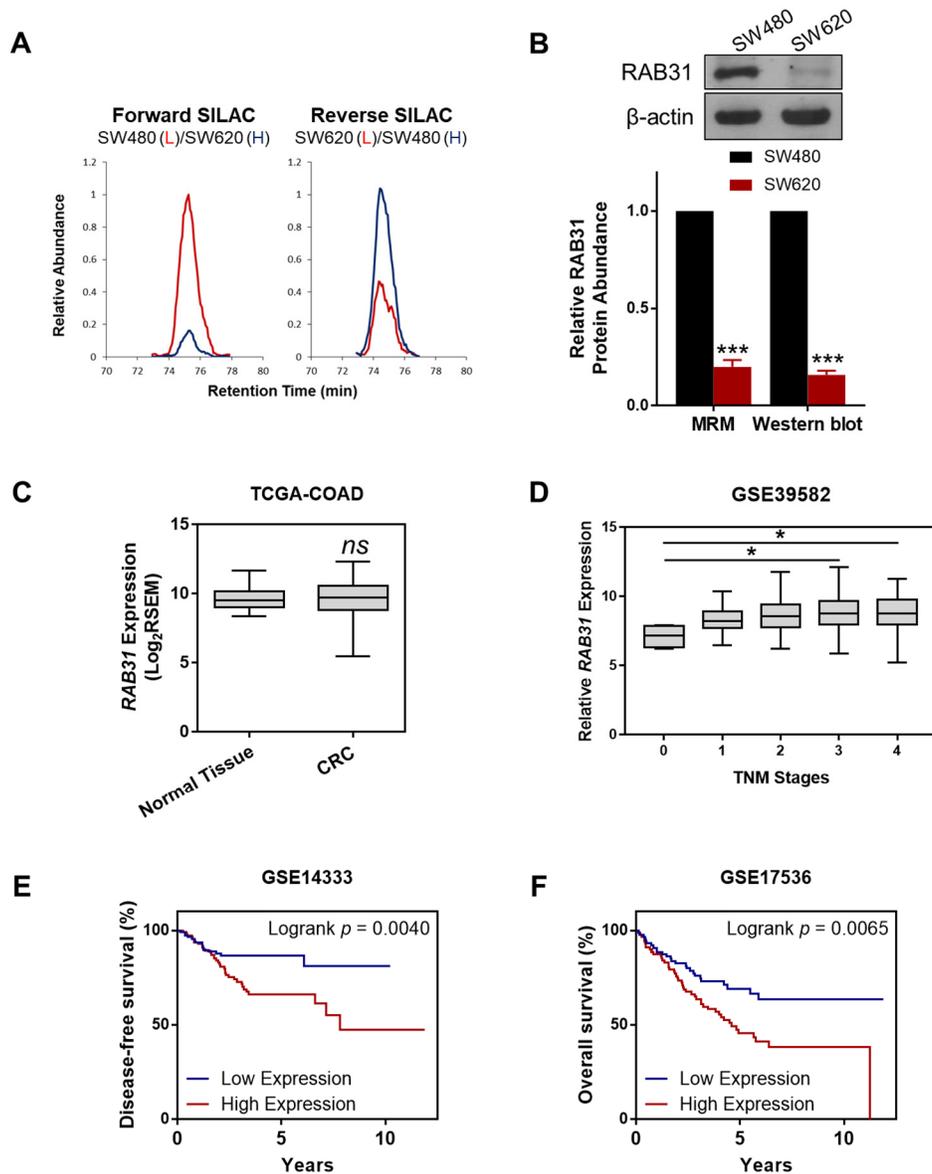
**Figure S4. Validation of differential expression of RAB27A and ARF4 in SW480/SW620 cells.** SICs for the quantification of tryptic peptides (A) SWIPEGVVR from RAB27A and (B) LGLQSLR from ARF4, in one forward and one reverse SILAC labeling experiments; Western blot validation of the differentially expressed (A) RAB27A and (B) ARF4 in SW480/SW620 cells, and quantitative comparison of protein ratios obtained from LC-MRM and Western blot analyses ( $n = 3$ ). The  $p$  values were calculated by using a paired two-tailed Student's  $t$  test (\*,  $0.01 \leq p < 0.05$ ; \*\*,  $0.001 \leq p < 0.01$ ; \*\*\*,  $p < 0.001$ ). The data represent the mean and standard deviation of results obtained from three parallel experiments.



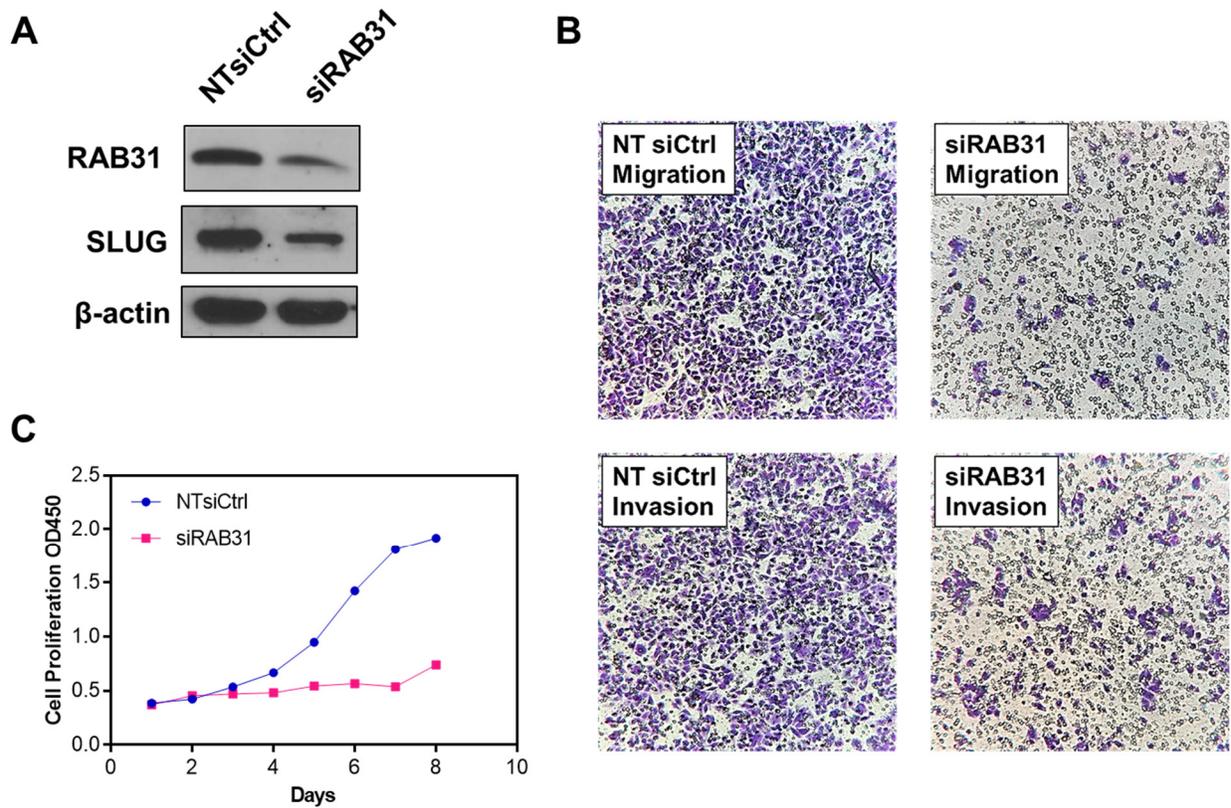
**Figure S5. Prognostic values of ARF4, RHOF and RAB6B in CRC patient cohorts.** Differential mRNA expression of *ARF4*, *RHOF* and *RAB6B* in the (A, D, G) GSE21510 (n = 44) (B, E, H) TCGA-COAD (n = 50) cohorts. The p values were calculated by using an unpaired two-tailed Student's *t* test (ns,  $p > 0.05$ ; \*,  $0.01 \leq p < 0.05$ ; \*\*,  $0.001 \leq p < 0.01$ ; \*\*\*\*,  $p < 0.0001$ ). Kaplan–Meier survival analysis of CRC patients stratified by the median mRNA expression levels of (C) *ARF4* and (F) *RHOF* in the GSE14333 cohort (n = 290) and (I) *RAB6B* in the GSE17536 cohort (n = 232). The log-rank (Mantel–Cox) test was performed to calculate the p values.



**Figure S6. Higher expression of *SAR1A* in CRC confers unfavorable patient prognosis.** Comparison of *SAR1A* expression levels in paired CRC tissues (CRC) with adjacent non-tumor tissues (Normal Tissue) in the (A) TCGA-COAD (n = 50) and (B) GSE21510 (n = 44) cohorts; The  $p$  values were calculated by using a paired two-tailed Student's  $t$  test (ns,  $p > 0.05$ ; \*\*,  $0.001 \leq p < 0.01$ ). Kaplan–Meier survival analysis of CRC patients stratified by the median *SAR1A* mRNA expression in the (C) GSE14333 (n = 290) and (D) GSE17536 (n = 232) cohorts. The log-rank (Mantel–Cox) test was performed to calculate the  $p$  values.



**Figure S7. Higher expression of *RAB31* is associated with worse CRC patient outcome and higher disease stages.** (A) SICs for the quantification of a tryptic peptide GSAAAVIVYDITK from RAB31; (B) Western blot validation of the differentially expressed RAB31 in SW480/SW620 cells, and quantitative comparison of protein ratios obtained from LC-MRM and Western blot analyses ( $n = 3$ ). The  $p$  values were calculated by using a paired two-tailed Student's  $t$  test (\*,  $0.01 \leq p < 0.05$ ; \*\*,  $0.001 \leq p < 0.01$ ). The data represent the mean and standard deviation of results obtained from three parallel experiments. (C) Comparison of *RAB31* expression levels in paired CRC tissues (CRC) with adjacent non-tumor tissues (Normal Tissue) in the TCGA-COAD cohort ( $n = 50$ ); (D) Correlation of *RAB31* mRNA expression with different CRC stages in the GSE39582 cohort ( $n = 585$ ). The  $p$  values were calculated by using an unpaired two-tailed Student's  $t$  test (*ns*,  $p > 0.05$ ; \*,  $0.01 \leq p < 0.05$ ). Kaplan–Meier survival analysis of CRC patients stratified by the median *RAB31* mRNA expression in (E) GSE14333 ( $n = 290$ ) and (F) GSE17536 ( $n = 232$ ) cohorts. The log-rank (Mantel–Cox) test was performed to calculate the  $p$  values.



**Figure S8. RAB31 knockdown modulates the migratory and invasive capacities and cell proliferation of SW480 cells.** (A) Western blot results showing the siRNA-mediated knock-down of RAB31 in SW480 cells and the resulting altered level of Slug protein; (B) Representative images depicting the migratory and invasive abilities of SW480 cells upon siRNA-mediated depletion of RAB31; (C) Cell proliferation of SW480 cells upon siRNA-mediated knockdown of RAB31.