Video Article

Analysis of LINE-1 Retrotransposition at the Single Nucleus Level

Pasano Bojang[1], Kenneth S. Ramos[1,2]

[1]Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, University of Arizona College of Medicine

[2]Center for Applied Genetics and Genomic Medicine, University of Arizona College of Medicine

Correspondence to: Pasano Bojang at pbojang@email.arizona.edu, Kenneth S. Ramos at ksramos@email.arizona.edu

URL: https://www.jove.com/video/53753
DOI: doi:10.3791/53753

Keywords: Biochemistry, Issue 110, Retrotransposon, *Line-1*, FISH, Ribonucleoprotein particles (RNP), HepG2, Neomycin, Chromosome Spreads

Date Published: 4/23/2016

Citation: Bojang, P., Ramos, K.S. Analysis of LINE-1 Retrotransposition at the Single Nucleus Level. *J. Vis. Exp.* (110), e53753, doi:10.3791/53753 (2016).

Abstract

Long interspersed nuclear element-1 (*Line-1* or *L1*) accounts for approximately 17% of the DNA present in the human genome. While the majority of *L1*s are inactive due to 5' truncations, ~80-100 of these elements remain retrotransposition competent and propagate to different locations throughout the genome via RNA intermediates. While older *L1s* are believed to target AT rich regions of the genome, the chromosomal targets of newer, more active *L1s* remain poorly defined. Here we describe fluorescence *in situ* hybridization (FISH) methodology that can be used to track patterns of *L1* insertion and rates of ectopic *L1* incorporation at the single nucleus level. In these experiments, fluorescein isothiocyanate/cyanine-3 (FITC/CY3) labeled neomycin probes were employed to track *L1* retrotransposition *in vitro* in HepG2 cells stably expressing ectopic *L1*. This methodology prevents errors in the estimation of rates of retrotransposition posed by toxicity and account for the occurrence of multiple insertions into a single nucleus.

Video Link

The video component of this article can be found at https://www.jove.com/video/53753/

Introduction

Human Long Interspersed Nuclear Element-1 (*Line-1* or *L1*) is an autonomous mobile element that propagates within the genome through a "copy and paste" retrotransposition mechanism. A typical human *L1* is ~6 kb long and consists of a 5'UTR (Untranslated Region) that serves as a promoter, two open reading frames (ORFs): *L1*-ORF1 and *L1*-ORF2, and a 3'UTR with a polyA tail. *L1*-ORF1 protein has three distinct domains: a coil-coil domain, RNA recognition Motif and C-terminal domain, while *L1*-ORF2 protein has endonuclease, reverse transcriptase and cysteine rich domains[1,2,3,4,5]. *L1*-ORF1 exhibits nucleic acid chaperone activities, while *L1*-ORF2 provides enzymatic activities, with both proteins required for retrotransposition[1,2,6].

A cycle of *L1* retrotransposition starts with transcription from the 5'UTR of *L1* bicistronic mRNA by RNA polymerase II, translocation into the cytoplasm, and translation. In the cytoplasm, *L1*-ORF1 exhibits either *cis*-binding where it packages its own RNA or *trans*-binding where it packages other RNAs (*i.e.,* SINE/SVAs/pseudogenes) to form a ribonucleoprotein particle (RNP) with *L1*-ORF2p[7,8]. RNPs translocate into the nucleus where the endonuclease activity of *L1*-ORF2p nicks genomic DNA (gDNA) to expose an OH^- group[9], that is in turn used by reverse transcriptase to prime and reverse synthesize RNA to DNA from the 3' end. During reverse synthesis, the second strand of DNA is nicked 7-20 bp from the original nick site to create staggered breaks which are filled to form the signature *L1* insertion sequences (*e.g.*, TTTTAA) called target site duplications (TSD)[9,10]. This process is known as target prime reverse transcription (TPRT) and leads to insertion of full or truncated copies of *L1*/other DNAs with TSDs at both ends of the inserted sequence. *L1* retrotransposition has also been shown to be mediated through TPRT-like non-homologous end-joining in some cells types[11].

The *L1* retrotransposition vector employed in our studies is non-episomal and consists of tagged *L1*-ORF1&2p driven by combined CMV-L1-5'UTR promoters (**Figure 1**)[12]. Earlier versions of this construct have been described in studies using yeast and human cell cultures[13,14,15]. Two distinct CMV promoters located at the 5' and 3' ends of the vector, with the 3' end placed in reverse orientation to drive expression of neomycin after splicing and integration. The retrotransposition indicator cassette at the 3' end consists of a neomycin gene inserted antisense to *L1*-ORFs and rendered inactive by separation into two halves by a globin intron with spliced donor (SD) and spliced acceptor (SA) sites (**Figure 1**). Upon integration into a chromosome, *L1* is transcribed from the common promoter to produce an mRNA that consists of a bicistronic mRNA and inactive neomycin mRNA. During RNA processing, the globin intron is spliced out of the neomycin gene to restore a fully functional neomycin gene. The hybrid mRNA is packaged into a RNP in *cis* and translocated into the nucleus where it is integrated into the genome as either full length or truncated insertions using TPRT.

Here we describe a methodology that uses fluorescence *in situ* hybridization (FISH) with probes specifically directed at the spliced neomycin gene (SNeo) to track *L1*-retrotransposiition patterns and insertion rates at the single nucleus level. The efficiency and specificity of detection was confirmed using retrotransposition competent and incompetent constructs and probes to detect SNeo or the neomycin and globin intron

junction[16]. This methodology accounts for some of the shortcomings of cell culture based retrotransposition assays, such as multiple insertions, colony resistance and favorable clone expansion.

Protocol

NOTE: All steps should be carried out at room temperature unless otherwise specified. Please refer to Reagents section for details on how to prepare individual reagents.

1. Labeling *L1* Probes

NOTE: Probes can be labeled by chemical or PCR labeling.

1. **Chemical labeling**
    1. Make 0.7-1% agarose gel in Tris-acetate-EDTA (TAE) buffer, heat in a microwave until melted, allow the gel to cool, and then add ethidium bromide (0.5 µg/ml). Pour into gel tray, add combs and allow the gel to solidify at room temperature.
    2. Label the SNeo probe (1-2 µg) with CY3 or FITC at 37 °C for 1 hr according to the manufacturer's protocol.
    NOTE: SNeo denotes the **S**pliced **Neo**mycin gene expressed after a full cycle of retrotransposition. The probe is ~1,000 bp in size. SNeo can be PCR amplified from any vector or from genomic DNA. See Table of Materials and Reagents for information on Streptavidin-CY3/FITC labeling reagents.
    3. Mix the labeled SNeo probe solution with 1x DNA loading buffer, load into each well and run at 75-100 volts for 15-20 min.
    4. Visualize the probe band using a Ultra-Violet (UV) Illuminator. Note that the size of the SNeo probe can be adjusted and that Lock nuclei acid (LNA) can be also be added (**Figure 2**).
    5. Cut labeled SNeo probe from the gel with a clean blade, solubilize and purify the SNeo probe from the gel using a PCR Clean-Up kit according to manufacturer's protocol.
    CAUTION: Cover every exposed part of the body with protective shields (*i.e.,* laboratory coats and UV clear face mask) when viewing and cutting the gel. See Table of Material and Reagents for information to gel purify PCR products.
    6. Re-run 5 µl of SNeo labeled probe with an un-labeled SNeo probe on a 0.7% agarose gel (see 1.1.1) to confirm size increase and loss of signal intensity (**Figure 2**).
    7. Quantify the amount of labeled SNeo probe by measuring the absorbance at 260 nm and dilute the SNeo probe to 10 ng/µl aliquots in nuclease free H_2O. Store aliquots at -20 °C.

2. **PCR Labeling (Alternative method for probe labeling)**
    1. Combine SNeo specific primers (5'-ggatagcattgggagatatacct-3' and 5'-attgaacaagatg gattgcacgc-3') with PCR reagents into a single tube: 13.6 µl nuclease free H_2O; 0.1 µl dATP, dCTP, dGTP (10 nM); 0.05 µl dTTP (10 nM); 0.05 µl dUTP-biotin/dUTP-FITC/CY3 (10 nM); 4 µl Go Tag 5x buffer; 0.4 µl Go Tag Polymerase; 1 µl SNeo template (10 ng). See Table of Materials and Reagents for information on PCR reagents.
    2. Adjust the amount of each reagent by multiplying by the total number of samples.
    3. Use the following PCR cycling parameters to amplify and label SNeo probes: 95 °C for 2 min; 35 cycles of 95 °C for 30 sec, 62 °C for 30 sec, 72 °C for 60 sec; 72 °C for 2 min; Hold at 4 °C indefinitely.
    NOTE: The annealing temperature can be adjusted or gradient PCR can be completed to determine optimal annealing temperatures for other probes.
    4. Make 0.7-1% gel as in section 1.1.1, load and visualize the probe band as in section 1.1.4.
    5. Cut and purify the labeled SNeo probe as in section 1.1.5.
    6. Re-run 5 µl of SNeo labeled probe with un-labeled SNeo probe to confirm size increase and loss of signal intensity (see **Figure 2**).
    7. Quantify the amount of labeled SNeo probe by measuring absorbance at 260 nm and dilute SNeo probe to 10 ng/µl aliquots in nuclease free H_2O.

2. Preparing Chromosome Spreads

1. Generate stable clones expressing vector backbone (control) or retrotransposition competent *L1* using standard selection methodologies.[12,16] While non-episomal reporters were used to correlate findings with studies of transcription, episomal vectors could also be used. Grow cells stably expressing control or L1 vector (1 x 10^6 cells per 10 cm plate, with adjustments in plate size made as needed) in complete growth media. For HepG2 cells, use RPMI-1600, 10% FBS and 200 µg/ml of hygromycin. Grow cultures to 70% confluence at 37 °C and 5% CO_2.
NOTE: The choice of growth medium is dependent on cell type. Given that the plasmids used here carry selection cassettes for both hygromycin and neomycin, stable selection of clones could also be done with neomycin after selection on hygromycin, but the amount of neomycin needs to be optimized to establish tolerance levels for each cell type.
2. Add colcemid (0.4 µg/ml) to culture medium and incubate for 90 min to arrest cells at metaphase. If using different cell types, determine the optimal concentration of colcemid and time of exposure (60, 90, and 120 min) for optimal metaphase arrest empirically.
3. Wash cells 2x with 10 ml of 1x Dulbecco's PBS (DPBS), trypsinize by adding 3-5 ml of 0.25% trypsin solution to the cells and incubate for 5 min to detach the cells. Inactivate trypsin with an equal amount of media containing 10% FBS.
4. Centrifuge the cells for 2 min at 1,000 x g at 4 °C, aspirate the medium from the cell pellet and wash the cells with DPBS. Aspirate all DPBS, leaving 200 µl of DPBS to re-suspend the cells. Ensure cells are mixed well and that all clumps are dispersed. Use flickering or gentle pipetting to mix and disperse clumps and avoid vortexing.
5. Add 5 ml of hypotonic solution (*i.e.,* 75 mM KCl) that is pre-warmed to 37 °C drop-wise while rotating the 15 ml tube horizontally and incubate at 37 °C for 20 min. See Table of Materials and Reagents for information on how to obtain and make hypotonic solution.
6. Centrifuge cells at 120 x g for 5 min at 4 °C and repeat steps 2.4 - 2.5 3x leaving approximately 200 µl of hypotonic solution to re-suspend the cells after each wash. Ensure that at the end of the 3 washes, the pellet is visibly white and swollen.

7. Re-suspend the pellet in 200 µl of Carnoy Fixative solution and make 1: 2, 1: 4, 1: 8, 1:16 dilutions in Carnoy Fixative solution. Drop 10 µl of each dilution onto a dry clean slide from approximately 1 cm above and immediately expose the spread-free side to hot steam from boiling water for 30 sec. See Table of Materials and Reagents for information on how to make Carnoy Fixative solution.
8. Dry the spreads at room temperature, stain with 0.1 µg/ml Hoechst-33342 by immersion for 15-20 min in Coplin Jar and wash 3x with DPBS. See Table of Materials and Reagents for information on how to make the Hoechst-33342 solution.
9. View spreads on fluorescence/phase-contrast microscope at 40X magnification. After optimizing spread preparation, spreads **need not be stained** with Hoechst-33342. View spreads on phase-contrast microscope at 10X magnification to determine spread quality and separation as outlined in the Results section (see **Figure 3**).
10. Select and circle good spreads with a Diamond Point Marker on the opposite side of the slide (*i.e.,* spread-free side).
11. Store spreads at -20 °C for up to a month. Avoid exposure to moisture.

3. Fluorescence *In Situ* Hybridization (FISH)

1. **Stabilizing and dehydrating spreads**
   NOTE: Equilibrate metaphase chromosome spreads to room temperature if stored at -20 °C. See Table of Materials and Reagents for information on how to obtain and make reagents.
   1. Incubate metaphase chromosome spreads with 200 µl RNase A for 1 hr at 37 °C.
   2. Wash slides in 2x-SSC buffer twice for 5 min each followed by rinsing with 10 mM HCl solution.
   3. Incubate spreads with 1% pepsin for 10 min at 37 °C, rinse with deionized H_2O and wash twice with 2x-SSC buffer for 5 min each.
   4. Incubate spreads with 4% paraformaldehyde for 10 min and wash in 2x SSC buffer twice for 5 min each.
   5. Dehydrate spread by incubating for 2 min in ethanol series: 70%, 80%, and 95% ethanol.
   6. Air-dry slides.

2. **Hybridization Spreads with *L1* -labeled retrotransposition probes (*i.e.,* SNeo)**
   CAUTION: For direct-labeled probes, keep away from light at all times. See Table of Materials and Reagents for information on how to obtain and make reagents.
   1. Add 30 ng of SNeo to hybridization buffer, heat at 72 °C for 10 min and cool for 2 min at room temperature.
   2. Add 30 µl of SNeo probe solution to each spread, cover with coverslip and seal the edges with rubber cement. Ensure no bubble formation.
   3. Heat the slide at 72 °C for 5 min on a heat block, gradually drop the temperature to 37 °C, and incubate overnight in a dark humidified chamber at 37 °C.

3. **Washing and Viewing (or Addition of Secondary Antibody)**
   CAUTION: For direct-labeled probes, keep away from light at all times.
   NOTE: See Table of Materials and Reagents for information on how to prepare. Use of sufficient volume to cover the entire chromosome spread is recommended. Remember that excess volume will be lost when coverslip is added.
   1. Immerse slides in 2x SSC buffer to remove coverslips.
   2. Wash slides by immersion in 2x-SSC buffer at 45 °C for 5 min.
   3. Wash slides by immersion in wash buffer at 45 °C for 5 min, 2x.
   4. Wash slides by immersion in 0.1x SSC buffer at 45 °C for 10 min.
   5. Wash slides by immersion in 2x SSC buffer at 45 °C for 10 min.
      NOTE: Place a Coplin Jar containing recommended buffer in a water bath, adjust temperature to 45 °C.
   6. Cool slides to room temperature and equilibrate slides in detection buffer.
   7. For direct labeled probes, skip to step 3.4.10.
   8. For indirect labeled probes, block in blocking buffer for 20-30 min and wash 3x in DPBS.
   9. Incubate with 50 µl of secondary antibody (*e.g.*, 5 µg/ml streptavidin-FITC, or αCY3 in blocking buffer) for 1 hr.
   10. Wash slides in 2x SSC for 5 min twice.
   11. Counterstain with Hoechst-33342 solution (0.1 µg/ml) for 10 min.
       NOTE: this staining is done to stain chromosomes for co-localization with probe.
   12. Wash in DPBS 3x, add a drop of mounting medium, place a coverslip and seal the edges with nail polish.
   13. Analyze *L1* retrotransposition using fluorescence microscope at 40X magnification.

Representative Results

A schematic diagram of the *L1* retrotransposition vector is presented in **Figure 1**. The vector consists of a neomycin gene in antisense orientation to *L1* ORFs that is interrupted by a globin intron in sense orientation and sandwiched by SD and SA sites. When stably integrated into a chromosome, *L1* mRNA is transcribed from the combined CMV and L1-5'UTR promoter (**Figure 1**). During RNA processing, the globin intron is spliced out of the neomycin gene. The Neo-*L1* RNA is packaged, translocated and integrated into the genome as either a full length or truncated insertion. As indicated, *L1* retrotransposition can be tracked by FISH using probes specific for spliced neomycin (**Figure 1**).

Figure 2 shows a schematic representation of the neomycin probe, with the labeled probed that is larger in size and fluorescence deficient due to differences in excitation wavelengths. Note that the CY3 and FITC excite at different wavelengths than ethidium bromide stained DNA.

The density of metaphase chromosome spreads was determined by diluting the original stock solution into 1:2, 1:4, 1:8 and 1:16 dilutions and each spread evaluated for density, distribution and distance of spreads from the nucleus at 10X magnification (**Figure 3A**). The results show high density, clumpy and burst out spreads (**Figure 3A-i-ii**), as well as evenly distributed and well-spaced spreads (**Figure 3A-iii-iv**). Low density spreads with even distribution were chosen for subsequent analysis.

Chromosome spreads were stained with Hoechst dye and spread quality evaluated based on the length of chromosomes, roundness of the spreads and inter-chromosome distance (**Figure 3B**). These indices were influenced by the length of time the cells were treated with colcemid, hypotonic solution, Carnoy Fixative solution and/or the manner in which the cells were busted to release chromosomes. If cells were incubated for a short period in hypotonic solution, chromosome spreads became tightly knotted and individual chromosomes were difficult to visualize (**Figure 3B-i**). On the other hand, longer incubation in hypotonic solution resulted in rupture of the nuclei, scattering of chromosomes, and/or loss of chromosomes (**Figure 3B-ii**). Longer incubations in colcemid increased the number of cells in metaphase, but lead to condensation of chromosomes (**Figure 3B-iii**). As such, a good quality chromosome spread requires optimal incubation periods in both colcemid and hypotonic KCl solution. In our hands, a 90 min incubation in colcemid, 20 in hypotonic solution at 37 °C and the use of a hot steam to burst the cells were found to be optimal conditions for generation of high quality spreads (**Figure 3B-iv**).

Chromosome spreads were stained for *L1* retrotransposition using probes that target SNeo. The probe was labeled with both FITC and CY3 to show that in both cases *L1* retrotransposition is seen only in cells expressing wild type *L1* (**Figure 4**). No staining was seen in cells expressing the vector backbone alone (**Figure 4**; compare top and bottom panels for each fluorophore). In our hands, probe signal was detected in 80-90% of nuclei, with the overwhelming majority of the signal representing single retrotransposition events. We only scored retrotransposition in nuclei with more than one insertion and spread quality was always examined before FISH.

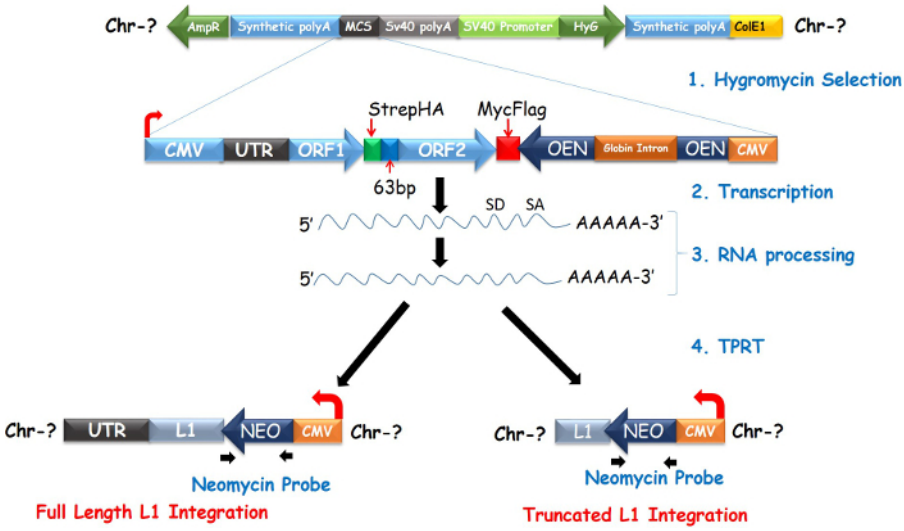

Figure 1. Schematic diagram of *L1* retrotransposition vector. Diagram depicts the process of target prime reverse transcription leading to full or truncated *L1* insertions. Please click here to view a larger version of this figure.

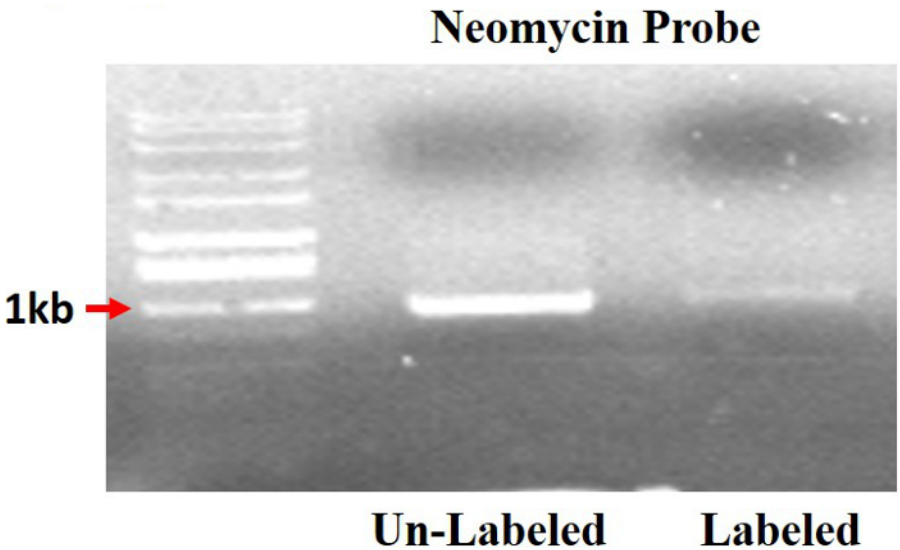

Figure 2. Labeled SNeo and unlabeled neomycin probes. The labeled probe exhibits less fluorescence and is larger in size compared to the unlabeled probe. Please click here to view a larger version of this figure.

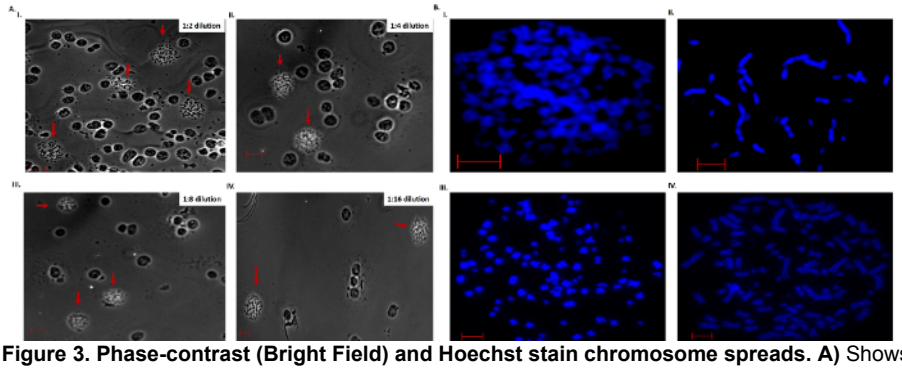

Figure 3. Phase-contrast (Bright Field) and Hoechst stain chromosome spreads. A) Shows dilution of spreads to obtain well-spaced preparations. Scale bar is 50 µm. **B)** Hoechst stain chromosome showing the influence of colcemid, hypotonic solution, Carnoy Fixative solution and the bursting on spread quality. Scale bar is 10 µm. Please click here to view a larger version of this figure.

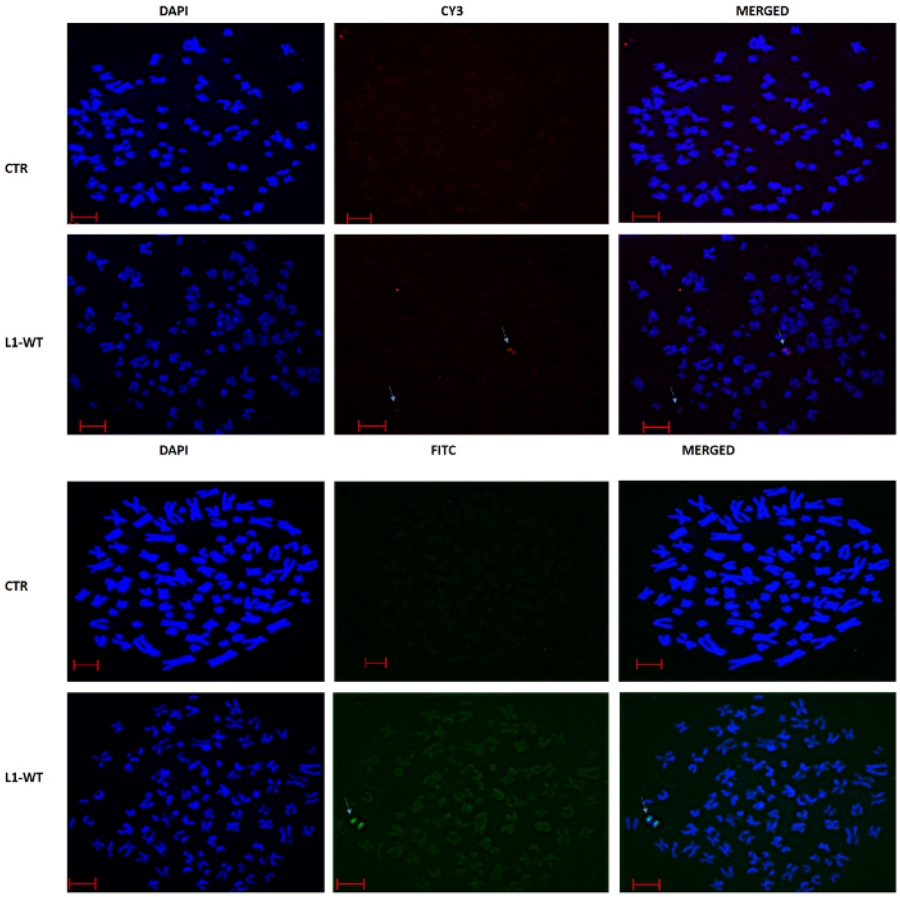

Figure 4. FISH analysis of *L1* retrotransposition. Staining of SNeo is absent in control cells, but present in cells expressing *L1* wildtype vector. Scale bar is 10 µm. Please click here to view a larger version of this figure.

Discussion

Methodologies such as whole genome sequencing, inverse PCR and southern blotting have been employed to study *L1* retrotransposition. Although these methodologies are extremely valuable in locating where *L1* insertions occur within genomes, a confounding challenge for all of them is the need for *in-silico* programing to reassemble sequences. The FISH methodology described here is designed to complement these methods, especially in the case of studies requiring analyses of ectopic *L1* retrotransposition in cultured cells. The approach can be used for both quantitative and qualitative evaluation of retrotransposition events and applied to interphase nuclei. This methodology boasts the accuracy of subsequent *in-silico* sequence alignment methods employed to determine ectopic *L1* insertion sites[16]. Alignment of repetitive sequences can be ambiguous due to the formation of homopolymers, while repetitive sequences can be lost during primer amplification of template resulting in underrepresentation of repetitive sequences. FISH methodology can overcome these problems because FISH probes can anneal to homopolymers and are unlikely to be biased against intact repeat sequences[16]. In addition, compared to cell culture-based retrotransposition assays, FISH can determine retrotransposition rates at the single nucleus level. As such, this approach prevents both under and over estimation

of retrotransposition rates as multiple insertions can occur into single nuclei[16]. Recent data obtained from NGS have confirmed that cell culture based methodologies underestimate retrotransposition frequencies[25].

It remains unclear where young *L1s* prefer to insert and whether a targeting mechanism exists to direct these insertions within the genome. Using the above methodology we have shown that *L1* inserts preferentially into gene poor regions of the genome[16]. Others have shown that the abundance of L1 sequences is increase in chromosomes with a lesser gene density[17,18]. Together, these findings suggest a regulated mode of insertion where threats to the cell are minimized by insertion of *L1* into "lesser active" regions of the genome. The pattern of transposable element movement in lower organisms has lent support to this interpretation. For example, fission yeast retrotransposon, Tf1, integrates 95% of the time upstream of gene promoters and the majority of these genes are involved in stress regulation[19,20]. In yeast cells that lack access to nitrogen, Ty5 integrates into the ORFs instead of heterochromatin regions[21]. Experiments in maize have shown that integration of DNA transposons lead to variegated corn color phenotypes and integration of *Hatvine1-rrm* DNA transposon into the promoter region of *VvTFL1A* gene has been shown to influence the branching pattern and the fruit size of grapevines[22,23].

The steps critical to the success of the FISH methodology detailed here are the generation of good chromosome spreads and proper labeling, purification and hybridization of probes. If probes are not cleaned properly, the resulting high background signal will make it difficult to resolve probe binding to chromosomes. Preferably, probes should be gel-purified to eliminate residual fluorescence from dNTPs. Also, the length of time that spreads are incubated in Carnoy Fixative solution is important to prepare good chromosome spreads. If too long, the cell membrane may burst prematurely and cause chromosome loss. If too short, it may be difficult to obtain suitable spreads because of deficient membrane rupture. The amount/concentration of formamide determines the focus of chromatids and therefore it is important to optimize empirically for best results. All washes should be thorough to ensure that all unbound probes and secondary antibodies are washed off.

In conclusion, the FISH methodology described here can be used to detect both full length and truncated ectopic *L1* retrotransposition events and can be applied to other retrotransposition vectors using green fluorescence protein (GFP) as a retrotransposition indicator cassette. Although full length *L1* insertions can be determined using this methodology, it will not discern new endogenous truncated insertions due to the number of truncated *L1*s within the genome. The accessibility of probe might also be restricted by increase heterochromatin formation[16,24]. However, the inclusion of LNA within FISH probes can be employed to enhance probe annealing. Detection of SNeo is contingent upon a full cycle of retrotransposition and insertions smaller than the probe/deletion can be missed.

For detailed descriptions of experiments in which the use of these vectors and the expression of L1 proteins and retrotransposition are characterized, please refer to published works[12,16].

Disclosures

The authors declare no potential competing interests.

Acknowledgements

This work was supported in part by grants from the National Institute of Environmental Health Sciences (ES014443 and ES017274) and AstraZeneca to KSR.

References

1. Mathias, S. L., Scott, A. F., Kazazian, H. H., Jr., Boeke, J. D., & Gabriel, A. Reverse transcriptase encoded by a human transposable element. *Science.* **254**, 1808-1810 (1991).
2. Feng, Q., Moran, J. V., Kazazian, H. H., Jr., & Boeke, J. D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell.* **87**, 905-916, (1996).
3. Clements, A. P., & Singer, M. F. The human LINE-1 reverse transcriptase:effect of deletions outside the common reverse transcriptase domain. *Nucl Acids Res.* **26**, 3528-3535 (1998).
4. Martin, S. L., Branciforte, D., Keller, D., & Bain, D. L. Trimeric structure for an essential protein in L1 retrotransposition. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13815-13820 (2003).
5. Khazina, E., *et al.* Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. *Nat Struct Mol Biol.* **18**, 1006-1014 (2011).
6. Martin, S. L., Li, J., & Weisz, J. A. Deletion analysis defines distinct functional domains for protein-protein and nucleic acid interactions in the ORF1 protein of mouse LINE-1. *J Mol Biol.* **304**, 11-20 (2000).
7. Martin, S. L. Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells. *Mol Cell Biol.* **11**(9), 4804-4807 (1991).
8. Hohjoh, H., & Singer, M. F. Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *EMBO J.* **15**, 630-639 (1996).
9. Cost, G. J., Feng, Q., Jacquier, A., & Boeke, J. D. Human L1 element target-primed reverse transcription in vitro. *EMBO J.* **21**, 5899-5910 (2002).
10. Szak, S. T., *et al.* Molecular archeology of L1 insertions in the human genome. *Genome Biol.* **3** (2002).
11. Sen, S. K., Huang, C. T., Han, K., & Batzer, M. A. Endonuclease-independent insertion provides an alternative pathway for L1 retrotransposition in the human genome. *Nucl Acids Res.* **35**, 3741-3751 (2007).
12. Bojang, P., Jr., Roberts, R. A., Anderton, M. J., & Ramos, K. S. Reprogramming of the HepG2 genome by long interspersed nuclear element-1. *Mol Oncol.* **7**, 812-825 (2013).
13. Boeke, J. D., Garfinkel, D. J., Styles, C. A., & Fink, G. R. Ty elements transpose through an RNA intermediate. *Cell.* **40**, 491-500 (1985).
14. Heidmann, T., Heidmann, O., & Nicolas, J. F. An indicator gene to demonstrate intracellular transposition of defective retroviruses. *PNAS.* **85**, 2219-2223 (1988).
15. Moran, J. V., *et al.* High frequency retrotransposition in cultured mammalian cells. *Cell.* **87**, 917-927 (1996).

16. Bojang, P., Jr., Anderton, M. J., Roberts, R. A., & Ramos, K. S. De novo LINE-1 retrotransposition in HepG2 cells preferentially targets gene poor regions of chromosome 13. *Genomics.* **104**, 96-104 (2014).
17. Rozen, S. *et al.* Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature.* **423**, 873-876 (2003).
18. Graves, J. A., Koina, E., & Sankovic, N. How the gene content of human sex chromosomes evolved. *Curr Op Gen Develop.* **16**, 219-224 (2006).
19. Leem, Y. E. *et al.* Retrotransposon Tf1 is targeted to Pol II promoters by transcription activators. *Mol Cell.* **30**, 98-107 (2008).
20. Guo, Y., & Levin, H. L. High-throughput sequencing of retrotransposon integration provides a saturated profile of target activity in Schizosaccharomyces pombe. *Genome Res.* **20**, 239-248 (2010).
21. Dai, J., Xie, W., Brady, T. L., Gao, J., & Voytas, D. F. Phosphorylation regulates integration of the yeast Ty5 retrotransposon into heterochromatin. *Mol Cell.* **27**, 289-299 (2007).
22. McClintock, C. B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. U.S.A.* **36**, 344-355 (1950).
23. Fernandez, L., Torregrosa, L., Segura, V., Bouquet, A., & Martinez-Zapater, J. M. Transposon-induced gene activation as a mechanism generating cluster shape somatic variation in grapevine. *Plant J Cell Mol Biol.* **61**, 545-557 (2010).
24. Garcia-Perez, J. L., *et al.* Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. *Nature.* **466** (7307), 769-73 (2010).
25. Rodic, N., *et al.* Retrotransposon insertions in the clonal evolution of pancreatic ductual adenocarcinoma. *Nat Med.* **21** (9) (2015).

