

The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes

SANDRA R. RICHARDSON,¹ AURÉLIEN J. DOUCET,¹
HUIRA C. KOPERA,¹ JOHN B. MOLDOVAN,²
JOSÉ LUIS GARCIA-PEREZ,³ and JOHN V. MORAN⁴

¹Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109; ²Cellular and Molecular Biology Graduate Program, University of Michigan Medical School, Ann Arbor, MI 48109; ³Department of Human DNA Variability, GENYO (Pfizer-University of Granada & Andalusian Regional Government Genomics & Oncology Center), 18016 Granada, Spain; ⁴Department of Human Genetics; Cellular and Molecular Biology Graduate Program, University of Michigan Medical School, Ann Arbor, Michigan 48109

ABSTRACT Transposable elements have had a profound impact on the structure and function of mammalian genomes. The retrotransposon Long INterspersed Element-1 (LINE-1 or L1), by virtue of its replicative mobilization mechanism, comprises ~17% of the human genome. Although the vast majority of human LINE-1 sequences are inactive molecular fossils, an estimated 80–100 copies per individual retain the ability to mobilize by a process termed retrotransposition. Indeed, LINE-1 is the only active, autonomous retrotransposon in humans and its retrotransposition continues to generate both intra-individual and inter-individual genetic diversity. Here, we briefly review the types of transposable elements that reside in mammalian genomes. We will focus our discussion on LINE-1 retrotransposons and the non-autonomous Short INterspersed Elements (SINEs) that rely on the proteins encoded by LINE-1 for their mobilization. We review cases where LINE-1-mediated retrotransposition events have resulted in genetic disease and discuss how the characterization of these mutagenic insertions led to the identification of retrotransposition-competent LINE-1s in the human and mouse genomes. We then discuss how the integration of molecular genetic, biochemical, and modern genomic technologies have yielded insight into the mechanism of LINE-1 retrotransposition, the impact of LINE-1-mediated retrotransposition events on mammalian genomes, and the host cellular mechanisms that protect the genome from unabated LINE-1-mediated retrotransposition events. Throughout this review, we highlight unanswered questions in LINE-1 biology that provide exciting opportunities for future research. Clearly, much has been learned about LINE-1 and SINE biology since the publication

of *Mobile DNA II* thirteen years ago. Future studies should continue to yield exciting discoveries about how these retrotransposons contribute to genetic diversity in mammalian genomes.

INTRODUCTION

Transposable elements (TEs) or “jumping genes” historically have been disparaged as a class of “junk DNA” in mammalian genomes (1, 2). The advent of whole genome DNA sequencing, in conjunction with molecular genetic, biochemical, and modern genomic and functional studies, is revealing that TEs are biologically important components of mammalian genomes. TEs are classified by whether they mobilize via a DNA or an RNA intermediate (detailed in reference 3). Classical DNA transposons, such as the maize Activator/Dissociation elements

Received: 31 October 2014, **Accepted:** 3 November 2014,
Published: 5 March 2015

Editors: Alan Lambowitz, University of Texas, Austin, TX, and Nancy Craig, Johns Hopkins University, Baltimore, MD

Citation: Richardson SR, Doucet AJ, Kopera HC, Moldovan JB, Garcia-Perez JL, Moran JV. 2014. The influence of LINE-1 and SINE retrotransposons on mammalian genomes. *Microbiol Spectrum* 3(2):MDNA3-0061-2014. doi:10.1128/microbiolspec.MDNA3-0061-2014.

Correspondence: John Moran, moranj@umich.edu

© 2014 American Society for Microbiology. All rights reserved.

originally discovered by Barbara McClintock, move via a DNA intermediate (4, 5). Their mobility (*i.e.*, transposition) can impact organism phenotypes such as corn kernel variegation. Retrotransposons, the predominant class of TEs in most mammalian genomes, mobilize via an RNA intermediate by a process termed retrotransposition (6).

The completion of the human genome reference sequence (HGR) (7, 8) confirmed the results of DNA hybridization-based re-annealing studies (9, 10) and revealed that retrotransposons have been a major force in shaping the structure and function of mammalian genomes. The mobility of non-long terminal repeat (non-LTR) retrotransposons, namely autonomously active Long INterspersed Element-1 sequences (LINE-1s, also known as L1s) and non-autonomous Short INterspersed Elements (SINEs, such as Alu and SINE-R/VNTR/Alu-like retrotransposons [SVA elements]) continues to generate both intra-individual and inter-individual genetic variation in the human population (reviewed in reference 11). Remarkably, LINE-1s and SINEs constitute at least one-third (~1 billion bp) of human genomic DNA (7).

Since the publication of *Mobile DNA II* in 2002, much has been learned about the mechanism of LINE-1 and SINE retrotransposition, the impact of these elements on the human genome, and how the cell defends itself from unabated retrotransposition. A number of outstanding reviews have been published that discuss advances in each of the above areas (for example see refs. [11–19]). Here, we provide a brief background about the types of TEs in the human genome. Our discussions then focus on major advances in LINE-1 and SINE biology that have occurred in the past 13 years.

TRANSPOSABLE ELEMENTS IN MAMMALIAN GENOMES

Overview

Genomes are not simply static catalogues of genes. Instead, they are ever changing, genetically dynamic entities. While a typical cellular gene resides at a discrete chromosomal locus, TEs are present in multiple copies that reside at numerous genomic locations. TEs can invade new chromosomal locations, often increasing their copy number in the genome. The evolutionary impact of TEs is readily apparent by examining recently completed mammalian whole genome DNA sequences. For example, at least 46% of human DNA, 31% of canine DNA, and 37% of mouse DNA are derived from TEs (7, 20, 21). Computer algorithms that allow the detection of ancient, highly mutated TEs suggest that TEs could possibly account for as much as 70% of human genomic DNA (22).

Most individual TE-derived sequences in mammalian genomes cannot mobilize because they have been riddled by mutations over the course of evolution; hence, they can be considered as molecular fossils (7, 20, 21). It is now well established that most mammalian genomes contain active LINE-1 and SINE retrotransposons. Other chapters in *Mobile DNA III* discuss the impact of DNA transposons and LTR retrotransposons on mammalian genomes. Here, we briefly discuss the structure and abundance of DNA transposons and LTR retrotransposons in the human genome. We focus the remainder of this chapter on how non-LTR retrotransposons mobilize and how the resultant insertions contribute to disease, genetic variation, and genomic evolution.

DNA transposons: abundance and structure in the human genome

DNA transposons can move (*i.e.*, transpose) to new genomic locations via a DNA intermediate (3) and they comprise approximately 3% of the human genome (7). In their simplest form, DNA transposons contain inverted terminal repeat sequences that surround an open reading frame that encodes a protein (*i.e.*, transposase) (3). During a round of transposition, transcription of a DNA transposon is initiated from sequences within the transposable element. The resultant mRNA is transported to the cytoplasm where it undergoes translation, leading to the synthesis of a transposase protein. A generic transposase protein, such as that encoded by the eukaryotic Sleeping Beauty DNA transposon (23), contains a nuclear localization signal as well as DNA binding and integrase activities. After translation, transposase is imported into the nucleus, where it binds either within or near the transposon inverted terminal repeat sequences to promote transposition by a “cut and paste” (or in some cases, a replicative “copy and paste”) mechanism. A given transposase protein can mobilize both protein coding (*i.e.*, autonomous) and non-protein coding (*i.e.*, non-autonomous) DNA transposons to new genomic locations (3). As a consequence of transposase activity, the newly inserted DNA transposon is generally flanked by short target-site duplications of a defined size for a given class of element (3, 24).

DNA transposons are active in numerous organisms, (*e.g.*, insertion sequences, or IS elements, in bacteria [reviewed in reference 25], P-elements in *Drosophila* [reviewed in reference 26], Activator/Dissociation elements in maize [reviewed in reference 27], and PiggyBac elements from the cabbage looper moth [28–30]). Due to their mutagenic potential, DNA transposons have been exploited as genetic tools for various molecular biological

applications (reviewed in references 31–33). Although it was assumed that DNA transposons were no longer active in mammalian genomes, studies have identified potentially active DNA transposons (e.g., *hobo-Ac-TAM* [hAT] and helitron elements) in the genomes of certain bat species, suggesting that their transposition may continue to generate genetic diversity in these species (34–36). In contrast, virtually all human DNA transposons are mutated, incapable of transposition, and can be considered as molecular fossils (7, 37). Despite this fact, it is clear that DNA transposons have made enduring contributions to mammalian genomes over evolutionary time and that the recurrent domestication of transposase-derived genes has led to genomic innovations (7, 38). For example, the recombination activating gene-1 (*RAG-1*) and *RAG-2*, which are critical for V(D)J recombination and immune system development, may have been co-opted from an ancient Transib DNA transposon (39). Moreover, recent studies also have revealed that the human *THAP9* gene encodes a protein that is capable of mobilizing *Drosophila* P elements to new genomic locations in both *Drosophila* and human cells (40).

LTR retrotransposons in the human and mouse genomes

Long terminal repeat retrotransposons (also known as endogenous retroviruses or ERVs) and their non-autonomous derivatives are present at >450,000 copies in the human genome and comprise approximately 8% of human DNA (7). ERVs resemble simple retroviruses in structure. They retrotranspose via a “copy and paste” mechanism, but generally contain a nonfunctional envelope gene or lack the gene completely, which relegates them to an intracellular fate (reviewed in reference 41).

Virtually all ERVs in the human genome have been rendered inactive by mutations and cannot undergo autonomous retrotransposition (7). Certain human-specific ERVs (HERVs) from the HERV-K subfamily (where the K denotes the lysine tRNA needed to prime (–) strand cDNA synthesis from an ERV RNA template) are polymorphic with respect to presence in the human population (42–46). This fact, coupled with the identification of polymorphic ERVs in both the chimpanzee and gorilla genomes (47), suggests that ERVs have been active since the divergence of humans and chimpanzees. Despite concentrated efforts, no one has reported the identification of an active HERV. However, recent studies have demonstrated that HERV-K proviruses that have been reanimated using recombinant DNA technology are infectious in cultured human cells (48, 49). Since retrovirus-encoded proteins can work efficiently *in trans*,

it is formally possible that *trans*-complementation might allow the assembly of functional virus-like particles from partially defective HERVs, allowing the generation of new retrotransposition events. Advances in DNA sequencing technologies may reveal rare, active HERV-K elements or *de novo* germline or somatic HERV-K retrotransposition events in individual human genomes.

LTR-retrotransposons are present at >600,000 copies in mouse DNA and comprise approximately 10% of the genome (21). In contrast to the human genome, the mouse genome contains multiple, active ERV subfamilies (50, 51). These include autonomously active MusD and intracisternal A particle (IAP) elements, as well as non-autonomous early transposons (ETns) and mammalian apparent LTR retrotransposons (MaLRs). It is estimated that ERV insertions are responsible for approximately 10% of spontaneously arising mouse mutations (reviewed in reference 51) (discussed in greater detail in other chapters of *Mobile DNA III*).

LINE-1 retrotransposons: abundance and structure

A brief overview of human LINE-1 evolution and nomenclature

LINE-1 retrotransposons have been amplifying in mammalian genomes for more than 160 million years (52–54). In humans, the vast majority of LINE-1 sequences have amplified since the divergence of the ancestral mouse and human lineages approximately 65 to 75 million years ago (7). As a consequence, LINE-1-derived sequences now account for approximately 17% of human genomic DNA (7) (Figure 1).

Sequence comparisons between individual genomic LINE-1 sequences and a consensus sequence derived from modern, active LINE-1s can be used to estimate the age of genomic LINE-1s. These analyses uncovered 16 LINE-1 primate-specific subfamilies (termed PA1 to PA16) (53, 55). These subfamilies have a monophyletic origin, suggesting that older LINE-1 subfamilies are replaced over evolutionary time by the emergence of new LINE-1 subfamilies—a phenomenon known as subfamily succession. Indeed, recent studies suggest that host proteins that restrict LINE-1 expression may drive subfamily succession (56, 57) (see below).

Functional analyses have revealed that only certain human-specific LINE-1s (termed L1Hs elements) from the PA1 subfamily remain retrotransposition-competent (58–61). The majority of active L1Hs elements belong to a small population of elements termed the transcribed-active subset (Ta-subset) (62, 63). Ta-subset LINE-1s

Non-LTR retrotransposons in the human genome

		size	copy number	% of genome	active subfamilies
LINE-1		~6 kb	516,000	17%	L1 PA-1, Ta-1, Ta-0, and pre-Ta
Alu		~300 bp	1,090,000	11%	Alu Y (Ya5, Yb8, and Yd8) (Alu Sx ?)
SVA		~2 kb	2,700	0.2%	SVA-E, SVA-F1 (SVA-D, SVA-F ?)
Processed pseudogene		variable	8,000 to 17,000		

Non-LTR retrotransposons in the mouse genome

		size	copy number	% of genome	active subfamilies
LINE-1		~7 kb	599,000	18%	L1-A, L1-T _F and L1-G _F
B1		~135 bp	564,000	2.7%	
B2		~200 bp	348,000	2.4%	
Processed pseudogene		variable	5,000		

contain a diagnostic 5'-ACA-3' trinucleotide sequence and G nucleotide in their 3' untranslated regions (UTRs) (at positions 5930–5932 and 6015, respectively, based on the sequence of L1.2, an active LINE-1: GenBank accession number M80343 [64]). The diagnostic ACA and G nucleotides allow the discrimination of Ta-subset LINE-1s from older, retrotransposition-incompetent LINE-1s, which generally contain a GAG trinucleotide at the same position of their respective 3' UTRs. The exploitation of the ACA trinucleotide present in Ta-subset LINE-1s has been instrumental in allowing the identification of polymorphic, human-specific LINE-1 insertion alleles in both the HGR and individual human genomic sequences (7, 63, 65–72).

The Ta-subset of LINE-1s can be further grouped into finer subdivisions (e.g., Ta-1, Ta-0, and pre-Ta) based on more subtle sequence distinctions (63). Of these, the Ta-1 subset contains the greatest number of active LINE-1s, followed by the Ta-0 and pre-Ta subfamilies (58, 59, 61, 63). Notably, LINE-1s from the Ta-subset are responsible

for most of the retrotransposition activity in modern human genomes (reviewed in reference 11).

Active human LINE-1s: abundance

The vast majority of LINE-1-derived sequences in the HGR predate the emergence of the human lineage, are “fixed” with respect to presence in the human population, and have been rendered inactive by 5' truncations, internal rearrangements (such as inversion/deletion events), and point mutations that prevent the production of active forms of the LINE-1 encoded proteins (ORF1p and ORF2p) (7, 73). The completion of the HGR, in combination with functional assays to assess retrotransposition potential in cultured human HeLa cells, has revealed that the average human genome contains approximately 80 to 100 LINE-1s that remain retrotransposition-competent (59, 61). Intriguingly, only a handful of Ta-subset LINE-1s (termed “hot L1s”) were found to account for the bulk of retrotransposition activity in a given genome (59). Subsequent

FIGURE 1 Non-long terminal repeat (LTR) retrotransposons of the human and mouse genomes. The top and bottom panels represent non-LTR retrotransposons in the human and mouse genomes, respectively. Each non-LTR retrotransposon is listed with its name, structure, average size, copy number, percentage of the genome reference sequence occupied by the element, and, if applicable, the active subfamilies (question marks [?] denote uncertainty in whether Alu Sx and SVA-D and F elements are active *in vivo*). Details of the structure and abbreviations for human and mouse Long INterspersed Element-1 retrotransposons (LINE-1s): Untranslated regions (UTRs) (gray boxes); sense and antisense internal promoters (black arrows); monomeric repeats (white triangles) are followed by an untranslated linker sequence (white box) just upstream of open reading frame 1 (ORF1) in the mouse 5' UTR; ORF1 (yellow box for human LINE-1; brown box for mouse LINE-1) includes a coiled-coil domain (CC), an RNA recognition motif (RRM), and a C-terminal domain (CTD); inter-ORF spacer (gray box between ORF1 and ORF2); ORF2 (blue boxes) includes endonuclease (EN), reverse transcriptase (RT), and cysteine-rich domains (C); poly (A) tract (A_n downstream of 3' UTR). For human Alu: 7SL-derived monomers (orange boxes); RNA polymerase III transcription start site (black arrow) and conserved *cis*-acting sequences required for transcription (A and B white boxes in left 7SL-derived monomer); adenosine-rich fragment (AAA gray box between left and right 7SL-derived monomers); terminal poly (A) tract (AAAA gray box); variable sized flanking genomic DNA (interrupted small gray box) followed by the RNA pol III termination signal (TTTT). For human SVA: hexameric CCCTCT repeat ((CCCTCT) $_n$ light green box); inverted Alu-like repeat (green box with backward arrows); GC-rich VNTR (striped green box); SINE-R sequence sharing homology with HERV-K10, (envelope [ENV] and LTR); cleavage polyadenylation specific factor (CPSF) binding site; terminal poly (A) tail (A_n). For human and mouse processed pseudogenes: spliced cellular mRNA with UTR (gray boxes) and coding ORF (red boxes for human and purple boxes for mouse, boxes are interrupted by exon–exon junctions [vertical black lines]). For mouse B1 and B2: 7SL-derived monomer (light orange box) or tRNA derived sequence (dark orange box); RNA pol III transcription start site (black arrow) and conserved *cis*-acting sequences required for transcription (A and B white boxes); terminal poly (A) tract (AAAA dark gray box); variable sized flanking genomic DNA (interrupted gray box) followed by the RNA polymerase III termination signal (TTTT). The 3' end of B2 also contains a non-tRNA derived sequence (3' domain light gray box). Mouse ID and B4 elements are not represented in the figure. References are provided in the text. [doi:10.1128/microbiolspec.MDNA3-0061-2014.f1](https://doi.org/10.1128/microbiolspec.MDNA3-0061-2014.f1)

genomic and computationally based studies have revealed that Ta-subset LINE-1s, which are present at low allele frequencies in the human population, are highly enriched with active LINE-1s (58). When considering that there are more than 7 billion people on the planet, these studies raise the possibility that there are perhaps millions of rare, active LINE-1 alleles in the human population (58). When combined with the number of active non-autonomous non-LTR retrotransposons, it is undeniable that there are many more active retrotransposons in the human population than previously thought.

Active Human LINE-1s: structure

A retrotransposition-competent human LINE-1 (RC-L1) is ~6 kb in length (64, 74) (Figure 1). An RC-L1 encodes a 5' UTR containing an internal RNA polymerase II sense strand promoter (75), as well as an antisense promoter of unknown function (76). The 5' UTR is followed by two open reading frames (ORF1 and ORF2) and a 3' UTR that terminates in a poly (A) tract (64). Human ORF1 encodes a 40 kDa RNA binding protein (ORF1p) that has nucleic acid chaperone activity (77–80). Human ORF2 encodes a 150 kDa protein (ORF2p) with demonstrated endonuclease (L1 EN) (83) and reverse transcriptase (L1 RT) activities (84). The ORF2p C-terminus contains a cysteine-rich domain of unknown function (60, 85). Functional studies have revealed that both ORF1p and ORF2p are required for retrotransposition (60, 83).

Mouse LINE-1s: abundance and structure

The mouse genome is replete with LINE-1-derived sequences and they comprise approximately 18% of mouse DNA (21) (Figure 1). As in the HGR, the vast majority of LINE-1s in the *Mus musculus* reference genome have been rendered inactive by mutation (21). Functional studies in cultured cells have revealed that approximately 3,000 mouse LINE-1s remain retrotransposition-competent (86).

Mouse LINE-1s are structurally similar to human LINE-1s; however, mouse LINE-1s contain different RNA polymerase II promoter structures at their respective 5' ends (Figure 1). The mouse LINE-1 promoter consists of a series of repeats and is followed by an untranslated linker sequence located immediately upstream of ORF1 (reviewed in reference 87). As in humans, mouse LINE-1 sequences can be stratified into subfamilies: at least five structurally different subfamilies (V, F, A, T_F, and G_F) of mouse LINE-1s exist, and they differ in the DNA sequences of the monomeric repeats present in their respective 5' UTRs (88). LINE-1s from

the V and F subfamilies are thought to be inactive (89), whereas LINE-1s from the A, T_F, and G_F subfamilies remain retrotransposition-competent (86, 90, 91).

Non-autonomous retrotransposons: prevalence and structure

The proteins encoded by LINE-1s (ORF1p and/or ORF2p) can act *in trans* to mobilize non-autonomous retrotransposons (*e.g.*, human Alu and SVA elements and mouse B1 and B2 elements) (92–96) and cellular mRNAs to new genomic locations, with the latter giving rise to processed pseudogenes (97, 98) (Figure 1). As expected, each of the above elements contains structural hallmarks that are consistent with being mobilized by the LINE-1 encoded proteins (reviewed in reference 11). They generally: (i) terminate with a poly (A) or adenosine-rich sequence; (ii) are flanked by target-site duplications that vary both in their size and sequence; and (iii) integrate at a LINE-1 endonuclease consensus site in genomic DNA (99). A brief description of each of the major types of non-autonomous retrotransposons present in the human and mouse genomes is provided below.

Human Alu elements

Alu elements represent the most abundant class of human non-autonomous retrotransposons (7). They are present at more than 1 million copies in the HGR and comprise approximately 11% of genomic DNA (7) (Figure 1). Alu elements are approximately 300 bp (100) in length and are derived from the 7SL RNA component of the signal recognition particle (101). A full-length Alu element exhibits a dimeric structure that consists of highly similar left and right monomers that are separated by an adenosine-rich linker sequence. The left Alu monomer contains conserved A and B box sequences that are required for RNA polymerase III dependent transcription (102). The right monomer lacks conserved A and B boxes and ends in a poly (A) tract. Genomic DNA flanking the 3' ends of active Alu elements contains an RNA polymerase III terminator sequence (*i.e.*, a stretch of four to six consecutive thymidines) (102, 103). Both the size of an Alu poly (A) tract and genomic sequences that reside downstream, and in some cases upstream, of an individual Alu element can influence its expression and retrotransposition potential (104–109).

Like LINE-1s, Alu elements can be stratified into three major subfamilies: Alu J, S, and Y (110–114); these subfamilies can be further grouped into finer subdivisions (103, 115). Alu J represents the oldest Alu lineage; its retrotransposition activity peaked approximately 65 million years ago (reviewed in reference 103). Alu S

represents the second oldest lineage; its retrotransposition activity peaked approximately 30 million years ago (reviewed in reference [103](#)). Alu Y elements represent the youngest Alu lineage (reviewed in reference [103](#)). The analysis of individual human DNA sequences, in conjunction with molecular genetics and genomics-based approaches, has revealed that certain Alu Y subdivisions (*e.g.*, Alu Ya5, Alu Yb8, and Alu Yd8) are polymorphic with respect to presence in the human population ([116](#), [117](#)). Some of these Alus may represent “source elements” or “master genes” that presently are amplifying in the human population. Indeed, the identical-by-descent mode of transmission, homoplasmy-free nature, and directionality of integration (*i.e.*, the absence of the element represents the ancestral state) have allowed the use of polymorphic Alu elements and LINE-1s as genetic markers in population-based and phylogenetic studies ([118–123](#)) (reviewed in reference [124](#)).

Although it is clear that the Alu Y subfamily is responsible for the bulk of Alu retrotransposition in the human genome, studies have demonstrated that an average human genome may contain thousands of active Alu “core elements” derived from both the Alu Y and Alu S subfamilies ([125](#)) and that interactions between Alu RNA and the signal recognition proteins (SRP9 and SRP14) are required to undergo efficient retrotransposition ([125](#), [126](#)).

Human SVA elements

SVA elements comprise an evolutionarily young, non-autonomous retrotransposon family that arose in primate lineages approximately 25 million years ago ([127–129](#)). SVA elements are present at approximately 2,700 copies in the HGR, comprising approximately 0.2% of human genomic DNA ([Figure 1](#)). A typical SVA element is approximately 2,000 bp in length and has a composite structure that consists of: (i) a hexameric CCCTCT repeat; (ii) an inverted Alu-like element repeat; (iii) a set of GC-rich variable nucleotide tandem repeats (VNTRs); (iv) a SINE-R sequence that shares homology with HERV-K10, an inactive LTR retrotransposon; and (v) a canonical cleavage polyadenylation specificity factor binding site that is followed by a poly (A) tract (reviewed in reference [129](#)) ([Figure 1](#)). This latter feature suggests that SVA elements are transcribed by RNA polymerase II; whether SVA elements contain an internal RNA polymerase II promoter is not yet known. SVA elements also can be stratified into subfamilies based on sequence similarities. For example, approximately 40% of SVA elements from the youngest SVA subfamilies (SVA-D, SVA-E, SVA-F, and SVA-F1) are polymorphic with respect to presence in the human population, suggesting that members of these

subfamilies may retain the ability to retrotranspose and contribute to human genetic diversity ([129](#)). Interestingly, recent reports suggest that certain SVA elements can influence reporter gene expression in *in vitro* assays ([56](#), [130](#)); so, in principle, SVA elements might alter the expression of neighboring genes.

Gibbons and certain other nonhuman primates contain composite retrotransposons called LINE-Alu-VNTR-Alu (LAVA) elements ([131–133](#)). LAVA elements are similar in structure to SVA elements, suggesting that they are mobilized *in trans* by the LINE-1 encoded proteins. In contrast to SVA elements, LAVA elements lack a HERV-K sequence and instead contain a sequence motif that consists of unique DNA, as well as sequences derived from ancient Alu S and LINE-1 elements ([131–133](#)). Interestingly, LAVA elements have undergone a massive expansion in gibbon genomes. Recent reports demonstrate that LAVA elements are concentrated at centromeres and may contribute to gibbon genomic plasticity ([134](#)).

Mouse B1 and B2 elements and other SINEs

The completion of the *Mus musculus* reference genome revealed the presence of various abundant classes of non-autonomous SINEs ([21](#)). The first, B1, is an ~135 bp, 7SL-derived monomeric SINE ([135](#), [136](#)) ([Figure 1](#)). B1 elements have amplified to more than 500,000 copies and comprise approximately 2.7% of mouse DNA ([21](#)). Like Alu elements, an active B1 element contains conserved A and B box sequences that are required for RNA polymerase III dependent transcription and end in a poly (A) tract that is flanked by genomic DNA sequences containing an RNA polymerase III terminator. The second, B2, is an ~200 bp, tRNA-derived SINE ([137](#)) ([Figure 1](#)). B2 elements have amplified to more than 300,000 copies and comprise approximately 2.4% of mouse DNA ([21](#)). The tRNA-derived region of B2 contains conserved A and B box sequences that are required for RNA polymerase III dependent transcription. The 3' end of B2 ends in a poly (A) tract and is flanked by genomic DNA sequences containing an RNA polymerase III terminator ([137](#)). A third class of mouse SINE, the ID elements, appears to be derived from a tRNA; an ID-element-derived RNA (BC1 RNA) is expressed abundantly in the nervous system ([137](#), [138](#)). A fourth class, known as B4 elements, appears to represent an ancient hybrid SINE derived from a fusion between B1 and ID elements and does not seem to be active in the mouse genome ([139](#), [140](#)). Although cell culture-based assays have demonstrated the existence of active B1 and B2 SINEs in the mouse genome ([93](#)), the number of active elements remains unknown.

Processed pseudogenes

LINE-1 ORF1p and ORF2p occasionally can act *in trans* to mobilize matured cellular mRNAs to new genomic locations, leading to the formation of processed pseudogenes that bear LINE-1 structural hallmarks (97, 98, 141, 142) (Figure 1). Recent studies have revealed that there may be at least 8,000 to 17,000 processed pseudogenes in the HGR (143–145). The majority of processed pseudogenes appear to be derived from housekeeping genes or ribosomal protein-encoding genes. Whole genome DNA sequence studies have further revealed the presence of segregating processed pseudogene insertions in individual genomes, indicating that processed pseudogene formation continues to contribute to inter-individual genomic diversity (146–150). Finally, several studies have demonstrated how processed pseudogenes can evolve to acquire a cellular function (reviewed in reference 151). In extreme cases, this process can lead to the evolution of new genes that participate in host defense against exogenous viruses. For example, independent LINE-1-mediated retrotransposition events of cyclophilin mRNA into the TRIM5 locus have led to the production of a TRIM5/cyclophilin fusion protein that can restrict the mobility of HIV in owl monkeys (152, 153).

Other cellular RNAs mobilized by the LINE-1-encoded proteins

The LINE-1-encoded proteins have been implicated in the mobility of other cellular RNAs, such as uracil-rich small nuclear RNAs (snRNAs) (*i.e.*, U6, U6atac, and to a lesser extent U1, U2, U4, and U4atac) and small nucleolar RNAs (snoRNAs) (*i.e.*, U3 snoRNA) (154–158). In contrast to the mechanism of processed pseudogene formation, the structures of chimeric U6/L1 pseudogenes suggest that U6 snRNA was reverse transcribed onto a 5' truncated LINE-1 cDNA during the process of LINE-1 integration (155) (see below). Chimeric U6/LINE-1 pseudogenes have been identified in many primate genomes (159). Moreover, the recapitulation of U6/LINE-1 chimeric pseudogene formation in cultured human cells strongly supports the hypothesis that U6/LINE-1 pseudogene formation is ongoing in the human population (156, 157).

LINE-1-mediated retrotransposition events and human disease

A brief historical perspective

Historically, it was assumed that LINE-1s could be dismissed as a class of repetitive “junk DNA”. This view

changed radically in 1988, when the Kazazian laboratory identified two independent, mutagenic LINE-1 insertions into the *Factor VIII* genes of unrelated boys afflicted with hemophilia A (160). Although those LINE-1 insertions were predicted to be inactive (*i.e.*, one was 5' truncated, while the other was both 5' truncated and internally rearranged), DNA sequencing revealed that each of the disease-producing insertions likely was derived from an active full-length progenitor LINE-1 that contained intact open reading frames. The derivation of an oligonucleotide probe that specifically recognized the disease-producing LINE-1 insertion and related elements subsequently led to the identification of a cohort of full-length LINE-1s with intact open reading frames (61, 64, 161). One of these LINE-1s (L1.2B), isolated from genomic DNA derived from the mother of one of the hemophiliac patients, was identical (over the ~3.8 kb length of insertion) to the disease-producing LINE-1 insertion in her son (64). The demonstration that an allele of L1.2 (L1.2A) could retrotranspose in cultured human cells subsequently confirmed that LINE-1s were active in the human genome (60, 162).

Disease-producing LINE-1-mediated insertions

Mutagenic LINE-1 retrotransposition events have been implicated in at least 25 cases of human disease, including Duchenne muscular dystrophy, hemophilia B, chronic granulomatous disease, X-linked retinitis pigmentosa, and β -thalassemia trait (reviewed in reference 14). In general, these LINE-1 retrotransposition events either disrupt coding exons or occur into introns, which can result in mis-splicing or exon skipping, and lead to the generation of null or hypomorphic expression alleles (reviewed in references 11, 151). As predicted by Maxine Singer during her studies of LINE-1 expression in human teratocarcinoma cell lines (62, 163), all but two of the 25 disease-producing LINE-1 insertions were derived from the LINE-1 Ta-subset. The remaining insertions were derived from the slightly older pre-Ta subset (which contains an ACG at nucleotide positions 5930–5932 as opposed to the ACA in L1.2 [64]). These observations, as well as functional studies (reviewed in reference 11), provide evidence that the LINE-1 Ta-subfamily comprises the bulk of active LINE-1s in the human genome and that some pre-Ta LINE-1s remain active.

LINE-1-mediated non-autonomous retrotransposition events also have caused sporadic cases of human disease. For example, deleterious Alu retrotransposition events are responsible for over 60 disease-producing mutations in man (14, 164). Almost all of the disease-producing insertions are derived from members of the

Alu Y subfamily, and the majority of these belong to the Alu Ya5 and Alu Yb8 subdivisions. Similarly, deleterious SVA retrotransposition events (derived from the SVA-E and SVA-F1 subfamilies) are responsible for at least 10 disease-producing mutations in man (14, 165–167). Recent studies have identified a processed pseudogene insertion into the *CYBB* locus that is responsible for a sporadic case of chronic granulomatous disease (168). *De novo* processed pseudogene insertions also have been identified in cancer genomes (169). Finally, there are four examples where human diseases have been caused by the insertion of poly (A) tracts, which likely are derived from severely truncated LINE-1-mediated retrotransposition events (reviewed in reference 14).

The above studies have revealed that LINE-1-mediated retrotransposition events are responsible for approximately 100 cases of sporadic human disease and provoke the following question: how often does LINE-1-mediated retrotransposition lead to human disease? A union of the above data, in conjunction with recent efforts to catalog the spectrum of mutations of the *NF1* gene causing the autosomal dominant disease neurofibromatosis, suggests that LINE-1-mediated retrotransposition events are responsible for approximately 1 in 250 disease-producing mutations in man (170).

LINE-1-mediated retrotransposition events as mutagens in other mammals

Deleterious LINE-1-mediated retrotransposition events are implicated as disease-producing mutations in nonhuman mammals. For example, LINE-1 retrotransposition events into genes are linked to mutagenic phenotypes in at least seven different mouse strains (171–177). Two of these mutagenic insertions, one into the gene encoding the glycine receptor β -subunit (*i.e.*, $L1_{spa}$) and another into the *reeler* gene (*i.e.*, $L1_{Orl}$), represent full-length LINE-1 insertions (171–173). Subsequent studies revealed that both $L1_{spa}$ and $L1_{Orl}$ could retrotranspose in cultured human and mouse cells, and led to the discovery of active LINE-1s from the T_F subfamily (90). Likewise, a mutagenic B1 insertion into the mouse *Atcay* locus is responsible for the jittery mouse, further demonstrating that B1 element retrotransposition continues to impact the mouse genome (178, 179).

LINE-1 retrotransposition events have been implicated in various dog phenotypes. For example, a LINE-1 retrotransposition event into the *c-myc* gene has been identified in a canine transmissible venereal tumor (180, 181), although it remains uncertain if the insertion event is involved in tumorigenesis. Likewise, LINE-1 retrotransposition events into the *Factor IX*, *dystrophin*, and

DLX6 genes are implicated in a mild case of hemophilia B in German Wirehaired Pointers, Duchenne-like muscular dystrophy in the Pembroke Welsh Corgi, and cleft palate and mandibular abnormalities in the Nova Scotia Duck Tolling Retriever (182–184).

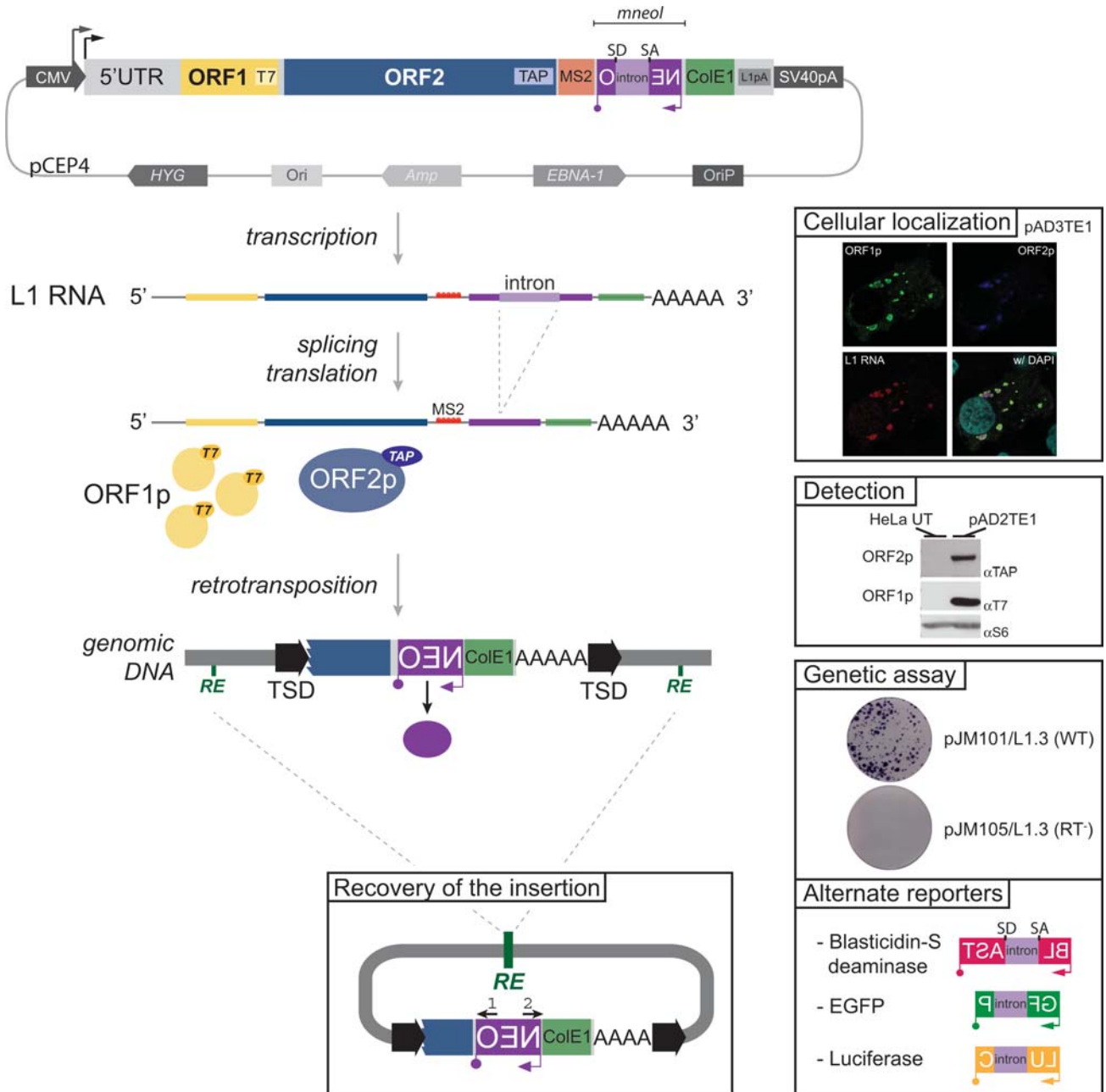
The retrotransposition of a canine tRNA-derived SINE (*i.e.*, SINEC_Cf) is responsible for various phenotypes in dog breeds. These include narcolepsy in Doberman Pinschers, centronuclear myopathy in Labrador Retrievers, “merle” coat color pigmentation in Shetland Sheepdogs, and progressive retinal atrophy in Tibetan Spaniels and Tibetan Terriers (185–188). Finally, the expression of an *FGF4* pseudogene is responsible for the “short-legged” phenotypes of at least 19 dog breeds (189). It will be interesting to determine the extent to which dog breeders have inadvertently selected for mutagenic LINE-1-mediated retrotransposition events that lead to “desirable” phenotypes within dog breeds. Indeed, selective breeding leads to genetic bottlenecks, which may allow the ready identification of retrotransposon insertions that dramatically affect phenotypic traits in mammals and other organisms.

MECHANISTIC STUDIES OF LINE-1 RETROTRANSPOSITION

An assay to study LINE-1 retrotransposition

Almost 20 years ago, a functional assay was developed to assess the retrotransposition potential of LINE-1s in cultured mammalian cells (60). The assay builds upon a rationale developed by Boeke and colleagues to demonstrate that the yeast Ty1 retrotransposon mobilizes via an RNA intermediate (6). Subsequent enhancements of the assay by the Heidmann and Curcio laboratories then led to the development of retrotransposition indicator cassettes that could only become activated for expression upon a successful round of retrotransposition (190, 191).

Briefly, the 3' UTR sequences of candidate full-length LINE-1s are tagged with a retrotransposition indicator cassette that consists of a backward copy of a neomycin phosphotransferase reporter gene equipped with its own promoter and polyadenylation signals (*i.e.*, the *mneoI* cassette; Figure 2). Importantly, the reporter gene is disrupted by an intron that resides in the same transcriptional orientation as the LINE-1 (60, 192, 193). This arrangement ensures that the expression of the neomycin phosphotransferase gene only occurs upon a successful round of LINE-1 retrotransposition, which ultimately leads to the generation of clonal foci that grow in the presence of the neomycin analog G418. Hence, the assay



allows a simple, yet powerful way to monitor LINE-1 retrotransposition efficiency by counting G418-resistant foci (60).

Since the inception of the cultured cell retrotransposition assay, a battery of retrotransposition indicator cassettes has been developed to assess LINE-1 retrotransposition by either exploiting drug selection (*i.e.*, a blasticidin resistance cassette) or screening for reporter gene activation (*i.e.*, green fluorescent protein and luciferase cassettes) (194–197) (Figure 2). Moreover, engineered LINE-1 elements, that contain epitope tags on the C-termini of ORF1p and ORF2p and an MS2 binding site in the LINE-1 mRNA, have allowed the direct detection of the LINE-1-encoded proteins and mRNA in cultured cells using both biochemical approaches and

fluorescence microscopy (82, 198–201) (Figure 2). Finally, retrotransposition indicator cassettes have been developed that allow the direct recovery of engineered LINE-1 retrotransposition events as autonomously replicating plasmids in *Escherichia coli* (156, 202, 203) (Figure 2).

In summary, the cultured cell retrotransposition assay, in conjunction with complementary molecular genetic and biochemical studies, has: (i) allowed the identification of active LINE elements from mammalian and vertebrate genomes (58–61, 86, 90, 91, 204, 205); (ii) shown that allelic heterogeneity affects LINE-1 retrotransposition (162, 206); (iii) facilitated experimental illumination of the LINE-1 retrotransposition mechanism (reviewed in reference 11); (iv) demonstrated that

FIGURE 2 Engineered Long Interspersed Element-1 (LINE-1) structure and cell based strategies to study retrotransposition. The LINE-1 expression vector consists of a retrotransposition-competent LINE-1 subcloned into pCEP4 (flanked by a CMV promoter and an SV40 polyadenylation signal). The pCEP4 vector is an episomal plasmid that has protein encoding (EBNA-1) and *cis*-acting (OriP) sequences necessary for replication in mammalian cells; it also has a hygromycin resistance gene (HYG) that allows for the selection of mammalian cells containing the vector, as well as a bacterial origin of replication (Ori) and ampicillin selection marker (Amp) for plasmid amplification in bacteria. The *mneol* reporter cassette, located in the LINE-1 3' UTR, contains the neomycin phosphotransferase gene (NEO, purple box, with its own promoter and polyadenylation signals, purple arrow and lollipop, respectively) in the opposite transcriptional orientation of LINE-1 transcription. The reporter gene is interrupted by an intron (light purple box) with splice donor (SD) and splice acceptor (SA) sites in the same transcriptional orientation as the LINE-1. This arrangement of the reporter cassette ensures that the reporter gene will only be expressed after a successful round of retrotransposition. *De novo* retrotransposition of the *mneol* reporter cassette will result in G418-resistant colonies that can be quantified—genetic assay panel with pJM101/L1.3 (wild-type [WT]) and pJM105/L1.3 (RT mutant [RT-]) LINE-1 constructs. Alternative reporters can be used instead of *mneol* to allow different drug-resistance, fluorescent, or luminescent read-outs (alternative reporters panel, with blasticidin-S deaminase [BLAST], enhanced green fluorescent protein [EGFP] or luciferase [LUC]) retrotransposition indicator cassettes. The addition of the ColE1 bacterial origin of replication (recovery of the insertion panel, green box) to a modified version of the *mneol* reporter cassette allows the recovery from cultured cell genomic DNA of engineered LINE-1 retrotransposition events as autonomously replicating plasmids in *Escherichia coli*. The insertions also can be characterized by inverse polymerase chain reaction using divergent oligonucleotide primers (recovery of the insertion panel, black arrows: 1 and 2) that anneal to the reporter gene. The use of epitope tags (T7-tag in C-terminus of ORF1, yellow box, and TAP-tag in C-terminus of ORF2, blue box) allow the immunoprecipitation (not shown) and detection of LINE-1 proteins by western blot and immunofluorescence (IF) (detection panel, with western blot data obtained with pAD2TE1, a vector expressing ORF1-T7p and ORF2-TAPP, compared to untransfected [UT] HeLa cells [82]). The addition of the RNA-stem loops that bind the bacteriophage MS2 coat protein (409) (orange box) in the 3' UTR of LINE-1 can be used to detect the cellular localization of LINE-1 RNA by fluorescent *in situ* hybridization (FISH). Both IF and FISH strategies can be combined to detect the subcellular localization of ORF1p, ORF2p, and LINE-1 RNA (cellular localization panel, with pAD3TE1 vector containing ORF1-T7p, ORF2-TAPP, and LINE-1 RNA-MS2 [82]). The images shown in the cellular localization and the detection box originally were published in (82). Additional references are provided in the text. [doi:10.1128/microbiolspec.MDNA3-0061-2014.f2](https://doi.org/10.1128/microbiolspec.MDNA3-0061-2014.f2)

LINE-1 retrotransposition generates genomic structural variation (156, 202, 203); (v) revealed that the LINE-1-encoded proteins (ORF1p and/or ORF2p) could act *in trans* to mediate SINE retrotransposition and processed pseudogene formation (92–98); and (vi) allowed the identification of host factors that may restrict and/or promote retrotransposition (see below). Clearly, the cultured cell assay has been and continues to be instrumental in allowing a deeper mechanistic understanding of LINE-1 biology.

Functional studies of LINE-1 retrotransposition

The LINE-1 5' UTR

Although the promoter structures of human and mouse LINE-1s differ (Figure 1), it is clear that the acquisition of an internal RNA polymerase II promoter has ensured that full-length retrotransposed LINE-1s retain the potential to undergo subsequent amplification in the genome. The human LINE-1 5' UTR is approximately 910 bp in length and contains an internal RNA polymerase II promoter that directs transcription of LINE-1 mRNA at or near the first nucleotide of the element (75, 207). Experimental studies have revealed that a YY1-binding site at the 5' end of the 5' UTR is critical for accurate transcriptional initiation and that most LINE-1 mRNAs contain a 7-methyl guanosine cap structure, which facilitates their translation (208, 209). The 5' UTR harbors *cis*-acting binding sites for the following transcription factors: Runx3, Sp1, and SRY-related (Sox) proteins (75, 207, 210, 211). Studies in cultured cells have revealed that mutations in these *cis*-acting sequences reduce LINE-1 transcription and retrotransposition. In addition, it is likely that other host factors bind the 5' UTR and regulate LINE-1 expression (see below).

In addition to containing a sense strand promoter, the human LINE-1 5' UTR contains a conserved RNA polymerase II antisense promoter (76). Transcription from the LINE-1 antisense promoter can lead to the generation of chimeric transcripts comprising LINE-1 sequences conjoined to sequences derived from the 5' genomic flank of a given LINE-1 locus (76, 212). These chimeric transcripts have been used as a proxy to identify transcriptionally active LINE-1 elements in human embryonic stem cells (213). Although mammalian bidirectional promoters have been identified to be the source of some noncoding RNAs (214), the function of the human LINE-1 antisense chimeric transcripts, if any, requires elucidation.

The mouse LINE-1 5' UTR consists of a series of up to seven and two-thirds monomeric repeats that are followed by an untranslated linker sequence immediately

upstream of ORF1 (87, 90). Reporter gene assays have revealed that 5' UTRs from the A, T_F, and G_F LINE-1 subfamilies remain transcriptionally active (90, 215, 216), whereas 5' UTRs from the V and F LINE-1 subfamilies generally lack transcription activity (89). Cell culture-based and biochemical assays have revealed that mRNAs derived from the A, T_F, and G_F subfamilies are enriched in ribonucleoprotein particles (RNPs) and that select LINE-1 elements from these remain retrotransposition-competent (86, 90, 91). Hence, it appears that the ability of mouse LINE-1s to capture new promoter sequences has, in part, led to their evolutionary success.

Mouse LINE-1s contain a transcriptionally active antisense RNA polymerase promoter (217, 218). Unlike human LINE-1s, the antisense promoter is located within ORF1. Transcription from the mouse antisense promoter leads to the generation of chimeric transcripts containing LINE-1 sequences conjoined to genomic sequences that flank the 5' end of the LINE-1 locus. Moreover, the overexpression of LINE-1 antisense mRNA could lead to a reduction of retrotransposition in cultured cells (218). Hence, it is intriguing to speculate that LINE-1 antisense mRNA may play a role in regulating mouse LINE-1 retrotransposition *in vivo*.

Recent studies indicate that a large fraction of mammalian long noncoding RNAs contain retrotransposon-derived sequences and that some are transcribed from LTR-retrotransposon-derived promoters (219, 220), (reviewed in reference 221). Intriguingly, select long noncoding RNAs are involved in maintaining the pluripotency of embryonic stem cells by yet unidentified mechanisms (222). It will be interesting to determine if the LINE-1 antisense promoter contributes to the transcriptional regulatory network regulating stem cell identity (reviewed in reference 223).

ORF1p

ORF1p is an ~40 kDa protein (also known as p40 for human LINE-1s) that is translated from LINE-1 mRNA by a traditional cap-dependent mechanism (77, 209, 224). Early biochemical studies demonstrated that mouse and human ORF1p resides in cytoplasmic RNPs and binds single-strand RNA in a sequence-independent manner (78, 225–228). Biochemical and genetic analyses clearly demonstrate that ORF1p is required for LINE-1 RNP formation, and that LINE-1 RNP formation is a necessary step in the retrotransposition process (198).

Structural studies have revealed that the N-terminus of LINE-1 ORF1p contains a coiled-coil domain, which facilitates trimerization of ORF1p molecules (79, 229–

232). The central region of ORF1p contains a non-canonical RNA recognition motif that, with assistance of its C-terminal domain, is required for ORF1p RNA binding (79, 231, 233). Notably, missense mutations in highly conserved amino acid residues in both the RNA recognition motif and C-terminal domain either abolish or adversely affect LINE-1 retrotransposition in cultured cells (60, 231).

Human and mouse ORF1p contain nucleic acid chaperone activities that can facilitate the re-annealing of single-strand DNAs *in vitro* (80, 234–236). Notably, several studies have shown that, despite the lack of sequence homology, proteins encoded by non-mammalian LINES also contain nucleic acid chaperone activity (237, 238). For example, ORF1p from a zebrafish LINE (ZfL2-1) has nucleic acid chaperone activity (239). It is hypothesized that this nucleic acid chaperone activity facilitates the initial steps of LINE-1 integration *in vivo*. Somewhat unexpectedly, the deletion of ORF1 does not abolish ZfL2-1 retrotransposition activity in cultured human cells (240). These data, coupled with the fact that Alu retrotransposition only requires the protein encoded by LINE-1 ORF2 (92), raise questions regarding how ORF1p nucleic acid chaperone activity participates in LINE-1 retrotransposition. The development of cell-free systems to monitor LINE-1 retrotransposition would allow a more rigorous examination of the ORF1p functions required for retrotransposition.

ORF2p

ORF2p is an ~150 kDa protein (81, 82, 200, 241) that contains endonuclease (L1 EN) (83) and reverse transcriptase (L1 RT) (84) activities that are critical for retrotransposition (60, 83). The L1 EN domain resides near the N-terminus of the protein, and bears similarity to apurinic/apyrimidinic endonucleases (APEs) (242–244). *In vitro* and bioinformatics analyses (71, 83, 99, 196, 245) suggest that L1 EN makes a single-strand endonucleolytic nick at a loosely defined consensus sequence in genomic DNA (5'-TTTT/A-3'; where the slash indicates the scissile phosphate), exposing a 5' phosphate and 3' hydroxyl group (83). Crystallographic studies suggest that L1 EN recognizes an extra helical “flipped” adenine residue 3' of the scissile bond to mediate cleavage using a mechanism similar to that employed by other APE proteins (244). In addition, it is likely that epigenetic modifications of target DNA (*e.g.*, nucleosome accessibility) might affect ORF2p accessibility and L1 EN cleavage activity (246).

The L1 RT domain is located downstream of the EN domain in ORF2p, and shares sequence similarity to

the RT domains encoded by telomerase, Penelope-like retrotransposons, group II introns, other non-LTR retrotransposons, LTR retrotransposons, and retroviruses (247–249). Biochemical and genetic assays originally were used to demonstrate that Ty1/LINE-1 ORF2p fusion proteins possess reverse transcriptase activity *in vitro* (84, 250). The subsequent purification of recombinant ORF2p produced in a baculovirus expression system revealed that full-length ORF2p could efficiently generate reverse transcripts from poly rA/oligo dT₁₂ primer template complexes, that L1 RT activity exhibited a preference for Mg²⁺ over Mn²⁺, and that L1 RT exhibited both RNA-dependent and DNA-dependent polymerase activities (251). Additional studies revealed that, like the RT encoded by the R2Bm retrotransposon (252), L1 RT is highly processive (when compared to Moloney murine leukemia virus RT) and lacks detectable RNase H activity (253).

L1 RT activity has been detected in LINE-1 RNP preparations derived from cells transfected with engineered LINE-1 expression vectors (199). Importantly, this work confirmed that ORF2p preferentially reverse transcribes its own mRNA template (*i.e.*, it exhibits *cis*-preference for its encoding RNA) and that point mutations in ORF1 and the L1 EN domain, which adversely affect LINE-1 retrotransposition, retain reverse transcriptase activity (199). Finally, these and subsequent studies confirmed previous inferences (156) that L1 RT can extend terminally mismatched primer–template complexes (199, 254). The latter property distinguishes L1 RT from Moloney murine leukemia virus and other retroviral reverse transcriptase enzymes.

Although LINE-1 ORF1p and L1 RT activity were readily detectable in RNP preparations, the detection of the LINE-1 ORF2p had been notoriously difficult. Epitope-tagging strategies have allowed the detection of ORF2p in whole cell extract and RNP preparations derived from cells transfected with engineered LINE-1 expression vectors (82, 200, 241). Using a similar strategy, immunofluorescence microscopy studies revealed that engineered ORF2p co-localizes in cytoplasmic foci with both ORF1p and LINE-1 mRNA (82, 200, 241) (Figure 2). Despite progress in detecting ORF2p in cultured cells, debate continues regarding the stoichiometry of ORF1p and ORF2p bound to LINE-1 mRNA (82, 200, 241). It appears that ORF1p is much more abundant in LINE-1 RNPs than ORF2p. Additionally, the composition of a functional LINE-1 RNP and how the LINE-1 mRNA transitions to a retrotransposition intermediate require elucidation. Clearly, the development of reconstituted *in vitro* target-site primed reverse

transcription reactions would greatly advance the understanding of the detailed molecular mechanism of LINE-1 retrotransposition.

LINE-1 ORF2p contains an ill-defined cysteine-rich domain (C-domain) at its C-terminus, which has been suggested to function as a zinc-knuckle domain (85). Consistent with its biological importance, cysteine to serine mutations in the C-domain interfere with LINE-1 RNP formation and strongly inhibit LINE-1 retrotransposition in cultured cells (60, 82). Recent studies indicate that a recombinant protein containing the last 180 amino acids of ORF2p exhibits nonsequence-specific RNA binding *in vitro*, and that cysteine to serine mutations in the C-domain do not adversely affect RNA binding (255). Hence, future studies are required to elucidate the exact function of the C-domain in LINE-1 retrotransposition.

Additional functional domains are likely to exist within LINE-1 ORF2p. Indeed, PCNA, which is the sliding clamp protein essential for DNA replication, recently was found to directly interact with ORF2p through a conserved sequence known as a PCNA interaction protein domain (PIP box), which is located between the L1 EN and L1 RT domains (200). Mutating the PIP box abolished LINE-1 retrotransposition (200); however, how PCNA functions in LINE-1 retrotransposition requires further elucidation.

How ORF2p is translated from bicistronic LINE-1 mRNA remains an active area of study and recent reports suggest that human and mouse LINE-1 ORF2p may be translated by distinct mechanisms (256, 257). In human LINE-1s, a 63-nucleotide spacer that contains two in-frame stop codons separates ORF1 and ORF2. Genetic studies in cultured cells suggest that ORF2p translation occurs by an unconventional termination/re-initiation mechanism where a translating ribosome must be able to scan from the stop codon of ORF1 to the start codon of ORF2 (256). Remarkably, studies show that human ORF2 can be translated in an AUG-independent manner (256). By comparison, evidence from luciferase reporter assays suggests that the presence of an internal ribosome entry site, which is located near the 3' end of mouse ORF1, is used to facilitate translation of mouse ORF2 (257). In addition, cell culture assays have revealed that mouse ORF2 may be translated in an AUG-independent manner (256, 257). Notably, it is unlikely that the 3' end of human ORF1 has an internal ribosome entry site (256). Indeed, it has been demonstrated that the sequence of the ORFs encoded by human and mouse LINE-1s can be subjected to substantial sequence changes by codon optimization without affecting retrotransposition in cultured cells (258,

259), suggesting that strict *cis*-acting sequences are not required for ORF2 translation.

It is unlikely that LINE-1s have evolved a novel mechanism to mediate ORF2 translation. Instead, we hypothesize that LINE-1s have evolved to exploit translation mechanisms inherent to their hosts to mediate ORF2 translation (256). Interestingly, recent ribosomal profiling studies have uncovered an increasing number of unannotated reading frames that reside 5' of annotated ORFs (260); some short ORFs also may be translated via an AUG-independent mechanism (261). Clearly, additional studies are warranted to elucidate the ORF2 translation mechanism and to determine if ORF2 translation differs among mammalian LINE-1s.

The LINE-1 3' UTR

The human LINE-1 3' UTR is ~206 bp in length and contains a conserved polypurine tract that is predicted to form a G-quadruplex structure (262). Intriguingly, the polypurine tract is not required for LINE-1 retrotransposition in cultured cells (60); but the polypurine tract can inhibit LINE-1 RT activity in *in vitro* biochemical assays (253). Despite its evolutionary conservation, how the polypurine tract functions in LINE-1 biology remains unknown.

LINE-1 3' UTRs contain a functional RNA polymerase II polyadenylation signal near their 3' ends. Experiments in cultured cells have revealed that the LINE-1 poly (A) signal is relatively weak, is often bypassed by RNA polymerase II, and that RNA polymerase II frequently uses canonical polyadenylation sites fortuitously present in 3' flanking genomic DNA sequences (263). The use of these genomic polyadenylation sequences can lead to the generation of chimeric LINE-1 transcripts containing genomic DNA sequences at their 3' end (see below). Finally, recent data suggest that the human and mouse 3' UTRs have promoter activity that leads to the generation of alternative LINE-1 transcripts in various tissues (264). The field awaits a better definition of this promoter activity and the role of the resultant transcripts in LINE-1 biology.

An overview of the LINE-1 replication pathway

LINE-1 retrotransposition occurs via a “copy and paste” process termed target-site primed reverse transcription (TPRT; Figure 3), a mechanism originally described by the Eickbush laboratory for the related site-specific non-LTR retrotransposon, R2Bm, from the silkworm *Bombyx mori* genome (265). After transcription from a chromosomal locus, a full-length bicistronic LINE-1 mRNA is exported to the cytoplasm. Upon translation, ORF1p and

ORF2p exhibit a strong *cis*-preference (97, 98) and bind to their respective encoding mRNA, forming an RNP (78, 82, 198, 227, 228). The LINE-1 RNP minimally consists of LINE-1 mRNA, multiple ORF1p trimers, and as few as one molecule of ORF2p (82, 256), but also likely contains numerous cellular proteins and RNAs (200, 201, 266) (Figure 3).

Intriguingly, subsequent studies revealed that ORF1p, ORF2p, and LINE-1 RNA accumulate in dense cytoplasmic foci, which are closely associated with stress granule proteins (82, 197). In yeast, proteins encoded by Ty1 and Ty3 retrotransposons are associated with cytoplasmic foci called processing bodies (P-bodies) and experiments suggest that P-body localization is important for RNP assembly and may represent a host mechanism that regulates retrotransposition (267–269).

How the LINE-1 RNP enters the nucleus is not fully understood. Experiments using modified second-generation adenoviral expression vectors containing an active human LINE-1 have demonstrated that LINE-1 retrotransposition can occur in G1/S arrested cells (270). Similarly, LINE-like sequences from *Candida albicans* (271, 272) and *Neurospora crassa* (273), which undergo closed mitosis, can retrotranspose independently of nuclear envelope breakdown. So, it does not appear that cell division is a requisite for LINE-1 retrotransposition. Some reports suggest that cell division augments the retrotransposition of engineered LINE-1s in cultured cells (274, 275). Notably, the cultured cell retrotransposition assay generally requires the detection of retrotransposition events as a function of reporter gene expression. Hence, as cells divide they may produce more of the reporter gene product, leading to an apparent increase in LINE-1 retrotransposition potential, thereby explaining the apparent discrepancies among the above studies.

Once in the nucleus, L1 EN makes a single-strand endonucleolytic nick in genomic DNA at a degenerate consensus sequence (5'-TTTT/A: where the “/” indicates the scissile phosphate), exposing a 3' hydroxyl group that serves as a primer for the reverse transcription of the LINE-1 mRNA by the L1 RT activity encoded in ORF2p (83, 276) (Figure 3). Whether the LINE-1 mRNA simply acts as a template for retrotransposition or whether it plays additional roles during TPRT requires more study. It is notable that codon-optimized synthetic mouse and human LINE-1s, in which ~25% of the nucleotide sequence has been replaced to increase the G-C content of LINE-1 RNA while retaining the amino acid sequence of the LINE-1 encoded proteins, can readily retrotranspose in cultured human cells (258, 259).

Studies in cultured human cells have revealed that the LINE-1 RT has a misincorporation error rate of ~1 in 6,500 bases (156). By using the binomial distribution, it has been estimated that ~40% of full-length LINE-1 retrotransposition events represent faithful copies of the progenitor LINE-1 element (~37% contain one mutation, and ~16% contain two mutations) (156). Subsequent steps in the retrotransposition process, including second-strand target-site DNA cleavage and second-strand LINE-1 cDNA synthesis, require additional investigation. By analogy to the evolutionarily related R2 retrotransposon of *Bombyx mori*, ORF2p may play a role in each of the above processes (277). The net result of TPRT is the integration of a new LINE-1 copy at a new chromosomal location (Figure 3).

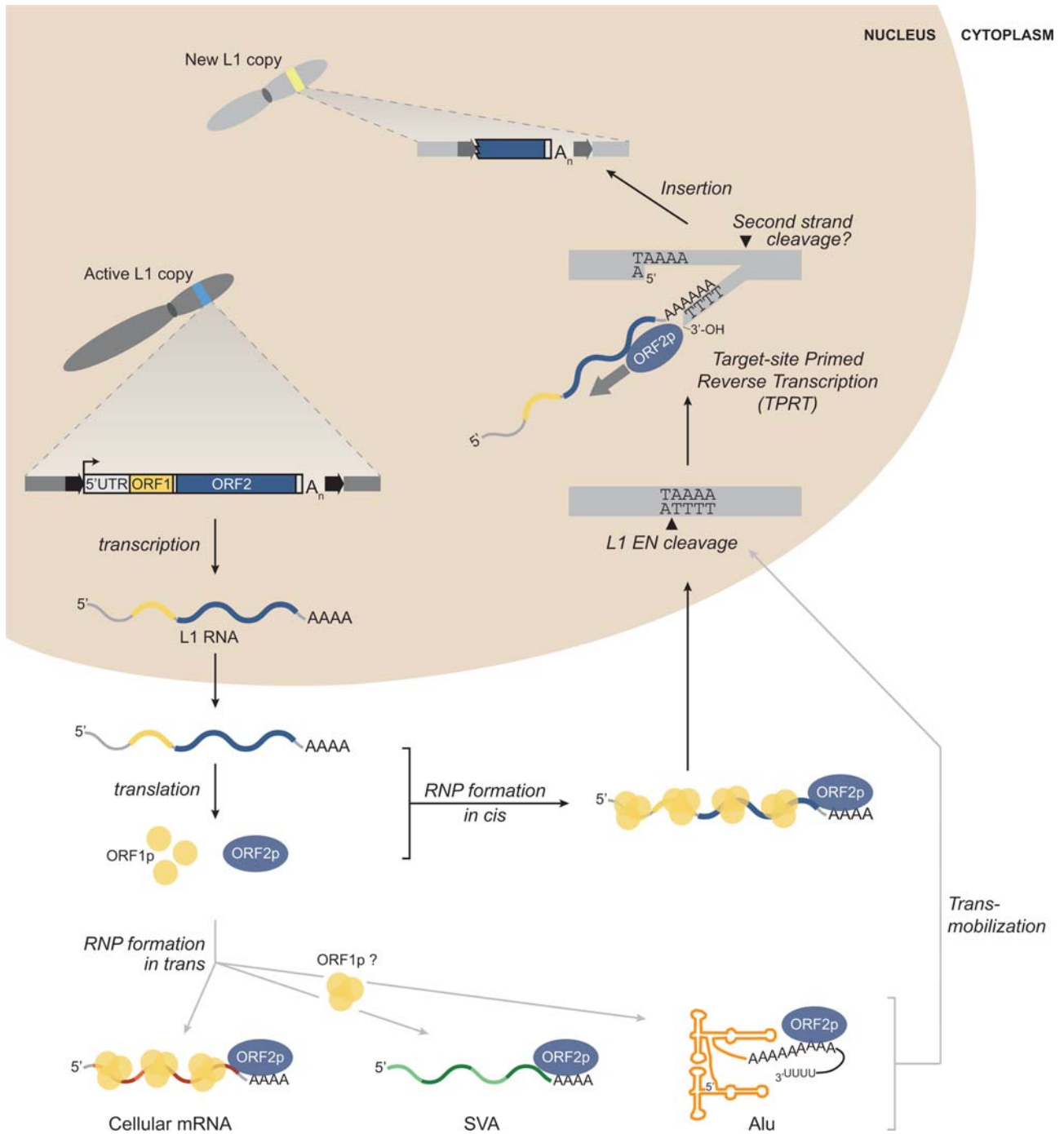
Genomic rearrangements generated during LINE-1 retrotransposition

LINE-1-mediated retrotransposition events are sometimes accompanied by intra-LINE-1 rearrangements (*e.g.*, 5' truncations and 5' truncations associated with inversion/deletion events) or genomic structural rearrangements (Figure 4). The features of these events suggest that host processes, such as DNA repair and/or DNA replication, may ultimately impact the structure of newly retrotransposed LINE-1s. Below we discuss some of these rearrangements.

Intra-LINE-1 alterations

The examination of the HGR reveals that ~30% to 35% of human-specific Ta-subset LINE-1 insertions are full-length, ~40% to 45% are truncated at their 5' ends, and ~25% contain internal rearrangements known as inversion/deletions (7, 63, 71, 278). The characterization of engineered LINE-1 retrotransposition events from cultured cells has led to the proposition that two pathways of LINE-1 retrotransposition exist: conventional and abortive retrotransposition (156). Conventional retrotransposition accounts for the generation of full-length LINE-1 insertions and can lead to the formation of new “master genes” that can serve as a source of retrotransposition events in subsequent generations (Figure 4A). In general, full-length LINE-1 insertions are characterized by typical LINE-1 structural hallmarks (*i.e.*, they terminate with a poly (A) tract; are flanked by variable size target-site duplications; and integrate at a LINE-1 endonuclease consensus site) (reviewed in reference 11).

The generation of 5' truncated LINE-1 elements is proposed to occur via abortive retrotransposition. Here, the L1 RT becomes dissociated from the (–) strand LINE-1



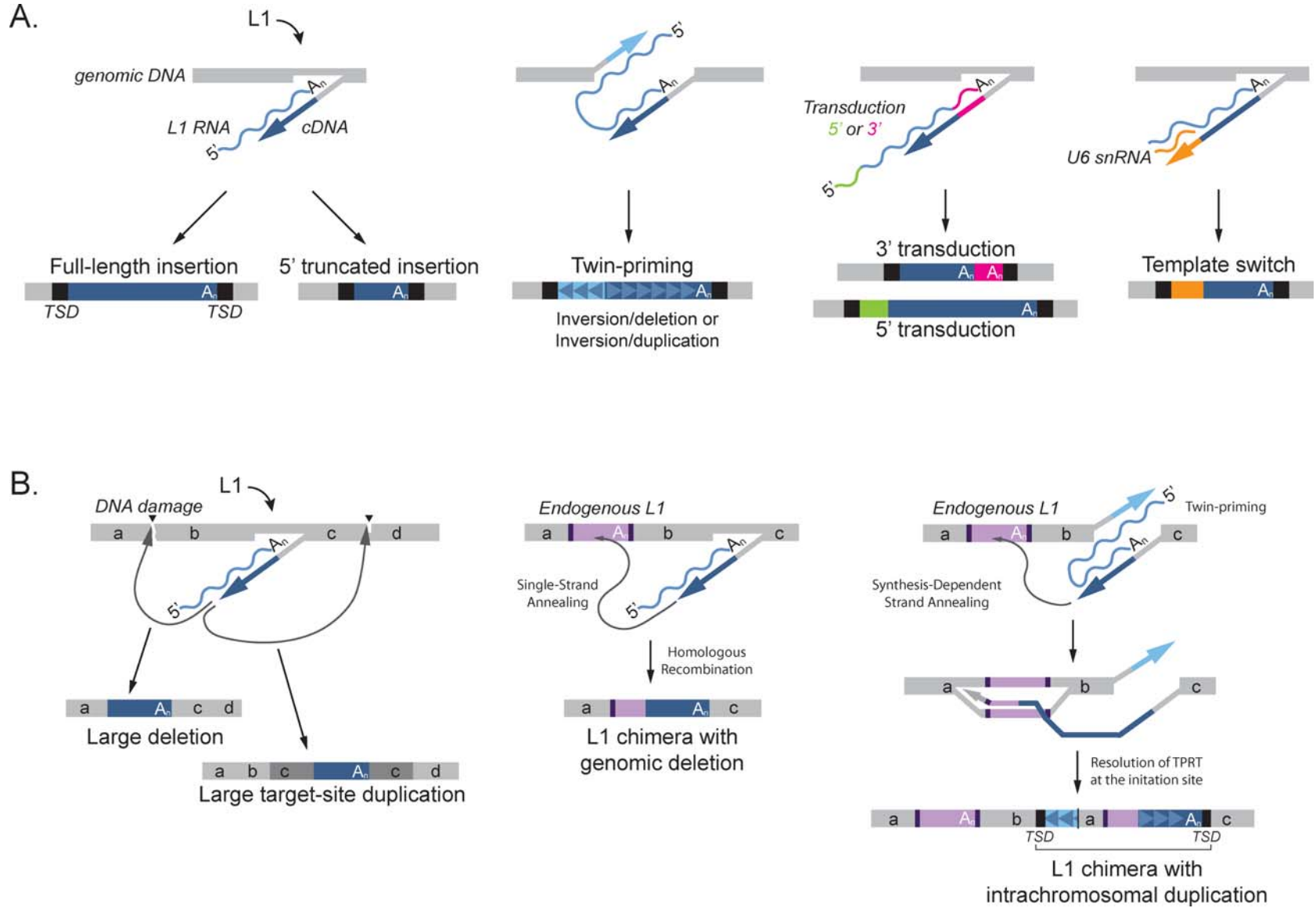
cDNA during TPRT. Annealing of the LINE-1 cDNA to top-strand genomic DNA then may specify the placement of top-strand (also referred to as second-strand) genomic DNA cleavage, generating a 3' hydroxyl group needed for DNA-dependent (+) strand LINE-1 cDNA synthesis (156, 279). How the L1 RT may become dissociated from the LINE-1 cDNA requires clarification; however, it is intriguing to speculate that the process of LINE-1 integration represents a battleground between LINE-1 and the host, and that the Y-branch intermediate generated during (–) strand LINE-1 cDNA synthesis may elicit a DNA repair response(s) by the host (156). Recent studies suggest that the ataxia telangiectasia mutated (ATM) and excision repair cross-complementation group 1 (ERCC1) proteins modulate LINE-1 retrotransposition (280–282); hence, it is reasonable to speculate that DNA repair pathways might influence the generation of 5' truncated LINE-1s.

The formation of inversion/deletion structures represents an alternative form of conventional retrotransposition termed “twin-priming” (278). Here, the LINE-1 mRNA anneals to single-strand DNA exposed at both the cleaved bottom- and top-strand genomic DNA sequences (Figure 4A). Template switching of the L1 RT during RNA-dependent (–) strand LINE-1 cDNA synthesis, or perhaps a second molecule of ORF2p, then allows the use of the 3' hydroxyl group generated at the top-strand

LINE-1 mRNA/genomic DNA duplex to serve as a primer for convergent RNA-dependent (–) strand LINE-1 cDNA synthesis. Microhomology-mediated annealing of the resultant cDNAs followed by the completion of cDNA synthesis (by either a LINE-1 RT DNA-dependent DNA polymerase activity or a host-encoded DNA polymerase) then can lead to the formation of inversion/deletion structures (Figure 4A). Notably, virtually all of the predictions of the twin-priming model have been confirmed by examining engineered LINE-1 integration events from cultured cells (156). How microhomology-mediated annealing occurs needs further study, although one can hypothesize that it is carried out by an alternative, microcomplementarity-mediated non-homologous end joining pathway of DNA repair (283).

The incorporation of untemplated nucleotides, presumably added after the completion of (–) strand LINE-1 cDNA synthesis by the LINE-1 RT, can result in short stretches of nontemplated sequence at the 5' genomic DNA/LINE-1 junction (156, 203, 208, 284), which may facilitate annealing of the LINE-1 cDNA to single-strand DNA exposed at the top-strand genomic DNA target-site (279, 285, 286). If so, we reason that the resultant LINE-1 cDNA/genomic DNA hybrid then may specify the placement of top-strand genomic DNA cleavage, generating the 3' hydroxyl group needed for DNA-dependent (+) strand cDNA synthesis by either the

FIGURE 3 Long Interspersed Element-1 (LINE-1) retrotransposition cycle. An active copy of LINE-1 is present at one chromosomal locus (light blue box in dark gray chromosome) and consists of a 5' untranslated region (UTR) (light gray box) with an internal promoter (thin black arrow), two open reading frames (ORF1, yellow box, and ORF2, blue box), a 3' UTR (light gray box) followed by a poly (A) tract (A_n) and is flanked by target-site duplications (thick black arrows). Transcription of LINE-1 occurs in the nucleus and produces a bicistronic RNA (wavy line). Upon translation in the cytoplasm, ORF1p and ORF2p (yellow circles and blue oval, respectively) bind back to their encoding RNA (*cis*-preference) to form a ribonucleoprotein particle (RNP) complex. ORF1p and/or ORF2p also can retrotranspose cellular RNAs (mRNA, SVA, and Alu, in red, green, and orange wavy lines, respectively). The retrotransposition of Alu RNA only requires ORF2p (92). There is some debate as to whether ORF1p augments Alu retrotransposition (410), and if SVA retrotransposition requires both ORF1p and ORF2p (94, 95). The LINE-1 RNP enters the nucleus where *de novo* insertion occurs by a mechanism termed target-site primed reverse transcription (TPRT). The ORF2p endonuclease activity makes a single-strand endonucleolytic nick at the genomic DNA target (L1 EN cleavage), at a loosely defined consensus site (5'-TTTT/A-3', with “/” indicating the scissile phosphate). The ORF2p RT activity then uses the exposed 3'-OH group to initiate first-strand LINE-1 cDNA synthesis using the bound RNA as a template. The final steps of TPRT (*i.e.*, top-strand cleavage, second-strand LINE-1 cDNA synthesis, and repair of the DNA ends) lead to the insertion of a *de novo* LINE-1 copy at a new chromosomal locus (light yellow box in light gray chromosome). The new LINE-1 copy is often 5' truncated, contains a variable-sized poly (A) tract (A_n), and generally is flanked by target-site duplications (thick gray arrows). Additional references are provided in the text. [doi:10.1128/microbiolspec.MDNA3-0061-2014.f3](https://doi.org/10.1128/microbiolspec.MDNA3-0061-2014.f3)



LINE-1 reverse transcriptase or a host-encoded DNA polymerase.

LINE-1-mediated transduction events

Active LINE-1s mobilize sequences that are derived from their 5' and 3' flanking genomic DNA by a process termed LINE-1-mediated transduction (Figure 4A). LINE-1s containing 5' transduction events occur when a cellular promoter, which resides upstream of an active genomic full-length LINE-1 copy, is used to initiate LINE-1 transcription. Retrotransposition of the chimeric 5' genomic/LINE-1 mRNA transcript then leads to the transduction of the 5' derived genomic DNA sequence to a new chromosomal location. If a conventional RNA polymerase II promoter in genomic DNA is used to initiate LINE-1 transcription, the resultant 5'

transduced LINE-1 will lack the genomic promoter and generally will be transcribed using the internal promoter present in the LINE-1 5' UTR in successive rounds of retrotransposition. Full-length LINE-1s containing 5' transduced genomic DNA sequences originally were detected in the HGR (7). A 5' transduction event is relatively rare and can only be identified by examining the sequences of full-length LINE-1s. Notably, the Nathans laboratory demonstrated that a full-length mouse LINE-1 insertion carrying a 28 bp 5' transduction led to the mis-splicing of the *Nr2e3* gene in a retinal degeneration 7-mouse model (287).

Due to the presence of inherently weak polyadenylation signals in their 3' UTRs, LINE-1s also can mobilize sequences that are derived from their 3' flanks, including exons, that range in size from tens of base pairs to at

FIGURE 4 Alterations generated upon Long Interspersed Element-1 (LINE-1) retrotransposition. (A) LINE-1 retrotransposition events can alter target-site genomic DNA. *De novo* insertion of LINE-1 occurs at a genomic DNA target (thick gray line). LINE-1 RNA (blue wavy line) is followed by a poly (A) tail (A_n); LINE-1 cDNA (blue arrow); and a new LINE-1 copy (blue box including a poly (A) tract (A_n)). Insertions can occur by either conventional (full-length, left) or abortive (5' truncated, right) retrotransposition and generally result in the formation of variable-length target-site duplications (TSD, black boxes). "Twin-priming" generates LINE-1 inversion/deletions or inversion/duplications (represented by opposing arrows in the new LINE-1 new copy). The priming of LINE-1 cDNA synthesis from the cleaved top-strand genomic DNA is represented by the light blue arrow. The transduction of genomic DNA sequences can occur when either 5' or 3' flanking genomic sequences are incorporated into LINE-1 RNAs and are mobilized by retrotransposition. The 5' and 3' transductions are depicted in both LINE-1 RNA (green or pink wavy lines) and the new LINE-1 copy (green or pink boxes). The 3' transduction events contain two poly (A) sequences (A_n). The LINE-1 enzymatic machinery also can mobilize small nuclear RNAs (snRNAs) such as U6 snRNA to new genomic locations. The proposed model involves an L1 RT template switch from LINE-1 RNA to the U6 snRNA (orange wavy line) to generate U6 cDNA (orange arrow) during target-site primed reverse transcription (TPRT). (B) LINE-1 retrotransposition events associated with genomic structural variation. LINE-1 RNA, cDNA, and a *de novo* LINE-1 insertion are depicted as in panel A. Lower case letters (a, b, c, or d) in genomic DNA (gray boxes) are used to depict deletions or duplications (by alteration of the alphabetical order). The resolution of TPRT at the site of DNA damage (left panel, black arrowhead upstream of the integration site) is hypothesized to result in a large genomic deletion (the loss of segment "b"), whereas the resolution of TPRT at a single-strand endonucleolytic nick downstream from the LINE-1 integration site (left panel, black arrowhead) is hypothesized to lead to a large target-site duplication (the duplication of segment "c"). The resolution of TPRT by single-strand annealing (middle) can lead to the generation of a chimeric LINE-1, where an endogenous LINE-1 (light purple box) is fused to a new LINE-1 (dark blue box); the formation of the chimera results in the loss of segment "b". Similarly, the resolution of "twin-priming" intermediates by synthesis-dependent strand annealing (right) can lead to the generation of an L1 chimera with an intrachromosomal duplication (the duplication of both segment "a" and the endogenous L1 sequence). The entire insertion is flanked by target-site duplications (black boxes). Notably, synthesis-dependent strand annealing can occasionally repair LINE-1 insertions generated in cultured cells by "twin-priming" (156). Details on how chimeric LINE-1 integration events are formed can be found elsewhere (156, 202, 203). Additional references are provided in the text. doi:10.1128/microbiolspec.MDNA3-0061-2014.f4

least 1.6 kb in length by 3' transduction (263, 288–290) (Figure 4A). This class of insertion is generated when RNA polymerase II bypasses the weak polyadenylation signal present at the 3' end of a full-length LINE-1 and instead uses a fortuitous polyadenylation signal in the 3' flanking genomic DNA. Retrotransposition of the resultant LINE-1/genomic hybrid mRNA leads to the insertion of the 3' flanking genomic DNA downstream of the new LINE-1 copy at a new chromosomal location. Due to the structure of most mammalian genes, which contain long introns with significant numbers of LINE-1 and SINE insertions, we speculate that LINE-1s have evolved to contain a weak polyadenylation signal to minimize premature polyadenylation of intron-containing genes (263) (reviewed in references 291–293).

The fact that LINE-1s can retrotranspose sequences derived from their 3' genomic flanks to new genomic locations was first appreciated while characterizing a mutagenic LINE-1 insertion into the dystrophin gene (290). The 3' transduction “genomic tag” in the mutagenic insertion then was used to isolate the likely progenitor LINE-1, named LRE2 (290). Experiments in cultured cells subsequently showed the LINE-1 3' transduction events occur frequently and could, in principle, mobilize exons and promoters to new genomic locations, providing a possible mechanism for exon shuffling (60, 263). Since that time, it has become apparent that ~20% to 25% of LINE-1 retrotransposition events are accompanied by 3' transduced sequences (288, 289). The presence of 3' transductions also have been used as “genomic tags” to identify progeny/offspring relationships among LINE-1 elements, to stratify full-length LINE-1 elements into subdivisions, and to identify clusters of LINE-1s that are actively mobilizing in the human population (58, 294).

The severe 5' truncation of the LINE-1/genomic mRNA upon TPRT can lead to the generation of “orphan transductions” that lack LINE-1 sequence (263). A mutagenic “orphan transduction” recently was identified as a cause of Duchene muscular dystrophy (295). Notably, and consistent with cultured cell studies, a recent study reported that ~24% of somatic LINE-1 retrotransposition events in tumors derived from 244 patients were accompanied by 3' transduction events, and that many of these events represented “orphan transductions” (296). Finally, it is noteworthy that transduction is not peculiar to LINE-1s. Both 5' and 3' transduction events have been observed with SVA retrotransposons (127, 297–299). Indeed, SVA retrotransposons provided a vehicle to shuffle the acyl-malonyl condensing enzyme-1 (AMAC) gene to three different locations in primate genomes (298).

LINE-1 target-site alterations: local alterations at the integration site

The mechanism of top-strand target-site cleavage at the genomic LINE-1 integration site requires elucidation. It is clear that the placement of second-strand DNA cleavage can influence the structure of the resultant LINE-1 integration events. After characterizing 100 engineered LINE-1 insertions in cultured cells, Gilbert and colleagues proposed a model that accounts for a number of observed target-site alterations (156, 202). In this model, top-strand DNA cleavage upstream from the bottom-strand endonucleolytic nick can lead to small deletions of genomic DNA at the LINE-1 integration site. Likewise, top-strand DNA cleavage directly opposite to bottom-strand cleavage can lead to LINE-1 integration events that lack target-site alterations, whereas top-strand cleavage downstream of the initial endonucleolytic nick can lead to the generation of target-site duplications (Figure 4A). Interestingly, ~10% of LINE-1 insertions in cultured cells were accompanied by large (greater than 50 bp) target-site duplications (TSDs); these large TSDs rarely are observed flanking LINE-1s in the human genome reference sequence (156, 202) (Figure 4B). Otherwise, the cultured cell retrotransposition assay largely recapitulates the spectrum of structural outcomes observed among endogenous germline retrotransposition events. However, the cellular milieu in which retrotransposition takes place, perhaps defined by the presence and activity of DNA repair machinery and other host factors, likely influences the range of possible LINE-1-mediated target-site alterations. Notably, the mechanisms described above also are likely to account for local target-site alterations accompanying SINE retrotransposition and processed pseudogene formation (reviewed in reference 11).

LINE-1-mediated retrotransposition target-site alterations: the generation of structural variants

In addition to minor target-site alterations, LINE-1 retrotransposition can lead to more substantial target-site genomic DNA modifications. The examination of LINE-1 integration events in cultured human cells has revealed that approximately 10% of retrotransposition events are accompanied by rearrangements of target-site DNA, creating genomic structural variation (156, 202, 203) (Figure 4B). Comparisons of pre-integration and post-integration sites in genomic DNA have revealed that the resolution of TPRT intermediates can lead to the generation of chimeric LINE-1 sequences (156, 202, 203). Single nucleotide polymorphism analyses revealed that the resultant integration events contain an

endogenous genomic LINE-1 fused to the engineered LINE-1 and concomitant genomic alterations (156, 202, 203) (Figure 4B).

The formation of chimeric LINE-1 elements can occur by various mechanisms. For example, the resolution of TPRT intermediates by single-strand annealing or synthesis-dependent strand annealing can lead to the formation of LINE-1 retrotransposition-mediated deletions or duplications, respectively (156, 202, 203) (Figure 4B). Importantly, large-scale genomic alterations observed in cultured cells are reflective of events occurring in humans. For example, a LINE-1 retrotransposition event was responsible for a 46 kb deletion in the *PDHX* gene of a human patient, resulting in pyruvate dehydrogenase deficiency (300).

Genomic alterations also can accompany the retrotransposition of both Alu and SVA elements. For example, an Alu retrotransposition event that occurred ~2.7 million years ago led to the deletion of an internal 92 bp exon within the *CMP-Neu5Ac* hydroxylase gene. As a result, humans are genetically deficient for *N*-glyconeuraminic acid (301). Similarly, an SVA retrotransposition event led to an ~14 kb deletion that resulted in the loss of the *HLA-A* gene in a cohort of Japanese families afflicted with leukemia (302). Finally, recent studies showed that two independent post-zygotic SVA retrotransposition events into the *NF1* gene were associated with large deletions of ~1 Mb and 867 kb, respectively (167).

On a larger scale, comparative genomics approaches between the human and chimpanzee reference sequences led to the identification of 50 LINE-1 retrotransposition events responsible for the deletion of ~18 kb from the human genome and ~15 kb from the chimpanzee genome (303). Similar approaches uncovered 33 Alu retrotransposition events that eliminated approximately 9 kb of human DNA (304). Hence, although relatively rare when compared with conventional retrotransposition events, LINE-1 retrotransposition-mediated deletion events continue to sculpt the landscape of the human genome.

Post-integration recombination events between genomic retrotransposons

The sheer mass of LINE-1 and Alu sequences in the genome also can provide substrates for post-integration recombination, generating structural variation in the human genome. For example, non-allelic homologous recombination, non-homologous DNA end joining, and other types of recombination events between genomic LINE-1 or Alu elements can lead to genomic alterations

that result in human disease (reviewed in references 11, 151, and 305–307). Clearly, these examples will continue to grow as individual whole genome DNA sequencing continues during the coming years.

Endonuclease-independent LINE-1 retrotransposition

LINE-1 retrotransposition by TPRT is usually initiated by the cleavage of genomic DNA by L1 EN. The examination of LINE-1 retrotransposition in Chinese hamster ovary (CHO) cells deficient in components of the non-homologous end joining pathway of DNA double-stranded break repair led to the discovery of an alternative integration pathway termed endonuclease-independent (ENi) LINE-1 retrotransposition (196). The ENi pathway of LINE-1 retrotransposition is reminiscent of a type of RNA-mediated DNA repair in which LINE-1 elements that lack L1 EN function presumably can use genomic lesions to initiate TPRT (196, 308). ENi retrotransposition events bear structural hallmarks distinct from canonical TPRT-mediated LINE-1 insertions in that they are frequently both 5' and 3' truncated, do not occur at typical LINE-1 endonuclease sites in genomic DNA, generally lack target-site duplications, and often are accompanied by the deletion of genomic DNA at the integration site (196). ENi LINE-1 retrotransposition events also occasionally are accompanied by the insertion of short cDNA fragments at both their 5' and 3' LINE-1/genomic DNA junctions, which appear to be derived from the reverse transcription of cellular mRNAs (196, 309).

Subsequent studies in DNA protein kinase catalytic subunit-deficient CHO cells revealed that ENi retrotransposition events could occur at dysfunctional telomeres, highlighting similarities between ENi retrotransposition and telomerase activity (309–312). Indeed, these results parallel the situation in *Drosophila* and bdelloid rotifer genomes, where “domesticated” retrotransposons function to maintain telomere length in place of a conventional telomerase activity (311, 313–315).

As with other phenomena discovered using engineered LINE-1 retrotransposons in the cultured cell retrotransposition assay, putative ENi retrotransposition events also have been identified in the human and mouse genomes (174, 316). Indeed, a likely ENi retrotransposition event into the *EYA1* gene was accompanied by an ~17 kb deletion, leading to a sporadic case of human oto-renal syndrome (317). It will be interesting to learn whether deficiencies in other DNA repair pathways lead to increased ENi LINE-1 retrotransposition.

Notably, some group II introns lack an EN domain and can use 3' hydroxyl groups at nascent DNA strands present at DNA replication forks to initiate retrotransposition (318). Moreover, it has been proposed that the L1 EN domain was acquired after the L1 RT domain during LINE-1 evolution (248). Together, these data indicate that the ENi pathway of LINE-1 retrotransposition may represent an ancient mechanism of LINE-1 retrotransposition before the acquisition of an APE-like endonuclease domain.

THE IDENTIFICATION OF HOST FACTORS THAT REGULATE LINE-1 RETROTRANSPOSITION

In the face of the deleterious consequences of LINE-1-mediated retrotransposition, the host cell appears to have evolved a variety of mechanisms to restrict LINE-1 activity (reviewed in reference 13). Several of these processes are briefly described below.

LINE-1 DNA methylation DNA methyltransferases

The methylation of CpG sequences in the LINE-1 5' UTR is associated with the suppression of LINE-1 expression in a variety of cell types (319). DNA methylation is established in primordial germ cells, is maintained throughout the life of an organism, and is thought to control the expression of LINE-1 in somatic tissues (320). *De novo* DNA methylation is catalyzed by the DNA methyltransferases Dnmt3a and Dnmt3b, but targeting of methylation to retrotransposon sequences requires the non-catalytic paralog DNA methyltransferase 3-like (Dnmt3L) (reviewed in reference 321). Notably, Dnmt3L-deficient male mice exhibit meiotic catastrophe, concomitant with an aberrant overexpression of LINE-1 and other transposon sequences (322). Similar phenotypes also are observed in Maelstrom-deficient male mice (323). However, the mechanistic link between the loss of transposon methylation, transcriptional reactivation, and meiotic catastrophe in germ cells requires further investigation. Finally, recent reports suggest that 5-hydroxymethylation of cytidine, which is a mark enriched on the LINE-1 5' UTR in pluripotent cells, might further regulate LINE-1 transcription (reviewed in reference 324).

piRNAs, PIWI proteins, and small RNAs derived from LINE-1

Other factors also are proposed to regulate *de novo* methylation of LINE-1 sequences in the developing germline (13). P-element induced wimpy testes (PIWI)

proteins, representing a subclade of the Argonaute family of small RNA binding proteins, interact with 26 to 31 nucleotide RNAs termed piwi-interacting RNAs (piRNAs) (reviewed in references 325 and 326). In the *Drosophila* germline, PIWI proteins and piRNAs comprise an effective transposon defense system, wherein transposon-derived piRNAs direct PIWI proteins to cleave active transposon transcripts (327, 328). The cleavage and subsequent processing of expressed transposon RNAs give rise to additional piRNAs that fuel additional rounds of this robust and adaptive transposon defense cycle, which is aptly named the “ping-pong” amplification cycle (329, 330) (reviewed in references 13, 326, 331 and 332).

Small piRNAs actively defend the mammalian germline from transposons (333) (reviewed in references 326 and 334). In male mice, a deficiency in either of two murine PIWI clade proteins, murine piwi (MIWI2) or miwi-like (MILI), results in a similar phenotype to Dnmt3L deficiency marked by meiotic catastrophe, aberrant retrotransposon transcriptional upregulation, and failure to establish methylation on retroelements in the genome (335, 336). A similar phenotype also is observed in MILI-interacting Tudor domain containing-1-deficient mice (337). MIWI2 and MILI are proposed to participate in a “ping-pong-like” amplification cycle in the male primordial germline to generate piRNAs against actively expressed retrotransposons. In addition, MIWI2, directed by piRNA, is proposed to participate in the targeting of *de novo* methylation to retrotransposon sequences (333, 335).

Several RNA-based restriction mechanisms also have been implicated in LINE-1 control. For example, the antisense promoter within the LINE-1 5' UTR can give rise to an antisense RNA transcript that may form a complex with the LINE-1 mRNA to create a substrate for siRNA biogenesis, resulting in siRNAs that target LINE-1 transcripts (338). In addition, the miRNA biogenesis factor Microprocessor/Drosha-DGCR8 has recently been demonstrated to specifically bind LINE-1, Alu, and SVA transcripts and cleave LINE-1 RNA *in vitro* (339). Accordingly, cultured cell experiments reveal that Microprocessor can restrict LINE-1 and Alu retrotransposition (339). In mouse embryonic stem cells, Dicer-dependent and Ago2-dependent RNAi mechanisms may possibly participate in LINE-1 regulation by limiting LINE-1 transcript accumulation (340, 341).

LINE-1 RNA transcription and splicing KRAB zinc-finger proteins

Recent studies have revealed that members of the Krüppel-associated box (KRAB) zinc-finger (KZNF) protein family can recruit KRAB-associated protein-1 (KAP1) and its

associated repressive complex to LINE-1 and SVA retrotransposons, which, in turn, inhibits their expression in embryonic stem cells (56, 57). The human KZNF protein, ZNF91, requires the SVA VNTR to bind and potently inhibit SVA expression (56). By comparison, ZNF93 binds to a sequence within the 5' UTR of older LINE-1 subfamilies (*i.e.*, L1PA3, L1PA4, and older elements) to inhibit their expression (56). Interestingly, the ZNF93 binding site has been deleted from the 5' UTR of active L1Hs elements, suggesting that L1Hs elements evaded the repressive effects of ZNF93 (56). Consistent with this idea, ZNF93 overexpression potently inhibited the retrotransposition of an engineered L1Hs element containing a reconstituted ZNF93 binding site in its 5' UTR (56). Indeed, these data suggest that select KZNF proteins may be locked in an evolutionary arms race with LINE-1 and SVA retrotransposons and that the ability of LINE-1 to evade ZNF93 binding may represent a mechanism to drive LINE-1 subfamily succession (56, 57).

LINE-1 RNA splicing and premature polyadenylation

TPRT often leads to the retrotransposition of a full-length human LINE-1 mRNA. Hence, it was somewhat surprising to find that full-length human LINE-1 RNAs contain a conserved splice donor sequence at position +98 of the 5' UTR (342). Recent studies have revealed that the splice donor site is functional and that its use can lead to the generation of shorter LINE-1 mRNAs, which may be compromised for retrotransposition (342, 343). Indeed, LINE-1 mRNA splicing may serve as a regulatory mechanism to restrict LINE-1 expression and/or retrotransposition in a tissue-specific manner.

A full-length LINE-1 also contains a number of potential polyadenylation signals, and their use can lead to the generation of shorter LINE-1 mRNAs that are compromised for retrotransposition (344). Indeed, it also is proposed that the adenosine-rich nature of the LINE-1 sense strand transcript, as well as the presence of the abovementioned polyadenylation signals, may act as “molecular rheostats” to fine-tune the expression of genes containing full-length LINE-1s (345).

Cellular RNA binding proteins

Recent evidence suggests that RNA binding proteins may play a direct role in modulating LINE-1 retrotransposition. For example, heterogeneous nuclear ribonucleoprotein L (hnRNPL), a protein that facilitates alternative splicing, was shown to associate with both mouse LINE-1 RNA (346) and human LINE-1 RNPs (201). The knock-down of hnRNPL led to an increase in the levels of mouse

LINE-1 RNA and ORF1p and an increase in the retrotransposition of engineered LINE-1 in cultured cells (346). In addition, the poly (A) binding protein C1 (PABPC1) interacts with the LINE-1 RNA and is required for LINE-1 RNP formation and efficient retrotransposition (347). Finally, cell culture assays have shown that human LINE-1 and mouse IAP mobility are restricted by the RNase activity of human RNase L, a member of the 2',5'-oligoadenylate (2-5A) synthetase (OAS)-RNase L system, initially described for restricting viral infections during the interferon antiviral response (348). Clearly, it will be interesting to determine how these RNA binding proteins affect LINE-1 retrotransposition at the mechanistic level.

Other proteins that restrict LINE-1 retrotransposition

APOBEC3 cytidine deaminases

The human *APOBEC3* gene family encodes seven proteins that can catalyze the deamination of cytidine to uridine residues in single-strand DNA substrates (reviewed in reference 349). Landmark findings by Malim and colleagues revealed that APOBEC3G potently restricts *vif*-deficient HIV infectivity (350). Subsequent studies demonstrated that several members of the APOBEC3 family, most potently APOBEC3A (A3A) and APOBEC3B (A3B), robustly inhibit LINE-1 and Alu retrotransposition in cultured cells (351–356). How A3B restricts LINE-1 retrotransposition remains unknown, although it may do so by both cytidine deamination-dependent and -independent pathways (351). Similarly, APOBEC3C (A3C) modestly inhibits LINE-1 retrotransposition in cultured cells by a cytidine deaminase-independent pathway (357). In contrast, a recent study demonstrated that A3A-mediated LINE-1 inhibition occurs by a deaminase-dependent pathway and that A3A can deaminate single-strand DNAs that are exposed transiently during LINE-1 TPRT (358). This mechanism is consistent with recent reports of APOBEC3-mediated deamination of single-strand genomic DNA in several types of cancer (359, 360).

Trex1/SAMHD1

Aicardi–Goutières syndrome (AGS) is a rare childhood inflammatory disorder that can lead to neurodevelopmental deficiencies (reviewed in references 361 and 362). AGS can be caused by mutations in several genes, including *Trex1*, *SAMHD1*, and *RNaseH2* (363–365). Intriguingly, the overexpression of *Trex1*, a 3' to 5' DNA exonuclease, can restrict both LINE-1 and IAP retrotransposition in cultured cells (366). Similarly, the overexpression of *SAMHD1*, a triphosphohydrolase that can reduce intracellular dNTP pools, can inhibit LINE-1

retrotransposition in cultured cells (367). Detailed studies are needed to determine the mechanism by which Trex1 and SAMHD1 inhibit LINE-1 retrotransposition. It also will be interesting to determine whether LINE-1 retrotransposition is elevated in tissues (e.g., brain regions [see below]) of AGS patients and if mutations in any of the three subunits of human RNaseH2 or other genes leading to AGS (i.e., *ADAR1* and *IFIH1*) regulate LINE-1 retrotransposition (368–370).

MOV10

Moloney leukemia virus 10 (MOV10) is an RNA helicase that inhibits the activity of several retroviruses including HIV-1 (reviewed in reference 371). MOV10 also restricts the retrotransposition of human LINE-1, Alu, SVA, and mouse IAP elements in cultured cells (372–374). The mechanism by which MOV10 inhibits LINE-1 retrotransposition requires elucidation. However, recent experiments have shown that MOV10 co-localizes with LINE-1 ORF1p in RNPs, directly binds LINE-1 RNA, and that mutations in the conserved helicase domains of MOV10 inhibit its ability to restrict LINE-1 retrotransposition (373–375). Notably, the MOV10-like-1 (MOV10L1) putative helicase also has been reported to play a role in piRNA-directed retrotransposon silencing in male germ cells (376, 377).

THE IMPACT OF LINE-1-MEDIATED RETROTRANSPPOSITION EVENTS ON THE GENOME

Developmental timing of LINE-1-mediated retrotransposition events

Despite the myriad of host defense mechanisms in place to restrict LINE-1 retrotransposition, *de novo* heritable LINE-1 insertions continue to occur and contribute to inter-individual human variation (reviewed in references 11 and 13). Although it remains unclear how frequently new, heritable LINE-1-mediated retrotransposition events arise in the human population, estimates suggest that ~1/100 humans harbors a *de novo* LINE-1 insertion and ~1/20 humans harbors a new Alu insertion (reviewed in reference 378). Moreover, the developmental timing of heritable retrotransposition events remains an area of active investigation. Retrotransposition in the germline lineage (i.e., gametes or precursor cells giving rise to gametes) presents an attractive and logical hypothesis for the generation of heritable insertions and is supported by evidence for full-length LINE-1 mRNA expression in the male germline and LINE-1 ORF1p expression in male and female germ cells (379–381). On the other hand, recent

studies employing transgenic animal models, human patient samples, and cell culture models have provided substantial evidence that pluripotent cells of the early embryo (i.e., embryonic stem cells) express LINE-1 mRNA and ORF1p, and provide a permissive milieu for accumulating potentially heritable *de novo* retrotransposition events (see below) (Figure 5).

Transgenic animal models

The first study to experimentally recapitulate LINE-1 retrotransposition *in vivo* used transgenic mice harboring an engineered human LINE-1 tagged with an *EGFP* retrotransposition indicator cassette (195). A pPolII promoter augmented the expression of the LINE-1, whereas a sperm-specific acrosin promoter drove the expression of the *EGFP* gene. The enhanced green fluorescent protein (EGFP) was equipped with a signal peptide, which would allow EGFP to localize to the acrosome of developing spermatids (382). The examination of 135 offspring derived from crosses of transgenic males to wild-type females allowed the identification of two *de novo* LINE-1 retrotransposition events. Interestingly, one of these animals contained a *de novo* LINE-1 retrotransposition event but lacked the transgene, suggesting that LINE-1 retrotransposition could occur in the male germline before the onset of meiosis II (382, 383).

Subsequent work, employing transgenic mouse and rat lines in which LINE-1 expression was driven by the native LINE-1 promoter, once again led to the identification of *de novo* LINE-1 insertions (384). As above, a number of offspring contained *de novo* LINE-1 retrotransposition events, but lacked the transgene. Unexpectedly, these animals were unable to pass the *de novo* retrotransposition event on to their progeny, suggesting that animals were mosaic—that is, they contained the *de novo* retrotransposition event in somatic cells, but not in their germline cells (384). How might such a scenario happen? The authors suggested that LINE-1 RNPs that are formed in either the male or female germline can be carried over to the zygote and subsequently can undergo retrotransposition in the early embryo. Clearly, these results require verification; however, this study provided tantalizing evidence that a substantial number of LINE-1 retrotransposition events may occur postzygotically.

It remains difficult to study LINE-1 retrotransposition in female oocytes. However, recent data suggest that LINE-1 expression and/or retrotransposition may lead to fetal oocyte attrition in mice (385). Hence, additional research in this area might provide mechanistic information about the timing of heritable LINE-1 retrotransposition events.

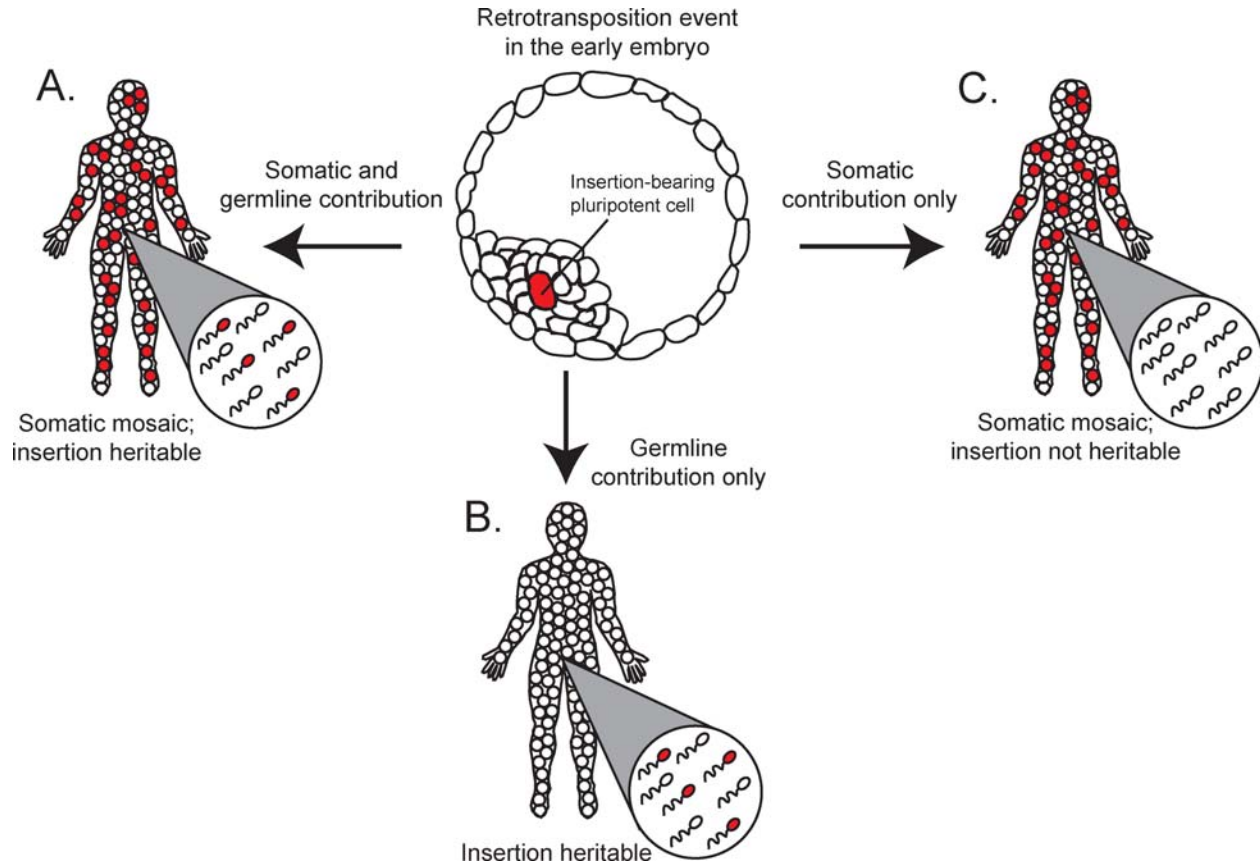


FIGURE 5 Hypothetical consequences of retrotransposition in pluripotent cells of the early embryo. (A) Cells harboring a *de novo* retrotransposition event could contribute both to the soma and germline, resulting in an individual with somatic as well as germline mosaicism and a heritable insertion. (B) Conceivably, cells harboring the retrotransposon insertion could contribute solely to the germline, giving rise to germline mosaicism, thereby rendering the insertion heritable. (C) Retrotransposon insertion-bearing cells could contribute to the somatic lineage but not to the germline, resulting in somatic mosaicism. Such an event would not be transmissible to the next generation. Red and white shaded circles in the human figures and sperm represent retrotransposon insertion-bearing and non-insertion-bearing cells in the soma and germline, respectively. (This figure was reproduced from Sandra Richardson's doctoral thesis [408]). Additional references are provided in the text. [doi:10.1128/microbiolspec.MDNA3-0061-2014.f5](https://doi.org/10.1128/microbiolspec.MDNA3-0061-2014.f5)

Human patients

Although approximately 100 cases of human genetic disease have been attributed to LINE-1-mediated retrotransposition events (reviewed in reference 14), it remains difficult to identify the developmental timing of a retrotransposition event. It was long assumed that LINE-1 actively retrotransposes in the male and female germline. Recent reports suggest that endogenous LINE-1s are expressed in human female oocytes and that engineered LINE-1 constructs readily retrotranspose in oocytes (381). In contrast, despite concerted efforts, it has been difficult to find evidence for endogenous LINE-1 retrotransposition in human sperm (386).

The identification of disease-producing LINE-1-mediated retrotransposition events has provided additional evidence for postzygotic, somatic insertions. In one instance, a LINE-1 retrotransposition event into the *CHM* gene led to X-linked choroideremia in a male patient (387). The examination of DNA from his family members revealed that his two sisters shared the same maternally inherited haplotype in the *CHM* region, but only one sister contained the retrotransposition event (387). Surprisingly, the mother of the patient was a somatic mosaic with respect to the LINE-1 retrotransposition event. These data, combined with the failure to pass the LINE-1 retrotransposition event to one of her

daughters (see above), provided unequivocal evidence that LINE-1 retrotransposition can occur in the early embryo to generate both somatic and germline mosaicism and that the resultant retrotransposition event can be passed on to subsequent generations (387) (Figure 5). Indeed, recent data suggest that SVA and processed pseudogene insertions also occur postzygotically (167, 168).

Experiments in developmentally relevant cultured cell lines

The notion that LINE-1-mediated retrotransposition events can occur early in development is supported by studies in cultured cells. For example, human embryonic stem cells express endogenous LINE-1 mRNA and ORF1p and can accommodate the retrotransposition of engineered human and mouse LINE-1s (213, 388). Similar results also have been obtained from studies conducted with human embryonic-carcinoma-derived cell lines (389) and induced pluripotent stem cells (390). Together, the above studies suggest that cells of the early embryo may represent an important developmental stage for heritable retrotransposition events in mammals.

Somatic LINE-1 retrotransposition

LINE-1 retrotransposition in the brain

In order to ensure its continued existence, it was widely believed that LINE-1 retrotransposition must occur in cells (*e.g.*, germ cells) that have the potential to contribute to subsequent generations. This view was radically overturned by the unexpected observations that: (i) the differentiation of adult rat hippocampal neural stem cells into neuronal precursor cells and neurons leads to an increase in LINE-1 transcript abundance, and (ii) engineered LINE-1 retrotransposons could retrotranspose in cultured rat neuronal precursor cells and in the brain of transgenic mice (391) (Figure 6A). Subsequent studies revealed that engineered LINE-1s could retrotranspose in both fetal and human embryonic stem cell-derived neuronal progenitor cells (392), that sensitive quantitative polymerase chain reaction-based experiments could detect an increase in LINE-1 copy number in several human brain regions (*e.g.*, hippocampus) when compared with matched liver or heart samples (392), and that engineered human LINE-1 retrotransposons exhibit enhanced somatic retrotransposition in the mouse models lacking either the methyl-CpG-binding protein 2 (*MeCP2*) or *ATM* genes (282, 393) (Figure 6).

The use of next-generation sequencing is now providing additional insights into LINE-1 retrotransposition in the brain. The development of retrotransposon-capture sequencing (RC-seq), a technique in which custom oligo-

nucleotide probes are used to enrich sequencing libraries for fragments containing LINE-1/genome junctions, demonstrated that endogenous LINE-1 insertions contribute to somatic mosaicism in the human brain (394) (Figure 6). The examination of DNAs derived from the hippocampus and caudate nucleus of three advanced-age post-mortem brain samples revealed ~7,700 potential somatic LINE-1 insertions and thousands of potential Alu and SVA somatic insertions (394). Sanger sequencing verified a handful of these putative insertions and showed that several insertions exhibited structural LINE-1 hallmarks.

One of the limitations of RC-seq and other methods that analyze bulk tissue samples is the inability to accurately quantify the extent of somatic LINE-1 mosaicism in the brain. Single-cell genomic analysis, a technically challenging undertaking, has the potential to provide key insights regarding this facet of neuronal retrotransposition. Indeed, a recent study using whole cell amplified genomic DNAs from 300 single neurons from the cerebral cortex and caudate nucleus led to the identification of somatic retrotransposition events in the brain (395). However, the level of LINE-1 retrotransposition, which was estimated to be ~0.6 insertions per cell (395), was lower than the level detected in previous studies (392, 394). Methodological differences, as well as differences in the brain regions that were analyzed (*e.g.*, the hippocampus, caudate nucleus, or cortex) may account for the different number of retrotransposition events detected in the above studies.

Clearly, more studies are needed to determine if the rate of LINE-1 retrotransposition varies in different brain regions or cell types, why neuronal precursor cells apparently are permissive for LINE-1 retrotransposition, and whether LINE-1 retrotransposition in the brain has biological consequences. However, a union of the above results indicates that LINE-1 can retrotranspose in the human brain. It also is notable that the mobility of TEs may not be restricted to the mammalian brain, as recent reports suggest that TEs also are active in the *Drosophila* brain (396, 397).

LINE-1 retrotransposition in cancer

The discovery of a mutagenic LINE-1 retrotransposition event into the adenomatous polyposis coli (*APC*) gene led to the realization that LINE-1 retrotransposition may have a role in tumorigenesis or tumor progression (398). In recent years, the creation of polymerase chain reaction amplicon libraries enriched for L1Hs elements and their associated genomic flanking DNA sequences, in conjunction with next-generation DNA sequencing, has led to the identification of *de novo* LINE-1-mediated retrotransposition events in a variety of tumor types,

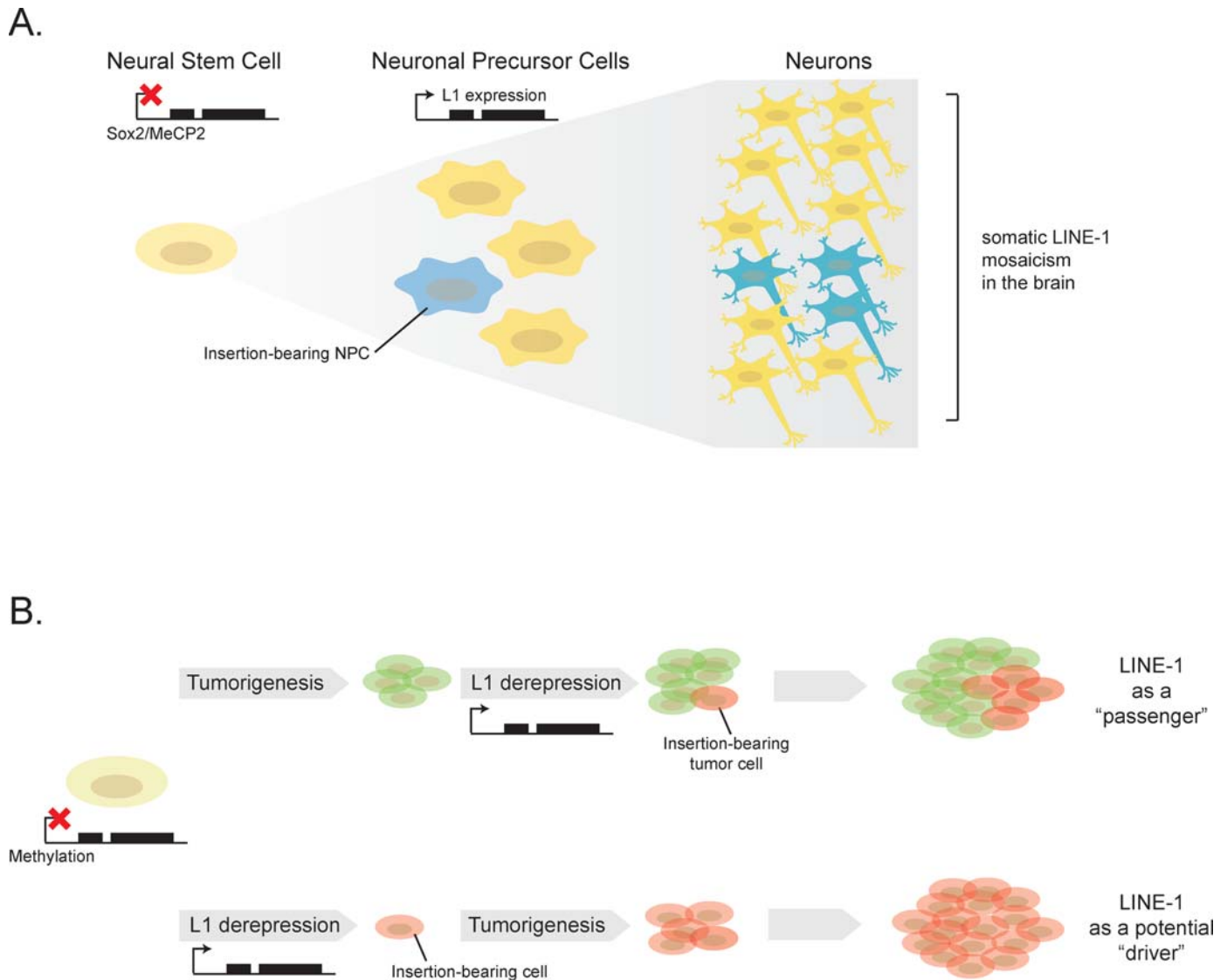


FIGURE 6 Long INterspersed Element-1 (LINE-1) retrotransposition in the brain and in cancer. (A) Model for how LINE-1 generates somatic mosaicism in the brain. Sox2, MeCP2, and promoter methylation (red X over the LINE-1 5' untranslated region [UTR]) are hypothesized to repress LINE-1 expression in neural stem cells (yellow cell). The differentiation of neural stem cells into neuronal precursor cells (NPCs) correlates with a reduction in LINE-1 promoter methylation and a derepression of LINE-1 expression, allowing a permissive milieu for retrotransposition (insertion-bearing NPC [blue cell]). Subsequent differentiation of NPCs into neurons leads to somatic LINE-1 mosaicism in the brain (insertion-bearing neurons [blue cells]). It is unknown whether LINE-1 retrotransposition occurs in postmitotic neurons. (B) Model for how LINE-1 may act as a "driver" or "passenger" mutation during cancer progression. In a somatic cell (yellow cell), LINE-1 expression generally is repressed by promoter methylation (red X over the LINE-1 5' UTR). After oncogenic transformation (top panel), the derepression of LINE-1 expression in some tumor cells (green cells), allows *de novo* LINE-1 retrotransposition events that act as "passenger" mutations (insertion-bearing tumor cell in red), leading to somatic mosaicism in the resultant tumor. Alternatively, tumorigenesis can be triggered by a *de novo* LINE-1 retrotransposition event that acts as a potential "driver" mutation (bottom panel), leading to the clonal amplification of the insertion-bearing cell (red cell). Additional references are provided in the text. [doi:10.1128/microbiolspec.MDNA3-0061-2014.f6](https://doi.org/10.1128/microbiolspec.MDNA3-0061-2014.f6)

including lung, liver, ovarian, colorectal, and prostate cancers (69, 296, 399–402). Whether these insertions represent “driver” or “passenger” mutations in tumors requires elucidation (403) (Figure 6B); however, it is clear that advances in DNA sequencing technology should allow an answer to this question in coming years.

CONCLUDING REMARKS

The discovery of TEs in the mid-twentieth century by Barbara McClintock elegantly demonstrated that TEs were arbiters of genetic diversity and that their mobility could influence gene expression and genome structure. It is now clear that TEs are a dynamic component of mammalian genomes that have a profound effect on genome evolution. LINE-1s and SINEs generate both intra-individual and inter-individual genetic variation and continue to sculpt the human genome. Advances made during the previous 13 years have elucidated mechanistic features of LINE-1-mediated retrotransposition, the impact of LINE-1-mediated retrotransposition on the human genome, and how the host defends itself from the onslaught of unabated LINE-1-mediated retrotransposition events. Future research should elucidate how, when, and where LINE-1-mediated retrotransposition events occur in humans, how epigenetic mechanisms act to regulate LINE-1 expression (for example see refs. [389, 404]), and how LINE-1s alter the epigenetic landscape of the genome (e.g., perhaps by playing a role in X-inactivation [405–407]). Indeed, we look forward to gaining a greater understanding about how LINE-1-mediated retrotransposition events contribute to diseases, such as cancer, and if they play a role in human neuronal plasticity. Although once disparaged as a class of “junk DNA”, it is now clear that a deep understanding of LINE-1 biology is paramount to our understanding of the evolutionary forces that have shaped human genomes.

ACKNOWLEDGMENTS

We thank Nancy Leff for outstanding editorial assistance during the preparation of this chapter. We thank Drs. Haig Kazazian, Mark Batzer, and Tomo-ichiro Miyoshi, and members of the Moran laboratory for helpful discussions and for help with editing this Chapter. S.R.R. was supported in part by a National Institutes of Health Training Grant (T32-GM07544). A.J.D. was supported in part by a Fondation pour la Recherche Médicale (FRM) fellowship. J.B.M. was supported in part by a National Institutes of Health Training Grant (T32 GM007315). The laboratory of J.L.G.-P. is supported by CICE-FEDER-P09-CTS-4980, FIS-FEDER-PI11/01489, the European Research Council (ERC-Consolidator ERC-STG-2012-233764), and by an International Early Career Scientist grant from the Howard Hughes Medical Institute (IECS-55007420). J.V.M. is supported by National

Institutes of Health Grant GM060518 and is an Investigator in the Howard Hughes Medical Institute. Portions of this chapter originally appeared in draft form in Dr. Sandra Richardson’s PhD thesis dissertation (408); they have been rewritten and adapted for this chapter. The funders had no role in the decision to submit this work for publication.

Conflicts of interest: J.V.M. is an inventor on the following patent: “Kazazian, H.H., Boeke, J.D., Moran, J.V., and Dombroski, B.A. Compositions and methods of use of mammalian retrotransposons. Application No. 60/006,831; Patent number 6,150,160; Issued November 21, 2000.” He has not made any money from this patent and discloses this information voluntarily.

REFERENCES

1. Orgel LE, Crick FH, Sapienza C. 1980. Selfish DNA. *Nature* 288:645–646.
2. Doolittle WF, Sapienza C. 1980. Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284:601–603.
3. Craig NL, Craigie R, Gellert M, Lambowitz AM. 2002. p 1–1204. *Mobile DNA II*. ASM Press, Washington, D.C.
4. McClintock B. 1950. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci U S A* 36:344–355.
5. Fedoroff N, Wessler S, Shure M. 1983. Isolation of the transposable maize controlling elements Ac and Ds. *Cell* 35:235–242.
6. Boeke JD, Garfinkel DJ, Styles CA, Fink GR. 1985. Ty elements transpose through an RNA intermediate. *Cell* 40:491–500.
7. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramsay J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsieck G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G,

- Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ. 2001. Initial sequencing and analysis of the human genome. *Nature* 409:860–921.
8. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hattton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. 2001. The sequence of the human genome. *Science* 291:1304–1351.
9. Waring M, Britten RJ. 1966. Nucleotide sequence repetition: a rapidly reassociating fraction of mouse DNA. *Science* 154:791–794.
10. Britten RJ, Kohne DE. 1968. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science* 161:529–540.
11. Beck CR, Garcia-Perez JL, Badge RM, Moran JV. 2011. LINE-1 elements in structural variation and disease. *Annu Rev Genomics Hum Genet* 12:187–215.
12. Belancio VP, Roy-Engel AM, Deininger PL. 2010. All y'all need to know 'bout retroelements in cancer. *Semin Cancer Biol* 20:200–210.
13. Levin HL, Moran JV. 2011. Dynamic interactions between transposable elements and their hosts. *Nat Rev Genet* 12:615–627.
14. Hancks DC, Kazazian HH, Jr. 2012. Active human retrotransposons: variation and disease. *Curr Opin Genet Dev* 22:191–203.
15. Richardson SR, Morell S, Faulkner GJ. 2014. L1 Retrotransposons and somatic mosaicism in the brain. *Annu Rev Genet*. 48:1–27.
16. Goodier JL, Kazazian HH, Jr. 2008. Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell* 135:23–35.
17. Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. *Nat Rev Genet* 9:397–405.
18. Babushok DV, Kazazian HH, Jr. 2007. Progress in understanding the biology of the human mutagen LINE-1. *Hum Mutat* 28:527–539.
19. Burns KH, Boeke JD. 2012. Human transposon tectonics. *Cell* 149:740–752.
20. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas EJ, 3rd, Zody MC, Mauceli E, Xie X, Breen M, Wayne RK, Ostrander EA, Ponting CP, Galibert F, Smith DR, DeJong PJ, Kirkness E, Alvarez P, Biagi T, Brockman W, Butler J, Chin CW, Cook A, Cuff J, Daly MJ, DeCaprio D, Gnerre S, Grabherr M, Kellis M, Kleber M, Bardeleben C, Goodstadt L, Heger A, Hitte C, Kim L, Koepfli KP, Parker HG, Pollinger JP, Searle SM, Sutter NB, Thomas R, Webber C, Baldwin J, Abebe A, Abouelleil A, Aftuck L, Ait-Zahra M, Aldredge T, Allen N, An P, Anderson S, Antoine C, Arachchi H, Aslam A, Ayotte L, Bachantsang P, Barry A, Bayul T, Benamara M, Berlin A, Bessette D, Blitshteyn B, Bloom T, Blye J, Boguslavskiy L, Bonnet C, Boukhgalter B, Brown A, Cahill P, Calixte N, Camarata J, Cheshatsang Y, Chu J, Citroen H, Collymore A, Cooke P, Dawoo T, Daza R, Decktor K, DeGray S, Dhargay N, Dooley K, Dorje P, Dorjee K, Dorris L, Duffey N, Dupes A, Egbremolen O, Elong R, Falk J, Farina A, Faro S, Ferguson D, Ferreira P, Fisher S, FitzGerald M, Foley K, Foley C, Franke A, Friedrich D, Gage D, Garber M, Gearin G, Giannoukos G, Goode T, Goyette A, Graham J, Grandbois E, Gyaltsen K, Hafez N, Hagopian D, Hagos B, Hall J, Healy C, Hegarty R, Honan T, Horn A, Houde N, Hughes L, Hunnicutt L, Husby M, Jester B, Jones C, Kamat A, Kanga B, Kells C, Khazanovich D, Kieu AC, Kisner P, Kumar M, Lance K, Landers T, Lara M, Lee W, Leger JP, Lennon N, Leuper L, LeVine S, Liu J, Liu X, Lokyitsang Y, Lokyitsang T, Lui A, Macdonald J, Major J, Marabella R, Maru K, Matthews C, McDonough S, Mehta T, Meldrim J, Melnikov A, Meneus L, Mihalev A, Mihova T, Miller K, Mittelman R, Mlenga V, Mulrain L, Munson G, Navidi A, Naylor J, Nguyen T, Nguyen N, Nguyen C, Nicol R, Norbu N, Norbu C, Novod N, Nyima T, Olandt P, O'Neill B, O'Neill K, Osman S, Oyono L, Patti C, Perrin D, Phunkhang P, Pierre F, Priest M, Rachupka A, Raghuraman S, Rameau R, Ray V, Raymond C, Rege F, Rise C, Rogers J, Rogov P, Sahalie J, Settupalli S, Sharpe T, Shea T, Sheehan M, Sherpa N, Shi J, Shih D, Sloan J, Smith C, Sparrow T, Stalker J, Stange-Thomann N, Stavropoulos S, Stone C, Stone S, Sykes S, Tchuinga P, Tenzing P, Tesfaye S, Thoulutsang D, Thoulutsang Y, Topham K, Topping I, Tsamla T, Vassiliev H, Venkataraman V, Vo A, Wangchuk T, Wangdi T, Weiland M, Wilkinson J, Wilson A, Yadav S, Yang S, Yang X, Young G, Yu Q, Zainoun J, Zembek L, Zimmer A, Lander ES. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438:803–819.
21. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyas E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, Karolchik D, Kasprzyk A, Kawai J, Keibler E, Kells C, Kent WJ, Kirby A, Kolbe DL, Korfi I, Kucherlapati RS, Kulbokas EJ, Kulp D, Landers T, Leger JP, Leonard S, Letunic I, Levine R, Li J, Li M, Lloyd C, Lucas S, Ma B, Maglott DR,

- Mardis ER, Matthews L, Mauceli E, Mayer JH, McCarthy M, McCombie WR, McLaren S, McLay K, McPherson JD, Meldrim J, Meredith B, Mesirov JP, Miller W, Miner TL, Mongin E, Montgomery KT, Morgan M, Mott R, Mullikin JC, Muzny DM, Nash WE, Nelson JO, Nhan MN, Nicol R, Ning Z, Nusbaum C, O'Connor MJ, Okazaki Y, Oliver K, Overton-Larty E, Pachter L, Parra G, Pepin KH, Peterson J, Pevzner P, Plumb R, Pohl CS, Poliakov A, Ponce TC, Ponting CP, Potter S, Quail M, Raymond A, Roe BA, Roskin KM, Rubin EM, Rust AG, Santos R, Sapojnikov V, Schultz B, Schultz J, Schwartz MS, Schwartz S, Scott C, Seaman S, Searle S, Sharpe T, Sheridan A, Shownkeen R, Sims S, Singer JB, Slater G, Smit A, Smith DR, Spencer B, Stabenau A, Stange-Thomann N, Sugnet C, Suyama M, Tesler G, Thompson J, Torrents D, Trevaskis E, Tromp J, Ucla C, Ureta-Vidal A, Vinson JP, Von Niederhausern AC, Wade CM, Wall M, Weber RJ, Weiss RB, Wendl MC, West AP, Wetterstrand K, Wheeler R, Whelan S, Wierzbowski J, Willey D, Williams S, Wilson RK, Winter E, Worley KC, Wyman D, Yang S, Yang SP, Zdobnov EM, Zody MC, Lander ES. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520–562.
22. de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. 2011. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* 7:e1002384.
23. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. 1997. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* 91:501–510.
24. Shapiro JA. 1979. Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. *Proc Natl Acad Sci U S A* 76:1933–1937.
25. Kleckner N. 1990. Regulation of transposition in bacteria. *Annu Rev Cell Biol* 6:297–327.
26. Rio DC. 2002. P transposable elements in *Drosophila melanogaster*, p 484–518. In Craig NL, Craigie R, Gellert M, Lambowitz AM (ed), *Mobile DNA II*. ASM Press, Washington, D.C.
27. Lazarow K, Doll ML, Kunze R. 2013. Molecular biology of maize Ac/Ds elements: an overview. *Methods Mol Biol* 1057:59–82.
28. Cary LC, Goebel M, Corsaro BG, Wang HG, Rosen E, Fraser MJ. 1989. Transposon mutagenesis of baculoviruses: analysis of Trichoplusia ni transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. *Virology* 172:156–169.
29. Fraser MJ, Cary L, Boonvisudhi K, Wang HG. 1995. Assay for movement of Lepidopteran transposon IFP2 in insect cells using a baculovirus genome as a target DNA. *Virology* 211:397–407.
30. Ding S, Wu X, Li G, Han M, Zhuang Y, Xu T. 2005. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* 122:473–483.
31. Rad R, Rad L, Wang W, Cadinanos J, Vassiliou G, Rice S, Campos LS, Yusa K, Banerjee R, Li MA, de la Rosa J, Strong A, Lu D, Ellis P, Conte N, Yang FT, Liu P, Bradley A. 2010. PiggyBac transposon mutagenesis: a tool for cancer gene discovery in mice. *Science* 330:1104–1107.
32. Ivics Z, Li MA, Mates L, Boeke JD, Nagy A, Bradley A, Izsvak Z. 2009. Transposon-mediated genome manipulation in vertebrates. *Nat Methods* 6:415–422.
33. Largaespada DA. 2009. Transposon mutagenesis in mice. *Methods Mol Biol* 530:379–390.
34. Ray DA, Feschotte C, Pagan HJ, Smith JD, Pritham EJ, Arensburger P, Atkinson PW, Craig NL. 2008. Multiple waves of recent DNA transposon activity in the bat, *Myotis lucifugus*. *Genome Res* 18:717–728.
35. Ray DA, Pagan HJ, Thompson ML, Stevens RD. 2007. Bats with hATs: evidence for recent DNA transposon activity in genus *Myotis*. *Mol Biol Evol* 24:632–639.
36. Mitra R, Li X, Kapusta A, Mayhew D, Mitra RD, Feschotte C, Craig NL. 2013. Functional characterization of piggyBat from the bat *Myotis lucifugus* unveils an active mammalian DNA transposon. *Proc Natl Acad Sci U S A* 110:234–239.
37. Pace JK, 2nd, Feschotte C. 2007. The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage. *Genome Res* 17:422–432.
38. Volf JN. 2006. Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. *Bioessays* 28:913–922.
39. Kapitonov VV, Jurka J. 2005. RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS Biol* 3:e181.
40. Majumdar S, Singh A, Rio DC. 2013. The human THAP9 gene encodes an active P-element DNA transposase. *Science* 339:446–448.
41. Bannert N, Kurth R. 2006. The evolutionary dynamics of human endogenous retroviral families. *Annu Rev Genomics Hum Genet* 7:149–173.
42. Belshaw R, Dawson AL, Woolven-Allen J, Redding J, Burt A, Tristem M. 2005. Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K(HML2): implications for present-day activity. *J Virol* 79:12507–12514.
43. Macfarlane C, Simmonds P. 2004. Allelic variation of HERV-K(HML-2) endogenous retroviral elements in human populations. *J Mol Evol* 59:642–656.
44. Moyes D, Griffiths DJ, Venables PJ. 2007. Insertional polymorphisms: a new lease of life for endogenous retroviruses in human disease. *Trends Genet* 23:326–333.
45. Shin W, Lee J, Son SY, Ahn K, Kim HS, Han K. 2013. Human-specific HERV-K insertion causes genomic variations in the human genome. *PLoS One* 8:e60605.
46. Hughes JF, Coffin JM. 2004. Human endogenous retrovirus K solo-LTR formation and insertional polymorphisms: implications for human and viral evolution. *Proc Natl Acad Sci U S A* 101:1668–1672.
47. Yohn CT, Jiang Z, McGrath SD, Hayden KE, Khaitovich P, Johnson ME, Eichler MY, McPherson JD, Zhao S, Paabo S, Eichler EE. 2005. Lineage-specific expansions of retroviral insertions within the genomes of African great apes but not humans and orangutans. *PLoS Biol* 3:e110.
48. Dewannieux M, Harper F, Richaud A, Letzelter C, Ribet D, Pierron G, Heidmann T. 2006. Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. *Genome Res* 16:1548–1556.
49. Lee YN, Bieniasz PD. 2007. Reconstitution of an infectious human endogenous retrovirus. *PLoS Pathog* 3:e10.
50. Stocking C, Kozak CA. 2008. Murine endogenous retroviruses. *Cell Mol Life Sci* 65:3383–3398.
51. Maksakova IA, Romanish MT, Gagnier L, Dunn CA, van de Lagemaat LN, Mager DL. 2006. Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. *PLoS Genet* 2:e2.
52. Burton FH, Loeb DD, Voliva CF, Martin SL, Edgell MH, Hutchison CA, 3rd. 1986. Conservation throughout mammalia and extensive protein-encoding capacity of the highly repeated DNA long interspersed sequence one. *J Mol Biol* 187:291–304.
53. Smit AF, Toth G, Riggs AD, Jurka J. 1995. Ancestral, mammalian-wide subfamilies of LINE-1 repetitive sequences. *J Mol Biol* 246:401–417.
54. Yang L, Bruntsfeld J, Scott L, Wichman H. 2014. Reviving the dead: history and reactivation of an extinct l1. *PLoS Genet* 10:e1004395.
55. Khan H, Smit A, Boissinot S. 2006. Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. *Genome Res* 16:78–87.
56. Jacobs FM, Greenberg D, Nguyen N, Haeussler M, Ewing AD, Katzman S, Paten B, Salama SR, Haussler D. 2014. An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature* [Epub ahead of print].
57. Castro-Diaz N, Ecco G, Coluccio A, Kapopoulou A, Yazdanpanah B, Friedli M, Duc J, Jang SM, Turelli P, Trono D. 2014. Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes Dev* 28:1397–1409.

58. Beck CR, Collier P, Macfarlane C, Malig M, Kidd JM, Eichler EE, Badge RM, Moran JV. 2010. LINE-1 retrotransposition activity in human genomes. *Cell* 141:1159–1170.
59. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH, Jr. 2003. Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A* 100:5280–5285.
60. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH, Jr. 1996. High frequency retrotransposition in cultured mammalian cells. *Cell* 87:917–927.
61. Sassaman DM, Dombroski BA, Moran JV, Kimberland ML, Naas TP, DeBerardinis RJ, Gabriel A, Swergold GD, Kazazian HH, Jr. 1997. Many human L1 elements are capable of retrotransposition. *Nat Genet* 16:37–43.
62. Skowronski J, Fanning TG, Singer MF. 1988. Unit-length line-1 transcripts in human teratocarcinoma cells. *Mol Cell Biol* 8:1385–1397.
63. Boissinot S, Chevret P, Furano AV. 2000. L1 (LINE-1) retrotransposon evolution and amplification in recent human history. *Mol Biol Evol* 17:915–928.
64. Dombroski BA, Mathias SL, Nanthakumar E, Scott AF, Kazazian HH, Jr. 1991. Isolation of an active human transposable element. *Science* 254:1805–1808.
65. Ovchinnikov I, Troxel AB, Swergold GD. 2001. Genomic characterization of recent human LINE-1 insertions: evidence supporting random insertion. *Genome Res* 11:2050–2058.
66. Sheen FM, Sherry ST, Risch GM, Robichaux M, Nasidze I, Stoneking M, Batzer MA, Swergold GD. 2000. Reading between the LINEs: human genomic variation induced by LINE-1 retrotransposition. *Genome Res* 10:1496–508.
67. Badge RM, Alisch RS, Moran JV. 2003. ATLAS: a system to selectively identify human-specific L1 insertions. *Am J Hum Genet* 72:823–838.
68. Ewing AD, Kazazian HH, Jr. 2010. High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome Res* 20:1262–1270.
69. Iskow RC, McCabe MT, Mills RE, Torene S, Pittard WS, Neuwald AF, Van Meir EG, Vertino PM, Devine SE. 2010. Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell* 141:1253–1261.
70. Boissinot S, Entezam A, Young L, Munson PJ, Furano AV. 2004. The insertional history of an active family of L1 retrotransposons in humans. *Genome Res* 14:1221–1231.
71. Myers JS, Vincent BJ, Udall H, Watkins WS, Morrish TA, Kilroy GE, Swergold GD, Henke J, Henke L, Moran JV, Jorde LB, Batzer MA. 2002. A comprehensive analysis of recently integrated human Ta L1 elements. *Am J Hum Genet* 71:312–326.
72. Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, Gibbs RA, Hurles ME, McVean GA. 2010. A map of human genome variation from population-scale sequencing. *Nature* 467:1061–1073.
73. Grimaldi G, Skowronski J, Singer MF. 1984. Defining the beginning and end of KpnI family segments. *EMBO J* 3:1753–1759.
74. Scott AF, Schmeckpeper BJ, Abdelrazik M, Comey CT, O'Hara B, Rossiter JP, Cooley T, Heath P, Smith KD, Margolet L. 1987. Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. *Genomics* 1:113–125.
75. Swergold GD. 1990. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol Cell Biol* 10:6718–6729.
76. Speck M. 2001. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol Cell Biol* 21:1973–1985.
77. Holmes SE, Singer MF, Swergold GD. 1992. Studies on p40, the leucine zipper motif-containing protein encoded by the first open reading frame of an active human LINE-1 transposable element. *J Biol Chem* 267:19765–19768.
78. Hohjoh H, Singer MF. 1996. Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *Embo J* 15:630–639.
79. Khazina E, Truffault V, Buttner R, Schmidt S, Coles M, Weichenrieder O. 2011. Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. *Nat Struct Mol Biol* 18:1006–1014.
80. Martin SL, Bushman FD. 2001. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. *Mol Cell Biol* 21:467–475.
81. Ergun S, Buschmann C, Heukeshoven J, Dammann K, Schnieders F, Lauke H, Chalajour F, Kilic N, Stratling WH, Schumann GG. 2004. Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. *J Biol Chem* 279:27753–27763.
82. Doucet AJ, Hulme AE, Sahinovic E, Kulpa DA, Moldovan JB, Kopera HC, Athanikar JN, Hasnaoui M, Bucheton A, Moran JV, Gilbert N. 2010. Characterization of LINE-1 ribonucleoprotein particles. *PLoS Genet* 6:e1001150.
83. Feng Q, Moran J, Kazazian H, Boeke JD. 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 87:905–916.
84. Mathias SL, Scott AF, Kazazian HH, Jr., Boeke JD, Gabriel A. 1991. Reverse transcriptase encoded by a human transposable element. *Science* 254:1808–1810.
85. Fanning T, Singer M. 1987. The LINE-1 DNA sequences in four mammalian orders predict proteins that conserve homologies to retrovirus proteins. *Nucleic Acids Res* 15:2251–2260.
86. Goodier JL, Ostertag EM, Du K, Kazazian HH, Jr. 2001. A novel active L1 retrotransposon subfamily in the mouse. *Genome Res* 11:1677–1685.
87. Furano AV. 2000. The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. *Prog Nucleic Acid Res Mol Biol* 64:255–294.
88. Sookdeo A, Hepp CM, McClure MA, Boissinot S. 2013. Revisiting the evolution of mouse LINE-1 in the genomic era. *Mob DNA* 4:3.
89. Adey NB, Tollefsbol TO, Sparks AB, Edgell MH, Hutchison CA, 3rd. 1994. Molecular resurrection of an extinct ancestral promoter for mouse L1. *Proc Natl Acad Sci U S A* 91:1569–1573.
90. Naas TP, DeBerardinis RJ, Moran JV, Ostertag EM, Kingsmore SF, Seldin MF, Hayashizaki Y, Martin SL, Kazazian HH. 1998. An actively retrotransposing, novel subfamily of mouse L1 elements. *EMBO J* 17:590–597.
91. DeBerardinis RJ, Goodier JL, Ostertag EM, Kazazian HH, Jr. 1998. Rapid amplification of a retrotransposon subfamily is evolving the mouse genome. *Nat Genet* 20:288–290.
92. Dewannieux M, Esnault C, Heidmann T. 2003. LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet* 35:41–48.
93. Dewannieux M, Heidmann T. 2005. L1-mediated retrotransposition of murine B1 and B2 SINEs recapitulated in cultured cells. *J Mol Biol* 349:241–247.
94. Raiz J, Damert A, Chira S, Held U, Klawitter S, Hamdorf M, Lower J, Stratling WH, Lower R, Schumann GG. 2012. The non-autonomous retrotransposon SVA is trans-mobilized by the human LINE-1 protein machinery. *Nucleic Acids Res* 40:1666–1683.
95. Hancks DC, Goodier JL, Mandal PK, Cheung LE, Kazazian HH, Jr. 2011. Retrotransposition of marked SVA elements by human L1s in cultured cells. *Hum Mol Genet* 20:3386–3400.
96. Hancks DC, Mandal PK, Cheung LE, Kazazian HH, Jr. 2012. The minimal active human SVA retrotransposon requires only the 5'-hexamer and Alu-like domains. *Mol Cell Biol* 32:4718–4726.
97. Esnault C, Maestre J, Heidmann T. 2000. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet* 24:363–367.
98. Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, Boeke JD, Moran JV. 2001. Human L1 retrotransposition: cis preference versus trans complementation. *Mol Cell Biol* 21:1429–1439.

99. Jurka J. 1997. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proc Natl Acad Sci U S A* 94:1872–1877.
100. Rubin CM, Houck CM, Deininger PL, Friedmann T, Schmid CW. 1980. Partial nucleotide sequence of the 300-nucleotide interspersed repeated human DNA sequences. *Nature* 284:372–374.
101. Ullu E, Tschudi C. 1984. Alu sequences are processed 7SL RNA genes. *Nature* 312:171–172.
102. Chu WM, Liu WM, Schmid CW. 1995. RNA polymerase III promoter and terminator elements affect Alu RNA expression. *Nucleic Acids Res* 23:1750–1757.
103. Batzer MA, Deininger PL. 2002. Alu repeats and human genomic diversity. *Nat Rev Genet* 3:370–379.
104. Chesnokov I, Schmid CW. 1996. Flanking sequences of an Alu source stimulate transcription in vitro by interacting with sequence-specific transcription factors. *J Mol Evol* 42:30–36.
105. Comeaux MS, Roy-Engel AM, Hedges DJ, Deininger PL. 2009. Diverse cis factors controlling Alu retrotransposition: what causes Alu elements to die? *Genome Res* 19:545–555.
106. Dewannieux M, Heidmann T. 2005. Role of poly(A) tail length in Alu retrotransposition. *Genomics* 86:378–381.
107. Goodier JL, Marais RJ. 1998. Terminator-specific recycling of a B1-Alu transcription complex by RNA polymerase III is mediated by the RNA terminus-binding protein La. *J Biol Chem* 273:26110–26116.
108. Liu WM, Schmid CW. 1993. Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation. *Nucleic Acids Res* 21:1351–1359.
109. Ullu E, Weiner AM. 1985. Upstream sequences modulate the internal promoter of the human 7SL RNA gene. *Nature* 318:371–374.
110. Jurka J, Smith T. 1988. A fundamental division in the Alu family of repeated sequences. *Proc Natl Acad Sci U S A* 85:4775–4778.
111. Slagel V, Flemington E, Traina-Dorge V, Bradshaw H, Deininger P. 1987. Clustering and subfamily relationships of the Alu family in the human genome. *Mol Biol Evol* 4:19–29.
112. Britten RJ, Baron WF, Stout DB, Davidson EH. 1988. Sources and evolution of human Alu repeated sequences. *Proc Natl Acad Sci U S A* 85:4770–4774.
113. Willard C, Nguyen HT, Schmid CW. 1987. Existence of at least three distinct Alu subfamilies. *J Mol Evol* 26:180–186.
114. Batzer MA, Deininger PL, Hellmann-Blumberg U, Jurka J, Labuda D, Rubin CM, Schmid CW, Zietkiewicz E, Zuckerkandl E. 1996. Standardized nomenclature for Alu repeats. *J Mol Evol* 42:3–6.
115. Batzer MA, Schmid CW, Deininger PL. 1993. Evolutionary analyses of repetitive DNA sequences. *Methods Enzymol* 224:213–232.
116. Carroll ML, Roy-Engel AM, Nguyen SV, Salem AH, Vogel E, Vincent B, Myers J, Ahmad Z, Nguyen L, Sammarco M, Watkins WS, Henke J, Makalowski W, Jorde LB, Deininger PL, Batzer MA. 2001. Large-scale analysis of the Alu Ya5 and Yb8 subfamilies and their contribution to human genomic diversity. *J Mol Biol* 311:17–40.
117. Witherspoon DJ, Zhang Y, Xing J, Watkins WS, Ha H, Batzer MA, Jorde LB. 2013. Mobile element scanning (ME-Scan) identifies thousands of novel Alu insertions in diverse human populations. *Genome Res* 23:1170–1181.
118. Minghetti PP, Dugaiczak A. 1993. The emergence of new DNA repeats and the divergence of primates. *Proc Natl Acad Sci U S A* 90:1872–1876.
119. Batzer MA, Stoneking M, Alegria-Hartman M, Bazan H, Kass DH, Shaikh TH, Novick GE, Ioannou PA, Scheer WD, Herrera RJ, et al. 1994. African origin of human-specific polymorphic Alu insertions. *Proc Natl Acad Sci U S A* 91:12288–12292.
120. Nikaido M, Rooney AP, Okada N. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: hippopotamuses are the closest extant relatives of whales. *Proc Natl Acad Sci U S A* 96:10261–10266.
121. Okada N. 1991. SINES. *Curr Opin Genet Dev* 1:498–504.
122. Ray DA, Xing J, Salem AH, Batzer MA. 2006. SINES of a nearly perfect character. *Syst Biol* 55:928–935.
123. Witherspoon DJ, Marchani EE, Watkins WS, Ostler CT, Wooding SP, Anders BA, Fowlkes JD, Boissinot S, Furano AV, Ray DA, Rogers AR, Batzer MA, Jorde LB. 2006. Human population genetic structure and diversity inferred from polymorphic L1(LINE-1) and Alu insertions. *Hum Hered* 62:30–46.
124. Konkel MK, Walker JA, Batzer MA. 2010. LINEs and SINES of primate evolution. *Evol Anthropol* 19:236–249.
125. Bennett EA, Keller H, Mills RE, Schmidt S, Moran JV, Weichenrieder O, Devine SE. 2008. Active Alu retrotransposons in the human genome. *Genome Res* 18:1875–1883.
126. Sarrowa J, Chang DY, Marais RJ. 1997. The decline in human Alu retroposition was accompanied by an asymmetric decrease in SRP9/14 binding to dimeric Alu RNA and increased expression of small cytoplasmic Alu RNA. *Mol Cell Biol* 17:1144–1151.
127. Ostertag EM, Goodier JL, Zhang Y, Kazazian HH, Jr. 2003. SVA elements are nonautonomous retrotransposons that cause disease in humans. *Am J Hum Genet* 73:1444–1451.
128. Wang H, Xing J, Grover D, Hedges DJ, Han K, Walker JA, Batzer MA. 2005. SVA elements: a hominid-specific retroposon family. *J Mol Biol* 354:994–1007.
129. Hancks DC, Kazazian HH, Jr. 2010. SVA retrotransposons: Evolution and genetic instability. *Semin Cancer Biol* 20:234–245.
130. Savage AL, Bubb VJ, Breen G, Quinn JP. 2013. Characterisation of the potential function of SVA retrotransposons to modulate gene expression patterns. *BMC Evol Biol* 13:101.
131. Hara T, Hirai Y, Baicharoen S, Hayakawa T, Hirai H, Koga A. 2012. A novel composite retrotransposon derived from or generated independently of the SVA (SINE/VNTR/Alu) transposon has undergone proliferation in gibbon genomes. *Genes Genet Syst* 87:181–190.
132. Ianc B, Ochis C, Persch R, Popescu O, Damert A. 2014. Hominoid Composite Non-LTR Retrotransposons-Variety, Assembly, Evolution, and Structural Determinants of Mobilization. *Mol Biol Evol* 31:2847–2864.
133. Carbone L, Harris RA, Mootnick AR, Milosavljevic A, Martin DI, Rocchi M, Capozzi O, Archidiacono N, Konkel MK, Walker JA, Batzer MA, de Jong PJ. 2012. Centromere remodeling in *Hoolock leuconedys* (Hylobatidae) by a new transposable element unique to the gibbons. *Genome Biol Evol* 4:648–658.
134. Carbone L, Harris RA, Gnerre S, Veeramah KR, Lorente-Galdos B, Huddleston J, Meyer TJ, Herrero J, Roos C, Aken B, Anaclerio F, Archidiacono N, Baker C, Barrell D, Batzer MA, Beal K, Blancher A, Bohron CL, Brameier M, Campbell MS, Capozzi O, Casola C, Chiatante G, Cree A, Damert A, de Jong PJ, Dumas L, Fernandez-Callejo M, Flicek P, Fuchs NV, Gut I, Gut M, Hahn MW, Hernandez-Rodriguez J, Hillier LW, Hubley R, Ianc B, Izsvak Z, Jablonski NG, Johnstone LM, Karimpour-Fard A, Konkel MK, Kostka D, Lazar NH, Lee SL, Lewis LR, Liu Y, Locke DP, Mallick S, Mendez FL, Muffato M, Nazareth LV, Nevenon KA, O'Bleness M, Ochis C, Odom DT, Pollard KS, Quilez J, Reich D, Rocchi M, Schumann GG, Searle S, Sikela JM, Skollar G, Smit A, Sonmez K, ten Hallers B, Terhune E, Thomas GW, Ullmer B, Ventura M, Walker JA, Wall JD, Walter L, Ward MC, Wheelan SJ, Whelan CW, White S, Wilhelm LJ, Woerner AE, Yandell M, Zhu B, Hammer MF, Marques-Bonet T, Eichler EE, Fulton L, Fronick C, Muzny DM, Warren WC, Worley KC, Rogers J, Wilson RK, Gibbs RA. 2014. Gibbon genome and the fast karyotype evolution of small apes. *Nature* 513:195–201.
135. Krayev AS, Kramerov DA, Skryabin KG, Ryskov AP, Bayev AA, Georgiev GP. 1980. The nucleotide sequence of the ubiquitous repetitive DNA sequence B1 complementary to the most abundant class of mouse fold-back RNA. *Nucleic Acids Res* 8:1201–1215.
136. Quentin Y. 1994. A master sequence related to a free left Alu monomer (FLAM) at the origin of the B1 family in rodent genomes. *Nucleic Acids Res* 22:2222–2227.

137. Daniels GR, Deininger PL. 1985. Repeat sequence families derived from mammalian tRNA genes. *Nature* 317:819–822.
138. DeChiara TM, Brosius J. 1987. Neural BC1 RNA: cDNA clones reveal nonrepetitive sequence content. *Proc Natl Acad Sci U S A* 84:2624–2628.
139. Lee IY, Westaway D, Smit AF, Wang K, Seto J, Chen L, Acharya C, Ankener M, Baskin D, Cooper C, Yao H, Prusiner SB, Hood LE. 1998. Complete genomic sequence and analysis of the prion protein gene region from three mammalian species. *Genome Res* 8:1022–1037.
140. Serdobova IM, Kramerov DA. 1998. Short retrotransposons of the B2 superfamily: evolution and application for the study of rodent phylogeny. *J Mol Evol* 46:202–214.
141. Weiner AM, Deininger PL, Efstratiadis A. 1986. Nonviral retrotransposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu Rev Biochem* 55:631–661.
142. Vanin EF. 1985. Processed pseudogenes: characteristics and evolution. *Annu Rev Genet* 19:253–272.
143. Torrents D, Suyama M, Zdobnov E, Bork P. 2003. A genome-wide survey of human pseudogenes. *Genome Res* 13:2559–2567.
144. Zhang Z, Harrison P, Gerstein M. 2002. Identification and analysis of over 2000 ribosomal protein pseudogenes in the human genome. *Genome Res* 12:1466–82.
145. Zhang Z, Harrison PM, Liu Y, Gerstein M. 2003. Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome. *Genome Res* 13:2541–58.
146. Ewing AD, Ballinger TJ, Earl D, Harris CC, Ding L, Wilson RK, Haussler D. 2013. Retrotransposition of gene transcripts leads to structural variation in mammalian genomes. *Genome Biol* 14:R22.
147. Schrider DR, Navarro FC, Galante PA, Parmigiani RB, Camargo AA, Hahn MW, de Souza SJ. 2013. Gene copy-number polymorphism caused by retrotransposition in humans. *PLoS Genet* 9:e1003242.
148. Abyzov A, Iskow R, Gokcumen O, Radke DW, Balasubramanian S, Pei B, Habegger L, Lee C, Gerstein M. 2013. Analysis of variable retro-duplications in human populations suggests coupling of retrotransposition to cell division. *Genome Res* 23:2042–2052.
149. Richardson SR, Salvador-Palomeque C, Faulkner GJ. 2014. Diversity through duplication: whole-genome sequencing reveals novel gene retrocopies in the human population. *Bioessays* 36:475–481.
150. Kazazian HH, Jr. 2014. Processed pseudogene insertions in somatic cells. *Mob DNA* 5:20.
151. Hulme AE, Kulpa DA, Garcia-Perez JL, Moran JV. 2006. p 35–72. In Lupski JR, Stankiewicz P (ed), *The impact of LINE-1 retrotransposition on the human genome*. Humana Press, Totowa, New Jersey.
152. Sayah DM, Sokolskaja E, Berthoux L, Luban J. 2004. Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature* 430:569–573.
153. Malfavon-Borja R, Wu LI, Emerman M, Malik HS. 2013. Birth, decay, and reconstruction of an ancient TRIMCyp gene fusion in primate genomes. *Proc Natl Acad Sci U S A* 110:E583–E592.
154. Buzdin A, Gogvadze E, Kovalskaya E, Volchkov P, Ustyugova S, Illarionova A, Fushan A, Vinogradova T, Sverdlov E. 2003. The human genome contains many types of chimeric retrogenes generated through in vivo RNA recombination. *Nucleic Acids Res* 31:4385–4390.
155. Buzdin A, Ustyugova S, Gogvadze E, Vinogradova T, Lebedev Y, Sverdlov E. 2002. A new family of chimeric retrotranscripts formed by a full copy of U6 small nuclear RNA fused to the 3' terminus of 11. *Genomics* 80:402–406.
156. Gilbert N, Lutz S, Morrish TA, Moran JV. 2005. Multiple fates of L1 retrotransposition intermediates in cultured human cells. *Mol Cell Biol* 25:7780–7795.
157. Garcia-Perez JL, Doucet AJ, Bucheton A, Moran JV, Gilbert N. 2007. Distinct mechanisms for trans-mediated mobilization of cellular RNAs by the LINE-1 reverse transcriptase. *Genome Res* 17:602–611.
158. Weber MJ. 2006. Mammalian small nucleolar RNAs are mobile genetic elements. *PLoS Genet* 2:e205.
159. Hasnaoui M, Doucet AJ, Meziane O, Gilbert N. 2009. Ancient repeat sequence derived from U6 snRNA in primate genomes. *Gene* 448:139–144.
160. Kazazian HH, Jr., Wong C, Yousoufian H, Scott AF, Phillips DG, Antonarakis SE. 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature* 332:164–166.
161. Dombroski BA, Scott AF, Kazazian HH, Jr. 1993. Two additional potential retrotransposons isolated from a human L1 subfamily that contains an active retrotransposable element. *Proc Natl Acad Sci U S A* 90:6513–6517.
162. Lutz SM, Vincent BJ, Kazazian HH, Jr., Batzer MA, Moran JV. 2003. Allelic heterogeneity in LINE-1 retrotransposition activity. *Am J Hum Genet* 73:1431–1437.
163. Skowronski J, Singer MF. 1985. Expression of a cytoplasmic LINE-1 transcript is regulated in a human teratocarcinoma cell line. *Proc Natl Acad Sci U S A* 82:6050–6054.
164. Wallace MR, Andersen LB, Saulino AM, Gregory PE, Glover TW, Collins FS. 1991. A de novo Alu insertion results in neurofibromatosis type 1. *Nature* 353:864–866.
165. Hassoun H, Coetzer TL, Vassiliadis JN, Sahr KE, Maalouf GJ, Saad ST, Catanzariti L, Palek J. 1994. A novel mobile element inserted in the alpha spectrin gene: spectrin dayton. A truncated alpha spectrin associated with hereditary elliptocytosis. *J Clin Invest* 94:643–648.
166. van der Klift HM, Tops CM, Hes FJ, Devilee P, Wijnen JT. 2012. Insertion of an SVA element, a nonautonomous retrotransposon, in PMS2 intron 7 as a novel cause of Lynch syndrome. *Hum Mutat* 33:1051–1055.
167. Vogt J, Bengesser K, Claes KB, Wimmer K, Mautner VF, van Minkelen R, Legius E, Brems H, Upadhyaya M, Hogel J, Lazaro C, Rosenbaum T, Bammert S, Messiaen L, Cooper DN, Kehrer-Sawatzki H. 2014. SVA retrotransposon insertion-associated deletion represents a novel mutational mechanism underlying large genomic copy number changes with non-recurrent breakpoints. *Genome Biol* 15:R80.
168. de Boer M, van Leeuwen K, Geissler J, Weemaes CM, van den Berg TK, Kuijpers TW, Warris A, Roos D. 2014. Primary immunodeficiency caused by an exonized retroposed gene copy inserted in the CYBB gene. *Hum Mutat* 35:486–496.
169. Cooke SL, Shlien A, Marshall J, Pipinikas CP, Martincorena I, Tubio JM, Li Y, Menzies A, Mudie L, Ramakrishna M, Yates L, Davies H, Bolli N, Bignell GR, Tarpey PS, Behjati S, Nik-Zainal S, Papaemmanuil E, Teixeira VH, Raine K, O'Meara S, Dodoran MS, Teague JW, Butler AP, Iacobuzio-Donahue C, Santarius T, Grundy RG, Malkin D, Greaves M, Munshi N, Flanagan AM, Bowtell D, Martin S, Larsimont D, Reis-Filho JS, Boussioutas A, Taylor JA, Hayes ND, Janes SM, Futreal PA, Stratton MR, McDermott U, Campbell PJ. 2014. Processed pseudogenes acquired somatically during cancer development. *Nat Commun* 5:3644.
170. Wimmer K, Callens T, Wernstedt A, Messiaen L. 2011. The NF1 gene contains hotspots for L1 endonuclease-dependent de novo insertion. *PLoS Genet* 7:e1002371.
171. Kingsmore SF, Giros B, Suh D, Bieniarz M, Caron MG, Seldin MF. 1994. Glycine receptor beta-subunit gene mutation in spastic mouse associated with LINE-1 element insertion. *Nat Genet* 7:136–141.
172. Mulhardt C, Fischer M, Gass P, Simon-Chazottes D, Guenet JL, Kuhse J, Betz H, Becker CM. 1994. The spastic mouse: aberrant splicing of glycine receptor beta subunit mRNA caused by intronic insertion of L1 element. *Neuron* 13:1003–1015.
173. Takahara T, Ohsumi T, Kuromitsu J, Shibata K, Sasaki N, Okazaki Y, Shibata H, Sato S, Yoshiki A, Kusakabe M, Muramatsu M, Ueki M, Okuda K, Hayashizaki Y. 1996. Dysfunction of the Orleans reeler gene arising from exon skipping due to transposition of a full-length copy of an active L1 sequence into the skipped exon. *Hum Mol Genet* 5:989–993.

174. Kojima T, Nakajima K, Mikoshiba K. 2000. The disabled 1 gene is disrupted by a replacement with L1 fragment in yotari mice. *Brain Res Mol Brain Res* 75:121–127.
175. Perou CM, Pryor RJ, Naas TP, Kaplan J. 1997. The bg allele mutation is due to a LINE1 element retrotransposition. *Genomics* 42:366–368.
176. Yajima I, Sato S, Kimura T, Yasumoto K, Shibahara S, Goding CR, Yamamoto H. 1999. An L1 element intronic insertion in the black-eyed white (Mitf[mi-bw]) gene: the loss of a single Mitf isoform responsible for the pigmentary defect and inner ear deafness. *Hum Mol Genet* 8:1431–1441.
177. Kohrman DC, Harris JB, Meisler MH. 1996. Mutation detection in the med and medJ alleles of the sodium channel Scn8a. Unusual splicing due to a minor class AT-AC intron. *J Biol Chem* 271:17576–17581.
178. Bomar JM, Benke PJ, Slattery EL, Puttagunta R, Taylor LP, Seong E, Nystuen A, Chen W, Albin RL, Patel PD, Kittles RA, Sheffield VC, Burmeister M. 2003. Mutations in a novel gene encoding a CRAL-TRIO domain cause human Cayman ataxia and ataxia/dystonia in the jittery mouse. *Nat Genet* 35:264–269.
179. Gilbert N, Bomar JM, Burmeister M, Moran JV. 2004. Characterization of a mutagenic B1 retrotransposon insertion in the jittery mouse. *Hum Mutat* 24:9–13.
180. Katzir N, Rechavi G, Cohen JB, Unger T, Simoni F, Segal S, Cohen D, Givol D. 1985. “Retroposon” insertion into the cellular oncogene c-myc in canine transmissible venereal tumor. *Proc Natl Acad Sci U S A* 82:1054–1058.
181. Choi Y, Ishiguro N, Shinagawa M, Kim CJ, Okamoto Y, Minami S, Ogiwara K. 1999. Molecular structure of canine LINE-1 elements in canine transmissible venereal tumor. *Anim Genet* 30:51–53.
182. Brooks MB, Gu W, Barnas JL, Ray J, Ray K. 2003. A Line 1 insertion in the Factor IX gene segregates with mild hemophilia B in dogs. *Mamm Genome* 14:788–795.
183. Smith BF, Yue Y, Woods PR, Kornegay JN, Shin JH, Williams RR, Duan D. 2011. An intronic LINE-1 element insertion in the dystrophin gene aborts dystrophin expression and results in Duchenne-like muscular dystrophy in the corgi breed. *Lab Invest* 91:216–231.
184. Wolf ZT, Leslie EJ, Arzi B, Jayashankar K, Karmi N, Jia Z, Rowland DJ, Young A, Safran N, Sliskovic S, Murray JC, Wade CM, Bannasch DL. 2014. A LINE-1 insertion in DLX6 is responsible for cleft palate and mandibular abnormalities in a canine model of Pierre Robin sequence. *PLoS Genet* 10:e1004257.
185. Downs LM, Mellersh CS. 2014. An Intronic SINE insertion in FAM161A that causes exon-skipping is associated with progressive retinal atrophy in Tibetan Spaniels and Tibetan Terriers. *PLoS One* 9:e93990.
186. Clark LA, Wahl JM, Rees CA, Murphy KE. 2006. Retrotransposon insertion in SILV is responsible for merle patterning of the domestic dog. *Proc Natl Acad Sci U S A* 103:1376–1381.
187. Lin L, Faraco J, Li R, Kadotani H, Rogers W, Lin X, Qiu X, de Jong PJ, Nishino S, Mignot E. 1999. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* 98:365–376.
188. Pele M, Tired L, Kessler JL, Blot S, Panthier JJ. 2005. SINE exonic insertion in the PTPLA gene leads to multiple splicing defects and segregates with the autosomal recessive centronuclear myopathy in dogs. *Hum Mol Genet* 14:1417–1427.
189. Parker HG, VonHoldt BM, Quignon P, Margulies EH, Shao S, Mosher DS, Spady TC, Elkhouloun A, Cargill M, Jones PG, Maslen CL, Acland GM, Sutter NB, Kuroki K, Bustamante CD, Wayne RK, Ostrander EA. 2009. An expressed fgf4 retrogene is associated with breed-defining chondrodysplasia in domestic dogs. *Science* 325:995–998.
190. Heidmann T, Heidmann O, Nicolas JF. 1988. An indicator gene to demonstrate intracellular transposition of defective retroviruses. *Proc Natl Acad Sci U S A* 85:2219–2223.
191. Curcio MJ, Garfinkel DJ. 1991. Single-step selection for Ty1 element retrotransposition. *Proc Natl Acad Sci USA* 88:936–940.
192. Wei W, Morrish TA, Alisch RS, Moran JV. 2000. A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events. *Anal Biochem* 284:435–438.
193. Freeman JD, Goodchild NL, Mager DL. 1994. A modified indicator gene for selection of retrotransposition events in mammalian cells. *Biotechniques* 17:46, 8–9, 52.
194. Xie Y, Rosser JM, Thompson TL, Boeke JD, An W. 2011. Characterization of L1 retrotransposition with high-throughput dual-luciferase assays. *Nucleic Acids Res* 39:e16.
195. Ostertag EM, Prak ET, DeBerardinis RJ, Moran JV, Kazazian HH, Jr. 2000. Determination of L1 retrotransposition kinetics in cultured cells. *Nucleic Acids Res* 28:1418–1423.
196. Morrish TA, Gilbert N, Myers JS, Vincent BJ, Stamato TD, Taccioli GE, Batzer MA, Moran JV. 2002. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat Genet* 31:159–165.
197. Goodier JL, Zhang L, Vetter MR, Kazazian HH, Jr. 2007. LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. *Mol Cell Biol* 27:6469–6483.
198. Kulpa DA, Moran JV. 2005. Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. *Hum Mol Genet* 14:3237–3248.
199. Kulpa DA, Moran JV. 2006. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat Struct Mol Biol* 13:655–660.
200. Taylor MS, Lacava J, Mita P, Molloy KR, Huang CR, Li D, Adney EM, Jiang H, Burns KH, Chait BT, Rout MP, Boeke JD, Dai L. 2013. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. *Cell* 155:1034–1048.
201. Goodier JL, Cheung LE, Kazazian HH, Jr. 2013. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids Res* 41:7401–7419.
202. Gilbert N, Lutz-Prigge S, Moran JV. 2002. Genomic deletions created upon LINE-1 retrotransposition. *Cell* 110:315–325.
203. Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, Boeke JD. 2002. Human l1 retrotransposition is associated with genetic instability in vivo. *Cell* 110:327–338.
204. Kajikawa M, Okada N. 2002. LINEs mobilize SINEs in the eel through a shared 3' sequence. *Cell* 111:433–444.
205. Sugano T, Kajikawa M, Okada N. 2006. Isolation and characterization of retrotransposition-competent LINEs from zebrafish. *Gene* 365: 74–82.
206. Seleme MC, Vetter MR, Cordaux R, Bastone L, Batzer MA, Kazazian HH, Jr. 2006. Extensive individual variation in L1 retrotransposition capability contributes to human genetic diversity. *Proc Natl Acad Sci U S A* 103:6611–6616.
207. Becker KG, Swergold GD, Ozato K, Thayer RE. 1993. Binding of the ubiquitous nuclear transcription factor YY1 to a cis regulatory sequence in the human LINE-1 transposable element. *Hum Mol Genet* 2:1697–1702.
208. Athanikar JN, Badge RM, Moran JV. 2004. A YY1-binding site is required for accurate human LINE-1 transcription initiation. *Nucleic Acids Res* 32:3846–3855.
209. Dmitriev SE, Andreev DE, Terenin IM, Olovnikov IA, Prassolov VS, Merrick WC, Shatsky IN. 2007. Efficient translation initiation directed by the 900-nucleotide-long and GC-rich 5' untranslated region of the human retrotransposon LINE-1 mRNA is strictly cap dependent rather than internal ribosome entry site mediated. *Mol Cell Biol* 27:4685–4697.
210. Tchenio T, Casella JF, Heidmann T. 2000. Members of the SRY family regulate the human LINE retrotransposons. *Nucleic Acids Res* 28:411–415.
211. Yang N, Zhang L, Zhang Y, Kazazian HH, Jr. 2003. An important role for RUNX3 in human L1 transcription and retrotransposition. *Nucleic Acids Res* 31:4929–4940.

212. Nigumann P, Redik K, Matlik K, Speek M. 2002. Many human genes are transcribed from the antisense promoter of L1 retrotransposon. *Genomics* 79:628–634.
213. Macia A, Munoz-Lopez M, Cortes JL, Hastings RK, Morell S, Lucena-Aguilar G, Marchal JA, Badge RM, Garcia-Perez JL. 2011. Epigenetic control of retrotransposon expression in human embryonic stem cells. *Mol Cell Biol* 31:300–316.
214. Uesaka M, Nishimura O, Go Y, Nakashima K, Agata K, Imamura T. 2014. Bidirectional promoters are the major source of gene activation-associated non-coding RNAs in mammals. *BMC Genomics* 15:35.
215. DeBerardinis RJ, Kazazian HH, Jr. 1999. Analysis of the promoter from an expanding mouse retrotransposon subfamily. *Genomics* 56:317–323.
216. Severynse DM, Hutchison CA, 3rd, Edgell MH. 1992. Identification of transcriptional regulatory activity within the 5' A-type monomer sequence of the mouse LINE-1 retroposon. *Mamm Genome* 2:41–50.
217. Zemojtel T, Penzkofer T, Schultz J, Dandekar T, Badge R, Vingron M. 2007. Exonization of active mouse L1s: a driver of transcriptome evolution? *BMC Genomics* 8:392.
218. Li J, Kannan M, Trivett AL, Liao H, Wu X, Akagi K, Symer DE. 2014. An antisense promoter in mouse L1 retrotransposon open reading frame-1 initiates expression of diverse fusion transcripts and limits retrotransposition. *Nucleic Acids Res* 42:4546–4562.
219. Fort A, Hashimoto K, Yamada D, Salimullah M, Keya CA, Saxena A, Bonetti A, Voineagu I, Bertin N, Kratz A, Noro Y, Wong CH, de Hoon M, Andersson R, Sandelin A, Suzuki H, Wei CL, Koseki H, Hasegawa Y, Forrest AR, Carninci P. 2014. Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. *Nat Genet* 46:558–566.
220. Kapusta A, Kronenberg Z, Lynch VJ, Zhuo X, Ramsay L, Bourque G, Young M, Feschotte C. 2013. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet* 9:e1003470.
221. Kapusta A, Feschotte C. 2014. Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications. *Trends Genet* 30:439–452.
222. Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, Young G, Lucas AB, Ach R, Bruhn L, Yang X, Amit I, Meissner A, Regev A, Rinn JL, Root DE, Lander ES. 2011. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 477:295–300.
223. Macia A, Blanco-Jimenez E, Garcia-Perez JL. 2014. Retrotransposons in pluripotent cells: Impact and new roles in cellular plasticity. *Biochim Biophys Acta* [Epub ahead of print].
224. Leibold DM, Swergold GD, Singer MF, Thayer RE, Dombroski BA, Fanning TG. 1990. Translation of LINE-1 DNA elements in vitro and in human cells. *Proc Natl Acad Sci U S A* 87:6990–6994.
225. Kolosha VO, Martin SL. 2003. High-affinity, non-sequence-specific RNA binding by the open reading frame 1 (ORF1) protein from long interspersed nuclear element 1 (LINE-1). *J Biol Chem* 278:8112–8117.
226. Kolosha VO, Martin SL. 1997. In vitro properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. *Proc Natl Acad Sci U S A* 94:10155–10160.
227. Martin SL. 1991. Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells. *Mol Cell Biol* 11:4804–4807.
228. Hohjoh H, Singer MF. 1997. Ribonuclease and high salt sensitivity of the ribonucleoprotein complex formed by the human LINE-1 retrotransposon. *J Mol Biol* 271:7–12.
229. Martin SL, Branciforte D, Keller D, Bain DL. 2003. Trimeric structure for an essential protein in L1 retrotransposition. *Proc Natl Acad Sci U S A* 100:13815–13820.
230. Basame S, Wai-lun Li P, Howard G, Branciforte D, Keller D, Martin SL. 2006. Spatial assembly and RNA binding stoichiometry of a LINE-1 protein essential for retrotransposition. *J Mol Biol* 357:351–357.
231. Khazina E, Weichenrieder O. 2009. Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. *Proc Natl Acad Sci U S A* 106:731–736.
232. Callahan KE, Hickman AB, Jones CE, Ghirlando R, Furano AV. 2012. Polymerization and nucleic acid-binding properties of human L1 ORF1 protein. *Nucleic Acids Res* 40:813–827.
233. Januszky K, Li PW, Villareal V, Branciforte D, Wu H, Xie Y, Feigon J, Loo JA, Martin SL, Clubb RT. 2007. Identification and solution structure of a highly conserved C-terminal domain within ORF1p required for retrotransposition of long interspersed nuclear element-1. *J Biol Chem* 282:24893–24904.
234. Martin SL. 2010. Nucleic acid chaperone properties of ORF1p from the non-LTR retrotransposon, LINE-1. *RNA Biol* 7:706–711.
235. Martin SL, Bushman D, Wang F, Li PW, Walker A, Cumiskey J, Branciforte D, Williams MC. 2008. A single amino acid substitution in ORF1 dramatically decreases L1 retrotransposition and provides insight into nucleic acid chaperone activity. *Nucleic Acids Res* 36:5845–5854.
236. Evans JD, Peddigari S, Chaurasiya KR, Williams MC, Martin SL. 2011. Paired mutations abolish and restore the balanced annealing and melting activities of ORF1p that are required for LINE-1 retrotransposition. *Nucleic Acids Res* 39:5611–5621.
237. Heras SR, Lopez MC, Garcia-Perez JL, Martin SL, Thomas MC. 2005. The L1Tc C-terminal domain from *Trypanosoma cruzi* non-long terminal repeat retrotransposon codes for a protein that bears two C2H2 zinc finger motifs and is endowed with nucleic acid chaperone activity. *Mol Cell Biol* 25:9209–9220.
238. Dawson A, Hartwood E, Paterson T, Finnegan DJ. 1997. A LINE-like transposable element in *Drosophila*, the I factor, encodes a protein with properties similar to those of retroviral nucleocapsids. *EMBO J* 16:4448–44455.
239. Nakamura M, Okada N, Kajikawa M. 2012. Self-interaction, nucleic acid binding, and nucleic acid chaperone activities are unexpectedly retained in the unique ORF1p of zebrafish LINE. *Mol Cell Biol* 32:458–469.
240. Kajikawa M, Sugano T, Sakurai R, Okada N. 2012. Low dependency of retrotransposition on the ORF1 protein of the zebrafish LINE, ZfL2-1. *Gene* 499:41–47.
241. Goodier JL, Mandal PK, Zhang L, Kazazian HH, Jr. 2010. Discrete subcellular partitioning of human retrotransposon RNAs despite a common mechanism of genome insertion. *Hum Mol Genet* 19:1712–1725.
242. Martin F, Maranon C, Olivares M, Alonso C, Lopez MC. 1995. Characterization of a non-long terminal repeat retrotransposon cDNA (L1Tc) from *Trypanosoma cruzi*: homology of the first ORF with the ape family of DNA repair enzymes. *J Mol Biol* 247:49–59.
243. Mol CD, Kuo CF, Thayer MM, Cunningham RP, Tainer JA. 1995. Structure and function of the multifunctional DNA-repair enzyme exonuclease III. *Nature* 374:381–386.
244. Weichenrieder O, Repanas K, Perrakis A. 2004. Crystal structure of the targeting endonuclease of the human LINE-1 retrotransposon. *Structure (Camb)* 12:975–986.
245. Cost GJ, Boeke JD. 1998. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. *Biochemistry* 37:18081–18093.
246. Cost GJ, Golding A, Schlissel MS, Boeke JD. 2001. Target DNA chromatinization modulates nicking by L1 endonuclease. *Nucleic Acids Res* 29:573–577.
247. Xiong Y, Eickbush TH. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J* 9:3353–3362.
248. Malik HS, Burke WD, Eickbush TH. 1999. The age and evolution of non-LTR retrotransposable elements. *Mol Biol Evol* 16:793–805.
249. Evgen'ev MB, Arkhipova IR. 2005. Penelope-like elements—a new class of retroelements: distribution, function and possible evolutionary significance. *Cytogenet Genome Res* 110:510–521.

250. Dombroski BA, Feng Q, Mathias SL, Sassaman DM, Scott AF, Kazazian HH, Jr., Boeke JD. 1994. An in vivo assay for the reverse transcriptase of human retrotransposon L1 in *Saccharomyces cerevisiae*. *Mol Cell Biol* 14:4485–4492.
251. Piskareva O, Denmukhametova S, Schmatchenko V. 2003. Functional reverse transcriptase encoded by the human LINE-1 from baculovirus-infected insect cells. *Protein Expr Purif* 28:125–130.
252. Bibillo A, Eickbush TH. 2002. High processivity of the reverse transcriptase from a non-long terminal repeat retrotransposon. *J Biol Chem* 277:34836–34845.
253. Piskareva O, Schmatchenko V. 2006. DNA polymerization by the reverse transcriptase of the human L1 retrotransposon on its own template in vitro. *FEBS Lett* 580:661–668.
254. Monot C, Kuciak M, Viollet S, Mir AA, Gabus C, Darlix JL, Cristofari G. 2013. The specificity and flexibility of L1 reverse transcription priming at imperfect T-tracts. *PLoS Genet* 9:e1003499.
255. Piskareva O, Ernst C, Higgins N, Schmatchenko V. 2013. The carboxy-terminal segment of the human LINE-1 ORF2 protein is involved in RNA binding. *FEBS Open Bio* 3:433–437.
256. Alisch RS, Garcia-Perez JL, Muotri AR, Gage FH, Moran JV. 2006. Unconventional translation of mammalian LINE-1 retrotransposons. *Genes Dev* 20:210–224.
257. Li PW, Li J, Timmerman SL, Krushel LA, Martin SL. 2006. The dicistronic RNA from the mouse LINE-1 retrotransposon contains an internal ribosome entry site upstream of each ORF: implications for retrotransposition. *Nucleic Acids Res* 34:853–864.
258. An W, Dai L, Niewiadomska AM, Yetil A, O'Donnell KA, Han JS, Boeke JD. 2011. Characterization of a synthetic human LINE-1 retrotransposon ORFeus-Hs. *Mob DNA* 2:2.
259. Han JS, Boeke JD. 2004. A highly active synthetic mammalian retrotransposon. *Nature* 429:314–318.
260. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324:218–223.
261. Slavoff SA, Mitchell AJ, Schwaid AG, Cabili MN, Ma J, Levin JZ, Karger AD, Budnik BA, Rinn JL, Saghatelian A. 2013. Peptidomic discovery of short open reading frame-encoded peptides in human cells. *Nat Chem Biol* 9:59–64.
262. Usdin K, Furano AV. 1989. The structure of the guanine-rich polypurine:polypyrimidine sequence at the right end of the rat L1 (LINE) element. *J Biol Chem* 264:15681–15687.
263. Moran JV, DeBerardinis RJ, Kazazian HH, Jr. 1999. Exon shuffling by L1 retrotransposition. *Science* 283:1530–1534.
264. Faulkner GJ, Kimura Y, Daub CO, Wani S, Plessy C, Irvine KM, Schroder K, Cloonan N, Steptoe AL, Lassmann T, Waki K, Hornig N, Arakawa T, Takahashi H, Kawai J, Forrest AR, Suzuki H, Hayashizaki Y, Hume DA, Orlando V, Grimmond SM, Carninci P. 2009. The regulated retrotransposon transcriptome of mammalian cells. *Nat Genet* 41:563–571.
265. Luan DD, Korman MH, Jakubczak JL, Eickbush TH. 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* 72:595–605.
266. Mandal PK, Ewing AD, Hancks DC, Kazazian HH, Jr. 2013. Enrichment of processed pseudogene transcripts in L1-ribonucleoprotein particles. *Hum Mol Genet* 22:3730–3748.
267. Dutko JA, Kenny AE, Gamache ER, Curcio MJ. 2010. 5' to 3' mRNA decay factors colocalize with Ty1 gag and human APOBEC3G and promote Ty1 retrotransposition. *J Virol* 84:5052–5066.
268. Larsen LS, Beliakova-Bethell N, Bilanchone V, Zhang M, Lamsa A, Dasilva R, Hatfield GW, Nagashima K, Sandmeyer S. 2008. Ty3 nucleocapsid controls localization of particle assembly. *J Virol* 82:2501–2514.
269. Larsen LS, Zhang M, Beliakova-Bethell N, Bilanchone V, Lamsa A, Nagashima K, Najdi R, Kosaka K, Kovacevic V, Cheng J, Baldi P, Hatfield GW, Sandmeyer S. 2007. Ty3 capsid mutations reveal early and late functions of the amino-terminal domain. *J Virol* 81:6957–6972.
270. Kubo S, Seleme MC, Soifer HS, Perez JL, Moran JV, Kazazian HH, Jr., Kasahara N. 2006. L1 retrotransposition in nondividing and primary human somatic cells. *Proc Natl Acad Sci U S A* 103:8036–8041.
271. Goodwin TJ, Ormandy JE, Poulter RT. 2001. L1-like non-LTR retrotransposons in the yeast *Candida albicans*. *Curr Genet* 39:83–91.
272. Dong C, Poulter RT, Han JS. 2009. LINE-like retrotransposition in *Saccharomyces cerevisiae*. *Genetics* 181:301–311.
273. Kinsey JA. 1993. Transnuclear retrotransposition of the Tad element of *Neurospora*. *Proc Natl Acad Sci U S A* 90:9384–9387.
274. Xie Y, Mates L, Ivics Z, Izsvak Z, Martin SL, An W. 2013. Cell division promotes efficient retrotransposition in a stable L1 reporter cell line. *Mob DNA* 4:10.
275. Shi X, Seluanov A, Gorbunova V. 2007. Cell divisions are required for L1 retrotransposition. *Mol Cell Biol* 27:1264–1270.
276. Cost GJ, Feng Q, Jacquier A, Boeke JD. 2002. Human L1 element target-primed reverse transcription in vitro. *EMBO J* 21:5899–5910.
277. Christensen SM, Eickbush TH. 2005. R2 target-primed reverse transcription: ordered cleavage and polymerization steps by protein subunits asymmetrically bound to the target DNA. *Mol Cell Biol* 25:6617–6628.
278. Ostertag EM, Kazazian HH, Jr. 2001. Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition. *Genome Res* 11:2059–2065.
279. Zingler N, Willhoeft U, Brose HP, Schoder V, Jahns T, Hanschmann KM, Morrish TA, Lower J, Schumann GG. 2005. Analysis of 5' junctions of human LINE-1 and Alu retrotransposons suggests an alternative model for 5'-end attachment requiring microhomology-mediated end-joining. *Genome Res* 15:780–789.
280. Gasior SL, Wakeman TP, Xu B, Deininger PL. 2006. The human LINE-1 retrotransposon creates DNA double-strand breaks. *J Mol Biol* 357:1383–1393.
281. Gasior SL, Roy-Engel AM, Deininger PL. 2008. ERCC1/XPF limits L1 retrotransposition. *DNA Repair (Amst)* 7:983–989.
282. Coufal NG, Garcia-Perez JL, Peng GE, Marchetto MC, Muotri AR, Mu Y, Carson CT, Macia A, Moran JV, Gage FH. 2011. Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. *Proc Natl Acad Sci U S A* 108:20382–20387.
283. Boboila C, Alt FW, Schwer B. 2012. Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. *Adv Immunol* 116:1–49.
284. Lavie L, Maldener E, Brouha B, Meese EU, Mayer J. 2004. The human L1 promoter: variable transcription initiation sites and a major impact of upstream flanking sequence on promoter activity. *Genome Res* 14:2253–2260.
285. Babushok DV, Ostertag EM, Courtney CE, Choi JM, Kazazian HH, Jr. 2006. L1 integration in a transgenic mouse model. *Genome Res* 16:240–250.
286. Ichiyanagi K, Nakajima R, Kajikawa M, Okada N. 2007. Novel retrotransposon analysis reveals multiple mobility pathways dictated by hosts. *Genome Res* 17:33–41.
287. Chen J, Rattner A, Nathans J. 2006. Effects of L1 retrotransposon insertion on transcript processing, localization and accumulation: lessons from the retinal degeneration 7 mouse and implications for the genomic ecology of L1 elements. *Hum Mol Genet* 15:2146–2156.
288. Goodier JL, Ostertag EM, Kazazian HH, Jr. 2000. Transduction of 3'-flanking sequences is common in L1 retrotransposition. *Hum Mol Genet* 9:653–657.
289. Pickeral OK, Makalowski W, Boguski MS, Boeke JD. 2000. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. *Genome Res* 10:411–415.

290. Holmes SE, Dombroski BA, Krebs CM, Boehm CD, Kazazian HH, Jr. 1994. A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. *Nat Genet* 7:143–148.
291. Moran JV. 1999. Human L1 retrotransposition: insights and peculiarities learned from a cultured cell retrotransposition assay. *Genetica* 107:39–51.
292. Moran JV, Gilbert N. 2002. Mammalian LINE-1 retrotransposons and related elements, p 836–869. In Craig NL, Craigie R, Gellert M, Lambowitz AM (ed), *Mobile DNA II*. ASM Press, Washington, D.C.
293. Eickbush T. 1999. Exon shuffling in retrospect. *Science* 283:1465–1467.
294. Macfarlane CM, Collier P, Rahbari R, Beck CR, Wagstaff JF, Igoe S, Moran JV, Badge RM. 2013. Transduction-Specific ATLAS Reveals a Cohort of Highly Active L1 Retrotransposons in Human Populations. *Hum Mutat* 34:974–985.
295. Solyom S, Ewing AD, Hancks DC, Takeshima Y, Awano H, Matsuo M, Kazazian HH, Jr. 2012. Pathogenic orphan transduction created by a nonreference LINE-1 retrotransposon. *Hum Mutat* 33:369–371.
296. Tubio JM, Li Y, Ju YS, Martincorena I, Cooke SL, Tojo M, Gundem G, Pipinikas CP, Zamora J, Raine K, Menzies A, Roman-Garcia P, Fullam A, Gerstung M, Shlien A, Tarpey PS, Papaemmanuil E, Knappskog S, Van Loo P, Ramakrishna M, Davies HR, Marshall J, Wedge DC, Teague JW, Butler AP, Nik-Zainal S, Alexandrov L, Behjati S, Yates LR, Bolli N, Mudie L, Hardy C, Martin S, McLaren S, O'Meara S, Anderson E, Maddison M, Gamble S, Foster C, Warren AY, Whitaker H, Brewer D, Eeles R, Cooper C, Neal D, Lynch AG, Visakorpi T, Isaacs WB, van't Veer L, Caldas C, Desmedt C, Sotiriou C, Aparicio S, Foekens JA, Eyfjord JE, Lakhani SR, Thomas G, Myklebost O, Span PN, Borresen-Dale AL, Richardson AL, Van de Vijver M, Vincent-Salomon A, Van den Eynden GG, Flanagan AM, Futreal PA, Janes SM, Bova GS, Stratton MR, McDermott U, Campbell PJ. 2014. Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science* 345:1251343.
297. Hancks DC, Ewing AD, Chen JE, Tokunaga K, Kazazian HH, Jr. 2009. Exon-trapping mediated by the human retrotransposon SVA. *Genome Res* 19:1983–1991.
298. Xing J, Wang H, Belancio VP, Cordaux R, Deininger PL, Batzer MA. 2006. Emergence of primate genes by retrotransposon-mediated sequence transduction. *Proc Natl Acad Sci U S A* 103:17608–17613.
299. Damert A, Raiz J, Horn AV, Lower J, Wang H, Xing J, Batzer MA, Lower R, Schumann GG. 2009. 5'-Transducing SVA retrotransposon groups spread efficiently throughout the human genome. *Genome Res* 19:1992–2008.
300. Mine M, Chen JM, Brivet M, Desguerre I, Marchant D, de Lonlay P, Bernard A, Ferec C, Abitbol M, Ricquier D, Marsac C. 2007. A large genomic deletion in the PDHX gene caused by the retrotranspositional insertion of a full-length LINE-1 element. *Hum Mutat* 28:137–142.
301. Chou HH, Hayakawa T, Diaz S, Krings M, Indriati E, Leakey M, Paabo S, Satta Y, Takahata N, Varki A. 2002. Inactivation of CMP-N-acetylneuraminic acid hydroxylase occurred prior to brain expansion during human evolution. *Proc Natl Acad Sci U S A* 99:11736–11741.
302. Takasu M, Hayashi R, Maruya E, Ota M, Imura K, Kougo K, Kobayashi C, Saji H, Ishikawa Y, Asai T, Tokunaga K. 2007. Deletion of entire HLA-A gene accompanied by an insertion of a retrotransposon. *Tissue Antigens* 70:144–150.
303. Han K, Sen SK, Wang J, Callinan PA, Lee J, Cordaux R, Liang P, Batzer MA. 2005. Genomic rearrangements by LINE-1 insertion-mediated deletion in the human and chimpanzee lineages. *Nucleic Acids Res* 33:4040–4052.
304. Callinan PA, Wang J, Herke SW, Garber RK, Liang P, Batzer MA. 2005. Alu retrotransposition-mediated deletion. *J Mol Biol* 348:791–800.
305. Chen JM, Cooper DN, Ferec C, Kehrer-Sawatzki H, Patrinos GP. 2010. Genomic rearrangements in inherited disease and cancer. *Semin Cancer Biol* 20:222–233.
306. Ade C, Roy-Engel AM, Deininger PL. 2013. Alu elements: an intrinsic source of human genome instability. *Curr Opin Virol* 3:639–645.
307. Deininger PL, Moran JV, Batzer MA, Kazazian HH, Jr. 2003. Mobile elements and mammalian genome evolution. *Curr Opin Genet Dev* 13:651–658.
308. Eickbush TH. 2002. Repair by retrotransposition. *Nat Genet* 31:126–127.
309. Morrish TA, Garcia-Perez JL, Stamato TD, Taccioli GE, Sekiguchi J, Moran JV. 2007. Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres. *Nature* 446:208–212.
310. Kopera HC, Moldovan JB, Morrish TA, Garcia-Perez JL, Moran JV. 2011. Similarities between long interspersed element-1 (LINE-1) reverse transcriptase and telomerase. *Proc Natl Acad Sci U S A* 108:20345–20350.
311. Belfort M, Curcio MJ, Lue NF. 2011. Telomerase and retrotransposons: reverse transcriptases that shaped genomes. *Proc Natl Acad Sci U S A* 108:20304–20310.
312. Eickbush TH. 1997. Telomerase and retrotransposons: which came first? *Science* 277:911–912.
313. Levis RW, Ganesan R, Houtchens K, Tolar LA, Sheen FM. 1993. Transposons in place of telomeric repeats at a *Drosophila* telomere. *Cell* 75:1083–1093.
314. Biessmann H, Valgeirsdottir K, Lofsky A, Chin C, Ginther B, Levis RW, Pardue ML. 1992. HeT-A, a transposable element specifically involved in “healing” broken chromosome ends in *Drosophila melanogaster*. *Mol Cell Biol* 12:3910–3918.
315. Gladyshev EA, Arkhipova IR. 2007. Telomere-associated endonuclease-deficient Penelope-like retroelements in diverse eukaryotes. *Proc Natl Acad Sci U S A* 104:9352–9357.
316. Sen SK, Huang CT, Han K, Batzer MA. 2007. Endonuclease-independent insertion provides an alternative pathway for L1 retrotransposition in the human genome. *Nucleic Acids Res* 35:3741–3751.
317. Morisada N, Rendtorff ND, Nozu K, Morishita T, Miyakawa T, Matsumoto T, Hisano S, Iijima K, Tranebjaerg L, Shirahata A, Matsuo M, Kusuhashi K. 2010. Branchio-oto-renal syndrome caused by partial EYA1 deletion due to LINE-1 insertion. *Pediatr Nephrol* 25:1343–1348.
318. Zhong J, Lambowitz AM. 2003. Group II intron mobility using nascent strands at DNA replication forks to prime reverse transcription. *EMBO J* 22:4555–4565.
319. Yoder JA, Walsh CP, Bestor TH. 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet* 13:335–340.
320. Bestor TH, Bourc'his D. 2004. Transposon silencing and imprint establishment in mammalian germ cells. *Cold Spring Harb Symp Quant Biol* 69:381–387.
321. Ooi SK, O'Donnell AH, Bestor TH. 2009. Mammalian cytosine methylation at a glance. *J Cell Sci* 122:2787–2791.
322. Bourc'his D, Bestor TH. 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* 431:96–99.
323. Soper SF, van der Heijden GW, Hardiman TC, Goodheart M, Martin SL, de Boer P, Bortvin A. 2008. Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. *Dev Cell* 15:285–297.
324. Branco MR, Ficz G, Reik W. 2012. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet* 13:7–13.
325. Siomi MC, Sato K, Pezic D, Aravin AA. 2011. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat Rev Mol Cell Biol* 12:246–258.
326. Aravin AA, Hannon GJ, Brennecke J. 2007. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 318:761–764.
327. Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. 2009. Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* 137:522–535.

328. Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. 2008. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 322:1387–1392.
329. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128:1089–103.
330. Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* 315:1587–1590.
331. Ghildiyal M, Zamore PD. 2009. Small silencing RNAs: an expanding universe. *Nat Rev Genet* 10:94–108.
332. Malone CD, Hannon GJ. 2009. Small RNAs as guardians of the genome. *Cell* 136:656–668.
333. Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ. 2008. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol Cell* 31:785–799.
334. Aravin AA, Bourc'his D. 2008. Small RNA guides for de novo DNA methylation in mammalian germ cells. *Genes Dev* 22:970–975.
335. Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ. 2007. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev Cell* 12:503–514.
336. 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. 2008. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev* 22:908–917.
337. Reuter M, Chuma S, Tanaka T, Franz T, Stark A, Pillai RS. 2009. Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. *Nat Struct Mol Biol* 16:639–646.
338. Yang N, Kazazian HH, Jr. 2006. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat Struct Mol Biol* 13:763–771.
339. Heras SR, Macias S, Plass M, Fernandez N, Cano D, Eyraes E, Garcia-Perez JL, Caceres JF. 2013. The Microprocessor controls the activity of mammalian retrotransposons. *Nat Struct Mol Biol* 20:1173–1181.
340. Ciaudo C, Jay F, Okamoto I, Chen CJ, Sarazin A, Servant N, Barillot E, Heard E, Voignat O. 2013. RNAi-dependent and independent control of LINE1 accumulation and mobility in mouse embryonic stem cells. *PLoS Genet* 9:e1003791.
341. Faulkner GJ. 2013. Retrotransposon silencing during embryogenesis: dicer cuts in LINE. *PLoS Genet* 9:e1003944.
342. Belancio VP, Hedges DJ, Deininger P. 2006. LINE-1 RNA splicing and influences on mammalian gene expression. *Nucleic Acids Res* 34:1512–1521.
343. Belancio VP, Roy-Engel AM, Deininger P. 2008. The impact of multiple splice sites in human L1 elements. *Gene* 411:38–45.
344. Perepelitsa-Belancio V, Deininger P. 2003. RNA truncation by premature polyadenylation attenuates human mobile element activity. *Nat Genet* 35:363–366.
345. Han JS, Szak ST, Boeke JD. 2004. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature* 429:268–274.
346. Peddigari S, Li PW, Rabe JL, Martin SL. 2013. hnRNPL and nucleolin bind LINE-1 RNA and function as host factors to modulate retrotransposition. *Nucleic Acids Res* 41:575–585.
347. Dai L, Taylor MS, O'Donnell KA, Boeke JD. 2012. Poly(A) binding protein C1 is essential for efficient L1 retrotransposition and affects L1 RNP formation. *Mol Cell Biol* 32:4323–4336.
348. Zhang A, Dong B, Doucet AJ, Moldovan JB, Moran JV, Silverman RH. 2014. RNase L restricts the mobility of engineered retrotransposons in cultured human cells. *Nucleic Acids Res* 42:3803–3820.
349. Chiu YL, Greene WC. 2008. The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu Rev Immunol* 26:317–353.
350. Sheehy AM, Gaddis NC, Choi JD, Malim MH. 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418:646–650.
351. Bogerd HP, Wiegand HL, Hulme AE, Garcia-Perez JL, O'Shea KS, Moran JV, Cullen BR. 2006. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc Natl Acad Sci U S A* 103:8780–8785.
352. Chen H, Lilley CE, Yu Q, Lee DV, Chou J, Narvaiza I, Landau NR, Weitzman MD. 2006. APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. *Curr Biol* 16:480–485.
353. Hulme AE, Bogerd HP, Cullen BR, Moran JV. 2007. Selective inhibition of Alu retrotransposition by APOBEC3G. *Gene* 390:199–205.
354. Muckenfuss H, Hamdorf M, Held U, Perkovic M, Lower J, Cichutek K, Flory E, Schumann GG, Munk C. 2006. APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J Biol Chem* 281:22161–22172.
355. Schumann GG. 2007. APOBEC3 proteins: major players in intracellular defence against LINE-1-mediated retrotransposition. *Biochem Soc Trans* 35:637–642.
356. Wissing S, Montano M, Garcia-Perez JL, Moran JV, Greene WC. 2011. Endogenous APOBEC3B restricts LINE-1 retrotransposition in transformed cells and human embryonic stem cells. *J Biol Chem* 286:36427–36437.
357. Horn AV, Klawitter S, Held U, Berger A, Vasudevan AA, Bock A, Hofmann H, Hanschmann KM, Trosmeier JH, Flory E, Jabulowsky RA, Han JS, Lower J, Lower R, Munk C, Schumann GG. 2014. Human LINE-1 restriction by APOBEC3C is deaminase independent and mediated by an ORF1p interaction that affects LINE reverse transcriptase activity. *Nucleic Acids Res* 42:396–416.
358. Richardson SR, Narvaiza I, Planegger RA, Weitzman MD, Moran JV. 2014. APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. *Elife* 3:e02008.
359. Nik-Zainal S, Wedge DC, Alexandrov LB, Petljak M, Butler AP, Bolli N, Davies HR, Knappskog S, Martin S, Papaemmanuil E, Ramakrishna M, Shlien A, Simonic I, Xue Y, Tyler-Smith C, Campbell PJ, Stratton MR. 2014. Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. *Nat Genet* 46:487–491.
360. Helleday T, Eshtad S, Nik-Zainal S. 2014. Mechanisms underlying mutational signatures in human cancers. *Nat Rev Genet* 15:585–598.
361. Crow YJ. 2013. Aicardi–Goutières syndrome. *Handb Clin Neurol* 113:1629–1635.
362. Crow YJ, Rehwinkel J. 2009. Aicardi–Goutières syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity. *Hum Mol Genet* 18:R130–R136.
363. Crow YJ, Leitch A, Hayward BE, Garner A, Parmar R, Griffith E, Ali M, Semple C, Aicardi J, Babul-Hirji R, Baumann C, Baxter P, Bertini E, Chandler KE, Chitayat D, Cau D, Dery C, Fazzi E, Goizet C, King MD, Klepper J, Lacombe D, Lanzi G, Lyall H, Martinez-Frias ML, Mathieu M, McKeown C, Monier A, Oade Y, Quarrell OW, Rittey CD, Rogers RC, Sanchis A, Stephenson JB, Tacke U, Till M, Tolmie JL, Tomlin P, Voit T, Weschke B, Woods CG, Lebon P, Bonthron DT, Ponting CP, Jackson AP. 2006. Mutations in genes encoding ribonuclease H2 subunits cause Aicardi–Goutières syndrome and mimic congenital viral brain infection. *Nat Genet* 38:910–916.
364. Crow YJ, Hayward BE, Parmar R, Robins P, Leitch A, Ali M, Black DN, van Bokhoven H, Brunner HG, Hamel BC, Corry PC, Cowan FM, Frints SG, Klepper J, Livingston JH, Lynch SA, Massey RF, Meritet JF, Michaud JL, Ponsot G, Voit T, Lebon P, Bonthron DT, Jackson AP,

- Barnes DE, Lindahl T. 2006. Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 cause Aicardi-Goutières syndrome at the AGS1 locus. *Nat Genet* 38:917–920.
365. Rice GI, Bond J, Asipu A, Brunette RL, Manfield IW, Carr IM, Fuller JC, Jackson RM, Lamb T, Briggs TA, Ali M, Gornall H, Couthard LR, Aeby A, Attard-Montalto SP, Bertini E, Bodemer C, Brockmann K, Brueton LA, Corry PC, Desguerre I, Fazzi E, Cazorla AG, Gener B, Hamel BC, Heiberg A, Hunter M, van der Knaap MS, Kumar R, Lagae L, Landrieu PG, Lourenco CM, Marom D, McDermott MF, van der Merwe W, Orcesi S, Prendiville JS, Rasmussen M, Shalev SA, Soler DM, Shinawi M, Spiegel R, Tan TY, Vanderver A, Wakeling EL, Wassmer E, Whittaker E, Lebon P, Stetson DB, Bonthron DT, Crow YJ. 2009. Mutations involved in Aicardi-Goutières syndrome implicate SAMHD1 as regulator of the innate immune response. *Nat Genet* 41:829–832.
366. Stetson DB, Ko JS, Heidmann T, Medzhitov R. 2008. Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell* 134:587–598.
367. Zhao K, Du J, Han X, Goodier JL, Li P, Zhou X, Wei W, Evans SL, Li L, Zhang W, Cheung LE, Wang G, Kazazian HH, Jr., Yu XF. 2013. Modulation of LINE-1 and Alu/SVA retrotransposition by Aicardi-Goutières syndrome-related SAMHD1. *Cell Rep* 4:1108–1115.
368. Rice GI, Kasher PR, Forte GM, Mannion NM, Greenwood SM, Szykiewicz M, Dickerson JE, Bhaskar SS, Zampini MA, Briggs TA, Jenkinson EM, Bacino CA, Battini R, Bertini E, Brogan PA, Brueton LA, Carpanelli M, De Laet C, de Lonlay P, del Toro M, Desguerre I, Fazzi E, Garcia-Cazorla A, Heiberg A, Kawaguchi M, Kumar R, Lin JP, Lourenco CM, Male AM, Marques W, Jr., Mignot C, Olivieri I, Orcesi S, Prabhakar P, Rasmussen M, Robinson RA, Rozenberg F, Schmidt JL, Steindl K, Tan TY, van der Merwe WG, Vanderver A, Vassallo G, Wakeling EL, Wassmer E, Whittaker E, Livingston JH, Lebon P, Suzuki T, McLaughlin PJ, Keegan LP, O'Connell MA, Lovell SC, Crow YJ. 2012. Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type I interferon signature. *Nat Genet* 44:1243–1248.
369. Oda H, Nakagawa K, Abe J, Awaya T, Funabiki M, Hijikata A, Nishikomori R, Funatsuka M, Ohshima Y, Sugawara Y, Yasumi T, Kato H, Shirai T, Ohara O, Fujita T, Heike T. 2014. Aicardi-Goutières syndrome is caused by IFIH1 mutations. *Am J Hum Genet* 95:121–125.
370. Rice GI, del Toro Duany Y, Jenkinson EM, Forte GM, Anderson BH, Ariaudo G, Bader-Meunier B, Baildam EM, Battini R, Beresford MW, Casarano M, Chouchane M, Cimaz B, Collins AE, Cordeiro NJ, Dale RC, Davidson JE, De Waele L, Desguerre I, Faivre L, Fazzi E, Isidor B, Lagae L, Latchman AR, Lebon P, Li C, Livingston JH, Lourenco CM, Mancardi MM, Masurel-Paulet A, McInnes IB, Menezes MP, Mignot C, O'Sullivan J, Orcesi S, Picco PP, Riva E, Robinson RA, Rodriguez D, Salvatici E, Scott C, Szybowska M, Tolmie JL, Vanderver A, Vanhulle C, Vieira JP, Webb K, Whitney RN, Williams SG, Wolfe LA, Zuberi SM, Hur S, Crow YJ. 2014. Gain-of-function mutations in IFIH1 cause a spectrum of human disease phenotypes associated with upregulated type I interferon signaling. *Nat Genet* 46:503–509.
371. Zheng YH, Jeang KT, Tokunaga K. 2012. Host restriction factors in retroviral infection: promises in virus–host interaction. *Retrovirology* 9:112.
372. Arjan-Odedra S, Swanson CM, Sherer NM, Wolinsky SM, Malim MH. 2012. Endogenous MOV10 inhibits the retrotransposition of endogenous retroelements but not the replication of exogenous retroviruses. *Retrovirology* 9:53.
373. Goodier JL, Cheung LE, Kazazian HH, Jr. 2012. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. *PLoS Genet* 8:e1002941.
374. Li X, Zhang J, Jia R, Cheng V, Xu X, Qiao W, Guo F, Liang C, Cen S. 2013. The MOV10 helicase inhibits LINE-1 mobility. *J Biol Chem* 288:21148–21160.
375. Gregersen LH, Schueler M, Munschauer M, Mastrobuoni G, Chen W, Kempa S, Dieterich C, Landthaler M. 2014. MOV10 is a 5' to 3' RNA helicase contributing to UPF1 mRNA target degradation by translocation along 3' UTRs. *Mol Cell* 54:573–585.
376. Frost RJ, Hamra FK, Richardson JA, Qi X, Bassel-Duby R, Olson EN. 2010. MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. *Proc Natl Acad Sci U S A* 107:11847–11852.
377. Zheng K, Xiol J, Reuter M, Eckardt S, Leu NA, McLaughlin KJ, Stark A, Sachidanandam R, Pillai RS, Wang PJ. 2010. Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway. *Proc Natl Acad Sci U S A* 107:11841–11846.
378. Cordaux R, Batzer MA. 2009. The impact of retrotransposons on human genome evolution. *Nat Rev Genet* 10:691–703.
379. Branciforte D, Martin SL. 1994. Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. *Mol Cell Biol* 14:2584–2592.
380. Trelogan SA, Martin SL. 1995. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. *Proc Natl Acad Sci U S A* 92:1520–1524.
381. Georgiou I, Noutsopoulos D, Dimitriadou E, Markopoulos G, Aperi A, Lazaros L, Vaxevanoglou T, Pantos K, Syrrou M, Tzavaras T. 2009. Retrotransposon RNA expression and evidence for retrotransposition events in human oocytes. *Hum Mol Genet* 18:1221–1228.
382. Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, Gerton GL, Kazazian HH, Jr. 2002. A mouse model of human L1 retrotransposition. *Nat Genet* 32:655–660.
383. Athanikar JN, Morrish TA, Moran JV. 2002. Of man in mice. *Nat Genet* 32:562–563.
384. Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, Kazazian HH, Jr. 2009. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes Dev* 23:1303–1312.
385. Malki S, van der Heijden GW, O'Donnell KA, Martin SL, Bortvin A. 2014. A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. *Dev Cell* 29:521–533.
386. Freeman P, Macfarlane C, Collier P, Jeffreys AJ, Badge RM. 2011. L1 hybridization enrichment: a method for directly accessing de novo L1 insertions in the human germline. *Hum Mutat* 32:978–988.
387. van den Hurk JA, Meij IC, Seleme MC, Kano H, Nikopoulos K, Hoefsloot LH, Sistermans EA, de Wijs IJ, Mukhopadhyay A, Plomp AS, de Jong PT, Kazazian HH, Cremers FP. 2007. L1 retrotransposition can occur early in human embryonic development. *Hum Mol Genet* 16:1587–1592.
388. Garcia-Perez JL, Marchetto MC, Muotri AR, Coufal NG, Gage FH, O'Shea KS, Moran JV. 2007. LINE-1 retrotransposition in human embryonic stem cells. *Hum Mol Genet* 16:1569–1577.
389. Garcia-Perez JL, Morell M, Scheys JO, Kulpa DA, Morell S, Carter CC, Hammer GD, Collins KL, O'Shea KS, Menendez P, Moran JV. 2010. Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. *Nature* 466:769–773.
390. Wissing S, Munoz-Lopez M, Macia A, Yang Z, Montano M, Collins W, Garcia-Perez JL, Moran JV, Greene WC. 2012. Reprogramming somatic cells into iPS cells activates LINE-1 retroelement mobility. *Hum Mol Genet* 21:208–218.
391. Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. 2005. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* 435:903–910.
392. Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M, O'Shea KS, Moran JV, Gage FH. 2009. L1 retrotransposition in human neural progenitor cells. *Nature* 460:1127–1131.
393. Muotri AR, Marchetto MC, Coufal NG, Oefner R, Yeo G, Nakashima K, Gage FH. 2010. L1 retrotransposition in neurons is modulated by MeCP2. *Nature* 468:443–446.
394. Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, De Sapio F, Brennan PM, Rizzu P, Smith S, Fell M, Talbot RT, Gustincich S, Freeman TC, Mattick JS, Hume DA, Heutink P, Carninci P, Jeddloh JA,

- Faulkner GJ. 2011. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* 479:534–537.
395. Evrony GD, Cai X, Lee E, Hills LB, Elhosary PC, Lehmann HS, Parker JJ, Atabay KD, Gilmore EC, Poduri A, Park PJ, Walsh CA. 2012. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* 151:483–496.
396. Li W, Prazak L, Chatterjee N, Gruninger S, Krug L, Theodorou D, Dubnau J. 2013. Activation of transposable elements during aging and neuronal decline in *Drosophila*. *Nat Neurosci* 16:529–531.
397. Perrat PN, DasGupta S, Wang J, Theurkauf W, Weng Z, Rosbash M, Waddell S. 2013. Transposition-driven genomic heterogeneity in the *Drosophila* brain. *Science* 340:91–95.
398. Miki Y, Nishisho I, Horii A, Miyoshi Y, Utsunomiya J, Kinzler KW, Vogelstein B, Nakamura Y. 1992. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res* 52:643–645.
399. Solyom S, Ewing AD, Rahrmann EP, Doucet T, Nelson HH, Burns MB, Harris RS, Sigmon DF, Casella A, Erlanger B, Wheelan S, Upton KR, Shukla R, Faulkner GJ, Largaespada DA, Kazazian HH, Jr. 2012. Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res* 22:2328–2338.
400. Lee E, Iskow R, Yang L, Gokcumen O, Haseley P, Luquette LJ, 3rd, Lohr JG, Harris CC, Ding L, Wilson RK, Wheeler DA, Gibbs RA, Kucherlapati R, Lee C, Kharchenko PV, Park PJ. 2012. Landscape of somatic retrotransposition in human cancers. *Science* 337:967–971.
401. Helman E, Lawrence MS, Stewart C, Sougnez C, Getz G, Meyerson M. 2014. Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. *Genome Res* 24:1053–1063.
402. Shukla R, Upton KR, Munoz-Lopez M, Gerhardt DJ, Fisher ME, Nguyen T, Brennan PM, Baillie JK, Collino A, Ghisletti S, Sinha S, Iannelli F, Radaelli E, Dos Santos A, Rapoud D, Guettier C, Samuel D, Natoli G, Carninci P, Ciccarelli FD, Garcia-Perez JL, Faivre J, Faulkner GJ. 2013. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. *Cell* 153:101–111.
403. Rodic N, Burns KH. 2013. Long interspersed element-1 (LINE-1): passenger or driver in human neoplasms? *PLoS Genet* 9:e1003402.
404. Bulut-Karslioglu A, De La Rosa-Velazquez IA, Ramirez F, Barenboim M, Onishi-Seebacher M, Arand J, Galan C, Winter GE, Engist B, Gerle B, O'Sullivan RJ, Martens JH, Walter J, Manke T, Lachner M, Jenuwein T. 2014. Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. *Mol Cell* 55:277–290.
405. Chow JC, Ciaudo C, Fazzari MJ, Mise N, Servant N, Glass JL, Attreed M, Avner P, Wutz A, Barillot E, Grealley JM, Voinnet O, Heard E. 2010. LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. *Cell* 141:956–969.
406. Bailey JA, Carrel L, Chakravarti A, Eichler EE. 2000. Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. *Proc Natl Acad Sci U S A* 97:6634–6639.
407. Lyon MF. 1998. X-chromosome inactivation: a repeat hypothesis. *Cytogenet Cell Genet* 80:133–137.
408. Richardson SR. 2013. A mechanistic examination of APOBEC3-mediated LINE-1 inhibition, Doctoral Dissertation Thesis, University of Michigan, Ann Arbor.
409. Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM. 1998. Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 2:437–445.
410. Wallace N, Wagstaff BJ, Deininger PL, Roy-Engel AM. 2008. LINE-1 ORF1 protein enhances Alu SINE retrotransposition. *Gene* 419:1–6.