Affinity Proteomics Reveals Human Host Factors Implicated in Discrete Stages of LINE-1 Retrotransposition

Martin S. Taylor,[1,7] John LaCava,[2,7] Paolo Mita,[3] Kelly R. Molloy,[4] Cheng Ran Lisa Huang,[3,5] Donghui Li,[3,5] Emily M. Adney,[3,5] Hua Jiang,[2] Kathleen H. Burns,[5,6] Brian T. Chait,[4] Michael P. Rout,[2] Jef D. Boeke,[3,*] and Lixin Dai[3,7,*]

[1]High Throughput Biology Center and Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
[2]Laboratory of Cellular and Structural Biology, The Rockefeller University, New York, NY 10021, USA
[3]High Throughput Biology Center and Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
[4]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, NY 10021, USA
[5]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
[6]Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
[7]These authors contributed equally to this work
*Correspondence: jboeke@jhmi.edu (J.D.B.), lxdai@jhmi.edu (L.D.)
http://dx.doi.org/10.1016/j.cell.2013.10.021

SUMMARY

LINE-1s are active human DNA parasites that are agents of genome dynamics in evolution and disease. These streamlined elements require host factors to complete their life cycles, whereas hosts have developed mechanisms to combat retrotransposition's mutagenic effects. As such, endogenous L1 expression levels are extremely low, creating a roadblock for detailed interactomic analyses. Here, we describe a system to express and purify highly active L1 RNP complexes from human suspension cell culture and characterize the copurified proteome, identifying 37 high-confidence candidate interactors. These data sets include known interactors PABPC1 and MOV10 and, with in-cell imaging studies, suggest existence of at least three types of compositionally and functionally distinct L1 RNPs. Among the findings, UPF1, a key nonsense-mediated decay factor, and PCNA, the polymerase-delta-associated sliding DNA clamp, were identified and validated. PCNA interacts with ORF2p via a PIP box motif; mechanistic studies suggest that this occurs during or immediately after target-primed reverse transcription.

INTRODUCTION

Long interspersed element-1s (L1s or LINE-1s) are autonomous retrotransposons that continuously re-enter host genomes (Ostertag and Kazazian, 2001). Although most L1 copies are functionally inactive, ~90 retrotransposition-competent L1s inhabit the human genome (Brouha et al., 2003). Somatic L1 insertions are rare in most tissues but are found in brain, testis, and many cancer types and have been implicated in carcinogenesis (Shukla et al., 2013). Most L1 insertions are detrimental to the host; thus, retrotransposition is tightly controlled by a network of host factors (Arjan-Odedra et al., 2012; Dai et al., 2012; Goodier et al., 2012; Niewiadomska et al., 2007; Peddigari et al., 2013; Suzuki et al., 2009). L1s are extinct in certain lineages (Khan et al., 2006; Malik et al., 1999) and may contribute to neural plasticity in others (Muotri et al., 2005).

The full-length L1 transcript is ~6 kb long, is bicistronic, and includes its promoter sequence within its 5′ UTR (Ostertag and Kazazian, 2001). The L1 messenger RNA (mRNA) encodes two nonoverlapping proteins, ORF1p and ORF2p, and includes a short 3′ UTR and a poly(A) tail (Belancio et al., 2007). L1 has a complex life cycle, beginning with transcription by RNA polymerase II (pol II). The 40 kDa ORF1p forms a homotrimeric protein with nucleic acid chaperone activities (Martin and Bushman, 2001). The ~150 kDa ORF2p is multifunctional with endonuclease (EN) (Feng et al., 1996) and reverse transcriptase (RT) activities (Mathias et al., 1991). ORF1p is efficiently translated, but ORF2p translation occurs at low levels through an unconventional mechanism (Alisch et al., 2006), making mechanistic research on ORF2p challenging. ORF2p exhibits preferential interaction with the L1 mRNA encoding it (cis-preference), forming a ribonucleoprotein (RNP) particle presumed to constitute a direct retrotransposition intermediate (Kulpa and Moran, 2006; Wei et al., 2001). L1 likely transposes via target-primed reverse transcription (TPRT) (Cost et al., 2002; Luan and Eickbush, 1995).

Being parasitic, L1 requires host factors to complete its life cycle (Beauregard et al., 2008). For example, several tissue-specific transcription factors promote L1 transcription by host-encoded pol II (e.g., Athanikar et al., 2004). Poly(A) binding proteins are important in RNP formation and retrotransposition

(Dai et al., 2012), and DNA repair factors affect L1 retrotransposition, most likely in TPRT (Suzuki et al., 2009). Host genomes have evolved mechanisms to restrict potentially harmful L1 retrotransposition, employing DNA methylation (Hata and Sakaki, 1997), RNAi pathways (Soifer et al., 2005; Yang and Kazazian, 2006), APOBEC3 cytidine deaminases (Niewiadomska et al., 2007), and RNA helicase MOV10 (Arjan-Odedra et al., 2012; Goodier et al., 2012). Two recent studies identified protein and RNA interactors in L1 RNP complexes via an ORF1 tag (Goodier et al., 2013; Mandal et al., 2013). However, although host factors have been found, technical limitations have hindered isolation of pure, active RNP complexes in analytically tractable quantities.

Although L1 is highly active in germline and brain, it is effectively repressed in most somatic cells (Muotri et al., 2005). To evaluate L1 activity in cell culture, it is often paired with a constitutive, high-level promoter (Moran et al., 1996). Fractions containing expressed L1 RNP complexes have been prepared by sucrose cushion velocity sedimentation of cell lysates (Kulpa and Moran, 2005, 2006). These provided valuable insights into mechanism but suffer from low purity, limiting their usefulness for studying L1-host interactomes. Traditional epitope tagging of ORF polypeptides has been used to study proteins copurifying with the L1 RNP (Doucet et al., 2010; Goodier et al., 2012, 2013), but these conferred insufficient yield and purity for comprehensive proteomics with high specific activity.

In contrast, new methods implementing cryomilling/rapid affinity capture provide outstanding purification that significantly improves interactomic surveys (Cristea et al., 2005; Domanski et al., 2012; Oeffinger et al., 2007). Intact cells are flash-frozen in liquid nitrogen and are milled in solid phase (cryomilling). Native intermolecular interactions are preserved during milling, producing a fine powder well-suited to subsequent affinity isolation of protein complexes (pullouts). Micron-scale antibody-conjugated magnetic beads enable rapid purifications with binding times as short as 5 min and facilitate binding of large complexes excluded from pores of traditional matrices (Oeffinger et al., 2007). When coupled with Isotopic Differentiation of Interactions as Random or Targeted (I-DIRT), stable interactions formed in vivo are distinguished from postextraction artifacts (Tackett et al., 2005). These techniques have been applied to the study of protein complexes from many systems (Di Virgilio et al., 2013; Domanski et al., 2012; Lee et al., 2008; Oeffinger et al., 2007).

In this study, we constructed a series of inducible L1 expression vectors containing either synthetic (ORFeus-Hs) (An et al., 2011) or native (L1RP) L1 (Kimberland et al., 1999) elements and screened epitope tags on both L1 ORFs for effects on retrotransposition. Active constructs were expressed in suspension culture at a large scale, and RNPs were purified using cryomilling and rapid capture, facilitating enzymatic, transcriptomic, and proteomic assays. Purified RNPs had ∼70-fold higher specific activity than those made by previous techniques. RNA sequencing (RNA-seq) revealed L1 RNA as an abundant component along with U6 small nuclear RNA (snRNA), commonly observed as part of hybrid elements with L1 (Buzdin et al., 2002). The ORF1p to ORF2p ratio was unexpectedly low in tandem purified L1 particles. Cell imaging revealed that most cells fail to express ORF2p. Mass spectrometric (MS) characterization identified 37 high-confidence interactors, including known interactors PABPC1 and MOV10, and revealed that ORF1 and ORF2 RNPs exhibit both overlapping and distinct interactomes. UPF1, a helicase critical for nonsense-mediated decay (NMD), and PCNA, the polymerase-associated sliding clamp, were further validated for roles in the L1 life cycle.

RESULTS

Producing Highly Purified L1-RNPs

To isolate L1 RNP complexes at high yield and purity, we used synthetic human L1 ORFeus-Hs, which has high activity and produces > 40-fold more RNA and 5- to 10-fold more ORF1p than a native counterpart, while encoding identical proteins (An et al., 2011). Uncontrolled L1 expression slows or eliminates growth; therefore, we leveraged expression from a tetracycline-regulated mini-CMV promoter (pTRE, Figure 1A) (O'Donnell et al., 2013). Due to a lack of appropriate antibodies, especially for ORF2p, we designed many epitope tag combinations and measured retrotransposition efficiency relative to untagged ORFeus-Hs (Figure 1B, Supplemental Information, and Tables S1, S2, and S3). Constructs pLD288 (ORF1p-Flag) and pLD401 (ORF2p-3xFlag) were best, as they combined high L1 activity with short tag sequences.

We introduced these constructs into adherent Tet-On HEK293T$_{LD}$ (Dai et al., 2012) and selected transfected cells for 1 week or more with puromycin. Doxycycline (Dox) concentration and times had similar effects in monolayer and suspension cultures (Figures 1D–1F); expression of both ORFs peaked 24 hr postinduction. Although ORF1p levels were relatively stable over the next 3 days, ORF2p levels decreased within 48 hr (Figures 1F, 2H, and S2).

Multiliter suspension cultures were induced, harvested, and cryomilled, and pullouts were optimized (Figures 2A–2D and S1 available online). Both L1 proteins were readily observed by Coomassie brilliant blue (CBB) staining (Figures 2E and 2F). We also tagged native L1RP, which has lower protein expression levels (Figure 1C). Surprisingly, the C-terminal Flag and 3xFlag tag combinations on ORF1p and ORF2p that worked well in ORFeus-Hs were poorly tolerated in L1RP. We designed and screened new tags and linkers; a single Myc-tag at the C terminus of ORF1p was tolerated (pMT293 and pMT303) but did not provide high-quality affinity isolations. By shortening the linker on ORF2p-3xFLAG, we preserved L1RP retrotransposition efficiency (pMT302 and pMT303). An improved αORF1p antibody obviated need for an epitope-tagged L1RP ORF1p, and thus, pMT302 (L1RP with ORF2p-3xFlag) was selected (Figures 2E and 2F).

Both synthetic (pLD401) and native (pMT302) constructs gave robust ORF2p expression (Figures 2G and 2H). Taking into account the different detection sensitivities exhibited by the two tags (Hernan et al., 2000), ORF1p is apparently expressed 1,000- to 10,000-fold higher than ORF2p (Figures 2G and S2). Comparing ORFeus-Hs and L1RP constructs, ORF1p levels differ by ∼2.5-fold, but ORF2p levels differ by ∼20- to 40-fold compared to a purified standard (Figures 2H, 2I, and S2).

The expression constructs, cell lines, and approach collectively constitute an analytical suite unhindered by typical
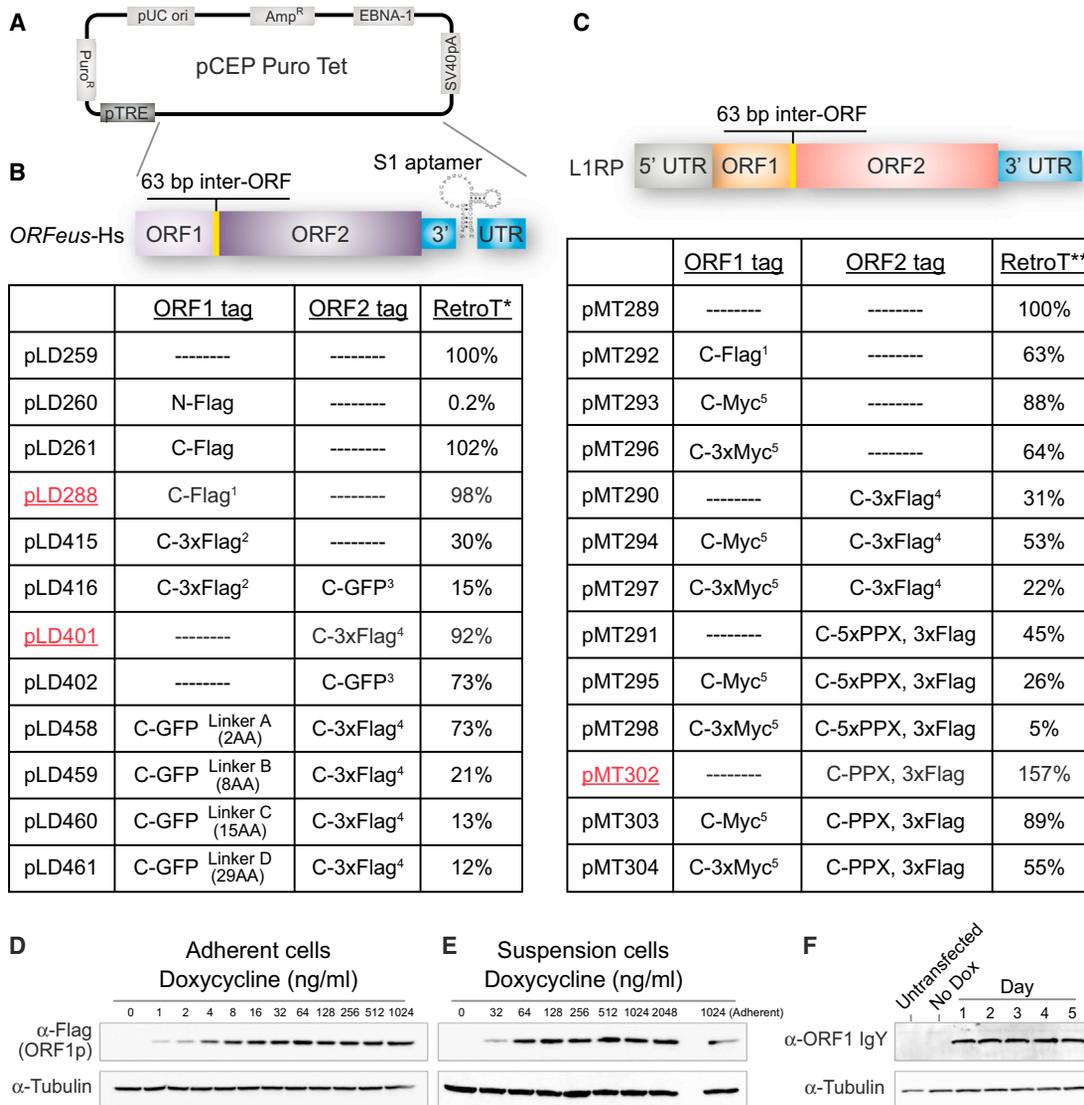
Figure 1. L1 Expression and Retrotransposition Constructs for Purification

	ORF1 tag	ORF2 tag	RetroT*
pLD259	--------	--------	100%
pLD260	N-Flag	--------	0.2%
pLD261	C-Flag	--------	102%
pLD288	C-Flag[1]	--------	98%
pLD415	C-3xFlag[2]	--------	30%
pLD416	C-3xFlag[2]	C-GFP[3]	15%
pLD401	--------	C-3xFlag[4]	92%
pLD402	--------	C-GFP[3]	73%
pLD458	C-GFP Linker A (2AA)	C-3xFlag[4]	73%
pLD459	C-GFP Linker B (8AA)	C-3xFlag[4]	21%
pLD460	C-GFP Linker C (15AA)	C-3xFlag[4]	13%
pLD461	C-GFP Linker D (29AA)	C-3xFlag[4]	12%

	ORF1 tag	ORF2 tag	RetroT**
pMT289	--------	--------	100%
pMT292	C-Flag[1]	--------	63%
pMT293	C-Myc[5]	--------	88%
pMT296	C-3xMyc[5]	--------	64%
pMT290	--------	C-3xFlag[4]	31%
pMT294	C-Myc[5]	C-3xFlag[4]	53%
pMT297	C-3xMyc[5]	C-3xFlag[4]	22%
pMT291	--------	C-5xPPX, 3xFlag	45%
pMT295	C-Myc[5]	C-5xPPX, 3xFlag	26%
pMT298	C-3xMyc[5]	C-5xPPX, 3xFlag	5%
pMT302	--------	C-PPX, 3xFlag	157%
pMT303	C-Myc[5]	C-PPX, 3xFlag	89%
pMT304	C-3xMyc[5]	C-PPX, 3xFlag	55%

(A) Backbone plasmid pCEP-puro-Tet.
(B) Constructs in *ORFeus*-Hs. Complete tag and linker sequences (protein and DNA) are detailed in Table S3. 1, Linker-Tev-V5-HA-Flag; 2, Linker-Tev-V5-HA-3xFlag; 3, Linker-LAP tag (His8-2xPrescission-GFP); 4, Linker-His8-2xPrescission-2xMyc-3xFlag. RetroT*: retrotransposition efficiency of all *ORFeus*-Hs-tagged constructs is relative to untagged L1 (pLD259), assigned as 100%.
(C) Constructs in L1RP. 5, Linker-Tev-Myc or Linker-Tev-3xMyc. RetroT**: retrotransposition efficiency of all tagged L1RP constructs were relative to untagged L1 (pMT289), assigned as 100%.
(D and E) L1 expression (pLD288) in adherent and suspension Tet-On HEK293$_{LD}$ cells at various Dox concentrations.
(F) Untagged L1 expression (pLD259) in suspension cells over time with 1 μg/ml Dox.
See also Figures 2H and S2.

limitations in protein abundance, detection sensitivity, or capture affinity.

L1-RNP Properties

Eluting native L1 RNPs by gentle displacement with excess 3xFlag peptide revealed both copurifying species and associated enzymatic activity. ORF2p shows an RNase sensitive association with ORF1p pullouts as determined by immunoblotting (Figure 2J). Initially, interactors were characterized by MALDI-TOF PMF (matrix-assisted laser desorption/ionization time of flight peptide mass fingerprinting) of selected protein bands observed after CBB staining. Numerous protein bands were more abundant or uniquely present in ORF pullouts as compared to mock pullouts (Figure 3A). These included PABPC1 and MOV10, recently characterized components of an otherwise sparsely defined L1-host interactome (Arjan-Odedra et al., 2012; Dai et al., 2012; Goodier et al., 2012).

We assayed RT activity in our preparations with an exogenous homopolymeric template/primer. Pullouts from both tagged ORF1p (pLD288)- and ORF2p (pLD401)-expressing cells were active; ORF2p preparations exhibited ∼40-fold more activity (Figure S3), requiring dilution to prevent saturation of the assay. Signal in both types of purified RNPs, but not controls, was approximately doubled by addition of RNase A/T1, suggesting that additional RT activity may be masked by RNA-based interactions in RNPs.

We also used the L1 element amplification protocol (LEAP) assay, a more stringent test measuring the ability of RNPs to reverse transcribe endogenous RNA upon addition of a specialized DNA primer. LEAP has demonstrated *cis*-preference of ORF2p in vitro; it is currently the best biochemical assay for functional coassembly of L1 RNA and proteins (Kulpa and Moran, 2006). Pullout RNPs exhibited 20-fold more LEAP activity than those isolated from traditional sucrose cushions (Figure 3B) but had 3.4-fold less L1 RNA, resulting in 67-fold higher specific activity (Figure 3C); product electrophoretic mobilities were similar. When signal was normalized to total RNA, pullout RNPs had >13,000-fold higher LEAP activity (Figure S3).

In RNA-seq of pullout RNPs, L1 RNA represented 8.3%–10.3% and 18.0%–28.2% of mappable reads for ORF1p and ORF2p, respectively. U6 snRNA, but not other snRNAs, was enriched 2.5- to 5.9-fold in ORF2p pullouts, but not enriched in ORF1p pullouts, relative to housekeeping mRNA controls. This finding mechanistically supports U6/L1 hybrid elements commonly found in the human genome (Buzdin et al., 2002). Interestingly, *Alu* RNA was not significantly enriched, possibly because *Alu* and ORF2p interactions are weaker or transient and may not have been retained in the stringent purification condition used. *Alu* RNA was identified in a similar study in which a crosslinking method was used (Mandal et al., 2013).

Relative Amounts of L1 Proteins

It is widely believed that L1 RNA is coated by ORF1p trimers to form an RNP containing one or two molecules of ORF2p, and this is supported by relative expression levels observed for ORF1p and ORF2p (Beck et al., 2011; Burns and Boeke, 2012; Doucet et al., 2010; Goodier and Kazazian, 2008). Each ORF1p trimer can protect ∼27 nucleotides, and trimers can be spaced every 45 nucleotides (Khazina et al., 2011). Thus a fully coated L1 RNA could contain as many as one ORF1p trimer every ∼75 nt to a maximum of 80 trimers or 240 ORF1p molecules per 6kb L1 RNA, giving an ORF1p:ORF2p ratio of up to 240:1. In contrast, our initial observations, based on CBB staining of ORF2-3xFlag pullouts in both *ORFeus*-Hs and L1RP, suggested a dramatically lower ORF1p:ORF2p molar ratio of ∼3:1 (Figure S1).

Due to overexpression, some of the affinity captured ORF2-3xFlag protein may not be bound within a physiological L1 RNP, and this may skew the ORF1p:ORF2p ratio. To mitigate this possibility, we performed tandem pullout experiments—purifying first by ORF2p-3xFlag, natively eluting the complexes with 3xFlag peptide and then repurifying by ORF1p affinity. Hence, every ORF2p observed in the elution from a tandem pullout was physically linked to ORF1p throughout. We measured the relative ratio within these more homogenous L1 preparations

containing both polypeptides (Figures 3D and S4). Based on both Sypro and CBB staining intensity, we estimate that the ratio of ORF1p:ORF2p in tandem affinity isolated RNPs is ∼6:1–9:1 for *ORFeus* and L1RP constructs (Figure S4). In an alternative approach, we utilized MS and implemented intensity-based absolute quantification (iBAQ) (Schwanhäusser et al., 2011; Smits et al., 2013) to assess stoichiometry of ORF proteins. We found ratios of 47:1 ± 2 for particles derived from *ORFeus* and 27:1 ± 7 for particles derived from L1RP (Table S4). Because the former method accounts only for full-length proteins and not partially degraded protein that survives the tandem pullout procedure and because the latter can detect peptides derived from either full-length or partially degraded proteins, these results are not incompatible. Some degradation likely occurs during the tandem pullout despite the use of protease inhibitors. The higher ORF1:ORF2p ratios determined by iBAQ may better reflect actual ratios within RNPs but are significantly lower than the ratio expected if L1 RNA were stably coated with ORF1p trimers.

To better understand the discrepancy between total and RNP ORF1p content, we examined intracellular distribution of the proteins by immunofluorescence in Tet-On HEK293T$_{LD}$, Tet-On HeLa, and standard HeLa cells. The predominant localization of both proteins was cytoplasmic, although we noted fewer foci and greater staining uniformity than reported elsewhere (Doucet et al., 2010; Goodier et al., 2007, 2010). Unexpectedly, we found that only ∼30% of cells expressing ORF1p coexpress ORF2p (Figure 3E). This trend held in both Tet-On HEK293T$_{LD}$ and HeLa, when the inducible Tet promoter was replaced with a CMV promoter, and in both L1RP and *ORFeus*-Hs constructs (Figures 3E and S5). In cells expressing both proteins, ORF1p and ORF2p colocalize. To verify that this ∼30% expression was due to L1 context and not an inherent property of ORF2p, we expressed ORF2p alone from a TRE promoter and found that >90% of cells express ORF2p (Figure S5, pLD561, diagrammed in Figure 5B).

MS of L1-RNPs

Bolstered by CBB stain detection of PABPC1 and MOV10, we implemented I-DIRT (Tackett et al., 2005) in order to distinguish proteins that tightly associate with L1 RNPs from nonspecific contaminants (Figure 4A). High-sensitivity LC-MS/MS (liquid chromatography electrospray ionization tandem MS) analyses of six independent ORF1p and ORF2p pullouts from *ORFeus*-Hs- or L1RP-expressing cells identified 37 specific interactors from 755 sequenced candidates (Figure S6 and Table S5). Results of a 5 min ORF1p-Flag (pLD288) pullout and a 30 min ORF2p-3xFlag (pLD401) pullout are shown in Figures 4B and 4C, respectively.

Although ORF1p and ORF2p reciprocally coprecipitate as expected, they also yield a number of differentially coprecipitated factors. For example, TROVE2, the Ro autoantigen known to bind misfolded RNA (Fuchs et al., 2006), was found with high specificity in ORF1p pullouts only. PCNA, found with high specificity but low abundance in only ORF2p I-DIRT analyses, was followed up as a novel potentially nuclear factor. Other nuclear factors include DNA repair, transcription and chromatin factors (PARP1, PURA, PURB, RUVBL1, NAP1, TOP1, and ZCCHC3
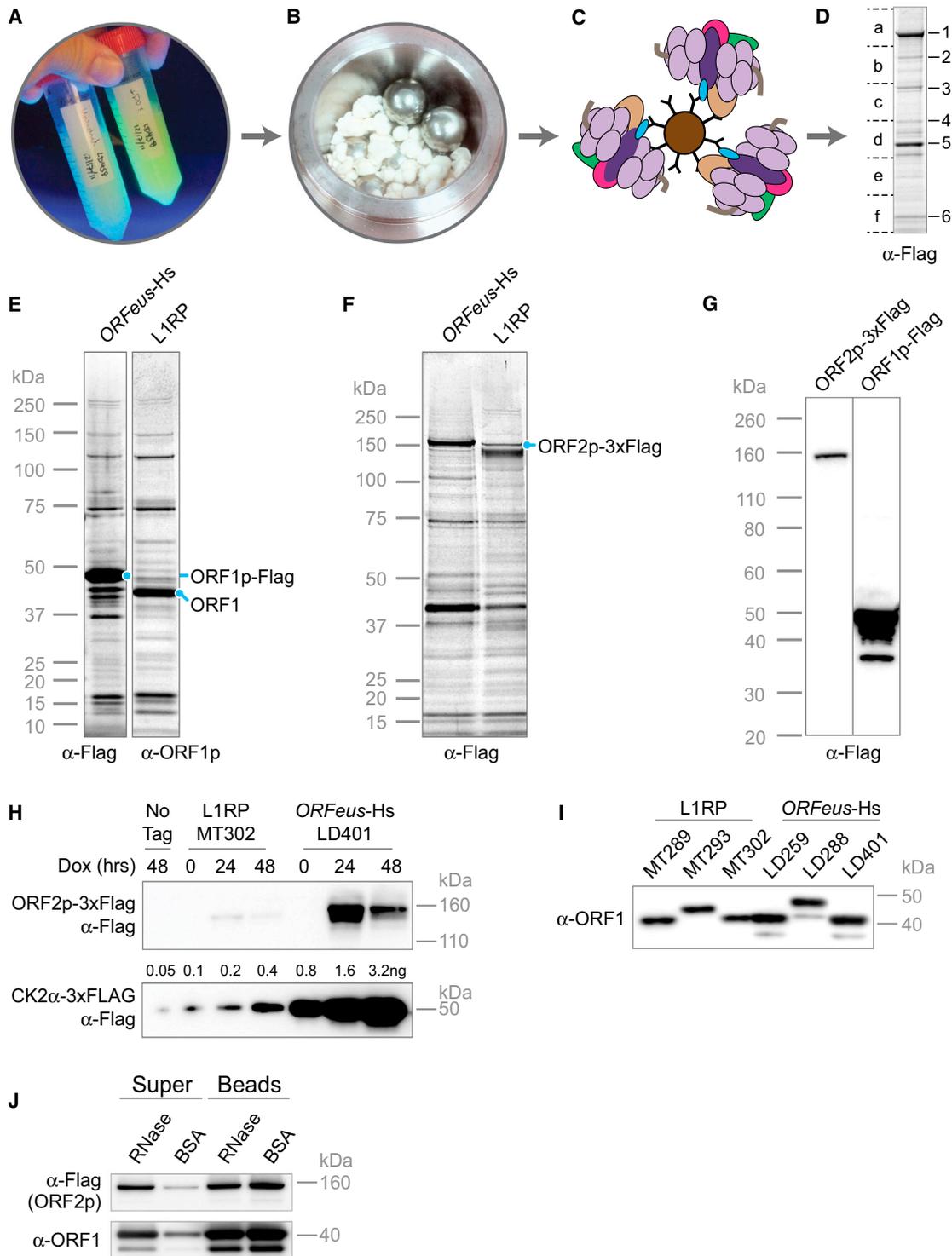
Figure 2. Interactomics Workflow Applied to L1 Produces Tractable Quantities of Both ORF Proteins from Both L1RP and *ORFeus*-Hs

(A) Transposon-expressing human cell cultures are grown in suspension and are induced for 24 hr.

(B) Frozen cell "BBs" (~3 mm diameter spherical globules of cells previously frozen in LN$_2$) are cryomilled using stainless steel ball bearings under liquid nitrogen, resulting in a fine powder stored at −80°C.

(C) Antibody-conjugated μm-scale magnetic medium is used for rapid pullout.

(D) Gel-based proteomics: purified proteins are subjected to MALDI-TOF-PMF for identification of major species (1–6) or LC-MS/MS (a–f).

(E) ORF1p and copurified factors for *ORFeus*-Hs (pLD288) and L1RP (pMT289) by ORF1p purification.

(legend continued on next page)

proteins), an apoptosis regulator (HAX1), and a nuclear import factor (IPO7). These results are consistent with the possibility of ORF1p and ORF2p having both shared and distinct host interactomes. Notably, PABPC1, PABPC4, MOV10, UPF1, and ZCCHC3 were specific interactors in both ORF1p and ORF2p RNP preparations. although all of these hits are candidates for mechanistic studies, we focused initial follow-up on UPF1 and PCNA. UPF1 was chosen because of its high abundance and obvious potential connections between bicistronic L1 mRNA and NMD.

UPF1 Interaction Requires ORF2 and Downregulates L1 Expression

We validated the L1/ UPF1 interaction by immunoblotting for UPF1 in ORF1p-Flag and ORF2p-3xFlag pullouts from cell lines expressing pLD288 and pLD401, respectively. UPF1 was detected in both ORF1p and ORF2p elutions and was copurified through tandem pullouts from pLD401-expressing cells (Figure 5A). The L1/ UPF1 interaction in purified fractions was sensitive to RNase treatment (Figures 6G and 6H, see below).

In addition to recognizing mRNAs bearing premature nonsense codons, UPF1 affects expression of intronless and bicistronic mRNAs (e.g., MAP3K14 and ARHGEF18) (Mendell et al., 2004). Within the L1 mRNA, a 63 nt inter-ORF region contains two additional stop codons in frame with ORF1 (Figure 5B). We hypothesized that the three inter-ORF stop codons in conjunction with a long 3′ UTR-mimic (∼4 kb ORF2 plus the native 3′ UTR) might be recognized by UPF1 (Hogg and Goff, 2010). It is known that UPF1 also binds to and scans 3′ UTRs of many cellular mRNAs (Shigeoka et al., 2012). We constructed several plasmids with different combinations of ORF, interORF, and 3′ UTR sequences to map interacting sequences (Figure 5B).

ORF1p affinity isolations demonstrate that removal of ORF2p results in higher yields of both ORF1p protein and L1 RNA but results in a dramatically reduced yield of UPF1; removal of the 3′ UTR had no effect (Figure S7 and 5C). In ORF2p affinity isolations, removal of the 3′ UTR similarly had no effect on the yield of UPF1, but, in contrast to our hypothesis, removal of ORF1p also had no effect on UPF1 yield (Supplemental Information and Figures 5D and S7). Although the yield of ORF2p in the three purifications differed slightly, the UPF1 and ORF2p signal intensities precisely correlate (r^2 = 0.995) (Figure S7G). In summary, the ORF2 coding sequence or ORF2p expression, rather than the interORF sequence, ORF1p sequence, bicistronic context, or 3′ UTR, enables the L1-UPF1 interaction.

Next, we investigated the impact of UPF1 on L1 expression by knocking down endogenous UPF1 by RNA interference. L1 mRNA and polypeptides were all upregulated in the knockdown cells. Surprisingly, L1 retrotransposition efficiency decreased

by ∼50% (Figures 5E–5G and S7C–S7F). To minimize the possibility of off-target RNAi effects, we coexpressed L1 with five different anti-UPF1 shRNAs. Four of these reduced retrotransposition significantly by as much as 90%, suggesting that UPF1 depletion decreases retrotransposition. Knockdown of two other proteins that act in concert with UPF1 in NMD, UPF2, and UPF3a also reduced retrotransposition and increased L1 mRNA levels (Table S6).

PCNA Binding to an ORF2 PIP Box Is Critical for Retrotransposition

PCNA was found as a low-abundance but high-specificity interactor of ORF2p, but not ORF1p. We identified a canonical PCNA-interacting protein (PIP) box (Qxx[V/L/M/I]xx[F/Y][F/Y]) in residues 407–415 of ORF2p, located between the EN and RT domains (Figure 6A). Within the PIP sequence, four amino acid residues (I407, I411, Y414, and Y415) are highly conserved across L1 ORF2s from diverse species, suggesting a crucial role for a PIP box in ORF2 binding to PCNA. Mutation of these residues to alanine (A) disrupted ORF2p/PCNA copurification, whereas mutation of Q408, less conserved in ORF2p, does not affect it (Figure 6B). Mutations that disrupt the ORF2p /PCNA interaction correspondingly decrease retrotransposition efficiency of L1, whereas the Q408A mutation shows no effect on retrotransposition (Figure 6C). Overall, these data support an ORF2p/PCNA interaction through a PIP box, and this interaction is critical for retrotransposition. We also knocked down endogenous PCNA using four small hairpin RNAs (shRNAs) and observed downregulation of L1 retrotransposition well correlated with decreased PCNA levels (r^2 = 0.93; Figure 6D). To investigate the L1 life cycle stage at which L1/PCNA interaction functions, we assayed for maintenance of this interaction in cells expressing an ORF2p EN domain catalytic mutant (H230A) (Figure 6E), which was previously shown to abolish DNA nicking and L1 retrotransposition (Feng et al., 1996). Copurification of PCNA was greatly reduced in RNPs isolated from the endonuclease domain mutant (Figure 6E). To rule out the possibility that mutations in the EN domain disrupt overall ORF2p folding, we repeated the experiment with six additional well-studied EN mutants (E43A, D145A, R155A, T192Y, I204Y, and D205G; Figures 6F and S7) (Repanas et al., 2007, 2009; Weichenrieder et al., 2004). We measured RT activity of these EN mutants, and all were highly active, suggesting proper folding (Figure S7). Similarly, a catalytic RT mutant (D702Y) also abolished ORF2p-PCNA interaction (Figure 6F). These observations suggest that PCNA may be recruited to ORF2 after completion of the target DNA nicking and during or after L1 complementary DNA (cDNA) synthesis steps.

Because PCNA is involved in DNA repair pathways, we also studied other L1 RNP factors with possible roles in DNA repair

---

(F) ORF2p and copurified factors for *ORFeus*-Hs (pLD401) and L1RP (pMT302) by ORF2p (anti-Flag) purification.

(G) Comparison of ORF1p-Flag (pLD288) and ORF2p-3xFlag (pLD401) expression levels in puro-selected adherent cells; 20μg protein loaded in each lane. Please see Figure S2 for detailed calculations.

(H) ORF2p expression levels in L1RP versus *ORFeus*-Hs. CK2α-3xFlag was loaded as control. 20 μg protein loaded in each lane.

(I) ORF1p expression levels in L1RP versus *ORFeus*-Hs. 2 μg protein loaded in each lane.

(J) ORF1p-ORF2p interactions are RNase sensitive. L1 was affinity purified with anti-ORF1 Dynabeads from pLD401 and then treated on-bead with RNases A/T1 or mock treated with BSA. "Super," proteins released into the supernatant; "beads," sample retained on beads after washing and then eluted with LDS sample buffer. See also Figures S1 and S2.
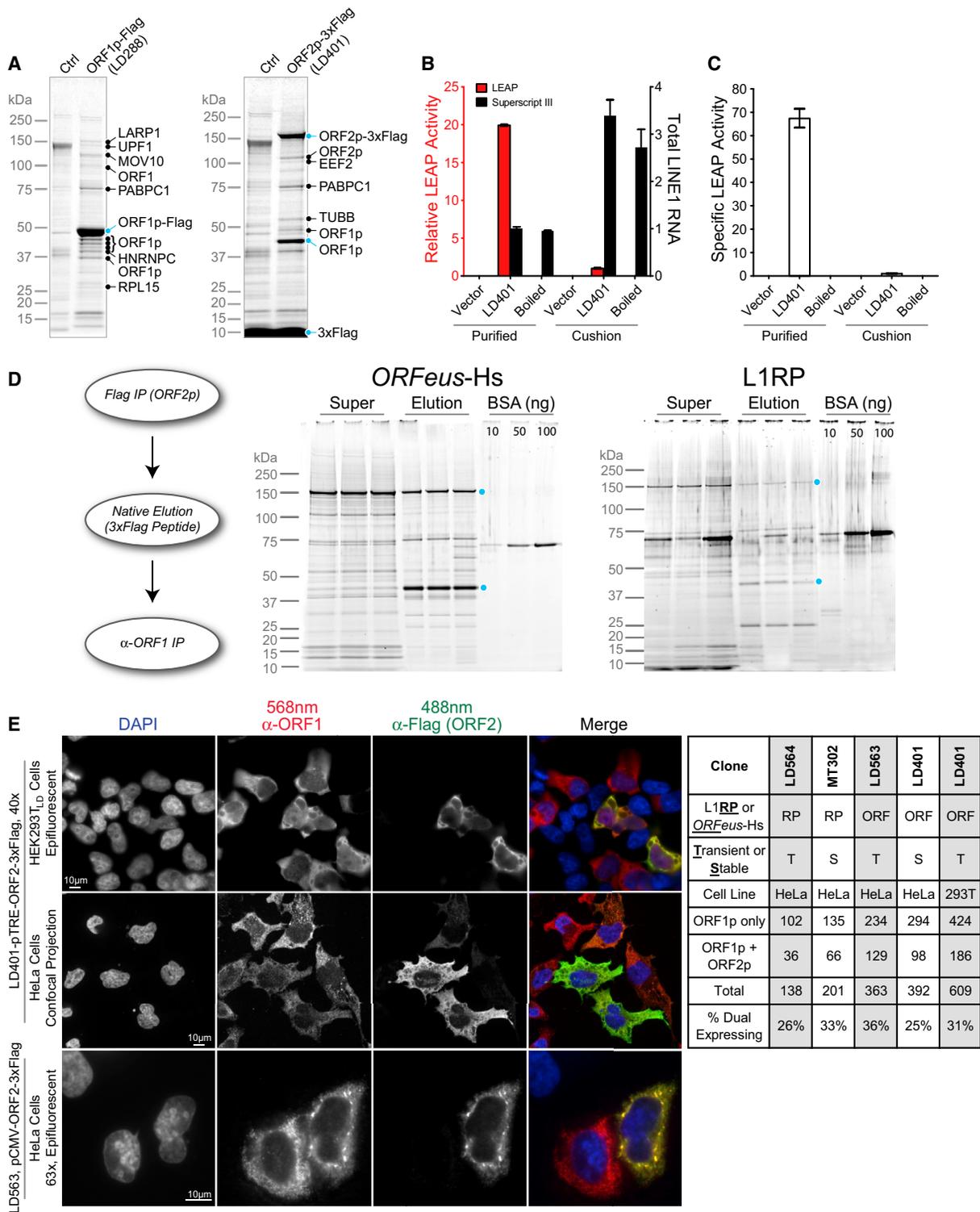
Figure 3. Characterization of Highly Purified L1 RNPs

(A) L1 interactors characterized by MALDI-TOF-PMF MS. Ctrl, empty vector-containing cells.

(B) Relative LEAP activity from affinity-purified RNPs and RNPs made by sucrose cushion velocity sedimentation. Vector, RNP purified from cells transfected with empty pCEP puro vector; boiled, RNP sample was boiled for 10 min at 100°C before the reaction. Data are represented as mean ± SEM.

(C) Specific LEAP activity was normalized to total L1 mRNA in each sample. LEAP activity of RNP prepared from sucrose cushion was assigned as 1. Data are represented as mean ± SEM.

(legend continued on next page)

in the L1 retrotransposition assay. We found that knocking down PARP1 and RUVBL1 reduced L1 activity (Table S6), which is consistent with possible roles as retrotransposition host factors.

Connectivity and Forms of L1 RNPs

To investigate the RNA-dependent connectivity within these L1 RNP particles, affinity-captured ORF1p-Flag and ORF2-3xFlag fractions were incubated with a mixture of RNases A and T1 or subjected to mock treatments while immobilized on magnetic beads. When ORF1p was immobilized, ORF1p itself and UPF1 and ORF2p were all RNase sensitive, suggesting interaction via L1 RNA (Figures 2J and 6G). PCNA was undetectable following ORF1p purification. When ORF2p was immobilized, ORF1p and UPF1 were RNase sensitive, but PCNA and ORF2p itself were insensitive (Figure 6H). PCNA was dramatically reduced following ORF1p purification (Figure 5A).

The combination of the RNase sensitivity data, the immunofluorescence data, and the differential purification leads us to propose (at least) these three distinct RNP forms (Figure 7A): (1) cytoplasmic and (2) nuclear RNPs in ORF1p+ORF2p-expressing cells and (3) cytoplasmic RNPs in cells expressing ORF1p alone. However, it is likely that this simplistic classification hides additional complexity in the form of additional RNP types.

DISCUSSION

As a result of its high copy number and the ability to mobilize itself and nonautonomous retroelements like *Alu*, *SVA*, and processed pseudogenes, L1 is considered the most successful parasite in the human genome. Viewed in this light, physical interactions between L1 proteins and host factors likely reflect a longstanding evolutionary battle between the parasite, which exploits opportunities to proliferate, and the host, which fights to suppress L1 accumulation. Dissecting the dynamics of this L1-host "arms race" will help to better decipher our genome dynamics in relation to cancer, brain development, reproduction, and potentially to other human diseases.

The construction of synthetic LINE-1 elements has facilitated L1 expression in mammalian cells for in vitro and in vivo studies (An et al., 2011; Han and Boeke, 2004). Another facilitator of this study was the adaptation of adherent cells to suspension culture, dramatically reducing space, cost, and labor needed to maintain 10^9–10^{10} cells. Exploration of epitope tags unexpectedly revealed that some tags compatible with *ORFeus*-Hs were not necessarily compatible with L1RP at the same position. This may reflect differences in local RNA structure, overall translational context, or complex effects on protein/RNP structure. Cryogenic milling of frozen cells allowed us to prepare large batches of frozen solid powder usable over long periods of experimentation. Experiments within and between batches were easily calibrated to use the same mass of cell powder in

each experiment. These advantages allowed us to minimize experimental variation, perform extensive optimization, and develop efficient pullout procedures, including only 5 min of binding.

Comparing native and synthetic elements, ORF1p levels differed by ∼2.5-fold, but we found that ORF2p levels differ by ∼20- to 40-fold, mirroring the ∼40-fold difference seen previously in the RNA levels (An et al., 2011). Given the modest increase in retrotransposition observed in *ORFeus*-Hs-expressing cells, bulk L1 expression may not limit retrotransposition in these models. With absolute efficiencies under 20% in even the most permissive cell lines, host defenses may result in seemingly stochastic expression of ORF2p in just 30% of cells, which may in part explain the limitation. Alternatively, lack of balance in expression of the RNA and the two ORFs may explain observed nonlinear increases in retrotransposition. Conversely, the sensitivity of the L1RP to tag sequences that are well tolerated in *ORFeus*-Hs may suggest that higher expression can compensate for impaired L1 protein function.

Unlike other non-LTR transposons, such as the R2 element, L1 ORF2p is translated sparingly. ORF2p detection has been challenging even by immunoblotting with the help of epitope tags. This is now routine, and moreover, our method provides L1 RNPs at previously unobtainable yield and purity. Most importantly, affinity-isolated RNPs were highly active in vitro in LEAP and RT assays. The high activity of the RNPs in vivo and in vitro should facilitate further biochemical dissection of the L1 retrotransposition mechanism.

Pullout experiments with ORF1 or ORF2 shed light on relative stoichiometry in RNPs. A low ORF1p:ORF2p ratio challenges widely held assumptions about RNP composition. RNase insensitivity of immobilized ORF2p in the context of RNase sensitivity of other RNP components is consistent with a single ORF2p copy per mRNA; however, we cannot rule out ORF2p dimerization, dual tethering to the affinity matrix, or RNase-insensitive interactions. Instead of a constant L1 protein ratio in each cell, imaging studies revealed that only a minority of cells expresses both proteins. These data suggest that the host may combat L1 by regulating ORF2p translation or stability to produce deleterious ORF1p:ORF2p ratios. This dual nature of cell types may be cell-cycle regulated, epigenetically controlled, or a response to other signals. In light of this, some interactors found bound only to ORF1p may represent interactions formed in cells not expressing ORF2p. Surprisingly, we do not observe nuclear localization of ORF2p by immunofluorescence. At least a fraction of it must be nuclear, given that the target DNA is located there and that ORF2p/PCNA interaction requires ORF2p EN and RT activities.

In this study, we identify 37 high-confidence interactors among 755 candidates. Many known interactors were recovered with confidence; some were not recovered, presumably because our methods enrich stable interactors. A recent study

---

(D) Tandem affinity purification of L1 RNPs. L1 RNPs were first purified from ORF2p (3xFlag) and then ORF1p. Sypro Ruby stained supernatant (super) and elution from triplicate ORF1p purifications are shown for both *ORFeus*-Hs and L1RP; ORF1p and ORF2p levels were quantified against a BSA standard (Figure S4).
(E) Indirect immunofluorescent imaging of ORF1p and ORF2p in HeLa and HEK293T cells. Ratio calculation for HeLa and HEK293T cells expressing ORF1p alone or both ORF1p and ORF2p are shown. Numbers in the table are counts of cells expressing ORF1p alone or both ORFs above background.
See also Figures S3, S4, and S5.
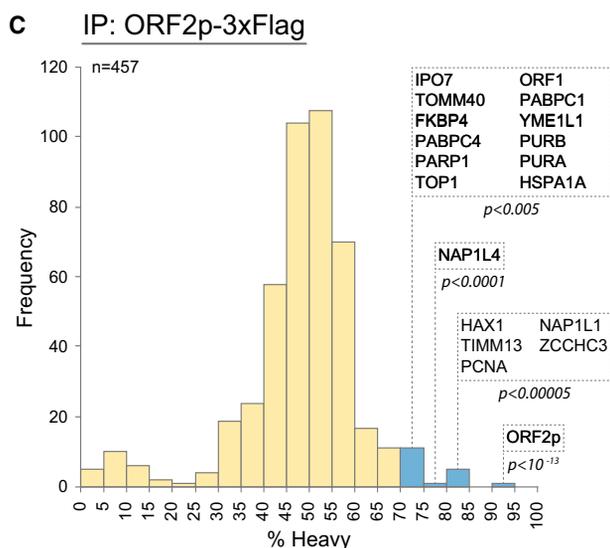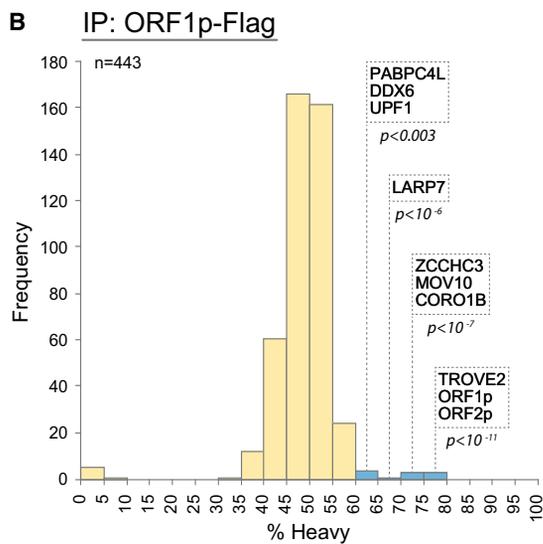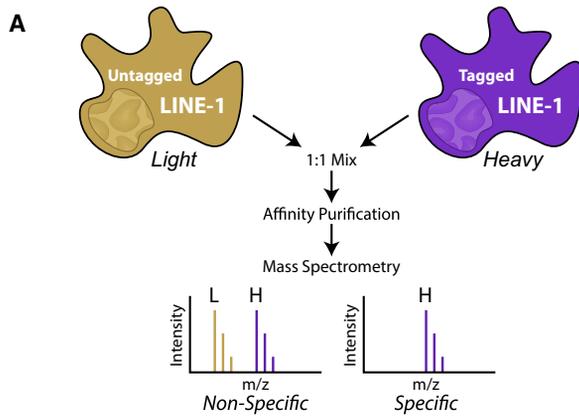
Figure 4. High-Confidence Interactors of ORF1p and ORF2p from I-DIRT Experiments

(A) I-DIRT methodology. Cells expressing a tagged L1 protein are grown in heavy-isotope media, whereas cells expressing untagged L1 proteins are

reported 96 L1 ORF1 interactors (Goodier et al., 2013) of which we found 69; seven (10%) were significant using the I-DIRT threshold employed (Table S7).

Models for UPF1's Role and L1 Expression Effects

NMD is an important pathway for maintenance of high-fidelity gene expression. In NMD, UPF1 is a central player, recruiting other factors to initiate targeted degradation of mRNAs containing premature termination codons. UPF1 also targets bicistronic and intronless mRNAs, and RNA sequencing of UPF1-depleted HeLa cells revealed hundreds of downregulated endogenous transcripts (Mendell et al., 2004; Tani et al., 2012). L1 bicistronic mRNA seemed reasonable as a target of UPF1. Consistent with this model, UPF1 copurification with ORF1p-Flag was greatly reduced when the ORF2p coding sequence was removed. However, UPF1 copurification with ORF2p-3xFLAG was retained on deletion of ORF1 and the inter-ORF region (Figures 5B–5D). Results from additional constructs ruled out the 3′ UTR as being critical for UPF1 binding (Figure 5C). Thus, ORF1, the stop codons, inter-ORF region, bicistronic transcript structure, and 3′ UTR are all dispensable for UPF1 recruitment. Therefore, UPF1 recognition of L1 RNA does not resemble canonical NMD. Additionally, RNA is critical to the interaction, as the ORF2p/UPF1 interaction exhibits sensitivity to RNase treatment (Figure 6G).

Taken together, these data suggest that L1-UPF1 interaction occurs through one of three modes: (1) through ORF2 RNA, (2) through the ORF2p-L1 RNA complex, or (3) through the ORF2 protein but stabilized when ORF2p is also bound to RNA. The first model seems unlikely because both *ORFeus* and L1RP recruit UPF1, even though their ORF2 RNA sequences are quite different. Indeed, UPF1 presents as a positive regulator of retrotransposition activity, despite negatively regulating the levels of L1 constituents (Figures 5E–5G), leaving open the possibility of UPF1 affecting other RNP-dependent processes. Furthermore, like the UPF1, UPF2 and UPF3a knockdowns paradoxically increased L1 RNA levels while decreasing retrotransposition. In conclusion, the data support a positive role for UPF1, UPF2, and UPF3a in L1 RNP function while at the same time demonstrating repression of protein and RNA levels through a mechanism distinct from the canonical NMD pathway.

Models for PCNA's Role in the L1 Life Cycle

We showed that the ORF2 protein binds to the sliding DNA clamp protein PCNA through a canonical PIP domain. Disruption of the conserved amino acid residues within the ORF2 PIP

grown in light-isotope media. After cryomilling, cells of both types are mixed, and L1 is affinity purified using the tag. Purified proteins enriched for heavy isotopes represent specific, stable in vivo interactions, whereas near-equal mixtures of light and heavy may be interactions formed during purification. Light and heavy represent stable isotope labeling of Arg and Lys residues.

(B) Interactors identified by ORF1p (Flag) IP with 5 min binding.

(C) Interactors identified by ORF2p (Flag) IP with 30 min binding. Histograms plot the number of recovered proteins by heavy isotope content. Yellow bars, nonsignificant groups; blue bars, statistically significant specificity after Benjamini-Hochberg correction for multiple hypothesis testing. p values were determined by Perseus software.

See also Figure S6 for results of all six I-DIRT experiments.
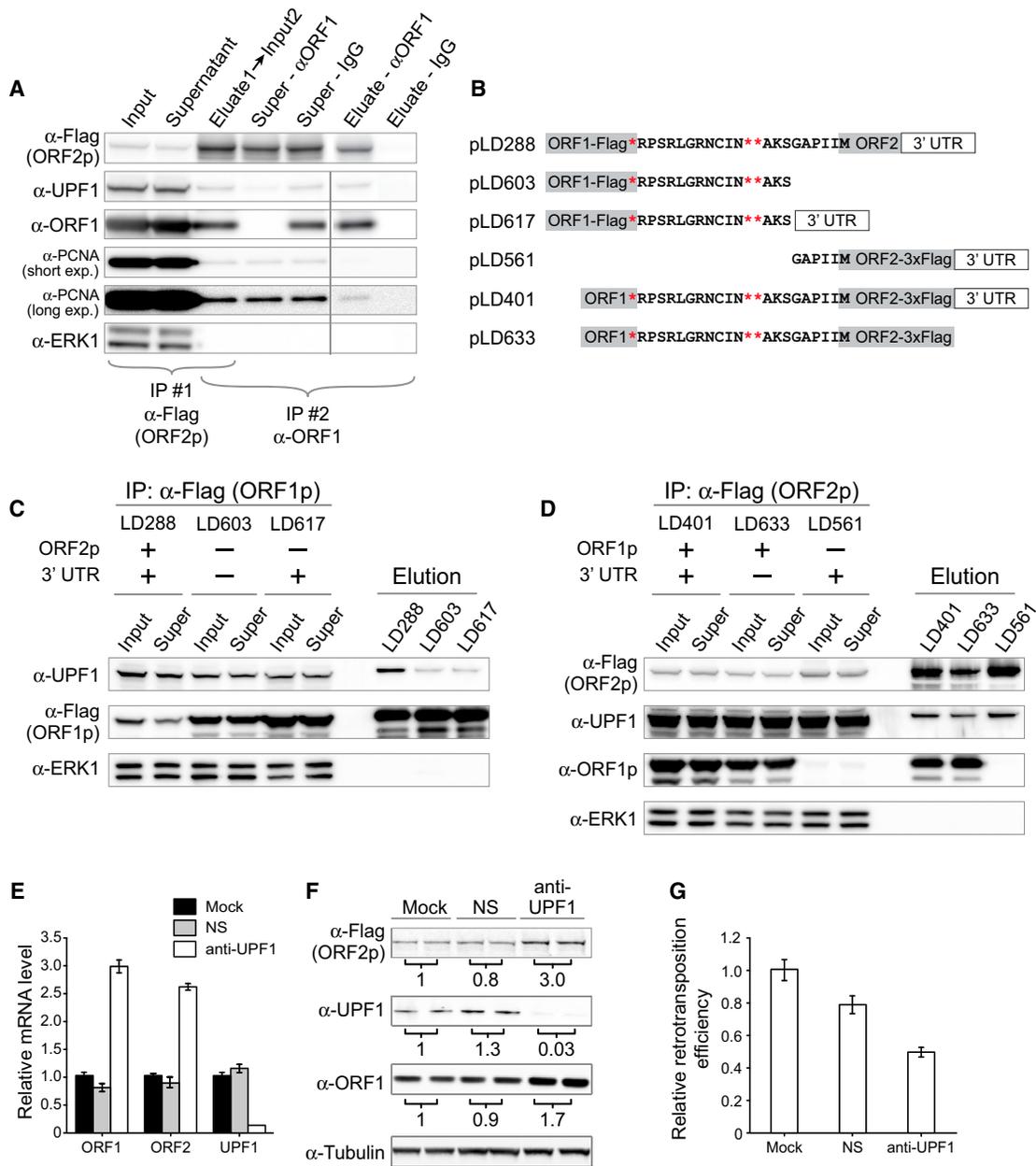
Figure 5. UPF1 Has Multiple Effects on L1 Expression and Retrotransposition

(A) UPF1 is copurified in particles containing both ORF1p and ORF2p. L1 from pLD401 was affinity purified sequentially using first anti-Flag and then, second, using either anti-ORF1 or, as a control, IgG-conjugated magnetic beads. Super, supernatant after purification.

(B) Constructs used to test copurification of UPF1.

(C) UPF1 binding to L1 RNPs was significantly reduced without ORF2 and is not affected by the 3′ UTR. Please see Figure S7 for detailed calculations.

(D) UPF1 and PCNA binding to the L1 RNP was unaffected by the absence of ORF1 (LD561) or 3′ UTR (pLD633).

(E) Knockdown of endogenous UPF1 increased L1 mRNA levels. Data are represented as mean ± SEM. NS, nonsilencing siRNA.

(F) Knockdown of endogenous UPF1 increased L1 protein levels. Protein levels were normalized to a tubulin loading control, and numbers below each panel are the average fold change relative to mock transfection from two independent experiments. Data are represented as mean ± SEM.

(G) Knockdown of endogenous UPF1 reduced L1 retrotransposition efficiency. The retrotransposition efficiency from mock-treated cells was assigned as 1, and the values are averages of four independent experiments.

Data are represented as mean ± SEM. Please also see Figure S7.

sequence abrogated ORF2p binding to PCNA and inhibited L1 retrotransposition. The well-established role of PCNA in DNA replication and DNA damage repair led to three hypotheses for

how PCNA supports L1 retrotransposition: (1) ORF2p may be recruited by PCNA onto genomic DNA (or led to it) and thereby enable efficient scanning of gDNA until a preferred
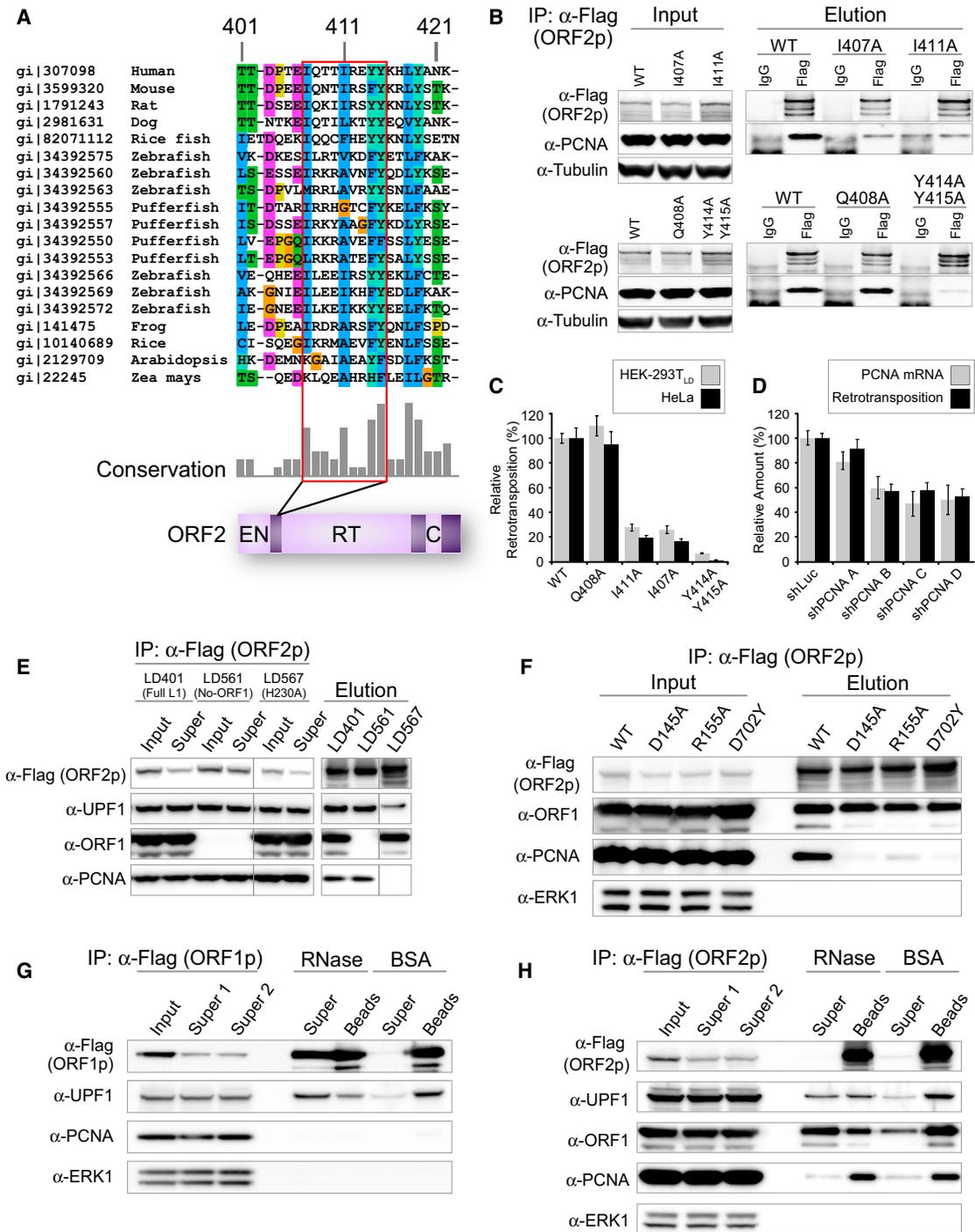
Figure 6. PCNA Specifically Interacts with a Conserved Region within ORF2p in an EN- and RT-Dependent Way

(A) Partial sequence alignment of LINE ORF2p from various species indicating the PIP box, a known PCNA interaction motif, outlined in red. PIP: QXX(V/L/M/I)XX(F/Y)(F/Y).

(B) PCNA specifically coimmunoprecipitates with ORF2p, and this interaction is dependent on the PIP box.

(C) The PIP box is important for L1 retrotransposition activity in both HeLa and HEK293T cells. The retrotransposition efficiency of the wild-type L1 was assigned as 1; values are the average of three independent experiments; SE is shown.

(D) Knockdown of endogenous PCNA decreased L1 retrotransposition, and data are represented as mean ± SEM.

(E) PCNA-ORF2p coimmunoprecipitation is abolished in the ORF2p EN mutant. UPF1, but not PCNA, binding is retained in a mutant ORF2p lacking endonuclease activity (H230A).

(legend continued on next page)
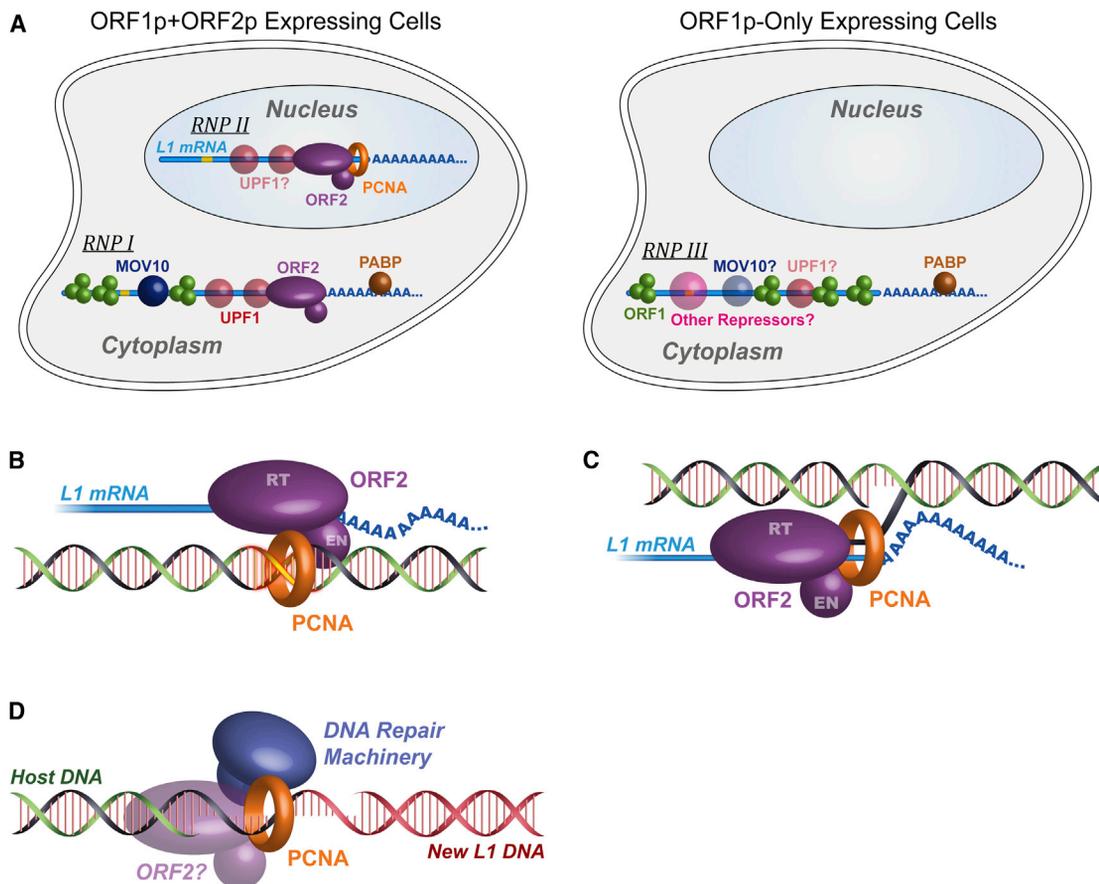
Figure 7. Proposed Models for L1 RNP Complex and Its Interaction with UPF1 and PCNA

(A) Proposed model of distinct L1 RNPs in the cytoplasm and nucleus. Green sphere, ORF1p; purple oval, ORF2p; red sphere, UPF1; blue sphere, MOV10; brown sphere, PABPC1, orange ring, PCNA; blue line, L1 mRNA; yellow line, inter-ORF region of the L1 RNA containing three in-frame stop codons.

(B) PCNA model 1: ORF2p uses PCNA to find insertion sites at pre-existing nicks or sites of DNA damage (yellow).

(C) PCNA model 2: PCNA is a processivity factor and may coordinate RNase activities on the newly synthesized DNA-RNA hybrid.

(D) PCNA model 3: PCNA is recruited by ORF2p to coordinate postinsertion DNA repair, perhaps including second-strand synthesis.

endonuclease site or a preformed nick is found in target DNA (Figure 7B); (2) PCNA may act as a processivity factor for L1 RT (Figure 7C); or (3) PCNA may be recruited by ORF2p to repair the "damage" in the form of nicks, gaps, or RNA flaps at junctions between retrotransposon and host DNAs (Figure 7D). These hypotheses, especially the second and third hypotheses, are not mutually exclusive.

According to the first model, PCNA may function as a scaffolding protein that "scans" genomic DNA either during DNA replication or during DNA damage repair to find permissive sites for L1 retrotransposition. It has been previously speculated (Cost et al., 2002; Farkash et al., 2006) that L1 may use DNA damage

sites as targets for retrotransposition (Coufal et al., 2011). Also, in cell lines deficient for nonhomologous end-joining repair (NHEJ), L1 can use DNA damage sites for retrotransposition, and this event is independent of the ORF2p endonuclease activity (Morrish et al., 2002). All of these findings make such a scanning model of some interest. However, in contrast with this model, our coimmunoprecipitation experiments using EN mutant or RT mutant elements clearly show that PCNA binding to ORF2p depends on EN and RT activities and are thus more consistent with its recruitment during or subsequent to TPRT. These observations are most consistent with a model in which PCNA is involved in a step after target DNA nicking and L1 cDNA

(F) PCNA-ORF2p coimmunoprecipitation is abolished in an RT mutant (D702Y) and an additional ORF2p EN mutant D145A; partially active EN mutant R155A has reduced PCNA recovery.

(G) The interaction between UPF1 and ORF1p is RNase sensitive. pLD288 purified with anti-Flag Dynabeads and treated on-bead with RNases A/T1 or mock treated with BSA. "Super," supernatants; "beads," samples retained on beads after washing and eluted with LDS sample buffer.

(H) ORF2p-PCNA and ORF2p-ORF2p interactions are RNase resistant, and ORF2p-ORF1p and ORF2p-UPF1 interactions are RNase sensitive. pLD401 material was purified with anti-Flag Dynabeads and treated as in (G).

See also Figure S7.

synthesis. All EN mutants maintained high RT activity, and several of them have crystal structures closely resembling the wild-type enzyme, minimizing the possibility that EN mutations disrupt overall folding of ORF2p and thereby the physical interaction between ORF2p and PCNA.

Further, PCNA is known to be capable of recruiting repair enzymes like lesion bypass polymerases and ligases (Beattie and Bell, 2011; Ulrich, 2011) and, most interestingly, has been specifically implicated in recruiting RNase H2 to RNA/DNA hybrids in genomic DNA (Bubeck et al., 2011). Mammalian L1 elements do not encode a recognizable RNase H activity, unlike most other retrotransposons (including non-LTR elements like L1 in other organisms such as the *Drosophila* I factor), even though the activity is presumably required to degrade the original RNA strand, allowing second DNA strand synthesis. Thus, RNase H binding activity of PCNA would be highly consistent with a role in DNA elongation during reverse transcription and provide a convenient link to the RNase H domain of RT that is "missing" from mammalian L1 elements. Although this model may be appealing, the lack of interaction with the RT mutant presents a challenge to it; however, it is formally possible that a few base pairs of synthesis are required to enable PCNA engagement.

Other copurified factors may be associated with nuclear L1 RNPs through PCNA or in relation to PCNA. PURA, PURB, MEPCE, and PCNA have been shown to be associated with PCNA clamp-loaders RFC1–5 (Havugimana et al., 2012; Kubota et al., 2013). Similarly, RUVBL1 are copurified ring-shaped proteins implicated in DNA damage response (Jha and Dutta, 2009; Rosenbaum et al., 2013). PARP1 may provide a link between nuclear and mitochondrial-derived interactors of ORF2p via DNA damage repair and apoptosis (Hong et al., 2004), and interestingly, PARP1 modifies both PCNA and TOP1, which are both high-confidence interactors of ORF2p (Simbulan-Rosenthal et al., 1999). We did not recover any APOBEC3 family cytidine deaminases, which are known inhibitors of LINE1 activity with mixed nuclear and cytoplasmic localizations (Bogerd et al., 2006); however, these proteins may not be expressed highly (or at all) in HEK293T cells, and this represents a general limitation of this study. Knockdown of endogenous PARP1 and RUVBL1 resulted in reduced retrotransposition, suggesting additional roles for other enzymes involved in generic DNA repair in the retrotransposition pathway.

Distinct RNPs at Different Life Cycle Stages?

Our data are consistent with isolation of L1 RNPs of at least three distinct types, (Figure 7A). First, some RNPs purified using ORF1p are likely isolated from cells not expressing ORF2p, perhaps a state in which the host has defeated L1 by preventing ORF2p expression (RNP III). Focusing on the tandem pullout (Figure 5A), we identify two additional distinct populations: (1) the elution, containing all detectable ORF1p, the majority of UPF1, and only a trace amount of PCNA (RNP I) and (2) the supernatant depleted for ORF1p but containing the vast majority of the PCNA and a minority—but nevertheless substantial amount—of UPF1 (RNP II).

Taken together with the idea that PCNA may be tightly and specifically bound to ORF2p only after initiation of TPRT, this suggests that these two populations of L1 RNPs, while both retaining ORF2p, interact with ORF1p, PCNA, and UPF1 in distinct ways—RNP I is the tandem-purified RNP, present in the cytoplasm and primarily containing ORF1p/ORF2p/UPF1 in which UPF1 is active in a negative regulatory capacity, and RNP II is a distinct complex in nuclei, primarily containing ORF2p/PCNA with lesser amounts of UPF1 and seemingly lacking ORF1p. This suggests that ORF1p may be absent in steps subsequent to EN activity. Interestingly, UPF1 has been linked to Pol δ in DNA replication/repair (Azzalin and Lingner, 2006) and may play a second role here as a permissive host factor; this could explain the unexpectedly reduced retrotransposition in the context of UPF1 knockdown.

EXPERIMENTAL PROCEDURES

Materials and experimental procedures are detailed in the Extended Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2013.10.021.

REFERENCES

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

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

Arjan-Odedra, S., Swanson, C.M., Sherer, N.M., Wolinsky, S.M., and Malim, M.H. (2012). Endogenous MOV10 inhibits the retrotransposition of endogenous retroelements but not the replication of exogenous retroviruses. Retrovirology *9*, 53.

Athanikar, J.N., Badge, R.M., and Moran, J.V. (2004). A YY1-binding site is required for accurate human LINE-1 transcription initiation. Nucleic Acids Res. *32*, 3846–3855.

Azzalin, C.M., and Lingner, J. (2006). The human RNA surveillance factor UPF1 is required for S phase progression and genome stability. Curr. Biol. *16*, 433–439.

Beattie, T.R., and Bell, S.D. (2011). The role of the DNA sliding clamp in Okazaki fragment maturation in archaea and eukaryotes. Biochem. Soc. Trans. *39*, 70–76.

Beauregard, A., Curcio, M.J., and Belfort, M. (2008). The take and give between retrotransposable elements and their hosts. Annu. Rev. Genet. *42*, 587–617.

Beck, C.R., Garcia-Perez, J.L., Badge, R.M., and Moran, J.V. (2011). LINE-1 elements in structural variation and disease. Annu. Rev. Genomics Hum. Genet. *12*, 187–215.

Belancio, V.P., Whelton, M., and Deininger, P. (2007). Requirements for poly-adenylation at the 3′ end of LINE-1 elements. Gene *390*, 98–107.

Bogerd, H.P., Wiegand, H.L., Hulme, A.E., Garcia-Perez, J.L., O'Shea, K.S., Moran, J.V., and Cullen, B.R. (2006). Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. Proc. Natl. Acad. Sci. USA *103*, 8780–8785.

Brouha, B., Schustak, J., Badge, R.M., Lutz-Prigge, S., Farley, A.H., Moran, J.V., and Kazazian, H.H., Jr. (2003). Hot L1s account for the bulk of retrotransposition in the human population. Proc. Natl. Acad. Sci. USA *100*, 5280–5285.

Bubeck, D., Reijns, M.A., Graham, S.C., Astell, K.R., Jones, E.Y., and Jackson, A.P. (2011). PCNA directs type 2 RNase H activity on DNA replication and repair substrates. Nucleic Acids Res. *39*, 3652–3666.

Burns, K.H., and Boeke, J.D. (2012). Human transposon tectonics. Cell *149*, 740–752.

Buzdin, A., Ustyugova, S., Gogvadze, E., Vinogradova, T., Lebedev, Y., and Sverdlov, E. (2002). A new family of chimeric retrotranscripts formed by a full copy of U6 small nuclear RNA fused to the 3′ terminus of l1. Genomics *80*, 402–406.

Cost, G.J., Feng, Q., Jacquier, A., and Boeke, J.D. (2002). Human L1 element target-primed reverse transcription in vitro. EMBO J. *21*, 5899–5910.

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Marchetto, M.C., Muotri, A.R., Mu, Y., Carson, C.T., Macia, A., Moran, J.V., and Gage, F.H. (2011). Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. Proc. Natl. Acad. Sci. USA *108*, 20382–20387.

Cristea, I.M., Williams, R., Chait, B.T., and Rout, M.P. (2005). Fluorescent proteins as proteomic probes. Mol. Cell. Proteomics *4*, 1933–1941.

Dai, L., Taylor, M.S., O'Donnell, K.A., and Boeke, J.D. (2012). Poly(A) binding protein C1 is essential for efficient L1 retrotransposition and affects L1 RNP formation. Mol. Cell. Biol. *32*, 4323–4336.

Di Virgilio, M., Callen, E., Yamane, A., Zhang, W., Jankovic, M., Gitlin, A.D., Feldhahn, N., Resch, W., Oliveira, T.Y., Chait, B.T., et al. (2013). Rif1 prevents resection of DNA breaks and promotes immunoglobulin class switching. Science *339*, 711–715.

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

Doucet, A.J., Hulme, A.E., Sahinovic, E., Kulpa, D.A., Moldovan, J.B., Kopera, H.C., Athanikar, J.N., Hasnaoui, M., Bucheton, A., Moran, J.V., and Gilbert, N. (2010). Characterization of LINE-1 ribonucleoprotein particles. PLoS Genet. *6*, e1001150.

Farkash, E.A., Kao, G.D., Horman, S.R., and Prak, E.T. (2006). Gamma radiation increases endonuclease-dependent L1 retrotransposition in a cultured cell assay. Nucleic Acids Res. *34*, 1196–1204.

Feng, Q., Moran, J.V., Kazazian, H.H., Jr., and Boeke, J.D. (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell *87*, 905–916.

Fuchs, G., Stein, A.J., Fu, C., Reinisch, K.M., and Wolin, S.L. (2006). Structural and biochemical basis for misfolded RNA recognition by the Ro autoantigen. Nat. Struct. Mol. Biol. *13*, 1002–1009.

Goodier, J.L., and Kazazian, H.H., Jr. (2008). Retrotransposons revisited: the restraint and rehabilitation of parasites. Cell *135*, 23–35.

Goodier, J.L., Zhang, L., Vetter, M.R., and Kazazian, H.H., 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.

Goodier, J.L., Mandal, P.K., Zhang, L., and Kazazian, H.H., Jr. (2010). Discrete subcellular partitioning of human retrotransposon RNAs despite a common mechanism of genome insertion. Hum. Mol. Genet. *19*, 1712–1725.

Goodier, J.L., Cheung, L.E., and Kazazian, H.H., Jr. (2012). MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. PLoS Genet. *8*, e1002941.

Goodier, J.L., Cheung, L.E., and Kazazian, H.H., Jr. (2013). Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. Nucleic Acids Res. *41*, 7401–7419.

Han, J.S., and Boeke, J.D. (2004). A highly active synthetic mammalian retrotransposon. Nature *429*, 314–318.

Hata, K., and Sakaki, Y. (1997). Identification of critical CpG sites for repression of L1 transcription by DNA methylation. Gene *189*, 227–234.

Havugimana, P.C., Hart, G.T., Nepusz, T., Yang, H., Turinsky, A.L., Li, Z., Wang, P.I., Boutz, D.R., Fong, V., Phanse, S., et al. (2012). A census of human soluble protein complexes. Cell *150*, 1068–1081.

Hernan, R., Heuermann, K., and Brizzard, B. (2000). Multiple epitope tagging of expressed proteins for enhanced detection. Biotechniques *28*, 789–793.

Hogg, J.R., and Goff, S.P. (2010). Upf1 senses 3'UTR length to potentiate mRNA decay. Cell *143*, 379–389.

Hong, S.J., Dawson, T.M., and Dawson, V.L. (2004). Nuclear and mitochondrial conversations in cell death: PARP-1 and AIF signaling. Trends Pharmacol. Sci. *25*, 259–264.

Jha, S., and Dutta, A. (2009). RVB1/RVB2: running rings around molecular biology. Mol. Cell *34*, 521–533.

Khan, H., Smit, A., and Boissinot, S. (2006). Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. Genome Res. *16*, 78–87.

Khazina, E., Truffault, V., Büttner, R., Schmidt, S., Coles, M., and Weichenrieder, O. (2011). Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. Nat. Struct. Mol. Biol. *18*, 1006–1014.

Kimberland, M.L., Divoky, V., Prchal, J., Schwahn, U., Berger, W., and Kazazian, H.H., Jr. (1999). Full-length human L1 insertions retain the capacity for high frequency retrotransposition in cultured cells. Hum. Mol. Genet. *8*, 1557–1560.

Kubota, T., Nishimura, K., Kanemaki, M.T., and Donaldson, A.D. (2013). The Elg1 replication factor C-like complex functions in PCNA unloading during DNA replication. Mol. Cell *50*, 273–280.

Kulpa, D.A., and Moran, J.V. (2005). Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. Hum. Mol. Genet. *14*, 3237–3248.

Kulpa, D.A., and Moran, J.V. (2006). Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. Nat. Struct. Mol. Biol. *13*, 655–660.

Lee, D.J., Busby, S.J.W., Westblade, L.F., and Chait, B.T. (2008). Affinity isolation and I-DIRT mass spectrometric analysis of the Escherichia coli O157:H7 Sakai RNA polymerase complex. J. Bacteriol. *190*, 1284–1289.

Luan, D.D., and Eickbush, T.H. (1995). RNA template requirements for target DNA-primed reverse transcription by the R2 retrotransposable element. Mol. Cell. Biol. *15*, 3882–3891.

Malik, H.S., Burke, W.D., and Eickbush, T.H. (1999). The age and evolution of non-LTR retrotransposable elements. Mol. Biol. Evol. *16*, 793–805.

Mandal, P.K., Ewing, A.D., Hancks, D.C., and Kazazian, H.H., Jr. (2013). Enrichment of processed pseudogene transcripts in L1-ribonucleoprotein particles. Hum. Mol. Genet. *22*, 3730–3748.

Martin, S.L., and Bushman, F.D. (2001). Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. Mol. Cell. Biol. *21*, 467–475.

Mathias, S.L., Scott, A.F., Kazazian, H.H., Jr., Boeke, J.D., and Gabriel, A. (1991). Reverse transcriptase encoded by a human transposable element. Science *254*, 1808–1810.

Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat. Genet. *36*, 1073–1078.

Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D., and Kazazian, H.H., Jr. (1996). High frequency retrotransposition in cultured mammalian cells. Cell *87*, 917–927.

Morrish, T.A., Gilbert, N., Myers, J.S., Vincent, B.J., Stamato, T.D., Taccioli, G.E., Batzer, M.A., and Moran, J.V. (2002). DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nat. Genet. *31*, 159–165.

Muotri, A.R., Chu, V.T., Marchetto, M.C., Deng, W., Moran, J.V., and Gage, F.H. (2005). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature *435*, 903–910.

Niewiadomska, A.M., Tian, C., Tan, L., Wang, T., Sarkis, P.T., and Yu, X.F. (2007). Differential inhibition of long interspersed element 1 by APOBEC3 does not correlate with high-molecular-mass-complex formation or P-body association. J. Virol. *81*, 9577–9583.

O'Donnell, K.A., An, W., Schrum, C.T., Wheelan, S.J., and Boeke, J.D. (2013). Controlled insertional mutagenesis using a LINE-1 (ORFeus) gene-trap mouse model. Proc. Natl. Acad. Sci. USA *110*, E2706–E2713.

Oeffinger, M., Wei, K.E., Rogers, R., DeGrasse, J.A., Chait, B.T., Aitchison, J.D., and Rout, M.P. (2007). Comprehensive analysis of diverse ribonucleoprotein complexes. Nat. Methods *4*, 951–956.

Ostertag, E.M., and Kazazian, H.H., Jr. (2001). Biology of mammalian L1 retrotransposons. Annu. Rev. Genet. *35*, 501–538.

Peddigari, S., Li, P.W., Rabe, J.L., and Martin, S.L. (2013). hnRNPL and nucleolin bind LINE-1 RNA and function as host factors to modulate retrotransposition. Nucleic Acids Res. *41*, 575–585.

Repanas, K., Zingler, N., Layer, L.E., Schumann, G.G., Perrakis, A., and Weichenrieder, O. (2007). Determinants for DNA target structure selectivity of the human LINE-1 retrotransposon endonuclease. Nucleic Acids Res. *35*, 4914–4926.

Repanas, K., Fuentes, G., Cohen, S.X., Bonvin, A.M., and Perrakis, A. (2009). Insights into the DNA cleavage mechanism of human LINE-1 retrotransposon endonuclease. Proteins *74*, 917–928.

Rosenbaum, J., Baek, S.H., Dutta, A., Houry, W.A., Huber, O., Hupp, T.R., and Matias, P.M. (2013). The emergence of the conserved AAA+ ATPases Pontin and Reptin on the signaling landscape. Sci. Signal. *6*, mr1.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

Shigeoka, T., Kato, S., Kawaichi, M., and Ishida, Y. (2012). Evidence that the Upf1-related molecular motor scans the 3′-UTR to ensure mRNA integrity. Nucleic Acids Res. *40*, 6887–6897.

Shukla, R., Upton, K.R., Muñoz-Lopez, M., Gerhardt, D.J., Fisher, M.E., Nguyen, T., Brennan, P.M., Baillie, J.K., Collino, A., Ghisletti, S., et al. (2013). Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. Cell *153*, 101–111.

Simbulan-Rosenthal, C.M., Rosenthal, D.S., Iyer, S., Boulares, H., and Smulson, M.E. (1999). Involvement of PARP and poly(ADP-ribosyl)ation in the early stages of apoptosis and DNA replication. Mol. Cell. Biochem. *193*, 137–148.

Smits, A.H., Jansen, P.W.T.C., Poser, I., Hyman, A.A., and Vermeulen, M. (2013). Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. Nucleic Acids Res. *41*, e28.

Soifer, H.S., Zaragoza, A., Peyvan, M., Behlke, M.A., and Rossi, J.J. (2005). A potential role for RNA interference in controlling the activity of the human LINE-1 retrotransposon. Nucleic Acids Res. *33*, 846–856.

Suzuki, J., Yamaguchi, K., Kajikawa, M., Ichiyanagi, K., Adachi, N., Koyama, H., Takeda, S., and Okada, N. (2009). Genetic evidence that the non-homologous end-joining repair pathway is involved in LINE retrotransposition. PLoS Genet. *5*, e1000461.

Tackett, A.J., DeGrasse, J.A., Sekedat, M.D., Oeffinger, M., Rout, M.P., and Chait, B.T. (2005). I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. J. Proteome Res. *4*, 1752–1756.

Tani, H., Imamachi, N., Salam, K.A., Mizutani, R., Ijiri, K., Irie, T., Yada, T., Suzuki, Y., and Akimitsu, N. (2012). Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. RNA Biol. *9*, 1370–1379.

Ulrich, H.D. (2011). Timing and spacing of ubiquitin-dependent DNA damage bypass. FEBS Lett. *585*, 2861–2867.

Wei, W., Gilbert, N., Ooi, S.L., Lawler, J.F., Ostertag, E.M., Kazazian, H.H., Boeke, J.D., and Moran, J.V. (2001). Human L1 retrotransposition: cis preference versus trans complementation. Mol. Cell. Biol. *21*, 1429–1439.

Weichenrieder, O., Repanas, K., and Perrakis, A. (2004). Crystal structure of the targeting endonuclease of the human LINE-1 retrotransposon. Structure *12*, 975–986.

Yang, N., and Kazazian, H.H., Jr. (2006). L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat. Struct. Mol. Biol. *13*, 763–771.