Mobile Genetic Elements 4, e29319; January–December 2014; © 2014 Landes Bioscience

Expression and detection of LINE-1 ORF-encoded proteins

Lixin Dai,[1] John LaCava,[2] Martin S Taylor,[3] and Jef D Boeke[4,*]

[1]High Throughput Biology Center and Department of Molecular Biology and Genetics; Johns Hopkins University School of Medicine; Baltimore, MD USA; [2]Laboratory of Cellular and Structural Biology; The Rockefeller University; New York, NY USA; [3]High Throughput Biology Center and Department of Pharmacology and Molecular Sciences; Johns Hopkins University School of Medicine; Baltimore, MD USA; [4]Institute for Systems Genetics; New York University School of Medicine; New York, NY USA

LINE-1 (L1) elements are endogenous retrotransposons active in mammalian genomes. The L1 RNA is bicistronic, encoding two non-overlapping open reading frames, *ORF1* and *ORF2*, whose protein products (ORF1p and ORF2p) bind the L1 RNA to form a ribonucleoprotein (RNP) complex that is presumed to be a critical retrotransposition intermediate. However, ORF2p is expressed at a significantly lower level than ORF1p; these differences are thought to be controlled at the level of translation, due to a low frequency ribosome reinitiation mechanism controlling *ORF2* expression. As a result, while ORF1p is readily detectable, ORF2p has previously been very challenging to detect in vitro and in vivo. To address this, we recently tested several epitope tags fused to the N- or C-termini of the ORF proteins in an effort to enable robust detection and affinity purification from native (L1RP) and synthetic (*ORFeus*-Hs) L1 constructs. An analysis of tagged RNPs from both L1RP and *ORFeus*-Hs showed similar host-cell-derived protein interactors. Our observations also revealed that the tag sequences affected the retrotransposition competency of native and synthetic L1s differently although they encode identical ORF proteins. Unexpectedly, we observed apparently stochastic expression of ORF2p within seemingly homogenous L1-expressing cell populations.

Keywords: LINE-1, mass spectrometry, I-DIRT, synthetic transposon, epitope tag

Submitted: 03/07/2014

Revised: 05/19/2014

Accepted: 05/22/2014

Published Online: 05/22/2014

Citation: Dai L, LaCava J, Taylor MS, Boeke JD. Expression and detection of LINE-1 ORF-encoded proteins. Mobile Genetic Elements 2014; 4:e29319; http://dx.doi.org/10.4161/mge.29319

*Correspondence to: Jef D Boeke; Email: Jef.Boeke@nyumc.org

Commentary to: Taylor MS, Lacava J, Mita P, Molloy KR, Huang CR, Li D, Adney EM, Jiang H, Burns KH, Chait BT, et al. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. Cell 2013; 155:1034–48; PMID:24267889; http://dx.doi.org/10.1016/j.cell.2013.10.021

Long Interspersed Element-1s (LINE-1s or L1s) are the only active endogenous autonomous transposons in the human genome.[1] Though most copies are inactive due to truncation, there are ~80–100 retrotransposition-competent L1s in any individual's diploid genome.[2] L1 genomic sequences are ~6 kb in length. The promoter region driving transcription of these elements is internal and encodes the 5′ UTR in the resulting bicistronic L1 RNA, which contains two non-overlapping open reading frames (*ORF1* and *ORF2*).[3] ORF1p is a ~40 kDa polypeptide that assembles as a homotrimeric protein and functions as a nucleic acid chaperone.[4,5] ORF2p is a ~150 kDa multidomain protein with endonuclease[6] and reverse transcriptase functions,[7] as well as a cysteine rich domain with unknown function.[8] A ribonucleoprotein (RNP) complex comprised of the L1 RNA and both ORF proteins is believed to form in the cytoplasm.[9,10] Upon subsequent translocation of the RNP to the nucleus, it may generate a new chromosomal insertion via a target primed reverse transcription mechanism (TPRT).[11,12]

Both ORF proteins are required for successful RNP formation and retrotransposition in a cell culture based assay.[13] However, the precise details of particle assembly and the timing and mechanisms of L1 molecular physiology have remained elusive. Hence, it has been of great interest to monitor both ORF proteins throughout the L1 life cycle. Because of its relatively high abundance, ORF1p (endogenous and overexpressed) from both human and mouse L1s have been readily detected by western blot and immunostaining using anti-ORF1p antibodies.[14-18] ORF1p has also been successfully epitope tagged with

e.g., T7 and FLAG tags.[10,19] On the other hand, detecting ORF2p has constituted a major challenge. Previously antibodies targeted to EN and C domains were shown to be able to detect ORF2p from human and rat L1s by western blot, albeit with low signal strength.[16,20-22] Short epitope tags have been less efficacious, although two longer C-terminal tags (GST and TAP) have been used to successfully detect ORF2p by western blot.[11,20] Recently we exploited a 3xFlag epitope tag with a long linker sequence at the C-terminus of ORF2p and were successful in readily detecting ORF2p.[23]

An obvious potential cause for ORF2p's recalcitrance to detection is its low expression level from native L1 RNA sequences. Endogenous L1 genes have high adenosine content (up to 40%) and as a result, can suffer from a transcription elongation defect.[24] Moreover, ORF2p is believed to be translated by an unconventional mechanism at ~one molecule per RNA.[23,25] Our group has synthesized two synthetic L1 elements (*ORFeus*-Mm and *ORFeus*-Hs) by codon optimizing L1s from mouse and human respectively.[17,26] These two elements have provided significantly enhanced levels of RNA transcription and ORF protein expression, and now serve as useful tools for the study of L1 molecular biology.

Translation of L1 ORFs

The two non-overlapping open reading frames comprising the L1 bicistron are separated by a 63 bp inter-ORF sequence. In contrast to viruses that translate their genes as a polyprotein or via translational frame shift,[27] ORF1p and ORF2p are believed to be translated sequentially through an unconventional mechanism.[25] It has been reported that there are internal ribosome entry sites (IRES) upstream of each mouse LINE-1 ORF, implying that they may be translated via an IRES-mediated translation initiation mechanism.[28] But recoded *ORFeus*-Mm, which lacks these IRES sequences, transposes at a much higher rate than endogenous mouse L1, suggesting that IRES-mediated translation is not the major mechanism for LINE-1 ORFs or recoded *ORFeus*-Mm adopts a different mechanism.[26] In human L1s, no such IRES has been identified, and the 3′-end of *ORF1* as well as the inter-ORF spacer are dispensable for ORF2p translation.[25] Furthermore, successful ORF2p translation is independent of its initiation codon AUG and of ORF1p functionality but requires completion of an upstream ORF (though it need not be *ORF1* specifically).[25] These properties of ORF2p are all in the context of the bicistronic transcript; when expressed as a monocistronic RNA from a conventional promoter, ORF2p is expressed at much higher levels.[23]

Until recently, it was widely postulated that the L1 RNA was fully coated with ORF1p trimers, but this hypothesis had never been tested.[20] A previous study on mouse ORF1p showed that each trimer captures 50 nt of RNA.[29] Based on the length of the RNA and the properties of the ORF1p timer,[4] we calculated that a fully coated L1 RNA would contain as many as ~240 ORF1p molecules;[23] and were such to be the case, it would support a model of ORF protein expression including numerous rounds of ORF1p translation and only infrequent ORF2p translation. Evidence suggests that upon successful translation, ORF2p joins the L1 RNA to form an RNP in *cis*;[9,30] therefore, it is conceivable that *ORF2* RNA sequences may no longer be available for further translation upon RNP complex formation.

As little direct evidence of L1 stoichiometry exists, we assayed ORF protein abundance at both cell-population and molecular levels. Our results revealed an unexpected, stochastic pattern of ORF2p expression within cultured cells. Nearly all cells in the population expressed ORF1p in abundance, but within this seemingly homogenous cell population, the majority of cells (~70%) failed to express ORF2p.[23] In the remaining ~30%, ORF2p expression is robust. ORF1p expression levels were found to be comparable in both ORF2p-expressing and non-expressing cells, thus, no connection between the level of ORF1p and ORF2p expression was found. In contrast, an ORF2p-only construct driven by a canonical Pol II promoter is expressed in > 95% of cells. We confirmed these observations using both *ORFeus*-Hs and L1RP in HeLa and HEK293T cell lines transfected under varying conditions, suggesting that this pattern of expression may be an intrinsic property of the human L1 bicistronic RNA.

At the level of the purified proteins, we observed some differences in the complement of proteins exhibiting specific co-enrichment with ORF1p and ORF2p. For example, TROVE2, known to bind to misfolded RNAs, and MEPCE, the 7SK snRNA methylphosphate capping enzyme[31] were only identified as specific interactors in fractions affinity purified by ORF1p. Given the broader expression of ORF1p within the cell population, fractions captured by ORF1p from whole cell extracts must contain an abundance of co-purifying material originating from cells not expressing ORF2p. These complexes from ORF1p-only expressing cells may either comprise a comparable subset of those present in ORF2p expressing cells or be distinct "non-functional" particles.

To directly examine the stoichiometry of ORF proteins in L1 RNPs, we established a two-dimensional affinity purification procedure. After first affinity purifying 3xFlag-tagged ORF2p from whole cell extracts, natively eluted complexes were subsequently further purified by ORF1p using an antibody against the native protein. Thus, in this tandem enrichment procedure we obtained a fraction of L1 particles i) containing an ORF1p population in physical association with the co-purifying ORF2p throughout the procedure, and ii) which was also separated from extraneous ORF2p. We evaluated the amounts of ORF1 and ORF2 proteins in these particles selected for the presence of both proteins. Measuring by image densitometry using two staining methods, the ratio was estimated at ~6:1–9:1, and by label-free mass spectrometry (iBAQ),[32] the ratio was estimated at between 27:1 and 47:1.[23]

In the first estimation only full-length protein signals were counted. The second estimation by mass spectrometry counts both full-length and fragmented proteins that fall below the level of detection by staining. Regardless of the method preferred, all values are much lower than the expected ORF1p:ORF2p ratio of ~240:1

if one assumes a single ORF2p per RNP. Potential explanations for this discrepancy include: the L1 RNA could form secondary and tertiary structure rather than being a linear molecule, and as such, the RNPs may not be fully coated by ORF1p; the overexpression of L1 may result in unnatural component stoichiometries; and/or L1 RNPs naturally contain > 1 ORF2p.

We assayed the lability of the ORF proteins in affinity captured particles after treatment with RNases and observed that ORF1p could be released regardless of the capture method.[23] These results suggest that a multiplicity of ORF1p trimers co-purify in an RNase sensitive way, e.g., by decorating co-purifying RNA, and that the maintenance of interactions between ORF1p and ORF2p is also dependent upon the presence of RNA. In contrast, we observed that ORF2p-3xFlag in L1 RNPs immobilized on anti-Flag affinity medium was insensitive to RNase treatment, while other components of the RNP were RNase sensitive.[23] These contrasting data suggest that the behavior of ORF2p-3xFlag is consistent with the one-ORF2p-per-RNP model. Nevertheless, these data do not exclude the possibility that multiple ORF2ps are incorporated within these RNPs and retained upon the capture medium through simultaneous direct affinity interactions, or that any released ORF2p-3xFlag is rapidly recaptured upon the affinity medium during RNase treatment.

Further study may be required to exclude the possibility that a biased or incomplete RNP was enriched in our tandem purification process. However, we noted that affinity purifying ORF2p-3xFlag from whole cell extract at a lower concentration of NaCl (300 mM as opposed to 500 mM) decreased both the yield of ORF2p-3xFlag as well as ORF1p without an apparent significant change in their relative ratios.[23] This demonstrates that a more thorough extraction was achieved at the higher salt concentration and that the stability of ORF1 was not noticeably affected, providing a higher total yield of complexes containing both proteins at a comparable relative ratio.

How host cells regulate the initiation of ORF2p translation, and the potential functions of any ORF1p-only RNPs, both remain unknown; but we speculate that these are related phenomenon. It is certainly conceivable that host defenses, trained on preventing L1 proliferation, include mechanisms specifically targeting ORF2p translation. This intriguing phenomenon is not yet remotely understood, and would be a worthwhile topic not only for those studying transposons but also for researchers working on eukaryotic translation mechanisms.

Detecting LINE-1 Elements: Differences between Native and Synthetic Constructs

In order to affinity purify LINE-1 RNP complexes we systematically tested various epitope tags at different locations of both ORFs from *ORFeus*-Hs.[23] To assess the biological impact of various tagging strategies, we employed a retrotransposition assay as a test of function because retrotransposition requires the activity of both proteins.[13] First we found that a single Flag tag, adding only 8 amino acids to the N-terminus of ORF1p, almost completely abolished L1 retrotransposition activity (0.2% of wild type). Since an N-terminus coiled coil domain is responsible for ORF1p trimerization, extra N-terminal tag sequence may disrupt proper trimerization, which is important to L1 activity.[4,5] Compared with the N-terminus, the C-terminus of ORF1p tolerates tags. A single Flag tag had no effect on L1 retrotransposition and a longer tag of 48 amino acids including a flexible linker, TEV protease cleavage site, V5,[33] HA[34] and Flag[35] only decreased L1 activity by 2%. This tagged ORF1p was detected by western blot using V5, HA and Flag antibodies without difficulty. When we slightly increased the tag size by replacing single Flag with 3xFlag, L1 activity quickly dropped to 30% of wild type. But a similar 3xFlag at the C-terminus of ORF2p only had minimal impact to L1 activity. The 3xFlag tag is reported by Sigma-Aldrich to be 20–200 times more sensitive than other systems and this enhancement now makes detection of ORF2p routine. In a separate study on a different protein complex, we found that when the salt concentration was increased from 100 to 300 mM NaCl during affinity purification, single Flag signal is either completely or almost completely lost (by Coomassie staining and Western) whereas 3xFlag tagged protein was robustly purified in a high salt buffer.[36] We also attempted to tag both ORFs with a fluorescent GFP tag and found that this longer tag inhibited L1 retrotransposition by ~25% on either ORF1p or ORF2p C-terminus. ORF1p-GFP fusion exhibited strong signal in-cell by immunostaining, but a small yet significant population of free GFP was revealed by western blotting, suggesting a cleavage between ORF1p and the tag sequence (unpublished observation). When the linker sequence between ORF1p and GFP increased by 6 amino acids, L1 transposition dropped to 23% of wild type. ORF2p-GFP fusion protein was barely detectable in vivo by microscopy or in vitro by western blotting.

Interestingly, native LINE-1 (L1RP) behaves differently when the same tag sequences were placed at the same location relative to the ORFs. For example, the 48 AA multi-tag (V5, HA, Flag) at the C-terminus of ORF1p reduced L1RP activity by more than 30%, which was deemed unsuitable for further study.[23] Similarly, the 3xFlag tag, successfully appended ORF2p in the in *ORFeus*-Hs construct, was only compatible with ORF2p expressed from L1RP once the linker region was reduced by 2/3. We speculate that the differential sensitivity of L1RP and *ORFeus*-Hs to tag sequence is most likely due to different expression levels, but cannot rule out effects due to RNA structure or sequence-specific binding elements.

Utilizing the epitope tags, we analyzed the sub-cellular localization of both L1 ORFs. Similar to several previous studies, we observed that both ORFs are epredominantly cytoplasmic. Although the presence of active retrotransposition in these cells demands that ORF2p must be in nucleus at a certain stage of the L1 life cycle, this population was not observed.[23] In contrast, the literature documents several cases in which both ORFs were detectable in both nucleus and cytoplasm.[21,22] The difference may also be due to different sensitivity or cross reactivity of

various antibodies, or to different behavior of endogenously expressed proteins vs. overexpressed proteins. Notably, when a massive overexpression system using a modified vaccinia virus/T7 RNA polymerase was used in HeLa cells,[21] a very different pattern was observed than with the lower level of overexpression in our studies.[23] In depth comparison, immunostaining and live cell imaging analysis are needed to better understand how L1 ORFs travel in the cell. An additional possibility is that the localization of the two ORFs may be cell-cycle dependent. Mitotic cells are poorly retained through most 2D immunofluorescent staining protocols due to reduced adherence, and cell division is required for L1 retrotransposition. We summarize the detection of L1 encoded ORFs in **Table 1**.

Synthetic vs. Native L1: Sequence and Context Dependent Expression Differences

The variable impact of these sequence insertions may reflect context-dependent expression differences between synthetic and endogenous L1s. By lowering adenosine content to 25% and codon optimization, a synthetic mouse LINE-1 element (*ORFeus*-Mm) displayed a > 200-fold increase in retrotransposition frequency. Using the same strategy, a synthetic human L1 (*ORFeus*-Hs) was generated based on the L1RP sequence, resulting in a ~40 fold increase in RNA and ~3–5-fold increase in ORF1p.[17] The 3xFlag tag allowed us to quantify the ORF2p expressed from *ORFeus*-Hs, which is ~40 fold of that from L1RP. In contrast to the drastic increase in retrotransposition frequency observed for *ORFeus*-Mm, *ORFeus*-Hs retrotransposes at essentially similar levels as L1RP (which is also comparable to that of *ORFeus*-Mm).

An explanation for the failure to observe increased transposition frequency when expressing L1 from *ORFeus*-Hs, as compared with what was obtained with expression from *ORFeus*-Mm, is a different baseline activity level: while both synthetic constructs transpose at similar levels, L1RP already actively transposes in cell culture based assays with an absolute efficiency of ~10–15% in HEK293T cells or ~5–10% in HeLa cells, as monitored by a GFPAI marker,[17] whereas the baseline activity of mouse L1 is ~0.1%. As mentioned above, we observed that in any snapshot in time, only ~30% of the cells transfected with L1 expression constructs expressed both ORF1p and ORF2p and the remainder expressed ORF1p only.[23] If these cell populations are stable, meaning that cells that fail to express ORF2p do not regain this ability later, this may account in-part for the limitation in retrotransposition frequency, but cannot account for all of it.

As indicated by mfold,[40] the extensive recoding of native human L1 during the construction of *ORFeus*-Hs most likely changed any specific sequence motif or RNA structure (secondary or tertiary) that may have resided within L1RP coding regions. In mass spectrometry-based proteomic experiments aimed at identifying stable constituents of L1 RNPs we identified comparable sets of proteins for both *ORFeus*-Hs and L1RP RNPs. This argues against gross differences in the synthetic and WT particles as a result of sequence specific protein interactors with RNA (at least among the most stable constituents). Additionally, the L1RP constructs contained the 5′UTR, whereas the *ORFeus*-Hs constructs did not, similarly arguing against sequence-specific binding partners in this region of the RNA. Therefore, it is plausible that the L1 life cycle is driven primarily by the L1-encoded ORFs, their binding partners, and RNA sequence/structure insensitive binding proteins, but we cannot exclude the existence of sequence/structure-specific host-encoded nucleic acid binding partners.

Though synthetic and native L1s formed similar RNP complexes, interesting differences exist. As mentioned earlier, epitope tags that behaved well in the *ORFeus*-Hs context didn't necessarily work efficiently in the L1RP context. It has been suggested that the abundance of certain tRNAs, local mRNA structure, and translation rate all possibly affect the overall folding, activity and turnover rate of final polypeptide product.[41-47] It is known that transient translational pausing or ribosomal attenuation aids co-translational folding. We observed that an *ORFeus*-native chimera L1, containing 3/4 synthetic and 1/4 native coding sequence within ORF2p, acquired higher transposition efficiency (1.4-fold) than fully synthetic L1.[17] Since the native L1 sequence contains rare codons, which are absent in the codon-optimized *ORFeus*-Hs, it is conceivable that slowing down of ribosomes near the end of ORF2p translation may help assure the correct folding of ORF2 protein or better engagement of ORF2p to L1 RNA. This could lead to a higher retrotransposition frequency for the chimeric L1 element. While the 3′ UTR sequence is dispensable for L1 activity, we observed that the endogenous 5′ UTR sequence may affect L1 activity, especially for native L1. When L1 sequences were driven by a CMV promoter without the 5′ UTR, L1RP transposes at a lower frequency (~1/3) compared with *ORFeus*-Hs, but transposes at a similar level when the 5′ UTR was included in the constructs, in either the presence or absence of CMV promoter.[17] So it is possible that there is some positive "crosstalk" between 5′ UTR and native coding RNA sequence that is absent from in *ORFeus*-Hs.

Summary and Conclusions

Routine detection of LINE-1 ORF2p has been a major challenge in the field due to the difficulty of generating reliable antibodies. Epitope tagging with 3xFlag facilitates this goal by its high sensitivity and small size, without adversely affecting human L1 retrotranposition activity. We believe that a comparable tagging strategy is likely applicable to ORF sequences from other species. Synthetic L1 elements have further aided this situation by significantly enhancing the L1 expression level and providing a point of comparison to endogenous L1s, together advancing our understanding of the fundamental properties of this active "jumping gene" that has contributed > 20% of our genome.[48] The ability to affinity purify highly active L1 RNP complexes has allowed us to catalog its composition, assess functional relationships between constituents, and will enable additional assays targeted at biochemical and structural characterization of L1 transposition pathway. The ability

Table 1. Summary of antibodies and epitope tags used for L1 encoded ORFs detection

	Antigen	Epitope location	Epitope	RNA/protein expression	Retrotransposition efficiency	Detection	Endogenous or overexpressed	Subcellular localization	Ref.
ORF1p (mouse)	Full-length protein					Western blot, immunostaining	Endo	Cyto	15
	Full-length protein					Western blot	Over	ND[a]	37
ORF1p (rat)	Fragment (1–143)					Western blot, immunoprecipitation	Endo	Cyto and nuclear	22
ORF2p (rat)	Fragment (292–480)					Western blot, immunoprecipitation	Endo	Cyto and nuclear	22
ORF1p (human)	Full-length protein					Western blot, immunostaining	Endo	ND[a]	14
	Full-length protein					Western blot, immunostaining	Over	Cyto	38
	Full-length protein					Western blot	Endo	ND[a]	37
	Fragment					Western blot, immunostaining	Over	Cyto	23
	Peptide (35–44)					Western blot, immunoprecipitation	Endo	Cyto	18, 23
	Peptide (318–338)					Western blot, immunostaining	Over	Cyto	21
	Full-length protein					Western blot, immunostaining	Endo	Cyto and nuclear	39
		N[b]	GFP	ND[a]	ND[a]	immunostaining	Over	Cyto, nucleolus	21
		C[c]	T7	No change	Almost the same as untagged	Western blot, immunostaining	Over	Cyto	10, 20, 21
		C[c]	Flag-HA	ND[a]	ND[a]	Western blot, immunoprecipitation	Over	ND[a]	19
		C[c]	Flag	No change	63% of untagged	Western blot	Over	ND[a]	23
		C[c]	Myc	ND[1]	88% of untagged	Western blot	Over	ND[a]	23
		C[c]	3 × Myc	ND[1]	64% of untagged	Western blot	Over	ND[a]	23
ORF1p (*ORFeus*-Hs)		N[b]	Flag	No change	0.2% of untagged	Western blot	Over	ND[a]	23
		C[c]	Flag	No change	102% of untagged	Western blot	Over	ND[a]	23
		C[c]	HA-V5-Flag	No change	98% of untagged	Western blot, immunoprecipitation	Over	Cyto	23
		C[c]	GFP	ND[a]	12–73% of untagged	Western blot	Over	ND[a]	23
		C[c]	3 × Flag	No change	30% of untagged	Western blot	Over	ND[a]	23
		C[c]	Myc	No change	80% of untagged	Western blot	Over	ND[a]	23
		C[c]	3 × Myc	No change	64% of untagged	Western blot	Over	ND[a]	23

Table 1. Summary of antibodies and epitope tags used for L1 encoded ORFs detection (Continued)

	Antigen	Epitope location	Epitope	RNA/protein expression	Retrotransposition efficiency	Detection	Endogenous or overexpressed	Subcellular localization	Ref.
ORF2p (human)	Peptide in EN domain (154–167)					Western blot, immunostaining	Over	Cyto, nucleolus	21
	Peptide in EN domain (48–63 and 152–166)					Western blot, immunostaining	Over	Cyto	16
	Peptide in C domain (1259–1275)					Western blot, immunostaining	Over	Cyto, nucleolus	21
	Full-length protein					Western blot, immunostaining	Endo	Cyto	39
		N[b]	Flag	ND[a]	ND[a]	Western blot, immunostaining	Over	Cyto, nucleolus	21
		C[c]	5 × PPX, 3 × Flag	ND[a]	31% of untagged	Western blot	Over	ND	23
		C[c]	1 × PPX, 3 × Flag	ND[a]	157% of untagged	Western blot, immunostaining, immunoprecipitation	Over	Cyto	23
		C[c]	Tap	ND[a]	Almost the same as untagged	Western blot, immunostaining	Over	Cyto	20
ORF2p (*ORFeus*-Hs)		C[c]	3 × Flag	ND[a]	92% of untagged	Western blot, immunostaining, immunoprecipitation	Over	Cyto	23
		C[c]	GFP	ND[a]	15% of untagged	Western blot	Over	ND[a]	23

[a]Not determined, [b]N-terminus, [c]C-terminus

to monitor LINE-1 ORFs in cell culture has raised new questions about sub-cellular localization. We have already shown that ORF2p functional mutants (EN- and RT-) failed to interact with PCNA at a late stage of the L1 life cycle.[23] By applying the same technique to other ORF1p/ORF2p mutants, we should be able to further dissect the L1-host interactome and catalog the relationships with respect to specific step(s) of the L1 lifecycle. A complementary approach includes coupling large scale RNAi with L1 induction in suspension cell culture, since we have been able to easily generate shRNA-L1 co-expression vectors.[38] This will allow us to test which L1 RNP constituents affect the remaining complement of proteins' abilities to co-purify with the L1 RNP, as well as to assay effects on retrotransposition activity in vivo and in vitro.

One mysterious step in the L1 life cycle is TPRT, which is mostly hypothesized based on the similarity of L1 structure to another non-LTR retrotransposon - R2 element, and because it explains the phenomenon of target site duplication.[12] Previously full-length L1 ORF2 protein could only be obtained as a purified fraction from insect cells, and only inefficiently catalyzed a TPRT reaction in vitro.[11] Furthermore, though it displayed active reverse transcriptase activity, endonuclease activity was not detected until a partial proteolytic digestion was performed.[11] Our affinity purified RNP complex contains high RT and LEAP activity and it will be very interesting to compare ORF2p EN function from both sources. We expect to be able to perform a more efficient in vitro TPRT reaction on account of the quantity and purity of the L1 RNP complex we can obtain by our recently described methods.[23]

It has been hypothesized that excess retrotransposon activity in testicular germ cell tumors amplifies DNA damaging effects of chemo/radiotherapy and is the reason for their high sensitivity to therapy (Lisa Cheng Ran Huang, personal communication). We believe that a comparable approach should be applied in other cell-types such as germ line, stem cells and cancer cells, where LINE-1 has been found to actively transpose,[13,49,50]—such data are certain to support important hypotheses pertaining to L1 function in carcinogenesis and pluripotency/cell differentiation.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

This work was supported by NIH grants U54-GM103520 and P50-GM107632.

References

1. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al.; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 2001; 409:860-921; PMID:11237011; http://dx.doi.org/10.1038/35057062
2. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH Jr. Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci U S A 2003; 100:5280-5; PMID:12682288; http://dx.doi.org/10.1073/pnas.0831042100
3. Scott AF, Schmeckpeper BJ, Abdelrazik M, Comey CT, O'Hara B, Rossiter JP, Cooley T, Heath P, Smith KD, Margolet L. Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. Genomics 1987; 1:113-25; PMID:3692483; http://dx.doi.org/10.1016/0888-7543(87)90003-6
4. Khazina E, Truffault V, Büttner R, Schmidt S, Coles M, Weichenrieder O. Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. Nat Struct Mol Biol 2011; 18:1006-14; PMID:21822284; http://dx.doi.org/10.1038/nsmb.2097
5. Martin SL, Branciforte D, Keller D, Bain DL. Trimeric structure for an essential protein in L1 retrotransposition. Proc Natl Acad Sci U S A 2003; 100:13815-20; PMID:14615577; http://dx.doi.org/10.1073/pnas.2336221100
6. Feng Q, Moran JV, Kazazian HH Jr., Boeke JD. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 1996; 87:905-16; PMID:8945517; http://dx.doi.org/10.1016/S0092-8674(00)81997-2
7. Mathias SL, Scott AF, Kazazian HH Jr., Boeke JD, Gabriel A. Reverse transcriptase encoded by a human transposable element. Science 1991; 254:1808-10; PMID:1722352; http://dx.doi.org/10.1126/science.1722352
8. Fanning T, Singer M. The LINE-1 DNA sequences in four mammalian orders predict proteins that conserve homologies to retrovirus proteins. Nucleic Acids Res 1987; 15:2251-60; PMID:3562227; http://dx.doi.org/10.1093/nar/15.5.2251
9. Kulpa DA, Moran JV. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. Nat Struct Mol Biol 2006; 13:655-60; PMID:16783376; http://dx.doi.org/10.1038/nsmb1107
10. Kulpa DA, Moran JV. Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. Hum Mol Genet 2005; 14:3237-48; PMID:16183655; http://dx.doi.org/10.1093/hmg/ddi354
11. Cost GJ, Feng Q, Jacquier A, Boeke JD. Human L1 element target-primed reverse transcription in vitro. EMBO J 2002; 21:5899-910; PMID:12411507; http://dx.doi.org/10.1093/emboj/cdf592
12. Luan DD, Korman MH, Jakubczak JL, Eickbush TH. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 1993; 72:595-605; PMID:7679954; http://dx.doi.org/10.1016/0092-8674(93)90078-5
13. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH Jr. High frequency retrotransposition in cultured mammalian cells. Cell 1996; 87:917-27; PMID:8945518; http://dx.doi.org/10.1016/S0092-8674(00)81998-4
14. Bratthauer GL, Cardiff RD, Fanning TG. Expression of LINE-1 retrotransposons in human breast cancer. Cancer 1994; 73:2333-6; PMID:8168038; http://dx.doi.org/10.1002/1097-0142(19940501)73:9<2333::AID-CNCR2820730915>3.0.CO;2-4
15. Martin SL, Branciforte D. Synchronous expression of LINE-1 RNA and protein in mouse embryonal carcinoma cells. Mol Cell Biol 1993; 13:5383-92; PMID:8395003
16. Ergün S, Buschmann C, Heukeshoven J, Dammann K, Schnieders F, Lauke H, Chalajour F, Kilic N, Strätling WH, Schumann GG. Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. J Biol Chem 2004; 279:27753-63; PMID:15056671; http://dx.doi.org/10.1074/jbc.M312985200
17. An W, Dai L, Niewiadomska AM, Yetil A, O'Donnell KA, Han JS, Boeke JD. Characterization of a synthetic human LINE-1 retrotransposon ORFeus-Hs. Mob DNA 2011; 2:2; PMID:21320307; http://dx.doi.org/10.1186/1759-8753-2-2
18. Rodić N, Sharma R, Sharma R, Zampella J, Dai L, Taylor MS, Hruban RH, Iacobuzio-Donahue CA, Maitra A, Torbenson MS, et al. Long interspersed element-1 protein expression is a hallmark of many human cancers. Am J Pathol 2014; 184:1280-6; PMID:24607009; http://dx.doi.org/10.1016/j.ajpath.2014.01.007
19. Goodier JL, Cheung LE, Kazazian HH Jr. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. Nucleic Acids Res 2013; 41:7401-19; PMID:23749060; http://dx.doi.org/10.1093/nar/gkt512
20. Doucet AJ, Hulme AE, Sahinovic E, Kulpa DA, Moldovan JB, Kopera HC, Athanikar JN, Hasnaoui M, Bucheton A, Moran JV, et al. Characterization of LINE-1 ribonucleoprotein particles. PLoS Genet 2010; 6; PMID:20949108; http://dx.doi.org/10.1371/journal.pgen.1001150
21. Goodier JL, Ostertag EM, Engleka KA, Seleme MC, Kazazian HH Jr. A potential role for the nucleolus in L1 retrotransposition. Hum Mol Genet 2004; 13:1041-8; PMID:15028673; http://dx.doi.org/10.1093/hmg/ddh118
22. Kirilyuk A, Tolstonog GV, Damert A, Held U, Hahn S, Löwer R, Buschmann C, Horn AV, Traub P, Schumann GG. Functional endogenous LINE-1 retrotransposons are expressed and mobilized in rat chloroleukemia cells. Nucleic Acids Res 2008; 36:648-65; PMID:18073200; http://dx.doi.org/10.1093/nar/gkm1045
23. Taylor MS, Lacava J, Mita P, Molloy KR, Huang CR, Li D, Adney EM, Jiang H, Burns KH, Chait BT, et al. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. Cell 2013; 155:1034-48; PMID:24267889; http://dx.doi.org/10.1016/j.cell.2013.10.021
24. Han JS, Szak ST, Boeke JD. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. Nature 2004; 429:268-74; PMID:15152245; http://dx.doi.org/10.1038/nature02536
25. Alisch RS, Garcia-Perez JL, Muotri AR, Gage FH, Moran JV. Unconventional translation of mammalian LINE-1 retrotransposons. Genes Dev 2006; 20:210-24; PMID:16418485; http://dx.doi.org/10.1101/gad.1380406
26. Han JS, Boeke JD. A highly active synthetic mammalian retrotransposon. Nature 2004; 429:314-8; PMID:15152256; http://dx.doi.org/10.1038/nature02535
27. Gale M Jr., Tan SL, Katze MG. Translational control of viral gene expression in eukaryotes. Microbiol Mol Biol Rev 2000; 64:239-80; PMID:10839817; http://dx.doi.org/10.1128/MMBR.64.2.239-280.2000
28. Li PW, Li J, Timmerman SL, Krushel LA, Martin SL. 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 2006; 34:853-64; PMID:16464823; http://dx.doi.org/10.1093/nar/gkj490
29. Basame S, Wai-lun Li P, Howard G, Branciforte D, Keller D, Martin SL. Spatial assembly and RNA binding stoichiometry of a LINE-1 protein essential for retrotransposition. J Mol Biol 2006; 357:351-7; PMID:16434051; http://dx.doi.org/10.1016/j.jmb.2005.12.063
30. Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, Boeke JD, Moran JV. Human L1 retrotransposition: cis preference versus trans complementation. [doi]. Mol Cell Biol 2001; 21:1429-39; PMID:11158327; http://dx.doi.org/10.1128/MCB.21.4.1429-1439.2001
31. Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, Thérien C, Bergeron D, Bourassa S, Greenblatt J, et al. Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. Mol Cell 2007; 27:262-74; PMID:17643375; http://dx.doi.org/10.1016/j.molcel.2007.06.027
32. Smits AH, Jansen PW, Poser I, Hyman AA, Vermeulen M. Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. Nucleic Acids Res 2013; 41:e28; PMID:23066101; http://dx.doi.org/10.1093/nar/gks941
33. Southern JA, Young DF, Heaney F, Baumgärtner WK, Randall RE. Identification of an epitope on the P and V proteins of simian virus 5 that distinguishes between two isolates with different biological characteristics. J Gen Virol 1991; 72:1551-7; PMID:1713260; http://dx.doi.org/10.1099/0022-1317-72-7-1551
34. Chen YT, Holcomb C, Moore HP. Expression and localization of two low molecular weight GTP-binding proteins, Rab8 and Rab10, by epitope tag. Proc Natl Acad Sci U S A 1993; 90:6508-12; PMID:7688123; http://dx.doi.org/10.1073/pnas.90.14.6508
35. Hernan R, Heuermann K, Brizzard B. Multiple epitope tagging of expressed proteins for enhanced detection. Biotechniques 2000; 28:789-93; PMID:10769759
36. Domanski M, Molloy K, Jiang H, Chait BT, Rout MP, Jensen TH, LaCava J. Improved methodology for the affinity isolation of human protein complexes expressed at near endogenous levels. Biotechniques 2012; 0:1-6; PMID:22668517
37. Sokolowski M, deHaro D, Christian CM, Kines KJ, Belancio VP. Characterization of L1 ORF1p self-interaction and cellular localization using a mammalian two-hybrid system. PLoS One 2013; 8:e82021; PMID:24324740; http://dx.doi.org/10.1371/journal.pone.0082021
38. Dai L, Taylor MS, O'Donnell KA, Boeke JD. Poly(A) binding protein C1 is essential for efficient L1 retrotransposition and affects L1 RNP formation. Mol Cell Biol 2012; 32:4323-36; PMID:22907758; http://dx.doi.org/10.1128/MCB.06785-11
39. Chen L, Dahlstrom JE, Chandra A, Board P, Rangasamy D. Prognostic value of LINE-1 retrotransposon expression and its subcellular localization in breast cancer. Breast Cancer Res Treat 2012; 136:129-42; PMID:23053642; http://dx.doi.org/10.1007/s10549-012-2246-7
40. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003; 31:3406-15; PMID:12824337; http://dx.doi.org/10.1093/nar/gkg595

41. Kubo M, Imanaka T. mRNA secondary structure in an open reading frame reduces translation efficiency in Bacillus subtilis. J Bacteriol 1989; 171:4080-2; PMID:2500423
42. Siller E, DeZwaan DC, Anderson JF, Freeman BC, Barral JM. Slowing bacterial translation speed enhances eukaryotic protein folding efficiency. J Mol Biol 2010; 396:1310-8; PMID:20043920; http://dx.doi.org/10.1016/j.jmb.2009.12.042
43. Yanagitani K, Kimata Y, Kadokura H, Kohno K. Translational pausing ensures membrane targeting and cytoplasmic splicing of XBP1u mRNA. Science 2011; 331:586-9; PMID:21233347; http://dx.doi.org/10.1126/science.1197142
44. Zhang G, Hubalewska M, Ignatova Z. Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. Nat Struct Mol Biol 2009; 16:274-80; PMID:19198590; http://dx.doi.org/10.1038/nsmb.1554
45. Novoa EM, Ribas de Pouplana L. Speeding with control: codon usage, tRNAs, and ribosomes. Trends Genet 2012; 28:574-81; PMID:22921354; http://dx.doi.org/10.1016/j.tig.2012.07.006
46. Saunders R, Deane CM. Synonymous codon usage influences the local protein structure observed. Nucleic Acids Res 2010; 38:6719-28; PMID:20530529; http://dx.doi.org/10.1093/nar/gkq495
47. Makhoul CH, Trifonov EN. Distribution of rare triplets along mRNA and their relation to protein folding. J Biomol Struct Dyn 2002; 20:413-20; PMID:12437379; http://dx.doi.org/10.1080/07391102.2002.10506859
48. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. Nat Rev Genet 2009; 10:691-703; PMID:19763152; http://dx.doi.org/10.1038/nrg2640
49. Garcia-Perez JL, Morell M, Scheys JO, Kulpa DA, Morell S, Carter CC, Hammer GD, Collins KL, O'Shea KS, Menendez P, et al. Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. Nature 2010; 466:769-73; PMID:20686575; http://dx.doi.org/10.1038/nature09209
50. Garcia-Perez JL, Marchetto MC, Muotri AR, Coufal NG, Gage FH, O'Shea KS, Moran JV. LINE-1 retrotransposition in human embryonic stem cells. Hum Mol Genet 2007; 16:1569-77; PMID:17468180; http://dx.doi.org/10.1093/hmg/ddm105