Chapter 12

Characterization of Engineered L1 Retrotransposition Events: The Recovery Method

David Cano, Santiago Morell, Andres J. Pulgarin, Suyapa Amador, and Jose L. Garcia-Pérez

Abstract

Long Interspersed Element class 1 retrotransposons (LINE-1 or L1) are abundant Transposable Elements in mammalian genomes and their mobility continues to impact the human genome. The development of engineered retrotransposition assays has been instrumental to understand how these elements are regulated and to identify domains involved in the process of retrotransposition. Additionally, the modification of a retrotransposition indicator cassette has allowed developing straightforward approaches to characterize the site of new L1 insertions in cultured cells. In this chapter, we describe a method termed "L1-recovery" that has been used to characterize the site of insertion on engineered L1 retrotransposition events in cultured mammalian cells. Notably, the recovery assay is based on a genetic strategy and avoids the use of PCR and thus reduces to a minimum the appearance of false positives/artifacts.

Key words LINE-1, Retrotransposon, Recovery, Engineered, Insertion, Deletion, Target site duplication

1 Introduction

Most mammalian genomes are characterized for the high prevalence of repeated DNA sequences. Among repeated DNA sequences, Transposable Elements (TEs) are repeated DNA sequences that can move within genomes (reviewed in Refs. [1–3]). TEs are very diverse in their structure and abundance depending on the genome that is examined [3]. In humans, between 45 and 70 % of the human genome is made of TEs [4, 5], and active TEs continue to impact our genome [1–3]. Long Interspersed Element class 1 retrotransposons (LINE-1 or L1) are very prevalent sequences in the human genome and up to 21 % of our genome is made of LINE derived sequences [5]. Although most L1 copies are inactive fossils accumulated during human genome evolution, an average human genome contains between 80 and 100 active

Jose L. Garcia-Pérez (ed.), *Transposons and Retrotransposons: Methods and Protocols*, Methods in Molecular Biology, vol. 1400, DOI 10.1007/978-1-4939-3372-3_12,

L1s (termed Retrotransposition Competent L1s or RC-L1s) [6, 7]. RC-L1s are retrotransposons that move using a copy and paste mechanism termed Target Primed Reverse Transcription (TPRT) [1–3, 8]. Active RC-L1s are 6-kb in length elements containing a 900-bp long 5′ Untranslated Region (UTR) that contains conserved sense and antisense promoter activities [9–11], two non-overlapping Open Reading Frames (ORF1 and ORF2) and end in a short 3′ UTR region with a weak polyadenylation sequence [12–15]. L1-ORF1p codes for a 40 kDa protein with RNA binding and nucleic acid chaperone activity [16–19]. L1-ORF2p codes for a 150 kDa protein with demonstrated Endonuclease (EN) and Reverse Transcriptase (RT) activities [20–22]. Both ORFs are strictly required for the mobilization of RC-L1, as demonstrated using an engineered L1 retrotransposition assay in cultured cells [14]. Retrotransposition starts with the generation of a full-length polyadenylated L1 mRNA from an active L1 located elsewhere in the genome. The L1 mRNA is exported to the cytoplasm where L1-ORF1p translation takes place by a cap-dependent mechanism [23]. Notably, the L1 mRNA is unusual because is bi-cistronic and studies in cultured cells have revealed that L1-ORF2p is translated by an unconventional termination/reinitiation mechanism [24]. L1-ORF2p translation is very inefficient when compared to L1-ORF1p translation, and although translation of the L1 mRNA might produce hundreds of L1-ORF1p molecules, as little as one L1-ORF2p molecule might be translated from the same L1 mRNA [1–3, 24]. Remarkably, both proteins seem to bind back strongly to the same L1 mRNA used as a template for translation, a concept termed *cis*-preference [25, 26]. The L1 mRNA and both encoded proteins form a high weight molecular complex termed L1 ribonucleoprotein particle (L1-RNPs) that is supposed to be a retrotransposition intermediate [27]. L1-RNPs can be visualized in cultured cells using epitope tags, and several studies have revealed that L1-RNPs are often located in dense cytoplasmic foci [27–29]. During retrotransposition, L1-RNPs enter the nucleus where, using the intermediate L1 mRNA as a template, a new insertion is generated in a different genomic location, by TPRT. During TPRT, the EN activity of L1-ORF2p is thought to mediate a single strand break on DNA, releasing a free 3′ OH group that is then used by the RT activity of L1-ORF2p as a primer to generate the first cDNA copy of the L1 mRNA attached to the genome of the cell [8]. Second strand cDNA synthesis is thought to occur using a similar mechanism, resulting in the insertion of a new L1 copy in a different genomic place. Because of the L1-EN cleaving mechanism, most de novo L1 insertions are flanked by 2–20 bp Target Site Duplication (TSD) sequences. However, L1 insertions can also generate small deletions or no duplication of sequence (i.e., blunt insertion), depending on how L1-EN cleaves both strands at the insertion site [30]. Notably, most de novo L1 insertions are 5′

truncated, by an ill-defined mechanism [31]; additionally, some L1 insertions generated inverted/deleted structures by a mechanism termed Twin Priming [32]. These characteristics associated with de novo L1 retrotransposition events are bona fide hallmarks often used to characterize a new L1 insertion in a given genome. Because of their repeated nature and their high prevalence in genomes, identifying new L1 insertions in genomes is a very complicated task equivalent to identify a needle in a haystack.

However, the mechanism of L1 retrotransposition by TPRT allowed the developing of a genetic based L1 retrotransposition assay that is based on the activation of a reporter gene (REP) only after a round of L1-retrotransposition in transfected cultured cells (Fig. 1) [14, 33].

Briefly, in 1996 Moran and colleagues exploit the existence of an intermediate L1 mRNA during L1 retrotransposition to design a reporter gene using an engineered intron with a configuration that allowed expressing a functional reporter product only after a round of bona fide retrotransposition ([14], a method in this book and reviewed in Refs. [34, 35]). Briefly, the modified reporter cassette consists of the ORF that codes for a given reporter gene equipped with an exogenous promoter and polyadenylation sequences cloned in an antisense configuration in the 3′ UTR of an RC-L1. Additionally, the reporter ORF is interrupted by an intron cloned in the same transcriptional orientation as the RC-L1 (Fig. 1). With this configuration, only mRNAs generated from the L1 promoter can remove the intron by *cis*-splicing and the resulting chimeric L1 mRNA go trough a round of L1 retrotransposition, resulting in the insertion of a reporter gene lacking the intron and thus allowing efficient expression of the reporter.

The retrotransposition assay has been instrumental to increase our knowledge of L1 biology. The original assay used the sequence of the neomycin phosphotransferase gene as a reporter (*mneoI* cassette [14, 33]) and cells harboring a new retrotransposition event could be selected in culture using the mammalian antibiotic neomycin or G418 (Fig. 1). Since then, a number of reporter ORFs have been used to developed different retrotransposition markers (also covered in this Book and reviewed extensively elsewhere). Notably, using the *mneoI* retrotransposition indicator cassette, new L1 insertions in the genome of cultured cells are tag with a unique spliced *mneoI* sequence, which further allowed to characterize L1 insertions using conventional library construction/screening or inverse PCR methods [14, 26, 36]. Indeed, the characterization of de novo engineered L1 insertions using this assay in cultured cells has allowed to demonstrated how L1 can delete genomic DNA upon insertion, can mediate major alterations at the insertion site, can retrotranspose in mammalian neuronal cells, how L1 can insert by a new mechanism of insertion independent of its endonuclease, etc. [13, 30, 37–42].

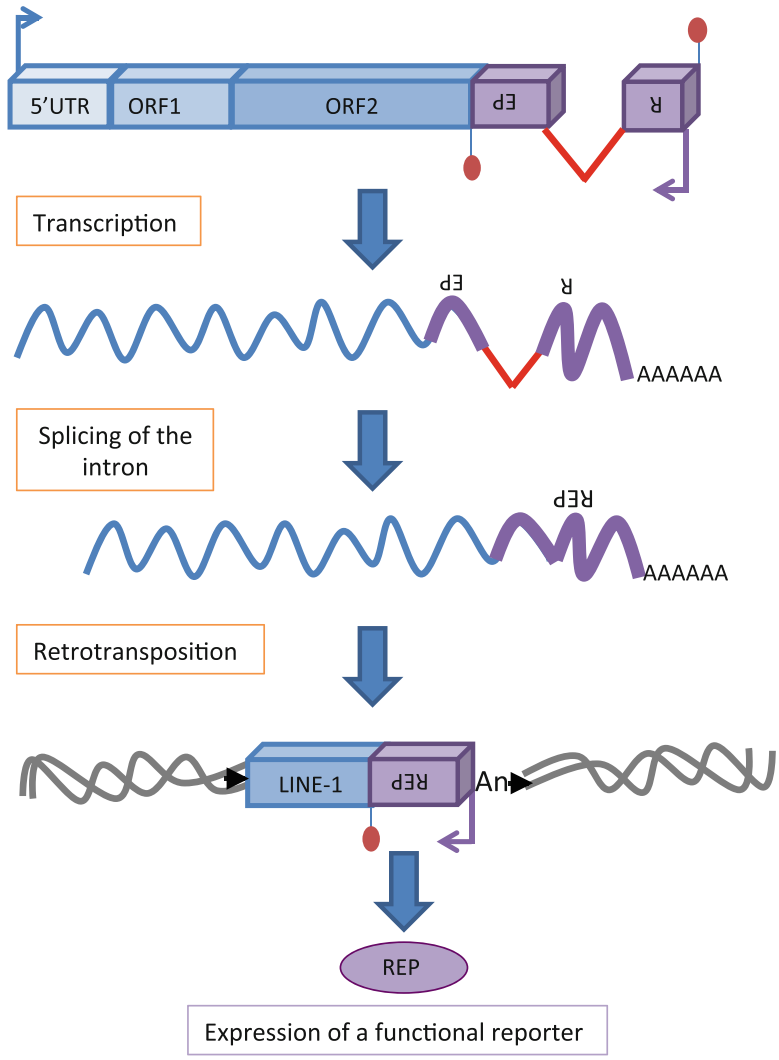

Fig. 1 Rationale of the L1-retrotransposition assay. Further details are provided in the text

In 2002, Gilbert and colleagues developed a modified *mneoI-based* retrotransposition indicator cassette that allows the recovery of engineered L1 insertions as autonomously replicating plasmids in bacteria [39] (Fig. 2).

Briefly, as the codifying sequence NEO confers G418 resistance in mammalian cells and kanamycin resistance in bacteria, Gilbert and colleagues modified the *mneoI* retrotransposition indicator cassette and included a prokaryotic promoter (EM7) and a Shine–Dalgarno sequence upstream the initiating AUG codon of NEO. Additionally, Gilbert and colleagues added a bacterial origin of replication (ColE1) downstream of the modified *mneoI* cassette and upstream of the L1 polyadenylation sequence. With this

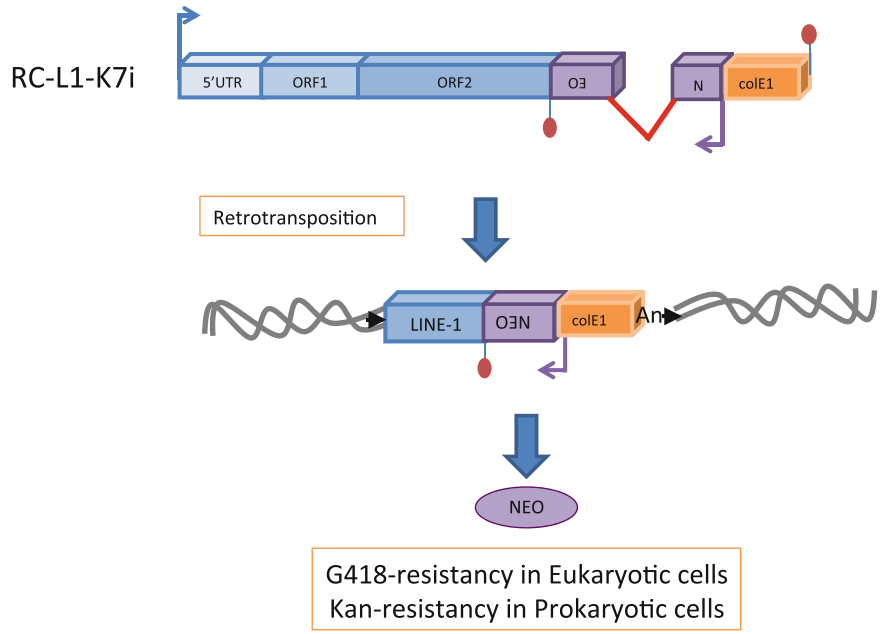

Fig. 2 Structure of the engineered L1-recovery plasmid. Shown is a scheme of a RC-L1 containing a recovery retrotransposition indicator cassette (termed K7i)

configuration, a new retrotransposition event will deliver into the genome of cells a spliced modified *mneoI* and a colE1 sequence, which is the basis of the generation of a prokaryotic plasmid (Fig. 2).

Thus, cells are first transfected with a plasmid containing an RC-L1 tagged with the recovery indicator cassette and cells containing a new retrotransposition evens can be selected with G418. Next, foci are expanded (either as single colonies or pools) and cells allowed to grow to obtain enough cells to isolate at least 20 μg of genomic DNA (gDNA). The gDNA of a colony is next digested with a restriction enzyme that cleaves frequently in the genome of the host cells (for human cells *Ssp*I, *Hind*III, *Bgl*II, etc.) and that ideally does not cut within the sequence of the transfected plasmid (or if it does, an enzyme that cleaves as closer as possible to nucleotide 1 of the transfected L1). The digested DNA is next ligated in very diluted conditions to favor intramolecular ligation events, resulting in the generation of a prokaryotic plasmid structure (Fig. 3).

This mix of ligated DNAs is next transformed to ultracompetent *E. coli* cells, and after recovery of transformed cells, the engineered L1 insertion plasmid is selected using kanamycin in bacteria (Fig. 3). These colonies will have replicated the engineered retrotransposed product plus flanking genomic DNA as a plasmid that can be extracted and purified using conventional miniprep purification of DNA. Finally, these plasmids can be sequenced

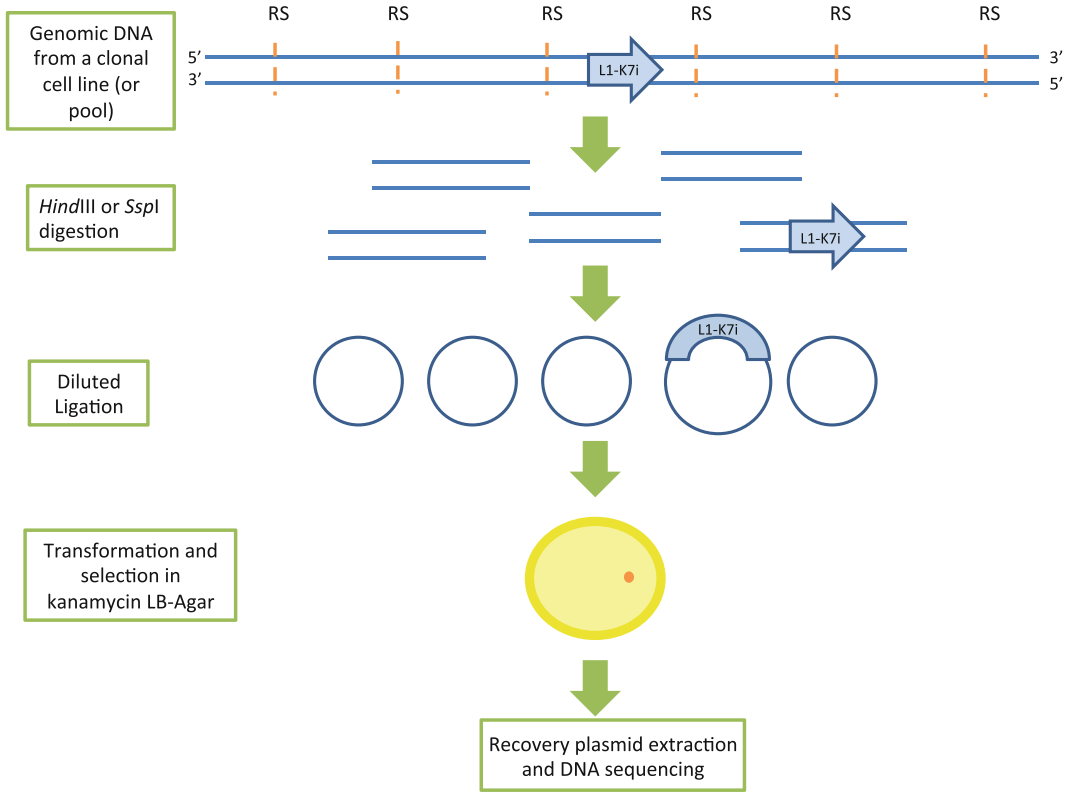

Fig. 3 Rationale of the recovery assay. Details are included in the text. R.S., restriction site

using Sanger methodology and the characteristics associated with new L1 retrotransposition events inferred at the nucleotide level (Fig. 3 and *see* [39]). In this chapter, we cover the recovery process, which is essentially the assay described by Gilbert and colleagues in 2002 with minor modifications that increase the efficiency of the recovery process. Notably, alternative methods based on similar methodologies have been also described [43], although they are not covered in this chapter.

2 Materials

2.1 Retrotransposition Assay and Establishment of Cell Lines

1. Cell culture medium for the cell line to be used (*see* **Note 1**).
2. 10 cm cell culture dishes.
3. Appropriate transfection reagent (Fugene 6 (Promega), Xtreme (Roche), Lipofectamine (Invitrogen), etc.; *see* **Note 1**).
4. 50 mg/ml geneticin (G418).
5. Trypsin 0.05 or 0.25 % (*see* **Note 1**).

6. 70 % ethanol.
7. 0.2 × 0.2 cm 3MM sterilized chromatography paper (Whatman).
8. Sterile small forceps.
9. 24-well tissue culture plates.
10. T75 tissue culture flasks.
11. Cryopreservation tubes and cryopreservation media (*see* **Note 1**).
12. 1× PBS.
13. An engineered L1 construct containing the recovery retrotransposition indicator cassette (as published by the Moran lab) [39–41]. The plasmid might be purified from bacteria using a Midi or Maxiprep purification kit (from local providers). The plasmid used for cell transfection should be in a supercoiled form (80 % or higher).

2.2 Genomic DNA Extraction

1. Blood and Cell Culture DNA midi kit (Qiagen) or a similar kit that provided high quality genomic DNA.
2. Cell lysis buffer: 10 mM Tris–HCl, pH 8.2, 10 mM EDTA, 200 mM NaCl, 0.5 % (w/v) SDS. To prepare 100 ml of lysis buffer: 0.12114 g Tris–HCl; 0.29225 g EDTA; 1.17 g NaCl; 0.5 g SDS (powder) Adjust pH with HCl.
3. 100 % ethanol.
4. 20 mg/ml proteinase K.
5. 25:24:1 phenol–chloroform–isoamyl alcohol.
6. TE buffer, pH 8.
7. Sterilized 1.7 ml microfuge tubes.
8. 70 % ethanol.

2.3 Recue of Retrotransposition Events

1. *Hind* III and *Ssp*I enzymes (New England Biolabs).
2. 10× Buffer 2 and *Ssp*I 10× buffer (New England Biolabs).
3. 100× BSA (New England Biolabs).
4. T4 DNA ligase (New England Biolabs).
5. 10× T4 DNA ligase buffer (New England Biolabs) (*see* **Note 2**).
6. Ultra 0.5 ml Centrifugal Filters Ultracel—100 K (Amicon).
7. Ultracompetent XL1-gold or XL1 Blue bacteria (*see* **Note 3**).
8. 15 ml Falcon tubes.
9. Liquid LB culture medium.
10. 140 mm bacterial plates.
11. 25 mg/ml kanamycin.

12. Wizard Plus SV Miniprep DNA purification system (Promega) or a similar kit from local suppliers.
13. *Eco*RI (New England Biolabs).
14. 10× *Eco*RI buffer (New England Biolabs).
15. 5 ml sterilized bacteria culture tubes.

2.4 Sequencing

1. Primers to sequence recovery plasmids (*see* Table 1).

Table 1
Sequence of primers commonly used to characterize L1 insertions using the recovery method

NEO210as	gaccgcttcctcgtgctttacg
Reco3	ctaaagtatatatgagtaacc
ORF2Jas	ggtgcgctgcacccactaatg
ORF2Ias	cacattttcttaatccagtc
ORF2Has	ctaactggtgtgagatgatatc
ORF2Gas	gaattgatttttgtataaggtg
ORF2Fas	cattttcacgatattgattcttcc
ORF2Eas	gtcccatcaatacctaatttattg
ORF2Das	gaattcggctgtgaatccatctgg
ORF2Cas	ttaattgtgatgttagggtgtc
ORF2Bas	gatttggggtggagagttctg
ORF2Aas	gccttctttgtctcttttgat
ORF1Cas	ctttccatgtttagcgcttcc
ORF1Bas	ggtcttttcacatagtcccat
ORF1Aas	cgttccttggaggaggagaggc
5′ UTRas	gcaggcaggcctccttgagctg

3 Methods

General note: standard good practices for handling cells should be used, including the use of a certified laminar flow biosafety hood and sterile materials and techniques for any reagent used. Note that mammalian cells are very sensitive to DNA transfection when infected by *Mycoplasma* spp. We routinely test for *Mycoplasma* spp. once a month. We also verified the nature of cell lines using short tandem repeat (STR) analyses to ensure cross-contamination with other cell lines.

3.1 Retrotransposition Assay

1. Trypsinize, count, and seed an appropriate number of cultured cells in a sterile 10 cm cell culture plate (*see* **Notes 1** and **4**).
2. Transfect cells with the L1 recovery plasmid (*see* **Notes 1** and **5**) 12–14 h after seeding.
3. After 12 h, change media in the 10 cm plates.
4. 72 h post-transfection, replace medium and add medium supplemented with G418 (*see* **Note 6**).
5. Allow G418 to select G418-resistant foci; the process take a minimum of 10 days but can change between cell lines (*see* **Notes 1** and **6**). Change the media every other day. *See* also **Note 7**.

3.2 Establishment of G418-Resistant Cell Lines (See Note 8)

1. Take 10–20 small sterilized 0.2 × 0.2 cm 3MM paper squares and soak them in Trypsin using a sterile 10 cm tissue culture plate (*see* **Note 1**).
2. Sterilize forceps by immersion in 70 % ethanol. Air-dry forceps.
3. Aided with the sterilized forceps, put a trypsin soaked 0.2 × 0.2 cm 3MM square paper on top of a colony. Incubate at RT for 1–3 min.
4. Using the forceps, put the piece of 3MM paper in a well of a 24-well plate containing 1 ml of complete media. Shake the plate right/left and up/down 2–3 times and incubate cells overnight in an incubator (*see* **Note 9**).
5. Repeat **steps 2–4** as many times as required.
6. Next day, remove papers using a vacuum sucker device and add 1 ml of complete cell culture medium to each well.
7. Expand G-418 foci following standard procedures until obtaining a T-75 full flask of cells (95 % confluent).
8. From the T-75 flask, trypsinize cells and freeze at least one vial of cells using an appropriate method for the cell line of choice (*see* **Note 1**). Collect the remaining cells using a Falcon 15 ml tube by centrifugation at 1500 × *g* for 5 min at 4 °C. Remove supernatant and store pellets at −80 °C.

3.3 Genomic DNA Isolation for the Rescue of Engineered LINE-1 Retrotransposition Events (See Note 10)

1. Extract genomic DNA from cell lines (*see* **Note 11**). The starting amount of cells required to obtain enough genomic DNA depends on the method used to isolate DNA (*see* **Note 12**).
2. If using a commercial DNA extraction kit, follow the manufacturer's instructions. Jump to **step 4**.
3. If using the homemade protocol, follow these steps:
   (a) Thaw frozen pellets on ice.
   (b) Resuspend the cell pellet in 0.5 ml lysis buffer.
   (c) Add 5 μl of 20 mg/ml proteinase k (*see* **Note 13**).

(d) Incubate for 3 h at 56 °C. If necessary, the digestion can be done overnight at 56 °C (*see* **Note 13**).

(e) Transfer the lysates into clean 1.7 ml tubes or 2 ml tubes depending on the volume of your lysate.

(f) Add 1 volume of phenol–chloroform–isoamyl alcohol and shake tubes vigorously.

(g) Centrifuge for 5 min at max speed at room temperature (RT) using a microfuge.

(h) Keep the aqueous phase (top part) and transfer it to a new 1.7 ml tube (*see* **Note 14**).

(i) Add 2 volumes of 100 % ethanol and place tubes at −20 °C for at least 3–4 h or overnight.

(j) Centrifuge tubes at max speed at 4 °C for 15 min. Discard the supernatant.

(k) Wash the pellet with 500 μl of 70 % ethanol and centrifuge at max speed for 10 min at 4 °C. Discard supernatant.

(l) Air-dry pellets (*see* **Note 15**) but do not over-dry.

(m) Add 200 μl of TE buffer, pH 8.0 and dissolve the pellet. If the pellet is resistant to solution, heat the tube at 56 °C for 1 h with strong agitation. If necessary, incubate it longer.

4. Quantify the concentration of your genomic DNA sample. We recommend using an spectrophotometer as measurements are more accurate that when using a NanoDrop or similar devices.

3.4 The Recovery Protocol (See Note 10)

Below, we described the recovery process divided in days.

Day 1

1. Genomic DNA digestion. Digest 8 μg of genomic DNA with either *Hind*III or *Ssp*I (*see* **Note 16**) at 37 °C overnight as follows (*see* **Note 17**):
   - 5 μl of *Hind*III or *Ssp*I.
   - 10 μl of buffer number 2 or 10 μl of *Ssp*I buffer.
   - 1 μl of BSA 100×.
   - 8 μg of genomic DNA.
   - Complete with DNA-free ddH_2O up to 100 μl (*see* **Notes 10** and **18**).

Day 2

1. In the morning, add 1 μl of either *Hind*III or *Ssp*I and incubate at 37 °C for 2 h. Heat at 65 °C to inactive *Hind*III or *Ssp*I enzymes for 25 min.
2. Set up ligation. Add the following reagents to the 100 μl digested DNA sample:

- 8 µl of T4 DNA ligase.
- 40 µl of 10× T4 DNA ligase buffer.
- 351 µl of ddH_2O (*see* **Notes 10** and **18**).

3. Incubate overnight at 16 °C for *Hind*III digested DNAs and at 25 °C for *Ssp*I digested DNAs (*see* **Note 19**).

Day 3

1. In the morning, add 1 µl of T4 DNA ligase and incubate tubes for at least 4 h at RT.
2. Concentrate DNAs as follows:
   - Transfer each ligation into a clean Amicon.
   - Centrifuge at 5000×*g* for 5 min at RT. Discard the flow-through.
   - Add 450 µl of DNA-free ddH_2O (*see* **Notes 10** and **18**).
   - Centrifuge at 5000×*g* for 5 min at RT. Note that a small volume of liquid will remain in the Amicon. Do not discard it.
   - Place the Amicon upside down and put it inside a new 1.7 ml tube (*see* **Note 20**).
   - Short spin. The volume of the concentrated DNA should be around 50–60 µl. If larger, recentrifuge at 5000×*g* for 5 min at RT.

3 Bacteria transformation.

- Slowly thaw on ice a 500 µl aliquot of XL1 Blue ultracompetent cells (*see* **Note 21**).
- Label 15 ml tubes and cool them on ice. Also, prewarm an aliquot of sterile LB medium at 37 °C (*see* **Note 22**).
- Add 1/3 of the concentrated DNA solution (15–20 µl) to the thawed bacteria. Transfer the mix slowly to a precooled 15 ml tube.
- Incubate on ice for 25–30 min.
- Heat shock at 42 °C for 38 s. CRITICAL: Do not exceed this time.
- Incubate on ice for 2 min.
- Add 1 ml of prewarmed LB (37 °C). Incubate the mix overnight at RT on a shaker but do not exceed 600 rpm (*see* **Note 23** and *see* Fig. 4).

Day 4

1. Pellet bacteria by centrifugation at 500×*g* for 8 min at RT.
2. Remove supernatant from tubes but keep around 200–300 µl.
3. Resuspend the bacterial suspension carefully with a pipette.

Fig. 4 A system to shake transformed bacteria. See text for further details

4. Plate the suspension of bacteria in 140 mm LB-Agar plates containing 25 μg/ml kanamycin. Incubate at 37 °C overnight (*see* **Note 24**).

Day 5

1. After 24 h, colonies should be already visible. In the afternoon, pick individual colonies using an sterilized tip and inoculate a tube containing 2–4 ml of LB with 25 μg/ml kanamycin. Plates with colonies can be also stored at 4 °C for a week.
2. Incubate tubes overnight at 37 °C on a shaker.

Day 6

1. Extract plasmid DNA from bacteria cultures using a miniprep kit. Follow manufacturer's instructions.
2. Digest a 5 μl aliquot of each recovered plasmid with *Eco*RI using a standard protocol (*see* **Note 25**).
3. Resolve digestions on 1–1.5 % agarose gels and identify particular restriction patterns (Fig. 5 and **Note 25**).

3.5 Sequencing Recovered Plasmids

1. Select the plasmids with a unique restriction pattern and sequence them with oligonucleotides Neo 210AS and Reco3 (*see* Table 1). *See* **Note 26**.

4 Notes

1. The recovery assay can be used in any cell line that support engineered L1 retrotransposition. The media, the transfection conditions, freezing conditions and other general methods must be optimized for each cell line.
2. To avoid degradation of ATP in cycles of thawing/freezing, aliquot T4DNA ligase 10× buffer in 100 μl aliquots.

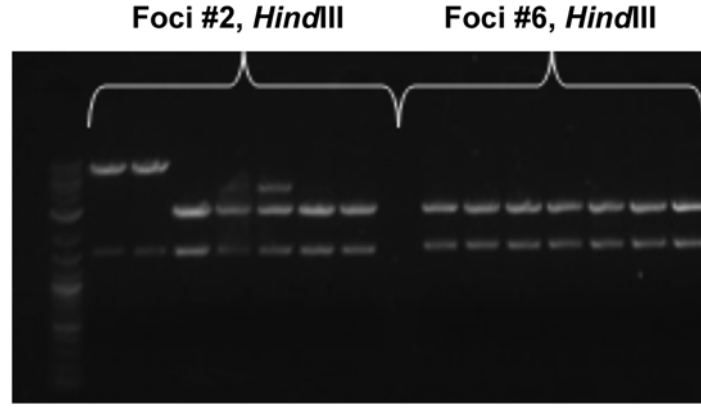

Fig. 5 Representative results from the recovery assay. Shown are results of plasmids isolated from two G418-resistant foci and digested with *Eco*RI. In cell line #2, two patterns can be observed which upon sequencing revealed that this foci contained at least two engineered L1 insertions

3. XL1-gold or XL1-blue ultracompetent bacteria can be prepared in house or purchased from local vendors. It is critical that bacteria are ultracompetent.
4. In our hands, we have been able to recover insertions from any cellular type than can support engineered L1 retrotransposition. The number of cells that need to be plated for each cell line might be determined experimentally. The goal of the assay is to have enough colonies dispersed in the 10 cm plate after G418 selection. For HeLa cells, we routinely seed $4–8 \times 10^4$ cells/10 cm plate.
5. Similarly, the transfection method or reagent that can be used depends on the cellular type used. We routinely use transfection reagents such as Fugene 6 or Xtreme 9, as both reagents allows transfecting cells plated at a low cellular density without elevated toxicity. For each transfection reagent, follow manufacturer's instructions.
6. The concentration of G418 used for selection should be empirically determined for each cell line. For HeLa cells, we routinely use 400 μg/ml.
7. It is informative to use an internal control in the retrotransposition assay, by using alleles of the engineered L1 plasmids containing missense mutations in domains involved in retrotransposition (EN and/or RT missense mutants).
8. To avoid cross-contamination in the establishment of G418 resistant cell lines, G418-foci should be distinct in plates, separated by at least 0.5 cm from nearby foci.
9. The well size depends on the ability of cells to grow at a low cellular density. Some cell lines can grow very well starting from few cells in a large culture dish while others are slower/

incapable of generating foci. Thus, the well size will vary among cell lines and it must be empirically determined.

10. The recovery process is very sensitive to cross-contamination with other plasmids that use kanamycin as selection. Unfortunately, these plasmids are very common in most labs and special attention should be paid to avoid cross-contamination. Use pipettes and a microfuge devoted only for the recovery process; alternatively, clean all used equipment with diluted bleach and 70 % ethanol.
11. DNA extraction can be performed either using a specific kit (i.e., Qiagen) or using a homemade protocol. In our hands, both protocols are equally efficient but, of course, the costs associated with the homemade protocol are lower.
12. If using a kit, just follow the manufacturer's instructions about required number of cells. If using the homemade protocol collect $\approx 5 \times 10^6$ (i.e., half of a T75 flask of cells at a 95 % confluence).
13. After 3 h, check the transparency of the lysate. If the lysate is dense/cloudy and it is difficult to pipette up and down, leave it overnight but re-add 0.5 ml of lysis buffer and 5 μl of 20 mg/ml proteinase k.
14. Two phases form: the top aqueous phase (contains the genomic DNA) and the lower phase (organic phase, undigested cellular components will be there). Be careful when collecting the top phase pipetting since it is very easy to take part of the organic phase. Pipette very carefully and do not try to recover the whole upper phase. 2/3 or 3/4 at the most might be sufficient to extract enough but purified genomic DNA.
15. In order to quickly dry pellets, if possible, leave tubes open inside a fuming hood to efficiently remove the remaining ethanol.
16. In our lab, we routinely conduct independent *Hind*III and SspI digestions to increase the efficiency of the recovery process.
17. We recommend starting in the afternoon (16.00 PM).
18. Be especially careful to avoid cross-contamination with kanamycin resistant-containing plasmids. We recommend acquire commercial ddH_2O purified water and prepare 5 ml aliquots.
19. *Hind*III produces overhang ends while *Ssp*I generates blunt ends.
20. We routinely use the collection tubes that are provided with the Amicon tubes.
21. It is important to use XL1-Blue cells, as they can replicate methylated genomic DNAs. Other strains of *E. coli* can be used, but make sure they are ultracompetent and that they can replicate methylated mammalian DNAs.

22. Avoid any cross-contamination with kanamycin resistant-containing plasmids. We recommend acquire commercial LB and prepare 5 ml aliquots.
23. In our lab, we use a very simple system: a polyspam rack attached to a Thermomixer with tape. A conventional shaker can also be used, but make sure temperature is not higher than 25 °C.
24. We recommend to seed transformations either late in the morning or soon in the afternoon, since recovery colonies need about 24 h to be visible by eye.
25. An *Eco*RI digestion will allow you to distinguish among different insertions. In fact, sequence those that do not have the same restriction pattern upon *Eco*RI digestion as they might be different insertions that occurred in the same colony.
26. Sequencing with Reco3 will reach the poly A tail of the L1 insertion. If the polyA tail it is not too long, the inferred DNA sequence might be sufficient to identify the 3′ region of the insertion. If Neo 210AS does not allow you to reach the genomic region at the 5′ end of the insertion, keep sequencing upstream with the rest of the primers that anneal on the LINE-1 sequence (Table 1).

Acknowledgments

David Cano, Santiago Morell, and Andres J. Pulgarin have contributed equally to this study and are listed alphabetically. We acknowledge current members of the J.L.G.-P. lab (Genyo) for valuable input during the course of the project. We also acknowledge Dr. John Moran (University of Michigan, USA) for helpful discussions. D.C. is supported by a FPU fellowship from the Government of Spain (MINECO, Ref AP2010-0135). J.L.G.P's lab is supported by CICE-FEDER-P09-CTS-4980, CICE-FEDER-P12-CTS-2256, Plan Nacional de I+D+i 2008-2011 and 2013-2016 (FIS-FEDER-PI11/01489 and FIS-FEDER-PI14/02152), PCIN-2014-115-ERA-NET NEURON II, the European Research Council (ERC-Consolidator ERC-STG-2012-233764) and by an International Early Career Scientist grant from the Howard Hughes Medical Institute (IECS-55007420).

References

1. Beck CR, Garcia-Perez JL, Badge RM, Moran JV (2011) LINE-1 elements in structural variation and disease. Annu Rev Genomics Hum Genet12:187–215.doi:10.1146/annurev-genom-082509-141802
2. Levin HL, Moran JV (2011) Dynamic interactions between transposable elements and their hosts. Nat Rev Genet 12(9):615–627. doi:10.1038/nrg3030
3. Macia A, Blanco-Jimenez E, Garcia-Perez JL (2015) Retrotransposons in pluripotent cells: impact and new roles in cellular plasticity. Biochim Biophys Acta 1849(4):417–426. doi:10.1016/j.bbagrm.2014.07.007

4. de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD (2011) Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet 7(12):e1002384. doi:10.1371/journal.pgen.1002384
5. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ (2001) Initial sequencing and analysis of the human genome. Nature 409(6822):860–921
6. Beck CR, Collier P, Macfarlane C, Malig M, Kidd JM, Eichler EE, Badge RM, Moran JV (2010) LINE-1 retrotransposition activity in human genomes. Cell 141(7):1159–1170
7. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH Jr (2003) Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci U S A 100(9):5280–5285
8. Luan DD, Korman MH, Jakubczak JL, Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72(4):595–605
9. Macia A, Munoz-Lopez M, Cortes JL, Hastings RK, Morell S, Lucena-Aguilar G, Marchal JA, Badge RM, Garcia-Perez JL (2011) Epigenetic control of retrotransposon expression in human embryonic stem cells. Mol Cell Biol 31(2):300–316
10. Speek M (2001) Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. Mol Cell Biol 21(6):1973–1985
11. Swergold GD (1990) Identification, characterization, and cell specificity of a human LINE-1 promoter. Mol Cell Biol 10(12):6718–6729
12. Boeke JD (1997) LINEs and Alus--the polyA connection. Nat Genet 16(1):6–7
13. Moran JV, DeBerardinis RJ, Kazazian HH Jr (1999) Exon shuffling by L1 retrotransposition. Science 283(5407):1530–1534
14. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH Jr (1996) High frequency retrotransposition in cultured mammalian cells. Cell 87(5):917–927
15. Scott AF, Schmeckpeper BJ, Abdelrazik M, Comey CT, O'Hara B, Rossiter JP, Cooley T, Heath P, Smith KD, Margolet L (1987) Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. Genomics 1(2):113–125

16. Hohjoh H, Singer MF (1996) Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. EMBO J 15(3):630–639
17. Hohjoh H, Singer MF (1997) Sequence-specific single-strand RNA binding protein encoded by the human LINE-1 retrotransposon. EMBO J 16(19):6034–6043
18. Khazina E, Weichenrieder O (2009) Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. Proc Natl Acad Sci U S A 106(3): 731–736
19. Martin SL, Bushman FD (2001) Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. Mol Cell Biol 21(2):467–475
20. Feng Q, Moran JV, Kazazian HH Jr, Boeke JD (1996) Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87(5):905–916
21. Martin F, Maranon C, Olivares M, Alonso C, Lopez MC (1995) Characterization of a non-long terminal repeat retrotransposon cDNA (L1Tc) from Trypanosoma cruzi: homology of the first ORF with the ape family of DNA repair enzymes. J Mol Biol 247(1):49–59
22. Mathias SL, Scott AF, Kazazian HH Jr, Boeke JD, Gabriel A (1991) Reverse transcriptase encoded by a human transposable element. Science 254(5039):1808–1810
23. Dmitriev SE, Andreev DE, Terenin IM, Olovnikov IA, Prassolov VS, Merrick WC, Shatsky IN (2007) Efficient translation initiation directed by the 900-nucleotide-long and GC-rich 5′ untranslated region of the human retrotransposon LINE-1 mRNA is strictly cap dependent rather than internal ribosome entry site mediated. Mol Cell Biol 27(13): 4685–4697
24. Alisch RS, Garcia-Perez JL, Muotri AR, Gage FH, Moran JV (2006) Unconventional translation of mammalian LINE-1 retrotransposons. Genes Dev 20(2):210–224
25. Esnault C, Maestre J, Heidmann T (2000) Human LINE retrotransposons generate processed pseudogenes. Nat Genet 24(4): 363–367
26. Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, Boeke JD, Moran JV (2001) Human L1 retrotransposition: cis preference versus trans complementation. Mol Cell Biol 21(4):1429–1439
27. Doucet AJ, Hulme AE, Sahinovic E, Kulpa DA, Moldovan JB, Kopera HC, Athanikar JN, Hasnaoui M, Bucheton A, Moran JV, Gilbert N (2010) Characterization of LINE-1 ribonucleoprotein particles. PLoS Genet 6(10):e1001150
28. Goodier JL, Zhang L, Vetter MR, Kazazian HH Jr (2007) LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. Mol Cell Biol 27(18):6469–6483
29. Taylor MS, Lacava J, Mita P, Molloy KR, Huang CR, Li D, Adney EM, Jiang H, Burns KH, Chait BT, Rout MP, Boeke JD, Dai L (2013) Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. Cell 155(5):1034–1048. doi:10.1016/j.cell.2013.10.021
30. Gilbert N, Lutz S, Morrish TA, Moran JV (2005) Multiple fates of l1 retrotransposition intermediates in cultured human cells. Mol Cell Biol 25(17):7780–7795
31. Coufal NG, Garcia-Perez JL, Peng GE, Marchetto MC, Muotri AR, Mu Y, Carson CT, Macia A, Moran JV, Gage FH (2011) Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. Proc Natl Acad Sci U S A 108(51):20382–20387. doi:10.1073/pnas.1100273108
32. Ostertag EM, Kazazian HH Jr (2001) Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition. Genome Res 11(12):2059–2065
33. Freeman JD, Goodchild NL, Mager DL (1994) A modified indicator gene for selection of retrotransposition events in mammalian cells. Biotechniques 17(1):46, 48-49, 52
34. Moran JV (1999) Human L1 retrotransposition: insights and peculiarities learned from a cultured cell retrotransposition assay. Genetica 107(1-3):39–51
35. Rangwala SH, Kazazian HH (2009) The L1 retrotransposition assay: a retrospective and toolkit. Methods 49(3):219–226
36. Wei W, Morrish TA, Alisch RS, Moran JV (2000) A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events. Anal Biochem 284(2):435–438
37. Cost GJ, Feng Q, Jacquier A, Boeke JD (2002) Human L1 element target-primed reverse transcription in vitro. EMBO J 21(21): 5899–5910
38. Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M, O'Shea KS, Moran JV, Gage FH (2009) L1 retrotransposition in human neural progenitor cells. Nature 460(7259):1127–1131

39. Gilbert N, Lutz-Prigge S, Moran JV (2002) Genomic deletions created upon LINE-1 retrotransposition. Cell 110(3):315–325
40. Morrish TA, Garcia-Perez JL, Stamato TD, Taccioli GE, Sekiguchi J, Moran JV (2007) Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres. Nature 446(7132):208–212
41. Morrish TA, Gilbert N, Myers JS, Vincent BJ, Stamato TD, Taccioli GE, Batzer MA, Moran JV (2002) DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nat Genet 31(2):159–165
42. Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH (2005) Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature 435(7044): 903–910
43. Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, Boeke JD (2002) Human l1 retrotransposition is associated with genetic instability in vivo. Cell 110(3):327–338