Heritable L1 retrotransposition in the mouse primordial germline and early embryo

Citation for published version:

Digital Object Identifier (DOI):
10.1101/gr.219022.116

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Genome Research

Publisher Rights Statement:
This article, published in Genome Research, is available under a Creative Commons License (Attribution 4.0 International), as described at http://creativecommons.org/licenses/by/4.0/.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Heritable L1 retrotransposition in the mouse primordial germline and early embryo

Sandra R. Richardson,1 Patricia Gerdes,1 Daniel J. Gerhardt,1,2 Francisco J. Sanchez-Luque,1,3 Gabriela-Oana Bodea,1 Martin Muñoz-Lopez,3 J. Samuel Jesuadian,1 Marie-Jeanne H.C. Kempen,1 Patricia E. Carreira,1 Jeffrey A. Jeddeloh,4 Jose L. Garcia-Perez,3,5 Haig H. Kazazian Jr.,6 Adam D. Ewing,1 and Geoffrey J. Faulkner1,7,8

1Mater Research Institute–University of Queensland, Woolloongabba QLD 4102, Australia; 2Invenra, Incorporated, Madison, Wisconsin 53719, USA; 3Department of Genomic Medicine, GENYO, Centre for Genomics and Oncological Research, Pfizer–University of Granada–Andalusian Regional Government, PTS Granada, 18016 Granada, Spain; 4Roche Sequencing Solutions, Incorporated, Madison, Wisconsin 53719, USA; 5Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, United Kingdom; 6Institute of Genetic Medicine and Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA; 7School of Biomedical Sciences, 8Queensland Brain Institute, University of Queensland, Brisbane QLD 4072, Australia

Long interspersed element 1 (LINE-1 or L1) is a mobile genetic element active in nearly all mammals (Furano 2000). L1 sequences mobilize via a copy-and-paste mechanism, termed retrotransposition (Moran et al. 1996), and comprise ~18% of mouse DNA (Waterston et al. 2002). Each mouse genome harbors ~3000 full-length retrotransposition-competent L1s (RC-L1s) belonging to three L1 subfamilies (T′p, GP, and A) as well as nearly 600,000 L1 copies rendered immobile by 5′ truncation and the accumulation of internal mutations (Fanning 1983; Loeb et al. 1986; Padgett et al. 1988; Kingsmore et al. 1994; Takahara et al. 1996; DeBerardinis et al. 1998; Naas et al. 1998; Goodier et al. 2001; Waterston et al. 2002; Sookdeo et al. 2013). The ongoing production of new, heritable RC-L1 copies is therefore essential to preserve L1 mobility over evolutionary time. It follows that L1 mRNA and protein are expressed during germline and early embryonic development (Martin and Branciforte 1993; Branciforte and Martin 1994; Trelogan and Martin 1995; Garcia-Perez et al. 2007; Soper et al. 2008; Malik et al. 2014), and numerous host mechanisms regulate L1 activity during these stages (Yoder et al. 1997; Bourc’his and Bestor 2004; Watanabe et al. 2006, 2008; Aravin et al. 2008; Soper et al. 2008; Rowe et al. 2010; Zamudio and Bourc’his 2010; Wissing et al. 2011; Castro-Díaz et al. 2014; Crichton et al. 2014).

The developmental timing of only two heritable human L1 insertions has been elucidated; one event likely occurred in the female germline (Brouha et al. 2002), and the other occurred in a pluripotent embryonic cell and resulted in maternal somatic and germline mosaicism (van den Hurk et al. 2007). This result is consistent with reports of L1 retrotransposition in cultured human embryonic stem cells and induced pluripotent stem cells (iPSCs) (Garcia-Perez et al. 2007; Wissing et al. 2012; Klawitter et al. 2016), although a study of mouse iPSCs revealed little endogenous retroelement activity (Quinlan et al. 2011). Studies of transgenic L1 reporter animals have demonstrated retrotransposition in the germline (Ostertag et al. 2002; An et al. 2006) and in the early embryo (Kano et al. 2009). Surprisingly, in the latter study, transmission of engineered L1 insertions from mosaic parental animals to offspring was never observed, suggesting somatic but not germline contribution of insertion-harboring embryonic cells (Kano et al. 2009). Overall, the frequency and developmental timing of heritable L1 retrotransposition in vivo remain unclear. Here, we overcome the rarity of phenotype-causing endogenous

© 2017 Richardson et al. This article, published in Genome Research, is available under a Creative Commons License (Attribution 4.0 International), as described at http://creativecommons.org/licenses/by/4.0/.
Results

We bred two- and three-generation pedigrees of C57BL/6J mice (Fig. 1), a strain known to accommodate recent L1 activity (Akagi et al. 2008), and used mRC-seq and whole-genome sequencing (WGS) to identify retrotransposon insertions absent from the C57BL/6J reference genome and not previously identified in an extensive analysis of polymorphic transposable element insertions across 13 commonly used inbred and four wild-derived mouse strains (Fig. 1; Supplemental Fig. S1; Methods; Nellaker et al. 2012). After sequencing analyses, we classified 28 L1 and short in- 


terspersed element (SINE) insertions present in at least one of the 


twenty mouse strain genomes and identified de novo L1 insertions (Fig. 1; Supplemental Fig. S1A; Mager and Stoye 2015; Richardson et al. 2015). We apply this technology to pedigrees of wild-type C57BL/6J mice to investigate the rate and developmental timing of heritable L1 insertions.

Next, we used PCR genotyping to investigate the development of each de novo L1 insertion (Supplemental Fig. S1B). Insertions #1 and #2 were identified by mRC-seq in mice SRAB2 and SRAB15, respectively (Fig. 2A; Supplemental Fig. S3; Supplemental Table 2). We did not detect these insertions by PCR genotyping in the somatic tissues of parental mice SRA and SRB; however, insertions #1 and #2 were detected by PCR in both testicles of paternal mouse SRB (Fig. 2A; Supplemental Fig. S3A). Inheritance of each insertion by only 1/20 offspring, coupled with their presence in both testicles of the paternal mouse, suggested germline-restricted mosaicism for insertions #1 and #2 in mouse SRB (Fig. 2B; Supplemental Fig. S3A).

In an inverse approach to identify germline-restricted mosaic insertions, we performed deep (~260×) mRC-seq on the germ cell fraction of each testicle of mouse SRCD14. We detected and PCR-validated insertion #7 in both testicles and did not detect it in the somatic tissues of SRCD14 (Fig. 2C,D; Supplemental Table 2). An insertion-specific genomic DNA qPCR assay targeting the 5′ L1-genome junction of insertion #7 revealed its prevalence of ∼11% and ∼4% in the germ cell fraction of the left and right testicle of SRCD14, respectively (Fig. 2E), and subsequent PCR genotyping demonstrated transmission of the insertion to 2/65 progeny of SRCD14 (F2-7 and F2-55) (Fig. 1; Supplemental Fig. S1G). Insertion #5 also occurred within an intronic L1mAS element, and two additional intergenic insertions landed within existing L1 repeats (Supplemental Table 4).

All 11 de novo insertions were S′ detected by mRC-seq and thus full-length (containing ≥1 T′ termini), reflecting depletion of mouse L1 3′ termini observed during Illumina sequencing (Supplemental Fig. S1D, Supplemental Tables 1, 2), possibly due to the GC-rich nature of the mouse L1 3′ end sequence (Chambers et al. 2015). Furthermore, the de novo L1 insertions had relatively long poly(A) tracts (average ~64 bp), reducing the likelihood that L1 3′ end sequence and flanking genomic DNA would be captured in a single sequencing read. However, 20–30× WGS, which, in principle, could allow detection of the 5′ junctions of 5′ truncated L1 insertions, applied to nine mouse genomes (SRE, SRF, and offspring SREF15-21) uncovered no 5′ truncated de novo L1 insertions (Fig. 1; Supplemental Tables 1, 2), and a previous analysis suggested that T′ L1s undergo 5′ truncation less frequently than other L1 elements (Hardies et al. 2000). Complete internal sequencing of nine de novo L1 insertions revealed intact ORFs and the absence of mutations in critical functional domain residues (Supplemental Fig. S2; Furano 2000).

We regard insertions detected in one or more offspring but absent from the corresponding parents as potentially de novo (Supplemental Fig. S1B). Using PCR and capillary sequencing, we validated 11 de novo L1 insertions (Table 1; Supplemental Table 2). All 11 were T′ subfamily elements, consistent with previous reports of disease-causing L1 insertions in mice wherein all insertions for which sufficient L1 sequence was present for subfamily distinction were identified as T′ elements (Kingsmore et al. 1994; Mulhardt et al. 1994; Kohrman et al. 1996; Takahara et al. 1996; Perou et al. 1997; Naas et al. 1998; Yajima et al. 1999; Cunilffe et al. 2001). De novo L1 insertions bore hallmarks of L1 retrotransposition by target-primed reverse transcription (TPRT), including insertion at sequences resembling the L1 endonuclease cleavage motif (5′-TTTT/AA-3′), the presence of 13- to 17-bp target-site duplications (TSDs), and 3′ poly(A) tracts (Table 1; Supplemental Figs. S1E, S4; Singer et al. 1983; Scott et al. 1987; Luan et al. 1993; Moran et al. 1996; Jurka 1997). The average GC content of de novo L1 insertion sites within a 50-bp and 20-kb window of the en-
Figure 1. Origin and transmission of de novo L1 insertions in mouse pedigrees. Above, two-generation pedigree originating from parental mice SRA/SRB. Below, three-generation pedigrees originating from parental mice SRC/SRD and SRE/SRF. F1 animals are designated by the parental pair from which they arose (AB, CD, EF); F2 animals are so indicated. De novo L1 T\textsubscript{\textsuperscript{f}} insertions \#1–\#11 are color coded. In subsequent figures, schematics of each insertion are likewise color coded. With respect to each insertion, filled shapes indicate heterozygous animals, vertical hatching indicates germline mosaicism, horizontal hatching indicates somatic mosaicism, and vertical and horizontal hatching together indicates both somatic and germline mosaicism. Circles represent female animals; squares represent males. Dashed lines indicate matings. Animals for which mRC-seq was performed on somatic tissue gDNA are indicated with an asterisk; animals for which whole-genome sequencing (WGS) was performed on somatic tissue gDNA are indicated with a plus sign. For mouse SRCD14, mRC-seq was performed on somatic tissues as well as the germ cell fraction of the left and right testicles. Polymorphic retrotransposon insertions are not depicted in this figure.
in germline-restricted genetic mosaicism and heritable de novo L1 insertions.

To investigate the capacity of de novo L1 insertions for subsequent retrotransposition, we cloned insertions #1 and #7 and tested their activity in a cultured cell retrotransposition assay (Fig. 2F; Fig. S6A,B, #6 (Supplemental Fig. S6C,D), and #11 ∼12q 3.5 TCTT/AG 16 ∼100 Male primordial germ line 1/20 (5%) ∼10q 5 TCTT/GT 14 ∼39 Female early embryo 4/7 (57%) ∼10q 5 TCTT/GT 14 ∼39 Female early embryo 4/7 (57%) ∼1q 4 TCTT/AC 17 ∼40 Male germ line 2/15 (13%) ∼1q 4 TCTT/AA 16 ∼40 Male germ line 2/15 (13%) ∼2q 3 CTTC/AA 13 ∼55 Female early embryo 0/5 (0%) ∼4.5 TCTT/AT 15 ∼40 Late germ line 11/23 (48%) ∼4.5 TCTT/AT 15 ∼40 Late germ line 11/23 (48%) ∼4.5 TTAT/GT 13 ∼60 Male early embryo 0/6 (0%) ∼4.5 TTAT/GT 13 ∼60 Male early embryo 0/6 (0%) ∼4.5 TTAT/GT 13 ∼60 Male early embryo 0/6 (0%) ∼4.5 TTAT/GT 13 ∼60 Male early embryo 0/6 (0%) ∼4.5 TTAT/GT 13 ∼60 Male early embryo 0/6 (0%) Table 1. Characteristics of 11 de novo L1 insertions identified in this study

All insertions are L1 T, elements. (T, monomers) Number of repetitive promoter units contained by each insertion. Cleavage indicates the L1 endonuclease cleavage motif, shown 5′ to 3′, (TSD) target-site duplication length, (Poly(A)) poly(A) tract length. Poly(A) tract lengths were determined by capillary sequencing and should be regarded as estimates. Insertions marked with an asterisk contain a 3′ transduction (see Supplemental Figs. S5, S4 for details). For each insertion, the deduced developmental origin and the frequency of transmission (i.e., the percentage of siblings heterozygous for the insertion) are indicated.

mouse SREF16 to mouse SRCD14, insertion #4 was transmitted to 11/19 offspring (Figs. 1, #5), consistent with mouse SREF16 being either consummately heterozygous for the insertion or mosaic with a high degree of germline prevalence. To distinguish these possibilities, we used a 5′ L1-genome junction qPCR assay for insertion #4, with heterozygous and wild-type F2 offspring as controls, and found that SREF16 contained ~1 copy each of insertion #4 and the genomic empty site across all tissues tested (Fig. 4C,D). Insertion #4 could not be detected in the gonads of SRE or SRF, despite sensitivity of the qPCR assay to 0.1% prevalence (Supplemental Fig. S7A). We therefore conclude that mouse SREF16 was heterozygous for insertion #4 and that this insertion arose sufficiently late during germline development of parental mouse SRE or SRF to preclude its detection in bulk gonad tissues, or possibly post-conception in SREF16 at the mature zygote stage.

Finally, insertion #9 was detected in 11/23 offspring of SRCD11 and SREF20 (Fig. 1; Supplemental Fig. S7B,C; Supplemental Table 2). This rate of transmission suggested either parental heterozygosity or mosaicism with a high degree of germline prevalence. However, we could not detect insertion #9 by nested PCR in the gonads of SRCD11 or SREF20 (Supplemental Fig. S7H). The developmental origins of insertion #9 are therefore unclear. We speculate that this event was germline-restricted mosaic in SRCD11 or SREF20, but that the subset of germ cells carrying insertion #9 was depleted by the time the gonads of these animals were harvested, at age 39 and 40 wk (Lei and Spradling 2013).

We next sought to identify the progenitor L1 elements responsible for heritable de novo insertions. During transduction of an RC-L1 element, the native L1 polyadenylation signal is occasionally bypassed in favor of a downstream genomic polyadenylation signal, and upon retrotransposition the nascent L1 insertion incorporates a genomic sequence tag, or 3′ transduction, that identifies the progenitor L1 element (Holmes et al. 1994; Moran et al. 1999, 1996; Goodier et al. 2000; Pickeral et al. 2000; Garcia-Perez et al. 2007; van den Hurk et al. 2007; Kano et al. 2013; Macfarlane et al. 2013). Two early PGC insertions (#2 and #7) (Fig. 2C; Supplemental Fig. S8A,B) and two presumably recent polymorphic L1 T, insertions differentially present/absent among our animals (Poly1L1T,3 and Poly1L1T,4) (Table 1; Supplemental Table 2; Supplemental Figs. S5, S8A,C) harbored RC-L1 elements, indicating the potential to serve as progenitor elements for subsequent retrotransposition events.

We traced the developmental timing of six de novo L1 insertions to the early embryo. As an illustrative example, insertion #5 was detected by mRC-seq with a single sequencing read in the brain of maternal mouse SRE and robustly detected in 4/7 offspring (Fig. 1; Supplemental Table 2). PCR genotyping revealed bands of varying intensity among the somatic tissues and ovaries of mouse SRE (Fig. 3A,B), and by genomic DNA qPCR, the prevalence of insertion #5 ranged from ~0.2% in brain to ~1.5% in the right ovary (Fig. 3C). Similarly, insertions #3 (Supplemental Fig. S6A,B), #6 (Supplemental Fig. S6C,D), and #11 (Supplemental Fig. S6E) were each identified by mRC-seq in multiple offspring and genotyped as mosaic in the respective maternal tissues. Insertions #8 (Supplemental Fig. S6F) and #10 (Supplemental Fig. S6G) were identified by mRC-seq and were confirmed as mosaic in the tissues of mouse SREF17 and mouse F2-145, respectively, but were not transmitted to the limited offspring produced by these animals (five and six progeny, respectively) (Supplemental Fig. S6F,G). Therefore, we can neither confirm nor rule out the contribution of insertions #8 and #10 to the germ lineage. Taken together and consistent with previous studies (Garcia-Perez et al. 2007; van den Hurk et al. 2007; Kano et al. 2009; Wissing et al. 2012; Klawitter et al. 2016), our results demonstrate that L1 retrotransposition occurs in pluripotent cells of the early embryo, generates somatic and germline genetic mosaicism, and can give rise to heritable de novo L1 insertions.

Of the remaining two de novo L1 insertions, insertion #4 likely represented a late germline event. This insertion was identified by mRC-seq in mouse SREF16, one of seven offspring of parental mice SRE and SRF, but was not detected in somatic and germ tissues of SRE or SRF (Fig. 4A,B; Supplemental Table 2). Upon crossing
Having observed transmission of germline-restricted and somatic/germline mosaic insertions in mouse, consistent with previous studies of human patients (Brouha et al. 2002; van den Hurk et al. 2007), we next investigated a previously reported mutagenic human L1 insertion for which the developmental timing had not been resolved. The JH-27 insertion, which occurred in exon 14 of the Factor VIII gene, was identified in 1988 as the causative mutation in a case of noninherited hemophilia and was the first such example exhibiting L1 mobility in modern humans (Supplemental Fig. S9A; Kazazian et al. 1988). We performed a S′ junction nested PCR (55 total cycles) specific for the JH-27 insertion on blood genomic DNA from the afflicted patient and his mother, with paternal DNA serving as a negative control. While a robust PCR product was detected in the patient DNA, the JH-27 insertion could not be detected in the maternal sample (Supplemental Fig. S9B). Given our mouse data indicating each heritable early embryonic L1 insertion was detectable in tissues derived from all three germ layers (Fig. 3; Supplemental Fig. S6), we suggest that insertion JH-27 was very likely maternal germline-restricted.

**Discussion**

We uncovered 11 de novo L1 insertions among 85 mouse genomes, providing an estimate of one new insertion per eight mice (11/85 = 0.13 or ~1/8). This figure is consistent with but more conservative than previous estimations that a new L1 insertion may arise in every two to three mice (Kazazian and Moran 1998; Kazazian 2000) and is much higher than estimates of one new L1 insertion per 100 live births in humans (Hancks and Kazazian 2012). Indeed, the technical hurdles limiting detection of S′ L1-germline junctions may have precluded identification of additional, S′ truncated de novo L1 insertions in our pedigrees. Furthermore, while 85 genomes constitute the largest cohort of individual mice examined for de novo L1 insertions to date, future examination of more animals may allow fine-tuning of this rate estimate. The structure of our breeding pedigrees allowed us to observe the transmission of two new L1 insertions, #4 and #5, from heterozygous animals to their offspring (Fig. 1). We found that transmission of these insertions to 58% and 52% of offspring, respectively, was not...
significant differences from the expected transmission rate of 50% ($\chi^2$ test with one degree of freedom, two-tailed P-values of 0.49 and 0.67, respectively). Thus, as expected, within a single generation we observe no evidence for positive or negative selection on a de novo L1 insertion. However, future studies tracking the transmission of de novo L1 insertions through many generations may reveal evidence for selection on particular insertions, perhaps dependent on their genomic locations and functional consequences.

We established for the first time that heritable endogenous L1 retrotransposition events arise in early PGCS, before the PIWI/piRNA retrotransposon defense pathway becomes active in male embryonic gonads (Aravin et al. 2008), resulting in germ-line-restricted genetic mosaicism (Fig. 5A). Alternatively, it is possible that these insertions arose earlier during embryonic development, in cells of the primitive ectoderm that had been set aside for the germline and did not contribute to the somatic lineages (Soriano and Jaenisch 1986). It is worth noting that all three de novo insertions traced to early PGCS occurred in male mice. This correlation may stem from the relatively small number of insertions identified, and examination of more genomes may reveal equivalent events in female animals.

Consistent with previous studies (van den Hurk et al. 2007), we traced L1 insertions to the pluripotent cells of the early embryo and, in contrast to experiments using transgenic L1 reporter mice (Kano et al. 2009), we demonstrated germline transmission of four early embryonic insertions (Fig. 5B). All four transmitted insertions arose in somatic/germline mosaic female animals; again, future studies employing more animals may reveal transmission of similar events from mosaic males to offspring. Notably, deep mRC-seq of individual tissues from mosaic maternal mice SRA and SRE produced few reads (1–2) for insertions #3, #5, and #11 (Table 1; Supplemental Tables 1, 5), suggesting that additional mosaic insertions may have fallen below the detection threshold of mRC-seq. Thus, it is possible that early embryonic retrotransposition frequently generates low-level somatic-restricted mosaicism (Kano et al. 2009), as well as low-level somatic and germline mosaicism for insertions that ultimately

**Figure 3.** Retrotransposition in the early embryo. (A) Schematic of insertion #5 in the antisense orientation within the second intron of the gene Ano4 on Chr 10, and the donor element of insertion #5 on Chr 1. Features are annotated as in Figure 2A. (B) Genotyping panel for insertion #5. SRE (maternal) and SRF (paternal) tissues are indicated. (Somatic) Mix of liver, heart, and brain genomic DNA (R.O., L.O.) right ovary, left ovary; (R.T., L.T.) right testicle, left testicle. (C) Top: Control assay demonstrating the ability of the qPCR assay to detect mosaicism for insertion #5. Data are reported as the mean and standard deviation of four technical replicates per reaction. Bottom: Prevalence of insertion #5 among the tissues of maternal mouse SRE. Liver genomic DNA from mouse SREF18, a heterozygote for insertion #5, is set to 100%. Mouse SREF15, which lacks insertion #5, is included as a negative control. Data are reported as the mean and standard deviation of three independent qPCR experiments, each comprising four technical replicates per reaction.

**Figure 4.** Retrotransposition in the late germline. (A) Diagram of insertion #4. Features are depicted as in Figure 2A. Red arrows indicate the position of 3′ junction genotyping primers. (B) Genotyping panel for insertion #4. SRE and SRF are P-generation mice, SREF16 is the F1 maternal mouse, and 72-73, 111-112, and 1-10 are the F2 offspring of SRE and SRECD14. (Somatic) Mix of liver, heart, and brain genomic DNA (R.O., L.O.) right ovary, left ovary; (R.T., L.T.) right testicle, left testicle. Variation in size of the 3′ junction genotyping primers reflects shortening of the poly(A) tract of insertion #4, as previously reported (Grandi et al. 2013). (C) Prevalence of insertion #4 among tissues of maternal mouse SREF16. X-axis, from left: tissues of F1 mouse SREF16, ovaries and testes of P mice SRE and SRF, liver of 11 F1 mice heterozygous for insertion #4, and liver of two F1 mice lacking insertion #4. The average value among the 11 heterozygous F1 mice is set as a copy number of 1. Data are reported as the mean and standard deviation of three independent qPCR experiments, each comprising four technical replicates per reaction. (Top) Schematic of the probe-based qPCR assay used to quantify insertion #4 and empty site prevalence. The L1 insertion is shown in red; positions of the forward primer (P1) junction-spanning filled site reverse primer (P2), empty site reverse primer (P3), and hydrolysis probe (green) are indicated.
De novo heritable L1 retrotransposition is a component of the ongoing evolutionary interplay between retroelements and mammalian genomes, the importance of which is exemplified by recent studies demonstrating the exaptation of LTR retrotransposon sequences and protein products for pluripotency maintenance and embryonic development (Wang et al. 2014; Grow et al. 2015). Furthermore, recent publications have implicated somatic retrotransposition in the brain as a feature of both normal neurobiology and neurological diseases (Muotri et al. 2005, 2010; Coufal et al. 2009, 2011; Baillie et al. 2011; Evrony et al. 2012, 2015; Bundo et al. 2014; Upton et al. 2015; Erwin et al. 2016). Early embryonic insertions contributing broad mosaicism to tissues including the brain likewise represent a component of genetic neurodiversity, and the prevalence and potential functional consequences of such insertions compared to those occurring specifically in cells of the neuronal lineage remain to be determined. Future studies employing single-cell genomic analyses, and other approaches, will likely further elucidate the scope and consequences of ongoing retrotransposition in the mammalian germline and early embryo (Malki et al. 2014).

Methods

Animals

All animal breeding and handling procedures were carried out in compliance with the guidelines set forth by the University of Queensland Animal Ethics Committee. To establish breeding pedigrees, adult wild-type C57BL/6 mice were ordered from the University of Queensland Biological Resources Facility (UQ-BRF), which in turn sources animals from the Animal Resources Center (ARC, Western Australia). The UQ-BRF and ARC provided helpful information regarding the source and breeding history of the animals used in this study. The “P-generation” mice used to initiate the breeding pedigrees in this study were no more than 10 generations removed from the Jackson Laboratory C57BL6/J strain.

mRC-seq library construction

Genomic DNA from animals and tissues of interest was used to construct Illumina libraries for mRC-seq as described in Shukla et al. (2013), except using insert sizes of 450 and 550 bp. Illumina libraries were constructed using the Illumina TruSeq DNA LT kit or the Illumina TruSeq Nano DNA LT kit according to the manufacturer’s instructions (Illumina). A detailed description of library preparation can be found in Supplemental Methods.

mRC-seq hybridization reactions

Hybridization reactions were performed as described in Shukla et al. (2013), except using a pool of biotinylated capture probes designed against mouse retrotransposons represented by L1 subfamilies T1, G1, and A, SINEs B1 and B2, and the LTR elements IAP and EFn (Supplemental Table 1). Illumina libraries were pooled to achieve a total mass of 1 μg (Supplemental Table 1). A detailed description of mRC-seq hybridization can be found in Supplemental Methods.
L1 insertions were absent from the parental mice. In addition, some insertions were called as potentially de novo if they appeared in one or more offspring from the same pedigree and were absent from the parental mice. In addition, some insertions were detected in doublet tissues. Such insertions were treated as potentially de novo if the alignment spanned ≥33 nt of one contig end. Reads passing this filter were then aligned to mm10 using LAST and formed into clusters following an existing strategy. Clusters with ≥3 reads were then manually inspected for evidence of chimerism and annotated as polymorphic if found in an existing database of mouse polymorphisms (Nellaker et al. 2012) or in all of the libraries from at least one founder animal.

Validation and structural characterization of de novo L1 insertions
Putative insertions were called as potentially de novo if they appeared in one or more offspring from the same pedigree and were absent from the parental mice. In addition, some insertions were detected in doublet tissues. Such insertions were treated as potentially de novo and chimeric in the maternal mouse. Reads were then manually inspected using SerialCloner (http://serialcloner.sourceforge.net/) and the BLAT tool on the UCSC Genome Browser (Kent 2002). Reads which clearly represented molecular chimeras and those which could not be manually assigned to a specific genomic location due to repeat content were disregarded. For putative insertions passing manual inspection, primers were designed in the putative 5' and 3' flanking genomic DNA. Oligonucleotide primers were ordered from Integrated DNA Technologies (IDT).

Empty-filled validation PCRs were carried out using primers specific to the 5' and 3' genomic sequence flanking putative insertions. Validation PCRs for 5' and 3' junctions were carried out using the appropriate flanking genomic primer paired with a primer internal to the L1 sequence; where necessary, hemi-nested and fully nested PCR reactions were carried out using appropriately designed genomic and L1-specific primers. The full details of PCR validation can be found in Supplemental Methods.

Plasmid constructs
pTNT201 (Naas et al. 1998), TGF21 (Goodier et al. 2001), pJM101/L1.3 (Dombroski et al. 1993; Sassaman et al. 1997), and pJM105/L1.3 (Wei et al. 2000) were described previously. Descriptions of these constructs can be found in Supplemental Methods.

Generation of mouse L1 reporter constructs
Insertion #1 and insertion #7 were PCR-amplified using the Roche Expand Long Template PCR system and cloned into retrotransposition indicator vectors using standard molecular biology techniques. Detailed descriptions of the cloning strategies can be found in Supplemental Methods.

Cultured cell retrotransposition assay
HeLa-JVM cells were seeded at 2 x 10^4 cells/well in 6-well plates and transfected using FuGENE HD Transfection Reagent (Promega) at a ratio of 3 µL to 1 µg plasmid DNA. G418 selection (400 µg/mL) was initiated at 72 h post-transfection and carried out for 10–12 d (Wei et al. 2000).

Assays for transfection efficiency were performed in parallel by cotransfection of pcAG-EGFP with L1 reporter plasmids. At 48 h post-transfection, cells were subjected to flow cytometry on a Cyan ADP Analyzer (Beckman-Coulter) at the Translational Research Institute Flow Cytometry Core. The percentage of EGFP-positive cells for each L1 reporter construct was used to normalize the G418-resistant colony counts obtained in the retrotransposition assay (Wei et al. 2000; Kepera et al. 2016). Full details of the cultured cell retrotransposition assay can be found in Supplemental Methods.

Mosaicism analysis qPCR
Quantitative PCR using genomic DNA as template was carried out using primers and dual-labeled PrimeTime qPCR probes (5' 6-FAM-ZEN-3' Iowa Black FQ) from IDT. Control reactions for DNA input were performed using a predesigned PrimeTime qPCR assay for the single-copy mouse gene RPP25 (Mm.PT.58.21641426.g), with a dual-labeled probe (5' 6-FAM™-ZEN-3' Iowa Black FQ). Reactions were run on a Roche LightCycler 480 II with the following cycling conditions: 95°C, 5 min, followed by 45 cycles of 95°C for 10 sec, and 57°C for 1 min, then melt curve (0.11°C per sec from 57°C to 95°C). Ct values were calculated on the LightCycler software using absolute quantification 2nd derivative max. Details of mosaicism analysis qPCRs can be found in Supplemental Methods.

Data access
mRC-seq and WGS data from this study have been submitted to the European Nucleotide Archive (ENA; http://www.ebi.ac.uk/ena) under project accession number PRJEB10299. Sanger trace files from this study have been submitted to the NCBI Trace Archive (http://www.ncbi.nlm.nih.gov/Traces/home/index.cgi) with Trace numbers (TI2344112704–TI2344112736 and TI234412752).
Competing interest statement

J.A.J. is employed by Roche Sequencing Solutions, Inc., and Roche Sequencing Solutions reagents were used in the study.

Acknowledgments

We thank Dr. John V. Moran for the constructs pJM101/L.1.3 and pJM105/L.1.3 and for the HeLa-JVM cell line. We thank Dr. John L. Goodier for assistance in providing samples and for helpful discussion. We thank Dr. John V. Moran and Dr. Sara R. Heras for critically reading the manuscript. We thank Ms. Carmen Salvador-Palomque and Dr. Santiago Morell for helpful contributions and discussion. G.J.F. acknowledges the support of an National Health and Medical Research Council (NHMRC) Research Fellowship (GNT1106214). Work in the Faulkner laboratory was funded by Australian NHMRC Project Grants GNT1044249, GNT1045991, GNT1067983, and GNT1106206, the Mater Foundation, and the European Union’s Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 259743 underpinning the MODHEP consortium. Work in the Kazanov laboratory was funded by an R01 (1R01GM099875-01) and a P50 from the National Institute of General Medical Sciences. Work in the Garcia-Perez laboratory was supported by CICE-FEDER-P12-CTS-2256, Plan Nacional de I+D+I 2008-2011 and Ministerio de Ciencia e Innovación 2013-2016 (FIS-FEDER-P11/01489 and FIS-FEDER-P14/02152), PCIN-2014-115-ERA-NET NEURON II, the European Research Council (ERC-Consolidator ERC-STG-2012-233764), by an International Early Career Scientist grant from the Howard Hughes Medical Institute (IECS-S5007420), and by The Wellcome Trust–University of Edinburgh Institutional Strategic Support Fund (ISF2). P.G. was supported by a University of Queensland International Postgraduate Research Scholarship (IPRS) and an Australian Postgraduate Award (APA). F.J.S.-L. was supported by the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. PIOF-GA-2013-623324. G.-O.B. was supported by a German Research Foundation (DFG)–Germany postdoctoral fellowship (BO4460/1-1). A.D.E. was supported by an Australian Research Council Discovery Early Career Researcher Award (DE150101117).


References


Hancks DC, Kazazian HH Jr. 2012. Active human retrotransposons: varia-


Holmes SE, Dombroski BA, Krebs CM, Boehm CD, Kazazian HH Jr. 1994. A newly acquired human L1 element from D246D2 on chro-

Jurka J. 1997. Sequence patterns indicate an enzymatic involvement in in-


Kohrman DC, Harris JB, Meisler MH. 1996. Mutation detection in the med-


Morgan JV, DeBerardinis RJ, Kazazian HH Jr. 1999. Exon shuffling by L1 ret-


Muratori AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. 2005. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotrans-


Naas TP, DeBerardinis RJ, Moran JV, Ostertag EM, Kingsmore SF, Seldin MF, Hayashizaki Y, Martin SL, Kazazian HH. 1998. An actively retrotranspos-

nerv skalfamily of mouse L1 elements. EMBO J 17: 590–597.

Developmental origins of heritable L1 insertions

Received December 4, 2016; accepted in revised form May 2, 2017.


Heritable L1 retrotransposition in the mouse primordial germline and early embryo

Sandra R. Richardson, Patricia Gerdes, Daniel J. Gerhardt, et al.

*Genome Res.* published online May 8, 2017
Access the most recent version at doi:10.1101/gr.219022.116

Supplemental Material [http://genome.cshlp.org/content/suppl/2017/06/28/gr.219022.116.DC1](http://genome.cshlp.org/content/suppl/2017/06/28/gr.219022.116.DC1)

P<Sup>2<Sup>
Published online May 8, 2017 in advance of the print journal.

Open Access Freely available online through the *Genome Research* Open Access option.

Creative Commons License This article, published in *Genome Research*, is available under a Creative Commons License (Attribution 4.0 International), as described at [http://creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/).

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Genome Research* go to: [http://genome.cshlp.org/subscriptions](http://genome.cshlp.org/subscriptions)

© 2017 Richardson et al.; Published by Cold Spring Harbor Laboratory Press