Research Article

LINE-1 Based Insertional Mutagenesis Screens to Identify Genes Involved in Embryonic Stem Cell Differentiation

Sung-Hun Lee and Danny Rangasamy*

John Curtin School of Medical Research, The Australian National University, Canberra, Australian Capital Territory 2601, Australia

Abstract

Background: Knowledge of the intrinsic properties of embryonic stem cells is essential before the possibility of using these cells for therapeutic purposes becomes a reality. Insertional mutagenesis screens are widely used to identify genes sufficient to confer a particular phenotype; however, applications of such approaches are limited to mouse models. Thus, there is a need to develop a new DNA transposon system that allows gene discovery approaches in stem cells.

Methods: This study was performed to explore the possibility of using the long interspersed nuclear element 1 (LINE -1) retrotransposon as a gene trap vector to identify genes involved in the maintenance and differentiation of mouse embryonic stem cells.

Results: We developed an episomal, nonviral LINE-1 retrotransposon system using the scaffold/matrix attachment regions in the backbone of our vector. This gene trap vector harbors a GFP marker whose expression occurs only after successful insertional mutagenesis. By utilizing this vector, coupled with GFP expression, we have successfully isolated four individual embryonic stem cell clones that display disrupted genes, including two known genes. We then confirmed the identity of these genes using an inverse PCR approach and verified their function in cell differentiation using shRNAs and undifferentiated markers of embryonic stem cells.

Conclusions: The ease of using this insertional mutagen and the simplicity of identifying the cells with disrupted genes by GFP expression make this LINE-1 vector a promising tool for embryonic stem cell and cancer stem cell gene discovery.

Keywords: Gene disruption; Small interference RNA; Retrotransposon; Cell differentiation

Introduction

Embryonic stem (ES) cells offer great hope for the future of medicine in such areas as cardiovascular disease, diabetes and Parkinson's disease. Of course, knowledge of the intrinsic properties of ES cells must be gained before the possibility of using these cells for therapeutic purposes becomes a reality. ES cells possess an unlimited capacity to be maintained as self-renewing undifferentiated cells. Their ability to grow for prolonged periods of time while maintaining a normal karyotype and pluripotency offers the enormous potential of differentiating them into any specific cell type of interest by manipulating their growth conditions. *In vitro*, ES cells are maintained in the presence of the cytokine leukemia inhibitory factor (LIF) as undifferentiated populations [1] that are capable of differentiating into various cell types. Upon removal of LIF, the ES cells can be easily induced to differentiate into spheroid cell aggregates termed embryoid bodies, recapitulating early developmental processes. These ES cell capabilities require the selective activation and repression of many genes or networks of regulatory genes [2,3]. Our current understanding of the genes that regulate the self-renewal and differentiation of stem cells is far from complete, and identifying these genes is critical for understanding the molecular basis of cell commitment. In addition, several genes known to play a role in human leukemogenesis have been identified in mouse ES cells, including *Notch1*, *Flt3*, *Lmo2* and *Nf1* [4]. Thus, identifying the genes that govern ES cells would provide new insights into proteins whose function was previously unknown.

Loss-of-function genetic screening is a powerful method for identifying the genes that are sufficient to confer a particular phenotype. A traditional approach for generating loss-of-function gene mutations has been the use of N-ethyl-N-nitrosourea (ENU) as a mutagen, but finding the causal point mutation in such systems is difficult and time consuming. Recently, RNAi has been shown to be effective in downregulating gene function; however, this strategy is labor intensive and is not a cost-effective approach. Alternatively, insertional mutagenesis by a retroviral vector derived from murine leukemia virus has proved to be a powerful gene trap method for generating loss-of-function mutations in ES cells but requires the generation of adult mice for functional analysis [5]. Several groups have recently used DNA transposons, such as *Sleeping Beauty (SB)* and *Drosophila PiggyBac*, for gene trap methods [6,7]. However, one limitation of the currently available *SB* transposons is the necessity of co-delivering the *SB* transposon with transposase-encoding DNA for the gene integration event. In addition, low transfer efficiency and a lack of sustained transposase expression are problems that have been reported to occur during cell culture [8,9]. Although the *SB* transposon is commonly used, other DNA transposons are currently being developed for gene discovery approaches. In this study, we explored the possibility of using long interspersed nuclear element 1 (LINE-1 or

L1 retrotransposon) as a tool to deliver a gene trap in ES cells because of its efficient disruption of gene function.

The L1 retrotransposon is an insertional mutagen that is capable of inserting its sequence into a gene and disrupting the gene's function in individual cells or whole animals. L1 offers potential advantages over the *SB* and *PiggyBac* transposons, as L1 performs the random disruption of genes at a high frequency of insertion throughout the genome [10,11]. In addition, L1 mobilizes itself to a new genomic location by a 'copy and paste' mechanism, which offers an infinite source of insertional mutagens for efficient gene knockout throughout the genome. Furthermore, the L1 insertion is stable and permanent in all of the progeny of integrated cells, and the inserted sequence itself serves as a molecular tag to identify the disrupted gene [12]. These inherent features of L1 make it a valuable tool in ES cell gene discovery applications. We have created an episomal, nonviral L1 retrotransposition system using the scaffold/matrix attachment regions (S/MARs) in the vector backbone [13] and evaluated its utility in identifying genes in mouse ES cells.

Here, we demonstrate that, by utilizing this vector and coupled with GFP expression, we have successfully isolated 4 individual ES cell clones out of 50 clones screened that display disrupted genes, including one novel gene. We confirmed the identity of all of these genes by an inverse PCR method and verified their function in cell differentiation using undifferentiated markers of ES cells. The ease of using this insertional mutagenesis approach and the simplicity of identifying cells with disrupted genes by GFP expression make this L1 system a potential tool for ES cell gene discovery. Similar to other systems, this vector can also be applied to different kinds of stem cells or cancer stem cells to identify the genes that are responsive to certain cell growth conditions or involved in cancer development.

Materials and Methods

L1 expression vector

The construction of the S/MAR-based L1 retrotransposition vector has been described previously [13]. The vector DNA was amplified in *E.coli* DH10B cells and isolated using an endotoxin-free Maxi prep kit (Qiagen).

ES cell culture and transfection

Mouse ES cells were grown on mitomycin C-treated embryonic fibroblast (MEF) cell layers and were cultured in complete knockout Dulbecco's modified Eagle's medium (Invitrogen) with 4.5 mg/ml glucose, 2 mM L-glutamine, 0.1 mM β-mercaptoethanol, 15% FCS and 1000 U/ml leukemia inhibitory factor (LIF). The same growth medium (without LIF and MEFs) was used to differentiate the ES cells. Fresh glutamine was added when the medium was replaced daily. ES cell transfection was performed using the Amaxa ES nucleofector kit. Briefly, 2×10^5 ES cells were suspended in 90 μl Nucleofection solution and 10 μl Nucleofection solution containing 5 μg L1 vector added. This mixture was transferred to an electroporation cuvette and electroporated using the A-23 program of the Nucleofector 1 device (Amaxa Biosystems). Approximately 500 μl of the pre-warmed medium was added to the cuvette, and the ES cells were transferred to a 10-cm culture plate containing an MEF feeder layer. Neomycin (G418) was added to the medium 24 h after transfection at a final concentration of 175 μg/ml. The ES cells were subjected to neomycin selection for 7 to 8 days until colonies approximately 1 mm in diameter appeared.

GFP expression of ES clones

Using fluorescence microscopy, G418-resistant GFP-positive colonies were individually picked, partially digested with trypsin and transferred to a 24-well plate seeded with MEF feeder layers. After growing for 3 days, half of the ES cells derived from each clone were frozen as a stock, and the remainder was split into two wells of 24-well plates. One portion of the cells was grown in the medium with LIF and MEFs, and the other was grown without LIF and MEFs for 5 to 6 days until the formation of embryoid bodies. The ES clones that failed to differentiate in the absence of LIF and MEFs were harvested for further analysis.

PCR analysis of ES cells with retrotransposon insertions

Genomic DNA was isolated from each candidate ES clone using a QIAamp DNA kit (Qiagen), and PCR was performed using the Geno-5' (5'-TATTGCCGATCCCCTCAGAAGA-3') and Geno-3' (5'-CAAGGACGACGGCAACTACAAG-3') primers to examine whether GFP expression was a result of retrotransposed L1 insertion. The amplification was performed in 50-μl reactions containing 5 μl of 10X PCR buffer with 2.5 mM $MgCl_2$, 1 μl of 10 mM dNTPs, 5 μM each primer and 250 ng genomic DNA. After an initial step at 95°C for 10 min, 30 cycles of amplification were performed (95°C for 30 s, 58°C for 15 s and 72°C for 2 min), followed by a final step at 72°C for 8 min. The amplified products were visualized on a 1.2% agarose gel. Genomic DNA from the wild-type ES cells and vector DNA were used as negative and positive controls, respectively.

Identification of the integration site by Inverse PCR

Approximately 1 μg genomic DNA was subjected to *Ssp*I restriction enzyme digestion at 37°C for 4 h. After heat inactivation, the digested DNA was treated with 20 units of T4 DNA ligase (NEB) for an overnight self ligation. About 250 ng of the ligation mixture was used directly as the DNA template for the inverse PCR reaction. The first round of the inverse PCR was performed using the following primers for the GFP expression: 5'-CTTGAAGAAGATGGTGCG-3' and 5'-ACAACCACTACCTGAGCACC-3'. We used a 5-μl aliquot from the first round of PCR as the template for the second round of nested PCR using the primers, 5'-TTGAAGAAGTCGTGCTGC-3' and 5'-AAAGACCCCAACGAGAAGCG-3'. The product was cloned directly into the Topo-XL (Invitrogen) vector and sequenced. The location of the L1 integration sites and disrupted genes were identified using the UCSC mouse genome analysis.

qRT-PCR analysis

Total RNA was isolated using the RNeasy Mini Kit (Qiagen), and 1 μg RNA was reverse transcribed using SuperScript III Reverse transcriptase (Invitrogen) with an oligo(dT)$_{18}$ primer. qRCRs were performed using SYBR Green with 200 nM for forward and reverse primers with the PRISM7700 Sequence Detection System. The primers used were as follows: for Tollip, forward 5'- GACGAGTG-GTATAGCCTGAGTG-3', reverse 5'- ACACAGTTGGCATCAG-GACCAC-3'; for Oct4 forward 5'- AACCTTCAGGAGATATG-CAAATCG-3', reverse 5'- TTCTCAATGCTAGTTCGCTTTCTCT-3'; for Nanog forward 5' TCTTCCTGGTCCCCACAGTTT-3', reverse 5'-GCAAGAATAGTTCTCGGGATGAA-3'; for GAPDH forward 5'-AGGTCGGTGTGAACGGATTTG-3', reverse 5'-GGGGTCGTT-GATGGCAACA-3'. The measured transcript levels were normalized to GAPDH. The samples were amplified in triplicates.

siRNA transfection and validation of mutations

Tollip-specific shRNAs (TGGACTCGTTCTACCTTGA, CTCT-GCCAAGATACTAGAA, and ATGGTGGTAACAAGCACTT in the pSiLv-mU6 vector were purchased from GeneCopoeia (product ID:MSH033160). Each shRNA was transfected independently into wild-type mouse ES cells using an Amaxa Nucleofector kit as described above. The depletion of Tollip mRNAs was confirmed by qRT-PCR. The expression of the mCherryFP gene, which is present in the vector, was used as a marker to identify the transfected ES colonies, which were transferred to 24-well plates and grown in medium with or without LIF and MEFs for 5 days. Differentiated and undifferentiated ES cells were subsequently subjected to immunofluorescence staining of Nanog, Oct4, and SSEA-1 pluripotency antibodies, as described by the manufacturer's instructions (Abcam).

Results and Discussion

Construction of an episome vector expressing the L1 retrotransposon that allows GFP expression and G418 drug selection
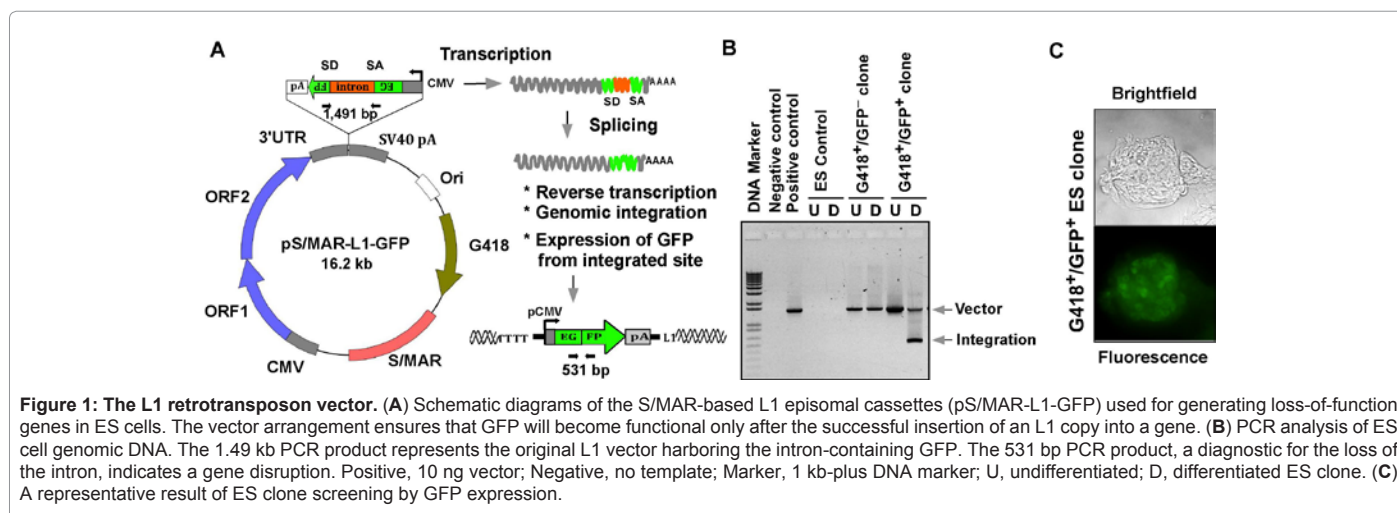
As an insertional mutagenesis system, we constructed an S/MAR-based L1 retrotransposon vector that consisted of a full-length active human L1 gene ($L1_{RP}$) under the control of the cytomegalovirus (CMV) immediate early promoter, which has been reported to be transcriptionally active in a variety of undifferentiated mouse ES cells [14]. The 3'-UTR of the L1 gene harbors a visual GFP marker disrupted by 960bp of a γ-globin intron in an antisense orientation (Figure 1A). GFP is co-transcribed as a single fusion transcript due to the presence of the splicing sites in the intron sequences. This arrangement ensures that GFP expression occurs only after L1 mobilization or insertional mutagenesis i.e., after L1 expression, γ-globin intron splicing, reverse transcription, and integration of the L1 copy into the genomic DNA (Figure 1B). Therefore, no GFP reporter is expressed unless the newly synthesized L1 is integrated into a new genomic location, and observation of the expression of GFP under a UV light would allow us to detect a real-time L1 disruptional event in living ES cells without cell staining (Figure 1C). In addition, this vector also contains a neomycin resistance gene that can be used to select transfected ES clones. To ensure the stability and the integrity of this vector, a DNA fragment containing scaffold/matrix attachment regions (S/MARs) was added to the 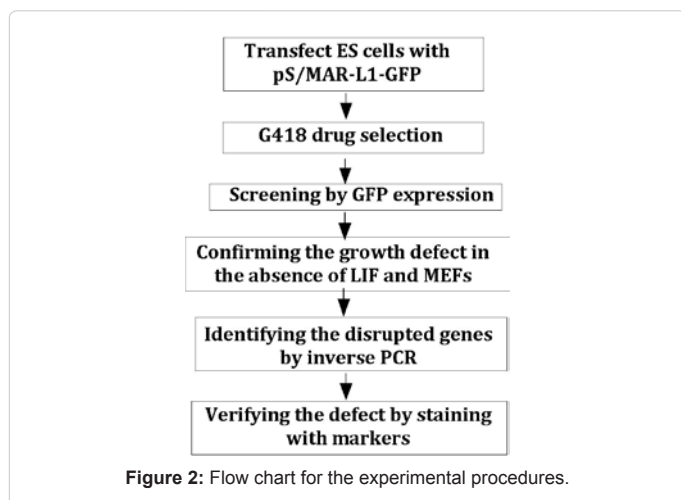vector backbone. We have recently shown that the inclusion of S/MARs effectively maintains a single copy of the vector in cells throughout multiple rounds of cell division without undergoing epigenetic silencing by DNA methylation or a loss of the vector, even in the absence of G418 drug selection [13]. We introduced the S/MAR-L1-GFP vector into mouse ES cells using nucleofection and, as determined by a fluorescence-activated cell sorting (FACS) analysis, achieved 2.4 ± 0.26% GFP-positive cells, compared to 1.7 ± 0.12% GFP-positive cells using an EBNA-based RP99-L1-GFP vector carrying the same L1-GFP expression cassette. This high level of retrotransposition from S/MAR-L1-GFP in ES cells is consistent with our previous studies [13], which demonstrated that the S/MAR-based L1 vector system increases the L1 retrotransposition events during the culturing of human somatic cells.

Selection of disrupted genes in ES clones that fail to differentiate

Undifferentiated ES cells contain reduced levels of global DNA methylation and active or open chromatin structures [15]. Under these conditions, the expression of the L1 retrotransposon results in the random insertion of a copy into the host genome. When L1 inserts into a gene, the protein encoded by that gene should be truncated and its function disrupted [16]. If the disrupted gene is essential for self-renewal and differentiation, then the ES cells may remain in an undifferentiated state even when subjected to differentiation-inducing conditions, such as the absence of LIF. Interestingly, L1 is rarely expressed in terminally differentiated cells. This approach should allow us to discover the potential genes that regulate ES cell renewal and differentiation.

To test this approach, we transfected mouse ES cells derived from C57BL/6 mice; the approach used to isolate the candidate gene in the ES cells is shown in Figure 2. An S/MAR-based L1 retrotransposition (pS/MAR-L1-GFP) was transfected into ES cells using the Amaxa ES Nucleofector Kit and was grown on a layer of feeder fibroblast cells (MEFs). The G418-resistant clonies were selected with 175 μg/ml neomycin for 7 days. Using fluorescence microscopy, each G418-resistant GFP-positive colony was picked individually and transferred to a 24-well plate seeded with MEF feeder layers. It should be emphasized that the GFP expression from the L1 vector occurs only after the successful insertion of the L1 copy into a new genomic location. In this study, we screened approximately 50 ES cell clones that displayed resistance to G418 in a single 10-cm plate. Approximately 85% of the G418-resistant clones we examined did not show any GFP

Figure 1: The L1 retrotransposon vector. (**A**) Schematic diagrams of the S/MAR-based L1 episomal cassettes (pS/MAR-L1-GFP) used for generating loss-of-function genes in ES cells. The vector arrangement ensures that GFP will become functional only after the successful insertion of an L1 copy into a gene. (**B**) PCR analysis of ES cell genomic DNA. The 1.49 kb PCR product represents the original L1 vector harboring the intron-containing GFP. The 531 bp PCR product, a diagnostic for the loss of the intron, indicates a gene disruption. Positive, 10 ng vector; Negative, no template; Marker, 1 kb-plus DNA marker; U, undifferentiated; D, differentiated ES clone. (**C**) A representative result of ES clone screening by GFP expression.

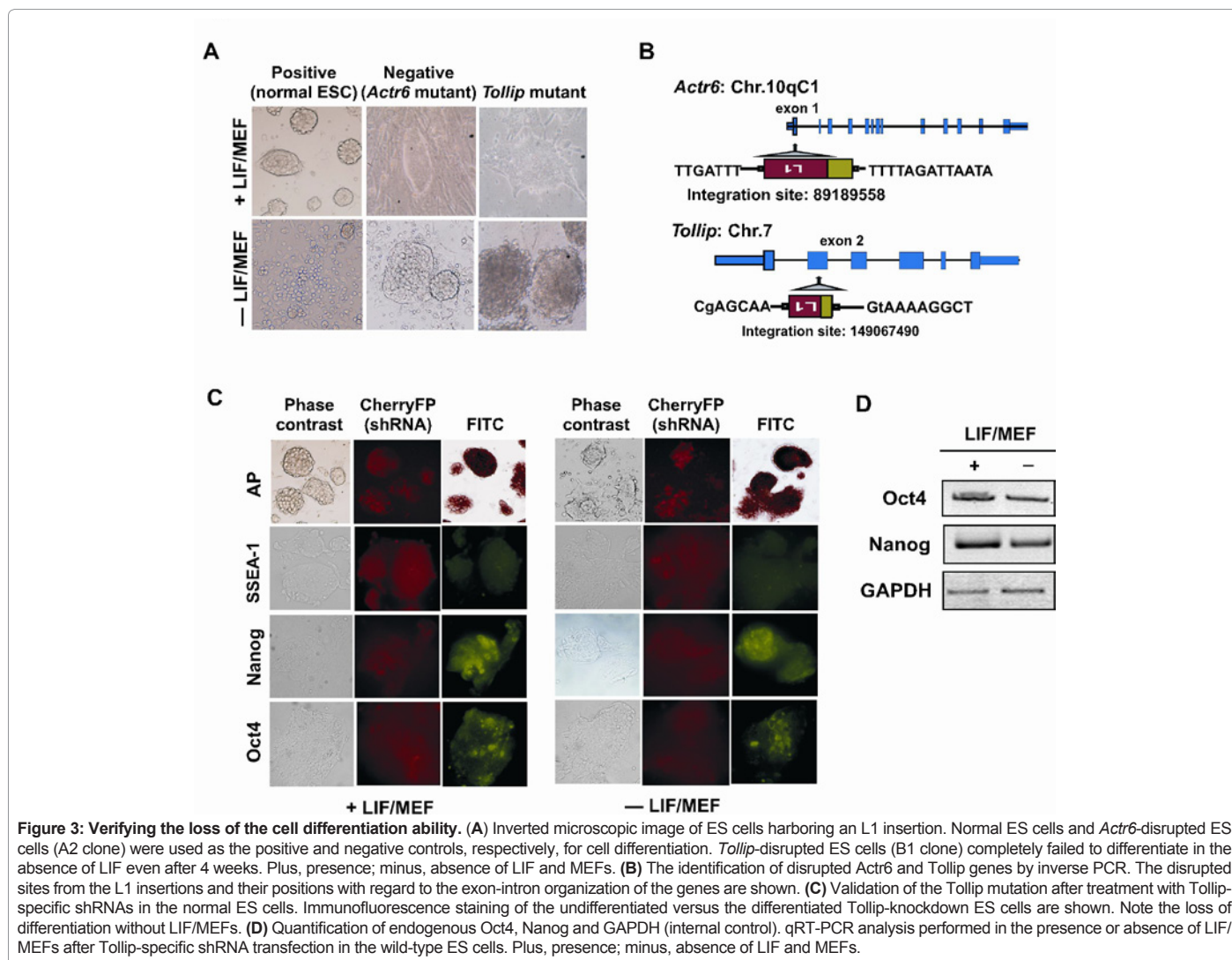**Figure 2:** Flow chart for the experimental procedures.

fluorescence while maintaining the L1 vector; thus, not all of the ES cells undergo the insertion of L1 integration event. This suggests that the efficiency of the S/MAR-based insertional mutagenesis in ES cells is relative low in comparison to standard retroviral-based insertional

vector approaches in which at least 10 insertional colonies have been reported in a single 10-cm dish [17]. In total, we identified only 9 GFP-positive, G418-resistant clones in a 10-cm plate. Each GFP-positive colony was maintained as an independent clone in culturing MEFs in replicates where one set was cryopreserved and the others were used to determine whether the cells retained the ability to differentiate. ES cells undergo *in vitro* cell differentiation in the absence of LIF and MEFs.

The selected GFP clones were grown in the absence of LIF and MEF feeder cells for 5 to 6 days, either in adherent monoculture or in suspension culture, for the formation of embryoid bodies (EBs). If a gene putatively involved in differentiation is disrupted by the L1 insertion, these cells will lose the ability to differentiate into the EBs. For comparison, untransformed wild-type ES cells or pEGFP-N1-expressing ES cells were used as the controls. The failure of the ES clones to differentiate was simply assessed by a morphological analysis using an inverted phase contrast microscope, thus narrowing down the number of potential ES clones into a manageable number for further genetic and phenotypic analysis. In the absence of LIF and MEFs, we found that 4 out of 9 individual GFP-positive cells remained undifferentiated. In contrast, the other five GFP-positive cells became differentiated within 4 days, similar to the wild-type ES cells without LIF and MEFs; we reasoned that these cases might be

Figure 3: Verifying the loss of the cell differentiation ability. (A) Inverted microscopic image of ES cells harboring an L1 insertion. Normal ES cells and *Actr6*-disrupted ES cells (A2 clone) were used as the positive and negative controls, respectively, for cell differentiation. *Tollip*-disrupted ES cells (B1 clone) completely failed to differentiate in the absence of LIF even after 4 weeks. Plus, presence; minus, absence of LIF and MEFs. **(B)** The identification of disrupted Actr6 and Tollip genes by inverse PCR. The disrupted sites from the L1 insertions and their positions with regard to the exon-intron organization of the genes are shown. **(C)** Validation of the Tollip mutation after treatment with Tollip-specific shRNAs in the normal ES cells. Immunofluorescence staining of the undifferentiated versus the differentiated Tollip-knockdown ES cells are shown. Note the loss of differentiation without LIF/MEFs. **(D)** Quantification of endogenous Oct4, Nanog and GAPDH (internal control). qRT-PCR analysis performed in the presence or absence of LIF/MEFs after Tollip-specific shRNA transfection in the wild-type ES cells. Plus, presence; minus, absence of LIF and MEFs.

due to the insertion of L1 into genes or noncoding intergenic regions that are not essential for ES cell differentiation. In fact, analyzing the L1 integration site of one of the GFP-positive ES differentiated cells showed that L1 was located in the intergenic region of chromosome 11 (112835881-112836059), a site with no genes nearby. A blast analysis of this genomic sequence showed that this region contained only a hypothetical protein, LOC72386, which is located 70 kb downstream of the L1 integration site (data not shown).

The intensity of the GFP fluorescence of the ES clone and the degree of cell differentiation upon ES cell differentiation varied among the candidate clones. One of the ES clones (named A8) that showed weak GFP expression was found to exhibit a delayed partial cell differentiation, whereas the other clones (named A2, B1 and B9) completely failed to undergo any cell differentiation even after 4 weeks of culturing in the absence of LIF (Figure 3A). To determine whether the defective ES cell clone was a result of the L1 insertion, we conducted a PCR-based analysis to confirm the presence of spliced GFP in these cells. The genomic DNA isolated from the defective ES clones was subjected to a GFP-specific PCR assay, as shown in Figure 1B. An oligonucleotide primer specific for the GFP gene was used to determine whether the intron had been removed by γ-splicing during the L1 insertional events; the unspliced PCR product is 1491 bp rather than the expected 531 bp in the case of a successful L1 insertion. This simple strategy allowed us to confirm the presence of retrotransposed L1 insertions in the defective ES cell colonies.

Identification of the disrupted candidate genes in the ES clones

To identify the disrupted genes in the ES clone, we used an inverse PCR analysis. Approximately 1 μg genomic DNA was isolated from the candidate ES cell clones and was digested with SspI followed by self-ligation with T4 DNA ligase. The resulting products were subjected to inverse PCR amplification using a primer complementary to the 3'-UTR of the L1 sequence (forward primer) and another primer corresponding to the downstream region of the GFP sequence (reverse primer). As the sequence of human L1 is completely different from the mouse version, the inverse primers should specifically amplify only the integrated L1 insert. The PCR product was subsequently cloned into the TOPO-XL vector for DNA sequencing, which enabled us to identify the flanking sequences of the disrupted genes in the genomic DNA using the UCSC mouse genome browser (Figure 3B). Using this strategy, we identified the L1 integration sites and their corresponding disrupted genes for the candidate ES clones that displayed the loss of ability to differentiate in the absence of LIF and MEFs, as illustrated in Table 1. Although L1 normally results in a single insertion event, we found that it occasionally integrates into multiple sites of the genome.

Two of the identified genes resulting from the L1 insertions are Arp6-actin-related protein (*Actr6*) (A2) and GATA-binding protein 4 (*Gata4*) (B9), which were previously shown to be necessary for ES cell differentiation. The Arp6 actin-related protein is a subunit

of chromatin-remodeling factor, Tip60. Using high-throughput RNAi screening, Tip60 was recently shown to be essential for ES cell development through its functional overlap with Nanog [18]. Depletion of Tip60 was also found to reduce the levels of the cell-cycle regulators and metabolic genes required for ES cell division. Gata4 is one of the well-characterised transcription factors that is expressed in ES cells [19]. The L1 integration site in the third ES clone (A8), which exhibited delayed partial cell differentiation, matched three different sites; one insertion was identified as the intron of the tripartite motif-containing 59 protein (*TRIM59*), and the remaining two sites occurred in the intergenic regions of chromosomes 7 and 11. Because this ES clone (A8) contains more than one L1 integration site, we did not pursue it further because the complexity and heterogeneity of multiple disruptions make them difficult to analyze. Interestingly, the L1 insertion site in the fourth clone (B1) is a candidate gene that has never been reported to play a role in ES cell differentiation. The L1 insertion is located in exon 2 of the Toll-interacting protein (*Tollip*) gene. Although little is known about this gene in ES cell differentiation, Tollip is known to interact with Toll-like receptors (TLR) to inhibit TLR-mediated signaling pathways [20,21]. In addition, Tollip plays a key role in cell-cell signaling, inflammatory cytokine production, and intracellular signaling pathways. Thus, further studies are required to delineate the exact function of this gene in stem cell differentiation. Nevertheless, the data presented here suggest that the L1 retrotransposon system can identify previously unknown genes involved in mouse ES cell differentiation.

Confirming the function of the disrupted genes in ES cell differentiation

To confirm whether ES cell differentiation is indeed dependent on the function of the identified genes, we resorted to the use of shRNAs against each gene in wild-type ES cells, followed by staining with alkaline phosphatase (AP), which is known to be expressed at a high level in the cell membrane of undifferentiated ES cells [22]. As the expression of AP is often used as an indicator of the undifferentiated state of ES cells, we utilized this simple staining approach to validate the undifferentiated state of the ES cell clones. Using wild-type ES cells, we introduced three Tollip-specific prevalidated shRNAs independently into the cells and, as determined by qRT-PCR analysis, achieved 78 ± 6% knockdown efficiency of the Tollip gene compared to the wild-type ES cells (data not shown). The ES cells transfected with each shRNA against the Tollip gene were cultured in the presence and absence of LIF, followed by AP staining. In this assay, the *Actr6* mutant ES cell (clone A2) was used as a known positive gene. The Tollip-knockdown ES cells showed high levels of AP activity in the presence or absence of LIF and MEFs similar to the expression pattern observed with other undifferentiated ES cell markers [23,24], including Oct4, Nanog, and SSEA-1 (Figure 3C). Additionally, we employed a semi-quantitative RT-PCR analysis to evaluate whether the ES cell markers showed similar expression levels in response to the removal of LIF. The expression levels of the Nanog and Oct4 markers remained similar in both the presence and

Trapped ES clone	Accession	L1 integrated site	Gene	Gene product
A2	NM_025914.2	Chr.10qC1: 89189558	Actr6	ARP6 actin-related protein
B9	NM_008092.3	Chr.14qD1:63819210	GATA4	GATA binding protein 4
A8	NM_025863.3	Chr.3qE1: 68838551	TRIM59	Tripartite motif containing 59*
B1	NM_023764.3	Chr.7qF5: 149067490	Tollip	Toll-interacting protein

*contained three L1 integration sites

Table 1: Genes identified by L1 insertional mutagenesis that displayed defect in ES cell differentiation.

absence of LIF, suggesting that the ES cell with defective the Tollip gene do not have the ability to differentiate (Figure 3D). As expected, under the standard *in vitro*-directed cell differentiation conditions, the Tollip-knockdown ES cells failed to undergo cell differentiation in the absence of LIF. Taken together, these results suggest that Tollip plays an essential role in ES cell differentiation.

At present, little is known about the function of Tollip or its regulatory pathways that direct stem cell differentiation. Tollip is known to act in the immune response to invading pathogens by controlling IRAK phosphorylation in the TLR and IL-1R signaling pathways [20]. Tollip is also known to suppress TLR-mediated NF-κB activity. A recent study in mesenchymal stem cells suggests that TLR4, which is the receptor for Tollip, inhibits the activation of transcription factor STAT3 (signal transducer and activator of transcription 3) and thereby exerts deleterious effects on stem cell proliferation [25]. As STAT3 mediates self-renewal and the maintenance of pluripotency in the absence of LIF, it is essential for ES cell differentiation [26]. Several molecules, such as Zfp57, GABP and β-catenin, have been identified as the factors involved in the LIF/STAT3 pathway. Thus, it is tempting to speculate that the disruption of Tollip may inactivate STAT3 through TLR receptors, resulting in the lack of proper ES cell differentiation. However, further studies are required to delineate the exact molecular mechanisms of this gene. Nonetheless, the data presented in this study show that the L1 retrotransposon approach can identify both novel genes and known genes that regulate ES cell differentiation.

Conclusion

Embryonic stem cells have a relatively stable genome and are highly amenable to loss-of-function genetic screens. The recent development of the L1 retrotransposon approach offers an efficient and highly versatile alternative for achieving the complete disruption of gene function. The ease of using this insertional mutagen and the simplicity of identifying the cells with disrupted genes by GFP expression make this L1 vector a promising tool to identify the genes that play roles in cell growth, morphology, differentiation and proliferation. Given that ES cells and other stem cells share many similarities, including the ability to self-renew, pluripotency and virtually identical chromatin states, this system can also be applied to identify critical genes in other stem cells, such as muscle stem cells, pluripotential stem cells or cancer stem cells.

References

1. Niwa H, Burdon T, Chambers I, Smith A (1998) Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev 12: 2048-2060.

2. Wang J, Rao S, Chu J, Shen X, Levasseur DN, et al. (2006) A protein interaction network for pluripotency of embryonic stem cells. Nature 444: 364-368.

3. Walker E, Ohishi M, Davey RE, Zhang W, Cassar PA, et al. (2007) Prediction and testing of novel transcriptional networks regulating embryonic stem cell self-renewal and commitment. Cell Stem Cell 1: 71-86.

4. Akagi K, Suzuki T, Stephens RM, Jenkins NA, Copeland NG (2004) RTCGD: retroviral tagged cancer gene database. Nucleic Acids Res 32: D523-527.

5. Carlson CM, Largaespada DA (2005) Insertional mutagenesis in mice: new perspectives and tools. Nat Rev Genet 6: 568-580.

6. An W, Boeke JD (2005) Transposon technology and vertebrate functional genomics. Genome Biol 6: 361.

7. Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, et al. (2009) piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature 458: 766-770.

8. Ikeda R, Kokubu C, Yusa K, Keng VW, Horie K, et al. (2007) Sleeping beauty transposase has an affinity for heterochromatin conformation. Mol Cell Biol 27: 1665-1676.

9. Jia F, Wilson KD, Sun N, Gupta DM, Huang M, et al. (2010) A nonviral minicircle vector for deriving human iPS cells. Nat Methods 7: 197-199.

10. Ostertag EM, Madison BB, Kano H (2007) Mutagenesis in rodents using the L1 retrotransposon. Genome Biol 8: S16.

11. Collier LS, Largaespada DA (2007) Transposons for cancer gene discovery: Sleeping Beauty and beyond. Genome Biol 8: S15.

12. Babushok DV, Ostertag EM, Courtney CE, Choi JM, Kazazian HH Jr (2006) L1 integration in a transgenic mouse model. Genome Res 16: 240-250.

13. Rangasamy D (2010) An S/MAR-based L1 retrotransposition cassette mediates sustained levels of insertional mutagenesis without suffering from epigenetic silencing of DNA methylation. Epigenetics 5: 601-611.

14. Ward CM, Stern PL (2002) The human cytomegalovirus immediate-early promoter is transcriptionally active in undifferentiated mouse embryonic stem cells. Stem Cells 20: 472-475.

15. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, et al. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125: 315-326.

16. An W, Han JS, Wheelan SJ, Davis ES, Coombes CE, et al. (2006) Active retrotransposition by a synthetic L1 element in mice. Proc Natl Acad Sci USA 103: 18662-18667.

17. Li X, Leder P (2007) Identifying genes preferentially expressed in undifferentiated embryonic stem cells. BMC Cell Biol 8: 37.

18. Fazzio TG, Huff JT, Panning B (2008) Chromatin regulation Tip(60)s the balance in embryonic stem cell self-renewal. Cell cycle 7: 3302-3306.

19. Soudais C, Bielinska M, Heikinheimo M, MacArthur CA, Narita N, et al. (1995) Targeted mutagenesis of the transcription factor GATA-4 gene in mouse embryonic stem cells disrupts visceral endoderm differentiation in vitro. Development 121: 3877-3888.

20. Burns K, Clatworthy J, Martin L, Martinon F, Plumpton C, et al. (2000) Tollip, a new component of the IL-1RI pathway, links IRAK to the IL-1 receptor. Nat Cell Biol 2: 346-351.

21. Zhang G, Ghosh S (2002) Negative regulation of toll-like receptor-mediated signaling by Tollip. J Biol Chem 277: 7059-7065.

22. Thomson JA, Kalishman J, Golos TG, Durning M, Harris CP, et al. (1995) Isolation of a primate embryonic stem cell line. Proc Natl Acad Sci USA 92: 7844-7848.

23. Park IH, Zhao R, West JA, Yabuuchi A, Huo H, et al. (2008) Reprogramming of human somatic cells to pluripotency with defined factors. Nature 451: 141-146.

24. Toyooka Y, Shimosato D, Murakami K, Takahashi K, Niwa H (2008) Identification and characterization of subpopulations in undifferentiated ES cell culture. Development 135: 909-918.

25. Wang Y, Abarbanell AM, Herrmann JL, Weil BR, Manukyan MC, et al. (2010) TLR4 inhibits mesenchymal stem cell (MSC) STAT3 activation and thereby exerts deleterious effects on MSC-mediated cardioprotection. PLoS One 5: e14206.

26. Matsuda T, Nakamura T, Nakao K, Arai T, Katsuki M, et al. (1999) STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. EMBO J 18: 4261-4269.