

Characterization of LINE-1 Ribonucleoprotein Particles

Aurélien J. Doucet^{1,2*}, Amy E. Hulme², Elodie Sahinovic¹, Deanna A. Kulpa², John B. Moldovan³, Huira C. Kopera², Jyoti N. Athanikar², Manel Hasnaoui¹, Alain Bucheton¹, John V. Moran^{2,3,4,5*}, Nicolas Gilbert^{1*}

1 Institut de Génétique Humaine, CNRS, UPR 1142, Montpellier, France, **2** Department of Human Genetics, University of Michigan Medical School, Ann Arbor, Michigan, United States of America, **3** Cellular and Molecular Biology Program, University of Michigan Medical School, Ann Arbor, Michigan, United States of America, **4** Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan, United States of America, **5** Howard Hughes Medical Institute, Chevy Chase, Maryland, United States of America

Abstract

The average human genome contains a small cohort of active L1 retrotransposons that encode two proteins (ORF1p and ORF2p) required for their mobility (*i.e.*, retrotransposition). Prior studies demonstrated that human ORF1p, L1 RNA, and an ORF2p-encoded reverse transcriptase activity are present in ribonucleoprotein (RNP) complexes. However, the inability to physically detect ORF2p from engineered human L1 constructs has remained a technical challenge in the field. Here, we have employed an epitope/RNA tagging strategy with engineered human L1 retrotransposons to identify ORF1p, ORF2p, and L1 RNA in a RNP complex. We next used this system to assess how mutations in ORF1p and/or ORF2p impact RNP formation. Importantly, we demonstrate that mutations in the coiled-coil domain and RNA recognition motif of ORF1p, as well as the cysteine-rich domain of ORF2p, reduce the levels of ORF1p and/or ORF2p in L1 RNPs. Finally, we used this tagging strategy to localize the L1-encoded proteins and L1 RNA to cytoplasmic foci that often were associated with stress granules. Thus, we conclude that a precise interplay among ORF1p, ORF2p, and L1 RNA is critical for L1 RNP assembly, function, and L1 retrotransposition.

Citation: Doucet AJ, Hulme AE, Sahinovic E, Kulpa DA, Moldovan JB, et al. (2010) Characterization of LINE-1 Ribonucleoprotein Particles. *PLoS Genet* 6(10): e1001150. doi:10.1371/journal.pgen.1001150

Editor: Gregory S. Barsh, Stanford University, United States of America

Received: February 10, 2010; **Accepted:** September 3, 2010; **Published:** October 7, 2010

Copyright: © 2010 Doucet et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: AJD was the recipient of fellowships from the French government (Ministère de l'Enseignement Supérieur et de la Recherche), from the Association pour la Recherche contre le Cancer (ARC), and from the Fondation Recherche Médicale (FRM). AEH and DAK were supported in part by a Michigan Predoctoral Training Grant from the National Institutes of Health (5T32GM07544). AEH also was supported by a Rackham Predoctoral Fellowship. JNA was supported in part by a NRSA postdoctoral fellowship (F32GM20859). HCK was supported in part by an American Cancer Society postdoctoral fellowship (PF-07-059-01-GMC). Work in the laboratory of NG is supported by the Centre National de Recherche Scientifique and the Agence Nationale de la Recherche (ANR-05-JCJC-0120). Work in the laboratory of JVM is supported by an NIH grant (GM060518). The University of Michigan Cancer Center Support Grant (5P30CA46592) helped defray some of the sequencing costs incurred during this study. JVM is an investigator of the Howard Hughes Medical Institute. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: John V. Moran is listed as an inventor of the following patent: "Compositions and methods of use of human retrotransposons: Application No. 60/006,831, Issued November, 2000." Moran has received no money from the patent and the patent does not influence any of the results/interpretations in the paper. This should not represent a conflict of interest, but is being voluntarily disclosed.

* E-mail: Nicolas.Gilbert@igh.cnrs.fr (NG); moranj@umich.edu (JVM); adoucet@umich.edu (AJD)

Introduction

Long Interspersed Element-1 (LINE-1 or L1) sequences comprise 17% of human DNA and represent the predominant class of autonomous retrotransposon-derived sequences in the genome [1]. Greater than 99.9% of L1 elements are molecular fossils that are no longer capable of mobilization (*i.e.*, retrotransposition) [1–3]. However, the average human genome still harbors a small cohort (approximately 80–100) of retrotransposition-competent L1s (RC-L1s) [4,5]. A wealth of experimental evidence suggests that ongoing RC-L1 retrotransposition has the potential to impact the genome by a myriad of mechanisms (reviewed in [6–8]).

A human RC-L1 is approximately 6 kb in length; it begins with a ~910 bp 5' untranslated region (UTR) that harbors an internal RNA polymerase II promoter [9–11], two non-overlapping open reading frames (ORF1 and ORF2), and ends with a 3' UTR that is followed by either a polyadenylic acid (poly A) or A-rich sequence (Figure 1A) [12,13]. Genetic and biochemical evidence suggest that the ORF1 and ORF2-encoded proteins (ORF1p and ORF2p, respectively) preferentially associate with their encoding mRNA *in cis* to form a ribonucleoprotein particle (RNP) that

probably is an intermediate in the retrotransposition process [14–19]. The resultant RNP then gains access to the nucleus, where L1 integration presumably occurs by target-site primed reverse transcription (TPRT) [20–23].

Studies conducted with mouse and human RC-L1s have uncovered a number of conserved domains within ORF1p that are important for retrotransposition. The amino acid sequence of the ORF1p amino-terminus is poorly conserved among mammalian L1s, but it is predicted to form a coiled-coil or α -helical domain that is important for ORF1p multimerization [15,24–27]. In human ORF1p, this region contains a putative leucine zipper (LZ) domain that is absent from other mammalian L1s, although a similar motif is present in the L1-like Swimmer element of teleosts [15,24,27–29]. The coiled-coil domain of ORF1p is followed by a RNA recognition motif (RRM) [30], and experiments in cultured human cells have shown that mutations in conserved residues of the RRM domain (*e.g.*, a N₁₅₇A/R₁₅₉A double mutant) adversely affect L1 retrotransposition and the formation of cytoplasmic structures known as ORF1 cytoplasmic foci [31].

The carboxyl-terminus of ORF1p contains amino acid residues that are conserved among mammalian L1s [24,27,32]. Biochem-

Author Summary

Long Interspersed Element-1 (LINE-1 or L1) sequences are the predominant class of autonomous retrotransposons in the human genome and comprise an astounding 17% of human DNA. Although the majority of L1s are considered to be “dead,” an average human genome contains ~80–100 active L1s. Active L1s encode two proteins (ORF1p and ORF2p) that are required for mobility (retrotransposition) by a “copy and paste” mechanism termed target-site primed reverse transcription. Prior experiments suggested that ORF1p, ORF2p reverse transcriptase activity, and L1 mRNA associate in ribonucleoprotein (RNP) particles and that RNP formation is a necessary step in L1 retrotransposition. However, the difficulty in detecting ORF2p from engineered human L1s has prevented a thorough understanding of its role in L1 retrotransposition. Here, we have exploited epitope and/or RNA-tagging strategies to detect and characterize a “basal” RNP complex from engineered human L1s. We also expanded on previous studies and characterized how mutations in conserved functional domains of ORF1p and ORF2p can adversely affect L1 RNP formation/function. Finally, our strategy allowed us to detect the L1-encoded proteins and L1 RNA in cytoplasmic foci. Thus, we have developed and employed a system to gain greater understanding of LINE-1 retrotransposition at the molecular level.

ical analyses have shown that mouse ORF1p homotrimers bind L1 RNA in a sequence independent manner [33,34]. Mutations of a conserved di-arginine motif (RR_{261–262} in human L1) can decrease ORF1p RNA binding or mouse ORF1p nucleic acid chaperone activity [33,35]. Similarly, studies using human L1s revealed that alanine mutations in conserved amino acid residues in the carboxyl terminus of ORF1p (RR_{261–262}, and YPAKLS_{282–287}, respectively) both compromise the ability of ORF1p to localize to RNPs and severely reduce L1 retrotransposition efficiency [16,32]. Thus, ORF1p is postulated to have critical functions at discrete steps in the retrotransposition pathway.

Biochemical and genetic studies have revealed that human and mouse ORF2 are translated by an unconventional mechanism [36–39]. It is hypothesized that as few as one or two molecules of ORF2p are translated per L1 RNA molecule, which could explain why it has been difficult to detect ORF2p produced from engineered L1s in cultured cells [37]. ORF2p contains endonuclease (EN) and reverse transcriptase (RT) activities that are critical for the target-site cleavage and reverse transcription steps of TPRT [22,23,32,40]. ORF2p also contains a conserved cysteine-rich (C) domain near its carboxyl-terminus [27,41]. Mutations in the C-domain adversely affect L1 retrotransposition [32]; however, the biochemical role of the C-domain in L1 retrotransposition remains poorly understood.

Epitope-tagging systems and enzymatic assays have been developed to facilitate detection of L1 ORF1p and ORF2p RT activity from engineered wild-type and mutant human L1s [16,17]. However, the inability to reliably and directly detect ORF2p from engineered human L1s in transfected cultured human cells has hindered progress in the field [37,42]. Here, we have devised an epitope and/or RNA-tagging system to show that ORF1p, ORF2p, and L1 RNA form a ribonucleoprotein complex, which may represent a minimal RNP retrotransposition intermediate. Consistent with previous studies, transient transfection/immunofluorescence-based experiments revealed that the L1-encoded proteins and L1 mRNA often form discrete cytoplasmic foci, and that many of these foci associate with stress granules [31]. Finally, we have

extended previous analyses [16,17] and demonstrate that mutations in conserved functional domains of ORF1p and/or ORF2p adversely affect L1 RNP formation, the reverse transcription of L1 RNA, and L1 cytoplasmic foci formation. Thus, we have developed a system that should allow a greater understanding of the L1 retrotransposition mechanism at the molecular level.

Results

A system to detect L1 ORF2p in cultured cells

Previous studies have examined the co-localization of L1 ORF1p and L1 RNA in RNPs derived from cells transfected with epitope-tagged wild-type or mutant human L1 expression constructs [16,17]. To physically detect L1 ORF2p, we modified existing L1 expression vectors (pJM101/L1.3 or pDK101) to contain either a 530 bp TAP tag or a 72 bp FLAG-HA tag on the carboxyl-terminus of ORF2p (Figure 1A; pAD2TE1 and pES2TE1) [43,44]. To facilitate the identification of L1 RNA, we also introduced a 1312 bp DNA fragment that contains 24 copies of a stem-loop sequence that can bind the phage MS2 protein into the L1 3'UTR (Figure 1A; pAD3TE1) [45,46]. As a control, we generated a plasmid that expresses TAP-tagged ORF2p from a monocistronic transcript (Figure 1A; pAD500).

L1 constructs were equipped with a retrotransposition indicator cassette (*mneoI*), subcloned into a pCEP4 episomal expression vector, and were assayed for retrotransposition in cultured human HeLa cells [32,47,48]. Inclusion of either the TAP or FLAG-HA epitope tag onto the carboxyl-terminus of ORF2p had little effect on the L1 retrotransposition efficiency when compared to a wild-type control construct lacking the tag (Figure 1B; pADO2Tt, pAD2TE1, and pES2TE1 *vs.* pJM101/L1.3). Similarly, the inclusion of the MS2 stem loop sequences into the L1 3'UTR did not dramatically affect L1 retrotransposition efficiency (Figure 1B; pADL1MT *vs.* pJM101/L1.3), although we did observe an approximate 2.7 fold reduction in L1 retrotransposition efficiency from a construct containing both the protein and MS2 tags (Figure 1B; pAD3TE1 *vs.* pJM101/L1.3). As a negative control, we demonstrated that a construct containing a missense mutation in the putative L1 RT active site (pAD135; D₇₀₂A) was defective for retrotransposition (Figure 1B). Thus, engineering epitope and/or RNA tags into the L1 expression vectors is compatible with retrotransposition in cultured cells.

Physical detection of ORF2p in HeLa cells

To detect the L1-encoded proteins from the engineered plasmids, we transfected each construct into HeLa cells and selected for cells containing the respective L1 expression vectors by exploiting the hygromycin B selectable marker on the pCEP4 episome (Figure 1A; see Materials and Methods). Consistent with previous studies [16,17,49], western blot analyses of whole cell lysates using antibodies directed against the ORF1p T7-epitope tag revealed the presence of a ~40 kDa protein from constructs containing the tag (Figure 2A, middle panel (α T7); pAD2TE1, pAD3TE1, and pES2TE1), but not from controls lacking the tag (Figure 2A; pJM101/L1.3 and pADO2Tt). We also could detect the ~40 kDa protein with polyclonal antibodies against endogenous human ORF1p (Figure 2B, α ORF1 panels; pAD2TE1, pJM101/L1.3, and pDK101) [50]. Notably, we observed a slight difference in the mobility of T7-tagged and untagged ORF1p (Figure 2B; right panel (α ORF1), pJM101/L1.3 *vs.* pDK101), which most likely is due to the additional amino acids imparted by the T7 epitope tag. Controls revealed that ORF1p was not detected from a construct that lacks ORF1 (Figure 2A and 2B; pAD500) or from a construct that contains a premature stop

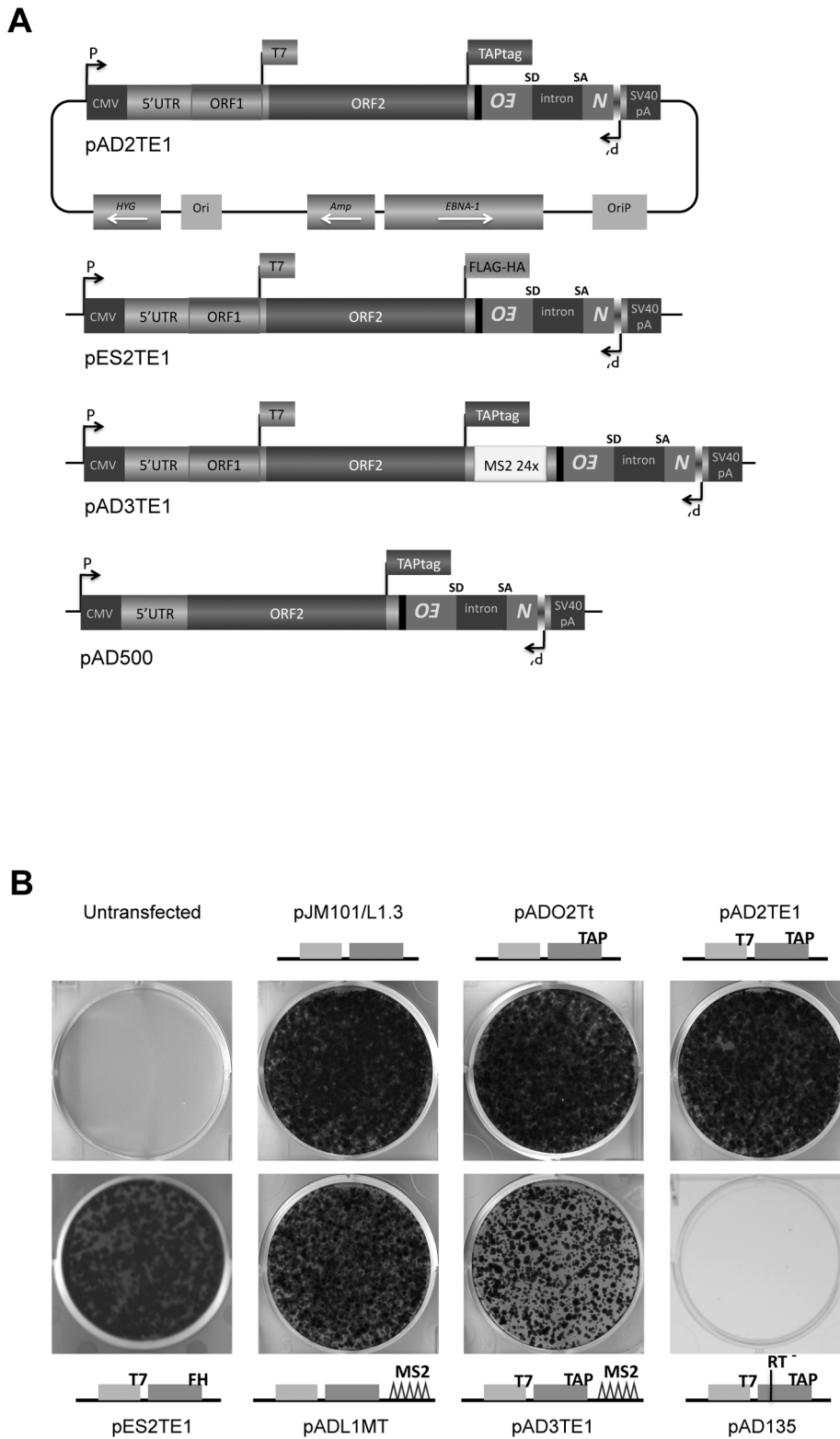


Figure 1. The retrotransposition efficiency of engineered L1s used in this study. *A.* A diagram of L1 plasmids used in this study: Each plasmid is a derivative of pJM101/L1.3 or pDK101 [4,16]. The constructs were tagged with the *mneoI* retrotransposition indicator cassette [32,47], and are expressed from the pCEP4 episomal vector (Invitrogen). Labeled rectangles indicate the relative positions of the L1 5'UTR, ORF1p and ORF2p. Labeled flags at the 3' ends of ORF1 and/or ORF2 are used to denote the epitope tag in the respective constructs. pAD3TE1 also contains 24 copies of a stem-loop sequence that can bind the phage MS2 protein (light rectangle labeled MS2 24x) [65]. pAD500 is a monocistronic ORF2p expression vector that lacks ORF1 as well as the inter-ORF spacer sequence. *B.* Representative results of the retrotransposition assay: L1 retrotransposition efficiency was assayed as described previously [32,48]. HeLa cells transfected with pJM101/L1.3 serve as a positive control. Untransfected HeLa cells and HeLa cells transfected with an RT mutant (pAD135; D702A) serve as negative controls. Cartoons of constructs used in the experiment are indicated in the figure. All constructs contain the *mneoI* retrotransposition indicator cassette. doi:10.1371/journal.pgen.1001150.g001

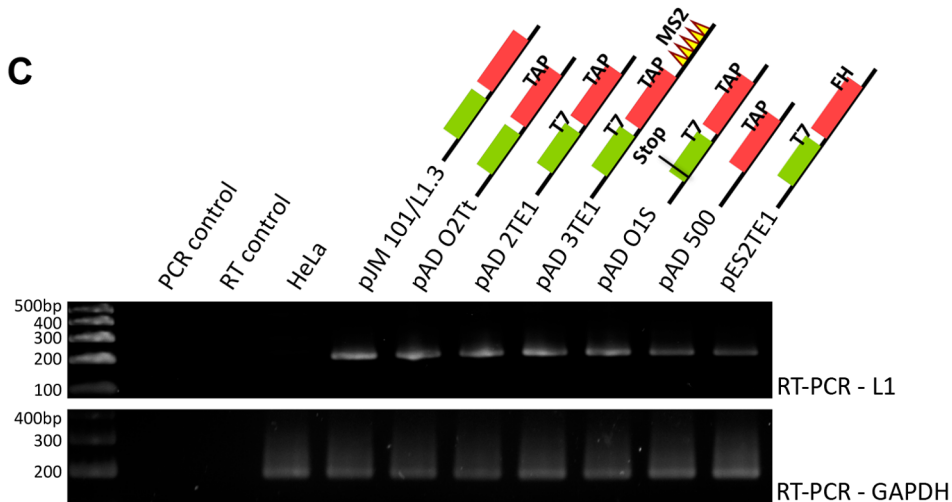
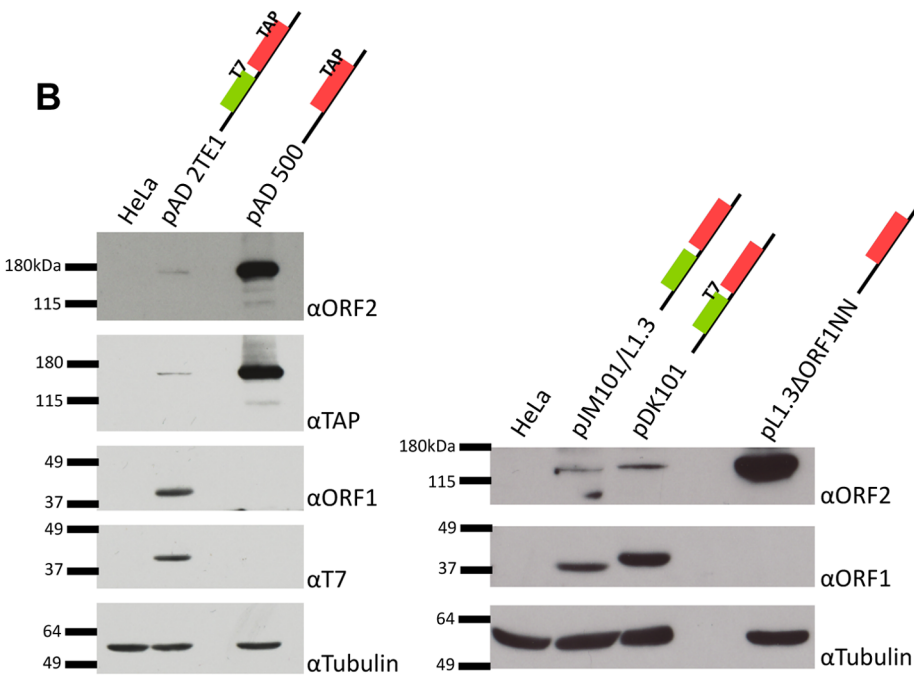
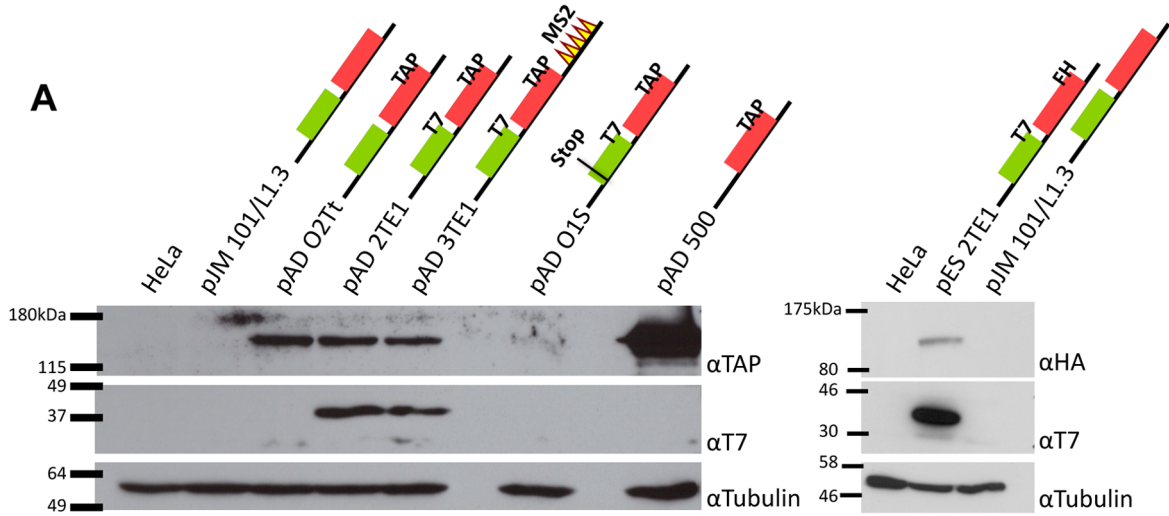


Figure 2. Detection of ORF1p and ORF2p from engineered L1 constructs. *A. Representative results from western blot analyses:* Whole cell lysates derived from untransfected HeLa cells or HeLa cells transfected with the indicated L1 expression constructs were subjected to western blot analyses. Top panels: western blots conducted with anti-TAP antibodies (α TAP; left side) or anti-HA antibodies (α HA; right side) to detect epitope-tagged ORF2p. Middle panels: western blots conducted with anti-T7 antibodies (α T7) to detect epitope-tagged ORF1p. Lower panels: western blots conducted with anti-tubulin antibodies (α Tubulin) served as a loading control. Molecular weight standards (Invitrogen, left side, and New England Biolabs, right side) are listed at the left of each series of gels. *B. Protein detection specificity:* Whole cell lysates derived from HeLa cells transfected with the indicated constructs were subjected to western blot analyses using antibodies against either endogenous ORF1p (α ORF1) or endogenous ORF2p (α ORF2). TAP-tagged ORF2p also was detected using an anti-TAP antibody. T7-tagged ORF1p also was detected with an anti-T7 antibody. Tubulin was detected using an anti-tubulin (α Tubulin) antibody and served as a loading control. Molecular weight standards (Invitrogen) are listed at the left of each series of gels. *C. RT-PCR analyses:* RT-PCR reactions using RNAs isolated from whole cell lysates derived from transfected cells revealed that L1 RNA was expressed from each of the constructs. GAPDH mRNA detection was used to assess the quality of the RNA preparations and as a loading control. Reactions without template (PCR control) or reverse transcriptase (RT control) were used as negative controls. DNA size markers (Invitrogen) are indicated at the left of the gel. Colored cartoons of the constructs used in the experiments are indicated next to their respective names. The black lines indicate the 5' and 3' UTRs. The green and red boxes indicate ORF1 and ORF2p respectively. When present, epitope tags are indicated. All constructs contain the *mneol* retrotransposition indicator cassette. doi:10.1371/journal.pgen.1001150.g002

codon in ORF1 (Figure 2A; pADO1S). Qualitative reverse transcriptase-PCR (RT-PCR) experiments further confirmed that L1 RNA was expressed from each of the transfected constructs (Figure 2C; see Oligonucleotides and RT-PCR sections in Materials and Methods for details).

To detect ORF2p expression, we conducted western blot analyses on whole cell lysates derived from transfected cells using antibodies directed against either the TAP or FLAG-HA epitope tag (Figure 2A). A \sim 170 kDa protein was detected from L1 constructs containing TAP-tagged ORF2p, but not from an untagged wild-type control (Figure 2A; left panel (α TAP); pADO2Tt, pAD2TE1, pAD3TE1, and pAD500 *vs.* pJM101/L1.3). The \sim 170 kDa product corresponds to the predicted size of ORF2p (\sim 150 kDa) plus the predicted size of the TAP tag (\sim 19 kDa) [12,43]. A \sim 170 kDa protein also was detected using antibodies against endogenous ORF2p (Figure 2B; left panel (α ORF2)) [42]. Similarly, a \sim 155 kDa protein was detected from L1 constructs containing FLAG-HA-tagged ORF2p, but not from the untagged wild-type control (Figure 2A; right panel (α HA); pES2TE1 *vs.* pJM101/L1.3). Consistent with previous genetic studies, TAP-tagged ORF2p expression was greatly diminished by introducing a stop codon in ORF1 (Figure 2A; pADO1S) and was most abundant when expressed from an ORF2p monocistronic expression vector (Figure 2A and 2B; pAD500) [37].

Epitope-tagged ORF2p localizes to ribonucleoprotein particles

To test whether ORF2p localizes to ribonucleoprotein particles (RNPs), we transfected HeLa cells with pAD2TE1, selected for transfected cells, and isolated RNPs by ultracentrifugation (see Materials and Methods) [16,17]. Western blotting revealed that ORF1p and ORF2p were readily detected in the RNP fraction (Figure 3A, top panel).

We next used the L1 Element Amplification Protocol (LEAP) assay to determine whether the RNP preparations contained an L1-specific reverse transcriptase activity [17]. Consistent with previous studies, a diffuse set of LEAP products that ranged in size from \sim 220 to \sim 400 bp was detected in pAD2TE1-derived RNPs, but not from pAD135-derived (D_{702A}; RT mutant) RNPs (Figure 3A, lower panel). Cloning and sequencing of the pAD2TE1-derived LEAP products confirmed that L1 reverse transcription generally initiated at variable sites within the L1 poly (A) tail, which accounts for variably-sized LEAP products (data not shown [17]).

Epitope-tagged ORF2p form a complex with ORF1p and their encoding RNA

To further verify that ORF1p, ORF2p, and L1 mRNA form an RNP, HeLa cells were transfected with either pES2TE1 or

pDK101. Whole cell extracts then were subjected to immunoprecipitation using an anti-FLAG M2 antibody fused to agarose beads (Figure 3B). Incubation of the beads with a FLAG peptide followed by western blot analysis revealed an enrichment of ORF1p and ORF2p in the pES2TE1, but not in the pDK101 immunoprecipitated reactions (Figure 3C). We sometimes detected a faint band of \sim 40 kDa in the pDK101 immunoprecipitated reactions upon longer film exposures, suggesting that some T7-tagged ORF1p may bind non-specifically to the anti-FLAG M2 agarose beads (data not shown). However, subsequent experiments/product characterization determined that the pES2TE1 immunoprecipitated fraction contained LEAP activity, whereas the pDK101 immunoprecipitated fraction lacked a readily detectable LEAP activity (Figure 3D).

Interestingly, we consistently observed less ORF1p associated with RNPs in immunoprecipitation experiments when compared to experiments conducted with whole cell lysates or crude RNPs (Figure 2A and Figure 3A). These data suggest either that ORF1p is less tightly associated with L1 mRNA than ORF2p in RNPs (which is consistent with previous observations [17]) and/or that a fraction of ORF1p is dissociated from L1 RNA during the immunoprecipitation process. Regardless, whereas previous studies showed that ORF1p, ORF2p RT activity, and L1 RNA co-localize to RNPs [16,17], we were able to demonstrate the physical association of these components in immunoprecipitation experiments.

Mutations in both ORF1p and ORF2p affect L1 RNP formation and/or function

Previous studies identified activities associated with ORF1p and ORF2p that are critical for L1 retrotransposition [22,30,32,35]. Here, we expanded on these analyses to determine whether mutations in the L1-encoded proteins affect their ability to localize to RNPs and/or impact L1 reverse transcriptase activity in the LEAP assay. We first tested mutants in the following functional domains of ORF1p: 1) the putative leucine zipper domain (pADLZC; L_{93,100,107,114V}); 2) the RNA-recognition motif (pAD113; NLR_{157-159ALA}); 3) the carboxyl-terminal nucleic acid binding domain (pAD105; RR_{261-262AA}); 4) an ORF1p mutation that affects mouse nucleic acid chaperone activity (pAD106; RR_{261-262KK}); and 5) a double mutant in the putative leucine zipper domain and carboxyl-terminal nucleic acid binding domain (pADL/R; L_{93,100,107,114V}/RR_{261-262AA}) (Figure 4A; see Materials and Methods) [16,30-32,35]. Each of these mutations, including the LZC mutation (pADLZC; L_{93,100,107,114V}), severely compromise L1 retrotransposition efficiency in HeLa cells (Figure S1A, S1B). The LZC mutant data are in agreement with a published report, which demonstrated a L_{93/100/114A} triple mutation inactivates L1 retrotransposition [31].

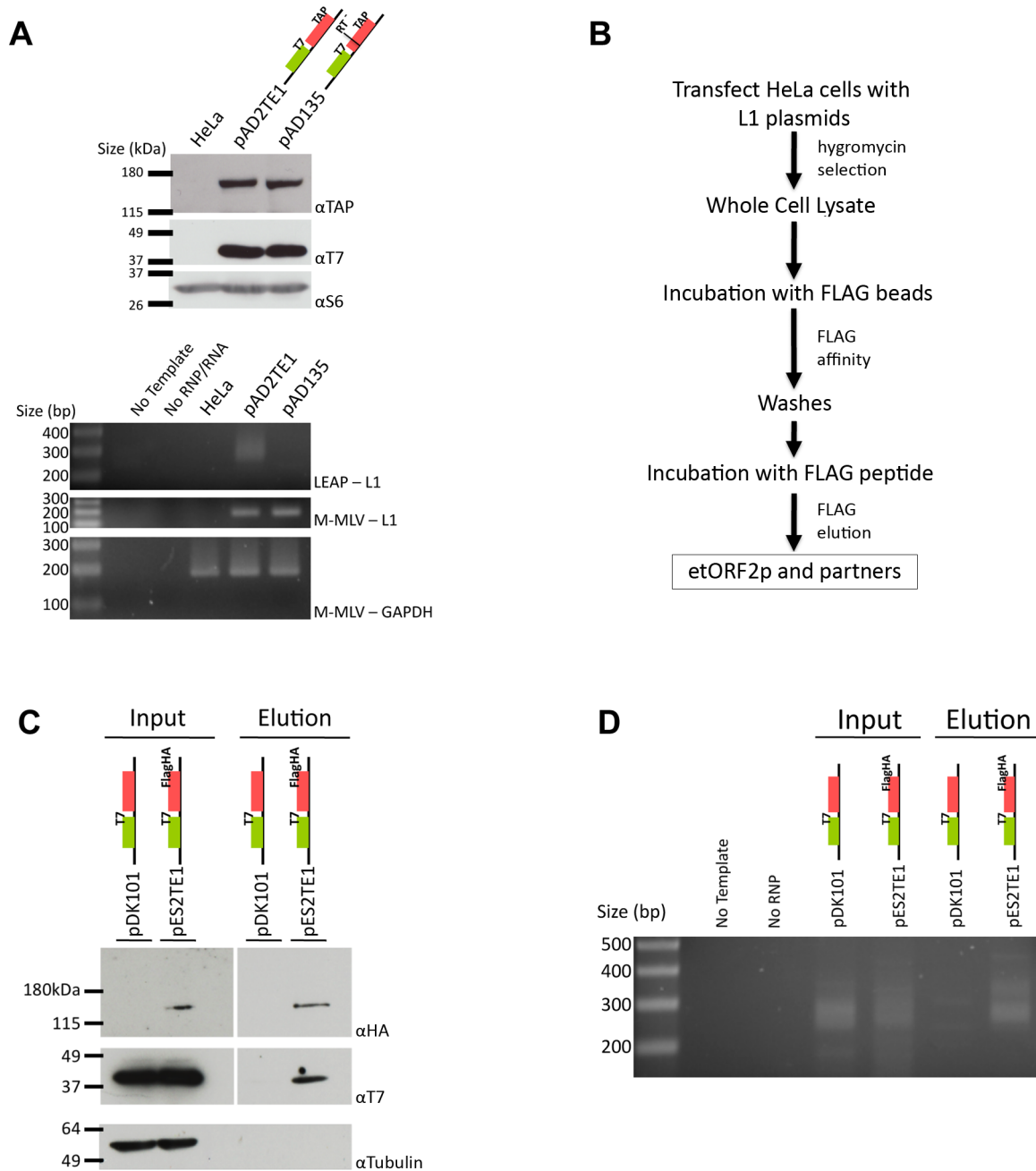


Figure 3. Biochemical identification of a basal L1 RNP complex. *A. L1 RNPs contain ORF1p, ORF2p, L1 RNA, and L1 reverse transcriptase activity:* RNP pellets were obtained from untransfected HeLa cells, or from HeLa cells transfected with wild-type (pAD2TE1) and reverse transcriptase deficient (pAD135) L1 constructs. As in Figure 2, tagged ORF1p and ORF2p were detected using anti-T7 (α T7) and anti-TAP (α TAP) respectively. Ribosomal S6 protein was detected using an anti-S6 (α S6) antibody and was used as an RNP loading control. Reverse transcriptase activity was detected using the LEAP assay as described previously [17]. Reactions without template (No Template) or RNPs (No RNP/RNA) were used as negative controls. Top panel: LEAP reactions (LEAP-L1). Middle panel: L1 RT-PCR reactions conducted with M-MLV reverse transcriptase control for the presence of L1 RNA in RNPs (M-MLV-L1). Bottom panel: GAPDH RT-PCR reactions conducted with M-MLV reverse transcriptase assess RNP RNA quality and serve as a RT-PCR internal control (M-MLV-GAPDH). *B. Flow chart of the L1 RNP immunoprecipitation reaction:* Whole cell extracts were prepared from HeLa cells transfected with either pDK101 or pES2TE1. Immunoprecipitation reactions were conducted by incubating the resultant lysates with agarose beads fused to an anti-FLAG M2 antibody. The elution of ORF2p from the beads was performed by FLAG peptide competition. Western blotting and LEAP assays were performed on aliquots of the whole cell extracts or the elution fractions to detect the L1-encoded proteins and L1-specific reverse transcriptase activity, respectively. *C. Co-immunoprecipitation of ORF1p and ORF2p:* Whole cell extract (input) and immunoprecipitated (elution) products from pDK101 or pES2TE1 transfected cells were subjected to western blotting to identify ORF2p (α HA; top panel), ORF1p (α T7 middle panel) or tubulin (α Tubulin, bottom panel). The femto ECL substrate (Pierce) was used to detect ORF1p and ORF2p. *D. A basal L1 RNP complex contains L1 RNA and retains L1 reverse transcriptase specific activity:* LEAP was performed on whole cell extracts (input) or immunoprecipitated (elution) products from pDK101 or pES2TE1 transfected cells. Reactions conducted without template (No Template) or without RNPs (No RNP) were used as negative controls. As in Figure 2, colored cartoons of the constructs are indicated in panels A, C and D. Molecular weight/DNA size markers (Invitrogen) are indicated at the left of the images. All constructs contain the *mneol* retrotransposition indicator cassette. doi:10.1371/journal.pgen.1001150.g003

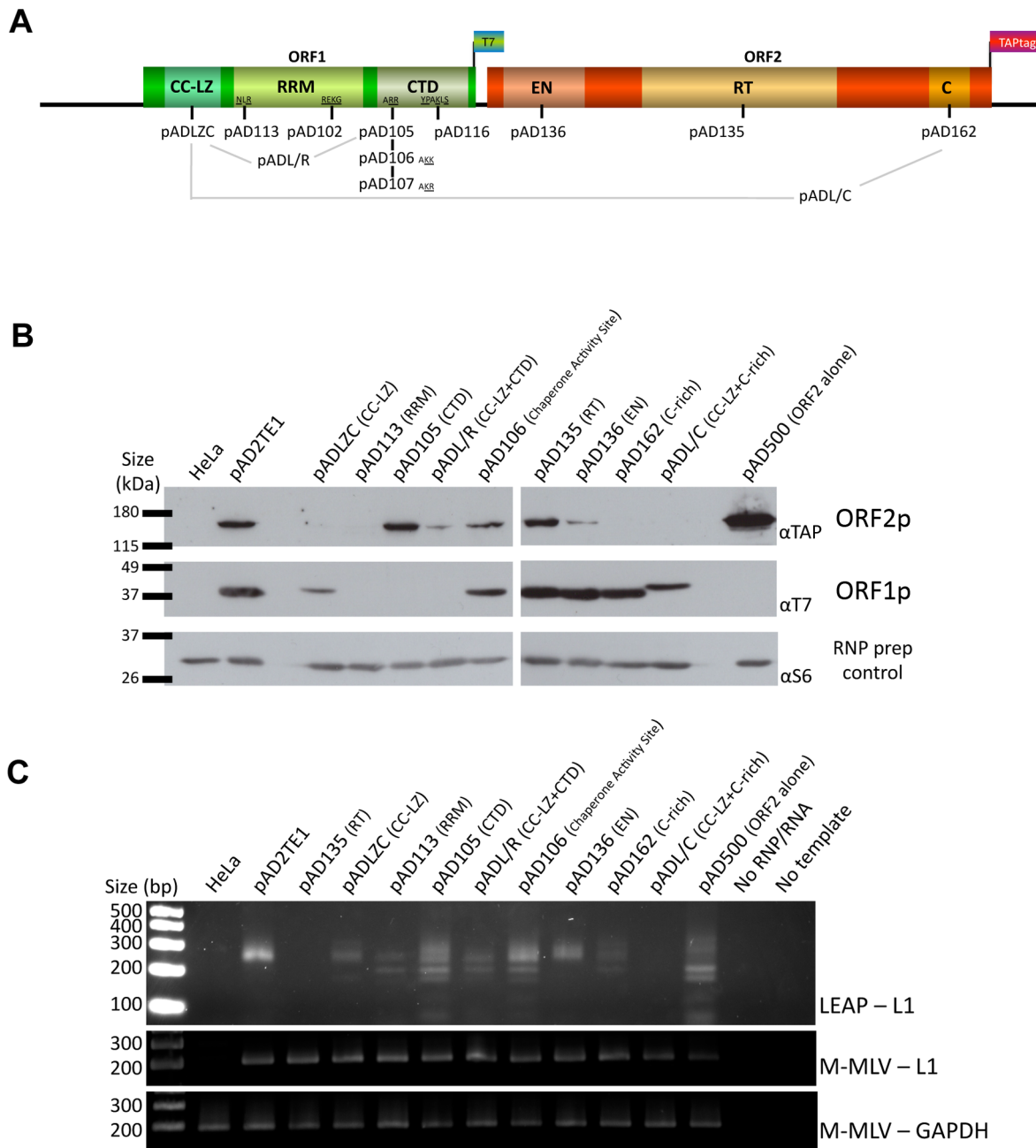


Figure 4. TAP tagged ORF2p and RT activity detection in RNP preparation. *A. Schematic representation of the amino acid mutation positions in L1 sequence:* The names of plasmids containing L1s with mutations in the ORF1p coiled-coil domain (CC-LZ), the ORF1p RNA recognition motif (RRM), and the ORF1p carboxyl-terminal (CTD) domain are indicated below the schematic. The names of plasmids containing mutations in the ORF2p endonuclease domain (EN), reverse transcriptase domain (RT) or cysteine-rich domain (C) also are shown. pADL/R is a double mutant that contains a putative leucine zipper mutation and a carboxyl-terminal domain mutation in ORF1p. pADL/C is a double mutant that contains a putative leucine zipper mutation in ORF1p and a C-domain mutation in ORF2p. The flags indicate the epitope tag present on ORF1 and ORF2. *B. Detection of ORF1p and ORF2p from mutant L1 constructs:* RNPs from HeLa cells transfected with a RC-L1 (pAD2TE1) or the indicated mutant L1 constructs (see Figure 4A) were analyzed by western blotting [16]. Tagged L1 proteins were detected as in Figure 3; ORF2p (top panel), ORF1p (middle panel). Ribosomal S6 protein detection was used as a loading control (bottom panel). Molecular weight markers (Invitrogen) are indicated at the left of the image. *C. L1 RT activity of RNP fractions detected by LEAP:* An aliquot from each of the indicated RNP preparations noted above was used to perform LEAP assays (see Figure 3) [17]. RNPs from pAD2TE1 served as a positive control. RNPs from untransfected HeLa cells or pAD135 (D₇₀₂A; RT mutant) transfected cells served as negative controls. Reactions without RNPs (No RNP/RNA) or template (No Template) also were used as negative controls. Top panel: LEAP reactions (LEAP-L1). Middle panel: L1 RT-PCR reactions conducted with M-MLV reverse transcriptase control for the presence of L1 RNA in the RNP fractions (M-MLV-L1). Bottom panel: GAPDH RT-PCR reactions conducted with M-MLV reverse transcriptase assess RNP RNA quality and serve as a RT-PCR internal control (M-MLV-GAPDH). DNA size markers (Invitrogen) are indicated at the left of the image. All constructs in panel B and C contain the *mneol* retrotransposition indicator cassette.
doi:10.1371/journal.pgen.1001150.g004

Multiple independent RNP preparations derived from cells transfected with each of the respective mutants were analyzed by western blotting to examine the presence and abundance of both ORF1p and ORF2p (Figure 4B). LEAP assays then were used to determine whether those RNPs contained an L1-specific reverse transcriptase activity (Figure 4C). Once again, control MLV RT-PCR-based experiments, using the same oligonucleotide primers employed in the LEAP assay, indicated that L1 RNA was present at roughly comparable levels in the RNP fraction of HeLa cells transfected with the mutant constructs (Figure 4C).

Consistent with previous data [16], a mutation in the ORF1p carboxyl-terminal domain (pAD105; RR_{261–262}AA) led to a severe reduction in the ability of ORF1p, but not ORF2p, to localize to the RNP fraction (Figure 4B). RNPs derived from pAD105-transfected cells had a readily detectable LEAP activity, although the constellation of LEAP products differed from those in the wild-type control, pAD2TE1, because they frequently initiated reverse transcription from within the 3' end of the L1 mRNA (Figure 4B and 4C; Figure S2A, S2B, S2C and S2D; pDK105 RR_{261–262}AA; data not shown). Similar data also were observed for an L1 containing a mutation in the carboxyl-terminal domain (pDK116; YPAKLS_{282–287}AAAALA) as well as for pAD500, a TAP-tagged ORF2p construct that lacks ORF1 (Figure 4B and Figure S2A and S2B). These findings support the hypothesis that ORF2p can preferentially associate with its encoding RNA independent of ORF1p binding and that the resultant RNPs retain LEAP activity [17]. Indeed, the constellation of LEAP products observed in the RR_{261–262}AA, YPAKLS_{282–287}AAAALA, and pAD500 mutants support our previous hypothesis that ORF1p binding to L1 mRNA possibly may restrict hybridization of the LEAP primer to the L1 poly (A) tail [17].

Mutations that affect mouse nucleic acid chaperone activity (pAD106; RR_{261–262}KK) had little effect on the ability of ORF1p and ORF2p to localize to RNPs or on LEAP activity (Figure 4B; [16,35]). We occasionally observed a greater abundance of the lower molecular weight LEAP products, when compared to our wild-type control, pAD2TE1 (Figure 4C). Indeed, closer inspection consistently revealed slightly higher levels of the major LEAP products (~220 to ~400 bp) and a slightly lower level of the shorter LEAP products from the RR_{261–262}KK mutant (pAD106 and pDK106; Figure 4C, Figure S2B and S2C) when compared to LEAP products derived from the RR_{261–262}AA (pAD105 and pDK105; Figure 4C, Figure S2B and S2C) and YPAKLS_{282–287}AAAALA mutants (pDK116; Figure S2B). Thus, although the L1 RT activity detected in the LEAP assay does not appear to require ORF1p, it is clear that specific mutations in ORF1p can affect the constellation of products observed in these assays.

Mutations in the putative ORF1p leucine zipper-binding domain (pADLZC; L_{93,100,107,114}V) reduced ORF1p and ORF2p localization in the RNP fraction and consistently exhibited lower qualitative levels of LEAP activity when compared to the wild-type control, pAD2TE1 (Figure 4B and 4C). Indeed, quantitative LEAP experiments conducted with pLZC-derived RNPs (a L_{93,100,107,114}V mutant that lacks an epitope tag on ORF2p) revealed a five to seven-fold reduction in LEAP activity when compared to a corresponding wild-type control (pDK101; Figure S2E). Subsequent data from LEAP experiments designed to detect variable length L1 cDNA products further suggest that the LZC mutation adversely affects early steps in the reverse transcription of L1 RNA and does not appear to affect L1 RT elongation (Figure S2F).

The putative leucine zipper domain-carboxyl terminal domain double mutant (ADL/R; L_{93,100,107,114}V/RR_{261–262}AA) shared biochemical characteristics of each single mutant. Similar to

pAD105; RR_{261–262}AA, ORF1p levels were severely reduced in pADL/R-derived RNPs. However, similar to the putative leucine zipper domain (pADLZC; L_{93,100,107,114}V) mutant, ORF2p levels, as well as LEAP activity, were reduced in pADL/R-derived RNPs when compared to the wild-type control, pAD2TE1. Moreover, the LEAP product profile in the double mutant resembled that in the pAD105 mutant (Figure 4B and 4C; Figure S2). Thus, the above data suggest that the LZC mutant adversely affects the accumulation and/or stability of L1 RNPs and that the reduction of ORF2p in RNPs likely contributes to the observed decrease in LEAP activity.

Mutations in the ORF1p RRM domain (pAD113; NLR_{157–159}ALA) also led to a severe reduction in the ability of ORF1p and ORF2p to localize to the RNP fraction of transfected cells (Figure 4B). Indeed, ORF2p only was observed upon over-exposure of the resultant western blots (data not shown). The reduced level of ORF2p in pAD113-derived RNPs also correlated with a decrease in LEAP activity when compared to the pAD2TE1 wild-type control (Figure 4B and 4C). Notably, it is unlikely that the NLR_{157–159}ALA mutation dramatically affects ORF2 translation because we can detect ORF2p from this mutant by immunofluorescence (see below). Moreover, preliminary data (n = 4 independent experiments) indicates that the NLR_{157–159}ALA mutant can serve as a “driver” in a genetic-based *trans*-complementation assay to mobilize a reporter gene (*ORF1mneoI*; [19]) at roughly 60 to 80% the level of the wild-type control, pAD2TE1-NT (Doucet et al., preliminary data). These data are consistent with previous genetic studies, which suggested that ORF1p binding to L1 RNA is not required for ORF2 translation [37]. Moreover, the data suggest that the NLR_{157–159}ALA mutations severely compromise the accumulation and/or stability of L1 RNPs (see Discussion).

We next tested mutants in the following functional domains of ORF2p for their effect on L1 RNP formation and L1 reverse transcriptase activity: 1) the L1 endonuclease domain (pAD136; H₂₃₀A); 2) the L1 reverse transcriptase domain (pAD135; D₇₀₂A); 3) the cysteine-rich domain (pAD162; CWWDC_{1143–1147}SWWDS) (Figure 4A) [19,22,32]. As expected, the L1 RT mutant (pAD135; D₇₀₂A) did not dramatically affect the ability of ORF1p or ORF2p to localize to RNPs, although it did abolish LEAP activity (Figure 4B and 4C) [17]. These data are consistent with previous suppositions that the D₇₀₂A mutant likely blocks the reverse transcription step in TPRT [17,32,40].

We repeatedly observed a slight reduction of ORF2p in RNPs derived from the tested endonuclease mutant, and this reduction correlates with a reproducible decrease in LEAP activity (Figure 4A; pAD136; H₂₃₀A). We also observed a severe reduction of ORF2p, as it was only detected upon longer film exposures (data not shown), and a strong decrease of LEAP activity in RNPs derived from the tested cysteine-rich domain mutant (pAD162; CWWDC_{1143–1147}SWWDS). Finally, the leucine zipper/C-domain double mutant (pADL/C; L_{93,100,107,114}V/CWWDC_{1143–1147}SWWDS) displayed both a reduction of ORF1p in RNPs and a concomitant decrease in LEAP activity (Figure 4B and 4C).

As additional controls for the above experiments, we demonstrated that mutant constructs containing a T7-epitope tag on ORF1p, but lacking an ORF2p epitope tag exhibited similar qualitative LEAP activities as the pAD2TE1 mutant based constructs (Figure S2). We also demonstrated that the amount of T7-tagged ORF1p and TAP-tagged ORF2p in whole cell lysates is similar to that in the RNP fraction for each of the pAD2TE1 mutant constructs, and that these proteins were not enriched in insoluble aggregates in the pellet obtained after cell lysis (data not shown). Thus, we conclude that mutations within discrete

functional domains of ORF1p and ORF2p have differential effects on L1 RNP formation/function.

Immunofluorescence detection of the L1-encoded proteins and L1 mRNA

Previous studies have shown that ORF1p often aggregates in cytoplasmic structures termed cytoplasmic foci [31]. Unlike the RNP assays described above (which detect the steady state amount of ORF1p and ORF2p in the RNP fraction of hygromycin resistant cells ~9 days post-transfection), the L1 cytoplasmic foci formation assays allows the opportunity to visually detect the L1-encoded proteins and/or L1 RNA when over-expressed ~48 hours post-transfection.

To test whether the ORF1p cytoplasmic foci also contain ORF2p and L1 RNA, we conducted immunofluorescence-based localization experiments in a U-2 OS human osteosarcoma cell line that can support the retrotransposition of engineered human L1 constructs (Figure S3A). Initial experiments conducted with pAD2TE1 revealed that ORF1p and ORF2p generally co-localized to discrete cytoplasmic foci 48 hours post-transfection, and that many foci were located near the periphery of the nucleus (Figure 5A). Time course analyses further demonstrated that cytoplasmic foci were apparent in ~50% of transfected cells as early as 12 hours post-transfection, and that ~90% of transfected cells displayed cytoplasmic foci 72 hours post-transfection (Figure S3B). ORF1p/ORF2p-containing cytoplasmic foci also were observed in U-2 OS cells transiently transfected with pAD2TE1-NT, which lacks the *mneoI* retrotransposition indicator cassette (Figure 5A) and with a pAD2TE1 derivative lacking the heterologous cytomegalovirus immediate early (CMV) promoter, although foci appeared 24–48 hours later as compared to cells transfected with the wild type control, pAD2TE1 (data not shown). ORF1p and ORF2p co-localization also was observed using an anti-HA antibody to detect ORF2p (Figure 5A; pES2TE1) or antibodies against endogenous ORF1p or ORF2p (Figure S3C; pES2TE1). Qualitatively similar results were obtained when pAD2TE1 was transiently transfected into HeLa or 143Btk cells, which also support L1 retrotransposition [32,51] (data not shown).

To test whether L1 RNA co-localizes with ORF1p and ORF2p to cytoplasmic foci, we transiently transfected pAD3TE1 into U-2 OS cells. *In situ* hybridization experiments using a fluorescently-labeled probe complementary to the MS2 stem loop structures in the L1 3'UTR revealed the presence of L1 RNA in cytoplasmic foci as well as in nuclei of transfected cells (Figure 5B and Figure S3D). The co-localization of ORF1p, ORF2p, and L1 RNA was confirmed by conducting co-transfection experiments with pAD3TE1 and a plasmid expressing a fluorescently labeled MS2 protein (Figure S3D), and by staining with antibodies against ORF1p and ORF2p (Figure S3C). As above, qualitatively similar results were obtained upon transient transfection of pAD3TE1 into HeLa or HEK293 cells, which also support L1 retrotransposition [32,51] (data not shown).

To determine whether L1 foci are associated with specific cytoplasmic substructures, we co-transfected U-2 OS cells with pAD2TE1 and plasmids that express GFP fusion proteins that can localize to processing bodies (*i.e.*, P-bodies) and/or stress granules. Consistent with previous analyses, ORF1p and ORF2p associated with an Ago2-GFP fusion protein that localizes both to P-bodies and stress granules (Figure 5C; panel 1) [31,52]. Refining this analysis revealed that ORF1p and ORF2p co-localized with the stress granule marker G3BP-GFP [53], but did not associate with the P-body marker DCP1 α -GFP [54] (Figure 5C; panel 2 and 3). By comparison, experiments conducted with fluorescently labeled antibodies specific for eIF3 and G3BP [53,55] revealed that stress

granules appear to closely associate with the L1 foci (Figure 5C, panel 4 and 5).

Together, the above data demonstrate that ORF1p, ORF2p, and L1 mRNA co-localize to cytoplasmic foci when over-expressed from a variety of engineered L1 episomal expression constructs and that many of these cytoplasmic foci associate with stress granules. However, future experiments are needed to determine whether cytoplasmic foci represent accumulation depots for L1 RNPs or if they play an important role in L1 retrotransposition.

Mutations in ORF1p and ORF2p adversely affect the formation of L1 cytoplasmic foci

We next examined if mutations in the L1-encoded proteins affect L1 cytoplasmic foci formation. Transient transfection of ORF1p mutant expression vectors into U-2 OS cells followed by immunofluorescence staining with anti-T7 and anti-TAP antibodies confirmed that ORF1p and ORF2p are expressed in these cells (Figure 6A). Consistent with previous studies, mutations in the ORF1p RRM domain (pAD113; NLR_{157–159}ALA) and carboxyl-terminal RNA binding domain (pAD105; RR_{261–262}AA) led to a reduction in the number of L1 cytoplasmic foci (Figure 6A and 6B) [31]. A reduction in the number of L1 cytoplasmic foci also was observed for an RRM domain mutant (pAD102; REKG_{235–238}AAAA), an additional carboxyl-terminal domain mutant (pAD116; YPAKLS_{282–287}AAAALA), and the putative leucine zipper domain/carboxyl-terminal RNA binding domain double mutant (pADL/R; L_{93,100,107,114}V/RR_{261–262}AA). By comparison, mutations in the putative ORF1p LZ domain (pADLZC; L_{93,100,107,114}V) or mutations that affect the nucleic acid chaperone activity of mouse ORF1p (pAD106; RR_{261–262}KK and pAD107; R₂₆₁K) had little effect on L1 cytoplasmic foci formation (Figure 6A and 6B), although we sometimes observed an apparent nucleolar localization of ORF1p in pADLZC transfected cells. None of the ORF2p mutations had a dramatic effect on L1 cytoplasmic foci formation (Figure 6A and 6B), although, we observed a diffuse nuclear localization of TAP-tagged ORF2p in cells transfected with either pAD162 or the putative leucine zipper domain/cysteine-rich domain double mutant (pADL/C; L_{93,100,107,114}V/CWWDC_{1143–1147}SWWDS) (Figure 6A).

The above data suggest that the ability of ORF1p to bind L1 RNA is critical for L1 cytoplasmic foci formation (Figure 6B). Consistent with this idea, we were able to detect L1 cytoplasmic foci, as well as diffuse ORF1p staining, in U-2 OS cells transiently transfected with a T7-tagged ORF1p expression vector (Figure 6C; pDK500). However, L1 cytoplasmic foci were not detected in U-2 OS cells transiently transfected with a TAP-tagged ORF2p expression vector (Figure 5D; pAD500). Thus, these data, as well as our previously published *trans*-complementation experiments [56], suggest that ORF1p interacts with its encoding RNA *in cis*, and that this association allows L1 cytoplasmic foci formation in the absence of ORF2p.

Discussion

ORF2p has been notoriously difficult to detect from engineered human L1s in cultured cells. It has been hypothesized that human ORF2p is translated at low levels when compared to ORF1p and/or may be an unstable protein, which might help explain why it has evaded detection [36–39,42]. Previous biochemical studies have identified human ORF2p from vascular endothelial cells *in vivo* [57] and have demonstrated that ORF2p RT activity co-localizes with ORF1p and L1 RNA in cytoplasmic RNPs derived from HeLa cells transfected with wild-type engineered human L1

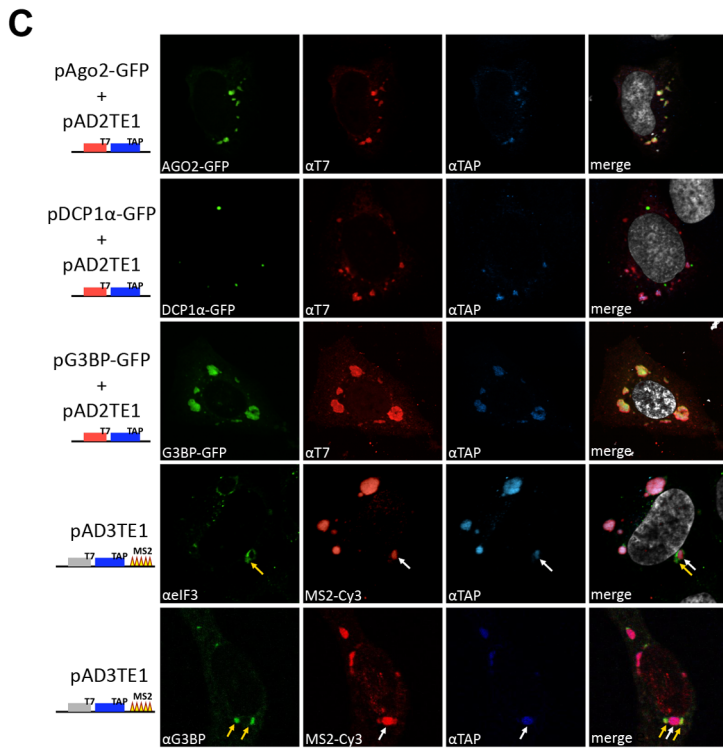
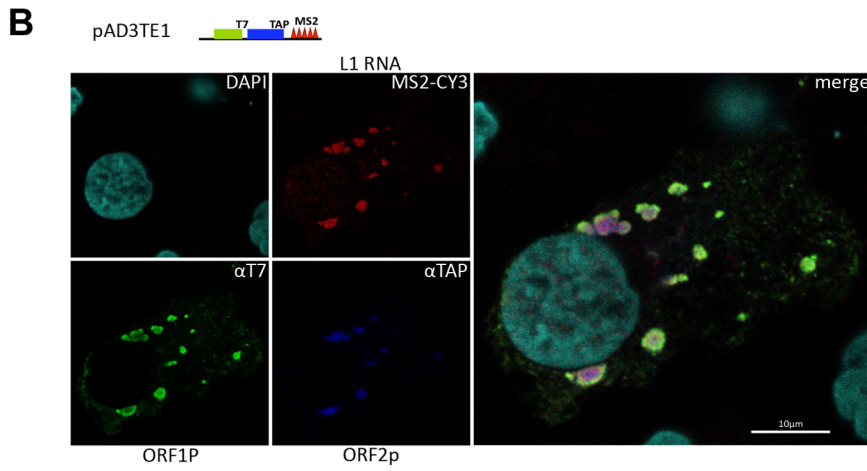
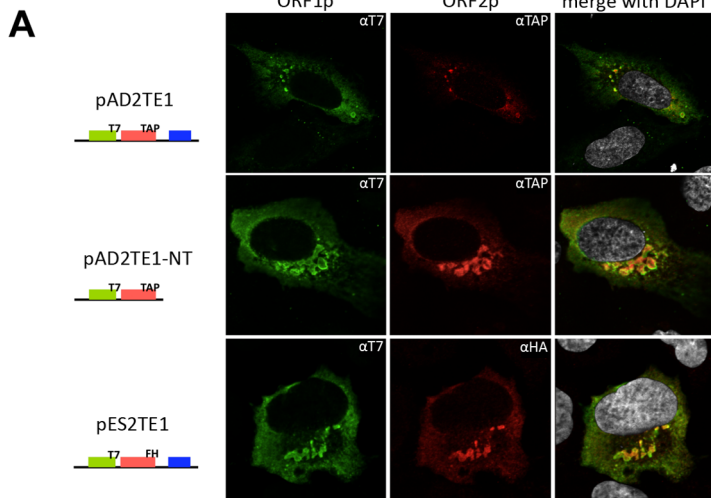


Figure 5. Cellular identification of L1 cytoplasmic foci. *A. Cellular localization of the L1-encoded proteins:* Immunofluorescence was conducted on pAD2TE1 transfected U-2 OS cells 48 hours post-transfection. T7-tagged ORF1p (green; left column) and TAP-tagged ORF2p (red; middle column) staining are shown for representative transfected cells. A merged image is shown in the rightmost column; DAPI (grey) was used to stain nuclear DNA. Cartoons of the constructs are indicated at the left of the micrographs. The blue rectangle in the constructs indicates the *mneol* cassette. *B. Cellular localization of L1-encoded proteins and RNA:* Immunofluorescence/RNA FISH was conducted on pAD3TE1 transfected U-2 OS cells 48 hours post-transfection. T7-tagged ORF1p (green), TAP-tagged ORF2p (blue), L1 RNA (red), and DAPI (turquoise) staining are indicated in left four micrographs. A merged image is shown in the rightmost panel. The cartoon of pAD3TE1 is shown above the micrographs. *C. L1 cytoplasmic foci are associated with stress granules:* Immunofluorescence/fluorescence microscopy was performed on U-2 OS cells co-transfected with pAD2TE1 and one of the following plasmids: 1) pAgo2-GFP (green staining, top row of images); 2) pDCP1 α -GFP (green staining, second row of images); 3) pG3BP-GFP (green staining, third row of images). T7-tagged ORF1p (red), and TAP-tagged ORF2p (blue) also are shown. A merged image is shown in the rightmost panels; DAPI (grey) was used to stain nuclear DNA. Immunofluorescence also was performed on U-2 OS cells transfected with pAD3TE1. Images using antibodies against the endogenous stress granule components eIF3 (α eIF3 (green)) and G3BP (α G3BP (green)) are shown. L1 RNA (red), and ORF2p (blue) also are indicated. Arrows indicate the association of L1 cytoplasmic foci (white) and stress granules (yellow). A merged image is shown in the rightmost columns; DAPI (grey) was used to stain nuclear DNA. All L1 constructs in panel B and C contain the *mneol* retrotransposition indicator cassette.
doi:10.1371/journal.pgen.1001150.g005

expression constructs [16,17]. Here, we have built on these studies and have combined epitope and RNA tagging strategies to physically detect L1 ORF1p, ORF2p and L1 mRNA in cytoplasmic RNPs.

Why our approach allows the ready detection of ORF2p expressed from engineered human L1 constructs requires further study. Experiments conducted with anti-TAP antibodies consistently yielded more robust detection of ORF2p when compared to anti-HA or anti-ORF2p antibodies. Thus, the inclusion of a large carboxyl-terminal tag, such as the TAP-tag, might stabilize ORF2p. However, since engineered L1 constructs containing either the TAP or HA epitope tags on the carboxyl-terminus of ORF2p remain retrotransposition-competent, the strategy described here allows a way to both directly study the expression of ORF1p and ORF2p from a bicistronic transcript and establishes an experimental platform to determine how each protein interacts with L1 RNA. Furthermore, this strategy now allows a comprehensive means to assess how mutations in ORF1p and/or ORF2p affect L1 RNP biogenesis and/or L1 retrotransposition.

Biochemical methods allowed us to assess how mutations in the L1-encoded proteins affect RNP function. For example, in agreement with previous studies, a mutation in the carboxyl-terminal domain of ORF1p (pAD105, RR_{261–262}AA) markedly reduced ORF1p levels in RNPs, but did not noticeably affect ORF2p accumulation or LEAP activity (Figure 6A; Figure S3) [16,17]. Similarly, we could detect ORF2p and LEAP activity in RNPs derived from cells transfected with a construct that lacks ORF1 (Figure 6A and 6B). Thus, we conclude that ORF2p can preferentially associate *in cis* with its encoding transcript to form an RNP independently of ORF1p RNA binding (Figure 7).

Our studies further suggest that an interplay exists between ORF1p, ORF2p and L1 RNA that is critical for proper L1 RNP formation/function (Figure 7). For example, mutations in the putative leucine zipper (pADLZC; L_{93,100,107,114}V) or RRM (pAD113; NLR_{157–159}ALA) domains led to a reduced amount of ORF2p in the RNP fraction, as well as a decrease in LEAP activity. The L_{93,100,107,114}V mutations reside in the N-terminal coiled-coil domain of ORF1p and could potentially alter the structure of the protein. Similarly, the NLR_{157–159}ALA mutations reside near coiled-coil domain/RRM junction and structural studies indicate that a hydrogen bond between N₁₅₇ and D₂₅₂ is important for correct folding of the RRM domain [30]. Thus, both of the above mutations may adversely affect the structural integrity of ORF1p, leading to the destabilization of the resultant L1 RNPs. Indeed, such a scenario could potentially account for the reduced levels of ORF2p in RNPs and/or L1 RT activity in these mutants (Figure 7). It is unlikely that the L_{93,100,107,114}V, L_{93,100,107,114}V/RR_{261–262}AA, and NLR_{157–159}ALA mutants significantly affect ORF2p translation, since our preliminary data

indicate that each mutant can serve as a “driver” in a genetic-based *trans*-complementation assay (Doucet, Hulme *et al.*, preliminary data).

As expected, a mutation in the endonuclease domain of ORF2p (pAD136; H₂₃₀A) had no discernable effect on the ability of ORF1p to accumulate in RNPs when compared to a wild-type control construct. However, this mutation consistently led to a slightly reduced amount of ORF2p in RNPs, which correlated with a lower LEAP activity [17]. These findings could potentially explain why the H₂₃₀A mutant consistently exhibited lower levels of endonuclease-independent L1 retrotransposition in Chinese Hamster Ovary cells that are deficient in the non-homologous end-joining pathway of DNA repair when compared to a D₂₀₅A endonuclease domain mutation [58].

Mutations in the cysteine-rich domain (pAD162; CWWDC_{1143–1147}SWWDS) did not have a major effect on the ability of ORF1p to accumulate in the RNP fraction. However, these mutations led to a reduced amount of ORF2p in the RNP fraction and a concomitant decrease in LEAP activity when compared to a wild-type control construct. How mutations in the C-domain affect ORF2p accumulation in RNPs requires further study; however, it is possible that these mutations alter the ability of ORF2p to interact with L1 RNA and/or host factors that are important for the biogenesis of L1 RNPs (Figure 7).

A second assay allowed us to determine how mutations in the L1-encoded proteins affect L1 protein expression and cytoplasmic foci formation shortly after transfection. First, we observed that ORF1p and ORF2p can be detected when transiently expressed from the wild-type and mutant L1 constructs used in the study. We next measured the ability of these proteins to form cytoplasmic foci. Consistent with previous studies, retrotransposition-defective L1s containing mutations in either the RRM (pAD113; NLR_{157–159}ALA, pAD102; REKG_{235–238}AAAA) or carboxyl-terminal domain of ORF1p (pAD105, RR_{261–262}AA; pAD116, YPAKLS_{282–287}AAAALA) reduced L1 cytoplasmic foci formation [31,42,59]. In contrast, mutations in the putative leucine zipper domain (pADLZC; L_{93,100,107,114}V) or mutations analogous to those that adversely affect the nucleic acid chaperone activity of mouse ORF1p (pAD106; RR_{261–262}KK), which are not predicted to inhibit L1 RNA binding, or mutations in ORF2p had little effect on L1 cytoplasmic foci formation [16,35]. Thus, unlike our biochemical assays, the L1 cytoplasmic foci formation assay does not allow us to readily assess ORF2p function. Instead, it provides a valuable tool to screen for ORF1p mutations that affect RNA binding or perhaps protein stability (Figure 7).

Consistent with previous studies, we found that L1 cytoplasmic foci are in close association with proteins that are components of stress granules (Figure 7) [31,59]. Interestingly, recent studies have shown an important role for another cytoplasmic structure (P-

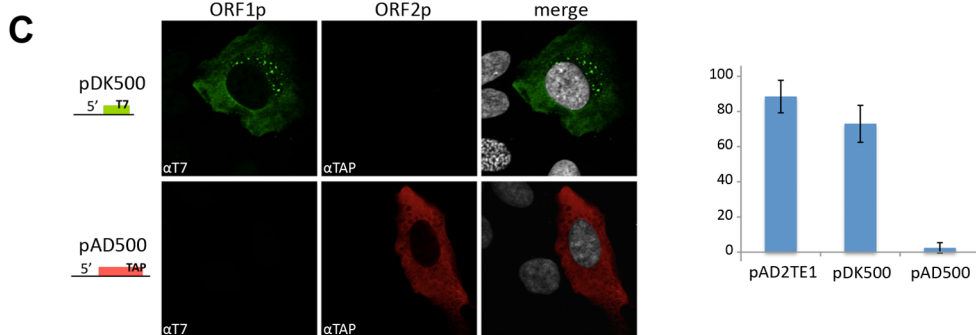
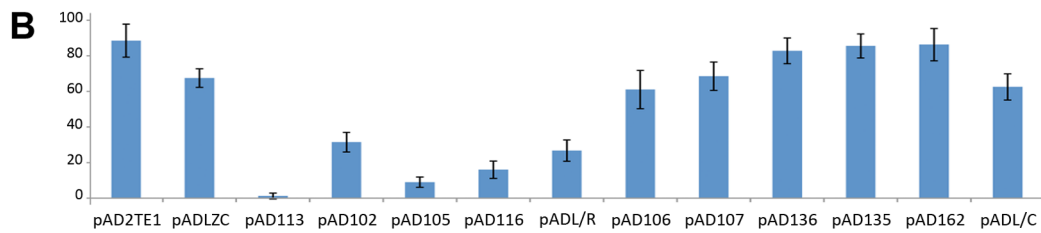
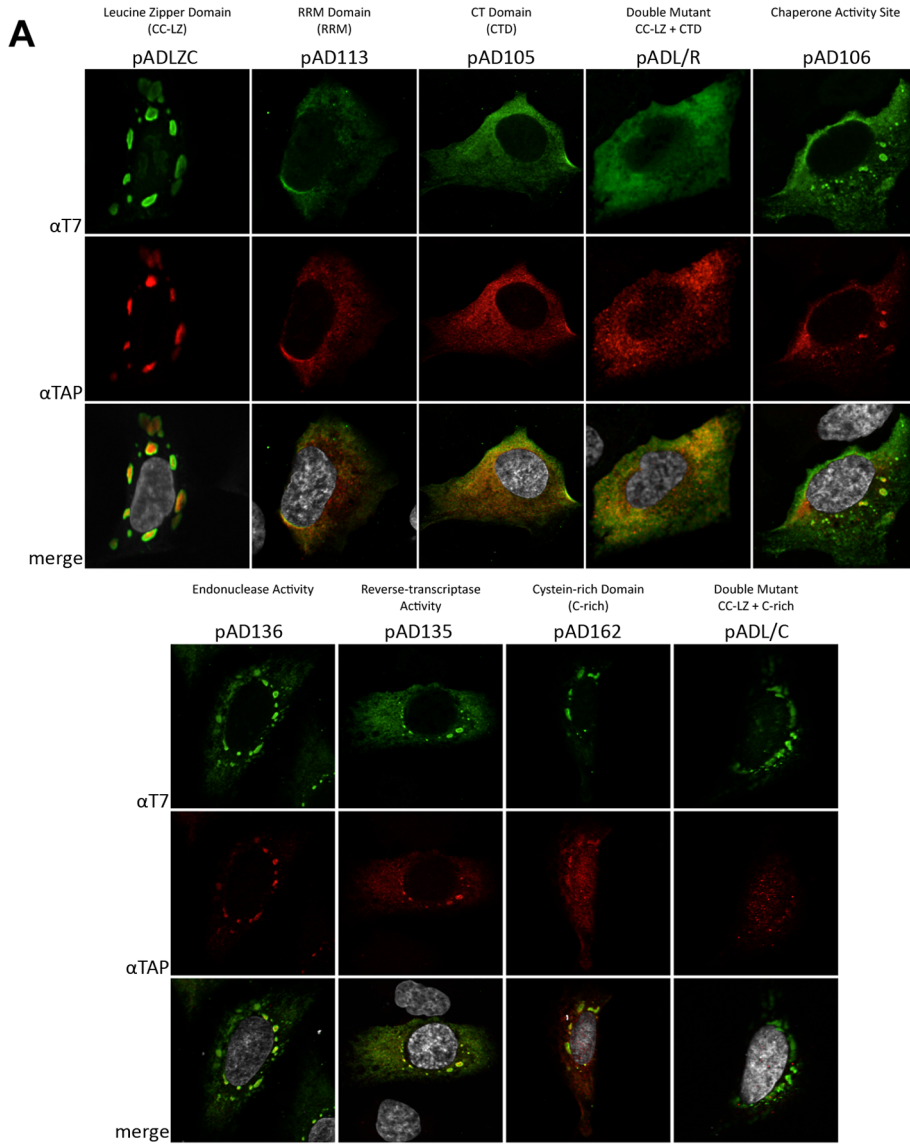


Figure 6. L1 cytoplasmic foci formation requires the nucleic acid binding domain of ORF1p. *A. L1 cytoplasmic foci formation requires the nucleic acid binding domain of ORF1p:* Immunofluorescence was performed on U-2 OS cells transfected with the indicated pAD2TE1-derived mutant plasmids (described in Figure 4A). T7-tagged ORF1p (green; top panels) and TAP-tagged ORF2p (red; middle panels) staining are shown for representative transfected cells. A merged image is shown in the bottom panels; DAPI (grey) was used to stain nuclear DNA. *B. Quantitative analyses of L1 cytoplasmic foci formation:* The number of U-2 OS transfected cells that contains L1 cytoplasmic foci were quantified. The name of the construct used for each transfection is indicated on the X-axis (described in Figure 4A). The percentage of transfected cells displaying L1 cytoplasmic foci is indicated on the Y-axis. pAD2TE1 serves as a positive control. The average of four independent experiments is indicated; error bars = standard deviation of the mean. *C. ORF1p is necessary and sufficient for L1 cytoplasmic foci formation:* Immunofluorescence was performed on U-2 OS cells transfected with pDK500 and pAD500. A cartoon of the constructs is shown at the left of the micrographs. T7-tagged ORF1p (green; left column) and TAP-tagged ORF2p (red; middle column) staining are shown for representative transfected cells. A merged image is shown in the rightmost column; DAPI (grey) was used to stain nuclear DNA. The graph indicates the percentage of cells exhibiting L1 cytoplasmic foci. The name of the construct is indicated on the X-axis. The percentage of transfected cells displaying L1 cytoplasmic foci is indicated on the Y-axis. pAD2TE1 serves as a positive control. Four independent analyses of 100 transfected cells were analyzed for each construct. Error bars = standard deviation of the mean. All L1 constructs contain the *mneoI* retrotransposition indicator cassette. doi:10.1371/journal.pgen.1001150.g006

bodies) for Ty3 and Ty1 retrotransposition in yeast [60–62]. Whether L1 cytoplasmic foci play an important role in L1 retrotransposition awaits further experimentation.

In sum, we have developed a powerful system to physically detect the proteins and RNA encoded by both retrotransposition-competent and mutant L1 constructs in RNP complexes, which now augments previous studies that were based on inferring the presence of ORF2p from its enzymatic activity. It is noteworthy that RNPs derived from the RC-L1s characterized in this study exhibit the biochemical properties predicted of a “basal” L1 retrotransposition intermediate. Thus, we speculate that at least some of the L1 cytoplasmic foci identified here could serve as *bona fide* L1 retrotransposition intermediates. Finally, we predict that the use of the L1 expression constructs developed here will allow a powerful means to identify host factors that play a role in L1 retrotransposition and predict that adaptations of this system will prove useful in identifying RNPs encoded by other non-LTR retrotransposons.

Materials and Methods

Oligonucleotides

Sequences of the oligonucleotides used in this study that have been published previously or are available upon request.

3'RACE adapter: 5'- GCGAGCACAGAATTAATACGACT-CACTATAGGTTTTTTTTTTTTTTVN-3'

3'RACE outer: 5'-GCGAGCACAGAATTAATACGACT-3'

GAPDH 3' end: 5'-GACCCTCACTGCTGGGGAGTCC-3'

Neo Promoter Sens (NPS): 5'-GGTTGCTGACTAATTGATGATGCATGC-3'

Neo8161S: 5'-CACATTCCACAGCTGATCGATACC-3'

L1 3'end: 5'-GGGTTTCGAAATCGATAAGCTTGGATCCA-GAC-3'

LEAP-86: 5'-CAAACCACAACACTAGAATGCAGTG-3'

LEAP-46: 5'-GTGAAATTTGTGATGCTATTGC-3'

Plasmid constructs

The following plasmids are based on the previously described pJM101/L1.3 and pDK101 constructs [4,16]. The amino acid and nucleotide numbers indicate the mutation position based on L1.3 accession number L19088 [63]. The constructs were cloned into the pCEP4 expression vector (Invitrogen) and contain the *mneoI* indicator cassette [32,47] in the L1 3'UTR unless otherwise indicated. PCR followed by subcloning was used to introduce the respective epitope tag sequences onto the 3' end of ORF2. As a result of this procedure, we deleted a portion of the L1 3'UTR (nts 5818 to 5953). Oligonucleotides used in our cloning strategies are available upon request.

pADO2Tt contains a Tandem Affinity Purification epitope tag (TAP tag) [43] on ORF2p and was cloned from the pZome-1-C vector (Euroscarf).

pAD2TE1 is derived from pDK101 (L1.3) [16] and contains both the T7 *gene 10* epitope tag on the carboxyl-terminus of ORF1p and a TAP tag on the carboxyl-terminus of ORF2p.

pAD2TE1-Δ2 is derived from pAD2TE1, but lacks CMV promoter and SV40 polyadenylation signal present in the original pCEP4 vector.

pAD2TE1-NT is identical to pAD2TE1, but lacks the *mneoI* indicator cassette.

pES2TE1 is identical to pAD2TE1, but contains a tandem affinity FLAG-HA tag on the carboxyl-terminus ORF2p [44].

pAD500 is derived from L1.3ΔORF1NN [37], and contains a TAP tag on the carboxyl-terminus of ORF2p.

pADL1MT is derived from pJM101/L1.3 and contains 24 repeats of the MS2 stem-loop (MS2 tag) upstream of the *mneoI* indicator cassette in the L1 3'UTR. The MS2 repeats were subcloned from the pTRIP vector [64].

pAD3TE1 is identical to pAD2TE1, but contains the MS2 tag in the 3'UTR (at the same position as in pADL1MT).

pADO1S is identical to pAD2TE1, but contains three stop codons in ORF1. The first two stop codons (R₇Stop; K₈Stop) were generated by introducing a thymidine at nucleotide position 928 to create a frameshift mutation and by mutating an A to a T at nucleotide position 930. The third stop codon is from the construct pJM108/L1.3 carrying the mutation S₁₁₉Stop [19,32].

pADLZC is identical to pAD2TE1, but contains four leucine to valine mutations (L_{93,100,107,114}V) in the ORF1p putative leucine zipper domain.

pAD102 is identical to pAD2TE1, but contains the REKG_{235–238} AAAA mutations in the ORF1p RRM domain [16,32].

pAD105 is identical to pAD2TE1, but contains the RR_{261–262}AA mutations in the ORF1p C-terminal domain [16,19,32].

pAD106 is identical to pAD2TE1, but contains the RR_{261–262}KK mutations in the ORF1p C-terminal domain [16].

pAD107 is identical to pAD2TE1, but contains the RR_{261–262}KR mutation in the ORF1p C-terminal domain [16].

pAD113 is identical to pAD2TE1, but contains the NLR_{157–159} ALA mutations in the ORF1p RRM domain [31].

pAD116 is identical to pAD2TE1, but contains the YPAKLS_{282–287}AAAAALA substitution in the ORF1p C-terminal domain [16,32].

pAD135 is identical to pAD2TE1, but contains the D₇₀₂A mutation in the putative ORF2p RT active site [19].

pAD136 is identical to pAD2TE1, but contains the H₂₃₀A mutation in the ORF2p EN domain [19].

pAD162 is identical to pAD2TE1, but contains the CWWDC_{1143–1147}SWWDS mutations in the ORF2p C-domain [32].

pADL/R is identical to pAD2TE1, but contains a putative leucine zipper domain as well as a C-terminal domain mutant (L_{93,100,107,114}V; RR_{261–262}AA) in ORF1p.

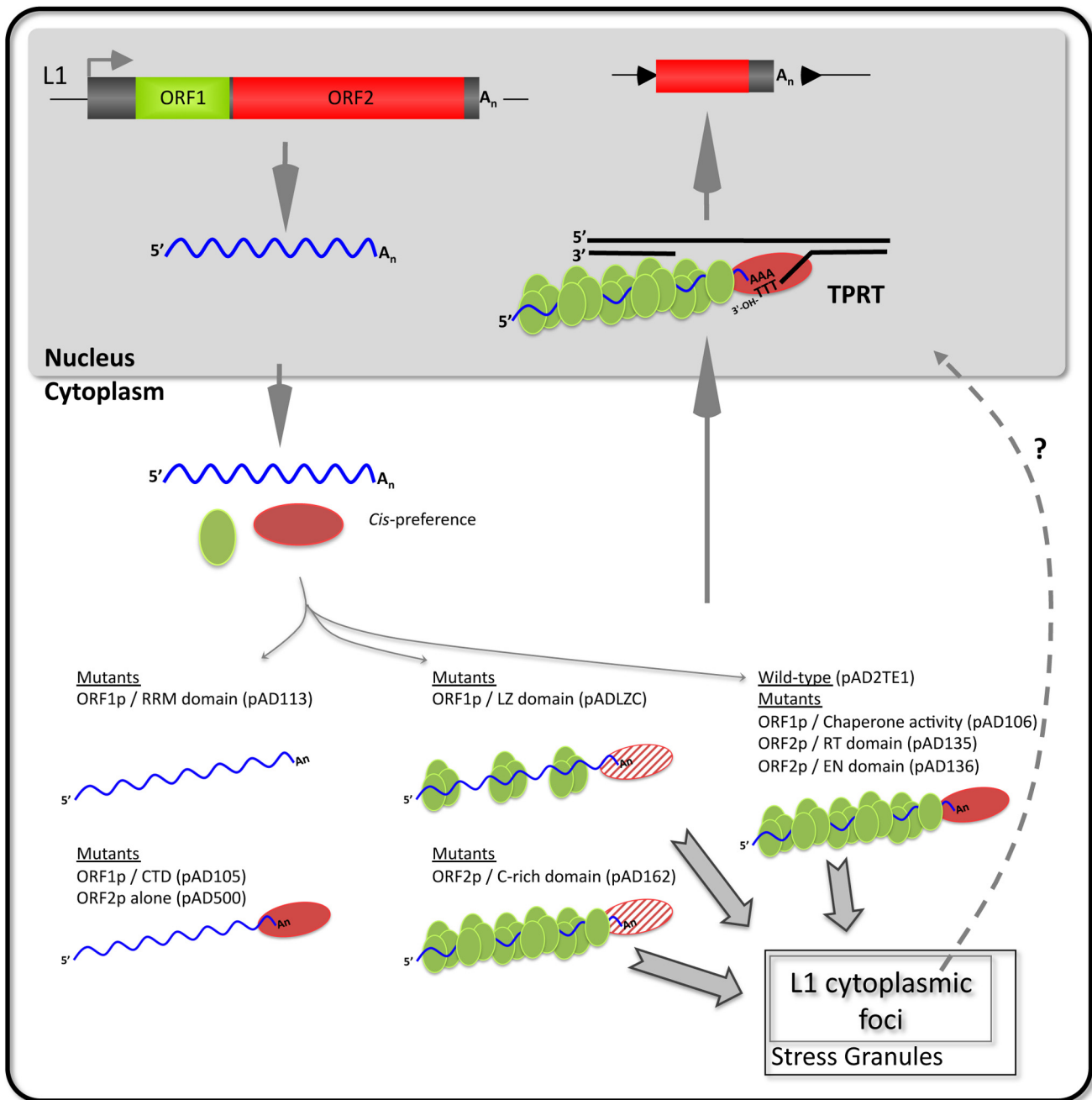


Figure 7. A working model of L1 cytoplasmic RNP formation. A hypothetical model based on our data that builds on previous models of L1 retrotransposition by target-site primed reverse transcription (TPRT; recently reviewed in [6,8]). ORF1p (green oval), ORF2p (red oval), and L1 RNA (wavy blue line) associate with their encoding mRNA via *cis*-preference to form a “basal” retrotransposition complex (right side, pAD2TE1). Mutations in ORF1p and/or ORF2p functional domains have different affects on L1 RNP formation and/or function (thin gray arrows). Mutations in the ORF1p RNA recognition motif (pAD113) disrupt L1 cytoplasmic foci formation and lead to a severe reduction of ORF1p and ORF2p in cytoplasmic RNP complexes (top left side). In some mutants (pAD105 and pAD500) ORF2p can still associate with L1 RNA in the absence of ORF1p RNA binding (bottom left side). Mutations in the putative ORF1p leucine zipper domain (pADLZC) lead to a reduction in ORF1p and ORF2p in RNPs (top center; the reduction in ORF2p is indicated by the striped red oval). Mutations in the ORF2p cysteine-rich domain (pAD162) still allow L1 cytoplasmic foci formation, but adversely affect ORF2p accumulation in RNPs (bottom center). Mutations that disrupt ORF1p nucleic acid chaperone activity (pAD106) or mutations in either the ORF2p endonuclease (pAD136) or reverse transcriptase (pAD135) domains form cytoplasmic RNPs containing ORF1p, ORF2p, and L1 RNA (right side of figure). These mutations probably adversely affect L1 retrotransposition downstream of RNP formation and/or during TPRT. Some RNP complexes localize to L1 cytoplasmic foci and frequently are found in association with stress granules (bold gray arrows). However, whether these foci play a role in L1 retrotransposition remains unknown (indicated by the dotted line and question mark). doi:10.1371/journal.pgen.1001150.g007

pADL/C is identical to pAD2TE1, but contains a putative leucine zipper domain mutation (L_{93,100,107,114V}) in ORF1p as well as a C-domain mutation (CWWDC₁₁₄₃₋₁₁₄₇SWWDS) in ORF2p.

LZC is derived from pDK101 and contains four leucine to valine mutations (L_{93,100,107,114V}) in the ORF1p putative leucine zipper domain.

LZ1/2 is derived from pDK101 and contains two leucine to valine mutations (L_{93,100V}) in the ORF1p putative leucine zipper domain.

LZ2/3 is derived from pDK101 and contains two leucine to valine mutations (L_{100,107V}) in the ORF1p putative leucine zipper domain.

LZ3/4 is derived from pDK101 and contains two leucine to valine mutations (L_{107,114V}) in the ORF1p putative leucine zipper domain.

pDK101, pDK102, pDK105, pDK106, pDK107, pDK108, pDK116, pDK135, and pDK500 were described previously [16].

pMS2-GFP-nls, pMS2-CFP, and pTRIP were generous gifts from Edouard Bertrand [64–66].

pAgo2-GFP and pDCP1 α -GFP were generous gifts from Gregory Hannon [54].

pG3BP-GFP was a generous gift from Jamal Tazi [53].

Cell culture

Cell lines were maintained in a tissue culture incubator (37°C at a 7% CO₂ level) in high glucose Dulbecco's modified Eagle medium (DMEM) without pyruvate (GIBCO), supplemented with 10% fetal bovine calf serum and 1X Penicillin-Streptomycin-Glutamine (GIBCO) as described previously [32].

The L1 retrotransposition assay

The cultured cell retrotransposition assay was conducted as described previously [32,48]. Briefly, 2×10^4 HeLa cells/well were plated in 6 well dishes. Within 24 hours, each well was transfected with 1 μ g of plasmid DNA (prepared with a Midiprep Plasmid DNA Kit (QIAGEN)) using FuGene-6 transfection reagent (Roche). Three days post-transfection, cells were grown in the presence of G418 (400 μ g/mL) to select for retrotransposition events. The media was changed daily. After ~12 days of selection, the resultant cells were washed with 1X Phosphate-Buffered Saline (PBS), fixed, and stained with crystal violet to visualize colonies. In parallel, HeLa cells were plated in 6 well dishes and transfected with 0.5 μ g of the same plasmids and hrGFP (Stratagene). Three days post-transfection cells were subjected to flow cytometry and the transfection efficiency was determined based on the number of GFP positive cells by FACS. In some experiments, 2×10^5 HeLa cells/well were transfected to monitor L1 retrotransposition.

Protein expression and western blot analysis

HeLa cells were transfected with a given L1 expression construct in T-25, T-75, or T-175 tissue culture flasks. Whole cell lysates then were prepared after 9 days of hygromycin selection as described previously [16]. The cells were washed in 1X PBS, scraped from plates in 1X PBS, and spun at 3,000 g for 5 minutes at 4°C. One volume of pelleted cells was lysed using two volumes of the following buffer: 1.5 mM KCl, 2.5 mM MgCl₂, 5 mM Tris-HCl, pH 7.5, 1% deoxycholic acid, 1% Triton X-100, 1X Complete Mini EDTA-free Protease Inhibitor Cocktail (Roche Applied Science). The cells were resuspended by gentle pipetting and incubated on ice for 10 minutes. The lysate was cleared by centrifugation at 3,000 g for 5 minutes at 4°C. Untransfected HeLa cell samples were obtained three days after plating. The

Bradford reagent (Bio-Rad) was used to determine the protein concentrations [67]. The same amount of total protein was separated by SDS-PAGE. BenchMark Pre-Stained Protein Ladder (Invitrogen) was used as a molecular weight marker. The proteins were detected by western blot using the following primary antibodies: mouse anti-T7-Tag (Novagen), rabbit anti-TAP (Open Biosystems), rat anti-HA (3F10 clone, Roche), mouse anti- α -tubulin (Sigma), rabbit anti-S6 (Cell Signaling Technology), rabbit anti-ORF1p (a generous gift from Thomas Fanning [50]) and rabbit α -ORF2p-N (a generous gift from John Goodier [42]). Goat anti-mouse, anti-rabbit and anti-rat HRP-conjugated secondary antibodies were purchased from GE/Amersham. Western blots were developed using either the pico or femto ECL substrate (Pierce) according to manufacturer's protocols.

RNA preparation and RT-PCR analysis

RNA isolation was performed with the RNeasy Kit (QIAGEN) coupled to an on-column DNase treatment (QIAGEN). Whole cell lysates (10–50 μ L) were used as starting material. The isolated RNAs were resuspended in Ultrapure distilled water (GIBCO) and quantified using a Nanodrop spectrophotometer (Thermo Scientific). For the LEAP assay controls, RNA was isolated from a 50 μ L RNP sample (1.5 mg/mL). RT-PCR was performed on 0.5 μ g total RNA, using the 3'RACE adapter primer (0.4 μ M) and M-MLV reverse transcriptase (200U) (Promega). The resultant cDNA products then were amplified by PCR using HotStart Pfu Turbo polymerase (Stratagene) with one primer specific to the transfected L1 constructs (L1 3' end) or GAPDH (GAPDH 3' end) and the 3'RACE outer primer, as described previously [17]. The PCR cycles were as follows: one cycle at 94°C for 3 minutes, then thirty five cycles of 94°C for 30 seconds, 58°C for 30 seconds and 72°C for 30 seconds. Then, a final extension was performed at 72°C for 10 minutes.

LEAP assay

The LEAP assay has been described previously [17]. Briefly, HeLa cells were plated at 6×10^6 cells/flask in T-175 flasks, and transfected within 24 hours with 30 μ g plasmid DNA (Midiprep Plasmid DNA Kit (QIAGEN)) using FuGene-6 transfection reagent (Roche). HeLa cells were grown in the presence of hygromycin from days 3 to 9 post-transfection (200 μ g/mL) to select for episome-containing cells. HeLa cells grown for three days in the absence of hygromycin served as an untransfected (naïve) control. On day 9, transfected cells and naïve HeLa cells were harvested, lysed, and the cleared whole cell lysates were centrifuged through an 8.5%/17% (w/v) sucrose cushion at 178,000 g for 2 hours. The resultant pellet was resuspended with 100 μ L dH₂O +1X Complete EDTA-free protease inhibitor cocktail (Roche). Bradford reagent (Bio-Rad) was used to determine protein concentration and this RNP sample was diluted to a final concentration of 1.5 mg/mL. An aliquot (1.5 μ g) of the RNP sample was added to 49 μ L of LEAP assay master mix (50 mM Tris-HCl (pH = 7.5), 50 mM KCl, 5 mM MgCl₂, 10 mM DTT, 0.4 μ M 3'RACE adapter primer, 20U RNasin (Promega), 0.2 mM dNTPs, and 0.05% (v/v) Tween 20) and was incubated at 37°C for 1 hour. LEAP cDNA products (1 μ L) were amplified in a standard 50 μ L PCR reaction containing 0.4 μ M of the 3'RACE outer primer and 0.4 μ M of one of the following forward primers: L1 3' end; Neo promoter sense; Neo8161S; LEAP-86; LEAP-46, using HotStart Pfu Turbo polymerase (Stratagene) according to the manufacturer's protocol (see Figure S2). The resultant products were visualized on 2% agarose gels. PCR products were isolated, cloned into the pCR-Blunt vector (Invitrogen), and sequenced to confirm their identity. The diffuse

profile of the amplification above 220 bp is explained by initiation of reverse transcription at many places on L1 poly (A). Lower bands, below 200 bp, are due to an internal initiation of reverse transcription 5' of the poly (A) tail [17].

Affinity purification

The affinity purification procedure described in Figure 3 was adapted from a published protocol [68]. To prepare the samples, HeLa cells were plated in T-175 flasks and transfected as described in the previous paragraph. Hygromycin selection on days 3 to 9 post-transfection was used to select for cells expressing the respective constructs. Using these conditions, one T-175 flask per plasmid was sufficient to yield enough cellular material (3 mg) for an experiment. Cells were washed, scraped from the flasks in 1X PBS, and centrifuged at 3,000 g for 5 minutes at 4°C. Cells were lysed by repeated pipetting with 3 volumes of IP FLAG buffer (0.1% NP-40, 100 mM KCl, 20 mM Tris-HCl pH 8, 1 mM DTT, 10% Glycerol, 1X complete EDTA free Protease inhibitor (Roche)) and incubated for 15 minutes on ice. The cellular debris was removed by centrifugation at 3,000 g for 5 minutes at 4°C. The protein concentration of the supernatant was quantified by a Bradford assay (Biorad Protein Assay).

For immunoprecipitation reactions, anti-FLAG beads (EZview Red ANTI-FLAG M2 Affinity Gel, Sigma) were equilibrated in 0.1M Glycine (pH 2.2) (5 µL for 100 µL of beads) for 5 minutes at room temperature. After addition of Tris-HCl (pH 8.0) (10 µL for 100 µL of beads), the beads were spun down for 3 minutes at 3000 rpm and then washed 3 times with IP FLAG buffer (mentioned above). For each condition, 3 mg of protein extract (input) was then incubated on a rotating wheel overnight at 4°C with 20 µL of the pre-equilibrated anti-FLAG beads. The next day, the beads were washed 5 times with 1 mL of IP FLAG Buffer for 10 minutes at 4°C. The beads were incubated 1 hour (at 4°C on the wheel) with 200 µL of IP FLAG buffer containing 200 µg/mL of 3X FLAG peptide (Sigma). The elution fraction then was collected and analyzed alongside the corresponding input fraction by western blotting (as described above in the dedicated section). The femto ECL substrate (Pierce) was used in the detection of both T7-tagged ORF1p and FLAG-HA-tagged ORF2p in this experiment. An aliquot (1 µL) of the input and elution samples then were used to perform the LEAP assay (see previous section for detailed protocol).

Quantitative real-time PCR

Quantitative PCR was performed on LEAP cDNA samples or M-MLV RT-PCR products using the 7300 Real Time PCR system (Applied Biosystems). For analysis, 1 µL of LEAP or M-MLV RT products was added to 19 µL of master mix (1X SYBR Green PCR Master Mix (Applied Biosystems), 500 nM L1 3' end primer, and 500 nM L1 Reverse primer), and amplified in a standard Q-PCR run of 45 cycles. The average cycle threshold (Ct) value for each experimental or control sample was calculated from three independent reactions within a Q-PCR run. The 'absolute quantitation by standard curve' method was used to determine the number of cDNA molecules in each LEAP RNP or RNA sample. A standard curve was generated using dilutions of a L1 LEAP product cloned into a plasmid, and a best fit line (log(molecules) versus average Ct value) for these standards was generated by linear regression. For each wild-type or mutant L1, RNA levels from three independent RNP samples were examined by at least one RT reaction and two Q-PCR runs. The level of LEAP activity in each wild-type or mutant L1 was determined from four independent RNP samples. These RNP samples were characterized by at least one and up to three independent LEAP

RT reactions and one or two independent Q-PCR runs. For LEAP activity, the negative control RT- (pDK135) gave a background amplification level of ~15–30 molecules of cDNA due to the presence of the transfected L1 plasmid in the RNP sample. This RT- background control was included in each Q-PCR run and the background amount of molecules was subtracted from each experimental sample in Figure S2 and when calculating fold changes.

Fluorescent In Situ Hybridization (FISH) and immunofluorescence

U-2 OS cells were plated at 10^5 onto sterile glass cover slips in 6 well tissue culture dishes. The following day, cells were transfected using 1 µg of purified plasmid DNA (Midiprep Plasmid DNA Kit, QIAGEN) and 3 µL of FuGene-6 Transfection Reagent (Roche Applied Science). The FISH protocol was adapted from the Robert Singer (Albert Einstein College of Medicine, New York) lab protocol (available at <http://www.singerlab.org/protocols>) and was modified to allow protein detection by immunofluorescence. Briefly, 48 h post-transfection, cells were washed twice with 1X PBS and fixed with 4% paraformaldehyde in 1X PBS for 10 minutes at room temperature. The fixed cells then were washed 2 additional times with 1X PBS. The fixed cells were permeabilized by treatment with 70% ethanol overnight at 4°C. The following day, cells were rehydrated with 1X saline-sodium citrate (SSC) and 10% formamide for 5 minutes at room temperature. To prepare the hybridization solution, a first mix containing 40 µg of *E.coli* tRNA (Sigma), 1X SSC, 10% formamide, and 7.5 ng of MS2-Cy3 probe (generous gift from Dr. Edouard Bertrand) was boiled for 1 minute at 100°C in order to denaturize the probe. The quantities of probe and tRNA are indicated for hybridization of one slide. A second mix was prepared with 10% dextran sulfate, 2 mM vanadyl-ribonucleoside complex (Sigma), and 0.02% RNase free BSA (Roche Applied Science). After probe denaturation, mixes 1 and 2 were combined to form the final hybridization solution. The re-hydrated cells were hybridized overnight at 37°C in 30 µL of this hybridization solution. Cells were then washed twice for 30 minutes at 37°C with 1X SSC, 10% formamide and 3% BSA and then were incubated with primary antibodies for 1 hour at 37°C. The cells were washed three times with 1X PBS and were incubated with secondary antibodies and 0.2 µg/mL 4',6'-diamidino-2-phenylindole (DAPI, Molecular Probes) for 30 minutes at 37°C and washed three times with 1X PBS. The primary and secondary antibodies were diluted in 1X PBS and 3% BSA and are as follows: anti-T7 (Novagen), anti-TAP (Open Biosystems), anti-HA (Roche), rabbit anti-ORF1p (a generous gift from Thomas Fanning) [50] and rabbit α -ORF2p-N (a generous gift from John Goodier) [42], anti-eIF3 (Santa Cruz BioTechnology), anti-G3BP (generous gift from Jamal Tazi), Alexa Fluor 488 anti-mouse and anti-rabbit (Invitrogen), Alexa Fluor 546 anti-mouse and anti-rabbit (Invitrogen), Cy3-conjugated anti-rat (Jackson Immuno Research) and Cy5-conjugated anti-mouse and anti-rabbit (Jackson Immuno Research). Cells were rinsed with water and mounted on slides with Vectashield (Vector Laboratories). Samples were then analyzed with appropriate fluorescent filters on DMRXA Leica microscope and images were captured using a Zeiss LSM510 META confocal microscope.

The above protocol was used for both RNA and protein detection analyses. In experiments where we only sought to detect RNA, the protocol was stopped after hybridization with the MS2-Cy3 probe and subsequent washes. Cover slides were stained with DAPI and mounted on slides as described above. In experiments where we only sought to detect protein, fixed cells were permeabilized by treatment with anhydrous methanol for 1

minute. After three washes with 1X PBS, the cells were incubated with 3% BSA in 1X PBS for 30 minutes. Antibody incubation and DAPI staining were performed as described above. We verified that the protein A domain contained in the TAP tag of ORF2p did not react with the secondary antibodies (data not shown).

Analysis of L1 cytoplasmic foci

In general, L1 cytoplasmic foci formation was measured 48 hours post-transfection. At least two independent series of slides were analyzed. Each analysis corresponds to 100 transfected cells that were quantified in a blinded manner. Cells in which we were able to distinguish a concentrated cytoplasmic signal from a diffuse cytoplasmic signal using a 63x or 100x objective (equivalent of 1 micrometer of diameter) were considered as L1 cytoplasmic foci.

Supporting Information

Figure S1 Retrotransposition assays with mutant L1 constructs. A. Retrotransposition assays with mutant L1 constructs: 2×10^4 HeLa cells were transfected with the indicated constructs. pJM101/L1.3 and pAD2TE1 were used as positive controls. All of the pAD-based constructs contain the ORF1p T7 epitope tag and the ORF2p TAP-tag except for pAD500, which lacks ORF1. pAD135 is an RT mutant (D702A), and serves as a negative control. B. Retrotransposition assay with leucine zipper domain mutants: 2×10^5 HeLa cells were transfected with the indicated pDK101-derived constructs. T7WT (pDK101) is a wild-type L1 (L1.3) that contains the T7 epitope tag on the carboxyl terminus of ORF1p. pDK101 was modified to create LZ1/2 (L93V, L100V), LZ2/3 (L100V, L107V), LZ3/4 (L107V, L114V), and LZC (L93V, L100V, L107V, L114V). Each of the mutations abolished L1 retrotransposition.

Found at: doi:10.1371/journal.pgen.1001150.s001 (0.57 MB TIF)

Figure S2 The effect of ORF1p mutations on LEAP activity. A. Results of western blot analyses: RNPs derived from wild-type (pDK101) and the indicated mutant constructs were subjected to western blot analyses with an anti-T7 antibody (α T7). An ~ 40 kDa band indicative of epitope-tagged ORF1p is shown. Untransfected HeLa cells served as a negative control. The ribosomal S6 protein was detected using an anti-S6 (α S6) antibody (bottom panel; RNP prep control) and served as a loading control. Molecular weight markers (Invitrogen) are indicated at the left of the gel. B. Results of LEAP assays: Top panel: An aliquot of the above RNPs was used to measure LEAP activity. RNPs derived from wild-type (pDK101) generate strong LEAP products of ~ 220 –400 bp and served as a positive control. Untransfected HeLa cells and an RT mutant (pDK135; D702A) serve as negative controls. LEAP products generated in ORF1p mutant RNPs are shown. Reactions conducted without template (No Template) or without RNPs (No RNP/No RNA) were used as negative controls. Middle and bottom panels: RT-PCR with M-MLV RT and primers specific to either the transfected L1 constructs or GAPDH confirmed the presence of L1 RNA in RNPs and the integrity of the RNA isolation procedure. DNA size markers (Invitrogen) are indicated at the left of the gel. C. LEAP products derived from the ORF1p RNA binding mutant (RR261-262AA) and putative chaperone mutant (RR261-262KK): Representative LEAP products derived from wild-type (pDK101), an ORF1p RNA binding mutant (pDK105; RR261-262AA), and a putative ORF1p nucleic acid chaperone activity mutant (pDK106; RR261-262KK) are depicted at the top gel. The middle and bottom gels are RT-PCR reactions conducted with M-MLV RT and primers specific to L1 and GAPDH transcripts, respectively. The black arrow on the

middle gel indicates the size of the specific L1 cDNA amplification products. Untransfected HeLa cells and a RT mutant (pDK135) served as negative controls. Additional negative controls include reactions conducted without template (No Template) as well as reactions conducted without RNPs or RNA (No RNP/No RNA). DNA size markers (Invitrogen) are indicated at the left of the gel. D. LZC RNPs have decreased reverse transcriptase activity: Top panel: RNPs derived from T7WT (pDK101) generate strong LEAP products of ~ 220 –400 bp. By comparison, LZC (see Figure S1B) had a less intense band at ~ 220 –400 bp. LEAP products also were seen in the RR261-262AA (pDK105) mutant and the LZC/RR261-262AA mutants. No product was seen in untransfected HeLa cells or for a L1 containing a RT active site mutation (pDK135; RT-). Middle panels: RT-PCR with M-MLV RT and primers specific to either the transfected L1 constructs or GAPDH confirmed the presence of L1 RNA in RNPs and the integrity of the RNA isolation procedure. No RNP/RT and dH₂O served as negative controls. DNA size markers (Invitrogen) are shown at the left side of the gel. Bottom panels: Western blot against the T7 epitope tag detects ORF1p (α T7). Untransfected HeLa cells served as a negative control. Western blot against ribosomal protein S6 was used as a loading control. Western blot against ribosomal protein S6 was used as a loading control (α S6). Molecular weight markers (Invitrogen) are indicated at the left of the blot. E. Quantitative PCR of LEAP cDNAs from ORF1p LZC and carboxyl-terminal nucleic acid binding domain mutants: Representative results of a Q-PCR run are shown. Standard deviations are indicated on the graph. Q-PCR was performed on four independent RNP preps for T7WT (pDK101), LZC, RR261-262AA (pDK105), and the LZC/RR261-262AA double mutant. Three of four preps showed a 5–7 fold decrease in LZC LEAP activity compared to wild-type. One RNP prep showed a 16–23 fold decrease in LZC LEAP activity compared to wild-type. F. LZC mutation does not appear to affect reverse transcriptase elongation: Top panel: PCR was performed on LEAP cDNAs using different primers pairs. The names of the primers and the approximate size of each product are indicated in the cartoon above the gel. LZC RNPs yield fewer products when compared to T7WT (pDK101) RNPs with all tested primer sets. Bottom panel: RT-PCR with M-MLV RT confirms the presence of L1 RNA in the RNP samples. DNA size markers (Invitrogen) are indicated at the left of the gel. The L1 constructs in all panels contain the mneoI retrotransposition indicator cassette.

Found at: doi:10.1371/journal.pgen.1001150.s002 (1.42 MB TIF)

Figure S3 L1 retrotransposition and L1 cytoplasmic foci formation in U-2 OS cells. A. L1 retrotransposition assays: 2×10^4 cells were transfected with the indicated L1 constructs. pJM101/L1.3 and pAD2TE1 were used as positive controls. Untransfected cells and pAD135 (RT mutant (D702A)) serve as negative controls. A cartoon of each L1 is shown above the tissue culture dishes. Green rectangle = ORF1; Red rectangle = ORF2. The relative positions of the T7 and TAP tags also are indicated. All constructs contain the mneoI retrotransposition indicator cassette. B. Time course analyses of L1 cytoplasmic foci formation: Cells were transfected with pAD2TE1. X-axis = time after transfection. Y-axis = percentage of transfected cells containing L1 cytoplasmic foci. For each time point, 100 transfected cells were analyzed for the presence of ORF1p and ORF2p in L1 cytoplasmic foci. Error bars = standard deviation ($n = 3$). C. Cytoplasmic localization of L1 proteins and RNA: Top panels: Cells were transfected with pAD3TE1. ORF2p was visualized with an anti-ORF2 antibody (α ORF2, green). ORF1p was visualized with an anti-T7 antibody (α T7, blue). L1 RNA was visualized with an MS2-Cy3 FISH probe (red). A merged image is shown in the rightmost column; DAPI (grey) was used to stain nuclear DNA. Middle panels: Cells were transfected with

pES2TE1. ORF2p was visualized with an anti-ORF2 antibody (α ORF2, green) and an anti-HA antibody (α HA, red). ORF1p was visualized with an anti-T7 antibody (α T7, blue). A merged image is shown in the rightmost column; DAPI (grey) was used to stain nuclear DNA. Bottom panels: Cells were transfected with pES2TE1. ORF1p was visualized with an anti-ORF1 antibody (α ORF1, green) and an anti-T7 antibody (α T7, blue). ORF2p was visualized with an anti-HA antibody (red). A merged image is shown in the rightmost column; DAPI (grey) was used to stain nuclear DNA. D. Localization of L1 RNA: Top panels: Cells were transfected with pADL1MT. L1 RNA was visualized with an MS2-Cy3 FISH probe (MS2-Cy3, red). DAPI was used to stain nuclear DNA (gray). A merged image is shown in the rightmost column. Middle panels: Cells were co-transfected with pAD3TE1 and pMS2-GFP-nls. Fluorescence was used to visualize MS2-GFP-nls (green). L1 RNA was visualized with an MS2-Cy3 FISH probe (MS2-Cy3, red). ORF1p was visualized with an anti-T7 antibody (α T7, blue). A merged image is shown in the rightmost column. Bottom panels: Cells were co-transfected with pAD3TE1 and pMS2-CFP. Fluorescence was used to visualize MS2-CFP (green). L1 RNA was visualized with an MS2-Cy3 FISH probe (MS2-Cy3, red). ORF2p was visualized with an anti-TAP antibody (α TAP, blue). A merged image is shown in the rightmost column. The use of MS2 binding protein system confirms the cytoplasmic localization of the L1 RNA observed by FISH as well as the colocalization of L1 RNA with ORF1p and ORF2p. The L1

constructs in all panels contain the mneoI retrotransposition indicator cassette.

Found at: doi:10.1371/journal.pgen.1001150.s003 (2.01 MB TIF)

Acknowledgments

We thank Christine Beck, José Luis Garcia Perez, Billy Giblin, Sandra Richardson, Séverine Chambeyron, Alain Pélisson, and Oliver Weichenrieder for helpful comments during the course of this study. We thank Edouard Bertrand, Thomas Fanning, John Goodier, Greg Hannon, and Jamal Tazi for generously providing reagents. We thank the IGH and University of Michigan Sequencing Cores for help with DNA sequencing. We are grateful to Eugenia Basyuk (FISH), Christelle Cayrou, Bijan Sobhian (purification protocols), and Séverine Chambeyron (fluorescence analysis) for assistance in designing/implementing experiments. We thank Nicole Lautredou (Montpellier RIO Imaging) and Chris Edwards (Microscopy and Image-analysis Laboratory at the University of Michigan, Department of Cell and Developmental Biology) for assistance with confocal microscopy.

Author Contributions

Conceived and designed the experiments: AJD AEH ES DAK JBM HCK JNA JVM NG. Performed the experiments: AJD AEH ES DAK JBM HCK JNA MH NG. Analyzed the data: AJD AEH ES DAK JBM HCK JNA JVM NG. Contributed reagents/materials/analysis tools: AB JVM NG. Wrote the paper: AJD AEH JVM NG.

References

- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860–921.
- Grimaldi G, Singer MF (1983) Members of the KpnI family of long interspersed repeated sequences join and interrupt alpha-satellite in the monkey genome. *Nucleic Acids Res* 11: 321–338.
- Kazazian HH, Jr., Moran JV (1998) The impact of L1 retrotransposons on the human genome. *Nat Genet* 19: 19–24.
- Sassaman DM, Dombroski BA, Moran JV, Kimberland ML, Naas TP, et al. (1997) Many human L1 elements are capable of retrotransposition. *Nat Genet* 16: 37–43.
- Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, et al. (2003) Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A* 100: 5280–5285.
- Babushok DV, Kazazian HH, Jr. (2007) Progress in understanding the biology of the human mutagen LINE-1. *Hum Mutat* 28: 527–539.
- Goodier JL, Kazazian HH, Jr. (2008) Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell* 135: 23–35.
- Cordaux R, Batzer MA (2009) The impact of retrotransposons on human genome evolution. *Nat Rev Genet* 10: 691–703.
- Swergold GD (1990) Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol Cell Biol* 10: 6718–6729.
- Athanikar JN, Badge RM, Moran JV (2004) A YY1-binding site is required for accurate human LINE-1 transcription initiation. *Nucleic Acids Res* 32: 3846–3855.
- Lavie L, Maldener E, Brouha B, Meese EU, Mayer J (2004) The human L1 promoter: variable transcription initiation sites and a major impact of upstream flanking sequence on promoter activity. *Genome Res* 14: 2253–2260.
- Scott AF, Schmeckpeper BJ, Abdelrazik M, Comey CT, O'Hara B, et al. (1987) Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. *Genomics* 1: 113–125.
- Dombroski BA, Mathias SL, Nanthakumar E, Scott AF, Kazazian HH, Jr. (1991) Isolation of an active human transposable element. *Science* 254: 1805–1808.
- Esnault C, Maestre J, Heidmann T (2000) Human LINE retrotransposons generate processed pseudogenes. *Nat Genet* 24: 363–367.
- Hohjoh H, Singer MF (1996) Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *Embo J* 15: 630–639.
- Kulpa DA, Moran JV (2005) Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. *Hum Mol Genet* 14: 3237–3248.
- Kulpa DA, Moran JV (2006) Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat Struct Mol Biol* 13: 655–660.
- Martin SL (1991) Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells. *Mol Cell Biol* 11: 4804–4807.
- Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, et al. (2001) Human L1 retrotransposition: cis preference versus trans complementation. *Mol Cell Biol* 21: 1429–1439.
- Luan DD, Korman MH, Jakubczak JL, Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* 72: 595–605.
- Kubo S, Selem MC, Soifer HS, Perez JL, Moran JV, et al. (2006) L1 retrotransposition in nondividing and primary human somatic cells. *Proc Natl Acad Sci U S A* 103: 8036–8041.
- Feng Q, Moran JV, Kazazian HH, Jr., Boeke JD (1996) Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 87: 905–916.
- Cost GJ, Boeke JD (1998) Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. *Biochemistry* 37: 18081–18093.
- Furano AV (2000) The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. *Prog Nucleic Acid Res Mol Biol* 64: 255–294.
- Kolosha VO, Martin SL (1997) In vitro properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. *Proc Natl Acad Sci U S A* 94: 10155–10160.
- Martin SL, Li J, Weisz JA (2000) Deletion analysis defines distinct functional domains for protein-protein and nucleic acid interactions in the ORF1 protein of mouse LINE-1. *J Mol Biol* 304: 11–20.
- Moran JV, Gilbert N (2002) Mammalian LINE-1 retrotransposons and related elements. In: Craig N, Craggie R, Gellert M, Lambowitz A, eds. *Mobile DNA II*. Washington, DC: ASM Press. pp 836–869.
- Duvernell DD, Turner BJ (1998) Swimmer 1, a new low-copy-number LINE family in teleost genomes with sequence similarity to mammalian L1. *Mol Biol Evol* 15: 1791–1793.
- Furano AV, Duvernell DD, Boissinot S (2004) L1 (LINE-1) retrotransposon diversity differs dramatically between mammals and fish. *Trends Genet* 20: 9–14.
- Khazina E, Weichenrieder O (2009) Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. *Proc Natl Acad Sci U S A* 106: 731–736.
- Goodier JL, Zhang L, Vetter MR, Kazazian HH, Jr. (2007) LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. *Mol Cell Biol* 27: 6469–6483.
- Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, et al. (1996) High frequency retrotransposition in cultured mammalian cells. *Cell* 87: 917–927.
- Basame S, Wai-Lun Li P, Howard G, Branciforte D, Keller D, et al. (2006) Spatial Assembly and RNA Binding Stoichiometry of a LINE-1 Protein Essential for Retrotransposition. *J Mol Biol* 357: 351–357.
- Martin SL, Branciforte D, Keller D, Bain DL (2003) Trimeric structure for an essential protein in L1 retrotransposition. *Proc Natl Acad Sci U S A* 100: 13815–13820.

35. Martin SL, Cruceanu M, Branciforte D, Wai-Lun Li P, Kwok SC, et al. (2005) LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. *J Mol Biol* 348: 549–561.
36. McMillan JP, Singer MF (1993) Translation of the human LINE-1 element, L1Hs. *Proc Natl Acad Sci U S A* 90: 11533–11537.
37. Alisch RS, Garcia-Perez JL, Muotri AR, Gage FH, Moran JV (2006) Unconventional translation of mammalian LINE-1 retrotransposons. *Genes Dev* 20: 210–224.
38. Li PW, Li J, Timmerman SL, Krushel LA, Martin SL (2006) The dicistronic RNA from the mouse LINE-1 retrotransposon contains an internal ribosome entry site upstream of each ORF: implications for retrotransposition. *Nucleic Acids Res* 34: 853–864.
39. Dmitriev SE, Andreev DE, Terenin IM, Olovnikov IA, Prassolov VS, et al. (2007) Efficient translation initiation directed by the 900-nucleotide-long and GC-rich 5' untranslated region of the human retrotransposon LINE-1 mRNA is strictly cap dependent rather than internal ribosome entry site mediated. *Mol Cell Biol* 27: 4685–4697.
40. Mathias SL, Scott AF, Kazazian HH, Jr., Boeke JD, Gabriel A (1991) Reverse transcriptase encoded by a human transposable element. *Science* 254: 1808–1810.
41. Fanning T, Singer M (1987) The LINE-1 DNA sequences in four mammalian orders predict proteins that conserve homologies to retrovirus proteins. *Nucleic Acids Res* 15: 2251–2260.
42. Goodier JL, Ostertag EM, Engleka KA, Seleme MC, Kazazian HH, Jr. (2004) A potential role for the nucleolus in L1 retrotransposition. *Hum Mol Genet* 13: 1041–1048.
43. Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, et al. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* 17: 1030–1032.
44. Nakatani Y, Ogryzko V (2003) Immunoaffinity purification of mammalian protein complexes. *Methods Enzymol* 370: 430–444.
45. Fouts DE, True HL, Celander DW (1997) Functional recognition of fragmented operator sites by R17/MS2 coat protein, a translational repressor. *Nucleic Acids Res* 25: 4464–4473.
46. Valegard K, Murray JB, Stonehouse NJ, van den Worm S, Stockley PG, et al. (1997) The three-dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein-RNA interactions. *J Mol Biol* 270: 724–738.
47. Freeman JD, Goodchild NL, Mager DL (1994) A modified indicator gene for selection of retrotransposition events in mammalian cells. *Biotechniques* 17: 46, 48–49, 52.
48. Wei W, Morrish TA, Alisch RS, Moran JV (2000) A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events. *Anal Biochem* 284: 435–438.
49. Holmes SE, Singer MF, Swergold GD (1992) Studies on p40, the leucine zipper motif-containing protein encoded by the first open reading frame of an active human LINE-1 transposable element. *J Biol Chem* 267: 19765–19768.
50. Leibold DM, Swergold GD, Singer MF, Thayer RE, Dombroski BA, et al. (1990) Translation of LINE-1 DNA elements in vitro and in human cells. *Proc Natl Acad Sci U S A* 87: 6990–6994.
51. Rangwala SH, Kazazian HH, Jr. (2009) The L1 retrotransposition assay: a retrospective and toolkit. *Methods* 49: 219–226.
52. Kedersha N, Stoecklin G, Ayodele M, Yacono P, Lykke-Andersen J, et al. (2005) Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J Cell Biol* 169: 871–884.
53. Tourriere H, Chebli K, Zekri L, Courselaud B, Blanchard JM, et al. (2003) The RasGAP-associated endoribonuclease G3BP assembles stress granules. *J Cell Biol* 160: 823–831.
54. Liu J, Valencia-Sanchez MA, Hannon GJ, Parker R (2005) MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* 7: 719–723.
55. Kedersha N, Chen S, Gilks N, Li W, Miller IJ, et al. (2002) Evidence that ternary complex (eIF2-GTP-tRNAⁱ(Met))-deficient preinitiation complexes are core constituents of mammalian stress granules. *Mol Biol Cell* 13: 195–210.
56. Garcia-Perez JL, Doucet AJ, Bucheton A, Moran JV, Gilbert N (2007) Distinct mechanisms for trans-mediated mobilization of cellular RNAs by the LINE-1 reverse transcriptase. *Genome Res*.
57. Ergun S, Buschmann C, Heukeshoven J, Dammann K, Schnieders F, et al. (2004) Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. *J Biol Chem* 279: 27753–27763.
58. Morrish TA, Gilbert N, Myers JS, Vincent BJ, Stamato TD, et al. (2002) DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat Genet* 31: 159–165.
59. Goodier JL, Mandal PK, Zhang L, Kazazian HH, Jr. (2010) Discrete subcellular partitioning of human retrotransposon RNAs despite a common mechanism of genome insertion. *Hum Mol Genet* 19: 1712–1725.
60. Beliakova-Bethell N, Beckham C, Giddings TH, Jr., Winey M, Parker R, et al. (2006) Virus-like particles of the Ty3 retrotransposon assemble in association with P-body components. *RNA* 12: 94–101.
61. Checkley MA, Nagashima K, Lockett SJ, Nyswaner KM, Garfinkel DJ (2010) P-body components are required for Ty1 retrotransposition during assembly of retrotransposition-competent virus-like particles. *Mol Cell Biol* 30: 382–398.
62. Dutko JA, Kenny AE, Gamache ER, Curcio MJ (2010) 5' to 3' mRNA decay factors colocalize with Ty1 gag and human APOBEC3G and promote Ty1 retrotransposition. *J Virol* 84: 5052–5066.
63. Dombroski BA, Scott AF, Kazazian HH, Jr. (1993) Two additional potential retrotransposons isolated from a human L1 subfamily that contains an active retrotransposable element. *Proc Natl Acad Sci U S A* 90: 6513–6517.
64. Boireau S, Maiuri P, Basyuk E, de la Mata M, Knezevich A, et al. (2007) The transcriptional cycle of HIV-1 in real-time and live cells. *J Cell Biol* 179: 291–304.
65. Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, et al. (1998) Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 2: 437–445.
66. Fusco D, Accornero N, Lavoie B, Shenoy SM, Blanchard JM, et al. (2003) Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr Biol* 13: 161–167.
67. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248–254.
68. Qi HH, Ongusaha PP, Myllyharju J, Cheng D, Pakkanen O, et al. (2008) Prolyl 4-hydroxylation regulates Argonaute 2 stability. *Nature* 455: 421–424.