Published online 10 November 2010

Nucleic Acids Research, 2011, Vol. 39, No. 3 ***e16***
doi:10.1093/nar/gkq1076

Characterization of L1 retrotransposition with high-throughput dual-luciferase assays

Yi Xie[1], James M. Rosser[1], Tina L. Thompson[1], Jef D. Boeke[2] and Wenfeng An[1,*]

[1]School of Molecular Biosciences and Center for Reproductive Biology, Washington State University, Pullman, WA 99164 and [2]The High Throughput Biology Center and Department of Molecular Biology and Genetics, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

Received September 12, 2010; Revised October 10, 2010; Accepted October 14, 2010

ABSTRACT

Recent studies employing genome-wide approaches have provided an unprecedented view of the scope of L1 activities on structural variations in the human genome, and further reinforced the role of L1s as one of the major driving forces behind human genome evolution. The rapid identification of novel L1 elements by these high-throughput approaches demands improved L1 functional assays. However, the existing assays use antibiotic selection markers or fluorescent proteins as reporters; neither is amenable to miniaturization. To increase assay sensitivity and throughput, we have developed a third generation assay by using dual-luciferase reporters, in which firefly luciferase is used as the retrotransposition indicator and Renilla luciferase is encoded on the same or separate plasmid for normalization. This novel assay is highly sensitive and has a broad dynamic range. Quantitative data with high signal-to-noise ratios can be obtained from 24- up to 96-well plates in 2–4 days after transfection. Using the dual-luciferase assays, we have characterized profiles of retrotransposition by various human and mouse L1 elements, and detailed the kinetics of L1 retrotransposition in cultured cells. Its high-throughput and short assay timeframe make it well suited for routine tests as well as large-scale screening efforts.

INTRODUCTION

The preponderance of long interspersed elements type 1 (LINE-1s or L1s) in the human genome was eloquently revealed by the human genome sequencing project: as the most abundant autonomous transposable element in the human genome, they account for ~17% of the human genome mass (1). L1s are arguably one of the major driving forces behind human genome evolution (2,3). They shape the genomic architecture through active transposition and ectopic recombination between existing non-allelic homologous elements. In addition to self-propagation, L1s can mobilize other transcribed DNA sequences (4,5). A prominent example is Alu elements, which have outnumbered L1s and account for ~11% of the human genome mass (6). A hominid-specific transposon family, SVA elements, is also believed to be *trans*-mobilized by the L1 machinery (7,8). L1-mediated insertional and post-insertional mutagenesis are also responsible for a wide spectrum of sporadic human diseases (9,10). A previous survey of full-length L1s from the reference human genome indicates that there are 80–100 retrotransposition-competent L1s in an average human individual, and that a few highly active (or hot) L1s account for the bulk of retrotransposition activities in the human population (11). Using genome-wide approaches, such as paired-end sequencing of fosmid libraries, microarray (TIP-chip) and deep sequencing (L1-Seq), recent studies provide an unprecedented view of the scope and impact of L1 activities on structural variations in the human genome (12–15). For example, a genome-wide study of six individuals identified 69 dimorphic full-length L1 elements that are absent from the reference human genome, and the majority of these L1s were categorized as hot L1s using a retrotransposition assay (12). Hundreds of novel L1 insertions were recovered from individual genomes in three independent genome-wide transposon mapping studies (13–15); accordingly, the frequency of L1 retrotransposition in the human germ line has been revised upward to one insertion in every 108 or 140 live births (13,14). One study also provides strong evidence that L1 can act as a potent endogenous mutagen in somatic tissues because abundant *de novo* L1 insertions were documented in human lung cancer genomes (15). The rapid identification of novel

*To whom correspondence should be addressed. Tel: +1 509 335 8296; Fax: +1 509 335 4159; Email: wenfeng_an@wsu.edu

L1 elements by these high-throughput approaches necessitates the development of improved L1 functional assays.

Cell-based L1 functional assays are essential tools for studying L1 biology (Table 1). Two types of L1 retrotransposition assays are currently available. The first type employs antibiotic selectable marker genes, such as the neomycin phosphotransferase gene (*neo*), and the assay readout is colony formation (16). Since its inception, the *neo*-based assay has been used in all aspects of L1 studies, leading to the identification of essential L1 sequences, other active L1s, and a wide range of effects that L1 retrotransposition exerts on mammalian genomes (17). However, this assay typically takes 2–4 weeks to complete, is labor-intensive and has very low throughput (16,18). The second type of L1 assay is based on the enhanced green fluorescence protein (EGFP) (19). It has a shorter assay timeframe, and usually can be finished in 7 days. However, the EGFP-based assay has a relatively narrow dynamic range and lacks sensitivity in higher density multi-well plates. In order to increase assay sensitivity and throughput, we have now developed a third-generation assay for L1 retrotransposition using a dual-luciferase assay scheme: the firefly luciferase (Fluc) gene is used as the reporter for L1 retrotransposition and the Renilla luciferase (Rluc) gene for transfection normalization. Quantitative data can be obtained from 24- to 96-well plates in 2–4 days after transfection by using common luminometry equipment. This novel assay is highly sensitive and has a wide dynamic range. Using this assay, we have characterized retrotransposition by various human and mouse L1 elements, including the kinetics of L1 retrotransposition.

MATERIALS AND METHODS

Plasmids

Intron-disrupted Fluc expression cassettes were derived from pGL4.13 (Promega), which contains a simian virus 40 (SV40) early enhancer/promoter, an optimized Fluc reporter gene and an SV40 late poly(A) signal. A 900-bp fragment of the human γ-globin intron [or a 133-bp fragment of the synthetic intron from pCI (Promega)] was introduced in its antisense orientation into pGL4.13 at nucleotide position 514 of the Fluc reporter gene through PCR-mediated site-directed mutagenesis, generating pWA292 (or pWA293). The SV40 late poly(A) signal in pWA292 (or pWA293) was removed by FseI(blunted)/BamHI digestion, and replaced by a BstEII(blunted)/BglII digested herpes simplex virus thymidine kinase (HSV-TK) poly(A) signal, resulting in pWA304 (or pWA303). The intron-disrupted Fluc cassette from pWA304 (or pWA303) was released as a BglII/EcoRV fragment, which was ligated either to a PmlI/BamHI digested CMVCAG-*ORFeus* plasmid to derive pWA346 (or pWA345), or to a NotI/BstZ17I fragment of 5′-UTR-$L1_{RP}$ from pJCC5-RPS (19) and a NotI/BamHI fragment from a CMV-less version of pCEP-Puro (20) to derive pWA355 (or pWA354). The functional $L1_{RP}$ in pWA355 (or pWA354) was replaced by the corresponding NotI/BamHI fragment of $L1_{RP}$/JM111 from pBS-$L1_{RP}$(JM111)-EGFP (19) to derive pWA366 (or pWA356). pWA357 is an $L1_{RP}$ vector with both CMV and 5′-UTR promoters; it was constructed by ligating a NotI/MluI fragment of 5′-UTR-$L1_{RP}$-Fluc from pWA355 to NotI/MluI linearized pCEP-Puro backbone. pWA358 is an $L1_{RP}$ vector with CMV promoter only; it was constructed by three-way ligation of PvuII/BamHI and BamHI/MluI fragments from pWA355 and a NgoMIV(blunted)/MluI fragment of pCEP-Puro. pWA359 is an $L1_{RP}$ vector with both CMV and CAG promoters; it was constructed by three-way ligation of PvuII/BamHI and BamHI/MluI fragments from pWA355 and an AscI(blunted)/MluI fragment of pWA346. pWA367 is an $L1_{RP}$ vector with CAG promoter only; it was constructed by transferring CAG-$L1_{RP}$ in pWA359 as a NotI/BamHI fragment to NotI/BamHI linearized pWA355.

Intron-disrupted Rluc expression cassettes were derived from pGL4.73 (Promega) that contains an SV40 early enhancer/promoter, an optimized Rluc reporter gene and an SV40 late poly(A) signal. A 900-bp fragment of the human γ-globin intron (or a 133-bp fragment of the synthetic intron from pCI) was introduced in its antisense orientation into pGL4.73 at nucleotide position 206 of the Rluc reporter gene through PCR-mediated site-directed mutagenesis, generating pWA296 (or pWA297). The SV40 late poly(A) signal in pWA296 (or pWA297) was removed by FseI(blunted)/BamHI digestion, and replaced by a BstEII(blunted)/BglII digested HSV TK poly(A) signal, resulting in pWA307 (or pWA306). The intron-disrupted Rluc cassette from pWA307 (or pWA306) was released as a SalI(blunted)/BglII fragment (or a BglII/HincII fragment), and ligated to a PmlI/BamHI digested CMVCAG-*ORFeus* plasmid to derive pWA349 (or pWA348).

pWA003 and pWA196 are two control vectors. pWA003 contains an EGFP expression cassette on a pCEP-Puro backbone. It was derived by ligating the EGFP gene as a NheI/BamHI fragment from pEGFP-C1 (Clontech) into NheI/BamHI linearized pCEP-Puro. pWA196 contains a 5′-UTR-$L1_{RP}$ element tagged by an EGFP retrotransposition indicator. It was created by ligating a NotI/BstZ17I fragment of 5′-UTR-$L1_{RP}$ from pJCC5-RPS and a SalI(blunted)/BamHI fragment of pBSKS-EGFP-INT (19) into NotI/BamHI linearized CMV-less version of pCEP-Puro.

pYX013, pYX014, pYX015, pYX016 and pYX017 are plasmids for single-vector assays, and they all contain an intact Rluc expression cassette on the vector backbone for

Table 1. Characteristics of different L1 reporter assays

Reporter	Protocol	Timeframe (days)[a]	Special instrument	Thoughput	Reference
neo	Standard	32–38	No	≥6 well	(16)
neo	Transient	15	No	≥6 well	(18)
EGFP	Standard	7	Flow cytometer	6 well	(19)
Fluc	Standard	2–4	Luminometer	24, 96 well	This study

[a]Days post transfection.

normalizing transfection efficiency. The control Rluc expression cassette was derived from pGL4.73 in two subcloning steps. The SV40 late poly(A) signal in pGL4.73 was first removed by FseI(blunted)/BamHI digestion, and replaced by an HSV TK poly(A) signal with BstEII(blunted)/BglII ends, resulting in pWA305. The SV40 early enhancer/promoter in pWA305 was then removed by FspI/HindIII digestion, and replaced by an HSV TK promoter as an FspI/HindIII fragment, resulting in pTT010. pWA003 was partially digested by EcoRI, and the EcoRI ends were blunted by Klenow and subsequently religated to create pYX011, which lacks the EcoRI site at nucleotide position 7965. The Rluc cassette was released from pTT010 as an FspI/PciI fragment and ligated into EcoRI(blunted)/PciI digested pYX011 to make pYX013. The 5′-UTR-$L1_{RP}$-Fluc element was released from pWA355 as a SacII/PvuI fragment, and ligated to two pYX013 fragments (SacII/MluI and MluI/PvuI) to create pYX014. The 5′-UTR-$L1_{RP}$/JM111-Fluc element was released from pWA366 as an SacII/PvuI fragment, and ligated to two pYX013 fragments (SacII/MluI and MluI/PvuI) to create pYX015. The CAG-$L1_{RP}$-Fluc element was released from pWA367 as a NotI/SwaI fragment, and ligated to NotI/SwaI linearized pYX014 to create pYX017. The CAG-*ORFeus*-Fluc element was released from pWA346 as a NotI/SwaI fragment, and ligated to NotI/SwaI linearized pYX014 to create pYX016.

Cell culture

HeLa cells were grown at 37°C in a humidified 5% CO_2 incubator in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum and 2 mM stable dipeptide of L-alanyl-L-glutamine. Cells were passaged every three days by using 0.25% porcine trypsin. All medium components for HeLa cell culture were purchased from Hyclone.

L1 retrotransposition assay

HeLa cells were seeded in 24-well plates at a density of 7×10^4 cells/well and grown up to 60–70% confluence. Endotoxin-free plasmid DNA was prepared with anion-exchange resin columns (Qiagen) and diluted in sterile 10 mM Tris–HCl buffer. Cells were transfected with Lipofectamine 2000 reagent (Invitrogen) following the manufacturer's protocol. Each well received 200 ng plasmid DNA in the single-vector assay, 0.6 µl Lipofectamine 2000 reagent and 100 µl GlutaMAX-I-supplemented Opti-MEM I reduced-serum medium (Invitrogen). In the two-vector assay, a 20:1 ratio of L1 plasmid (200 ng) and the Rluc control plasmid pGL4.73 (10 ng) was used. Puromycin (Sigma-Aldrich) was added in complete medium at a final concentration of 2.5 µg/ml at 24 h post-transfection. Cells were harvested 4 days post-transfection for luciferase activity analysis unless otherwise indicated. To perform single-vector assays in 96-well plates, HeLa cells were seeded at a density of 1×10^4 cells/well and grown up to 80–85% confluence. One-hundred-nanogram plasmid DNA, 0.3 µl Lipofectamine 2000 reagent and 50 µl GlutaMAX-I-supplemented Opti-MEM I reduced-serum medium were used per well.

Nucleoside reverse transcriptase inhibitors dosing and kinetics assays

HeLa cells were seeded in 24-well plate and transfected with L1 plasmids as described above. All nucleoside reverse transcriptase inhibitors (NRTIs) were acquired from Sigma-Aldrich, including 2′-3′-didehydro-2′-3′-dideoxythymidine (d4T), 2′-3′-dideoxy-3′-thiacytidine (3TC) and 2′3′-dideoxyinosine (ddI). To determine effective concentrations, NRTIs were added to the culture medium at 2 h before transfection, and were replenished every 24 h onwards until the endpoint of an experiment. To determine L1 retrotransposition kinetics, cells in parallel wells were treated with NRTIs at 2 h before transfection, or 16, 24, 28, 32, 40 and 48 h post-transfection, respectively. Untreated cells were used as controls for full activity. Puromycin was supplemented during each medium exchange starting from 24 h post-transfection. Cells were harvested 4 days post-transfection for luminescence analysis.

Dual-luciferase luminescence measurement

Luminescence was measured using the Dual-Luciferase Reporter Assay System (Promega) following the manufacture's instructions. For assays in 24-well plates, 100 µl Passive Lysis Buffer was used to lyse cells in each well; for assays in 96-well plates, 25 µl was added per well. For all assays, 20 µl lysate was transferred to a solid white 96-well plate, mixed with 100 µl Luciferase Assay Reagent II and Fluc activity was read on a GloMax-Multi Detection System (Promega). Rluc activity was subsequently read after mixing 100 µl Stop & Glo Reagent into the cell lysate containing Luciferase Assay Reagent II.

Data normalization and statistical analysis

To facilitate comparison of L1 activities from different elements, we calculated a normalized luminescence ratio (NLR) for each vector (see 'Results' section). We routinely used the retrotransposition-defective $L1_{RP}$/JM111 as the reference Fluc vector. In all our tests, Fluc signals were indistinguishable between $L1_{RP}$/JM111 and EGFP-expressing vectors carrying the same vector backbone. However, Rluc signals were typically 1.5- to 2.0-fold lower for EGFP-expressing vectors (i.e. pWA003 and pYX013) than those for $L1_{RP}$/JM111 in the same assay (data not shown). Such differences in Rluc signals can explain why EGFP-expressing control vectors had NLR values slightly higher than expected (in the range of 1.4–2.2 when $L1_{RP}$/JM111 was used as the reference vector). The relative reduction of Rluc signals in cells transfected with EGFP-expressing vectors is likely caused by promoter interference, in which a strong promoter interferes with the test promoter by competing for general or specific transcription factors. We used a constant amount of DNA for different transfections, but 1.7- to 2-fold more DNA was used for the control plasmids due to their smaller sizes. In fact, Rluc signals from $L1_{RP}$/JM111 were highly similar to those from

retrotransposition-competent L1 vectors. Thus, a retrotransposition-defective L1 is best suited to serve as the reference Fluc vector. However, this does not exclude the use of non-L1 vectors, such as pWA003 or pYX013, as reference Fluc vectors as long as the same vector is used as the reference for all assay conditions. To compare retrotransposition rates between different conditions, we used two-tailed Student's *t*-test.

Genomic DNA PCR and quantitative PCR

To confirm intron removal from *de novo* L1 insertions, HeLa cells transfected with pYX013, pYX014 or pYX015 were harvested at 48 h post-transfection. Genomic DNA was extracted with a DNeasy Tissue Kit (Qiagen). pGL4.13 plasmid DNA was used as an intron- control; pWA292 was used as an intron+ control. The following primers specific to the Fluc coding sequence were used with ExTaq polymerase (Takara): Fluc-P1 5′-CAACGTGCAAAAGAAGCTACC-3′ and Fluc-P2 5′-ATGACTGAATCGGACACAAGC-3′. The PCR cycling parameters consist of an initial denaturation at 94°C for 2 min, 30 cycles of denaturing at 94°C for 15 s, annealing at 60°C for 30 s and extending at 72°C for 30 s, and a final extension step at 72°C for 7 min. PCR product was resolved on a 1% agarose gel containing 0.5 μg/ml ethidium bromide, and documented with a Gel Doc XR system (Bio-Rad). To determine the copy number of *de novo* L1 insertions, we performed quantitative PCR (qPCR) according to a previously established protocol (21). For each genomic DNA sample, L1 insertions were amplified in triplicate with primers and an exon-exon junction probe specific to the Fluc gene. Human GAPDH was amplified in parallel wells as an endogenous control for normalization. The following primers and probes were used: GAPDH forward 5′-CCCACTCCTC CACCTTTGAC-3′, GAPDH reverse 5′-TGTTGCTGTA GCCAAATTCGTT-3′, GAPDH probe 5′-AAGCTCATT TCCTGGTATGA-3′; Fluc forward 5′-GCAAAAGAAG CTACCGATCATACA-3′, Fluc reverse 5′-GAAGCTCT CGGGCACGAA-3′, Fluc exon-exon junction probe 5′-C TTCCCACCTGCCACC-3′. qPCR data were analyzed by the comparative C_T method, and normalized to signals from cells transfected with CAG-*ORFeus* and harvested at 48 h post-transfection. Each data point represents average ± SE from four separate transfections.

RESULTS

A two-vector dual-luciferase assay for L1 retrotransposition

We implemented a dual-luciferase scheme for assaying L1 retrotransposition. The Fluc gene was used to report L1 activity (Figure 1A). As in previous L1 reporter assays, the coding sequence of the Fluc reporter gene was interrupted by an antisense intron. As expected, no Fluc signal was detected above background when a stand-alone intron-disrupted Fluc reporter cassette was transfected into HeLa cells (data not shown). The Fluc cassette was first tested as a reporter for $L1_{RP}$. $L1_{RP}$ is the causative full-length L1 insertion isolated from a patient with *retinitis pigmentosa* (22). It remains one of the most active endogenous L1 elements identified thus far and has been featured in a wide range of L1 studies. The Fluc cassette was inserted in the 3′-untranslated region (3′-UTR) of $L1_{RP}$ in the antisense orientation relative to L1 expression. Such a sequence arrangement ensures that Fluc signals arise in transfected cells only after the donor L1 has undergone one round of retrotransposition

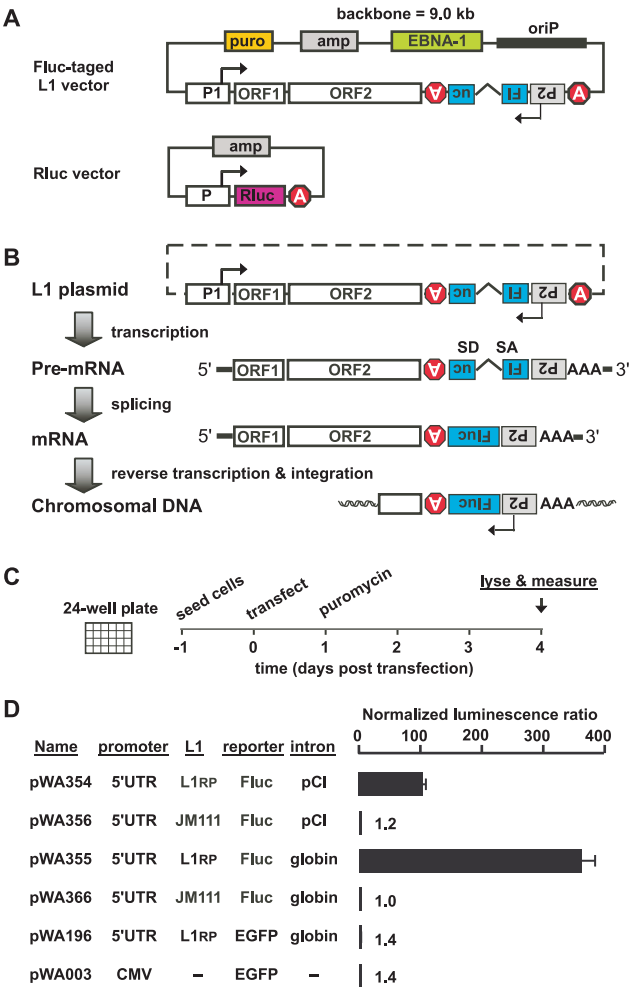

Figure 1. A two-vector dual-luciferase assay for L1 retrotransposition. (**A**) Diagrams of a Rluc control vector (3.9 kb) and an L1 vector tagged by an intron-disrupted Fluc reporter (18.3 kb for pWA355). L1 is expressed from a pCEP4-based episomal plasmid, and marked by a Fluc gene interrupted by an antisense intron (also known as a retrotransposition reporter). The Fluc cassette has its own promoter (P2) and polyadenylation signal, and is expressed from the antisense strand relative to the L1 promoter (P1). (**B**) The rationale of L1 retrotransposition assay. The level of L1 retrotransposition is detected as Fluc activity. No Fluc signal is expected in transfected cells unless the donor L1 has undergone one round of retrotransposition (i.e. transcription, splicing, reverse transcription and integration into the genome). Adapted from Moran *et al.* (16). (**C**) A typical workflow in 24-well plates. (**D**) Detection of L1 retrotransposition with a two-vector dual-luciferase assay. Either $L1_{RP}$ or $L1_{RP}$/JM111 (a mutated $L1_{RP}$ carrying two missense mutations in ORF1) was used as the donor element. Two different introns were used to construct the Fluc reporter cassette. L1 activity is expressed as NLRs relative to the background signal from pWA366. Error bars represent mean ± SE ($N = 4$).

(i.e. transcription, splicing, reverse transcription and integration into the genome) (Figure 1B). The Rluc gene was used to normalize transfection efficiency by cotransfection (Figure 1A). A standard assay was performed in 24-well plates, and could be completed in 4 days post-transfection (Figure 1C). Briefly, HeLa cells were cotransfected with a Fluc-tagged L1 vector and a Rluc control vector. Transfected cells were selected with puromycin at 24 h post-transfection, and lysed for luminescence analysis at 4 days post-transfection. Luminescence was measured in a 96-well plate, and could be completed within 30 min starting from cell lysis.

We first assessed the background luminescence for both Fluc and Rluc activities by using several control vectors, including an $L1_{RP}$ tagged with an EGFP retrotransposition indicator (pWA196) and an EGFP expression vector in the same plasmid backbone (pWA003); both showed background level of luminescence that was indistinguishable from mock transfected cells (Figure 1D). Such background luminescence was consistent from experiment to experiment, and ranged from 100 to 300 relative light units (RLUs). Cells cotransfected with the Rluc control vector yielded Rluc signals >1000-fold above the background. Fluc signals from cells transfected by $L1_{RP}$ vectors (pWA355) were >100-fold above the background (Figure 1D). To confirm that the observed Fluc signals were originated from L1 retrotransposition, we tagged $L1_{RP}$/JM111, which carried missense mutations in the ORF1 coding sequence. These mutations have previously been shown to completely abolish L1 retrotransposition (16). As expected, Fluc readings from $L1_{RP}$/JM111 vectors (such as pWA366) were indistinguishable from background luminescence (Figure 1D). To facilitate comparison of L1 activities from different elements, we calculated a NLR for each vector. $L1_{RP}$/JM111 was used as the reference Fluc vector (the choice of reference vector is discussed in Materials and methods section). For example, $NLR_{pWA355} = (Fluc_{pWA355}/Rluc_{pWA355}) / (Fluc_{pWA366}/Rluc_{pWA366})$. In a representative experiment, an NLR value of 363 was obtained from the 5′-UTR-$L1_{RP}$ element (Figure 1D), indicating that the 5′-UTR-$L1_{RP}$ element retrotransposed at a level that was 363-fold higher than that observed with a retrotransposition-defective L1.

Choice of introns for L1 reporter

We tested two different introns for their suitability to disrupt the luciferase reporter gene. The first intron is from the human γ-globin gene, which has been previously used in *neo*- and EGFP-based reporters (16,19). The second is a small synthetic intron from plasmid pCI, which was chosen for its smaller size. Their effects on L1 activity were tested by using 5′-UTR-$L1_{RP}$ as the donor element (Figure 1D). A 3.5-fold higher Fluc signal was observed when the γ-globin intron was used in the Fluc reporter than when the pCI intron was used, presumably because the γ-globin intron had a higher splicing efficiency (Figure 1D, comparing pWA355 to pWA354). These two introns were also tested as part of the Fluc reporter cassette when a hyperactive synthetic mouse L1, *ORFeus* (23,24), was used as the donor element (Figure 2A). The hyperactive nature of *ORFeus* was indeed confirmed by the new luciferase-based assay: *ORFeus* generated Fluc signals at more than 8000-fold above the background (Figure 2A). Interestingly, the choice of intron did not significantly affect the overall L1 activity when *ORFeus* was used as the donor: a Fluc cassette containing the pCI intron yielded a 12% higher signal than the γ-globin-intron-containing Fluc cassette (Figure 2A). Thus the intron efficiency appears to be influenced by long range sequence context. This may be due to the extent to which cryptic splice sites within each element contribute to aberrant splice variants that reduce overall retrotransposition efficiency. We decided to use the γ-globin intron for later experiments because it tends to generate higher Fluc signals when the overall L1 activity is in the lower range, as is the case when 5′-UTR-$L1_{RP}$ was used as the donor (Figure 1D). This conclusion was also supported by the following experiments when *ORFeus* was used as the donor element and Rluc was used as the retrotransposition reporter (Figure 2B, discussed below).

Choice of luciferase genes as L1 reporter

In dual-luciferase applications, Fluc is preferred over Rluc as the 'experimental' reporter, while Rluc is reserved as the 'control' reporter for signal normalization, because the dynamic range of Fluc is one order of magnitude higher than that of Rluc. However, a unique feature of L1 retrotransposition assays is that signals from the experimental reporter can only be generated from *de novo* L1 insertions that contain an intact (i.e. full length) experimental reporter cassette. Since the majority of L1 insertions are 5′ truncated (25,26), the length of a reporter gene could potentially be a critical factor in determining the overall L1 activity that can be observed. As the Rluc gene (936 bp) is 1.8-fold shorter than the Fluc gene (1653 bp), we decided to test whether Rluc was more efficient than Fluc as an L1 reporter for the hyperactive *ORFeus* element (Figure 2B). As compared to

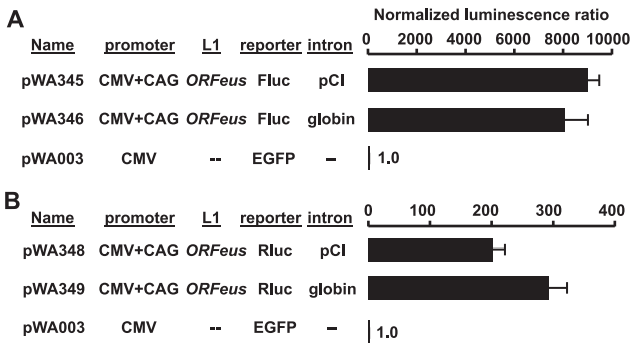

Figure 2. Evaluation of different luciferase genes and introns for L1 reporters. (**A**) Use of Fluc as the L1 reporter and Rluc as the cotransfection control. Either pCI or γ-globin intron was used to disrupt the Fluc reporter cassette. (**B**) Use of Rluc as the L1 reporter and Fluc as the cotransfection control. Either pCI or γ-globin intron was used to disrupt the Rluc reporter cassette. In both (A and B), CMV-CAG-*ORFeus* was used as the donor L1. L1 activity is expressed as NLRs relative to the background signal from pWA003. Error bars represent mean ± SE ($N = 4$).

Fluc-based reporters, a dramatic decrease in the experimental signal was observed for Rluc-based reporters: a 45-fold decrease when the pCI intron was used to disrupt the reporter gene (comparing pWA345 and pWA348) and a 28-fold decrease when the γ-globin intron was used (comparing pWA346 and pWA349) (Figure 2). Thus, we confirmed that Fluc was superior to Rluc when used to report L1 activity, despite its longer coding sequence.

The utility of two-vector dual-luciferase assay for detecting L1 activities with different promoters

To demonstrate the sensitivity of the newly established two-vector dual-luciferase assay, we tested whether it could distinguish retrotransposition levels of several L1 retrotransposition vectors with different promoters or tandem combinations of promoters. We compared $L1_{RP}$ elements under the regulation of its native promoter (5′-UTR), and several heterologous promoters, including CMV, CAG and tandem promoters (Figure 3). Among all promoters tested, the lowest L1 activity was observed from the CMV promoter (NLR = 69), and the highest activity was from the CAG promoter (NLR = 1404), representing a 20-fold activity range. Other promoter(s) displayed intermediate L1 activities at the following order: CAG > CMV+CAG > CMV+5′-UTR > 5′-UTR > CMV (Figure 3).

A single-vector dual-luciferase assay for L1 retrotransposition

Thus far, we have demonstrated dual-luciferase assays for L1 retrotransposition by using two separate plasmids: an L1 plasmid tagged with a Fluc reporter, and a control plasmid encoding Rluc. To test any L1 plasmid in the two-vector assay, cotransfection with the control plasmid at a consistent molar ratio of 20:1 (L1-Fluc vector versus the Rluc control) is required. To further simplify experimental protocol and minimize experimental variation, we tested a single-vector dual-luciferase strategy (Figure 4). In the single-vector configuration, a Rluc expression cassette was incorporated into the backbone of the L1 assay vector (Figure 4A). We modified the Rluc cassette by using a weaker HSV-TK promoter (the control

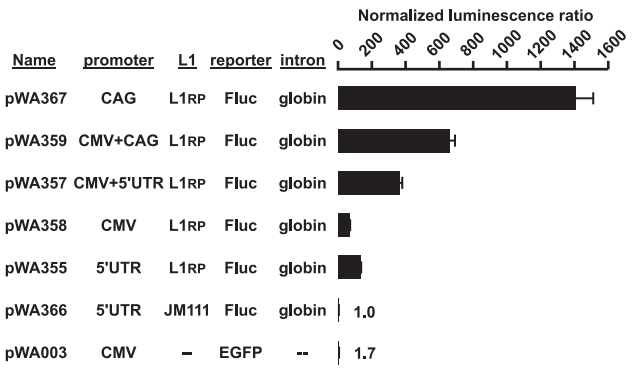

Figure 3. Effect of different promoters on L1 retrotransposition. L1 activity is expressed as NLRs relative to the background signal from pWA366. Error bars represent mean ± SE ($N = 6$).

Rluc vector pGL4.73 used in the two-vector assay carries an SV40 promoter). The promoter exchange yielded optimal Rluc signals for the one-vector assay as the Rluc cassette was present at a molar ratio of 1:1 relative to the L1 element. In addition, sequence elements responsible for episomal maintenance of pCEP-Puro derived plasmids were removed from the single assay vector. These sequences include the EBNA-1 gene and Epstein-Barr virus replication origin *ori*P, which were dispensable for *neo*-based L1 assays (20). Consequently, the overall size of the assay vector was reduced by 3.2 kb (comparing Figure 1A to Figure 4A).

We tested the single-vector dual-luciferase scheme by comparing L1 activities from several different L1 elements (Figure 4B). The 5′-UTR-$L1_{RP}$ element (pYX014) displayed an NLR signal of 383-fold above background, comparable to the two-vector assay (see pWA355 in Figure 1D). Substitution of the L1 5′-UTR by a strong CAG promoter increased retrotransposition 2.7-fold (pYX017, NLR = 1017). Further substitution of $L1_{RP}$ by *ORFeus* resulted in another 45-fold increase on overall retrotransposition activity (pYX016, NLR = 46067).

The superlative activity from the CAG-*ORFeus* element made it possible to measure the dynamic range of our luciferase-based L1 assays. Serial dilutions of CAG-*ORFeus* transfected cells yielded linear signals for both Fluc and Rluc in the range from 800 to 500 000 cells (Figure 4C). This is equivalent to a minimum of 600-fold dynamic range. The weakest Fluc signals from this experiment were obtained from 800-cell samples, and remained significantly above the background luminescence (~53-fold) at an average of 10 558 RLUs. To further confirm that Fluc signals result from L1 retrotransposition, we monitored intron removal by genomic DNA PCR (Figure 4D). Genomic DNA from transfected HeLa cells were amplified by primers flanking the γ-globin intron in the Fluc reporter cassette (Figure 4A). An intronless band of 250 bp was present in cells transfected with a retrotransposition-competent 5′-UTR-$L1_{RP}$ element (pYX014) but absent in cells transfected with an $L1_{RP}$/JM111 element (pYX015), indicating that a functional Fluc gene was reconstituted as the result of retrotransposition (Figure 4D).

Using reverse transcriptase inhibitors to evaluate the kinetics of L1 retrotransposition

The kinetics of L1 retrotransposition has been recently investigated by using a *neo*-based L1 reporter assay and timed d4T treatment (27). As an NRTI, d4T inhibits reverse transcription after being incorporated into the growing cDNA chain because the phosphorylated d4T lacks a 3′-hydroxyl group. Thus, in such experiments, the timing of L1 retrotransposition is defined as the time required for L1 reverse transcriptase to complete reverse transcription after the donor L1 is transfected into cells. Note this operational definition does not cover the remaining steps of L1 retrotransposition, such as the second strand synthesis and subsequent integration of a new copy of L1 into the genome. To evaluate the kinetics

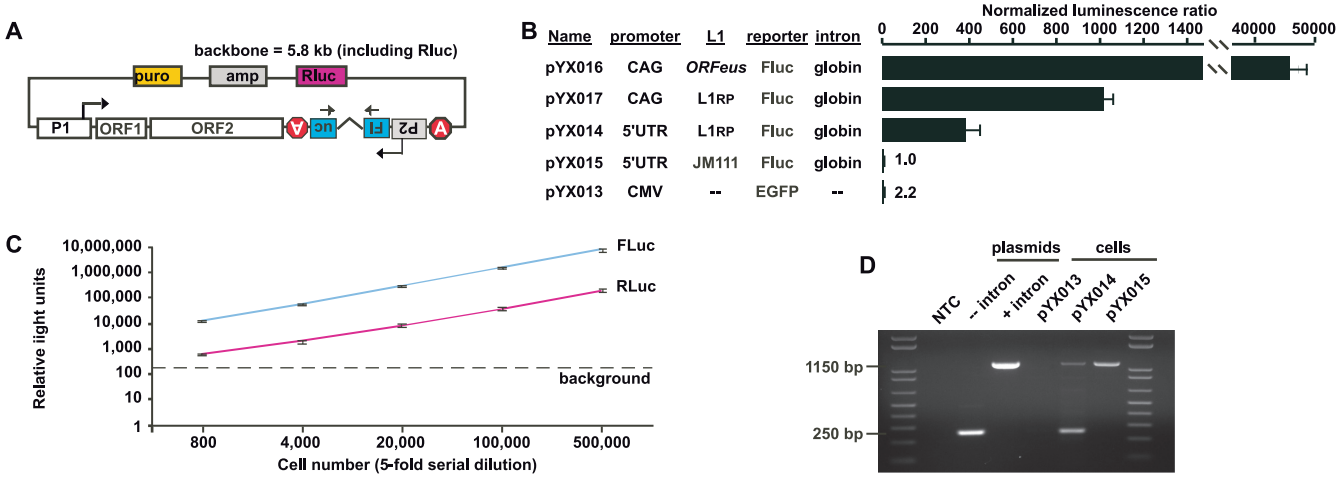

Figure 4. A single-vector dual-luciferase assay for L1 retrotransposition. (**A**) Diagram of a dual-luciferase assay vector in a single-vector format (15.1 kb for pYX014). (**B**) Evaluation of L1 vectors with different coding sequences and promoters. L1 activity is expressed as NLRs relative to the background signal from pYX015. Error bars represent mean ± SE ($N = 4$). (**C**) Assay dynamic range. HeLa cells were transfected with pYX016. Cells were counted before lysis. Cell lysates were serially diluted by 5-fold and measured for luminescence. Data represent mean ± SE ($N = 3$). The level of background signals is indicated (ranging from 100 to 300 relative lights units). (**D**) Confirmation of retrotransposition by PCR. HeLa cells were transfected by each L1 plasmid, and genomic DNA was amplified by an intron-flanking primer pair (shown as arrows in A). The presence of a band of 250 bp is diagnostic for intron removal; the intron-containing donor DNA is amplified as a band of 1150 p. NTC, no template control. Fluc plasmids with or without the intron are used as controls. Molecular weight was indicated by 1 kb Plus DNA Ladder (Invitrogen).

of L1 retrotransposition with the luciferase-based reporter assay, we first determined effective inhibitory concentrations of several NRTIs in our assay with 5′-UTR-L1$_{RP}$ as the donor element. In such tests, NRTI treatment started at 2 h before transfection and was maintained throughout the assay (Figure 5A). Consistent with previous reports, 50 μM d4T showed potent inhibition of L1 retrotransposition and reduced Fluc signals to background. Two other NRTIs, 3TC and ddI, were even more potent than d4T (Supplementary Figure S1A). In addition, no overt cytotoxicity was witnessed at these testing conditions (not shown).

We determined retrotransposition kinetics of 5′-UTR-L1$_{RP}$ by adding 50 μM 3TC at different time points before or post-transfection (Figure 5A). For each time point assayed, the drug was replenished daily until the cells were lysed at 96 h post-transfection for luminescence measurement. Relative to the −2 h control, a 9-fold increase of L1 activity was first detected at the 28-h time point (Figure 5B, $P = 0.14$). The level of L1 activity reached statistical significance at the 32-h time point (15-fold; $P < 0.05$). The overall kinetics of L1 retrotransposition by 5′-UTR-L1$_{RP}$ were similar when cells were treated with d4T (Supplementary Figure S1B). In both cases, a significant fraction of L1 retrotransposition events had occurred by 48 h post-transfection (39% for 3TC treated cells and 28% for d4T), the latest time point that cells were treated with either NRTI.

We also examined the kinetics of L1 retrotransposition in cultured cells for CAG-L1$_{RP}$ and CAG-*ORFeus* elements (Figure 5B). These two vectors had substantially higher activities than 5′-UTR-L1$_{RP}$ in the absence of NRTI treatment (Figure 4A); nevertheless, 50 μM of 3TC could effectively reduce L1 signals to below 0.1% relative to the untreated control (Supplementary

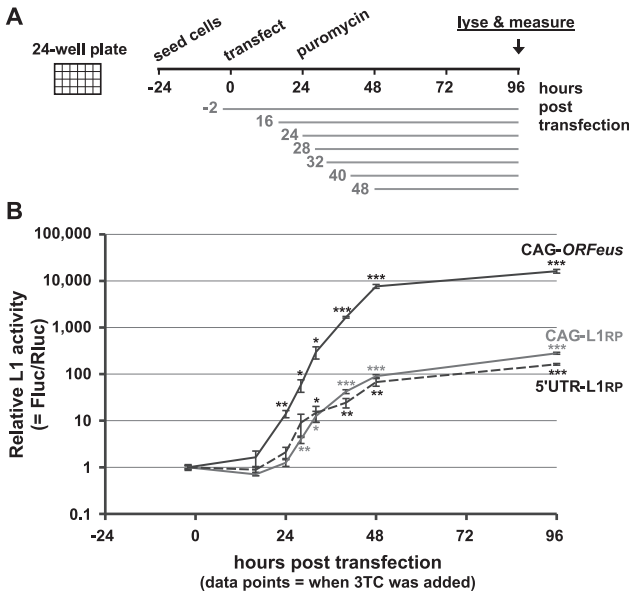

Figure 5. Kinetics of L1 retrotransposition. (**A**) Procedure for NRTI treatment. L1 assays were performed following the standard assay procedure except the NRTI treatment: HeLa cells were treated with an NRTI at −2 (i.e. 2 h before transfection), 16, 24, 28, 32, 40 or 48 h post-transfection. The duration of drug treatment is indicated by horizontal lines below the time axis. (**B**) Measurement of kinetics with 3TC. Cells were transfected with 5′-UTR-L1$_{RP}$ (pYX014), CAG-L1$_{RP}$ (pYX017), CAG-*ORFeus* (pYX016) or retrotransposition-incompetent 5′-UTR-L1$_{RP}$/JM111 (pYX015). Reverse transcription was inhibited at each time point by 50 μM 3TC. Luminescence was measured at 96 h post-transfection. L1 activity for each condition is first calculated as normalized Fluc/Rluc values, and then expressed as relative activities by setting Fluc/Rluc values from the −2 h control as 1. Note the 96-h data points represent cells not treated with 3TC. Error bars represent mean ± SE ($N = 4$). Retrotransposition at each data point was compared to the −2 h control by two-tailed Student's *t*-test (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

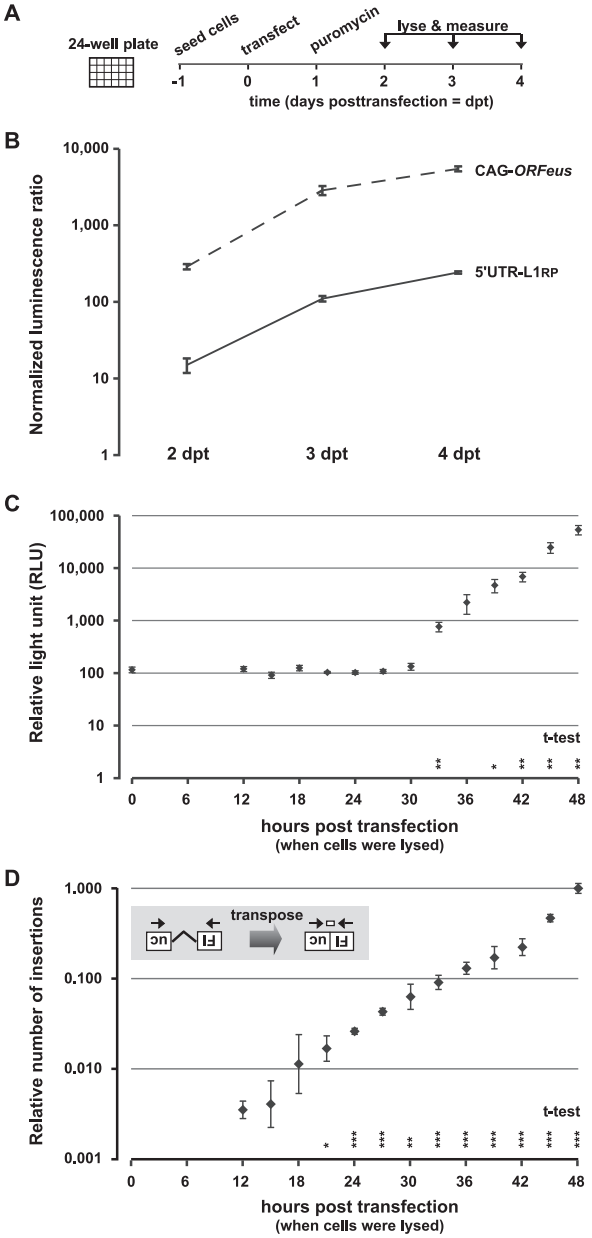

Figure 6. The time course of L1 retrotransposition. (**A**) Procedure for time course experiments. Cells were harvested at 2, 3 or 4 days post-transfection, lysed and measured for luminescence. (**B**) Detection of 5′-UTR- $L1_{RP}$ and CAG-*ORFeus* activities between 2 and 4 days post-transfection. L1 activity is expressed as NLRs relative to the background signal from pYX015. Error bars represent mean ± SE ($N = 4$). (**C**) Detection of CAG-*ORFeus* activities before 48 h post-transfection. HeLa cells were transfected with pYX016 at time zero, and lysed for luminometry analysis at indicated time points. Raw Fluc signals are plotted as RLUs for each time point. Untransfected HeLa cells were used as a control for background luminescence (the 0-h time point). Fluc signals from each time point were compared to time zero by two-tailed Student's *t*-test; statistically significant *P*-values are indicated over the *x*-axis (*$P < 0.05$, **$P < 0.01$). Error bars represent mean ± SE ($N = 4$). (**D**) Quantification of *de novo* L1 insertions by qPCR. Genomic DNA was extracted from an aliquot of cells described in C, and the relative number of L1 insertions was determined by a TaqMan-based qPCR assay. The relative positions of Fluc primers (arrows) and probe (an open rectangle between the primers) are illustrated (inset). qPCR signals were normalized by setting signals from the 48-h time point to 1. The normalized signals from each time point were then compared to the 12-h time point by two-tailed Student's *t*-test; *P*-values are indicated over the *x*-axis (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$). Error bars represent mean ± SE ($N = 4$).

Figure S1C). Both CAG-$L1_{RP}$ and CAG-*ORFeus* elements displayed highly similar kinetics as compared to 5′-UTR-$L1_{RP}$ under 3TC treatment (Figure 5B). For CAG-*ORFeus*, reproducible signals were detected as early as at the 24 h time point (14-fold increase relative to the −2 h control; $P < 0.01$). Overall, L1 retrotransposition conforms to exponential kinetics that can be divided into three phases: a lag phase encompassing the first 24 h, an exponential phase between 24 and 48 h (R^2 ranges from 0.943 to 0.984), and a steady phase after 48 h (Figure 5B).

The time-course of L1 retrotransposition in single-vector assays

In our standard protocol, the frequency of retrotransposition is determined at 4 days post-transfection. We consistently obtain robust NLR signals at 4 days post-transfection in both two-vector and one-vector assays for all donor L1 vectors tested: ~300 for 5′-UTR-$L1_{RP}$, ~1000 for CAG-$L1_{RP}$ and ~10 000 for CAG-*ORFeus*. To determine whether it would be feasible to detect L1 activities at even earlier time points, we harvested cells at 2, 3 and 4 days post-transfection in a one-vector assay (Figure 6A). Indeed, using 5′-UTR-$L1_{RP}$ as the donor element, we were able to detect significant L1 activities starting at 2 days post-transfection (NLR = 15, $P < 0.01$; Figure 6B). L1 activity increased to 110-fold above background at 3 days post-transfection and to 242-fold above background at 4 days post-transfection. Similar time course was observed for the hyperactive CAG-*ORFeus* element (Figure 6B). Not surprisingly, CAG-*ORFeus* displayed strong retrotransposition signals at all three time points. At 2 days post-transfection, its NLR has reached 288-fold above background ($P < 0.0001$).

We reasoned that the strong signal at 2 days post-transfection from CAG-*ORFeus* should permit further characterization of the time course of L1 retrotransposition within the first 48 h. Thus, we transfected CAG-*ORFeus* vector (pYX016) into HeLa cells, lysed and measured luciferase reporter activities at 3-h intervals from 12 to 48 h post-transfection. We plotted raw Fluc signals as RLUs over time (Figure 6C). No significant Fluc signals were detected in the first 30 h; thereafter, Fluc signals increased exponentially ($R^2 = 0.972$ for signals between 30 and 48 h). In fact, reproducible signals were first detected at 33 h post-transfection ($P < 0.01$). To control for well-to-well variations in cell number, we normalized the raw Fluc signals to the number of cells in each well (Supplementary Figure S2A). Again, significant signals first appeared at 33 h ($P < 0.05$), and exponentially increased thereafter ($R^2 = 0.979$ for normalized signals between 30 and 48 h).

In the previous section on retrotransposition kinetics, we have shown that a minimum of 24 h are required for CAG-*ORFeus* activity to be detected at a statistically significant level (Figure 5B). In that experiment, L1 reverse transcription was stopped at 24 h post-transfection but the cells were grown for additional 72 h in the presence of 3TC. This delay in luciferase measurement not only gave

ample time for the Fluc reporter gene to be expressed from the *de novo* L1 insertions, but might have also amplified the total signal output due to cell division. In contrast, the minimum of 33 h determined by the time-course experiment are the time required for DNA uptake, L1 retrotransposition (i.e. from transcription, translation, reverse transcription to integration), and reporter gene expression (i.e. transcription and translation of Fluc) (Figure 6C). Thus, the 9-h lag in the time-course experiment reflects the time that is needed for the reporter gene to be expressed from L1 insertions.

Reporter genes for L1 retrotransposition are prone to epigenetic silencing [(28); J. Li, M. Kannan, J. D. Boeke and D. E. Symer, personal communication]. Thus, our measurement of L1 kinetics and/or retrotransposition time course may be confounded by some unknown characteristics in reporter gene expression. To corroborate our kinetics data, we measured the number of *de novo* L1 insertions in cells harvested from above time-course experiment by a previously established qPCR method, which is totally independent of reporter gene expression. Overall, we observed an exponential increase of *de novo* L1 insertions during the assay timeframe (Figure 6D, $R^2 = 0.985$). Statistically significant signals were first detected by qPCR at 21 h post-transfection (normalized activity = 1.7%, $P < 0.05$), and a normalized activity of 9.1% was reached by 33 h post-transfection. In addition, the qPCR data displayed a good correlation with the retrotransposition data from 33 to 48 h post-transfection (Supplementary Figure S2B, $R^2 = 0.946$).

Miniaturization of single-vector assays into a 96-well format

Ninety-six-well plates or higher density format have become the standard platform for assay automation. With robust signals from 24-well plates, we evaluated the performance of the new single-vector dual-luciferase assay in a 96-well format (Figure 7). NLR signals comparable to those acquired from 24-well plates were obtained for all L1 elements tested: 193-fold above background for 5′-UTR-L1$_{RP}$, 1066-fold for CAG-L1$_{RP}$ and 26 594-fold for CAG-*ORFeus* (Figure 7; $R^2 = 0.996$ between 24- and 96-well plates).

DISCUSSION

We have developed novel dual-luciferase reporter assays for L1 retrotransposition. In contrast to existing *neo*- or EGFP-based assays, the new luciferase-based assays have several significant improvements. First, it has the shortest assay timeframe. The standard luciferase assay developed here requires only 4 days from transfection to data collection (Table 1). Our time-course results indicate that it can be shortened to 3 days from transfection for the standard 5′-UTR-L1$_{RP}$ element and to 2 days from transfection for the hyperactive CAG-*ORFeus* element, while maintaining a signal of >100-fold above background (Figure 6B). Second, the assay requires minimal hands-on time beyond transfection. Selection with puromycin is simply achieved by medium exchange; quantitative data can be rapidly collected from >96 samples within a half hour. Third, the assay displays superlative signal-to-noise ratios. The native 5′-UTR-L1$_{RP}$ element consistently yields NLR signals >100-fold above background. In contrast, the EGFP-based assay is often compromised by a high background due to autofluorescence: at 6 days post-transfection, ~2.5% of cells transfected with 5′-UTR-L1$_{RP}$ were EGFP-positive, while 0.02–0.26% of cells were considered positive from control transfections with the vector backbone (19). Fourth, the assay has a very broad dynamic range. A minimum 600-fold linear dynamic range was confirmed by serial dilution experiments. The actual dynamic range may be larger because L1 activities can be reproducibly detected in the range from NLR = 15 (at 2 days post-transfection for 5′-UTR-L1$_{RP}$; Figure 6B) to over 46 000 (at 4 days post-transfection for CAG-*ORFeus*; Figure 4A). Lastly, the luciferase-based retrotransposition assay can be performed either in a two-vector format or in a single-vector format. The single-vector assay format offers an additional advantage because the internal Rluc control is encoded on the same vector. This eliminates variations due to imprecise pipetting or quantification of absolute concentrations of different plasmids.

We investigated the kinetics of L1 retrotransposition in cell culture using three different approaches. We first used NRTIs to measure the timing of L1 retrotransposition, defined operationally as the time required for L1 reverse transcriptase to complete reverse transcription after the donor L1 is transfected into cells. In a previous study, a *neo*-based L1 reporter assay was employed and d4T was used to inhibit L1 reverse transcription at different time points post-transfection (27). Few G418-resistant colonies were observed when d4T was added at 32 h post-transfection, and ~20% of L1 retrotransposition (relative to untreated controls) had occurred when d4T was added at 42 h post-transposition (27). With its unprecedented sensitivity and dynamic range, our luciferase-based assay provided a high-resolution view of L1 kinetics, which can be summarized into three distinct phases relative to the assay timeframe (Figure 5). The first phase shows inactivity (from 0 to 24 h post-transfection) as minimal L1 signal was detected when cells were treated prior to 24 h post-transfection. The second phase lasts 24 h (from 24 to 48 h post-transfection) during which L1 activity increased exponentially. In fact, reproducible and statistically significant signals could be detected from CAG-*ORFeus* at 24 h post-transfection. The third phase spans 48 h (from 48 to 96 h post-transfection), when L1 retrotransposition returned to a slower rate and eventually reached 100% activity. These three distinct phases were observed for all three L1 donor elements when treated with 3TC (Figure 5B) as well as for 5′-UTR-L1$_{RP}$ when treated with d4T (Supplementary Figure S1B), suggesting that the overall kinetics of retrotransposition is invariant among L1 elements with different intrinsic activities. We confirmed the timing and kinetics of L1 retrotransposition by performing time-course experiments (Figure 6). The earliest time point when significant Fluc signals can be detected for CAG-*ORFeus* was 33 h post-transfection;

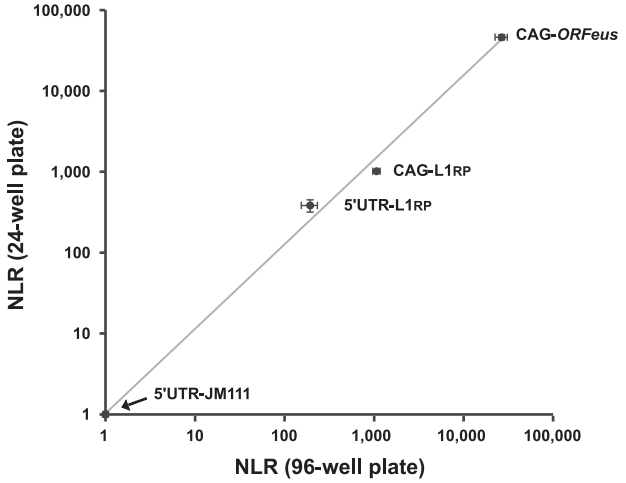

Figure 7. Miniaturization of single-vector dual-luciferase L1 assays. L1 vectors with different coding sequences and promoters, including pYX016 (CAG-*ORFeus*), pYX017 (CAG-$L1_{RP}$), pYX014 (5′-UTR-$L1_{RP}$) and pYX015 (5′-UTR-JM111), were evaluated in 24-well (Figure 4A) and 96-well plates, respectively. L1 activities are expressed as NLRs relative to the background signal from pYX015, and are graphed as a scatter plot. The *x*- and *y*-axes represent data from 96- and 24-well plates, respectively. Error bars represent mean ± SE ($N = 10$ for 96-well data; $N = 4$ for 24-well data). A power trendline is shown ($y = 1.060x^{1.042}$; $R^2 = 0.996$).

thereafter, Fluc signals increased exponentially. The 9-h delay in detecting CAG-*ORFeus* activity reflects the time that is needed for the reporter gene to be expressed from L1 insertions (see Results section). Indeed, when *de novo* L1 insertions were quantified directly by a sensitive qPCR method, significant signals can be detected as early as at 21 h post-transfection (Figure 6D). Thus, the high sensitivity and the broad dynamic range of our luciferase-based assays afford us an intimate view of kinetics of retrotransposition from plasmid-borne donor L1s.

The current study has generated a repertoire of L1 retrotransposition vectors that are equipped with dual-luciferase reporters. These vectors feature two different L1 donor elements, including a native human $L1_{RP}$ (22) and a synthetic mouse L1 *ORFeus* (23,24), and a collection of promoter elements for each donor element. *ORFeus* vectors consistently display higher retrotransposition frequencies than $L1_{RP}$ vectors as previously reported (24) (45-fold higher when both donor elements were regulated by the CAG promoter; Figure 4B). In fact, CAG-*ORFeus* is the most active retrotransposition vector in our dual-luciferase reporter assay: it yields an NLR signal of >100 even at 2 days post-transfection (Figure 6B), and is thus well suited for applications that demand short assay timeframes. We also compared a series of $L1_{RP}$ vectors that are regulated by single or tandem combinations of different promoters including the native human L1 promoter 5′-UTR and heterologous promoters CMV and CAG (Figure 3B). Overall retrotransposition frequencies vary by 20-fold among these vectors in the following order: CAG > CMV+CAG > CMV+5′-UTR > 5′-UTR > CMV. It should be noted here that the promoter strength is not necessarily the only determinant of the overall retrotransposition output. For example, the 5′-UTR of human L1.3 harbors cryptic splice sites that may affect the processing of full-length L1 transcripts (29). However, the potential contribution from such alternative splicing events has not been evaluated in our assay. In addition, previous studies using the first-generation *neo*-based assay suggest that the overall effect of different promoters on retrotransposition is context dependent. When the native human L1.2 element was tested, the exact order of CMV+5′-UTR > 5′-UTR > CMV was observed (16). In contrast, the 5′-UTR promoter from two other human L1 elements, L1.3 and $L1_{RP}$, yielded higher frequencies of retrotransposition than the tandem CMV+5′-UTR promoter (22). Thus, to facilitate cross-study comparison of retrotransposition results, we recommend the use of singular promoters. As a general guideline, CAG is the promoter of choice if the highest activity is desired, CMV is an attractive alternative as it has a much smaller size (0.6 kb for CMV versus 1.7 kb for CAG), and finally the use of 5′-UTR is required if the regulation of native L1 promoter is the focus of the study.

In summary, the newly established luciferase-based assays for L1 retrotransposition are rapid, sensitive and highly efficient. We have shown that high-throughput testing of L1 retrotransposition can be performed in 24-well or 96-well plates. High signal-to-noise ratios were obtained from 96-well plates even with standard 5′-UTR-$L1_{RP}$ elements, suggesting it has the potential to be further miniaturized into 384-well or higher density formats. Since the initial demonstration of retrotransposition by a cloned L1 in cultured cells, retrotransposition assays have been playing a pivotal role in unraveling mechanistic aspects of L1 movement. Aided by its short assay timeframe, the luciferase-based assay is expected to facilitate routine tests for L1 function as well as large-scale screening efforts, such as genome-wide screening of host factors or compounds, which modulate L1 retrotransposition.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Dr Haig Kazazian for providing $L1_{RP}$ and $L1_{RP}$/JM111 plasmids and Dr Ronald Brosemer for critical reading of the article.

FUNDING

Start-up funds from Washington State University (to W.A.); National Cancer Institutes (P01 CA16519 to J.D.B.). Funding for open access charge: Start-up funds from the Washington State University.

Conflict of interest statement. None declared.

REFERENCES

1. 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.
2. Kazazian,H.H. Jr (2004) Mobile elements: drivers of genome evolution. *Science*, **303**, 1626–1632.
3. Cordaux,R. and Batzer,M.A. (2009) The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.*, **10**, 691–703.
4. Esnault,C., Maestre,J. and Heidmann,T. (2000) Human LINE retrotransposons generate processed pseudogenes. *Nat. Genet.*, **24**, 363–367.
5. Wei,W., Gilbert,N., Ooi,S.L., Lawler,J.F., Ostertag,E.M., Kazazian,H.H., Boeke,J.D. and Moran,J.V. (2001) Human L1 retrotransposition: cis preference versus trans complementation. *Mol. Cell Biol.*, **21**, 1429–1439.
6. Dewannieux,M., Esnault,C. and Heidmann,T. (2003) LINE-mediated retrotransposition of marked Alu sequences. *Nat. Genet.*, **35**, 41–48.
7. Ostertag,E.M., Goodier,J.L., Zhang,Y. and Kazazian,H.H. Jr (2003) SVA elements are nonautonomous retrotransposons that cause disease in humans. *Am. J. Hum. Genet.*, **73**, 1444–1451.
8. 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.
9. 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.
10. Chen,J.M., Stenson,P.D., Cooper,D.N. and Ferec,C. (2005) A systematic analysis of LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease. *Hum. Genet.*, **117**, 411–427.
11. Brouha,B., Schustak,J., Badge,R.M., Lutz-Prigge,S., Farley,A.H., Moran,J.V. and Kazazian,H.H. Jr (2003) Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl Acad. Sci. USA*, **100**, 5280–5285.
12. 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.
13. Huang,C.R., Schneider,A.M., Lu,Y., Niranjan,T., Shen,P., Robinson,M.A., Steranka,J.P., Valle,D., Civin,C.I., Wang,T. *et al.* (2010) Mobile interspersed repeats are major structural variants in the human genome. *Cell*, **141**, 1171–1182.
14. Ewing,A.D. and Kazazian,H.H. Jr (2010) High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome Res.*, **20**, 1262–1270.
15. Iskow,R.C., McCabe,M.T., Mills,R.E., Torene,S., Pittard,W.S., 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.
16. Moran,J.V., Holmes,S.E., Naas,T.P., DeBerardinis,R.J., Boeke,J.D. and Kazazian,H.H. Jr (1996) High frequency retrotransposition in cultured mammalian cells. *Cell*, **87**, 917–927.
17. Rangwala,S.H. and Kazazian,H.H. Jr (2009) The L1 retrotransposition assay: a retrospective and toolkit. *Methods*, **49**, 219–226.
18. Wei,W., Morrish,T.A., Alisch,R.S. and Moran,J.V. (2000) A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events. *Anal. Biochem.*, **284**, 435–438.
19. Ostertag,E.M., Prak,E.T., DeBerardinis,R.J., Moran,J.V. and Kazazian,H.H. Jr (2000) Determination of L1 retrotransposition kinetics in cultured cells. *Nucleic Acids Res.*, **28**, 1418–1423.
20. An,W., Davis,E.S., Thompson,T.L., O'Donnell,K.A., Lee,C.Y. and Boeke,J.D. (2009) Plug and play modular strategies for synthetic retrotransposons. *Methods*, **49**, 227–235.
21. Rosser,J.M. and An,W. (2010) Repeat-induced gene silencing of L1 transgenes is correlated with differential promoter methylation. *Gene*, **456**, 15–23.
22. Kimberland,M.L., Divoky,V., Prchal,J., Schwahn,U., Berger,W. and Kazazian,H.H. Jr (1999) Full-length human L1 insertions retain the capacity for high frequency retrotransposition in cultured cells. *Hum. Mol. Genet.*, **8**, 1557–1560.
23. An,W., Han,J.S., Wheelan,S.J., Davis,E.S., Coombes,C.E., Ye,P., Triplett,C. and Boeke,J.D. (2006) Active retrotransposition by a synthetic L1 element in mice. *Proc. Natl Acad. Sci. USA*, **103**, 18662–18667.
24. Han,J.S. and Boeke,J.D. (2004) A highly active synthetic mammalian retrotransposon. *Nature*, **429**, 314–318.
25. Grimaldi,G., Skowronski,J. and Singer,M.F. (1984) Defining the beginning and end of KpnI family segments. *EMBO J.*, **3**, 1753–1759.
26. Szak,S.T., Pickeral,O.K., Makalowski,W., Boguski,M.S., Landsman,D. and Boeke,J.D. (2002) Molecular archeology of L1 insertions in the human genome. *Genome Biol.*, **3**, research0052.
27. Kroutter,E.N., Belancio,V.P., Wagstaff,B.J. and Roy-Engel,A.M. (2009) The RNA polymerase dictates ORF1 requirement and timing of LINE and SINE retrotransposition. *PLoS Genet.*, **5**, e1000458.
28. Garcia-Perez,J.L., Morell,M., Scheys,J.O., Kulpa,D.A., Morell,S., Carter,C.C., Hammer,G.D., Collins,K.L., O'Shea,K.S., Menendez,P. *et al.* (2010) Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. *Nature*, **466**, 769–773.
29. Belancio,V.P., Hedges,D.J. and Deininger,P. (2006) LINE-1 RNA splicing and influences on mammalian gene expression. *Nucleic Acids Res.*, **34**, 1512–1521.