Supporting Information

Cook et al. 10.1073/pnas.1416869112

SI Materials and Methods

Plasmids.

pGEX-Pin1. Pin1 was amplified by PCR from a mammalian GST-Pin1 expression plasmid, Addgene plasmid 19027 (a gift from Michael B. Yaffe, Koch Institute, Massachusetts Institute of Technology, Cambridge, MA) (1), using the forward primer CGCGGATCCGAATGGCGGACGAGGAGAAGCTGCCG and reverse primer GCCGGAATTCTCACTCAGTGCGGAGGATGATGTG, and placed in-frame into pGEX-5x-2 (GE Healthcare) digested in the multiple cloning site (MCS) with BamH1 and EcoRI.

pRTC2-puro. This vector is based on the original retrotransposition reporter (2), but it contains the highly active modern human L1.3 element (3). The pRTC2-puro backbone was derived from pCEP4 (Life Technologies), in which the hygromycin resistance gene was replaced by the CMV promoter-driven *pac* gene cassette that encodes puromycin *N*-acetyl-transferase. This cassette was isolated from the pSELECT-puro-mcs vector (InvivoGen) by PCR using Phusion High-Fidelity polymerase (New England Biolabs) and blunt-end ligated into the NruI site that had been retained from the starting pCEP4 vector. In addition to the *pac* gene, pRTC2-puro differs from the original reporter (2) by the following modifications: the CMV promoter that drove sense transcription of the L1 element was replaced with the SV40 early promoter, the antisense G418 gene was relocated from its original position within the L1 3′ UTR to downstream of it, and its SV40 early promoter was replaced with the Rous sarcoma virus LTR. Mutations in ORF1 were generated by site-directed mutagenesis using the Agilent QuikChange II Kit on a vector that contained just the 5′ UTR and ORF1. The BsiWI/AgeI restriction enzyme fragment, containing the 5′ UTR and ORF1 sequence, was isolated and then inserted into the corresponding sites of pRTC2-puro.

pORF1-FLAG. This mammalian expression plasmid was constructed with pcDNA3.1(+)-puro (from the Don Ganem laboraratory, University of California San Francisco). ORF1-Flag amplicons, containing a 5′ BamH1-Kozak sequence and 3′ EcoRI-FLAG sequence, were generated by PCR with a high-fidelity polymerase from WT or mutant ORF1 pRCT2 templates with the forward primer CGCGGATCCGCAATGGGGAAAAAACAGAAC and reverse primer GCCGGAATTCCTACTTGTCGTCGTCGTCCTTATAATCCATTTTGGCATG. The PCR fragment was inserted into the MCS of pcDNA3.1(+)-puro. Some mutants were made using WT pORF1-FLAG as a template for site-directed mutagenesis with the Agilent QuikChange II Kit.

All mutations were verified by DNA sequencing, and plasmid DNA was purified using the endotoxin-free plasmid DNA purification kit, NucleoBond Xtra Midi EF (Macherey-Nagel).

Cell Culture. HeLa cells were grown in DMEM (Life Technologies) with high glucose and pyruvate, supplemented with 10% FBS, 100 U/mL penicillin, and 100 μg/mL streptomycin.

Protein Purification. GST and GST-Pin1 were expressed and purified from Rosetta (DE3) cells (Novagen) transformed with pGEX or pGEX-Pin1. Cells were grown in an overnight culture of 25 mL LB medium containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol at 37 °C on a rotary shaker at 250 rpm. The next day, the culture was expanded 50-fold into LB with the indicated antibiotics and grown at 37 °C, at 250 rpm, to an OD_{600} of ∼0.6, at which time cultures were induced with 1 mM isopropyl-beta-D-thiogalactopyranoside and incubated for an additional 3 h. Cells were centrifuged, after which the cell pellets were frozen at −80 °C, thawed, and resuspended in PBS(+) [PBS supplemented with 1 mM DTT and 0.5 mM phenylmethylsulfonyl fluoride (PMSF)] with 1× cOmplete EDTA-free protease inhibitor mixture (Roche). The suspension was then sonicated for ∼2 min. Triton X-100 (1%) was added, and lysates were centrifuged at 14,500 × *g* for 25 min at 4 °C. The supernatants were applied to glutathione agarose beads (Sigma-Aldrich) that had been previously washed with PBS(+). The suspension was gently rotated for 2 h at 4 °C, after which the agarose beads were washed four times with PBS(+). Protein was eluted with 5 mM L-glutathione (Sigma-Aldrich) reduced in 50 mM Tris·Cl pH 8.8, and 1 mM DTT during gentle rotation for 1 h at 4 °C. The eluent was dialyzed twice against 50 mM Tris·Cl pH 8.0, 100 mM KCl, 12.5 mM $MgCl_2$, 5 μM $ZnCl_2$, 1 mM EDTA, 20% glycerol, 0.025% Tween, and 1mM DTT. Proteins were quantified after denaturing gel electrophoresis by comparison with known amounts of BSA) after staining with Coomassie G-250 (Novex SimplyBlue SafeStain; Life Technologies).

Untagged ORF1p was purified from H5 insect cells as described previously (4). ORF1p-Flag was purified from HeLa cells as follows: T175 flasks were seeded with 8×10^6 cells in 30 mL of cell culture media and transfected 24 h later with 35 μg of pORF1-Flag constructs using 100 μL of FuGENE6 (Promega) diluted in serum-free media (SFM) according to the manufacturer's protocol. Cells were grown for 48 h, scraped into PBS, centrifuged at 300 × *g*, and lysed in 50 mM Tris·Cl pH 7.4, 650 mM NaCl, 1 mM EDTA, 1% Triton X-100, cOmplete EDTA-free protease inhibitor mixture (Roche), PhosSTOP phosphatase inhibitor mixture (Roche), 100 μM leupeptin, and 1 mM PMSF. Lysates were sonicated in a Bioruptor (Diagneode) and centrifuged at 17,000 × *g* for 15 min at 4 °C. The supernatants were applied to 175 μL of ANTI-FLAG M2 Affinity Gel Resin (Sigma-Aldrich) prepared in accordance with the manufacturer's instructions, and rotated at 4 °C for 4 h. The resin was washed twice with Tris-buffered saline (TBS), 500 mM NaCl, and the foregoing protease and phosphatase inhibitors, then twice without inhibitors. ORF1p-Flag was eluted with 200 μL of 150 ng/μL 1× Flag peptide in TBS with added 500 mM NaCl. The purified proteins were electrophoresed on denaturing gels with varying amounts of insect-purified ORF1p of known concentration, which served as the standard. The proteins were stained with Coomassie G-250, the gels were scanned, and the protein bands were quantified using ImageJ.

LC-MS/MS Analysis. ORF1p-Flag, excised from a Coomassie-stained polyacrylamide gel, and ORF1p purified from H5 cells were analyzed by the Mass Spectrometry and Proteomics Resource of the W.M. Keck Foundation Biotechnology Resource Laboratory at Yale University, New Haven, CT. Samples were digested with trypsin, and phosphorylated peptides were enriched using a titanium dioxide column (TiO_2; GlySci). The enriched (EN) and flow-through (FT) fractions were analyzed by LC-MS/MS using an LTQ Orbitrap Elite equipped with a Waters nanoAcquity (75 μm × 250 mm eluted at 300 nL/min) ultra-performance liquid chromatography (UPLC) system as described previously (5). Data were processed using MASCOT Distiller, and protein identification was performed using the Mascot search algorithm (Matrix Science) (6). Phosphorylation sites were scored based on the Mascot Delta Scoring system (7) and PhosphoRS (8).

Retrotransposition. Six-well plates were seeded in two sets of duplicates at 2×10^5 HeLa cells per well, grown for 24 h, and

transfected with 1 μg of pRTC-puro using 3 μL of FuGENE6 (Promega). At 24 h posttransfection, one set of duplicates was used to determine transfection efficiency by treatment with 10 μg/mL puromycin (InvivoGen) for 24 h, which is sufficient to kill all nontransfected cells. Following puromycin selection, cells were allowed to recover for ∼48 h, fixed with 2% formaldehyde and 0.2% glutaraldehyde in PBS, washed with PBS, and then stained with KaryoMAX Giemsa stain (Life Technologies). The second set of duplicates was treated with 400 μg/mL G418 (Life Technologies) beginning 72 h after transfection for ∼10 d, at which time cells were fixed and stained as above. The relative retrotransposition efficiencies were quantified as described in the main text. Statistical analyses were performed with R (9).

Western Blot Analysis. Six-well plates were seeded with 6×10^5 HeLa cells per well and transfected with 1 μg of pORF1-Flag constructs using 3 μL of FuGENE6 (Promega). Cells were grown for 48 h, washed twice with ice-cold PBS, and then lysed with ice-cold lysis buffer used for ORF1p-Flag purification. Lysates were sonicated and centrifuged at 17,000 x *g* for 15 min at 4 °C. The protein concentration of the supernatants was quantified using the Bradford (BioRad) reagent. Equivalent amounts of supernatant protein were subjected to denaturing gel electrophoresis, transferred to nitrocellulose membranes, and probed with ANTI-FLAG M2 monoclonal antibody (Sigma-Aldrich). Blots were developed with SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) and exposed on HyBlot ES autoradiography film (Denville Scientific). Blots were then treated with Restore PLUS Western Blot Stripping Buffer (Thermo Scientific) according to the manufacturer's protocol and reprobed with antitubulin. To determine transfection efficiencies, parallel 12-well plates were seeded with 3×10^5 HeLa cells, which were then transfected with 0.5 μg of the pORF1-Flag constructs using 1.5 μL of FuGENE6 (Promega), after which the cells were treated with puromycin as described for the retrotransposition assays.

FRET-Based RNA Annealing Assay. We assayed the ability of ORF1p-Flag proteins to anneal two complementary 21-mer RNA oligonucleotides, one labeled with the cyanine fluorophore Cy3 and the other labeled with Cy5: 5′-Cy3-ACUGCUAGAGAUUUUCCACAU-3′ and 5′-Cy5-AUGUGGAAAAUCUCUAGCAGU-3′ (10). Reactions (20 μL) initially contained 20 nM ORF1p-Flag proteins (in terms of monomer), 100 nM Cy3-RNA, and 2× FRET buffer (100 mM Tris·Cl pH 7.4, 300 mM NaCl, 6 mM $MgCl_2$, and 2 mM DTT), which were incubated at 30 °C for ∼5 min in a 96-well half-area, low-binding black plate (Corning; 3993). Annealing was started by injection of 20 μL of 100 nM Cy5-RNA solution in water for a final concentration of 10 nM ORF1p and 50 nM of each RNA oligonucleotide in 1× FRET buffer. Cy3 was excited every 0.7 s at 535 nm (25 nm band pass), and emissions were read at 590/25 (Cy3) and 680/30 (Cy5). On annealing to its Cy5-labeled complement, the Cy3 emission is quantitatively absorbed by Cy5, whereupon Cy5 emits at 680 nm. The FRET ratio (Em_{680}/Em_{590}; *y* axis) was plotted as a function of time (*x* axis), and the annealing rate constant (k_{ann}) was obtained using KaleidaGraph 4.1 to apply a least squares fit to the equation, $y = a + A^*\ [1–1/(e^{k^*t} + 1)]$, where *a* is the *y*-axis offset, *A* is the maximal FRET ratio, *t* is time, and *k* is k_{ann}. All reactions were done in triplicate. Preliminary titration experiments showed that 10 nM ORF1p WT resulted in 50% maximal activity. We used a Synergy2 Microplate Reader running Gen5 data analysis software (Biotek Instruments), which had been fitted with a red photomultiplier tube and a xenon flash lamp.

GST Pull-Down Assay. HeLa cells were seeded at 2×10^6 cells in 100-mm culture dishes and transfected 24 h later with 8 μg of pORF1-Flag constructs using 24 μL of FuGENE6 (Promega) diluted in SFM according to the manufacturer's protocol. Cells were grown for 48 h, washed twice with ice-cold PBS, then lysed with 500 μL of ice-cold pull-down lysis buffer (25 mM Hepes pH 7.9, 150 mM NaCl, 25 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 5 μM $ZnCl_2$, 10% glycerol, 0.15% Nonidet P-40) supplemented with cOmplete EDTA-free protease inhibitor mixture (Roche), PhosSTOP phosphatase inhibitor mixture (Roche), and 100 μM leupeptin, 1 mM PMSF, and 2 mM DTT. Lysates were sonicated in a Bioruptor (Diagneode), cleared by centrifugation at 17,000 × *g* for 15 min at 4 °C, and quantified with a Bradford assay (BioRad). GST or GST-Pin1 proteins (500 nM) were immobilized onto 15 μL of 50% glutathione agarose solution (Sigma-Aldrich) in pull-down lysis buffer via gentle rotation for 2 h at 4 °C, then washed twice with pull-down lysis buffer. Cleared HeLa lysates (1.5 mg) were applied to immobilized GST or GST-Pin1, incubated with gentle rotation for 5 h, washed four times with pull-down lysis buffer (washes 1 and 2 supplemented as above; washes 3 and 4 supplemented only with 2 mM DTT). The proteins were resolved using denaturing electrophoresis, transferred to nitrocellulose membranes, and probed with ANTI-FLAG M2 monoclonal antibody (Sigma-Aldrich). Blots were developed with SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) and exposed on HyBlot ES autoradiography film (Denville Scientific).

SI Results

SI LC-MS/MS ORF1p Data. The LC-MS/MS results for ORF1p purified from H5 insect cells and ORF1p-Flag purified from HeLa cells were generated by the Mass Spectrometry and Proteomics Resource of the W.M. Keck Foundation Biotechnology Resource Laboratory at Yale University. The data are available at the Yale Protein Expression Database (YPED), an open source repository, at yped.med.yale.edu/repository (project name ORF1p). Datasets for the TiO_2 EN and FT MS fractions are shown for each ORF1p construct. Information on score values can be found at medicine.yale.edu/keck/proteomics/literature/literature.aspx.

1. Yaffe MB, et al. (1997) Sequence-specific and phosphorylation-dependent proline isomerization: A potential mitotic regulatory mechanism. *Science* 278(5345):1957–1960.
2. Moran JV, et al. (1996) High-frequency retrotransposition in cultured mammalian cells. *Cell* 87(5):917–927.
3. Sassaman DM, et al. (1997) Many human L1 elements are capable of retrotransposition. *Nat Genet* 16(1):37–43.
4. Callahan KE, Hickman AB, Jones CE, Ghirlando R, Furano AV (2012) Polymerization and nucleic acid-binding properties of human L1 ORF1 protein. *Nucleic Acids Res* 40(2):813–827.
5. Krishnan N, et al. (2012) The prolyl isomerase Pin1 targets stem-loop binding protein (SLBP) to dissociate the SLBP-histone mRNA complex linking histone mRNA decay with SLBP ubiquitination. *Mol Cell Biol* 32(21):4306–4322.
6. Hirosawa M, Hoshida M, Ishikawa M, Toya T (1993) MASCOT: Multiple alignment system for protein sequences based on three-way dynamic programming. *Comput Appl Biosci* 9(2):161–167.
7. Savitski MM, et al. (2011) Confident phosphorylation site localization using the Mascot Delta Score. *Mol Cell Proteomics* 10(2):M110.003830.
8. Taus T, et al. (2011) Universal and confident phosphorylation site localization using phosphoRS. *J Proteome Res* 10(12):5354–5362.
9. R Core Team (2014) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria).
10. Rajkowitsch L, Schroeder R (2007) Coupling RNA annealing and strand displacement: A FRET-based microplate reader assay for RNA chaperone activity. *Biotechniques* 43(3):304–310, 306, 308 passim.
11. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675.

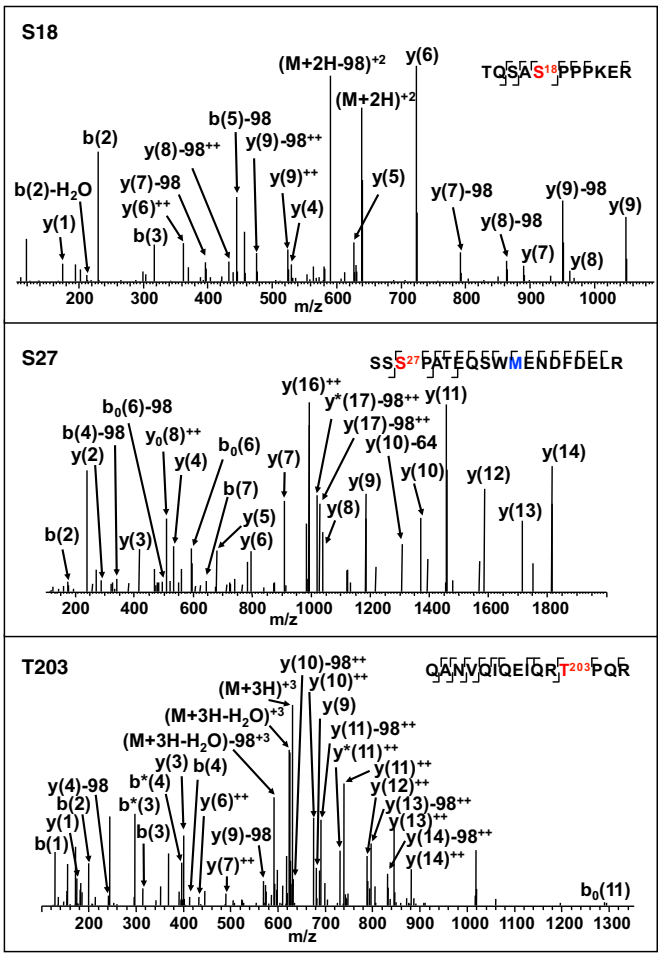


Fig. S1. Verification example of MS/MS spectra analyses to confirm peptide identification and phosphorylation site assignments for ORF1p-Flag S18, S27, and T203 purified from HeLa cells. Fragmentation patters (b and y ions) and their corresponding water loss (denoted by a subscript "0"); -NH3 loss (denoted by "*"); or phosphor loss (denoted by "−98") are assigned based on MS/MS matches to theoretical predicted fragmentation patters corresponding to the peptide (with mass, "M") of interest. The "$^{++}$" indicates a doubly charged peptide or fragment mass. The sequence inset shows the corresponding b ions (from N terminus to cut site) and y ions (from C terminus to fragment site).

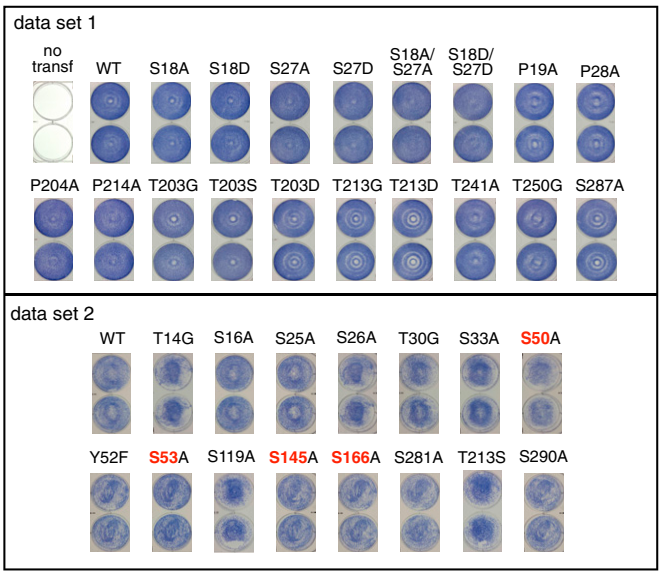
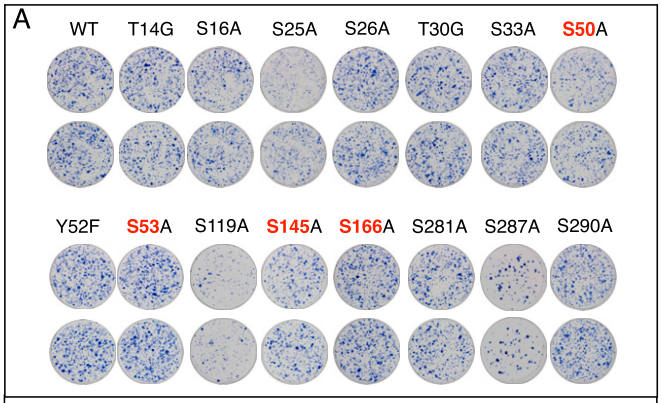


Fig. S2. Transfection efficiency. Giemsa-stained cells from transfections performed in parallel with each retrotransposition assay showing the relative transfection efficiencies of each ORF1p mutant L1 construct. Cells were treated with puromycin for 24 h starting at 1 d posttransfection as described in the main text.

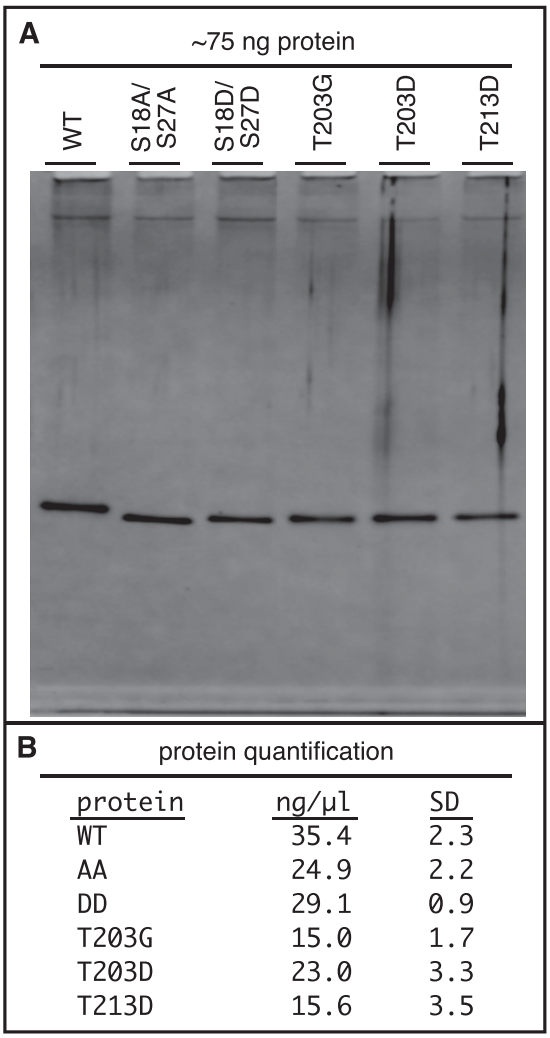


B

relative colony area

sample	well #1	well #2	mean	sample	well #1	well #2	mean
WT	0.94	0.97	0.96	Y52F	1.13	1.13	1.13
T14G	0.95	0.80	0.88	S53A	1.25	1.27	1.26
S16A	0.85	0.86	0.86	S119A	0.26	0.27	0.27
S25A	0.66	0.93	0.79	S145A	0.80	0.89	0.85
S26A	1.16	1.11	1.14	S166A	1.01	1.00	1.01
T30G	0.90	0.93	0.91	S281A	0.82	0.75	0.78
S33A	0.95	0.89	0.92	S287A	0.24	0.21	0.22
S50A	0.70	0.72	0.71	S290A	0.97	1.06	1.01

Fig. S3. Effects of mutations of non-PDPK sites in ORF1p on L1 retrotransposition. (*A*) Wells showing Giemsa-stained foci generated from HeLa cells transfected with the WT L1 retrotransposition vector or indicated ORF1 mutants. Two of four WT samples are shown. The residues indicated in red are the high-confidence phospho sites in HeLa cells. (*B*) The percent colony area for the two WT wells shown in *A* are 37.2 (*Top*) and 38.44 (*Bottom*); percent colony area for the remaining two WT samples were 43.7 and 38.6, for an overall WT mean of 39.5 ± 3.9. The relative colony areas for each mutant are normalized to the WT mean of 39.5.



protein	ng/µl	SD
WT	35.4	2.3
AA	24.9	2.2
DD	29.1	0.9
T203G	15.0	1.7
T203D	23.0	3.3
T213D	15.6	3.5

Fig. S4. Purified ORF1p-FLAG proteins. (*A*) Approximately 75 ng of each purified ORF1p-Flag protein is shown in a silver-stained polyacrylamide gel (Pierce, Thermo Scientific). (*B*) Relative concentrations of ORF1p-Flag proteins were determined with five independent gels, each containing a standard curve of ORF1p purified from insect cells and stained with Coomassie G-250 (Novex SimplyBlue SafeStain; Life Technologies); gels were imaged and quantified using the ImageJ Gel Analysis Module (11).

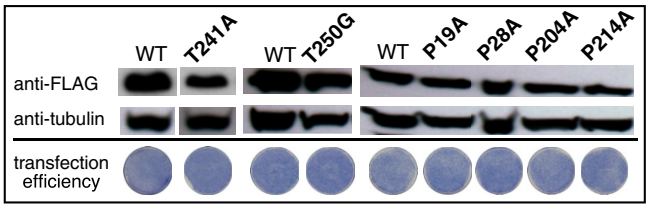


Fig. S5. Mutant ORF1p proteins are stably expressed in HeLa cells. Western blot analysis showing expression of ORF1p-Flag WT and mutant constructs in HeLa cells. Total amounts of whole-cell lysates loaded for each set of WT and mutants: T241, 185 µg; T250, 125 µg; prolines, 185 µg. Tubulin served as a loading control, and the transfection efficiency was assessed by puromycin selection.