KU LEUVEN

FACULTY OF MEDICINE
BIOMEDICAL SCIENCES

Identification of LEDGF/p75 as novel host factor of L1 retrotransposition

The L1 retrotransposition assay: a tool to study the interactome of retrotransposons

Thesis submitted in partial fulfilment of the requirements for the degree of Master of Biomedical Sciences by

Saskia LESIRE

Supervisor: Prof. dr. Zeger DEBYSER
Second supervisor: dr. Frauke CHRIST
Instructor: Paulien VAN DE VELDE

Department of Pharmaceutical and Pharmacological Sciences
Laboratory of Molecular Virology and Gene Therapy

Leuven, 2019-2020

This Master's Thesis is an exam document. Possibly assessed errors were not corrected after the defense. In publications, references to this thesis may only be made with written permission of the supervisor(s) mentioned on the title page.

The preparation and completion of this master's thesis partly took place during the period when COVID-19 measures were in effect in Belgium (started on 13 March 2020). Within the framework of the measures taken by KU Leuven, it was not possible to complete the data collection for this master's thesis and the processing therefore had to be based on limited results. This aspect needs to be taken into account when consulting this document.

KU LEUVEN

FACULTY OF MEDICINE
BIOMEDICAL SCIENCES

Identification of LEDGF/p75 as novel host factor of L1 retrotransposition

The L1 retrotransposition assay: a tool to study the interactome of retrotransposons

Thesis submitted in partial fulfilment of the requirements for the degree of Master of Biomedical Sciences by

Saskia LESIRE

Supervisor: Prof. dr. Zeger DEBYSER
Second supervisor: dr. Frauke CHRIST
Instructor: Paulien VAN DE VELDE

Department of Pharmaceutical and Pharmacological Sciences
Laboratory of Molecular Virology and Gene Therapy

Leuven, 2019-2020

Na 5 jaar studeren had ik mij het einde toch wel anders voorgesteld. Afstuderen in tijden van een pandemie, het zal je maar overkomen. Het neemt niet weg dat de voorbije 5 jaren fantastisch waren.

Dankjewel Prof. Dr. Zeger Debyser, niet alleen voor alle begeleiding gedurende dit jaar, maar ook om mij als master studente te verwelkomen in je lab. Je enthousiasme voor onderzoek is aanstekelijk en zorgt ervoor dat ik telkens het beste van mezelf wil geven. Op één jaar tijd heb ik zo veel geleerd en zo veel leuke collega's leren kennen dat ik mezelf alleen maar gelukkig kan prijzen dat ik hier terecht ben gekomen. Thank you Dr. Frauke Christ for all your support and input on this project during the meetings. Bedankt Paulien voor de fantastische begeleiding, op korte tijd ben je erin geslaagd om mij talloze vaardigheden aan te leren. Ik heb het niet genoeg gezegd, maar bedankt voor al je geduld toen ik weer eens fout pipetteerde. Bedankt Babs voor alle hulp en de toffe babbels waardoor ik me meteen welkom voelde. Dankjewel Yannick om mij vanaf dag 1 te betrekken in je onderzoek en me te begeleiden bij het schrijven van deze thesis. Thank you Rodrigo for helping me with the practical and theoretical work for this project, and for 'babysitting' me when no one else was around. En bedankt Dr. Chris Van den Haute om telkens tijd te maken om me verder te helpen.

Nooit had ik gedacht dat ik tijdens mijn studies zo veel geweldige mensen zou leren kennen. Ik had het geluk om in een fantastisch jaar terecht te komen met medestudenten die zorgden dat ik gemotiveerd bleef doorheen deze zware studies. Ik leerde hier mijn 'GHB-XL familie' kennen die zonder twijfel vrienden voor het leven zullen blijven. En nog een shoutout naar de 'BMW Girls', een groep buitengewoon straffe madammen met wie ik tijdens deze 5 jaar de heel toffe en soms ook minder toffe momenten deelde.

Dankjewel Fabian om mij op het einde van een lange dag een dikke knuffel te geven en mij van een citroentaartje te voorzien op de nodige momenten. En zus, bedankt om elke week te vragen hoe het met de thesis stond, ook al heb je nog steeds geen flauw idee waar dit allemaal over gaat.

Et merci papa, pour tout!

Table of contents

List of abbreviations

AD	Alzheimer's disease
AGS	Aicardi-Goutières syndrome
ALS	Amyotrophic lateral sclerosis
AmpR	Ampicillin resistance
ASD	Autism spectrum disorders
AT	Ataxia-telangiectasia
ATM	Ataxia telangiectasia mutated
BsdR	Blasticidin resistance
CC	Coiled coil
CMV	Cytomegalovirus
cPPT	Central polypurine tracks
CRD	Cysteine-rich domain
CTD	C-terminal domain
DMEM	Dulbecco's modified Eagle's medium
DSB	Double stranded break
EBNA-1	Epstein-Barr nuclear antigen 1
EGFP	Enhanced green fluorescent protein
EN	Endonuclease domain
ESC	Embryonic stem cell
FCS	Fetal bovine calf serum
Fluc	Firefly luciferase
FTD	Frontotemporal dementia
FTLD	Frontotemporal lobe degeneration
Gag	Group-specific antigen

GAR	Purine-rich region
gDNA	Genomic DNA
HDAC1	Histone deacetylase 1
HIV-1	Human immunodeficiency virus type 1
HR	Homologous recombination
ID	Interdomain
IFN	Interferon
iPSC	Induced pluripotent stem cell
IRES	Internal ribosome entry site
JM111	99-JM111-EGFP-Puro control plasmid
L1	Long interspersed nuclear element
$L1_{RP}$	99-L1RPS-EGFP-Puro plasmid
LEDGF/p75	Lens epithelium-derived growth factor/p75
LINE-1	Long interspersed nuclear element
LTR	Long terminal repeat
MBD	Methyl binding domain
MeCP2	Methyl CpG-binding protein 2
miR	Micro RNA
miRNA	Micro RNA
NeoR	Neomycin resistance
NPC	Neuronal progenitor cell
NTD	N-terminal domain
ORF	Open reading frame
ORF1p	ORF1 protein
ORF2p	ORF2 protein

Ori	*E. coli* origin of replication
OriP	Epstein-Barr virus origin of replication
pCMV	Early cytomegalovirus promoter
PEI	Polyethylenimine
piRISC	piRNA-induced silencing complex
piRNA	PIWI-interacting RNA
Pol	Polymerase
PolyA	Polyadenylation
PuroR	Puromycin resistance
R	Repeat element
RISC	RNA-induced silencing complex
RNAi	RNA interference
RNP	Ribonucleoprotein
RRE	Rev response element
RRM	RNA recognition motif
RT	Reverse transcriptase domain
RT-qPCR	Real-time quantitative polymerase chain reaction
RTi	Reverse transcriptase inhibitors
RTT	Rett syndrome
SA	Splice acceptor
Scr	Scrambled
SD	Standard deviation
SFFV	Spleen focus forming virus
shRNA	Short hairpin RNA
siRNA	Small interfering RNA

SIV	Simian immunodeficiency virus
TE	Transposable element
TNPO1	Transportin-1
TPRT	Target-primed reverse transcription
TRD	Transcriptional repression domain
TRN-SR2	Transportin SR-2
U3	3’ unique element
U5	5’ unique element
UTR	Untranslated region
VSV-G	Vesicular stomatitis virus glycoprotein
WPRE	Woodchuck hepatitis virus posttranscriptional regulatory element
ZeoR	Zeocin resistance

Abstract

Retrotransposons are a subgroup of transposable elements that expand the genome by a copy-and-paste mechanism through an RNA intermediate. Long interspersed nuclear elements (LINE-1 or L1) are the only autonomous retrotransposons in the human genome still active today. L1 retrotransposons resemble retroviruses with respect to their structure and function in search of their own survival by (re)-integration. Retrotransposons contribute to somatic mosaicism, are implicated in memory formation and act as potent regulatory elements. However, their biological significance remains poorly understood. Our cells have evolved various defense mechanisms to restrict aberrant L1 activity. L1 elements have evolved to evade these mechanisms as a result of the dynamic interaction between L1 and their host. Growing evidence suggests a role for L1 in healthy somatic tissue as well as in various diseases. L1 retrotransposons have been associated with a wide range of human diseases, such as Rett syndrome, schizophrenia and cancer, making them potential pathogens and targets for new therapeutics. Therefore, a thorough understanding of the L1 interactome is required. In this study we optimized the L1 retrotransposition assay in order to study host and restriction factors of L1 retrotransposition. The $L1_{RP}$-EGFP reporter plasmid was used to perform the L1 retrotransposition assay in HEK293T and HeLa P4 cells. We confirmed that HEK293T cells are capable of supporting high frequency L1 retrotransposition from an episomal $L1_{RP}$-EGFP plasmid. Up to 84% EGFP-positive cells were observed using flow cytometry for the EGFP readout, reflecting high retrotransposition rates in HEK293T cells. The assay was validated using MeCP2 depletion and an expected increase in retrotransposition rate was observed. HeLa P4 cells did not seem to support L1 retrotransposition from the $L1_{RP}$-EGFP plasmid based on our results. Only very low retrotransposition rates were observed that lack biological relevance. The second objective was to study the role of LEDGF/p75 in L1 retrotransposition. Therefore, we used the optimized conditions of the L1 retrotransposition assay to study L1 retrotransposition in a LEDGF/p75 depletion cell line. A decrease in retrotransposition rate was observed upon depletion of LEDGF/p75. These preliminary results suggest a role for LEDGF/p75 as host factor of L1 retrotransposition. Future studies will have to investigate whether LEDGF/p75 exerts a direct or indirect effect on L1 retrotransposition.

1 Introductory overview of the literature

1.1 Introduction

The evolution of the human genome is a dynamic process. Copies of DNA sequences have accumulated over time and occupy a large part of our genome. These copies or repeats are called transposons, also known as 'jumping genes'. In 1948, geneticist Barbara McClintock was the first to describe these mobile DNA elements in the context of maize kernels. She discovered transposable elements (TEs) that were able to move and reinsert themselves into the genome (1). Her work was initially received with a lot of skepticism since at the time the genome was believed to be static. Nevertheless, research performed in the years to follow underlined the importance of mobile DNA in the evolution of the mammalian genome and the emergence of new human-specific genes (2). Approximately 45 % of the human genome is originally derived from TEs, yet this number is presumed to be an underestimation since over time diverged TEs may have become unrecognizable (2). For a long time, TEs have been considered as 'junk DNA' since they did not have any apparent cellular function at first sight. However, over the past few decades it has become clear that TEs had and still have a fundamental role in genome evolution and the origin of genetic disorders (3).

TEs can be classified as DNA transposons or retrotransposons. DNA transposition occurs via a cut-and-paste mechanism, in contrast to retrotransposition that occurs via a copy-and-paste mechanism through an RNA intermediate (4). Retrotransposons can be further categorized in long terminal repeat (LTR) or non-LTR retrotransposons. Long interspersed nuclear elements 1 (L1 or LINE-1) are part of the non-LTR family of retrotransposons. Full length L1s are the only autonomous elements in the mammalian genome (5, 6). They contain all the information required to replicate and reinsert themselves into the host genome (7). Initial sequencing of the human genome revealed that L1 sequences occupy 17 % of the genomic DNA (8). Most retrotransposition events result in the integration of a 5' truncated L1 copy, generating a defective L1 that is no longer able to jump and that remains inactive in the genome (4). However, 80 to 100 L1 elements in the human genome are estimated to be retrotransposition competent, of which a small part is even highly active (9). These 'hot' L1s play an important role in further genome expansion and genome plasticity (9).

1.2 Physiological role of L1 retrotransposons

L1 activity is tissue specific and occurs at different stages during development and adult life. Starting from the idea that L1 transposons are 'molecular parasites' continuously looking for their own survival and re-integration, their purpose would be to be preserved and expanded in genomes over generations (10). This would only make sense if new L1 insertions could be transmitted to a new generation. Therefore, L1 retrotransposition can only exert an effect on genome evolution if it is active in germline cells or in embryonic stem cells (ESC). Only then new genomic integrations can be passed on to the new generation (10). On the other hand, TEs are involved in the emergence of new human specific genes during evolution (11). The majority of pseudogenes in the human genome is derived from L1 retrotransposons (12). These are called processed pseudogenes because they lack intronic regions and are a result of *trans*-mobilization of cellular mRNAs. A large number of processed pseudogenes are derived from mRNAs that are highly expressed in ESCs. In support of this evolutionary hypothesis, activity of L1 retrotransposition has been confirmed during gametogenesis and embryogenesis in Mammalia, including in human ESC (13-19). During embryogenesis, a wave of hypomethylation occurs that derepresses the L1 promoter, resulting in elevated L1 mRNA levels (20). Additionally, human ESC have a more open chromatin structure, facilitating retrotransposition. It remains unclear whether L1 has a functional impact during embryogenesis and other physiological processes or whether L1 retrotransposition is a type of genomic 'noise'.

ESCs give rise to different cell types. When a retrotransposition event occurs before the ESC becomes a distinct lineage, the L1 insertion will be present in all the cells of the individual (4). If retrotransposition occurs after differentiation, individuals may contain a heterogenous cell population with respect to the genetic content. L1 retrotransposition is one of the mechanisms known to create somatic mosaicism (20). Somatic mosaicism is a term used for multiple cell populations that are genetically different within an individual. During development or in adult life genetic alterations can occur that accumulate in a subset of somatic cells. This results in genetic diversity and phenotypic heterogeneity within one tissue of an individual (4). The number of somatic mutations accumulate over time but are restricted to the lifetime of an individual since, in contrast to germline cells, L1-mediated insertions in somatic cells cannot be passed on to the progeny.

L1 retrotransposition is not solely restricted to early stages of development. L1 activity has been described in neuronal progenitor cells (NPC) during adult neurogenesis and in mature non-dividing neurons of the human brain (21, 22). There is no clear explanation that supports the role of L1 in the human brain. However, L1 retrotransposition in NPCs was found to be the mediator of neuronal somatic mosaicism (23). This effect is clearly seen in the adult hippocampus where L1 mobility in progenitor cells leads to neuronal mosaicism among hippocampal neurons (24). These findings give rise to the question how somatic mosaicism in the brain can impact neurobiological function. The genomic diversity between neurons can give rise to unique transcriptomes that may ultimately influence neuronal phenotype and function. L1 retrotransposons can thus be seen as regulatory elements. Various studies have shown somatic L1 insertions in protein-coding genes and in regions implicated in gene regulation (23-25). When these insertions occur in or near genes expressed in neurons they may influence neuronal circuits and lead to altered cognition and behavior (25). This has been proven in the hippocampus where L1-mediated mosaicism was shown to be involved in memory formation (26). Interestingly, environmental factors such as voluntary exercise, stress and alcohol consumption can influence L1 activity in the adult brain (27-30). Taken together, these studies suggest that retrotransposition in the human brain occur much more often than initially anticipated. Why some neuronal cell types appear to accommodate higher levels of L1 retrotransposition compared to other somatic cell types remains a mystery. Additionally, it remains unclear whether all neuronal cell types are capable of supporting L1 retrotransposition. Further research needs to be conducted to fully understand the effect of L1 retrotransposition in the healthy or diseased human brain.

1.3 L1 structure and biology

An active L1 sequence is ~6 kb in length and consists of a promoter in the 5' untranslated region (UTR), two non-overlapping open reading frames (ORF), ORF1 encoding ORF1 protein (ORF1p) and ORF2 encoding ORF2 protein (ORF2p), and a 3' UTR containing a polyadenylation (polyA) signal (Figure 1) (6). The 5' UTR is CpG-rich and contains a sense and an antisense promoter that can be methylated to regulate L1 expression (31). ORF1p is an RNA-binding protein that has a chaperone activity to stabilize the L1 mRNA during a retrotransposition event (32). ORF2p consists of three domains: the zinc finger domain binds the 3' end of L1 mRNA and brings it into proximity of the genomic integration site, the endonuclease domain nicks the AT-consensus sequence, and the reverse transcriptase domain

transcribes the RNA into a cDNA sequence (33, 34). An additional primate-specific ORF0 has been discovered in the 5' UTR that is transcribed from the antisense promoter and that is able to generate fusion proteins with downstream cellular genes (35). ORF0 also has the ability to enhance L1 mobility, yet it appears that ORF0 is not essential for L1 retrotransposition (35). It has been suggested that ORF0 is related to L1 self-regulation by expressing antisense RNAs triggering an RNA interference (RNAi) response (15, 35, 36).

*Figure 1: **Structure of human L1**. A full-length L1 has a size of ~6 kb. The 5' UTR contains a sense and antisense promoter. ORF0 is located in the 5' UTR and is transcribed by the antisense promoter. ORF1 consists of a coiled coil (CC) domain, an RNA recognition motif (RRM) and a C-terminal domain (CTD). ORF2 consists of an endonuclease domain (EN), a reverse transcriptase domain (RT) and cysteine-rich domains (CRD). ORF1 and ORF2 are separated by a noncoding spacer region. The 3' UTR contains a polyA signal.*

A model for L1 retrotransposition is explained in Figure 2. L1 is transcribed by RNA polymerase II and exported to the cytoplasm where ORF1 and ORF2 are translated by ribosomes (37, 38). The ORF1p and ORF2p proteins associate with L1 mRNA to form ribonucleoprotein (RNP) particles. They show a *cis*-preference, meaning that they prefer to mobilize the L1 transcript from which they are encoded (39). *Trans*-mobilization of cellular mRNA happens in very low levels and gives rise to processed pseudogenes (40). The cellular cofactor PABPC1 is involved in the assembly and trafficking of L1-RNPs (41). After import of L1-RNPs into the nucleus the L1 mRNA is reverse transcribed and integrated into the genome by a mechanism called target-primed reverse transcription (TPRT) (42, 43). The endonuclease domain of ORF2p cleaves the antisense strand of genomic DNA in the AT-consensus sequence and uses the free 3'OH to synthesize cDNA. This way the reverse transcriptase of ORF2p generates an RNA:cDNA hybrid that is linked to the genomic DNA. The mechanism by which the second DNA strand is generated as well as the mechanism by which the 5' end is integrated is not well understood. The entire L1 mRNA needs to be reverse transcribed in order to make another fully functional L1 (5). Most L1 sequences in the genome are 5' truncated, making them unable to start a new cycle of retrotransposition (42).

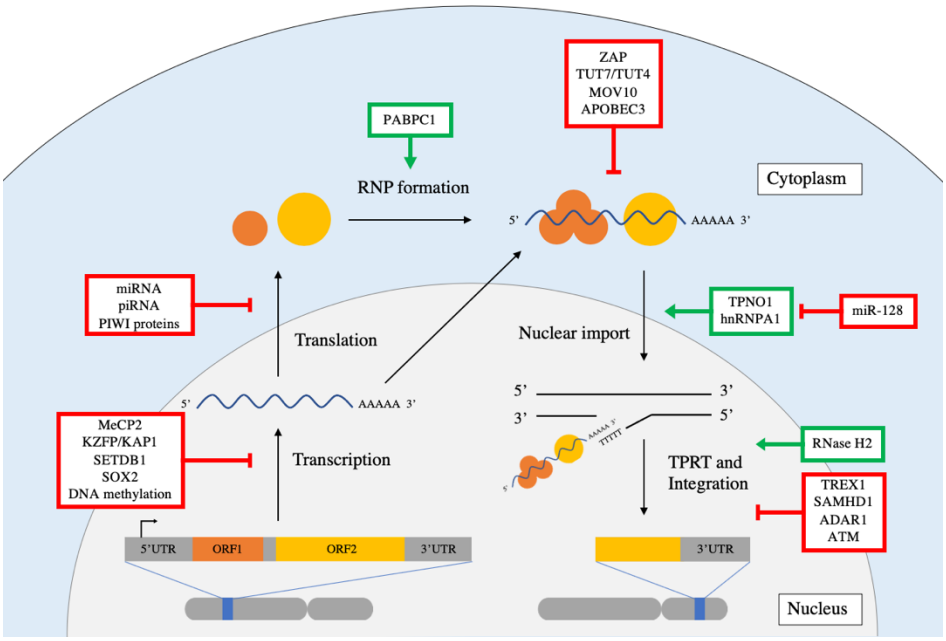


Figure 2: ***L1 retrotransposition cycle with cellular host and restriction factors.*** *L1 is transcribed by RNA polymerase II from its endogenous promoter located in the 5' UTR. ORF1 and ORF2 are translated in the cytosol by ribosomes. Ribonucleoprotein (RNP) particles are formed in the cytoplasm and are composed of ORF1p in trimer, ORF2 and L1 mRNA. L1-RNPs are imported into the nucleus where L1 mRNA is reverse transcribed and integrated in the genome by target-primed reverse transcription (TPRT). Most newly integrated L1s are retrotransposition defective due to 5' truncation. Host factors needed for L1 retrotransposition are indicated by green boxes. Red boxes indicate restriction factors that regulate and restrict L1 retrotransposition.*

1.4 Host and restriction factors of L1 retrotransposition

Host cells have developed various defense mechanisms to repress aberrant activity of TEs (Figure 2). Many known regulators of TEs show similarities with responses to retroviral pathogens. Transposons, on the other hand, have evolved to limit genome damage, resulting in a dynamic interaction between TEs and their host (44, 45). Here I provide an overview of host and restriction factors of L1 retrotransposition following the retrotransposition cycle.

1.4.1 Transcriptional regulation

Epigenetic modification is an important mechanism that restricts L1 expression (6). Transcriptional repression of TEs can be induced by DNA methylation or by histone modifications (20, 46). Methyl CpG-binding protein 2 (MeCP2) is a nuclear protein well known to regulate transcription of L1 through binding of the methylated promoter in the 5' UTR (Figure 3) (47-49). In literature the involvement of MeCP2 in gene silencing is reported as a result of the interaction with transcriptional co-repressors (50). The membrane binding domain (MBD) of MeCP2 binds the methylated CpG islands in the L1 promoter and the transcriptional repression domain (TRD) is required for interaction with co-repressors and transcriptional

repression (49). ChIP-sequencing analysis in mouse neurons revealed that MeCP2 is widely distributed throughout the neuronal genome and that its binding follows the methyl-CpG distribution suggesting it can exert its repressive function on the wide range of L1 sequences throughout the genome (51).

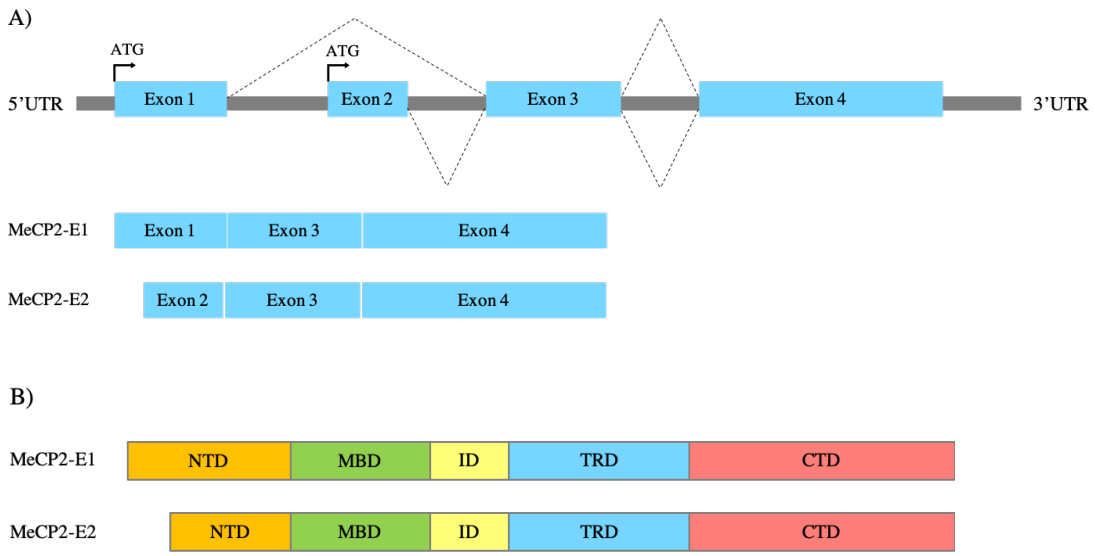


Figure 3: ***Protein domains of MeCP2-E1 and MeCP2-E2.*** *A) The MeCP2 gene contains 4 exons that code for the 2 protein isoforms MeCP2-E1 and MeCP2-E2 of which MeCP2-E1 is the most abundant form (52). They are alternatively spliced and differ at the N-terminal region. MeCP2-E1 is encoded by exons 1, 3 and 4, MeCP2-E2 is encoded by exons 2, 3 and 4. B) MeCP2 consists of an N-terminal domain (NTD) at the 5' end and a C-terminal domain (CTD) at the 3' end. The NTD is followed by the methyl binding domain (MBD) that binds to a single methyl-CpG pair (53). The interdomain (ID) connects the MBD to the transcriptional repression domain (TRD). The TRD mediates the link between MeCP2 and its co-repressors (50).*

Various other transcription factors can regulate L1 transcription in a context specific manner (54). The KZFP/KAP1 complex is known to control L1 expression in human and mouse ESCs by recruiting SETDB1 histone methyltransferase leading to the formation of heterochromatin and transcriptional repression (55). Throughout further development permanent silencing can be obtained by methylation via DNA methyltransferases (4). During neurogenesis L1 transcription is repressed by a complex formed between SOX2 and histone deacetylase 1 (HDAC1). During the transformation from neuronal stem cells to NPCs derepression of the SOX2/HDAC1 complex occurs followed by the activation of the Wnt-signaling pathway that induces L1 expression (56).

1.4.2 Post-transcriptional regulation

The cell has various post-transcriptional responses to restrict aberrant L1 activity (57). After transcription, alternative splicing and pre-mature polyadenylation of the L1 transcript contribute to pre-mRNA regulation of L1 by removing parts of the ORFs or 5' UTR (58, 59). Small interfering RNAs (siRNA) are also important post-transcriptional regulators that can bind complementary L1 mRNAs and target them towards degradation by endonucleolytic cleavage (60). Mainly PIWI-interacting RNAs (piRNA) are known to repress TEs in germline cells (61). piRNAs associate with PIWI proteins to form piRNA-induced silencing complexes (piRISC) (62). piRISCs can regulate TEs on transcriptional level by inducing chromatin modifications as well as on post-transcriptional level for complementary TE sequences by way of RNAi. Analogously, TE-derived miRNAs can regulate TEs by way of RNAi (63).

After translation of L1 mRNA ORF1p will form a trimeric structure that binds L1 mRNA (64). ORF1p and ORF2p do not directly interact but appear to be associated through their interaction with L1 mRNA (65). The cell has several defense mechanisms to prevent the formation of the L1-RNP or to destabilize it. The zinc-finger protease ZAP affects post-transcriptional processing L1 mRNAs (66, 67). ZAP prevents accumulation of ORF1p and ORF2p by interacting with L1 mRNA in the cytoplasm and thereby it restricts L1 retrotransposition.

MOV10 is an ATP-dependent RNA helicase that is known to control retroviral infections but also strongly inhibits endogenous retroelements (68-70). It displays broad RNA-binding properties and 5' to 3' dsRNA unwinding activity (71). Although many mechanisms have been proposed by which MOV10 could restrict L1 retrotransposition, the exact mechanism(s) remain(s) unclear. MOV10 was found to associate with the RNA-induced silencing complex (RISC) through binding of the key component AGO2 (72). This led to the hypothesis that MOV10 could restrict L1 retrotransposition by mediating the interaction between L1 mRNA and AGO2 of the RISC (70). However, depletion of AGO2 did not affect the ability of MOV10 to restrict L1 retrotransposition, suggesting that the residual amount of AGO2 would be sufficient to mediate restriction of L1 or that MOV10 regulation of L1 occurs independently of the RISC pathway (70). The same study showed the interaction between MOV10 and ORF1 of the L1 mRNA, consistent with a previous study that found an interaction between MOV10 and L1-RNPs (69). Likewise, AGO2 was found to colocalize with ORF1 of the L1 mRNA in cytoplasmic granules (73). This led to the hypothesis that MOV10 sequesters L1 mRNA leading

to the recruitment of L1-RNPs to stress granules where siRNA pathways induce silencing or degradation of the L1-RNP (69, 71).

Another major mechanism by which MOV10 regulates L1 retrotransposition is its functional co-operation with the uridyltransferases TUT4 and TUT7 (74). MOV10 has shown to counteract the chaperone activity of ORF1p on L1 mRNA enabling access of the uridyltransferases to the L1 mRNA. It was proposed that uridylation by TUT4 in the cytoplasm destabilizes L1 mRNA whereas uridylation of the L1 mRNA in the cytoplasm by TUT7 inhibits reverse transcription activity of ORF2p after re-entry into the nucleus (74). Additionally, MOV10L1, a homologue of MOV10 has been shown to control the expression of retrotransposons in germline cells through association with PIWI-proteins (75, 76).

Several members of the APOBEC3 cytidine deaminase family are known to repress L1 retrotransposition but they differ in specificity (77-82). Certain APOBEC proteins are also potent restriction factors for retroviral reverse transcription by editing single stranded DNA. It appears that inhibition of L1 retrotransposition by APOBEC proteins is mediated by a mechanism independent of cytidine deamination (79, 83). Other mechanisms were proposed including sequestration of L1-RNPs in cytoplasmic complexes and their subsequent targeting to stress granules and P-bodies followed by degradation via RNAi (60). This hypothesis is supported by the observation that APOBEC3G colocalizes with RNPs in stress granules and P-bodies (84). APOBEC3G was found to competitively inhibit the interaction between MOV10 and AGO2 of the RISC suggesting that APOBEC3G disturbs the MOV10-mediated RISC assembly subsequently counteracting gene silencing (85). The impact on L1 retrotransposition of this mechanism remains to be elucidated.

1.4.3 Nuclear import of L1-RNPs

During the life cycle of L1 retrotransposition, the L1-RNP needs to be efficiently imported back into the nucleus. Nuclear breakdown during mitosis appears to be an important entry mechanism in analogy with some exogenous retroviruses (86). A study by Mita *et al*. confirmed that nuclear entry of the L1-RNP mainly occurs during the S-phase of the cell cycle in rapidly dividing cancer cells (87). In contrast, it was observed that retrotransposition also occurs, albeit at lower rates, in mature nondividing human neurons suggesting the presence of other mechanisms of nuclear import independent of the cell cycle (22). The nuclear import factor Transportin-1 (TNPO1) appears to be yet another common host factor in the replication of L1

retroviruses, more specifically human immunodeficiency virus type 1 (HIV-1) (88). TNPO1 was shown to import L1-RNPs into the nucleus, thereby catalyzing an essential step in the L1 retrotransposition cycle. The restriction factor miR-128 has three ways to repress L1 retrotransposition. It depletes TNPO1 mRNA resulting in decreased levels of TNPO1 in the nuclear membrane, thereby repressing nuclear import of L1-RNPs (88). Secondly, it decreases levels of hnRNPA1, an RNA-binding protein that mediates shuttling through the nuclear pore complex by interacting with TNPO1 (89). hnRNPA1 has been described as part of the L1-RNP complex by interacting with ORF1p through an RNA-bridge (90). Finally, miR-128 also represses L1 by directly binding to L1 mRNA and targeting it towards degradation via the RISC (91). miRNAs regulating multiple targets in the same cellular pathway is not a novel concept (89). miR-128 also appears to have binding sites in all members of the transportin family. Interestingly, Transportin-SR2 (TRN-SR2) is known as an essential nuclear import factor of HIV-1, but its role in nuclear import of L1-RNPs has not been investigated (92). Additionally, TNPO1 has been reported to act as a second import factor of HIV-1 making the analogy between nuclear import of HIV-1 and L1 even more interesting to explore (93).

1.4.4 Regulation of L1 TPRT

The majority of novel L1 insertions is 5' truncated resulting in a defective L1 sequence that is unable to undergo a novel cycle of retrotransposition (43). 5' truncation may be explained by the inability of reverse transcriptase to fully copy the L1 mRNA before the complex dissociates or by host defense mechanisms that occur during TPRT (43). L1 insertions in the genome are also often found inverted. This is explained by a mechanism called twin priming that always includes inversion and 5' truncation of L1 (94). Twin priming limits the insertion of full-length L1 sequences thereby inhibiting the replication of retrotransposition competent L1s.

Various host and restriction factors of L1 retrotransposition are known to act during TPRT in order to, respectively, enable or prevent the integration of new L1 sequences. One of the host factors that is required for efficient integration of L1 is RNase H2 (95, 96). RNase H2 is a nuclear trimeric enzyme that is responsible for RNA:DNA hybrid degradation and plays an essential role in the removal of accidentally incorporated ribonucleotides in the genomic DNA (97). A model was proposed in which RNase H2 degrades the L1 mRNA in the RNA:cDNA hybrid generated during TPRT. RNase H2 only appears to be an essential host factor for endogenous retroelements that do not have their own RNase H2 domain, these include L1

retrotransposons, in contrast to HIV-1 where reverse transcriptase also has an RNase H2 activity (98). This hypothesis was contradicted by a study that reported that RNase H2 and nuclear MOV10 interact in an RNA-dependent manner to repress L1 retrotransposition by inhibiting the formation of the RNA:cDNA hybrid during TPRT (99, 100). A possible explanation for this apparent paradox is that RNase H2 as a host factor enables L1 retrotransposition but that it cannot perform its function when it is associated with MOV10.

Various restriction factors of L1 retrotransposition are located in the nucleus. SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase and a ribonuclease that is predominantly expressed in the nucleus (101). It is a known inhibitor of retroviruses and endogenous retroelements (102, 103). SAMHD1 is thought to inhibit reverse transcription of retroviruses by depleting the dNTP pool, thereby inhibiting viral replication (104, 105). A study observed that the dNTP hydrolase activity of SAMHD1 is also necessary to restrict L1 retrotransposition (106). Additionally, a direct interaction between SAMHD1 and ORF2p was observed that is regulated by phosphorylation of SAMHD1. The authors suggest a model in which unphosphorylated SAMHD1 binds ORF2p in the nucleus and inhibits reverse transcription by locally depleting the dNTP pool through its enzymatic dNTP hydrolase activity. This was supported by an earlier study suggesting that SAMHD1 is required for the inhibition of ORF2p reverse transcription in the nucleus (102).

Another restriction factor of L1 retrotransposition in the nucleus is TREX1 (107). TREX1 is an abundant 3' to 5' DNA exonuclease that is ubiquitously expressed in the cell (108). TREX1 has antiviral DNase activities by targeting reverse transcribed viral cDNA. It thereby prevents the accumulation of viral cDNA in the cytosol and the innate immune responses to DNA species in the cytoplasm (60). TREX1 has also shown to metabolize reverse transcribed L1 cDNA (107). Since TREX1 is ubiquitously expressed in the cell, it is possible that it acts during TPRT in the nucleus, although it also opens the possibility for another mechanism of reverse transcription in the cytoplasm (109).

These studies prove the presence of a complex network of host factors that regulates both retroviruses and endogenous retroelements. Various known host and restriction factors show an essential function in the regulation of retroviral activity notwithstanding that this is not always the same function that is observed in the regulation of endogenous retroelements. For some cellular factors it is even unclear whether it originally defended the host against retroviruses or

whether it defended the host against endogenous retroelements (110). Nevertheless, the majority of host control mechanisms remain to be discovered.

1.5 L1 insertions as a cause of human disease

Kazazian and colleagues were the first to report that L1 retrotransposition can cause insertional mutagenesis leading to disease (111). In 1988 they reported a patient with hemophilia A with a new exonic L1 insertion in the factor VIII gene. Since then, more than 100 heritable disease cases were reported to be caused by L1 insertion such as Duchenne muscular dystrophy, cystic fibrosis and breast cancer (6, 112).

L1 insertions can affect the genome in a multitude of ways (43). Insertions into exons can introduce nonsense codons or cause exon skipping. This will have a direct impact on protein structure. Insertions into introns can also be mutagenic by generating new splice sites causing exon skipping or inserting polyA signals causing premature termination of transcription (59). Intronic insertions can also decrease stability of pre-mRNA, thereby lowering the expression levels of the gene. Mainly insertion of the 5' UTR L1 promoters impact the expression of nearby genes (31, 113, 114). The 5' UTR of L1 contains a sense and antisense promoter that can interfere with normal gene expression of adjacent cellular genes (113). Transcription from the antisense promoter can lead to the formation of chimeric antisense transcripts containing a part of the L1 5' UTR and gene sequences flanking the 5' end of L1 (31, 113). The L1 promoters can also act as promoters for regulatory miRNAs, thereby indirectly affecting the expression levels of other genes (115, 116). A study observed that L1 integrations primarily occur antisense to gene introns in hippocampal neurons (24). A mechanistic explanation for this preference is still lacking.

Integrated L1 elements do not need to be fully functional to exhibit cell toxicity and DNA damage (6, 117). Truncated L1s that are retrotransposition defective may still encode one or part of the functional domains of L1 (117). It was shown that L1 sequences containing a premature stop codon can still encode part of the ORF2 leading to a truncated ORF2p that contains a full endonuclease domain (118). This endonuclease activity may generate retrotransposition-independent nicks and contribute to formation of double stranded breaks (DSB) in genomic DNA (119).

1.6 Pathologies associated with L1 retrotransposition

There is accumulating evidence that TEs also play a role in more complex diseases. Changes in the mobility of retrotransposons has been observed in a range of inflammatory, neurological and neurodegenerative diseases (54). The question rises how endogenous retroelements are implicated in the pathophysiological pathways leading to the pathogenesis or aggravation of these diseases. What is the role of endogenous retroelements when there is an innate inflammatory response, but no viral infection is detected? Misregulation of retrotransposons may have a larger impact on human disease than initially anticipated (54). In the following paragraph a concise overview is given of pathologies that have already been associated with altered L1 retrotransposition.

1.6.1 Rett syndrome and MeCP2 duplication syndrome

Rett syndrome (RTT) is a neurodevelopmental disorder that affects approximately 1 in 10 000 young girls (120, 121). RTT patients appear to develop normally until 6-18 months of age but gradually start to lose their acquired abilities (122). Loss-of-function mutations in the MeCP2 gene were found to be the main cause of RTT (123). The MeCP2 gene is located on the X chromosome and is in consequence subjected to X-inactivation that influences the phenotypic severity of RTT (121). MeCP2 is ubiquitously expressed throughout the human body where the highest expression levels are seen in the brain, particularly in neurons (50). Most RTT causing mutations in MeCP2 are found in the MBD or in the TRD interfering with the ability of MeCP2 to bind methylated DNA or cofactors, respectively (48, 124).

Strict regulation of MeCP2 during neurodevelopment and adulthood is necessary given that both up- and downregulation of MeCP2 causes neurological dysfunctions (125). In contrast to RTT, MeCP2 gain-of-function mutations in MeCP2 duplication syndrome are also detrimental for life quality (126). These patients present with neurological dysfunctions similar to RTT (122, 125). This underlines the importance of an appropriate balance in MeCP2 expression. A study that used NPCs derived from induced pluripotent stem cells (iPSC) and human tissue of RTT patients revealed that RTT patients who carry a MeCP2 mutation have increased susceptibility for L1 retrotransposition (47). Another study found more somatic L1 insertions in the brain of RTT patients compared to non-brain tissue of the same patient (127). This is not surprising considering MeCP2 is a known epigenetic regulator of L1 and is mainly expressed in the brain (47). The high rates of L1 retrotransposition in RTT patients leads to an increased

rate of somatic mutations in the brain of these patients, leaving us with the question whether L1 retrotransposition is the cause or a consequence of RTT disease progression.

1.6.2 Autism spectrum disorders

Autism spectrum disorders (ASD) are developmental disorders that affect behavior and communication, but the etiology of ASD remains unknown. Various genetic and environmental factors have been identified that are involved in the occurrence of ASD (128). RTT was once considered part of ASD. Together with the evidence that L1 is responsive to environmental stressors, it prompted researchers to study the role of L1 retrotransposition in ASD and other neurobehavioral disorders (129). Indeed, L1 expression was found significantly increased in the brain of ASD patients (130). This has been related to lower methylation levels of the L1 promoter and consequently less repression by MeCP2 in the brain of these patients (130, 131). In fact, another study found an association between L1 insertions and dysregulated genes that are related to ASD (132). Further research will be needed to investigate which neuronal circuits may be affected by L1 insertions in the brain of these patients

1.6.3 Schizophrenia

More and more evidence accumulates on the association between somatic mutations and the pathophysiological mechanisms underlying neurobehavioral and psychiatric disorders (133). Schizophrenia is a complex neuropsychiatric disorder that affects around 0.5 - 1 % of the global population (134). Starting from the hypothesis that L1 retrotransposition may not only be involved in Mendelian disorders, Bundo *et al*. studied L1 retrotransposition in the pathophysiology of mental disorders (135). They found an increased L1 content in brain samples from established schizophrenia mice models as well as in neurons from schizophrenia patients. They observed that L1 insertions were concentrated in genomic loci related to schizophrenia and synaptic function. Their results indicated that both genetic components and environmental factors during early stages of development could increase susceptibility to schizophrenia and play a role in pathophysiology. Another study used next-generation sequencing to detect L1 insertions in schizophrenia patient samples (136). One third of these insertions were detected in the coding sequence of protein-coding genes implicated in the pathogenesis schizophrenia. These findings were confirmed by another study where a significant increase in novel intragenic L1 insertions was observed in dorsolateral prefrontal cortex neurons of schizophrenia patients (137). Taken together, these studies indicate that L1

retrotransposition during embryogenesis and adult neurogenesis may contribute to the etiology of schizophrenia. Additionally, intragenic L1 insertions may be inherited and may form an increased risk for schizophrenia by disrupting gene function and subsequently lead to neurodevelopmental changes (137). Furthermore, several studies observed a decrease in L1 methylation levels, although other studies did not find the same results possibly due to differences in demographic and clinical variables (138-142). Nevertheless, an increase in L1 activity as a result of hypomethylation has been observed in other mental disorders including post-traumatic stress disorder, bipolar disorder and major depressive disorder (6).

1.6.4 Aicardi-Goutières syndrome

Aicardi-Goutières syndrome (AGS) is an inherited neuroinflammatory disorder associated with constitutive upregulation of type I interferon (IFN) production (143). Patients suffer from early-onset encephalopathy that resembles a congenital HIV-1 infection and also shows similarities with the autoimmune disease systemic lupus erythematosus. AGS can be caused by mutations in different genes that are mainly involved in nucleic acid metabolism and signaling, including TREX1, SAMHD1, RNase H2, ADAR and IFIH1. Loss-of-function mutations in these proteins cause a disturbance of endogenous nucleic acid pathways triggering an innate immune response that is normally induced by exogenous nucleic acids (143). There are two possible sources of endogenous nucleic acids that can elicit such a response: 1) nucleic acids arising as a result of chronic DNA damage, or 2) nucleic acids derived from retrotransposons (144).

TREX1, SAMHD1 and ADAR1 are all negative regulators of L1 retrotransposition (102, 107, 145). Mutations in these AGS-associated proteins may comprise their ability to regulate both endogenous and exogenous nucleic acids. Based on these observations, a study was performed in TREX1-deficient mice that were given a combination of HIV-1 reverse transcriptase inhibitors (RTi) (146). They were able to rescue the inflammatory phenotype of the mice, supporting the hypothesis that RTi can reduce IFN signaling in AGS patients by inhibiting the reverse transcription of endogenous retrotransposons. A phase II clinical trial was conducted for AGS patients who received combinations of RTi (147). A reduction of IFN levels and a decrease in IFN stimulated genes was observed after 12 months but returned to pre-treatment levels after discontinuing the therapy.

Mutations in the three genes encoding the RNase H2 subunits are the most common cause of AGS (148, 149). At first, it was suggested that RNase H2 would inhibit L1 retrotransposition,

similar to TREX1, SAMHD1 and ADAR1 (144). However, results of a study indicated that RNase H2 activity promotes L1 retrotransposition, questioning the proposed mechanism for AGS pathogenesis centered around accumulation of endogenous retroelements (95). Instead they proposed that genome instability and the accumulation of RNA:cDNA hybrids in the nucleus might be the underlying cause of inflammation in AGS caused by RNase H2 mutations. Further research is required to determine the relative importance of endogenous retroelements in AGS. Other AGS disease-causing genes such as IFIH1, TMEM173 and ISG15 may also have a role in metabolism of retroelements, since many were already found to modulate or repress exogenous retroviruses (143, 150-153). More research is needed to elucidate their potential roles in the process of retrotransposition.

1.6.5 Ataxia-telangiectasia

Ataxia-telangiectasia (AT) is a neurodegenerative disease that is characterized by progressive cerebellar degeneration (154). AT is caused by inactivating mutations in the ataxia telangiectasia mutated (ATM) gene that encodes a serine/threonine kinase that responds to DNA damage. Mutations in ATM leads to the inability to repair particular types of DSBs resulting in DNA mutagenesis (155). Coufal and colleagues observed an increase in L1 retrotransposition in ATM deficient cells suggesting a role of ATM in the repression of L1 retrotransposition (156). Their finding was supported by the observation of an increased L1 content in the genomic DNA of postmortem hippocampal brain tissues of AT patients. They suggest that ATM acts as repair factor of the dsDNA break generated during TPRT. Consequently, loss of ATM would lead to more or longer L1 insertions.

1.6.6 Neurodegenerative diseases and ageing

Interest around the impact of retrotransposons in neurodegenerative disorders is increasing. Altered activity of retrotransposons has been observed in various neurodegenerative disorders and ageing (54). Most of the evidence exists for amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) (157). ALS is a neuromuscular disease that is characterized by degeneration of the upper and lower motor neurons in the corticospinal tract (157). FTD belongs to the same disease spectrum as ALS and involves frontotemporal lobe degeneration (FTLD) in the brain leading to progressive changes in personality, behavior, language dysfunction and executive deficits (157). ALS and FTD are often associated with accumulation of RNA-binding protein TDP-43 that forms ubiquitinated inclusions in diseased neurons. A study showed that

TDP-43 binds to L1 mRNA transcripts, and that this binding is reduced in FTLD patients carrying a TDP-43 mutation (158). Another study found that loss of TDP-43 is associated with increased chromatin accessibility around L1 in post-mortem brain samples of ALS-FTD patients, leading to increased L1 retrotransposition (159). These results stand in contrast to another study that did not find evidence for misregulation of L1 expression in sporadic ALS nor did they find an effect of TDP-43 on L1 retrotransposition (160). More research is needed to elucidate the contribution that TEs may have in TDP-related neurodegenerative disorders.

Alzheimer's disease (AD) is a neurodegenerative disease and the most common cause of dementia that is characterized by a progressive loss of memory and cognitive abilities (161). There was a growing interest to study the role of TEs in AD after the observation of somatic and germline mosaicism in patients with sporadic early-onset AD (162). Bollati *et al*. investigated the methylation levels of L1 in AD patients and compared these to healthy controls (163). They observed an increased L1 content in AD patients. Higher levels of L1 methylation were seen in the group of patients who scored better on cognitive tests than in the group that performed worse. These results were contradicted by another study that did not observe differences in L1 methylation levels between AD patients and healthy controls (164). These negative results were also confirmed by another study that did not observe a difference in L1 copy number between AD patients and age-matched healthy controls (165). Yet, a more recent study showed that tau activates TEs leading to an altered transcriptional landscape in AD (166). It was suggested that this aberrant activity of TEs may be damaging to neurons, contributing to the progression of AD. It would be interesting to continue investigations down this road to elucidate the role of TEs in AD considering previous studies that showed L1 retrotransposition in hippocampal neurons is able to induce somatic mosaicism that can influence memory formation and possibly other cognitive functions (26).

Ageing is a major risk factor in various neurodegenerative diseases. Accumulation of age-related mutations makes it intriguing to hypothesize that L1 retrotransposition may be a source of this DNA damage (6, 119). L1 retrotransposition may cause somatic mutations in mature nondividing neurons leading to somatic mutations that accumulate over the course of human ageing and even cause neurodegeneration (167). Li *et al*. used *Drosophila* as a model to study the activity of TEs during ageing by comparing the brain tissues at different ages (168). They found increasing activity of TEs with age suggesting that TEs may contribute to age-dependent neuronal decline.

1.6.7 Cancer

L1 retrotransposons constitute a source of mutations in our genome. They exert adverse effects on genome stability. Therefore, control of L1 retrotransposons is important to maintain genome integrity. It is becoming more and more evident that L1 as a source of mutations can be problematic as it has been shown that L1 insertions can participate in the origin and evolution of cancer (169). L1 insertions in or near tumor suppressor genes or oncogenes can contribute to tumor development, progression and metastasis. The number of L1 insertions vary among tumors, with some having more than 50 somatic insertions and other types of cancers completely lacking L1 insertions (170). Mechanisms such as epigenetic modification that repress L1 retrotransposition are often dysregulated in cancer (169). Studies revealed a correlation between hypomethylation of the L1 promoter and increased L1 expression in certain types (171, 172).

Thousands of somatic mutations caused by L1 insertions have been identified in many tumor types. The majority of these mutations are passenger mutations (173). The question remains how often L1 insertions drive tumorigenesis in humans. A better understanding of L1 retrotransposition in healthy somatic tissue is required in order to better understand the role of L1 in carcinogenesis.

1.7 The cell culture-based L1 retrotransposition assay

The L1 retrotransposition assay is widely used as a reliable method to study L1 retrotransposition. This method allows to study the retrotransposition process and host factors that regulate L1 retrotransposition in cell cultures as well as in animal models (174, 175). The experimental system was first described by Moran *et al*. to study L1 retrotransposition in cultured mammalian cell lines (176). They designed a plasmid that contained a human L1 element and cloned a reporter cassette in the 3' UTR of the L1. The reporter consisted of a neomycin resistance (NeoR) gene in the antisense direction interrupted by a γ-globin intron flanked by a splice donor and acceptor site in the opposite transcriptional direction. The cassette also contained a polyA signal at the 5' end of the reporter gene and a heterologous promoter at the 3' end of the reporter gene. After transcription the L1-NeoR transcript is spliced, reverse transcribed and integrated into the genomic DNA. Cells transfected with the plasmid will only become G418-resistant after a full cycle of retrotransposition.

Other L1 retrotransposition plasmids have been developed containing different reporter cassettes for the detection of L1 mobilization. Ostertag *et al*. made use of enhanced green fluorescent protein (EGFP) to measure relative L1 retrotransposition rates (177). They developed a plasmid that contains the human $L1_{RP}$ tagged with an EGFP cassette in the 3' UTR (Figure 4). In this plasmid EGFP is flanked by a cytomegalovirus (CMV) promoter and a polyA signal. Transcription of $L1_{RP}$ is driven by the endogenous L1 promoter in the 5' UTR. EGFP positive cells can be visually detected by FACS analysis. The EGFP positive cells represent the cells in which a retrotransposition event has taken place. The integrated copy number of the EGFP gene can also be determined by quantitative Real-Time PCR on the genomic DNA of transfected cells (178). This way an estimation of L1 retrotransposition can be made regardless of EGFP gene expression.

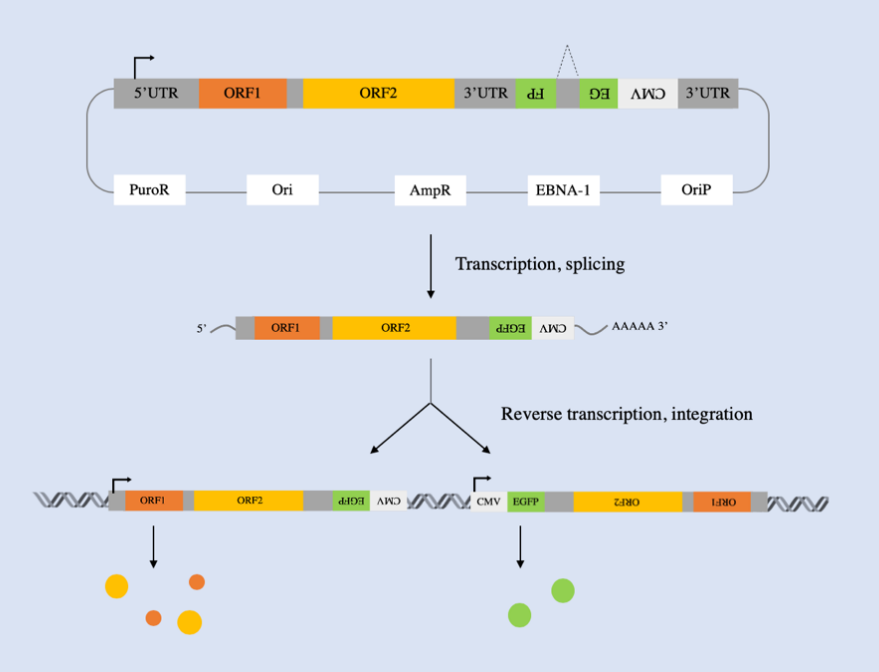


Figure 4: ***The L1 retrotransposition assay with an EGFP reporter cassette.*** *The active human $L1_{RP}$ is cloned in a pCEP4 backbone. $L1_{RP}$ consists of 2 ORFs flanked by a 5' UTR and a 3' UTR in which a reporter cassette is cloned. The EGFP reporter gene is interrupted by a γ-globin intron that is positioned in the opposite transcriptional direction. Cells will express EGFP when the transcript undergoes splicing, reverse transcription and integration in the genomic DNA where it can be expressed from the CMV promoter. The plasmid contains a puromycin resistance gene (PuroR) for selection of transfected cells. Epstein-Barr virus nuclear antigen 1 (EBNA-1) and Epstein-Barr virus origin of replication (oriP) allow replication of episomes in primate cells. The plasmid can also replicate in prokaryotic cells through the E. coli origin of replication (ori) and ampicillin resistance gene (AmpR). Integrated L1 elements that are fully functional can undergo multiple cycles of retrotransposition.*

1.8 Hypothesis

L1 retrotransposition in the human genome has a multitude of consequences but the biological significance is not well-understood. Retrotransposons are known to contribute to somatic mosaicism and to act as potent regulatory elements in the human genome although many aspects on their function and effect on gene regulation remain unclear. Growing evidence suggests a role for L1 in healthy somatic tissue as well as in various diseases. L1 retrotransposons have been associated with a wide range of human diseases, making them potential pathogens and targets for new therapeutics. Therefore, thorough understanding of their biology is required. Fundamental research aims to identify new host factors required for L1 retrotransposition.

In my Master thesis I explore host factors of L1 retrotransposition. Since many known regulators of L1 retrotransposition show similarities with responses to retroviral pathogens, we hypothesized that lens epithelium-derived growth factor/p75 (LEDGF/p75) can act as a host factor of L1 retrotransposition. Given that LEDGF/p75 interacts with MeCP2, a known regulator of L1 retrotransposition, and that LEDGF/p75 is an essential host factor for HIV-1 integration, we hypothesized that LEDGF/p75 may play a direct or an indirect role in L1 retrotransposition.

2 Research objectives

2.1 General objective

L1 is the only autonomous retrotransposons in the human genome. It is estimated that 100 L1 elements are still active in our genome, a small part of which is even highly active and plays an important role in genome plasticity and expansion. L1 activity has shown to be implicated in neurophysiological processes such as memory formation in neuronal progenitor cells. On the other hand, imbalance in L1 activity has been observed in human diseases including neurodevelopmental disorders and cancer. A thorough understanding of the L1 interactome is essential to understand its implications in health and disease.

2.2 Specific research objectives

Objective 1: To optimize the L1 retrotransposition assay in two different cell lines.

The goal is to reproduce the L1 retrotransposition assay in our lab using a $L1_{RP}$-EGFP reporter plasmid and to optimize the conditions to acquire a robust readout. Two different cancer cell lines, HEK293T and HeLa P4 cells, will be used to perform this assay. HEK293T cells and HeLa cells have been used before in this assay (106, 176). The L1 retrotransposition assay will be validated using MeCP2 depletion cell lines.

Objective 2: To study the role of LEDGF/p75 in L1 retrotransposition.

LEDGF/p75 is an essential host factor involved in the integration of HIV-1 (179). The goal is to determine whether LEDGF/p75 is also part of the L1 retrotransposition interactome, using the optimized L1 retrotransposition assay.

3 Materials and methods

3.1 Cell culture

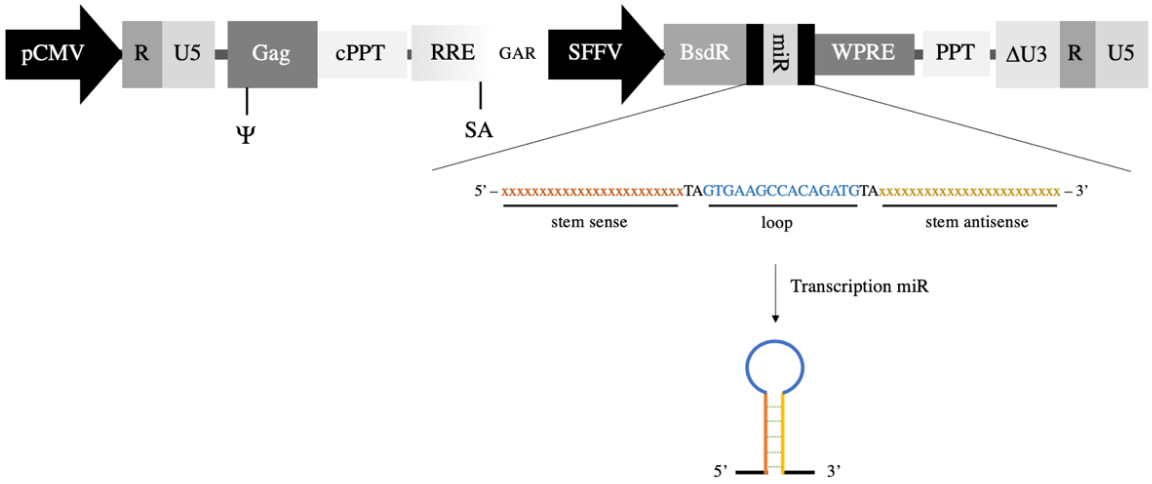
HEK293T and HeLa P4 cells were grown in a humified atmosphere containing 5 % CO_2 at 37°C. HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) with GlutaMAX [Gibco] supplemented with 5 % fetal bovine calf serum (FCS; [Gibco]) and 50 µg/mL gentamicin [Gibco], hereafter referred to as DMEM complete medium. HeLa P4 cells were cultured in DMEM complete medium supplemented with 0.5 mg/mL geneticin [Gibco]. Cultured cells were routinely checked for mycoplasma using the PlasmoTest Mycoplasma Detection kit [Invivogen] to ensure lack of contamination.

3.2 Plasmids

Transfer plasmids encoding a single miR30-based short hairpin RNA (shRNA) were created for viral vector production as previously described (180). DNA sequences encoding four different MeCP2-specific miRNA-based shRNAs (TRCN0000330972, TRCN0000330971, TRCN0000379510, TRCN0000021241; Table 1) were cloned into a pGAE backbone containing a spleen focus forming viral (SFFV) promoter driving the miRNA and a blasticidin resistance (BsdR) gene and a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) for simian immunodeficiency virus (SIV) vector production (Figure 5). All shRNAs target both isoforms of MeCP2 for knock-down. A miRNA containing a scrambled sequence or a previously established shRNA against firefly luciferase (Fluc) were used as non-targeting controls. Digestion of the pGAE backbone with E*sp3I* allowed ligation of the annealed oligonucleotide pairs. All cloning steps DNA sequence verified.

For the L1 retrotransposition assay the retrotransposition competent plasmid 99-L1RPS-EGFP-Puro ($L1_{RP}$) and the retrotransposition defective control plasmid 99-JM111-EGFP-Puro (JM111), a kind gift of Prof. T. Gramberg (Friedrich-Alexander-University Erlangen-Nürnberg, Germany) were used as described previously (177). The L1 sequence is based on the human $L1_{RP}$ sequence and cloned into a pCEP4 backbone containing a PuroR gene (Figure 4).

A) pGAE_miR transfer plasmid



B) pSIV3+ packaging plasmid

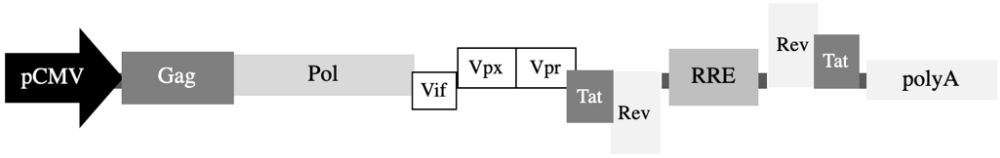


Figure 5: ***Schematic representation of pGAE_miR30 transfer plasmids and the pSIV3+ packaging construct for SIV vector production****. Promoters are symbolized by black arrows. Genes and cis-elements are depicted by boxes. A) The pGAE transfer plasmid contains a spleen focus-forming virus promoter (SFFV) driving a blasticidin resistance gene (BsdR) and a miR30-based shRNA gene (miR) to target MeCP2 and to create a potent knock-down through RNAi. The miR is followed by a woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) for stabilization of the mRNA. The miR consists of a complementary sense and antisense sequence connected by a loop sequence. After transcription of the miR these complementary sequences will form the stem of the miR. The stem is flanked by a 5' and 3' miR30-specific backbone sequence. SIV vectors are designed to be self-inactivating by a partially deleted 3' unique element (ΔU3) in the 3' long terminal repeat (LTR). B) Schematic representation of the pSIV3+ packaging plasmid. Early cytomegalovirus promoter (pCMV); repeat element (R); 5' unique element (U5): group-specific antigen (Gag); packaging signal (Ψ); central (cPPT) and 3' polypurine tracks (PPT); Rev response element (RRE); the SIV Rev/Tat splice acceptor (SA) is extended to the purine-rich region (GAR); polymerase (Pol); polyadenylation signal (polyA). Figure adapted from Mangeot et al. (181) and Osorio et al. (180).*

Table 1: ***List of MeCP2-specific miR30-based shRNAs.***

Name	Target site	Sequence
MeCP2-specific		
MeCP2_miR_1	5' UTR	AAGAAGGTATATTGCTGTTGACAGTGAGCGGTGACAAAG CTTCCCGATTAACTAGTGAAGCCACAGATGTAGTTAATCG GGAAGCTTTGTCAGTGCCTACTGCCTCGGACTTCAAGGG
MeCP2_miR_2	ORF	AAGAAGGTATATTGCTGTTGACAGTGAGCGCGAGAGCGC AAAGACATTGTTTTAGTGAAGCCACAGATGTAAAACAAT GTCTTTGCGCTCTCCTGCCTACTGCCTCGGACTTCAAGGG
MeCP2_miR_3	ORF	AAGAAGGTATATTGCTGTTGACAGTGAGCGTGCCGTGAA GGAGTCTTCTATCTAGTGAAGCCACAGATGTAGATAGAA GACTCCTTCACGGCTTGCCTACTGCCTCGGACTTCAAGGG
MeCP2_miR_4	ORF	AAGAAGGTATATTGCTGTTGACAGTGAGCGCCTGGGAAG TATGATGTGTATTTAGTGAAGCCACAGATGTAAATACAC ATCATACTTCCCAGCTGCCTACTGCCTCGGACTTCAAGGG
Controls		
Ctrl_miR_scr	Non-targeting	AAGAAGGTATATTGCTGTTGACAGTGAGCGACGCTCTAA AGTGGAGTTGATTTAGTGAAGCCACAGATGTAAATCAAC TCCACTTTAGAGCGATGCCTACTGCCTCGGACTTCAAGGG
Ctrl_miR_fluc	Fluc	AAGAAGGTATATTGCTGTTGACAGTGAGCGGCGCTGAGT ACTTCGAAATGTCTAGTGAAGCCACAGATGTAGACATTTC GAAGTACTCAGCGTTGCCTACTGCCTCGGACTTCAAGGG

3.3 Viral vector production

SIV vectors were prepared in 10 cm cell culture dishes with the triple transient transfection method using linear polyethylenimine (PEI; [Polysciences]) as described previously with minor modifications (182). For vector productions, 6 x 10^6 HEK293T cells were seeded in Opti-MEM [Gibco] supplemented with 2 % FCS [Gibco]. 24 h after seeding, the cells were triple transfected with 5 µg envelope plasmid pMD.G (183) that encodes vesicular stomatitis virus glycoprotein (VSV-G), 15 µg pSIV3+ packaging plasmid, a kind gift from Prof. D. Nègre (Université De Lyon, France), and 15 µg pGAE transfer plasmid. Medium was replaced 24 h after transfection by Opti-MEM [Gibco]. Supernatant was collected 48 h and 72 h after transfection and filtrated over a 0.45 µm pore-size syringe filter [Merck]. The vector was concentrated using a Vivaspin 15 50,000 MW cut-off [Sartorius] and stored at -80°C.

3.4 Generation of stable knock-down cell lines

To generate stable MeCP2 KD cell lines, HEK293T or HeLa P4 cells were seeded at 2 x 10^4 cells/well in 96-well plates and transduced for 3 days in DMEM complete medium. Cells were transduced with serial dilutions of vector. After transduction cells were continuously kept under selection in DMEM complete medium with 5 μg/mL blasticidin. KD was confirmed on Western Blot and/or RT-qPCR.

A monoclonal HEK293T LEDGF/p75 KD zeocin resistant (ZeoR) cell line was previously established in the host lab by transduction with an SIV-based vector encoding a LEDGF/p75-specific miR30-based shRNA (184). Cells were kept under continuous selection in DMEM complete medium supplemented with 100 μg/mL zeocin.

3.5 Western blot

Cells were plated in 6-well plates at 1 x 10^6 cells/well in DMEM complete medium. After 24 h cells were washed twice with PBS and lysed in 1 % SDS. Protein concentrations of whole cell extracts were determined using the BCA protein assay [ThermoFisher Scientific]. Cell extracts containing 20 μg of total protein were separated on 12.5 % SDS-PAGE in-house gels and electroblotted on Polyvinylidene difluoride membranes [Bio-rad]. Membranes were blocked in PBS with 0.1 % Triton-X100 and 5 % milk subsequently incubated with 1:1000 rabbit MeCP2-specific antibodies [Proteintech]. Mouse β-tubulin-specific antibodies [Abcam] were used to confirm equal loading. Detection was performed using secondary horseradish peroxidase-conjugated goat anti-rabbit [Thermofisher] or goat anti-mouse antibodies [Dako], chemiluminescence (Clarity ECL [Bio-rad]) and the LAS-4000 imaging system [Fujifilm].

3.6 Cell count

HeLa P4 or HEK293T cells were plated at 25 000 cells/well in a 24-well plate in DMEM complete medium supplemented with the appropriate antibiotic to keep cells under selection. Cells were trypsinized and counted after 24 h, 48 h, 72 h, and 96 h using the TC20 Automated Cell Counter [Bio-rad].

Analogously, HeLa P4 or HEK293T cells were plated at 6000 cells/well in a 24-well plate in DMEM complete medium supplemented with the appropriate selection antibiotic. Cells were

trypsinized and counted at day 1, 4, 5 and 6 after seeding using the TC20 Automated Cell Counter [Bio-rad].

3.7 L1 retrotransposition assay

For the L1 retrotransposition assay cells were seeded in 48-well dishes at a density of 3.5 x 10^4 cells/well in DMEM complete medium. After 24 h cells were transfected with the $L1_{RP}$ or JM111 plasmid using the FuGENE 6 transfection reagent [Promega]. For 1 µg plasmid DNA 4.5 µL FuGENE was used in a final volume of 100 µL Opti-MEM [Gibco]. An additional transfection mix was prepared to determine transfection efficiency consisting of 0.5 µg of a GFP mock reporter plasmid, 1 µg of $L1_{RP}$ or JM111 plasmid and 6 µL FuGENE in a final volume of 100 µL Opti-MEM [Gibco]. The transfection mixes were added to the cells in a ratio of 1:11 with Opti-MEM. Transfection was stopped after 24 h. Transfection efficiency was determined 2 days post-transfection and selection started from this moment onwards using DMEM complete medium supplemented with 1 µg/mL puromycin. Samples were taken at various time points after transfection by trypsinizing the cells and collecting them for flow cytometry. Cells were fixed using 4 % paraformaldehyde.

3.8 Flow cytometry

The percentage of green fluorescent cells in the samples was measured using the Guava EasyCyte HT with 488 nm laser [Luminex]. Data was analyzed using the InCyte Software for Guava EasyCyte HT Systems software package provided with the instrument. The forward scatter versus side scatter was used to gate for living cells.

3.9 RT-qPCR

Real-time quantitative PCR (RT-qPCR) was performed to determine mRNA levels. Total RNA was isolated using the total RNA Mini Kit [Bio-rad]. RNA concentrations were measured photometrically using the NanoPhotometer [Implen] and adjusted to 50 ng/µL for reverse transcription using the High-capacity cDNA Archive Kit [Applied Biosystems]. RT-qPCR was performed using TaqMan probes and the iQ supermix [Bio-rad] or SYBR Green Master Mix [ThermoFisher Scientific] to determine melting curves. The Lightcycler 480 [Roche] was used as detection system. Primers used to determine mRNA levels of MeCP2 and LEDGF/p75 are described in Table 2. β-actin was used as a housekeeping control to normalize mRNA levels.

Table 2: ***List of RT-qPCR primers.***

Gene	Forward primer	Reverse primer	Probe
RT-qPCR on mRNA			
MeCP2	CACGGAAGCTTAAGCAAAGG	CTGGAGCTTTGGGAGATTG	GGTAGGCGACACATCCCTGG
LEDGF/p75	GAACTTGCTTCACTTCAGGTCACA	TCGCCGTATTTTTTTCAGTGTAGT	TGCAACAAGCTCAGAAACACACAGAGATGA
β-actin	CACTGAGCGAGGCTACAGCTT	TTGATGTCGCGCACGATTT	ACCACCACGGCCGAGCGG
RT-qPCR on genomic DNA			
EGFP *(178)*	GGTCACGAACTCCAGCAG	CAGAAGAACGGCATCAAGG	/
β-actin gDNA	TCACCCACACTGTGCCCATCTACGA	CAGCGGAACCGCTCATTGCCAATGG	ATGCCCTCCCCCATGCCATCCTGCGT

To determine the integrated copy number of episomal-derived EGFP a RT-qPCR was performed on genomic DNA on day 13 post-transfection. Genomic DNA was extracted using the DNA Extraction Kit [Bio-rad] with DNase-free RNase. DNA concentrations were measured using the NanoPhotometer [Implen] and adjusted to 70 ng/µL. RT-qPCR was performed using the SYBR Green Master Mix [ThermoFisher Scientific] and the Lightcycler 480 [Roche] as detection system. The primers flank the γ-globin intron interrupting EGFP and are listed in Table 2. β-actin was used as a housekeeping control.

3.10 Statistical analysis

Statistical analysis was performed using the GraphPad Prism 8.0 software package and Excel. Results are expressed as mean ± standard deviation (SD).

4 Results

4.1 Optimization of the L1 retrotransposition assay

4.1.1 Generation and validation of MeCP2 KD HEK293T and HeLa P4 cell lines

We started by creating stable polyclonal MeCP2 KD cell lines in two immortalized cancer cell lines, HeLa P4 and HEK293T. Four different shRNA sequences from the GenomeRNAi database were used to create viral vectors expressing a single miR30-based shRNA that target MeCP2 and induce potent gene silencing (180). MeCP2 is a well-known repressor of L1 retrotransposition, and therefore an excellent target to validate the L1 retrotransposition assay (47). The cells were transduced with a serial dilution of vectors in case vector toxicity would be observed at high vector concentrations. No vector toxicity was observed in the cells transduced with undiluted or diluted vectors. RT-qPCR was performed to confirm MeCP2 depletion in cells transduced with undiluted and 1/5 diluted vector. MeCP2 KD was observed in all cell lines on RNA level (Figure 6). Table 3 shows the percentages of KD compared to WT.

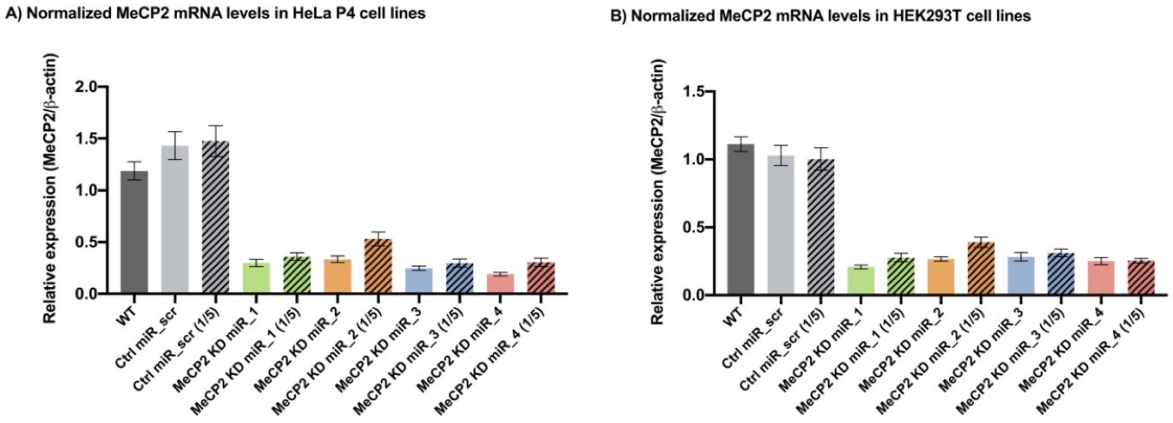


Figure 6: ***Validation of MeCP2 KD cell lines by RT-qPCR.*** *HeLa P4 (A) and HEK293T (B) cells were transduced with four different vectors, each expressing a MeCP2-specific miRNA or a scrambled control. MeCP2 mRNA levels were determined in cells transduced with the undiluted and 1/5 diluted vector concentrations from a serial dilution. After transduction cells were continuously kept under blasticidin selection. MeCP2 levels were determined 1 month after transduction. MeCP2 expression levels were normalized for β-actin expression levels. RT-qPCR was performed in technical triplicates. Data are presented as mean ± SD of a single test.*

Table 3: ***Percentages of MeCP2 KD compared to WT determined by RT-qPCR.***

Cell line	HEK293T		HeLa P4	
Vector dilution	**Undiluted**	**1/5 dilution**	**Undiluted**	**1/5 dilution**
MeCP2_miR_1	81 %	75 %	75 %	70 %
MeCP2_miR_2	76 %	65 %	72 %	55 %
MeCP2_miR_3	75 %	72 %	79 %	75 %
MeCP2_miR_4	78 %	77 %	84 %	75 %

MeCP2 KD cell lines were further validated on western blot (Figure 7). We only continued with the cells transduced with the undiluted vector concentration since they showed a more potent KD in preliminary RT-qPCR data. Endogenous MeCP2 could be detected at a molecular weight of 75 kDa in WT and control cell lines, however this was less clear for HeLa P4 cells. MeCP2 was clearly depleted in all HeLa P4 and HEK293T MeCP2 KD cell lines (Figure 7). Multiple other bands were detected above and below the expected MeCP2 band at 75 kDa in all samples. It has been suggested that post-transcriptional processing of MeCP2 results in multiple molecular forms that are detected as MeCP2 immunoreactive bands on western blots (185). However, many of these bands showed an equally strong signal in all cell lines, suggesting certain aspecific interactions of the MeCP2 antibody.

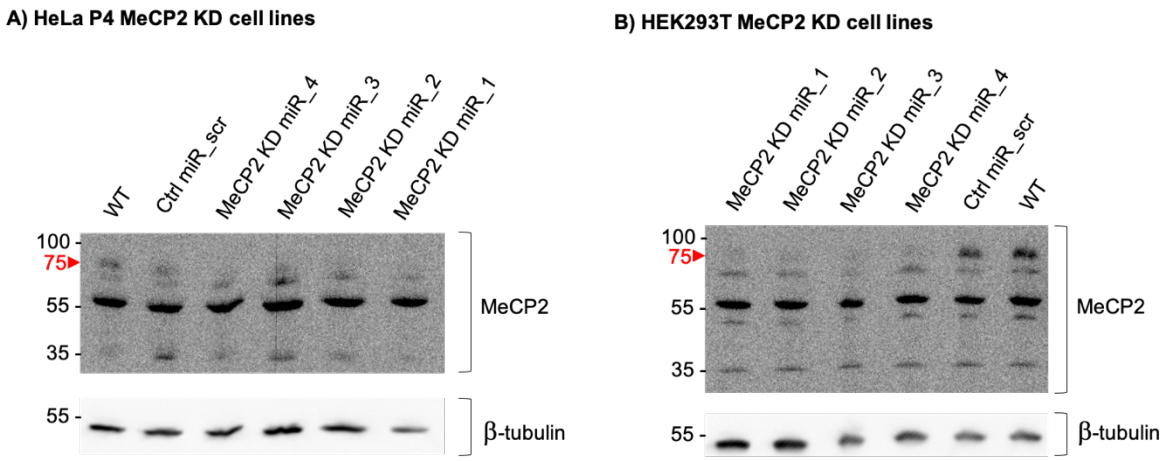


Figure 7: ***Western blot of MeCP2 KD HeLa P4 and HEK293T cell lines****. MeCP2 protein levels in HeLa P4 (A) and HEK293T (B) cells were determined on western blot 1 month after transduction. Transduced cells were continuously kept under blasticidin selection. MeCP2 was detected in WT and control cells as an immunoreactive band of 75 kDa. Multiple aspecific bands were observed in all cell lines. β-tubulin was used as loading control.*

The 75 kDa immunoreactive band of MeCP2 on western blot appeared to be less strong in HeLa P4 cells compared to HEK293T cells. We reasoned that lower levels of endogenous MeCP2 in HeLa P4 could influence the interpretation of the L1 retrotransposition assay when comparing WT and MeCP2 KD cells. Therefore, we compared endogenous MeCP2 mRNA levels between WT HeLa P4 and HEK293T cells by RT-qPCR. MeCP2 expression levels appeared to be two-fold higher in HEK293T compared to HeLa P4 cells (Figure 8).

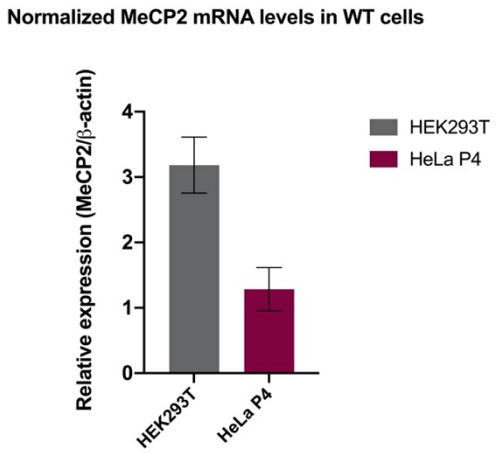


*Figure 8**: Comparison of endogenous MeCP2 mRNA levels between WT HEK293T and HeLa P4 cells.** *WT HEK293T cells show a two-fold higher MeCP2 mRNA level compared to WT HeLa P4 cells. MeCP2 expression levels were normalized for β-actin expression levels. RT-qPCR was performed in technical triplicates. Data are presented as mean ± SD of a single test.*

We continued to validate the MeCP2 KD miR_1 and MeCP2 KD miR_4 cell lines in HeLa P4 and HEK293T. MiR_1 and miR_4 target the 5'UTR and ORF of MeCP2, respectively. Both showed a MeCP2 KD of around 80 % on RT-qPCR in both HEK293T and HeLa P4 cells 1 month after transduction (Table 3). We repeated the RT-qPCR to confirm MeCP2 KD in these cell lines 2 months after transduction and included the LEDGF/p75 KD HEK293T and LEDGF KD HeLa P4 cell lines (Figure 9). An equally potent MeCP2 KD was observed in both MeCP2 KD HeLa P4 cell lines at 2 months post-transduction compared to 1 month post-transduction (Figure 9B). However, a less potent MeCP2 KD was observed in both MeCP2 KD HEK293T cell lines (Figure 9A). We observed a minor increase in MeCP2 mRNA levels in HEK293T MeCP2 KD cells 2 months post-transduction compared to 1 month post-transduction. Therefore, we also determined MeCP2 levels in MeCP2 KD HEK293T cell lines 6 months after transduction (Figure 10A). We observed that over time MeCP2 mRNA levels increased again resulting in only 46 % and 67 % KD for MeCP2 KD miR_1 and MeCP2 KD miR_4 HEK293T cells, respectively. Possibly this problem rises from the variability in potency of gene silencing in different cells of a polyclonal cell line. Cells with a more potent MeCP2 KD may have a disadvantage and are overgrown by cells with a less potent MeCP2 KD. To avoid this problem in the future, it is recommended to make a monoclonal cell line.

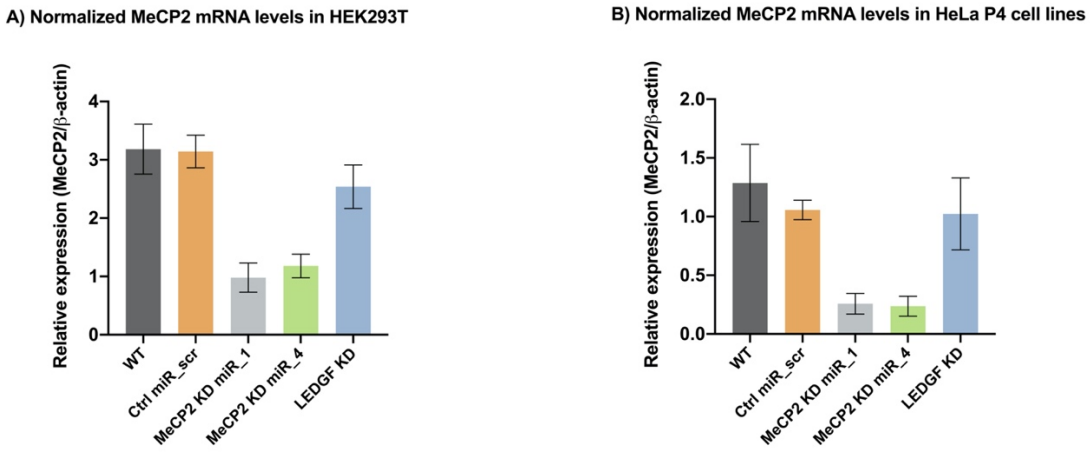


Figure 9: ***MeCP2 mRNA levels in HEK293T and HeLa P4 cell lines determined by RT-qPCR 2 months post-transduction.*** *MeCP2 mRNA levels were measured in MeCP2 KD HEK293T (A) and HeLa P4 (B) cell lines 2 months after transduction with a vector expressing a MeCP2-specific miRNA. MeCP2 mRNA levels were also verified in LEDGF/p75 KD cell lines. MeCP2 expression levels were normalized for β-actin expression levels. RT-qPCR was performed in technical triplicates. Data are presented as mean ± SD of a single test.*

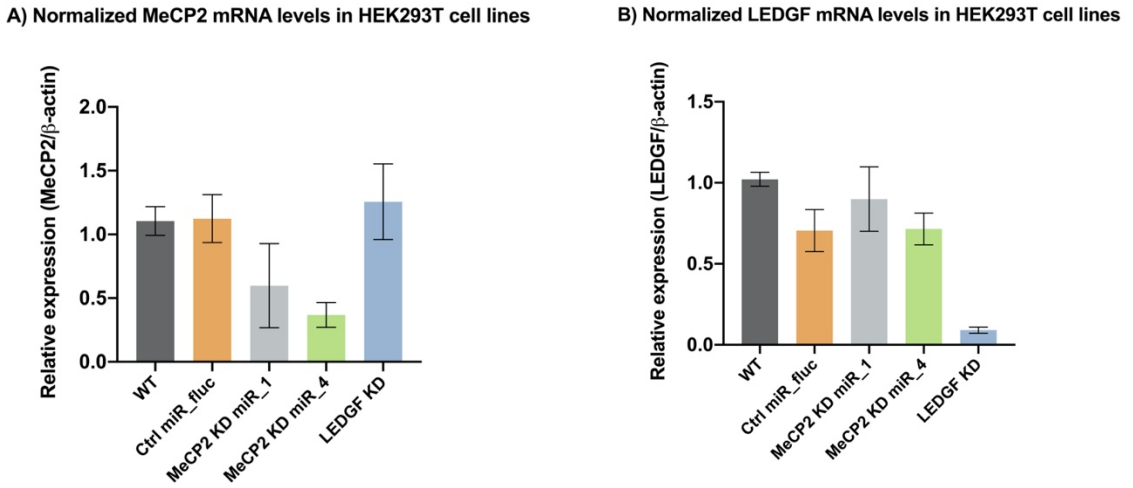


Figure 10: ***MeCP2 and LEDGF/p75 expression levels in HEK293T cell lines determined by RT-qPCR 6 months post-transduction.*** *A) MeCP2 mRNA levels were measured 6 months after transduction with a vector expressing a MeCP2-specific miRNA. MeCP2 KD was less potent after 6 months compared to the KD observed 1 month after transduction and only obtained 64 % KD for MeCP2 KD miR_1 and 57 % KD for MeCP2 KD miR_4. B) LEDGF/p75 mRNA levels were measured in MeCP2 and LEDGF/p75 KD cells. A potent KD of 90 % was observed in a previously established LEDGF/p75 KD HEK293T cell line from the host lab. LEDGF/p75 mRNA levels did not show a decrease in MeCP2 KD cells. Expression levels were normalized for β-actin expression levels. RT-qPCR was performed in technical triplicates. Data are presented as mean ± SD of two biological replicates.*

MeCP2 is a transcriptional repressor that is known to regulate transcription of a large number of genes. We investigated whether MeCP2 KD also influences the expression of LEDGF/p75 by performing RT-qPCR (Figure 10 B). A clear LEDGF/p75 KD of 90 % was observed in a previously established and validated LEDGF/p75 KD HEK293T cell line from the host lab. No decrease in LEDGF/p75 mRNA levels was seen in MeCP2 KD cell lines compared to WT and control (Figure 10B). On the other hand, LEDGF/p75 is also an important regulator of gene transcription that is known to interact with MeCP2. We investigated whether depletion of LEDGF/p75 influenced expression of MeCP2. LEDGF/p75 KD cells showed comparable MeCP2 levels to WT and control cells (Figure 9, Figure 10A). This would exclude the possibility of an indirect effect of LEDGF/p75 on L1 retrotransposition through transcriptional regulation of MeCP2.

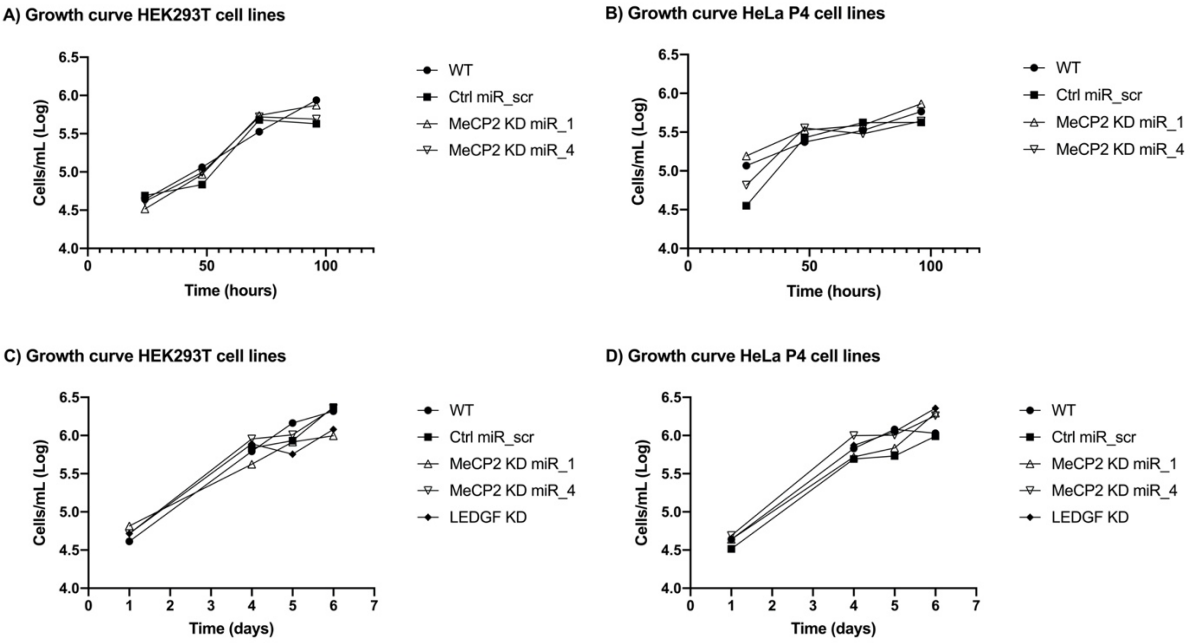


Figure 11: ***Growth curves of MeCP2 KD and LEDGF KD HEK293T and HeLa P4 cell lines****. (A and B) Cells were plated at a density of 25 000 cells/well in a 24-well plate. Cells were trypsinized and counted 24 h, 48 h, 72 h and 96 h after seeding. To confirm cell proliferation rates (C and D), cells were plated at a density of 6000 cells/well in a 24-well plate. Cells were trypsinized and counted on day 1, 4, 5 and 6 after seeding. The TC20 Automated Cell Counter [Bio-rad] was used to determine the total cell count. Cells were actively proliferating during the exponential growth phase in all conditions. Simple linear regression after logarithmic transformation was performed to determine the slopes of the growth curves. Slopes of the curves were not significantly different suggesting that the MeCP2 KD and LEDGF KD cell lines proliferate at the same rate as WT cells. No replicates were counted.*

Growth curves were made to ensure that the cell lines did not show different cell proliferation rates when performing the L1 retrotransposition assay. Cells were counted at four time points

during the exponential growth phase in which cultured cells actively proliferate (Figure 9). MeCP2 KD and LEDGF KD HEK293T and HeLa P4 cell lines were counted 24 h, 48 h, 72 h and 96 h after plating the cells (Figure 9A, 9B). Cell proliferation rates were confirmed by plating the cells at a lower density and counting them at day 1, 4, 5 and 6 after seeding (Figure 9C, 9D). The slopes were determined using simple linear regression after logarithmic transformation. Differences between the slopes were not significant suggesting that all cell lines proliferated at comparable rates.

4.1.2 The L1 retrotransposition assay in WT HEK293T and HeLa P4 cells

My next objective was to set up the L1 retrotransposition assay in HEK293T and HeLa P4 cells. The L1 retrotransposition assay was originally designed by the Kazazian laboratory to measure retrotransposition rates in cultured mammalian cells (176). The $L1_{RP}$ plasmid contains an active human L1 element that was found in exon 1 of the retinitis pigmentosa-2 gene of an X-linked retinitis pigmentosa patient (186). As a negative control the JM111 plasmid was used that encodes a mutated L1 element containing two missense mutations in ORF1 that abolish retrotransposition (176). An EGFP reporter cassette was introduced to detect real-time retrotransposition rates. After transfection with the $L1_{RP}$ plasmid or the negative control JM111 plasmid, cells can be analyzed for retrotransposition activity by flow cytometry.

Misregulation of L1 retrotransposition has been shown in various cancer cells, making it relevant to study retrotransposition in immortalized cancer cell lines. The assay requires to be optimized for each cell line to reach a robust readout. HEK293T cells have been used before in the L1 retrotransposition assay to study host factors, but yielded only low percentages of L1-EGFP positive cells (106). HeLa cells are more commonly used in the L1 retrotransposition assay (176, 187). In the host lab (Laboratory of Virology and Gene Therapy) the HeLa P4 cell line is routinely used to study HIV-1 replication. For these reasons we aimed to optimize the L1 retrotransposition assay in HEK293T and HeLa P4 cells.

We started by comparing WT HEK293T and HeLa P4 cells transfected with the retrotransposition competent $L1_{RP}$ plasmid over time (Figure 12). The cells were continuously kept under selection from 2 days post-transfection onwards. Samples were taken at multiple time points after transfection to determine the percentage of EGFP-positive cells by flow cytometry. Cells expressing EGFP represent the number of cells in which a retrotransposition event has taken place. Retrotransposition activity in HEK293T cells increased with time after

transfection and reached a plateau phase 15 days post-transfection. However, the percentages of EGFP positive cells in HeLa P4 cells remained low and did not exceed 9 % at any time point. In order to investigate whether these low percentages of retrotransposition are biologically relevant, $L1_{RP}$-transfected cells need to be compared to a negative control. The JM111 plasmid encodes a retrotransposition defective L1 element that is routinely used as negative control to determine background fluorescence.

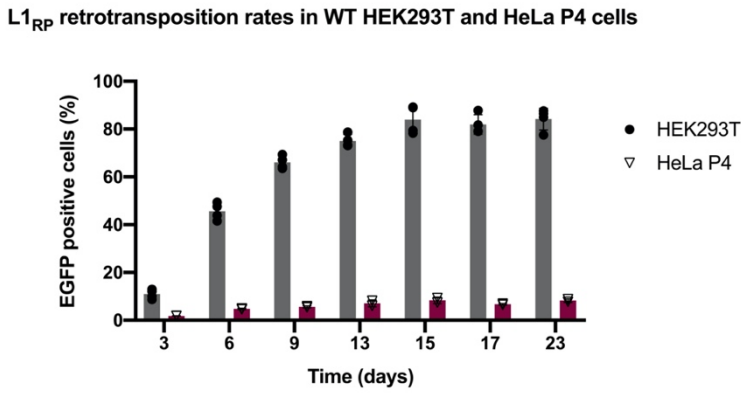


Figure 12: ***L1 retrotransposition rates in WT HEK293T and HeLa P4 cells.*** *WT HEK293T and HeLa P4 cells were transfected with the $L1_{RP}$-EGFP reporter plasmid and selected with puromycin from 2 days post-transfection onwards. Samples were taken at different time points and EGFP-positive cells were quantified by flow cytometry. Data represent a single test. Error bars represent the SD.*

4.1.3 Evaluation of the L1 retrotransposition assay using MeCP2 KD in HeLa P4 cells

We next investigated whether we could see a clear difference between WT and MeCP2 KD HeLa P4 cells even within these low percentages of EGFP-positive cells. Samples were taken at different time points after transfection and selection. Percentages of EGFP-positive cells were again very low for both WT and MeCP2 KD cells and did not exceed 6 % in any of the cell lines at any time point (Figure 13). No clear difference could be observed between WT and MeCP2 KD $L1_{RP}$ transfected cells. Additionally, cells transfected with $L1_{RP}$ did not clearly exceed the background retrotransposition levels of cells transfected with the retrotransposition defective JM111 plasmid. Due to the lack of biological relevance of these low percentages of EGFP-positive cells, no statistical analysis was performed.

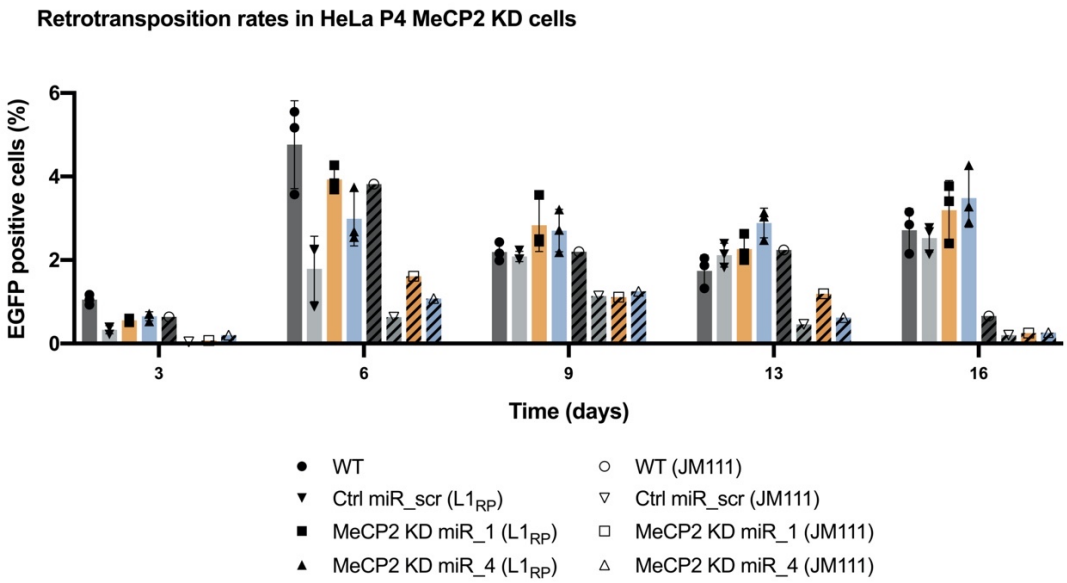


Figure 13: ***L1 retrotransposition rates in MeCP2 KD HeLa P4 cells.*** *MeCP2 KD cells were transfected with the retrotransposition competent $L1_{RP}$ or retrotransposition defective JM111 plasmid and were selected from 2 days post-transfection onwards. Samples were taken at multiple time points post-transfection and the percentage of EGFP-positive cells was determined by flow cytometry. Transfections with the $L1_{RP}$-plasmid were performed in triplicate. Data represent a single test. Error bars represent the SD.*

During the course of the experiment, we wanted to investigate whether increased retrotransposition rates in $L1_{RP}$-transfected cells induced cell toxicity compared to JM111-transfected cells. Therefore, we counted cells between day 13 and 17 post-transfection in parallel to the L1 retrotransposition assay (Figure 14). From a retrospective point of view, we expected no difference in cell toxicity between $L1_{RP}$- and JM111-transfected cells since no increased retrotransposition rates were observed in $L1_{RP}$-transfected cells compared to JM111-transfected cells using EGFP expression as readout (Figure 12). Cell counts showed large variability in the biological replicates of $L1_{RP}$-transfected cells across all cell lines (Figure 14). Since measuring retrotransposition-induced cell toxicity was not the primary aim of this experiment, only biological replicates from $L1_{RP}$ transfected cells were counted. No clear difference in cell growth were observed between $L1_{RP}$- and JM111-transfected cells in all of the cell lines.

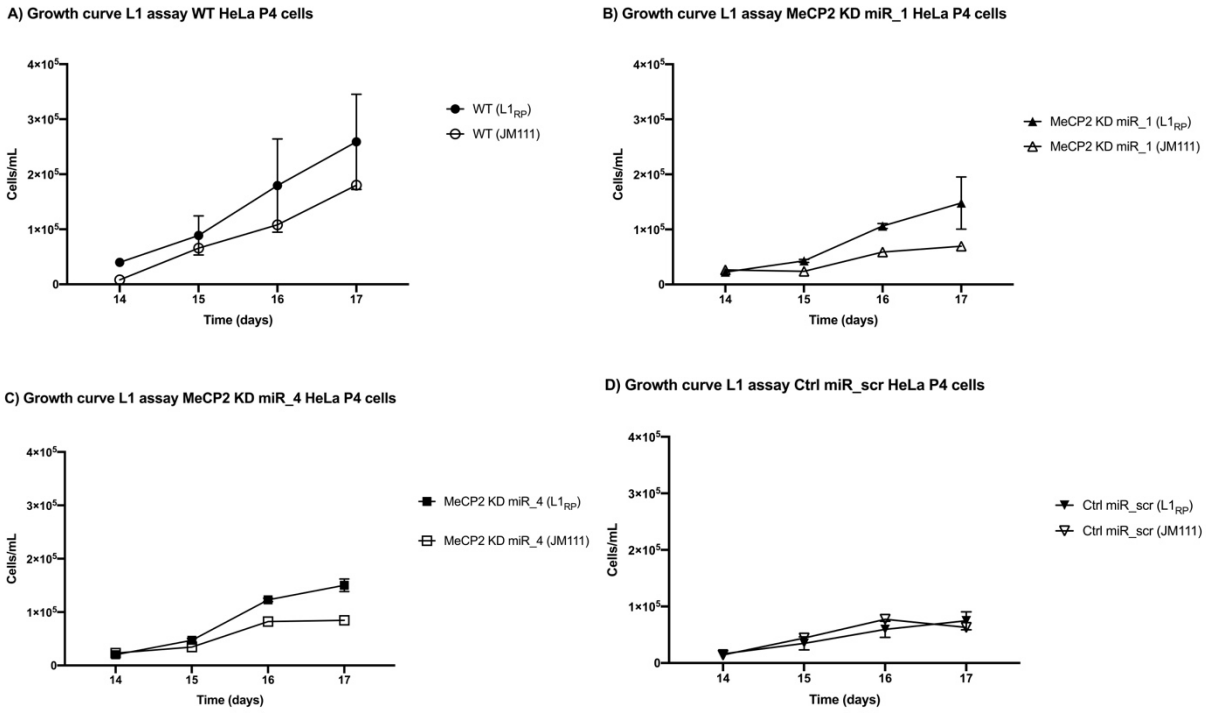


Figure 14: ***Growth curves of HeLa P4 cells during the L1 retrotransposition assay to determine retrotransposition induced cell toxicity****. Cells were plated at a density of 25 000 cells/well at day 13 of the L1 retrotransposition assay and were counted the following 4 days. Biological replicates of $L1_{RP}$-transfected cells were counted. Large variances in cell counts were measured. No clear difference in cell growth could be observed between $L1_{RP}$- and JM111-transfected cells. Error bars represent the SD.*

The L1 retrotransposition assay in HeLa P4 cells was not performant in our hands. Only very low retrotransposition rates were observed that lack biological relevance. We did not continue to work with HeLa P4 cells in the L1 retrotransposition assay.

4.1.4 Validation of the L1 retrotransposition assay using MeCP2 KD in HEK293T cells

Next, the L1 retrotransposition assay was validated in HEK293T cells by comparing WT and MeCP2 KD cells (Figure 15). High percentages of EGFP-positive cells were reached already on day 6 in $L1_{RP}$-transfected cells (Figure 15A). The percentage of EGFP-positive cells in the negative control JM111-transfected cells remained low in all cell lines. MeCP2 KD cells showed higher retrotransposition rates compared to WT cells. However, after day 6 the cells were in very poor condition. Repeating the assay resulted in only low percentages of EGFP-positive cells and the same poor cell condition after day 7 (Figure 15B). Surprisingly, also the cells transduced with a miRNA containing a scrambled control sequence showed increased retrotransposition rates. This will be discussed below.

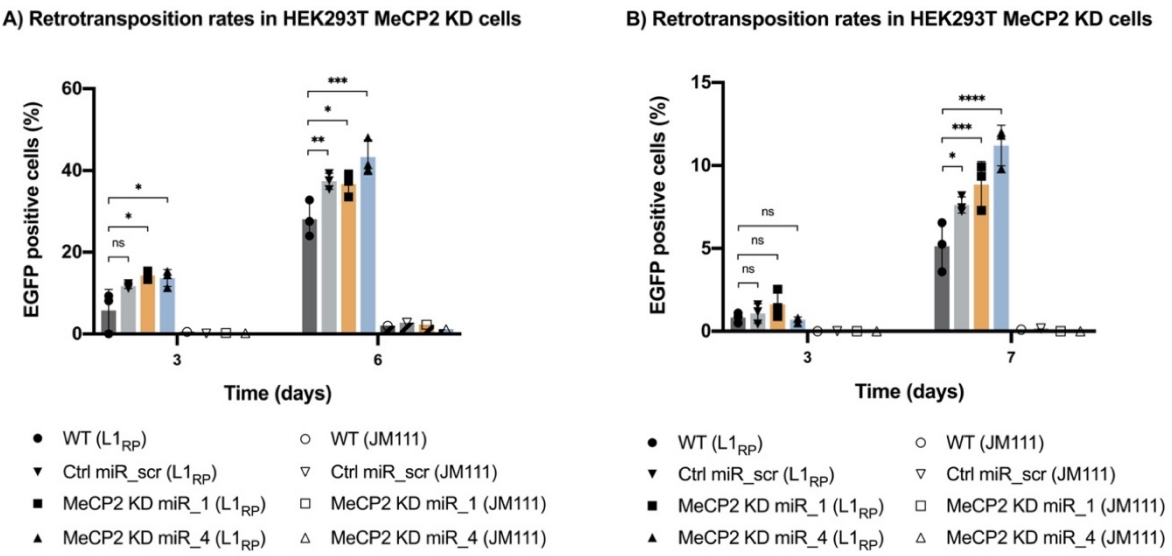


Figure 15: ***L1 retrotransposition rates in MeCP2 KD HEK293T cells.*** *A) Cells were transfected with the $L1_{RP}$ plasmid or with the negative control JM111 plasmid both containing an EGFP reporter and were kept under continuous puromycin selection from day 2 post-transfection onwards. Samples were taken on day 3 and day 6 post-transfection. EGFP-positive cells were detected by flow cytometry. B) The L1 assay was repeated using the same conditions. Samples were taken at day 3 and day 7 post-transfection. Transfections with the $L1_{RP}$-plasmid were performed in triplicate. Data represent a single test. Error bars represent the SD. Statistical analysis was done using two-way ANOVA followed by Dunnett's multiple comparison test vs. WT. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; ns, not significant.*

The assay was repeated with and without antibiotic selection to see whether the cell death occurred because the cells suffered to much from the puromycin selection (Figure 16). During the course of the experiment, we noticed that the transfected cells were very sensitive to trypsinization. This was presumably also the cause of cell death observed in the previous experiments. The same transfection mixes were used for both the selected and non-selected conditions and transfection efficiencies were measured 2 days post-transfection. When keeping the cells under continuous selection, high percentages of EGFP-positive cells were obtained in all $L1_{RP}$-transfected cell lines (Figure 16A). Retrotransposition rates in $L1_{RP}$-transfected MeCP2 KD cells were significantly increased compared to WT cells. Not selecting the transfected cells resulted in a peak percentage of EGFP-positive cells on day 6 and consequently a gradual decrease in retrotransposition rates (Figure 16B). This can be explained by the fact that transfected cells have a disadvantage and are overgrown by non-transfected cells. However, on day 6 a significant increase in retrotransposition rate can be observed in MeCP2 KD cells compared to WT cells indicating that differences between cell lines can be observed even without selecting the cells. It must be noted that this experiment was performed with MeCP2

KD cell lines that had a less potent MeCP2 KD (see also Figure 10A). It appears that the effect of MeCP2 KD on L1 retrotransposition can still be observed even when MeCP2 is not fully depleted.

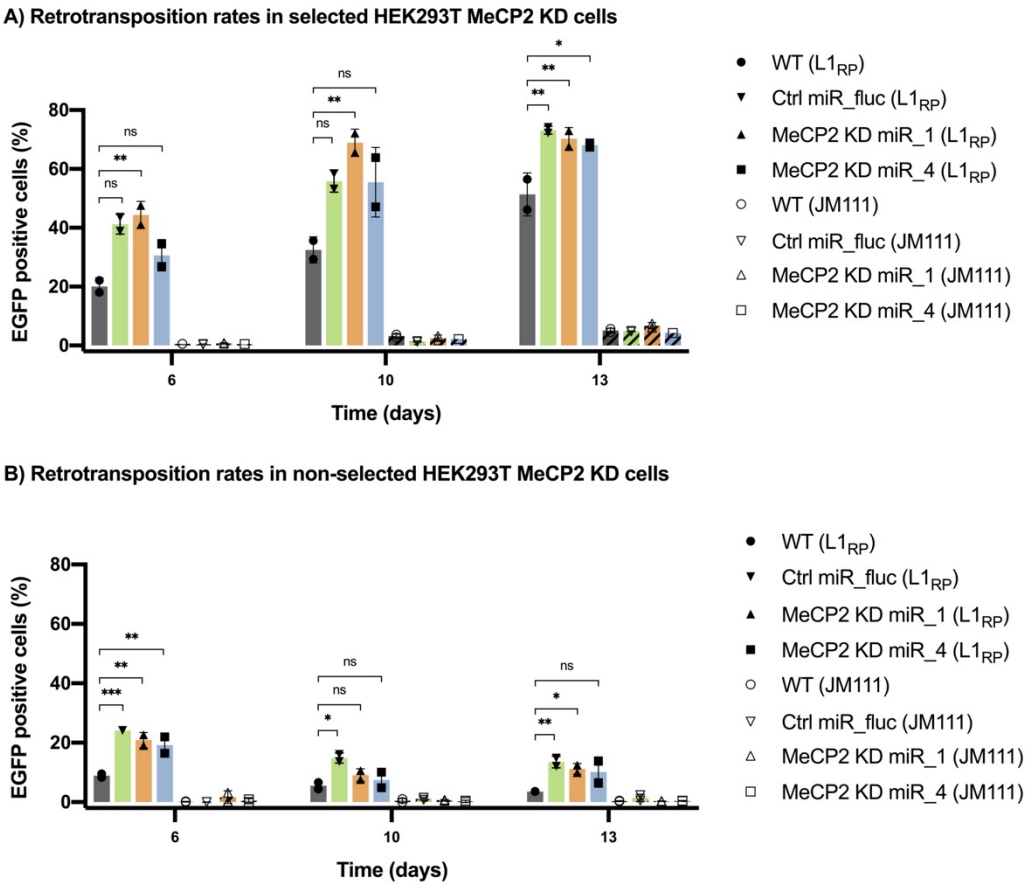


Figure 16: ***The L1 retrotransposition assay in HEK293T cells with or without puromycin selection****. WT and MeCP2 KD HEK293T cells were transfected with the retrotransposition competent $L1_{RP}$ plasmid or with the defective JM111 plasmid both containing an EGFP reporter. EGFP-positive cells were detected by flow cytometry. Cells were continuously kept under puromycin selection from day 2 onwards (A) or were not selected after transfection (B). Transfections were performed in duplicates. C) Transfection efficiencies were determined by co-transfecting the $L1_{RP}$ or JM111 plasmid with a GFP mock plasmid and measured 2 days post-transfection. Data represent a single test. Error bars represent the SD. Statistical analysis for selected cells was done using two-way ANOVA followed by Dunnett's multiple comparison test vs. WT. Statistical analysis for non-selected cells was done using mixed effects analysis followed by Dunnett's multiple comparisons test vs. WT. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns, not significant.*

During this experiment we also monitored the transfection efficiencies by co-transfecting the cells in parallel with the $L1_{RP}$ plasmid or the JM111 plasmid and with a GFP mock plasmid. (Figure 16C). Transfection efficiencies were measured 2 days post-transfection using flow cytometry to determine the percentage of GFP-positive cells. All cell lines showed a very high transfection efficiency between 80 % and 90 % for both the $L1_{RP}$ and the JM111 plasmid.

A problem we came across in the L1 retrotransposition assay was the increased retrotransposition rate observed in the control cell line transduced with vector carrying a non-targeting scrambled miRNA (see also Figure 15). We created a new control cell line transduced with a vector carrying a miRNA that targets Fluc. When performing the L1 retrotransposition assay, increased retrotransposition rates were repeatedly observed in the miR_fluc control cell line transfected with the $L1_{RP}$ plasmid (Figure 16 and 17A). This effect has been observed before when using non-targeting siRNA controls in the L1 retrotransposition assay (188). Therefore, rescue by reintroducing MeCP2 into the MeCP2 KD cells would be a better control. In this repeat MeCP2 KD cell lines showed again an increase in L1 retrotransposition rates compared to WT cells (Figure 17A). All cell lines showed very high transfection efficiencies exceeding 95 % for both the $L1_{RP}$ and the JM111 plasmid (Figure 17B).

In parallel to flow cytometry to detect EGFP expression, we used RT-qPCR to detect integrated EGFP copy numbers on genomic DNA (Figure 17C). On day 13, before reaching the plateau phase, the cells were trypsinized and genomic DNA was extracted and used as template for qPCR amplification. The primers flank the γ-globin intron that disrupts the EGFP reporter gene so that only the spliced and integrated EGFP sequence of 206 bp is amplified (178). The amplification product of an unspliced EGFP cassette is 1109 bp long and is too large to be amplified by RT-qPCR. The dissociation curve showed a single peak that corresponds to the amplification of a single product (Figure 17D). The results of RT-qPCR confirm the results obtained by flow cytometry. MeCP2 KD cells transfected with the $L1_{RP}$ plasmid showed an increase in EGFP genomic copy number compared to JM111 transfected WT cells. The recurrent problem with the miR_fluc control cell line was also observed with RT-qPCR.

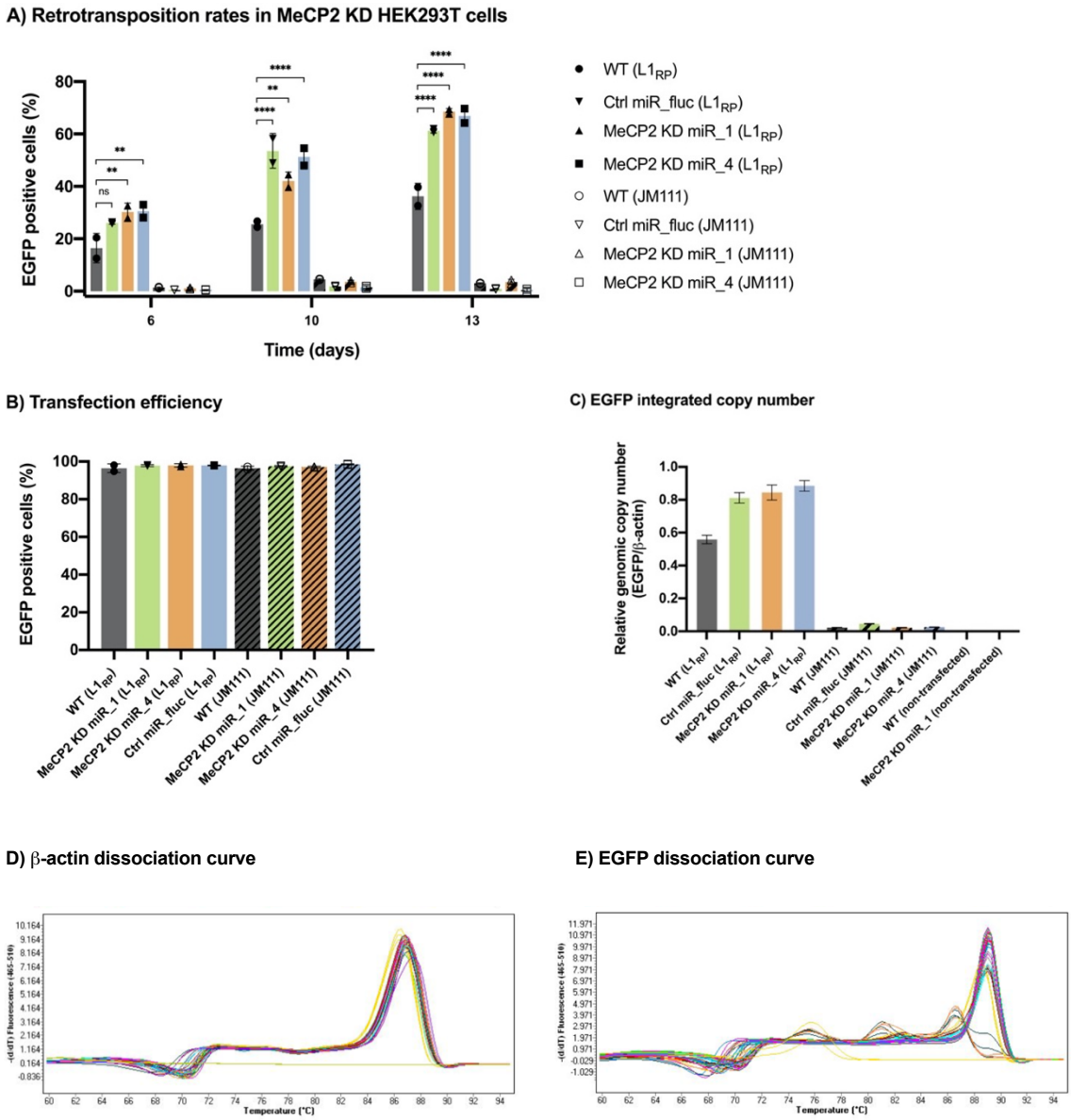


Figure 17: ***L1-EGFP reporter readout in MeCP2 KD HEK293T cell by flow cytometry or RT-qPCR.*** *WT and MeCP2 KD HEK293T cells were transfected with the retrotransposition competent $L1_{RP}$ plasmid or with the defective JM111 plasmid both containing an EGFP reporter cassette. Transfections were performed in duplicate. Cells were continuously kept under puromycin selection from day 2 onwards. A) EGFP-positive cells were detected by flow cytometry at multiple time points post-transfection. B) Transfection efficiency was determined by co-transfecting the $L1_{RP}$ or JM111 plasmid with a GFP mock plasmid and the percentage EGFP-positive cells was measured 2 days post-transfection. C) Integrated EGFP copy numbers were determined by RT-qPCR using EGFP primers flanking the γ-globin intron that interrupts the EGFP gene in the reporter cassette. RT-qPCR was performed in technical triplicates. D) Genomic copy numbers were normalized for β-actin. The dissociation curve of β-actin shows a single peak. E) The dissociation curve of EGFP shows a single peak indicating the amplification of the 206 bp intronless EGFP sequence. Data represent a single test. Error bars represent the SD. Statistical analysis was done using two-way ANOVA followed by Dunnett's multiple comparison test vs. WT. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; ns, not significant.*

These results confirm that HEK293T cells can be used to perform the L1 retrotransposition assay. The optimal EGFP readout of the assay by flow cytometry can be obtained from samples between day 6 and day 13 post-transfection when a linear increase in retrotransposition rates is observed. The assay was validated with two different MeCP2 KD cell lines and by using two different methods for EGFP readout: flow cytometry for EGFP expression and RT-qPCR for integrated EGFP copy numbers. We were not able to establish a control cell line with vectors carrying a non-targeting miRNA. This may be due to a problem with non-targeting controls as previously described for siRNAs in the L1 retrotransposition assay (188). The assay should be further validated with MeCP2 WT and mutant overexpression cell lines.

4.2 Role of LEDGF/p75 in L1 retrotransposition

LEDGF/p75 is an essential host factor for the integration of HIV-1. An earlier study by the host lab described the interaction between MeCP2 and LEDGF/p75 in a cancer cell line (189). We hypothesized that LEDGF/p75 is also a host factor of L1 retrotransposition in a direct or indirect manner. We used the optimized L1 retrotransposition assay for HEK293T cells to study the role of LEDGF/p75 in L1 retrotransposition. A previously established LEDGF/p75 KD cell line from the host lab was used in the assay to study the effect of LEDGF/p75 depletion on retrotransposition rates. Therefore, validation of the LEDGF/p75 KD cells by RT-qPCR prior to the assay was performed to confirm depletion of LEDGF/p75 mRNA (Figure 17).

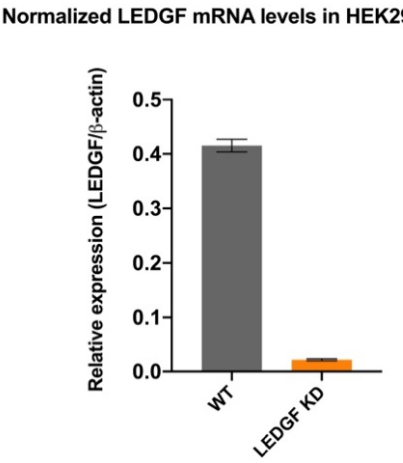


Figure 18: ***Validation of a LEDGF/p75 KD HEK293T cell line by RT-qPCR****. LEDGF/p75 KD cells express a miRNA that targets LEDGF/p75 mRNA transcripts; they were continuously kept under zeocin selection. LEDGF/p75 expression levels were determined by RT-qPCR and were normalized for β-actin expression levels. RT-qPCR was performed in technical triplicates. Data are presented as mean ± SD of two biological replicates.*

We compared WT and LEDGF/p75 KD HEK293T cells transfected with the $L1_{RP}$ plasmid and followed the transfection rates over time (Figure 18). The cells were selected with puromycin from day 2 post-transfection and were continuously kept under selection. Samples were taken at multiple time points after transfection to determine the percentage of EGFP-positive cells by

flow cytometry reflecting the retrotransposition rates. A decrease in retrotransposition rates was observed in LEDGF/p75 KD cells compared to WT. These results indicate that LEDGF/p75 has a direct or indirect effect as host factor of L1 retrotransposition.

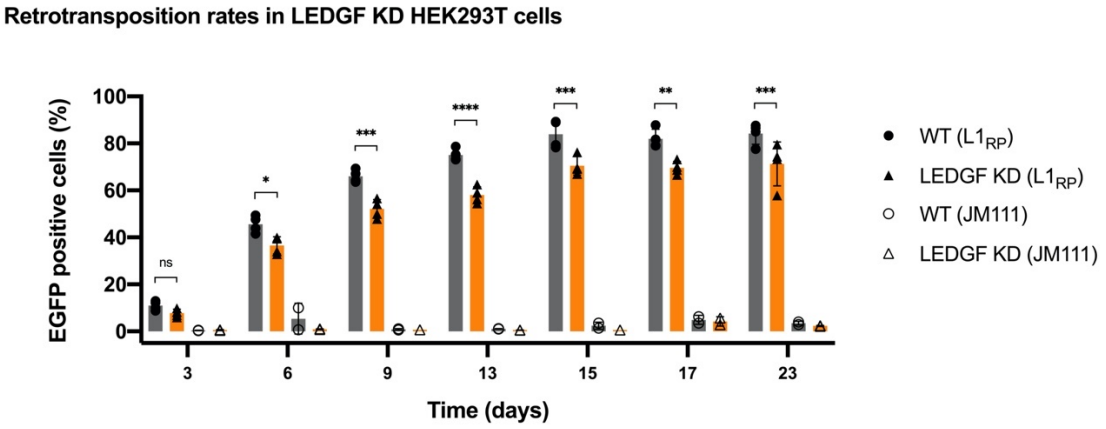


Figure 19: ***The L1 retrotransposition assay in LEDGF/p75 KD HEK293T cells****. WT and LEDGF/p75 KD cells were transfected with the retrotransposition competent $L1_{RP}$ or defective JM111 control plasmid both containing an EGFP reporter cassette. Cells were continuously kept under puromycin selection from day 2 onwards. EGFP-positive cells were detected by flow cytometry at multiple time points post-transfection. Data represent a single test. Error bars represent the SD. Statistical analysis was done using two-way ANOVA followed by Sidak's multiple comparison test. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; ns, not significant.*

These preliminary results suggest that LEDGF/p75 is a novel host factor that enables L1 retrotransposition in HEK293T cells. The assay should be repeated, and percentages of EGFP-positive cells should be normalized for transfection efficiency. An LEDGF/p75 overexpression cell line should be included as a control. We would expect that LEDGF/p75 overexpression rescues retrotransposition. It remains to be determined whether the observed decrease in retrotransposition rates in LEDGF/p75 KD cells is due to a direct or indirect effect of LEDGF/p75.

5 Discussion

5.1 HEK293T cells support high frequency L1 retrotransposition

We confirm that HEK293T cells are capable of supporting high frequency L1 retrotransposition from an episomal plasmid expressing a L1-reporter construct. For the L1 retrotransposition assay we used the engineered $L1_{RP}$ element with an EGFP-reporter cassette to study host factors of L1 retrotransposition (177). We showed that HEK293T cells can be used in the L1 retrotransposition assay and observed higher retrotransposition rates in HEK293T cells than previously described (106, 190). We observed up to 84% EGFP-positive cells using flow cytometry for EGFP readout, which reflects high retrotransposition rates in HEK293T cells. After transfection with the $L1_{RP}$-EGFP plasmid L1 retrotransposition rates increased over time and reached a plateau phase after 15 days. During the linear increase in retrotransposition rates, an optimal time frame for EGFP readout by flow cytometry is created to compare conditions. We validated the L1 retrotransposition assay using MeCP2 KD cells and saw an expected increase in L1 retrotransposition (Figure 15, 16 and 17). MeCP2 is well known to regulate transcription of L1 through binding of the L1 promoter, thereby making it an excellent target to validate the assay. Despite the decreasing potency of MeCP2 gene silencing in the MeCP2 KD cell lines over time, higher retrotransposition rates could still be observed in MeCP2 depleted cells (Figure 16 and 17). Possibly a larger effect could be observed with a more potent MeCP2 KD. Monoclonal cell lines would be recommended when performing the assay in the future.

We measured transfection efficiencies during the course of two experiments with MeCP2 KD HEK293T cells (Figure 16C and 17B). We could observe that transduction of miRNAs had little or no impact on the transfection efficiencies of the $L1_{RP}$-EGFP plasmid. In case of different transfection efficiencies across cell lines and plasmids, normalization of the percentage EGFP-positive cells is required. Small changes in transfection mixture can affect transfection efficiency and can eventually impact the result of the L1 retrotransposition assay. According to the original protocol in HeLa cells, GFP expression to measure transfection efficiency should not interfere with the EGFP from the reporter plasmids 1 to 3 days post-transfection since a delay of at least 48h was observed in the detection of retrotransposition (175). However, we observed a percentage of EGFP-positive cells of around 10 % in $L1_{RP}$-transfected cells at day 3 post-transfection in HEK293T cells (Figure 12, 15 and 19). In cells with higher retrotransposition rates, such as HEK293T cells, the accumulation of EGFP to a detectable

threshold may be quicker, leading to a shorter delay in EGFP expression. If this is the case, EGFP expression could interfere with GFP expression to measure transfection efficiency. A better solution would be to measure transfection efficiency by co-transfecting the cells with a plasmid encoding another fluorescent reporter that does not interfere with EGFP, such as mCherry.

Based on our results, HeLa P4 cells do not seem to support high frequency retrotransposition from the $L1_{RP}$-EGFP plasmid (Figure 13). The stable transformation of HeLa P4 cells for the expression of CD4 is not necessarily the cause of this, although this possibility may not be excluded. A more probable explanation can be found in the variation between HeLa populations in different laboratories (191). L1 retrotransposition rates appear to depend on clonal variation in HeLa cells with some populations supporting high levels of retrotransposition and other populations completely lacking the ability to support L1 retrotransposition. These differences possibly arise from passaging mixed populations of HeLa cells and selecting for genetic and epigenetic changes that favor or restrict L1 retrotransposition (191). However, the possibility that the observed differences in HeLa cell lines supporting retrotransposition may be caused by more reversible epigenetic changes cannot be excluded (191). Additionally, it was also observed that the retrotransposition potential fluctuates over time in HeLa cells when passaging the cells, possibly also due to the polyclonal nature of the cell lines since HeLa cells are transformed cells with an unstable genome (191). As a result, retrotransposition rates can vary between experiments. A proper reference control to normalize the retrotransposition rates would be recommended to standardize between experimental variation. Nevertheless, HeLa cells are commonly used in the L1 retrotransposition assay. Two subcloned HeLa strains, HeLa-JVM and HeLa-HA are known to support L1 retrotransposition and are therefore mostly used to perform the L1 retrotransposition assay. HeLa cells do not express high levels of endogenous L1 elements, but certain subclones appear to support episomal L1 expression more robustly than the expression of endogenous L1 elements (176, 192, 193). Epigenetic silencing of endogenous L1 promoters may be the cause of this discrepancy since L1 promoters are sufficient to support retrotransposition in HeLa cells (176). Taken together, these observations in HeLa cells also imply that similar clonal variation may be found in different populations of HEK293T cells.

In the original protocol of the L1 retrotransposition assay with the human $L1_{RE2}$ and a NeoR reporter cassette, high retrotransposition frequencies were reported in HeLa cells (176). They

compared it to the human $L1_{1.2}$ element with a NeoR reporter cassette and observed lower retrotransposition rates in cells transfected with the $L1_{1.2}$ plasmid. Retrotransposition frequencies were dependent on the type of L1 element in addition to the cell type. The low retrotransposition rates we observed in HeLa cells may also be the result of the $L1_{RP}$ element that only has low frequency retrotransposition levels in HeLa cells. Low retrotransposition rates were also observed in the original protocol with the $L1_{RP}$-EGFP plasmid (177). From a personal communication with Prof. J.V. Moran (University of Michigan), it appears that $L1_{RE3}$ is now used more often in HeLa cells. $L1_{RE3}$ is the most active human L1 identified to date and shows high retrotransposition frequencies in the L1 retrotransposition assay with HeLa cells (194).

During the course of the L1 retrotransposition assay in HeLa P4 cells, we started to investigate whether increased retrotransposition rates in $L1_{RP}$-transfected cells induced retrotransposition-induced cell toxicity compared to JM111-transfected cells by counting cells in parallel to the L1 retrotransposition assay (Figure 14). No difference in cell toxicity could have been expected from a retrospective point of view considering the L1 assay in HeLa P4 cells was not performant in our hands. Indeed, no clear difference in cell proliferation rates was observed between cells transfected with the retrotransposition competent $L1_{RP}$ plasmid and cells transfected with the negative control JM111 plasmid. Data were too inconclusive to draw any conclusion on retrotransposition-induced cell toxicity. Biological replicates showed large variability in cell counts. Other methods may be more suitable to measure retrotransposition-induced cell toxicity during the L1 retrotransposition assay such as live-cell analysis. An optimized method for measuring cell toxicity during the L1 retrotransposition assay may be useful to determine the role of L1 host factors in retrotransposition-induced cell toxicity. We did not investigate this. However, it remains interesting for future research whether increased retrotransposition rates may induce cell toxicity.

We also observed that non-targeting miRNAs elicit an apparent effect on retrotransposition rates. This raises concern about the use of non-targeting controls in the L1 retrotransposition assay as it would not be appropriate to use these as normalizing controls. We studied two non-targeting controls, namely a miRNA with a scrambled sequence and a miRNA targeting Fluc, and saw a strong increase in retrotransposition rates in cells transduced with these miRNAs compared to non-transduced cells (Figure 15, 16 and 17). A similar observation was made in a study using non-targeting control siRNAs to study L1 retrotransposition (188). Off-target effects are a major drawback of RNAi and non-targeting controls are often used to correct for

this based on the assumption that non-targeting controls have equivalent off-target effects. However, it appears that problems arise when these two methods are used together. Without a proper control, we cannot exclude the possibility that the observed differences in retrotransposition rates in cells with a targeting miRNA are also caused by off-target effects. To confirm that the effect on L1 retrotransposition is caused by KD of the gene of interest, the assay needs to be performed with rescue cell lines as a control.

The L1 retrotransposition assay is a reliable method to study L1 retrotransposition and is widely used in the field of TEs. Although EGFP readout through flow cytometry is a convenient method to asses retrotransposition rates, some issues make this method variable. For the EGFP to be detected, the entire and quite large EGFP cassette must be reverse transcribed and integrated into the genomic DNA before EGFP can be expressed. Various host defense mechanisms may prevent this; mainly 5' truncation of the L1-EGFP can pose a problem. EGFP may also not be expressed due to silencing of the CMV promoter that drives EGFP, or due to integration of EGFP into a dense chromatin structure that does not allow expression (195). Additionally, EGFP needs to accumulate to reach the flow cytometry detection threshold. For these reasons, quantification of integrated EGFP copies by RT-qPCR may be a more sensitive method to detect retrotransposition events from an episomal L1 plasmid. RT-qPCR can detect integrated EGFP in genomic regions where transcription is repressed. It can also detect truncated and/or inversed EGFP that cannot be transcribed but that still contain the primer binding sites. Flow cytometry can determine the percentage of cells in which at least one retrotransposition event has taken place, whereas RT-qPCR can detect multiple retrotransposition events in a cell and can determine absolute integrated copy numbers, which makes this method of detection more quantitative. However, a major disadvantage of using RT-qPCR for EGFP readout is that you cannot measure transfection efficiency and consequently you cannot normalize the results.

We compared both methods for EGFP readout and obtained similar results (Figure 17). An increase in retrotransposition rate was observed in MeCP2 KD cells compared to WT both in flow cytometry data and in RT-qPCR data. Both methods worked well in our hand. The advantages and disadvantages for the methods of EGFP readout need to be weighed out and depend on whether the experimental condition influences the readout. For example, when treatment and control are being compared, the treatment should not influence the CMV promoter since this could influence EGFP readout by flow cytometry. The same applies to host

factors that influence the CMV promoter. When a host factor that interferes with the CMV promoter is overexpressed, higher percentages of EGFP-positive cells will be obtained while there is not necessarily an increase in retrotransposition rates. In this case it would be recommended to use RT-qPCR for EGFP readout.

5.2 L1 retrotransposition is decreased in LEDGF/p75 depleted cells

We used the optimized conditions of the L1 retrotransposition assay in HEK293T cells to investigate the effect of LEDGF/p75 depletion on L1 retrotransposition. Retrotransposition rates in LEDGF/p75 depleted cells were decreased compared to WT cells (Figure 19). These preliminary results suggest that LEDGF/p75 is a host factor of L1 retrotransposition in a direct or indirect way. However, we must keep in mind that the decrease of L1 retrotransposition observed in LEDGF/p75 depleted cells may be caused by off-target effects of the miRNA as described earlier.

One of the possibilities is that LEDGF/p75 interferes with the repressive function of MeCP2. LEDGF/p75 was originally discovered as transcriptional activator that interacts with the general transcription machinery (196). LEDGF/p75 can bind DNA through a tripartite element that consists of a nuclear localization signal and two copies of the AT-hook motif (197). The PWWP domain of LEDGF/p75 is known to interact with the TRD of MeCP2 (189, 198). While MeCP2 is primarily known as a transcriptional repressor, MeCP2 can also act as a transcriptional activator when it interacts with certain co-activators such as CREB (50). Transcriptional regulation by a LEDGF/p75/MeCP2 complex could not only have implications for the regulation of L1, but also for the regulation of a wide range of other MeCP2 regulated genes. LEDGF/p75 is a stress response protein that is thought to transcriptionally activate genes to promote cell survival under stress (189). Interestingly, studies showed that TEs in other species can be activated by stress and that stress can redirect integration of TEs to alternative sites in the genome (44). This is consistent with the findings that L1 is responsive to environmental factors such as stress (28).

LEDGF/p75 is an essential host factor that tethers the HIV-1 integrase to the host genome (199). If we make the analogy with HIV-1, LEDGF/p75 may have a similar function for the integration of L1 retrotransposons. Retrotransposons have evolved to limit damage to the host. Since the integration into the host genome has mostly little or no impact, it was suggested that this specific integration is not only the result of the nicking specificity of L1 endonuclease, but that

integration site specificity is also guided by a tethering host factor (200). This hypothesis was supported by a study that suggested that L1 endonuclease determines the integration site selection based on the nicking specificity of the AT-consensus sequence, but that additional host factors could also influence the integration site selection (201). However, this hypothesis is contradicted by the Darwinian point of view L1 integration sites are mainly determined by the post-integration process of selection leading to the removal of deleterious L1 insertions (202, 203). L1 endonuclease only has a weak specificity to cleave AT-consensus sequences and as a result L1 elements are found interspersed throughout the genome. These findings contradict the presence of a mechanism that tethers L1 to specific sites in the genome (44).

LEDGF/p75 is part of the homologous recombination (HR) DNA repair machinery that operates during the S and G2 phases of the cell cycle (204). When DNA damage occurs, LEDGF/p75 binds chromatin through its PWWP domain and recruits the DNA damage response protein CtIP to DSBs. During HIV-1 integration, a double stranded virus-host DNA intermediate is created, but the mechanism by which the virus-host DNA intermediate is subsequently repaired remains unclear. It was speculated the LEDGF/p75 HR DNA repair machinery would repair DSBs created by the insertion of HIV-1 into the genomic DNA (204). This would mean that LEDGF/p75 does not only tether HIV-1 to the genomic DNA but also that LEDGF/p75 subsequently recruits CtIP to repair the DBS caused by integration of HIV-1. Therefore, a fraction of the virus-host DNA intermediate would have to enter the S phase unrepaired. Earlier studies already suggested that DNA repair host factors may be implicated in the post-integration process of HIV-1 (205-207). In analogy with HIV-1, DSBs are created when L1 integrates into genomic DNA. After TPRT a double stranded L1-host DNA intermediate is created that needs to be integrated into the genomic DNA. As for HIV-1, the mechanism by which the L1-host DNA gap is repaired remains unknown. In a similar way to HIV-1 we could speculate that the LEDGF/p75 HR DNA repair machinery repairs the L1-host DNA gap. In this case LEDGF/p75 would be a host factor that facilitates the integration of L1 into the host genome by recruiting HR DNA damage repair factors. Consistent with this hypothesis, L1 retrotransposition has shown to occur in dividing cells, mainly during the S phase of the cell cycle, the cell cycle phase in which the HR DNA repair machinery is active (87). However, various DSB repair factors that are active during the S and G2 phase of the cell cycle were found to be potent inhibitors of L1 retrotransposition (208). The authors suggested that DNA end resection complexes form a physical barrier to inhibit TPRT. Many ambiguities

still exist on the role of DNA damage response factors in L1 retrotransposition. Further studies will be needed to evaluate their role in L1 retrotransposition especially during cell division. This will also have implications for certain cancers in which HR pathways are altered and an increase in L1 retrotransposition is observed.

5.3 Limitations of the study and future perspectives

Certain limitations of the study hamper strong conclusions about the role of LEDGF/p75 in L1 retrotransposition. First, due to time limitations, we were not able to perform the assay with the appropriate rescue cell lines as controls. In future experiments, it will be important to include these as a control to exclude that observed differences are caused by off-target effects. Second, transfection efficiency should always be monitored. A suitable detection method in HEK293T cells still needs to be optimized. We propose a detection method using a different fluorescent protein to avoid interference with the EGFP reporter. Third, working with monoclonal cell lines will reduce the variability between experiments. This will not only reduce the variability in retrotransposition rates within a population of cells, but this will also allow to maintain a potent gene silencing when working with miRNA transduced cells. Fourth, it may be more relevant to study L1 retrotransposition in neuronal cell lines. NPCs and neuroblastoma cells have shown to support L1 retrotransposition and could be used in the L1 retrotransposition assay (47, 178). Some regulators of L1 function in a cell-type specific manner, making neuronal cell lines more relevant to study L1 retrotransposition in biology and pathology of the human brain. When working with other cell lines such as neuronal cell lines, other and more active L1 elements may have to be taken into consideration depending on how well the cell line supports retrotransposition. Various other L1 elements exist that have been used in the L1 retrotransposition assay such as the human $L1_{RE3}$ or the synthetic $L1_{ORFeus}$ (209).

Since the L1 retrotransposition assay is a functional assay, it may also be used in the future to study the functional effect of MeCP2 mutants. This will help us understand the molecular function of MeCP2 and the underlying pathogenic implications of MeCP2 mutants in RTT. Furthermore, future studies will have to investigate whether LEDGF/p75 exerts a direct or indirect effect on L1 retrotransposition. It would be interesting to investigate whether the effect on L1 retrotransposition is caused by an interaction with MeCP2 or by a MeCP2-independent mechanism.

References

1. McClintock B. Controlling elements and the gene. Cold Spring Harb Symp Quant Biol. 1956;21:197-216.
2. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. Nat Rev Genet. 2009;10(10):691-703.
3. Garcia Perez JL. Transposons and retrotransposons : methods and protocols. Springer. 2016.
4. Singer T, McConnell MJ, Marchetto MC, Coufal NG, Gage FH. LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? Trends Neurosci. 2010;33(8):345-54.
5. Thomas CA, Paquola AC, Muotri AR. LINE-1 retrotransposition in the nervous system. Annu Rev Cell Dev Biol. 2012;28:555-73.
6. Suarez NA, Macia A, Muotri AR. LINE-1 retrotransposons in healthy and diseased human brain. Dev Neurobiol. 2018;78(5):434-55.
7. Thomas CA, Muotri AR. LINE-1: creators of neuronal diversity. Front Biosci (Elite Ed). 2012;4:1663-8.
8. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860-921.
9. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, et al. Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci U S A. 2003;100(9):5280-5.
10. Human retrotransposons in health and disease. Springer. 2019.
11. Muotri AR, Marchetto MC, Coufal NG, Gage FH. The necessary junk: new functions for transposable elements. Hum Mol Genet. 2007;16 Spec No. 2:R159-67.
12. Ohshima K, Hattori M, Yada T, Gojobori T, Sakaki Y, Okada N. Whole-genome screening indicates a possible burst of formation of processed pseudogenes and Alu repeats by particular L1 subfamilies in ancestral primates. Genome Biol. 2003;4(11):R74.
13. Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, Gerton GL, et al. A mouse model of human L1 retrotransposition. Nat Genet. 2002;32(4):655-60.
14. Lazaros L, Kitsou C, Kostoulas C, Bellou S, Hatzi E, Ladias P, et al. Retrotransposon expression and incorporation of cloned human and mouse retroelements in human spermatozoa. Fertil Steril. 2017;107(3):821-30.
15. Malki S, van der Heijden GW, O'Donnell KA, Martin SL, Bortvin A. A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. Dev Cell. 2014;29(5):521-33.
16. Wang L, Dou K, Moon S, Tan FJ, Zhang ZZ. Hijacking Oogenesis Enables Massive Propagation of LINE and Retroviral Transposons. Cell. 2018;174(5):1082-94 e12.
17. Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, et al. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. Genes & development. 2009;23(11):1303-12.
18. Garcia-Perez JL, Marchetto MC, Muotri AR, Coufal NG, Gage FH, O'Shea KS, et al. LINE-1 retrotransposition in human embryonic stem cells. Hum Mol Genet. 2007;16(13):1569-77.
19. van den Hurk JAJM, Meij IC, del Carmen Seleme M, Kano H, Nikopoulos K, Hoefsloot LH, et al. L1 retrotransposition can occur early in human embryonic development. Human Molecular Genetics. 2007;16(13):1587-92.
20. De S. Somatic mosaicism in healthy human tissues. Trends Genet. 2011;27(6):217-23.
21. Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, et al. L1 retrotransposition in human neural progenitor cells. Nature. 2009;460(7259):1127-31.
22. Macia A, Widmann TJ, Heras SR, Ayllon V, Sanchez L, Benkaddour-Boumzaouad M, et al. Engineered LINE-1 retrotransposition in nondividing human neurons. Genome Res. 2017;27(3):335-48.

23. Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature. 2005;435(7044):903-10.
24. Upton KR, Gerhardt DJ, Jesuadian JS, Richardson SR, Sanchez-Luque FJ, Bodea GO, et al. Ubiquitous L1 mosaicism in hippocampal neurons. Cell. 2015;161(2):228-39.
25. Han JS, Szak ST, Boeke JD. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. Nature. 2004;429(6989):268-74.
26. Bachiller S, Del-Pozo-Martin Y, Carrion AM. L1 retrotransposition alters the hippocampal genomic landscape enabling memory formation. Brain Behav Immun. 2017;64:65-70.
27. Muotri AR, Zhao C, Marchetto MC, Gage FH. Environmental influence on L1 retrotransposons in the adult hippocampus. Hippocampus. 2009;19(10):1002-7.
28. Ponomarev I, Rau V, Eger EI, Harris RA, Fanselow MS. Amygdala transcriptome and cellular mechanisms underlying stress-enhanced fear learning in a rat model of posttraumatic stress disorder. Neuropsychopharmacology. 2010;35(6):1402-11.
29. Hunter RG, Murakami G, Dewell S, Seligsohn M, Baker ME, Datson NA, et al. Acute stress and hippocampal histone H3 lysine 9 trimethylation, a retrotransposon silencing response. Proc Natl Acad Sci U S A. 2012;109(43):17657-62.
30. Ponomarev I, Wang S, Zhang L, Harris RA, Mayfield RD. Gene coexpression networks in human brain identify epigenetic modifications in alcohol dependence. J Neurosci. 2012;32(5):1884-97.
31. Nigumann P, Redik K, Matlik K, Speek M. Many human genes are transcribed from the antisense promoter of L1 retrotransposon. Genomics. 2002;79(5):628-34.
32. Kolosha VO, Martin SL. In vitro properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. Proc Natl Acad Sci U S A. 1997;94(19):10155-60.
33. Mathias SL, Scott AF, Kazazian HH, Jr., Boeke JD, Gabriel A. Reverse transcriptase encoded by a human transposable element. Science. 1991;254(5039):1808-10.
34. Feng Q, Moran JV, Kazazian HH, Jr., Boeke JD. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell. 1996;87(5):905-16.
35. Denli AM, Narvaiza I, Kerman BE, Pena M, Benner C, Marchetto MC, et al. Primate-specific ORF0 contributes to retrotransposon-mediated diversity. Cell. 2015;163(3):583-93.
36. Yang N, Kazazian HH, Jr. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat Struct Mol Biol. 2006;13(9):763-71.
37. Alisch RS, Garcia-Perez JL, Muotri AR, Gage FH, Moran JV. Unconventional translation of mammalian LINE-1 retrotransposons. Genes Dev. 2006;20(2):210-24.
38. Kulpa DA, Moran JV. Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. Hum Mol Genet. 2005;14(21):3237-48.
39. Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, et al. Human L1 retrotransposition: cis preference versus trans complementation. Mol Cell Biol. 2001;21(4):1429-39.
40. Garcia-Perez JL, Doucet AJ, Bucheton A, Moran JV, Gilbert N. Distinct mechanisms for trans-mediated mobilization of cellular RNAs by the LINE-1 reverse transcriptase. Genome Res. 2007;17(5):602-11.
41. Dai L, Taylor MS, O'Donnell KA, Boeke JD. Poly(A) binding protein C1 is essential for efficient L1 retrotransposition and affects L1 RNP formation. Mol Cell Biol. 2012;32(21):4323-36.
42. Cost GJ, Feng Q, Jacquier A, Boeke JD. Human L1 element target-primed reverse transcription in vitro. EMBO J. 2002;21(21):5899-910.
43. Ostertag EM, Kazazian HH, Jr. Biology of mammalian L1 retrotransposons. Annu Rev Genet. 2001;35:501-38.

44. Levin HL, Moran JV. Dynamic interactions between transposable elements and their hosts. Nat Rev Genet. 2011;12(9):615-27.
45. Mita P, Boeke JD. How retrotransposons shape genome regulation. Curr Opin Genet Dev. 2016;37:90-100.
46. Hatanaka Y, Inoue K, Oikawa M, Kamimura S, Ogonuki N, Kodama EN, et al. Histone chaperone CAF-1 mediates repressive histone modifications to protect preimplantation mouse embryos from endogenous retrotransposons. Proc Natl Acad Sci U S A. 2015;112(47):14641-6.
47. Muotri AR, Marchetto MC, Coufal NG, Oefner R, Yeo G, Nakashima K, et al. L1 retrotransposition in neurons is modulated by MeCP2. Nature. 2010;468(7322):443-6.
48. Nan X, Campoy FJ, Bird A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell. 1997;88(4):471-81.
49. Yu F, Zingler N, Schumann G, Stratling WH. Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic Acids Res. 2001;29(21):4493-501.
50. Della Ragione F, Vacca M, Fioriniello S, Pepe G, D'Esposito M. MECP2, a multi-talented modulator of chromatin architecture. Brief Funct Genomics. 2016;15(6):420-31.
51. Skene PJ, Illingworth RS, Webb S, Kerr AR, James KD, Turner DJ, et al. Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. Mol Cell. 2010;37(4):457-68.
52. Kriaucionis S, Bird A. The major form of MeCP2 has a novel N-terminus generated by alternative splicing. Nucleic Acids Res. 2004;32(5):1818-23.
53. Lewis JD, Meehan RR, Henzel WJ, Maurer-Fogy I, Jeppesen P, Klein F, et al. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell. 1992;69(6):905-14.
54. Saleh A, Macia A, Muotri AR. Transposable Elements, Inflammation, and Neurological Disease. Front Neurol. 2019;10:894.
55. Castro-Diaz N, Ecco G, Coluccio A, Kapopoulou A, Yazdanpanah B, Friedli M, et al. Evolutionally dynamic L1 regulation in embryonic stem cells. Genes Dev. 2014;28(13):1397-409.
56. Kuwabara T, Hsieh J, Muotri A, Yeo G, Warashina M, Lie DC, et al. Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. Nat Neurosci. 2009;12(9):1097-105.
57. Pizarro JG, Cristofari G. Post-Transcriptional Control of LINE-1 Retrotransposition by Cellular Host Factors in Somatic Cells. Front Cell Dev Biol. 2016;4:14.
58. Belancio VP, Hedges DJ, Deininger P. LINE-1 RNA splicing and influences on mammalian gene expression. Nucleic Acids Res. 2006;34(5):1512-21.
59. Belancio VP, Roy-Engel AM, Deininger P. The impact of multiple splice sites in human L1 elements. Gene. 2008;411(1-2):38-45.
60. Goodier JL. Restricting retrotransposons: a review. Mob DNA. 2016;7:16.
61. Toth KF, Pezic D, Stuwe E, Webster A. The piRNA Pathway Guards the Germline Genome Against Transposable Elements. Adv Exp Med Biol. 2016;886:51-77.
62. Russell SJ, Stalker L, LaMarre J. PIWIs, piRNAs and Retrotransposons: Complex battles during reprogramming in gametes and early embryos. Reprod Domest Anim. 2017;52 Suppl 4:28-38.
63. Roberts JT, Cardin SE, Borchert GM. Burgeoning evidence indicates that microRNAs were initially formed from transposable element sequences. Mob Genet Elements. 2014;4:e29255.
64. Khazina E, Truffault V, Buttner R, Schmidt S, Coles M, Weichenrieder O. Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. Nat Struct Mol Biol. 2011;18(9):1006-14.

65. Taylor MS, LaCava J, Mita P, Molloy KR, Huang CR, Li D, et al. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. Cell. 2013;155(5):1034-48.
66. Moldovan JB, Moran JV. The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition. PLoS Genet. 2015;11(5):e1005121.
67. Goodier JL, Pereira GC, Cheung LE, Rose RJ, Kazazian HH, Jr. The Broad-Spectrum Antiviral Protein ZAP Restricts Human Retrotransposition. PLoS Genet. 2015;11(5):e1005252.
68. Arjan-Odedra S, Swanson CM, Sherer NM, Wolinsky SM, Malim MH. Endogenous MOV10 inhibits the retrotransposition of endogenous retroelements but not the replication of exogenous retroviruses. Retrovirology. 2012;9:53.
69. Goodier JL, Cheung LE, Kazazian HH, Jr. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. PLoS Genet. 2012;8(10):e1002941.
70. Li X, Zhang J, Jia R, Cheng V, Xu X, Qiao W, et al. The MOV10 helicase inhibits LINE-1 mobility. J Biol Chem. 2013;288(29):21148-60.
71. Gregersen LH, Schueler M, Munschauer M, Mastrobuoni G, Chen W, Kempa S, et al. MOV10 Is a 5' to 3' RNA helicase contributing to UPF1 mRNA target degradation by translocation along 3' UTRs. Mol Cell. 2014;54(4):573-85.
72. Meister G, Landthaler M, Peters L, Chen PY, Urlaub H, Luhrmann R, et al. Identification of novel argonaute-associated proteins. Curr Biol. 2005;15(23):2149-55.
73. Goodier JL, Zhang L, Vetter MR, Kazazian HH, Jr. 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. 2007;27(18):6469-83.
74. Warkocki Z, Krawczyk PS, Adamska D, Bijata K, Garcia-Perez JL, Dziembowski A. Uridylation by TUT4/7 Restricts Retrotransposition of Human LINE-1s. Cell. 2018;174(6):1537-48 e29.
75. Frost RJ, Hamra FK, Richardson JA, Qi X, Bassel-Duby R, Olson EN. MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. Proc Natl Acad Sci U S A. 2010;107(26):11847-52.
76. Zheng K, Xiol J, Reuter M, Eckardt S, Leu NA, McLaughlin KJ, et al. Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway. Proc Natl Acad Sci U S A. 2010;107(26):11841-6.
77. Chen H, Lilley CE, Yu Q, Lee DV, Chou J, Narvaiza I, et al. APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. Curr Biol. 2006;16(5):480-5.
78. Esnault C, Heidmann O, Delebecque F, Dewannieux M, Ribet D, Hance AJ, et al. APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. Nature. 2005;433(7024):430-3.
79. Muckenfuss H, Hamdorf M, Held U, Perkovic M, Lower J, Cichutek K, et al. APOBEC3 proteins inhibit human LINE-1 retrotransposition. J Biol Chem. 2006;281(31):22161-72.
80. Bogerd HP, Wiegand HL, Hulme AE, Garcia-Perez JL, O'Shea KS, Moran JV, et al. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. Proc Natl Acad Sci U S A. 2006;103(23):8780-5.
81. Kinomoto M, Kanno T, Shimura M, Ishizaka Y, Kojima A, Kurata T, et al. All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. Nucleic Acids Res. 2007;35(9):2955-64.
82. Schumann GG. APOBEC3 proteins: major players in intracellular defence against LINE-1-mediated retrotransposition. Biochem Soc Trans. 2007;35(Pt 3):637-42.
83. Stenglein MD, Harris RS. APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. J Biol Chem. 2006;281(25):16837-41.

84. Goila-Gaur R, Strebel K. HIV-1 Vif, APOBEC, and intrinsic immunity. Retrovirology. 2008;5:51.
85. Liu C, Zhang X, Huang F, Yang B, Li J, Liu B, et al. APOBEC3G inhibits microRNA-mediated repression of translation by interfering with the interaction between Argonaute-2 and MOV10. J Biol Chem. 2012;287(35):29373-83.
86. Xie Y, Mates L, Ivics Z, Izsvak Z, Martin SL, An W. Cell division promotes efficient retrotransposition in a stable L1 reporter cell line. Mob DNA. 2013;4(1):10.
87. Mita P, Wudzinska A, Sun X, Andrade J, Nayak S, Kahler DJ, et al. LINE-1 protein localization and functional dynamics during the cell cycle. Elife. 2018;7.
88. Idica A, Sevrioukov EA, Zisoulis DG, Hamdorf M, Daugaard I, Kadandale P, et al. MicroRNA miR-128 represses LINE-1 (L1) retrotransposition by down-regulating the nuclear import factor TNPO1. J Biol Chem. 2017;292(50):20494-508.
89. Fung L, Guzman H, Sevrioukov E, Idica A, Park E, Bochnakian A, et al. miR-128 Restriction of LINE-1 (L1) Retrotransposition Is Dependent on Targeting hnRNPA1 mRNA. Int J Mol Sci. 2019;20(8).
90. Goodier JL, Cheung LE, Kazazian HH, Jr. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. Nucleic Acids Res. 2013;41(15):7401-19.
91. Hamdorf M, Idica A, Zisoulis DG, Gamelin L, Martin C, Sanders KJ, et al. miR-128 represses L1 retrotransposition by binding directly to L1 RNA. Nat Struct Mol Biol. 2015;22(10):824-31.
92. Christ F, Thys W, De Rijck J, Gijsbers R, Albanese A, Arosio D, et al. Transportin-SR2 imports HIV into the nucleus. Current biology : CB. 2008;18(16):1192-202.
93. Fernandez J, Machado AK, Lyonnais S, Chamontin C, Gartner K, Leger T, et al. Transportin-1 binds to the HIV-1 capsid via a nuclear localization signal and triggers uncoating. Nat Microbiol. 2019;4(11):1840-50.
94. Ostertag EM, Kazazian HH, Jr. Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition. Genome Res. 2001;11(12):2059-65.
95. Benitez-Guijarro M, Lopez-Ruiz C, Tarnauskaite Z, Murina O, Mian Mohammad M, Williams TC, et al. RNase H2, mutated in Aicardi-Goutieres syndrome, promotes LINE-1 retrotransposition. EMBO J. 2018;37(15).
96. Bartsch K, Knittler K, Borowski C, Rudnik S, Damme M, Aden K, et al. Absence of RNase H2 triggers generation of immunogenic micronuclei removed by autophagy. Hum Mol Genet. 2017;26(20):3960-72.
97. Reijns MA, Jackson AP. Ribonuclease H2 in health and disease. Biochem Soc Trans. 2014;42(4):717-25.
98. Beilhartz GL, Gotte M. HIV-1 Ribonuclease H: Structure, Catalytic Mechanism and Inhibitors. Viruses. 2010;2(4):900-26.
99. Choi J, Hwang SY, Ahn K. Interplay between RNASEH2 and MOV10 controls LINE-1 retrotransposition. Nucleic Acids Res. 2018;46(4):1912-26.
100. Skariah G, Seimetz J, Norsworthy M, Lannom MC, Kenny PJ, Elrakhawy M, et al. Mov10 suppresses retroelements and regulates neuronal development and function in the developing brain. BMC Biol. 2017;15(1):54.
101. Laguette N, Benkirane M. How SAMHD1 changes our view of viral restriction. Trends Immunol. 2012;33(1):26-33.
102. Zhao K, Du J, Han X, Goodier JL, Li P, Zhou X, et al. Modulation of LINE-1 and Alu/SVA retrotransposition by Aicardi-Goutieres syndrome-related SAMHD1. Cell Rep. 2013;4(6):1108-15.

103. Hu S, Li J, Xu F, Mei S, Le Duff Y, Yin L, et al. SAMHD1 Inhibits LINE-1 Retrotransposition by Promoting Stress Granule Formation. PLoS Genet. 2015;11(7):e1005367.
104. Gramberg T, Kahle T, Bloch N, Wittmann S, Mullers E, Daddacha W, et al. Restriction of diverse retroviruses by SAMHD1. Retrovirology. 2013;10:26.
105. Lahouassa H, Daddacha W, Hofmann H, Ayinde D, Logue EC, Dragin L, et al. SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. Nat Immunol. 2012;13(3):223-8.
106. Herrmann A, Wittmann S, Thomas D, Shepard CN, Kim B, Ferreiros N, et al. The SAMHD1-mediated block of LINE-1 retroelements is regulated by phosphorylation. Mob DNA. 2018;9:11.
107. Stetson DB, Ko JS, Heidmann T, Medzhitov R. Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell. 2008;134(4):587-98.
108. Mazur DJ, Perrino FW. Identification and expression of the TREX1 and TREX2 cDNA sequences encoding mammalian 3'-->5' exonucleases. J Biol Chem. 1999;274(28):19655-60.
109. Jones RB, Song H, Xu Y, Garrison KE, Buzdin AA, Anwar N, et al. LINE-1 retrotransposable element DNA accumulates in HIV-1-infected cells. J Virol. 2013;87(24):13307-20.
110. Sawyer SL, Emerman M, Malik HS. Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. PLoS Biol. 2004;2(9):e275.
111. Kazazian HH, Jr., Wong C, Youssoufian H, Scott AF, Phillips DG, Antonarakis SE. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature. 1988;332(6160):164-6.
112. Chen JM, Stenson PD, Cooper DN, Ferec C. A systematic analysis of LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease. Hum Genet. 2005;117(5):411-27.
113. Speek M. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. Mol Cell Biol. 2001;21(6):1973-85.
114. Matlik K, Redik K, Speek M. L1 antisense promoter drives tissue-specific transcription of human genes. J Biomed Biotechnol. 2006;2006(1):71753.
115. Piriyapongsa J, Marino-Ramirez L, Jordan IK. Origin and evolution of human microRNAs from transposable elements. Genetics. 2007;176(2):1323-37.
116. Qin S, Jin P, Zhou X, Chen L, Ma F. The Role of Transposable Elements in the Origin and Evolution of MicroRNAs in Human. PLoS One. 2015;10(6):e0131365.
117. Kines KJ, Sokolowski M, deHaro DL, Christian CM, Belancio VP. Potential for genomic instability associated with retrotranspositionally-incompetent L1 loci. Nucleic Acids Res. 2014;42(16):10488-502.
118. Gasior SL, Wakeman TP, Xu B, Deininger PL. The human LINE-1 retrotransposon creates DNA double-strand breaks. J Mol Biol. 2006;357(5):1383-93.
119. Wallace NA, Belancio VP, Deininger PL. L1 mobile element expression causes multiple types of toxicity. Gene. 2008;419(1-2):75-81.
120. Balachandar V, Dhivya V, Gomathi M, Mohanadevi S, Venkatesh B, Geetha B. A review of Rett syndrome (RTT) with induced pluripotent stem cells. Stem Cell Investig. 2016;3:52.
121. Leonard H, Cobb S, Downs J. Clinical and biological progress over 50 years in Rett syndrome. Nat Rev Neurol. 2017;13(1):37-51.
122. Picard N, Fagiolini M. MeCP2: an epigenetic regulator of critical periods. Curr Opin Neurobiol. 2019;59:95-101.
123. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet. 1999;23(2):185-8.

124. Yusufzai TM, Wolffe AP. Functional consequences of Rett syndrome mutations on human MeCP2. Nucleic Acids Res. 2000;28(21):4172-9.
125. Gulmez Karaca K, Brito DVC, Oliveira AMM. MeCP2: A Critical Regulator of Chromatin in Neurodevelopment and Adult Brain Function. Int J Mol Sci. 2019;20(18).
126. Ramocki MB, Tavyev YJ, Peters SU. The MECP2 duplication syndrome. Am J Med Genet A. 2010;152A(5):1079-88.
127. Zhao B, Wu Q, Ye AY, Guo J, Zheng X, Yang X, et al. Somatic LINE-1 retrotransposition in cortical neurons and non-brain tissues of Rett patients and healthy individuals. PLoS Genet. 2019;15(4):e1008043.
128. Karimi P, Kamali E, Mousavi SM, Karahmadi M. Environmental factors influencing the risk of autism. J Res Med Sci. 2017;22:27.
129. McPartland J, Volkmar FR. Autism and related disorders. Handb Clin Neurol. 2012;106:407-18.
130. Shpyleva S, Melnyk S, Pavliv O, Pogribny I, Jill James S. Overexpression of LINE-1 Retrotransposons in Autism Brain. Mol Neurobiol. 2018;55(2):1740-9.
131. Mitchell MM, Woods R, Chi LH, Schmidt RJ, Pessah IN, Kostyniak PJ, et al. Levels of select PCB and PBDE congeners in human postmortem brain reveal possible environmental involvement in 15q11-q13 duplication autism spectrum disorder. Environ Mol Mutagen. 2012;53(8):589-98.
132. Tangsuwansri C, Saeliw T, Thongkorn S, Chonchaiya W, Suphapeetiporn K, Mutirangura A, et al. Investigation of epigenetic regulatory networks associated with autism spectrum disorder (ASD) by integrated global LINE-1 methylation and gene expression profiling analyses. PLoS One. 2018;13(7):e0201071.
133. Nishioka M, Bundo M, Iwamoto K, Kato T. Somatic mutations in the human brain: implications for psychiatric research. Mol Psychiatry. 2019;24(6):839-56.
134. Smigielski L, Jagannath V, Rössler W, Walitza S, Grünblatt E. Epigenetic mechanisms in schizophrenia and other psychotic disorders: a systematic review of empirical human findings. Molecular Psychiatry. 2020.
135. Bundo M, Toyoshima M, Okada Y, Akamatsu W, Ueda J, Nemoto-Miyauchi T, et al. Increased l1 retrotransposition in the neuronal genome in schizophrenia. Neuron. 2014;81(2):306-13.
136. Guffanti G, Gaudi S, Klengel T, Fallon JH, Mangalam H, Madduri R, et al. LINE1 insertions as a genomic risk factor for schizophrenia: Preliminary evidence from an affected family. Am J Med Genet B Neuropsychiatr Genet. 2016;171(4):534-45.
137. Doyle GA, Crist RC, Karatas ET, Hammond MJ, Ewing AD, Ferraro TN, et al. Analysis of LINE-1 Elements in DNA from Postmortem Brains of Individuals with Schizophrenia. Neuropsychopharmacology. 2017;42(13):2602-11.
138. Jiang T, Zong L, Zhou L, Hou Y, Zhang L, Zheng X, et al. Variation in global DNA hydroxymethylation with age associated with schizophrenia. Psychiatry Res. 2017;257:497-500.
139. Li S, Yang Q, Hou Y, Jiang T, Zong L, Wang Z, et al. Hypomethylation of LINE-1 elements in schizophrenia and bipolar disorder. J Psychiatr Res. 2018;107:68-72.
140. Melas PA, Rogdaki M, Osby U, Schalling M, Lavebratt C, Ekstrom TJ. Epigenetic aberrations in leukocytes of patients with schizophrenia: association of global DNA methylation with antipsychotic drug treatment and disease onset. FASEB J. 2012;26(6):2712-8.
141. Misiak B, Szmida E, Karpinski P, Loska O, Sasiadek MM, Frydecka D. Lower LINE-1 methylation in first-episode schizophrenia patients with the history of childhood trauma. Epigenomics. 2015;7(8):1275-85.

142. Shimabukuro M, Sasaki T, Imamura A, Tsujita T, Fuke C, Umekage T, et al. Global hypomethylation of peripheral leukocyte DNA in male patients with schizophrenia: a potential link between epigenetics and schizophrenia. J Psychiatr Res. 2007;41(12):1042-6.
143. Crow YJ, Manel N. Aicardi-Goutieres syndrome and the type I interferonopathies. Nat Rev Immunol. 2015;15(7):429-40.
144. Volkman HE, Stetson DB. The enemy within: endogenous retroelements and autoimmune disease. Nature immunology. 2014;15(5):415-22.
145. Orecchini E, Doria M, Antonioni A, Galardi S, Ciafre SA, Frassinelli L, et al. ADAR1 restricts LINE-1 retrotransposition. Nucleic Acids Res. 2017;45(1):155-68.
146. Beck-Engeser GB, Eilat D, Wabl M. An autoimmune disease prevented by anti-retroviral drugs. Retrovirology. 2011;8:91.
147. Rice GI, Meyzer C, Bouazza N, Hully M, Boddaert N, Semeraro M, et al. Reverse-Transcriptase Inhibitors in the Aicardi-Goutieres Syndrome. N Engl J Med. 2018;379(23):2275-7.
148. Crow YJ, Leitch A, Hayward BE, Garner A, Parmar R, Griffith E, et al. Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection. Nat Genet. 2006;38(8):910-6.
149. Coffin SR, Hollis T, Perrino FW. Functional consequences of the RNase H2A subunit mutations that cause Aicardi-Goutieres syndrome. J Biol Chem. 2011;286(19):16984-91.
150. Gorman JA, Hundhausen C, Errett JS, Stone AE, Allenspach EJ, Ge Y, et al. The A946T variant of the RNA sensor IFIH1 mediates an interferon program that limits viral infection but increases the risk for autoimmunity. Nat Immunol. 2017;18(7):744-52.
151. Morales DJ, Lenschow DJ. The antiviral activities of ISG15. J Mol Biol. 2013;425(24):4995-5008.
152. Rodriguez-Garcia E, Olague C, Rius-Rocabert S, Ferrero R, Llorens C, Larrea E, et al. TMEM173 Alternative Spliced Isoforms Modulate Viral Replication through the STING Pathway. Immunohorizons. 2018;2(11):363-76.
153. Zheng Y, Lorenzo C, Beal PA. DNA editing in DNA/RNA hybrids by adenosine deaminases that act on RNA. Nucleic Acids Res. 2017;45(6):3369-77.
154. Rothblum-Oviatt C, Wright J, Lefton-Greif MA, McGrath-Morrow SA, Crawford TO, Lederman HM. Ataxia telangiectasia: a review. Orphanet J Rare Dis. 2016;11(1):159.
155. Blocher D, Sigut D, Hannan MA. Fibroblasts from ataxia telangiectasia (AT) and AT heterozygotes show an enhanced level of residual DNA double-strand breaks after low dose-rate gamma-irradiation as assayed by pulsed field gel electrophoresis. Int J Radiat Biol. 1991;60(5):791-802.
156. Coufal NG, Garcia-Perez JL, Peng GE, Marchetto MC, Muotri AR, Mu Y, et al. Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. Proc Natl Acad Sci U S A. 2011;108(51):20382-7.
157. Couratier P, Corcia P, Lautrette G, Nicol M, Marin B. ALS and frontotemporal dementia belong to a common disease spectrum. Rev Neurol (Paris). 2017;173(5):273-9.
158. Li W, Jin Y, Prazak L, Hammell M, Dubnau J. Transposable elements in TDP-43-mediated neurodegenerative disorders. PLoS One. 2012;7(9):e44099.
159. Liu EY, Russ J, Cali CP, Phan JM, Amlie-Wolf A, Lee EB. Loss of Nuclear TDP-43 Is Associated with Decondensation of LINE Retrotransposons. Cell Rep. 2019;27(5):1409-21 e6.
160. Pereira GC, Sanchez L, Schaughency PM, Rubio-Roldan A, Choi JA, Planet E, et al. Properties of LINE-1 proteins and repeat element expression in the context of amyotrophic lateral sclerosis. Mob DNA. 2018;9:35.
161. Masters CL, Bateman R, Blennow K, Rowe CC, Sperling RA, Cummings JL. Alzheimer's disease. Nat Rev Dis Primers. 2015;1:15056.

162. Beck JA, Poulter M, Campbell TA, Uphill JB, Adamson G, Geddes JF, et al. Somatic and germline mosaicism in sporadic early-onset Alzheimer's disease. Hum Mol Genet. 2004;13(12):1219-24.
163. Bollati V, Galimberti D, Pergoli L, Dalla Valle E, Barretta F, Cortini F, et al. DNA methylation in repetitive elements and Alzheimer disease. Brain Behav Immun. 2011;25(6):1078-83.
164. Hernandez HG, Mahecha MF, Mejia A, Arboleda H, Forero DA. Global long interspersed nuclear element 1 DNA methylation in a Colombian sample of patients with late-onset Alzheimer's disease. Am J Alzheimers Dis Other Demen. 2014;29(1):50-3.
165. Protasova MS, Gusev FE, Grigorenko AP, Kuznetsova IL, Rogaev EI, Andreeva TV. Quantitative Analysis of L1-Retrotransposons in Alzheimer's Disease and Aging. Biochemistry (Mosc). 2017;82(8):962-71.
166. Guo C, Jeong HH, Hsieh YC, Klein HU, Bennett DA, De Jager PL, et al. Tau Activates Transposable Elements in Alzheimer's Disease. Cell Rep. 2018;23(10):2874-80.
167. Verheijen BM, Vermulst M, van Leeuwen FW. Somatic mutations in neurons during aging and neurodegeneration. Acta Neuropathol. 2018;135(6):811-26.
168. Li W, Prazak L, Chatterjee N, Gruninger S, Krug L, Theodorou D, et al. Activation of transposable elements during aging and neuronal decline in Drosophila. Nat Neurosci. 2013;16(5):529-31.
169. Kemp JR, Longworth MS. Crossing the LINE Toward Genomic Instability: LINE-1 Retrotransposition in Cancer. Front Chem. 2015;3:68.
170. Kazazian HH, Jr., Moran JV. Mobile DNA in Health and Disease. N Engl J Med. 2017;377(4):361-70.
171. Belancio VP, Roy-Engel AM, Deininger PL. All y'all need to know 'bout retroelements in cancer. Semin Cancer Biol. 2010;20(4):200-10.
172. Alves G, Tatro A, Fanning T. Differential methylation of human LINE-1 retrotransposons in malignant cells. Gene. 1996;176(1-2):39-44.
173. Scott EC, Devine SE. The Role of Somatic L1 Retrotransposition in Human Cancers. Viruses. 2017;9(6).
174. Rangwala SH, Kazazian HH, Jr. The L1 retrotransposition assay: a retrospective and toolkit. Methods. 2009;49(3):219-26.
175. Kopera HC, Larson PA, Moldovan JB, Richardson SR, Liu Y, Moran JV. LINE-1 Cultured Cell Retrotransposition Assay. Methods Mol Biol. 2016;1400:139-56.
176. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH, Jr. High frequency retrotransposition in cultured mammalian cells. Cell. 1996;87(5):917-27.
177. Ostertag EM, Prak ET, DeBerardinis RJ, Moran JV, Kazazian HH, Jr. Determination of L1 retrotransposition kinetics in cultured cells. Nucleic Acids Res. 2000;28(6):1418-23.
178. Del Re B, Marcantonio P, Capri M, Giorgi G. Evaluation of LINE-1 mobility in neuroblastoma cells by in vitro retrotransposition reporter assay: FACS analysis can detect only the tip of the iceberg of the inserted L1 elements. Exp Cell Res. 2010;316(20):3358-67.
179. Vandekerckhove L, Christ F, Van Maele B, De Rijck J, Gijsbers R, Van den Haute C, et al. Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. J Virol. 2006;80(4):1886-96.
180. Osorio L, Gijsbers R, Oliveras-Salva M, Michiels A, Debyser Z, Van den Haute C, et al. Viral vectors expressing a single microRNA-based short-hairpin RNA result in potent gene silencing in vitro and in vivo. J Biotechnol. 2014;169:71-81.
181. Mangeot PE, Duperrier K, Negre D, Boson B, Rigal D, Cosset FL, et al. High levels of transduction of human dendritic cells with optimized SIV vectors. Mol Ther. 2002;5(3):283-90.

182. Geraerts M, Michiels M, Baekelandt V, Debyser Z, Gijsbers R. Upscaling of lentiviral vector production by tangential flow filtration. J Gene Med. 2005;7(10):1299-310.
183. Naldini L, Blomer U, Gage FH, Trono D, Verma IM. Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. Proc Natl Acad Sci U S A. 1996;93(21):11382-8.
184. Vranckx LS, Demeulemeester J, Saleh S, Boll A, Vansant G, Schrijvers R, et al. LEDGIN-mediated Inhibition of Integrase-LEDGF/p75 Interaction Reduces Reactivation of Residual Latent HIV. EBioMedicine. 2016;8:248-64.
185. Bueno C, Tabares-Seisdedos R, Moraleda JM, Martinez S. Rett Syndrome Mutant Neural Cells Lacks MeCP2 Immunoreactive Bands. PLoS One. 2016;11(4):e0153262.
186. Schwahn U, Lenzner S, Dong J, Feil S, Hinzmann B, van Duijnhoven G, et al. Positional cloning of the gene for X-linked retinitis pigmentosa 2. Nat Genet. 1998;19(4):327-32.
187. Metzner M, Jack HM, Wabl M. LINE-1 retroelements complexed and inhibited by activation induced cytidine deaminase. PLoS One. 2012;7(11):e49358.
188. Cook PR, Tabor GT. Deciphering fact from artifact when using reporter assays to investigate the roles of host factors on L1 retrotransposition. Mob DNA. 2016;7:23.
189. Leoh LS, van Heertum B, De Rijck J, Filippova M, Rios-Colon L, Basu A, et al. The stress oncoprotein LEDGF/p75 interacts with the methyl CpG binding protein MeCP2 and influences its transcriptional activity. Mol Cancer Res. 2012;10(3):378-91.
190. MacLennan M, Garcia-Canadas M, Reichmann J, Khazina E, Wagner G, Playfoot CJ, et al. Mobilization of LINE-1 retrotransposons is restricted by Tex19.1 in mouse embryonic stem cells. Elife. 2017;6.
191. Streva VA, Faber ZJ, Deininger PL. LINE-1 and Alu retrotransposition exhibit clonal variation. Mob DNA. 2013;4(1):16.
192. Swergold GD. Identification, characterization, and cell specificity of a human LINE-1 promoter. Mol Cell Biol. 1990;10(12):6718-29.
193. Leibold DM, Swergold GD, Singer MF, Thayer RE, Dombroski BA, Fanning TG. Translation of LINE-1 DNA elements in vitro and in human cells. Proc Natl Acad Sci U S A. 1990;87(18):6990-4.
194. Brouha B, Meischl C, Ostertag E, de Boer M, Zhang Y, Neijens H, et al. Evidence consistent with human L1 retrotransposition in maternal meiosis I. Am J Hum Genet. 2002;71(2):327-36.
195. Garcia-Perez JL, Morell M, Scheys JO, Kulpa DA, Morell S, Carter CC, et al. Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. Nature. 2010;466(7307):769-73.
196. Ge H, Si Y, Roeder RG. Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. EMBO J. 1998;17(22):6723-9.
197. Turlure F, Maertens G, Rahman S, Cherepanov P, Engelman A. A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. Nucleic Acids Res. 2006;34(5):1653-65.
198. Li R, Dong Q, Yuan X, Zeng X, Gao Y, Chiao C, et al. Misregulation of Alternative Splicing in a Mouse Model of Rett Syndrome. PLoS Genet. 2016;12(6):e1006129.
199. Cherepanov P, Maertens G, Proost P, Devreese B, Van Beeumen J, Engelborghs Y, et al. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. J Biol Chem. 2003;278(1):372-81.
200. Weichenrieder O, Repanas K, Perrakis A. Crystal structure of the targeting endonuclease of the human LINE-1 retrotransposon. Structure. 2004;12(6):975-86.

201. Repanas K, Zingler N, Layer LE, Schumann GG, Perrakis A, Weichenrieder O. Determinants for DNA target structure selectivity of the human LINE-1 retrotransposon endonuclease. Nucleic Acids Res. 2007;35(14):4914-26.
202. Sultana T, van Essen D, Siol O, Bailly-Bechet M, Philippe C, Zine El Aabidine A, et al. The Landscape of L1 Retrotransposons in the Human Genome Is Shaped by Pre-insertion Sequence Biases and Post-insertion Selection. Mol Cell. 2019;74(3):555-70 e7.
203. Flasch DA, Macia A, Sanchez L, Ljungman M, Heras SR, Garcia-Perez JL, et al. Genome-wide de novo L1 Retrotransposition Connects Endonuclease Activity with Replication. Cell. 2019;177(4):837-51 e28.
204. Daugaard M, Baude A, Fugger K, Povlsen LK, Beck H, Sorensen CS, et al. LEDGF (p75) promotes DNA-end resection and homologous recombination. Nat Struct Mol Biol. 2012;19(8):803-10.
205. Katz RA, Greger JG, Skalka AM. Effects of cell cycle status on early events in retroviral replication. J Cell Biochem. 2005;94(5):880-9.
206. Sinclair A, Yarranton S, Schelcher C. DNA-damage response pathways triggered by viral replication. Expert Rev Mol Med. 2006;8(5):1-11.
207. Skalka AM, Katz RA. Retroviral DNA integration and the DNA damage response. Cell Death Differ. 2005;12 Suppl 1:971-8.
208. Mita P, Sun X, Fenyo D, Kahler DJ, Li D, Agmon N, et al. BRCA1 and S phase DNA repair pathways restrict LINE-1 retrotransposition in human cells. Nat Struct Mol Biol. 2020;27(2):179-91.
209. An W, Dai L, Niewiadomska AM, Yetil A, O'Donnell KA, Han JS, et al. Characterization of a synthetic human LINE-1 retrotransposon ORFeus-Hs. Mob DNA. 2011;2(1):2.