L1 Retrotransposons and Somatic Mosaicism in the Brain

Sandra R. Richardson,[1] Santiago Morell,[1] and Geoffrey J. Faulkner[1,2]

[1]Mater Research Institute, The University of Queensland, Translational Research Institute, Woolloongabba QLD 4102, Australia; email: faulknergj@gmail.com

[2]School of Biomedical Sciences, The University of Queensland, Brisbane QLD 4072, Australia

Annu. Rev. Genet. 2014. 48:1–27

First published online as a Review in Advance on July 14, 2014

The *Annual Review of Genetics* is online at genet.annualreviews.org

This article's doi: 10.1146/annurev-genet-120213-092412

Keywords

LINE-1, transposon, retrotransposition, neuron, neurogenesis

Abstract

Long interspersed element 1 (LINE-1 or L1) retrotransposons have generated one-third of the human genome, and their ongoing mobility is a source of inter- and intraindividual genetic diversity. Although retrotransposition in metazoans has long been considered a germline phenomenon, recent experiments using cultured cells, animal models, and human tissues have revealed extensive L1 mobilization in rodent and human neurons, as well as mobile element activity in the *Drosophila* brain. In this review, we evaluate the available evidence for L1 retrotransposition in the brain and discuss mechanisms that may regulate neuronal retrotransposition in vivo. We compare experimental strategies used to map de novo somatic retrotransposition events and present the optimal criteria to identify a somatic L1 insertion. Finally, we discuss the unresolved impact of L1-mediated somatic mosaicism upon normal neurobiology, as well as its potential to drive neurological disease.

INTRODUCTION

Transposable elements are found in virtually all eukaryotic life. Their activity rearranges and adds to the genetic instructions encoded by DNA and, as such, is a key component of genotypic variation and evolutionary selection. That transposons, retrotransposons, and other mobile DNA contribute to phenotype is clear; striking examples of variegation in plants and animals, such as morning glory flower pigmentation (56), agouti mouse coat color (94), and, in arguably the prototypical case, maize kernel mosaicism explained by Barbara McClintock (85), are all manifestations of transposable element activity (reviewed in 120). As established by Kazazian et al. (62) in 1988, retrotransposition continues to occur in humans and, as shown repeatedly then and since, is a notable source of mutagenesis leading to disease (10, 20, 34, 53). Thus, transposable elements are an integral part of genome evolution in humans and other species, with the capacity to fundamentally impact a wide range of biological phenomena.

The autonomous retrotransposon long interspersed element 1 (LINE-1 or L1) is arguably the most impactful—and still active—human transposable element. First, L1 sequences account for approximately 17% of our DNA. Second, the L1-encoded enzymatic machinery mobilizes nonautonomous retroelements, such as *Alu* and SINE-VNTR-*Alu* (SVA), and generates processed pseudogenes (31). L1-mediated retrotransposition has therefore resulted in the accumulation of at least one-third of the human genome (26, 69) and continues to shape its landscape, as witnessed by the estimated 400 million polymorphic retrotransposon insertions in the global human population (reviewed in 8, 34, 110).

L1 Retrotransposition Mechanism

A retrotransposition-competent human L1 is approximately 6 kb in length (28, 115) (**Figure 1*a***). The L1 5′ end comprises a 5′ untranslated region (UTR) that harbors an internal promoter and an antisense promoter of unknown function (122, 125). The L1 internal promoter contains *cis*-acting binding sites for several transcription factors, including YY1 and RUNX3 as well as SOX family transcription factors (5, 9, 68, 89, 129, 136). L1 encodes two open reading frames (28, 115), ORF1 and ORF2, the protein products of which are ORF1p, a ~40-kDa nucleic acid–binding protein (55, 63, 81), and ORF2p, a ~150-kDa protein with demonstrated endonuclease (EN) and reverse transcriptase (RT) activities (29, 30, 36, 82, 128). Both proteins are required for L1 retrotransposition (93). The 3′ end of the L1 element consists of a 3′ UTR, followed by a poly-A tail thought to facilitate efficient L1 translation and reverse transcription (28, 115). New L1 insertions are typically flanked by variable length target-site duplications (TSDs), which are a structural hallmark of the L1 integration process (48) and distinguish retrotransposition from other types of genomic rearrangement (e.g., translocation).

The process of retrotransposition begins with transcription of a full-length L1 from its internal promoter (**Figure 1*b***). The resulting L1 mRNA is exported to the cytoplasm, where it is translated by an unconventional termination-reinitiation mechanism (1, 27). Multiple ORF1p molecules and potentially as few as one ORF2p molecule bind back to their encoding L1 mRNA in a phenomenon known as *cis* preference, giving rise to the L1 ribonucleoprotein particle (RNP), a hypothesized retrotransposition intermediate (31, 55, 66, 67, 80, 133). The L1 RNP then enters the nucleus by a mechanism that is not completely understood but that can take place independently of cell division (65).

In the nucleus, the L1-encoded EN activity creates a single-strand nick in genomic DNA, with a loose preference for 5′-TTTT/A-3′ motifs (60). The nick liberates a free 3′ hydroxyl residue, which is in turn used as a primer from which the L1 RT initiates reverse transcription of its associated

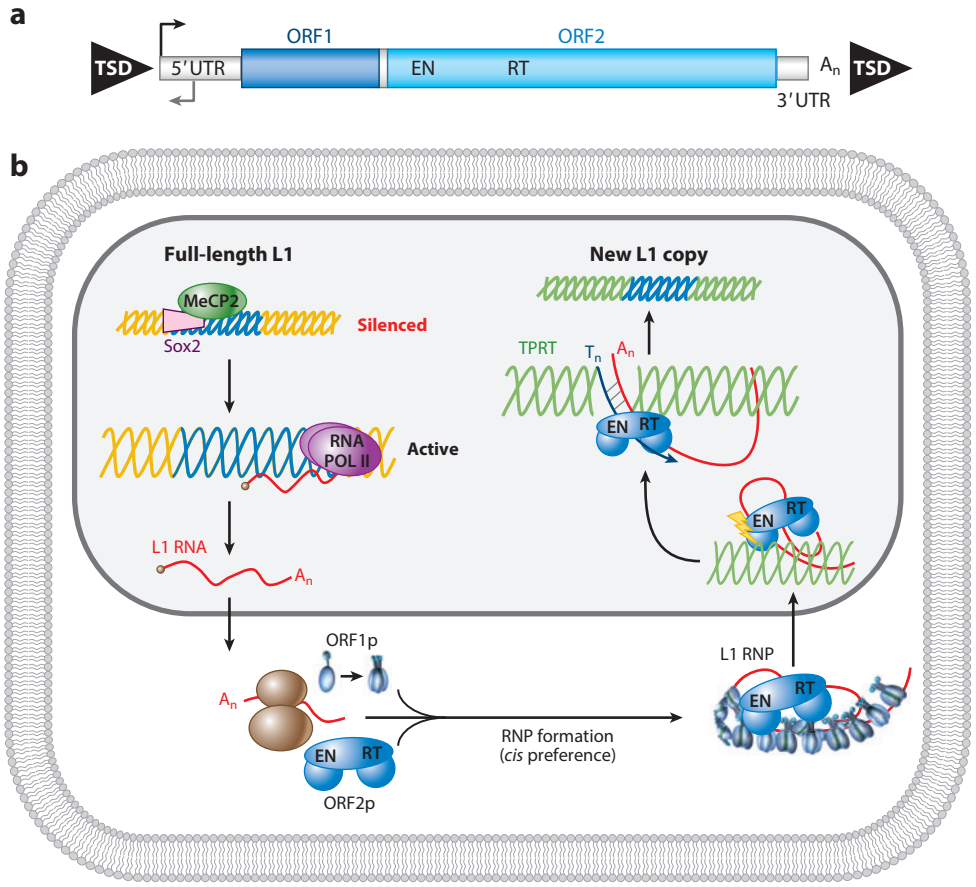


Figure 1

(*a*) The structure of a retrotransposition-competent L1 (long interspersed element 1). The L1 5′ UTR (untranslated region) contains an internal promoter (*black arrow*). L1 harbors two open reading frames: ORF1 (*dark blue rectangle*) and ORF2 (*light blue rectangle*); ORF2 encodes L1 endonuclease (EN) and reverse transcriptase (RT) activities. The L1 sequence terminates in a 3′ UTR and a poly-A tail (A_n). L1s in the genome are frequently flanked by target-site duplications (TSDs; *black triangles*). (*b*) L1 retrotransposition mechanism. L1 expression in neuronal cell types is dynamically regulated by factors such as MeCP2 and Sox2. In a round of retrotransposition, the L1 RNA is transcribed by RNA polymerase II and then translated, giving rise to multiple copies of ORF1p and as few as one copy of ORF2p. Both proteins associate with their encoding RNA (*cis* preference) to assemble the L1 ribonucleoprotein particle (RNP). The L1 RNP enters the nucleus, where the L1 endonuclease activity nicks the genomic DNA at the consensus 5′-TTTT/A-3′ to liberate a 3′ hydroxyl residue from which the L1 reverse transcriptase initiates reverse transcription of its associated L1 mRNA. This process, termed target-site-primed reverse transcription (TPRT), generates a new, frequently 5′ truncated L1 insertion.

mRNA (21, 22, 36). In vitro studies have elucidated a complex set of rules governing target-site selection due to L1 RT preference for genomic poly-T tracts and perfect terminal complementarity for L1 RNA-genomic DNA complexes (91). This process, known as target-site-primed reverse transcription (TPRT), was first elucidated by biochemical studies of the *Bombyx mori* non-LTR retrotransposon R2 (77). Second-strand target DNA cleavage usually takes place some distance downstream of first-strand cleavage, giving rise to the aforementioned TSDs. Second-strand

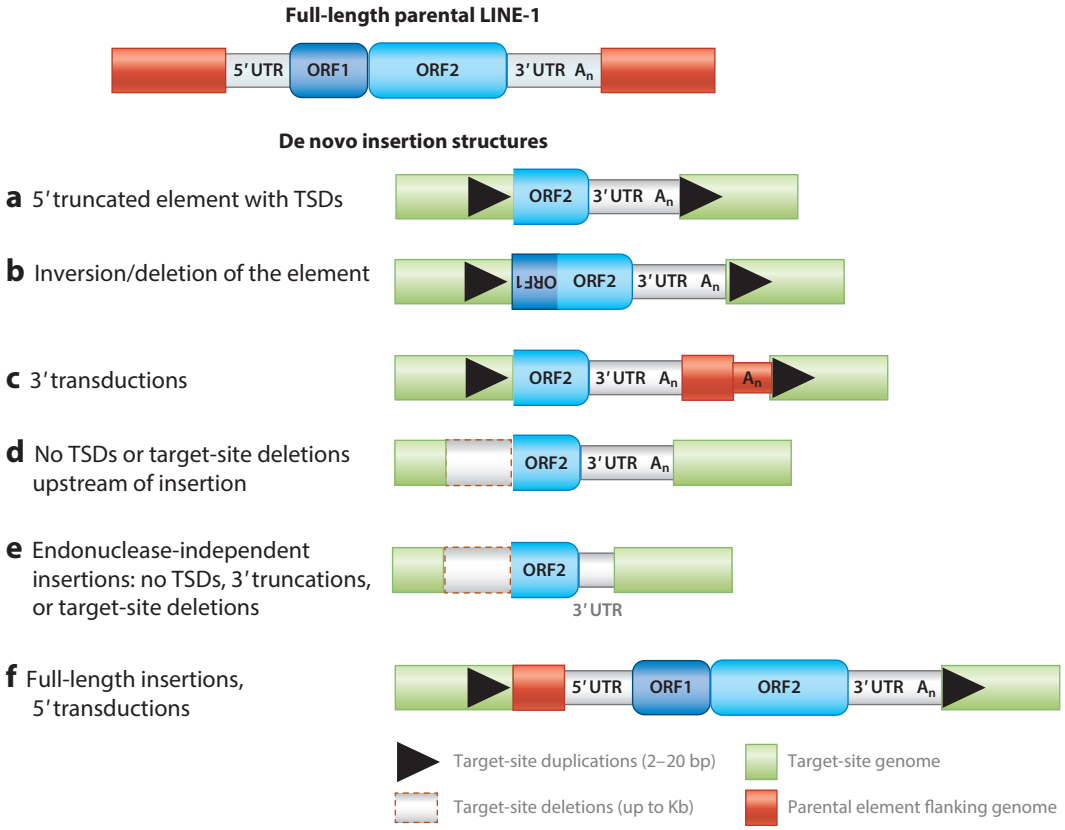


Figure 2

Features of L1 (long interspersed element 1) retrotransposition events. Above: a full-length parental L1. (*a*) A typical L1 insertion is 5′ truncated, ends in a poly-A tail (A_n), and is flanked by target-site duplications (TSDs; *black triangles*). (*b*) L1 insertions frequently contain internal inversions and deletions of the retroelement sequence. (*c*) Bypass of the L1 polyadenylation signal in favor of a strong genomic polyadenylation signal can lead to 3′ transductions of genomic DNA from the donor L1 locus. (*d*) The position of second-strand cleavage can lead to insertions lacking TSDs or small deletions of target-site DNA. (*e*) Endonuclease-independent insertions lack TSDs, are frequently 3′ truncated, and are associated with deletion of target-site DNA. (*f*) Transcription of a parental element from an upstream genomic promoter can lead to full-length insertions with 5′ transductions of the genomic sequence. Abbreviation: UTR, untranslated region.

cleavage can also occur directly opposite the first-strand cleavage site, leading to blunt insertions lacking TSDs, or upstream of the first-strand cleavage site, leading to small deletions of target-site DNA (40, 41, 45) (**Figure 2**). The subsequent steps of insertion formation, including second-strand DNA synthesis and integration, are not completely understood. The majority of new L1 insertions are 5′ truncated (48, 69) and often contain internal rearrangements, such as inversions and deletions of the L1 sequence (40, 41) (**Figure 2**).

Host factors almost certainly play roles in L1 retrotransposition by interacting with the L1 mRNA-encoded proteins. Recent work by Taylor et al. (128) identified 37 proteins that interact with the L1 RNP, including the polymerase-δ-associated sliding clamp PCNA (proliferating cell nuclear antigen) and the nonsense-mediated decay factor UPF1. PCNA interacts with L1 ORF2p via a PIP-box motif and is proposed to influence L1 integration during or immediately after TPRT

(128). In addition, Goodier et al. (44) identified 96 L1 ORF1p-interacting proteins, 19 of which could restrict L1 retrotransposition in a cultured cell assay by greater than 60%.

Previous studies have demonstrated that cellular factors, including members of the APOBEC3 family of cytidine deaminases (13, 14, 18, 64, 97, 103, 113, 123), the nuclease Trex1 (124), the antiretroviral factor SAMHD1 (138), and the Mov10 putative helicase (4, 43, 75), can restrict retrotransposition in cultured cells and may play a role in regulating L1 activity in vivo. RNA-based mechanisms have also been implicated in L1 regulation. Heras et al. (54) demonstrated that the miRNA biogenesis factor Microprocessor/Drosha-DGCR8 specifically binds L1, *Alu*, and SVA retroelement RNAs, and can cleave L1 RNA in vitro. Indeed, experiments in cultured cells revealed that Microprocessor negatively regulates L1 and *Alu* retrotransposition. In addition, Ciaudo et al. (19) demonstrated a role for both Dicer-dependent and Ago2-dependent RNAi in L1 regulation in mouse embryonic stem cells. Along with epigenetic silencing, these mechanisms defend a host genome from the likely deleterious consequences of unrestrained retrotransposition (120).

Under certain conditions, L1 is capable of generating new insertions by a noncanonical, EN-independent (ENi) pathway (23, 95, 96). ENi retrotransposition has been demonstrated in Chinese hamster ovary cells deficient in both p53 and components of the nonhomologous end joining pathway of DNA repair. In these cells, EN-deficient L1s can retrotranspose efficiently, presumably by exploiting pre-existing genomic DNA lesions resulting from delayed or impaired DNA repair. ENi retrotransposition events are often distinguished by a lack of TSDs as well as by L1 3′ truncation and deletion of flanking genomic DNA (96) (**Figure 2**). Notably, the human genome reference sequence incorporates 21 L1 insertions bearing structural features consistent with ENi retrotransposition, indicating that ENi retrotransposition can occur in the human germline (116). Thus, although in wild-type cells the vast majority of L1 retrotransposition occurs via TPRT, exceptional cases of ENi mobilization do occur, especially in cellular environments associated with elevated DNA damage.

Consequences of Retrotransposition for the Host Genome

In addition to duplication and deletion of target-site DNA, L1 retrotransposition is frequently associated with various genomic alterations, including addition of nontemplated nucleotides and, occasionally, large-scale rearrangements, such as chromosomal translocations (40, 41, 90, 126, 127). In 10–20% of cases, L1 retrotransposition is associated with 3′ transduction, in which transcription of the donor element bypasses the canonical L1 polyadenylation signal and uses a genomic signal some distance downstream, causing the non-L1 sequence to be retrotransposed to a new genomic location (47, 92, 109) (**Figure 2**). Furthermore, L1 retrotransposition proximal to and within coding regions can disrupt gene function and regulation. Exonic retrotransposition events can act as insertional mutagens (62); additionally, insertions within introns can induce missplicing or premature polyadenylation due to cryptic splice and polyadenylation signals within the AT-rich L1 sequence (11, 107). Intronic insertions can also impact RNA polymerase processivity through host genes, which has led to the hypothesis that L1 can act as a molecular rheostat to effect subtle changes in gene expression levels (52) and even engage in gene breaking (134). The L1 antisense promoter and other transcription initiation sites in the L1 3′ end can generate transcripts with the potential to impact regulation of adjacent genes (35, 83, 104, 122, 135). A recent study has implicated L1-derived stable nuclear RNA in regulating chromatin state, suggesting an expanded impact of L1 activity on global gene expression (50). L1 insertions are also subject to epigenetic regulation; in some cases, e.g., in PA-1 embryonal carcinoma cells, epigenetic marks may be targeted specifically to nascent L1 insertions during TPRT (38). Epigenetic silencing of L1 insertions may impact the expression of nearby genes if chromatin modifications spread from the L1

sequence into surrounding DNA, as seen for LTR retrotransposons (111). L1 retrotransposition can therefore impact the host genome, epigenome, and transcriptome via numerous routes, any of which may be sufficient to subtly or grossly alter organismal phenotype.

Retrotransposition in vivo: Where and When?

L1 has historically been viewed as a molecular parasite that must be transmitted to subsequent generations to ensure its propagation. As such, the host germline has long been considered as the primary milieu for retrotransposition (reviewed in 73). Recent studies using tissue culture systems, animal models, and human samples have reinforced this view, revealing that L1 retrotransposition

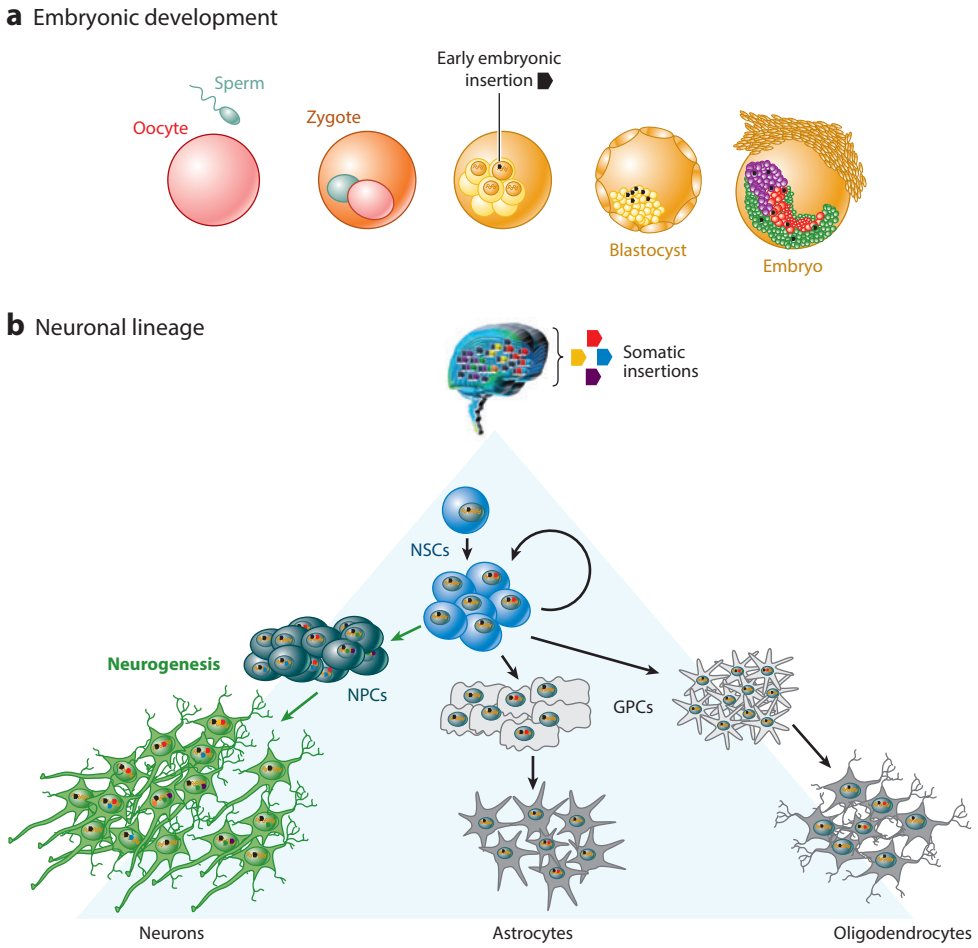


Figure 3

L1 (long interspersed element 1) activity during development. (*a*) Embryogenesis. L1 retrotransposition in the early embryo can generate somatic mosaicism (*black pentagons*) and can contribute to the germ lineage. (*b*) Neuronal lineage. L1 is active in both neural stem cells (NSCs) (*blue*) and neural precursor cells (NPCs) (*blue-green*), thus playing a role of potential importance in genetic mosaicism and perhaps neuronal plasticity in mature neurons (*bright green*). Glial precursor cells (GPCs) support markedly less L1 activity than NPCs. Each colored pentagon represents a different somatic insertion.

can occur in the mammalian germline and in the early embryo prior to germline specification (15, 37, 61, 105, 131, 132) (**Figure 3*a***). By contrast, L1 sequences are normally transcriptionally repressed in somatic cells by DNA methylation, with these patterns established in the primordial germline (12). However, in 1992 Miki et al. (88) linked a case of colorectal cancer to a de novo somatic retrotransposition event in the archetypal tumor suppressor gene APC, implicating a likely somatic retrotransposition as an agent of oncogenesis and also demonstrating incomplete epigenetic suppression of retrotransposition in somatic cells. Furthermore, recent studies have uncovered L1-mediated retrotransposition in a variety of tumor types, including lung, liver, colorectal, prostate, and ovarian cancers (57, 71, 118, 121). Thus, deregulated L1 retrotransposition in the soma has the potential to alter cellular phenotype, as evidenced by overt consequences such as tumorigenesis. More subtle, as yet undescribed effects of somatic retrotransposition are also likely to exist.

One of the most intriguing recent findings in L1 biology is that elevated levels of retrotransposition can occur in normal cells in the mammalian brain. In this review, we present the seminal works leading to the discovery of neuronal retrotransposition, and discuss mechanisms by which neuronal retrotransposition is regulated. We then present criteria that should be met in order to classify a potential somatic insertion as a bona fide retrotransposition event, and examine recent reports that have exploited high-throughput sequencing strategies to map and characterize such insertions in neuronal cells. Finally, we highlight important unanswered questions regarding the timing, frequency, impact on human disease, and ultimately the normal physiological role, if any, of neuronal retrotransposition.

L1 RETROTRANSPOSITION IN NEURONS

The first evidence that L1 retrotransposition occurs in the mammalian brain came from studies employing an engineered L1-EGFP (enhanced green fluorescent protein) reporter transgene in cultured cells and animal models. Neural progenitor cells (NPCs), which can be derived in vitro from several brain regions, including the hippocampus and subventricular zone, are multipotent cells capable of giving rise to diverse cell types of the neuronal lineage (87) (**Figure 3*b***). In 2005, Muotri et al. (99) reported the unexpected observation that L1 mRNAs were abundant in NPCs derived in vitro from adult hippocampus neural stem cells. L1 expression was further demonstrated to be repressed in neural stem cells by the transcription factor Sox2. Intriguingly, a decrease in Sox2 expression during lineage commitment correlated with derepression of the L1 promoter and increased L1 transcription in NPCs. The L1-EGFP reporter readily mobilized in NPCs in vitro, and the resulting events sometimes occurred within genes and had the capacity to alter target gene expression. Transgenic mice harboring a human L1-EGFP reporter exhibited neuronal retrotransposition events consistent with L1 mobilization during both embryonic and adult neurogenesis. These observations led to the hypothesis that cells of the neuronal lineage can accommodate L1 retrotransposition and that L1 activity may contribute to genomic and functional diversity among individual neurons (119).

In 2009, Coufal et al. (24) extended the above rodent studies to humans, demonstrating that the L1-EGFP reporter can retrotranspose in NPCs isolated from the human fetal brain and in NPCs derived from human embryonic stem cells (hESCs). Endogenous L1 copies were shown to be hypomethylated in the fetal brain relative to skin, suggesting that endogenous L1s could be transcriptionally active in the human brain and that this process was regulated by Sox2 and MeCP2. To estimate endogenous L1 copy-number variation (CNV) in human tissue, Coufal et al. (24) developed a multiplex TaqMan quantitative polymerase chain reaction (qPCR) assay and applied this approach to a wide range of human organs and brain subregions from several postmortem

donors. This assay suggested that, on average, the adult hippocampus contained approximately 80 more L1 ORF2 copies per cell than heart or liver, with substantial variability between individuals. The hippocampus also exhibited elevated L1 CNV compared with other brain regions, a result explained by the presence of the subgranular zone, a major neurogenic niche, in the hippocampus. Although lacking sequence-based characterization of endogenous retrotransposition events from the brain, these L1 CNV data provided tantalizing evidence that L1 mobilization caused extensive somatic mosaicism in the human brain.

These experiments were a fundamental advance beyond the results of Muotri et al. (99) because (*a*) active retrotransposons are far more common in the mouse genome than in the human genome (>3,000 versus ~100 RC-L1s, respectively) (7, 16, 46, 102) and (*b*) the L1-EGFP reporter may not have accurately recapitulated endogenous L1 activation rates because of differential epigenetic suppression of an engineered human L1 transgene in mouse cells. Observation of L1 CNV in human tissues therefore excluded the conclusion that somatic L1 retrotransposition was an artifact of the L1-EGFP system, or a phenomenon restricted to rodents, and demonstrated that L1 was likely active across a broad spectrum of mammalian neurons.

L1 Regulation in the Neuronal Lineage

The finding that neuronal cell types are permissive for L1 retrotransposition raises the question of how L1 circumvents suppression in these cells. Muotri et al. (99) found that Sox2 interacts with the L1 promoter and represses L1 expression in rat neural stem cells (NSCs) and upon differentiation into NPCs, a decrease in Sox2 levels permits L1 transcriptional activation and presumably retrotransposition. Subsequent work by Coufal et al. (24) in human fetal NSCs and hESC-derived NPCs corroborated this result, indicating that L1 regulation by Sox2 is conserved in rodents and humans. Kuwabara et al. (68) found that the L1 promoter, as well as the L1 ORF2 sequence, contains overlapping Sox2 and T-cell factor/lymphoid enhancer factor (TCF/LEF) binding sites (Sox/LEF). Sox2, a negative regulator of neuronal differentiation, was demonstrated to repress promoter activity from these sites, whereas Wnt3a and β-catenin signaling increased L1 promoter activity. Moreover, the transcription factor NeuroD1, which promotes neuronal differentiation in adult hippocampal neural progenitors, is also regulated by Sox2 and Wnt/β-catenin signaling through Sox/LEF sites in its promoter (68). These results are consistent with derepression of L1 transcription concurrent with commitment of NPCs to the neuronal lineage and, strikingly, Sox2 regulation of the core neurogenesis pathway incorporating NeuroD1 is inseparable from Sox2 regulation of L1. If we also consider the expression and function of Sox2 in male gametes, it is plausible that establishment of the Sox2 regulatory program acts to limit L1 mobilization in the germline while not inhibiting retrotransposition during neurogenesis. Whether the latter property of Sox2 affects evolutionary fitness is unknown.

DNA methylation is another critical component of L1 repression in somatic cells. The X-linked DNA methyl-binding protein MeCP2 has been demonstrated to associate with the L1 promoter and repress L1 expression in cultured cells (137). Coufal et al. (24) investigated MeCP2-mediated L1 regulation in the brain and found higher levels of MeCP2 association with the L1 promoter in NSCs than in hESC-derived NPCs, suggesting a role for MeCP2 in modulating L1 activity during neuronal development. Indeed, using the L1-EGFP reporter, Muotri et al. (100) demonstrated that MeCP2 represses L1 promoter activity in a methylation-dependent manner. Notably, L1 regulation was specific to MeCP2, as perturbation of the DNA methyl-binding protein MDB1 did not affect L1 expression. Studies of MeCP2 knockout mice harboring the L1-EGFP reporter revealed an elevated rate of neuronal L1 retrotransposition in vivo in the absence of MeCP2. Endogenous L1s also appeared to undergo elevated rates of retrotransposition in the absence of

MeCP2, as a single-cell genomic qPCR assay revealed an increase in L1 copy number in MeCP2 knockout mouse neuroepithelial cells, but not fibroblasts, compared with control animals.

Germline and de novo MeCP2 mutations in humans cause Rett Syndrome (RTT), a condition characterized by profound neurodevelopmental abnormality. The disease affects ~1/10,000 females but is rarely seen in males, as mutations in the X-linked MeCP2 are hemizygous lethal (3). Given the importance of MeCP2 in L1 suppression, L1 may be transcriptionally more active in RTT neurons, and thus RTT neurons may harbor a far higher somatic L1 mobilization rate than seen for healthy controls. How increased L1 activity is related to the neurological symptoms of RTT is unknown, but it is possible that elevated levels of L1 transcription could interfere with normal cellular processes or that an increased rate of retrotransposition events into neuronally expressed genes could have deleterious effects on neuronal function.

To study the relationship between MeCP2 deficiency, RTT, and L1 mobilization, Muotri et al. (100) derived NPCs from induced pluripotent stem cells (iPSCs) previously reprogrammed from RTT patient fibroblasts. They found that MeCP2 mutant NPCs accommodated significantly more retrotransposition of the L1-EGFP reporter than NPCs derived from control iPSCs. Furthermore, qPCR analysis showed a significantly higher L1 copy number in postmortem human brain tissue from RTT patients compared with controls. These experiments conclusively indicated that MeCP2 regulates L1 expression and retrotransposition activity in the mammalian brain. However, as for the L1 CNV data reported by Coufal et al. (24), these results must ultimately be corroborated by the genomic mapping of endogenous L1 insertions to be certain that L1 CNV is associated with L1 copies integrated into the genome and not with the accumulation of extrachromosomal L1 DNA via a largely uncharacterized mechanism (51, 59, 124). It is also important to note that, despite obvious differences in how well neurological phenotypes can be assessed in humans and animals, conditional restoration of MeCP2 function in MeCP2 mutant mice appears to ameliorate neurological dysfunction (49). Thus, although the role of MeCP2 in L1 regulation is clear, it remains unknown whether elevated L1 activity contributes to RTT etiology.

Regulation of L1 retrotransposition in neuronal cell types may not be limited to transcriptional and epigenetic control. Ataxia telangiectasia mutated (ATM) is a serine/threonine kinase that functions as a sensor of DNA damage (117). As its name reflects, ATM is mutated in the autosomal recessive disorder ataxia telangiectasia, which is characterized by progressive neuronal degeneration, variable immunodeficiency, ocular telangiectasias, and cancer susceptibility (2). Coufal et al. (23) demonstrated in 2011 that a human L1-EGFP can retrotranspose with increased efficiency in the brains of ATM-deficient mice as well as with human NPCs derived from hESCs in which ATM expression had been knocked down by RNAi. Furthermore, endogenous L1 CNV detection by qPCR revealed higher levels of the L1 ORF2 sequence in hippocampal samples from ataxia telangiectasia patients compared with hippocampal samples from normal matched controls. Experiments in non-neuronal cultured cell types (HeLa and HCT116 cells) revealed that L1 may generate more or possibly longer insertions in the absence of ATM, consistent with ATM recognition of the L1 TPRT intermediate as DNA damage, and resultant abrogation or truncation of the nascent L1 insertion. Thus, in normal cells, ATM is predicted to limit L1 retrotransposition, whereas in ataxia telangiectasia patients, loss of ATM function may allow elevated rates of retrotransposition or longer L1 insertions. Mapping of increased retrotransposition events using sequence-based approaches from the brains of ataxia telangiectasia patients relative to controls would provide definitive evidence for a role of ATM in regulating neuronal retrotransposition. Furthermore, mapping and characterization of retrotransposition events in ATM-deficient neurons may provide clues about the relationship, if there is one, between increased L1 retrotransposition and the progressive neurodegeneration observed in ataxia telangiectasia patients.

The studies discussed above have begun to uncover the mechanisms responsible for L1 regulation in the mammalian brain. It is very likely that additional host factors affect neuronal retrotransposition. Considering that L1 is usually repressed in somatic tissues, the appropriate line of inquiry may focus on factors that have previously been demonstrated to regulate retroelement activity in cells that occasionally accommodate retrotransposition, such as the early embryo (25, 37, 131). For example, the epigenetic regulator TRIM28/KAP1 has previously been demonstrated to regulate LTR retrotransposons in mouse embryonic stem cells (114). Furthermore, deletion of TRIM28/KAP1 in the forebrain of adult mice leads to stress-related behavioral abnormalities in learning and memory (58). It would therefore be interesting to determine whether TRIM28/KAP1 has a role in regulating retrotransposons in the mouse brain and whether neurological abnormalities associated with forebrain-specific TRIM28/KAP1 deletion are related to increased transcription and mobilization of retrotransposons. Similarly, other epigenetic effectors, such as the histone lysine methyltransferase SETDB1 (84), the histone deacetylase 1 (HDAC1) (112), the polycomb repressive complexes PRC1 and PRC2 (72), and the lysine-specific demethylase KDM1A/LSD1 (78), have been demonstrated to affect LTR retrotransposon activity in mouse embryonic stem cells. Whether these factors also modulate retrotransposon activity in the brain presents an interesting line of future inquiry.

MAPPING DE NOVO RETROTRANSPOSON INSERTIONS: CRITERIA AND DOCUMENTED EXAMPLES

Engineered L1s undergoing retrotransposition in cultured neuronal cell types and in the brains of transgenic animals, coupled with qPCR detection of L1 CNV in the human brain, constituted compelling yet incomplete evidence for endogenous L1 retrotransposition in mammalian neurons. In the following sections, we review recent studies in which deep-sequencing technologies and high-throughput analysis have been employed to identify and characterize somatic retrotransposon insertions, providing critical proof of bona fide retrotransposition in neurons. When considering such studies, however, it is important to bear in mind the unique challenges associated with mapping somatic retrotransposition events. New somatic insertions must be identified among the hundreds of thousands of copies already residing in the genome. Furthermore, somatic retrotransposon insertions are expected to be present in only a subset of cells and may even be unique to an individual cell. Determination of the extent of such somatic mosaicism requires analysis of single cell genomes, which can be achieved by whole-genome amplification (WGA). However, WGA may in turn introduce artifacts, such as chimeric sequences. In all cases, it is imperative that rigorous standards are upheld when calling and validating somatic transposon insertions from large sequencing data sets. Fortunately, retrotransposition events are usually accompanied by certain structural hallmarks that can be used to discern true insertions from other forms of genomic rearrangements or artifacts that may, at first pass, mimic bona fide transposable element activity.

L1-mediated retrotransposition, which includes mobilization of L1 as well as the nonautonomous retrotransposons *Alu* and SVA, accounts for all transposable element activity in humans. The ideal characterization of a de novo somatic L1-mediated retrotransposition event would comprise the following (**Figure 4**):

1. Mapping of L1-genome junctions at both the 5′ and 3′ flanks of the insertion, to single-nucleotide resolution.
2. Identification of structural features consistent with mobilization by retrotransposition: target-site duplications, poly-A tail, and 5′ truncation.
3. Insertion at a nucleotide motif resembling the loose L1 EN cleavage consensus site 5′-TTTT/A-3′.

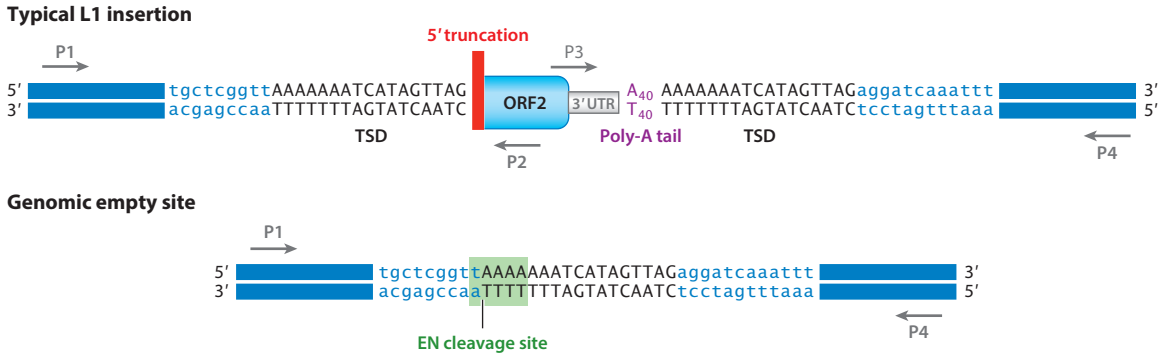


Figure 4

A fully-characterized typical L1 (long interspersed element 1) insertion. Hallmarks of retrotransposition include 5′ truncation (*red*), a poly-A tail (*purple*), and target-site duplications (TSDs) (*black*). Surrounding genomic DNA not contributing to the TSD is represented in blue. Below, the genomic empty site is depicted; the canonical L1 endonuclease cleavage motif (5′-TTTT/A-3′) is emphasized in green. Typical validation primers are depicted as gray arrows: P1 and P4 distinguish empty versus filled sites and can be used to confirm the empty site in control tissues. P1 and P2 amplify the 5′ L1-genome junction; P3 and P4 amplify the 3′ L1-genome junction. Abbreviation: EN, endonuclease.

4. Validation of both 5′ and 3′ L1-genome junctions by PCR and capillary sequencing.
5. Confirmation that the insertion is absent from matched control tissue.

Although it is preferable that all reported somatic insertions fulfill these specifications, certain technical and biological constraints make this task difficult or impossible to achieve in some cases. L1 insertions are frequently associated with deletions of target-site genomic DNA, as well as internal deletions and rearrangements of the L1 sequence itself (40, 41, 90, 126, 127) (**Figure 2**). Therefore, one end of an insertion may be easily identified from sequencing reads, but mapping the other end may be considerably more difficult. In the case of large deletions of genomic DNA or chromosomal translocations associated with retrotransposition events, mapping both L1-genome junctions of a new insertion may be virtually impossible. L1-mediated retrotransposition often occurs into highly repetitive genomic regions, which can confound mapping strategies and make PCR-based validation of insertions challenging. PCR validation can also be difficult for the 3′ ends of insertions because of the presence of the poly-A tail, as Baillie et al. (6) found. Transductions of 3′ flanking genomic DNA and, less frequently, 5′ flanking genomic DNA can occur during L1 mobilization (**Figure 3**). A transduced sequence presents an additional challenge to mapping strategies, but, excitingly, 3′ transductions provide a powerful tool to identify highly active progenitor elements that give rise to de novo insertions (79). Another point to consider is that not all retrotransposition events generate target-site duplications; a substantial percentage of insertions are blunt and can be associated with deletions of target-site DNA (40, 41). Indeed, as noted above, ENi L1 insertions typically lack TSDs, are frequently associated with L1 3′ truncations and deletion of target-site DNA, and are not expected to occur at the L1 EN consensus cleavage site (95, 96).

With the above limitations in mind, what constitutes sufficient evidence for a somatic insertion? Calling insertions based on discordant paired-end reads without detailed resolution of at least one L1-genome junction, because of a gap in the read pairs, is not ideal. Considering the prevalence of retroelement sequences in the genome, it is possible that such read pairs could indicate other types of genomic rearrangements or indeed could simply be DNA chimeras produced as an artifact of sequencing. We argue that for each putative insertion, at least one L1-genome junction should be identified at single-nucleotide resolution, from which the presence of an L1 EN cleavage consensus site can be discerned. Validation by PCR and capillary sequencing, although the strongest

evidence to verify a putative insertion, cannot practically be applied to every insertion from data sets numbering in the thousands. Therefore, a reasonable expectation is that a random subset of insertions should be chosen for PCR validation so that a false-positive rate can be determined. Validation PCRs performed in parallel on matched control tissues are also necessary to conclusively confirm that the insertion represents a somatic rather than germline event. These criteria may appear prescriptive. However, the spatiotemporal boundaries of reported somatic L1 activity are expanding rapidly and, critically, incorporate human diseases in which somatic L1 mobilization could be considered as an etiological factor targeted for clinical intervention. In this setting, stringent requirements for reporting new L1 insertions are arguably both necessary and appropriate.

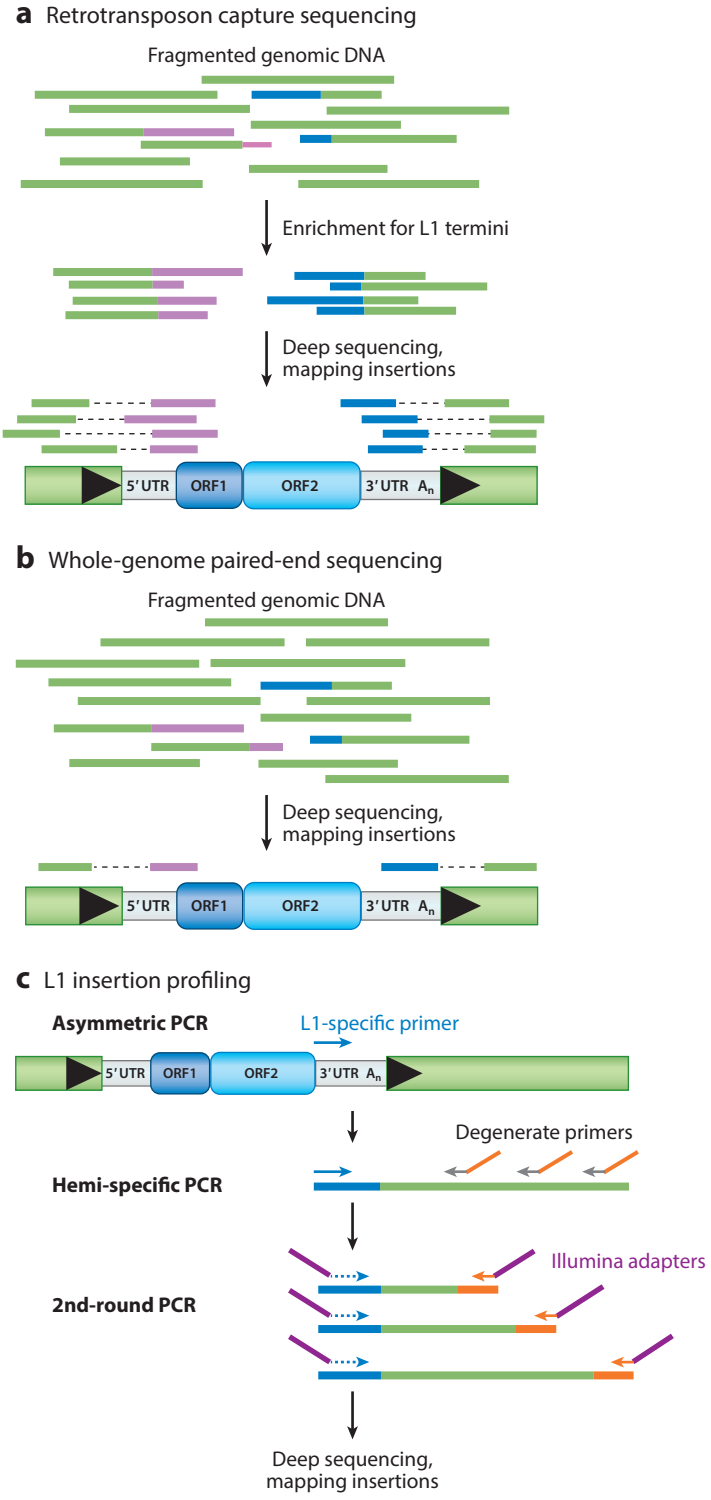
Mapping Neuronal L1 Retrotransposition Events by Retrotransposon Capture Sequencing

The first study to successfully map and characterize bona fide L1 insertions from human tissue employed a high-throughput approach for detection of endogenous retrotransposition events termed retrotransposon capture sequencing (RC-seq) (6) (**Figure 5*a***). In RC-seq, genomic DNA is captured using custom arrays targeting the termini of full-length L1, *Alu*, and SVA consensus sequences, and then is deeply sequenced to generate paired-end reads spanning retrotransposon-genome junctions. These read pairs are then mapped to the reference genome to identify loci containing known and novel insertions, following the premise that L1 insertion heterogeneity in the brain can be overcome through targeted L1 sequencing.

To identify brain-specific somatic L1 insertions, Baillie et al. (6) used genomic DNA extracted from five different brain regions of three elderly postmortem donors without pathological signs of neurological disease. Quantitative PCR detection of L1 CNV confirmed elevated L1 copy number in the hippocampus relative to other brain regions in two donors, consistent with previous observations (24). RC-seq was then performed on genomic DNA from the hippocampus (highest L1 copy number) and caudate nucleus (lowest L1 copy number) from all three individuals. Putative insertions present in more than one individual or brain region, existing catalogs of retroelement polymorphisms or RC-seq previously performed on an exclusive cohort of pooled genomic DNA extracted from human blood donors (as other tissues from the brain donors were not available), were designated as germline. Strikingly, only 8.4% of *Alu* insertions and 1.9% of L1 insertions

Figure 5

Strategies for sequencing somatic L1 (long interspersed element 1) insertions. (*a*) Retrotransposon capture sequencing (RC-seq). RC-seq selectively enriches randomly fragmented genomic DNA for 5′ and 3′ L1-genome junctions. Illumina adapters are ligated to enriched libraries and fragments are deeply sequenced. Paired-end reads are tiled across L1-genome junctions. The library insert size used by Baillie et al. (6) allowed a gap between paired-end reads; subsequent iterations of the RC-seq protocol require overlapping read pairs and therefore allow single-nucleotide resolution of L1-genome junctions (118). (*b*) Whole-genome sequencing (WGS) without L1 enrichment. WGS is carried out on fragmented genomic DNA. Putative L1 insertions are identified as paired-end reads, wherein one end aligns with the reference genome, and the other aligns with the retroelement sequence. Genomic DNA is shown in green; the 5′ termini and 3′ termini of L1 insertions are represented in light purple and blue, respectively. L1 target-site duplications are represented as black triangles. (*c*) L1 insertion profiling (L1-IP). L1-IP employs a hemi-specific PCR (polymerase chain reaction) scheme to amplify the 3′ flanking regions of L1 insertions. Asymmetric PCR with an L1-specific primer targeting active subfamily L1s (*blue arrow*) is followed by hemi-specific PCR reaction using degenerate primers (*gray arrows*) with linker sequences (*orange lines*). A second round of PCR with a second L1-specific primer (*dashed blue arrow*) introduces Illumina sequencing adapters (*purple lines*) to facilitate deep sequencing and mapping of putative insertions. Abbreviation: UTR, untranslated region.



detected from the brain were determined to be germline insertions, suggesting a vast number of potential brain-specific somatic insertions: 7,743 L1, 13,692 *Alu*, and 1,350 SVA insertions. The majority of the 29 selected insertions could be validated by PCR targeting the L1 5′ end-insertion site junction, consistent with confident identification of new retrotransposon insertions as laid out in **Figure 4**. For three insertions (one L1 and two *Alu*), both the 5′ and 3′ L1-insertion site junctions could be discerned from RC-seq reads. Among these, one L1 and one *Alu* insertion were flanked by TSDs, indicative of retrotransposition by canonical TPRT. Consistent with retrotransposon evolution and mobilization in the germline, more than 80% of somatic L1 and *Alu* insertions arose from the most recently active L1 and *Alu* subfamilies. However, the lack of a nonbrain control tissue, such as liver, meant that only insertions present in one brain region and not the other were termed somatic, effectively excluding from reporting any somatic insertions present in multiple brain regions.

L1 and *Alu* insertions in the brain occurred into genes with a higher frequency than known germline insertions. This observation supported the notion that intragenic insertions are generally deleterious, and therefore those occurring in the germline would quickly be lost through negative selection, whereas in the brain there is no opportunity for this selection to occur (130). However, the frequency of intragenic insertions in the brain was also higher than random expectation, and, indeed, L1 insertions occurred preferentially into genes highly expressed in the brain. Genes related to neurogenesis and synaptic function were shown to be favored for L1 insertion by gene ontology analysis, even when the higher average length of genes active in the brain was accounted for. Taken together, these results suggest that L1 mobilization in the brain occurs more readily into actively expressed and euchromatinized genes. These loci favor L1 EN target-site nicking, and, similarly, L1 integration here is more likely to impact host gene expression (52), as compared with a gene not expressed in the brain. Considering the numerous examples of how L1 insertions can impact gene function and regulation, it is tempting to speculate that mosaicism generated by somatic retrotransposition in the brain can alter the functional output of individual neurons, and perhaps contribute to neuronal plasticity.

Baillie et al. (6) detected somatic L1 retrotransposition in the brains of individuals who were of advanced age (~92 years on average) but otherwise presented no neuropathological abnormalities. If L1 retrotransposition is a normal part of mammalian brain physiology, what, if any, relationship is there between L1 retrotransposition and human neurological disease? As described above, Muotri et al. (100) approached this question with their discovery of increased L1 retrotransposition in mice deficient in MeCP2, which is mutated in RTT. Coufal et al. (23) similarly uncovered evidence for increased neuronal retrotransposition in ataxia telangiectasia patients. Future studies will no doubt shed light on the relationship between increased neuronal retroelement expression, retrotransposition, and neuronal dysfunction in RTT and ataxia telangiectasia.

In a very recent study, Bundo et al. (17) sought a link between increased L1 copy number and major mental health disorders, including schizophrenia. Schizophrenia is a psychiatric disorder characterized by psychosis and includes delusions, hallucinations, disorganized speech and behavior, impaired cognition, and altered emotional reactivity (98, 106). Schizophrenia rarely, if ever, exhibits Mendelian inheritance, and many genomic loci and copy-number variants, a prominent example being 22q11 deletion, have been associated with the disease (98). Bundo et al. (17) employed quantitative PCR and demonstrated an increase in L1 ORF2 sequences in neurons from the prefrontal cortex of schizophrenia patients compared with control patients. This result was recapitulated in mice and macaques exposed to treatments mimicking environmental risk factors for perturbed early neuronal development. In addition, iPSC-derived neurons from schizophrenia patients with the 22q11 deletion exhibited an elevated L1 copy number in the L1 CNV assay pioneered by Coufal et al. (24). Strikingly, however, whole-genome sequencing (WGS) of the

brain and liver tissue of schizophrenia patients, performed without enrichment for L1 sequences (**Figure 5*b***), revealed no significant increase in the number of brain-specific L1 insertions in schizophrenia patients compared with controls. The authors report that brain-specific L1s had a significantly higher tendency to retrotranspose into neuronal- and mental disorder–associated genes in schizophrenia patient brains than in control patients. However, examination of their WGS data set reveals that nearly half of the putative L1 insertions arose from older L1 subfamilies not currently active in the human genome, rather than from young, highly active elements, which is not consistent with de novo L1 retrotransposition. Furthermore, insertions were detected based only on discordant paired-end reads and therefore were not characterized at single-nucleotide resolution, and validation by PCR and capillary sequencing (**Figure 4**) was not performed. In sum, the relationship, if one exists, between L1 activity in the brain and the pathogenesis of schizophrenia requires extensive additional study before definitive conclusions can be drawn.

Detection of an increased L1 copy number by qPCR, in the absence of a corresponding increase in mapped L1 insertions in the brain, suggests that phenomena other than de novo L1 retrotransposition may contribute to L1 CNV in disease. Intriguingly, a recent study reported a massive increase in L1 cDNA in HIV-infected cells compared with control cells (59). A report by Han & Shao (51) has also demonstrated that TPRT events can undergo alternate pathways of resolution, leading to the generation of extrachromosomal circular DNA products bearing L1 sequence. Thus, one may speculate that under conditions of cellular stress, L1 reverse transcription could take place ectopically, perhaps by using an alternate cellular nucleic acid as a primer. Likewise, initiated TPRT events could resolve without integration of a new L1 copy into the genome. To further speculate, it is possible that cellular abnormalities associated with schizophrenia, or perhaps the drugs used in its treatment, could trigger aberrant L1 reverse transcription initiation or outcomes, thus accounting for an increase in the L1 copy number without a corresponding increase in genomic L1 integration. Indeed, a recent report used L1 CNV qPCR and detected a startling 255 L1 copy/cell increase in wild-type mouse embryonic stem cells and an ~860 copy/cell increase in *Dicer*-deficient mouse embryonic stem cells over 20 passages (19). As the mutational load exacted by hundreds of genomic insertions would likely be intolerable by embryonic stem cells, the generation of extrachromosomal L1 DNA may provide an alternate explanation for these results.

Similarly, large-scale genomic CNV between control tissues and brain, between brain subregions, and between individual neurons, is another potential confounding factor in the assessment of L1 CNV in disease. Indeed, recent reports employing WGA strategies have detected CNVs in individual neurons and other cell types (42, 86), although further work is required to definitively address whether these events represent technical artifacts of WGA. This phenomenon could result in the uneven deletion or duplication of resident L1 insertions, leading to differences in relative differences in the L1 copy number not arising from retrotransposition. Therefore, as stressed above, it is critical that qPCR-based L1 CNV assays are corroborated by the mapping of insertions bearing hallmarks of retrotransposition by TPRT.

THE EXTENT OF NEURONAL MOSAICISM: SINGLE-CELL APPROACHES

Detection of bona fide somatic L1 insertions from bulk human brain tissue, complemented by earlier observations, provides definitive evidence that the neuronal genome is a somatic mosaic due to L1 mobilization. However, these analyses do not directly address the key question of how complex that mosaicism is, i.e., how frequently does L1 mobilize in the neuronal lineage? Low read depth for brain-specific L1 insertions relative to germline insertions uncovered by Baillie

et al. (6) suggests that each insertion is present in only a subset of neurons, but does not constitute a quantitative measure of the prevalence of a given L1 insertion in the brain. It is possible that, owing to L1 mobilization in the latter phases of neurogenesis, each individual neuron is genetically unique with regard to its cohort of loci containing somatic L1 insertions. It is also possible that L1 insertions arising during embryogenesis could be present in large subpopulations of neurons throughout the brain. Indeed, both of these scenarios could be true, painting a complex picture of L1-driven neuronal mosaicism in the adult brain.

The extent to which individual neurons share or differ in L1 content can be resolved by single cell–based approaches. Analyzing the genome of single cells presents several technical challenges, as WGA is required to generate sufficient material for analysis. The individual cell in question is consumed during the amplification process, and no starting material remains from which to validate results obtained from amplified DNA. Single-cell analysis experiments must therefore be carefully designed and subject to rigorous quality-control measures. Data from single-cell analysis must also be interpreted with technical caveats, such as WGA bias and the likely presence of molecular artifacts, in mind (70).

In the first study to successfully perform single-neuron analysis of L1 insertions, Evrony et al. (32) isolated single nuclei from the caudate nucleus and pyramidal nuclei from the frontal cortex of three individuals by fluorescence-activated cell sorting (FACS) and performed multiple displacement amplification (MDA) to obtain 15–20 μg of DNA from each of 300 single nuclei. Putative L1 insertions were identified using a modified version of the method of Ewing & Kazazian (33), which entails hemi-nested PCR amplification of L1-genome junctions, using primers annealing specifically to the 3′ end of human-specific L1s in combination with degenerate primers to anneal to genomic flanking DNA (**Figure 5*c***). For this study, the method of Ewing et al. (33) was adapted for high-throughput, multiplexed sequencing and, along with the MDA, required an estimated equivalent of ~60 cycles of PCR for library preparation. In the amplified genomes of individual neurons, greater than 80% of known reference insertions could be detected on average and for single-copy X-linked insertions in nuclei from a male donor, this figure was approximately 75%. For detection of somatic L1 insertions from individual neurons, the authors used a sensitivity threshold at which only 50% of known reference insertions can be detected. The frequency of unique L1 insertions was estimated at less than one per neuron, with 82% of neurons harboring no unique insertions. PCR validation of L1 3′ end-genome junctions (**Figure 4**) yielded results consistent with a bona fide L1 retrotransposition event for only 5 out of 81 putative insertions, suggesting a high rate of false positives due to MDA chimeras and technical artifacts. Among the five validated insertions, one could be fully characterized and bore the hallmarks of L1 retrotransposition; owing to a 5′ transduction, this insertion could be traced to a progenitor element on chromosome 8. This insertion was present in two cortical nuclei and could be amplified from bulk samples. The other four insertions could be characterized only at their 3′ L1-genome junction and appeared to be specific to the individual neuron in which they were detected.

The work of Evrony et al. (32) represented a major methodological advance for studies of L1 and for single-cell genomics in general. Evrony et al. (32) identified and validated the first somatic full-length L1 insertion found in the brain, emphatically demonstrating that L1-mediated mosaicism can occur in neurons. Notably, the estimates of L1 mobilization frequency produced by Evrony et al. (32) were more than two orders of magnitude lower than the qPCR-based estimates of Coufal et al. (24) (0.6 versus 80 per neuron, respectively). The reasons for this disparity remain to be ascertained. However, even if the former figure is provisionally accepted and extrapolated, 6×10^{10} somatic L1 insertions would be present in each human brain, a figure contrasting with the aforementioned estimate of 4×10^{8} polymorphic retrotransposon insertions in the entire human population and, in our opinion, at odds with Evrony et al.'s (32) description of somatic

L1 mobilization as rare. It is also important to note that the figures reported by Coufal et al. (24) specified the hippocampus as a hotspot of L1 retrotransposition, whereas Evrony et al. (32) studied cortical neurons. Moreover, chimera formation during MDA is an established source of artifacts (70). To exclude events representing localized chimeras between existing L1 insertions and nearby genomic DNA, Evrony et al. (32) only considered somatic L1 insertions greater than 20 kb from known L1 reference genome and polymorphic insertions. Finally, the true-positive rate from the 3′ L1-genome junction PCR is critical to estimating the somatic L1 mobilization rate with this approach. It is possible that MDA amplification is biased in favor of chimeric sequences, which could lead to a higher read count for chimeras relative to bona fide L1 insertions, thus skewing the pool of potential insertions selected for PCR validation in favor of false positives. Therefore, prioritizing insertions for validation on the basis of high read count could artificially elevate the false-positive rate. In sum, although the identification by Evrony et al. (32) of a full-length L1 insertion bearing the hallmarks of TPRT was a landmark achievement, substantial additional work is still required to accurately quantify somatic L1 mobilization in the human brain.

NEURONAL RETROTRANSPOSITION IN *DROSOPHILA*

Retrotransposition in neurons, particularly in brain regions critical for learning and memory, was very recently revealed to be not exclusive to mammals. Perrat et al. (108) uncovered evidence for increased transposable element expression in the *Drosophila* mushroom body, a brain structure required for olfactory memory. Specifically, transposon transcripts were highly expressed in αβ neurons, which are proposed to be required for storage and recall of memories, as contrasted to α′β′ or γ neurons, which are proposed to be required for acquisition and stabilization of memories. Transposable element transcripts expressed in αβ neurons include LTR and non-LTR retrotransposons, as well as DNA transposons. To identify somatic mobilization events, Perrat et al. (108) employed WGS on genomic DNA extracted from αβ neurons, other brain tissue, and embryos genetically identical to the flies used in the neuron analysis. They reported 215 insertions present in αβ neurons but absent from embryos and the *Drosophila* reference genome, and 200 such insertions in other brain tissue, 19 of which overlapped with the set identified in αβ neurons. Putative insertions were identified by discordant paired-end reads, which do not allow single-nucleotide resolution of retroelement-genome junctions or identification of the structural hallmarks of transposable element integration (**Figure 4**). Furthermore, sequencing depth was ∼3.1 times for a sample of 5,000 pooled αβ neurons, requiring extrapolation to assess the per-neuron transposition rate. Future validation of putative insertions by PCR and capillary sequencing, perhaps combined with single-neuron analysis, are required to definitively assess how frequently endogenous transposable element insertions occur in the *Drosophila* brain.

Evidence for transposon activity in the *Drosophila* brain was also recently uncovered by Li et al. (74), who found that some transposable elements exhibited age-dependent transcriptional upregulation. The authors used a *gypsy-TRAP* reporter system, in which insertions of the LTR retrotransposon *gypsy* into an engineered reporter construct activate GFP expression in the affected cell, to demonstrate retrotransposition in mushroom body neurons of aged flies but not their younger counterparts. The *gypsy*-TRAP system specifically detects *gypsy* retrotransposition events occurring into a transgenic preferred target locus and therefore does not provide information about the activity of other transposable elements or the full spectrum of genomic insertion sites targeted by transposons. Nevertheless, this experiment clearly demonstrates that neuronal retrotransposition occurs in *Drosophila* in an age-dependent manner. This result raises the question of whether transposable element insertions also accumulate in the mammalian brain as a function of advancing age.

Both of the above studies also investigated how transposable elements are regulated in the *Drosophila* brain. Prompted by the observation that the translocated *Stellate* locus *STE12DOR*, which is usually repressed by piRNAs (PIWI-interacting RNAs), was upregulated in αβ neurons relative to other brain tissues, Perrat et al. (108) investigated whether the piRNA pathway is involved in differential transposon regulation among neuronal types in the *Drosophila* brain. Indeed, they found that αβ neurons do not express the PIWI clade proteins Aub or Ago3, consistent with higher transposon expression compared with PIWI-expressing α′β′ and γ neurons. Furthermore, analysis of RNA isolated from the heads of PIWI-mutant, piRNA-defective flies revealed elevated expression levels of some transposons. Consistent with previous reports (39), Perrat et al. (108) also found that heads from endo-siRNA (small interfering RNA)-defective *ago2* and *dcr-2* mutant flies contained elevated levels of certain transposon transcripts. Taken together, these results suggest that the endo-siRNA pathway and the piRNA pathway, which was previously demonstrated to repress transposons in the *Drosophila* germline, may both contribute to transposon regulation in the *Drosophila* brain.

In parallel, Li et al. (74) investigated the functional consequences of transposon derepression in the *Drosophila* brain by employing *ago2* mutant flies in which endo-siRNA-mediated transposon repression is released in somatic tissues. In *ago2* mutant heads, R2 and gypsy transposon transcripts were elevated relative to controls, and transcript levels from *ago2* mutant young flies were comparable to those observed in wild-type aged flies. Elevated *gypsy* transcript levels were accompanied by the detection of de novo gypsy insertions in the *ovo* locus, an established preferred *gypsy* integration target. To investigate a functional link between elevated transposon activity and age-dependent neuronal impairment, *ago2* mutant flies were subjected to learning and memory tests. Young *ago2* mutant flies exhibited memory defects that could be rescued by *ago2* transgene expression and that worsened with increasing age. *Ago2* mutant flies also had shorter life spans. Disruption in neurons of *loki*, the *Drosophila* Chk2 ortholog that mediates DNA damage-induced apoptosis, increased life span in wild-type flies, and also partially relieved age-dependent memory loss. Although links between increased transposon activity, memory loss, and mortality uncovered by this study are correlative and await more direct mechanistic studies of the consequences of transposon activity on neuronal function and comprehensive mapping of transposon insertions in the *Drosophila* brain, the evidence for a functional consequence of neuronal transposition is nevertheless intriguing.

REMAINING QUESTIONS AND FUTURE DIRECTIONS

Perhaps the most interesting question regarding neuronal retrotransposition, and also the most difficult to answer, is whether these events play a role in normal brain function. In order to elucidate a function for retrotransposition in the brain, it will be important to accomplish a thorough characterization of the rate, developmental timing, and genomic impact of neuronal retrotransposition. From this information, testable hypotheses about the function of neuronal retrotransposition may emerge. Below, we highlight fundamental questions that should be resolved in the near future.

What Is the Rate of Neuronal Retrotransposon Insertions?

On the basis of a quantitative PCR assay, Coufal et al. (24) estimated that the adult human hippocampus contains approximately 80 new L1 copies per cell, suggesting a strikingly high rate of somatic retrotransposition in the brain. Baillie et al. (6) identified 7,743 neuronal L1 insertions

from RC-seq analysis of bulk DNA, yielding a conservative estimate of 0.04 unique insertions per neuron. Evrony et al. (32) employed single-cell analysis and arrived at a figure of <0.6 unique somatic insertions per neuron; when taking into account only insertions that could be validated by PCR, this figure drops to 0.04 unique insertions per neuron. There are several potential reasons for the discrepancy between estimates based on qPCR and those based on numbers of mapped insertions. As discussed above, qPCR assays detect relative L1 DNA content but do not discriminate between bona fide genomic L1 insertions and other sources of increased L1 DNA, such as genomic instability not arising from retrotransposition and perhaps the accumulation of extrachromosomal L1 reverse transcription products (59, 124). It is therefore possible that qPCR-based assays could produce an overestimate of L1 retrotransposition. However, estimating the actual rate of L1 retrotransposition using insertions detected in bulk DNA by the method of Baillie et al. (6) is fraught with uncertainty because detection is probably not at saturation and insertions present in only one or a small subset of neurons are likely to evade detection. Indeed, it is impossible to determine how many insertions are missed in this method. Analysis of single cells will ultimately produce the most reliable quantification of neuronal retrotransposition. Evrony et al. (32) have taken the first step toward this end, and as methods for WGA from single cells improve and strategies for detection become more sensitive, an accurate estimate of the rate of neuronal retrotransposition will ultimately emerge.

What Is the Developmental Timing and Cell-Type Specificity of Neuronal Retrotransposition?

The developmental timing and cell-type specificity of neuronal retrotransposition are important subjects for future investigation. Neuronal retrotransposition events may accumulate during a specific developmental stage, such as embryonic development, or the generation of neuronal somatic mosaicism may be an ongoing process throughout the life of an organism. It is important to determine whether retrotransposition occurs mainly in dividing neuronal precursor cells, as previous studies have suggested (24, 99), or whether new events can occur in fully differentiated neurons. The answers to such questions will contribute to a complete picture of retrotransposition-derived neuronal mosaicism. Furthermore, this knowledge will direct speculation as to whether retrotransposition generates pre-existing neuronal diversity that can be exploited during learning and memory formation, or whether new neuronal retrotransposition events occur in response to external stimuli. Indeed, Muotri et al. (101) found that voluntary exercise is correlated with an increase in hippocampal retrotransposition in mice, suggesting that external stimuli may indeed lead to increased retrotransposition and therefore neuronal genomic diversity.

What Are the Functional Consequences of Neuronal Retrotransposition?

Muotri et al. (99) demonstrated that an engineered L1 insertion could alter gene expression and cell fate in rat NPCs cultured in vitro. Future studies should focus on how such alterations to target gene expression occur on a mechanistic level. Numerous studies have put forth mechanisms by which L1 insertions can disrupt transcript integrity and expression levels of target genes, and L1 insertions are also associated with epigenetic alterations to target DNA. It will be interesting to determine the prevalence of these various alterations in neuronal retrotransposition, and whether these changes alter the expression and integrity of target genes. Addressing these questions on a per-neuron basis presents a formidable technical challenge, as current techniques do not allow analysis of genomic DNA and RNA expression from the same cell.

How Does Neuronal Retrotransposon Activity Relate to Human Disease?

Understanding a potential role for neuronal retrotransposition in normal physiology will be aided by examining the relationship between retrotransposition and brain disorders and aging. The observations that L1 is highly expressed and retrotransposition occurs with high frequency in apparently normal human and rodent brains indicate that a certain level of L1 activity is tolerable and may even have a physiological function. However, demonstrated correlations between L1 upregulation and RTT in humans and dysregulation of transposable elements by *ago2* deficiency and memory impairment in flies suggest that too much neuronal transposition may have negative consequences (74, 100). Similarly, numerous cellular factors, including the APOBEC3 cytidine deaminases (13, 14, 18, 64, 97, 103, 123), the Aicardi-Goutieres syndrome gene products Trex1 (124) and SAMHD1 (138), and the Mov10 putative helicase (4, 43, 75), have been demonstrated to restrict L1 retrotransposition in vitro. It would be interesting to determine whether deficiency in such factors in humans or rodents in vivo leads to an increase in neuronal retrotransposition and whether this activity has any impact on brain function.

Given recent reports demonstrating somatic L1 mobilization in cancer (57, 71, 118, 121) and the emergence of the brain as a major site of somatic retrotransposition in mammals, the logical follow-up question is whether L1 mobilization occurs in brain tumors. Strikingly, examination of 5 medulloblastoma and 5 glioblastoma genomes by Iskow et al. (57) and 19 glioblastoma genomes by Lee et al. (71) revealed no tumor-specific retroelement insertions. One explanation for this result is that retrotransposition in the brain appears to occur primarily in the neuronal lineage (99) (**Figure 3*b***) and may not, therefore, be prevalent in tumors arising from the glial cell lineage (76). Alternatively, if retrotransposition is indeed a normal or even necessary factor in brain physiology, host mechanisms may direct new insertions to specific regions of the neuronal genome, precluding potentially oncogenic retrotransposition events.

To conclude, the surprising discovery in recent years that somatic retrotransposition can occur in the metazoan brain provides fuel for speculation regarding the functional impact of transposable elements upon both normal and abnormal physiology. A thorough characterization of the rate, timing, consequences, and regulation of neuronal retrotransposition is, however, required before a functional role for retrotransposition in the brain can be conclusively discerned. In any case, that genetic mosaicism associated with the mobilization of transposable elements is now a major focus of research is somewhat ironic, considering that the original characterization of transposition in maize by McClintock (85) nearly 60 years ago described a somatic phenomenon. There, a phenotypic outcome was clear. Here, the functional consequences of L1 activity in neurons may prove comparatively elusive and yet, arguably, equally important for our understanding of the genetic basis for life.

SUMMARY POINTS

1. Recent studies have demonstrated that mobile elements are active and generate somatic mosaicism in the mammalian and *Drosophila* brain.
2. Certain regulatory factors, including the methyl-binding protein MeCP2 and the master cell cycle regulator ATM, play a role in limiting retrotransposon activity in the mammalian brain. Deficiencies in such factors suggest a correlation between deregulated neuronal retrotransposition and neurological disease.

3. Advances in sequencing technology have allowed mapping and characterization of somatic retrotransposon insertions. However, sequencing data must be carefully analyzed to distinguish bona fide retrotransposon insertions from other genomic rearrangements or technical artifacts.

FUTURE ISSUES

1. What are the rate, developmental timing, and cell-type specificity of neuronal retrotransposition?
2. How are elevated levels of retroelement expression and mobilization observed in neurological diseases, such as RTT, related to disease etiology?
3. Does retrotransposition in the brain, and the resultant somatic mosaicism, play a beneficial or vital role in normal brain physiology?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Jose Garcia-Perez and Adam Ewing for reading the manuscript and constructive comments, and apologize to colleagues whose work could not be cited because of space limitations. G.J.F. acknowledges the support of an Australian NHMRC Career Development Fellowship (GNT1045237), NHMRC Project Grants GNT1042449, GNT1045991, GNT1052303, GNT1067983, and GNT1068789 and the European Union's Seventh Framework Program (FP7/2007-2013) under grant agreement No. 259743 underpinning the MODHEP consortium.

LITERATURE CITED

1. Alisch RS, Garcia-Perez JL, Muotri AR, Gage FH, Moran JV. 2006. Unconventional translation of mammalian LINE-1 retrotransposons. *Genes Dev.* 20:210–24
2. Ambrose M, Gatti RA. 2013. Pathogenesis of ataxia-telangiectasia: the next generation of ATM functions. *Blood* 121:4036–45
3. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. 1999. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* 23:185–88
4. 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
5. Athanikar JN, Badge RM, Moran JV. 2004. A YY1-binding site is required for accurate human LINE-1 transcription initiation. *Nucleic Acids Res.* 32:3846–55
6. Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, et al. 2011. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* 479:534–37
7. Beck CR, Collier P, Macfarlane C, Malig M, Kidd JM, et al. 2010. LINE-1 retrotransposition activity in human genomes. *Cell* 141:1159–70

8. Beck CR, Garcia-Perez JL, Badge RM, Moran JV. 2010. LINE-1 elements in structural variation and disease. *Annu. Rev. Genomics Hum. Genet.* 12:187–215
9. 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.* 2:1697–702
10. Belancio VP, Hedges DJ, Deininger P. 2008. Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. *Genome Res.* 18:343–58
11. Belancio VP, Roy-Engel AM, Deininger P. 2008. The impact of multiple splice sites in human L1 elements. *Gene* 411:38–45
12. Bestor TH, Bourc'his D. 2004. Transposon silencing and imprint establishment in mammalian germ cells. *Cold Spring Harb. Symp. Quant. Biol.* 69:381–87
13. Bogerd HP, Wiegand HL, Doehle BP, Lueders KK, Cullen BR. 2006. APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. *Nucleic Acids Res.* 34:89–95
14. Bogerd HP, Wiegand HL, Hulme AE, Garcia-Perez JL, O'Shea KS, et al. 2006. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc. Natl. Acad. Sci. USA* 103:8780–85
15. Brouha B, Meischl C, Ostertag E, de Boer M, Zhang Y, et al. 2002. Evidence consistent with human L1 retrotransposition in maternal meiosis I. *Am. J. Hum. Genet.* 71:327–36
16. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, et al. 2003. Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl. Acad. Sci. USA* 100:5280–85
17. Bundo M, Toyoshima M, Okada Y, Akamatsu W, Ueda J, et al. 2014. Increased L1 retrotransposition in the neuronal genome in schizophrenia. *Neuron* 81:306–13
18. Chen H, Lilley CE, Yu Q, Lee DV, Chou J, et al. 2006. APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. *Curr. Biol.* 16:480–85
19. Ciaudo C, Jay F, Okamoto I, Chen CJ, Sarazin A, et al. 2013. RNAi-dependent and independent control of LINE1 accumulation and mobility in mouse embryonic stem cells. *PLoS Genet.* 9:e1003791
20. Cordaux R, Batzer MA. 2009. The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* 10:691–703
21. 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* 37:18081–93
22. Cost GJ, Golding A, Schlissel MS, Boeke JD. 2001. Target DNA chromatinization modulates nicking by L1 endonuclease. *Nucleic Acids Res.* 29:573–77
23. Coufal NG, Garcia-Perez JL, Peng GE, Marchetto MC, Muotri AR, et al. 2011. Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. *Proc. Natl. Acad. Sci. USA* 108:20382–87
24. Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, et al. 2009. L1 retrotransposition in human neural progenitor cells. *Nature* 460:1127–31
25. de Boer M, van Leeuwen K, Geissler J, Weemaes CM, van den Berg TK, et al. 2014. Primary immunodeficiency caused by an exonized retroposed gene copy inserted in the *CYBB* gene. *Hum. Mutat.* 35:486–96
26. 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
27. Dmitriev SE, Andreev DE, Terenin IM, Olovnikov IA, Prassolov VS, et al. 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.* 27:4685–97
28. Dombroski BA, Mathias SL, Nanthakumar E, Scott AF, Kazazian HH Jr. 1991. Isolation of an active human transposable element. *Science* 254:1805–8
29. Doucet AJ, Hulme AE, Sahinovic E, Kulpa DA, Moldovan JB, et al. 2010. Characterization of LINE-1 ribonucleoprotein particles. *PLoS Genet.* 6:e1001150
30. Ergun S, Buschmann C, Heukeshoven J, Dammann K, Schnieders F, et al. 2004. Cell type–specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. *J. Biol. Chem.* 279:27753–63

31. Esnault C, Maestre J, Heidmann T. 2000. Human LINE retrotransposons generate processed pseudogenes. *Nat. Genet.* 24:363–67
32. Evrony GD, Cai X, Lee E, Hills LB, Elhosary PC, et al. 2012. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* 151:483–96
33. Ewing AD, Kazazian HH Jr. 2010. High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome Res.* 20:1262–70
34. Faulkner GJ. 2011. Retrotransposons: mobile and mutagenic from conception to death. *FEBS Lett.* 585:1589–94
35. Faulkner GJ, Kimura Y, Daub CO, Wani S, Plessy C, et al. 2009. The regulated retrotransposon transcriptome of mammalian cells. *Nat. Genet.* 41:563–71
36. Feng Q, Moran JV, Kazazian HH Jr, Boeke JD. 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 87:905–16
37. Garcia-Perez JL, Marchetto MC, Muotri AR, Coufal NG, Gage FH, et al. 2007. LINE-1 retrotransposition in human embryonic stem cells. *Hum. Mol. Genet.* 16:1569–77
38. Garcia-Perez JL, Morell M, Scheys JO, Kulpa DA, Morell S, et al. 2010. Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. *Nature* 466:769–73
39. Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, et al. 2008. Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* 320:1077–81
40. Gilbert N, Lutz S, Morrish TA, Moran JV. 2005. Multiple fates of L1 retrotransposition intermediates in cultured human cells. *Mol. Cell. Biol.* 25:7780–95
41. Gilbert N, Lutz-Prigge S, Moran JV. 2002. Genomic deletions created upon LINE-1 retrotransposition. *Cell* 110:315–25
42. Gole J, Gore A, Richards A, Chiu YJ, Fung HL, et al. 2013. Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. *Nat. Biotechnol.* 31:1126–32
43. Goodier JL, Cheung LE, Kazazian HH Jr. 2012. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. *PLoS Genet.* 8:e1002941
44. Goodier JL, Cheung LE, Kazazian HH Jr. 2013. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids Res.* 41:7401–19
45. Goodier JL, Kazazian HH Jr. 2008. Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell* 135:23–35
46. Goodier JL, Ostertag EM, Du K, Kazazian HH Jr. 2001. A novel active L1 retrotransposon subfamily in the mouse. *Genome Res.* 11:1677–85
47. Goodier JL, Ostertag EM, Kazazian HH Jr. 2000. Transduction of 3′-flanking sequences is common in L1 retrotransposition. *Hum. Mol. Genet.* 9:653–57
48. Grimaldi G, Skowronski J, Singer MF. 1984. Defining the beginning and end of KpnI family segments. *EMBO J.* 3:1753–59
49. Guy J, Gan J, Selfridge J, Cobb S, Bird A. 2007. Reversal of neurological defects in a mouse model of Rett syndrome. *Science* 315:1143–47
50. Hall LL, Carone DM, Gomez AV, Kolpa HJ, Byron M, et al. 2014. Stable C0T-1 repeat RNA is abundant and is associated with euchromatic interphase chromosomes. *Cell* 156:907–19
51. Han JS, Shao S. 2012. Circular retrotransposition products generated by a LINE retrotransposon. *Nucleic Acids Res.* 40:10866–77
52. Han JS, Szak ST, Boeke JD. 2004. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature* 429:268–74
53. Hancks DC, Kazazian HH Jr. 2012. Active human retrotransposons: variation and disease. *Curr. Opin. Genet. Dev.* 22:191–203
54. Heras SR, Macias S, Plass M, Fernandez N, Cano D, et al. 2013. The Microprocessor controls the activity of mammalian retrotransposons. *Nat. Struct. Mol. Biol.* 20:1173–81
55. Hohjoh H, Singer MF. 1996. Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *EMBO J.* 15:630–39
56. Iida S, Morita Y, Choi JD, Park KI, Hoshino A. 2004. Genetics and epigenetics in flower pigmentation associated with transposable elements in morning glories. *Adv. Biophys.* 38:141–59

57. Iskow RC, McCabe MT, Mills RE, Torene S, Pittard WS, et al. 2010. Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell* 141:1253–61
58. Jakobsson J, Cordero MI, Bisaz R, Groner AC, Busskamp V, et al. 2008. KAP1-mediated epigenetic repression in the forebrain modulates behavioral vulnerability to stress. *Neuron* 60:818–31
59. Jones RB, Song H, Xu Y, Garrison KE, Buzdin AA, et al. 2013. LINE-1 retrotransposable element DNA accumulates in HIV-1-infected cells. *J. Virol.* 87:13307–20
60. Jurka J. 1997. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proc. Natl. Acad. Sci. USA* 94:1872–77
61. Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, et al. 2009. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes Dev.* 23:1303–12
62. Kazazian HH Jr, Wong C, Youssoufian 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* 332:164–66
63. Khazina E, Weichenrieder O. 2009. Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. *Proc. Natl. Acad. Sci. USA* 106:731–36
64. Kinomoto M, Kanno T, Shimura M, Ishizaka Y, Kojima A, et al. 2007. All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. *Nucleic Acids Res.* 35:2955–64
65. Kubo S, Seleme MC, Soifer HS, Perez JL, Moran JV, et al. 2006. L1 retrotransposition in nondividing and primary human somatic cells. *Proc. Natl. Acad. Sci. USA* 103:8036–41
66. Kulpa DA, Moran JV. 2005. Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. *Hum. Mol. Genet.* 14:3237–48
67. Kulpa DA, Moran JV. 2006. *cis*-Preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat. Struct. Mol. Biol.* 13:655–60
68. Kuwabara T, Hsieh J, Muotri A, Yeo G, Warashina M, et al. 2009. Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. *Nat. Neurosci.* 12:1097–105
69. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. 2001. Initial sequencing and analysis of the human genome. *Nature* 409:860–921
70. Lasken RS, Stockwell TB. 2007. Mechanism of chimera formation during the multiple displacement amplification reaction. *BMC Biotechnol.* 7:19
71. Lee E, Iskow R, Yang L, Gokcumen O, Haseley P, et al. 2012. Landscape of somatic retrotransposition in human cancers. *Science* 337:967–71
72. Leeb M, Pasini D, Novatchkova M, Jaritz M, Helin K, Wutz A. 2010. Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev.* 24:265–76
73. Levin HL, Moran JV. 2011. Dynamic interactions between transposable elements and their hosts. *Nat. Rev. Genet.* 12:615–27
74. Li W, Prazak L, Chatterjee N, Gruninger S, Krug L, et al. 2013. Activation of transposable elements during aging and neuronal decline in *Drosophila*. *Nat. Neurosci.* 16:529–31
75. Li X, Zhang J, Jia R, Cheng V, Xu X, et al. 2013. The MOV10 helicase inhibits LINE-1 mobility. *J. Biol. Chem.* 288:21148–60
76. Liu C, Sage JC, Miller MR, Verhaak RG, Hippenmeyer S, et al. 2011. Mosaic analysis with double markers reveals tumor cell of origin in glioma. *Cell* 146:209–21
77. 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
78. Macfarlan TS, Gifford WD, Agarwal S, Driscoll S, Lettieri K, et al. 2011. Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. *Genes Dev.* 25:594–607
79. Macfarlane CM, Collier P, Rahbari R, Beck CR, Wagstaff JF, et al. 2013. Transduction-specific ATLAS reveals a cohort of highly active L1 retrotransposons in human populations. *Hum. Mutat.* 34:974–85
80. Martin SL. 1991. Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells. *Mol. Cell. Biol.* 11:4804–7
81. Martin SL, Bushman FD. 2001. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. *Mol. Cell. Biol.* 21:467–75

82. Mathias SL, Scott AF, Kazazian HH Jr, Boeke JD, Gabriel A. 1991. Reverse transcriptase encoded by a human transposable element. *Science* 254:1808–10
83. Matlik K, Redik K, Speek M. 2006. L1 antisense promoter drives tissue-specific transcription of human genes. *J. Biomed. Biotechnol.* 2006:71753
84. Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H, et al. 2010. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* 464:927–31
85. McClintock B. 1950. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. USA* 36:344–55
86. McConnell MJ, Lindberg MR, Brennand KJ, Piper JC, Voet T, et al. 2013. Mosaic copy number variation in human neurons. *Science* 342:632–37
87. McKay R. 1997. Stem cells in the central nervous system. *Science* 276:66–71
88. Miki Y, Nishisho I, Horii A, Miyoshi Y, Utsunomiya J, et al. 1992. Disruption of the *APC* gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res.* 52:643–45
89. Minakami R, Kurose K, Etoh K, Furuhata Y, Hattori M, Sakaki Y. 1992. Identification of an internal *cis*-element essential for the human L1 transcription and a nuclear factor(s) binding to the element. *Nucleic Acids Res.* 20:3139–45
90. Mine M, Chen JM, Brivet M, Desguerre I, Marchant D, et al. 2007. A large genomic deletion in the *PDHX* gene caused by the retrotranspositional insertion of a full-length LINE-1 element. *Hum. Mutat.* 28:137–42
91. Monot C, Kuciak M, Viollet S, Mir AA, Gabus C, et al. 2013. The specificity and flexibility of L1 reverse transcription priming at imperfect T-tracts. *PLoS Genet.* 9:e1003499
92. Moran JV, DeBerardinis RJ, Kazazian HH Jr. 1999. Exon shuffling by L1 retrotransposition. *Science* 283:1530–34
93. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH Jr. 1996. High frequency retrotransposition in cultured mammalian cells. *Cell* 87:917–27
94. Morgan HD, Sutherland HG, Martin DI, Whitelaw E. 1999. Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* 23:314–18
95. Morrish TA, Garcia-Perez JL, Stamato TD, Taccioli GE, Sekiguchi J, Moran JV. 2007. Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres. *Nature* 446:208–12
96. Morrish TA, Gilbert N, Myers JS, Vincent BJ, Stamato TD, et al. 2002. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat. Genet.* 31:159–65
97. Muckenfuss H, Hamdorf M, Held U, Perkovic M, Lower J, et al. 2006. APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J. Biol. Chem.* 281:22161–72
98. Mulle JG. 2012. Schizophrenia genetics: progress, at last. *Curr. Opin. Genet. Dev.* 22:238–44
99. Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. 2005. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* 435:903–10
100. Muotri AR, Marchetto MC, Coufal NG, Oefner R, Yeo G, et al. 2010. L1 retrotransposition in neurons is modulated by MeCP2. *Nature* 468:443–46
101. Muotri AR, Zhao C, Marchetto MC, Gage FH. 2009. Environmental influence on L1 retrotransposons in the adult hippocampus. *Hippocampus* 19:1002–7
102. Naas TP, DeBerardinis RJ, Moran JV, Ostertag EM, Kingsmore SF, et al. 1998. An actively retrotransposing, novel subfamily of mouse L1 elements. *EMBO J.* 17:590–97
103. Niewiadomska AM, Tian C, Tan L, Wang T, Sarkis PT, Yu XF. 2007. Differential inhibition of long interspersed element 1 by APOBEC3 does not correlate with high-molecular-mass-complex formation or P-body association. *J. Virol.* 81:9577–83
104. Nigumann P, Redik K, Matlik K, Speek M. 2002. Many human genes are transcribed from the antisense promoter of L1 retrotransposon. *Genomics* 79:628–34
105. Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, et al. 2002. A mouse model of human L1 retrotransposition. *Nat. Genet.* 32:655–60
106. Owen MJ, Williams HJ, O'Donovan MC. 2009. Schizophrenia genetics: advancing on two fronts. *Curr. Opin. Genet. Dev.* 19:266–70
107. Perepelitsa-Belancio V, Deininger P. 2003. RNA truncation by premature polyadenylation attenuates human mobile element activity. *Nat. Genet.* 35:363–66

108. Perrat PN, DasGupta S, Wang J, Theurkauf W, Weng Z, et al. 2013. Transposition-driven genomic heterogeneity in the *Drosophila* brain. *Science* 340:91–95
109. Pickeral OK, Makalowski W, Boguski MS, Boeke JD. 2000. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. *Genome Res.* 10:411–15
110. Ray DA, Batzer MA. 2011. Reading TE leaves: new approaches to the identification of transposable element insertions. *Genome Res.* 21:813–20
111. Rebollo R, Karimi MM, Bilenky M, Gagnier L, Miceli-Royer K, et al. 2011. Retrotransposon-induced heterochromatin spreading in the mouse revealed by insertional polymorphisms. *PLoS Genet.* 7:e1002301
112. Reichmann J, Crichton JH, Madej MJ, Taggart M, Gautier P, et al. 2012. Microarray analysis of LTR retrotransposon silencing identifies Hdac1 as a regulator of retrotransposon expression in mouse embryonic stem cells. *PLoS Comput. Biol.* 8:e1002486
113. 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
114. Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, et al. 2010. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* 463:237–40
115. Scott AF, Schmeckpeper BJ, Abdelrazik M, Comey CT, O'Hara B, et al. 1987. Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. *Genomics* 1:113–25
116. 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.* 35:3741–51
117. Shiloh Y. 2001. ATM (ataxia telangiectasia mutated): expanding roles in the DNA damage response and cellular homeostasis. *Biochem. Soc. Trans.* 29:661–66
118. Shukla R, Upton KR, Munoz-Lopez M, Gerhardt DJ, Fisher ME, et al. 2013. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. *Cell* 153:101–11
119. Singer T, McConnell MJ, Marchetto MC, Coufal NG, Gage FH. 2010. LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? *Trends Neurosci.* 33:345–54
120. Slotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* 8:272–85
121. Solyom S, Ewing AD, Rahrmann EP, Doucet T, Nelson HH, et al. 2012. Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res.* 22:2328–38
122. Speek M. 2001. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol. Cell. Biol.* 21:1973–85
123. Stenglein MD, Harris RS. 2006. APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. *J. Biol. Chem.* 281:16837–41
124. Stetson DB, Ko JS, Heidmann T, Medzhitov R. 2008. Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell* 134:587–98
125. Swergold GD. 1990. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol. Cell. Biol.* 10:6718–29
126. Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, et al. 2002. Human L1 retrotransposition is associated with genetic instability in vivo. *Cell* 110:327–38
127. Takasu M, Hayashi R, Maruya E, Ota M, Imura K, et al. 2007. Deletion of entire HLA-A gene accompanied by an insertion of a retrotransposon. *Tissue Antigens* 70:144–50
128. Taylor MS, Lacava J, Mita P, Molloy KR, Huang CR, et al. 2013. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. *Cell* 155:1034–48
129. Tchenio T, Casella JF, Heidmann T. 2000. Members of the SRY family regulate the human LINE retrotransposons. *Nucleic Acids Res.* 28:411–15
130. Upton KR, Baillie JK, Faulkner GJ. 2011. Is somatic retrotransposition a parasitic or symbiotic phenomenon? *Mob. Genet. Elements* 1:279–82
131. van den Hurk JA, Meij IC, Seleme MC, Kano H, Nikopoulos K, et al. 2007. L1 retrotransposition can occur early in human embryonic development. *Hum. Mol. Genet.* 16:1587–92
132. van den Hurk JA, van de Pol DJ, Wissinger B, van Driel MA, Hoefsloot LH, et al. 2003. Novel types of mutation in the choroideremia (CHM) gene: a full-length L1 insertion and an intronic mutation activating a cryptic exon. *Hum. Genet.* 113:268–75

133. Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, et al. 2001. Human L1 retrotransposition: *cis* preference versus *trans* complementation. *Mol. Cell. Biol.* 21:1429–39
134. Wheelan SJ, Aizawa Y, Han JS, Boeke JD. 2005. Gene-breaking: a new paradigm for human retrotransposon-mediated gene evolution. *Genome Res.* 15:1073–78
135. Yang N, Kazazian HH Jr. 2006. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat. Struct. Mol. Biol.* 13:763–71
136. Yang N, Zhang L, Zhang Y, Kazazian HH Jr. 2003. An important role for RUNX3 in human L1 transcription and retrotransposition. *Nucleic Acids Res.* 31:4929–40
137. Yu F, Zingler N, Schumann G, Stratling WH. 2001. Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. *Nucleic Acids Res.* 29:4493–501
138. Zhao K, Du J, Han X, Goodier JL, Li P, et al. 2013. Modulation of LINE-1 and Alu/SVA retrotransposition by Aicardi-Goutieres syndrome-related SAMHD1. *Cell Rep.* 4:1108–15

Annual Review of Genetics

Volume 48, 2014

Contents

Errata

An online log of corrections to *Annual Review of Genetics* articles may be found at http://www.annualreviews.org/errata/genet