Supplementary figure legends

Supplementary Figure 1. Optimization of ORF1p and ORF2p purification (related to Figure 2)

All gels stained with CBB. All buffers contain 20 mM HEPES-Na pH 7.4 with 1% v/v Triton X-100 supplemented protease inhibitors unless otherwise indicated. (A) Exploration of conditions for pLD288, ORF1p-Flag. The indicated salts were used. In the Tween lane, Tween-20 detergent was used in place of Triton X-100. pCEP-Puro represents empty vector control. (B) Exploration of conditions for pLD401, ORF2p-3xFlag. Leftmost four lanes optimize the ratio of powder used (mg) to anti-Flag DBs (µl). In the CHAPS lane, 5mM CHAPS detergent (0.3%) was used in place of Triton X-100. (C) 500mM NaCl instead of 300mM NaCl results in higher yield and reduced background for both ORF1p and ORF2p. Increased yield is due to improved extraction from powder into lysate (not shown). (D) Optimization of native elution with 3xFlag peptide or protease cleavage for ORF1p and ORF2p. Buffers all contain 500mM NaCl. (E) Two samples of 3xFlag-peptide eluted ORF1p and ORF2p (pLD401, see panel D) were quantified by CBB staining densitometry. The numbers shown diagonally in red indicated measured and standard values. After correction for molecular weights, the molar ratio of ORF1p:ORF2p is 2.8 for both samples. (F) Optimization of the amount of anti-flag DynaBeads used in an ORF1p purification from pLD288 (*ORFeus*-Hs). 50 mg powder was used for each experiment. A pooled lysate was made and aliquotted and the indicated amount of DynaBeads were added for a 30-minute binding, then the samples were eluted with LDS sample buffer. (G) Optimization of binding time for pLD288 with 50mg powder each and 50µl anti-Flag DynaBeads. A pooled lysate was made and aliquotted for each experiment and samples were eluted with LDS sample buffer.

Supplementary Figure 2. Quantitation of ORF1p:ORF2p expression in HEK293T cells shows a ratio of 1,000-10,000:1. (related to Figure 1 and 2)

(A) Tet-On HEK293T$_{LD}$ cells were transfected with pLD401 or pLD288 and puromycin selected, then lysed after 24 hrs with or without doxycyline induction as indicated. Equal protein amounts were loaded and blotted with anti-Flag. (B) Purified CK2α-3xFlag or CK2β-Flag was loaded in the indicated amount and non-saturated chemiluminescent signals were quantified. Numbers shown diagonally in red indicated measured and standard values. By multiplying the ratio of proteins in (A) by the ratio of sensitivities of the tags here, ~7,500x as much ORF1p is present in cell extracts than ORF2p; based on error inherent in these measurements the actual range for ORF1p:ORF2p we estimate is 1,000-10,000:1. This does not include correction for different transfer efficiencies at different molecular weights, which would result in a lower ratio. (C) Puromycin-selected Tet-On HEK293T$_{LD}$ cells containing the indicated constructs were induced with doxycycline, as indicated, lysed, and equal protein amounts were loaded and blotted. Non-saturated chemiluminescent signals were quantified; numbers shown diagonally in red indicated measured and standard values. The experiments shown in A-C represent three separate transfections on different days. (D) Four different amounts of pMT302 lysate were loaded to provide a standard curve for ORF1p and tubulin antibodies. After normalizing to the tubulin loading control, the relative amounts of each protein were computed. As compared to L1RP, *ORFeus*-Hs produces ~2.5-fold more ORF1p and 20-40 fold more ORF2p.

Supplementary Figure 3. RNPs Purified using ORF1p and ORF2p are active in reverse transcriptase assays. (related to Figure 3)

(A) One µL of affinity purified L1 RNP or the commercial reverse transcriptase SuperScript III (Invitrogen) was added in a 20 µl reaction mixture. Before the assay, purified L1 RNP from pLD401 was diluted 1:5 and SuperScript III was diluted 1:10. For the treatment of RNase, 10 units of RNase A/T1 was added to the reaction mix. (B) Realtime RT-PCR amplification curve. (C) The arrow indicates that an equal amount of final PCR products were loaded on ethidium bromide stained agarose gel. LEAP and total L1 cDNA reactions was carried out as described using 1 µl of RNP prepared by either affinity-purification or sucrose cushion velocity sedimentation such that the input represented an equal number of cells. Total cDNA was made using Superscript III (SS3) reverse transcriptase and an equal amount of boiled RNP, and then both LEAP and SS3 products were PCR amplified together. In addition, we noted a different product size from LEAP and SS3, as seen previously (Kulpa and Moran, 2006). We cut and cloned various products from both LEAP and Superscript and then sequenced ~50 clones. Clones differed only by poly(A) tail length; for Superscript III, the poly(A) tail was 17±5.4 nt long (min 11, max 32), and for LEAP it was 89.2 ± 54.8 nt long (min 14, max 211); some of these were likely longer but limited by the sequencing reaction and perhaps by transformation limitations of *E. Coli*. (D) Total RNA was extracted using Trizol from RNP preps starting from the same amount of cells using two different methods (pullout and sucrose cushion). Relative LEAP signal and L1 RNA levels are from Figure 3B; RNA and LEAP activity enrichment was calculated based on the amount of total RNA.

Supplementary Figure 4. Quantitation of ORF1p:ORF2p in tandem purified L1 RNPs. (related to Figure 3)

RNPs were tandem affinity purified from LD401 or MT302, both containing ORF2p-3xFlag, first using anti-Flag with 3xFlag peptide elution and second with anti-ORF1p, as diagrammed. Supernatant (Super) and elution from the anti-ORF1p purification are shown for three separate purifications done from one common solubilized powder. Gels were stained with Sypro Ruby or colloidal Coomassie blue as indicated and quantified against the standard curve of BSA. Ratios of ORF1p:ORF2p were then calculated, correcting for the proteins' different molecular weights.

Supplementary Figure 5. Indirect Immunofluorescence in HeLa cells of ORF1p and ORF2p from L1RP and *ORFeus*-Hs cells under different promoters. (related to Figure 3)

Anti-ORF1 antibody used in all panels is rabbit monoclonal JH73. For constructs with a CMV promoter and pLD288, Hela or Tet-On HEK293T_{LD} cells were plated on coverslips and transfected 8-16 hours after plating. 24 hours later, cells were fixed, permeabilized, and stained with anti-Flag and anti-ORF1 antibodies and then fluorescent-conjugated secondary antibodies. For remaining constructs with a pTRE promoter, Hela or Tet-On HEK293T_{LD} cells were transfected and puro selected for three days. Then puro resistant cells were plated on coverslips and induced by adding Dox 8-16 hours after plating. For Tet-On HEK293T_{LD} cells, fibronectin-coated coverslips were used.

In the left panels, images were captured using a 40x objective and normalized identically for each construct; cells expressing ORF1p alone or both ORF1p and ORF2p at levels over background were counted. No cells were found to express ORF2p alone, except for pLD561, in which no cells were found to express ORF1p. In the right panels,

within each cell type, the same exposure and normalization was used for each particular antibody. Representative fields are shown. Images all taken using an epiflourescent microscope.

Supplementary Figure 6. High confidence interactors of ORF1p and ORF2p from six I-DIRT experiments. (related Figure 4)

Histograms plot the number of recovered proteins as compared to their heavy isotope content for six individual I-DIRT experiments. See Table S1 for details. Yellow bars represent non-significant groups, blue bars represent statistically significant specificity (i.e. percentage of heavy isotope content) after Benjamini–Hochberg correction for false discovery. wP-values were determined by (see Experimental Procedures). Note that due to binning, the leftmost blue bar may contain some non-significant proteins; only significant proteins are listed. ‡ Because the distribution is non-normal for the anti-ORF1p pullout for L1RP ORF1p-Flag, no p-Values can be generated and the %heavy is not normalized. An arbitrary cutoff of >71% heavy was assigned. LD288.1, LD288.2, LD401.1, and LD401.2 notation represent our first and second I-DIRT experiments, respectively, with tagged ORF1p and ORF2p, respectively. The second experiment was further optimized with respect to binding time and lysate to Dynabead ratios.

Supplementary Figure 7. Additional data supporting that L1 interact with UPF1 through the ORF2 sequence or ORF2p, and that the PCNA-L1 interaction was abolished in EN and RT mutants. (related to Figure 5 and 6)

(A) Affinity purification of L1 RNP was done using the powders of pLD288 (full-length, ORF1-Flag), pLD603 (ORF1-Flag only, no 3' UTR), pLD617 (ORF1-Flag only, with 3' UTR), pLD401 (full-length, ORF2-3XFlag) and pLD633 (full-length, ORF2-3XFlag, no 3' UTR). Eluted samples were split into two halves. Five μg total RNA

extracted from HEK293T cells transfected with pCEP puro eGFP were added to one half of the RNP samples as an extraction internal control. Then total RNA was prepared from the RNPs using Trizol reagent. Realtime RT-PCR was done to quantify the L1 mRNA level with primers within both the ORF1 and ORF2 sequences with and one pair for eGFP as internal control. Triplicate samples were performed and standard error is shown. pLD288 and pLD401 signals were assigned as 1. The Y axis is $\log_2$. (B) The second half of the RNP samples for pLD288 and pLD603 was resolved on parallel SDS PAGE gels and stained by Sypro and Colloidal Coomassie Brilliant Blue (CBB) to quantify the ORF1p in each sample vs a standard curve of BSA. The values in red are the average of three different lanes. (C) Knockdown of endogenous UPF1 increased cellular L1 mRNA levels. The L1 expression plasmid (pLD401) was transfected into Tet-On HEK293T_{LD} cells and selected with puromycin for more than three days without adding dox. Then siRNA targeting UPF1 was introduced into the puro^R cell pool pre-transfected with the L1 vector. Two days later, doxycycline (50 ng/ml) was added to induce L1 expression, and L1 RNA and protein levels were assayed by realtime PCR and immunoblotting. Triplicate samples were performed and standard error is shown. (D) Knockdown of endogenous UPF1 increased L1 protein level in the cell. Triplicate samples were performed and standard error is shown. (E) Knockdown of endogenous UPF1 reduced L1 retrotransposition efficiency. Anti-UPF1 siRNA was introduced into Tet-On HEK293T_{LD} cells. Two d later, pLD478 (*ORFeus*-Hs with GFPAI) was transfected into the siRNA pre-transfected cell pool and then we proceeded with the standard retrotransposition assay. The retrotransposition efficiency from mock transfection was assigned as 1 and all values were the mean of four independent experiments. Triplicate samples were performed and standard error is shown. (F) Knockdown of UPF1 at different time points. Cell lysate was prepared at different time points after siRNA transfection and probed for

UPF1. Actin was used as loading control. (G) ORF2p and UPF1 immunoblot signals were quantified from the elution of the anti-Flag ORF2p pullout in Figure 5D. The signals correlate with r^2=0.995. (H) PCNA-L1 interaction is reduced in EN- and RT- mutants. Tet-On HEK293T_{LD} cells were transfected with different constructs and L1 RNP was purified using anti-Flag conjugated Dynabeads. pLD401 was used as WT control, and was also purified using IgG conjugated Dynabeads as a negative control. (I) EN- and RT- mutants are deficient for L1 retrotransposition but maintained high RT activity (except D702Y) and L1 RNA binding capacity. pLD401 was used as WT control. RT activity was measured by *in vitro* homopolymer RT assay. RNA binding activity was measured by Trizol extraction of RNA from RNP samples followed by realtime RT-PCR. All results were done in triplicates and normalized to the ORF2p level shown in H. Standard error is shown.

Supplementary Table 1. Comprehensive mass spectrometry data for all six I-DIRT experiments. Provided with description as a separate Microsoft Excel formatted file.

Supplementary Table 2. DNA plasmids used in this study. Provided with description as a separate Microsoft Excel formatted file.

Supplementary Table 3. Primers used in this study

Name	Sequence
JB11560	5'- GCGAGCACAGAATTAATACGACTCACTATAGGTTTTTTTTTTTT-3'
JB11564	5'- GCGAGCACAGAATTAATACGACTC-3'
JB14067	5'- GGATCCAGACATGATAAGATACATTGATGA-3'
JB12931	5'- GCTCGTCGTCGACAACGGCT-3',
JB12932	5'- CAAACATGATCTGGGTCATCTTCTC-3'
JB13421	5'- GACATCGGCAAGGTGTGG-3'
JB13422	5'- AGGCCTTCCATCTGTTGCT-3'
JB13415	5'- GCTGGATGGAGAACGACTTC-3'
JB13416	5'- TTCAGCTCCATCAGCTCCTT-3'
JB13417	5'- CTGATCAGCCGCATCTACAA-3'
JB13418	5'- TGGTCTTGATCTGCATCTCG-3'
JB15869	5'-AAGGAGCTGATGGAGCTGAA-3'
JB14116	5'- CGGGATCCGTATACGCGGCCGCTCCAGGCGATCTGACGGTTCACTAA-3'
JB14135	5'- ATCGTGTACAAGGCGTATCACGAGGCCCTTTCGTCTTCA-3'
JB15970	5'- CGCTTCACGTAGTCCCAGAT-3'
JB15871	5'- GCTGGATGGAGAACGACTTC-3'
JB15872	5'- CCTTGGTCTTCAGCTCCATC-3'
JB16703	5'- AACTGGCAAGACAGCTCTGG-3'
JB16704	5'- CCCATGGGATTCTCTGTCTC-3'
JB16665	5'- ATGGAGGCGAAGACAAGAGA-3'
JB16666	5'- TTTTCCTCTTCCCGCAAAC-3'
JB16661	5'- AGACAAAACAGGCGCATTCT-3'
JB16662	5'- GCCCATGTTCTTCTGGTGTT-3'
JB16607	5'- GAGGTGGATGGGTTCTCTGA-3'
JB16608	5'- GCACCTGGCCCTTTTCTATC-3'
JB16603	5'- TGTAAACCTGCAGAGCATGG-3'
JB16604	5'- TTCTCCTGGTTTGGTGCTTC-3'
JB15433	5'- CTGTGGAAGGAAAACCCTTCT-3'
JB15434	5'- GCAAAGTGAAGTAGGCGATTCT-3'
JB16671	5' -CGAATCGGTGCCCACT -3'
JB16672	5' -CTGGCCAGCAATAAGGACTG -3'
JB13766	5'- ACGTAAACGGCCACAAGTTC-3'
JB13767	5'- AAGTCGTGCTGCTTCATGTG-3'

Supplementary Table 4. Tag sequences

Construct	Tag Location	Tag Sequence (Protein)	Tag Sequence (DNA)
pLD260	ORF1p-N	MDYKDDDDK	ATGGATTACAAGGATGACGATGACAAG
pLD261	ORF1p-C	DYKDDDDK*	GATTACAAGGATGACGATGACAAGTAA
pLD288 pMT292	ORF1p-C	KLGGGSGGGSENLYFQGGKPIPN PLLGLDSTYPYDVPDYADYKDDD DK*	AAGCTTGGCGGCGGCAGCGGCGGCGGCAGCGAAAACCTGT ATTTTCAGGGCGGTAAGCCTATCCCTAACCCTCTCCTCGG TCTCGATTCTACGTACCCATACGATGTTCCAGATTACGCT GATTACAAGGATGACGATGACAAGTAA
pLD401 pLD458-61 pMT290 pMT294 pMT297	ORF2p-C	GGGSGGGSHHHHHHHHLEVLFQG PLEVLFQGPLKEQKLISEEDLEQ KLISEEDLDYKDHDGDYKDHDID YKDDDDK*	GGCGGCGGCAGCGGCGGCGGCAGCCATCACCATCACCATC ACCATCACCTGGAAGTTCTGTTCCAGGGGCCCCTGGAAGT TCTGTTCCAGGGGCCCCTTAAGGAACAGAAGCTAATCTCA GAAGAAGACCTGGAACAGAAGCTAATCTCAGAAGAAGACC TGGACTACAAAGACCATGACGGTGATTATAAAGATCATGA CATCGATTACAAGGATGACGATGACAAGTAGATTTAAAT →3′ UTR
pLD415 pLD416	ORF1p-C	KLGGGSGGGSENLYFQGGKPIPN PLLGLDSTYPYDVPDYA DYKDHDGDYKDHDIDYKDDDDK*	AAGCTTGGCGGCGGCAGCGGCGGCGGCAGCGAAAACCTGT ATTTTCAGGGCGGTAAGCCTATCCCTAACCCTCTCCTCGG TCTCGATTCTACGTACCCATACGATGTTCCAGATTACGCT GACTACAAAGACCATGACGGTGATTATAAAGATCATGACA TCGATTACAAGGATGACGATGACAAGTAA
pLD402	ORF2p-C	GGGSGGGSHHHHHHHHLEVLFQG PLEVLFQGP[eGFP from LAP tag]*	GGCGGCGGCAGCGGCGGCGGCAGCCATCACCATCACCATC ACCATCACCTGGAAGTTCTGTTCCAGGGGCCCCTGGAAGT TCTGTTCCAGGGGCCC[eGFP from LAP tag]TAGATTTAAAT →3′ UTR
pLD458	ORF1p-C	Linker A: KL [eGFP from LAP tag]*	AAGCTT[eGFP from LAP tag]TAA
pLD459	ORF1p-C	Linker B: GDGAGLKL[eGFP from LAP tag]*	AAGCTT[eGFP from LAP tag]TAA
pLD460	ORF1p-C	Linker C: GGGSGGGSENLYFQG[eGFP from LAP tag]*	GGCGGCGGCAGCGGCGGCGGCAGCGAAAACCTGTATTTTC AGGGC [eGFP from LAP tag]TAA
pLD461	ORF1p-C	Linker D: GGGSGGGSENLYFQGGGGSGGGS [eGFP from LAP tag]*	GGCGGCGGCAGCGGCGGCGGCAGCGAAAACCTGTATTTTC AGGGCGGCGGCGGCAGCGGCGGCGGCAGC [eGFP from LAP tag]TAA
pMT293-5	ORF1p-C	GGGSGGGSENLYFQGEQKLISEE DL*	GGCGGCGGCAGCGGCGGCGGCAGCGAAAACCTGTATTTTC AGGGCGAACAGAAGCTAATCTCAGAAGAAGACCTGTAA
pMT296-8	ORF1p-C	GGGSGGGSENLYFQGEQKLISEE DLEQKLISEEDLEQKLISEEDL*	GGCGGCGGCAGCGGCGGCGGCAGCGAAAACCTGTATTTTC AGGGCGAACAGAAGCTAATCTCAGAAGAAGACCTGGAACA GAAGCTAATCTCAGAAGAAGACCTGGAACAGAAGCTAATC TCAGAAGAAGACCTGTAA
pMT291 pMT295 pMT298	ORF2p-C	GGGSGGGSLEVLFQGPLEVLFQG PLEVLFQGPLEVLFQGPLEVLFQ GPDYKDHDGDYKDHDIDYKDDDD K*	GGCGGCGGCAGCGGCGGCGGCAGCCTGGAAGTTCTGTTCC AGGGGCCCCTGGAAGTTCTGTTCCAGGGGCCCCTGGAAGT TCTGTTCCAGGGGCCCCTGGAAGTTCTGTTCCAGGGGCCC CTGGAAGTTCTGTTCCAGGGGCCCGACTACAAAGACCATG ACGGTGATTATAAAGATCATGACATCGATTACAAGGATGA CGATGACAAGTAA
pMT302-4 pMT30 pMT304	ORF2p-C	GGGSGGGSLEVLFQGPDYKDHDG DYKDHDIDYKDDDDK*	GGCGGCGGCAGCGGCGGCGGCAGCCTGGAAGTTCTGTTCC AGGGGCCCGACTACAAAGACCATGACGGTGATTATAAAGA TCATGACATCGATTACAAGGATGACGATGACAAGTAA

Key:

AAGCTT - HindIII site (KL)
GGGSGGGS - Flexible Linker
ENLYFQG - TEV protease
GKPIPNPLLGLDST - V5 Tag
YPYDVPDYA - HA tag
DYKDDDDK - Flag tag
HHHHHHHH - His8

LEVLFQGP Prescission (3C) ProteaseP(PPX)
EQKLISEEDL - Myc
DYKDHDGDYKDHDIDYKDDDDK - 3xFlag
ATTTAAAT - SwaI site
GDGAGL - Flexible linker commonly used with fluorescent proteins.

[eGFP from LAP tag] encodes:
LKVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDH
MKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADK
QKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLKEFVTAAGITLGMDE
LYK*

Supplementary Table 5. Determination of ORF1:ORF2 ration by Intensity-based Absolute Quantification. Provided with description as a separate Microsoft Excel formatted file.

Supplementary Table 6. L1 retrotransposition defects when target genes are knocked down.

Target gene	**Validated? (By Sigma)**	**Relative L1 retrotranspostion [1]**	**Target gene knock down vs control[1]**	**shRNA clone ID**
PARP1	Yes	38% ± 6%	53% ±20%	TRCN0000007929
PCNA	Yes	91% ± 6%	82% ± 9%	TRCN0000003862
PCNA	No	57% ± 8%	60% ± 7%	TRCN0000003861
PCNA	No	58% ± 6%	47% ±10%	TRCN0000003863
PCNA	Yes	53% ± 6%	50% ±12%	TRCN0000003864
RUVBL1	No	16% ± 4%	45% ± 6%	TRCN0000018910
RUVBL1	Yes	80% ± 12%	77% ± 17%	TRCN0000018912
RUVBL1	Yes	42% ± 22%	70% ± 8%	TRCN0000018913
RUVBL1	Yes	56% ± 12%	35% ± 3%	TRCN0000018914
RUVBL2	Yes	7% ± 4%	2% ± 0.1%	TRCN0000051566
RUVBL2	Yes	4% ± 3%	3% ± 0.5%	TRCN0000051567
UPF2	No	26% ± 6%	75% ± 5%	TRCN0000157489
UPF2	Yes	51% ± 11%	61% ± 13%	TRCN0000157152
UPF3A	No	27% ± 1%	35% ± 8%	TRCN0000154972

[1]Anti-luciferase shRNA was used as control and assigned as 100%. All experiments were done in triplicate standard error is shown.

Supplementary Table 7. "Scorecard" comparing recovered interactors with host proteins identified in previous studies. Provided with description as a separate Microsoft Excel formatted file.

Abbreviations

pTRE, tetracycline regulated promoter

SV40, SV40 polyadenylation signal

EBNA-1, Epstein-Barr nuclear antigen 1

Amp^R, ampicillin resistance

pUC ori, pUC origin of replication

$puro^R$, puromycin resistance

S1 aptamer, streptavidin aptamers

MALDI-TOF PMF, Matrix-assisted laser desorption/ionization time of flight peptide mass fingerprinting

LC-MS/MS, liquid chromatography electrospray ionization tandem MS

Supplementary Methods

Antibodies

Monoclonal M2 anti-Flag antibody was purchased from Sigma-Aldrich (St. Louis, MO). Polyclonal rabbit anti-human UPF1 antibodies was a gift from Dr. Joan Steitz. IRDye® 800CW Goat anti-Mouse IgG (H+L) and IRDye® 680RD Goat anti-Rabbit IgG (H+L) were purchased from LI-COR Biosciences. Mouse monoclonal anti-human L1 ORF1 clone 4H1 was purchased from ABmart. Rabbit monoclonal anti-ORF1 was a gift from Dr. Jeff Han. Mouse IgG antibody was purchased from Santa Cruz Biotechnology, Inc (Dallas, TX). Alexa Fluor-conjugated goat anti-mouse and anti-rabbit IgG were purchased from Invitrogen.

Isotope labeling of cell culture

For isotope labeling experiments, Tet-free FBS was dialyzed to remove amino acids for 24 hrs against 10-volumes of PBS with 2 buffer changes using 3500MWCO dialysis tubing (Pierce, Rockford, IL). Lysine-arginine-free D-MEM medium (Invitrogen, Carlsbad, CA) was then supplemented with 10% of this dialyzed serum, 4mM L-glutamine, and heavy stable isotopes (heavy) as follows: 800µM $^{13}C_6{}^{15}N_4$-Arginine (Arg10) and 400µM $^{13}C_6{}^{15}$N2-Lysine (Lys8) (Cambridge Isotope Laboratories, Andover, MA) or their naturally occurring (light) alternatives (Sigma-Aldrich, St. Louis, MO). 200mg/L light proline was added (Lossner et al., 2011) and the medium was sterilized by filtration. Cultures were grown in labeled medium supplemented with 1µg/mL puromycin with for more than 10 divisions, expanding to thirty 15cm dishes. Cells were induced for 24hrs with 1µg/ml doxycycline, detached, washed in PBS, and extruded into liquid nitrogen. Heavy amino acid incorporation was confirmed as described in supplementary methods.

Cryogenic milling and affinity RNP purification

Cryogenic breakage of human cells, anti-Flag and anti-GFP affinity isolations, and subsequent SDS-PAGE and MALDI-MS analyses were achieved essentially as previously described (Domanski et al., 2012). The cryogenic milling assembly utilized external Teflon insulation that facilitated longer, uninterrupted milling intervals at cryogenic temperature, and ensured that liquid nitrogen remained present inside the milling chambers. See supplementary methods for more details.

Affinity RNP purification

Affinity isolation / MADLI-MS of ORF1-Flag utilized 250 mg LD288 cell powder and 50 µl anti-Flag M2 conjugated Dynabeads slurry (αFlag-DBs); and ORF2-3xFlag utilized 500 mg LD401 and 25 µl αFlag-DBs. Due to the extreme differences in expression levels between the two proteins, up to 20-fold more cell powder per µl DBs was used to purify ORF2p-3xFlag than ORF1p-Flag. In both cases an identical mock experiment utilizing the parental HEK293T cell line transfected with vector only (pCEP-Puro) was carried out in parallel. In all cases the extract was prepared 1:4 w:v in a 20 mM HEPES-Na buffered solution at pH 7.4 with 500 mM NaCl and 1% v/v Triton X-100 (supplemented with Roche Complete EDTA-free protease inhibitors) and batch binding was carried out 1 hr at 4°C. The post-batch binding, affinity isolated material was washed three times with 1 ml extraction buffer and then eluted by 30 min incubation of the beads at RT with 25 µl of a 1 mg/ml solution of 3xFlag peptide in the stated extraction buffer. Each eluate was combined with reducing LDS sample buffer and alkylated with iodoacetamide prior to running on a 4-12% Bis-Tris NuPAGE gel, followed by staining and select band excision. Affinity isolation of ORF1p when using Dynabeads conjugated with mouse monoclonal anti-ORF1 (clone 4H1) was performed analogously, using 50 µL Dynabeads to 200 mg cell powder and samples were eluted with LDS sample buffer at 70 °C for 5’ with mixing. Eluted fractions with combined with reducing agent and subjected to SDS-PAGE. When employed as the second step in tandem purifications, 10 µL (MT302) or 20µL (LD401) of anti-ORF1 DBs were used (see below).

Tandem affinity capture was implemented to selectively purify L1 particles containing both ORF proteins. First 3xFlag-tagged ORF2 was isolated from cell extracts

(MT302 and LD401) and eluted using 3xFlag peptide. ORF1 containing particles were subsequently captured from the eluted fraction using anti-ORF1 antibody conjugated Dynabeads and eluted using LDS (for SDS-PAGE) or 8M guanidinium-HCl (for MS). Relative levels of ORF1 and ORF2 proteins in polyacrylamide gels were assessed by Coomassie Blue and/or Sypro Ruby (Life Technologies, Grand Island, NY) staining using image densitometry as described in (LaCava et al., 2013); and were assessed in solution by MS using iBAQ (Schwanhaeusser et al., 2011; Smits et al., 2013). See supplementary methods for more details.

Affinity Capture for I-DIRT

Affinity capture preceding I-DIRT (Tackett et al., 2005) LC-MS/MS analyses were carried out using the following tagged cell lines: LD288 (ORF1-Flag, *ORFeus*-Hs), LD401 (ORF2-3xFlag, *ORFeus*-Hs), and MT302 (ORF2-3xFlag, L1RP); with their cognate untagged cell lines: LD259 (*ORFeus*, for LD288 and LD401) and LD289 (L1RP, for MT302). The tagged cell lines were grown heavy medium, whereas the untagged cell lines were grown in light medium (see Cell Lines). For MT293 (ORF1-myc, L1RP), a single round of purification with anti-myc DBs was not successful, and we instead used mouse-anti-human-ORF1p DBs and matched heavy-labeled MT293 with pCEP-Puro (empty vector) transfected cells. All cell types were cryogenically milled separately then appropriate pairs were mixed 1:1 by mass as a liquid N_2 slurry and thoroughly stirred (the N_2 allowed to dissipate overnight at -80°C) prior to dispensing cell powder for extraction and affinity capture. For LD288/LD259 – 100 mg of pre-mixed powder was extracted 1:4 w:v as described above using 50 µl αFlag-DBs; for LD401/LD259 – 400 mg mixed powder, 10 µl αFlag-DBs; for MT302/LD289 – 500 mg mixed powder, 10 µl αFlag-DBs;

for MT293/pCEP-Puro – 100 mg mixed powder, 20 µl αORF1-DBs (mouse anti-ORF1 clone 4H1). Binding time ranged from 5 min to 1 hr. Washing, elution and SDS-PAGE carried out as described above. Regions of each gel lane were excised and processed as described below.

Mass spectrometric and bioinformatic analysis for I-DIRT

Entire gel lanes were sliced at 1 mm intervals and slices were pooled into 5 or 6 samples by molecular weight. Proteins were subjected to in-gel trypsin digestion and analyzed by LC-MS/MS (see supplementary method provided for evaluating heavy isotope incorporation in cultured cells). Protein identification and isotopic ratio calculations were performed using MaxQuant (Cox and Mann, 2008). Raw data were searched against a database of human protein sequences, as well as a decoy database of reversed protein sequences, with Arg10 and Lys8 as potential heavy labels. Statistical analyses of the MaxQuant output table "proteinGroups.txt" were performed with the Perseus software (version 1.1.1.34) accompanying MaxQuant. The H/(H+L) values reported here are based on protein isotopic ratios that have been normalized to correct for unequal mixtures of heavy- and light-labeled cryogenic cell powders prior to affinity purification. The significance of outlier normalized protein ratios was calculated in Perseus by the method described in (Cox and Mann, 2008). Further details concerning the determination of high-confidence interactors are provided in supplementary methods.

Tagged ORF Construction

The ORF1p N-terminus was incompatible with tagging (pLD260, Figure 1B), and a 3xFlag tag at the ORF1p C-terminus (pLD415/6) severely interfered with L1 activity. In

contrast, a single Flag-tag at the ORF1p C-terminus (pLD261, pLD288) had no detectable effect on L1 retrotransposition. A C-terminal eGFP tag was relatively well tolerated, suggesting tag sequence matters more than tag size. We did not attempt to tag the N-terminus of ORF2p due to concerns regarding potential interference with its unconventional translation mechanism, which exhibits sensitivity to the length of the inter-ORF sequence (Alisch et al., 2006). The ORF2p C-terminus was relatively tolerant of both 3xFlag- and eGFP-tags, which affected overall L1 activity by <10% and <30%, respectively. A doubly tagged construct (pLD458, ORF1-eGFP; ORF2p-3xFlag) was also competent for retrotransposition (73% of WT). L1 activity quickly decreased when the length of the linker between ORF1p and eGFP increased (pLD459-61). The constructs pLD288 (ORF1p-Flag) and pLD401 (ORF2p-3xFlag) were chosen for further analysis because they exhibited the highest L1 activity with relatively short tag sequences in *ORFeus*-Hs.

After success in purification of *ORFeus*-Hs constructs, we moved to L1RP, first cloning the tag sequences from pLD288 (pMT292) and pLD401 (pMT290). We found that these tags were less well tolerated in L1RP. A shorter tag sequence (pMT302) fixed this problem for ORF2p. We sought to develop a short tag sequence on ORF1p orthogonal to Flag and selected the myc tag, in part because of availability the hybridoma for the 9E10 clone. While we found the myc tag was well tolerated on ORF1p, it did not provide high quality affinity isolations, perhaps due to phosphorylation of the myc epitope. Development of an improved mouse αORF1p antibody obviated the need for an epitope tagged ORF1p in L1RP and thus pMT302 (L1RP w/ ORF2p-3xFlag) was selected for further study.

Plasmids

The pCEP-puro plasmid was modified by replacing the CMV promoter with the TRE promoter, PCR amplified (JB14116, JB14135) using pTRE-Tight (Clontech, Mountain View, CA) as a template. L1 expression constructs without retrotransposition markers were all cloned into this pCEP-puro-Tet plasmid. L1 constructs with GFPAI markers were made by subcloning the L1 sequence into pLD189 (Dai et al., 2012). L1 constructs with BSDAI markers were made by subcloning the L1 sequence into pLD166, which contains a CMV promoter and BSDAI marker within the pCEP-puro vector. Tagged *ORFeus*-Hs L1 constructs were made by starting with pLD259 by using a combination of templateless, nested, and fusion PCR techniques with tag sequences built into synthetic oligos or Ultramer oligos (IDT, Coralville, Iowa), together with restriction digests at *Not* I, *Asc* I, *Bsr* GI, and *BstZ17* I restriction sites. The eGFP sequence from the LAP Tag system (a gift from Dr. Ina Poser) was used for GFP fusion proteins. L1RP in pCEP-Puro-Tet (pMT289) was made by first cutting the 5' 4.5kB from pWA172 (An et al., 2011) using *Not* I and *Bam* HI and then installing it in pCEP-Puro-Tet. The remaining 3' fragment was installed in the *Bam* HI site of this intermediate plasmid by creating a 5'*Bam* HI-*Bgl* II-3' terminated fragment, with or without tags, using a combination of templateless, nested, and fusion PCR techniques using pWA172 and tagged *ORFeus*-Hs constructs as templates and screening for orientation. All constructs were fully verified by DNA sequencing, and the tag sequences are shown in Table S4. pLD603 was made by cutting pLD288 with *Asc* I and *Bam* HI, blunting, and re-circularizing. pLD617 was made by replacing ORF2 sequence with L1 3' UTR sequence. pLD561 was made analogously from pLD401 using *Not* I and *Asc* I sites. pLD633 was made by removing 3' UTR

sequence from pLD401. pLD567 (ORF2-H230A modification of pLD401) and pLD618 – pLD624 were made from pLD401 using MISO Mutagenesis as reported in Figure 2B of (Mitchell et al., 2013). pDL124 was made by inserting a LIC cassette (Dai et al., 2012) into pLD258 between $Puro^R$ gene and pTRE promoter. Individual TRC shRNA clones were PCR and inserted into pDL124 by ligation independent cloning (Dai et al., 2012).

Cell culture

HeLa and HEK-293T cells were maintained in Dulbecco's Modified Eagle Medium (D-MEM, high glucose) supplemented with 10% fetal bovine serum (FBS), 4 mM L-glutamine and 100 units/ml penicillin/streptomycin (Invitrogen, Carlsbad, CA). The Tet-On $HEK293_{LD}$ cell line used was reported previously (Dai et al., 2012). Tet-On HeLa cells were from the Clontech (Mountain View, CA) and both adherent Tet-On lines were maintained as above but instead using 10% tetracycline-free FBS (Clontech, Mountain View, CA).

Tet-On $HEK293_{LD}$ cells were grown in suspension culture using Freestyle 293 medium (Gibco) supplemented with 2 mM L-glutamine and 1% tetracycline-free FBS in an atmosphere containing 5% or 8% CO_2 as described (Muller et al., 2005), with modification. Cells are maintained between $0.1\text{-}5\times10^6$/ml at 130 rpm in square or rectangular bottles from 60-1000mL made from glass (Corning, Tewksbury, MA) or polyethylene (Kartell, Noviglio, Italy) in a volume up to 30-40% of the bottle's indicated capacity. Adaptation to suspension was by transitioning cells from adherent growth media to 90% suspension media over two passages followed by selective passaging of non-clumping cells in suspension. Initially, each construct was transfected and puromycin selected (1µg/ml) in adherent culture for 7 days and then adapted to suspension. This

process was subsequently simplified with improved expression by transfection in suspension using Polyethylenimine "Max" high potency linear PEI (Polysciences) followed by 7-d puromycin selection (2.5µg/ml) and then subsequent reduction of puromycin concentration to 0.5µg/ml. Transfection mixtures were prepared with 1µg plasmid and 3µg PEI-Max per ml culture, mixed in 1/20th culture volume Hybridoma medium (Gibco) and incubated for 20 min before addition to cells. Cultures were grown to approximately 3-4×10^{6}/ml and induced with 1µg/ml doxycycline for 24 h and, after removal of 1ml analytical samples for blotting or microscopy, were centrifuged at 500×g for 10 min, transferred in PBS to a capped 30cc syringe with plunger removed (BD Biosciences, San Jose, CA) and centrifuged at 250×g for 10 min inside a 50mL conical tube. The supernatant was aspirated and then the cell pellet extruded into liquid nitrogen in 50mL conical tubes to form "BB's" (small spherical globules of frozen cells) or "noodles" (long thin tubules of frozen cells). (http://lab.rockefeller.edu/rout/assets/media/noodles.mp4). Typical yields were 10-15g wet cell weight (WCW) per liter of medium.

L1 retrotransposition assays

L1 construct retrotransposition assays were conducted in Tet-On HEK293T$_{LD}$ (eGFP-AI) cells if not specifically stated as described previously (Moran et al., 1995). Briefly, Tet-On HEK293T$_{LD}$ cells were seeded in 6-well plates (2×10^{5} cells/well) the day before transfection. The next day, each well was transfected with 1 µg plasmid using Fugene HD (Promega, Madison WI) according to the manufacturer's protocol. The day after transfection, cells were trypsinized and transferred to a 6 cm plate with 4 ml DMEM medium containing puromycin (1 µg/mL). After 3 days of puromycin selection, L1

retrotransposition frequency of constructs with the eGFP-AI reporter was evaluated by FACS using a Becton Dickinson LSR II sorter. L1 retrotransposition assays using plasmids with pTRE promoter were done in Tet-On HEK293T_{LD} cells (Dai et al., 2012). The day after transfection, puromycin (1 µg/mL) and doxycycline (500 ng/ml) were added. Five days after transfection, retrotransposition efficiency was measured by FACS. For plasmids containing the BSD-AI reporter, Hela cells were used. Puromycin resistant cells were trypsinized and seeded in 10 cm plates with Blasticidin (5 µg/mL). Two weeks later, $Blastcidin^R$ colonies were fixed and stained with crystal violet.

Cryogenic Milling

For small-scale preparations (1-3 g WCW) we used a 50 ml jar with 2 x 20 mm steel balls. The program included 3 x 3 m cycles at 400 RPM with a direction change every 30 s; between each 3 m cycle the jar assembly was re-cooled with liquid N_2. Large-scale samples (≳3 ~ 30g) were milled in a 125 ml jar using 5 x 20 mm steel balls with the same program as above. All resulting cell powders (powder) were transferred to N_2 cooled polypropylene tubes and stored at -80°C. Affinity isolations for mass spectrometry and western blot analyses utilized these powders.

MALDI/PMF

Gel bands were excised, cubed, destained with 50% v/v acetonitrile in 50 mM ammonium bicarbonate, dehydrated with acetonitrile, and digested overnight with trypsin (Promega). Tryptic peptides were extracted with POROS20 C18 beads (Krutchinsky et al., 2001) and desalted on Zip Tips (0.6 µL C18 resin, Millipore, Billerica MA), from which they were directly eluted onto a MALDI target in 3 µl of 25% saturated 4-HCCA

matrix solution in 70% acetonitrile, 0.1% TFA. Each spot was briefly washed with 10 µl cold 0.1% TFA. Data were collected on a prOTOF 2000 MALDI O-TOF mass spectrometer (Perkin Elmer, Waltham, MA). Monoisotopic centroid masses were calculated from these data using the software M/Z (Genomic Solutions, Bath, UK). Mass calibration was performed using the trypsin autolysis product at 2211 m/z. Mass lists were filtered in PeakErazor (Hjernø and Højrup, 2004) to remove common contaminants. The remaining masses were searched in ProFound (Zhang and Chait, 2000) with the following settings: taxonomy: H. sapiens; protein Mass: 0 – 600 kDa; modifications: $+C_2H_3ON$@C (Complete); +O@M (Partial); 10 ppm mass tolerance. All other settings were left as default. Protein identifications of expectation value lower than 0.05 were considered indicative of the presence of that protein in the excised band.

Sample preparation for relative quantification using iBAQ

Tandem-purified L1 particles intended for relative quantification using MS were eluted from anti-ORF1 beads using 8M guanidinium-HCl in 40 mM Tris-HCl, pH 8.0, at 70 °C for 5 min with mixing and then subsequently prepared for MS using a modified FASP procedure (Wiśniewski et al., 2009) without urea. For analysis of L1 particles derived from LD401, 0.5% w/v PEG was used to block filters (Wiśniewski et al., 2011). Once applied to the filters, samples were washed at least 3 times with 50 mM ammonium bicarbonate and then digested overnight using 500 ng of MS grade trypsin (Promega). Digests were collected by centrifugation and an additional 50 µl of 0.5 M NaCl was passed across the filter and combined with the tryptic digest.

The combined fraction was acidified, desalted on StageTips, and analyzed by LC-MS as described for "Evaluating heavy isotope incorporation." Using MaxQuant (version

1.2.2.5) (Cox and Mann, 2008), raw data were searched against a database of human protein sequences (ipi.HUMAN.v3.68.fasta), as well as a decoy database of reversed protein sequences, with iBAQ and "match between runs" selected. The "proteingroups.txt" output file was filtered to remove a) contaminants b) proteins for which the posterior error probability of the identification exceeded that of any protein identified from the decoy database of reversed protein sequences and c) proteins not identified in all three replicates. The relative abundance of LINE-1 proteins was determined by dividing the iBAQ intensity of ORF1 by that of ORF2, as in (Smits et al., 2013).

In vitro RT assay

The RT assays for ORF2 RT were performed in a 20 μl reaction mixtures containing 50 mM Tris-Cl (pH 8.0), 50 mM KCl, 5 mM $MgCl_2$, 10 mM DDT, 0.01 U template/primer, 1 μl [α-^{32}P] dTTP or [α-^{32}P] dCTP (final concentration 0.17 μM, 3000 Ci/mmol, 10 mCi/ml) at 37°C for 30 min. Then the mixture was spotted on DE81 paper and washed three times with 2×SSC buffer for a total time of 30 min. The DE81 paper was dried and counted by scintillation counter Beckman LS6000SC.

LEAP

L1 element amplification protocol (LEAP) was performed according to Kulpa *et al*. (2006). Cell powder from ~250 million Tet-on 293T_{LD} cells (250mg) was lysed with 3 ml buffer (1.5 mM KCl, 2.5 mM $MgCl_2$, 5 mM Tris-HCl pH 7.5, 1% deoxycholic acid, 1% Triton X-100, 1× protease inhibitor cocktail (Roche complete tablets EDTA-free)) for 5 min on ice. The lysate was clarified by centrifugation at 3,000 g for 5 min at 4°C, and the

supernatant was transferred to an 8.5% to 17% sucrose cushion. The gradient was spun at 39,000 rpm (SW40.1 rotor, 178,000 g) for 2 h at 4°C. The pellet was resuspended in 100 µL 5 mM Tris-HCl (pH 7.5) with 1× protease inhibitor, and 50% glycerol and stored at −80°C. Affinity purified RNP prepared from 250mg of the same powder was eluted with 3xFlag peptide in 50uL as below and also stored at −80°C with 50% glycerol, resulting in identical volume as the sucrose cushion preparation.

For the LEAP reaction, 2µL of RNP was added to 50 µL of 50 mM Tris-HCl pH 7.5, 50 mM KCl, 5 mM $MgCl_2$, 10 mM dithiothreitol, 0.4 µM 3′ LEAP primer (JB11560; Table S3), 20 units RNasin, 0.2 mM dNTPs and 0.05% Tween 20, and incubated at 37°C for 1 h. To make total L1 cDNA, 2 µL of pre-boiled RNP and 0.25 µL SuperScript® III Reverse Transcriptase were added to the same reaction mix and incubated at 50°C for 1 h. Reaction products were then amplified by PCR as previously reported or by quantitative real-time PCR (qPCR). For qPCR amplification, in triplicate 5µL of a 1:10 diluted LEAP product was assayed in 20 µL containing 0.25 µL each 10µM primers JB11564 and JB14067 (Table S3), and 10 µL 2x Fast SYBR® Green Master Mix (Applied Biosystems, Carlsbad, CA) using a Step One Plus instrument with relative cDNA levels calculated by the $2^{-\Delta\Delta_{CT}}$ method. Products from the qPCR run were compared to those from a parallel PCR reaction to verify that there were no amplification artifacts contributing to signal.

Real time RT-PCR

Total RNA was extracted using RNeasy mini kit (Qiagen, Valencia, CA) according to the manufacturer's protocol. One µg RNA was used for cDNA synthesis using

Superscript® III reverse transcription kit (Invitrogen, Carlsbad, CA) in a 20 µL reaction. Real time PCR was performed using a 1µL cDNA sample as template in a 20 µL reaction on a Step One Plus instrument (Applied Biosystems, Carlsbad, CA). Primers used for real time PCR were: Beta actin gene: JB12931, JB12932; *puro* gene: JB13421, JB13422; *ORFeus*-Hs L1 ORF1: JB13415, JB13416; JB15869, JB15970; JB15871 JB15872; *ORFeus*-Hs L1 ORF2: JB13417, JB13418; PCNA: JB16603, JB16604; UPF1: JB15433, JB15434; UPF2: JB16661, JB16662; UPF3A: JB16665, JB16666; PARP1: JB16607, JB16608; RUVBL1: JB16703, JB16704; RUVBL2: JB16671, JB16672; GFP: JB13766, JB13767. The gene knockdown efficiency was calculated by the $2^{-\Delta\Delta CT}$ method after normalization to the actin mRNA level.

Evaluating heavy isotope incorporation

To measure the degree of heavy amino acid incorporation in the proteomes of cells cultured in medium containing Lys8 and Arg10, cleared cell lysates of each strain (LD288, LD401, MT302, and MT293) were separated by SDS-PAGE. A ~42 kDa band was excised from each of the four gel lanes. Bands were subjected to in-gel digestion with trypsin, and peptides were extracted with POROS beads (Life Technologies, Grand Island, NY) and desalted on StageTips (Rappsilber et al., 2007). Peptides eluted from StageTips were concentrated by SpeedVac and pressure-loaded onto a PicoFrit column (New Objective, Woburn, MA) self-packed with 6 cm of reverse-phase C18 material (YMC*Gel ODS-A, YMC, Allentown, PA) and gradient eluted (Solvent A = 0.1 M acetic acid, Solvent B = 70% acetonitrile in 0.1M acetic acid, flow rate 200 nL/min) into an LTQ-Orbitrap-Velos mass spectrometer (Thermo Fisher Scientific, Waltham, MA). The instrument was operated in data-dependent mode, such that the 20 most abundant

ions in each full scan were sequentially fragmented by CID. Ions were set on a dynamic exclusion list after selection for MS/MS to avoid repeated sequencing of the same species.

Using MaxQuant (version 1.2.2.5) (Cox and Mann, 2008), raw data were searched against a database of human protein sequences (ipi.HUMAN.v3.68.fasta), as well as a decoy database of reversed protein sequences, with Arg10 and Lys8 as potential heavy labels. Default parameters were left in place, except the "re-quantify" option was not utilized. More than 300 proteins were identified. The "peptide.txt" output file was used to calculate heavy amino acid incorporation (Geiger et al., 2011). Peptides corresponding to known exogenous contaminants such as keratin and trypsin were first removed. Lysine- and arginine-containing peptides were considered separately. The "ratio H/L" for each peptide was converted to H/(H+L) using the equation H/(H+L) = ("ratio H/L") / (1+ "ratio H/L"). The median values of H/(H+L) for lysine- and arginine-containing peptides were 0.92 and 0.98, respectively, for each of the four cell lines. These values underestimate incorporation, as a peptide "ratio H/L" cannot be calculated by MaxQuant when there is no detectable signal at the mass-to-charge ratio of the light partner, as would be expected in the case of full incorporation of heavy amino acid.

Determination of high-confidence interactors from I-DIRT experiments

The MaxQuant output file "proteingroups.txt" was filtered to remove a) known exogenous contaminants b) proteins for which the posterior error probability of the identification exceeded that of any protein identified from the decoy database of reversed protein sequences and c) proteins with fewer than 3 ratio counts. Ratio counts indicate the number of independent pieces of information (different peptide sequences, or different

charge states or chemical modifications of the same peptide sequence) used to calculate the heavy isotope content of each protein. Significance testing in Perseus was carried out on log-transformed normalized H/L ratios. The p-values reported are designated as "significance A" in (Cox and Mann, 2008) and are determined by a one-sided test. Using a Benjamini-Hochberg correction for multiple hypothesis testing, the list of significant hits was truncated with a cutoff false discovery rate value of 0.0875.

PCNA Immunoprecipitation assay from fresh cell lysates

Immunoprecipitation (IP) assays to analyze PCNA binding to wild type and mutated ORF2p have been performed using Tet-On $HEK293_{LD}$ (Dai et al., 2012) cells freshly lysed in Triton-buffer (50mM Hepes pH7.5, 150mM NaCl, 1mM EDTA, 1mM EGTA, 10% glycerol, 1% Triton X-100, 25mM NaF 10mM $ZnCl_2$ freshly supplemented with protease and phosphatase inhibitors (Roche, Indianapolis, IN). Upon lysis, cell lysates are normalized for their protein content by Bradford assay and about 1 mg of proteins are used to immunoprecipitate ORF2 using Flag-M2 conjugated dynabeads. The same amount of proteins was used for the normal mouse IgG antibody control. For the latter 10 μg of normal mouse IgG and 20 μl of protein G dynabeads were incubated with the cell lysate and the samples have been process in parallel with the Flag IP samples. Incubation of the lysate with the antibodies was performed for 1 h at 4°C under rotation. After incubation beads have been washed 5 times with 0.8ml of Extraction buffer (20mM Hepes, 1% Triton X-100, 300mM NaCl) and then resuspended in LDS 1X sample buffer (Invitrogen, Carlsbad, CA). Immune complexes were eluted from beads by shaking the beads in sample buffer for 10 min at room temperature. Supernatant was then collected and 5% β-mercaptoethanol added. Samples were then heated for 10 min at 70°C and then

stored at -20°C. Samples were run on NuPage 4-20% bis-tris precast gels (Invitrogen, Carlsbad, CA) and transferred on PVDF-F membrane then processed for Western blotting and ODYSSEY® CLx (LI-COR, Lincoln, NE) infrared imaging.

Immunoblot assays

The anti-ORF1 antibody used for all immunoblotting in this paper is mouse monoclonal clone 4H1. For measurement of protein expression in small-scale expression experiments (when cryomilling was not feasible), transfected Tet-On HEK293T_{LD} cells were lysed in RIPA buffer (150mM NaCl, 50mM Tris pH 8.0, 1% NP-40, 0.5% Na-Deoxycholate, 0.1% SDS) supplemented with protease inhibitors and Benzonase nuclease and spun for 15 minutes at 20,000 × g to clarify the cell lysate. Clarified cell lysate (2-25 µg) was supplemented with LDS sample buffer (Life Technologies) and 1.25% (v/v) β-mercaptoethanol and then heated for 10 min at 70°C before separation on 4-12% Bis-Tris SDS-PAGE gels . After transferring to PVDF membranes, membranes were probed with primary antibody in 5% blocking reagent. Western blots were developed with Luminata Forte Western HRP substrate (Millipore, Billerica MA), detected using a Fuji LAS4000 (GE HealthCare, Pittsburgh PA) instrument or the LI-COR imaging system. Only the non-saturated signals were quantified, and the background was subtracted. Results were normalized by using the tubulin controls as a reference, and are presented as fold difference relative to control. For the infrared imaging system, membranes were incubated in 1:1 LI-COR Odyssey™ Blocking buffer and TBS (10 mM Tris-HCl, pH 7.4, 0.5 M NaCl) for 1 hour. Primary antibodies were then probed in 1:1 LI-COR Odyssey™ Blocking buffer and TBS plus Tween-20 (0.2% v/v) within cold room over night. After washing 3 x 10min with TBS-Tween, the

membranes were probed with secondary antibodies in 1:1 LI-COR Odyssey Blocking buffer and TBS-Tween containing 0.01% SDS for 1.5h. Membranes were washed 5 x 6min in TBS-Tween, and then scanned with the infrared imaging system.

Indirect Immunofluorescence Microscopy

About 200,000 puromycin-selected or untransfected Tet-On HEK293T_{LD} or HeLa cells were plated on glass coverslips and induced or transfected 8-16 hours after plating. For Tet-On HEK293T_{LD} cells, coverslips were pre-coated with 10 µg/ml fibronectin in PBS (Life Technologies) for 1-2 hours. 24 hours later, cells were fixed in 3% paraformaldehyde for 10 min. Fixative was then quenched using PBS containing 10mM glycine and 0.2% Na-Azide (PBS/gly). The cells were permeabilized for 3 min in 0.5% Triton X-100 and washed twice with PBS/gly. Staining with primary and secondary antibodies was done for 20 minutes at room temperature by inverting coverslips onto Parafilm containing 45µl drops of PBS/gly supplemented with 1% BSA and appropriate antibodies. Antibodies used were Mouse anti-Flag M2 (1:500), Rabbit anti-ORF1 JH73 (1:4000), Alexa Fluor® 488 conjugated anti-mouse IgG (1:1000), and Alexa Fluor 568 conjugated anti-rabbit IgG (1:1000). DNA was stained prior to imaging with Hoechst 33285 (0.1 µg/ml). Epiflourescent images were collected using an Axioscop microscope (Zeiss, Jena, Germany) equipped for epifluorescence using an ORCA-03G CCD camera (Hamamatsu, Japan). Confocal images were acquired with a Cascade QuantEM 512SC camera (Photometrics) attached to a Zeiss AxioImager with Yokogawa spinning disk confocal scanner and Slidebook software (Intelligent Imaging Innovations).

Supplementary references

Alisch, R.S., Garcia-Perez, J.L., Muotri, A.R., Gage, F.H., and Moran, J.V. (2006). Unconventional translation of mammalian LINE-1 retrotransposons. Genes & development *20*, 210-224.

An, W., Dai, L., Niewiadomska, A.M., Yetil, A., O'Donnell, K.A., Han, J.S. and Boeke, J.D. (2011) Characterization of a synthetic human LINE-1 retrotransposon ORFeus-Hs. Mobile DNA 2, 2.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature biotechnology *26*, 1367-1372.

Dai, L., Taylor, M.S., O'Donnell, K.A., and Boeke, J.D. (2012). Poly(A) binding protein C1 is essential for efficient L1 retrotransposition and affects L1 RNP formation. Molecular and cellular biology.

Domanski, M., Molloy, K., Jiang, H., Chait, B.T., Rout, M.P., Jensen, T.H., and LaCava, J. (2012). Improved methodology for the affinity isolation of human protein complexes expressed at near endogenous levels. BioTechniques 0, 1-6.

Geiger, T., Wisniewski, J.R., Cox, J., Zanivan, S., Kruger, M., Ishihama, Y., and Mann, M. (2011). Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. Nature protocols *6*, 147-157.

Hjernø, K., and Højrup, P. (2004). PeakErazor: A Windows-based program for improving peptide mass searches. In Methods in Proteome and Protein Analysis, R.M. Kamp, J.J. Calvete, and T. Choli-Papadopoulou, eds. (New York: Springer Verlag), pp. 359-370.

Krutchinsky, A.N., Kalkum, M., and Chait, B.T. (2001). Automatic identification of proteins with a MALDI-quadrupole ion trap mass spectrometer. Analytical chemistry *73*, 5066-5077.

Kulpa, D.A., and Moran, J.V. (2006). Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. Nature structural & molecular biology *13*, 655-660.

LaCava, J., Chandramouli, N., Jiang, H., and Rout, M.P. (2013). Improved native isolation of endogenous protein a-tagged protein complexes. BioTechniques 54, 213-216.

Lossner, C., Warnken, U., Pscherer, A., and Schnolzer, M. (2011). Preventing arginine-to-proline conversion in a cell-line-independent manner during cell cultivation under stable isotope labeling by amino acids in cell culture (SILAC) conditions. Analytical biochemistry 412, 123-125.

Mitchell, L.A., Cai, Y., Taylor, M.S., Noronha, A.M., Chuang, J., Dai, L., and Boeke, J.D. (2013). Multichange Isothermal Mutagenesis: a new strategy for multiple site-directed mutations in plasmid DNA. ACS Synth Biol 2, 473-477.

Muller, N., Girard, P., Hacker, D.L., Jordan, M., and Wurm, F.M. (2005). Orbital shaker technology for the cultivation of mammalian cells in suspension. Biotechnology and bioengineering *89*, 400-406.

Schwanhaeusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature 473, 337-342.
Smits, A.H., Jansen, P.W.T.C., Poser, I., Hyman, A.A., and Vermeulen, M. (2013). Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. Nucleic Acids Research *41*, e28.
Tackett, A.J., DeGrasse, J.A., Sekedat, M.D., Oeffinger, M., Rout, M.P., and Chait, B.T. (2005). I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. Journal of proteome research 4, 1752-1756.
Wiśniewski, J.R., Ostasiewicz, P., and Mann, M. (2011). High recovery FASP applied to the proteomic analysis of microdissected formalin fixed paraffin embedded cancer tissues retrieves known colon cancer markers. Journal of proteome research *10*, 3040-3049.
Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. Nature methods *6*, 359-362.
Zhang, W., and Chait, B.T. (2000). ProFound: an expert system for protein identification using mass spectrometric peptide mapping information. Analytical chemistry *72*, 2482-2489.