

Article

Large Deletions, Cleavage of the Telomeric Repeat Sequence, and Reverse Transcriptase-Mediated DNA Damage Response Associated with Long Interspersed Element-1 ORF2p Enzymatic Activities

Kristine J. Kines ^{1,†}, Mark Sokolowski ^{1,†}, Cecily DeFreece ², Afzaal Shareef ¹ , Dawn L. deHaro ¹ and Victoria P. Belancio ^{1,*}

¹ Department of Structural and Cellular Biology, Tulane School of Medicine, Tulane Cancer Center, New Orleans, LA 70112, USA

² Department of Biology, Xavier University of Louisiana, New Orleans, LA 70125, USA

* Correspondence: vperpe@tulane.edu

† These authors contributed equally to this work.

Abstract: L1 elements can cause DNA damage and genomic variation via retrotransposition and the generation of endonuclease-dependent DNA breaks. These processes require L1 ORF2p protein that contains an endonuclease domain, which cuts genomic DNA, and a reverse transcriptase domain, which synthesizes cDNA. The complete impact of L1 enzymatic activities on genome stability and cellular function remains understudied, and the spectrum of L1-induced mutations, other than L1 insertions, is mostly unknown. Using an inducible system, we demonstrate that an ORF2p containing functional reverse transcriptase is sufficient to elicit DNA damage response even in the absence of the functional endonuclease. Using a TK/Neo reporter system that captures misrepaired DNA breaks, we demonstrate that L1 expression results in large genomic deletions that lack any signatures of L1 involvement. Using an in vitro cleavage assay, we demonstrate that L1 endonuclease efficiently cuts telomeric repeat sequences. These findings support that L1 could be an unrecognized source of disease-promoting genomic deletions, telomere dysfunction, and an underappreciated source of chronic RT-mediated DNA damage response in mammalian cells. Our findings expand the spectrum of biological processes that can be triggered by functional and nonfunctional L1s, which have impactful evolutionary- and health-relevant consequences.

Keywords: LINE-1; ORF2p; retroelements; DNA damage; H2AX; telomeres; genomic instability



Citation: Kines, K.J.; Sokolowski, M.; DeFreece, C.; Shareef, A.; deHaro, D.L.; Belancio, V.P. Large Deletions, Cleavage of the Telomeric Repeat Sequence, and Reverse Transcriptase-Mediated DNA Damage Response Associated with Long Interspersed Element-1 ORF2p Enzymatic Activities. *Genes* **2024**, *15*, 143. <https://doi.org/10.3390/genes15020143>

Academic Editors: Mark A. Batzer and David A. Ray

Received: 22 December 2023

Revised: 17 January 2024

Accepted: 18 January 2024

Published: 23 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Long Interspersed Element-1 (L1) retrotransposons are currently active in most mammalian genomes [1–7]. L1 expression is reported to occur in the germ line, during embryogenesis, and in most adult somatic tissues [8–12]. The significant upregulation of L1 mRNA and protein expression, as well as retrotransposition, has been observed in human tumors and induced pluripotent stem cells (iPSCs) [7,13–20]. This increase is attributed to the reactivation of the L1 promoter associated with genome-wide hypomethylation and epigenetic reprogramming [21–24]. The epigenetic control of L1 promoter activity is likely one of the reasons underlying cell type-, organ-, sex-, and age-specific differences in the spectrum of endogenous L1 loci expressed in human, mouse, and rat organs [9,25–27]. This is further supported by a comparative analysis of expressed and silent L1 loci in human cell lines, which identified specific epigenetic marks associated with expressed L1 loci and the requirement of long-range interactions with enhancers for optimal L1 promoter function [28,29]. These analyses have also established that both retrotransposition-competent and -incompetent endogenous L1 loci are expressed [9,25–27]. Studies of retrotransposition-competent L1 loci have dominated the field from the time of its conception, with only a

handful of studies recognizing the potential of expressed retrotransposition-incompetent L1 loci to impact host genome stability or cell function [30–34].

A full-length L1 locus, harboring an intact promoter in its 5'UTR, generates L1 mRNAs containing two open reading frames 1 and 2 (ORF1 and ORF2) (Figure 1A) (reviewed in [35,36]). L1 mRNA translation results in the production of ORF1 and ORF2 proteins (ORF1p and ORF2p), which are required for L1 retrotransposition [37–48]. ORF1p is a structural protein that performs obligatory functions involving trimerization, RNA binding, nucleic acid chaperon function [38,48–50], and contains amino acids involved in its nuclear localization [51]. ORF2p is a multifunctional protein containing endonuclease (EN) and reverse transcriptase (RT) enzymatic activities [42,43,52–56]. ORF2p also contains Cryptic, Z, and C-terminal regions, which do not possess any known enzymatic activities but are required for ORF2p function in retrotransposition [39,40,53,57,58]. The EN domain is structurally and functionally related to the cellular APE1 protein, which is an endonuclease that is essential for DNA repair via the BER pathway [43]. Based on the *in vitro* studies of the prototypic R2 element and L1 ORF2p, the EN domain of the ORF2p is responsible for the generation of DNA nicks used by the RT domain to initiate cDNA synthesis leading to the production of *de novo* L1 inserts [53,59–62]. ORF2p-dependent cDNA synthesis requires reverse transcriptase, Cryptic, Z, and C-terminal regions [40,42,47,53]. The process of cDNA synthesis by ORF2p has been linked to the induction of inflammatory response through generation of cytoplasmic L1 cDNA [63].

L1 contributes to genomic instability via retrotransposition and the generation of DNA double-strand breaks (DSBs) [64,65]. The outcome of L1 retrotransposition is *de novo* L1 inserts, some of which are mutagenic because they disrupt gene function or cause large deletions at their integration site [7,36,64,66,67] (Figure 1B). While efficient L1 retrotransposition relies on the function of both the endonuclease and the reverse transcriptase domains, experimental evidence shows that the two domains can function independently of each other and even outside of the context of the full-length ORF2p molecule [30,31,40]. When purified from bacteria, the EN domain cleaves DNA substrates *in vitro* and, when expressed in mammalian cells, causes significant DNA damage [32,43,68] (Figure 1B). The endonuclease-independent retrotransposition of defective L1 elements in mammalian telomeres [30] and in some DNA repair-deficient cells [31] has been reported (Figure 1B), demonstrating that the RT domain can function in the context of the ORF2p containing nonfunctional EN domain. Additionally, ORF2p deficient in EN function or completely lacking the EN domain supports residual levels of Alu retrotransposition in cell culture-based assays [40].

The involvement of the EN domain in causing DNA breaks and eliciting cellular DNA damage response (DDR) has been reported using the detection of phosphorylated H2AX, COMET, or toxicity assays [32,65,69,70]. The histone H2AX is a fairly common component of the chromatin of mammalian cells, with about 1 in 40 to 1 in 4 histones being H2AX, depending on the cell type [71]. Following a double-stranded break (DSB) in the DNA, H2AX phosphorylation occurs rapidly and spreads around the break site [71–74]. H2AX phosphorylation is used in clinical (reviewed in [75]) and *in vitro* (reviewed in [76]) studies to characterize the presence of DSBs. It is well established that L1 ORF2p and the endonuclease domain can cause DNA breaks and DNA damage response (Figure 1B) [32,65]. Mammalian cells have multiple DNA repair pathways that handle the repair of DSBs, leading to faithful or mutagenic repair, which produces small or large deletions/insertions or complex chromosomal rearrangements [77,78]. Whether any of these types of DSB misrepair happen at the L1-induced DSBs is not known (Figure 1B).

In addition to supporting residual retrotransposition without functional EN, the L1 RT domain has been linked to triggering an inflammatory response through generation of cytoplasmic cDNA (Figure 1B) [63]. The mechanism of this process is not completely understood in light of existing biochemical studies [62,79]. The proposed mechanism involves the promiscuous ability of the ORF2p RT to utilize various forms of nucleic acids as a primer for cDNA synthesis [53,63]. With many nonfunctional L1 loci retaining

either functional EN or RT enzymatic activities being expressed in normal and cancer cells [9,26,27], gaining a better understanding of the potential of these loci to elicit a DNA damage response and their contribution to genomic instability is necessary.

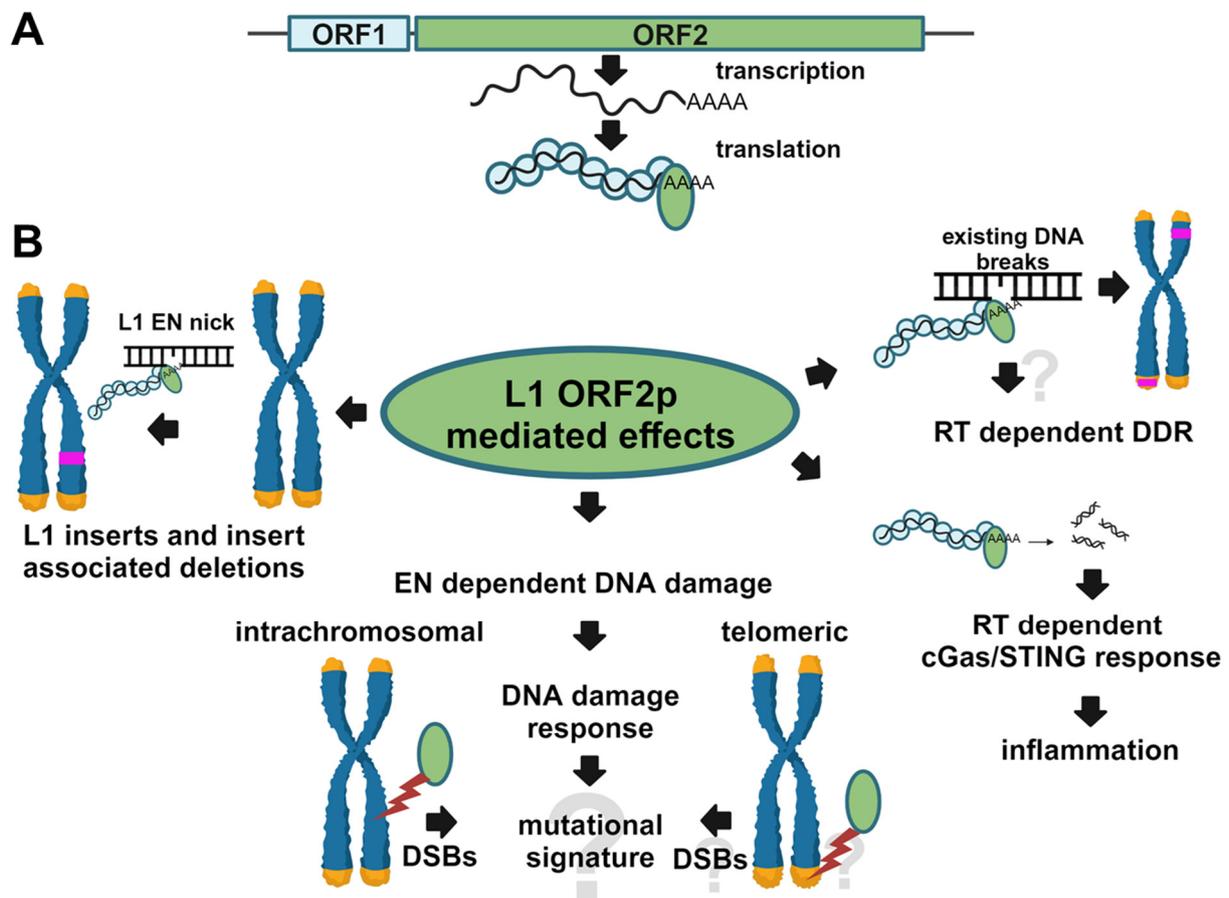


Figure 1. An overview of the L1 replication cycle and consequences of L1 ORF2p enzymatic functions. (A) Pol-II-driven transcription of a full-length functional L1 locus produces a full-length bicistronic mRNA containing open reading frames (ORFs) encoding ORF1 (blue) and ORF2 (green) proteins. The two proteins and the L1 mRNA assemble into L1 ribonucleoprotein (RNP) complexes capable of mobilization. The ORF2p contains an endonuclease (EN) domain that initiates L1 integration by nicking genomic DNA. It also contains a reverse transcriptase (RT) domain, which is responsible for DNA primer-mediated cDNA synthesis during L1 integration. (B) Expression of functional L1 in human cells results in de novo L1 inserts (pink), which sometimes cause large genomic rearrangements (left). This process relies on both EN and RT functions. The L1 expression also leads to the formation of EN-dependent DNA double-strand breaks (shown with a red bolt), which trigger DNA damage response (center). Other than L1 inserts, the spectrum of mutations resulting from these DSBs and whether L1 EN can cut telomeric sequences is not known. Among RT-dependent outcomes are EN-independent L1 intrachromosomal and telomeric integrations (shown with pink bands, bottom right) and inflammatory response triggered by cGas/STING activation by RT-generated cDNA. Whether L1 RT-mediated cDNA synthesis can trigger DNA damage response (DDR) is not known (top right). Question marks indicate unknown outcomes.

The DNA damage caused by individual endogenously expressed L1 loci is difficult to reliably measure *in vivo* due to the lack of sensitive and quantitative approaches and very low expression levels of endogenous L1 elements [25,26]. Therefore, L1-associated DDR has been investigated using various versions of transiently transfected L1 or ORF2p expression plasmids [32,65,70]. A major disadvantage of this approach is that transient transfections themselves may cause a DDR-like response in mammalian cells [80,81], creat-

ing a background signal that often masks cellular response to the L1-associated damage. We capitalized on the reported inducible L1 system that eliminates the need for transient transfections to trigger measurable L1 expression [57]. Using this approach, we generated plasmids that express defective L1 elements, ORF2, and EN-containing truncated proteins harboring inactivating mutations within their EN and/or RT domains. Our findings demonstrate that the induced expression of either an EN-mutant full-length L1 element or an EN-mutant ORF2 leads to the accumulation of phosphorylated H2AX (γ H2AX) signal. Our findings provide evidence that retrotransposition-incompetent L1 loci, specifically those that retain RT function, can contribute to the DDR, which may contribute to the observed RT-dependent inflammatory response triggered by L1. We also report a Thymidine Kinase Neomycin resistance fusion reporter cassette designed to capture novel types of L1-associated mutations. This system shows that exposure to functional L1 can lead to de novo genomic deletions that lack any hallmarks of L1 contribution. Finally, we report that the L1 endonuclease domain purified from bacterial cells can efficiently cut mammalian telomeric sequences in vitro, establishing the foundation for future studies involving its potential contribution to telomere shortening occurring during normal aging and short telomeres typically present in cancer cells.

2. Materials and Methods

2.1. Generation of L1-, ORF2-, and EN-Containing Plasmids

EN (D205A) (5'-CCTACAGCAAGATCGCCCCACATCGTGGGCAG-3' and 5'-CTGCCC-ACG ATGTGGGCGATCTTGCTGTAGG-3')- and RT (D702A) (5'-GAGCCTGTTCGCCGCC-GAC ATGATCGTGT-3' and 5'-ACACGATCATGTCCGGCGCGAACAGGCTC-3')-inactivating mutations were introduced into the previously reported L1 expression plasmid [57] using site-directed mutagenesis as previously described [32]. The human codon-optimized ORF2 sequence [57,70] was PCR-amplified using these L1-containing plasmids for subcloning to generate plasmids containing wild-type and mutant ORF2 sequences that can be used in the reported inducible system.

2.2. Cell Culture

FLP-InTM-293 (Invitrogen, Carlsbad, CA, USA) cells were cultured in HyClone Dulbecco's modified Eagle's medium with 10% fetal bovine serum (Invitrogen) and maintained under 6% CO₂ at 37 °C. Tet-on HEK 293 cells were cultured as previously described [32,57]. The cells were seeded (1 million cells per T25 flask) 16–18 h prior to induction with doxycycline (dox) or transfection (normal growth media was replaced 3 h post-transfection for all experiments) [57].

2.3. Generation of Doxycycline-Inducible 293T Cell Lines

Cell lines harboring wild-type or mutant L1-, ORF2-, and EN-containing plasmids were generated as previously described [57]. At least two independently generated inducible cell lines were used for each experiment. The induction of L1, ORF2, or ORF2-truncated fragments was initiated by doxycycline (1 μ g/mL) treatment as described [57].

2.4. Protein Harvest for Western Blot Analysis

Cells were washed with 1X PBS (137 mM NaCl (Sigma S9888), 2.7 mM KCl (Sigma P4505), 10 mM Na₂HPO₄ (Sigma S3264) and 2 mM KH₂PO₄ (Sigma P9791), pH = 7.4) and harvested using 300 μ L of TLB SDS (50 mM Tris, 150 mM NaCl, 10 mM EDTA, 0.5% sodium dodecyl sulfate, TritonX-100 0.5% *v/v*, Halt Protease inhibitor 10 μ L/mL, phosphatase inhibitors 2 and 3 (Sigma, Livonia, MI, USA), pH = 7.2) per T25 flask. The total lysate samples were sonicated three times for 10 s at 12 watts RMS using a 3 mm wide Qsonica Microson homogenizer with Microson ultrasonic cell disruptor XL2000 (Microson, Brussels, Belgium). The samples were centrifuged at 21,130 \times *g* for 15 min at 4 degrees Celsius. The resulting supernatant was transferred to a new microcentrifuge tube. The protein concentration for each sample was determined using 595 nm wavelength OD values against

a Bovine Serum Albumin (BSA) standard. An equal amount of protein (15–30 micrograms) was loaded for each sample.

2.5. Western Blot Analysis

For each sample, 15–20 µg of protein was combined with 2x Laemmli buffer to obtain the final concentration of 1X, 1.6 µL β-mercaptoethanol, and heated at 100 degrees Celsius for 5 min prior to loading. The samples were fractionated on a Bis-Tris 4–12% Midi gel or Novex Tris-Glycine 4% Mini gel (Invitrogen) and transferred to a nitrocellulose membrane (iBlot2 system: Invitrogen). The membrane was first incubated for 1 h in the 5% milk/PBS-Tween buffer (0.1% *v/v* Tween 20 (Sigma P2287), 137 mM NaCl (Sigma S9888), 2.7 mM KCl (Sigma P4505), 10 mM Na₂HPO₄ (Sigma S3264) and 2 mM KH₂PO₄ (Sigma P9791), pH = 7.4), and then overnight at 4 degrees Celsius with 1:5000 dilution of hORF1 (custom polyclonal rabbit antibody: TGNSKTQSASPPPK antibody [46] or 1:10,000 Flag-Tag-Specific (Sigma Monoclonal Mouse Anti-Flag M2: F3165) antibody, 1:100,000 γH2AX (Santa Cruz, CA, USA; sc-101696) antibody, or ORF2-specific antibodies [32] in 3% milk/PBS-Tween buffer. Following the overnight incubation, the membranes were washed for 5 min 3 times using PBS-Tween buffer and incubated at room temperature for 1 h with 1:5000 dilution of horseradish peroxidase-conjugated secondary antibodies (HRP-donkey anti-rabbit (Santa Cruz: sc2317)) in 3% milk/PBS-Tween buffer. The membranes were washed one time for 15 min with PBS-Tween buffer and then twice for 5 min using PBS-Tween buffer. The development was carried out using a 5 min incubation with a Bio-Rad Clarity Kit (Bio-Rad, Hercules, CA, USA). The signal was detected using a Chemi-Doc XRS+ Molecular Imager (Bio-Rad). anti-GAPDH antibodies (Santa Cruz: sc25778, 1:5000 dilution) were used to detect GAPDH using the incubation conditions described above.

2.6. LEAP Assay

Analysis of the RT function was performed using LEAP as previously described [40,79]. In short, HEK 293 cells were transiently transfected with indicated ORF2 expression plasmids; 48 h later cytoplasmic RNP-containing fraction was extracted and purified using sucrose gradient. The resulting RNP-containing pellet was used for testing the ORF2 RT-mediated cDNA synthesis using a primer specific to the 3' end of the ORF2 expression plasmid and a primer specific to the LEAP primer.

2.7. Retrotransposition Assay

The retrotransposition assay was performed in HeLa cells as previously described using AluNeo reporter plasmid [82].

2.8. Analysis of L1-Induced Mutations

The TKNeo reporter cassette was synthesized by GenScript. An annotated TK/Neo reporter sequence is shown in Supplemental Materials (File S1). The resulting DNA was then cloned in the pEF5/FRT/V5-DEST destination vector of the Flp-In system (Invitrogen). An amount of 0.6 µg of this plasmid was co-transfected with 5.4 µg of the pOG44 Flp-recombinase plasmid expressing the Flp recombinase using 12 µL of Plus and 36 µL of Lipofectamine into the FRT containing HEK 293-FRT cells seeded 18 h prior (5,000,000 cells). Then, 48 h later, cells were reseeded, and hygromycin selection was initiated 24 h later. Homologous recombination between the FRT sites results in the site and direction-specific integration of a single copy of our TKNeo cassette into the cellular genome. Stable integration of our pEF5/FRT/V5-DEST TK/Neo cassette confers hygromycin resistance to the cells. Hygromycin-resistant HEK 293-FRT colonies were picked and expanded. Selection was used to select cells containing a single copy of the TK/Neo cassette. Their genomic DNA was used to screen clones containing TK/Neo cassette by site-specific PCR with primers complementary to the SV40 promoter and Hygromycin. Selected positive clones were expanded in the presence of neomycin to eliminate any cells that may have undergone any L1-induced or random mutagenesis during normal cell proliferation. 500,000–1,000,000 cells

was seeded and transfected the next day with 0.4–1 µg of either ISce-I-, L1-, or ORF2p-expressing plasmids containing EN- or RT-inactivating mutations. An amount of 4–8 µL of Plus Reagent and 12 µL of Lipofectamine were used. Gancyclovir selection (25 µM) was initiated 2–4 days post-transfection and was maintained for 14 days or until cell collection.

2.9. Site-Specific PCR Analysis

500,000 Neomycin-resistant TKNeo cells were seeded in T75 flasks and transiently transfected ~18–20 h later with 0.5–2 µg of L1 or ISceI expression plasmids (4 µL of plus and 12 µL of Lipofectamine) to induce DSBs. Selection with 1 µM gancyclovir was initiated two to four days post-transfection and maintained for 2 weeks to allow for colony formation. DNA from L1 and ISceI exposed cells was collected at 2, 3, 4, 5, or 14 days after transfection. Site-specific PCR analysis was performed using 100 ng of DNA template, 5% DMSO, 1 µL of a forward and 1 µL of a reverse primer, and 45 µL of Platinum PCR Supermix. PCR cycling conditions: 94C for 1 min, [94C for 30 s, 55C for 30 s, 68C for 45 s] X 33 or 35 cycles, 72C for 10 min. The primers used for PCR are as follows:

pEF1355F primer (5'-GAGGGGTTTTATGCGATGGAGTTT-3');
 TK/Neo291R primer (5'-AGCAGTTGCGTGGTGGTGGTTTTTC-3');
 pEF5F primer (5'-CGCCCACAGTCCCCGAGAA-3');
 200F primer (5'-CCAGTGTGGTGGAAATTCTGCAGA-3');
 200R primer (5'-GCAGACGCATGTTGATGGCA-3');
 F3 primer (5'-CAGGGTTATTGTCTCATGAGCGGA-3');
 R3 primer (5'-ATGCGGCATCAGAGCAGATTGT-3');
 F6 primer (5'-GCTATGACTGGGCACAACAGACA-3');
 NeoR primer (5'-TGCCTCGTCCTGCAGTTCAT-3').

Primer sequences are annotated in the TK/Neo reporter sequence in the Supplemental Materials, except for the F3/R3 primer pair because they are complementary to the sequence in the Flp-In cassette between the Amp gene and pEF5 promoter.

2.10. Endonuclease In Vitro Cleavage Assay

Endonuclease in vitro cleavage assay was performed as previously described using *Escherichia coli* (*E. coli*) purified L1 endonuclease domain [68]. Primer sequences L1 WT 5'-AGGCTTAAAATT AAAATTTAAAATTTAAAAGGC-3', TEL WT 5'-AGGCTTAGGGTTAGGGTTAGGGTTA GGGAGGC-3', TEL MUT 5'-AGGCTCCGGTCCGGTCCGGTCCGGGAGGC-3' and their reverse complement sequences were used to generate double-stranded DNA substrate. Either the top or bottom strand was 5' labeled with AlexaFluor 488 (IDT).

2.11. Assessment of Disease-Causing Mutations and the Number of L1 EN Sites in Human Genes

Gene sequences were extracted from the human genome assembly. The number of L1 endonuclease recognition sites was determined using the DNASTAR Lasergene program. The number of L1 sites per kb was determined using the following formula: total # of L1 sites/total length of a gene. The number of total disease-causing mutations for each gene was determined using The Human Gene Mutation database (HGMD) <https://www.hgmd.cf.ac.uk/ac/index.php>. (accessed on 117 October 2017) The percent of genomic deletions per gene was calculated using the following formula: (# deletions/# total mutations) X100. Pearson correlation was performed using GraphPad Prism.

3. Results

3.1. RT-Dependent Accumulation of γ H2AX Signal upon Induction of L1 Expression

Previously reported functional codon-optimized (L1.3) inducible human L1 (L1 EN+ RT+) [57] was modified to contain inactivating mutations in either the EN (D205A), RT (D702A), or both domains [47]. These constructs are referred to as L1 EN–RT+, L1 EN+RT–, and L1 EN–RT–, respectively. Tet HEK 293T cells harboring these respective constructs were generated and induced by the addition of doxycycline (Dox, 1 µL/mL) as previously reported [57]. The accumulation of L1 ORF1 and ORF2 proteins, as well as the γ H2AX

signal, was monitored by Western blot analysis of total protein lysates collected from these cells at 3, 24, and 48 h post-induction (Figure 2). Western blot analysis performed using anti-ORF1p-specific antibodies, as previously described, demonstrated that all four constructs produce similar levels of L1 ORF1 protein with low levels of the protein being detected as early as 3 h post-induction (Figure 2A, ORF1p panel). Similar to the ORF1p, low levels of the ORF2 protein were also detected at 3 h post-induction using Western blot analysis with Flag-specific antibody (Figure 2A, ORF2 panel). Using this approach, several variants of the ORF2p were detected. This observation is in agreement with previous reports [32]. The detected variants are consistent with the ORF2 proteins that have lost their N-terminal sequences based on the size of detected bands and the presence of the Flag-tag on the C-terminus of the protein. A different trend in the ORF1p and ORF2p accumulation is observed between constructs expressing ORF2p with functional versus nonfunctional RT. Expression of ORF2p containing functional RT peaked at 24 h post-induction, while expression of ORF2p containing nonfunctional RT was comparable at 24 and 48 h post-induction (Figure 2).

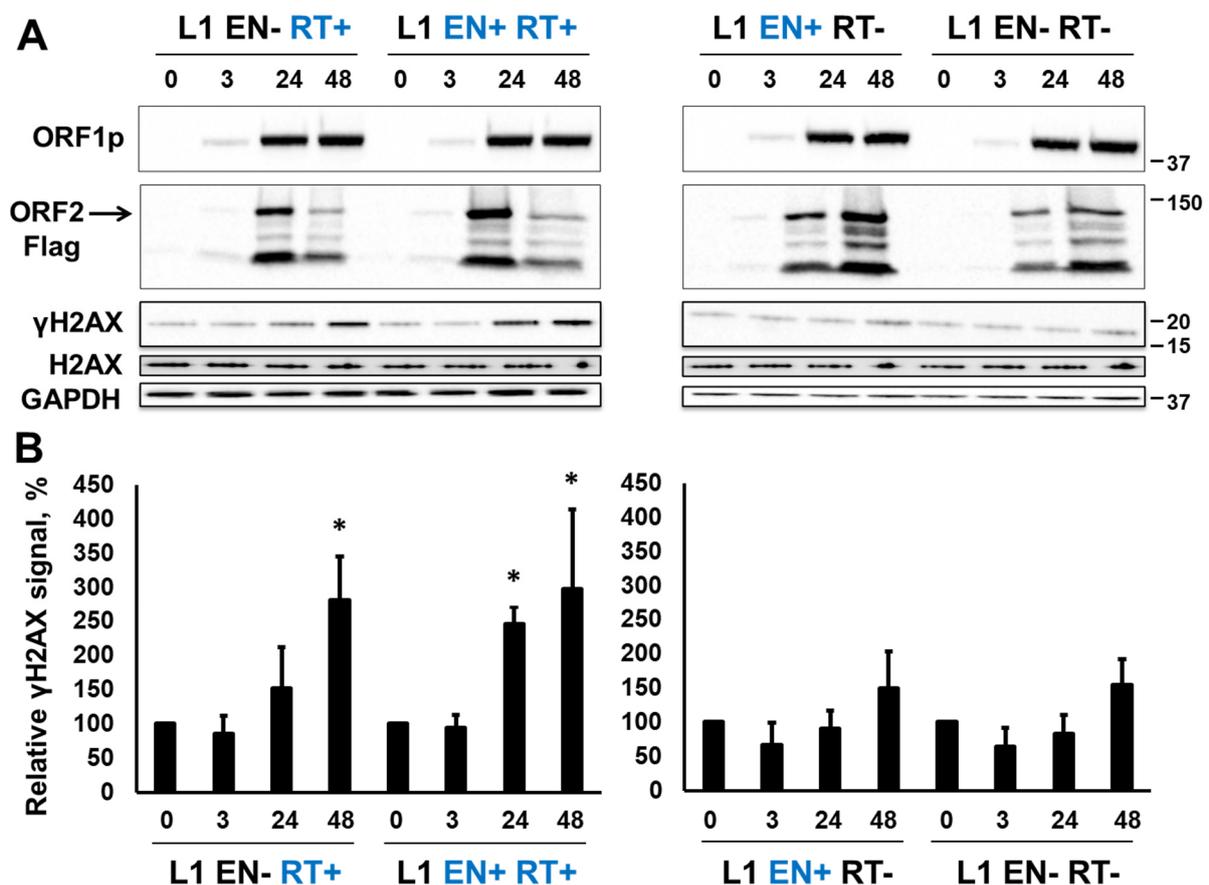


Figure 2. Expression of retrotransposition-incompetent L1 elements triggers DNA damage response. (A) Western blot analysis of L1 ORF1 and ORF2 protein expression and γ H2AX and H2AX signal in HEK 293 cells at 3, 24, and 48 h post-induction of L1 expression with Doxycycline. EN+ (blue) and EN- correspond to either functional or mutant endonuclease domains. RT+ (blue) and RT- correspond to the functional and mutant reverse transcriptase domain. Anti-ORF1-specific Ab was used to detect ORF1p. Anti-Flag-tag-specific antibodies were used to detect ORF2 protein. A band corresponding to the full-length ORF2p is indicated by the arrow. Anti-phosphoH2AX- and H2AX-specific antibodies were used to detect the γ H2AX and H2AX signals, respectively. (B) Quantitation of the γ H2AX signal relative to the 0-time point. Asterisks denote statistically significant values relative to the 0-time point (N = 3, Student's *t*-test, $p < 0.05$).

It was previously reported that the expression of functional L1 elements in mammalian cells results in H2AX phosphorylation [65], which is indicative of the cellular response to DNA damage, such as DNA double-strand breaks and other types of lesions. Western blot analysis using anti-139 phospho-serine H2AX antibodies demonstrated accumulation of the γ H2AX signal in cells induced to express L1 EN+RT+ or L1 EN-RT+ full-length elements (Figure 2). No significant increase in the γ H2AX signal was observed when L1 EN+RT- or L1 EN-RT- full-length elements were expressed (Figure 2B).

3.2. RT-Dependent Accumulation of γ H2AX Signal upon Induction of ORF2p Expression

To determine whether this pattern of γ H2AX signal accumulation is unique to the full-length L1 element or whether it is specific to the L1 ORF2 protein, Tet HEK 293T cells harboring inducible constructs designed to express full-length functional human ORF2 protein (ORF2 EN+RT+) or ORF2 proteins containing inactivating mutations in their EN (D2015A), RT (D702A), or both the EN and the RT domains were generated. These constructs are referred to as ORF2 EN-RT+, ORF2 EN+RT-, and ORF2 EN-RT-, respectively. Tet HEK 293T cells were induced to express the respective ORF2 proteins via the addition of doxycycline (1 μ L/mL) as previously described [57]. The accumulation of these ORF2 proteins, as well as the γ H2AX signal, was analyzed using the total protein lysates collected from these cells at 3, 24, and 48 h post-induction. The expression of ORF2 proteins was detected using Western blot analysis with anti-ORF2p antibodies (Figure 3). DNA from cells harboring various ORF2 constructs was extracted and used for ORF2-specific PCR, followed by sequencing to confirm the wild-type or mutant status of EN and/or RT domains for each inducible cell line. Similar to the results observed with the full-length L1, expression of the EN+RT+ and EN-RT+ ORF2 proteins resulted in the significant accumulation of the γ H2AX signal detected at 24 and 48 h after the addition of Dox (Figure 3A, ORF2 and γ H2AX panels, and Figure 3B). No increase above the background γ H2AX signal was detected over the 48 h time course in the HEK 293 cells harboring the ORF2 EN+RT- or EN-RT- constructs.

To confirm that the RT function remains intact when the EN domain is inactivated, a LEAP assay was performed as previously described (37). This approach demonstrated that both the EN+RT+ and the EN-RT+ ORF2 proteins were able to synthesize cDNA when an exogenous primer was included in the reaction (Figure S1) and support varying levels of Alu retrotransposition in HeLa cells (Figure S2).

Although the induction of L1 EN+RT- or the ORF2 EN+RT- expression did not elicit significant DNA damage response ($p = 0.06$ or $p = 0.07$, respectively) as demonstrated by the lack of accumulation of the γ H2AX signal (Figures 2 and 3), the induction of the EN domain (aa 1–239) expression resulted in a significant increase in γ H2AX signal as early as 2 h post-induction (Figure 4). As expected, the induction of the catalytically inactive EN domain did not elicit an increase in the γ H2AX signal. We also confirmed that the Dox-driven expression of the C-terminally truncated ORF2p fragments, containing the endonuclease domain, Cryptic region, and Z-domain resulted in accumulation of the γ H2AX signal only when the fragment, containing the functional endonuclease, was expressed (Figure 5, EN+CryZ). Consistent with our previous observations [32], the γ H2AX signal in response to the expression of the EN+CryZ protein was reduced and delayed compared to the signal observed when the EN+ protein was expressed. No increase in the γ H2AX signal was observed when the ORF2p fragment containing the nonfunctional EN (EN-CryZ) was expressed (Figure 5, EN-CryZ). The detection of the γ H2AX signal as early as 2 h post-induction demonstrates the advantage of the inducible expression over transient transfections where the nonspecific DNA damage response masks cellular response to the L1-induced DNA damage during this time frame. The findings at 19 and 24 h are consistent with our previous results obtained using transient transfection of expression plasmids containing above mentioned fragments of the human ORF2 sequence [32].

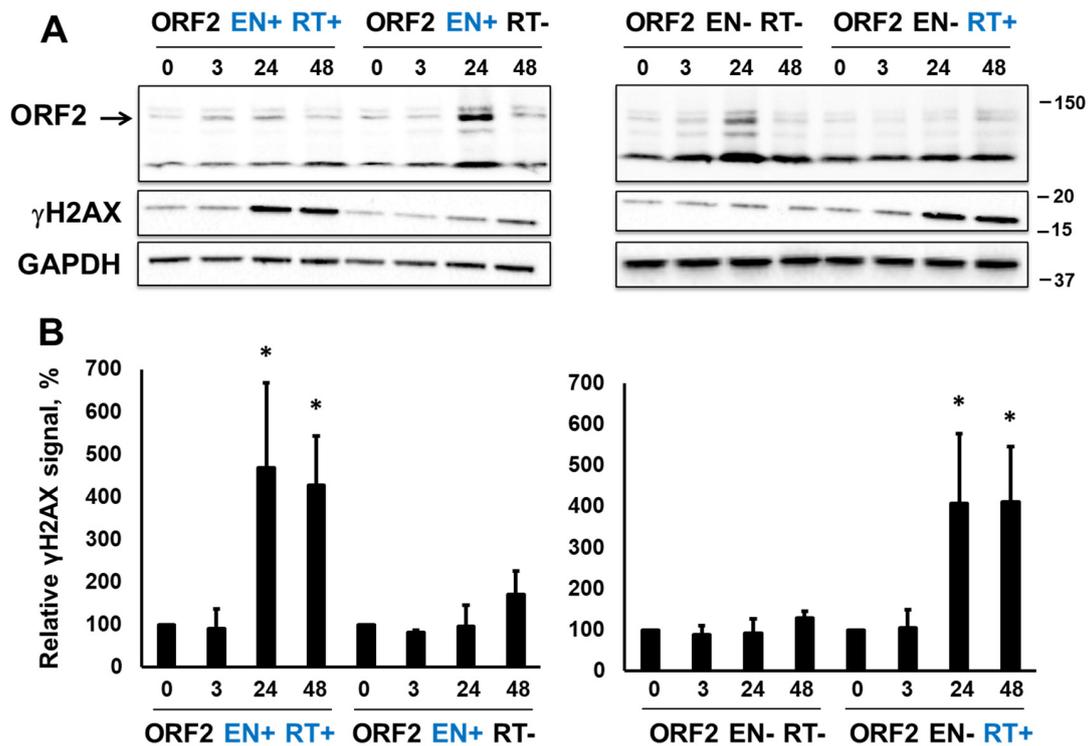


Figure 3. Expression of retrotransposition-incompetent ORF2 proteins triggers DNA damage response. (A) Western blot analysis of the ORF2 protein expression and γ H2AX signal in HEK 293 cells at 3, 24, and 48 h post-induction of ORF2p expression with doxycycline. EN+ (blue) and EN− correspond to either functional or mutant endonuclease domains. RT+ (blue) and RT− indicate either functional or mutant reverse transcriptase domain. ORF2-specific antibodies were used to detect the ORF2 protein. A band corresponding to the ORF2p is indicated by the arrow. (B) Quantitation of the γ H2AX signal relative to the 0-time point. Asterisks denote statistically significant values relative to the 0 time point (N = 3, Student's *t*-test, $p < 0.05$).

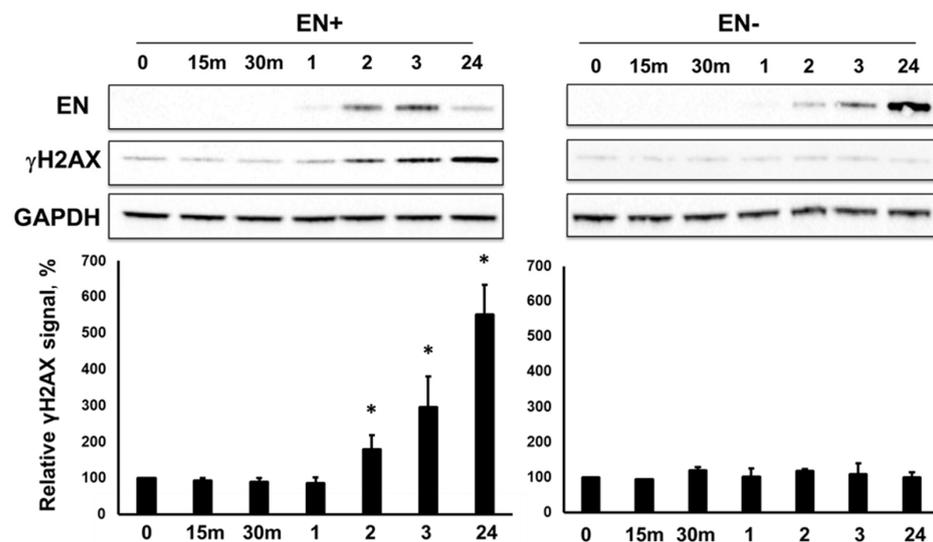


Figure 4. Expression of inducible EN protein triggers DNA damage response. Western blot analysis of the L1 endonuclease domain (EN) protein expression and γ H2AX signal in HEK 293 cells at 15 minutes (15 m), 30 minutes (30 m), 1, 2, 3, and 24 h post-induction of endonuclease expression with Doxycycline. EN+ and EN− indicate either a functional or mutant endonuclease. Anti-ORF2-specific antibodies were used to detect the EN+ and EN− proteins. Asterisks denote statistically significant values relative to the 0-time point (N = 3, Student's *t*-test, $p < 0.05$).

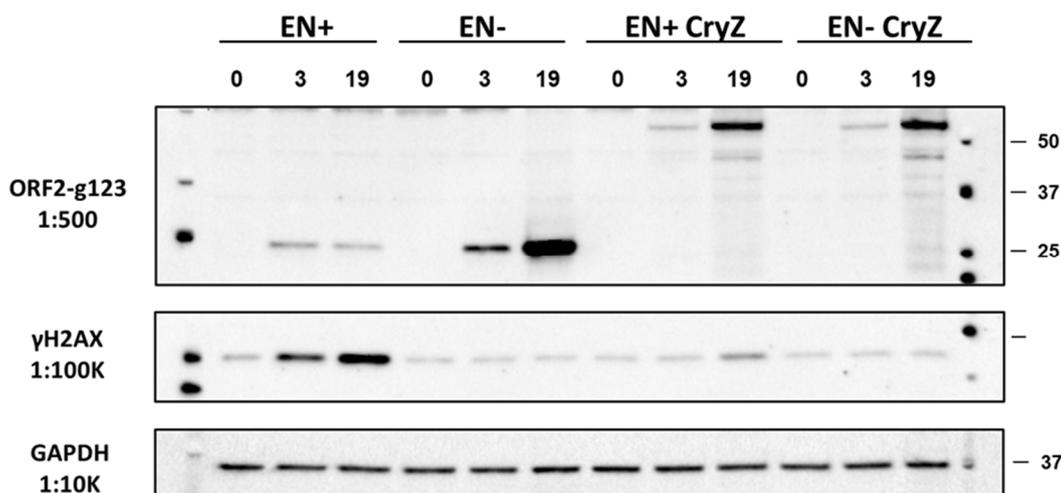


Figure 5. Inducible expression of C-terminally truncated ORF2p in HEK 293 cells. Constructs containing functional (EN+) or nonfunctional (EN−) endonuclease portions with or without the Cryptic (Cry) region of the human ORF2 sequence were generated for their use in the doxycycline-inducible system. Doxycycline treatment of HEK 293 cells containing these constructs resulted in the expression of corresponding proteins as detected by Western blot analysis with anti-ORF2p antibodies at 3 or 19 h after induction. Anti-phosphoH2AX-specific antibodies were used to detect the γ H2AX signal. GAPDH detection was used as a loading control.

3.3. L1 Expression Causes Large Genomic Deletions That Do Not Contain Any L1 Sequences

In addition to canonical L1 inserts that do not create any major structural rearrangements at their integration sites, *de novo* L1 integration sporadically causes large genomic deletions in cultured cells and *in vivo* [64,67]. These events are attributed to L1 because they contain L1 sequences and hallmarks of *de novo* L1 integration. To investigate the possibility that L1 could cause other types of genomic changes, including deletions or insertions that do not contain any signatures of L1 retrotransposition, we developed a reporter cassette designed to capture products of misrepaired L1-induced DSBs (Figure 6A). A series of L1 EN sites (5'-TTTTAA-3') were included downstream of the translation initiation codon (ATG) of the thymidine kinase/neomycin (TK/Neo) fusion gene (Figures 6A and S3). An I-SceI site is included as a positive control. This reporter cassette was integrated as a single copy into an FRT site that was previously inserted in the genome of HEK 293 FRT cells (Figures 6A,B, and S3). Neomycin selection during propagation of clones with stably integrated TK/Neo cassette was used to prevent the accumulation of frameshift background mutations in this reporter sequence. Ganciclovir administration after transfection with plasmids expressing L1 (functional or EN mutant) or I-SceI allowed the survival of cells that have lost TK function due to the misrepair of DNA breaks (Figure S3). Formation of ganciclovir resistant (GanR) colonies resulting from the I-SceI induced DSBs was used as a positive control, L1 containing an inactivating mutation in the EN domain was used to control for the L1 endonuclease contribution to the observed signal (Figure 6C). This approach determined that the expression of functional L1 increases the frequency of GanR colonies compared to the expression of L1 containing an EN-inactivating mutation (Figure 6C).

To detect outcomes of misrepaired events within the TK/Neo cassette, we performed site-specific PCR using primers flanking the I-SceI, L1 site sequences in the TK/Neo cassette and DNA extracted from pooled GanR colonies (Supplemental Figure S4B). This approach identified 25 unique repair products out of 40 analyzed sequences recovered from the I-SceI-induced GanR colonies. These were deletions ranging from 1 to 8 bp or insertions ranging from 120 to 386 bp. This result is consistent with previous findings of I-SceI-induced small deletions and insertions [83]. All sequences recovered from the pooled L1EN− or L1 EN+ induced GanR colonies (N = 20 each) were either of the wild type or contained point mutations representing background. Site-specific PCR analysis of DNA extracted from

identified small deletions and insertions only in the PCR products generated using DNA extracted from cells transfected with the I-SceI expression plasmid.

No such events are detected in the PCR products generated using DNA extracted from cells exposed to L1 EN⁻, L1EN⁺, or ORF2EN⁺, i.e., these PCR products contained wt TK/Neo sequences or had random background point mutations inactivating TK function. This observation is consistent with our findings reflected in Supplemental Figure S4.

In addition to the 2.7 kb PCR product, this approach also detected smaller PCR products that are consistent with internal deletions in the TK/Neo cassette (Figure 6D). These products are observed in the cells transfected with plasmids expressing wild-type L1, ORF2p, or I-SceI. Sequencing analysis of smaller PCR products recovered from cells transfected with the plasmids expressing functional L1 identified deletions ranging from 300 to 2400 bp (Figures 6 and 7A). 56% of these deletions (5/9) contained small non-template insertions (1–32 bp), which are also observed in 25% (2/8) of deletions recovered from the cells exposed to I-SceI. None of the deletions recovered from cells transfected with the wild-type L1 contain L1 sequences or A tail that could be consistent with L1 integration.

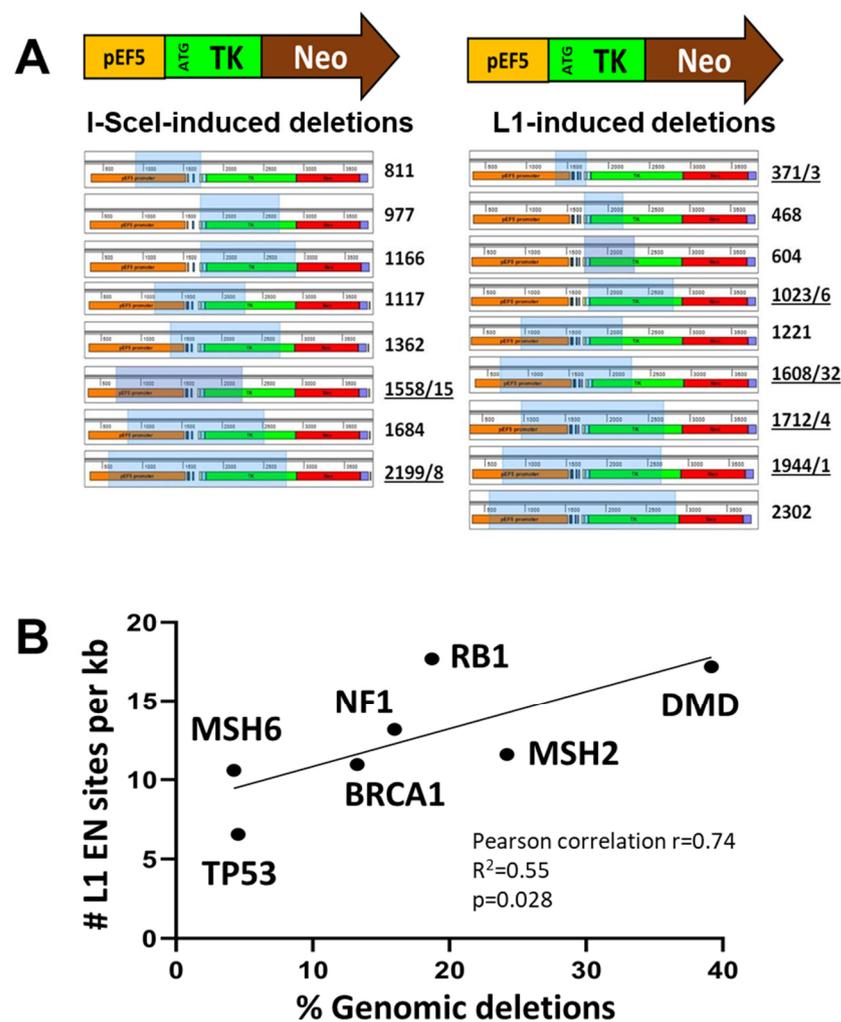


Figure 7. L1 expression causes large genomic deletions. (A) A schematic of the TK/Neo reporter cassette and deletions recovered from I-SceI (left) or L1-induced (right) cells. Numbers on the left of each panel indicate the size of deletion/number of non-template base pairs (underlined when occur together). Light blue boxes indicate deleted TK/Neo sequences. (B) Percent of genomic deletions correlates with the relative number of canonical L1 endonuclease sites. The number of canonical L1 endonuclease (EN) sites was identified in seven human genes and normalized to the gene lengths. The number of disease-causing deletions determined for the same genes correlates with the relative number of L1 EN sites in these genes (Pearson correlation $r = 0.74$). Kb is kilobase.

These findings support that L1 expression could be an underlying cause of at least some large genomic deletions that are currently not attributed to L1 because they do not contain any signatures of L1 involvement. To investigate a potential relationship between L1 EN sites and disease-causing deletions, we determined the frequency of canonical L1 EN sites in seven disease-relevant human genes: NF1, TP53, BRCA1, RB1, MSH6, MSH2, and DMD. The density of L1 EN sites per kb of gene sequence in these seven human genes strongly correlates with the relative frequency of large deletions cataloged for these genes in the human gene mutation database (Figure 7B, Pearson correlation $r = 0.74$, $R^2 = 0.55$, $p = 0.028$).

3.4. L1 Endonuclease Cuts Telomeric Sequences In Vitro

In addition to L1 endonuclease cleavage sites scattered throughout genomic DNA, telomeric repeat sequence (5'-TTAGGG-3') is one of the noncanonical L1 endonuclease cleavage sites based on the reported L1 endonuclease consensus sequence [43,84]. To test whether L1 EN can cut this sequence, we used our previously reported in vitro endonuclease cleavage assay [68]. Functional L1 endonuclease domain purified from *E. coli* was used to test its ability to cut double-strand (ds) DNA containing canonical L1 endonuclease cleavage sites (5'-TTTT/AA-3'), wild-type (5'-TTAGGG-3') or mutant (5'-TCCGGG-3') telomeric repeats. The experiment was carried out with dsDNA containing 5'-end Alexa labeling on either top or bottom strands (Figure 8A). This approach shows that the L1 endonuclease efficiently cuts dsDNA containing canonical L1 EN sites and wild-type telomeric repeats (Figure 8B,C).

A time course analysis shows that within 10 min of the substrate incubation with L1 endonuclease over 90% of the dsDNA containing canonical L1 sites is cut at various sites (Figure 8B). Efficient cleavage occurs on both the top and bottom strands of the substrate (Figure 8B). Under the same experimental conditions, the TEL WT substrate is cleaved more efficiently on the bottom strand than the top strand, with 90% and 50% of the respective substrates being cut within 10 min of incubation with the L1 EN domain (Figure 8C, the chi-square statistic is 50.106. The p -value is <0.00001 . The result is significant at $p < 0.05$). Chi-square analysis determined that the purified L1 endonuclease cuts the bottom strand of the TEL WT substrate as efficiently as it cuts the bottom strand of the L1 WT substrate (the chi-square statistic is 2.834. The p -value is 0.24244. The result is not significant at $p < 0.05$). However, the L1 WT top strand sequence is a significantly better L1 EN substrate than the top strand sequence of the TEL WT dsDNA (The chi-square statistic is 69.0724. The p -value is <0.00001 . The result is significant at $p < 0.05$). While some cleavage of the TEL MUT substrate is observed after 10 min of incubation with the L1 EN, most of this DNA remains intact after 40 min of exposure to the enzyme compared to the L1 WT and TEL WT substrates (Figure 8, 78% top strand label versus 1 and 34%, respectively). Chi-square analysis confirmed that the TEL WT sequence is a significantly better substrate for L1 EN cleavage than the TEL MUT sequence (top strand label, the chi-square statistic is 15.9621. The p -value is 0.001154. The result is significant at $p < 0.05$, and bottom strand label, the chi-square statistic is 53.2662. The p -value is <0.00001 . The result is significant at $p < 0.05$). Despite being an overall poor substrate, the bottom strand sequence of the TEL MUT dsDNA is a better substrate for the L1 EN cleavage than the top strand of the TEL MUT substrate (Figure 8D, the chi-square statistic is 13.1233. The p -value is 0.004377. The result is significant at $p < 0.05$). This result is consistent with the findings observed for the TEL WT substrate (Figure 8C).

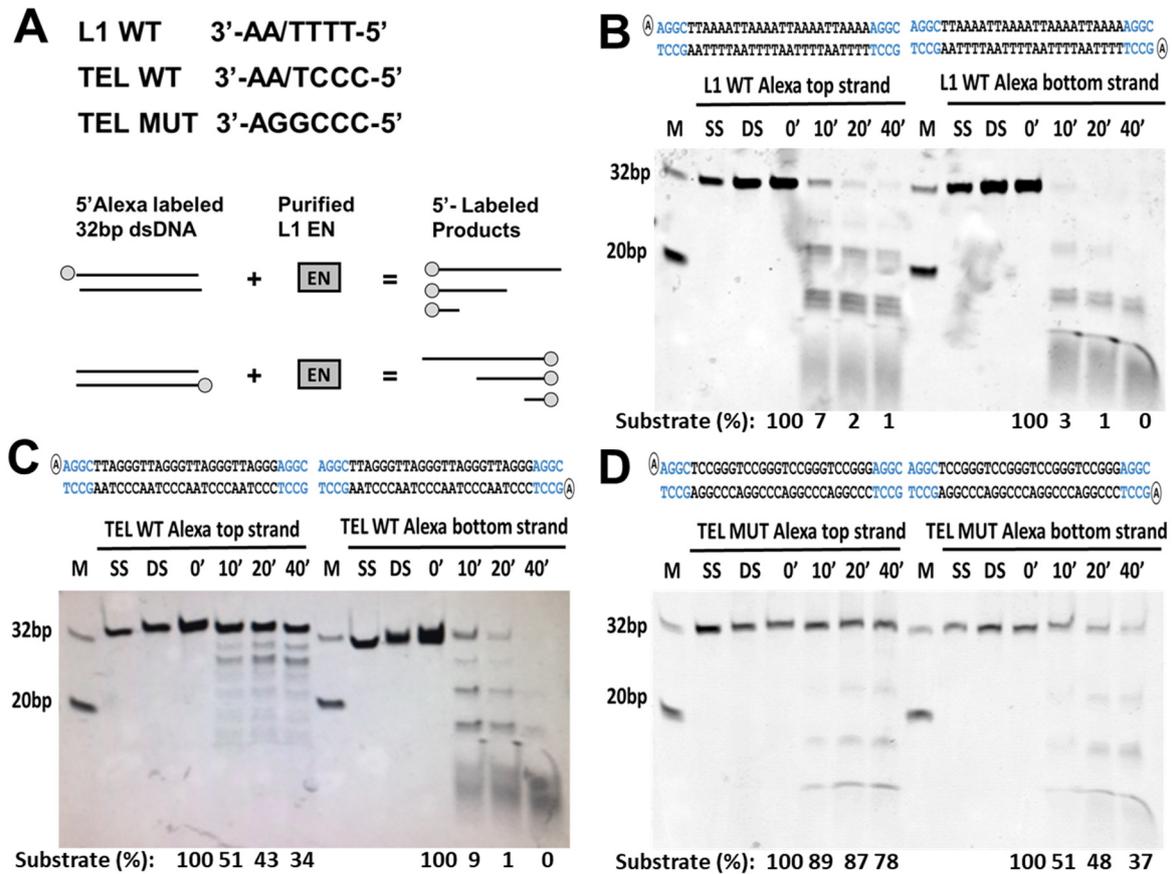


Figure 8. L1 endonuclease cuts telomeric repeat sequence in vitro. (A) Experimental design shows bottom strand sequences that are used to generate 5' Alexa labeled 32 bp long double-stranded DNA. L1 WT is the L1 endonuclease consensus site, TEL WT is a wild wild-type telomeric repeat sequence, and TEL MUT is a mutant telomeric sequence. Overall, 5' top or bottom strand labeled dsDNA was used as a substrate for a purified wild-type L1 endonuclease domain (EN) in an in vitro cleavage assay. Blue indicates L1 or TEL unrelated extra sequence included in design. Products were assessed after 10, 20, or 40 min of incubation. Resulting products are shown for L1 WT (B), TEL WT (C), and TEL MUT (D) substrates. Primer sequences are shown at the top of each panel. M is a 32 and 20 bp marker, SS single strand DNA, DS double-strand DNA, 0', 10', 20', and 40' are incubation times in minutes. Numbers at the bottom (Substrate %) represent the quantitation of the remaining substrate relative to the initiation time point (0').

4. Discussion

It has been demonstrated that L1 expression in the germ line, during embryogenesis, and in adult somatic tissues contributes to genomic instability via the generation of de novo L1 inserts [64,85,86]. L1 mRNA and protein expression, as well as L1 retrotransposition, are significantly increased in most human tumors and cancer cell lines [16,20]. Although both functional and nonfunctional L1 loci are expressed in human, mouse, and rat cells and organs [9,10,27,87], the potential impact from the retrotransposition-incompetent L1 loci has only recently been recognized [30–33]. Similarly, recent findings support that L1 retrotransposition is only one of the potentially many ways expression of functional L1 loci may affect host genome stability or cellular function [63,65,88,89].

We identified that the induced expression of retrotransposition-incompetent L1 or ORF2p containing a nonfunctional EN domain elicits the accumulation of γ H2AX signal in cultured mammalian cells. This phenomenon is associated with the functional RT of the L1 ORF2p (Figures 2 and 3) because γ H2AX signal accumulation is observed when the expression of EN+RT+ or EN-RT+, but not EN+RT- ORF2p, is induced (Figures 2 and 3). These findings suggest that the EN function may be limited in the full-length ORF2p in

order to be used in a coordinated manner during other steps of the retrotransposition cycle. This possibility is supported by the robust accumulation of the γ H2AX signal in cells expressing functional L1 endonuclease domain (Figures 4 and 5). This response to EN expression is notably dampened when an ORF2p fragment containing the EN, Cry, and Z sequences is expressed, even though the ENCryZ protein accumulates to higher steady-state levels than the EN protein by 19 h post-induction (Figure 5). This observation is consistent with our previously published data [32,69] that the Cryptic sequence within the human ORF2p may suppress L1 endonuclease function via a yet unknown mechanism. The lack of detectable increase in the γ H2AX signal in cells expressing L1 or ORF2p EN+RT- (Figures 2 and 3) could be explained by either rapid repair of the EN-induced nicks in the absence of the functional RT or by the aforementioned suppression of EN activity in the context of the full-length ORF2 protein lacking a functional RT as previously suggested [69]. It is worth noting that the two scenarios are not mutually exclusive.

We speculate that the RT-induced DNA damage response may result from the ORF2p's ability to utilize the EN-induced or existing DNA breaks for cDNA synthesis. This possibility is supported by the published observations that EN-deficient L1 elements retrotranspose in NHEJ DNA-repair-deficient cells [31] as well as our observations that EN-RT+ ORF2p retains its ability to make cDNA, using ORF2 mRNA, and supports residual Alu retrotransposition in HeLa cells (Supplemental Figure S2 and [31,40]). This finding is also consistent with the ability of truncated ORF2 proteins lacking the EN domain to drive low levels of Alu retrotransposition in HeLa cells [40]. The EN-RT+ ORF2p association with existing single or double-strand DNA breaks and resulting cDNA synthesis could lead to DNA damage response to resolve these stalled retrotransposition intermediates. Alternatively, an association of the EN-mutant ORF2p with existing nicks or DSBs may impede their repair, leading to detectable DNA damage response. This DDR can help complete or interfere with the initiated integration [90–92]. The discovery of the RT-mediated DDR is also consistent with the previous observation that only when the RT function is mutated, the ORF2 mediated toxicity is completely eliminated [70]. The finding that RT alone is sufficient to trigger DDR is yet another way by which retrotransposition-incompetent L1 loci may perturb cellular function. The consequences of this outcome could be chronic DDR that may contribute to the inflammatory response attributed to the L1 RT-mediated cDNA synthesis (Figure 1) because chronic DNA damage can lead to an inflammatory response [93].

L1 elements containing mutant EN can retrotranspose in telomeres of cells deficient in DNA-PKcs function [30]. It is established that L1 EN cuts AT-rich sequences in vitro and in vivo [43]; however, whether L1 EN can cut telomeric repeats is not known. It is also not known whether L1 EN-induced DNA breaks result in mutations, and if so, the spectrum of these mutations is likewise unknown. We provide evidence that L1 expression leads to a measurable increase in mutagenic events (Figures 4, S4 and S5) and that misrepair of L1-induced DNA breaks in HEK 293 cells produces large genomic deletions (Figures 6 and 7). In contrast to the deletions associated with L1 integration, these deletions do not contain any traceable L1 sequences that would indicate L1 involvement. We cannot rule out the possibility that these events result from aborted L1 retrotransposition intermediates, repair of which eliminated all hallmarks of L1 involvement. However, it is also possible that the repair of L1-induced DNA breaks occurred prior to or without initiation of cDNA synthesis. Our findings support that the generation of these deletions is L1-EN-dependent (Figure 6C,D). In contrast to the I-SceI-induced DSBs, which frequently result in small deletions/insertions, small in/dels have not been detected at the L1 sites in our reporter cassette. There are several plausible explanations for this difference in the I-SceI and L1 mutational spectrum. One possibility is the difference in the nature of the breaks induced by these two enzymes. I-SceI produces a 5' overhang DSBs. The structure of the L1-induced breaks in vivo is not known. L1-generated DNA nicks may become DSBs during cell division. Alternatively, L1-induced DNA breaks can be bulky DNA lesions containing newly synthesized cDNA. Another possibility for the difference in the I-SceI

and L1 mutation spectrum is the number of proximal breaks. Only one copy of the I-SceI recognition sequence is present in our reporter system and the genome. In contrast, multiple canonical L1 EN recognition sites are introduced in the TK/Neo reporter sequence. Additionally, multiple canonical and noncanonical L1 EN recognition sites are present in the sequence surrounding our reporter cassette. Thus, it is possible that deletions detected after exposure to L1 result from the misrepair of multiple L1-induced DNA breaks in the region rather than from a single L1-induced DNA break. While further studies to understand the mechanism underlying L1-induced “genomic scars” are needed, our findings establish the possibility that chronic L1 expression in the germline, as well as normal and cancer cells, can contribute to genomic deletions that have no traceable marks of L1 involvement.

These findings call for the reassessment of an already significant L1 contribution to evolution [94,95] and human disease [64,66]. For example, up to 35% of somatically acquired mutations in the genome of an average human cancer sample are represented by large or focal deletions [96], some of which are responsible for the origin or progression of the disease [97]. DSB repair induced genomic deletions underly chromoplexy, which is recognized to drive prostate cancer progression to its aggressive form [98]. None of these deletions are currently attributed to L1, but based on our findings some of them could be generated by L1 because L1 expression in cancer correlates with copy number alterations [99]. A higher genomic content of full-length L1 loci is reported in African American (AA) populations compared to other populations [100]. Coincidentally, AAs have a higher rate and/or more aggressive cancers compared to the Caucasian population [101]. L1-induced deletions could also occur in the germline, where high levels of L1 expression are observed. Cytogenetically detectable germline deletions occur once in every 7000 live births [102]. They could also contribute to a significant number of idiopathic infertility cases and spontaneous abortions. Supporting the possibility of L1 contribution to disease-causing deletions, we observe a strong correlation between the density of L1 sites within seven human genes and the frequency of the disease-causing deletions in these genes (Figure 7B). In addition to genomic deletions, we demonstrate that L1 endonuclease can cut telomeric sequences *in vitro* almost as efficiently as it cleaves canonical L1 sites (Figure 8). While it remains uncertain whether L1 can cut telomeres *in vivo*, this observation provides strong support for testing the hypothesis that L1-induced breaks may be contributing to telomere shortening during aging. If it were true, L1-induced DNA breaks at telomeres might be one of the reasons behind the short telomeres observed in telomerase-expressing cancer cells where L1 expression is typically elevated [103].

5. Conclusions

In summary, we provide evidence that in addition to retrotransposition, the expression of functional and nonfunctional L1 loci may adversely affect host genome stability or cellular function by triggering RT-mediated DNA damage response, causing large genomic deletions or telomere dysfunction. Our findings expand the spectrum of mutations associated with L1 expression by identifying deletions that do not contain any hallmarks of L1 involvement. These results support the need to understand the extent of L1 contribution to genomic deletions that drive genome evolution, accumulate with age, and contribute to human diseases, including cancer.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes15020143/s1>.

Author Contributions: Conceptualization, K.J.K., M.S., C.D. and V.P.B.; methodology, K.J.K., M.S., C.D. and V.P.B.; software, K.J.K., M.S., C.D., A.S., D.L.d. and V.P.B.; formal analysis, K.J.K., M.S., C.D. and V.P.B.; writing—original draft preparation, K.J.K., M.S., C.D. and V.P.B.; writing—review and editing, K.J.K., M.S., C.D., A.S., D.L.d. and V.P.B.; project administration, V.P.B.; funding acquisition, M.S. and V.P.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the Louisiana State Board of Regents Graduate Research Fellowship (to M.S.) and by the Brown Foundation and the National Institutes of Health R01AG057597, R21AG055387, and R21ES027797 (to V.P.B.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article and Supplementary Materials.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921.
- Waterston, R.H.; Lindblad-Toh, K.; Birney, E.; Rogers, J.; Abril, J.F.; Agarwal, P.; Agarwala, R.; Ainscough, R.; Alexandersson, M.; An, P. Initial sequencing and comparative analysis of the mouse genome. *Nature* **2002**, *420*, 520–562. [[PubMed](#)]
- Gardner, E.J.; Lam, V.K.; Harris, D.N.; Chuang, N.T.; Scott, E.C.; Pittard, W.S.; Mills, R.E.; Devine, S.E. The Mobile Element Locator Tool (MELT): Population-scale mobile element discovery and biology. *Genome Res.* **2017**, *27*, 1916–1929. [[CrossRef](#)] [[PubMed](#)]
- Evrony, G.D.; Cai, X.; Lee, E.; Hills, L.B.; Elhosary, P.C.; Lehmann, H.S.; Parker, J.J.; Atabay, K.D.; Gilmore, E.C.; Poduri, A. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* **2012**, *151*, 483–496. [[CrossRef](#)] [[PubMed](#)]
- Richardson, S.R.; Gerdes, P.; Gerhardt, D.J.; Sanchez-Luque, F.J.; Bodea, G.O.; Muñoz-Lopez, M.; Jesuadian, J.S.; Kempen, M.H.C.; Carreira, P.E.; Jeddeloh, J.A. Heritable L1 retrotransposition in the mouse primordial germline and early embryo. *Genome Res.* **2017**, *27*, 1395–1405. [[CrossRef](#)]
- Platt, R.N., 2nd; Vandeweghe, M.W.; Ray, D.A. Mammalian transposable elements and their impacts on genome evolution. *Chromosome Res.* **2018**, *26*, 25–43. [[CrossRef](#)]
- Scott, E.C.; Gardner, E.J.; Masood, A.; Chuang, N.T.; Vertino, P.M.; Devine, S.E. A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer. *Genome Res.* **2016**, *26*, 745–755. [[CrossRef](#)]
- Belancio, V.P.; Roy-Engel, A.M.; Pochampally, R.R.; Deininger, P. Somatic expression of LINE-1 elements in human tissues. *Nucleic Acids Res.* **2010**, *38*, 3909–3922. [[CrossRef](#)]
- Stow, E.C.; Kaul, T.; deHaro, D.L.; Dem, M.R.; Beletsky, A.G.; Morales, M.E.; Du, Q.; LaRosa, A.J.; Yang, H.; Smither, E. Organ-, sex- and age-dependent patterns of endogenous L1 mRNA expression at a single locus resolution. *Nucleic Acids Res.* **2021**, *49*, 5813–5831. [[CrossRef](#)]
- Ergün, S.; Buschmann, C.; Heukeshoven, J.; Dammann, K.; Schnieders, F.; Lauke, H.; Chalajour, F.; Kilic, N.; Strätling, W.H.; Schumann, G.G. Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. *J. Biol. Chem.* **2004**, *279*, 27753–27763. [[CrossRef](#)]
- Martin, S.L. Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells. *Mol. Cell Biol.* **1991**, *11*, 4804–4807. [[PubMed](#)]
- Branciforte, D.; Martin, S.L. Developmental and cell type specificity of LINE-1 expression in mouse testis: Implications for transposition. *Mol. Cell Biol.* **1994**, *14*, 2584–2592. [[PubMed](#)]
- Harris, C.R.; Normart, R.; Yang, Q.; Stevenson, E.; Haffty, B.G.; Ganesan, S.; Cordon-Cardo, C.; Levine, A.J.; Tang, L.H. Association of nuclear localization of a long interspersed nuclear element-1 protein in breast tumors with poor prognostic outcomes. *Genes. Cancer* **2010**, *1*, 115–124. [[CrossRef](#)] [[PubMed](#)]
- Lee, E.; Iskow, R.; Yang, L.; Gokcumen, O.; Haseley, P.; Luquette, L.J., 3rd; Lohr, J.G.; Harris, C.C.; Ding, L.; Wilson, R.K. Landscape of somatic retrotransposition in human cancers. *Science* **2012**, *337*, 967–971. [[CrossRef](#)] [[PubMed](#)]
- Iskow, R.C.; McCabe, M.T.; Mills, R.E.; Torene, S.; Pittard, W.S.; Neuwald, A.F.; Van Meir, E.G.; Vertino, P.M.; Devine, S.E. Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell* **2010**, *141*, 1253–1261. [[CrossRef](#)] [[PubMed](#)]
- Rodić, N.; Sharma, R.; Sharma, R.; Zampella, J.; Dai, L.; Taylor, M.S.; Hruban, R.H.; Iacobuzio-Donahue, C.A.; Maitra, A.; Torbenson, M.S. Long interspersed element-1 protein expression is a hallmark of many human cancers. *Am. J. Pathol.* **2014**, *184*, 1280–1286. [[CrossRef](#)] [[PubMed](#)]
- Rodić, N.; Steranka, J.P.; Makohon-Moore, A.; Moyer, A.; Shen, P.; Sharma, R.; Kohutek, Z.A.; Huang, C.R.; Ahn, D.; Mita, P. Retrotransposon insertions in the clonal evolution of pancreatic ductal adenocarcinoma. *Nat. Med.* **2015**, *21*, 1060–1064. [[CrossRef](#)]
- Solyom, S.; Ewing, A.D.; Rahrman, E.P.; Doucet, T.; Nelson, H.H.; Burns, M.B.; Harris, R.S.; Sigmon, D.F.; Casella, A.; Erlanger, B. Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res.* **2012**, *22*, 2328–2338. [[CrossRef](#)]
- Tubio, J.M.C.; Li, Y.; Ju, Y.S.; Martincorena, I.; Cooke, S.L.; Tojo, M.; Gundem, G.; Pipinikas, C.P.; Zamora, J.; Raine, K. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science* **2014**, *345*, 1251343. [[CrossRef](#)]
- Taylor, M.S.; Wu, C.; Fridy, P.C.; Zhang, S.J.; Senussi, Y.; Wolters, J.C.; Cajuso, T.; Cheng, W.C.; Heaps, J.D.; Miller, B.D. Ultrasensitive detection of circulating LINE-1 ORF1p as a specific multi-cancer biomarker. *Cancer Discov.* **2023**, *13*, 2532–2547. [[CrossRef](#)]

21. Swergold, G.D. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol. Cell. Biol.* **1990**, *10*, 6718–6729. [[PubMed](#)]
22. Hata, K.; Sakaki, Y. Identification of critical CpG sites for repression of L1 transcription by DNA methylation. *Gene* **1997**, *189*, 227–234. [[CrossRef](#)] [[PubMed](#)]
23. Liang, G.; Chan, M.F.; Tomigahara, Y.; Tsai, Y.C.; Gonzales, F.A.; Li, E.; Laird, P.W.; Jones, P.A. Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Mol. Cell Biol.* **2002**, *22*, 480–491. [[CrossRef](#)]
24. Takai, D.; Yagi, Y.; Habib, N.; Sugimura, T.; Ushijima, T. Hypomethylation of LINE1 retrotransposon in human hepatocellular carcinomas, but not in surrounding liver cirrhosis. *Jpn. J. Clin. Oncol.* **2000**, *30*, 306–309. [[CrossRef](#)] [[PubMed](#)]
25. Deininger, P.; Morales, M.E.; White, T.B.; Baddoo, M.; Hedges, D.J.; Servant, G.; Srivastav, S.; Smither, M.E.; Concha, M.; DeHaro, D.L. A comprehensive approach to expression of L1 loci. *Nucleic Acids Res.* **2017**, *45*, e31. [[CrossRef](#)]
26. Philippe, C.; Vargas-Landin, D.B.; Doucet, A.J.; van Essen, D.; Vera-Otarola, J.; Kuciak, M.; Corbin, A.; Nigumann, P.; Cristofari, G. Activation of individual L1 retrotransposon instances is restricted to cell-type dependent permissive loci. *elife* **2016**, *5*, e13926. [[CrossRef](#)]
27. Stow, E.C.; Baddoo, M.; LaRosa, A.J.; LaCoste, D.; Deininger, P.; Belancio, V. SCIFER: Approach for analysis of LINE-1 mRNA expression in single cells at a single locus resolution. *Mob. DNA* **2022**, *13*, 21. [[CrossRef](#)]
28. Freeman, B.; White, T.; Kaul, T.; Stow, E.C.; Baddoo, M.; Ungerleider, N.; Morales, M.; Yang, H.; Deharo, D.; Deininger, P. Analysis of epigenetic features characteristic of L1 loci expressed in human cells. *Nucleic Acids Res.* **2022**, *50*, 1888–1907. [[CrossRef](#)]
29. Taylor, D.; Lowe, R.; Philippe, C.; Cheng, K.C.L.; Grant, O.A.; Zabet, N.R.; Cristofari, G.; Branco, M.R. Locus-specific chromatin profiling of evolutionarily young transposable elements. *Nucleic Acids Res.* **2021**, *50*, e33. [[CrossRef](#)]
30. Morrish, T.A.; Garcia-Perez, J.L.; Stamato, T.D.; Taccioli, G.E.; Sekiguchi, J.; Moran, J.V. Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres. *Nature* **2007**, *446*, 208–212. [[CrossRef](#)]
31. Morrish, T.A.; Gilbert, N.; Myers, J.S.; Vincent, B.J.; Stamato, T.D.; Taccioli, G.E.; Batzer, M.A.; Moran, J.V. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat. Genet.* **2002**, *31*, 159–165. [[CrossRef](#)] [[PubMed](#)]
32. Kines, K.J.; Sokolowski, M.; deHaro, D.L.; Christian, C.M.; Belancio, V.P. Potential for genomic instability associated with retrotranspositionally-incompetent L1 loci. *Nucleic Acids Res.* **2014**, *42*, 10488–10502. [[CrossRef](#)] [[PubMed](#)]
33. Sokolowski, M.; Chynces, M.; deHaro, D.; Christian, C.M.; Belancio, V.P. Truncated ORF1 proteins can suppress LINE-1 retrotransposition in trans. *Nucleic Acids Res.* **2017**, *45*, 5294–5308. [[CrossRef](#)]
34. Kines, K.J.; Sokolowski, M.; deHaro, D.L.; Christian, C.M.; Baddoo, M.; Smither, M.E.; Belancio, V.P. The endonuclease domain of the LINE-1 ORF2 protein can tolerate multiple mutations. *Mob. DNA* **2016**, *7*, 8. [[CrossRef](#)] [[PubMed](#)]
35. Belancio, V.P.; Hedges, D.J.; Deininger, P. Mammalian non-LTR retrotransposons: For better or worse, in sickness and in health. *Genome Res.* **2008**, *18*, 343–358. [[CrossRef](#)] [[PubMed](#)]
36. Payer, L.M.; Burns, K.H. Transposable elements in human genetic disease. *Nat. Rev. Genet.* **2019**, *20*, 760–772. [[CrossRef](#)]
37. Alisch, R.S.; Garcia-Perez, J.L.; Muotri, A.R.; Gage, F.H.; Moran, J.V. Unconventional translation of mammalian LINE-1 retrotransposons. *Genes. Dev.* **2006**, *20*, 210–224. [[CrossRef](#)]
38. Callahan, K.E.; Hickman, A.B.; Jones, C.E.; Ghirlando, R.; Furano, A.V. Polymerization and nucleic acid-binding properties of human L1 ORF1 protein. *Nucleic Acids Res.* **2012**, *40*, 813–827. [[CrossRef](#)]
39. Christian, C.M.; Sokolowski, M.; deHaro, D.; Kines, K.J.; Belancio, V.P. Involvement of Conserved Amino Acids in the C-Terminal Region of LINE-1 ORF2p in Retrotransposition. *Genetics* **2017**, *205*, 1139–1149. [[CrossRef](#)]
40. Christian, C.M.; deHaro, D.; Kines, K.J.; Sokolowski, M.; Belancio, V.P. Identification of L1 ORF2p sequence important to retrotransposition using Bipartite Alu retrotransposition (BAR). *Nucleic Acids Res.* **2016**, *44*, 4818–4834. [[CrossRef](#)]
41. Adney, E.M.; Ochmann, M.T.; Sil, S.; Truong, D.M.; Mita, P.; Wang, X.; Kahler, D.J.; Fenyö, D.; Holt, L.J.; Boeke, J.D. Comprehensive Scanning Mutagenesis of Human Retrotransposon LINE-1 Identifies Motifs Essential for Function. *Genetics* **2019**, *213*, 1401–1414. [[CrossRef](#)] [[PubMed](#)]
42. Martin, S.L.; Li, J.; Epperson, L.E.; Lieberman, B. Functional reverse transcriptases encoded by A-type mouse LINE-1: Defining the minimal domain by deletion analysis. *Gene* **1998**, *215*, 69–75. [[CrossRef](#)] [[PubMed](#)]
43. Feng, Q.; Moran, J.V.; Kazazian, H.H.; Boeke, J.D. Human L1 Retrotransposon Encodes a Conserved Endonuclease Required for Retrotransposition. *Cell* **1996**, *87*, 905–916. [[CrossRef](#)] [[PubMed](#)]
44. Cost, G.J.; Boeke, J.D. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. *Biochemistry* **1998**, *37*, 18081–18093. [[CrossRef](#)] [[PubMed](#)]
45. Mathias, S.L.; Scott, A.F.; Kazazian, H.H.; Boeke, J.D.; Gabriel, A. Reverse transcriptase encoded by a human transposable element. *Science* **1991**, *254*, 1808. [[CrossRef](#)] [[PubMed](#)]
46. Sokolowski, M.; deHaro, D.; Christian, C.M.; Kines, K.J.; Belancio, V.P. Characterization of L1 ORF1p self-interaction and cellular localization using a mammalian two-hybrid system. *PLoS ONE* **2013**, *8*, e82021. [[CrossRef](#)] [[PubMed](#)]
47. Moran, J.V.; Holmes, S.E.; Naas, T.P.; DeBerardinis, R.J.; Boeke, J.D.; Kazazian, H.H., Jr. High frequency retrotransposition in cultured mammalian cells. *Cell* **1996**, *87*, 917–927. [[CrossRef](#)]
48. Martin, S.L.; Bushman, F.D. Nucleic Acid Chaperone Activity of the ORF1 Protein from the Mouse LINE-1 Retrotransposon. *Mol. Cell. Biol.* **2001**, *21*, 467. [[CrossRef](#)]
49. Martin, S.L.; Li, J.; Weisz, J.A. Deletion analysis defines distinct functional domains for protein-protein and nucleic acid interactions in the ORF1 protein of mouse LINE-1. *J. Mol. Biol.* **2000**, *304*, 11–20. [[CrossRef](#)]

50. Basame, S.; Wai-lun Li, P.; Howard, G.; Branciforte, D.; Keller, D.; Martin, S.L. Spatial assembly and RNA binding stoichiometry of a LINE-1 protein essential for retrotransposition. *J. Mol. Biol.* **2006**, *357*, 351–357. [[CrossRef](#)]
51. Freeman, B.T.; Sokolowski, M.; Roy-Engel, A.M.; Smither, M.E.; Belancio, V.P. Identification of charged amino acids required for nuclear localization of human L1 ORF1 protein. *Mob. DNA* **2019**, *10*, 20. [[CrossRef](#)] [[PubMed](#)]
52. Cost, G.J.; Golding, A.; Schlissel, M.S.; Boeke, J.D. Target DNA chromatinization modulates nicking by L1 endonuclease. *Nucleic Acids Res.* **2001**, *29*, 573–577. [[CrossRef](#)] [[PubMed](#)]
53. Baldwin, E.T.; van Eeuwen, T.; Hoyos, D.; Zalevsky, A.; Tchesnokov, E.P.; Sánchez, R.; Miller, B.D.; Di Stefano, L.H.; Ruiz, F.X.; Hancock, M. Structures, functions, and adaptations of the human LINE-1 ORF2 protein. *Nature* **2023**. [[CrossRef](#)] [[PubMed](#)]
54. Dombroski, B.A.; Feng, Q.; Mathias, S.L.; Sassaman, D.M.; Scott, A.F.; Kazazian, H.H., Jr.; Boeke, J.D. An in vivo assay for the reverse transcriptase of human retrotransposon L1 in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* **1994**, *14*, 4485–4492. [[PubMed](#)]
55. Piskareva, O.; Denmukhametova, S.; Schmatchenko, V. Functional reverse transcriptase encoded by the human LINE-1 from baculovirus-infected insect cells. *Protein Expr. Purif.* **2003**, *28*, 125–130. [[CrossRef](#)] [[PubMed](#)]
56. Weichenrieder, O.; Repanas, K.; Perrakis, A. Crystal structure of the targeting endonuclease of the human LINE-1 retrotransposon. *Structure* **2004**, *12*, 975–986. [[CrossRef](#)] [[PubMed](#)]
57. Taylor, M.S.; LaCava, J.; Mita, P.; Molloy, K.R.; Huang, C.R.; Li, D.; Adney, E.M.; Jiang, H.; Burns, K.H.; Chait, B.T. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. *Cell* **2013**, *155*, 1034–1048. [[CrossRef](#)]
58. Clements, A.P.; Singer, M.F. The human LINE-1 reverse transcriptase: effect of deletions outside the common reverse transcriptase domain. *Nucleic Acids Res.* **1998**, *26*, 3528–3535. [[CrossRef](#)]
59. Burke, W.D.; Malik, H.S.; Jones, J.P.; Eickbush, T.H. The domain structure and retrotransposition mechanism of R2 elements are conserved throughout arthropods. *Mol. Biol. Evol.* **1999**, *16*, 502–511. [[CrossRef](#)]
60. Thawani, A.; Ariza, A.J.F.; Nogales, E.; Collins, K. Template and target site recognition by human LINE-1 in retrotransposition. *Nature* **2023**, *81*. [[CrossRef](#)]
61. Eickbush, T.H.; Jamburuthugoda, V.K. The diversity of retrotransposons and the properties of their reverse transcriptases. *Virus Res.* **2008**, *134*, 221–234. [[CrossRef](#)] [[PubMed](#)]
62. Viollet, S.; Monot, C.; Cristofari, G. L1 retrotransposition: The snap-velcro model and its consequences. *Mob. Genet. Elem.* **2014**, *4*, e28907. [[CrossRef](#)] [[PubMed](#)]
63. De Cecco, M.; Ito, T.; Petrashen, A.P.; Elias, A.E.; Skvir, N.J.; Criscione, S.W.; Caligiana, A.; Broccoli, G.; Adney, E.M.; Boeke, J.D. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature* **2019**, *566*, 73–78. [[CrossRef](#)] [[PubMed](#)]
64. Rodriguez-Martin, B.; Alvarez, E.G.; Baez-Ortega, A.; Zamora, J.; Supek, F.; Demeulemeester, J.; Santamarina, M.; Ju, Y.S.; Temes, J.; Garcia-Souto, D. Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. *Nat. Genet.* **2020**, *52*, 306–319. [[CrossRef](#)]
65. Gasior, S.L.; Wakeman, T.P.; Xu, B.; Deininger, P.L. The human LINE-1 retrotransposon creates DNA double-strand breaks. *J. Mol. Biol.* **2006**, *357*, 1383–1393. [[CrossRef](#)] [[PubMed](#)]
66. Cajuso, T.; Sulo, P.; Tanskanen, T.; Katainen, R.; Taira, A.; Hänninen, U.A.; Kondelin, J.; Forsström, L.; Välimäki, N.; Aavikko, M. Retrotransposon insertions can initiate colorectal cancer and are associated with poor survival. *Nat. Commun.* **2019**, *10*, 4022. [[CrossRef](#)]
67. Gilbert, N.; Lutz-Prigge, S.; Moran, J.V. Genomic deletions created upon LINE-1 retrotransposition. *Cell* **2002**, *110*, 315–325. [[CrossRef](#)]
68. Sokolowski, M.; DeFreece, C.B.; Servant, G.; Kines, K.J.; deHaro, D.L.; Belancio, V.P. Development of a monoclonal antibody specific to the endonuclease domain of the human LINE-1 ORF2 protein. *Mob. DNA* **2014**, *5*, 29. [[CrossRef](#)]
69. Christian, C.M.; Kines, K.J.; Belancio, V.P. The importance of L1 ORF2p cryptic sequence to ORF2p fragment-mediated cytotoxicity. *Mob. Genet. Elem.* **2016**, *6*, e1198300. [[CrossRef](#)]
70. Wallace, N.A.; Belancio, V.P.; Deininger, P.L. L1 mobile element expression causes multiple types of toxicity. *Gene* **2008**, *419*, 75–81. [[CrossRef](#)]
71. Rogakou, E.P.; Pilch, D.R.; Orr, A.H.; Ivanova, V.S.; Bonner, W.M. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* **1998**, *273*, 5858–5868. [[CrossRef](#)] [[PubMed](#)]
72. Fernandez-Capetillo, O.; Celeste, A.; Nussenzweig, A. Focusing on foci: H2AX and the recruitment of DNA-damage response factors. *Cell Cycle* **2003**, *2*, 426–427. [[CrossRef](#)] [[PubMed](#)]
73. Celeste, A.; Petersen, S.; Romanienko, P.J.; Fernandez-Capetillo, O.; Chen, H.T.; Sedelnikova, O.A.; Reina-San-Martin, B.; Coppola, V.; Meffre, E.; Difilippantonio, M.J. Genomic instability in mice lacking histone H2AX. *Science* **2002**, *296*, 922–927. [[CrossRef](#)]
74. Sedelnikova, O.A.; Rogakou, E.P.; Panyutin, I.G.; Bonner, W.M. Quantitative detection of (125)IdU-induced DNA double-strand breaks with gamma-H2AX antibody. *Radiat. Res.* **2002**, *158*, 486–492. [[CrossRef](#)] [[PubMed](#)]
75. Valdíglesias, V.; Giunta, S.; Fenech, M.; Neri, M.; Bonassi, S. γ H2AX as a marker of DNA double strand breaks and genomic instability in human population studies. *Mutat. Res.* **2013**, *753*, 24–40. [[CrossRef](#)]
76. Kuo, L.J.; Yang, L.X. Gamma-H2AX—A novel biomarker for DNA double-strand breaks. *Vivo* **2008**, *22*, 305–309.
77. Cisneros-Aguirre, M.; Ping, X.; Stark, J.M. To indel or not to indel: Factors influencing mutagenesis during chromosomal break end joining. *DNA Repair.* **2022**, *118*, 103380. [[CrossRef](#)]

78. Pannunzio, N.R.; Watanabe, G.; Lieber, M.R. Nonhomologous DNA end-joining for repair of DNA double-strand breaks. *J. Biol. Chem.* **2018**, *293*, 10512–10523. [[CrossRef](#)]
79. Kopera, H.C.; Flasch, D.A.; Nakamura, M.; Miyoshi, T.; Doucet, A.J.; Moran, J.V. LEAP: L1 Element Amplification Protocol. *Methods Mol. Biol.* **2016**, *1400*, 339–355.
80. Igoucheva, O.; Alexeev, V.; Yoon, K. Differential cellular responses to exogenous DNA in mammalian cells and its effect on oligonucleotide-directed gene modification. *Gene Ther.* **2006**, *13*, 266–275. [[CrossRef](#)]
81. Huerfano, S.; Ryabchenko, B.; Forstová, J. Nucleofection of expression vectors induces a robust interferon response and inhibition of cell proliferation. *DNA Cell Biol.* **2013**, *32*, 467–479. [[CrossRef](#)]
82. Dewannieux, M.; Esnault, C.; Heidmann, T. LINE-mediated retrotransposition of marked Alu sequences. *Nat. Genet.* **2003**, *35*, 41–48. [[CrossRef](#)] [[PubMed](#)]
83. Honma, M.; Sakuraba, M.; Koizumi, T.; Takashima, Y.; Sakamoto, H.; Hayashi, M. Non-homologous end-joining for repairing I-SceI-induced DNA double strand breaks in human cells. *DNA Repair.* **2007**, *6*, 781–788. [[CrossRef](#)]
84. Repanas, K.; Zingler, N.; Layer, L.E.; Schumann, G.G.; Perrakis, A.; Weichenrieder, O. Determinants for DNA target structure selectivity of the human LINE-1 retrotransposon endonuclease. *Nucleic Acids Res.* **2007**, *35*, 4914–4926. [[CrossRef](#)] [[PubMed](#)]
85. Kazazian, H.H., Jr.; Wong, C.; Youssoufian, H.; Scott, A.F.; Phillips, D.G.; Antonarakis, S.E. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature* **1988**, *332*, 164–166. [[CrossRef](#)] [[PubMed](#)]
86. Brouha, B.; Schustak, J.; Badge, R.M.; Lutz-Prigge, S.; Farley, A.H.; Moran, J.V.; Kazazian, H.H., Jr. Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5280–5285. [[CrossRef](#)]
87. Kirilyuk, A.; Tolstonog, G.V.; Damert, A.; Held, U.; Hahn, S.; Löwer, R.; Buschmann, C.; Horn, A.V.; Traub, P.; Schumann, G.G. Functional endogenous LINE-1 retrotransposons are expressed and mobilized in rat chloroleukemia cells. *Nucleic Acids Res.* **2008**, *36*, 648–665. [[CrossRef](#)]
88. Cao, X.; Zhang, Y.; Payer, L.M.; Lords, H.; Steranka, J.P.; Burns, K.H.; Xing, J. Polymorphic mobile element insertions contribute to gene expression and alternative splicing in human tissues. *Genome Biol.* **2020**, *21*, 185. [[CrossRef](#)]
89. Belancio, V.P.; Roy-Engel, A.M.; Deininger, P. The impact of multiple splice sites in human L1 elements. *Gene* **2008**, *411*, 38–45. [[CrossRef](#)]
90. Gasior, S.L.; Roy-Engel, A.M.; Deininger, P.L. ERCC1/XPF limits L1 retrotransposition. *DNA Repair.* **2008**, *7*, 983–989. [[CrossRef](#)]
91. Yamaguchi, K.; Kajikawa, M.; Okada, N. LINE retrotransposition and host DNA repair machinery. *Mob. Genet. Elem.* **2015**, *5*, 92–97. [[CrossRef](#)] [[PubMed](#)]
92. Suzuki, J.; Yamaguchi, K.; Kajikawa, M.; Ichiyangi, K.; Adachi, N.; Koyama, H.; Takeda, S.; Okada, N. Genetic evidence that the non-homologous end-joining repair pathway is involved in LINE retrotransposition. *PLoS Genet.* **2009**, *5*, e1000461. [[CrossRef](#)] [[PubMed](#)]
93. Pezone, A.; Olivieri, F.; Napoli, M.V.; Procopio, A.; Avvedimento, E.V.; Gabrielli, A. Inflammation and DNA damage: Cause, effect or both. *Nat. Rev. Rheumatol.* **2023**, *19*, 200–211. [[CrossRef](#)] [[PubMed](#)]
94. Lee, J.; Han, K.; Meyer, T.J.; Kim, H.S.; Batzer, M.A. Chromosomal inversions between human and chimpanzee lineages caused by retrotransposons. *PLoS ONE* **2008**, *3*, e4047. [[CrossRef](#)]
95. Cordaux, R.; Batzer, M.A. The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* **2009**, *10*, 691–703. [[CrossRef](#)] [[PubMed](#)]
96. Beroukhim, R.; Mermel, C.H.; Porter, D.; Wei, G.; Raychaudhuri, S.; Donovan, J.; Barretina, J.; Boehm, J.S.; Dobson, J.; Urashima, M. The landscape of somatic copy-number alteration across human cancers. *Nature* **2010**, *463*, 899–905. [[CrossRef](#)]
97. Cai, Y.; Sablina, A.A. Cancer-associated chromosomal deletions: Size makes a difference. *Cell Cycle* **2016**, *15*, 2850–2851. [[CrossRef](#)]
98. Baca, S.C.; Prandi, D.; Lawrence, M.S.; Mosquera, J.M.; Romanel, A.; Drier, Y.; Park, K.; Kitabayashi, N.; MacDonald, T.Y.; Ghandi, M. Punctuated evolution of prostate cancer genomes. *Cell* **2013**, *153*, 666–677. [[CrossRef](#)]
99. McKerrow, W.; Wang, X.; Mendez-Dorantes, C.; Mita, P.; Cao, S.; Grivainis, M.; Ding, L.; LaCava, J.; Burns, K.H.; Boeke, J.D. LINE-1 expression in cancer correlates with p53 mutation, copy number alteration, and S phase checkpoint. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2115999119. [[CrossRef](#)]
100. Chuang, N.T.; Gardner, E.J.; Terry, D.M.; Crabtree, J.; Mahurkar, A.A.; Rivell, G.L.; Hong, C.C.; Perry, J.A.; Devine, S.E. Mutagenesis of human genomes by endogenous mobile elements on a population scale. *Genome Res.* **2021**, *31*, 2225–2235. [[CrossRef](#)]
101. Stevens, C.; Hightower, A.; Buxbaum, S.G.; Falzarano, S.M.; Rhie, S.K. Genomic, epigenomic, and transcriptomic signatures of prostate cancer between African American and European American patients. *Front. Oncol.* **2023**, *13*, 1079037. [[CrossRef](#)] [[PubMed](#)]
102. Jacobs, P.A.; Browne, C.; Gregson, N.; Joyce, C.; White, H. Estimates of the frequency of chromosome abnormalities detectable in unselected newborns using moderate levels of banding. *J. Med. Genet.* **1992**, *29*, 103–108. [[CrossRef](#)] [[PubMed](#)]
103. Ardeljan, D.; Taylor, M.S.; Ting, D.T.; Burns, K.H. The Human Long Interspersed Element-1 Retrotransposon: An Emerging Biomarker of Neoplasia. *Clin. Chem.* **2017**, *63*, 816–822. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.