Published in final edited form as:

Gene. 2010 May 15; 456(1-2): 15–23. doi:10.1016/j.gene.2010.02.005.

Repeat-induced gene silencing of L1 transgenes is correlated with differential promoter methylation

James M. Rosser and Wenfeng An*

School of Molecular Biosciences, Washington State University, Pullman, WA 99164, USA; Center for Reproductive Biology, Washington State University, Pullman, WA 99164, USA

Abstract

Recent transgenic studies on L1 retrotransposons have afforded exciting insights into L1 biology, and a unique opportunity to model their function and regulation *in vivo*. Thus far, the majority of the transgenic L1 mouse lines are constructed via pronuclear microinjection, a procedure that typically results in the integration of tandem arrayed transgenes. Transgene arrays are susceptible to repeat-induced gene silencing (RIGS) in both plants and animals. In order to examine the potential impact of RIGS on L1 retrotransposition, we derived a cohort of animals carrying reduced copies of *ORFeus* transgene at the same genomic locus by Cre-mediated recombination. The copy number reduction of *ORFeus* transgenes did not decrease the overall retrotransposition activity. Using a sensitive and reproducible quantitative PCR assay, an average frequency of 0.45 insertions per cell was observed for animals carrying the donor transgene at a single copy, representing a 9-fold increase of retrotransposition frequency on a per-copy basis. DNA methylation analyses revealed that the observed retrotransposition activity was correlated with differential CpG methylation at the heterologous promoter: the promoter region was largely methylated in animals with the high-copy array but significantly hypomethylated in animals with the single-copy counterpart. In contrast, the ORF2 region, which represents the body of the *ORFeus* transgene, and the 3' end of the transgene showed high level of methylation in both high-copy and single-copy samples. The observed methylation patterns were metastable across generations. In summary, our data suggest that tandem arrayed L1 transgenes are subject to RIGS, and transgenes present at a single copy in the genome are thus recommended for modeling L1 in animals.

Keywords

LINE-1; Retrotransposon; Transgenic mouse; DNA methylation; Copy number

1. Introduction

Transposable elements of plants and animals constitute a major driving force for their genome evolution, and have become indispensable for genome engineering (Kazazian, 2004; Mates et al., 2007). Based on the mode of transposition, transposable elements are classified as either DNA transposons or retrotransposons. DNA transposons mobilize by a “cut-and-paste” mechanism with a minimal gain of transposon copy number. Two DNA

© 2010 Elsevier B.V. All rights reserved.

*Corresponding author. School of Molecular Biosciences, Washington State University, PO Box 647520, Pullman, WA 99164-7520, USA. Tel.: +1 509 335 8296; fax: +1 509 335 4159. wenfeng_an@wsu.edu (W. An)..

Appendix A. Supplementary data

Supplementary data associated with this article, can be found in the online version, at doi:10.1016/j.gene.2010.02.005.

transposons derived from heterologous species, *Sleeping Beauty* and *piggyBac*, have recently been adapted for insertional mutagenesis in mice (Ding et al., 2005; Collier and Largaespada, 2007; Takeda et al., 2007). Retrotransposons, including LTR and non-LTR retrotransposons, are able to increase their copy number through an RNA intermediate by a “copy-and-paste” mode of replication. Long interspersed elements type 1 (LINE-1 or L1) are autonomous non-LTR retrotransposons, present at extremely high-copy numbers in most mammalian genomes. Both human and mouse genomes harbor more than a half million copies of L1, but the vast majority of these L1 sequences are deficient in remobilization due to 5' truncation and internal mutations. A functional full-length L1 contains an internal Pol II promoter at its 5' UTR, encodes two open reading frames (ORF1 and ORF2), and carries a polyadenylation signal in its 3'UTR. The multiplicative nature of L1 retrotransposition makes it an attractive alternative to DNA transposons for genome manipulation in model vertebrate animals such as mice and rats (Bestor, 2005).

The potential of L1-mediated genome manipulation has been explored by transgenesis (Ostertag et al., 2002; Prak et al., 2003; Muotri et al., 2005, 2009; An et al., 2006, 2008; Babushok et al., 2006; Kano et al., 2009). In these studies, L1 elements are engineered to carry a utility cassette (retrotransposition reporters or gene trapping elements) and introduced into mice as donor transgenes. Several L1 elements are featured, including two human L1 isolates, RP (Ostertag et al., 2002; Prak et al., 2003; Muotri et al., 2005, 2009; Kano et al., 2009) and LRE3 (Babushok et al., 2006; Kano et al., 2009), and the synthetic mouse L1 element *ORFeus* (An et al., 2006, 2008). Endogenous L1s are known to be expressed in germ cells (Branciforte and Martin, 1994; Trelogan and Martin, 1995; Ergun et al., 2004). Similarly, in adult transgenic animals carrying human L1 transgenes under the regulation of their 5'UTR promoter, L1 transgene transcripts are readily detected by RT-PCR in both male and female gonads but rarely in any other tissues (Ostertag et al., 2002; Kano et al., 2009). In addition, abundant L1 transgene transcripts are detected not only from donor-positive embryos but also from donor-negative, preimplantation embryos, potentially through RNA carryover during meiosis (Kano et al., 2009). However, despite prevalent RNA expression in both germ cells and early embryos, retrotransposition events from human L1 transgenes are largely confined to somatic tissues (Kano et al., 2009). In particular, studies with human L1RP transgenes indicate that these transgenes can retrotranspose in neural progenitor cells during both embryonic and adult neurogenesis (Muotri et al., 2005, 2009). L1 elements have also been placed under the regulation of heterologous, constitutively expressing promoters (Ostertag et al., 2002; Prak et al., 2003; An et al., 2006, 2008). Accordingly, L1 transgene transcripts are present in all tissues (Ostertag et al., 2002) and retrotransposition can be detected in both mouse germline and somatic tissues (An et al., 2006, 2008).

Thus far, the majority of the transgenic L1 mouse lines are constructed via pronuclear microinjection, a procedure that typically results in the integration of multiple-copy transgenes as tandem arrays at single sites (Palmiter and Brinster, 1986; Bishop and Smith, 1989). Independent mouse lines made with the same transgene frequently differ in expression based on their location in the genome. Such position effects on transgene expression can complicate the interpretation of transgenic studies (Martin and Whitelaw, 1996; Dobie et al., 1997). One subcategory of position effects involves the observation that the activity of a transgene is not proportional to the number of transgene copies at a discrete integration site. This phenomenon, termed repeat-induced gene silencing (RIGS), has been demonstrated by varying the transgene copy number at a given chromosomal locus in multiple species, including *Arabidopsis* (Assaad et al., 1993; Ye and Signer, 1996), *Drosophila* (Dorer and Henikoff, 1994; Sabl and Henikoff, 1996) and mice (Garrick et al., 1998). The silencing of tandem arrayed transgenes appears to be intrinsic to the array and is not attributable to position effects of nearby sequences (Henikoff, 1998). Local

heterochromatin formation is thought to be responsible for RIGS on tandem arrayed transgenes as the reduction in transgene copy number is accompanied by increased steady-state mRNA (Assaad et al., 1993; Ye and Signer, 1996; Garrick et al., 1998), higher chromatin accessibility (Ye and Signer, 1996; Garrick et al., 1998) and decreased cytosine methylation (Assaad et al., 1993; Garrick et al., 1998). The identification of DNA methyltransferase 1 and chromatin-remodeling enzymes in a recent genome-wide screen for modifiers of transgene variegation further supports the role of DNA methylation and heterochromatin formation in silencing tandem arrayed transgenes (Ashe et al., 2008).

The transgene copy number is unknown for most of the L1 mouse models (Ostertag et al., 2002; Prak et al., 2003; Muotri et al., 2005, 2009; Babushok et al., 2006; Kano et al., 2009), and the effect of RIGS on L1 retrotransposition in these mouse models has yet to be determined. Thus far, the highest retrotransposition activity is seen in animals carrying a tandem array of *ORFeus* transgenes under the regulation of a heterologous constitutive promoter (An et al., 2006). In that study, somatic retrotransposition was detected in all donor-positive animals and the germline retrotransposition frequency was around 1 in every 3 germ cells (An et al., 2006). To directly address the potential effect of RIGS in L1 transgene activity, here we derived a cohort of animals carrying reduced copies of *ORFeus* transgene at the same genomic locus by Cre-mediated recombination. We found that animals carrying the single-copy donor transgene displayed a slightly higher overall activity than the parental high-copy animals. We further demonstrated that retrotransposition activity was correlated with differential CpG methylation at the promoter region. The implications of these findings on L1 transgenic studies were discussed.

2. Methods

2.1. Transgenic mice

All animal protocols have been approved by the Institutional Animal Care and Use Committee. Parental high-copy *ORFeus* transgenic mice were described previously (An et al., 2006). CAG-*cre* mice (Sakai and Miyazaki, 1997) were provided by Charles Hawkins; wild-type C57BL/6J mice were acquired from Jackson Laboratories (Maine, USA). Genomic DNA was extracted from various mouse tissues using the Puregene genomic DNA purification kit (Qiagen). Tail samples were collected at weaning (3–5 weeks of age) with the exception of those shown in Fig. 4. For Fig. 4, tissues (including tail) were taken from animals of 13–15 months of age. The genetic background of each mouse line is described in the Discussion section. Interindividual genetic variation at different generations was calculated according to Lee Silver's Mouse Genetics (<http://www.informatics.jax.org/silver/>).

2.2. Plasmids

Plasmid dilution series were used to measure real-time quantitative PCR (qPCR) dynamic range. Plasmid pRSVGFPuvINT contained the retrotransposition reporter cassette in a pBluescriptSK(–) (Stratagene) backbone (An et al., 2008), and was used in donor copy number qPCR. Plasmid pRSVGFPuv (a gift from Jeffery Han) was used in retrotransposition qPCR. pRSVGFPuv was similar to pRSVGFPuvINT except it carried an intact GFP coding sequence, not interrupted by an antisense intron.

2.3. Quantitative PCR

Donor transgene copy number and retrotransposition frequency (i.e., the number of insertions per cell) were determined by qPCRs against the human γ -globin intron and *gfp* exon-exon junction, respectively. GAPDH was amplified in separate wells as an endogenous control. A minimum of three technical replicates was conducted against each target for each

sample. All gDNA samples were diluted to a working concentration of 1 ng/μl. Plasmid standard series containing 5, 50, 500, or 5000 copies of linearized plasmids per microliter were prepared with 1 ng/μl wt gDNA. Two microliters of template gDNA was amplified in a final volume of 20 μl with TaqMan Gene Expression Master Mix on a 7500 Fast Real Time PCR cycler (Applied Biosystems). The PCR cycling parameters consist of 50°C for 20 s, 95 °C for 10 min, and 40 cycles of 94 °C for 15 s and 60 °C for 60 s. The following primers and probes were used: GAPDH forward CAAATTCAACGGCACAGTCAAG, GAPDH reverse CGCTCCTGGAAGATGGTGAT, GAPDH probe AATGGGAAGCTTGTCATC; GFP forward GTTGTCTGGGCAGCAGCAC, GFP reverse AACTTCAAGATCCGCCACAACAT, GFP probe CCTTCTATGAATGTTACTGTTTT; and intron forward ATGTTTCAGCACTGTTGCCTTTAG, intron reverse CCTGCTGTGCTCAGATCAATACTC, intron probe CTCGAGGCAACTTAGACAA. qPCR data were analyzed by the comparative C_T method (Schmittgen and Livak, 2008), and reported as values normalized to calibrator gDNA samples in which the quantity of PCR target had been independently determined. The donor copy qPCR was calibrated to a gDNA sample from mouse line 2G6, which carried a single-copy *ORFeus^L* transgene (An et al., 2008). The retrotransposition qPCR was calibrated to a gDNA sample carrying a single germline insertion (B1045), and further verified by a group of gDNA samples carrying varied number of germline and somatic insertions. The plasmid standard series were used to measure the sensitivity and dynamic range of each qPCR assay. For both qPCR assays, reproducible qPCR signals could be obtained from 10 copies of plasmid DNA. Signals from plasmid dilution series displayed strong linear correlation (the coefficient of determination, $R^2 = 0.9999$ for donor qPCR, and $R^2 = 0.9997$ for retrotransposition qPCR). However, they were not used to interpolate the absolute quantity of PCR target in gDNA samples because of the observed subtle differences in amplification efficiency between calibrator gDNA samples and plasmid standards (explained below). The mass of a diploid mouse genome is estimated to be 6 pg (assuming each haploid mouse genome is about 2.7 billion base pairs in size). Thus, the 2 ng gDNA in each qPCR reaction was derived from approximately 333 cells. In donor qPCR, the calibrator sample 2G6 was set to 1 (i.e., this is the experimental signal when ~333 copies of *genomic* donor target served as the template), and the normalized signal for 100 copies of pRSVGFPuvINT plasmid was 0.23 (i.e., this translates into a relative signal of 0.77 for 333 copies of *plasmid-borne* donor target). In retrotransposition qPCR, the calibrator sample B1045 was set to 1, and the normalized signal for 100 copies of pRSVGFPuv plasmid was 0.46 (i.e., this translates into a relative signal of 1.53 for 333 copies of *plasmid-borne* retrotransposition target). In both assays, the difference in signal strength between the calibrator and the corresponding plasmid standard was within 2-fold (a calibrator/plasmid signal ratio of 0.77 for donor qPCR and 1.53 for retrotransposition qPCR), which may reflect some difference in amplification efficiency for gDNA and plasmid targets. Detailed instructions for preparing plasmid standard series and calculating gDNA mass can be found online (http://www.appliedbiosystems.com/support/tutorials/pdf/quant_pcr.pdf).

2.4. Bisulfite sequencing

gDNA was processed with an EpiTect Bisulfite Kit (Qiagen). The following primers were used for target amplification with ExTaq HS polymerase (Takara): CAG forward CCTAACATTATACCCAATACATAACCTT, CAG reverse GGGGYGAAGGTAAYGTAGYGATTTT, ORF2 forward TTATTTGAAGGTTTTGGAGAAGAAG, ORF2 reverse TTCAACAAATCCACCTAATCCTAAT, GFP forward TTTGGAGGTTAGGTTGGAGTTT, and GFP reverse AACCATTCACCACATTACTAACAAC. The PCR cycling parameters consist of an initial denaturation at 94 °C for 2 min, 10 cycles of denaturing at 94 °C for 15 s, annealing at 60 °C

(−0.5 °C/cycle) for 30 s and 72 °C for 60 s, 30 cycles of 94 °C for 15 s, 55 °C for 30 s and 72 °C for 60 s, and a final extension step at 72 °C for 7 min. An aliquot of PCR product was resolved on a 1% agarose gel, and, if necessary, was gel purified by a QIAquick Gel Extraction Kit (Qiagen). Amplicons were TA-cloned using StrataClone PCR Cloning Kit (Stratagene). Competent cells were transformed with DNA mix, and plated without prior amplification to minimize the chance of analyzing sibling clones. White colonies were randomly selected for bidirectional sequencing analysis using universal vector primers. The efficiency of sodium bisulfite conversion, DNA methylation rate and standard error between bisulfite sequences for each target region were calculated using the Quantification Tool for Methylation Analysis software (Kumaki et al., 2008). The difference in DNA methylation rates between samples was evaluated by two-tailed Mann–Whitney *U*-test, which compares the difference of the average methylation level between two groups, regardless of the CpG methylation pattern.

3. Results

3.1. Generation of mouse lines with reduced copy number of *ORFeus* transgenes

Previously, we characterized a transgenic mouse line carrying a highly active *ORFeus* transgene as a tandem array (An et al., 2006) (Fig. 1A). Each unit copy of this transgene contains a single loxP site between the heterologous CAG promoter and L1 coding sequences (Fig. 1B). To compare the effect of donor copy number on retrotransposition activity, we generated mice carrying reduced copy number of *ORFeus* transgenes by Cre-mediated recombination (Fig. 1D). Transgenic G0 mice were first bred to a Cre-expressing mouse line, CAG-*cre* (Sakai and Miyazaki, 1997). The *cre* recombinase gene in this mouse line is regulated by the same ubiquitous CAG promoter as in the *ORFeus* transgene. In principle, Cre-mediated excision of intervening *ORFeus* copies could occur as early as at the zygote stage and in any cell types of a doubly transgenic G1 mouse. However, we did not select such G1 mice for our analyses because we have previously shown that a significant proportion of G1 animals underwent incomplete or mosaic recombination of a floxed cassette by this CAG-*cre* line (An et al., 2008). Instead, we backcrossed doubly transgenic G1 mice to wild-type C57BL/6J (B6) animals in order to segregate the CAG-*cre* transgene from the *ORFeus* transgene. G2 mice were genotyped for both *ORFeus* and CAG-*cre* transgenes as previously reported (An et al., 2008). *ORFeus*-positive, CAG-*cre*-negative G2 mice were selected for further analyses (Supplemental Fig. 1).

To determine the donor copy number, we developed a donor transgene-specific quantitative PCR (qPCR) assay. We used a Taqman probe and primers specific to the human γ -globin intron sequence that disrupts the *gfp* reporter gene (Fig. 1E). *ORFeus* insertions resulting from retrotransposition would not be detected as they lack the intronic sequence (Fig. 1C). By this method, the parental G0 animals were estimated to contain seven copies of *ORFeus* transgene (Fig. 1E, animals B1108 and B1109). Five G2 animals appeared to carry the transgene at a single copy (Fig. 1E, animals B1266, B1324, B1354, B1372, B1395) and one G2 animal had two copies (Fig. 1E, animal B1452).

3.2. High level of retrotransposition in G2 animals

The retrotransposition activities of *ORFeus* transgenes were examined in high-copy and low-copy mice by two different approaches. First, we performed retrotransposition profiling by an inverse PCR-based technique, which provides a semi-quantitative evaluation of overall transposition activities (An et al., 2006, 2008). In this assay, the number of bands per inverse PCR reaction for each sample represents the relative abundance of retrotransposition insertions. Our inverse PCR results suggested comparable levels of overall retrotransposition

between G0 and G2 samples. However, we were unable to derive quantitative data by this approach due to its semi-quantitative nature and potential PCR artifacts (data not shown).

To quantitatively compare retrotransposition between high-copy and low-copy samples, an insertion-specific Taqman-based qPCR approach was developed. We designed a probe that spans the splicing junction of the *gfp* reporter gene (Fig. 2A). The specificity of this qPCR strategy was verified by three negative and two positive control samples (Fig. 2B). The negative controls included wild-type gDNA (wt), no template control (NTC) and a gDNA sample (LSL) that carried a conditionally inactive, single-copy *ORFeus*^{LSL} transgene (An et al., 2008). The *ORFeus*^{LSL} transgene is identical in sequence to *ORFeus*^L except the presence of a loxP- β -geo-stop-loxP (i.e., LSL) cassette in between the CAG promoter and *ORFeus* coding sequences. No amplification was detected from sample LSL as *ORFeus* was not expressed in the presence of the LSL cassette (An et al., 2008). The first positive control (L-high) was a gDNA sample from the active, high-copy *ORFeus*^L mouse line (An et al., 2006). The second positive control (L) was from an animal carrying an activated, single-copy *ORFeus*^L transgene after Cre-mediated excision of the LSL cassette from *ORFeus*^{LSL} (An et al., 2008). The sensitivity and dynamic range of this qPCR assay were determined by using serial dilutions of a linearized GFP plasmid in wt gDNA (Fig. 2C). Quantitative signals could be reproducibly detected in a reaction containing as few as 10 GFP DNA molecules spiked in 2 ng of wt gDNA.

Retrotransposition events *in vivo* can be generally categorized as either germline or somatic insertions. By definition, germline insertions refer to retrotransposition events that occur in the germ cell lineage and thus can be transmitted to subsequent generations; somatic insertions refer to retrotransposition event in non-germ cells at any time point during development, which cannot pass to the next generation. From the perspective of quantification, a germline insertion is expected to generate a normalized signal of 1/cell since it is present at one copy in every single cell in the sample. In contrast, each somatic insertion is expected to generate a signal of <1/cell unless this insertion is uniformly present in every single cell in the sample. The readout of our qPCR approach is a cumulative signal from all insertions (germline and somatic) that are of sufficient length and can be amplified by our qPCR primers (Fig. 2A).

To accurately determine retrotransposition frequencies, we further calibrated the qPCR assay by comparing qPCR results with Southern blot analysis for a group of mouse samples carrying varied number of germline and somatic insertions (Fig. 2C,D). Sample B0385 was donor-positive but did not appear to carry any germline insertions according to Southern blot analysis, and its retrotransposition frequency was determined to be 0.48 events per cell. Samples B0386 and B0389 were donor-negative, and each contained a single germline insertion. Their retrotransposition frequencies were determined to be around one event per cell. Samples B0383 and B0388 were donor-containing, and both carried the same germline insertion. Their overall retrotransposition frequencies were 1.6 and 1.75 events per cell, respectively. Lastly, sample B0384 was donor-positive and carried two separate germline insertions, and its retrotransposition frequency was 2.91 per cell. The high concordance between the determined retrotransposition frequencies for these samples and data from Southern blot analysis validated our qPCR assay.

Retrotransposition frequencies for G0 and G2 samples were determined and calibrated by using samples each containing a single germline insertion event (Supplemental Fig. 2). On average, the single-copy G2 samples (0.45 events/cell) displayed an insignificant 1.2-fold increase in retrotransposition than the high-copy G0 samples (0.36 events/cell; Mann–Whitney–Wilcoxon test, $p>0.05$; Fig. 2E). However, the G0 group showed substantially higher variation among individual animals as compared with the G2 group (F -test, $p<0.001$;

Fig. 2E). In addition, the lone 2-copy G2 sample (B1452) had a retrotransposition frequency comparable to the lower range of G0 samples (0.26 events/cell; Supplemental Fig. 2). These qPCR data suggest that the animals examined here carried only somatic retrotransposition events since the normalized qPCR signal was <1 events/cell for all these samples.

3.3. Differential CpG methylation at the promoter sequence

Previous studies indicated that RIGS of tandem repeated transgenes was positively correlated with increased DNA methylation of transgene in high-copy arrays (Assaad et al., 1993; Garrick et al., 1998). To investigate the underlying mechanism that confers differential *ORFeus* activities among high- and low-copy *ORFeus* lines, we checked DNA methylation by bisulfite sequencing. Regions representing the promoter (CAG), the transgene body (ORF2) and the 3' end (GFP) of *ORFeus^L* transgene were selected for DNA methylation analysis (Fig. 1B). Two out of these three regions are specific to the donor transgene by design: the CAG promoter sequence is not part of the RNA transcript and is not expected to be present in any insertions; the forward primer for GFP amplicon resides in the intron, which will be removed during RNA processing and thus absent in insertions. In contrast, ORF2 primers will amplify both donor and insertion sequences. However, the majority of ORF2 amplicons should originate from donor elements for the following two reasons. First, most *de novo* L1 insertions are 5' truncated and <1 kb in length (An et al., 2006; Babushok et al., 2006), and ORF2 amplicon is situated 5.7 kb upstream from the polyadenylation site. Second, our qPCR results indicated that the average number of somatic insertions per cell was <0.5 in these samples but each cell had at least one copy of donor element (7 copies/cell for G0 samples).

Primers specific to each bisulfite converted target sequence were evaluated on gDNA samples containing either high- or low-copy *ORFeus* transgenes. Plasmid DNA of the transgene served as non-methylated DNA control. In all cases, specific amplicons were obtained with bisulfite treated DNA only (Supplemental Fig. 3A). The efficiency of bisulfite treatment was monitored by the C-to-T conversion rate in CpH dinucleotides (H=A, T, or C). An overall CpH cytosine conversion rate of 97.6% was achieved (Supplemental Fig. 3B). Reads with conversion rate below 95% were excluded from further analysis.

We first analyzed CpG methylation pattern in tail gDNA from G0 (Fig. 3A) and G2 animals (Fig. 3B), and observed distinct methylation patterns among three target regions. The GFP region was almost fully methylated in all animals (at an average methylation rate of 98.6%) and showed little variation from animal to animal. In comparison, the ORF2 region was methylated to a lesser degree (87.0% on average) and showed a wider range of methylation among animals (from 54.2% for B1395 to 99.5% for B1266). However, a pairwise comparison of ORF2 methylation level between parental G0 animals and G2 progeny indicated no statistical significance (Fig. 3C, $p>0.05$). The CAG promoter region displayed the most profound difference in DNA methylation between G0 and G2 animals. The average methylation level averaged 40.1% among all samples. Both G0 animals had consistently high level of methylation (65.1% and 69.4%). The 2-copy G2 sample had a similar, high level of methylation (B1452, 61.8%). The methylation level of CAG region varied greatly among five single-copy G2 samples, ranging from 4.1% to 60.9%. Statistical significance was detected for four of these five single-copy G2 samples when their methylation levels were compared with those of their G0 parents ($p<0.05$ or $p<0.01$, Fig. 3C).

Differential methylation at the promoter region between high-copy and low-copy animals was not limited to the tail tissue. We surveyed DNA methylation at all three target regions for multiple tissues from both high-copy and low-copy animals (Supplemental Fig. 4). Two high-copy animals analyzed showed little difference in DNA methylation in four different tissues (tail, brain, liver and testis; Fig. 4A). DNA methylation was compared between tail

and liver samples for selected G2 samples and the level of DNA methylation was largely maintained between assayed samples for individual animals (Fig. 4B).

We further confirmed that the observed differential methylation at the promoter region between high-copy and low-copy animals was maintained across generations. The methylation pattern showed little change from generation to generation across three generations of high-copy line (Supplemental Fig. 5). For low-copy samples, the overall methylation pattern was also maintained from G2 to G3 generation (Fig. 5).

3.4. Promoter hypomethylation at an independent, single-copy donor locus

Thus far, we have shown that differential DNA methylation at the promoter region exists between high- and low-copy animals, and that the pattern of DNA methylation is not tissue-specific and maintained across generations. However, this analysis was limited to one single genomic locus. To find out whether promoter hypomethylation occurs generally in single-copy transgenes, we also checked the DNA methylation pattern in a different transgenic line carrying an identical single-copy *ORFeus*^L transgene. This line was derived from a conditionally inactive, single-copy *ORFeus*^{LSL} transgene by Cre-mediated deletion of the LSL cassette (An et al., 2008). Active single-copy *ORFeus*^L line was generated by crossing to a germline specific *Zp3-cre* mouse line as previously described (Fig. 6A). In *Zp3-cre* mouse line, Cre expression is controlled by the mouse *zona pellucida* glycoprotein 3 (*Zp3*) promoter, and restricted to growing oocytes (de Vries et al., 2000). Thus, no Cre-mediated deletion is expected in somatic cells of doubly heterozygous female G1' animals. G2' animals that had undergone complete LSL excision and their G3' progeny were analyzed for DNA methylation. The results mirrored findings in Section 3.3: both ORF2 and GFP regions were highly methylated, but all samples were hypomethylated at the CAG promoter (14.7% on average).

4. Discussion

Several lines of evidence from this study suggest that tandem arrayed L1 transgenes with a heterologous promoter are subject to RIGS. First, the reduction of transgene copy number to single copy at the same genomic locus does not reduce the overall retrotransposition activity for *ORFeus* transgenes. An average frequency of 0.45 insertions per cell was observed for animals carrying a single-copy donor transgene, representing the highest somatic retrotransposition frequency hitherto reported. As a comparison, the somatic retrotransposition frequency ranged from 0.002 to 0.01 insertions per cell in a previous study with an L1RP mouse line (Kano et al., 2009). Second, an animal that carried two tandem copies of this transgene possessed a distinctly lower level of activity relative to its single-copy counterparts. Although this result is limited to a single animal, it is consistent with previous findings that RIGS operates on a minimum of two tandem copies (Assaad et al., 1993; Dorer and Henikoff, 1994). Third, as discussed below, differential CpG methylation at the promoter region is correlated with the observed retrotransposition activities between high-copy and single-copy animal groups. Fourth, this differential methylation is heritable across generations. Last, a similarly low level of methylation at the promoter was observed for an identical single-copy transgene at a different genomic locus, suggesting that differential methylation at the promoter is not restricted to a single genomic locus.

Differential methylation at the CAG promoter provides a plausible explanation for why the retrotransposition activity observed in single-copy donor samples is not reduced relative to that in high-copy samples. Previous studies on RIGS typically analyzed DNA methylation by using methylation sensitive restriction enzymes (Assaad et al., 1993; Garrick et al., 1998), an approach that surveys a limited number of CpG sites and does not delimit the differentially methylated regions. In the current study, three separate regions that represent

the promoter (CAG), the body (ORF2) and the 3' end (GFP) of the transgene were examined at high resolution by bisulfite sequencing. Interestingly, we observed heavy methylation in both ORF2 (87.0%) and GFP (98.6%) in both single-copy and high-copy mouse samples. The hypermethylation in the body and 3' end of the *ORFeus* transgene is not surprising as recent genome-scale surveys of cytosine methylation reveal high level of methylation in the gene body and 3' end of highly expressed human genes (Ball et al., 2009; Rauch et al., 2009). For the *ORFeus* transgene, no statistical significance was found in the level of gene body methylation between high-copy and single-copy samples. In contrast, the methylation level differed significantly at the promoter region between single-copy and high-copy samples for four out of the five single-copy samples. Therefore, in this study the promoter region is the primary target for RIGS-related differential methylation on the *ORFeus* transgene. However, the impact of differential DNA methylation on other stages of L1 life cycle, such as transcription, reverse transcription and integration has yet to be determined.

Different levels of interindividual variability on DNA methylation exist between high-copy and low-copy L1 transgenic animals. In the CAG promoter region, DNA methylation levels in the tail sample at the time of weaning vary greatly from 4.1% to 60.9% among single-copy G2 animals, and to a much lesser degree from 65.1% to 79% among high-copy parental animals (Fig. 3 and Supplemental Fig. 4). Interindividual variability on DNA methylation of human L1 5'UTR is frequently observed in normal human tissue samples (Chalitchagorn et al., 2004). Thus, we examined the potential role of genetic variation on DNA methylation among our mouse samples. The founder of the high-copy *ORFeus* animals was created in a B6SJLF1 background, and had since been backcrossed to wild-type B6 mice for five consecutive generations to produce G0 high-copy individuals. At this point, G0 animals were ~96.8% identical to B6. The CAG-*cre* mice were of pure B6 background. It took two rounds of breeding to derive G2 animals with reduced donor copy number, which is equivalent to 7 rounds of backcrossing. Thus, G2 animals were expected to be ~99.2% identical to B6. By the same logic, G3 animals were ~99.6% identical to B6. To place these numbers into perspective, a recent estimate based on comparing full human genomes from unrelated individuals is that genomes between human individuals are ≤99.5% identical to each other (Levy et al., 2007). Clearly, these G2 and G3 animals carried a level of genetic variation equal to or greater than that found in human populations, and such genetic variation could potentially contribute to the observed interindividual variations on DNA methylation. In addition, the observed interindividual variability may reflect the intrinsic, metastable nature of epialleles (Rakyan et al., 2002). Support for this epiallele hypothesis comes from an independent single-copy mouse line, in which the methylation levels on the CAG promoter region varied from 7.0% to 28.1% (Fig. 6), representing a more modest range of interindividual variation as compared to the single-copy animals above. In this second *ORFeus* mouse line, both the parental *ORFeus^{LSL}* mouse and *Zp3-cre* mouse were of pure B6 background; the G2' and G3' animals were genetically identical. However, to truly comprehend the role of genetic variation on L1 promoter methylation, further studies with L1 transgenes under the regulation of the endogenous 5'UTR promoter are needed.

The overall retrotransposition activity in single-copy animals is increased only slightly relative to high-copy samples, but such difference actually reflects an approximately 9-fold increase of retrotransposition activity on a per-copy basis for the single-copy transgene. The effect of RIGS on the high-copy *ORFeus* transgene is thus manifested as retrotransposition activities that are not proportional to the copy number of transgenes. The observed activity in high-copy animals is in agreement with the methylation data. The promoter region was heavily methylated in most of the bisulfite PCR amplicons from high-copy samples, but in each sample surveyed there existed one or more amplicons that were nearly devoid of methylation (see Fig. 3A). These hypomethylated transgene copies are the likely source for

the observed retrotransposition activity in high-copy samples. One possibility is that in every cell one or more *ORFeus* copies from the tandem array have escaped RIGS-related DNA methylation. However, a hallmark of RIGS is the cellular mosaicism in transgene expression (Martin and Whitelaw, 1996; Dobie et al., 1997). Thus, a more plausible explanation is that the tandem arrayed *ORFeus* transgenes are stochastically silenced at the cellular level—hypermethylated in some cells but hypomethylated in others. Cellular mosaicism in transgene expression may also underlie the high variability in *ORFeus* activity at the organismal level between high-copy animals (Fig. 2E).

It remains an open question as to whether RIGS applies to tandem arrayed L1 transgenes under the regulation of endogenous L1 promoters. DNA methylation is necessary for heritable silencing and regulation of L1 retrotransposons (Ooi et al., 2009). Earlier work has established that the 5'UTR and the body of endogenous L1s are heavily methylated in sperm and adult somatic tissues (Monk et al., 1987; Sanford et al., 1987; Nur et al., 1988; Lane et al., 2003), and modestly methylated in oocytes and pre-implantation embryos (Sanford et al., 1987; Howlett and Reik, 1991; Lane et al., 2003). In germ cell lineage, methylation of 5'UTR occurs in a developmentally coordinated fashion: starting from a high level of 65% at E11.5, it is gradually demethylated to approximately 17% by E13.5, and becomes fully remethylated in fetal testis by E17.5 (Hajkova et al., 2002; Lane et al., 2003; Lees-Murdock et al., 2003). This *de novo* DNA methylation of L1 retrotransposons in male germ cells is likely mediated through a piRNA pathway (Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008). Of particular interest, L1-derived piRNAs are enriched in the 5'UTR region in fetal premeiotic germ cells, suggesting that piRNAs are processed from full-length, potentially active copies (Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008). Several critical players from this *de novo* methylation pathway have been identified by recent loss of function genetic studies, including DNMT3L (Bourc'his and Bestor, 2004; Kato et al., 2007), MILI (Aravin et al., 2007; Kuramochi-Miyagawa et al., 2008), MIWI2 (Carmell et al., 2007; Kuramochi-Miyagawa et al., 2008) and MAEL (Soper et al., 2008). In light of these new findings, it is tempting to ponder how L1 can be best modeled by engineered transgenes. Endogenous L1s are present in mammalian genomes largely as dispersed individual copies. The number of full-length elements is in the order of 6000 to 7000 in both human and mouse genomes (Penzkofer et al., 2005; Khan et al., 2006). Among these, ~100 and ~3000 copies are potentially active for retrotransposition in human (Brouha et al., 2003) and mouse genomes (Goodier et al., 2001), respectively. Therefore, if the 5'UTR is the primary target for piRNA-guided *de novo* methylation and silencing of L1 retrotransposons, it is reasonable to expect that a single-copy L1 transgene with a mouse L1 5'UTR promoter will be regulated similarly as compared to endogenous L1s. How tandem arrayed 5'UTR-L1 transgenes respond to DNA methylation is debatable due to the lack of molecular insights on RIGS. Studies on transgene silencing in *Drosophila* support a homologous pairing model of heterochromatin formation for RIGS: homologous pairing of multiple closely linked copies produces folded structures that are recognized by heterochromatin-specific proteins (Dorer and Henikoff, 1994; Dorer, 1997). It is unknown whether small RNAs play any role in RIGS on tandem arrayed transgenes. Heterochromatin formation may be a common theme between RIGS and piRNA-mediated silencing of transposable elements as a recent study in *Drosophila* uncovered a tantalizing link between heterochromatin and piRNA pathways (Halic and Moazed, 2009; Klattenhoff et al., 2009). Nevertheless, when interpreting data from L1 transgenic studies it is prudent to bear in mind the potential implications of RIGS phenomenon on tandem arrayed transgenes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Jef Boeke for providing high-copy *ORFeus* transgenic mice, Charles Hawkins for the CAG-*cre* mice, Jeffery Han for plasmids and comments, and Julie Stanton for editorial help. This work was supported by start-up funds from Washington State University.

Abbreviations

B6	C57BL/6
CAG	a composite CMV immediate-early enhancer/modified chicken β -actin promoter
Kb	kilobase(s)
gDNA	genomic DNA
LINE-1 or L1	long interspersed element type 1
non-LTR	non-long terminal repeat
LSL	loxP- β -geo-stop-loxP cassette
ORF	open reading frame
RIGS	repeat-induced gene silencing
qPCR	quantitative polymerase chain reaction
UTR	untranslated region
wt	wild type
Zp3	mouse zona pellucida glycoprotein 3 promoter

References

- An W, Han JS, Wheelan SJ, Davis ES, Coombes CE, Ye P, Triplett C, Boeke JD. Active retrotransposition by a synthetic L1 element in mice. *Proc. Natl. Acad. Sci. U. S. A.* 2006; 103:18662–18667. [PubMed: 17124176]
- An W, Han JS, Schrum CM, Maitra A, Koentgen F, Boeke JD. Conditional activation of a single-copy L1 transgene in mice by Cre. *Genesis*. 2008; 46:373–383. [PubMed: 18615728]
- Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science*. 2007; 316:744–747. [PubMed: 17446352]
- Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell*. 2008; 31:785–799. [PubMed: 18922463]
- Ashe A, Morgan DK, Whitelaw NC, Bruxner TJ, Vickaryous NK, Cox LL, Butterfield NC, Wicking C, Blewitt ME, Wilkins SJ, Anderson GJ, Cox TC, Whitelaw E. A genome-wide screen for modifiers of transgene variegation identifies genes with critical roles in development. *Genome Biol*. 2008; 9:R182. [PubMed: 19099580]
- Assaad FF, Tucker KL, Signer ER. Epigenetic repeat-induced gene silencing (RIGS) in *Arabidopsis*. *Plant. Mol. Biol.* 1993; 22:1067–1085. [PubMed: 8400126]
- Babushok DV, Ostertag EM, Courtney CE, Choi JM, Kazazian HH Jr. L1 integration in a transgenic mouse model. *Genome Res*. 2006; 16:240–250. [PubMed: 16365384]
- Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, Xie B, Daley GQ, Church GM. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat. Biotechnol.* 2009; 27:361–368. [PubMed: 19329998]
- Bestor TH. Transposons reanimated in mice. *Cell*. 2005; 122:322–325. [PubMed: 16096053]

- Bishop JO, Smith P. Mechanism of chromosomal integration of microinjected DNA. *Mol. Biol. Med.* 1989; 6:283–298. [PubMed: 2695741]
- Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature.* 2004; 431:96–99. [PubMed: 15318244]
- Branciforte D, Martin SL. Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. *Mol. Cell Biol.* 1994; 14:2584–2592. [PubMed: 8139560]
- Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH Jr. Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl. Acad. Sci. U. S. A.* 2003; 100:5280–5285. [PubMed: 12682288]
- Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev. Cell.* 2007; 12:503–514. [PubMed: 17395546]
- Chalitchagorn K, Shuangshoti S, Hourpai N, Kongruttanachok N, Tangkijvanich P, Thong-ngam D, Voravud N, Sriuranpong V, Mutirangura A. Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis. *Oncogene.* 2004; 23:8841–8846. [PubMed: 15480421]
- Collier LS, Largaespada DA. Transposons for cancer gene discovery: Sleeping Beauty and beyond. *Genome Biol.* 2007; 8(1):S15. [PubMed: 18047692]
- de Vries WN, Binns LT, Fancher KS, Dean J, Moore R, Kemler R, Knowles BB. Expression of Cre recombinase in mouse oocytes: a means to study maternal effect genes. *Genesis.* 2000; 26:110–112. [PubMed: 10686600]
- Ding S, Wu X, Li G, Han M, Zhuang Y, Xu T. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell.* 2005; 122:473–483. [PubMed: 16096065]
- Dobie K, Mehtali M, McClenaghan M, Lathe R. Variegated gene expression in mice. *Trends Genet.* 1997; 13:127–130. [PubMed: 9097721]
- Dorer DR. Do transgene arrays form heterochromatin in vertebrates? *Transgenic Res.* 1997; 6:3–10. [PubMed: 9032972]
- Dorer DR, Henikoff S. Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell.* 1994; 77:993–1002. [PubMed: 8020105]
- Ergun S, Buschmann C, Heukeshoven J, Dammann K, Schnieders F, Lauke H, Chalajour F, Kilic N, Stratling WH, Schumann GG. Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. *J. Biol. Chem.* 2004; 279:27753–27763. [PubMed: 15056671]
- Garrick D, Fiering S, Martin DI, Whitelaw E. Repeat-induced gene silencing in mammals. *Nat. Genet.* 1998; 18:56–59. [PubMed: 9425901]
- Goodier JL, Ostertag EM, Du K, Kazazian HH Jr. A novel active L1 retrotransposon subfamily in the mouse. *Genome Res.* 2001; 11:1677–1685. [PubMed: 11591644]
- Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J, Surani MA. Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* 2002; 117:15–23. [PubMed: 12204247]
- Halic M, Moazed D. Transposon silencing by piRNAs. *Cell.* 2009; 138:1058–1060. [PubMed: 19766558]
- Henikoff S. Conspiracy of silence among repeated transgenes. *Bioessays.* 1998; 20:532–535. [PubMed: 9723001]
- Howlett SK, Reik W. Methylation levels of maternal and paternal genomes during preimplantation development. *Development.* 1991; 113:119–127. [PubMed: 1764989]
- Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, Kazazian HH Jr. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes Dev.* 2009; 23:1303–1312. [PubMed: 19487571]
- Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y, Okano M, Li E, Nozaki M, Sasaki H. Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum. Mol. Genet.* 2007; 16:2272–2280. [PubMed: 17616512]
- Kazazian HH Jr. Mobile elements: drivers of genome evolution. *Science.* 2004; 303:1626–1632. [PubMed: 15016989]

- Khan H, Smit A, Boissinot S. Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. *Genome Res.* 2006; 16:78–87. [PubMed: 16344559]
- Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, Seitz H, Zamore PD, Weng Z, Theurkauf WE. The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell.* 2009; 138:1137–1149. [PubMed: 19732946]
- Kumaki Y, Oda M, Okano M. QUMA: quantification tool for methylation analysis. *Nucleic Acids Res.* 2008; 36:W170–W175. [PubMed: 18487274]
- Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Totoki Y, Toyoda A, Ikawa M, Asada N, Kojima K, Yamaguchi Y, Ijiri TW, Hata K, Li E, Matsuda Y, Kimura T, Okabe M, Sakaki Y, Sasaki H, Nakano T. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev.* 2008; 22:908–917. [PubMed: 18381894]
- Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J, Reik W. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis.* 2003; 35:88–93. [PubMed: 12533790]
- Lees-Murdock DJ, De Felici M, Walsh CP. Methylation dynamics of repetitive DNA elements in the mouse germ cell lineage. *Genomics.* 2003; 82:230–237. [PubMed: 12837272]
- Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G, Lin Y, MacDonald JR, Pang AW, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers YH, Frazier ME, Scherer SW, Strausberg RL, Venter JC. The diploid genome sequence of an individual human. *PLoS Biol.* 2007; 5:e254. [PubMed: 17803354]
- Martin DI, Whitelaw E. The vagaries of variegating transgenes. *Bioessays.* 1996; 18:919–923. [PubMed: 8939070]
- Mates L, Izsvak Z, Ivics Z. Technology transfer from worms and flies to vertebrates: transposition-based genome manipulations and their future perspectives. *Genome Biol.* 2007; 8(1):S1. [PubMed: 18047686]
- Monk M, Boubelik M, Lehnert S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development.* 1987; 99:371–382. [PubMed: 3653008]
- Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature.* 2005; 435:903–910. [PubMed: 15959507]
- Muotri AR, Zhao C, Marchetto MC, Gage FH. Environmental influence on L1 retrotransposons in the adult hippocampus. *Hippocampus.* 2009; 19:1002–1007. [PubMed: 19771587]
- Nur I, Pascale E, Furano AV. The left end of rat L1 (L1Rn, long interspersed repeated) DNA which is a CpG island can function as a promoter. *Nucleic Acids Res.* 1988; 16:9233–9251. [PubMed: 2459662]
- Ooi SK, O'Donnell AH, Bestor TH. Mammalian cytosine methylation at a glance. *J. Cell Sci.* 2009; 122:2787–2791. [PubMed: 19657014]
- Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, Gerton GL, Kazazian HH Jr. A mouse model of human L1 retrotransposition. *Nat. Genet.* 2002; 32:655–660. [PubMed: 12415270]
- Palmiter RD, Brinster RL. Germ-line transformation of mice. *Annu. Rev. Genet.* 1986; 20:465–499. [PubMed: 3545063]
- Penzkofer T, Dandekar T, Zemojtel T. L1Base: from functional annotation to prediction of active LINE-1 elements. *Nucleic Acids Res.* 2005; 33:D498–D500. [PubMed: 15608246]
- Prak ET, Dodson AW, Farkash EA, Kazazian HH Jr. Tracking an embryonic L1 retrotransposition event. *Proc. Natl. Acad. Sci. U. S. A.* 2003; 100:1832–1837. [PubMed: 12569170]
- Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. Metastable epialleles in mammals. *Trends Genet.* 2002; 18:348–351. [PubMed: 12127774]
- Rauch TA, Wu X, Zhong X, Riggs AD, Pfeifer GP. A human B cell methylome at 100-base pair resolution. *Proc. Natl. Acad. Sci. U. S. A.* 2009; 106:671–678. [PubMed: 19139413]
- Sabl JF, Henikoff S. Copy number and orientation determine the susceptibility of a gene to silencing by nearby heterochromatin in *Drosophila*. *Genetics.* 1996; 142:447–458. [PubMed: 8852844]

- Sakai K, Miyazaki J. A transgenic mouse line that retains Cre recombinase activity in mature oocytes irrespective of the cre transgene transmission. *Biochem. Biophys. Res. Commun.* 1997; 237:318–324. [PubMed: 9268708]
- Sanford JP, Clark HJ, Chapman VM, Rossant J. Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes Dev.* 1987; 1:1039–1046. [PubMed: 3428592]
- Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C (T) method. *Nat. Protoc.* 2008; 3:1101–1108. [PubMed: 18546601]
- Soper SF, van der Heijden GW, Hardiman TC, Goodheart M, Martin SL, de Boer P, Bortvin A. Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. *Dev. Cell.* 2008; 15:285–297. [PubMed: 18694567]
- Takeda J, Keng VW, Horie K. Germline mutagenesis mediated by Sleeping Beauty transposon system in mice. *Genome Biol.* 2007; 8(1):S14. [PubMed: 18047691]
- Trelogan SA, Martin SL. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 1995; 92:1520–1524. [PubMed: 7878012]
- Ye F, Signer ER. RIGS (repeat-induced gene silencing) in *Arabidopsis* is transcriptional and alters chromatin configuration. *Proc. Natl. Acad. Sci. U. S. A.* 1996; 93:10881–10886. [PubMed: 8855276]

Fig. 1.

Generation of mice carrying reduced number of *ORFeus^L* transgenes. (A) The parental mouse line carries a tandem array of *ORFeus^L* transgene, shown as a head-to-tail array (An et al., 2006). The loxP site is illustrated as an open triangle. (B) After Cre-mediated recombination, intervening copies are excised, resulting in a transgene array with reduced copy number. A single-copy transgene is illustrated. Each copy is composed of a CAG promoter, followed by *ORFeus* coding sequences and a *gfp* based reporter cassette. The regions for DNA methylation analysis are indicated as blue lines above the vector diagram. (C) A 5' truncated insertion resulting from *ORFeus* retrotransposition is illustrated. Note the insertion is intronless as the intronic sequence has been spliced out from *ORFeus* transcript prior to reverse transcription and integration. (D) The breeding scheme to reduce transgene copy number. (E) Determination of transgene copy number by donor-specific qPCR. The Taqman probe (purple line) and primers (black arrowheads) are specific to the intron that disrupts the *gfp* reporter (inset). The transgene copy number was reduced to either 1 or 2 copies in six G2 mice. Wild-type tail gDNA samples spiked with indicated number of pRSVGFPuvINT plasmid were used to measure assay sensitivity and dynamic range. Error bars represent standard deviations of normalized transgene copy number from three technical replicates for each sample. Values for individual samples were normalized to that of a sample known to contain a single-copy transgene (2G6).

Fig. 2.

Quantitation of retrotransposition frequencies by qPCR. (A) qPCR strategy. The primers (black arrowheads) anneal to both the intron-containing donor element (top) and an insertion (bottom), but the Taqman probe (red line) is specific to insertions as it spans the exon–exon junction of the *gpf* reporter gene. (B) Confirmation of qPCR specificity by multiple control samples. At the completion of qPCR cycling, triplicate reactions for each control sample were resolved on a 2% agarose gel. GFP, retrotransposition-specific signal; GAPDH, endogenous control. Control tail gDNA samples L-high, LSL and L, detailed in text; wt, wild-type C57BL/6 tail gDNA sample; NTC, no template control; M, 100 bp DNA molecular weight ladder (New England Biolabs). (C) Calibration of retrotransposition qPCR with mouse tail gDNA and plasmid DNA samples. The presence/absence of the high-copy donor transgene and the number of germline insertions were examined independently by Southern blot (panel D). Wt tail gDNA samples spiked with indicated number of pRSVGFPuv plasmid were used to measure assay sensitivity and dynamic range. Error bars represent standard deviations of normalized retrotransposition frequencies from three technical replicates for each sample. Values for individual samples were normalized to that of a sample containing a single germline insertion (B1045, Supplemental Fig. 2). (D) Southern blot analysis of *ORFeus* donor and insertions among a litter of animals. The parental animal (B0131) carried the high-copy transgene array. Tail gDNA was digested with PstI, which cuts the transgene internally, and hybridized to a radiolabeled probe that targets the 3' end of the transgene as previously described (An et al., 2006). Donor transgene fragment migrated as an intense ~4.5 kb band on the gel. Plus signs indicate putative germline insertions. Animals that were surveyed in panel C are highlighted in blue. DNA migration positions are indicated (1 kb ladder, New England Biolabs). (E) Comparison of retrotransposition frequencies between G0 and single-copy G2 animals. Average frequencies were calculated for each group (G0, G2, and germline) from raw data presented in Supplemental Fig. 2, and normalized to that of the germline group (set to 1). Error bars

represent standard errors of normalized retrotransposition frequencies from corresponding biological replicates for each group. Normalized retrotransposition frequencies for each biological sample are also presented to show the distribution (black dots).

Fig. 3.

Differential promoter methylation between high-copy and low-copy *ORFeus* transgenes in tail gDNA. (A) DNA methylation profiles for two parental G0 mice at three regions of the transgene (CAG, ORF2, and GFP). Tail gDNA was used. Mouse generation and the transgene copy number of each sample are indicated in parentheses. Each horizontal line represents an independent bisulfite PCR amplicon. Methylated and unmethylated CpG sites are diagrammed as filled and open circles, respectively. (B) DNA methylation profiles for six G2 mice. Tail gDNA was used. The samples are arranged in ascending order by the overall methylation rate at the CAG promoter. (C) Comparison of DNA methylation rates between G0 and G2 animals. DNA methylation rates are derived from panels A and B. Error bars represent standard errors of DNA methylation rate between individual sequences for each target region. The difference in DNA methylation between G0 and G2 animals is evaluated by Mann–Whitney *U*-test (* $p < 0.05$; ** $p < 0.01$).

Fig. 4.

Comparison of DNA methylation rates at the CAG promoter among different tissues in high- (A) and low-copy animals (B). DNA methylation rates are derived from panels A and B of Supplemental Fig. 4, respectively. All gDNA samples shown were collected from animals at 13–15 months of age. Error bars represent standard errors of DNA methylation rate between individual sequences for each target region. The difference in DNA methylation between tissues is evaluated by Mann–Whitney *U*-test (* $p < 0.05$).

Fig. 5.

Cross-generation DNA methylation analysis in low-copy transgenes. (A) Two G3 animals each were selected for three G2 parental animals and analyzed for DNA methylation in three regions of the transgene (CAG, ORF2, and GFP). Tail gDNA was used for B1324, B1452 and their progeny. Liver gDNA was used for B1372 (at 13 months of age) and its progeny (both at 5 months of age). (B) Methylation rates are presented in histogram. The difference in DNA methylation between G2 and G3 generations is evaluated by Mann–Whitney *U*-test. Note B1508 is the only animal that shows statistical significance in such pairwise comparison (** $p < 0.01$).

Fig. 6.

DNA methylation profiles of a single-copy *ORFeus^L* transgene at an independent genomic locus. (A) Breeding strategy. This single-copy *ORFeus^L* line has been previously reported (An et al., 2008). It was derived from a single-copy *ORFeus^{LSL}* line, in which the promoter and coding sequences were separated by a loxP- β -geo-stop-loxP (LSL) cassette. To differentiate from Fig. 1, mouse generations are designated as G0', G1', etc. (B) DNA methylation profile for animals from different generations. Tail gDNA was used. (C) Methylation rates presented in histogram.