Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition

John L. Goodier*, Ling E. Cheung and Haig H. Kazazian Jr

McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University School of Medicine

Received November 16, 2012; Revised May 7, 2013; Accepted May 12, 2013

ABSTRACT

LINE1s occupy 17% of the human genome and are its only active autonomous mobile DNA. L1s are also responsible for genomic insertion of processed pseudogenes and >1 million non-autonomous retrotransposons (Alus and SVAs). These elements have significant effects on gene organization and expression. Despite the importance of retrotransposons for genome evolution, much about their biology remains unknown, including cellular factors involved in the complex processes of retrotransposition and forming and transporting L1 ribonucleoprotein particles. By co-immunoprecipitation of tagged L1 constructs and mass spectrometry, we identified proteins associated with the L1 ORF1 protein and its ribonucleoprotein. These include RNA transport proteins, gene expression regulators, post-translational modifiers, helicases and splicing factors. Many cellular proteins co-localize with L1 ORF1 protein in cytoplasmic granules. We also assayed the effects of these proteins on cell culture retrotransposition and found strong inhibiting proteins, including some that control HIV and other retroviruses. These data suggest candidate cofactors that interact with the L1 to modulate its activity and increase our understanding of the means by which the cell coexists with these genomic ‘parasites’.

INTRODUCTION

A recent study proposes that over two-thirds of the human genome are repetitive DNA, and that most of this derives from transposable elements (1). Long INterspersed Element-1s (LINE1s) are the major class of retrotransposons, and they move in a ‘copy and paste’ manner involving reverse transcription of an RNA intermediate and insertion of its cDNA copy at a new chromosomal location. RNA of the element and its associated proteins are transported into the nucleus where it is believed reverse transcription and integration occurs in a single step on the DNA itself by a process known as target-primed reverse transcription. In this model, element-encoded endonuclease nicks the bottom strand of target DNA to expose a 3'-hydroxyl that primes reverse transcription of the L1 RNA template. Second-strand cDNA synthesis follows and the integrant is resolved in a manner poorly understood (2). Short target site duplications flank the element at the insertion site.

LINE1 (L1) retrotransposons, the only remaining active autonomous mobile DNA in humans, underwent a massive mammalian expansion, and ~500,000 copies occupy at least 17% of our genome. L1-encoded proteins have also been responsible for the genomic insertion intrans of thousands of processed pseudogenes, and >1 million Alu and 3000 SINE-VNTR-Alu (SVA) non-autonomous retrotransposons (3–5). Most L1s are dead molecular fossils, truncated, rearranged or mutated. However, at least 100 remain potentially active in any human individual (6,7). Up to 5% of newborn children are estimated to have a new retrotransposon insertion, and to date, there are 97 known human disease-causing germ line insertions of L1s, Alus and SVAs (8,9). Recent investigations, driven in part by advances in high-throughput sequencing, suggest that rates of retrotransposition may be much higher during early embryogenesis and in certain somatic cell types, such as neuronal progenitor cells, stem cells and some cancers (10–16).

Retrotransposons are active participants in reorganizing their resident genomes. Retrotransposition occasionally generates target site deletions, or adds non-retrotransposon DNA to the genome by processes termed 5'- and 3'-transduction. Recombination between non-homologous retrotransposons causes deletions, duplications or rearrangements of gene sequence. Ongoing retrotransposition salts genomes with novel splice sites, polyadenylation signals and promoters, and so builds
new transcription modules. Transposable elements have also greatly contributed to chromosome architecture and cell evolution [see reviews in (9,17,18)].

The L1 expresses a 6-kb bicistronic RNA that encodes the 40 kDa Open Reading Frame-1 RNA-binding protein (ORF1p) of essential but uncertain function, and a 150 kDa ORF2 protein with endonuclease and reverse transcriptase (RT) activities (Figure 1A). Retrotransposition is a complex process involving transcription of the L1, transport of its RNA to the cytoplasm, translation of the bicistronic RNA, formation of a ribonucleoprotein (RNP) particle, its re-import to the nucleus and target-primed reverse transcription at the integration site. The complexities of retrotransposition, and the various known mechanisms of transposable element control, predict a complex interaction between L1s and cellular proteins. These proteins and their roles in retrotransposition are poorly understood. A few transcription factors that interact with L1s have been identified (19–23). Non-homologous end-joining repair proteins, such as XRCC1, Ku70 and DNA-PK, have been implicated in resolution of the L1 integrate at the time of insertion (24–26). Yeast two hybrid, affinity purification and shRNA screening studies have identified a few proteins that interact with the L1 and affect retrotransposition, including HNRNPL, NCL, MOV10 and PABPC1 (23,27–29). In addition, the cell has evolved a number of proteins that stand against unrestricted retrotransposition, including the APOBEC3 family of cytosine deaminases, adenosine deaminase ADAR1, chromatin-remodeling factors and members of the piRNA pathway for post-transcription gene silencing that function in the male germ line [reviewed in (30–32)].

The L1 RNP complex is the fundamental unit of retrotransposition. Ultra-centrifugation, immunofluorescence (IF), immunoprecipitation (IP) and cell culture assays indicate that the basal L1 RNP comprises ORF1p and ORF2p preferentially bound to their own encoding L1 RNA (33–35). Cellular factors modulating retrotransposition are likely associated with the basal RNP. We, therefore, predicted that valuable information might be gleaned about the many steps of retrotransposition by identifying cellular macromolecules that interact with the L1 RNP and affect its efficiency of retrotransposition. We epitope-tagged ORF1p in an L1 construct and, following its IP from human cells, identified associated proteins by mass spectrometry (MS). Preliminary examination of these results identified MOV10 helicase as interacting with the L1 complex to inhibit its retrotransposition (29). We now present a detailed investigation of the protein interactome of ORF1p and its L1 RNP complex. Association of many of these proteins with the L1/ORF1p RNP was confirmed by direct co-IP and subcellular co-localization in cytoplasmic granules. When overexpressed, many of these proteins significantly reduce retrotransposition in cell culture. Our results will guide future investigations of the L1 life cycle, from transcription to reinsertion in the genome, and of strategies the cell has evolved to coexist with these genomic mutagens.

### MATERIALS AND METHODS

#### Cloning of plasmid constructs

To generate pc-L1-1FH, we introduced by Kunkel mutagenesis (36) tandem FseI-PacI-SgrAI restriction enzyme sites to replace the stop codon of ORF1 of L1RP (37) cloned in the modified pBS KS- vector jCC5-RPS. A double FLAG-Hemaglutinin (HA)-tag was extracted by polymerase chain reaction (PCR) from plasmid pOZ-FH-N [gift of Y. Nakatani, Harvard (38)] and cloned between the FseI and SgrAI sites. The tagged L1 was then inserted into the vector pcDNA6 myc/his B (Invitrogen) between the NotI and SacII sites.

pCEP-EGFP, pc-L1-RP, ORF1-GFP-L1-RP, FLAG-tagged FL-CDSA (DBPA) and RFP-tagged HNRPNA1 have been described (27,34). The 6myc-p72 (full-length DDX17 isoform 1), mouse pUB-GFP-ZBP1 (GFP-mIGF2BP1) and pcDNA 3.1/V5-His CyclT1 (CCNT) were gifts from R. Janknecht [University of Oklahoma; (39)], S. Sim [Yale University; (40)] and A. Gatignol (Lady Davis Institute for Medical Research, Montéal), respectively. Ultimate ORF cDNA clones (Invitrogen), available to us from a cDNA library maintained by the ChemCore of Johns Hopkins University School of Medicine, were V5-tagged on their N-termini by shuttling them from pENTR221 vector into pcDNA3.1/nV5-DEST using Gateway Technology (Invitrogen). Ultimate ORF Clone ID numbers are shown in Supplementary Table S2. Selected cDNAs (SRP14, PURA, IVNS1ABP, NAT10 and DDX5) were amplified by PCR from a CytoTrap XR human testes library (Stratagene), cloned into the Gateway pDONR vector and shuttled into pcDNA3.1/nV5-DEST.

#### Immunoprecipitation

For MS sequence determination, HEK293T cells were transfected in T75 flasks with 15 μg of pc-L1-1FH or pcDNA6 myc/his B empty vector and expanded for 44 h. For each IP reaction, ~3 × 10^7 cells were harvested with phosphate-buffered saline (PBS)/2 mM ethylenediaminetetraacetic acid and pelleted, and 2.0 ml of cytoplasmic extracts was prepared with NE-PER Nuclear and Cytoplasmic Extraction Kit (Pierce) according to the manufacturer’s protocol, with the exception that the recommended detergent concentration was reduced by half (0.25% NP40). Nuclear pellets were washed three times with Dulbecco’s phosphate-buffered saline (DBPS; Invitrogen), lysed in 900 μl of nuclear extraction reagent and final nuclear extracts were diluted 1.5 times with DBPS. Buffers were supplemented with protease inhibitor cocktail, phosphatase inhibitor cocktails II and III, vanadyl ribonucleoside complexes, phenylmethylsulfonyl fluoride (Sigma) and RNasin (Promega). Extracts were incubated with 60 μl of anti-FLAG M2 affinity gel (Sigma) with rotation for 3 h, washed 5 × (5 min each) with buffer A [160 mM NaCl/50 mM Tris–HCl (pH 7.5)/1 mM ethylenediaminetetraacetic acid/0.2% NP40], and eluted with 150 μg/ml FLAG peptide (Sigma). Eluates (90 μl per sample) were recovered with Costar Spin-X Centrifuge Tube Filters (Corning) and mixed with 3×
Figure 1. pc-L1-1FH immunoprecipitates basal L1 RNP complexes from 293T cell lysates after α-FLAG agarose purification. (A) Structure of FLAG-HA-tagged pc-L1-1FH cloned in vector pcDNA6 myc/his B. RT: ORF2 reverse transcriptase domain; EN: endonuclease domain; P_{CMV}: CMV promoter; BGH An: bovine growth hormone polyadenylation signal. (B) FLAG-tagged ORF1p expressed from the construct pc-L1-1FH binds α-FLAG agarose independent of RNase digestion (lanes 5 and 8), but untagged ORF1p (construct pc-L1-RP) will not bind (lane 6). (C and D)
sodium dodecyl sulfate (SDS)/β-mercaptoethanol loading buffer. A purification factor table for immunoprecipitation of FLAG-tagged ORF1p expressed from pc-L1-1FH is shown in Supplementary Table S3. Total protein concentrations were determined by bicinchoninic acid protein assay. A standard curve was calculated from dilutions of bovine serum albumin. ORF1p purification was assessed by ImageJ software (NIH) analysis of band intensities on western blots.

For each co-immunoprecipitation of Figure 3A, extracts from 3.5 × 10^8 293T cells transfected with L1 and test protein constructs in T75 flasks were prepared in 370 μl of buffer A supplemented as aforementioned. Lysates containing test proteins of predominantly nuclear localization were sonicated. Reactions transfected with pc-L1-1FH were immunoprecipitated with 30 μl of anti-FLAG M2 affinity gel and eluted as aforementioned. RNase inhibitors were omitted from samples treated with 15 μg/ml DNase-free RNase (Roche). To detect their interaction with endogenous ORF1p, V5-tagged proteins were transfected in 7 × 10^5 2102Ep embryonal carcinoma cells (Figure 3B), immunoprecipitated on Protein G Agarose/Salmon Sperm DNA (Upstate) and eluted in SDS loading buffer with heating at 95°C, followed by addition of β-mercaptoethanol.

**MS sequencing and database analyses**

MS sequencing was performed by the Johns Hopkins Mass Spectrometry and Proteomics Facility. Proteins in solution were reduced and alkylated with iodoacetamide and proteolyzed with sequencing grade modified porcine trypsin (Promega). Proteins in gel bands were proteolyzed with trypsin as previously described (41). Protein identification by liquid chromatography tandem mass spectrometry (LCMS/MS) was performed using LTQ ion trap MS (Thermo Fisher Scientific) interfaced with a 2D nanoLC system (Eksigent). Peptides were fractionated by reverse-phase high-performance liquid chromatography on a 75 μm × 100 mm C18 column with a 10 μm emitter using a 0–60% acetonitrile/0.1% formic acid gradient >90 min at 300 nl/min. Peptide sequences were identified using Proteome Discoverer and Mascot software (Matrix Science) to search the NCBI nr 167 database, with oxidation on methionine and carbamidomethylation on cysteine as variable modifications. Mascot search result *.dat files were processed in Scaffold (Proteome Software, Inc.) to validate protein and peptide identifications. A protein required ≥90% probability for peptide annotation and the identification of at least two unique peptides to be included in this study. Any protein also detected in the empty vector reaction was discarded as a false positive. Core ribosomal proteins and suspected contaminant proteins (including keratins, annexin A1, plakoglobin, desmoplakin, peroxiredoxin2 and methylsoy subunit plCln) detected in pre-run or post-run bovine serum albumin controls were also omitted from analyses.

Additional analyses used DAVID v6.7 [Database for Annotation, Visualization and Integrated Discovery; (42)] and STRING v9 [Search Tool for the Retrieval of Interacting Genes/Proteins; (43)]. The Nucleolar Proteome Database is available at http://www.lamondlab.com/NoPDB (44).

**Yeast two-hybrid analyses**

The Cytotrap system (Stratagene) screen of a human testis cDNA library has been described (27). A human HeLa cDNA library (HL4000AA) was also screened in yeast strain Y187 using the GAL4 Matchmaker Two Hybrid System II according to the manufacturer’s protocols (Clontech). As baits, full-length L1 ORF1 and the N-terminal 717 nucleotides of ORF2 (EN domain) were cloned in the vector pAS2-1.

**Cell culture and retrotransposition assays**

Human 2102Ep embryonal carcinoma cells (a gift from P. K. Andrews, University of Sheffield), HeLa-HA cells (a gift from J. Moran, University of Michigan) and HEK293T cells (ATCC) were grown in Dulbecco’s modified Eagle’s medium with 10% FBS (Hyclone), GlutaMax and Pen-Strep (Invitrogen). All transfections were seeded in 6-well plates. The following day, 1.0 μg of 99-PUR-RPS-EGFP, a plasmid containing L1 RP and EGFP retrotransposition reporter cassette, was co-transfected together with 0.5 μg of empty vector (pcDNA3, Invitrogen) or test plasmid (four replicate wells each). Five days post-transfection, cells having a retrotransposition event, and hence expressing EGFP, were

![Figure 1](https://i.imgur.com/1234567.png) Continued

Detection of L1 proteins in the RNP IP. Lanes 1–4: input lysates; lanes 6–9: immunoprecipitates; lanes 1, 2, 6 and 7: cytoplasmic fractions; lanes 3, 4, 8 and 9: nuclear fractions. (C) FLAG-HA-tagged ORF1p, detected by α-FLAG antibody. Putative ORF1p dimer and trimer bands are visible in IP samples (lanes 7 and 9). The reason for their absence in lysate samples is unclear (lanes 2 and 4). IP purification factors were determined for immunoprecipitation of FLAG-tagged ORF1p expressed from pc-L1-1FH with pc-L1-1FH (lanes 3 and 5), but not in reactions with the empty vector (lanes 2 and 4). RT-control: the RT incubation step was omitted and 2 μl of pc-L1-1FH immunoprecipitate was added directly to the PCR reaction. No PCR product was detected (lanes 1 and 6). The assay is described in Kulpa and Moran (51). (F) L1 RNA detected by RT–PCR (lanes 3 and 5). RT-: RT enzyme was omitted from the cDNA synthesis step using pc-L1-1FH. (H) The purity of nuclear and cytoplasmic whole-cell lysate fractions is shown by western blotting. α-HDAC1 is a strictly nuclear protein (54) and α-MEK1/2 is cytoplasmic (55). (I) Immunoprecipitated samples resolved on silver-stained polyacrylamide gels. To support protein identification data from complex IP samples, selected prominent band regions were excised for additional MS sequencing. Both cytoplasmic (left) and nuclear (right) IP fractions are shown.
assayed by flow cytometry. Gating exclusions were based on background fluorescence of plasmid 99-PUR-JM111-EGFP, which has a double-point mutation in ORF1 that abolishes retrotransposition (46). Within each experiment, results were normalized to fluorescence of 99-PUR-RPS-EGFP co-transfected with pcDNA3 empty vector.

To control for any differences in transfection efficiencies of test proteins or off-target effects on EGFP reporter expression, we followed the strategy of Wei et al. (47). Each test plasmid (0.5 μg) was co-transfected in quadruplicate wells of 12-well plates with pCEP-EGFP (0.5 μg), a construct that constitutively expresses EGFP from a cytomegalovirus (CMV) promoter. Four days post-transfection, EGFP fluorescence was determined by flow cytometry. These results were then used to correct retrotransposition rates (by dividing each retrotransposition reading by the average of four replicate pCEP-EGFP readings for the corresponding test protein; Figure 6B).

To determine potential cell toxicity because of transfection of test proteins, 15,000 293T cells were seeded in 75 m² of Dulbecco’s modified Eagle’s complete medium in 96-well plates. The next day, transfection reactions prepared with 100 ng of test plasmid, 0.3 μl of Fugene HD and 25 μl of Opti-MEM Reduced Serum Medium (Invitrogen) were added to each well. After 3 days, a MultiTox-Fluor Multiplex Cytotoxicity Assay Kit (Promega) was used to assay cell toxicity. This assay simultaneously measures cell viability and cytotoxicity in a single-reagent reaction, permitting ratios of live to dead cell readings to be calculated.

**Immunofluorescence and Western blotting**

Immunofluorescence techniques have been described previously (27). Purified AH40.1 polyclonal antibody against ORF1 protein was a gift from T. Fanning (Armed Forces Institute of Pathology, MD, USA) and M. Singer (Carnegie Institution of Washington) (48). Polyclonal affinity-purified antibody α-ORF2-N (154–167) was described (27). Other antibodies included ms-FLAG-tag (Sigma), ms-α-V5-tag (Invitrogen), ms-α-Myc-tag (Invitrogen), ms-α-fibrillarin (38F3; Santa Cruz), rb-α-HEXIMI [a gift from J. Yik, Berkeley; (49)], gt-α-LARP1 (SDIX), ms-α-nucleolin/C23 (4E2; Santa Cruz), ms-α-nucleophosmin/B23 (Zymed), ms-α-PABPC1 (clone 10E10; G. Dreyfuss, University of Pennsylvania), ms-α-Ro60 (S. Wolin, Yale University) and rb-α-UAP56 [M. Green, University of Massachusetts Medical Center; (50)]. Donkey Cy3-, DyLight 488- or DyLight 549-conjugated and horseradish peroxidase-conjugated secondary antibodies were from Jackson ImmunoResearch Laboratories. Slides were examined with a Nikon A1 confocal laser microscope system, and images were analyzed with the NIS-Elements software platform.

All protein samples were resolved on NuPAGE 4–12% Bis–Tris polyacrylamide gels (Invitrogen) using 3 × SDS/β-mercaptoethanol loading buffer. Western blot antibody incubation was in DPBS/0.05% Tween 20/5% dry milk overnight at 4°C. Membranes were washed 3× for 10 min with DPBS/0.05% Tween 20, exposed to secondary antibody at room temperature for 2 h, and washed again. Detection used SuperSignal West Pico Chemiluminescent Substrate (Pierce) and Hyperfilm ECL (GE Healthcare).

**Reverse transcriptase assay and RT–PCR**

Reverse transcriptase analysis followed the LEAP protocol of Kulpa and Moran (51). Two microliters of IP sample was added to each cDNA extension reaction. cDNA PCR amplification used Expand Long Template PCR System (Roche) and 2 μl of the extension reaction. Products were separated on 2% agarose gels and confirmed by sequencing. Primers used in the assay are noted in Supplementary Table S1.

For RT PCR analyses, RNA from total cell lysates was extracted using the RNA Easy Mini kit (Qiagen). Immunoprecipitates were used without extraction. Samples were treated with TURBO DNase (Invitrogen), and cDNA was generated using an iScript Select cDNA Synthesis Kit for L1 RNA (Figure 1) or iScript cDNA Synthesis Kit (Bio-Rad) for small RNAs and mRNAs (Figure 4). Standard PCR reactions used the Expand Long Template PCR System (Roche) and primers listed in Supplementary Table S1.

**RESULTS**

**Tagged L1 constructs successfully immunoprecipitate the basal L1 RNP**

Endogenous full-length Short INterspersed Element (SINE) and LINE transcripts are difficult to detect in most cell types, although expression levels increase under some conditions of stress (52). L1-encoded proteins, especially ORF2p, are also suppressed in most cells (53). Low expressivity and the confounding presence of numerous retrotransposons within transcribed mRNAs make biochemical study of endogenous L1 elements difficult.

Faced with the difficulties of investigating RNA and proteins encoded by endogenous L1s, we previously developed expression constructs of L1RP [an L1 highly active in cell culture, (37)] containing ORF1 fused with T7, TAP or FLAG epitope tags (27,53). For the present study, L1rp with tandem HA-FLAG tag fused to the C-terminus of ORF1 (construct pc-L1-1FH) or ORF2 (pc-L1-2FH) was cloned into a vector that replicates in HEK293T cells (Figure 1A). Tagging ORF1 in this manner diminishes (to 30%) but maintains cell culture retrotransposition of L1RP in an EGFP-based reporter assay (29,45).

pc-L1-1FH allows the immunoprecipitation on α-FLAG agarose of basal L1 RNP complexes from cell lysates; untagged ORF1p expressed from an otherwise identical construct (pc-L1-RP) does not bind agarose (Figure 1B, lanes 5, 6 and 8). After transfection in 293T cells and purification, we detected in cytoplasmic and nuclear immunoprecipitates both ORF1 and ORF2 proteins (Figures 1C and 1D), robust ORF2p RT activity (Figure 1E) as determined by an in vitro PCR-based assay (51), and L1 RNA (Figure 1F). Western blot analyses of immunoprecipitated tagged ORF1p consistently yielded an expected band of 45 kDa plus minor
bands of ~90 and 140 kDa (Figure 1C, lanes 7 and 9), sizes consistent with persistent multimers resistant to SDS loading buffer containing β-mercaptoethanol. ORF1p trimerization has been characterized (56,57), and Hohjoh and Singer (58) indicated that ORF1p multimers might form by forces other than disulfide bonds. Multimers of ORF1p explain the detection of ORF1p peptides in excised gel fragments 45 kDa (Figure 1I).

Western blot analyses of RNP complexes formed by pc-L1-2FH detected tagged ORF2 protein in large-scale preparations of cytoplasmic and nuclear immunoprecipitates (Figure 1G). However, it was not possible using pc-L1-2FH to purify sufficient complexes to reliably analyze their protein composition by MS; therefore, we used pc-L1-1FH to co-IP cellular proteins associated with LINE1.

**MS analysis of the L1 RNP reveals a tight network of RNA-binding proteins**

We transfected pc-L1-1FH and empty vector in parallel in 293T cells and performed α-FLAG-tag IP from nuclear and cytoplasmic fractions. Western blotting confirmed significant fraction purity (Figure 1H). Both complex samples and eight selected band regions excised from preparative gels were analyzed by liquid chromatography tandem MS. Silver-stained electrophoretic gels showed relatively few bands in empty vector lanes compared with those in pc-L1-1FH lanes (Figure 1I). Excluding 36 ribosomal proteins (listed in Supplementary Table S2), 96 proteins with two or more predicted peptides and unique to pc-L1-1FH isolates were identified. Forty proteins were identified in the nuclear fraction only, 34 in the cytoplasmic fraction only and 22 in both fractions (Table 1 and Supplementary Table S2). As expected, exogenous ORF1 protein expressed from pc-L1-1FH was predominant in MS analyses of both cytoplasmic and nuclear IP reactions (Figure 1C and Supplementary Table S2). Although ORF2p is clearly visible on Western blots (Figure 1D), its peptide sequences were not found among the MS reads.

The list of 96 proteins was used to query the Gene Ontology (GO) database (Figure 2A and Supplementary Table S4). A majority of these proteins bind nucleic acids. The L1 ORF1p interactome contains many heterogeneous nuclear ribonucleoproteins (hnRNPs), splicing factors, chaperones, transport and localization proteins, and transcription/translation factors. Querying the STRING database of experimentally confirmed protein–protein interactions returns a tight network of interactions dominated by RNP proteins (Figure 2B).

cDNAs for 69 of the 96 putative L1-interacting proteins were subcloned from a cDNA library into expression vectors or were obtained as gifts. Most cDNAs-encoded proteins with N-terminal V5-epitope-tags. After their co-transfection in 293T cells, 41 of the 69 proteins tested were found to co-IP specifically with the FLAG-tagged ORF1p L1 RNP (Figure 3A). Nine additional proteins bound non-specifically to the affinity agarose. Nine other proteins were expressed at significantly higher levels in the presence of co-transfected pc-L1-1FH compared with the empty vector; therefore, their specific co-IP with the L1 could not reliably be confirmed (marked as '?' in Table 1). Predictably, most interactions were RNA-dependent and disappeared with RNase treatment (Figure 3A).

We next assayed the ability of a subset of the V5-tagged proteins to co-IP endogenous ORF1p from 2102Ep cells, a human embryonal carcinoma line that expresses L1 ORF1p at particularly high levels (48). After IP on α-V5-bound agarose, overexposure of Western blots showed that many proteins tested (16/21), but not empty vector, obviously co-IPed endogenous ORF1p (Figure 3B). Finally, IP of pc-L1-1FH on α-FLAG agarose was also able to pull-down selected endogenous proteins from 293T cells (Figure 3C).

Collectively, the co-IP data strongly support the MS sequencing results and validate the use of tagged-L1 constructs for discovering cellular proteins that potentially associate with the L1.

Further confirming their association with the L1, a number of the proteins identified in the present study were also detected during yeast two-hybrid cDNA library screens. As previously described (27), a CytoTrap human testes cDNA library (Stratagene) was queried using ORF1p as bait. This system targets test proteins to the cell membrane for detection. Six of nine proteins previously recovered by the CytoTrap screen were also found by the present study to associate with the L1 (CSDA, DDX21, DDX39A, HNRNPA1, SERBP1 and YBX1). Although recovered from yeast, TLS-FUS, signal recognition particle (SRP)72 and SRSF7 were not detected by L1 IP; however, SRP14 and several other serine/arginine-rich (SR) splicing factors (SRSF1, SRSF6 and SRSF10) were found ([27]; J. Goodier, unpublished data). We also performed a GAL4 Matchmaker II (Clontech) two-hybrid analysis of a HeLa cell cDNA library using ORF1p or the endonuclease domain (EN) of ORF2p as bait. The GAL4 system queries interactions by targeting proteins to the nucleus. The strongest and most frequently interacting ‘hit’ was HEXIM1, which was recovered with ORF2p EN. HEXIM1 negatively regulates transcription by sequestering positive transcription elongation factor b (P-TEFb), together with 7SK small nuclear (sn) RNA, CDK9, CCNT1 (cyclin T1), LARP7 and MEPCE in an inactive RNP (49,59). Notably, HEXIM1, MEPCE, CCNT1 and 7SK RNA were all also recovered with pc-L1-1FH during the present study (Table 1 and Figure 4). ORF1p bait mainly bound endogenous ORF1-related clones from the Matchmaker II library (five unique clones from a total of eight). No other hits overlapped with the present study.

**Non-L1 RNAs associate with the ORF1p complex**

Some proteins detected by our MS sequencing preferentially associate with small RNAs that are overrepresented in the human genome as pseudogene families. For example, SR140, SART1 and SR splicing factors bind small spliceosomal RNAs, U1 to U6; these small nuclear (sn) RNAs all have processed pseudogenes, >60% of which are related to U6 (60,61). Autoantigens TROVE2 (Ro60) and SSB (La) form a core complex with hY small
| Gene symbol | Alternate symbol | Protein name | Co-IPs with pc-L1-1FH | With ORF1p in granules | Nuclear (Nu) / cytoplasmic (C) extracts |
|-------------|-----------------|--------------|------------------------|-------------------------|----------------------------------------|
| C3orf26     | MGC4308         | Chromosome 3 open reading frame 26 | Y(a)                 | N                        | Nu                                     |
| C1orf79     | COPR5           | Chromosome 17 open reading frame 79, cooperator of PRMT5 | Beads             |Nd                       | Nu                                     |
| C22orf28    | HSPC117         | Chromosome 22 open reading frame 28, tRNA-splicing ligase RtcB homolog | N                     | Y                        | C                                      |
| CAT         |                 | Catalase     | nc                     | nc                       | C                                      |
| CCNT1       | CYCT1           | Cyclin T1    | ?                      | N                        | Nu                                     |
| CCT2        | TCP-1-β         | Chaperonin containing TCP1, subunit 2 (β) | N                     | N                        | C                                      |
| CCT4        | TCP1-Δ          | Chaperonin containing TCP1, subunit 4 (Δ)/stimulator of TAR RNA binding. | nc                    | nc                       | C                                      |
| CCT6B       | TCP-1-Ω         | Chaperonin containing TCP1, subunit 6B (Ω) | N                     | nd                       | C                                      |
| CCT8        | TCP-1-θ         | Chaperonin containing TCP1, subunit 8 (θ) | N                     | nd                       | C                                      |
| CDC5L       |                 | CDC5 cell division cycle 5-like (S. pombe) | ?                     | N                        | Nu                                     |
| CSDA        | DBPA            | Cold shock domain protein A | Y(27)                | Y                        | Nu/C                                   |
| DARS        |                 | Aspartyl-tRNA synthetase | nc                    | nc                       | C                                      |
| DDX17       | p82             | DEAD (Asp-Glu-Ala-Asp) box polypeptide 17, isoform p82 variant | Y(a)                 | N                        | Nu                                     |
| DDX21       | RH-H/GuA        | DEAD (Asp-Glu-Ala-Asp) box polypeptide 21 | Y(a)                 | N                        | Nu                                     |
| DDX23       | PRPF28          | DEAD (Asp-Glu-Ala-Asp) box polypeptide 25 | nc                    | nc                       | C                                      |
| DDX39A      | BAT1, UHR49     | DEAD (Asp-Glu-Ala-Asp) box polypeptide 39 | Y(a,b,c)             | N                        | Nu                                     |
| DDX5        | p68             | DEAD (Asp-Glu-Ala-Asp) box polypeptide 5 | Y(a)                 | N                        | Nu                                     |
| DHX9        | DDX9, RHA       | DEAH (Asp-Glu-Ala-His) box polypeptide 9 | ?                     | o                        | Nu                                     |
| DKC1        |                 | DKC1 dyskeratosis congenita 1, dykerin | N                     | nd                       | C                                      |
| EIF4B       |                 | Eukaryotic translation initiation factor 4B | beads                | Y                        | Nu/C                                   |
| ELAVL1      | HUR             | Embryonic lethal abnormal vision n, Drosophila like 1/Hu antigen R | Y(a,b)               | Y                        | Nu                                     |
| FAM120A     | C9orf10         | Family with sequence similarity 120A/oxidative stress-associated Src activator | nc                   | nc                       | Nu                                     |
| FAM98A      |                 | Family with sequence similarity 98, member A | Y(a)                 | Y                        | C                                      |
| FBL         |                 | Fibrillarin   | ?                      | N                        | Nu                                     |
| H1FX        |                 | Histone H1x  | Y(a,b)                | N                        | Nu                                     |
| HEXIM1      | HISI            | Hexamethylene bis-acetamide inducible 1 | Y(a,c)               | N                        | C                                      |
| HIST1H1B    | H1B             | Histone cluster 1, H1b | nc                    | nc                       | Nu/C                                   |
| HIST1H1C    | H1.2, H1C       | Histone 1 H1C | Y(a)                 | N                        | Nu/C                                   |
| HIST1H1E    | H1.4, H1E       | Histone cluster 1, H1e | nc                    | nc                       | Nu                                     |
| HNRNP1      |                 | Heterogeneous nuclear ribonucleoprotein A1 | Y(a,b)               | Y                        | Nu                                     |
| HNRNP2A     |                 | Heterogeneous nuclear ribonucleoprotein A2/B1, transcript variant Bl | Y(a)                 | N                        | Nu                                     |
| HNRNPAB     | ABBP1           | Heterogeneous nuclear ribonucleoprotein A/B/APOBEC1-binding protein 1 | Y(a)                 | Y                        | Nu                                     |
| HNRNPC      | HNRNP C1/C2    | Heterogeneous nuclear ribonucleoprotein C (C1/C2) | Y(a,b)               | N                        | Nu                                     |
| HNRNP3      |                 | Heterogeneous nuclear ribonucleoprotein H3 | ?                     | N                        | Nu                                     |
| HNRNPK      |                 | Heterogeneous nuclear ribonucleoprotein K | ?                     | N                        | Nu/C                                   |
| HNRNPL      |                 | Heterogeneous nuclear ribonucleoprotein L, | Y(a)                 | N                        | Nu                                     |
| HNRNPR      |                 | Heterogeneous nuclear ribonucleoprotein R | ?                     | N                        | Nu                                     |
| HNRNPV      | SAF-A           | Heterogeneous nuclear ribonucleoprotein U (scaffold attachment factor A) | Y(a,b)               | N                        | Nu                                     |
| IGF2BP1     | IMP1, ZBP1     | Insulin-like growth factor 2 mRNA binding protein 1 | Y(a)                 | Y                        | Nu/C                                   |
| IGF2BP2     | IMP2            | Insulin-like growth factor-binding protein 2, 36 kDa | nc                   | nc                       | C                                      |
| ILF2        | NF43            | Interleukin enhancer-binding factor 2 | Y(a)                 | o                        | Nu                                     |
| ILF3        | NFAR, NF90      | Interleukin enhancer-binding factor 3 | Y(a,b)               | N                        | Nu                                     |
| IVNS1ABP    |                 | Influenza virus NS1A-binding protein | Beads                | Y                        | C                                      |
| KPN2A       |                 | Karyopherin z (RAG cohort 1, importin z) | Y(a)                 | nd                       | C                                      |
| KRII        |                 | KR11 homolog (S.cerevisiae) | nc                   | nc                       | Nu                                     |
| LARP1       |                 | La ribonucleoprotein domain family, member 1 | Y(c)                 | Y                        | Nu/C                                   |
| LUC7L3      | CROP            | Cisplatin resistance-associated-overexpressed protein | nc                   | nc                       | C                                      |
| MARS        |                 | Methionyl-tRNA synthetase | nc                   | nc                       | C                                      |
| MATR3       |                 | Matrin 3 | nc                     | nc                       | Nu                                     |
| MEPCE       | BCDIN3          | 7SK snRNA methylphosphate capping enzyme | Y(a)                 | N                        | C                                      |
| MOV10       | K1AA1631        | Moloney leukemia virus 10, homolog (mouse) | Y(a,b)               | Y                        | Nu/C                                   |
| NAT10       |                 | N-acyltransferase 10 (GCN5-related) | ?                     | N                        | Nu                                     |
| NCL         |                 | Nucleolin   | Y(a,c)                | N                        | Nu/C                                   |
| NOP56       | NOL5A           | Nucleolar protein 56 | Y(a)                 | N                        | Nu                                     |
| NPM1        | B23             | Nucleophosmin 1 | Y(c)                 | N                        | Nu                                     |
| NUSAP1      |                 | Nucleolar and spindle-associated protein 1 | ?(y,b)               | N                        | C                                      |
| PABPC1      |                 | Poly(A)-binding protein, cytoplasmic 1 | Y(a,c)               | Y                        | Nu/C                                   |

(continued)
cytoplasmic (sc) RNAs, which also comprise pseudogene families (60,61). hnRNPK and nucleolin are subunits of TROVE2 particles containing hY1 and hY3 RNAs (62,63). TROVE2, SSB and SRP14 also bind RNA of Alu retrotransposons, which comprise 10% of the human genome, and HEXIM1, MEPCE and CCNT1 complex with 7SK snRNA (49,59). The human 7SK pseudogene family (64) includes 160 members with obvious short target site duplications, a hallmark of L1-mediated insertion (J. Goodier, unpublished data).

We performed conventional RT–PCR to assay for the presence of selected RNA species in tagged-ORF1p RNP immunoprecipitates. L1 RNA and all small RNAs tested were associated with pc-L1-1FH, but were absent or barely detectable in empty vector control samples (Figure 4). Alu RNA was also detected in RNP samples; however, it is difficult to conclude these represent bona fide Pol III-derived Alu transcripts, as many mRNAs also contain embedded Alus. Some but not all mRNAs tested were detected in the pc-L1-1FH sample, including \( \alpha \)-actin and Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) mRNA. Interestingly, humans have 62 and mice 331 GAPDH-associated pseudogenes, one of the largest numbers among all proteins, and a phenomenon linked with high-mRNA expression and recent bursts of mammalian retrotransposition (4,65).

### Table 1. Continued

| Gene symbol | Alternate symbol | Protein name | Co-IPs with pc-L1-1FH | With ORF1p in granules | Nuclear (Nu) / cytoplasmic (C) extracts |
|-------------|-----------------|--------------|-----------------------|------------------------|----------------------------------------|
| PABPC4      |                 | Poly(A)-binding protein, cytoplasmic 4 | nc                     | nc                     | Nu/C                                   |
| PCBP2       | HNRPN2          | Poly(rC)-binding protein 2 | Y(a)                  | Y                      | Nu/C                                   |
| PDIA3       |                 | Protein disulfide isomerase family A, member 3 | nc                     | nc                     | C                                      |
| PPAN        |                 | Peter pan homolog (Drosophila) | nc                     | nc                     | C                                      |
| PPMA1B      |                 | Protein phosphatase, Mg2+/Mn2+ dependent, 1B | nc                     | nc                     | C                                      |
| PRF4        |                 | PRP4 pre-mRNA processing factor 4 homolog (yeast)/ U4/U6 small nuclear ribonucleoprotein Prp4 | N                      | nd                     | C                                      |
| PRPF19      |                 | PRP19/PS04 pre-mRNA processing factor 19 homolog (S.cerevisiae) | N                      | N                      | C                                      |
| PRPF31      |                 | PRP31 pre-mRNA processing factor 31 homolog (S.cerevisiae)/ U4/U6 small nuclear ribonucleoprotein Prp31 | beads                  | Y                      | C                                      |
| PTBP1       | HNRPN1          | Polypyrimidine tract binding protein 1 | nc                     | nc                     | C                                      |
| PURA        |                 | Purine-rich element-binding protein A | Y(a,b)                 | N                      | Nu                                     |
| RALY        | HNRPNCL2        | RNA-binding protein, autoantigenic (hnRNP-associated with lethal yellow homolog (mouse)) | Y(a,b)                 | N                      | Nu                                     |
| RMBX        | HNRPNG          | RNA-binding motif protein, X-linked | Y(a)                  | N                      | Nu                                     |
| RO1K1       |                 | RIO kinase 1 (yeast) | Y(a)                  | Y                      | C                                      |
| RNMT        | RG7MT1          | RNA (guanine-7-) methyltransferase | nc                     | nc                     | Nu                                     |
| SART1       |                 | Squamous cell carcinoma antigen recognized by Tcells/U4/U6.5 tri-snRNP-associated protein 1 | beads                  | N                      | Nu/C                                   |
| SERBP1      | PAIRBP1, CGI-55 | SERPINE1 mRNA-binding protein 1 | Y(27)                 | Y                      | Nu                                     |
| SF3B1       | PRPF10          | Splicing factor 3B, subunit 1 | nc                     | nc                     | Nu/C                                   |
| SF3B3       | SAP130          | Splicing factor 3B, subunit 3 | Y(a,b)                 | N                      | Nu/C                                   |
| SNRP70      |                 | U1 Small nuclear ribonucleoprotein 70 kDa | Y(a)                  | N                      | C                                      |
| SNRP31      |                 | Small nuclear ribonucleoprotein polypeptide D3 polypeptide 18 kDa | Beads                  | N                      | C                                      |
| SNX8        |                 | Sorting nexin 8 | nc                     | nc                     | C                                      |
| SPN1        |                 | Spindlin 1 | nc                     | nc                     | Nu/C                                   |
| SR140       | U2SURP          | U2 snRNP-associated SURP domain containing | Beads                  | N                      | Nu                                     |
| SRP14       |                 | Signal recognition particle 14 kDa | Y(a)                  | N                      | Nu                                     |
| SRFS6       | ASF, SF2, SF2p33 | Serine/arginine-rich splicing factor 1 isoform 1 | Y(a,b)                 | Y                      | Nu/C                                   |
| SRFS6       | SFR56, SRF55    | Splicing factor, arginine/serine-rich 6 | Y(a,b)                 | N                      | C                                      |
| SRFS10      | FUSIP1, SRp38   | Serine/arginine-rich splicing factor 10 | Y(a)                  | Y                      | Nu/C                                   |
| SSB         | La, LARP3       | Sjogren syndrome antigen B (autoantigen La) | Y(c)                  | N                      | Nu/C                                   |
| STAA1       |                 | Staufen, RNA-binding protein, homolog 1 (Drosophila) | Y(a)                  | Y                      | C                                      |
| STAA2       |                 | Staufen, RNA-binding protein, homolog 2 (Drosophila) | Y(a,b)                 | Y                      | C                                      |
| STK38       |                 | Serine/threonine protein kinase 38 | Beads                  | Y                      | Nu/C                                   |
| SYNCRIP     |                 | Synaptotagmin binding, cytoplasmic RNA interacting protein | N                     | Y                      | Nu/C                                   |
| TAB1        | MAP3K7IP1       | TGF-\( \beta \) activated kinase 1/MAP3K7-beta binding protein 1 | ?                     | N                      | C                                      |
| TOP1        |                 | DNA topoisomerase 1 | nc                     | nc                     | Nu/C                                   |
| TRA2A       |                 | Transformer-2 \( \beta \) homolog (Drosophila) | Y(a)                  | N                      | Nu/C                                   |
| TRA2B       | SFRS10          | Transformer 2 \( \beta \) homolog (Drosophila) | Y(a)                  | N                      | Nu/C                                   |
| TROVE2      | Ro60, SSA       | 60kDa Ro protein, Sjogren syndrome antigen A2 | Y(a,c)                 | Y                      | Nu/C                                   |
| XRCCT5      | KU80            | X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining) | Beads                  | N                      | Nu/C                                   |
| YBX1        | YBI, NSEPI      | Y box-binding protein 1 | Y(a,b,c)               | o                      | Nu/C                                   |
Figure 2. Proteins identified with the L1 form a tight network of interactions dominated by RNA-binding proteins. (A) Pie chart of results of DAVID (Database for Annotation, Visualization and Integrated Discovery) analysis showing selected functional categories for the 96 candidate proteins (42). Protein counts for each category are shown within the slices. Protein names are listed in Supplementary Table S4. (B) STRING (Search tool for the retrieval of interacting genes/proteins)-derived network of known protein–protein interactions among the 96 candidate proteins. The confidence view is shown in Jensen et al. (43).
It is reasonable to suggest that the association of RNA-binding proteins with the L1 RNP make their bound RNAs preferred targets for pseudogene formation. Of course, expression levels of these RNAs may differ significantly between tumor cells, normal somatic cells and the germ line where retrotransposition acts to establish pseudogenes in the inherited human genome. RNA-binding proteins are also known to associate with

![Figure 3. Ectopically expressed and endogenous proteins associate with L1 complexes in multiple cell lines. (A) V5-, 6xMyc- or GFP-tagged proteins exogenously expressed in 293T cells specifically co-immunoprecipitate with pc-L1-1FH, but not empty vector (pcDNA6 myc/his B) [IP: α-FLAG affinity gel, western blotting (WB): α-V5, α-Myc or α-GFP]. IP reactions were in the presence or absence of 15 μg/ml RNase (lanes 3–5). Lysate input samples are also shown (lanes 1 and 2). Several protein panels are reproduced from Goodier et al. (29). GFP-miGF2BP1 is derived from mouse.](image-url)

(continued)
abundant RNA species in cell extracts after cell lysis; therefore, observed interactions may not recapitulate the in vivo state of ORF1p RNP complexes (66). Furthermore, some small RNAs may be overrepresented as retrogenes simply because of high-RNA copy number. Application to multiple cell lines of recently developed techniques such as PAR-CLIP (67), which combines photo-activatable cross-linking and high-throughput RNA-sequencing, will yield a more comprehensive and quantitative enumeration of the RNA composition of human L1 RNPs.

ORF1p co-localizes in cytoplasmic granules with many RNP-associated proteins

Stress granules (SGs) are cytoplasmic aggregates that contain stalled 48S pre-initiation complexes and are induced by a range of stresses (68). SGs frequently associate with P-bodies (PBs), another class of cytoplasmic granule rich in factors of RNA decay, including those of the RNA-induced silencing complex (RISC). We previously demonstrated that exogenously expressed ORF1p co-localizes with L1 RNA and ORF2p in SGs [(34), confirmed by (35)]. Endogenous ORF1p also forms cytoplasmic granules in the absence of external stress applied to the cell (53).

As described earlier in the text, a majority of proteins we tested directly co-IP with the ORF1p-tagged L1 (Figure 3). We next co-transfected in 293T cells tagged cDNA constructs and ORF1-EGFP-L1-RP, a plasmid containing CMV promoter, ORF1 C-terminally tagged with EGFP, followed by intact downstream L1 sequence (34). One-third of the tagged proteins tested
(23/69) at least partially co-localized with ORF1p-EGFP in cytoplasmic granules in the absence of any applied stress (Figure 5A–W). Panels X and Y show co-localization in granules of ORF1p with endogenous LARP1 and PCBP2. Similarly, we previously reported that endogenous ORF1p co-localizes with endogenous HNRNPA1 and PABPC1 in granules of chemically stressed 2102Ep cells (27). Several additional proteins we report here as associated with the ORF1p RNP have also been detected in cytoplasmic granules by other studies but not this one, perhaps because of differences in cell type (Supplementary Table S5). Indeed, many protein components of SGs and PBs have been identified [reviewed in (68,69)]. To the best of our knowledge, seven proteins of Figure 5 have not been previously reported in cytoplasmic granules, including C22orf28, FAM98A, HNRNPAB, IVNS1APB1, PRPF31, RIOK1 and STK38.

We also previously reported that both ORF1p and ORF2p enter nuclei and concentrate in nucleoli of a minor percentage of cells. A functional nucleolar localization signal was mapped to the EN domain of ORF2p (34,53). These data predict that some nucleolar proteins will associate with those L1 RNPs residing in nucleoli. It is also possible some protein(s) bind L1 RNPs to cause their active transport to nucleoli. We examined the Nucleolar Proteome Database (44) and DAVID analyses and found that 47 of the 96 non-ribosomal proteins that co-IP with the L1 RNP have been reported in nucleoli. These include such canonical nucleolar proteins as nucleolin, fibrillarin and nucleophosmin (Supplementary Table S6). Figure 5Z illustrates co-localization of endogenous fibrillarin and ORF1p in nucleoli of 2102Ep cells. High-resolution microscopy suggests that ORF1p and fibrillarin partially partition into different nucleolar compartments (inset). Fibrillarin concentrates in the dense fibrillar compartment (70).

**L1 RNP-associated proteins inhibit cell culture retrotransposition**

Cellular proteins with roles in retrotransposition, either in its promotion or inhibition, likely associate with the L1 RNP during some stage of its lifecycle. Therefore, we assayed 70 tagged cDNA clones for effect of their overexpression on L1 retrotransposition in a widely used cell culture assay developed in our laboratory (45,46). An EGFP reporter cassette, interrupted by an intron in opposite transcriptional orientation and inserted into the 3' UTR of L1RP (construct 99-PUR-RPS-EGFP), is expressed only when the L1 transcript is spliced, reverse transcribed, its cDNA inserted in the genome and the EGFP reporter gene expressed from its own promoter. We co-transfected in 293T cells 99-PUR-RPS-EGFP, together with empty vector or tagged cDNA constructs, and 5 days post-transfection assayed for fluorescent (i.e. retrotransposition-positive) cells by flow cytometry (Figure 6A).

The duration of the cell culture retrotransposition assay is necessarily long: in 293T cells first retrotransposition events appear after 24 hours, and slowly accumulate over the course of the experiment. Therefore, it is important to determine whether co-expression of proteins have off-target effects that might bias results. Potential effects include cytotoxicity or altered EGFP reporter gene expression caused by ectopic expression of the test protein, or differences in plasmid transfection efficiencies. To reveal any biases, we followed the approach of Wei et al. (47) and, in conjunction with retrotransposition assays, co-transfected pCEP-EGFP, a vector that constitutively expresses EGFP, together with empty vector or tagged
test constructs, and after 4 days performed flow cytometry. Twelve constructs reduced or increased expression from pCEP-EGFP by ≥20%. We, therefore, adjusted the retrotransposition assay results to account for loss/gain of EGFP expression caused by the test-proteins (Figure 6B).

We also used the MultiTox-Fluor Multiplex Cytotoxicity Assay kit (Promega) to directly assay cell toxicity caused by expression of the test proteins. This dual-detection assay generates a ratio of live to dead cell readings, so normalizing for cell number. Most test proteins showed little effect on the viability of 293T cells, and cell toxicity cannot generally account for observed decreases in retrotransposition (Figure 6C).

Proteins tested inhibited retrotransposition or left it little changed (Figure 6B). Fifteen reduced retrotransposition to <25% of wild-type levels. Together these proteins are involved in a range of cellular functions, including RNA splicing modulation, gene expression and RNA transport.

**DISCUSSION**

Little is known of the cellular proteins involved in LINE1 retrotransposition. Our work is the first to comprehensively identify macromolecules associated with the L1 RNP and to assess effects of their overexpression on cell culture retrotransposition. We report the isolation and analysis of human LINE1 RNP complexes and reveal candidate interacting non-ribosomal proteins. Associations with the L1 immunoprecipitates of many of the detected proteins were confirmed by co-IP and subcellular co-localization.

A significant percentage of the proteins identified may be common to many RNP complexes. For example, 18 of
the proteins shown in Table 1 are shared with proteins associated with Ago1 and Ago2 RISC RNP complexes (71). We predict that some proteins we have identified are integral components of the L1 RNP with roles in retrotransposition function and control. Other proteins may be non-specific binders, or tethered to non-L1 RNAs captured within the RNP. As with any affinity purification, association of certain proteins may be indirect because of their presence in other co-purifying multi-protein complexes that contain only one or two proteins that directly interact with the L1 RNP. Some interactions may be cell-line specific or may not represent what occurs during in vivo development. Moreover, an unknown amount of tagged ORF1p, free in solution and not part of retrotransposition-competent RNPs, will have co-purified with immunoprecipitates. Only ongoing functional analyses of individual members of the L1 interactome will ascertain their relevance to L1 biology.

However, the discovery of numerous interacting proteins suggests that the processes of L1 translation, RNP formation, RNA import and export, and genomic insertion recruit a wide spectrum of cellular cofactors. It is of considerable interest that many of the proteins that co-IP with the L1 also co-localize with ORF1p in cytoplasmic RNA granules. The observation that ORF1p foci form constitutively in the absence of exogenous stress prompted us to suggest that the cell recognizes overexpression of the L1 itself as a stress and targets L1 RNPs to granules where they can be sequestered or shunted to P-bodies for degradation, thereby mitigating mutagenic effects of retrotransposition (27). The data of Figure 5 describe candidate proteins that may mediate sequestration of L1 RNPs [one possible example being MOV10, which closely co-localizes with exogenous and endogenous ORF1p in cytoplasmic granules and strongly restricts retrotransposition; (29)]. However, a
direct role for cytoplasmic granules in retrotransposition remains unclear, and studies to date have dealt only with retroviral-like elements. P-body components are in fact important for retrotransposition of Ty1 and Ty3 yeast Long terminal repeat (LTR) retrotransposons (72,73). Contrarily, P-body disruption increases retrotransposition of rodent intracisternal A particle (IAP) elements (74).

The Nucleolar Proteome Database and DAVID reveal that one-half of pc-L1-1FH-associated non-ribosomal proteins can be found in nucleoli (Supplementary Table S6). As with stress granules, a role for nucleoli in L1 retrotransposition is unknown. Both ORF1 and ORF2 proteins have been detected in nucleoli. Many small RNAs that form retropseudogene families spend at least part of their life cycle in nucleoli (53,75); perhaps it is in nucleoli that these small RNAs are captured in L1 RNPs and carried to sites of genome insertion. Interestingly, arthropod non-LTR R1 and R2 retrotransposons have evolved to insert into nuclear 28S rRNA genes (76). Nucleolar proteins that associate with the L1 RNP may provide clues to investigate the role of nucleoli in retrotransposition.

Functional categories within our list of L1 RNP-associated proteins erect signposts to further research. For example, enzymes involved in post-translational modification were detected. Little is known of modifications to retrotransposon proteins and their effects on retrotransposition. Phosphorylation of Ty5 integrase regulates its integration into heterochromatin (77). Hohjoh and Singer (58) reported phosphorylation of L1 ORF1p, but they did not identify the modified residue. We detected with the ORF1p RNP one Ser/Thr protein phosphatase, PPM1B and two Ser/Thr kinases: RIOK1, a protein of unknown function, and STK38, a negative regulator of MAP3K1/2 signaling (78).

Several proteins are involved in the chromosome organization. NAT10 acetylates histones and microtubules and may affect telomerase subunit assembly and localization (79). NUSAP1 is essential for bundling and stabilizing spindle microtubules and may promote their targeting to chromosomes (80). Multiple isoforms of H1 linker histone also co-IP with ORF1p RNPs and inhibit cell culture retrotransposition when overexpressed (Figures 3 and 6B). Recently it was shown that inhibition of Drosophila dH1 increases expression of all classes of mobile DNA, including R1 and R2 retrotransposons (81).

In association with the pc-L1-1FH, we detected RNA and RNP transport proteins, including DDX39A, HNRNPA2B1, HNRNPA1, KPNA2 (importin α1), staufens (STAU1 and STAU2), SRP14, YBX1 and zipcode-binding proteins (IGF2BP1 and IGF2BP2) (Supplementary Table S4). Mouse IGF2BP1 and IGF2BP2, for example, has been shown to mediate nuclear export of the TROVE2 (Ro60)/Y3 snRNA complex (40); both Ro60 and hY3 RNA were detected with the L1 RNP (Figures 3 and 4). Staufens are double-stranded RNA-binding proteins required for microtubule-dependent transport of neuronal RNA through dendrites, and may be involved in RNA export from the nucleus (82). Importin α, in association with importin β, binds proteins at their nuclear localization signals and facilitates RNP transport to the nuclear pore and across the nuclear envelope (83). And SRP14, as part of the signal recognition particle, targets selected proteins to the endoplasmic reticulum. Binding of SRP 9/14 proteins to RNA of Alu elements [themselves evolved from SRP (7SL) RNA] precedes and is likely necessary for efficient L1-mediated Alu retrotransposition (84,85). By studying transport proteins, we may begin to understand how the cell controls movement of L1 RNPs out of the nucleus, into granules, or back into the nucleus to sites of genomic insertion.

We also identified factors involved in alternative mRNA splicing (Supplementary Table S4). TRA2B, for example, regulates tissue-specific and alternative splicing and is essential for mouse development (86). Recently, it was found that HNRNPC, a global suppressor of splicing, binds splice sites within Alu RNAs and competes with splicing factor U2AF65 to prevent aberrant exonization (87). Several SR proteins, and especially SRSF1, mediate not only constitutive and alternative splicing, but other aspects of mRNA elongation, export, translation and decay (88). It is not known whether these factors associate directly with L1 RNA or rather with other mRNAs copurifying with the L1 RNP complex. However, although the L1 lacks introns, its mRNA contains numerous, although mostly weakly predicted splice sites, can undergo extensive and complex splicing in cells and produces products that differ significantly in abundance and pattern between different tumor cell lines. It has been proposed that internal splicing of full-length L1 transcripts adds another level of control against retrotransposon insertion (89,90).

Phylogenetic analyses suggest that eukaryote non-LTR retrotransposons predate all LTR retrotransposons, which in turn gave rise to retroviruses through the acquisition of an envelope (env) gene (91–94). Predictably, some cellular mechanisms of retroviral and retrotransposon inhibition are shared. MOV10 helicase, for example, originally identified as an inhibitor of HIV, strongly restricts L1 retrotransposition in cell culture (29,95). APOBEC3 family members, also first known to inhibit some HIV strains by hypermutation of minus-strand DNA, were subsequently found to be potent inhibitors of retrotransposition in cells. Intriguingly, suppression of L1 and Alu retrotransposons appears not to be accompanied by APOBEC-induced mutation, but it may involve sequestering retrotransposon RNAs in high-molecular weight complexes, possibly stress granules [reviewed in (31,96)]. These multiple modes of activity predict the existence of auxiliary proteins that modulate APOBEC3 function. Our study did not detect APOBECs, probably because they are expressed at low levels in 293T cells (97). However, 35 proteins are shared between the APOBEC3G and our immunoprecipitated L1 complexes (98). APOBEC1, which deaminates apolipoprotein B mRNA, also reduces mobility of mouse L1s by a deamination-independent mechanism (99). Interestingly, two APOBEC1-associated proteins were detected in our L1 RNP reactions: (i) APOBEC1-binding protein-2 (ABBP-2/HNRPA8), a class II DnaJ homolog required for efficient apoB mRNA editing (100) and (ii) SYNCRIP (HNRPQ), which
inhibits apoB mRNA editing (101). Perhaps one of these proteins recruits the L1 to APOBEC complexes.

Figure 6B shows 19 proteins whose expression reduces cell culture retrotransposition by >60%. In some cases, protein overexpression may disrupt other cellular processes, and only indirectly limit L1 activity. Some retrotransposition effects may be cell line specific. With these caveats in mind, this study presents candidate factors worthy of further investigation for key roles in the biology of retrotransposition. We have shown that RISC pathway–associated helicase MOV10 strongly restricts all active human non-LTR retrotransposons in cells (29). Dai et al. (23) reported that PolyA-binding protein C1 (PABPC1) is part of the L1 RNP complex, and important for its formation and retrotransposition. Recently Peddigari et al. (28) used affinity capture to show that four proteins, hnRNPL, hnRNPR, SYNCRIP and NCL, specifically bind a cis-acting ORF2 IRES sequence in mouse L1 RNA. Notably, their knockdown of endogenous hnRNPL increased retrotransposition up to 10-fold in the cell culture assay, whereas knockdown of NCL decreased retrotransposition; SYNCRIP and hnRNPR had no effect. All four of these proteins were also identified in our study, and overexpression of human hnRNPL reduced retrotransposition almost 95% in 293T cells (the three other proteins had no effect; Figure 6B). Such studies justify our approach of cataloguing proteins interacting with the L1 RNP as a way to discover cellular factors involved in modulating retrotransposition.

We performed a literature search to ascertain whether any of the proteins shown in Table 1 also interact with HIV retrovirus. Twenty-one of the 150 top-ranked cellular proteins identified by a proteomics screen as associated with HIV-1 Rev also co-IPed with pc-L1-1FH, including RNA helicases DDX5, DHX9, DDX17 and DDX21. Silencing of DDX3 in HeLa CD4+ cells greatly reduced HIV p24 release and cellular Gag and Env, whereas conversely, silencing DDX5 increased levels of p24 and Gag (102). A functional genomic screen for host proteins that affect HIV-1 infection duplicates three of the four proteins identified by a proteomics screen as associated with HIV-1 replication identified only CCT2 (104), whereas a third (105) identified none of our proteins. In addition to MOV10, several proteins that strongly inhibit retrotransposition (Figure 6) also interact with infectious viruses. PURA binds the HIV-1 trans-activation response (TAR) element to activate transcription and interacts with Rev to enhance nuclear export of unspliced viral RNA (106,107). YBX1 also binds TAR and Tat protein to increase viral promoter activity (108). ELAVL1 binds HIV-1 RT and modulates reverse transcription in infected cells (109). A dominant negative truncation mutant of HNRNPU binds the 3' LTR of HIV-1 mRNA and blocks its accumulation in the cytoplasm (110). Inhibition of HNRNPA1 in T lymphocytes causes an increase of proviral transcription and increased viral production (111). PCBP2 downregulates vesicular stomatitis virus gene expression and is involved in the interferon-α response against hepatitis C virus (112). Finally, IVNS1APB, a protein of unknown function, interacts with the influenza A virus NS1 protein (113). We suggest that by studying factors controlling endogenous retrotransposition, new insights may also emerge into the control of viral infections.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online: Supplementary Tables 1–6, Supplementary Figure 1 and Supplementary References [114–131].

ACKNOWLEDGEMENTS
Special thanks to Dr Robert Cole and the Mass Spectrometry and Proteomics Facility, and the HIT Center ChemCore of Johns Hopkins University. The authors also thank the S. Leach laboratory for use of their confocal microscope.

FUNDING
National Institutes of Health (NIH) [R01 HL083017-01A2 and R01 GM099875-01A1 to H.H.K.]. Funding for open access charge: NIGMS and NIH.

Conflict of interest statement. None declared.

REFERENCES
1. De Koning,A.P.J., Gu,W., Castoe,T.A., Batzer,M.A. and Pollock,D.D. (2011) Repetitive elements may comprise over two-thirds of the human genome. PloS Genet., 7, e1002384.
2. Luan,D.D., Korman,M.H., Jakubczak,J.L. and Eickbush,T.H. (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell, 72, 595–605.
3. Lander,E.S., Linton,L.M., Birren,B., Nusbaum,C., Zody,M.C., Baldwin,J., Devon,K., Dewar,K., Doyle,M., FitzHugh,W. et al. (2001) Initial sequencing and analysis of the human genome. Nature, 409, 860–921.
4. Zhan,Z., Harrison,P.M., Liu,Y. and Gerstein,M. (2003) Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome. Genome Res., 13, 2541–2558.
5. Wang,H., Xing,J., Grover,D., Hedges,D.J., Han,K., Walker,J.A. and Batzer,M.A. (2005) SVA elements: a hominid-specific retroposon family. J. Mol. Biol., 354, 994–1007.
6. Brochu,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.
7. Beck,C.R., Collier,P., Macfarlane,C., Malig,M., Kidd,J.M., Eichler,E.E., Badge,R.M. and Moran,J.V. (2010) LINE-1 retrotransposition activity in human genomes. Cell, 141, 1159–1170.
8. Cordaux,R., Hedges,D.J., Herke,S.W. and Batzer,M.A. (2006) Estimating the retrotransposition rate of human Alu elements. Gene, 373, 134–137.
9. Hancks,D.C. and Kazazian,H.H. Jr (2012) Active human retrotransposons: variation and disease. Curr. Opin. Genet. Dev., 22, 191–203.
10. Muotri,A.R., Chu,V.T., Marchetto,M.C.N., Muotri,A.R., Coufal,N.G., Gage,F.H., O’Shea,K.S. and Moran,J.V. (2007) LINE-1
retrotransposition in human embryonic stem cells. *Hum. Mol. Genet.*, 16, 1569–1577.

12. Solyom,S., Ewing,A.D., Rahmann,F., Doucet,T., Nelson,H., Burns,B., Harris,R., Sigmund,D., Casella,A., Erlanger,B. et al. (2009) Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res.*, 22, 2328–2228.

13. Baillie,J.K., Barnett,M.W., Upton,K.R., Gerhardt,D.J., Richmond,T.A., De Sapio,F., Brennan,P.M., Rizzu,P., Smith,S., Fell,M. et al. (2011) Somatic retrotransposition alters the genetic landscape of the human brain. *Nature*, 479, 534–537.

14. Lee,E., Iskow,R., Yang,L., Gokceumen,O., Haseley,P., Luquette,L.J. 3rd, Lohr,G.H., Harris,C.C., Ding,L., Wilson,R.K. et al. (2012) Landscape of somatic retrotransposition in human cancers. *Science*, 337, 967–971.

15. Evrny,G.D., Cai,X., Lee,E., Hills,L.B., Eholsayy,P.C., Lehmann,H.S., Parker,J.J., Atabay,K.D., Gilmore,E.C., Poduri,A. et al. (2012) Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell*, 151, 483–496.

16. Iskow,R.C., McCabe,M.T., Mills,R.E., Torene,S., Pittard,W.S., Neuward,A.F., Van Meir,E.G., Vertino,P.M. and Devine,S.E. (2010) Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell*, 141, 1253–1261.

17. 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.

18. Burns,K.H. and Boeke,J.D. (2012) Human transposon tectonics. *Cell*, 149, 740–750.

19. Becker,K.G., Swergold,G.D., Ozato,K. and Thayer,R.E. (1993) Binding of the ubiquitous nuclear transcription factor YY1 to a cis regulatory sequence in the human LINE-1 transposable element. *Hum. Mol. Genet.*, 2, 1697–1702.

20. Tchenio,T., Casella,J.F. and Heidmann,T. (2000) Members of the SRY family regulate the human LINE retrotransposons. *Nucleic Acids Res.*, 28, 411–415.

21. Yang,N., Zhang,L., Zhang,Y. and Kazazian,H.H. Jr (2003) An important role for RUNX3 in human L1 transcription and retrotransposition. *Science*, 298, 162, 1–14.

22. Harris,C.R., Dewan,A., Zupnick,A., Normart,R., Gabriel,A., Prives,C., Levine,A.J. and Hoh,J. (2009) p53 responsive elements inactivating genes interact with LINE-1 elements in structural variation and disease. *Hum. Mol. Genet.*, 18, 1697–1702.

23. Sim,S., Yao,J., Weinberg,D.E., Niessen,S., Yates,J.R. 3rd and Moran,J.V. (2000) A novel RNA helicase is a potent inhibitor of retrotransposition in cells. *Mol. Cell.*, 5, 3857–3865.

24. Huang,D.W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, 4, 44–57.

25. Neuwald,A.F., Van Meir,E.G., Vertino,P.M. and Devine,S.E. (2010) Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell*, 141, 1253–1261.

26. Gasior,S.L., Roy-Engel,A.M. and Deininger,P.L. (2008) ERCC1/XPF limits L1 retrotransposition. *Cell Genes Dev.*, 22, 1864–1872.

27. Rebold,D.M., Swergold,G.D., Singer,M.F., Thayer,R.E., Dombroski,B.A. and Fanning,T.G. (1999) Translation of LINE-1 retrotransposon RNA despite a common mechanism of genome insertion. *Hum. Mol. Genet.*, 8, 435–438.

28. Fleckner,J., Zhang,M., Valcárcel,J. and Green,M.R. (1997) A novel ribonucleoprotein particle form accommodates LINE-1 retrotransposons. *Analytical Biochem.*, 254, 435–438.

29. Tackett,H.B. and Dombroski,B.A. (1990) Translation of LINE-1 retrotransposon RNA without phenotypic selection. *Genes Dev.*, 4, 1041–1048.

30. Mooney,S.M., Grande,P.K., Zhang,L. and Janneck,R. (2010) Sumoylation of p68 and p72 RNA helicases affects protein stability and transactivation potential. *Biochemistry*, 49, 1–10.

31. Sim,S., Yoo,J., Weinberg,D.E., Niessen,S., Yates,J.R. 3rd and Moran,J.V. (2002) T抢险inolin (2012) Discrete subcellular partitioning of human retrotransposon RNAs despite a common mechanism of genome insertion. *Hum. Mol. Genet.*, 19, 1717–1725.

32. Doucet,A.J., Hulme,A.E., Sahinovic,E., Kulpa,D.A., Moldovan,J.B., Kopena,H.C., Athanakar,J.H., Ito,M. et al. (2010) Characterization of LINE-1 ribonucleoprotein particles. *PLoS Genet.*, 6, e1001150.

33. Kunkel,T.A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA.*, 82, 488–492.

34. Goodier,J.L., Ostertag,E.M., Engleka,K.A., Selem,M.C. and Kazazian,H.H. Jr (2004) A potential role for the nucleolus in L1 retrotransposition. *Hum. Mol. Genet.*, 13, 1041–1048.
54. De Ruijter,A.J.M., van Gennip,A.H., Caron,H.N., Kemp,S. and van Kuilenburg,A.B.P. (2003) Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem. J., 370, 737–749.

55. Fukuda,M., Gotoh,I., Gotoh,Y. and Nishida,E. (1996) Cytoplasmic localization of mitogen-activated protein kinase kinase directed by its NH2-terminal, leucine-rich short amino acid sequence, which acts as a nuclear export signal. J. Biol. Chem., 271, 20024–20028.

56. Martin,S.L., Branciforte,D., Keller,D. and Bain,D.L. (2003) Trimeric structure for an essential protein in L1 retrotransposition. Proc. Natl Acad. Sci. USA, 100, 13813–13820.

57. Khazina,E., Truffault,V., Bittner,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.

58. Hohjoh,H. and Singer,M.F. (1996) Cytoplasmic ribonucleaseprotein complexes containing human LINE-1 protein and RNA. EMBO J., 15, 630–639.

59. Peterlin,B.M., Brogie,J.E. and Price,D.H. (2012) 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. Wiley Interdiscip. Rev. RNA, 3, 92–105.

60. Perreault,J., Neofitou,E., Courin,A., Bacher,E., Perreault,J.P. and Boire,G. (2005) Retropseudogenes derived from the human Ro/SS-A autoantigen-associated hY RNAs. Nucleic Acids Res., 33, 2032–2041.

61. Wolin,S.L. and Steitz,J.A. (1984) The Ro small cytoplasmic RNA-binding protein and microRNA target sites by PAR-CLIP. J. Mol. Biol., 278, 2071–2078.

62. Fabini,G., Rajmakers,R., Hayer,S., Fouraux,M.A., Prujin,G.J. and Steiner,G. (2001) The heterogeneous nuclear ribonucleoproteins I and K interact with a subset of the ro ribonucleoprotein-associated Y RNAs in vitro and in vivo. J. Biol. Chem., 276, 20711–20718.

63. Fouraux,M.A., Bouvet,P., Verkaart,S., van Venrooj,W.J. and Prujin,G.J.M. (2002) Nucleolin associates with a subset of the human Ro ribonucleoprotein complexes. J. Mol. Biol., 320, 475–488.

64. Murphy,S., Altruda,F., Ullu,E., Tripodi,M., Silengo,L. and Melli,M. (1984) DNA sequences complementary to human 7SK RNA show structural similarities to the short mobile elements of the mammalian genome. J. Mol. Biol., 177, 575–590.

65. Liu,Y.J., Zheng,D., Balasubramanian,S., Carriero,N., Khurana,E., Pruijn,G.J.M. and Cartron,J-P. (2010) Nucleolin associates with a subset of the human Ro ribonucleoprotein complexes. J. Biol. Chem., 285, 129–141.

66. Fukuda,M., Gotoh,I., Gotoh,Y. and Nishida,E. (1996) Cytoplasmic ribonucleoprotein complexes: more than splicing factors. EMBO J., 15, 3353–3362.

67. Dutko,J.A., Kenny,A.E., Gamache,E.R. and Curcio,M.J. (2010) Prussia is a non-SAP in linking microtubules to mitotic chromosomes. Curr. Biol., 20, 230–236.

68. Vujatovic,O., Zaragoza,K., Vaquero,A., Reina,O., Bernue´s,J. and Azorin,F. (2012) Drosophila melanogaster linker histone hD1 is required for transcriptome-wide identification of RNA splicing and influences on mammalian gene expression. Genome Res., 22, 1512–1521.

69. Miki,T., Takano,K. and Yoneda,Y. (2005) The role of mammalian Staufen on mRNA trafficking: a view from its nuclease domain. Proc. Natl Acad. Sci. USA, 102, 1187–1197.

70. Bennett,E.A., Keller,H., Mills,R.E., Schmidt,S., Weichenrieder,O. and Devine,S.E. (2008) Active Alu retrotransposons in the human genome. Genome Res., 18, 1875–1883.

71. Belancio,V.P., Hedges,D.J. and Deininger,P. (2008) Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. Genome Res., 18, 343–358.

72. Xu,J., Chang,D.Y. and Maria,R.J. (1995) Human signal recognition particle (SRP): Alu-associated protein also binds Alu interspersed repeat sequence RNAs. Characterization of human SRP RNA. J. Biol. Chem., 270, 10179–10186.

73. Malik,H.S. and Eickbush,T.H. (2001) Phylogenetic analysis of ribonucleas H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses. Genome Res., 11, 1187–1197.

74. Malik,H.S., Henikoff,S. and Eickbush,T.H. (2000) Posed for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. Genome Res., 10, 1307–1318.
106. Chepenik, L.G., Tretiakova, A.P., Krachmarov, C.P., Kaminski, R., Darbinian, N., Sawaya, B.E., Slonina, D., Amini, S., Zhou, H., Xu, M., Huang, Q., Gates, A.T., Zhang, X.D., Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Liu, L., Oliveira, N.M.M., Cheney, K.M., Pade, C., Dreja, H., Lau, P.P., Villanueva, H., Kobayashi, K., Nakamuta, M., Kress, E., Baydoun, H.H., Bex, F., Gazzolo, L. and Duc Dodon, M. (2006) Inhibition of HIV-1 gene binding protein, Pur-alpha, binds HIV-1 TAR RNA and

105. Liu, L., Oliveira, N.M.M., Cheney, K.M., Pade, C., Dreja, H., Lau, P.P., Villanueva, H., Kobayashi, K., Nakamuta, M., Kress, E., Baydoun, H.H., Bex, F., Gazzolo, L. and Duc Dodon, M. (2006) Inhibition of HIV-1 gene binding protein, Pur-alpha, binds HIV-1 TAR RNA and

104. Zhou, H., Xu, M., Huang, Q., Gates, A.T., Zhang, X.D., Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Liu, L., Oliveira, N.M.M., Cheney, K.M., Pade, C., Dreja, H., Lau, P.P., Villanueva, H., Kobayashi, K., Nakamuta, M., Kress, E., Baydoun, H.H., Bex, F., Gazzolo, L. and Duc Dodon, M. (2006) Inhibition of HIV-1 gene binding protein, Pur-alpha, binds HIV-1 TAR RNA and

103. Brass, A.L., Dykhooorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., Lieberman, J. and Elledge, S.J. (2008) Identification of host proteins required for HIV infection through a functional genomic screen. Science, 319, 921–926.

102. Liu, L., Oliveira, N.M.M., Cheney, K.M., Pade, C., Dreja, H., Bergin, A.-M.H., Bordgoff, V., Beach, D.H., Bishop, C.L., Dittmar, M.T. et al. (2011) A whole-genome screen for HIV restriction factors. Retrovirology, 8, 94.

101. Chepenik, L.G., Tretiakova, A.P., Krachmarov, C.P., Johnson, E.M. and Khaliili, K. (1998) The single-stranded DNA binding protein, Pur-alpha, binds HIV-1 TAR RNA and activates HIV-1 transcription. Gene, 210, 37–44.

100. Liu, L., Oliveira, N.M.M., Cheney, K.M., Pade, C., Dreja, H., Bergin, A.-M.H., Bordgoff, V., Beach, D.H., Bishop, C.L., Dittmar, M.T. et al. (2011) A whole-genome screen for HIV restriction factors. Retrovirology, 8, 94.

99. Chiu, Y.-L., Witkowska, H.E., Hall, S.C., Santiago, M., Soros, V.B., Vetter, M.L., Johnson, E.M., Rappaport, J., Khalili, K. and Darbinyan, A. (2008) HIV-1 intron-containing mRNA. J. Cell. Biochem., 103, 2165–2182.

98. Chiu, Y.-L., Witkowska, H.E., Hall, S.C., Santiago, M., Soros, V.B., Vetter, M.L., Johnson, E.M., Rappaport, J., Khalili, K. and Darbinyan, A. (2008) HIV-1 intron-containing mRNA. J. Cell. Biochem., 103, 2165–2182.

97. Vetter, M.L., Johnson, E.M. and Khalili, K. (1998) The single-stranded DNA binding protein-2, modulates apolipoprotein B mRNA editing. J. Biol. Chem., 273, 32262–32270.

96. Vetter, M.L., Johnson, E.M., Antons, A.K., Rappaport, J., Khalili, K. and Darbinyan, A. (2008) HIV-1 intron-containing mRNA. J. Cell. Biochem., 103, 2165–2182.

95. Arjan-Odedra, S., Swanston, 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.

94. Lemay, J., Maidou-Peindara, P., Bader, T., Ennifar, E., Rain, J.-C., Vallejo-Alvarez, R. and Eversole, A. (2009) Thapsigargin, arsenite and heat-shock treatments. Exp. Cell Res., 315, 968–980.

93. Chepenik, L.G., Tretiakova, A.P., Krachmarov, C.P., Johnson, E.M. and Khaliili, K. (1998) The single-stranded DNA binding protein, Pur-alpha, binds HIV-1 TAR RNA and activates HIV-1 transcription. Gene, 210, 37–44.

92. Cameron, B.M., Furtado, A.J., Schieve, J.M. and Malim, M.H. (2012) Endogenous MOV10 inhibits the retrotransposition of endogenous retroelements but not the replication of exogenous retroviruses. Retrovirology, 9, 53.

91. Lemay, J., Maidou-Peindara, P., Bader, T., Ennifar, E., Rain, J.-C., Benarous, R. and Liu, L.X. (2008) HuR interacts with human immunodeficiency virus type 1 reverse transcriptase, and modulates reverse transcription in infected cells. Retrovirology, 5, 47.

90. Valente, S.T. and Goff, S.P. (2006) Inhibition of HIV-1 gene expression by a fragment of hnrNPA U. Mol. Cell, 23, 597–605.

89. Lemay, J., Maidou-Peindara, P., Bader, T., Ennifar, E., Rain, J.-C., Benarous, R. and Liu, L.X. (2008) HuR interacts with human immunodeficiency virus type 1 reverse transcriptase, and modulates reverse transcription in infected cells. Retrovirology, 5, 47.

88. Anarsi, S.A., Safak, M., Gallia, G.L., Sawaya, B.E., Amini, S. and Khaliili, K. (1999) Interaction of YB-1 with human immunodeficiency virus type 1 Tat and TAR RNA modulates viral promoter activity. J. Gen. Virol., 80(10), 2629–2638.

87. Lemay, J., Maidou-Peindara, P., Bader, T., Ennifar, E., Rain, J.-C., Benarous, R. and Liu, L.X. (2008) HuR interacts with human immunodeficiency virus type 1 reverse transcriptase, and modulates reverse transcription in infected cells. Retrovirology, 5, 47.

86. Valente, S.T. and Goff, S.P. (2006) Inhibition of HIV-1 gene expression by a fragment of hnrNPA U. Mol. Cell, 23, 597–605.

85. Anarsi, S.A., Safak, M., Gallia, G.L., Sawaya, B.E., Amini, S. and Khaliili, K. (1999) Interaction of YB-1 with human immunodeficiency virus type 1 Tat and TAR RNA modulates viral promoter activity. J. Gen. Virol., 80(10), 2629–2638.

84. Lemay, J., Maidou-Peindara, P., Bader, T., Ennifar, E., Rain, J.-C., Benarous, R. and Liu, L.X. (2008) HuR interacts with human immunodeficiency virus type 1 reverse transcriptase, and modulates reverse transcription in infected cells. Retrovirology, 5, 47.

83. Valente, S.T. and Goff, S.P. (2006) Inhibition of HIV-1 gene expression by a fragment of hnrNPA U. Mol. Cell, 23, 597–605.

82. Anarsi, S.A., Safak, M., Gallia, G.L., Sawaya, B.E., Amini, S. and Khaliili, K. (1999) Interaction of YB-1 with human immunodeficiency virus type 1 Tat and TAR RNA modulates viral promoter activity. J. Gen. Virol., 80(10), 2629–2638.