LINE-1 retrotransposition in human embryonic stem cells

Jose L. Garcia-Perez1, Maria C.N. Marchetto3, Alysson R. Muotri3, Nicole G. Coufal3, Fred H. Gage3, K. Sue O’Shea2 and John V. Moran1,*

1Departments of Human Genetics and Internal Medicine, 2Department of Cell and Developmental Biology, University of Michigan Medical School, 1241 E. Catherine Street, Ann Arbor, MI 48109, USA and 3Laboratory of Genetics, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA

Received December 14, 2006; Revised March 6, 2007; Accepted April 17, 2007

LINE-1 elements comprise approximately 17% of human DNA and their mobility continues to impact genome evolution. However, little is known about the types of non-transformed cells that can support LINE-1 retrotransposition. Here, we show that human embryonic stem cells express endogenous LINE-1 elements and can accommodate LINE-1 retrotransposition in vitro. The resultant retrotransposition events can occur into genes and can result in the concomitant deletion of genomic DNA at the target site. Thus, these data suggest that LINE-1 retrotransposition events may occur during early stages of human development.

INTRODUCTION

Long interspersed element-1 (LINE-1 or L1) is an abundant retrotransposon that comprises approximately 17% of human DNA (1). The average human genome contains approximately 80–100 retrotransposition-competent LINE-1 elements (RC-L1s) (2). RC-L1 mobility in both germ and somatic cells has resulted in a variety of genetic disorders, including hemophilia A, muscular dystrophy and colon cancer [reviewed in (3,4)]. The LINE-1-encoded proteins are also responsible for the mobilization of Alu elements, certain short interspersed elements (SINEs) and the formation of processed pseudogenes, which together comprise approximately 13% of human DNA (3–7). Thus, LINE-1-mediated retrotransposition events are responsible for at least one billion bases in human DNA and have had a tremendous impact on human genome evolution.

Human RC-L1s are approximately 6 kb in length and contain a 5′-untranslated region (UTR), two non-overlapping open reading frames (ORF1 and ORF2) and a 3′-UTR that ends in a poly (A) tail (8). Genetic and biochemical studies indicate that the ORF1- and ORF2-encoded proteins (ORF1p and ORF2p) exhibit a strong cis-preference and preferentially associate with their encoding RNA to form a ribonucleoprotein particle (6,9,10), which is a likely retrotransposition intermediate (11,12). The resultant ribonucleaseprotein particle is transported to the nucleus (13) where LINE-1 retrotransposition is thought to occur by a mechanism termed target-site primed reverse transcription (TPRT) (10,14–16). Genetic and biochemical studies indicate that conventional TPRT requires a nucleic acid chaperone activity contained within ORF1p (17,18) as well as endonuclease and reverse transcriptase activities contained within ORF2p (15,19).

Although it is estimated that that up to 5% of newborns may contain a de novo L1-mediated retrotransposition event (4,20), relatively little is known about the developmental timing or the cell types that accommodate LINE-1 retrotransposition in humans. In vivo studies using mouse models indicate that LINE-1 expression and/or retrotransposition can occur in male and female germ cells, during early development, and in selected somatic tissues (12,21–23). By comparison, an in vitro retrotransposition assay has been used to demonstrate LINE-1 retrotransposition in a variety of human and rodent transformed cell lines (19,24–28), in rat neuronal progenitor cells (23) and at a relatively low level in primary human fibroblasts (13,29). Here, we demonstrate that human embryonic stem (hES) cells can accommodate the retrotransposition of engineered LINE-1 elements in vitro. These data suggest that LINE-1 retrotransposition events may occur at early stages in human embryogenesis and that some individuals in the population may be genetic mosaics with respect to their LINE-1 content.
RESULTS AND DISCUSSION

To analyze if undifferentiated hES cells express endogenous LINE-1 elements, we isolated ribonucleoprotein particles (RNPs) from federally approved H9 hES cells (30). Endogenous ORF1p and LINE-1 mRNA were present in those RNPs, and sequencing of the LINE-1 RT–PCR products showed that the RNAs were derived from both active and inactive L1 subfamilies (e.g. L1HS, L1PA2, L1PA3, L1PA4, L1PA6 and L1PA7; Fig. 1A), which is consistent with results from similar experiments conducted with human teratocarcinoma cells (Fig. 1A) (31). Similar RNA expression data were also

Figure 1. LINE-1 elements can retrotranspose in undifferentiated hES cells. (A) Left panel: Western blot analysis using a polyclonal antibody against ORF1p in ribonucleoprotein particles derived from H9 passage (p) 41 cells. N-tera 2D1 human teratocarcinoma cells are shown as a positive control (31). HeLa cells are used as a negative control. MW, molecular weight size standards. Right panel: RT–PCR analysis of LINE-1 RNA in H9p41 cell lines. C− denotes an RT–PCR reaction lacking input RNA. MW denotes 1 kb size standard. (B) Rationale of the LINE-1 retrotransposition assay (19). Details are described in the text. (C) (1) An undifferentiated Cyth25 colony containing a LINE-1 retrotransposition event from an RC-L1 equipped with the EGFP indicator cassette. (2) An undifferentiated H1p75 hES colony harboring a LINE-1 retrotransposition event from an RC-L1 equipped with the NEO indicator cassette. (3) An undifferentiated H9p60 hES colony harboring a LINE-1 retrotransposition event from an RC-L1 equipped with the NEO indicator cassette. (D) (see next page) PCR confirms precise splicing of the intron from the REP indicator cassette (19,21). Cartoons of the unspliced (1243 bp using the EGFP indicator cassette or 1361 bp using the NEO indicator cassette) and spliced (342 bp using the EGFP indicator cassette or 468 bp using the NEO indicator cassette) products are shown beside the gels. Controls for the spliced and unspliced products are indicated by C+ in the top gel and −intron (−I) and +intron (+I) in the bottom gel.
obtained from BG01, H1, H7, HSF-6 cells and from a non-
federally approved cell line (Cyth25; Supplementary Material,
Fig. S1A).

To determine if a LINE-1 element could retrotranspose in
hES cells, we transfected undifferentiated hES cells with
either a human RC-L1 (2) whose expression was driven
from its native 5'-UTR or a synthetic mouse RC-L1 (27)
whose expression was augmented by the cytomegalovirus
immediate early (CMV) promoter. These RC-L1s were tag-
ged with a retrotransposition indicator cassette in their respect-
ive 3'-UTRs, which consists of a backward copy of a reporter
gene (REP) containing its own promoter and polyadenylation
signal (Fig. 1B) (19,21). The reporter gene is also interrupted
by an intron in the same transcriptional orientation as the
RC-L1. This arrangement ensures that the reporter cassette
will only become activated if spliced RC-L1 mRNA under-
goes a successful round of retrotransposition (Fig. 1B).

RC-L1s containing either an enhanced green fluorescent
protein (EGFP) or neomycin phosphotransferase (NEO)
retrotransposition indicator cassette were transfected into
er four different hES cell lines (H1, H9, BG01 and Cyth25;
Fig. 1C). A low level of retrotransposition was reproducibly
observed in early and late passage hES cells transfected with
the RC-L1s. However, we never detected retrotransposition
from mutant LINE-1 constructs that contain either a missense
mutation in the reverse transcriptase active site (JM101/L1.3/
D702A) or two missense mutations in the ORF1 RNA-binding
domain (JM111/L1 RP/RR261-262AA). As expected, PCR
experiments conducted using genomic DNA templates iso-
lated from hES cell lines containing a retrotransposition
event confirmed the precise splicing of the intron from the
retrotransposition indicator cassette (Fig. 1D).

Additional control experiments demonstrated that the
G418-resistant cell lines express both OCT4 and TRA1–60
(30) (Fig. 2A; Supplementary Material, Fig. S1B–D), and
that they could be differentiated into embryoid bodies
(Fig. 2B; Supplementary Material, Fig. S1E) that express
mRNAs diagnostic for the three germ layers (Fig. 2C).
Together, these data demonstrate that an engineered LINE-1,
whose expression is driven by its own promoter, is able to
retrotranspose in undifferentiated hES cells in vitro.

We next characterized seven retrotransposition events from
hES cells. Insertion (A) was generated from a synthetic mouse
LINE-1 element whose expression was driven from a heter-
ologous CMV promoter, whereas insertions (B–G) were gen-
erated from two different human RC-L1s [LRE3 (insertions
B–F) or L1 RP (insertion G)] whose expression was driven
from their endogenous 5'-UTR (Fig. 3). The analysis of the
pre- and post-integration sites revealed structural hallmarks
of LINE-1 retrotransposition (Fig. 3, A–G, Table 1; Sup-
plementary Material, Fig. S2). Each insertion was 5'-truncated
and ended in a poly (A) tail. Five of seven events were flanked
by a variable sized target site duplication that ranged in size
from 3 to 16 bp, and at least six of the events integrated into
a sequence that closely resembles a consensus LINE-1 endo-
nuclease cleavage site (15,32) (5'-TTTT/A, Table 1; Sup-
plementary Material, Fig. S2). Three insertions occurred into
the introns of known (Fig. 3; insertions B and F) or hypothe-
sical genes (Fig. 3; insertion G). Two other events were associ-
ated with genomic deletions at the target site of either
Figure 2. hES cell lines containing a de novo LINE-1 retrotransposition event express multipotent markers and can be differentiated into embryoid bodies that express mRNAs diagnostic for the three major germ layers. (A) (1) Phase contrast micrograph of G418-resistant H9p60 foci. (2) Hoechst staining of the same cells. (3) OCT4 staining. (4) TRA1–60 staining. Bars = 100 μm. (B) Embryoid bodies derived from H1p75 hES cells containing a LINE-1 retrotransposition event after 9 days in culture. Bar = 500 μm. (C) (see next page) RT–PCR experiments showing expression of endoderm [α-fetoprotein (AFP), and GATA-4], mesoderm [myogenic factor 5 (MYF-5)] and ectoderm [keratin-18 (KRT-18) and Musashi 1] specific genes in embryoid bodies derived from various cell lines containing LINE-1 retrotransposition events (listed at the top of the figure). LINE-1 and β-actin RT–PCR products are shown as controls. MW, molecular weight markers. Water denotes (negative) control reactions lacking input RNA. hES denotes control reactions with total RNA from undifferentiated H1 cells.
approximately 939 or 25 bp, respectively (Fig. 3, insertions E and F). Interestingly, insertions C and D occurred into a similar region of genomic DNA and were only separated by 344 bp (Fig. 3), suggesting that this sequence on chromosome 15 may be a preferred site of LINE-1 integration and/or expression of the retrotransposition indicator cassette in hES cells. Together, these results are consistent with previously published studies using transformed human cell lines and rat neural progenitor cells (19,23,32).

Our data indicate that genetically engineered human and mouse LINE-1 elements can retrotranspose in hES cells. Technical issues (e.g. the difficulty in transfecting hES cells (10%), their low cloning efficiency (<1%) and their tendency to differentiate) make it difficult to precisely determine the LINE-1 retrotransposition efficiency at this time; however, it appears to be at least an order of magnitude lower than what has been observed in transformed cell lines (19,24–28). In addition to the technical challenges involved in working with hES cells, we can envision at least two reasons to account for this apparently lower retrotransposition efficiency. First, LINE-1 retrotransposition events in hES cells may integrate into non-permissive chromatin contexts or may be subject to epigenetic silencing leading to the inactivation of the reporter cassette either during or soon after retrotransposition (23). Secondly, hES cells may express endogenous proteins that serve to protect the genome from de novo retrotransposition events. Clearly, additional studies and technical advances in the manipulation of hES cells are necessary to address these interesting possibilities.

In sum, our results show that endogenous LINE-1 elements are expressed in hES cells, that retrotransposition events derived from engineered L1s can occur into genes, and that some of those insertions are accompanied by concomitant deletions of target site nucleotides. Although in vitro experiments conducted in isolated hES cells do not fully reflect the environment in utero, the available data suggest that LINE-1 retrotransposition events may occur at early stages in human embryogenesis and that some individuals in the population may be genetically mosaic with respect to their LINE-1 content. Indeed, the accompanying paper by van den Hurk et al. provides compelling genetic experiment that support our findings. Thus, in addition to the previously described mechanisms of genome instability in hES cells (33), we speculate that LINE-1 retrotransposition could contribute to hES genome fluidity and variability.

**MATERIALS AND METHODS**

**Culturing conditions**

The following hES cell lines [H1, H7, H9 (30), BG01 (34), HSF-6 (35) and Cyth 25 (36,37)] (Cythera Inc., San Diego, USA) were grown in gelatin-coated plates on mouse embryonic fibroblasts (MEFs, passage 3) derived from CF-1 mice. MEFs were mitotically inactivated by γ-irradiation with a 2100 Cesium source indicator with 3000–3200 rads, and then were seeded onto gelatin-coated plates (Sigma) at a density of 25 000 cells/cm²; MEFs were routinely used within 3 days after being irradiated. Culture medium for hES cells consists of: DMEM F12 (Invitrogen) supplemented with 4 ng/ml of β-FGF (fibroblast growth factor), 20% KO serum replacement, 1 mM L-glutamine, 50 μM β-mercaptoethanol and 0.1 mM non-essential amino acids. Cells were passaged by manual dissection every 4 days and the medium was replaced daily. Frozen stocks of karyotyped hES cells were used for approximately 30 passages.

**Plasmid construction**

Cloning strategies are available upon request. All plasmids used in this study contain the indicated fragments of L1 DNA cloned into pCEP4 (Invitrogen) or pBSKS-II (Stratagene). cep99-gfp-L1RP-hygro has been described previously (38). It contains an 8.7-kb NotI–BamHI fragment containing a full-length L1RP element (39), and the EGFP retrotransposition indicator cassette. It is cloned in a modified version of pCEP that lacks the CMV promoter. cep99-gfp-JM111-hygro is a derivative of cep99-gfp-L1 RP-hygro that contains the double missense mutation RR261-262AA in ORF1p (38). ks-99-gfp-LRE3-sv(+) contains an 8.9-kb NotI–SalI fragment containing a full-length LRE3 element (40), the EGFP retrotransposition indicator cassette and the SV40 late
polyadenylation signal. It is cloned in pBSKS-II. The LRE3 element was a kind gift from Dr Haig Kazazian.

ks-99-gfp-JM111-sv(+) contains an 8.9-kb NotI–SalI fragment containing a full-length L1RP element that contains the double missense mutation RR261-262AA in ORF1p (38), the EGFP retrotransposition indicator cassette and the SV40 late polyadenylation signal. It is cloned in pBSKS-II.

pcepL1SM contains a codon optimized full-length mouse element (derived from L1spa) containing the mneoI indicator cassette (27). The plasmid was a kind gift from Dr Jef Boeke.

ks-101-LRE3-sv(+) contains an 8.7-kb NotI–SalI fragment containing a full-length LRE3 element, the NEO retrotransposition indicator cassette and the SV40 late polyadenylation signal. It is cloned in pBSKS-II.

Table 1. The structural hallmarks of LINE-1 retrotransposition events generated in hES cells

<table>
<thead>
<tr>
<th>Clone</th>
<th>Chromosome</th>
<th>Type</th>
<th>Length</th>
<th>Cleavage</th>
<th>TSD (bp)</th>
<th>Poly(A) (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H153-SM</td>
<td>15</td>
<td>5'-truncated + 1 bp at 5' end</td>
<td>1115</td>
<td>TCTT/A</td>
<td>3</td>
<td>~45</td>
</tr>
<tr>
<td>H175-L3–1</td>
<td>4</td>
<td>5'-truncated</td>
<td>331</td>
<td>TTTT/A</td>
<td>4</td>
<td>~39</td>
</tr>
<tr>
<td>H175-L3–2</td>
<td>15</td>
<td>5'-truncated</td>
<td>247</td>
<td>AATT/A</td>
<td>15</td>
<td>~64</td>
</tr>
<tr>
<td>H175-L3–3</td>
<td>15</td>
<td>5'-truncated</td>
<td>355</td>
<td>TTTT/A</td>
<td>16</td>
<td>~58</td>
</tr>
<tr>
<td>H175-L3–4</td>
<td>7</td>
<td>5'-truncated 939 bp deletion</td>
<td>845</td>
<td>N.A. Del. 939 bp</td>
<td>~24</td>
<td></td>
</tr>
<tr>
<td>H176-L3</td>
<td>11</td>
<td>5'-truncated 25 bp deletion</td>
<td>570</td>
<td>TTTT/G</td>
<td>–</td>
<td>~24</td>
</tr>
<tr>
<td>Cyth25-RP-1</td>
<td>6</td>
<td>5'-truncated</td>
<td>964</td>
<td>TTTT/A</td>
<td>5</td>
<td>~82</td>
</tr>
</tbody>
</table>

Column 1, the name of the cell line; column 2, the chromosomal location of the insertion; column 3, sequence characteristics of the insertion; column 4, length of the insertion (excluding the retrotransposition indicator cassette); column 5, the sequence of the LINE-1 endonuclease bottom strand cleavage site; column 6, target site alterations associated with retrotransposition; column 7, length of the poly (A) tail. N.A., not analyzed. Please see Supplementary Material, Figure S2 for additional details.

Figure 3. LINE-1 retrotransposition events in hES cells exhibit hallmarks of retrotransposition. Schematics of seven LINE-1 retrotransposition events are shown (A–G). The names of the cell lines are indicated at the top of each panel. The chromosomal location of each insertion is indicated. Insertions that occurred into known or hypothetical genes are indicated, as are those that resulted in target site deletions (red rectangles; Insertions E & F). Orange arrowheads indicate target site duplications flanking the LINE-1 element. The length of the poly (A) tail is indicated. The numbering indicates the truncation site in the LINE-1 and is based on the L1.2 reference sequence (accession number M80343). Additional experimental details are provided in Supplementary Material, Figure S2.
ks-105-L1.3-sv(+) contains an 8.7-kb NotI–SalI fragment containing a full-length L1.3 element that contains a missense mutation in the RT active site (D702A) (25), the NEO indicator cassette and the SV40 late polyadenylation signal. It is cloned in pBSKS-II.

DNA preparation
LINE-1 plasmid DNAs were purified on Qiagen midi prep columns (Qiagen). DNAs for transfection experiments were checked for supercoiled by electrophoresis on 0.7% agarose-ethidium bromide gels. Only highly supercoiled preparations of DNA (>90%) were used for transfection. BAC clones were purchased from BACPAC Resources (Children's Hospital and Research Center, Oakland, CA, USA). BAC DNAs were purified on Qiagen midi prep columns according to the manufacturer's protocol.

RT–PCR analysis
Total RNA was extracted with Trizol (Invitrogen) and treated with RNase-free RQ-1 DNase I (Promega). RNA (1 μg) was reverse-transcribed with MMLV RT (25 U, Promega) primed with Oligo dT 12–18 (Invitrogen) for 1 h at 42°C. Rabbit polyclonal anti-ORF1 antibody (a generous gift from Dr Thomas Fanning) was used at a 1/10,000 dilution, followed by a goat anti-rabbit HRP conjugate secondary antibody at a 1/20,000 dilution (Abcam).

hES cell transfection and retrotransposition assay
Cells were transfected with either FuGene 6 (Roche) as previously described (42) or by Nucleofection using the V-Kit solution (Amaxa) and the A-23 program following manufacturer's instructions. Transient transfection using LINE-1 elements containing the 5′EGFP retrotransposition indicator cassette in a pBSKS(−) backbone were performed with approximately 2–4×10⁶ hES cells, which were then grown on plates containing irradiated MEFs. Cells were re-fed daily, and the appearance of EGFP (+) cells was monitored using fluorescence microscopy.

Immunocytochemistry
hES cells grown on MEFs or matrigel-coated dishes were fixed with paraformaldehyde (2% in PBS 1×) for at least 1 h at 4°C. Immunocytochemistry was performed as previously described (41). The Oct4 goat polyclonal antibody (Santa Cruz) was used at a 1/100 dilution. The Tra1–60 mouse monoclonal antibody (Chemicon) was used at a 1/50 dilution. Secondary antibodies were purchased from Jackson ImmunoResearch, and were used at a 1/200 dilution. Hoechst was purchased from Sigma and used at a 1/10,000 dilution.

Preparation of whole cell lysates, sucrose cushions and western blot analysis
Approximately 8×10⁶ hES cells were manually harvested, washed with 1× PBS and lysed with 1 ml of 1.5 mM KCl, 2.5 mM MgCl₂, 5 mM Tris–HCl pH 7.4, 1% deoxycholic acid, 1% Triton X-100, 1× Complete Mini EDTA-free protease inhibitor cocktail (Roche). After a 10-min incubation on ice, the cell debris was removed by centrifugation at 3000g at 4°C for 5 min. A sucrose cushion was prepared with 8.5 and 17% w/v sucrose in 80 mM NaCl, 5 mM MgCl₂, 20 mM Tris–HCl pH 7.5, 1 mM DTT, 1× Complete Mini EDTA-free protease inhibitor cocktail. Lysates were spun at 178,000 g at 4°C for 2 h using a Sorvall SW-41 rotor. After centrifugation, the pellet material was resuspended in 50 μl of purified water with 1× Complete Mini EDTA-free protease inhibitor cocktail (Roche). Total protein concentration was determined by Bradford assay (BioRad) according to the manufacturer's instructions, and 11 μg of sample were loaded on a pre-cast 10% SDS–PAGE gel (BioRad). Rabbit polyclonal anti-ORF1 antibody (a generous gift from Dr Thomas Fanning) was used at a 1/10,000 dilution, followed by a goat anti-rabbit HRP conjugate secondary antibody at a 1/20,000 dilution (Abcam).

Differentiation of hES cells into embryoid bodies
hES cells grown on MEFs were manually seeded onto non-coated 60 mm dishes in hES media lacking β-FGF [Dulbecco’s modified Eagle’s medium (DMEM) F12 (Invitrogen) supplemented with 20% knockout (KO) serum replacement, 1 mM L-glutamine, 50 μM β-mercaptoethanol and 0.1 mM non-essential amino acids]. The resultant embryoid bodies were re-fed every other day with the same medium and total
Supplementary Material

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

We are indebted to Nancy Leff, Crystal Pacut and Mariana Bormann in the University of Michigan Stem Cell Core for their excellent technical assistance. We thank Drs Gary Smith and Luis Villa for sharing their hES cell freezing protocol prior to publication, and Dr Deanna Kulpa for help in the preparation of LINE-1 RNPs. We thank Dr Haig Kazazian for providing the pJCC5/LRE3 plasmid, Dr Jef Boeke for providing the synthetic mouse LINE-1 construct and Dr Thomas Fanning for providing the polyclonal ORF1 antibody. J.V.M. is supported by grants GM60518 and a pilot project from GM069985 (K.S.O. is the P.I.) from the National Institutes of Health. F.H.G. is supported by grants GM60518 and a pilot project from MEC, Spain). M.C.M.N. is supported by a George E. Smith and Luis Villa for sharing their hES cell freezing protocol.

Conflict of Interest statement. None declared.

REFERENCES


