METHODS AND RESOURCES

Cellular labeling of endogenous retrovirus replication (CLEVR) reveals de novo insertions of the gypsy retrotransposable element in cell culture and in both neurons and glial cells of aging fruit flies

Yung-Heng Chang[1], Richard M. Keegan[2], Lisa Prazak [3], Josh Dubnau [1,2]*

1 Department of Anesthesiology, Stony Brook School of Medicine, Stony Brook, New York, United States of America, 2 Program in Neuroscience, Department of Neurobiology and Behavior, Stony Brook University, Stony Brook, New York, United States of America, 3 Biology, Farmingdale State College, Farmingdale, New York, United States of America

* joshua.dubnau@stonybrook.edu

Abstract

Evidence is rapidly mounting that transposable element (TE) expression and replication may impact biology more widely than previously thought. This includes potential effects on normal physiology of somatic tissues and dysfunctional impacts in diseases associated with aging, such as cancer and neurodegeneration. Investigation of the biological impact of mobile elements in somatic cells will be greatly facilitated by the use of donor elements that are engineered to report de novo events in vivo. In multicellular organisms, reporter constructs demonstrating engineered long interspersed nuclear element (LINE-1; L1) mobilization have been in use for quite some time, and strategies similar to L1 retrotransposition reporter assays have been developed to report replication of Ty1 elements in yeast and mouse intracisternal A particle (IAP) long terminal repeat (LTR) retrotransposons in cultivated cells. We describe a novel approach termed cellular labeling of endogenous retrovirus replication (CLEVR), which reports replication of the gypsy element within specific cells in vivo in *Drosophila*. The gypsy-CLEVR reporter reveals gypsy replication both in cell culture and in individual neurons and glial cells of the aging adult fly. We also demonstrate that the gypsy-CLEVR replication rate is increased when the short interfering RNA (siRNA) silencing system is genetically disrupted. This CLEVR strategy makes use of universally conserved features of retroviruses and should be widely applicable to other LTR retrotransposons, endogenous retroviruses (ERVs), and exogenous retroviruses.

Introduction

Nearly 50% of human DNA content, and equivalently vast fractions of most other animal and plant genomes, consist of sequences derived from transposable elements (TEs) [1]. TEs are

selfish genetic elements whose primary adaptation is to copy themselves in germline tissue, thereby passing on de novo copies. Mutations generated by germline transposition are a significant source of genetic variability, with obvious impact on phenotypic diversity and on evolutionary adaptation [2, 3]. Although replication in the germline is the means by which TEs generate inheritable de novo copies, it is increasingly clear that endogenous TEs also replicate in somatic tissues. The evidence for somatic transposition is particularly strong for Class 1 TEs, also known as retrotransposable elements (RTEs) [4–21]. RTEs replicate through an RNA intermediate by the use of reverse transcription and insertion of de novo cDNA copies. Somatic replication of RTEs has the potential for tremendous impact both on normal biological properties of tissues and on human health [2, 16–21].

Historically, most studies of the mechanisms of transposition have focused on events in the germline for two main reasons. First, replication in the germline rather than somatic cells is the evolutionary adaptation that has permitted TEs to maintain their presence in the genome. Thus, from a conceptual point of view, it made great sense to study TE action there. The other reason most research focused on germline events, however, is technical. Detection of de novo replication in the germline is experimentally tractable because new insertions can be passed on to offspring, where they are present in every cell. As a result, classical molecular approaches such as Southern blots, as well as more modern genomic approaches, are each feasible means to characterize new insertions. In contrast, detection of de novo inserts in somatic tissues is far more challenging than in the germline because each new TE insertion will be present in a single cell or, at best, in one clonally related cell lineage if the insertions are followed by cell divisions. The discovery that long interspersed nuclear element (LINE)-like RTEs are able to replicate during neural development and in pluripotent stem cells in rodents [5–7, 12, 14, 22], as well as somatic cells during *Drosophila* embryonic development [15], however, provided the impetus to more closely examine TE replication in soma. It is increasingly clear that TEs also can be actively expressed and even mobile in somatic tissues such as the brain during normal development [5–7, 9–14, 23–26], during aging in a variety of cell types and species [4, 27–33], in a variety of neurodegenerative diseases and in animal models of human neurodegeneration [34–44], and in cancer [17, 45–54].

Single-cell whole genome sequencing approaches have provided the resolution to detect rare de novo insertions within individual somatic cells. As a whole, such studies have clearly supported the idea that some RTEs are capable of replication during brain development as well as in the context of cancer progression, although the rate of such de novo events and whether they have functional impacts remain controversial or unresolved [9, 10, 18, 24, 25, 46, 55–57]. But these approaches are limited in some important practical and conceptual ways. First, single-cell sequencing is relatively expensive and is low throughput. In addition, the significant molecular biological and computational challenges that come with this approach have led to widely different estimates of mobilization frequency. But perhaps more importantly, detection of individual de novo replication events in single cells does not afford a platform to conduct structure/function manipulations of RTEs to dissect the biological mechanisms. Nor does it provide the opportunity to visualize or manipulate the cells in which the RTEs are replicating within a tissue.

A different approach to this problem is to engineer genetic reporters of RTE replication [4, 5, 58, 59]. Such reporters offer the advantage that they provide the means for structure-function studies to uncover *Cis*-acting sites, investigation of impact of environmental or genetic perturbations, and to uncover the type and fate of the cells in which RTEs replicate within living tissues. Such reporters have been described for Ty long terminal repeat (LTR) retrotransposons in yeast [58], mammalian LINE-1 (L1) retrotransposons in cell culture and in transgenic mice, and for intracisternal A particle (IAP) elements in cell culture [5, 59–61].

Each of the above reporters were constructed using the insertion of heterologous introns within the engineered RTE. The intron can be removed by splicing when the RTE passes through an RNA intermediate, thereby activating a reporter upon de novo integration of replicated cDNA copies into host DNA. The L1-GFP reporter mouse has revealed that de novo mobilization events take place during mouse development [5] and provided the means to investigate the impact of behavioral and genetic perturbations [6, 7]. We previously described an indirect mobilization reporter for the gypsy-LTR retrotransposon/endogenous retrovirus (ERV). This so-called gypsy-TRAP made use of a known chromosomal hot spot for gypsy integration to turn off a Gal80 repressor upon insertion of a de novo retro-element into the hot spot cassette [4]. When present together with both a Gal4 activator and a Gal4-responsive green fluorescent protein (GFP), this was sufficient to report integrations that disrupted the Gal80 cassette. This system has been successfully used to reveal de novo transposition during aging and in the context of neurodegeneration [4, 8, 27, 28, 38]. But there are several critical shortcomings of the gypsy-TRAP approach. First, it does not selectively report insertions of any single RTE. In fact, it can be activated by insertions from any gypsy family member or indeed any RTE that shares that hot spot. Second, it does not provide the means to manipulate the biology of the donor element. Third, because the gypsy-TRAP utilizes a repressor of Gal4 function, it precludes use of the powerful Gal4 system that is the primary tool for modern *Drosophila* genetics.

Although most of the work on somatic transposition has focused on LINE elements, an emerging literature has suggested the possibility of a far broader impact of RTEs within somatic tissues, with the possibility that both LINE-like and LTR RTEs, and ERVs may mobilize [4–14, 26–32, 34–43, 45–54, 57]. But aside from the gypsy-TRAP, whose utility is limited, no in vivo reporter system has yet been developed to monitor mobilization of donor LTR retrotransposons or ERVs in any multicellular organism. Using the *Drosophila* ERV gypsy as a model, we have developed a novel reporter to reveal de novo mobilization both in cell culture and in somatic brain tissue. The gypsy–cellular labeling of endogenous retrovirus replication (CLEVR) reporter becomes hyper-active with disruption to the short interfering RNA (siRNA) silencing system, which is known to help silence endogenous gypsy expression in somatic cells. Using this reporter system, we also confirm our previous report of age-dependent increases in mobilization in neurons and demonstrate similar effects in glial cells. Because the CLEVR reporter system capitalizes on highly conserved mechanisms of replication, it likely will be generalizable to LTR retrotransposons, ERVs, and even exogenous retroviruses in a variety of animal and plant species.

Results

Engineering of the *gypsy-CLEVR* reporter

To generate a reporter capable of tracing gypsy retrotransposition in specific cell lineages, we capitalized on conserved features of LTR-RTE/ERV replication (Fig 1A). Specifically, we took advantage of two universal features of retrovirus replication [62, 63]. First, because the 5′UTR that is part of the promoter sequence at the 5′ end of the 5′LTR is not transcribed, the corresponding sequences at the 5′ end of the 3′LTR in the genomic RNA normally is used as a template to complete cDNA synthesis of the 5′ end of the full-length transcript. Second, because the portion of the U5 region of the 3′LTR following the transcription termination signal is not transcribed, the corresponding portion of the 5′LTR is used as a template for replication of this region of the 3′LTR. As a result, each of the LTR sequences of the fully replicated cDNA are hybrids of the nucleotide sequence from the two LTR repeats. The CLEVR reporter takes advantage of this. We placed a promoter without a reporter within the U5 region of the 5′LTR
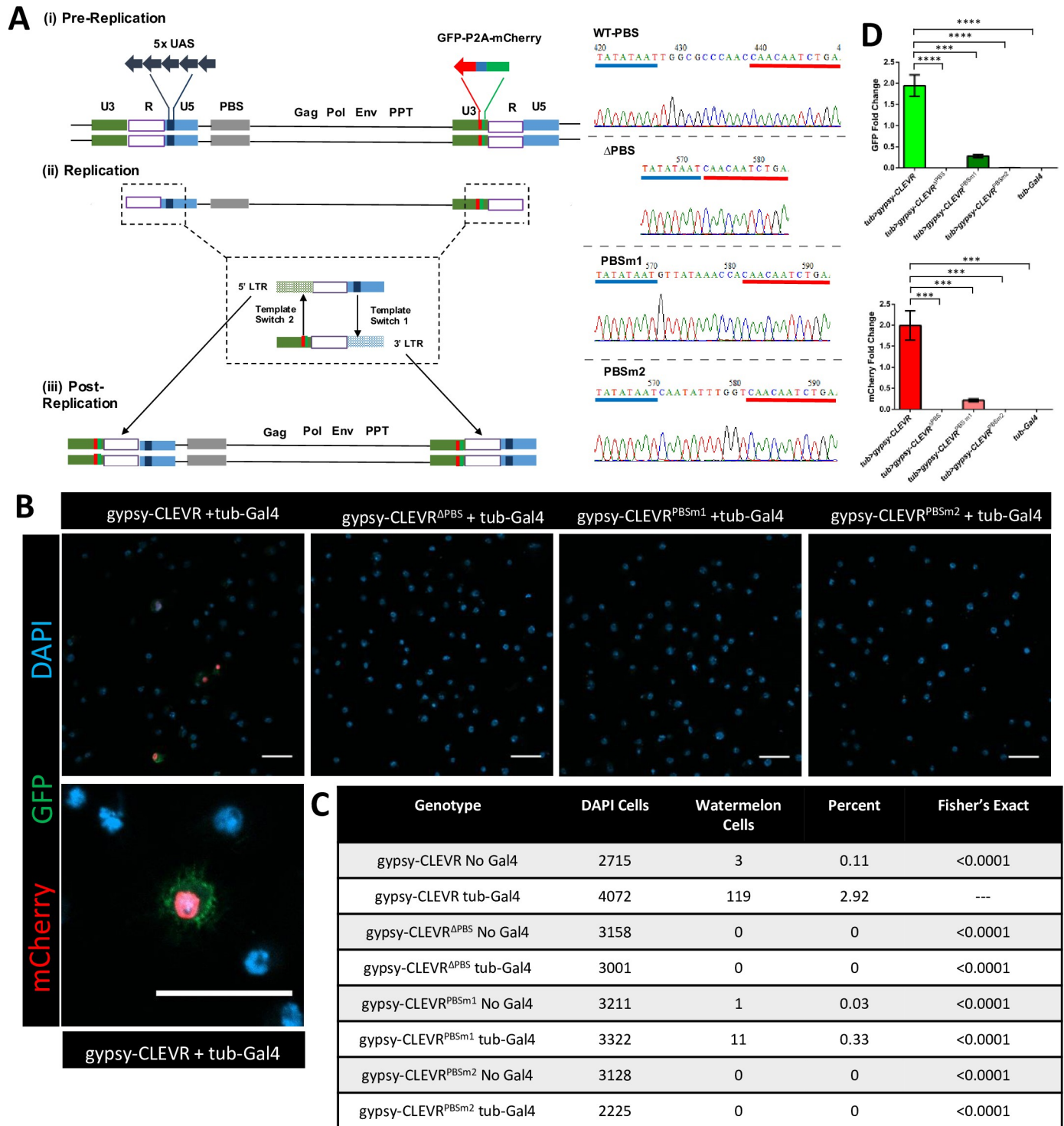
Fig 1. gypsy-CLEVR genetic structure and functional test in *Drosophila* S2 cells. (A) Schematic showing the replication cycle of the gypsy-CLEVR reporter. (i) gypsy-CLEVR contains 5XUAS and the Watermelon reporter that are inserted into the U5 region of 5′LTR and U3 region within 3′LTR, respectively, with reverse orientation to gypsy element. (ii) During normal replication of LTR-RTEs and ERVs, a series of template switches transfers DNA sequence information from between subregions of the 3′ and 5′LTRs. These two template switches cause a rearrangement in the replicated cDNA such that the WM reporter is placed downstream to the UAS regulatory element. (iii) The full-length cDNA reintegrates into the host genome at a distinct location. If that cell expresses Gal4, it will trigger expression of WM by binding to UAS regulatory element. The wild-type (WT), deletion (ΔPBS), and mutations (PBSm1 and PBSm2) of the primer binding site (PBS) variants of gypsy-CLEVR are shown, along with sequence traces confirming their fidelity. (B) Fluorescent images showing that WM positive cells are detected for gypsy-CLEVR with tub-Gal4, but not for deleted (gypsy-CLEVR^ΔPBS) or mutated (gypsy-CLEVR^PBSm1 and gypsy-CLEVR^PBSm2) gypsy-CLEVR

and we positioned a reporter with no promoter into the U3 region of the 3′LTR. Thus, after full replication, the reinserted de novo element should have rearranged the heterologous promoter so that it is adjacent to the reporter at both the 5′ and 3′LTRs.

As a heterologous promoter, we chose to use the Gal4-responsive 5XUAS enhancer, which is sufficient to drive strong expression under the control of the Gal4 transcription factor. This promoter is compact in size and provides maximal functional utility because its use taps into a wealth of extant *Drosophila* strains with Gal4 expression restricted to virtually any desired cell type or developmental stage. Importantly, the Gal4-responsive promoter faces to the left, opposite to the orientation of the gypsy promoter that lies in the LTR. Therefore, gypsy expression falls under its own control. For a reporter, we wished to provide the means to clearly mark nuclei in order to count cells and identify the location of their soma within tissues. But we also wanted to be able to see cell morphology, for which it is helpful to highlight the plasma membrane. This is particularly important in brain tissue, where glia and neurons can have extraordinarily complex morphologies. To accomplish both of these goals, we designed Watermelon (WM), a dual reporter module that encodes both a myristylated GFP that tethers to the membrane (myr-GFP-V5) and nuclear localized mCherry (H2B-mCherry-HA). The dual reporter expression relies on introduction of the porcine teschovirus-1 2A (P2A) self-cleaving peptide [64, 65] sequence (myr-GFP-V5-P2A-H2B-mCherry-HA, S1 Fig). When this dual reporter is expressed, it simultaneously marks the plasma membrane with GFP and the cell nucleus with mCherry.

To test the functionality of the dual reporter, we created an upstream activating sequence (UAS)-WM construct, which we co-transfected with actin-Gal4 into *Drosophila* S2 cells. Indeed, this marked transfected cells with brilliant green membrane fluorescence surrounding bright red fluorescence in the nuclei (S2A Fig). We also generated UAS-WM transgenic flies, which we tested by crossing with several tissue-specific Gal4 drivers. In the presence of *hedgehog* (*hh*)-*Gal4* (S2B Fig), the WM reporter marked posterior compartment cells [66] of wing discs during the larval stage with mCherry in the nucleus and GFP on the membrane. We also tested this dual reporter in the *Drosophila* adult brain by expressing this transgene in subperineurial glia (SPG) under the control of the *moody-Gal4*. Here too, the nuclei were marked by strong mCherry fluorescence that co-localized with the glial nuclear Repo marker in SPG glial cells. The SPG glial membranes were painted with strong GFP, revealing the characteristic mesh-like architecture of this cell type (S2C Fig) [67]. These results demonstrate the functionality of the WM reporter to simultaneously mark cell nuclei (red) and membranes (green).

We next incorporated the dual label into the gypsy-CLEVR design (Fig 1A) and tested the reporter in S2 cells. When the gypsy-CLEVR construct was co-transfected along with tubulin-Gal4, strong WM fluorescent labeling was readily detected in approximately 3% of cells (Fig 1B and 1C). Within each labeled cell, the nuclei were painted with mCherry (which co-localized with DAPI) and were encircled by green fluorescence on the membrane. This is consistent

with the idea that the gypsy-CLEVR reporter undergoes the rearrangement predicted during replication, such that the promoter and reporter are placed adjacent to each other in the full-length cDNA. This rearrangement would then permit Gal4-mediated expression of the WM dual reporter. As predicted, WM labeling was not detected in the absence of Gal4 (Fig 1C). To rule out the possibility that the reporter can somehow be expressed in the absence of gypsy replication, we conducted a series of further experimental manipulations.

We first tested the impact of deleting (ΔPBS) or mutating (PBSm1 and PBSm2) the primer binding site (PBS), a *Cis*-acting motif essential for retrovirus replication (Fig 1A). In contrast with the gypsy-CLEVR construct that contains the intact PBS, co-transfections of tubulin-Gal4 with the gypsy-CLEVR$^{\Delta PBS}$ did not yield any labeled cells (Fig 1B and 1C). Similarly, introduced mutations in the nucleotide sequence of the PBS greatly reduced (PBSm1; 0.33% of cells labeled) or eliminated (PBSm2; 0% of cells labeled) the reporter expression (Fig 1B and 1C). To reinforce this finding, we performed reverse-transcription quantitative PCR (RT-qPCR) to check the relative expression of GFP and mCherry signals derived from gypsy-CLEVR and its PBS modified versions (gypsy-CLEVR$^{\Delta PBS}$, gypsy-CLEVR^{PBSm1}, gypsy-CLEVR^{PBSm2}) (Fig 1D). In order to detect the relative expression levels of GFP and mCherry, qPCR primer sets targeting the GFP and mCherry regions of WM transcript were separately designed and tested. To demonstrate that these primer sets functionally detect WM reporter, mRNA extracted from cells with no vector, tub-Gal4 alone, UAS-WM alone, and tub-Gal4 combined with UAS-WM reporter were used to examine the relative expression of each of the fluorescent proteins encoded by WM. As expected, GFP and mCherry were only detected in tub-Gal4 and UAS-WM (tub>WM) co-transfected cells, but not cells with no vector, or with tub-Gal4 alone or UAS-WM alone as negative controls (S2D Fig). This result demonstrates the GFP and mCherry primer sets can robustly and specifically detect WM signals. Next, we used these GFP and mCherry primer sets to detect WM expression levels in gypsy-CLEVR and its PBS modified versions (gypsy-CLEVR$^{\Delta PBS}$, gypsy-CLEVR^{PBSm1}, and gypsy-CLEVR^{PBSm2}). We again compared reporter levels with co-transfection of the CLEVR-constructs and tub-Gal4 versus tub-Gal4 alone as a negative control. Here too, we see no detectable signal from WM in the tub-Gal4 alone group (Fig 1D). But in cells that were co-transfected with tub-Gal4 and the gypsy-CLEVR (tub>gypsy-CLEVR), both the GFP and mCherry transcripts were robustly detected. With each of these primer sets, the reporters were greatly reduced and, in fact, were barely detectable in the gypsy-CLEVR$^{\Delta PBS}$ (tub>gypsy-CLEVR$^{\Delta PBS}$) and gypsy-CLEVR^{PBSm2} (tub> gypsy-CLEVR^{PBSm2}) groups (Fig 1D). WM signals from gypsy-CLEVR^{PBSm1} (tub> gypsy-CLEVR^{PBSm1}) also were significantly reduced when compared with gypsy-CLEVR (GFP: 14.4% and mCherry: 11.1%) (Fig 1D). Notably, this quantitative effect seen with RT-qPCR for the PBSm1 mutation is consistent with the quantification of the number of cells exhibiting fluorescent reporter expression via confocal microscopy.

Molecular confirmation of *gypsy-CLEVR* retrotransposition in vivo

In order to test the efficacy and specificity of the gypsy-CLEVR reporter in vivo, we next introduced gypsy-CLEVR and its PBS modified variants (gypsy-CLEVR$^{\Delta PBS}$, gypsy-CLEVR^{PBSm1}, and gypsy-CLEVR^{PBSm2}) into transgenic flies. We first sought to establish molecular confirmation that the donor RTE of the gypsy-CLEVR reporter was capable of replication, which should lead to the predicted rearrangement of the LTR sequences. The gypsy-CLEVR with an intact PBS and the variants with deletion or mutation of the essential PBS were tested in parallel. For each case, adult flies were collected directly after eclosion and were then aged for 7 days. Genomic DNA was extracted from these flies, and nested PCR was performed to examine gypsy-CLEVR rearrangement. We did not detect a visible signal in the first round of PCR (gel region

indicated by green arrowhead in Fig 2A). The DNA content in the indicated region (predicted size is 4,039 nucleotides) of the gel was extracted and used as template to perform a second round of PCR. A PCR product of the predicted size (around 239 nucleotides) was only detected in the intact PBS gypsy-CLEVR tissue but not in non-transgenic control animals, and not in any of the three gypsy-CLEVR PBS-modifying variants (region indicated by red arrowhead in Fig 2A). This PCR product was purified, subcloned, and sequenced. The sequencing results from three independent clones confirmed the presence of the predicted rearrangement in the gypsy-CLEVR containing flies, and the sequence of this clone matched the predicted sequence expected to form after gypsy replication (S3 Fig). Together, these results provide strong molecular confirmation that the gypsy-CLEVR reporter is capable of replicating in vivo, and such replication is only detected at the PCR level in the presence of an intact PBS. We next tested whether the WM reporter could also be detected in these flies, and whether it too required an intact PBS.

Because of our previous finding that *gypsy* expression and replication is increased with age [4] and because expression may be highest in glia [35], we first tested whether we could detect gypsy-CLEVR replication in glial cells of aged flies by crossing the reporter flies with the pan-glial driver, repo-Gal4. This Gal4 strain is well characterized to express at high levels in virtually all glial cells. In order to independently mark all glial and neuronal nuclei, we used an antibody against the Repo nuclear protein that marks all glial nuclei, and an antibody against the nuclear Elav protein that marks all neuronal nuclei. As a second, independent label to unambiguously label glial nuclei, we added a UAS-driven nuclear LacZ transgene. This nuclear LacZ is driven from the same repo-Gal4 driver that activates the mCherry reporter, but unlike WM, its expression does not require replication of the gypsy-CLEVR construct. In 7-day-old animals, we detect strong mCherry fluorescence in many glial nuclei (Fig 2B and 2C), and we detect GFP membrane fluorescence surrounding the nuclear cherry (S4 Fig). By confocal imaging throughout the whole brain, we determined that most of the gypsy-CLEVR–positive cells were located near the brain surface, especially in glia near the anterior regions (S5 Fig). We quantified the total number of WM-labeled glia in anterior brain sections and found that, as was the case in cell culture, deletion or mutation of the PBS leads to a dramatic decrease in the number of labeled cells (Fig 2B and 2C), demonstrating that a *Cis*-acting motif that is essential for retrovirus replication also is required to activate the reporter. Although the replication of this construct is by design Gal4 independent, expression of the reporter is engineered to require Gal4. Indeed, we rarely detected any gypsy-CLEVR–positive cells in animals that did not contain the repo-Gal4 driver, even when these flies were aged for 30 days (S6 Fig), a time point when endogenous gypsy is expressed at high levels [4, 35]. These rare labeled cells in flies that lack Gal4 likely derive from de novo replication events that insert nearby active genomic loci, where the rearranged UAS-WM may be expressed by trapping nearby enhancers, thereby obviating the need for Gal4.

Age-dependent increase of gypsy retrotransposition in glial and neuronal cells

Our previous work using the gypsy-TRAP demonstrated an age-dependent increase in gypsy replication in neurons [4], and similar effects of age were subsequently documented with that reporter in other tissues [8, 27, 28, 38]. We therefore tested whether gypsy replication in glia was similarly age dependent and could be detected using the gypsy-CLEVR reporter. We quantified the total number of gypsy-CLEVR–positive glia in anterior brain sections and found that the number of labeled glia (32.1 ± 2.7 per anterior brain) is extremely low in young animals, just 2 days after eclosion (Fig 3A and 3B). This number greatly increases by 7 days
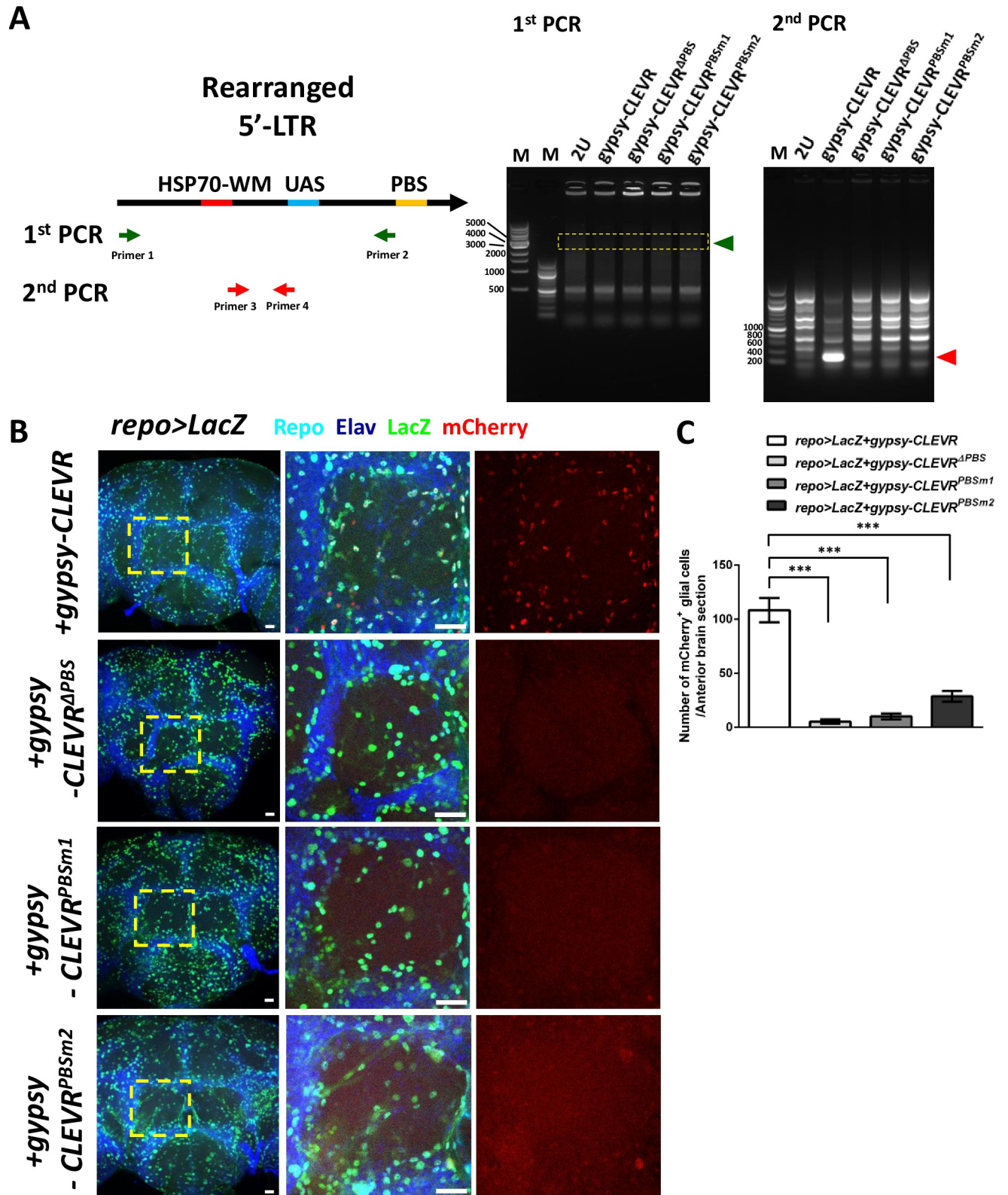
Fig 2. gypsy-CLEVR retrotransposition in glial cells in vivo. (A) Schematic illustration of LTR genetic structure after gypsy-CLEVR retrotransposition. Products of first and second rounds of nested PCR were amplified from genomic DNA extracted from 7-day-old transgenic flies containing either wild-type gypsy-CLEVR, gypsy-CLEVR$^{\Delta PBS}$, gypsy-CLEVR^{PBSm1}, or gypsy-CLEVR^{PBSm2}. Results from the first round of PCR (green primer set combining

Primer 1 with Primer 2 in schematic) do not detect a product. A second round of nested PCR (red primer set combining Primer 3 with Primer 4 in schematic) was conducted from product isolated from the highlighted gel region (green arrowhead at first-round PCR). Products of the correct size were detected (red arrowhead at second-round PCR) after this second round of PCR using nested primers. This product was only detected from flies with the intact gypsy-CLEVR transgene. PCR products were sequence verified (S3 Fig). (B) Gypsy-CLEVR replication was detected with confocal microscopy in adult brain glia by crossing flies containing the gypsy-CLEVR constructs with a strain containing the pan-glial repo-Gal4 and a UAS-nuclear-LacZ. Gypsy-CLEVR replication was detected using the mCherry nuclear signal (red) as a reporter. Glial nuclei were independently labeled both with a Repo antibody (cyan) and with an antibody against β-galactosidase (green). Neuronal nuclei were counterstained using the Elav antibody (dark blue). Optical sections from anterior brain regions are shown (left panels) for 7-day-old gypsy-CLEVR (wild-type, PBS deletion, and mutations) transgenic flies. High-magnification micrographs (from the region indicated by yellow dashed box) are shown to the right. Scale bar = 20 μm. (C) Quantification of the number of mCherry-labeled glial cells per anterior brain section (mean ± SEM) reveals that the mCherry labeling of glia requires an intact PBS (mCherry-labeled glia for each group: gypsy-CLEVR: 108.3 ± 11.2, $n = 16$; gypsy-CLEVR$^{\Delta PBS}$: 5.3 ± 2.0, $n = 11$; gypsy-CLEVR^{PBSm1}: 10.0 ± 2.7, $n = 10$; and gypsy-CLEVR^{PBSm2}: 28.7 ± 5.0, $n = 7$) (data in S3 Data). Number of labeled glial cells is significantly higher with the wild-type gypsy-CLEVR construct (***$p < 0.001$, unpaired t test). CLEVR, cellular labeling of endogenous retrovirus replication; LTR, long terminal repeat; PBS, primer binding site; UAS, upstream activating sequence; 2U, wilt type strain W(iso)CJ1.

https://doi.org/10.1371/journal.pbio.3000278.g002

(118.8 ± 18.1 per anterior brain) and further increases by 30 days (404.6 ± 16.4 per anterior brain) after eclosion (Fig 3A and 3B), consistent with the previous evidence that gypsy replication is age dependent in other cell types [4, 8, 27, 28, 38]. Moreover, this age-dependent increase in gypsy-CLEVR labeling could be blocked by introducing a Gal4-driven RNAi transgene against gypsy (UAS-gypsy-IR) [42] that we previously demonstrated was capable of interfering with the age-dependent increase in gypsy expression [35]. Addition of this UAS-gypsy-IR transgene was sufficient to block the age-dependent increase in gypsy-CLEVR–labeled glia at 2-day- (21.1 ± 3.5 per anterior brain), 7-day- (54.4 ± 5.9 per anterior brain), and 30-day- (189.6 ± 10.6 per anterior brain) old animals (Fig 3A and 3B). Thus, gypsy-CLEVR labeling requires gypsy expression as predicted.

We next tested the effects of age on gypsy-CLEVR labeling in neurons. Although our recent RNAseq experiments are consistent with the interpretation that gypsy levels are higher in glia than in neurons [35], we previously demonstrated, using the gypsy-TRAP reagent, that gypsy is able to replicate in Kenyon cells (KCs), the intrinsic neurons of the mushroom body (MB), and that the number of de novo events increases with advancing age [4]. To test whether the gypsy-CLEVR reporter could also reveal age-dependent gypsy replication in these neurons, we used the MB247-Gal4 line, which expresses Gal4 in approximately 800 KC neurons per brain hemisphere [68]. Indeed, when we crossed MB247-Gal4 to gypsy-CLEVR reporters (Fig 4), we detected rare de novo events in 2-day-old animals (Fig 4A and 4G), and the population of gypsy-CLEVR–expressing cells was significantly increased in 30-day-old animals (Fig 4D and 4G). This result is consistent with our previous findings using the gypsy-TRAP reporter [4]. Importantly, the gypsy-CLEVR reporter is more sensitive (e.g., we detect events in younger animals) and more versatile in its applications. Together with the findings described above, and previously, these results demonstrate that gypsy retrotransposon activity is increased during aging in *Drosophila* adult brains, both in glia and neurons, and our gypsy-CLEVR reporter is capable to sensitively reveal these de novo retrotransposition events in vivo.

Argonaute-2 keeps gypsy replication in check in aging neurons

Previous literature has established that endogenous siRNAs loaded onto *Drosophila* argonaute-2 (dAgo2) plays a key role in transposon silencing in somatic cells [69–72], including brain [4]. Mutations in dAgo2 lead to increased expression of a suite of TEs in various somatic tissues [69–72] and lead to precocious expression in the brains of young animals of several RTEs, including gypsy [4]. Our previous findings on the impact of dAgo2 on brain aging were consistent with the hypothesis that elevated gypsy expression could lead to an increased rate of mobilization, but we were not able to test this directly with the gypsy-TRAP due to the practicalities of moving that multicomponent system into homozygous dAgo2 mutant animals. We
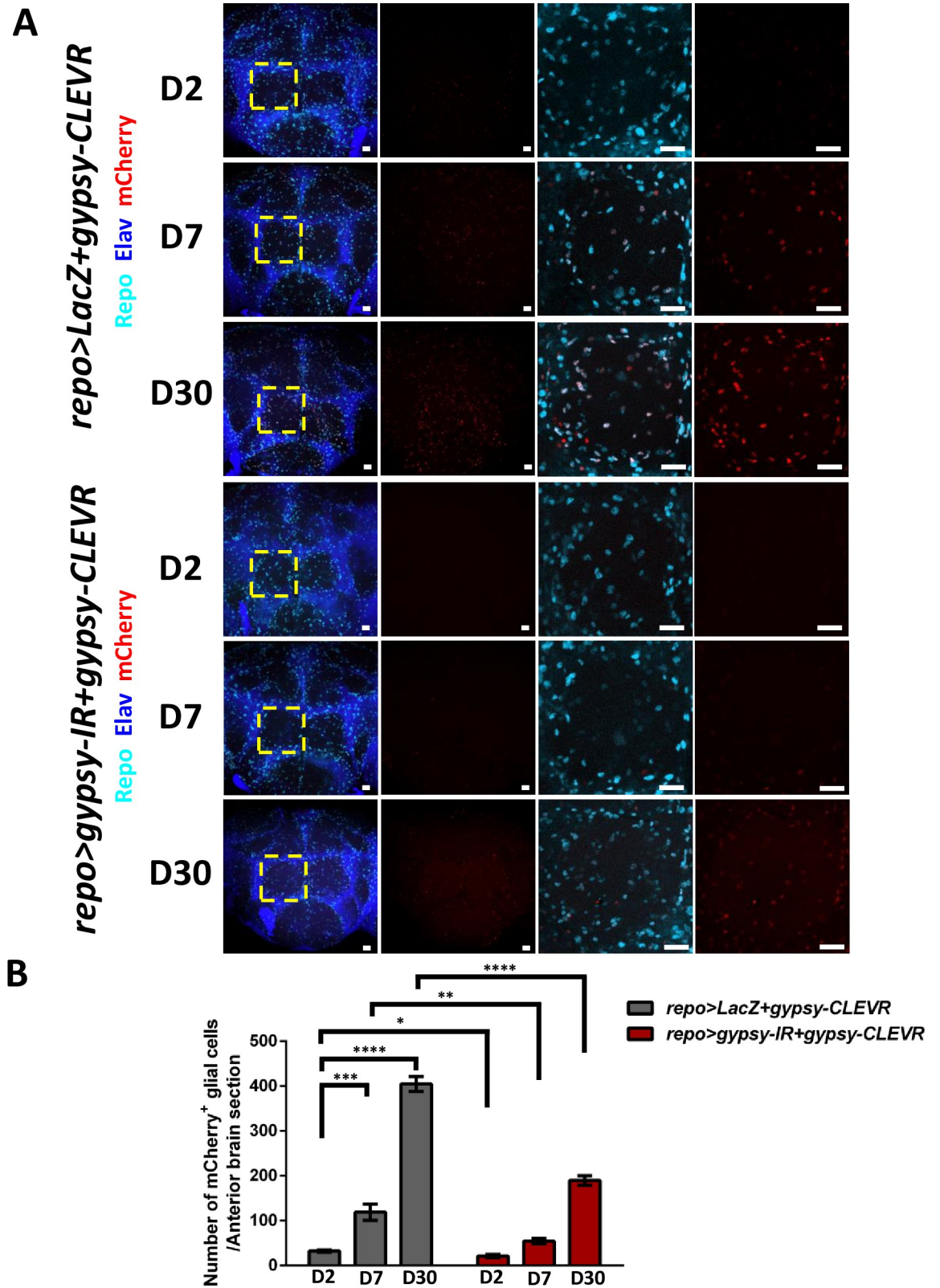
Fig 3. gypsy-CLEVR activity in glial cells is increased with age. (A) Optical sections are shown from anterior brain regions from 2-day-, 7-day-, and 30-day-old gypsy-CLEVR transgenic flies without and with a gypsy-IR transgene. Glial nuclei are independently labeled with antibodies against the glial Repo marker (cyan). Neuronal nuclei are labeled with antibodies against the Elav marker

(blue), and gypsy-CLEVR replication is labeled with mCherry (red). High-magnification images (right panels) from within the regions marked with yellow dashed box (left panels) are shown for each experimental group. Scale bar = 20 μm. (B) Numbers of mCherry-labeled glial cells (mean ± SEM) within anterior brain sections are shown for each experimental group (*repo>LacZ+gypsy-CLEVR*, D2: 32.1 ± 2.7, *n* = 8; D7: 118.8 ± 18.1, *n* = 8; D30: 404.6 ± 16.4, *n* = 10; *repo>gypsy-IR+gypsy-CLEVR*, D2: 21.1 ± 3.5, *n* = 7; D7: 54.4 ± 5.9, *n* = 11; D30: 189.6 ± 10.6, *n* = 12) (see S4 Data) (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$, ****$p < 0.0001$, n.s., not statistically significant, unpaired *t* test). CLEVR, cellular labeling of endogenous retrovirus replication; D, day; IR, inverted repeat.

https://doi.org/10.1371/journal.pbio.3000278.g003

took advantage of the simplicity and sensitivity of the gypsy-CLEVR system to test whether mutations in dAgo2 impacted gypsy replication rate per se, rather than just expression.

We used the same MB247-Gal4 driver to trace gypsy-CLEVR replication in MB KCs in *ago2* mutant animals. We find that in young flies at day 2, there is no significant increase in numbers of gypsy-CLEVR–positive KC neurons in dAgo2 mutant (*ago2$^{51B/414}$* and *ago2$^{414/414}$*) versus wild-type (Fig 4A–4C and 4G) animals. But in brains from 30-day-old animals, we detected a significant increase in numbers of gypsy-CLEVR–positive KC neurons in both the *ago2$^{51B/414}$* and the *ago2$^{414/414}$* mutant genotypes compared with wild-type controls (Fig 4D–4G). These data support the conclusion that the increase in gypsy expression levels in dAgo2 mutants leads to an increased rate of retrotransposition during aging in neurons.

Discussion

The CLEVR reporter system relies on universally conserved features of retroviral replication to activate genetic reporter expression after successful mobilization of the donor element. We demonstrate that the gypsy-CLEVR construct reports gypsy replication events both in cell culture and in several cell types in the adult *Drosophila* central nervous system. We present five separate lines of evidence that the activation of this reporter cassette is a consequence of replication of the gypsy reporter element. First, we found that three different mutations to the PBS each are sufficient to disrupt reporter activation both in cell culture and in vivo (the PBS is a universally conserved and essential *Cis*-acting site that is required to mediate priming of the first-strand cDNA synthesis by tRNA [tRNA-Lys in the case of gypsy]). This provides strong evidence that the activation of the dual reporter cannot occur in the absence of replication by some artifactual cause such as recombination. Second, we were able to verify with PCR and sequencing that the predicted replication-dependent rearrangement to place the UAS enhancer adjacent to the WM reporter actually takes place in vivo and is not detectable when the PBS is disrupted. Third, using an RNAi transgene that targets gypsy sequences, we demonstrated that activation of the reporter requires expression of gypsy. Fourth, we found that the activation of the gypsy-CLEVR reporter is age dependent in neurons and in glial cells. This dovetails with a previous literature using the gypsy-TRAP reporter, in which it was observed that gypsy de novo insertions accumulate with age in neurons [4, 38], adipose tissue [27, 28], and intestinal stem cells [8]. We further show evidence that the number of glial cells labeled by the gypsy-CLEVR reporter also increases with age. Fifth, we demonstrate using dAgo2 mutants that the gypsy-CLEVR reporter replication is inhibited by the siRNA surveillance system that normally stifles expression of endogenous gypsy elements. Taken together, these data strongly support the conclusion that CLEVR is both a sensitive and specific tool to reveal ERV replication in vivo.

The gypsy-CLEVR reporter offers significant advantages over our previously reported gypsy-TRAP. The TRAP reporter relied on using a DNA fragment containing a hot spot for gypsy family integration events, tethered to a Gal80 repressor in such a way that de novo integrations into that cassette would be likely to disrupt the Gal80 expression. In the presence of a Gal4 and a UAS reporter, it was possible to reveal de novo events [4, 8, 27, 28, 38]. While this earlier reporter has utility, the CLEVR reporter offers a number of key improvements. First,
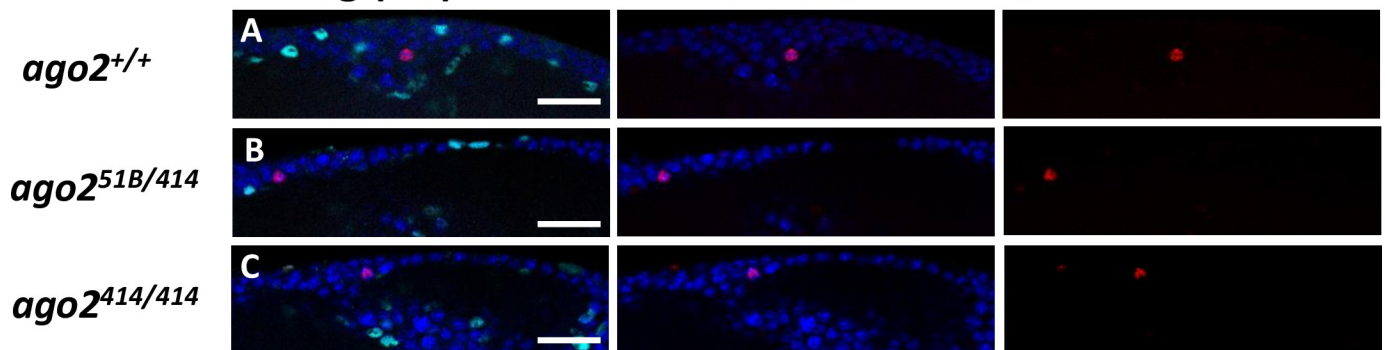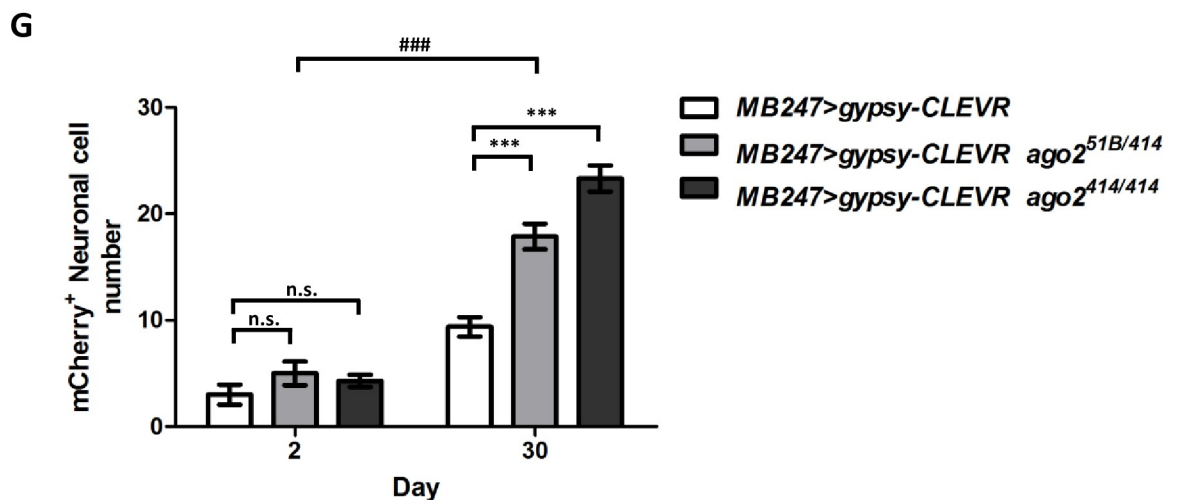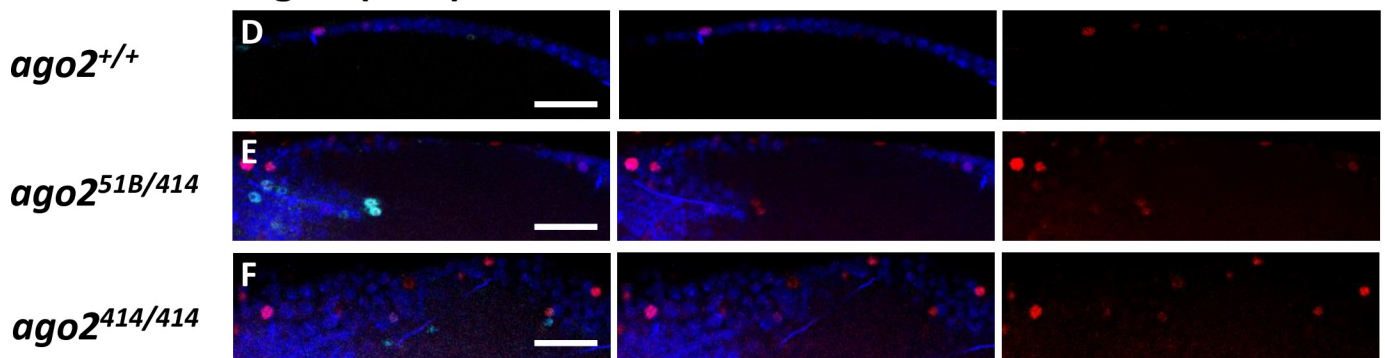
Fig 4. gypsy-CLEVR replication within KC neurons of the MB is increased with age, and with mutations in dAgo2. gypsy-CLEVR reporter was tested in flies that contain the MB247-Gal4, which expresses in approximately 800 KC neurons of the *Drosophila* MB. Replication of the gypsy-CLEVR reporter was revealed by mCherry fluorescence (red). Neuronal nuclei were independently labeled with antibodies against the Elav marker (blue), and glial nuclei were labeled with antibodies against the Repo marker (cyan). Numbers of neurons labeled by the gypsy-CLEVR reporter were significantly higher in brains from 30-day-old versus 2-day-old animals (A, D, G). Mutations in dAgo2 also significantly increased the number of mCherry-labeled nuclei in 30-day-old animals (D versuss E and F) but not in 2-day-old animals (A versus B and C). Scale bar = 20 μm. (G) Quantification of the effects of age and genotype reveals a significant effect of age and of mutations in dAgo2 (mean ± SEM shown in bar graphs). Gypsy-CLEVR–labeled cells for 2-day-old animals: wild-type: 3.8 ± 0.6, $n = 11$; $ago2^{51B/414}$: 5.0 ± 1.1, $n = 7$; $ago2^{414/414}$: 4.3 ± 0.6, $n = 7$ and for 30-day-old animals, wild-type: 9.4 ± 0.9, $n = 24$; $ago2^{51B/414}$: 17.9 ± 1.2, $n = 30$; $ago2^{414/414}$: 23.3 ± 1.2, $n = 26$) (data in S5 Data). n.s., not statistically significant (unpaired t test), $***p < 0.001$ (unpaired t test), $###p < 0.001$ (two-way ANOVA). ago2, argonaute2; CLEVR, cellular labeling of endogenous retrovirus replication; D, day; dAgo2, *Drosophila* argonaute-2; KC, Kenyon cell; MB, mushroom body.

https://doi.org/10.1371/journal.pbio.3000278.g004

the gypsy-TRAP does not distinguish which element has been inserted. In principle, any gypsy family member that shares a preference for the same hot spot can contribute labeled cells. Second, the gypsy-TRAP only captures the fraction of events that disrupts the particular cassette, and what fraction of the events are revealed is impossible to decipher. Third, the gypsy-CLEVR system is far less cumbersome technically because it does not rely on three separate components like the gypsy-TRAP. Fourth, the gypsy-TRAP precludes the use of the Gal4 system to separately manipulate other pathways; e.g., we were able to use UAS-RNAi against gypsy to demonstrate specificity, which would not be possible with the gypsy-TRAP. Fifth, the CLEVR system affords the ability to conduct structure-function studies, such as manipulations of the PBS used here.

These features of the new reporter have already permitted us to investigate new aspects of gypsy biology. For example, we were able to demonstrate that dAgo2 not only represses gypsy expression in somatic cells, as previously shown [4, 69–72], but that this has functionally relevant impact on the ability of gypsy to mobilize. This would have been difficult to test with the gypsy-TRAP because it would have necessitated combining five separate genetic components into one animal.

The CLEVR design in principle also should provide the modularity needed to flexibly interrogate a diverse suite of research questions. In our case, we chose to use the UAS enhancer in the 5′LTR because it provided us the means to query gypsy donor mobilization in different cell types and under different conditions. For example, we were able to separately query glial cells or KC neurons by selecting different Gal4 drivers. In principle, however, any relatively small promoter sequence could be used. For example, if we wanted to use the Gal4 system to separately manipulate a genetic pathway of interest and then query the mobilization of gypsy within MB KCs, we could introduce the 247–base pair fragment of the dMef gene, which is sufficient to drive Gal4 in KCs (and is the basis of the MB247-Gal4 line that we used here). CLEVR is also flexible with respect to the reporter in the 3′LTR. Here, we used a membrane GFP and a nuclear mCherry separated by the P2A sequence. It would be trivial to substitute reporters with different functional effects. For example, one could substitute GCaMP to image calcium or Channelrhodopsins for optogenetic manipulation to investigate functional correlates of gypsy mobilization in neurons. Finally, the CLEVR reporter relies on conserved features that are shared across all LTR-RTEs ERVs and exogenous retroviruses. Given the growing interest in the impact of ERVs on both normal and dysfunctional aspects of biology, the ability to follow insertion of proviruses within tissues should have palpable impact.

Materials and methods

Constructs

To generate UAS-myr-GFP-V5-P2A-H2B-mCherry-HA (referred to as Watermelon based on colors of fluorescence and abbreviated as WM), the myr-GFP and H2B-mCherry were separately amplified from pJFRC12-10xUAS-IVS-myr::GFP [73] and pEV-12xCSL-H2B-mCherry [74] vectors, and V5 and HA tags were separately added to the C termini of myr-GFP and H2B-mCherry by PCR. To synthesize the final UAS-WM, the myr-GFP-V5 and H2B-mCherry-HA products were linked with P2A sequence [64, 65] and then inserted into the MCS of pUAST with NotI and XhoI digestion (see S1 Fig for full WM sequences). The gypsy backbone was obtained as a gift from V. Salenko [75]. In order to generate gypsy-CLEVR reporter, gypsy backbone that was placed in the pCaSpeR5 transforming vector [76], and the 5xUAS regulatory element from pUAST plasmid were inserted into the BglII site of the U5 domain within gypsy 5′LTR, with antisense orientation relative to gypsy backbone. To synthesize the final gypsy-CLEVR reporter, the hsp70 promoter and the SV40 poly-A tail from

pUAST were separately fused upstream and downstream of the WM reporter, and this final chimera was also placed in antisense orientation to gypsy transcript into the XhoI site of the U3 domain within 3′LTR (see Fig 1A for gypsy-CLEVR structure in detail). The PBS sequences (5′-TGGCGCCCAAC-3′) of gypsy-CLEVR were deleted to generate the PBS deletion version (gypsy-CLEVR$^{\Delta PBS}$) and separately mutated to generate two variant constructs, gypsy-CLEVR^{PBSm1} (5′-GTTATAAACCA-3′) and gypsy-CLEVR^{PBSm2} (5′-CAATATTTGGT-3′).

Transgenic flies

All Gal4-drivers, UAS transgenes, and mutants were backcrossed to our laboratory wild-type strain, Canton-S derivative w^{1118} (*isoCJ1*), at least five generations. *MB247-Gal4*, *repo-Gal4*, *UAS-gypsy-IR*, *ago2^{414}* and *ago2^{51B}* [4, 35], *hh-Gal4* [66], and *moody-Gal4* [67] were gifts from published resources as indicated. The *UAS-nls-lacZ* was obtained from Bloomington Drosophila Stock Center. Transgenic flies of *UAS-WM* and *gypsy-CLEVR* were generated by BestGene (CA).

Immunostaining of S2 cells and fly brains

Drosophila S2 cells (R69007, Thermo Fisher Scientific) were cultured in Schneider's *Drosophila* Media (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (Thermo Fisher Scientific) and penicillin-streptomycin-glutamine (Thermo Fisher Scientific) in 75 cm^2 flasks. The actin-Gal4 [77] and tubulin-Gal4 [78] plasmids were used as in previous publications. Cells were transfected with 1 µg of each plasmid DNA with the Effectence transfection kit (Qiagen). After 48 hours transfection, cells were fixed in 4% paraformaldehyde and mounted on coverslips coated in 0.5mg/mL Concanavalin A and ProLong Diamond Antifade Mountant with DAPI (Thermo Fisher Scientific).

Fly adult brains and third instar wing discs were dissected, fixed, and immunostained as previously described [4, 77]. The LacZ primary antibody was used at 1:500 dilution (A-11132, Thermo Fisher Scientific). Repo (8D12) and Elav (7E8A10) co-staining were performed using a 1:10 dilution (Developmental Studies Hybridoma Bank). DyLight 405, Alexa Fluor 488, and Alexa Fluor 647 conjugated secondary antibodies were used at 1:100 dilution against Repo and Elav antibodies and at 1:500 against LacZ (Jackson ImmunoResearch). All S2 cell and fly brain images were acquired on a Zeiss LSM 800 confocal microscope and processed in the Zeiss ZEN software package.

Quantification of gypsy-CLEVR–positive cells in fly brain

The gypsy-CLEVR signal in every fly brain was collected through a series Z section with glial Repo and neuronal Elav co-labeling. These two markers provided landmarks of brain architecture, permitting us to define the equivalent brain sections between fly brains. The majority of gypsy-CLEVR signal from glia during aging was from the more anterior sections (as shown in S5 Fig). The average number of gypsy-CLEVR–positive cells within the anterior sections of brains from different experimental groups was used to quantify the relative change of gypsy-CLEVR reporter activity. A similar strategy was also performed to quantify KCs, but focusing on MB247-Gal4–expressing cells within the fly central brain.

RT-qPCR

Cells were transfected as above. Forty-eight hours after transfection, cells were harvested, and total RNA from each group was extracted using Trizol reagent (Invitrogen). Total RNA

extractions were then treated with RQ1 DNase (Promega), followed by reverse transcription. One microgram of the total RNA was used to synthesize cDNA via SuperScript IV VILO Master Mix (Invitrogen). TaqMan Gene Expression Master Mix was used to perform the qPCR analysis on an Applied Biosystems StepOnePlus Real Time PCR System. FAM Reporters and MGB Quencher were used in all TaqMan assays and each primer set, with their assay IDs as follows: GFP (AP324DD, forward primer: 5′-CCGGAGGGATACGTTCAAGA-3′; reverse primer: 5′- CGAGTGTGTCGCCTTCAAAT-3′; probe: 5′-CCGCTCGCGTCTTATAGTTGC CGT-3′), mCherry (APZTF7H, forward primer: 5′-CAGAGGCTGAAGCTGAAGGA-3′; reverse primer: 5′-GCTTCTTGGCCTTGTAGGTG-3′; probe: 5′-CGGCGGCCACTACGA CGCTG-3′) and Act5C (Dm02361909_s1) (Applied Biosystems). To obtain the relative fold change of GFP and mCherry from WM, values from each experimental group were normalized to those of tub-Gal4 and UAS-WM co-transfection (for S2D Fig) or tub-Gal4 with gypsy-CLEVR co-transfection (for Fig 1D).

Nested PCR and sequencing

Genomic DNA was extracted from 20 flies of wild-type, gypsy-CLEVR, and gypsy-CLEVR PBS, modifying variants at day 7 by PureLink Genomic DNA Kit (Thermo Fisher Scientific). The extracted genomic DNA was followed by two rounds of standard PCR in the nested fashion. Primer 1 (5′-ACAATGTATTGCTTCGTAGC-3′) and primer 2 (5′-AGATTGTTGGT TGGGCGCCA-3′) were used in the first-round PCR, and the extract from first-round PCR (green arrowhead in Fig 2A) was amplified by primer 3 (5′-AAACTTAGTTTTTCAATATTG-3′) and primer 4 (5′-GAGCGGAGACTCTAGCAGAT-3′). The predicted size range of the PCR product (red arrowhead in Fig 2A) from the second round of PCR was extracted from the gel and cloned by TOPO-TA cloning kit (Thermo Fisher Scientific). Sanger sequencing was performed by DNA Sequencing Facility at Stony Brook School of Medicine.

Statistical analysis

Cell culture data were analyzed using a Fisher exact test variant of the chi-squared analysis in order to obtain a p-value for significance. The statistical data from aging fly and genetic manipulation were analyzed by GraphPad Prism software. The unpaired t test was used to compare the different groups within the same time point, and two-way ANOVA was used to analyze the difference between groups with aging effect and genetic manipulation. Statistical values are presented as mean ± SEM throughout.

Supporting information

S1 Fig. Sequence information of WM transgene. Sequence of WM (myr-GFP-V5-P2A-H2B-mCherry-HA). Myr-GFP sequence was shown in green, with the V5 tag in dark green. P2A sequence is shaded with yellow. H2B-mCherry is shaded red, with HA tag in pink. Myr-GFP, myristylated-GFP; P2A, porcine teschovirus-1 2A; WM, Watermelon.
(TIF)

S2 Fig. Functional test of WM transgene in S2 cell culture and different developmental stages of different tissues. (A) *Drosophila* S2 cells were co-transfected with actin-Gal4 and UAS-WM (myr-GFP in green, H2B-mCherry in red, and DAPI in blue). Scale bar = 10 μm. (B) Wing discs were dissected from flies, with *hh-Gal4* driving *UAS-WM* (*hh>WM*) at the third instar stage (myr-GFP in green, H2B-mCherry in red and DAPI in blue). Posterior cells expressing WM within the yellow dashed box are shown in high magnification. A 3D image is shown from Z series sections of the wing disc. Scale bar = 20 μm. (C) *Moody Gal4* driving

UAS-WM in SPG (*moody>WM*). Brains were dissected from adult flies and stained with glial marker Repo (magenta). Scale bar = 20 μm. (D) An RT-qPCR approach is used to confirm the functionality of GFP and mCherry primer sets. Each value from these experimental groups was further normalized to the tub>WM group (tub-Gal4 and UAS-WM co-transfection) in order to get the relative fold change. The relative fold change of both GFP (tub>WM: 0.81 ± 0.09, tub-Gal4: undetectable, UAS-WM: 0.02 ± 0.00, No vector: undetectable) and mCherry (tub>WM: 0.96 ± 0.12, tub-Gal4: undetectable, UAS-WM: 0.03 ± 0.00, No vector: undetectable) are compared (data in S2 Data). n = 3 biological replicates (****$p < 0.0001$, unpaired t test). GFP, green fluorescent protein; *hh*, *hedgehog*; myr-GFP, myristilated-GFP; RT-qPCR, reverse-transcription quantitative PCR; SPG, subperineurial glia; tub, tubulin promoter; UAS, upstream activating sequence; WM, Watermelon.
(TIF)

S3 Fig. Sequence confirmation of gypsy-CLEVR retrotransposition. Sequencing comparison of nested PCR products from three different batches of aged gypsy-CLEVR flies, with comparison to the predicted sequence of gypsy-CLEVR rearrangement after retrotransposition. CLEVR, cellular labeling of endogenous retrovirus replication.
(TIF)

S4 Fig. The dual reporter labeling of nuclei and membrane can reveal cell morphology in vivo of cells in which gypsy-CLEVR replication has occurred. The gypsy-CLEVR was separately crossed with repo-Gal4 or MB247-Gal4. Adult brains from these crosses were dissected and labeled with glial marker (Repo in cyan) and both gypsy-CLEVR reporters, GFP (green, membrane) and mCherry (red, nuclei). Scale bar = 10 μm. CLEVR, cellular labeling of endogenous retrovirus replication; GFP, green fluorescent protein.
(TIF)

S5 Fig. Distribution of gypsy-CLEVR label in glial cells throughout the adult fly brain. Optical sections of 7-day-old adult fly are shown from anterior, central, and posterior regions. Glial nuclei are labeled with the pan-glial marker Repo (cyan) and neuronal marker Elav (blue). Glial nuclei are independently labeled with UAS-nuclear-LacZ (green). gypsy-CLEVR reporter replication is revealed with nuclear mCherry (red). Highest levels of gypsy-CLEVR replication are seen in anterior sections. Scale bar = 20 μm. CLEVR, cellular labeling of endogenous retrovirus replication; UAS, upstream activating sequence.
(TIF)

S6 Fig. Gal4 is required to activate WM reporter of gypsy-CLEVR. Gypsy-CLEVR transgenic flies that do not contain any Gal4 line were aged until 30 days, at which time high levels of gypsy expression and replication have taken place. Anterior brain sections of these aged gypsy-CLEVR adults contain few if any mCherry-labeled nuclei (red). Glial nuclei were counterstained with the Repo marker (cyan) and neuronal marker Elav (blue). Scale bar = 20 μm. CLEVR, cellular labeling of endogenous retrovirus replication; WM, Watermelon.
(TIF)

S1 Data.
(XLSX)

S2 Data.
(XLSX)

S3 Data.
(XLSX)

S4 Data.
(XLSX)

S5 Data.
(XLSX)

Acknowledgments

Author Contributions

Conceptualization: Josh Dubnau.

Formal analysis: Yung-Heng Chang, Richard M. Keegan.

Funding acquisition: Josh Dubnau.

Investigation: Yung-Heng Chang, Richard M. Keegan, Josh Dubnau.

Methodology: Yung-Heng Chang, Richard M. Keegan, Lisa Prazak, Josh Dubnau.

Project administration: Josh Dubnau.

Supervision: Josh Dubnau.

Visualization: Yung-Heng Chang, Richard M. Keegan, Josh Dubnau.

Writing – original draft: Yung-Heng Chang, Richard M. Keegan, Josh Dubnau.

Writing – review & editing: Josh Dubnau.

References

1. Huang CR, Burns KH, Boeke JD. Active transposition in genomes. Annu Rev Genet. 2012; 46:651–75. https://doi.org/10.1146/annurev-genet-110711-155616 PMID: 23145912; PubMed Central PMCID: PMCPMC3612533.

2. Babushok DV, Kazazian HH Jr. Progress in understanding the biology of the human mutagen LINE-1. Hum Mutat. 2007; 28(6):527–39. https://doi.org/10.1002/humu.20486 PMID: 17309057.

3. Kazazian HH Jr. Mobile elements: drivers of genome evolution. Science. 2004; 303(5664):1626–32. https://doi.org/10.1126/science.1089670 PMID: 15016989.

4. Li W, Prazak L, Chatterjee N, Gruninger S, Krug L, Theodorou D, et al. Activation of transposable elements during aging and neuronal decline in Drosophila. Nat Neurosci. 2013; 16(5):529–31. https://doi.org/10.1038/nn.3368 PMID: 23563579; PubMed Central PMCID: PMCPMC3821974.

5. Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature. 2005; 435(7044):903–10. https://doi.org/10.1038/nature03663 PMID: 15959507.

6. Muotri AR, Zhao C, Marchetto MC, Gage FH. Environmental influence on L1 retrotransposons in the adult hippocampus. Hippocampus. 2009; 19(10):1002–7. https://doi.org/10.1002/hipo.20564 PMID: 19771587; PubMed Central PMCID: PMCPMC2758700.

7. Muotri AR, Marchetto MC, Coufal NG, Oefner R, Yeo G, Nakashima K, et al. L1 retrotransposition in neurons is modulated by MeCP2. Nature. 2010; 468(7322):443–6. https://doi.org/10.1038/nature09544 PMID: 21085180; PubMed Central PMCID: PMCPMC3059197.

8. Sousa-Victor P, Ayyaz A, Hayashi R, Qi Y, Madden DT, Lunyak VV, et al. Piwi Is Required to Limit Exhaustion of Aging Somatic Stem Cells. Cell Rep. 2017; 20(11):2527–37. https://doi.org/10.1016/j.celrep.2017.08.059 PMID: 28903034; PubMed Central PMCID: PMCPMC5901960.

9. Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, De Sapio F, et al. Somatic retrotransposition alters the genetic landscape of the human brain. Nature. 2011; 479(7374):534–7. https://doi.org/10.1038/nature10531 PMID: 22037309; PubMed Central PMCID: PMCPMC3224101.

10. Richardson SR, Morell S, Faulkner GJ. L1 retrotransposons and somatic mosaicism in the brain. Annu Rev Genet. 2014; 48:1–27. https://doi.org/10.1146/annurev-genet-120213-092412 PMID: 25036377.

11. Upton KR, Gerhardt DJ, Jesuadian JS, Richardson SR, Sanchez-Luque FJ, Bodea GO, et al. Ubiquitous L1 mosaicism in hippocampal neurons. Cell. 2015; 161(2):228–39. https://doi.org/10.1016/j.cell.2015.03.026 PMID: 25860606; PubMed Central PMCID: PMCPMC4398972.

12. Klawitter S, Fuchs NV, Upton KR, Munoz-Lopez M, Shukla R, Wang J, et al. Reprogramming triggers endogenous L1 and Alu retrotransposition in human induced pluripotent stem cells. Nat Commun. 2016; 7:10286. https://doi.org/10.1038/ncomms10286 PMID: 26743714; PubMed Central PMCID: PMCPMC4729875.

13. Kubo S, Seleme MC, Soifer HS, Perez JL, Moran JV, Kazazian HH Jr., et al. L1 retrotransposition in nondividing and primary human somatic cells. Proc Natl Acad Sci U S A. 2006; 103(21):8036–41. https://doi.org/10.1073/pnas.0601954103 PMID: 16698926; PubMed Central PMCID: PMCPMC1472425.

14. Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, et al. L1 retrotransposition in human neural progenitor cells. Nature. 2009; 460(7259):1127–31. https://doi.org/10.1038/nature08248 PMID: 19657334; PubMed Central PMCID: PMCPMC2909034.

15. Eickbush MT, Eickbush TH. Retrotransposition of R2 elements in somatic nuclei during the early development of Drosophila. Mob DNA. 2011; 2(1):11. https://doi.org/10.1186/1759-8753-2-11 PMID: 21958913; PubMed Central PMCID: PMCPMC3190326.

16. Bedrosian TA, Linker S, Gage FH. Environment-driven somatic mosaicism in brain disorders. Genome Med. 2016; 8(1):58. https://doi.org/10.1186/s13073-016-0317-9 PMID: 27215330; PubMed Central PMCID: PMCPMC4877941.

17. Burns KH. Transposable elements in cancer. Nat Rev Cancer. 2017; 17(7):415–24. https://doi.org/10.1038/nrc.2017.35 PMID: 28642606.

18. Carreira PE, Richardson SR, Faulkner GJ. L1 retrotransposons, cancer stem cells and oncogenesis. FEBS J. 2014; 281(1):63–73. https://doi.org/10.1111/febs.12601 PMID: 24286172; PubMed Central PMCID: PMCPMC4160015.

19. Dubnau J. The Retrotransposon storm and the dangers of a Collyer's genome. Curr Opin Genet Dev. 2018; 49:95–105. https://doi.org/10.1016/j.gde.2018.04.004 PMID: 29705598; PubMed Central PMCID: PMCPMC5975205.

20. Kazazian HH Jr., Moran JV. Mobile DNA in Health and Disease. N Engl J Med. 2017; 377(4):361–70. https://doi.org/10.1056/NEJMra1510092 PMID: 28745987; PubMed Central PMCID: PMCPMC5980640.

21. Reilly MT, Faulkner GJ, Dubnau J, Ponomarev I, Gage FH. The role of transposable elements in health and diseases of the central nervous system. J Neurosci. 2013; 33(45):17577–86. https://doi.org/10.1523/JNEUROSCI.3369-13.2013 PMID: 24198348; PubMed Central PMCID: PMCPMC3818539.

22. Macia A, Widmann TJ, Heras SR, Ayllon V, Sanchez L, Benkaddour-Boumzaouad M, et al. Engineered LINE-1 retrotransposition in nondividing human neurons. Genome Res. 2017; 27(3):335–48. https://doi.org/10.1101/gr.206805.116 PMID: 27965292; PubMed Central PMCID: PMCPMC5340962.

23. Denli AM, Narvaiza I, Kerman BE, Pena M, Benner C, Marchetto MC, et al. Primate-specific ORF0 contributes to retrotransposon-mediated diversity. Cell. 2015; 163(3):583–93. https://doi.org/10.1016/j.cell.2015.09.025 PMID: 26496605.

24. Evrony GD, Cai X, Lee E, Hills LB, Elhosary PC, Lehmann HS, et al. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. Cell. 2012; 151(3):483–96. https://doi.org/10.1016/j.cell.2012.09.035 PMID: 23101622; PubMed Central PMCID: PMCPMC3567441.

25. Lodato MA, Woodworth MB, Lee S, Evrony GD, Mehta BK, Karger A, et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. Science. 2015; 350(6256):94–8. https://doi.org/10.1126/science.aab1785 PMID: 26430121; PubMed Central PMCID: PMCPMC4664477.

26. Hazen JL, Faust GG, Rodriguez AR, Ferguson WC, Shumilina S, Clark RA, et al. The Complete Genome Sequences, Unique Mutational Spectra, and Developmental Potency of Adult Neurons Revealed by Cloning. Neuron. 2016; 89(6):1223–36. https://doi.org/10.1016/j.neuron.2016.02.004 PMID: 26948891; PubMed Central PMCID: PMCPMC4795965.

27. Wood JG, Jones BC, Jiang N, Chang C, Hosier S, Wickremesinghe P, et al. Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in Drosophila. Proc Natl Acad Sci U S A. 2016; 113(40):11277–82. https://doi.org/10.1073/pnas.1604621113 PMID: 27621458; PubMed Central PMCID: PMCPMC5056045.

28. Jones BC, Wood JG, Chang C, Tam AD, Franklin MJ, Siegel ER, et al. A somatic piRNA pathway in the Drosophila fat body ensures metabolic homeostasis and normal lifespan. Nat Commun. 2016; 7:13856. https://doi.org/10.1038/ncomms13856 PMID: 28000665; PubMed Central PMCID: PMCPMC5187580.

29. Maxwell PH, Burhans WC, Curcio MJ. Retrotransposition is associated with genome instability during chronological aging. Proc Natl Acad Sci U S A. 2011; 108(51):20376–81. https://doi.org/10.1073/pnas.1100271108 PMID: 22021441; PubMed Central PMCID: PMCPMC3251071.

30. Maxwell PH. What might retrotransposons teach us about aging? Curr Genet. 2016; 62(2):277–82. https://doi.org/10.1007/s00294-015-0538-2 PMID: 26581630; PubMed Central PMCID: PMCPMC5120397.

31. De Cecco M, Criscione SW, Peterson AL, Neretti N, Sedivy JM, Kreiling JA. Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. Aging (Albany NY). 2013; 5(12):867–83. https://doi.org/10.18632/aging.100621 PMID: 24323947; PubMed Central PMCID: PMCPMC3883704.

32. Savva YA, Jepson JE, Chang YJ, Whitaker R, Jones BC, St Laurent G, et al. RNA editing regulates transposon-mediated heterochromatic gene silencing. Nat Commun. 2013; 4:2745. https://doi.org/10.1038/ncomms3745 PMID: 24201902; PubMed Central PMCID: PMCPMC3992701.

33. De Cecco M, Ito T, Petrashen AP, Elias AE, Skvir NJ, Criscione SW, et al. L1 drives IFN in senescent cells and promotes age-associated inflammation. Nature. 2019; 566(7742):73–8. https://doi.org/10.1038/s41586-018-0784-9 PMID: 30728521.

34. Douville R, Liu J, Rothstein J, Nath A. Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis. Ann Neurol. 2011; 69(1):141–51. https://doi.org/10.1002/ana.22149 PMID: 21280084; PubMed Central PMCID: PMCPMC3052883.

35. Krug L, Chatterjee N, Borges-Monroy R, Hearn S, Liao WW, Morrill K, et al. Retrotransposon activation contributes to neurodegeneration in a Drosophila TDP-43 model of ALS. PLoS Genet. 2017; 13(3): e1006635. https://doi.org/10.1371/journal.pgen.1006635 PMID: 28301478; PubMed Central PMCID: PMCPMC5354250.

36. Li W, Jin Y, Prazak L, Hammell M, Dubnau J. Transposable elements in TDP-43-mediated neurodegenerative disorders. PLoS ONE. 2012; 7(9):e44099. https://doi.org/10.1371/journal.pone.0044099 PMID: 22957047; PubMed Central PMCID: PMCPMC3434193.

37. Guo C, Jeong HH, Hsieh YC, Klein HU, Bennett DA, De Jager PL, et al. Tau Activates Transposable Elements in Alzheimer's Disease. Cell Rep. 2018; 23(10):2874–80. https://doi.org/10.1016/j.celrep.2018.05.004 PMID: 29874575.

38. Sun W, Samimi H, Gamez M, Zare H, Frost B. Pathogenic tau-induced piRNA depletion promotes neuronal death through transposable element dysregulation in neurodegenerative tauopathies. Nat Neurosci. 2018; 21(8):1038–48. https://doi.org/10.1038/s41593-018-0194-1 PMID: 30038280; PubMed Central PMCID: PMCPMC6095477.

39. Li W, Lee MH, Henderson L, Tyagi R, Bachani M, Steiner J, et al. Human endogenous retrovirus-K contributes to motor neuron disease. Sci Transl Med. 2015; 7(307):307ra153. https://doi.org/10.1126/scitranslmed.aac8201 PMID: 26424568.

40. Kaneko H, Dridi S, Tarallo V, Gelfand BD, Fowler BJ, Cho WG, et al. DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. Nature. 2011; 471(7338):325–30. https://doi.org/10.1038/nature09830 PMID: 21297615; PubMed Central PMCID: PMCPMC3077055.

41. Dewannieux M, Harper F, Richaud A, Letzelter C, Ribet D, Pierron G, et al. Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. Genome Res. 2006; 16 (12):1548–56. https://doi.org/10.1101/gr.5565706 PMID: 17077319; PubMed Central PMCID: PMCPMC1665638.

42. Tan H, Qurashi A, Poidevin M, Nelson DL, Li H, Jin P. Retrotransposon activation contributes to fragile X premutation rCGG-mediated neurodegeneration. Hum Mol Genet. 2012; 21(1):57–65. https://doi.org/10.1093/hmg/ddr437 PMID: 21940752; PubMed Central PMCID: PMCPMC3235010.

43. Blaudin de The FX, Rekaik H, Peze-Heidsieck E, Massiani-Beaudoin O, Joshi RL, Fuchs J, et al. Engrailed homeoprotein blocks degeneration in adult dopaminergic neurons through LINE-1 repression. EMBO J. 2018; 37(15). https://doi.org/10.15252/embj.201797374 PMID: 29941661; PubMed Central PMCID: PMCPMC6068427.

44. Tam OH, Rozhkov NV, Shaw R, Kim D, Hubbard I, Fennessey S, et al. Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia. bioRxiv. 2019:574509. https://doi.org/10.1101/574509

45. Wylie A, Jones AE, D'Brot A, Lu WJ, Kurtz P, Moran JV, et al. p53 genes function to restrain mobile elements. Genes Dev. 2016; 30(1):64–77. https://doi.org/10.1101/gad.266098.115 PMID: 26701264; PubMed Central PMCID: PMCPMC4701979.

46. Yang N, Kazazian HH, Jr. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat Struct Mol Biol. 2006; 13(9):763–71. https://doi.org/10.1038/nsmb1141 PMID: 16936727.

47. Doucet-O'Hare TT, Rodic N, Sharma R, Darbari I, Abril G, Choi JA, et al. LINE-1 expression and retrotransposition in Barrett's esophagus and esophageal carcinoma. Proc Natl Acad Sci U S A. 2015; 112 (35):E4894–900. https://doi.org/10.1073/pnas.1502474112 PMID: 26283398; PubMed Central PMCID: PMCPMC4568228.

48. Tang Z, Steranka JP, Ma S, Grivainis M, Rodic N, Huang CR, et al. Human transposon insertion profiling: Analysis, visualization and identification of somatic LINE-1 insertions in ovarian cancer. Proc Natl Acad Sci U S A. 2017; 114(5):E733–E40. https://doi.org/10.1073/pnas.1619797114 PMID: 28096347; PubMed Central PMCID: PMCPMC5293032.

49. Shukla R, Upton KR, Munoz-Lopez M, Gerhardt DJ, Fisher ME, Nguyen T, et al. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. Cell. 2013; 153(1):101–11. https://doi.org/10.1016/j.cell.2013.02.032 PMID: 23540693; PubMed Central PMCID: PMCPMC3898742.

50. Carreira PE, Ewing AD, Li G, Schauer SN, Upton KR, Fagg AC, et al. Evidence for L1-associated DNA rearrangements and negligible L1 retrotransposition in glioblastoma multiforme. Mob DNA. 2016; 7:21. https://doi.org/10.1186/s13100-016-0076-6 PMID: 27843499; PubMed Central PMCID: PMCPMC5105311.

51. Nguyen THM, Carreira PE, Sanchez-Luque FJ, Schauer SN, Fagg AC, Richardson SR, et al. L1 Retrotransposon Heterogeneity in Ovarian Tumor Cell Evolution. Cell Rep. 2018; 23(13):3730–40. https://doi.org/10.1016/j.celrep.2018.05.090 PMID: 29949758.

52. Schauer SN, Carreira PE, Shukla R, Gerhardt DJ, Gerdes P, Sanchez-Luque FJ, et al. L1 retrotransposition is a common feature of mammalian hepatocarcinogenesis. Genome Res. 2018; 28(5):639–53. https://doi.org/10.1101/gr.226993.117 PMID: 29643204; PubMed Central PMCID: PMCPMC5932605.

53. Solyom S, Ewing AD, Rahrmann EP, Doucet T, Nelson HH, Burns MB, et al. Extensive somatic L1 retrotransposition in colorectal tumors. Genome Res. 2012; 22(12):2328–38. https://doi.org/10.1101/gr.145235.112 PMID: 22968929; PubMed Central PMCID: PMCPMC3514663.

54. Doucet-O'Hare TT, Sharma R, Rodic N, Anders RA, Burns KH, Kazazian HH Jr. Somatically Acquired LINE-1 Insertions in Normal Esophagus Undergo Clonal Expansion in Esophageal Squamous Cell Carcinoma. Hum Mutat. 2016; 37(9):942–54. https://doi.org/10.1002/humu.23027 PMID: 27319353; PubMed Central PMCID: PMCPMC5548391.

55. Treiber CD, Waddell S. Resolving the prevalence of somatic transposition in Drosophila. Elife. 2017; 6. https://doi.org/10.7554/eLife.28297 PMID: 28742021; PubMed Central PMCID: PMCPMC5553932.

56. Bodea GO, McKelvey EGZ, Faulkner GJ. Retrotransposon-induced mosaicism in the neural genome. Open Biol. 2018; 8(7). https://doi.org/10.1098/rsob.180074 PMID: 30021882; PubMed Central PMCID: PMCPMC6070720.

57. Perrat PN, DasGupta S, Wang J, Theurkauf W, Weng Z, Rosbash M, et al. Transposition-driven genomic heterogeneity in the Drosophila brain. Science. 2013; 340(6128):91–5. https://doi.org/10.1126/science.1231965 PMID: 23559253; PubMed Central PMCID: PMCPMC3887341.

58. Boeke JD, Garfinkel DJ, Styles CA, Fink GR. Ty elements transpose through an RNA intermediate. Cell. 1985; 40(3):491–500. PMID: 2982495.

59. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH Jr. High frequency retrotransposition in cultured mammalian cells. Cell. 1996; 87(5):917–27. PMID: 8945518.

60. Heidmann O, Heidmann T. Retrotransposition of a mouse IAP sequence tagged with an indicator gene. Cell. 1991; 64(1):159–70. PMID: 1846087.

61. Ostertag EM, Prak ET, DeBerardinis RJ, Moran JV, Kazazian HH Jr. Determination of L1 retrotransposition kinetics in cultured cells. Nucleic Acids Res. 2000; 28(6):1418–23. https://doi.org/10.1093/nar/28.6.1418 PMID: 10684937; PubMed Central PMCID: PMCPMC111040.

62. Stoye JP. Studies of endogenous retroviruses reveal a continuing evolutionary saga. Nat Rev Microbiol. 2012; 10(6):395–406. https://doi.org/10.1038/nrmicro2783 PMID: 22565131.

63. Volkman HE, Stetson DB. The enemy within: endogenous retroelements and autoimmune disease. Nat Immunol. 2014; 15(5):415–22. https://doi.org/10.1038/ni.2872 PMID: 24747712; PubMed Central PMCID: PMCPMC4131434.

64. Ryan MD, Flint M. Virus-encoded proteinases of the picornavirus super-group. J Gen Virol. 1997; 78 (Pt 4):699–723. https://doi.org/10.1099/0022-1317-78-4-699 PMID: 9129643.

65. Szymczak AL, Workman CJ, Wang Y, Vignali KM, Dilioglou S, Vanin EF, et al. Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. Nat Biotechnol. 2004; 22(5):589–94. https://doi.org/10.1038/nbt957 PMID: 15064769.

66. Zhang L, Ren F, Zhang Q, Chen Y, Wang B, Jiang J. The TEAD/TEF family of transcription factor Scalloped mediates Hippo signaling in organ size control. Dev Cell. 2008; 14(3):377–87. https://doi.org/10.1016/j.devcel.2008.01.006 PMID: 18258485; PubMed Central PMCID: PMCPMC2292673.

67. Schwabe T, Bainton RJ, Fetter RD, Heberlein U, Gaul U. GPCR signaling is required for blood-brain barrier formation in drosophila. Cell. 2005; 123(1):133–44. https://doi.org/10.1016/j.cell.2005.08.037 PMID: 16213218.

68. Schwaerzel M, Heisenberg M, Zars T. Extinction antagonizes olfactory memory at the subcellular level. Neuron. 2002; 35(5):951–60. PMID: 12372288.

69. Chung WJ, Okamura K, Martin R, Lai EC. Endogenous RNA interference provides a somatic defense against Drosophila transposons. Curr Biol. 2008; 18(11):795–802. https://doi.org/10.1016/j.cub.2008.05.006 PMID: 18501606; PubMed Central PMCID: PMCPMC2812477.

70. Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, et al. An endogenous small interfering RNA pathway in Drosophila. Nature. 2008; 453(7196):798–802. https://doi.org/10.1038/nature07007 PMID: 18463631; PubMed Central PMCID: PMCPMC2895258.

71. Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, et al. Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science. 2008; 320(5879):1077–81. https://doi.org/10.1126/science.1157396 PMID: 18403677; PubMed Central PMCID: PMCPMC2953241.

72. Kawamura Y, Saito K, Kin T, Ono Y, Asai K, Sunohara T, et al. Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature. 2008; 453(7196):793–7. https://doi.org/10.1038/nature06938 PMID: 18463636.

73. Pfeiffer BD, Ngo TT, Hibbard KL, Murphy C, Jenett A, Truman JW, et al. Refinement of tools for targeted gene expression in Drosophila. Genetics. 2010; 186(2):735–55. https://doi.org/10.1534/genetics.110.119917 PMID: 20697123; PubMed Central PMCID: PMCPMC2942869.

74. Sprinzak D, Lakhanpal A, Lebon L, Santat LA, Fontes ME, Anderson GA, et al. Cis-interactions between Notch and Delta generate mutually exclusive signalling states. Nature. 2010; 465(7294):86–90. https://doi.org/10.1038/nature08959 PMID: 20418862; PubMed Central PMCID: PMCPMC2886601.

75. Salenko VB, Kotnova AP, Karpova NN, Lyubomirskaya NV, Ilyin YV. Polymorphism of full-length copies of the transposable element MDG4 (gypsy) cloned from the Drosophila melanogaster strain G32. Dokl Biochem Biophys. 2007; 412:29–32. PMID: 17506349.

76. Le T, Yu M, Williams B, Goel S, Paul SM, Beitel GJ. CaSpeR5, a family of Drosophila transgenesis and shuttle vectors with improved multiple cloning sites. Biotechniques. 2007; 42(2):164, 6. https://doi.org/10.2144/000112386 PMID: 17373479; PubMed Central PMCID: PMCPMC1955475.

77. Chang YH, Sun YH. Carrier of Wingless (Cow), a secreted heparan sulfate proteoglycan, promotes extracellular transport of Wingless. PLoS ONE. 2014; 9(10):e111573. https://doi.org/10.1371/journal.pone.0111573 PMID: 25360738; PubMed Central PMCID: PMCPMC4216105.

78. Galletti M, Riccardo S, Parisi F, Lora C, Saqcena MK, Rivas L, et al. Identification of domains responsible for ubiquitin-dependent degradation of dMyc by glycogen synthase kinase 3beta and casein kinase 1 kinases. Mol Cell Biol. 2009; 29(12):3424–34. https://doi.org/10.1128/MCB.01535-08 PMID: 19364825; PubMed Central PMCID: PMCPMC2698736.