

LINE-1 retrotransposons in cancer: Investigating human cervical carcinogenesis and developing an innovative transgenic mouse model for measuring somatic LINE-1 activity

Ph.D. Thesis

Réka Karkas, MD

Supervisor: Lajos Mátés, Ph.D.



**HUN
REN**



Doctoral School of Multidisciplinary Medical Sciences

Albert Szent-Györgyi Medical School

University of Szeged

Szeged

2025

LIST OF PUBLICATIONS

The publications served as the basis of the Ph.D. thesis:

I. Réka Karkas, Khaloon Sadiq Ahmed Abdullah*, Laszlo Kaizer, Ádám Ürmös, Lilla Tiszlavicz, Tibor Pankotai, Istvan Nagy, Lajos Mátés, Farkas Sükösd:

LINE-1 ORF1p is a Promising Biomarker in Cervical Intraepithelial Neoplasia Degree Assessment International. Journal of Gynecological Pathology (2024) DOI: 10.1097/PGP.0000000000001035 (Q2)

II. Anna Georgina Kopasz, Dávid Zsolt Pusztai, **Réka Karkas**, Liza Hudoba, Khaldoon Sadiq Ahmed Abdullah, Gergely Imre, Gabriella Pankotai-Bodó, Ede Migh, Andrea Nagy, András Kriston, Péter Germán, Andrea Bakné Drubi, Anna Molnár, Ildikó Fekete, Virág Éva Dani, Imre Ocsovszki, László Géza Puskás, Péter Horváth, Farkas Sükösd, Lajos Mátés

A versatile transposon-based technology to generate loss- and gain-of-function phenotypes in the mouse liver BMC BIOLOGY 20 : 1 Paper: 74 , 17 p. (2022) BMC Biology 20, 74 (2022). DOI: 10.1186/s12915-022-01262-x (D1)

III. Lajos Mátés, **Réka Karkas**, Sadiq Ahmed Abdullah Khaldoon, Andrea Mátés-Nagy, Gergely Imre, Anna Georgina Kopasz, Ildikó Fekete, Péter Horváth, Farkas Sükösd
Measurement of somatic Ll retrotransposition activity

Patent, (2023) International Application Number: PCT/HU2023/050014

Other publications published during the Ph.D. scholarship:

IV. Gergely Imre, Bertalan Takács, Erik Czipa, Andrea Bakné Drubi, Gábor Jaksa, Dóra Latinovics, Andrea Nagy, **Réka Karkas**, Liza Hudoba , Bálint Márk Vásárhelyi, Gabriella Pankotai-Bodó, András Blastyák, Zoltán Hegedűs, Péter Germán, Balázs Bálint, Khaldoon Sadiq Ahmed Abdullah, Anna Georgina Kopasz, Anita Kovács, László G. Nagy, Farkas Sükösd, Lajos Mátés *Prolonged activity of the transposase helper may raise safety concerns during DNA transposon-based gene therapy. Molecular Therapy Methods Clin. Dev.* 29 pp. 145-159. , 15 p. (2023) DOI: 10.1016/j.omtm.2023.03.003 (Q1)

V. Lajos Mátés, Anna Georgina Kopasz, **Réka Karkas**, Khaloon Sadiq Ahmed Abdullah, Andrea Mátés-Nagy, Gergely Imre, Ildikó Fekete, Péter Horváth, Farkas Sükösd:

Versatile vector system for directed gene expression modifications

Patent, (2023) International Application Number: PCT/HU2023/050013

LIST OF ABBREVIATIONS

APOBEC3 - Apolipoprotein B mRNA Editing Enzyme Catalytic Subunit 3
 bGH - Bovine Growth Hormone
 BRCA1 - Breast Cancer 1
 CDS - Coding Sequence
 CIN - Cervical Intraepithelial Neoplasia
 CMV - Cytomegalovirus
 DMSO - Dimethyl Sulfoxide
 DMBA/TPA - 7,12-Dimethylbenzo-Anthracene/ 12-o-Tetradecylphorbol -13-Acetate
 EBSS - Earle's Balanced Salt Solution
 EGFP - Enhanced Green Fluorescent Protein
 EEF1A1 - Eukaryotic Translation Elongation Factor 1 Alpha 1
 EMMA - European Mouse Mutant Archive
 EN - Endonuclease
 ESCRT - Endosomal Sorting Complex Required for Transport
 FACS - Fluorescence-Activated Cell Sorting
 FAH - Fumarylacetoacetate Hydrolase
 FICZ - 6-Formylindolo[3,2-b]carbazole
 FFPE - Formalin-Fixed Paraffin-Embedded
 GFP - Green Fluorescent Protein
 HADHA - Hydroxyacyl-CoA Dehydrogenase Trifunctional Multienzyme Complex Alpha Subunit
 HADHB - Hydroxyacyl-CoA Dehydrogenase Trifunctional Multienzyme Complex Beta Subunit
 Hpd - Hydroxyphenylpyruvate Dioxygenase
 HSVTK - Herpes Simplex Virus Thymidine Kinase
 IHC - Immunohistochemistry
 ITR - Inverted Terminal Repeat
 IVIS - In Vivo Imaging System

KPNA - Karyopherin-Alpha (KPNA family of nuclear transport proteins)
 L1 - LINE-1 Retrotransposons
 LTR - Long Terminal Repeat
 Luc - Luciferase
 MCS - Multiple Cloning Site
 MeIQx - 2-Amino-3,8-Dimethylimidazo [4,5-f]Quinoxaline
 MOV10 - Moloney Leukemia Virus 10
 MRN - Mre11-Rad50-Nbs1 Complex
 NHEJ - Non-Homologous End Joining
 NTBC - Nitisinone
 ORF1p - Open Reading Frame 1 Protein
 ORF2p - Open Reading Frame 2 Protein
 PCNA - Proliferating Cell Nuclear Antigen
 qPCR - Quantitative Polymerase Chain Reaction
 RAD18 - Radiation-sensitive 18
 RISC - RNA-Induced Silencing Complex
 RNP - Ribonucleoprotein Complex
 ROS - Reactive Oxygen Species
 RT - Reverse Transcriptase
 RT-qPCR - Reverse Transcription Quantitative Polymerase Chain Reaction
 SB - Sleeping Beauty (Transposon System)
 SB100 - Hyperactive Sleeping Beauty Transposase
 SINEs - Short Interspersed Nuclear Elements
 SIRT6 - Sirtuin 6
 T2A - Thosea asigna virus 2A
 TMA - Tissue Microarray
 TREX1 -Three Prime Repair Exonuclease1
 TSD - Target Site Duplication
 TT1 - Type I Tyrosinemia
 TT3 - Type III Tyrosinemia
 UTR - Untranslated Region
 WT - Wild Type

TABLE OF CONTENTS

List of publications.....	2
List of abbreviations.....	3
Abstract.....	5
1 Introduction.....	7
1.1 Overview of Transposable Elements.....	7
1.2 Retrotransposons.....	8
1.2.1 LINE-1 Retrotransposons.....	8
1.2.2 Consequences of germline retrotransposition activity of LINE-1s.....	11
1.2.3 Consequences of somatic retrotransposition activity of LINE-1s.....	12
1.2.4 Regulation of retrotransposition activity of LINE-1s.....	13
1.2.5 Factors affecting LINE-1 activity.....	18
1.2.6 Assays to Investigate LINE-1 Activity.....	19
2 Aims of the thesis.....	22
3 Materials and methods.....	23
3.1 Preparation of the Human Tissue Specimens.....	23
3.2 Animal Care and Maintenance.....	23
3.3 Immunohistochemistry.....	24
3.4 Creation of plasmid constructions.....	24
3.5 Hydrodynamic tail vein injection.....	25
3.6 Administration of screened compounds for LINE-1 inductor properties.....	26
3.7 Stereomicroscope Imaging.....	26
3.8 In vivo bioluminescence imaging.....	27
3.9 Liver perfusion and hepatocyte isolation.....	27
3.10 FACS-Based Measurement of EGFP-Positive Hepatocytes.....	27
3.11 qPCR Strategy for Detecting Intronless ORFeus Copies from Retrotransposition Events.....	28
3.12 Data visualization and statistics.....	28
4 Results.....	29
4.1 Investigation of LINE-1 ORF1 Protein Expression in FFPE Human Specimens.....	29
4.2 Measurement of somatic LINE1 retrotransposition activity in mice.....	32
4.2.1 Creation of the mouse model of human type I tyrosinemia.....	33
4.2.2 Therapeutic gene delivery-coupled somatic transgenesis in the mouse liver.....	34
4.2.3 Assessment of Chemical-Induced LINE-1 Retrotransposition Activity.....	39
4.2.4 Further methods for quantifying reporter retrotransposition events.....	42
5 Discussion.....	43
6 Conclusion.....	48
7 Acknowledgements.....	49
8 Funding.....	50
9 Bibliography.....	51

ABSTRACT

LINE-1 retrotransposons are the only active transposable elements in the human genome. They encode ORF1 and ORF2 proteins, essential for their “copy-and-paste” self-propagation. These elements entered the genomes of mammalian ancestors around 70 million years ago and have since successfully co-evolved with us. LINE-1 elements now constitute 17% of the human genome, with approximately 500,000 copies, of which 100-150 are retrotransposition-competent. LINE-1 retrotransposons have a role in genomic evolution by fostering genetic diversity through mechanisms like creating gene duplications, exon shuffling, and alternative splicing by inserting themselves into different genomic locations.

Previously, LINE-1 activity was believed to be limited to the germline. However, studies using immunostainings for the LINE-1 ORF1 protein have shown that carcinomas often display LINE-1 element expression (transcription of LINE-1 mRNA and translation of ORF1p), while normal human tissues exhibit virtually zero activity. New-generation sequencing has revealed that somatic retrotransposition events can be confirmed in a significant proportion of cancers, at a rate that closely matches the positivity rate for ORF1p staining. We conducted a comprehensive analysis of ORF1 expression during the stages of cervical carcinogenesis by examining 143 FFPE specimens of the human uterine cervix. ORF1p expression was evaluated and compared with conventional markers, Ki67 and p16. ORF1p immunopositivity was detected in 110 out of 112 neoplastic and dysplastic specimens (cervical intraepithelial neoplasia grade I to grade III [CIN I-III] and invasive cancers), whereas 19 of 24 normal cervical specimens completely lacked ORF1p expression. ORF1p expression showed a progressive increase with advancing CIN grades. CIN I was characterized by weaker ORF1p expression in the lower one or two-thirds of the cervical epithelium, whereas CIN II demonstrated moderate to strong ORF1p expression spanning the lower two-thirds, while CIN III was characterized by pronounced transepithelial ORF1p immunopositivity. Invasive cancer cases showed widespread ORF1p expression. This makes ORF1p an ideal candidate early diagnostic marker for cervical neoplasia.

LINE-1 retrotransposition has played a prominent role in the genomic evolution of species over millions of years by promoting genetic diversity and structural variation; meanwhile, it also contributes to tumor evolution within the lifespan of an individual organism. New-generation sequencing has uncovered several LINE-1-induced tumor driver

mutations, making it not just a marker but, in certain cases, the driver of carcinogenesis. Due to these potentially harmful effects, somatic cells have numerous mechanisms to suppress LINE-1 activity. If these defenses are compromised, LINE-1 activation can initiate tumor evolution. Environmental chemicals might impair these defense mechanisms, leading to LINE-1 activation and subsequent tumorigenesis. To date, investigating this subject has posed considerable difficulties, motivating the development of an *in vivo* model that facilitates this research.

The ORFeus reporter system is a synthetic LINE-1 linked to an EGFP reporter cassette, which is particularly suitable for measuring LINE-1 retrotransposition frequency. We utilized the non-autonomous version of the reporter, which contains only the ORF1 sequence, and the ORF2 protein provided endogenously by the cell. If the element goes through successful retrotransposition, the EGFP reporter gene is expressed, and it persists in all progeny of the affected cell. In principle, tumor-derived cell lines could also be used to test somatic LINE1 retrotransposition activity, however, results generated using such cells do not necessarily reflect the *in vivo* situation due to the partial functioning of LINE-1 defense mechanisms in tumor cell lines. To address this, we developed a mouse model enabling the measurement of somatic LINE1 activity, *in vivo*, which was previously unattainable, since the application of conventional germline transgenic mice is problematic due to early embryonic retrotransposition, resulting in inherent GFP positivity in all progeny cells.

Our strategy is to somatically introduce the reporter at later developmental stages, targeting the liver in a mouse strain with type I tyrosinemia. These mice lack the functional *Fah* gene, essential for tyrosine metabolism, thus, its reintroduction rescues hepatocytes from lethal toxicity. Liver regeneration is initiated from the genetically corrected liver cells, which allows us to create a transgenic organ for the "gene of interest" introduced alongside the *Fah* therapeutic gene. Using the Sleeping Beauty transposon system, we integrated the *Fah* gene along with the ORFeus reporter into the genomes of the hepatocytes of mice lacking *Fah* gene. We monitored LINE-1 activity by treating these reporter-bearing mice with specific chemicals. We observed increased reporter activation by Decitabine, a DNA hypomethylating agent, and FICZ, a tryptophan metabolite produced by the microbiome. Our results indicate that these substances can influence somatic LINE-1 activity, validating our technological toolkit, which can provide a novel method for assessing the unknown genotoxic potential of environmental chemicals. This approach may help reveal previously unrecognized etiological factors contributing to the development of cancers with unclear origins, shedding light on the "black box" of cancer etiology.

1 INTRODUCTION

Transposable elements (TEs), once considered "junk" DNA, were historically underappreciated. They were presumed to be unimportant, and their high copy number and variability caused unique technical challenges upon investigation since they occupy almost half (~45%) of the human genome. However, recent advancements in genomic technologies and growing interest in the human mobilome have revealed significant roles in genome evolution, diversity, and disease. Transposable elements are now recognized for their roles in shaping the genome and contributing to various pathologies, including cancer, highlighting their importance in human health and disease [1], [2].

1.1 Overview of Transposable Elements

TEs are broadly classified into two major classes based on their transposition mechanism: Class I (retrotransposons), which move via an RNA intermediate through a copy-and-paste mechanism, and Class II (DNA transposons), which transpose directly as DNA using a cut-and-paste mechanism. This classification reflects the fundamental differences in how TEs propagate and influence genome architecture. Class II elements represent a smaller (~4%) proportion of the human genome, and they are no longer functional, not capable of transposition [3]. Although there are examples of human genes that have been adopted by DNA transposons, such as recombinase-activating gene (RAG) proteins, which have been taken from DNA transposons to produce antibody diversity. They facilitate V(D)J recombination for antigen recognition by lymphocytes [4]. Another important aspect of DNA transposons, as discussed later in this thesis, is their utility in genetic engineering, since their reconstructed, sequence-optimized versions have become powerful biotechnological tools. For example, the Sleeping Beauty transposon, reverse-engineered from salmonid fish [5], and the PiggyBac transposon system, domesticated from the cabbage looper moth, are widely used in genetic engineering applications [6] and are even being investigated in clinical trials for gene therapy, demonstrating their potential to revolutionize medical treatments and research [7], [8]. Their high efficiency, the stable chromosomal gene transfer they mediate, their structural simplicity, large cargo capacity, and non-viral nature, thus less immunogenicity, make them valuable tools both for research and therapeutic applications [9], [10].

1.2 Retrotransposons

Recognizing their abundance and significant role in shaping the human genome, my thesis focuses mainly on retrotransposons (Class I elements). Retrotransposons can be further divided into two further main groups: LTR (long terminal repeat) retrotransposons and non-LTR retrotransposons. LTR retrotransposons have long terminal repeats (LTR) are characterized by the presence of LTR sequences at both ends, which play a crucial role in their transcription and integration at each end and They share structural and functional similarities with retroviruses. However, retroviruses differ from them by lacking the *envelope* (ENV) gene that enables retroviruses to bind target cells and penetrate the cell membrane [11]. LTR retrotransposons contain open reading frames (ORFs) similar to retroviruses, such as *pol* and *gag* ORFs, and encode enzymes required for transposition. Endogenous retroviruses (ERVs), are the most ubiquitous LTR retrotransposons in humans. These ERV sequences are the consequence of vertical transmission from ancestral exogenous viral infections in germline cells [12].

Non-LTR retrotransposons (polyA retrotransposons, or target-primed (TP) retrotransposons) have a structure resembling protein-coding genes, with a promoter at the beginning and a polyA sequence at the end. They are well-known for their enormous success in populating the mammalian genome [13], [14]. The two major types of non-LTR retrotransposons are the long interspersed nuclear elements (LINEs), comprising 17% of the human genome, while short interspersed nuclear elements (SINEs) occupy approximately 13% [15].

SINEs are non-autonomous retrotransposons; they are transcribed by RNA polymerase III, and they are dependent on the machinery of other mobile elements for their retrotranscription because they lack a functional reverse transcriptase [16]. Despite their non-autonomous nature, their most abundant representatives, Alu elements, have managed to propagate to over 1 million copies. These elements have been shown to modulate gene expression at the post-transcriptional level through their involvement in alternative splicing, RNA editing, and translation regulation [17].

1.2.1 LINE-1 Retrotransposons

LINE-1 (long interspersed nuclear element-1) is a type of non-LTR retrotransposon that represents approximately 17% of the human genome. These elements entered the genomes of

mammalian ancestors around 70 million years ago and have since co-evolved with their hosts. LINE-1 elements are encoding all the necessary machinery for their own “copy and paste” transposition [12]. Each LINE-1 element is about 6,000 base pairs long and contains a 5’ untranslated region (5’ UTR) with endogenous promoter activity, two open reading frames (ORFs: ORF1 and ORF2) that encode proteins essential for its mobilization, and it ends with a 3’ UTR containing polyA [13] (**Fig. 1**).



Figure 1. The schematic structure of human LINE-1 retrotransposons. 5’ untranslated region (5’ UTR) with endogenous promoter activity (red arrow) ORF1 (orange) possesses a leucine zipper (LZ) domain utilized in protein-protein interactions and a nucleic acid binding domain (BD). ORF2 (blue) has an apurinic/apyrimidinic endonuclease (EN) domain and a reverse-transcriptase (RT) domain. LINE-1 sequences end with a polyA tail (yellow). Dark blue rectangles represent target site duplications on each side of typical insertions [18].

Transcription of the LINE-1 RNA starts from a strong promoter in the 5’ UTR. The transcript leaves the nucleus and is translated in the cytoplasm. The product of the two ORFs are the ORF1 protein (ORF1p), which is a 40 kDa-sized RNA-binding protein with chaperone activity [19], and the ORF2p is a 150 kDa protein with endonuclease and reverse transcriptase features. After their translation, these proteins form a ribonucleoprotein complex (RNP) [20] with their template transcript at the cytoplasm (**Fig. 2**) and are transported back to the nucleus passively in dividing cells by the breakdown of the nuclear membrane [21]. However, reports indicate that L1 retrotransposition also occurs in nondividing cells, which supports the potential role of active nuclear transport of L1 RNP [22]. It is suggested that nuclear pore complexes have a role in nucleocytoplasmic trafficking of the LINE-1 RNPs since ORF1p interacts with all members of the karyopherin- α (KPNA) family [23]. Evidence also suggests that the membrane-associated endosomal sorting complex required for transportation (ESCRT) is also participating in the nuclear import of RNPs [24].

There is a cell cycle bias for localization and functional dynamics of retrotransposition events: the RNPs can effectively reach the nucleus during mitosis, while the insertion events happen mostly at the S phase [21], [25].

At the nucleus, ORF2p facilitates the reverse transcription of LINE-1 RNA to cDNA and its integration to a new genomic location throughout a process called target-primed reverse transcription (TPRT) (**Fig. 2**). During the TPRT, the endonuclease domain creates a nick at the target site. The target DNA sequence contains a polyT tract, which base pairs with the polyA sequence of the LINE-1 RNA. Then, the reverse transcriptase domain of the ORF2 synthesizes a DNA copy using the RNA template. The second DNA strand is then created using the cDNA as a template; moreover, the target DNA sequences at both ends are filled in and a target site duplication (TSD) is generated [26], [27].

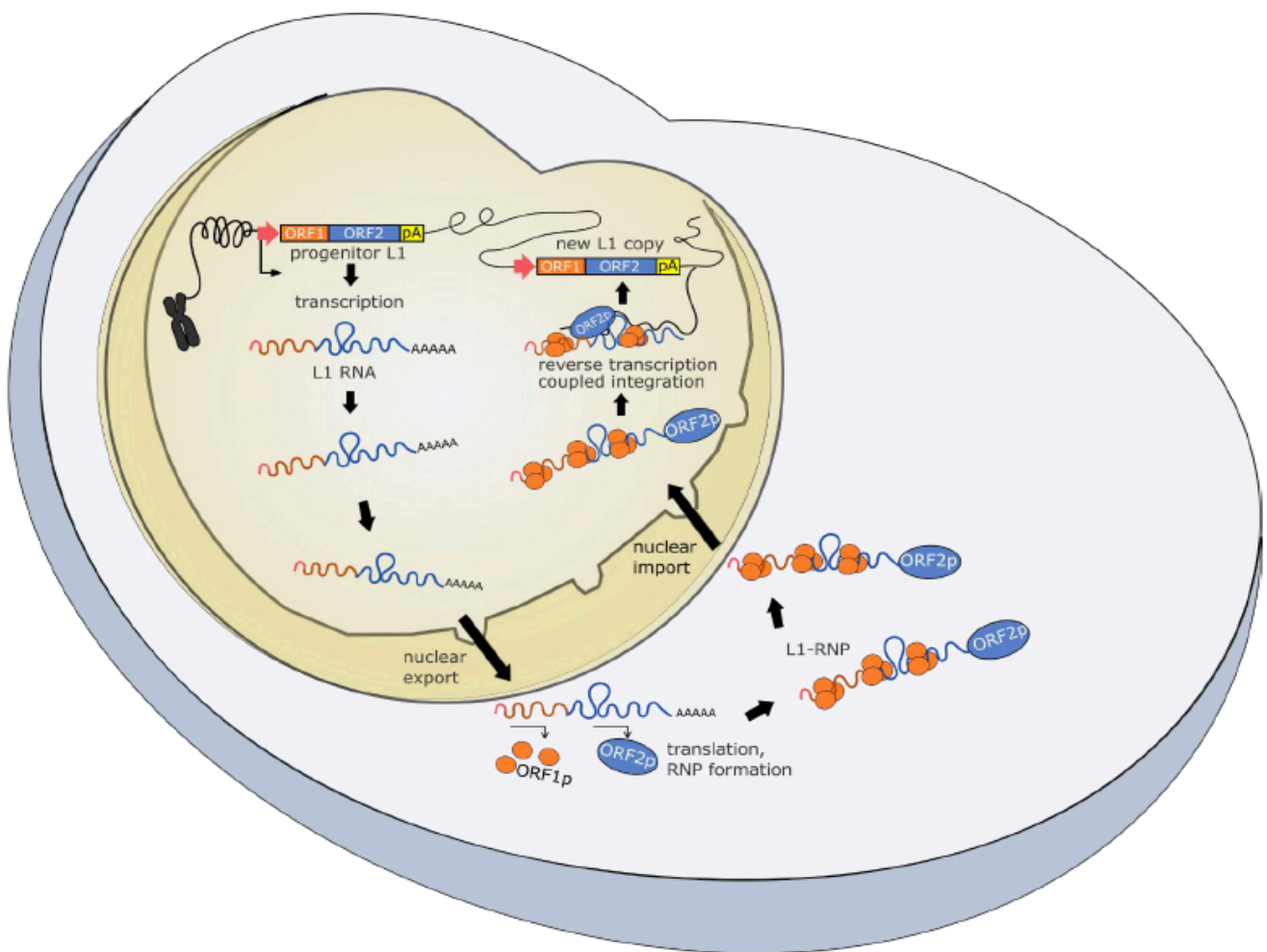


Figure 2. Schematic representation of the retrotransposition cycle. Firstly, the progenitor LINE-1 mRNA is transcribed and then transported to the cytoplasm, where ORF1 proteins (orange) and ORF2 proteins (blue) are translated. The LINE-1 mRNA immediately forms a ribonucleoprotein (RNP) complex with their protein products. The RNP complex is transported into the nucleus, where the target primed reverse transcription (TPRT) occurs, and the produced cDNA is incorporated into a new genomic position by the ORF2p.

The fundamental characteristics of LINE-1 biology are identical in mice and humans. LINE-1 elements comprise approximately 17% and 20% of the human and mouse genomes, respectively [28], [29]. Both species allow for minimal activity of LINE-1 elements in the germline; however, this activity is inhibited in the soma under physiological conditions, except for some epithelial and neuronal cells [30]. However, during tumorigenic processes, LINE-1 activity is activated in somatic cells. Moreover, LINE-1 elements from mice are capable of functioning in human cells, and vice versa [31].

1.2.2 Consequences of germline retrotransposition activity of LINE-1s

The human genome currently contains more than 500,000 copies of LINE-1 elements as a result of their successful multiplication abilities. However, most of these copies are truncated and incapable of retrotransposition. Only a small subset, approximately 100–150 LINE-1 loci, remain active in the average human genome. Among these, there are a few highly active copies, which make up around 5–10% of active elements, known as "hot-LINE-1s," which are particularly prone to retrotransposition [32]. These active LINE-1 elements are usually transcriptionally repressed. Under non-pathological settings, a well-regulated transposition activity is present but limited to the genome of germ cells. These *de novo* retrotransposition events in the germline are able to cause heritable genomic variations and, in rare cases, hereditary genetic diseases [33].

LINE-1 contribute to evolution by promoting the formation of new gene combinations and protein variants, thereby increasing genetic diversity. One key mechanism is exon retroshuffling, which can happen if a LINE-1 insertion within a gene is transcribed, and its own weak polyadenylation signal is bypassed and the transcription is terminated at the pA signal of the gene, instead. This process can lead to the incorporation of exons from the original gene into a new genomic location during a subsequent retrotransposition event, resulting in the creation of a novel gene with potentially altered functionality [34], [35].

The first published example of a LINE-1 insertion causing disease was reported in 1988 by Kazazian et al. A *de novo* LINE-1 insertion disrupted the factor VIII gene, leading to hemophilia A [36]. This discovery provided crucial evidence that transposable elements could actively contribute to genetic diseases. Up till now, there are at least 100 examples in the literature for single-gene diseases caused by retrotransposon insertions [37]. According to

genomic comparisons, the rate of new germline retrotransposition events is estimated to be roughly one new insertion per 63 newborns [38], [39].

1.2.3 Consequences of somatic retrotransposition activity of LINE-1s

To preserve the integrity of the genome of somatic cells in differentiated tissues, LINE-1 retrotransposition is rigorously suppressed by many defense mechanisms at each step of the retrotransposition cycle, keeping its activity virtually zero. Somatic cells are only permissive with some retrotransposition activity during early embryonic development [40], [41]. In differentiated cells, LINE-1 activity is present only sporadically, McKerrow and their colleagues observed some weak activity in normal epithelial cells and neurons, which increased modestly with age [30]. On the other hand, if this activity is dysregulated, it could generate significant genomic instability. Under undesirable circumstances, such as chronic inflammation, infections, and aging, the defense mechanisms can weaken and permit LINE-1 to get transcribed, translated, and even reverse-transcribed and integrated [42].

The activity of LINE-1 elements has profound effects on the human genome. By inserting themselves into new locations, they can disrupt gene function, alter gene expression, and contribute to genomic instability. LINE-1 insertions can occur within genes, potentially leading to loss-of-function mutations or altered splicing patterns. They can also insert themselves into regulatory regions, affecting the expression of nearby genes. LINE-1 integrations can cause complex translocations, large-scale duplications, and the deletion of megabase-scale chromosomal regions, which may result in the loss of tumor-suppressor genes [42], [43].

The expression of ORF2 alone could induce genomic instability due to its intrinsic activity, independent of successful LINE-1 integration. Its endonuclease activity may lead to chronic DNA damage by introducing DSBs and mobilizing Alu elements, further exacerbating genomic instability [44], [45]. Oncogenes can also be amplified through breakage-fusion-bridge cycles that are generated by somatic retrotransposition [46]. There are numerous published examples about cancer-causing somatic insertions of LINE-1 sequences in a wide range of tumors, and recently a robust study has been published by the framework of the Pan-Cancer Analysis of Whole Genomes (PCAWG) project, in which 2,954 cancer genomes were analyzed, and the findings were supplemented with rearrangement,

transcriptomic, and copy-number data. The results of the study validate the significance of the cancer-driving rearrangements caused by retrotransposition events [43].

In addition to self-mobilization, LINE-1 encoded proteins can facilitate the retrotransposition of other non-autonomous elements in trans, such as Alu and SVA elements, as well as processed pseudogenes. This capability of LINE-1 proteins significantly contributes to genomic diversity and structural variation [47], [48]. However, the LINE-1-encoded proteins have a strong *cis* preference to their encoding mRNA. This preference may allow LINE-1s to stay retrotransposition competent even in the excess of nonfunctioning LINE-1 and other cellular RNA sequences [49].

1.2.4 Regulation of retrotransposition activity of LINE-1s

As mentioned above, the genome integrity would be seriously threatened by the unrestricted retroposition of LINE-1 elements; several defense mechanisms have developed to keep them under tight control. At every step of ontogenesis, a variety of cellular mechanisms act to strictly regulate the activity of LINE-1 retrotransposons [50], [51]. Highlighted factors are summarized in **Fig. 3**.

In the genome of germline cells the most important factors are Piwi-interacting RNAs (piRNAs), which silence transposons both transcriptionally and post-transcriptionally. Nuclear PIWI proteins are directed by piRNAs to nascent transposon transcripts and produce heterochromatin on LINE-1 sequences by DNA or histone methylation, which is inherited by somatic cells. On the other hand, piRNA-guided PIWI slicing acts in the cytoplasm, where PIWI proteins are orchestrated by piRNAs to cut target transposon transcripts [52], [53].

In somatic cells, the most critical LINE-1 regulatory mechanisms are operating at the transcriptional level, including DNA methylation and heterochromatinization patterns inherited from the germline. Beyond this, there are additional mechanisms in somatic cells, such as longevity-regulating proteins from the Sirtuin family. They also participate in the epigenetic transcriptional repression of LINE-1s. Specifically, SIRT6 is able to mono-ADP-ribosylate the nuclear corepressor KAP1, which contributes to transcriptionally repressive heterochromatin formation on LINE-1 sequences [54]. Meanwhile, SIRT7 facilitates the anchoring of LINE-1s to the nuclear lamina [55].

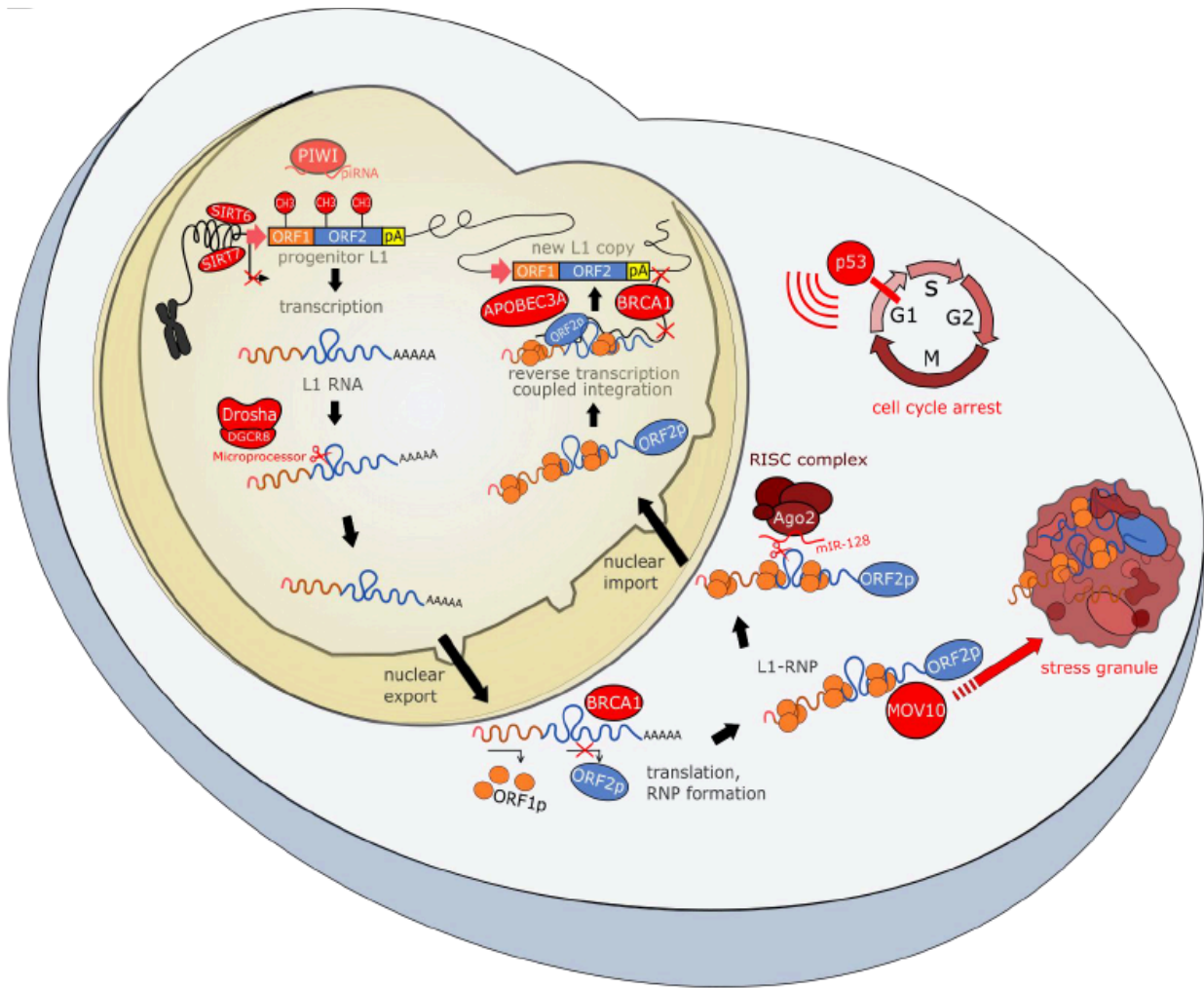


Figure 3. Schematic representation of the major cellular mechanisms acting against LINE-1 retrotransposition. At the nucleus, progenitor LINE-1 DNA, heterochromatinization, DNA methylation, and other factors reduce chromatin availability. LINE-1 RNA is depleted by microRNA pathways. At the cytoplasm, translation can be inhibited, or the LINE-1 RNP can be transferred to stress granules. *De novo* integrations are avoided by numerous factors involved in DNA repair. LINE-1 components can be detected by sensors, which can lead to cell cycle arrest.

The nucleosomal and remodeling deacetylase multiprotein complex (NuRD), stabilized by retinoblastoma (Rb) protein, also acts by decreasing chromatin availability on LINE-1 sequences. This complex promotes heterochromatinization by binding the promoter in the 5' UTR of the LINE-1 element [56].

Several factors are acting on the restriction of LINE-1 transcripts at the post-transcriptional level. There are LINE-1-specific mechanisms, while others are related to interferon-induced mechanisms in response to foreign nucleic acids. RNA interference (RNAi) machinery plays

a crucial role in suppressing LINE-1 activity. MicroRNA-128 (miR-128) acts as a key effector molecule, interacting with Argonaute (Ago) protein to form the RNA-induced silencing complex (RISC). miR-128 binds to the ORF2 region of the LINE-1 transcript, leading to its degradation [57]. Furthermore, the Microprocessor complex, responsible for processing precursor miRNAs into mature forms, might directly cleave a pri-miRNA-like hairpin structure within the LINE-1 5'UTR, repressing its activity [42], [58].

Due to their shared evolutionary origins, there is significant overlap in the regulatory mechanisms regulating retroelements and the cellular response to viral infections. Both can stimulate IFN signaling, which activates intracellular antiviral defense pathways [59]. With aging, in senescent cells, LINE-1 elements become transcriptionally derepressed and activate a type-I interferon (IFN-I) response [60]. While inhibition of LINE-1 reverse transcription in mice downregulated IFN-I activation and age-associated inflammation [61].

RNAse L (ribonuclease L), activated by the interferon response, significantly reduces ORF1 and ORF2 protein expression by cleaving LINE-1 mRNA. This cleavage activates RIG-I (retinoic acid-inducible gene I) and MDA5 (melanoma differentiation-associated protein 5) sensors, further stimulating interferon production and immune responses [62].

The GAIT (gamma-interferon-activated inhibitor of translation) complex, which includes EPRS (glutamyl-prolyl tRNA synthetase), also represses mRNA translation by interacting with translation initiation factors to block the translation process [63]. Additionally, the three-prime repair exonuclease I (TREX1) enzyme, activated by IFN1, facilitated the proteasome-mediated protein degradation of ORF1 and attenuated the L1-induced DNA nicking, resulting in altered subcellular localization of ORF1p [60], [64].

MOV10 (Moloney leukemia virus 10 homolog) helicase is another strong LINE-1 inhibitor. MOV10 interacts with other antiviral proteins to transfer the LINE-1 RNPs to processing bodies (P-bodies) that play a role in the storage, translational repression, or decay of mRNAs. MOV10 also has an inhibitory effect on LINE-1 at the late phase of the retrotransposition; it prevents RNA-DNA heteroduplex formation [65].

TUT4/7 uridylyltransferases are cooperating with MOV10 function. TUT7 adds uridine residues that inhibit the reverse transcription of LINE-1 mRNAs in the nucleus, whereas uridylation by TUT4 destabilizes mRNAs in the cytoplasm [66].

SAMHD1 (SAM domain and HD domain-containing protein 1) is another antiviral factor that inhibits LINE-1 by interaction with ORF2, preventing reverse transcription and reducing ORF2 protein expression. Inhibitory activity by SAMHD1 is regulated by phosphorylation, and it functions in both the nucleus and cytoplasm, where it contributes to the formation of stress granules [67].

The APOBEC3 (apolipoprotein B mRNA editing enzyme catalytic subunit 3) protein family, including hA3A (human APOBEC3A), hA3B (human APOBEC3B), hA3C (human APOBEC3C), hA3F (human APOBEC3F), and hA3DE (human APOBEC3DE), inhibits LINE-1 through deamination-dependent and independent mechanisms. For example, hA3A is responsible for the conversion of cytosine to uracil in the first strand of the LINE-1 cDNA transcript, which ultimately results in destruction [68]. On the other hand, hA3C and hA3DE block the process by interacting with LINE-1 RNP and ORF1 [67[69], [70], [71].

The final stage of LINE-1 retrotransposition, when the integration into the host genome happens, is another critical target where defense mechanisms act. There is a significant amount of crosstalk between these processes and various DNA repair pathways.

Studies have shown that proteins involved in repairing single- and double-stranded DNA breaks also play a role in inhibiting retrotransposition. For instance, defects in non-homologous end-joining (NHEJ) repair factors, such as DNA-PKCs and XRCC4, can increase LINE-1 retrotransposition [72].

Furthermore, factors involved in nucleotide excision repair (NER), which is responsible for repairing single-stranded breaks, also impact retrotransposition. They limit insertional mutations and prevent potentially mutagenic large genomic duplications at retrotransposon insertion sites [73].

Postreplicative DNA repair pathways also contribute significantly to LINE-1 inhibition. The ATM (Ataxia Telangiectasia Mutated) gene activation triggers a cascade of repair factors, including CHK2, p53, BRCA1, and the MRN complex, which are involved in repairing post-replicative DNA damage and regulating the cell cycle. ATM deficiency leads to increased retrotransposition activity, resulting in a higher copy number and greater length of the *de novo* insertions [74].

Factors like FANCD2 and SLX4, associated with Fanconi Anemia, also affect LINE-1 activity. SLX4 hinders the spread of the LINE-1 retrotransposition complex at the replication fork, while FANCD2 activates SLX4 and other repair factors, including the translesion synthesis factor RAD18, a ubiquitin ligase that inhibits LINE-1 activity by two mechanisms. On the one hand, RAD18 monoubiquitinylates PCNA (Proliferating Cell Nuclear Antigen), an essential accessory protein for DNA replication [75]. This modification disrupts the interaction of PCNA and ORF2p, which is crucial for LINE-1 integration. On the other hand, RAD18 interacts with ORF1p, potentially sequestering it into P bodies and stress granules [76], [77].

The DNA replication fork reversal, often triggered by replication stress, inadvertently creates a more permissive environment for LINE-1 activity. BRCA1 plays a key role in numerous DNA repair pathways. This ubiquitin ligase is able to suppress LINE-1 with two distinct mechanisms: it ensures rapid processing of double-strand breaks at replication forks by orchestrating the RPA proteins to create a protective coating of DNA ends. While in the cytoplasm, BRCA1 can inhibit the translation of ORF2 through binding to its mRNA, thus interfering with the formation of functional LINE-1 RNP [78].

Cell cycle factors also play a complex role in regulating LINE-1 activity [79]. While p53, a well-known tumor suppressor, can directly promote LINE-1 repression by binding to its promoter, it also indirectly suppresses LINE-1 activity in germ cells through the piRNA pathway and by enhancing repressive histone modifications on LINE-1 elements. It was also shown that p53 restricts the growth of LINE-1 expressing primary human RPE cells [76]. Additionally, cell cycle regulators like p21 and p27, often activated by p53, can inhibit LINE-1 integration by interacting with the ORF2 protein. Conversely, the MYC proto-oncogene, frequently dysregulated in cancer, can repress LINE-1 transcription [50].

Overall, the cellular defense system against LINE-1 elements is a complex and dynamic network with various layers of control, safeguarding genome integrity [50].

Most somatic LINE-1 control mechanisms are shared between mice and humans. For example, SIRT6 [54], SIRT7 [55], the Microprocessor complex [58], and MOV 10 helicase [80], [81] regulate LINE-1 expression in both species.

1.2.5 Factors affecting LINE-1 activity

The interplay between cellular defense mechanisms and LINE-1 retrotransposons is remarkably intricate and susceptible to disruption by a variety of internal and external factors. Environmental exposures, including therapeutic drugs, radiation (both UV radiation and radiotherapy), occupational hazards (chemical and biological), and other yet-to-be-identified agents, can directly or indirectly modulate LINE-1 activity [82].

These agents might interact directly with LINE-1 components throughout the whole retrotransposition cycle, directly influencing LINE-1 expression and retrotransposition rates. Conversely, indirect effects arise when certain factors interfere with the above-mentioned cellular pathways that modulate LINE-1 activity. Essentially, any external stressor or chemical impacting the body has the potential to affect LINE-1 retrotransposition and cause derepression of it [82].

In these cases, the resulting increase in LINE-1 activity could potentially lead to cancer "driver" mutations. This phenomenon likely contributes significantly to the incidence of sporadic cancers, which account for a large proportion of cancer cases globally. Epidemiological studies suggest that environmental factors play a more significant role in tumorigenesis than genetics [83]. For instance, studies on twins indicate that the environment contributes to sporadic cancer occurrence by 58–82%, compared to a 27–42% genetic contribution in cancers like colorectal, breast, and prostate cancers [50].

Numerous carcinogenic environmental factors increase retrotransposon activity in cultured cells, such as, benzo(a)pyrene, an environmental pollutant from tobacco smoke and fossil fuel combustion, which is a contributor to tumorigenesis in lung and esophageal cancers [84]. Metals like cadmium, chromium VI, and nickel also increase LINE-1 retrotransposition rates [85]. Oxidative stress, which is another common feature in tumors, can increase retrotransposition rates and chromosomal instability, potentially contributing to cancer development [86]. Thus, environmental factors that contribute to cancer development may also activate LINE-1 retrotransposition, increasing the likelihood of oncogenic mutations and tumorigenesis [82].

1.2.6 Assays to Investigate LINE-1 Activity

Quantifying *de novo* LINE-1 retrotransposition events is challenging considering the great abundance of LINE-1 elements in the genome. Understanding the strategies to investigate LINE-1 activity is crucial for interpreting the existing data.

The methylation status of LINE-1 promoters can be investigated by bisulfite sequencing. It is widely accepted as an indicator of global DNA hypomethylation. Global hypomethylation is frequently observed in many types of cancers [87], [88]. Although this method covers only a small portion of active LINE-1 elements, it does not always correlate with increased LINE-1 expression because the methylation status of inactive LINE-1 elements does not impact overall LINE-1 activity. Since most LINE-1 promoters in the genome belong to inactive elements, it is essential to distinguish between active and inactive loci [89].

The LINE-1 promoter-luciferase assay measures LINE-1 promoter activity in cell cultures by transfecting a plasmid containing the LINE-1 5' UTR promoter sequence upstream of a firefly luciferase reporter gene. The assay primarily evaluates the transcriptional activity of LINE-1 promoters [21], [90]. This enables screening of different compounds that are increasing the LINE-1 promoter activity, but it fails to detect the effects of those compounds that are causing global hypomethylation as the LINE-1 promoter-Luciferase gene is introduced on a plasmid. Therefore, substances or factors causing global hypomethylation may remain undetected. Furthermore, this only allows us to infer the presence of LINE1 mRNA and does not necessarily indicate that transposition is occurring.

Another popular approach for mammalian tissue culture experiments is the use of ORFeus-type reporters. These are codon-optimized, highly active engineered LINE-1 elements for mouse and human (*ORFeus-Mm*, and *ORFeus-Hs*, respectively) [91], [92]. In these assays, a reporter gene (most commonly EGFP) is inserted into the 3' UTR of the optimized LINE-1 sequence, but in an antisense direction to the LINE-1 itself. An intronic sequence is inserted in sense direction to interrupt the reporter gene. In order to remove the intron, the LINE-1 RNA needs to undergo splicing. After that, if the reconstituted reporter is reverse transcribed, integrated into the genome, and antisense expression occurs, the reassembled reporter can be visualized. **(Fig. 4)** The retrotransposed reporter exhibits strong and permanent marker (EGFP) expression in the affected cell and in all its daughter cells. This marker (EGFP) positivity can be quantified by FACS [92].

Colony-forming assays are also used to evaluate the effects of different chemicals. In this setup, an antisense antibiotic (e.g., Neomycin) resistance gene is inserted downstream to the ORF1 and ORF2 sequences, and it is also interrupted by a sense intron. The number of colonies after drug treatment represents the number of retrotransposition events, since only those colonies are viable, which have undergone retrotransposition [93].

LINE-1 ORF1 and LINE-1 ORF2 proteins of humans and mice can be used interchangeably. On top of that, the retrotransposition activity of chimeric human-mouse LINE-1 elements is not noticeably decreased in either host [31].

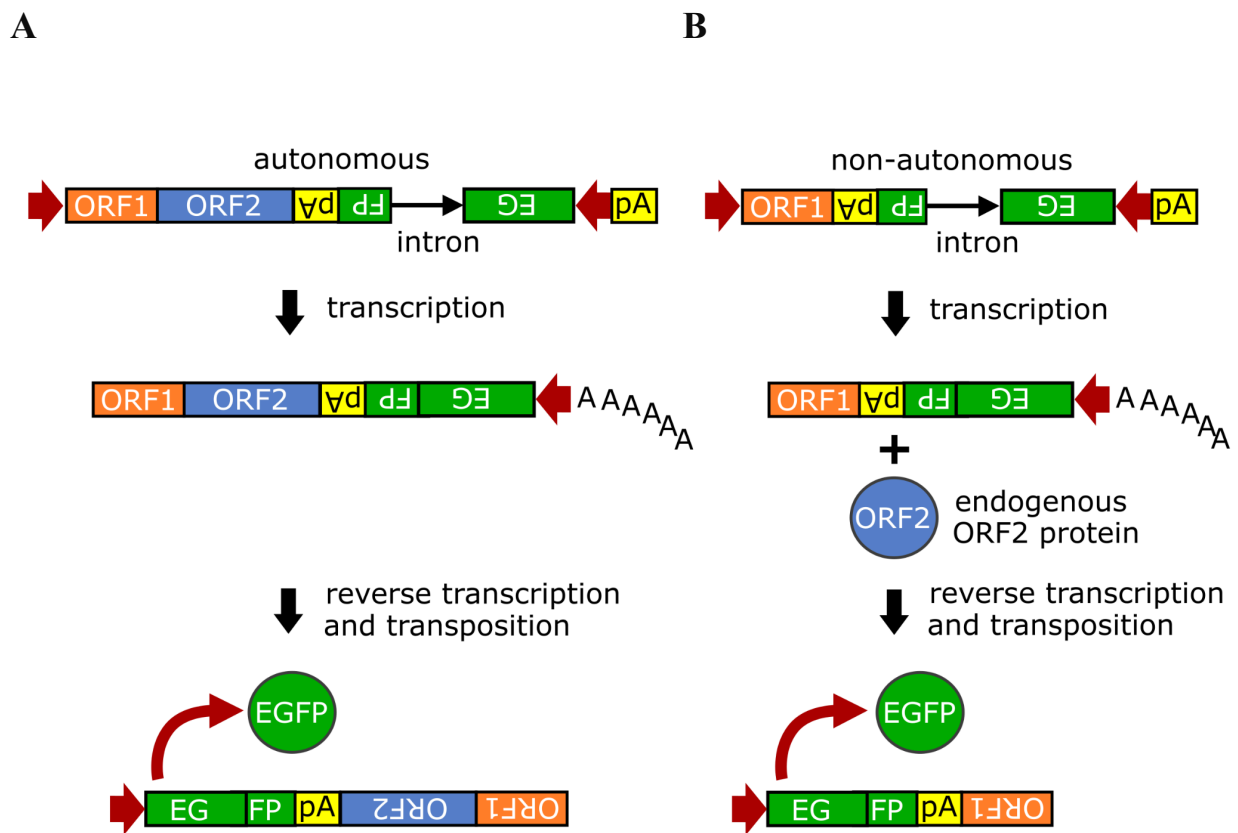


Figure 4. Schematic representation of the architecture of the ORFeus-type LINE-1 reporter cassettes. The autonomous reporter variant (A) contains both ORF1 and ORF2 sequences, while the non-autonomous reporter variant (B) contains only ORF1, while ORF2 has to be provided from an endogenous source to complete retrotransposition.

The lineage-tracing nature of the ORFeus reporters can be a convenient feature during the readout of *in vitro* assays, but poses a great challenge in utilizing them in *in vivo* systems, since LINE-1 retrotransposons are physiologically active in gametes and early embryonic cells, so the germline-transgenic application of the ORFeus reporter would result in early

activation, which would lead to widespread EGFP positivity in the animal even before complete development.

Therefore, detecting somatic LINE-1 activity *in vivo* remains a significant challenge. The only existing germline-transgenic mouse model was published by Okudaira and their colleagues. They microinjected human ORFeus reporter containing sequence into the pronuclei of fertilized eggs from mice. Strains were selected for low spontaneous background LINE-1 retrotransposition. Retrotransposition was detected using a PCR-based assay [94], [95].

They performed DMBA/TPA-induced skin tumorigenesis by topically applying chemical carcinogens. Their findings demonstrating EGFP immunohistochemical positivity reflect only the expression of the ORFeus reporter due to its large EGFP Exon 1 (which can be expressed before reverse transcription) , rather than indicating ORFeus retrotransposition. This IHC positivity is likely a result of global promoter hypomethylation associated with tumorigenesis, rather than successful retrotransposition events. Additionally, the mouse strains, selected for poor germline retrotransposition efficiency, only show somatic expression of the ORFeus reporter following DMBA/TPA induction, further highlighting the context-specific nature of the observed EGFP expression. It is also challenging to distinguish whether the measured retrotransposition events happened somatically or whether they were the results of early embryonic retrotransposition events.

Based on these complexities and limitations, the development of somatically transgenic mice is necessary. Such a model would provide a novel alternative for accurately investigating somatic LINE-1 activity *in vivo*.

2 AIMS OF THE THESIS

Despite considerable progress in cancer research, understanding the exact causes and mechanisms behind sporadic cancer development remains an enormous challenge. This thesis focuses on a lesser-explored area of cancer biology: the involvement of LINE-1 retrotransposons in the initiation of cancer. LINE-1 sequences are the most abundant group of mobile genetic elements in the human genome, and their uncontrolled activity acts as a somatic mutator. This work aims to deepen our understanding of the dynamics of LINE-1 activity *in vivo*, as well as its implications for carcinogenesis.

In the present thesis, our aims were to

- i) investigate the expression of the LINE-1 ORF1 protein in human cervical specimens during the different stages of cervical carcinogenesis and evaluate its potential as a biomarker for the diagnosis of cervical intraepithelial neoplasia, by setting up a new scoring system to evaluate LINE-1 ORF1p immunostainings of human uterine cervical epithelia, incorporating both the staining intensity and the extent of epithelial involvement,
- ii) to develop a novel technological platform to overcome the problems of the existing assays investigating somatic retrotransposition rate by establishing the first somatically transgenic mouse model for monitoring LINE-1 activity,
- iii) furthermore, harnessing this technological platform to investigate unconventional genotoxic effects of chemicals on somatically transgenic hepatocytes of the LINE-1 reporter-bearing mouse, and to validate the method as a chemical risk assessment technology, enhancing the currently available toxicological toolkit.

3 MATERIALS AND METHODS

3.1 Preparation of the Human Tissue Specimens

Archived pathological samples were used to create tissue microarrays (TMAs). 112 human formalin-fixed paraffin-embedded (FFPE) specimens were included: CIN I (n = 20), CIN II (n = 46), CIN III (n = 14), and cervical cancer (n = 32). Controls consisted of 31 non-dysplastic cervical tissues from patients who had hysterectomies for uterine leiomyomas or prolapse, categorized into normal (n = 24) and atrophic epithelium (n = 7). Cores were selected and verified using H&E staining. Tumor specimens were assessed for tumor grade, size, lymph node involvement, and metastasis status (pTNM). Patient data were anonymized. Ages ranged from 22 to 83 years (mean age 52), and samples were collected from 2018 to 2022 at the Clinical Center of Albert Szent-Györgyi Medical School, University of Szeged, Hungary. The Institutional Committee of Science and Research Ethics, Budapest, Hungary, approved the study. (Ref: BM/3049/2023).

3.2 Animal Care and Maintenance

Mice were bred and housed in the Central Animal House at the Biological Research Centre in Szeged, Hungary. Their pathogen-free status was confirmed every three months in accordance with FELASA guidelines [96]. The mice were maintained under a 12-hour light/dark cycle at a temperature of 22 °C, with *ad libitum* access to water and standard rodent chow. All experimental procedures were conducted following protocols approved by the Institutional Animal Care and Use Committee of the Biological Research Centre. Ethical authorization number: XVI./801/2018

The Fah mutant strain used in this study, C57BL/6N-Fah^{tm1(NCOM)Mfgc/Biat}, is deposited in the European Mouse Mutant Archive (EMMA) under EM:10787. Fah^{-/-} mice received Orfadin® (Nitisinone, NTBC) at a concentration of 8 mg/l in their drinking water. NTBC administration was discontinued immediately after hydrodynamic injection. Wild-type C57BL/6NTac mice were purchased from Taconic Biosciences.

3.3 Immunohistochemistry

Immunostaining was conducted on the Bond Max Autostainer (Leica Biosystems) using the Bond Polymer Detection System (Vision BioSystems). The antibodies used for the human uterine cervix FFPE specimens included:

- anti-LINE-1 ORF1p (clone 4H1, MABC1152, Millipore),
- anti-p16 (clone MX007, MAD-000690QD-12, Master Diagnostica),
- anti-Ki67 (clone SP6, #10080, Histopathology Ltd.).

Mice were euthanized at 5 weeks, 3 months, or 5 months post-injection for the somatic gene delivery study. Mice were sacrificed at 3 months post-injection for the LINE-1 *in vivo* retrotransposition study. Liver tissues were collected, overnight fixed in 4% formalin, embedded in paraffin, and sectioned into 5- μ m slices. The EnVision FLEX Mini Kit (DAKO) was used to perform the IHC protocol with antigen retrieval conducted in a PT Link machine (DAKO). Primary antibodies included:

- rabbit polyclonal anti-FAH (ThermoFisher Scientific, PA5-42049, 1:100)
- rabbit polyclonal anti-mCherry (GeneTex, GTX128508, 1:400)
- rabbit monoclonal anti-LINE-1 ORF1p [EPR21844-108] (Abcam, ab216324, 1:500)

Polyclonal goat anti-rabbit-HRP secondary antibody (DAKO, P0448) was applied for 30 minutes. Visualization employed the EnVision FLEX DAB+ Chromogen System (DAKO, GV825). Slides were counterstained with hematoxylin for 5 minutes, mounted, and scanned using a Panoramic Digital Slide Scanner (3D Histech).

3.4 Creation of plasmid constructions

The empty *pbiLiv-miR* vector was cloned in pUC57 plasmid backbone (GeneScript). This includes the bidirectional promoter of the alpha (HADHA) and beta (HADHB) subunits of the human hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex.

The HADHA side of the bidirectional promoter drives the expression of the mCherry fluorescent marker gene, which is disrupted by the first intron of the human eukaryotic translation elongation factor 1 alpha 1 (*EEF1A1*). The mCherry gene is connected to the

mouse fumaryl-aceto-acetate dehydrogenase (*Fah*) CDS by a T2A self-cleaving peptide to provide bicistronic expression, then a bGH polyadenylation signal terminates the expression unit.

The HADHB side of the bidirectional promoter is followed by a multiple cloning site (MCS) and a bovine growth hormone (bGH) polyadenylation signal. The entire structure is flanked by inverted terminal repeats of SB transposon.

The *pbiLiv-miR-Luc* construct was created by amplifying the firefly Luciferase gene from the pGL3-Basic plasmid (Promega) by PCR and inserting it into the NheI/PacI sites of the multiple cloning site (MCS) in *pbiLiv-miR*.

To generate *pbiLiv-miR-mORFeus* variants, the following expression units were inserted to the MCS in the B side of the HADHA/B promoter: TF monomers, ORF1 coding sequence, 3'UTR, hsvTK polyA, EGFP (in reverse orientation), interrupted by the hGamma Globin intron, CMV promoter, and SV40 polyA.

To gain higher cargo capacity, we replaced the SB ITRs with the ITRs of the PiggyBac transposon system. During gene delivery, we co-injected them with the hyperactive piggyBac transposase (hyPBBase) [97].

3.5 Hydrodynamic tail vein injection

The NucleoBond Xtra Maxi Plus EF Kit (Macherey-Nagel) was used to prepare plasmids for hydrodynamic tail vein injection in accordance with the manufacturer's instructions. The plasmid DNA was diluted in sterile Ringer's solution, which comprised 0.9% NaCl, 0.03% KCl, and 0.016% CaCl₂. 4 µg of the transposase helper plasmid (SB100 or hyPBBase) and 50 µg of the transposon-containing plasmid were administered for each animal.

The procedure was performed on mice aged 6–8 weeks old. Before the injection, the tails of the animals were soaked in warm water to dilate the tail veins. The animals were weighted, and a volume that equaled 10% of the body weight was administered through the lateral tail vein (*vena caudalis lateralis*) in 5-8 seconds to reach hydrodynamic effect [98], [99].

3.6 Administration of screened compounds for LINE-1 inductor properties

After hydrodynamic injections and NTBC withdrawal, we started the administration of the screened compounds for 3 months in the scheme detailed in **Table 1**. The same amount of vehicle (DMSO/corn oil) was injected into the animals of the vehicle control groups.

Compound	Vehicle	Dose	Route of administration	Frequency	References
Decitabine (MedChemExpress)	DMSO	1 mg/kg	intraperitoneal injection	2/week	[100]
Efavirenz (MedChemExpress)	DMSO	10 mg/kg	intraperitoneal injection	2/week	[101]
MeIQX (Abcam)	DMSO	5mg/kg	intraperitoneal injection	2/week	[102], [103]
FICZ (MedChemExpress)	corn oil	5mg/kg	intraperitoneal injection	2/week	[104]

Table 1. Details of the administered substances

3.7 Stereomicroscope Imaging

Whole mouse liver images were captured using an Olympus SZX12 fluorescence stereozoom microscope. The microscope was equipped with a 100-W mercury lamp and specialized filter sets to selectively excite and emit GFP and mCherry fluorescence.

3.8 *In vivo* bioluminescence imaging

In vivo bioluminescence imaging was conducted with an IVIS Lumina III instrument (PerkinElmer). Mice were injected with 150 mg/kg luciferase substrate, intraperitoneally (XenoLight D-Luciferin-K⁺ salt, PerkinElmer), then anesthetized using isoflurane (Isoflurin, Vetpharma) using the RAS-4 Rodent Anesthesia System (PerkinElmer). Imaging was started 10 minutes post-injection, capturing emitted photons with exposure times ranging from 1 to 10 seconds. The Living Image software (PerkinElmer) was used to quantify the average radiance within a defined circular region of interest, measured in photons per second per cm² per steradian (p/s/cm²/sr).

3.9 Liver perfusion and hepatocyte isolation

Livers were harvested for hepatocyte isolation under sodium pentobarbital (Nembutal, Sigma Aldrich) anesthesia. Hepatocytes were isolated using a three-step collagenase perfusion method. Initially, mice were perfused through the *superior vena cava* with calcium-free Earle's balanced salt solution (EBSS) containing EGTA. This was followed by a washout with calcium-containing EBSS, and then perfusion with EBSS containing 0.5 g/l Collagenase Type IV (Sigma Aldrich). Digested livers were extracted and placed in ice-cold washing buffer (0.01 mM HEPES, 140 mM NaCl, 7 mM KCl, pH 7.2), and all further steps were performed on ice.

The liver capsule was opened up to release cells into the washing buffer through gentle manual agitation. The cell suspension was filtered through a 100 µm filter to remove undigested tissue and debris. Cells were centrifuged at 1000 rpm for 4 minutes at 4°C. The resulting pellet was resuspended in washing buffer, mixed with an equal volume of Percoll solution (Sigma Aldrich), and centrifuged again at 1000 rpm for 4 minutes at 4°C. The hepatocyte-containing pellet was washed with washing buffer and centrifuged at 1000 rpm for 4 minutes at 4°C. Cell counts were obtained using a Bürker chamber, while cell viability was evaluated through the trypan blue exclusion test [105].

3.10 FACS-Based Measurement of EGFP-Positive Hepatocytes

Hepatocytes (2x10⁶/ml) isolated from mouse livers were suspended in PBS and filtered through a 100 µm mesh to prevent clumping. EGFP fluorescence was analyzed using a BD

FACSAria™ Fusion Flow Cytometer (Becton Dickinson) with standard flow cytometry protocols. Data analysis was performed using BD FACSDiva™ Software.

3.11 qPCR Strategy for Detecting Intronless ORFeus Copies from Retrotransposition Events

Genomic DNA was extracted from liver tissue with Proteinase K and incubated at 55°C overnight. DNA was extracted using phenol:chloroform, then precipitated with isopropanol, washed in 70% ethanol, air-dried, and resuspended in TE buffer.

To quantify retrotransposition events of the synthetic LINE-1 element, we performed qPCR analysis targeting the exon-exon junctions of the spliced, intronless EGFP.

We used PerfeCTa SYBR Green SuperMix (Quantabio). The cycling protocol included an initial denaturation at 95°C for 7 minutes, followed by 4 cycles of 10 seconds at 95°C and 15 seconds at 66°C, decreasing by 1°C per cycle (without data acquisition). This was followed by 40 cycles of 5 seconds at 95°C and 10 seconds at 62°C. The primers used were:

- Olfr16-Fwd: GAGTTCGTCTTCCTGGGATTC
- Olfr16-Rev:TAATGATGTTGCCAGCCAGA [106]
- GFP-Fwd: AAGCAGAAGAACGGCATCAAGGT
- GFP-EJ-Rev: TGGTAGTGGTCGGCCAGCTGC

All qPCR reactions were conducted in triplicates on a Rotor-Gene Q instrument (Qiagen), using 87 ng of genomic DNA (gDNA). Data analysis was carried out with Rotor-Gene Q software (Qiagen), and relative expression changes were calculated using the Δ CT method [107]. Results were normalized against the internal control gene Olfr16.

3.12 Data visualization and statistics

The data visualization was performed using R and GraphPad Prism software (version 8.4.3 for Windows, GraphPad Software). IHC marker expression of cervical sample groups were assessed using Fisher's exact test and Pearson's chi-squared test, as appropriate. Planned comparisons between the control group and specific drug treatment groups were conducted using independent samples t-tests. To account for multiple comparisons, the Bonferroni correction was applied, adjusting the significance level accordingly. The associations were considered statistically significant for $P < 0.05$ in all analyzed cases.

4 RESULTS

4.1 Investigation of LINE-1 ORF1 Protein Expression in FFPE Human Specimens

The emerging role of ORF1p as a biomarker in various cancers has created increasing attention, with its expression identified as a hallmark in a wide range of malignancies [108], including breast [109], ovarian [110], bladder [86], esophageal [111], and pancreatic [112] cancers. Some studies have shown that L1 upregulation occurs at early stages of tumor development, such as in Barrett's esophagus [111]. To further investigate this tendency, we have screened hundreds of various types of tumor samples from different histological grades and clinical stages using the high-throughput TMA technology. We observed that ORF1p expression intensifies with advancing stages of tumorigenesis, supporting the hypothesis that L1 retrotransposition contributes to tumor progression.

We were the first to examine ORF1p expression in the human uterine cervix on a comprehensive set of FFPE samples using immunohistochemistry. We examined 112 (FFPE) samples from various stages of cervical intraepithelial neoplasia (CIN) and cervical cancer, as well as 31 non-dysplastic cervical tissue samples. ORF1p expression increased with higher CIN grades and was minimal or absent in normal cervical tissue. Additionally, we performed immunostaining for conventional diagnostic markers: Ki-67 IHC to assess cell proliferation, and p16, which is a marker for oncogenic HPV infection, with overexpression serving as a surrogate marker for high-grade CIN, helping to differentiate between low-grade lesions and more severe dysplasia. Ki67 and p16 expression patterns deviated from normal at the CIN II stage, while ORF1p showed abnormal expression starting from the CIN I stage, with immunopositivity rising with CIN progression.

The extent of ORF1p positivity (involving the lower one-third, two-third, or whole epithelial thickness) correlated with the CIN grade (**Fig. 5**) To quantify this, we developed a scoring system based on a modified version of the Immunoreactivity Scale (IRS) by Remmele and Stegner [113]. The final ORF1p score was obtained by multiplying the staining intensity by the extent of epithelial involvement. The scoring system was applied to assess ORF1p immunopositivity (**Table 2.**). p16 expression was quantified by the percentage of positive cells, while Ki67 expression was assessed by counting positive nuclei in 200 epithelial cells

and assigning grades based on the percentage [114], [115]. Negative controls were performed with subclass-matched IgG1K. Two pathologists independently scored the sections, blinded to clinical data, and reached consensus on any discrepancies.

Stained Epithelial Thickness	Proportion of Positivity (PP) Score	Intensity of Staining (SI)	SI Score	IRS Classification	IRS Scores (PP × SI)
No staining	0	No reaction	0	Negative	0
Basal one-third	1	Mild reaction	1	Positive, mild immunoreaction	1-2
Basal two-thirds	2	Moderate reaction	2	Positive, moderate immunoreaction	3-4
Transepithelial	3	Strong reaction	3	Positive, strong immunoreaction	6-9

Table 2. Modified version of the immunoreactive scale (IRS) scoring system for the evaluation of ORF1p immunohistochemistry in cervical specimens

All CIN I, II, and III cases showed ORF1p positivity, with higher scores linked to higher CIN grades. (See details in **Fig. 6**) A sharp contrast emerges between dysplastic and non-dysplastic lesions, with the former predominantly exhibiting strong or moderate ORF1p expression, while the latter (Atrophy, Normal cervical epithelium) displays considerably fewer positive samples, with weak expression levels. Notably, most non-dysplastic epithelia lack ORF1p expression. Strong ORF1p immunopositivity was also observed in most invasive cancer cases, though two samples lacked ORF1p expression. Statistical analysis confirmed significantly higher ORF1p immunoreaction scores in dysplastic and neoplastic lesions compared to normal tissues (Fisher's exact test, $P < 0.0001$) (**Fig. 6**) No association was found between ORF1p immunopositivity and patient age, tumor grade, or pTNM stage.

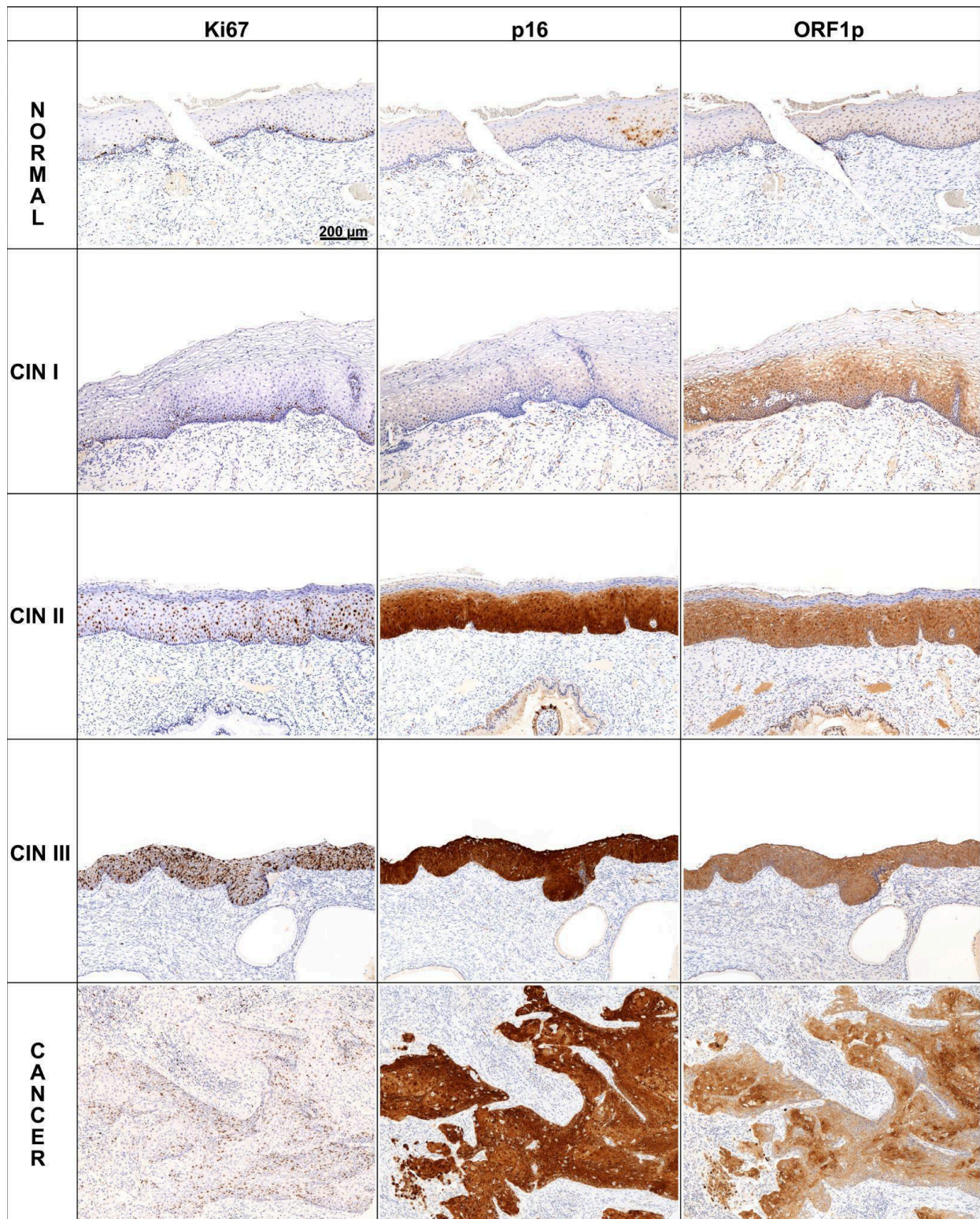


Figure 5. Photomicrographs from representative immunostainings to investigate ORF1p, p16, and Ki67 expression in normal cervix and various grades of cervical epithelial lesions as they advance towards invasive cancer. Sections of FFPE specimens from normal cervical epithelium, cervical epithelial neoplasia (CIN I, CIN II, and CIN III), and cancer are shown. The scale bar demonstrates 200 μ m.

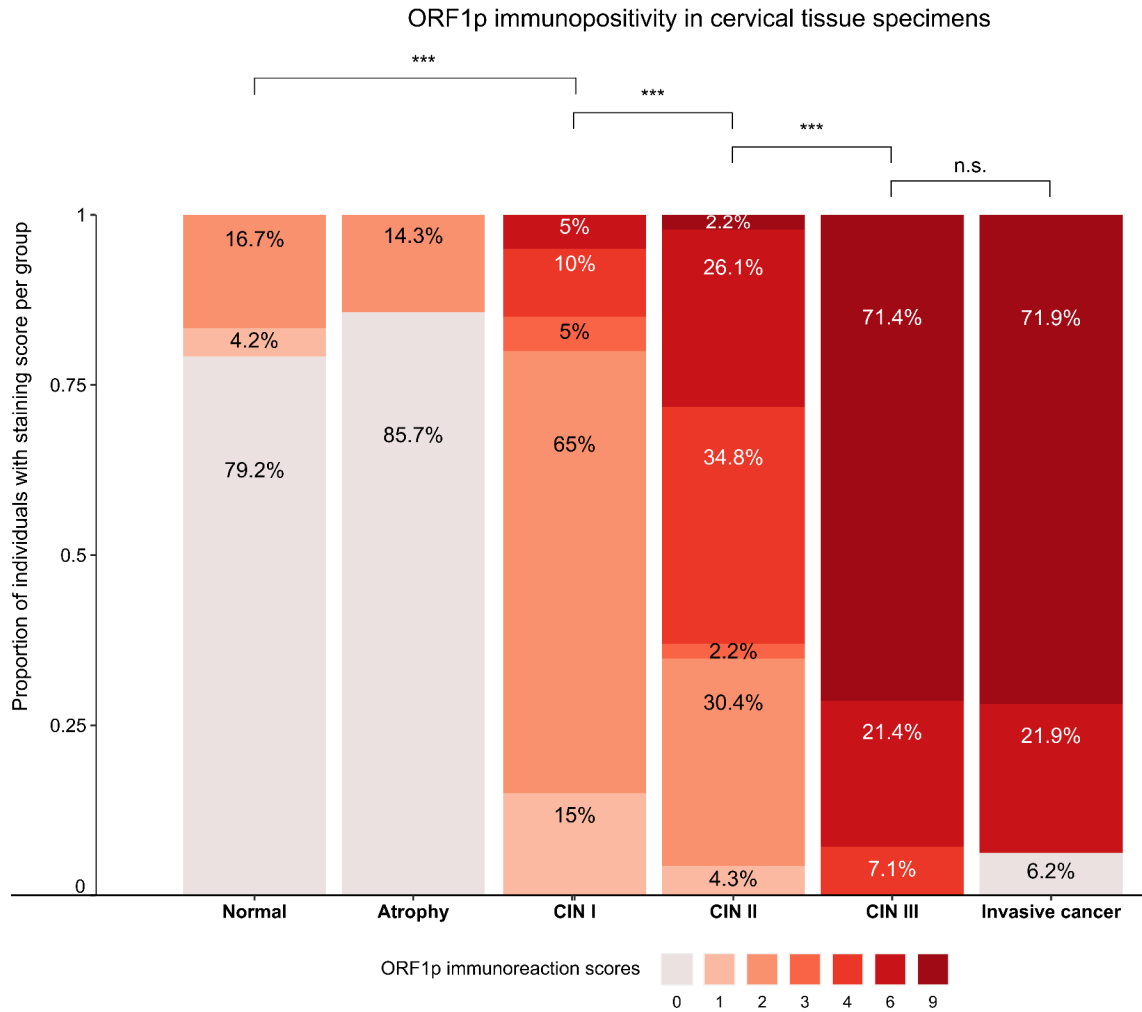


Figure 6. ORF1p IRS scores in Cervical Specimens. Statistical significance assessed via χ^2 tests is indicated by asterisks on the bars. *** indicates $p \leq 0.001$, n.s. = not significant

These findings underscore the significance of LINE-1 retrotransposon and suggest that abnormal somatic L1 activity could be an early driver, making it worthwhile to investigate it, using an *in vivo* model. Therefore, we established an *in vivo* somatic reporter system in mice. This required the creation of a somatic transgenesis system that enables stable transgene expression. The developed somatic L1 reporter system can be applied to screen for chemicals capable of inducing somatic L1 activity.

4.2 Measurement of somatic LINE1 retrotransposition activity in mice

The use of transgenic mouse models for measuring LINE1 retrotransposition activity is challenging due to the lineage-tracing nature of the ORFeus reporter system and the inherent germline and early embryonic L1 activity. The germline-transgenic use of the ORFeus

reporter would cause early reporter activation, resulting in widespread EGFP expression throughout the animal even before it fully develops. As a result, the seemingly straightforward application of germline transgenic mouse models to create a measuring system is not feasible.

To overcome this limitation, we first required a reliable somatic transgenesis platform that could provide stable transgene expression while avoiding germline transmission. We successfully established such a system using the mouse model of type I tyrosinemia. Employing this approach, we generated somatic L1 reporter transgenic mice, which enabled us to test our drug candidates effectively in a controlled somatic context.

4.2.1 Creation of the mouse model of human type I tyrosinemia

We utilized the preclinical mouse model of human type I tyrosinemia (TT1). TT1 is a fatal inherited disease caused by germline mutations in the fumarylacetoacetate hydrolase (*Fah*) gene, leading to progressive liver dysfunction due to the accumulation of toxic metabolites like succinylacetone. In the absence of the functional *Fah* enzyme, untreated mice develop liver failure. **(Fig. 7)** We created a TT1 mouse model by knocking out the *Fah* gene, by deleting exons 2 to 5 (C57BL/6N-*Fah*^{tm1(NCOM)Mfgc/Biat}). For further details, see our publication by *Imre et al.*, 2023 [116].

To manage the lethal condition, we used Nitisinone (NTBC), a drug that inhibits tyrosine catabolism upstream of FAH, preventing the buildup of harmful metabolites and converting TT1 to type III tyrosinemia (TT3), which is a milder form of the disease. Before delivering the therapeutic *Fah* minigene, NTBC was administered to prevent acute toxicity. **(Fig. 7)** Immediately after gene delivery, NTBC was withdrawn, enabling the selection of hepatocytes carrying the therapeutic *Fah* gene, thus creating a transgenic liver repopulated almost entirely by genetically modified cells.

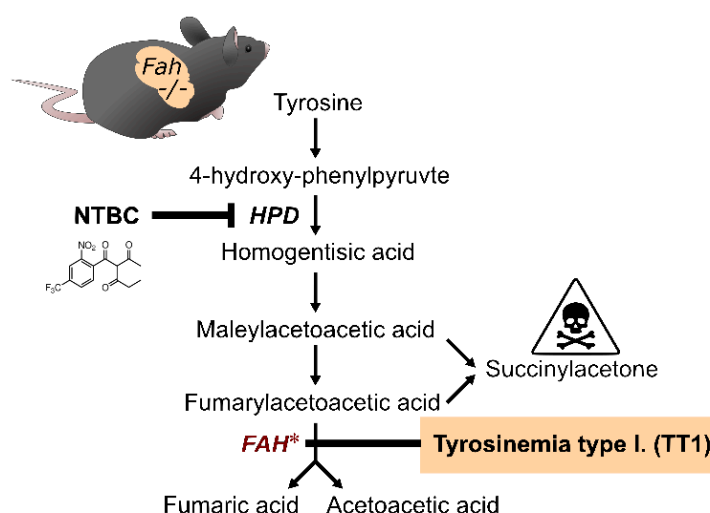


Figure 7. The tyrosine catabolic pathway. It is affected by the mutation of the enzyme fumarylacetoacetate hydrolase (*Fah*). The lethal phenotype can be rescued and transformed into a milder form of the disease by blocking an upstream enzyme in the pathway called hydroxyphenylpyruvate dioxygenase (*Hpd*) using an orally active, competitive, and reversible inhibitor, NTBC (2-(2-nitro-4-trifluoromethyl benzoyl)-1,3-cyclohexanedione) [117].

4.2.2 Therapeutic gene delivery-coupled somatic transgenesis in the mouse liver

To achieve robust gene delivery, we utilized hydrodynamic injection of plasmid DNA constructs to target hepatocytes by enhancing membrane permeability [118], [119]. We utilized a hyperactive transposase helper, SB100 [120], and an advanced transposon variant with T2 Inverted Terminal Repeat (ITR) structure [121], aiming for stable transgene expression through chromosomal integration by the “cut-and-paste” reaction catalyzed by the SB100 transposase [122]. However, the inhibitory effects of exogenous DNA sensors affect transgene expression despite chromosomal gene transfer [98]. These systems present a significant challenge to achieving permanent and stable gene expression. The primary obstacle is the activity of cytoplasmic DNA sensors, which detect exogenous DNA and trigger silencing and elimination processes that can hinder gene expression.

To counteract these negative processes, we employed a strategy involving positive selection for our transgene. Specifically, we utilized a therapeutic minigene for *Fah*, designed to apply selective pressure in favor of our transgene of interest. To accomplish this, we placed the *Fah* minigene on a transposon along with the transgene of interest and linked their expression

using a bidirectional promoter. This dual approach ensures that the selective advantage conferred by the *Fah* minigene directly supports the stable expression of the transgene, thereby overcoming the silencing effects imposed by the cellular defense mechanisms.

We aimed to achieve permanent transgene expression by replacing *Fah*^{-/-} hepatocytes with *Fah*^{+/+} cells by multinodular repopulation, benefiting from the growth advantage of the corrected, *Fah* expressing cells [123]. We harnessed this feature to create our transgenic gene-delivery system by including a *Fah* coding sequence (CDS) to our SB transposon construct and tested it in our *Fah* mutant mouse strain.

To enable targeted gene expression modifications, we created an SB transposon-based cloning system (**Fig. 8A**) that co-expresses two linked transcripts, ensuring that transgene expression is linked to the *Fah* selection marker. Transcript A contains the genes for the selection (*Fah*) and visualization marker (mCherry), connected by a “self-cleaving” T2A peptide for bicistronic expression [124]. Transcript B encodes the protein to be studied (in this case, firefly luciferase [*Luc*]) without the necessity for tagging. We utilized the human HADHA/B promoter, which facilitates a balanced bidirectional expression. This small (390 bp) promoter regulates the expression of the human hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex alpha (HADHA) and beta (HADHB) subunits [125].

The transposon and the SB100 transposase helper plasmid were mixed and administered via hydrodynamic tail vein injection to target the liver of *Fah*^{-/-} mice. After NTBC withdrawal, immunohistochemical investigations showed that almost all hepatocytes become *Fah* positive, due to intensive multi-nodular repopulation over 3 months post-injection (**Fig. 8B**)

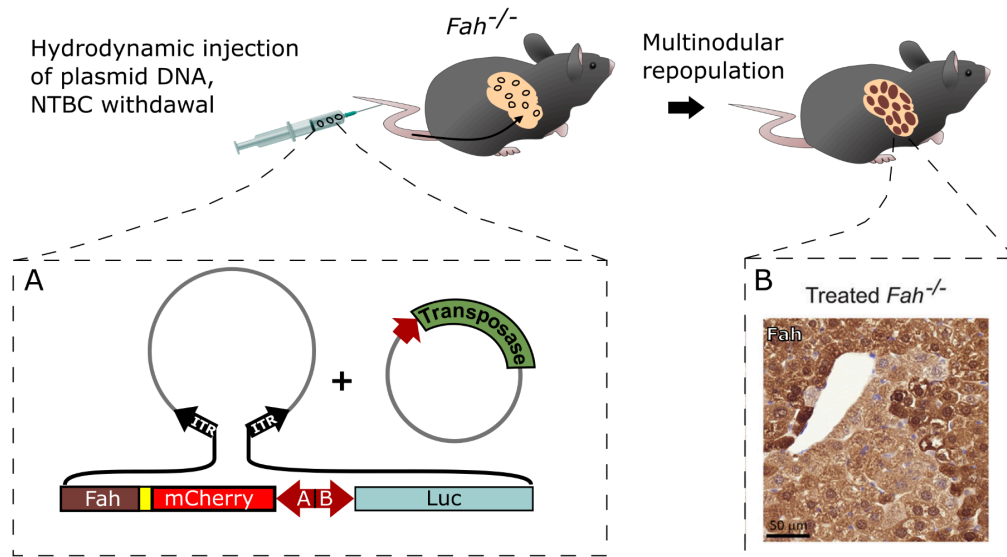


Figure 8. Hydrodynamic gene delivery of the (A) *pbiLiv-miR-Luc* containing *Fah*, *mCherry*, and *Luc* genes flanked by SB ITRs along with SB100 hyperactive transposase helper plasmid. (B) 3 months after gene delivery, *Fah* IHC showed multinodular repopulation of the treated cells. Scale Bar: 50μm.

We performed quantitative reverse transcription PCR (RT-qPCR) to quantify the expression levels of transcripts A and B driven by the HADHA and HADHB regions of the promoter. The RT-qPCR data revealed robust yet physiological gene expression levels, with transcripts A and B exhibiting nearly identical expression levels. (*Discussed further in the PhD thesis of Anna Georgina Kopasz.*)

To investigate the impact of multi-nodular repopulation on long-term transgene expression in *Fah*-deficient livers, we injected the constructs into both wild-type (WT) and *Fah* KO mice, along with the SB100 transposase helper. Transgene expression was tracked over time using the IVIS Lumina III imaging system (PerkinElmer) to detect bioluminescence *in vivo* (**Fig 9A**). Initially, *Luc* expression was slightly higher in WT mice during the first week post-injection. However, from day 28 onwards, this pattern changed. By 3 months, the average bioluminescence intensity in WT mice dropped to an almost undetectable level (387 times drop), while in *Fah* KO mice it reached a 61 times increment compared to the initial level (**Fig. 9B, 9C**).

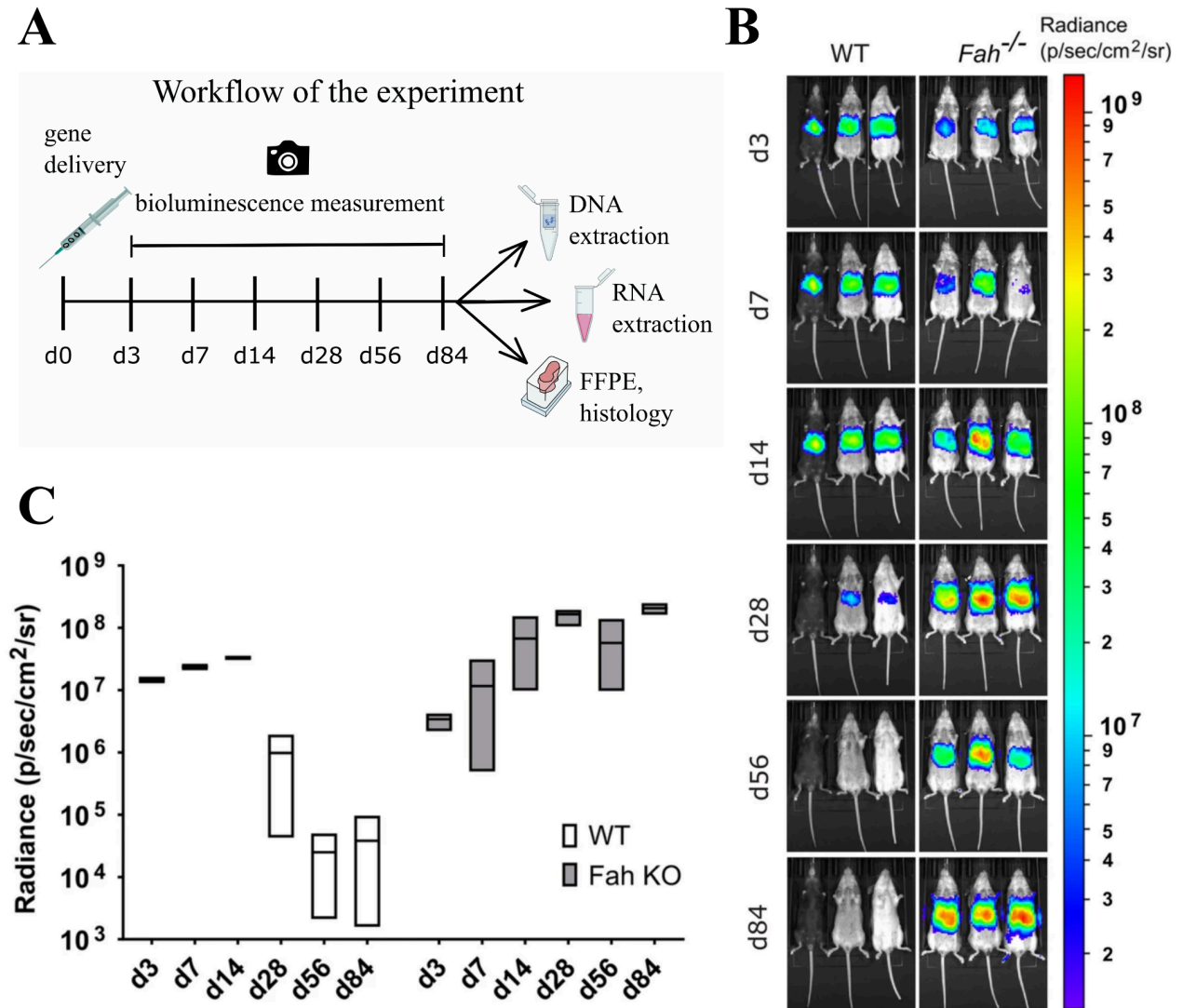


Figure 9. (A) Experimental setup of the somatic gene delivery experiments, (B) Bioluminescence imaging of *Fah*^{-/-} and WT mice was conducted after gene delivery using the IVIS Lumina III imaging system. Signals were captured at 3, 7, 14, 28, 56, and 84 days post-injection. (C) Bioluminescence changes were tracked over 3 months, with average radiance (photons/second/cm²/steradian [p/s/cm²/sr]) plotted on box plots showing the range and mean (n = 6)

Using the insights from these experiments, we could update our genomic technology “toolbox” to be capable of assessing somatic LINE-1 retrotransposition activity in transgenic liver cells of mice. We modified the vector system described above by inserting the ORFeus reporter element to the B side of the HADHA/B bidirectional promoter. We named it *pbiLiv-miR-mORFeus*, which was utilized to generate transgenic livers that permanently express LINE-1 reporter elements (**Fig. 10A**). To facilitate genomic integration of the

expression cassette, we switched to PiggyBac transposon ITRs and hyperactive transposase helper (hyPBase) to gain high cargo capacity for the transposition of this large-size (~10 kilobase long) expression unit.

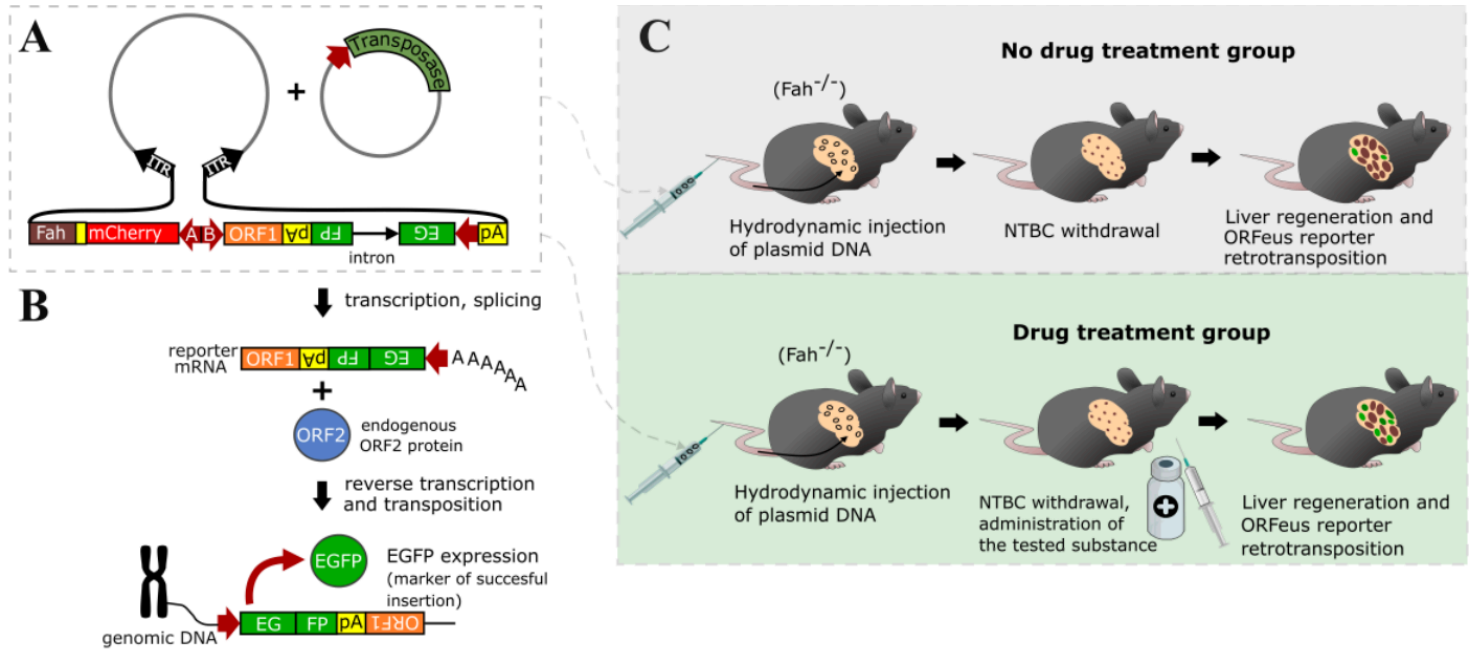


Figure 10. Workflow of measuring LINE-1 retrotransposition (A) *pbiLiv-miR-mORFeus* plasmid construct containing the *Fah* selection marker linked to the fluorescent mCherry marker on the A side of the HADHA/B bidirectional promoter and the non-autonomous LINE-1 ORFeus reporter on the B side. (B) Operation of the ORFeus reporter variant (C) Gene delivery and liver regeneration in the “No drug treatment group” and workflow of the experiment in the “Drug treatment group”

The initial validation of the system involved hydrodynamic co-administration of the *pbiLiv-miR-mORFeus* plasmid with the hyPBase helper plasmid to *Fah*^{-/-} mice, then we discontinued NTBC administration (**Fig. 10C**). In several hepatocytes, successful chromosomal integration was catalyzed by the transposase. Due to the selective growth advantage of these *Fah*-expressing cells, over 3 months, almost the entire population of hepatocytes expressed the ORFeus reporter (**Fig. 11**).

The presence of the *Fah* therapeutic gene and the ORF1p in the treated *Fah*^{-/-} livers 3 months post-gene delivery were confirmed by IHC (**Fig. 11**).

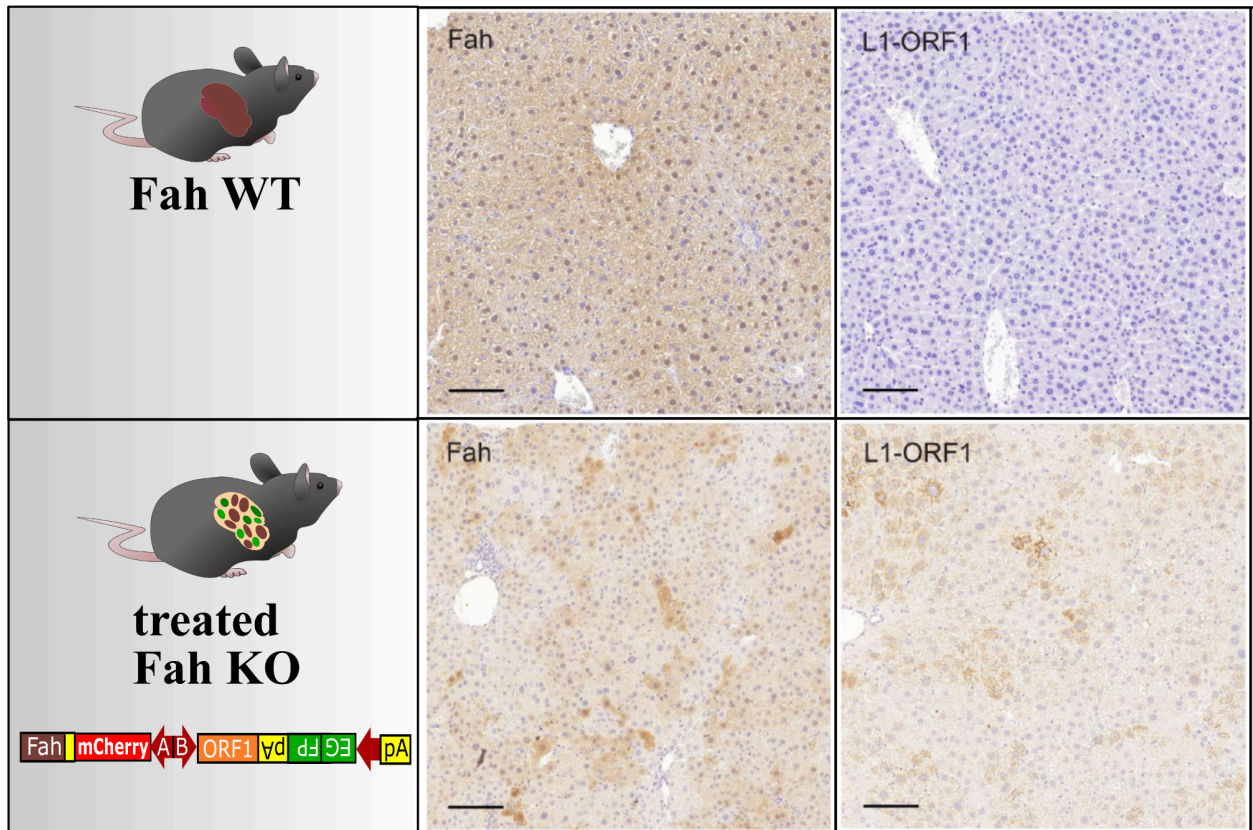


Figure 11. Evaluation of LINE-1 reporter expression. Liver sections from age-matched WT mice show equally distributed *Fah* immunopositive cells but no ORF1p immunoreaction. Liver sections from *Fah* KO ORFeus reporter-bearing mouse demonstrate effective hepatic expression of the ORFeus reporter 3 months post-injection. Negative selection against LINE-1-ORF1 is indicated by fewer clusters of highly immunoreactive (dark brown) hepatocytes and larger, less positive areas (light brown). Scale bars represent 100 μ m.

4.2.3 Assessment of Chemical-Induced LINE-1 Retrotransposition Activity

To validate our somatic LINE-1 activity measurement system, we measured the impact of exogenous factors like therapeutic drugs or foodborne chemicals on LINE-1 activity by administering various substances to our reporter-bearing transgenic mice.

We have chosen versatile chemicals to evaluate their LINE-1 inducing risk since there is evidence in the literature about their LINE-1 modulating properties. We administered these compounds for *pbiLiv-miR-mORFeus* injected *Fah*^{-/-} animals. Our candidates included: Decitabine (DAC); (5-aza-2'-deoxycytidine), a known DNA methylation inhibitor [126]. MeIQx (2-Amino-3,8-Dimethylimidazole[4,5.f]Quinoxaline) is heterocyclic amine, a food-borne carcinogen that forms during cooking meat at high temperatures [95], [103]. FICZ (6- Formylindolo[3,2-b]carbazole), a naturally occurring tryptophan-derived compound

produced by the microbiome [127]. Efavirenz (EFV) is a non-nucleoside reverse transcriptase inhibitor (NNRTI), antiretroviral agent for the treatment of HIV-1 infection [128].

A cohort of *Fah*^{-/-} animals (n = 50) were hydrodynamically injected with *pbiLiv-miR-mORFeus* and hyPBase plasmids, followed by NTBC withdrawal. These mice were divided into two groups: the first group received drug treatment (**detailed in Materials and Methods**), while the remaining animals served as untreated controls (**Fig. 10C, 10D**).

After 3 months of drug administration, the mice were euthanized. Upon macrovisualization, we calculated the number of EGFP-positive hepatocyte colonies in drug-treated mice compared to the non-treated controls. Fluorescence stereomicroscopy images highlighted a marked increase in the density of EGFP-positive cell colonies, in the livers of the FICZ-treated animals (n = 10), suggesting that FICZ induces LINE-1 retrotransposition activity in somatic cells (**Fig. 12H**). Notably, forced expression of the ORFeus reporter induced retrotransposition events in the untreated control group as well, evidenced by the presence of a small number of EGFP-positive hepatocytes (**Fig. 12D**).

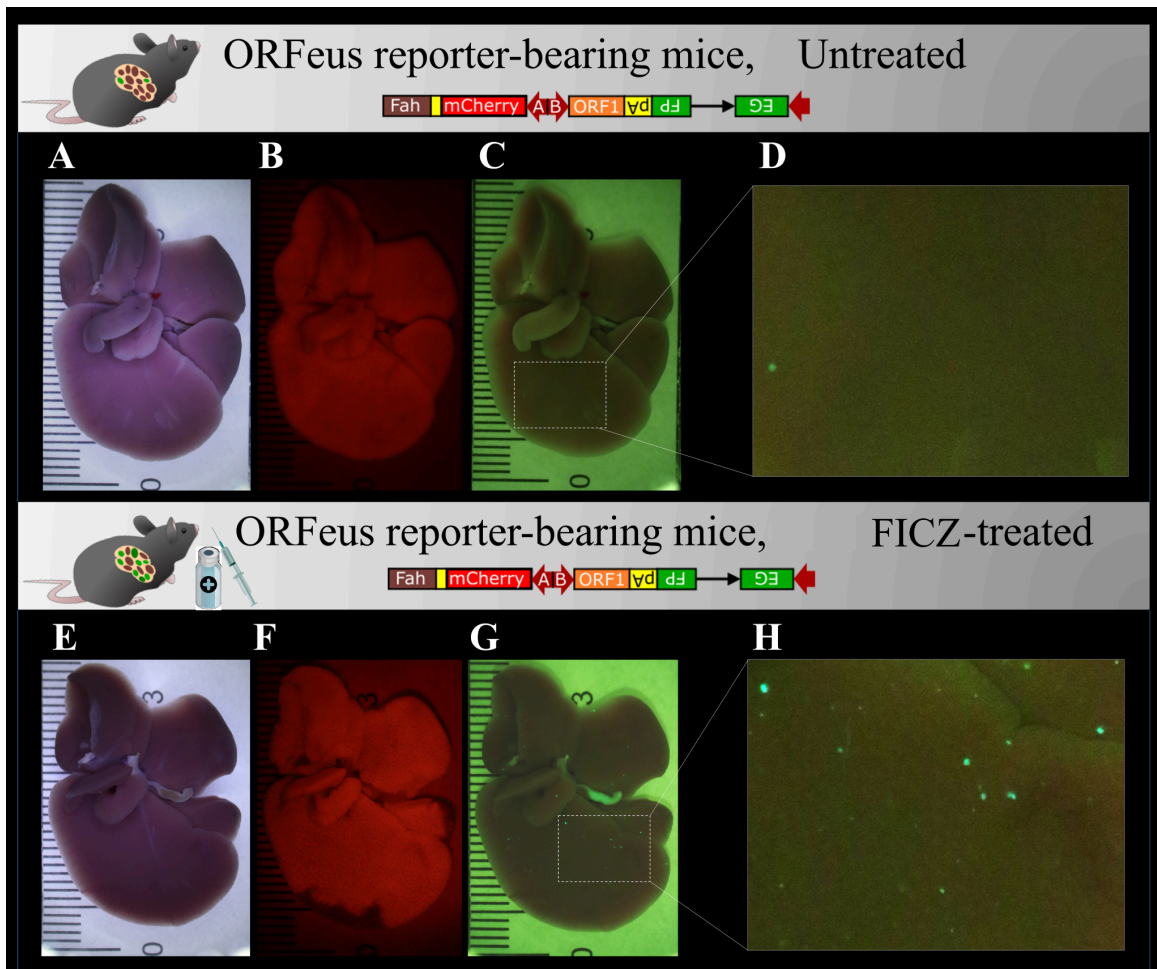


Figure 12. Macrovisualization of untreated and FICZ-treated ORFeus reporter-bearing mouse livers, 3 months after gene delivery, NTBC withdrawal, and FICZ administration. Stereomicroscopy images of untreated mouse livers (A) brightfield (B) mCherry fluorescence (C) GFP fluorescence (D) magnified area of GFP fluorescence. (E) brightfield (F) mCherry fluorescence (G) GFP fluorescence (H) magnified area of GFP fluorescence observed in the FICZ-treated mouse liver.

	Ctrl	Efavirenz	Decitabine	MelQX	FICZ
Mean	32,5	21,9	39,2	54,3	62,7
Median	32	17,5	43,5	53	61
SD	24,2	20,7	17,2	35,2	19,1
t-statistic	-	1,48	-0,349	-1,382	-2,871
p-value	-	0,156	0,731	0,184	0.010

Table 3. Results of the stereomicroscopic macrovisualization of untreated and drug treated ORFeus reporter-bearing mouse livers, 3 months after gene delivery. Independent samples t-tests, with Bonferroni correction to control for multiple comparisons, were used. The adjusted alpha level for statistical significance was 0.013.

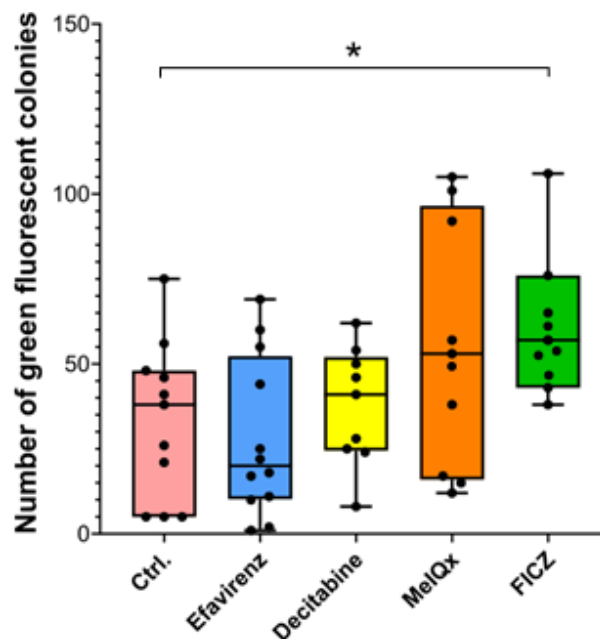


Figure 13. Box plots illustrating the results of stereomicroscopic macrovisualisation of EGFP-positive colonies.

Macrovisualization of EGFP-positive foci through a comprehensive analysis of the superior and inferior liver surfaces revealed significantly increased colony count ($p = 0.01$) in FICZ-treated animals, whereas Efavirenz treatment resulted in a slight, but not significant reduction in fluorescent colony numbers. (Table 3., Fig. 13).

4.2.4 Further methods for quantifying reporter retrotransposition events

To quantify EGFP-positive cells carrying ORFeus retrotransposition events we created a FACS-based pipeline. We performed hepatocyte isolation by hepatic perfusion on two mice from the Decitabine-treated and untreated experimental groups to evaluate this strategy. FACS analysis demonstrated a higher proportion of EGFP-positive cells in the liver samples from Decitabine-treated mice, confirming the increased retrotransposition activity observed in the fluorescence microscopy images. The WT control sample contained an order of magnitude fewer GFP-positive cells, which are potentially contaminating particles, compared to liver samples exhibiting visible GFP-positive colonies. This finding reinforces that our assay accurately measures GFP-positive cells that have undergone reporter retrotransposition (**Fig 14A**).

Additionally, an already existing strategy is qPCR analysis targeting intronless EGFP sequences, which provides quantitative evidence of retrotransposition events by amplifying the intronless copies of EGFP produced by the ORFeus reporter. The treated group exhibited higher levels of intronless EGFP DNA compared to the control group, reinforcing the conclusion that Decitabine enhances LINE-1 retrotransposition activity (**Fig 14B**).

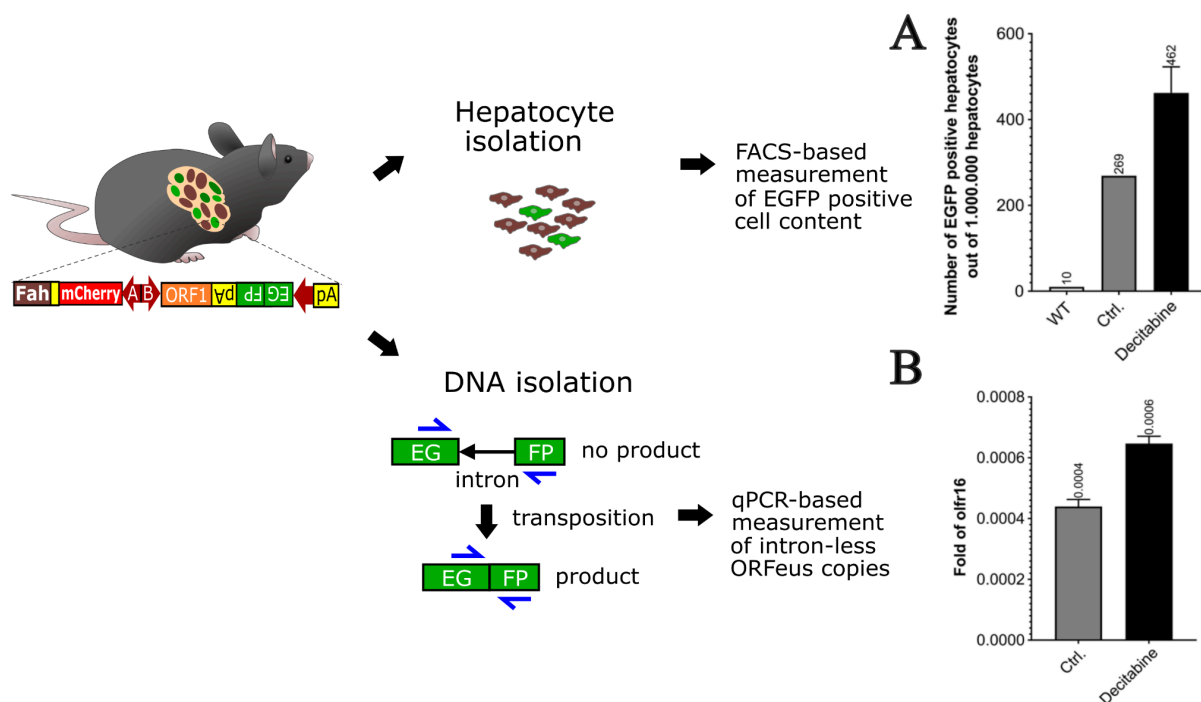


Figure 14. Quantitation of retrotransposition events in the untreated and Decitabine-treated ORFeus reporter bearing mouse livers, 3 months after gene delivery, NTBC withdrawal and Decitabine administration (A) results of FACS analysis after liver perfusion, (B) results of qPCR measurement of retrotransposed ORFeus copies, normalized to Olfr 16, and presented as mean $\pm$ SD. Control samples were obtained from non-drug treated *Fah*^{-/-} mice carrying the ORFeus reporter.

5 DISCUSSION

LINE-1 elements are autonomous retrotransposons whose activity has been linked to genomic instability. The investigation of LINE-1 activity has gained significant momentum, revealing its role not just as a biomarker but also as a contributor to tumorigenesis. This thesis summarized different important aspects of LINE-1 biology. First, we demonstrated an illustrative example of how LINE-1-related proteins are expressed during carcinogenesis from very early stages. Subsequently, we have developed a novel somatic transgenesis platform by which we could create a viable animal model to assess somatic LINE-1 activity, offering a powerful tool for toxicology and cancer research.

We investigated the expression patterns of ORF1p and made a side-by-side comparison with the traditionally used diagnostic markers p16, and Ki67 in cervical precursor lesions and cancer. ORF1p is a promising early marker whose expression pattern can differentiate between various grades of cervical intraepithelial neoplasia through distinct staining patterns. In CIN I, ORF1p expression is primarily localized to the lower third of the cervical epithelium. In CIN II, ORF1p expression expands to the lower two-thirds of the cervical epithelium, while in CIN III and invasive cancer, it reaches the whole epithelial thickness. **(Fig. 6)** The most significant clinical advantage of ORF1p as a marker is its capability to reliably differentiate normal cervical epithelium from CIN I lesions.

The retrotransposition activity of endogenous LINE-1 elements seems to play a significant role in the initial stages of cervical tumorigenesis. This is evidenced by the lack of ORF1p immunopositivity in normal cervical tissues, contrasted by the presence of ORF1 expression in all CIN I lesions, highlighting its early involvement in the process **(Fig. 5)**. Derepressed LINE-1 activity is present in over 50% of cancers and is recognized as a hallmark of cancer [108], contributing to genomic instability and genetic heterogeneity. Unique somatic LINE-1 integration events were identified as causative factors of many cancer driver events in multiple types of tumors [43].

The utility of ORF1p as a pan-tumor biomarker is emerging [108], not only as an IHC-based marker, but preliminary research has indicated that ORF1p is detectable from serum and other body fluids. *Taylor et al.* found that ORF1p detection in plasma may be useful for detecting and monitoring many cancers [129]. *Sato et al.* found ORF1p in ascites and plasma samples from patients with high-grade serous ovarian cancer [110]. These findings suggest its

potential in non-invasive cancer screening, early detection, or even treatment monitoring. However, further large-scale studies are essential to fully validate the diagnostic and prognostic utility of ORF1p across diverse clinical settings.

Based on the absence of ORF1p in normal tissues and its extensive presence in dysplastic lesions, the opportunity for targeting LINE-1 arises. In a phase 2 clinical study, suppressing LINE-1 activity has shown therapeutic benefit in the treatment of metastatic colorectal cancer by using 3TC (Lamivudine), which is a nucleoside reverse transcriptase inhibitor (NRTI) used in anti-HIV therapy. In a different phase 2 trial, efavirenz was evaluated as a possible therapeutic agent for metastatic castration-resistant prostate cancer. This finding indicates that future cancer therapies might be augmented by drugs that specifically target LINE-1 activity [130].

It is now a well-established fact that novel somatic LINE-1 insertions can generate tumor driver events potentially responsible for sporadic cancer cases where no clear cause or known trigger can be identified. We hypothesize that numerous environmental factors can affect LINE-1 activity, although their impact has not yet been investigated.

Environmental factors can trigger a range of cellular responses. One challenge in evaluating the effects of xenobiotics or other environmental factors lies in distinguishing direct effects from indirect ones. Much of the available data comes from tissue culture experiments, where any adverse effects on cell viability caused by the tested compound could skew the results. On the other hand, the majority of cell lines are tumor-derived and long-cultured, so their LINE-1 defense systems can be compromised and not represent the *in vivo* circumstances [131].

Additionally, the difficulty in assessing LINE-1 somatic retrotransposition creates a significant gap in knowledge of how exposure to various environmental factors influences LINE-1 activity in humans. Conventionally used genotoxicity screening methods currently include the Ames test [132], various mammalian cell culture-based methods; such as gene mutation assays, chromosomal aberration tests, micronucleus tests, and mammalian alkaline comet assay [133], [134]. These methods, despite their high sensitivity, often suffer from low specificity, and tend to identify only certain point mutations or large-scale genomic changes. Although these tests can detect a range of DNA damage, if we cannot accurately measure LINE-1 activity in somatic cells, our ability to assess the full genotoxic potential of

environmental exposures remains limited. For example, while some chemicals may not directly damage DNA, they might still induce retrotransposition events that disrupt genomic integrity indirectly.

Therefore, the development of more sensitive and specific methods to detect LINE-1 somatic retrotransposition, as explored in this work, will lead to a more comprehensive understanding of genotoxicity and the underlying mechanisms. This advancement could fill a critical gap in toxicology studies, enhancing our ability to evaluate and mitigate risks associated with environmental and occupational exposures.

A compound that stimulates LINE-1 might promote the demethylation of the LINE-1 promoter, increase transcription, and inhibit the defense mechanisms. It is plausible that future research will reveal that the risk of developing an LINE-1-induced disease depends both on the presence of "hot" LINE-1s in the genome of individuals and their history of exposure to certain xenobiotics that enhance LINE-1 activity. It is important to note that different LINE-1 loci in various tissues may respond uniquely to physiological or environmental triggers, adding another layer of complexity [29], [135], [136], [137].

We established an *in vivo* somatic transgenic mouse model for assessing LINE-1 retrotransposition and monitoring the effects of different exposures. To develop an effective animal model, we utilized innovative strategies. First, we developed a robust platform for introducing loss-of-function and gain-of-function modifications into the genome of experimental animals, since conventional methods to permanently manipulate the somatic genome face challenges. In particular, transcriptional repression hinders transgene expression [115][118], [138]. It is caused by the formation of repressive heterochromatin on plasmid DNA. This process can be triggered by cytoplasmic DNA sensors and may occur through CpG methylation or independent pathways [139], leading to a quick decline in transgene expression even with SB transposon-based systems. Previous studies have shown that transgene expression stabilizes at about 1% of its initial level weeks after transposition [98]. Our experiments show an even lower stabilization of Luc expression at 0.26% in wild-type mice.

However, using Fah-deficient livers, we achieved long-lasting and physiological gene expression throughout almost the entire liver. In these Fah-deficient mice, Luc expression increased up to 61 times over three months. This increase is likely due to the positive

selection pressure on *Fah* marker expression, and by liver regeneration by the propagation of *Fah* expressing hepatocytes. Although initial activation of DNA sensors cannot be avoided, the *Fah* KO mice showed a temporary decrease in transgene expression around day 56, which eventually ceased.

The SB transposon system has been utilized before for the treatment of TT1 by supplying therapeutic *Fah* gene copies [140]. Our versatile gene expression modification system technology has updated the existing methods by using the hyperactive SB100 transposase for a more effective chromosomal transgene delivery. We aimed for tag-free protein expression linked to the selection marker, using the HADHA/B bidirectional promoter for balanced expression. This promoter allows marker-linked expression of untagged native or mutant proteins, achieving a nearly 1/1 transcript ratio, which is crucial for producing protein complexes.

This platform enabled us to insert the synthetic LINE-1 reporter element, ORFeus, to the first side of the promoter, while keeping *Fah* as a positive selection marker gene on the opposite side of the promoter, which facilitates the selection of successfully transfected cells. This configuration ensures that the expression of the LINE-1 reporter is tightly linked to the integration and activity of the transgene within the hepatocytes, enabling the monitoring of LINE-1 activity in a physiologically relevant context.

The implementation of this vector into our *Fah* KO mice allows for the assessment of LINE-1 retrotransposition events. The retrotransposition reporter protein (EGFP) encoded by the ORFeus element is only expressed upon successful retrotransposition, providing a direct readout of LINE-1 activity. The nearly 100 million hepatocytes available in a single experimental animal provide a substantial pool of cells for functional testing. This high number of testable cells allows for the assessment of chemical compounds and their impact on LINE-1 activity as a potential genotoxic effect.

There are multiple approaches available to analyze our retrotransposition assay. The initial method involves stereomicroscopic visualization of EGFP-positive colonies on the liver surface, providing insight into early retrotransposition events that occurred within the first few weeks of drug treatment. These early events result in visible colony formation during multi-nodular repopulation, while later events remain undetectable with this approach. To estimate the overall number of affected cells, including those from later events, FACS and

qPCR-based methods are employed. This potentially explains why Decitabine treatment led to an obvious increase in EGFP-positive cells, as measured by FACS and confirmed by qPCR, even though macrovisualization did not show a marked increase in EGFP positive colony numbers. Our future direction is to develop a new generation of ORFeus reporters that will enable IHC-based readouts of reporter activation, offering a precise quantitation opportunity, giving a more detailed understanding of the dynamics of drug effects. It is crucial to highlight that using the current ORFeus reporter system, every cell expressing the reporter shows EGFP positivity in IHC with all commercially available EGFP antibodies we have tested. This raises an important consideration: despite the claims made by the authors in the Okudaira et al. study [94], [95], their results actually demonstrate L1 reporter expression rather than true L1 reporter activation with *de novo* retrotransposition events.

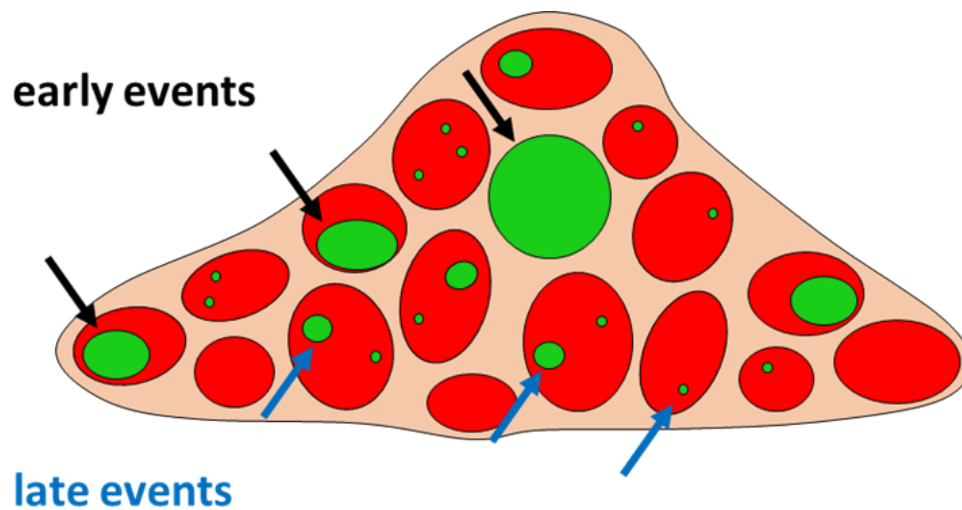


Figure 15. Schematic representation of an ORFeus reporter bearing *Fah*^{-/-} liver gone through multi-nodular repopulation. Red circles represent repopulated, mCherry expressing liver areas, Green ellipses represent EGFP positive cell colonies carrying retrotransposition events.

We anticipate our mouse model as a valuable tool in toxicology for evaluating the genotoxic potential of indirect mutagens via the activation of LINE-1 retrotransposons, thus identifying new environmental risk factors and enhancing public health safety. It is important to acknowledge that the study relies on a mouse model, which presents inherent limitations due to species-specific differences. For example, metabolic rates, immune system responses, hormone regulation, and drug metabolism could differ between human and laboratory mice. The controlled environment in which laboratory mice are kept, does not replicate the diverse and complex environments of humans. Factors such as diet, stress, and exposure to pathogens

can influence disease development and progression differently, requiring careful consideration when extrapolating findings. On the other hand, the fundamental biology of LINE-1 elements is largely conserved between mice and humans. Both species exhibit minimal LINE-1 activity in the germline, with somatic activity typically suppressed under physiological conditions. However, during tumorigenesis, LINE-1 activity is reactivated in somatic cells, and the shared mechanisms of LINE-1 control in both species, along with the functional interchangeability of human and mouse LINE-1 proteins, make our LINE-1 activity-based murine toxicology assay as relevant as conventional rodent toxicology studies.

Our findings provide an initial validation of the model system, demonstrating its potential. However, the study was conducted on a limited set of chemicals, and a broader screening with more compounds and a wider range of doses would allow for a more comprehensive evaluation. Investigating dose-dependent effects in greater detail could strengthen our conclusions. Additionally, we expect that the IHC-based readout will provide more quantitative and statistically robust data.

Overall, the study demonstrates the effectiveness of the ORFeus reporter system implemented in a somatically transgenic mouse model for the effective detection of somatic LINE-1 retrotransposition activity and highlights the potential genotoxic effects of chemicals. These findings pave the way for further research into the mechanisms by which environmental factors influence genomic stability and cancer susceptibility.

6 CONCLUSION

Our results highlight the potentially widespread and important role of somatic L1 activity in the development of sporadic cancers. The results of this study have significant implications for the field of toxicology and chemical risk assessment. The developed technology platform allows for a novel approach to evaluating the genotoxic effects of chemicals that might not be detectable through traditional methods. The ability to measure LINE-1 retrotransposition activity in somatic cells provides a new dimension to understand how environmental factors contribute to cancer development. This innovative tool could be used to screen various chemicals for their potential to induce genomic instability, thereby aiding in the identification of new environmental risks and contributing to public health safety.

7 ACKNOWLEDGEMENTS

I would like to take this opportunity to express my heartfelt gratitude to the many individuals who have contributed to my journey throughout this research.

First and foremost, I would like to thank my supervisor, **Lajos Mátés**, and his partner in both life and profession, **Andrea Mátés-Nagy**, whose guidance and expertise have shaped my research and academic growth. Thank you for welcoming me into your lab back in 2014 as a first-year medical student searching for a mentor and navigating my path toward a research career.

During this process, I was lucky enough to have beautiful lab members by my side. Thank you, **Liza Hudoba**, for taking me under your wing and teaching me the ins and outs of research, all while becoming a cherished friend along the way.

András Blastyák, your endless supply of ideas, creativity, and unrelenting curiosity, combined with your love for science and benchwork, have kept me inspired and grounded in this field.

A special thanks goes to **Imre Gergely**; you were always there in stressful times and offered a listening ear whenever I needed it. Our shared sense of humor (*around 4 PM*) brightened even the toughest days.

I must acknowledge the contributions of **Ildikó Fekete**, whose technical expertise in the lab was crucial for the success of our experiments. You turned the lab into a true "nest", a place that was always comforting to return to.

I want to express my gratitude to **Anna Georgina Kopasz**, **Dávid Pusztai**, **Andrea Bakné Drubi**, **Erzsébet Fehérné Juhász** and **Khaldoon Sadiq** for their support throughout the years working together. I would like to thank the new members of the lab, **May Raya**, **Kitti Törőcsik**, **Olívia Verebi**, and **Dominik Sándor Kocsis** for carrying our work forward. Your dedication and enthusiasm ensure that the impact of our research will continue for years to come.

I am grateful to the staff of the Central Animal House for their conscientious care and maintenance, ensuring an appropriate environment for our animal experiments.

I would like to thank the colleagues from the Department of Pathology, Szeged: **Pankotai-Bodó Gabriella**, **Ivett Csikós**, **Sarolta Bankó**, **Ádám Ürmös**, **Tibor Kun**, **Beáta Balogh**, **Krisztián Daru**, and **Mihály Dezső**. My thesis would not have been possible without your collaboration.

I am immensely grateful to **Farkas Sükösd**, for fostering an environment of collaboration and support. Your faith in my abilities inspires me every day. You taught me; *“It's not the destination that matters, but the people you share the journey with.”*

Additionally, I wish to express my gratitude to **Ágnes Czibula**, **Dávid Szüts**, and **Zoltán Villányi** for providing constructive feedback on my work. Your insights helped to enhance the quality of my thesis.

I would like to thank the entire Biological Research Center, especially the Institute of Genetics, for providing a space where I had the privilege to study molecular biology, surrounded by inspiring minds.

Last but certainly not least, I would like to thank my everloving family - whose unconditional love and support have been a source of strength throughout this journey. I am also thankful to my wonderful friends who have always lifted my spirits and taken care of my mental health with kind words and meaningful conversations. Your support has meant the world to me.

8 FUNDING

The project was, in part, funded by the Hungarian Academy of Sciences Momentum II Programme (LP2015-5/2015) , the GINOP (GINOP-2.3.2-15-2016-00024) and 2021-1.1.4-GYORSÍTÓSAV-2022-00018 project of the National Research, Development and Innovation Fund, Hungary. Réka Karkas was supported by the ÚNKP-23-3 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development, and Innovation Fund, Hungary.

9 BIBLIOGRAPHY

- [1] K. H. Burns, “Transposable elements in cancer,” *Nat. Rev. Cancer*, vol. 17, no. 7, pp. 415–424, Jul. 2017, doi: 10.1038/nrc.2017.35.
- [2] M. Lee, S. F. Ahmad, and J. Xu, “Regulation and function of transposable elements in cancer genomes,” *Cell. Mol. Life Sci.*, vol. 81, no. 1, p. 157, Mar. 2024, doi: 10.1007/s00018-024-05195-2.
- [3] E. T. L. Prak and H. H. K. Jr, “Mobile Elements and the Human Genome,” *Nat. Rev. Genet.*, vol. 1, no. November, 2000.
- [4] D. B. Roth and N. L. Craig, “VDJ Recombination: A Transposase Goes to Work,” *Cell*, vol. 94, no. 4, pp. 411–414, Aug. 1998, doi: 10.1016/S0092-8674(00)81580-9.
- [5] Z. Ivics, P. B. Hackett, R. H. Plasterk, and Z. Izsvák, “Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells,” *Cell*, vol. 91, no. 4, pp. 501–510, 1997, doi: 10.1016/S0092-8674(00)80436-5.
- [6] K. Yusa, “piggyBac Transposon,” *Microbiol. Spectr.*, vol. 3, no. 2, pp. MDNA3-0028–2014, Apr. 2015, doi: 10.1128/microbiolspec.MDNA3-0028-2014.
- [7] C. F. Magnani *et al.*, “Sleeping Beauty-engineered CAR T cells achieve antileukemic activity without severe toxicities,” *J. Clin. Invest.*, vol. 130, no. 11, pp. 6021–6033, 2020, doi: 10.1172/JCI138473.
- [8] Y. Zhang *et al.*, “Phase I clinical trial of EGFR-specific CAR-T cells generated by the piggyBac transposon system in advanced relapsed/refractory non-small cell lung cancer patients,” *J. Cancer Res. Clin. Oncol.*, vol. 147, no. 12, pp. 3725–3734, 2021, doi: 10.1007/s00432-021-03613-7.
- [9] A. Nagy, G. Imre, and L. Mátés, “Transposon-based gene delivery vectors for gene therapy,” *Acta Biol. Szeged.*, vol. 59, no. April 2002, pp. 247–260, 2015.
- [10] I. Grabundzija *et al.*, “Comparative analysis of transposable element vector systems in human cells,” *Mol. Ther. J. Am. Soc. Gene Ther.*, vol. 18, no. 6, pp. 1200–1209, Jun. 2010, doi: 10.1038/mt.2010.47.
- [11] T. Wicker *et al.*, “A unified classification system for eukaryotic transposable elements,” *Nat. Rev. Genet.*, vol. 8, no. 12, pp. 973–982, Dec. 2007, doi: 10.1038/nrg2165.
- [12] R. Cordaux and M. A. Batzer, “The impact of retrotransposons on human genome evolution,” *Nat. Rev. Genet.*, vol. 10, no. 10, pp. 691–703, Oct. 2009, doi: 10.1038/nrg2640.
- [13] V. P. Belancio, D. J. Hedges, and P. Deininger, “Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health,” *Genome Res.*, vol. 18, no. 3, pp. 343–358, Mar. 2008, doi: 10.1101/gr.5558208.
- [14] E. R. Havecker, X. Gao, and D. F. Voytas, “The diversity of LTR retrotransposons,” *Genome Biol.*, vol. 5, no. 6, p. 225, 2004, doi: 10.1186/gb-2004-5-6-225.
- [15] X.-O. Zhang, H. Pratt, and Z. Weng, “Investigating the Potential Roles of SINEs in the Human Genome,” *Annu. Rev. Genomics Hum. Genet.*, vol. 22, no. Volume 22, 2021, pp. 199–218, Aug. 2021, doi: 10.1146/annurev-genom-111620-100736.
- [16] P. Deininger, “Alu elements: know the SINEs,” *Genome Biol.*, vol. 12, no. 12, p. 236, Dec. 2011, doi: 10.1186/gb-2011-12-12-236.
- [17] “Alu elements as regulators of gene expression | Nucleic Acids Research | Oxford Academic.” Accessed: Sep. 14, 2024. [Online]. Available: <https://academic.oup.com/nar/article/34/19/5491/3112003>
- [18] J. S. Han, “Non-long terminal repeat (non-LTR) retrotransposons: mechanisms, recent developments, and unanswered questions,” *Mob. DNA*, vol. 1, p. 15, May 2010, doi: 10.1186/1759-8753-1-15.
- [19] S. L. Martin and F. D. Bushman, “Nucleic acid chaperone activity of the ORF1 protein

- from the mouse LINE-1 retrotransposon,” *Mol. Cell. Biol.*, vol. 21, no. 2, pp. 467–475, 2001, doi: 10.1128/MCB.21.2.467-475.2001.
- [20] H. Hohjoh and M. F. Singer, “Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA,” *EMBO J.*, vol. 15, no. 3, pp. 630–639, Feb. 1996.
- [21] Y. Xie, L. Mates, Z. Ivics, Z. Izsvák, S. L. Martin, and W. An, “Cell division promotes efficient retrotransposition in a stable L1 reporter cell line,” *Mob. DNA*, vol. 4, no. 1, p. 10, Mar. 2013, doi: 10.1186/1759-8753-4-10.
- [22] S. Kubo *et al.*, “L1 retrotransposition in nondividing and primary human somatic cells,” *Proc. Natl. Acad. Sci.*, vol. 103, no. 21, pp. 8036–8041, May 2006, doi: 10.1073/pnas.0601954103.
- [23] B. T. Freeman, M. Sokolowski, A. M. Roy-Engel, M. E. Smither, and V. P. Belancio, “Identification of charged amino acids required for nuclear localization of human L1 ORF1 protein,” *Mob. DNA*, vol. 10, no. 1, p. 20, May 2019, doi: 10.1186/s13100-019-0159-2.
- [24] A. V. Horn, I. Celic, C. Dong, I. Martirosyan, and J. S. Han, “A conserved role for the ESCRT membrane budding complex in LINE retrotransposition,” *PLOS Genet.*, vol. 13, no. 6, p. e1006837, Jun. 2017, doi: 10.1371/journal.pgen.1006837.
- [25] “LINE-1 protein localization and functional dynamics during the cell cycle | eLife.” Accessed: Sep. 14, 2024. [Online]. Available: <https://elifesciences.org/articles/30058>
- [26] A. Thawani, A. J. F. Ariza, E. Nogales, and K. Collins, “Template and target-site recognition by human LINE-1 in retrotransposition,” *Nature*, vol. 626, no. 7997, pp. 186–193, Feb. 2024, doi: 10.1038/s41586-023-06933-5.
- [27] G. J. Cost, Q. Feng, A. Jacquier, and J. D. Boeke, “Human L1 element target-primed reverse transcription in vitro,” *EMBO J.*, vol. 21, no. 21, pp. 5899–5910, Nov. 2002, doi: 10.1093/emboj/cdf592.
- [28] Mouse Genome Sequencing Consortium *et al.*, “Initial sequencing and comparative analysis of the mouse genome,” *Nature*, vol. 420, no. 6915, pp. 520–562, 2002, doi: 10.1038/nature01262.
- [29] H. H. Kazazian and J. V. Moran, “Mobile DNA in Health and Disease,” *N. Engl. J. Med.*, vol. 377, no. 4, pp. 361–370, Jul. 2017, doi: 10.1056/NEJMr1510092.
- [30] W. McKerrow *et al.*, “LINE-1 retrotransposon expression in cancerous, epithelial and neuronal cells revealed by 5' single-cell RNA-Seq,” *Nucleic Acids Res.*, vol. 51, no. 5, pp. 2033–2045, Mar. 2023, doi: 10.1093/nar/gkad049.
- [31] B. J. Wagstaff, M. Barnerssoi, and A. M. Roy-Engel, “Evolutionary conservation of the functional modularity of primate and murine LINE-1 elements,” *PloS One*, vol. 6, no. 5, p. e19672, May 2011, doi: 10.1371/journal.pone.0019672.
- [32] B. Brouha *et al.*, “Hot L1s account for the bulk of retrotransposition in the human population,” *Proc. Natl. Acad. Sci.*, vol. 100, no. 9, pp. 5280–5285, Apr. 2003, doi: 10.1073/pnas.0831042100.
- [33] C. R. Beck *et al.*, “LINE-1 Retrotransposition Activity in Human Genomes,” *Cell*, vol. 141, no. 7, pp. 1159–1170, Jun. 2010, doi: 10.1016/j.cell.2010.05.021.
- [34] T. Eickbush, “Exon Shuffling in Retrospect,” *Science*, vol. 283, no. 5407, pp. 1465–1467, Mar. 1999, doi: 10.1126/science.283.5407.1465.
- [35] J. V. Moran, R. J. DeBerardinis, and H. H. Kazazian, “Exon shuffling by L1 retrotransposition,” *Science*, vol. 283, no. 5407, pp. 1530–1534, Mar. 1999, doi: 10.1126/science.283.5407.1530.
- [36] H. H. Kazazian, C. Wong, H. Youssoufian, A. F. Scott, D. G. Phillips, and S. E. Antonarakis, “Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man,” *Nature*, vol. 332, no. 6160, pp. 164–166, Mar. 1988, doi: 10.1038/332164a0.

- [37] D. C. Hancks and H. H. Kazazian, “Active human retrotransposons: variation and disease,” *Curr. Opin. Genet. Dev.*, vol. 22, no. 3, pp. 191–203, Jun. 2012, doi: 10.1016/j.gde.2012.02.006.
- [38] J. Xing *et al.*, “Mobile elements create structural variation: Analysis of a complete human genome,” *Genome Res.*, vol. 19, no. 9, pp. 1516–1526, Jan. 2009, doi: 10.1101/gr.091827.109.
- [39] J. Feusier *et al.*, “Pedigree-based estimation of human mobile element retrotransposition rates,” *Genome Res.*, vol. 29, no. 10, pp. 1567–1577, Oct. 2019, doi: 10.1101/gr.247965.118.
- [40] H. Kano *et al.*, “L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism,” *Genes Dev.*, vol. 23, no. 11, pp. 1303–1312, Jan. 2009, doi: 10.1101/gad.1803909.
- [41] S. R. Richardson *et al.*, “Heritable L1 retrotransposition in the mouse primordial germline and early embryo,” *Genome Res.*, vol. 27, no. 8, pp. 1395–1405, 2017, doi: 10.1101/gr.219022.116.
- [42] L. Xiao-Jie, X. Hui-Ying, X. Qi, X. Jiang, and M. Shi-Jie, “LINE-1 in cancer: multifaceted functions and potential clinical implications,” *Genet. Med. Off. J. Am. Coll. Med. Genet.*, vol. 18, no. 5, pp. 431–439, May 2016, doi: 10.1038/gim.2015.119.
- [43] B. Rodriguez-Martin *et al.*, “Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition,” *Nat. Genet.*, vol. 52, no. 3, pp. 306–319, 2020, doi: 10.1038/s41588-019-0562-0.
- [44] K. J. Kines, M. Sokolowski, D. L. deHaro, C. M. Christian, and V. P. Belancio, “Potential for genomic instability associated with retrotranspositionally-incompetent L1 loci,” *Nucleic Acids Res.*, vol. 42, no. 16, pp. 10488–10502, Sep. 2014, doi: 10.1093/nar/gku687.
- [45] K. J. Kines *et al.*, “The endonuclease domain of the LINE-1 ORF2 protein can tolerate multiple mutations,” *Mob. DNA*, vol. 7, p. 8, Apr. 2016, doi: 10.1186/s13100-016-0064-x.
- [46] E. C. Scott and S. E. Devine, “The Role of Somatic L1 Retrotransposition in Human Cancers,” *Viruses*, vol. 9, no. 6, p. 131, May 2017, doi: 10.3390/v9060131.
- [47] M. Dewannieux, C. Esnault, and T. Heidmann, “LINE-mediated retrotransposition of marked Alu sequences,” *Nat. Genet.*, vol. 35, no. 1, pp. 41–48, Sep. 2003, doi: 10.1038/ng1223.
- [48] R.-L. Troskie, G. J. Faulkner, and S. W. Cheetham, “Processed pseudogenes: A substrate for evolutionary innovation,” *BioEssays*, vol. 43, no. 11, p. 2100186, 2021, doi: 10.1002/bies.202100186.
- [49] W. Wei *et al.*, “Human L1 Retrotransposition: cisPreference versus trans Complementation,” *Mol. Cell. Biol.*, vol. 21, no. 4, pp. 1429–1439, Feb. 2001, doi: 10.1128/MCB.21.4.1429-1439.2001.
- [50] M. S. Protasova, T. V. Andreeva, and E. I. Rogaev, “Factors Regulating the Activity of LINE1 Retrotransposons,” *Genes*, vol. 12, no. 10, Art. no. 10, Oct. 2021, doi: 10.3390/genes12101562.
- [51] M. Bodak, J. Yu, and C. Ciaudo, “Regulation of LINE-1 in mammals,” *Biomol. Concepts*, vol. 5, no. 5, pp. 409–428, Oct. 2014, doi: 10.1515/bmc-2014-0018.
- [52] D. M. Ozata, I. Gainetdinov, A. Zoch, D. O’Carroll, and P. D. Zamore, “PIWI-interacting RNAs: small RNAs with big functions,” *Nat. Rev. Genet.*, vol. 20, no. 2, pp. 89–108, Feb. 2019, doi: 10.1038/s41576-018-0073-3.
- [53] A. A. Aravin *et al.*, “A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice,” *Mol. Cell*, vol. 31, no. 6, pp. 785–799, Sep. 2008, doi: 10.1016/j.molcel.2008.09.003.

- [54] M. Van Meter *et al.*, “SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age,” *Nat. Commun.*, vol. 5, p. 5011, Sep. 2014, doi: 10.1038/ncomms6011.
- [55] B. N. Vazquez *et al.*, “SIRT7 mediates L1 elements transcriptional repression and their association with the nuclear lamina,” *Nucleic Acids Res.*, vol. 47, no. 15, pp. 7870–7885, Sep. 2019, doi: 10.1093/nar/gkz519.
- [56] D. E. Montoya-Durango, K. A. Ramos, P. Bojang, L. Ruiz, I. N. Ramos, and K. S. Ramos, “LINE-1 silencing by retinoblastoma proteins is effected through the nucleosomal and remodeling deacetylase multiprotein complex,” *BMC Cancer*, vol. 16, no. 1, p. 38, Jan. 2016, doi: 10.1186/s12885-016-2068-9.
- [57] J. L. Goodier, L. Zhang, M. R. Vetter, and H. H. Kazazian, “LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex,” *Mol. Cell. Biol.*, vol. 27, no. 18, pp. 6469–6483, Sep. 2007, doi: 10.1128/MCB.00332-07.
- [58] S. R. Heras *et al.*, “The Microprocessor controls the activity of mammalian retrotransposons,” *Nat. Struct. Mol. Biol.*, vol. 20, no. 10, pp. 1173–1181, Oct. 2013, doi: 10.1038/nsmb.2658.
- [59] X. Zhang, R. Zhang, and J. Yu, “New Understanding of the Relevant Role of LINE-1 Retrotransposition in Human Disease and Immune Modulation,” *Front. Cell Dev. Biol.*, vol. 8, no. August, pp. 1–10, 2020, doi: 10.3389/fcell.2020.00657.
- [60] Q. Yu *et al.*, “Type I Interferon Controls Propagation of Long Interspersed Element-1 *,” *J. Biol. Chem.*, vol. 290, no. 16, pp. 10191–10199, Apr. 2015, doi: 10.1074/jbc.M114.612374.
- [61] M. De Cecco *et al.*, “L1 drives IFN in senescent cells and promotes age-associated inflammation,” *Nature*, vol. 566, no. 7742, pp. 73–78, Feb. 2019, doi: 10.1038/s41586-018-0784-9.
- [62] A. Zhang, B. Dong, A. J. Doucet, J. B. Moldovan, J. V. Moran, and R. H. Silverman, “RNase L restricts the mobility of engineered retrotransposons in cultured human cells,” *Nucleic Acids Res.*, vol. 42, no. 6, pp. 3803–3820, Apr. 2014, doi: 10.1093/nar/gkt1308.
- [63] P. Kapasi *et al.*, “L13a Blocks 48S Assembly: Role of a General Initiation Factor in mRNA-specific Translational Control,” *Mol. Cell*, vol. 25, no. 1, pp. 113–126, Jan. 2007, doi: 10.1016/j.molcel.2006.11.028.
- [64] P. Li *et al.*, “Aicardi–Goutières syndrome protein TREX1 suppresses L1 and maintains genome integrity through exonuclease-independent ORF1p depletion,” *Nucleic Acids Res.*, vol. 45, no. 8, pp. 4619–4631, May 2017, doi: 10.1093/nar/gkx178.
- [65] J. L. Goodier, L. E. Cheung, and H. H. Kazazian, “MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells,” *PLoS Genet.*, vol. 8, no. 10, p. e1002941, 2012, doi: 10.1371/journal.pgen.1002941.
- [66] Z. Warkocki, P. S. Krawczyk, D. Adamska, K. Bijata, J. L. Garcia-Perez, and A. Dziembowski, “Uridylation by TUT4/7 Restricts Retrotransposition of Human LINE-1s,” *Cell*, vol. 174, no. 6, pp. 1537–1548.e29, Sep. 2018, doi: 10.1016/j.cell.2018.07.022.
- [67] J. Du *et al.*, “Nucleocytoplasmic shuttling of SAMHD1 is important for LINE-1 suppression,” *Biochem. Biophys. Res. Commun.*, vol. 510, no. 4, pp. 551–557, Mar. 2019, doi: 10.1016/j.bbrc.2019.02.009.
- [68] S. R. Richardson, I. Narvaiza, R. A. Planegger, M. D. Weitzman, and J. V. Moran, “APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition,” *eLife*, vol. 3, p. e02008, Apr. 2014, doi: 10.7554/eLife.02008.
- [69] A. V. Horn *et al.*, “Human LINE-1 restriction by APOBEC3C is deaminase independent and mediated by an ORF1p interaction that affects LINE reverse transcriptase activity,”

- Nucleic Acids Res.*, vol. 42, no. 1, pp. 396–416, 2014, doi: 10.1093/nar/gkt898.
- [70] Y. Liang, X. Qu, N. M. Shah, and T. Wang, “Towards targeting transposable elements for cancer therapy,” *Nat. Rev. Cancer*, vol. 24, no. 2, pp. 123–140, Feb. 2024, doi: 10.1038/s41568-023-00653-8.
 - [71] M. Kinomoto *et al.*, “All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition,” *Nucleic Acids Res.*, vol. 35, no. 9, pp. 2955–2964, May 2007, doi: 10.1093/nar/gkm181.
 - [72] T. A. Morrish *et al.*, “DNA repair mediated by endonuclease-independent LINE-1 retrotransposition,” *Nat. Genet.*, vol. 31, no. 2, pp. 159–165, Jun. 2002, doi: 10.1038/ng898.
 - [73] G. Servant *et al.*, “The Nucleotide Excision Repair Pathway Limits L1 Retrotransposition,” *Genetics*, vol. 205, no. 1, pp. 139–153, 2017, doi: 10.1534/genetics.116.188680.
 - [74] N. G. Coufal *et al.*, “Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells,” *Proc. Natl. Acad. Sci.*, vol. 108, no. 51, pp. 20382–20387, Dec. 2011, doi: 10.1073/pnas.1100273108.
 - [75] K. Watanabe, S. Tateishi, M. Kawasuji, T. Tsurimoto, H. Inoue, and M. Yamaizumi, “Rad18 guides pol η to replication stalling sites through physical interaction and PCNA monoubiquitination,” *EMBO J.*, vol. 23, no. 19, pp. 3886–3896, Sep. 2004, doi: 10.1038/sj.emboj.7600383.
 - [76] D. Ardeljan *et al.*, “Cell fitness screens reveal a conflict between LINE-1 retrotransposition and DNA replication,” *Nat. Struct. Mol. Biol.*, vol. 27, no. 2, pp. 168–178, Feb. 2020, doi: 10.1038/s41594-020-0372-1.
 - [77] Y. Ariumi *et al.*, “DNA repair protein Rad18 restricts LINE-1 mobility,” *Sci. Rep.*, vol. 8, no. 1, p. 15894, Oct. 2018, doi: 10.1038/s41598-018-34288-9.
 - [78] P. Mita *et al.*, “BRCA1 and S phase DNA repair pathways restrict LINE-1 retrotransposition in human cells,” *Nat. Struct. Mol. Biol.*, vol. 27, no. 2, pp. 179–191, Feb. 2020, doi: 10.1038/s41594-020-0374-z.
 - [79] W. McKerrow *et al.*, “LINE-1 expression in cancer correlates with p53 mutation, copy number alteration, and S phase checkpoint,” *Proc. Natl. Acad. Sci.*, vol. 119, no. 8, p. e2115999119, Feb. 2022, doi: 10.1073/pnas.2115999119.
 - [80] Y. Guan *et al.*, “The MOV10 RNA helicase is a dosage-dependent host restriction factor for LINE1 retrotransposition in mice,” *PLoS Genet.*, vol. 19, no. 5, p. e1010566, May 2023, doi: 10.1371/journal.pgen.1010566.
 - [81] X. Li *et al.*, “The MOV10 helicase inhibits LINE-1 mobility,” *J. Biol. Chem.*, vol. 288, no. 29, pp. 21148–21160, Jul. 2013, doi: 10.1074/jbc.M113.465856.
 - [82] B. Del Re and G. Giorgi, “Long INterspersed element-1 mobility as a sensor of environmental stresses,” *Environ. Mol. Mutagen.*, vol. 61, no. 4, pp. 465–493, 2020, doi: 10.1002/em.22366.
 - [83] P. Boffetta and F. Nyberg, “Contribution of environmental factors to cancer risk,” *Br. Med. Bull.*, vol. 68, pp. 71–94, 2003, doi: 10.1093/bmp/ldg023.
 - [84] I. Teneng, D. E. Montoya-Durango, J. L. Quertermous, M. E. Lacy, and K. S. Ramos, “Reactivation of L1 retrotransposon by benzo(a)pyrene involves complex genetic and epigenetic regulation,” *Epigenetics*, vol. 6, no. 3, pp. 355–367, Mar. 2011, doi: 10.4161/epi.6.3.14282.
 - [85] S. P. Kale, L. Moore, P. L. Deininger, and A. M. Roy-Engel, “Heavy Metals Stimulate Human LINE-1 Retrotransposition,” *Int. J. Environ. Res. Public. Health*, vol. 2, no. 1, pp. 14–23, May 2005.
 - [86] P. Whongsiri *et al.*, “LINE-1 ORF1 Protein Is Up-regulated by Reactive Oxygen Species and Associated with Bladder Urothelial Carcinoma Progression,” *Cancer*

- Genomics Proteomics*, vol. 15, no. 2, pp. 143–151, Mar. 2018.
- [87] M. Kaneko, M. Kotake, H. Bando, T. Yamada, H. Takemura, and T. Minamoto, “Prognostic and predictive significance of long interspersed nucleotide element-1 methylation in advanced-stage colorectal cancer,” *BMC Cancer*, vol. 16, p. 945, Dec. 2016, doi: 10.1186/s12885-016-2984-8.
 - [88] P. A. Jones and S. B. Baylin, “The fundamental role of epigenetic events in cancer,” *Nat. Rev. Genet.*, vol. 3, no. 6, pp. 415–428, Jun. 2002, doi: 10.1038/nrg816.
 - [89] D. J. Weisenberger *et al.*, “Analysis of repetitive element DNA methylation by MethyLight,” *Nucleic Acids Res.*, vol. 33, no. 21, pp. 6823–6836, 2005, doi: 10.1093/nar/gki987.
 - [90] N. Terasaki *et al.*, “In Vitro Screening for Compounds That Enhance Human L1 Mobilization,” *PLoS ONE*, vol. 8, no. 9, p. e74629, Sep. 2013, doi: 10.1371/journal.pone.0074629.
 - [91] J. S. Han and J. D. Boeke, “A highly active synthetic mammalian retrotransposon,” *Nature*, vol. 429, no. 6989, pp. 314–318, May 2004, doi: 10.1038/nature02535.
 - [92] W. An *et al.*, “Characterization of a synthetic human LINE-1 retrotransposon ORFeus-Hs,” *Mob. DNA*, vol. 2, no. 1, p. 2, Feb. 2011, doi: 10.1186/1759-8753-2-2.
 - [93] H. C. Kopera, P. A. Larson, J. B. Moldovan, S. R. Richardson, Y. Liu, and J. V. Moran, “LINE-1 Cultured Cell Retrotransposition Assay,” *Methods Mol. Biol. Clifton NJ*, vol. 1400, pp. 139–156, 2016, doi: 10.1007/978-1-4939-3372-3_10.
 - [94] N. Okudaira *et al.*, “Involvement of retrotransposition of long interspersed nucleotide element-1 in skin tumorigenesis induced by 7,12-dimethylbenz[a]anthracene and 12-O-tetradecanoylphorbol-13-acetate,” *Cancer Sci.*, vol. 102, no. 11, pp. 2000–2006, 2011, doi: 10.1111/j.1349-7006.2011.02060.x.
 - [95] N. Okudaira *et al.*, “Long interspersed element-1 is differentially regulated by food-borne carcinogens via the aryl hydrocarbon receptor,” *Oncogene*, vol. 32, no. 41, pp. 4903–4912, Oct. 2013, doi: 10.1038/onc.2012.516.
 - [96] FELASA working group on revision of guidelines for health monitoring of rodents and rabbits *et al.*, “FELASA recommendations for the health monitoring of mouse, rat, hamster, guinea pig and rabbit colonies in breeding and experimental units,” *Lab. Anim.*, vol. 48, no. 3, pp. 178–192, Jul. 2014, doi: 10.1177/0023677213516312.
 - [97] K. Yusa, L. Zhou, M. A. Li, A. Bradley, and N. L. Craig, “A hyperactive piggyBac transposase for mammalian applications,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 4, pp. 1531–1536, 2011, doi: 10.1073/pnas.1008322108.
 - [98] J. B. Bell, K. M. Podetz-Pedersen, E. L. Aronovich, L. R. Belur, R. S. McIvor, and P. B. Hackett, “Preferential delivery of the Sleeping Beauty transposon system to livers of mice by hydrodynamic injection,” *Nat. Protoc.*, vol. 2, no. 12, pp. 3153–3165, 2007, doi: 10.1038/nprot.2007.471.
 - [99] I. Portier *et al.*, “High and long-term von Willebrand factor expression after Sleeping Beauty transposon-mediated gene therapy in a mouse model of severe von Willebrand disease,” *J. Thromb. Haemost. JTH*, vol. 16, no. 3, pp. 592–604, Mar. 2018, doi: 10.1111/jth.13938.
 - [100] Y.-S. Huang *et al.*, “Pharmacological modulation of T cell immunity results in long-term remission of autoimmune arthritis,” *Proc. Natl. Acad. Sci.*, vol. 118, no. 19, p. e2100939118, May 2021, doi: 10.1073/pnas.2100939118.
 - [101] L. Bertrand, L. Dygert, and M. Toborek, “Antiretroviral Treatment with Efavirenz Disrupts the Blood-Brain Barrier Integrity and Increases Stroke Severity,” *Sci. Rep.*, vol. 6, no. 1, p. 39738, Dec. 2016, doi: 10.1038/srep39738.
 - [102] G. Nauwelaers *et al.*, “DNA adduct formation of 4-aminobiphenyl and heterocyclic aromatic amines in human hepatocytes,” *Chem. Res. Toxicol.*, vol. 24, no. 6, pp.

- 913–925, Jun. 2011, doi: 10.1021/tx200091y.
- [103] H. Ohgaki, H. Hasegawa, M. Suenaga, S. Sato, S. Takayama, and T. Sugimura, “Carcinogenicity in mice of a mutagenic compound, 2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline (MeIQx) from cooked foods,” *Carcinogenesis*, vol. 8, no. 5, pp. 665–668, May 1987, doi: 10.1093/carcin/8.5.665.
- [104] M. Keshavarzi, M. J. Khoshnoud, A. Ghaffarian Bahraman, and A. Mohammadi-Bardbori, “An Endogenous Ligand of Aryl Hydrocarbon Receptor 6-Formylindolo[3,2-b]Carbazole (FICZ) Is a Signaling Molecule in Neurogenesis of Adult Hippocampal Neurons,” *J. Mol. Neurosci. MN*, vol. 70, no. 5, pp. 806–817, May 2020, doi: 10.1007/s12031-020-01506-x.
- [105] W. Strober, “Trypan Blue Exclusion Test of Cell Viability,” *Curr. Protoc. Immunol.*, vol. 111, p. A3.B.1–A3.B.3, Nov. 2015, doi: 10.1002/0471142735.ima03bs111.
- [106] C. D’Hulst, I. Parvanova, D. Tomoiaga, M. L. Sapar, and P. Feinstein, “Fast quantitative real-time PCR-based screening for common chromosomal aneuploidies in mouse embryonic stem cells,” *Stem Cell Rep.*, vol. 1, no. 4, pp. 350–359, 2013, doi: 10.1016/j.stemcr.2013.08.003.
- [107] K. J. Livak and T. D. Schmittgen, “Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method,” *Methods San Diego Calif*, vol. 25, no. 4, pp. 402–408, 2001, doi: 10.1006/meth.2001.1262.
- [108] N. Rodić *et al.*, “Long interspersed element-1 protein expression is a hallmark of many human cancers,” *Am. J. Pathol.*, vol. 184, no. 5, pp. 1280–1286, 2014, doi: 10.1016/j.ajpath.2014.01.007.
- [109] L. Chen, J. E. Dahlstrom, A. Chandra, P. Board, and D. Rangasamy, “Prognostic value of LINE-1 retrotransposon expression and its subcellular localization in breast cancer,” *Breast Cancer Res. Treat.*, vol. 136, no. 1, pp. 129–142, 2012, doi: 10.1007/s10549-012-2246-7.
- [110] S. Sato *et al.*, “LINE-1 ORF1p as a candidate biomarker in high grade serous ovarian carcinoma,” *Sci. Rep.*, vol. 13, no. 1, 2023, doi: 10.1038/s41598-023-28840-5.
- [111] T. T. Doucet-O’Hare *et al.*, “LINE-1 expression and retrotransposition in Barrett’s esophagus and esophageal carcinoma,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 35, pp. E4894–E4900, 2015, doi: 10.1073/pnas.1502474112.
- [112] D. Ardeljan, M. S. Taylor, D. T. Ting, and K. H. Burns, “The human long interspersed element-1 retrotransposon: An emerging biomarker of Neoplasia,” *Clin. Chem.*, vol. 63, no. 4, pp. 816–822, 2017, doi: 10.1373/clinchem.2016.257444.
- [113] W. Remmele and H. E. Stegner, “[Recommendation for uniform definition of an immunoreactive score (IRS) for immunohistochemical estrogen receptor detection (ER-ICA) in breast cancer tissue],” *Pathol.*, vol. 8, no. 3, pp. 138–140, May 1987.
- [114] R. Klaes *et al.*, “Overexpression of p16INK4A as a specific marker for dysplastic and neoplastic epithelial cells of the cervix uteri,” *Int. J. Cancer*, vol. 92, no. 2, pp. 276–284, 2001, doi: 10.1002/ijc.1174.
- [115] E. J. Nam *et al.*, “Expression of the p16INK4a and Ki-67 in relation to the grade of cervical intraepithelial neoplasia and high-risk human papillomavirus infection,” *J. Gynecol. Oncol.*, vol. 19, no. 3, pp. 162–168, Sep. 2008, doi: 10.3802/jgo.2008.19.3.162.
- [116] G. Imre *et al.*, “Prolonged activity of the transposase helper may raise safety concerns during DNA transposon-based gene therapy,” *Mol. Ther. Methods Clin. Dev.*, vol. 29, pp. 145–159, Mar. 2023, doi: 10.1016/j.omtm.2023.03.003.
- [117] S. Lindstedt, E. Holme, E. A. Lock, O. Hjalmarson, and B. Strandvik, “Treatment of hereditary tyrosinaemia type I by inhibition of 4-hydroxyphenylpyruvate dioxygenase,” *Lancet Lond. Engl.*, vol. 340, no. 8823, pp. 813–817, Oct. 1992, doi:

- 10.1016/0140-6736(92)92685-9.
- [118] M. Lecocq, F. Andrianaivo, M.-T. Warnier, S. Wattiaux-De Coninck, R. Wattiaux, and M. Jadot, "Uptake by mouse liver and intracellular fate of plasmid DNA after a rapid tail vein injection of a small or a large volume," *J. Gene Med.*, vol. 5, no. 2, pp. 142–156, Feb. 2003, doi: 10.1002/jgm.328.
 - [119] G. Zhang *et al.*, "Hydroporation as the mechanism of hydrodynamic delivery," *Gene Ther.*, vol. 11, no. 8, pp. 675–682, Apr. 2004, doi: 10.1038/sj.gt.3302210.
 - [120] L. Mátés *et al.*, "Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates," *Nat. Genet.*, vol. 41, no. 6, pp. 753–761, Jun. 2009, doi: 10.1038/ng.343.
 - [121] Z. Cui, A. M. Geurts, G. Liu, C. D. Kaufman, and P. B. Hackett, "Structure-function analysis of the inverted terminal repeats of the sleeping beauty transposon," *J. Mol. Biol.*, vol. 318, no. 5, pp. 1221–1235, May 2002, doi: 10.1016/S0022-2836(02)00237-1.
 - [122] K. A. Skipper, P. R. Andersen, N. Sharma, and J. G. Mikkelsen, "DNA transposon-based gene vehicles - scenes from an evolutionary drive," *J. Biomed. Sci.*, vol. 20, no. 1, p. 92, 2013, doi: 10.1186/1423-0127-20-92.
 - [123] K. Overturf *et al.*, "Hepatocytes corrected by gene therapy are selected in vivo in a murine model of hereditary tyrosinaemia type I," *Nat. Genet.*, vol. 12, no. 3, pp. 266–273, Mar. 1996, doi: 10.1038/ng0396-266.
 - [124] A. L. Szymczak and D. A. A. Vignali, "Development of 2A peptide-based strategies in the design of multicistronic vectors," *Expert Opin. Biol. Ther.*, vol. 5, no. 5, pp. 627–638, May 2005, doi: 10.1517/14712598.5.5.627.
 - [125] K. E. Orii *et al.*, "Genes for the human mitochondrial trifunctional protein alpha- and beta-subunits are divergently transcribed from a common promoter region," *J. Biol. Chem.*, vol. 274, no. 12, pp. 8077–8084, Mar. 1999, doi: 10.1074/jbc.274.12.8077.
 - [126] H. I. Saba, "Decitabine in the treatment of myelodysplastic syndromes," *Ther. Clin. Risk Manag.*, vol. 3, no. 5, pp. 807–817, Oct. 2007.
 - [127] N. Okudaira *et al.*, "Induction of long interspersed nucleotide element-1 (L1) retrotransposition by 6-formylindolo[3,2-b]carbazole (FICZ), a tryptophan photoproduct," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 43, pp. 18487–18492, Oct. 2010, doi: 10.1073/pnas.1001252107.
 - [128] M. A. Schneider, A. A. Buzdin, A. Weber, P.-A. Clavien, and P. Borger, "Combination of Antiretroviral Drugs Zidovudine and Efavirenz Impairs Tumor Growths in a Mouse Model of Cancer," *Viruses*, vol. 13, no. 12, p. 2396, Dec. 2021, doi: 10.3390/v13122396.
 - [129] M. S. Taylor *et al.*, "Ultrasensitive Detection of Circulating LINE-1 ORF1p as a Specific Multicancer Biomarker," *Cancer Discov.*, vol. 13, no. 12, pp. 2532–2547, Dec. 2023, doi: 10.1158/2159-8290.CD-23-0313.
 - [130] M. Rajurkar *et al.*, "Reverse Transcriptase Inhibition Disrupts Repeat Element Life Cycle in Colorectal Cancer," *Cancer Discov.*, vol. 12, no. 6, pp. 1462–1481, Jun. 2022, doi: 10.1158/2159-8290.CD-21-1117.
 - [131] J. V. Moran, S. E. Holmes, T. P. Naas, R. J. DeBerardinis, J. D. Boeke, and H. H. Kazazian, "High frequency retrotransposition in cultured mammalian cells," *Cell*, vol. 87, no. 5, pp. 917–927, 1996, doi: 10.1016/S0092-8674(00)81998-4.
 - [132] D. M. Maron and B. N. Ames, "Revised methods for the Salmonella mutagenicity test," *Mutat. Res.*, vol. 113, no. 3–4, pp. 173–215, May 1983, doi: 10.1016/0165-1161(83)90010-9.
 - [133] OECD, *Test No. 489: In Vivo Mammalian Alkaline Comet Assay*. Paris: Organisation for Economic Co-operation and Development, 2016. Accessed: Sep. 15, 2024. [Online]. Available:

- https://www.oecd-ilibrary.org/environment/test-no-489-in-vivo-mammalian-alkaline-comet-assay_9789264264885-en
- [134] J. Menz *et al.*, “Genotoxicity assessment: opportunities, challenges and perspectives for quantitative evaluations of dose-response data,” *Arch. Toxicol.*, vol. 97, no. 9, pp. 2303–2328, Sep. 2023, doi: 10.1007/s00204-023-03553-w.
 - [135] C. Philippe *et al.*, “Activation of individual L1 retrotransposon instances is restricted to cell-type dependent permissive loci,” *eLife*, vol. 5, p. e13926, doi: 10.7554/eLife.13926.
 - [136] E. C. Scott, E. J. Gardner, A. Masood, N. T. Chuang, P. M. Vertino, and S. E. Devine, “A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer,” *Genome Res.*, vol. 26, no. 6, pp. 745–755, Jun. 2016, doi: 10.1101/gr.201814.115.
 - [137] C. M. Ade, G. Servant, M. E. Morales, and A. M. Roy-Engel, “Environment, Cellular Signaling, and L1 Activity,” in *Human Retrotransposons in Health and Disease*, G. Cristofari, Ed., Cham: Springer International Publishing, 2017, pp. 157–194. doi: 10.1007/978-3-319-48344-3_7.
 - [138] H. Herweijer, G. Zhang, V. M. Subbotin, V. Budker, P. Williams, and J. A. Wolff, “Time course of gene expression after plasmid DNA gene transfer to the liver,” *J. Gene Med.*, vol. 3, no. 3, pp. 280–291, 2001, doi: 10.1002/jgm.178.
 - [139] B. L. Hodges, K. M. Taylor, M. F. Joseph, S. A. Bourgeois, and R. K. Scheule, “Long-term transgene expression from plasmid DNA gene therapy vectors is negatively affected by CpG dinucleotides,” *Mol. Ther. J. Am. Soc. Gene Ther.*, vol. 10, no. 2, pp. 269–278, 2004, doi: 10.1016/j.ymthe.2004.04.018.
 - [140] P. K. Held, E. C. Olivares, C. P. Aguilar, M. Finegold, M. P. Calos, and M. Grompe, “In vivo correction of murine hereditary tyrosinemia type I by phiC31 integrase-mediated gene delivery,” *Mol. Ther. J. Am. Soc. Gene Ther.*, vol. 11, no. 3, pp. 399–408, Mar. 2005, doi: 10.1016/j.ymthe.2004.11.001.