

BRIEF REPORT

Open Access



# Co-expression of distinct L1 retrotransposon coiled coils can lead to their entanglement

Nikola A. Mizgier<sup>1</sup>, Charlie E. Jones<sup>1</sup> and Anthony V. Furano<sup>1\*</sup>

## Abstract

L1 (LINE1) non-LTR retrotransposons are ubiquitous genomic parasites and the dominant transposable element in humans having generated about 40% of their genomic DNA during their ~100 million years (Myr) of activity in primates. L1 replicates in germ line cells and early embryos, causing genetic diversity and defects, but can be active in some somatic stem cells, tumors and during aging. L1 encodes two proteins essential for retrotransposition: ORF2p, a reverse transcriptase that contains an endonuclease domain, and ORF1p, a coiled coil mediated homo trimer, which functions as a nucleic acid chaperone. Both proteins contain highly conserved domains and preferentially bind their encoding transcript to form an L1 ribonucleoprotein (RNP), which mediates retrotransposition. However, the coiled coil has periodically undergone episodes of substantial amino acid replacement to the extent that a given L1 family can concurrently express multiple ORF1s that differ in the sequence of their coiled coils. Here we show that such distinct ORF1p sequences can become entangled forming heterotrimers when co-expressed from separate vectors and speculate on how coiled coil entanglement could affect coiled coil evolution.

**Keywords** L1 retrotransposon, L1 ORF1p, Coiled coils, Coiled coil evolution

## Introduction

L1 non-LTR retrotransposons, which replicate by copying their RNA transcript into genomic DNA, reside in most eukaryotic genomes [1] and are active in nearly all vertebrates where they share the same general structure: a 5' untranslated region (UTR), which has regulatory functions; two protein encoding sequences (ORF1 and ORF2) and a 3' UTR (Fig. 1, top panel) [2]. All vertebrate ORF1 proteins (ORF1p) contain a coiled-coil domain shown in mouse and humans to mediate trimerization of the ORF1p monomer, which is required for high affinity binding to single stranded nucleic acid (ssNA) and NA-chaperone activity that are essential

for retrotransposition [3–6]. The ORF2 encoded protein which contains highly conserved reverse transcriptase and endonuclease domains catalyzes L1 replication [7–9].

Continued emergence, amplification and senescence of successive L1 families accompanied the evolution of most mammals. Each family generated hundreds to many thousands of copies that were retained in the genome. Most L1 copies were defective upon insertion or not being under purifying selection have decayed by the accumulation of random mutations. Although the sequences of these copies have diverged from their once active progenitor, its sequence can be inferred from these fossils, which thereby provide a record of the evolutionary antecedents of the modern families active in present day species, which in humans is designated L1Pa1 (or L1Hs). L1Pa1 is the product of a primate-specific L1

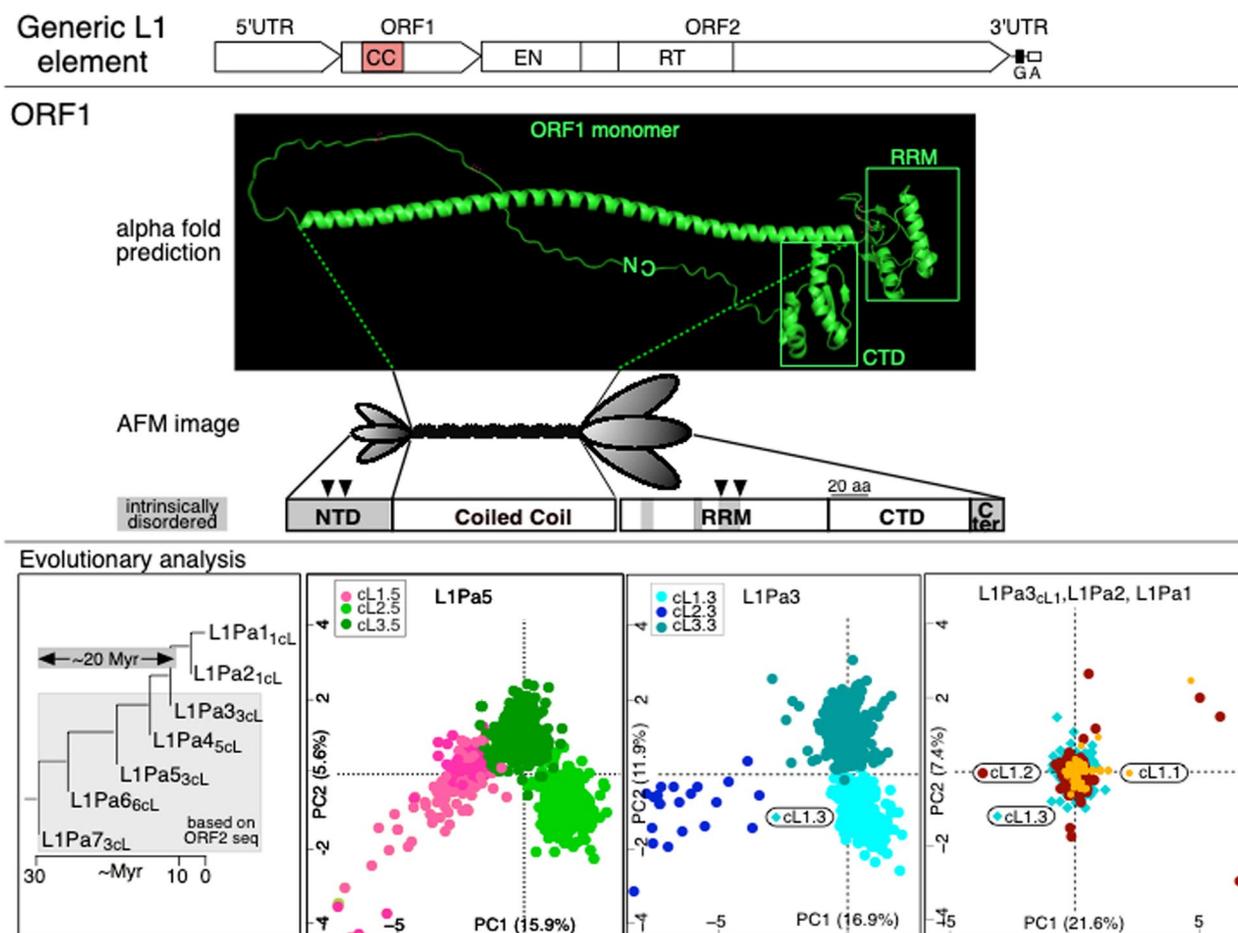
\*Correspondence:

Anthony V. Furano  
anthony.furano@nih.gov

<sup>1</sup> Laboratory of Cellular and Molecular Biology, NIDDK, National Institutes of Health, Bethesda, MD 20892, USA



© This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.



**Fig. 1** ORF1 structure and coiled coil variants

lineage that emerged 80–120 million years ago (MYA) and consists of a largely single lineage of ~16 distinct L1 clades (L1Pa16 – L1Pa1) based on ORF2 and 3' UTR sequences [10, 11].

In contrast to most of the L1 sequence, the 5' UTR and amino-terminal half of ORF1 are evolutionarily labile [2, 12], especially the coiled coil that has been subject to repeated episodes of extensive amino acid substitution [13]. The last such episode in primates lasted for ~20 Myr during the evolution of L1Pa7-L1Pa3 (i.e., from ~30–10 MYA, Fig. 1, bottom left panel). Unexpectedly, each L1 family that was active during this episode concurrently expressed three or more distinct ORF1ps based on their coiled coil sequences, which due to sequence divergence, are now represented by clusters (cL) of related sequences that can be resolved by metric multidimensional scaling (MMDS) [13].

The top panel shows a generic L1 element depicting to scale the coiled coil domain (CC) in ORF1, the

endonuclease (EN) and reverse transcriptase (RT) domains in ORF2, and the G-rich (G) and A-rich (A) motifs in the 3' UTR [2, 13]. The middle ORF1 panel shows the AlphaFold predicted structure for the ORF1p monomer, <https://alphafold.ebi.ac.uk/entry/Q9UN81> [14], and below that a rendering of the atomic force microscopic (AFM) image of the ORF1p trimer [15], its predicted secondary structure ( $\alpha$  helix,  $\beta$  sheet, white boxes) and intrinsically disordered regions, grey boxes [16], and essential phosphorylation sites (arrow heads) [17] - adapted from Fig. 1A of [18]. The bottom Evolutionary analysis panel was adapted from Fig. 3 of [13]. The number of distinct coiled coils (which present as clusters of related divergent sequences, nCL – see text) is indicated for each family on the phylogenetic tree that was derived from ORF2 sequences shown in the left most panel. The consensus sequence derived from cluster 3 of L1Pa5 (cL3.5) encodes the coiled coil of 555p [13].

The left most bottom panel of Fig. 1 shows the number of coiled coil clusters (nCL) in the L1Pa7–L1Pa1 families detected by MMDS, and the other panels show the results of cluster analyses for the L1Pa5, L1Pa3, L1Pa2 and L1Pa1 families. The latter two families diverged about 7 MYA and the consensus coiled coil sequences derived from each cluster are identical and marked the end of this episode of coiled coil change. When active, the L1Pa5 family expressed 3 ORF1 sequences with distinct coiled coil sequences, arbitrarily designated: cL1.5, cL2.5 and cL3.5. Consistent with work by others [19–24] we proposed and provided experimental evidence indicating that the enlarged functional coiled coil sequence space can buffer the effect of inactivating epistatic mutations in the coiled coil [13].

These findings also raised the question of whether co-expressed distinct coiled coils, which could have occurred during the replication and amplification of multi-coiled coil L1 families (e.g., L1Pa5, see Discussion in ref. 13), could become entangled and form heterotrimers. A recent proteome-wide ribosome profiling study showed that coiled coil formation between nascent peptides on neighboring (i.e., cis) ribosomes was the most common mechanism that mediated homomer (dimer) formation of numerous proteins [25]. They found no evidence for *trans*-mediated interactions (i.e., between nascent peptides undergoing translation on separate mRNAs) either on a proteome level search or when examined experimentally by testing for interaction between the lamin C 350 amino acid “rod” region, 320 amino acids of which form 3 coiled coils connected by short linker sequences [26]. In the case of distinct L1Pa5 coiled coil sequences, entanglement would have to not only be mediated in

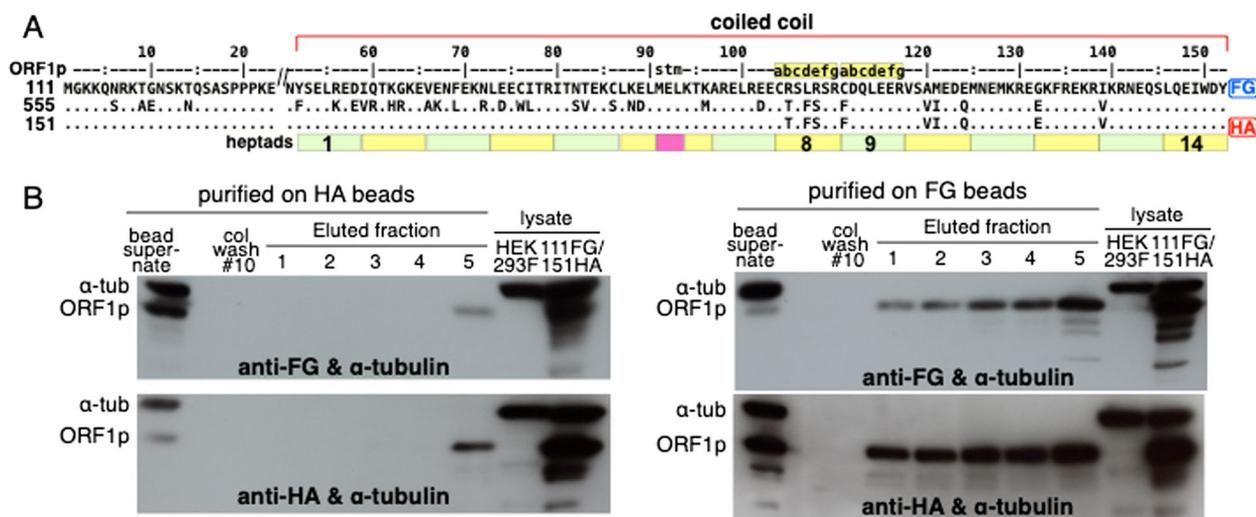
*trans*, but also involve  $\alpha$ -helices that differed in primary sequence.

### Results and discussion

To provide the simplest experimental test for entanglement we fused either Flag (FG) or HA (a peptide derived from the human influenza antigen) epitopes to the C-terminus of various ORF1 coding sequences and expressed them in HEK293F cells [27] using a pcDNA3-based expression vector as described in the Materials and Methods and the legends to the various Figures. To prevent non-specific aggregation of trimers in cell lysates all buffers contained 10  $\mu$ M of a 52 nt deoxynucleotide of T (dT<sub>52</sub>) [3], see Materials and Methods.

Panel A of Fig. 2 shows the amino acid sequences of the NTD and coiled coil of 111p, 151p and 555p, which are encoded respectively by the modern L1Pa1-ORF1, a mosaic L1Pa1/L1Pa5-ORF1, and L1Pa5-ORF1 resuscitated from the now extinct ancestral L1Pa5 family (corresponds to cL3.5, Fig. 1). The purified proteins are essentially indistinguishable with respect to ssNA affinity, and a FRET based NA chaperone assay (annealing/strand exchange) [3, 28, 29]. However, whereas 555p displays about 80% of the 111p activity in a cell culture based retrotransposition assay the mosaic 151p is inactive due to its inability to form tightly compacted complexes on ssNA [29].

(A) Alignment of 555p and 151p to the 111p reference sequence. FG and HA refer respectively to the Flag and HA peptide epitope tags, dots indicate identity. (B) Western blots of proteins that were captured by anti-HA (left) or anti-FG agarose beads (right) from cleared lysates prepared from HEK293F cells that had



**Fig. 2** Entanglement of coiled coils that differ at nine positions

been co-transfected with pcDNA3.1(+)-puro based vectors expressing either ORF1p-111-3XFG (111-FG) or ORF1p-151-HA (151-HA) as described in the Materials and Methods. The upper blots were challenged with anti-FG and the lower blots with anti-FA and both with anti  $\alpha$ -tubulin. In each instance lanes 1–10 of the blotted gel, starting at the left were loaded with: (1) the bead supernatant, (2) stained protein marker (stained bands not visible on these black and white autoradiograms), (3) column wash #10, (lanes,4–8) eluted fractions 1–5, and lanes 9 & 10 respectively, cleared lysate from non-transfected HEK293F and the doubly transfected cells. See Material & Methods, *Protein purification on affinity agarose beads* for the meaning of these sample designations. Alpha tubulin ( $\alpha$  tub) has a molecular weight of 50 kDa; the human ORF1p monomer has a calculated molecular weight of 40 kDa but migrates as a  $\geq 41$  kDa protein on denaturing acrylamide gels: (ref. 3 – Figs. 2 and 3 and ref. 29 – Fig. 1f).

The 111-FG and 151-HA proteins differ at 9 coiled coil positions (Fig. 2A) and cleared lysates from HEK293F cells that had been co-transfected with expression vectors for these proteins were incubated with anti-FG or anti-HA agarose beads (hereafter beads). After elution of bound proteins by their cognate epitopes, the eluates were denatured and subject to electrophoresis, under denaturing conditions. This treatment converts the ~120 kD ORF1p trimers to their constituent ~40 kD monomers. The gels were blotted to filters and exposed to anti-FG or anti-HA antibodies as described in the [Material and Methods](#). Figure 2B shows that protein recovered from either the HA-beads (left panels) or FG-beads (right panels) contained both ORF1p-FG monomers and ORF1p-HA monomers (top and bottom panels respectively).

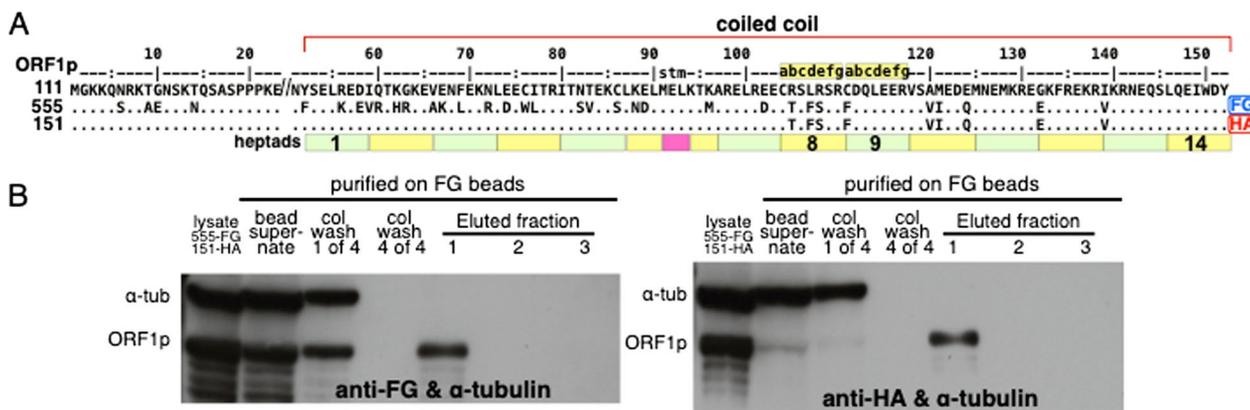
Figure 3 shows the results when co-expressed proteins that differ at 21 coiled coil positions were purified on just

anti-FG beads, which bind its epitope far more efficiently than the anti-HA beads (*cf.* left and right panels in Fig. 2). Figure 3B shows that both HA-tagged and FG-tagged ORF1p monomers were recovered from the ORF1p trimers recovered from the anti-FG beads.

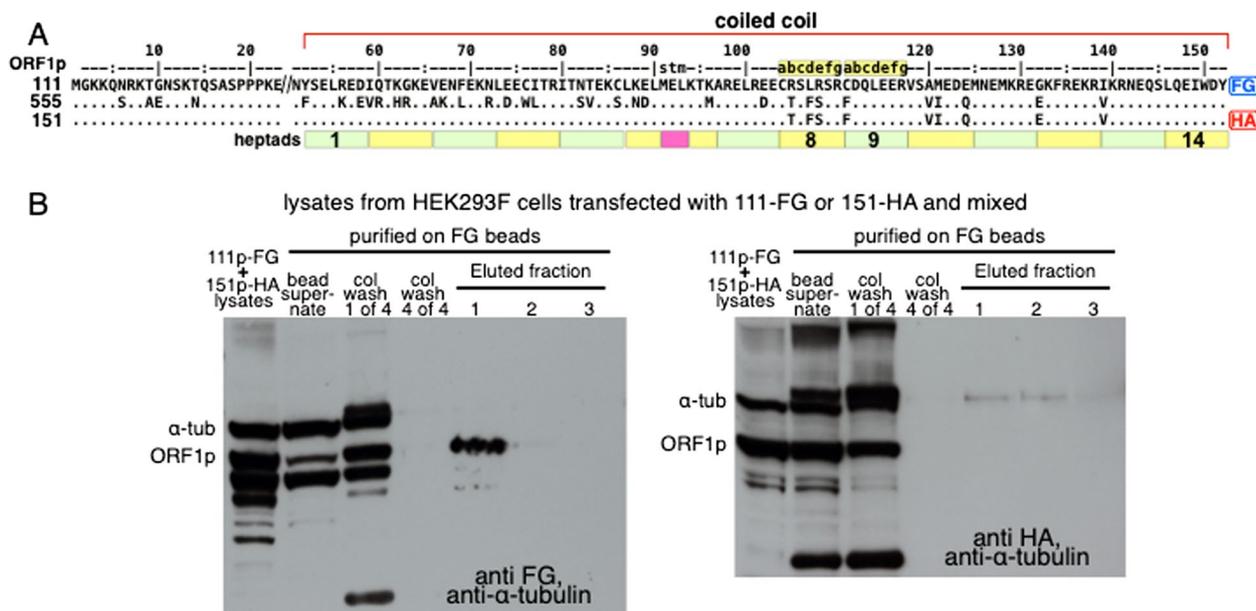
A. Alignment of 555-FG and 151-HA to 111p, showing the twenty-one amino acid differences between the coiled coils of 555p and 151p. B. Western blots of proteins captured by anti-FG agarose beads from cleared lysates prepared from HEK293F cells that had been transfected with expression vectors for both 555-FG and 151-HA as described in the Materials and Methods. In this case the FG-bound proteins were recovered from the anti-FG beads by batch elution. The blots were challenged with either anti-FG or anti-HA, and in both cases by anti- $\alpha$ -tubulin. In both instances lanes 1-7 of the blotted gel starting at the left were loaded with: (1) cleared lysate of the doubly transfected cells, lane (2) bead supernatant, lanes (3 & 4) respectively with column washes #1 and #4, lanes (5-7), eluted fractions 1-3. See Material & Methods, *Protein purification on affinity agarose beads* for the meaning of these sample designations. See legend to Fig. 2 for the molecular weights of  $\alpha$  tubulin and the ORF1p monomer.

To support the foregoing conclusions, we carried out several control experiments critical to their interpretation, which are presented in Additional File 1r. The data herein show that extracts of non-transfected HEK293F cells or those transfected with empty pcDNA.1(+) expression vector are nonreactive to anti-HA or anti-FG antibodies. In addition, we detected no cross activity between these antibodies.

Figure 4 shows that co-expression is required for entanglement. As described in [Materials and Methods](#) cleared lysates of HEK293F that had been transfected with 111-FG or 151-HA expression vectors were mixed for 2 h at



**Fig. 3** Entanglement of coiled coils that differ at twenty-one positions



**Fig. 4** Entanglement requires co-expression

25° and then challenged with anti-FG beads, the eluate of which was processed as described for the experiments in Figs. 2 and 3. In contrast to the co-transfection results, only 111-FG was recovered from the anti-FG resin (cf., left and right panels of Fig. 4), indicating that heterotrimer formation depends on co-expression.

(A) Alignment of 111p-FG and 151p-HA as described for Fig. 2. (B) Western blots of proteins captured by anti-FG agarose beads from cleared lysates prepared from HEK293F cells that had been transfected with expression vectors for either 555-FG or 151-HA and then mixed before processing as described in the *Mixed lysate experiment* section of the Materials and Methods (the lane designations are the same as for Fig. 3). Again, the blots were challenged with either anti-FG or anti-HA, and in both cases by anti-α-tubulin. In contrast to the co-transfection results, no HA-tagged proteins were recovered from the FG-beads.

Although these studies were prompted by our evolutionary evidence for concurrent expression of L1 elements encoding distinct coiled coils they do not speak to the issue of whether ORF1p expressed from full length L1 elements can also become entangled, and we have not attempted to pursue this experimentally. However, our findings indicate that the FG- and HA-tagged ORF1p are translated in close enough approximation to assemble hybrid trimers. A theoretical model of transposable element (TE) evolution posited that a variant TE could be successful if it “...incorporates heritable phenotypic mutations in the elements.” [30]. This

formulation embodies the concept of cis preference, which was supported by subsequent experimental evidence [31–33], showing that L1 proteins (ORF1p and ORF2p) bind to their encoding transcript, to form an L1RNP retrotransposition intermediate. In this context *cis* is more general than and distinct from its use to describe the interaction of nascent peptides undergoing translation on ribosomes bound to the same transcript that leads to coiled coil mediated homomer (dimer) formation [25] or in the case of ORF1p, trimer formation.

As multiple trimers are present in a single L1RNP [15], it is difficult to visualize how trimers once formed by interaction of nascent ORF1p monomers emerging from neighboring ribosomes would then bind to their encoding transcript, as this would presumably block further translation. As shown here trimers can be assembled from monomers that differ in their coiled coil sequence. Achieving this during co-translation would require interaction between nascent monomers emerging from ribosomes in *trans* (i.e., on distinct mRNAs). This phenomenon was not observed in a recent proteome wide study in vivo [25]. However, translation and synthesis of transcripts that encode components of multi-protein complexes are often coupled in so-called “translation factories”, e.g., [34–36]. Additionally, cytoplasmic ORF1p-containing foci can be generated in mouse cells from endogenous L1 elements [37] and in human cells from ORF1p expression vectors and L1 retrotransposition reporter plasmids [38–40] in the form of stress granules

and P-bodies. Thus, the avidity of ORF1p for its encoding transcript does not seem to be a barrier to its repeated translation.

ORF1p trimers can accommodate monomers with distinctive coiled coil amino acid sequences as would be the case for L1Pa7-L1Pa3 (Fig. 1). However, retrotransposition competence of ORF1p is very sensitive to the protein sequence of the coiled coil [13, 29]. Thus, such hybrid trimers would likely be inert, which would select against the survival of L1 families that encode multiple coiled coils. Thus, the advantage of expanded coiled coil sequence space to buffer negative epistasis in the coiled coil as expounded and reviewed in ref. 13 could be subverted by the biophysical dynamics of coiled coil synthesis and assembly.

This conflict could provide a mechanistic, albeit speculative, explanation for repeated episodes of substantial amino acid substitutions followed by strong sequence conservation in the coiled coil, that is a hallmark of L1 evolution.

## Materials and methods

### Plasmids and protein expression

We expressed various epitope-tagged versions of three ORF1 proteins, denoted 111p, 555p and 151p inserted into the BamH1 and EcoR1 sites of pORF1-Flag as previously described (Supplementary Information, [17]), using an expression vector that was derived from pcDNA3.1(+)-Puro (from the Don Ganem laboratory, University of California San Francisco). In this construct the gene for neomycin resistance had been replaced by one for puromycin resistance. Because its DNA sequence had not been reported, we determined the sequence of the relevant region of the plasmid and deposited it to Addgene as pORF1-Flag (Addgene Deposit 190565). We had earlier added the FLAG<sup>®</sup> (Millipore Sigma) peptide (FG, DYKDDDDK) to 111p and 151p using the methods presented in the Supplementary Information of reference [17]. Here we converted 111pFG to 111p3XFG or to 111pHA (YPYDVPDYA) and converted 151pFG to 151pHA using the NEBaseChanger tool. All DNA edits were verified by DNA sequencing. The ORF1p coding sequences are shown in Fig. S2 of Additional File 1.

### Buffers and other reagents

TBS is 50 mM Tris, 150 mM NaCl. Where indicated this buffer was supplemented with 10  $\mu$ M dT<sub>52</sub> (a 52 nt deoxynucleotide of T), which eliminates aggregation of ORF1p trimers that can occur to varying extents at <0.5 M NaCl [3]. PBS is phosphate buffered saline (0.15 M NaCl). Cell lysis buffer is M-Per (ThermoFisher) adjusted to 0.15 M NaCl, 10  $\mu$ M dT<sub>52</sub>, 5  $\mu$ g/ml leupeptin (Millipore Sigma), and 1 mini tablet of Pierce<sup>™</sup> Protease Inhibitor

(ThermoFisher) / 10 ml of buffer. Anti-FG M2 and anti-HA agarose beads were purchased from Millipore Sigma. 3XFG and HA peptides were purchased from Sigma-Aldrich and dissolved respectively at 100  $\mu$ g/ml and 500  $\mu$ g/ml in TBS.

### Expression of ORF1p

Proteins were synthesized in HEK293F [27] cells grown without serum in suspension using the medium and transfection reagents supplied in the GIBCO Expi293<sup>™</sup> Expression System as described in its accompanying protocols. Generally, 15 ml of cells ( $3 \times 10^6$ /ml) were transfected with 15  $\mu$ g of plasmid complexed with 30  $\mu$ l Expifectamine293 transfection reagent. After 3 days the cells were harvested by centrifugation at 1000xg and after washing the pellets with PBS the cells were lysed in 1 ml cell lysis buffer/100 mg wet weight of cell pellet. After gentle shaking for 10 min, the lysates were cleared of cell debris by centrifugation for 15 min at 14,000xg at 4°. Aliquots were stored at -20° or applied to either anti-FG or anti-HA agarose beads as described below or in the figure legends.

### Purification of ORF1p trimers on affinity agarose beads

Beads were washed sequentially with 10 volumes TBS pH7.4 (to remove glycerol storage buffer), 10 volumes 0.1 M glycine-HCL (to clear the antibody binding sites), and 10 volumes TBS, pH 7.4 to remove glycine-HCL and restore the pH to 7.4. Equal volumes of beads and cleared lysate were mixed by rotating 2–4 h or overnight at 4°. The slurry was added to a gravity flow disposable column, which was washed with 10 bed volumes of TBS-10  $\mu$ M dT<sub>52</sub>, collecting 1 ml fractions, and then eluted with TBS-10  $\mu$ M dT<sub>52</sub> containing either 100  $\mu$ g/ml 3XFG peptide or 500  $\mu$ g/ml HA peptide in TBS-10  $\mu$ M dT<sub>52</sub> collecting 1 ml fractions. For the experiments shown in Figs. 3 and 4, we added the lysate bead slurry to Bio-Rad Micro Bio-Spin columns and washed the beads and eluted the bound proteins using a batch procedure: The agarose beads were resuspended in 1 ml wash buffer and recovered by centrifugation. The beads were then repeatedly washed with 1 ml elution buffer, collected by centrifugation saving the washes for further analysis. The bound proteins were eluted by suspending the beads in 1 ml of either of the above peptide-containing elution buffers. The beads were regenerated with 0.1 M glycine-HCL, pH 3.5.

### Western blots

Samples were added to 0.25 volume of NuPAGE 4X LDS Sample Buffer containing 20 mM DTT, heated for 10 min at 70°, and loaded onto a 1.0 mm NuPAGE 10% Bis-Tris polyacrylamide gel. This treatment resolves the trimers into their constituent monomers.

After electrophoresis under denaturing conditions in NuPAGE™ MOPS 0.1% SDS Running Buffer (200 V, ~40 min), the proteins were transferred to PVDF membranes using an iBlot. This instrument and all the reagents are Invitrogen™ products. The PVDF membranes were blocked with Thermo-Fisher SuperBlock T20 in 1x TBS and challenged with Abcam rabbit anti-beta tubulin polyclonal antibody and either Sigma-Aldrich mouse monoclonal anti-FG or Sigma-Aldrich mouse monoclonal anti-HA, overnight on a rocker at 4°. The PVDF membranes were washed 3 times (10 min/wash on a rocker) with 1xTBS/0.05% Tween-20 and then challenged with Sigma-Aldrich HRP-anti-mouse IgG for 1.5 h on a rocker at room temperature. The membranes were washed as described above, and then 4 times with ~25 ml 1x TBS to remove the Tween-20. Following the rinses, the membranes were incubated with Thermo-Fisher SuperSignal Pico West chemiluminescent substrate; the blots were exposed to film and developed.

#### Mixed lysate experiment

We prepared cleared lysates as described above from cultures of HEK293F cells that had been separately transfected with pcDNA3.1-111-3XFG DNA or pcDNA3.1-151-HA DNA. Forty µl samples from each cleared lysate were combined and mixed with 40 µl of wash buffer (1x TBS-10 µM dT<sub>52</sub>). Also, forty µl of cleared lysate from either ORF1-111-3XFG or ORF1-151-HA transfected cells were diluted with 80 µl wash buffer. The three solutions were incubated at 25° C for thirty minutes, with occasional mixing. Thereupon the samples were transferred to 1.5 ml tubes containing 100 µl of Anti-FG M2 agarose beads (Millipore-Sigma) and incubated at 25° C for 2 h, with constant mixing in an Eppendorf Thermomixer R, after which the tubes were centrifuged at 1000xg for 5 min. The binding step supernatants were saved, and the agarose beads were collected by centrifugation, resuspended in wash buffer and centrifuged for 5 min at 1000xg. The wash was repeated four times saving the wash supernatants. Protein bound by the anti-Flag M2 agarose beads was eluted by incubation in 100 µl elution buffer (1x-TBS-10 µM dT<sub>52</sub>, containing 100 µg/ml 3XFG peptide) for 30 min on a Thermomixer shaker. After incubation, the bead slurry was transferred to a Micro Bio-spin columns and the eluate was collected by a 1 min centrifugation at 1000xg. This elution step was carried out three times. The binding step supernatants, wash supernatants and eluates were evaluated by western blot after denaturation of the ORF1p trimers into their constituent monomers as described above.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13100-023-00303-8>.

### Additional file 1.

#### Acknowledgements

The authors thank Dr. Nicholas Guydosh (LBG/NIDDK/NIH) for his informative comments and suggestions.

#### Authors' contributions

NAM & CEJ contributed equally to designing, implementing and performing the experiments, and preparation of the paper. AVF conceived of the project, provided overall direction and wrote the paper.

#### Funding

Open Access funding provided by the National Institutes of Health (NIH) This research was supported by the Intramural Research Program of the NIH, The National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).

#### Availability of data and materials

All data generated or analyzed for this study are included in the submitted manuscript and its supplementary information file.

#### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

All authors consent to publication.

#### Competing interests

None of the authors declare a competing interest.

Received: 31 August 2023 Accepted: 21 September 2023

Published online: 20 October 2023

#### References

- Ivancevic AM, Kortschak RD, Bertozzi T, Adelson DL. LINES between species: evolutionary dynamics of LINE-1 retrotransposons across the eukaryotic tree of life. *Genome Biol Evol.* 2016;8(11):3301–22.
- Boissinot S, Sookdeo A. The evolution of line-1 in vertebrates. *Genome biology and evolution.* 2016;8:3485–507.
- Callahan KE, Hickman AB, Jones CE, Ghirlando R, Furano AV. Polymerization and nucleic acid-binding properties of human L1 ORF1 protein. *Nucleic Acids Res.* 2012;40(2):813–27.
- Khazina E, Truffault V, Buttner 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(9):1006–U64.
- Khazina E, Weichenrieder O. Non-LTR retrotransposons encode non-canonical RRM domains in their first open reading frame. *Proc Natl Acad Sci USA.* 2009;106(3):731–6.
- 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(24):13815–20.
- Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH Jr. High frequency retrotransposition in cultured mammalian cells. *Cell.* 1996;87(5):917–27.
- Feng Q, Moran JV, Kazazian HH Jr, Boeke JD. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell.* 1996;87(5):905–16.
- Mathias SL, Scott AF, Kazazian HHJ, Boeke JD, Gabriel A. Reverse transcriptase encoded by a human transposable element. *Science.* 1991;254(5039):1808–10.

10. Khan H, Smit A, Boissinot S. Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. *Genome Res.* 2006;16(1):78–87.
11. Smit AFA, Tóth G, Riggs AD, Jurka J. Ancestral, mammalian-wide subfamilies of LINE-1 repetitive sequences. *J Mol Biol.* 1995;246:401–17.
12. Furano AV. The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. *Progress in Nucleic Acids Research & Molecular Biology.* 2000;64:255–94.
13. Furano AV, Jones CE, Periwal V, Callahan KE, Walsler J-C, Cook PR. Cryptic genetic variation enhances primate L1 retrotransposon survival by enlarging the functional coiled coil sequence space of ORF1p. *PLoS Genet.* 2020;16(8):e1008991.
14. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596(7873):583–9.
15. 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(2):351–7.
16. Buchan DW, Minnici F, Nugent TC, Bryson K, Jones DT. Scalable web services for the PSIPRED protein analysis Workbench. *Nucleic Acids Res.* 2013;41(Web Server issue):W349–57.
17. Cook PR, Jones CE, Furano AV. Phosphorylation of ORF1p is required for L1 retrotransposition. *Proc Natl Acad Sci U S A.* 2015;112(14):4298–303.
18. Furano AV, Cook PR. The challenge of ORF1p phosphorylation: Effects on L1 activity and its host. *Mob Genet Elements.* 2016;6(1):e1119927.
19. Burch CL, Chao L. Epistasis and its relationship to canalization in the RNA virus phi 6. *Genetics.* 2004;167(2):559–67.
20. Desai MM, Weissman D, Feldman MW. Evolution can favor antagonistic epistasis. *Genetics.* 2007;177(2):1001–10.
21. Hayden EJ, Ferrada E, Wagner A. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature.* 2011;474:92.
22. Hou J, van Leeuwen J, Andrews BJ, Boone C. Genetic network complexity shapes background-dependent phenotypic expression. *Trends Genet.* 2018;34(8):578–86.
23. Montville R, Froissart R, Remold SK, Tenaillon O, Turner PE. Evolution of Mutational Robustness in an RNA virus. *PLoS Biol.* 2005;3(11):e381.
24. Wagner A. Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society B: Biological Sciences.* 2008;275(1630):91–100.
25. Bertolini M, Fenzl K, Kats I, Wruck F, Tippmann F, Schmitt J, et al. Interactions between nascent proteins translated by adjacent ribosomes drive homomer assembly. *Science.* 2021;371(6524):57–64.
26. Herrmann H, Aebi U. Intermediate filaments: structure and assembly. *Cold Spring Harb Perspect Biol.* 2016;8(11):a018242.
27. Vink T, Oudshoorn-Dickmann M, Roza M, Reitsma J-J, de Jong RN. A simple, robust and highly efficient transient expression system for producing antibodies. *Methods.* 2014;65(1):5–10.
28. Callahan KE. Structure and function of the First Open Reading Frame (ORF1) protein encoded by the human LINE-1 Retrotransposon. Washington DC: Georgetown; 2012.
29. Naufer MN, Callahan KE, Cook PR, Perez-Gonzalez CE, Williams MC, Furano AV. L1 retrotransposition requires rapid ORF1p oligomerization, a novel coiled coil-dependent property conserved despite extensive remodeling. *Nucleic Acids Res.* 2016;44(1):281–93.
30. Kaplan N, Darden T, Langley CH. Evolution and extinction of transposable elements in mendelian populations. *Genetics.* 1985;109(2):459–80.
31. Esnault C, Maestre J, Heidmann T. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet.* 2000;24(4):363–7.
32. Kulpa DA, Moran JV. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat Struct Mol Biol.* 2006;13(7):655–60.
33. Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, et al. Human L1 retrotransposition: cis preference versus trans complementation. *Mol Cell Biol.* 2001;21(4):1429–39.
34. Wilk R, Hu J, Blotsky D, Krause HM. Diverse and pervasive subcellular distributions for both coding and long noncoding RNAs. *Genes Dev.* 2016;30(5):594–609.
35. Pichon X, Bastide A, Safieddine A, Chouaib R, Samacoits A, Basyuk E, et al. Visualization of single endogenous polysomes reveals the dynamics of translation in live human cells. *J Cell Biol.* 2016;214(6):769–81.
36. Kamenova I, Mukherjee P, Conic S, Mueller F, El-Saafin F, Bardot P, et al. Co-translational assembly of mammalian nuclear multisubunit complexes. *Nat Commun.* 2019;10(1):1740.
37. Martin SL, Branciforte D. Synchronous expression of LINE-1 RNA and protein in mouse embryonal carcinoma cells. *Mol Cell Biol.* 1993;13(9):5383–92.
38. Goodier JL, Mandal PK, Zhang L, Kazazian HH. Jr. Discrete subcellular partitioning of human retrotransposon RNAs despite a common mechanism of genome insertion. *Hum Mol Genet.* 2010;19(9):1712–25.
39. Goodier JL, Zhang L, Vetter MR, Kazazian HH. Jr. 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.* 2007;27(18):6469–83.
40. 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(10):1041–8.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

