Establishing cellular models to investigate the role of LINE-1 elements in X-chromosome inactivation and genome stability

Inês Cardial Mendes Dias

Instituto Superior Técnico, Lisbon, Portugal
November 2021

Abstract

Retrotransposons represent a considerable part of the mouse genome, and, among them, long interspersed nuclear element-1 (LINE-1) has remained highly active and capable of retrotransposition. Although this can lead to genome instability and disease, we are now able to appreciate that LINE-1 elements are major drivers of genome evolution and an alternative way of controlling gene expression. Despite their abundance and evidence suggesting the importance of transposable elements (TEs) in rewiring gene regulatory networks, these elements remain frequently "ignored" in genomic studies, due to their repetitive nature in the genome. In this project, cellular models to investigate the role of LINE-1 elements in X-chromosome inactivation and genome stability, during differentiation and in differentiated cells, were established. Engineered transcriptional activation and repression systems of LINE-1 elements that were already established in female mouse embryonic stem cells (mESCs) and female neural progenitor cells (NPCs) were partially optimized and a similar LINE-1 transcriptional activation system was implemented in female mouse embryonic fibroblasts (NIH-3T3 cells). The obtained results indicate that genome-wide perturbation of LINE-1 elements can be achieved with these systems, although future improvements must be performed at the level of differentiated cells.

Keywords

Transposable elements, LINE-1, CRISPR-Cas9, gene expression, X-chromosome inactivation, genome stability.

1. Introduction

1.1. LINE-1 elements in the mouse genome

Transposable elements (TEs) are DNA sequences with the ability to change their position in the genome, comprising the most abundant class of repeat sequences. TEs can be grouped into two classes, DNA transposons (class II) that show no recent evidence of transposition events, and retrotransposons (class I) (Bourque et al., 2018). Unlike most TEs that are truncated or inverted, thus showing no evidence for recent transposition events, full L1 elements seem to have remained highly active in the mouse lineage, containing at least 3,000 full-length elements that are potentially capable of retrotransposition (Goodier et al., 2001).

1.1.1. Structure, transposition mechanism and evolution

L1 elements are non-long terminal repeat (LTR) retrotransposons, which are known to rely on a unique transposition mechanism called target site-primed reverse transcription (TPRT) (Luan et al., 1993). Full-length mouse L1 elements are around 7-kb length, and they all have 5' untranslated region (UTR), two open reading frames (ORFs), and 3' poly(A) tail. The 5' UTR functions as a promoter and is composed of monomers, which are tandemly repeated sequences of ~200 bp situated upstream of single-copy, nonmonomeric sequence (Goodier et al., 2001). The ORFs encode for two proteins that are necessary for retrotransposition (Moran et al., 1996): ORF2p is a 150 kDa protein with endonuclease (Feng et al., 1996) and reverse transcription (Mathias et al., 1991) activities required for TPRT, and ORF1p is a 40 kDa nucleic acid chaperone.

Early phylogenetic analyses propose that mouse L1 evolution has been dominated by a single lineage (Adey et al., 1994), and that the least conserved regions within these elements are the 5´UTR as monomers differ in number and sequence among L1 elements, and their promoter activity is proportional to the number of monomers (DeBerardinis & Kazazian Jr., 1999). Although early L1 phylogenies usually show a cascade structure where a single family is active until a new family emerges and replaces the pre-existing one, it is now known that, in some instances, several lineages may co-exist until one becomes extinct; in fact, this is what is currently happening in the mouse genome, where two L1 lineages with non-homologous 5' UTR sequences have been active at the same time: L1MdA and L1MdTf/Gf (Sookdeo et al., 2013).

1.1.2. Function and regulation

The activation and retrotransposition of L1 elements can lead to the modification of the genome in a variety of ways, either beneficial or detrimental. Besides leading to gene inactivation through insertion into exons, L1 insertions into the genome are often associated with genome instability (Gilbert et al., 2002; Symer et al., 2002), and homologous recombination between dispersed L1 copies can lead to chromosomal rearrangements (Burwinkel & Kilimann, 1998). They can also modify cellular transcription acting as alternative promoters or enhancers, creating or disrupting polyadenylation sites, or splicing sites (Han et al., 2004). Although the misregulation of L1 elements can lead to diseases like cancer (Xiao-Jie et al., 2015) and neurodegeneration (Thomas et al., 2012), they have also been considered one of the strongest drivers of genome evolution and an alternative way of controlling gene expression.

The strongest and most significant mechanism of L1 regulation is possibly related to their chromatin structure, whereby L1 acquires a silent chromatin configuration. The main silencing signature present on most types of TEs is DNA methylation (Meissner et al., 2008), which is related to transcriptional repression. Among these, mouse L1 elements contain particularly high levels of DNA methylation and hypermethylated canonical promoters (Meissner et al., 2008). Nevertheless, DNA methylation is an epigenetic signature, which means that it is not always present on DNA and there are developmental time windows during which L1 elements are not methylated or not fully methylated. The most significant changes in the DNA methylation pattern occur when the genome undergoes global epigenetic reprogramming – during the formation of the primordial germ cells (PGCs) when almost a complete demethylation of the DNA occurs, and after fertilization when the newly formed zygote undergoes global DNA demethylation, reaching the lowest point in the early blastocyst (Mayer et al., 2000; Smith et al., 2012). Thus, the time period following fertilization in mammals is a great window of opportunity for the activation of L1 elements, given the chromatin context in the embryo. This led to hypothesize if the L1 reactivation during this time window was just a result of the chromatin state or whether it could play a role in development.

1.1.3. A possible role in development: X-chromosome inactivation

One of the most appealing hypotheses for the contribution of L1 elements in early development is that they could play a role in the inactivation of the X chromosome in females. X-chromosome inactivation (XCI) consists in the silencing of one X chromosome in females to compensate the differences in X-linked gene dosage between XX females and XY males (Lyon, 1961, 1962). Among eutherians, the mouse has been a preferred model to study XCI due to the use of murine embryonic stem cells (mESCs), which are t a useful model since in undifferentiated mESCs (as in cells from the ICM) both female X chromosomes

are active, but during differentiation (as in the embryo development), the silencing of one X chromosome is triggered. Thus, random XCI can be simulated during *in vitro* differentiation of mESCs, allowing the successive steps of XCI to be followed (Chaumeil et al., 2004; Wutz & Jaenisch, 2000). In the mouse, the process of XCI is initiated by the X-inactivation center (Xic), which includes a lncRNA called Xist (X-inactive-specific-transcript). The Xist transcript coats the future Xi chromosome in *cis* and triggers gene silencing. However, it has been shown that the low levels of Xist RNA in embryos and in differentiating female ES cells are incompatible with models that require Xist RNA to cover the entire Xi chromosome (Buzin et al., 1994). Here, chromosomal elements present on the X chromosome may exist to facilitate the propagation of the inactive state along the entire Xi; given the significant enrichment of L1 elements on the X chromosome, L1 elements were thought to play a role (Lyon, 1998), but the underlying mechanism was still largely unknown. Nevertheless, it was proposed by (Chow et al., 2010) that L1 elements may facilitate XCI at different levels, with silent LINEs participating in assembly of a heterochromatic nuclear compartment induced by Xist, and active LINEs facilitating silencing of certain regions of the Xi.

1.2. Functional approaches to address the role of TEs

Despite their abundance and evidence suggesting the importance of TEs in rewiring gene regulatory networks, these elements remain frequently "ignored" in genomic studies in part because of their repetitive nature, which makes them challenging to map onto a reference genome. To the best of my knowledge, only five genome-wide studies regarding the function and regulation of TEs in the genome were published until the present date (Fuentes et al., 2018; Honda et al., 2020; Jachowicz et al., 2017; Liu et al., 2017; Percharde et al., 2018). This knowledge gap is associated with technical challenges, given that a single gRNA is often insufficient for robust gene activation/silencing by CRISPRa/CRISPRi. TE families are often present in hundreds or thousands of copies, which are highly repetitive, but sufficiently sequence-divergent to prevent their recognition by a single short-sequence-dependent factor, such as a zinc finger protein or a CRISPR guide RNA (Fuentes et al., 2018).

1.2.1. Aim of this study

As stated before, there is a clear knowledge gap regarding systematic studies on TE function and consequent impact in genome stability. Thus, the main objective of this project is to establish functional cellular models that could be useful to address the impact of young LINE-1 elements genome-wide, during differentiation and in differentiated cells. Since functional perturbation systems of L1 expression in undifferentiated mouse embryonic stem cells had already been established in the lab, the implementation of such systems during differentiation and in differentiated cells would allow us to follow the cellular developmental stages and, thus, study the impact of L1 elements during this period.

2. Materials and methods

2.1. Cell culture

ES cells were cultured in 0.1% gelatin-coated plates. To maintain them in an undifferentiated state, ES cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) with sodium pyruvate (Gibco by Life Technologies), supplemented with 15% (v/v) heat-inactivated ES-grade fetal bovine serum (FBS) (Gibco by Life Technologies), 1% (v/v) 200 mM L-glutamine (Gibco by Life Technologies), 0.2% (v/v) 50 mM β-mercaptoethanol (Sigma Aldrich) and 0.01% (v/v) leukemia inhibitory factor (LIF) (Merck). The previous complete mixture will be referred as ES medium. Undifferentiated ES cells were passaged between 1:8 and 1:10 every two days.

ES cells were randomly differentiated through a process of LIF withdrawal. For this purpose, cells were cultured with Dulbecco's Modified Eagle's Medium (DMEM) with sodium pyruvate (Gibco by Life Technologies), supplemented with 10% (v/v) fetal bovine serum (FBS) (Gibco by Life Technologies), 1% (v/v) 200 mM L-glutamine (Gibco by Life Technologies) and 0.2% (v/v) 50 mM β-mercaptoethanol (Sigma Aldrich). The previous complete mixture will be referred as differentiation medium. For each differentiation experiment, 5×10^5 cells/well were plated.

Previously established NPCs were cultured in 0.1% gelatin-coated plates with N2B27 medium supplemented with murine epidermal growth factor (EGF) (Peprotech) and human fibroblast growth factor-basic (FGF) (Peprotech), at a final concentration of 10 ng/ml each. The previous complete mixture will be referred as NPC medium. N2B27 medium consists in a mixture 1:1 of Dulbecco's Modified Eagle's Medium (DMEM)/F12 (Gibco by Life Technologies) and Neurobasal Medium (Gibco by Life Technologies), supplemented with 1% (v/v) L-glutamine 200 mM (Gibco by Life Technologies), 1% (v/v) B27 (Gibco by Life Technologies), 0.5% (v/v) N2 (Millipore) and 0.2% (v/v) β-mercaptoethanol 50 mM (Sigma Aldrich). NPCs were passaged between 1:2 and 1:4 every 2-3 days.

NIH-3T3 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco by Life Technologies) supplemented with 10% (v/v) fetal bovine serum (FBS) (Gibco by Life Technologies), 1% (v/v) 200 mM L-glutamine (Gibco by Life Technologies) and 1% (v/v) penicillin-streptomycin antibiotics solution (Gibco by Life Technologies). NIH-3T3 cells were passaged between 1:8 and 1:10 every two days.

2.2. Transfection

Transfection of plasmids was performed with Lipofectamine 3000 Transfection Reagent (Invitrogen). A total amount of 2.5 μg of plasmid DNA was transfected in each experiment.

2.3. Plasmids and genetic constructs

A number of genetic constructs in the form of recombinant plasmids, already available in the host lab, were used. The plasmids pX330-sgTIGRE-Cas9 (Addgene, #92144) and pEN366-dCas9-VPR (a kind gift from Michel Wassef, Institut Curie) were used for CRISPR-mediated homologous recombination: pX330-sgTIGRE-Cas9 expresses a single guide RNA (sgRNA) for the TIGRE safe harbor locus, under the control of a human U6 (hU6) promoter, and a nuclease Cas9 protein, under the control of a constitutive hybrid human cytomegalovirus enhancer/chicken β-actin promoter (CBh); pEN366-dCas9-VPR expresses a nuclease-dead Cas9 (dCas9) protein fused to a VPR activation complex (composed by three transcriptional activation domains – VP64, p65 and Rta), under the control of a Tet-On 3G tetracycline-inducible expression system. The plasmid pLK01-sgTfmono2-3 expresses two sgRNAs for the 5'UTR monomers of L1Md-Tf elements, respectively under the control of hU6 and mouse U6 (mU6) promoters.

2.4. Bacterial transformation

The plasmids were amplified in competent *Escherichia coli* DH5α bacteria, following a heat shock protocol. The amplified plasmid DNA was separated and purified using the NZYMiniprep kit (Nzytech).

2.5. DNA extraction for genotyping

To genotype cell clones from a tissue culture 96-well plate, a quick cell lysis was performed to release DNA from cells, using a lysis buffer composed of 100 mM tris hydrochloride (Tris-HCl) pH 8.0, 200 mM sodium chloride (NaCl), 0.5% (v/v) Tween20 and freshly added Proteinase K (Nzytech) to a final concentration of 0.2 mg/ml. When higher quality DNA samples were required, a phenol/chloroform DNA purification protocol was followed.

2.6. RNA extraction

Total RNA isolation was performed using TRIzol reagent (Life Technologies). The obtained RNA samples were diluted with DNase/RNase free water to a final concentration of 200 ng/μl to

decrease the variability among samples in downstream applications.

2.7. Reverse transcription

To convert the previously extracted RNA into complementary DNA (cDNA), NZY First-Strand cDNA Synthesis Kit (Nzytech) was used. RNase H (Nzytech) was added to degrade the RNA that was bond to cDNA. The obtained cDNA samples were diluted 1:5 with DNase/RNase free water before being used in downstream applications.

2.8. Quantitative PCR

Quantitative PCR (qPCR) was used in association with reverse transcription. For this purpose, 2 μl of diluted sample (cDNA or negative control) were used in each reaction, as template for the quantitative PCR. Negative control samples consisted in 800 ng of RNA diluted in DNase/RNase free water to a final volume of 100 μl. Additionally, 5 μl of PowerUp SYBR Green Master Mix 2X (Thermo Fisher Scientific), 0.25 μl of forward/reverse primers mix (at 10 μM each) and 2.75 μl of DNase/RNase free water were added to the template, making a final volume of 10 μl per reaction. Each reaction was carried out in triplicates, except for the negative control samples, for which only one reaction was carried out per primers pair. A relative quantification assay following the $2^{-\Delta\Delta C_t}$ method was performed to determine the changes in gene expression of each sample relative to a reference sample.

2.9. Immunofluorescence

For immunofluorescence experiments, cells were grown on sterile 22x22 mm coverslips for 48 hours. Fixation was performed with a solution of 3.7% (v/v) paraformaldehyde (PFA) in PBS 1X. The cells were permeabilized with ice-cold solution of 0.5% (v/v) Triton-X100 in PBS 1X and blocked with a solution of 1% (v/v) BSA in PBS 1X (blocking solution). Incubation with the appropriate primary antibodies diluted in blocking solution was performed in a humid light-tight box for 1 hour at RT. Incubation with the appropriate conjugated secondary antibodies diluted in blocking solution was performed also in a humid light-tight box for 45 min at RT. For nuclear staining purposes, coverslips were incubated in the dark with 1 μg/ml solution of 4',6-diamidino-2-phenylindole (DAPI) (Sigma Aldrich) in PBS 1X. As primary antibodies, mouse anti-Cas9 (clone 8C1-F10, Active Motif, #61957) and rabbit anti-Cas9 (from the recombinant antibody platform at Institut Curie, Paris, France) were used, at a 1:100 dilution in blocking solution. As secondary antibodies, goat anti-mouse-Dy488 (Bethyl Laboratories, A90-244D2) and goat anti-rabbit-Cy3 (Bethyl Laboratories, A120-201C3) were used at a 1:200 dilution in blocking solution. Images were acquired using Zeiss Cell Observer, a widefield fluorescence microscope, and then, image analysis was performed on ImageJ software.

3. Results

3.1. CRISPRa-mediated activation levels of L1Md-Tf expression are not maintained during mES cells differentiation nor in differentiated NPCs

Before I came into the lab, a doxycycline-inducible CRISPR activation (CRISPRa) system had already been established in female mES cells (**Figure 1D**). In the presence of doxycycline, this system encodes for a dCas9 protein fused to a VPR transcriptional activation complex (including VP64, p65 and Rta transcriptional activators). Later, this dCas9-VPR complex binds two guide RNAs that are complementary to monomers 2 and 3 of the 5'UTR region of L1Md-Tf elements and, in theory, this should lead to genome-wide activation/increase of L1Md-Tf expression.

Doxycycline-treated (1 μg/ml of doxycycline for 48 hours) and untreated samples of B7, B4 and D4 clones were used. B7 is a control clone that lacks the guide RNAs for 5'UTR monomers of L1Md-Tf elements. B4 and D4 are two clones that contain a functional CRISPRa system, including the guide RNAs for 5'UTR monomers of L1Md-Tf elements. Although in undifferentiated mES cells (**Figure 1A**) it was possible to observe a great increase in expression of dCas9-VPR transgene (~ 200 to 600-fold change) and in genome-wide L1Md-Tf expression (~ 3-fold change) when doxycycline was present, these expression levels were not maintained during mES cells differentiation (**Figure 1B**). In fact, after 4 days of mES cells differentiation, the expression levels of dCas9-VPR transgene become much lower (~ 50 to 100-fold change) than in undifferentiated mES cells and the expression levels of L1Md-Tf elements no longer change in the presence of doxycycline. Furthermore, differentiated NPCs (**Figure 1C**) showed even lower values of dCas9-VPR transgene expression (~ 2.5-fold change) and, although there is some increase in the expression levels of L1Md-Tf elements (~ 1.5-fold change) in the presence of doxycycline, these are much lower than in undifferentiated mES cells. This decrease in expression of the dCas9-VPR transgene and the guide RNAs is not desirable, since it impairs the final goal of the CRISPRa system which is to impact the transcriptional levels of L1Md-Tf elements, by increasing their expression.

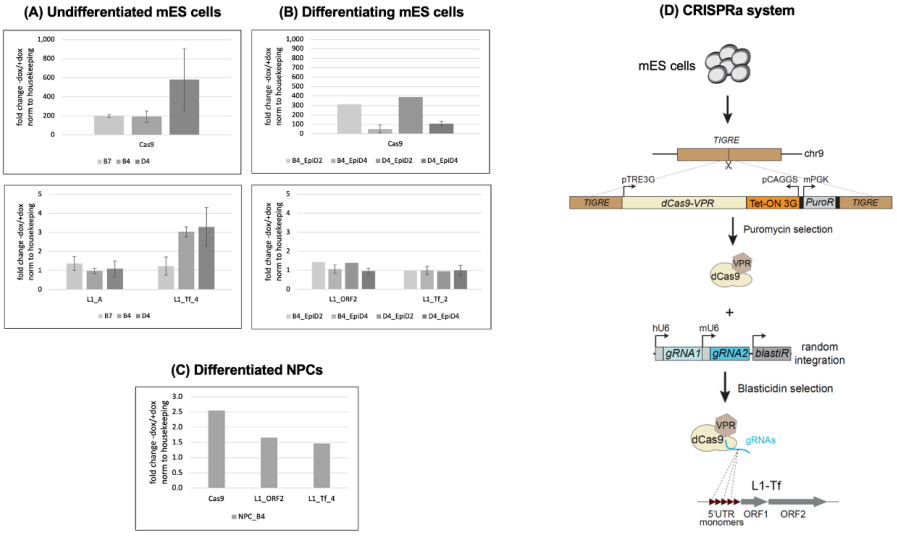

Figure 1: CRISPRa-mediated activation levels of L1Md-Tf expression are not maintained during mES cells differentiation nor in differentiated NPCs. Relative expression levels of dCas9-VPR transgene and L1Md-Tf elements after 48h treatment with 1 μg/ml of doxycycline in **(A)** undifferentiated female mES cells **(B)** 2-day and 4-day differentiated mES cells and **(C)** differentiated NPCs. Expression levels were obtained by qPCR and analyzed with the $2^{-\Delta\Delta C_t}$ method, using beta-actin as housekeeping gene. Data is normalized to the housekeeping gene and represented as mean of the fold change over the control (doxycycline-untreated cells), +/- standard deviation of the mean (n=3 biological replicates). Cas9, L1_A, L1_ORF2 and L1_Tf

correspond to primer pairs that amplify a genomic region of the dCas9-VPR transgene, L1Md-A elements, ORF2 coding region of L1 elements and L1Md-Tf elements, respectively. EpiD2 and EpiD4 are mES cell clones (B4 or D4) after 2 and 4 days of mESC differentiation, respectively. **(D)** L1Md-Tf activation system implemented in female mES cells.

3.2. KRAB-ZFP-mediated repression levels of L1Md-Tf expression are not maintained during mES cells differentiation

The same tendency was observed on a doxycycline-inducible transcriptional repression system (**Figure 2C**) that had already been established in female mES cells. In the presence of doxycycline, this system encodes for zinc finger proteins (ZFP) with an engineered DNA binding domain that is complementary to 5'UTR monomers of L1Md-Tf elements, fused to a Kruppel associated box (KRAB) transcriptional repressor. In theory, this system should lead to genome-wide repression/decrease of L1Md-Tf expression.

Doxycycline-treated (1 μg/ml of doxycycline for 48 hours) and untreated samples of B5, A5 and A6 clones were used. B5 is a control clone that carries mutations in the DNA binding domain of ZFP, preventing it from binding to 5'UTR monomers of L1Md-Tf elements. A5 and A6 clones are two clones with a functional KRAB-ZFP repression system for L1Md-Tf elements. Similarly to the CRISPRa system, although in undifferentiated mES cells (**Figure 2A**) we are able to observe a great increase in expression of KRAB-ZFP transgene (~ 25 to 65-fold change) and a considerable decrease in genome-wide L1Md-Tf expression (~ 0.2 to 0.4-fold change) when doxycycline was present, these expression levels were not maintained during mES cells differentiation (**Figure 2B**). After 4 days of mES cells differentiation, the expression levels of KRAB-ZFP transgene become lower (~ 20-fold change) than in undifferentiated mES cells and the expression levels of L1Md-Tf elements no longer change in the presence of doxycycline, which indicates that the repression system is no longer able to interfere with the transcriptional levels of these elements.

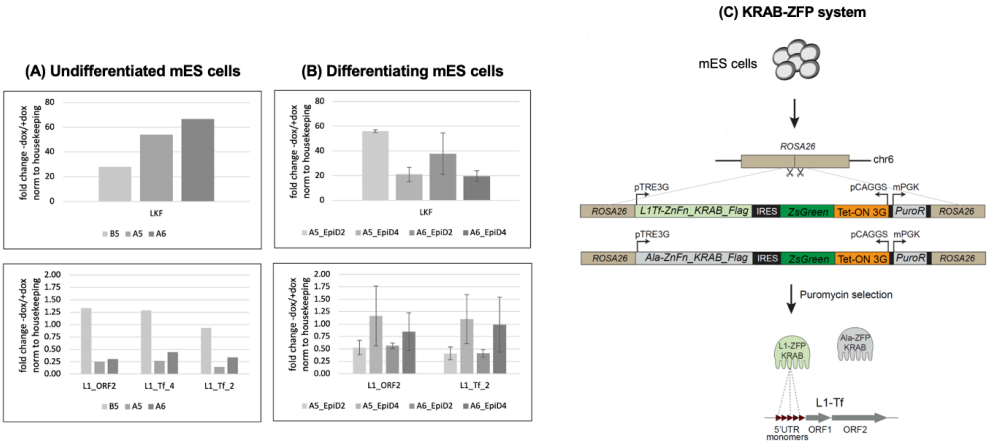

Figure 2: KRAB-ZFP-mediated repression levels of L1Md-Tf expression are not maintained during mES cells differentiation. Relative expression levels of KRAB-ZFP transgene and L1Md-Tf elements after 48h treatment with 1 μg/ml of doxycycline in **(A)** undifferentiated female mES cells and **(B)** 2-day and 4-day differentiated mES cells. Expression levels were obtained by qPCR and analyzed with the $2^{-\Delta\Delta C_t}$ method, using beta-actin as housekeeping gene. Data is normalized to the housekeeping gene and represented as mean of the fold changes over the control (doxycycline-untreated cells), +/- standard deviation of the mean (n=3 biological replicates). LKF, L1_ORF2 and L1_Tf correspond to primer pairs that amplify a genomic region of the KRAB-ZFP transgene, ORF2 coding region of L1 elements and L1Md-Tf elements, respectively. EpiD2 and EpiD4 are mES cell clones (A5 or A6) after 2 and 4 days of differentiation, respectively. **(C)** L1Md-Tf repression system implemented in female mES cells.

Considering the above-mentioned results, we wanted to test if we could improve both systems by creating an open chromatin configuration near the transgenes of interest (dCas9-VPR transgene and guide RNAs for the activation system; KRAB-ZFP transgene for the repression system) and thus, establish functional cellular models in which we would be able to interfere with the expression levels of L1Md-Tf elements during differentiation and in differentiated cells, such as NPCs. For that purpose, we decided to grow mES cells (undifferentiated and during differentiation) and differentiated NPCs with the drugs that had previously been used for selection purposes. Since the resistance marker for the selection drugs is located adjacent of the transgenes of interest (puromycin for dCas9-VPR and KRAB-ZFP transgenes and blasticidin for the sgRNAs; **Figure 1D** and **Figure 2C**), we hypothesize that by growing the cells with these drugs we will force the expression of the resistance genes and thus, create an open chromatin configuration in that region that could also favor the transcription of the adjacent genes – the transgenes of interest – since they would benefit from the imposed open chromatin configuration.

3.3. CRISPRa-mediated activation levels of L1Md-Tf expression do not change after treatment with selection drugs, in undifferentiated mES cells

Undifferentiated mES cells were grown in ES medium supplemented with 1 μg/ml of puromycin and 10 μg/ml of blasticidin. Doxycycline-treated (1 μg/ml of doxycycline for 48 hours) and untreated samples of the mES B7, B4 and D4 clones were then used for expression analysis, at both the RNA and the protein levels. At the RNA level (**Figure 3**), although it was possible to observe a great increase in expression of dCas9-VPR transgene in the 3 clones analyzed (~ 120 to 270-fold change) and in genome-wide L1Md-Tf expression in B4 and D4 clones (~ 4-fold change) when doxycycline was present, these expression levels are not considerably different from the expression levels obtained in undifferentiated mES cells that were grown in ES medium without selection drugs (**Figure 1A**).

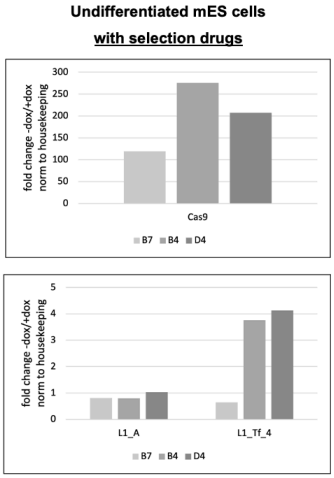

Figure 3: CRISPRa-mediated activation levels of L1Md-Tf expression do not increase after treatment with selection drugs, in undifferentiated mES cells. Relative expression levels of dCas9-VPR transgene and L1Md-Tf elements after 48h treatment with 1 μg/ml of doxycycline in undifferentiated female mES cells that were grown in ES medium supplemented with 1 μg/ml of puromycin and 10 μg/ml of

blasticidin. Expression levels were obtained by qPCR and analyzed with the $2^{-\Delta\Delta Ct}$ method, using Gapdh as housekeeping gene. Data is normalized to the housekeeping gene and represented as fold change over the control (doxycycline-untreated cells). Cas9, L1_A and L1_Tf correspond to primer pairs that amplify a genomic region of the dCas9-VPR transgene, L1Md-A elements and L1Md-Tf elements, respectively.

At the protein level, immunofluorescence was performed to further assess the impact of growing undifferentiated mES cells in ES medium supplemented with 1 μg/ml of puromycin and 10 μg/ml of blasticidin. Microscope slides were prepared with anti-Cas9 primary antibody (**Figure 4B**) and the percentage of Cas9-positive cells was determined for both conditions (with and without selection drugs) (**Figure 4C**). The obtained preliminary results suggest that dCas9 nuclease is being translated in presence of doxycycline, but no considerable enrichment seems to be observed when growing undifferentiated mES cells with selection drugs.

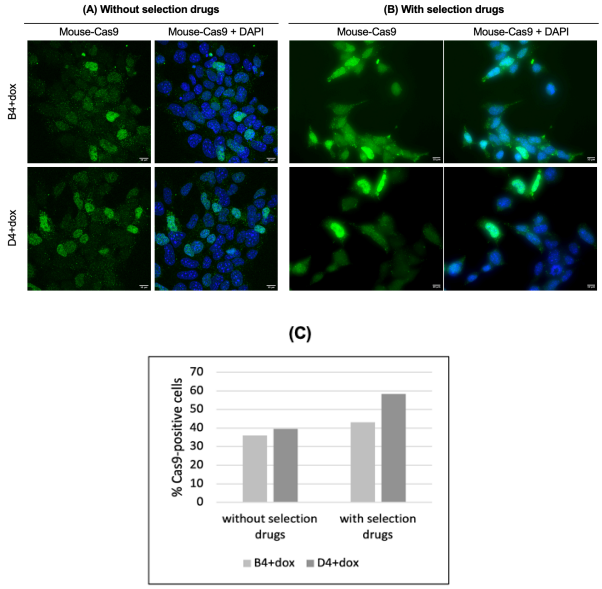

Figure 4: Undifferentiated mES cells population is not enriched in dCas9 nuclease after treatment with selection drugs. Representative immunofluorescence images of doxycycline-treated (1 μg/ml for 48h) mES cells that were grown in ES medium **(A)** without and **(B)** with 1 μg/ml of puromycin and 10 μg/ml of blasticidin. **(C)** Percentage of Cas9-positive cells in the mES cells population, with and without selection drugs. DAPI is presented here as a nuclear marker. Scale bars: 10 μm.

3.4. KRAB-ZFP-mediated repression of L1Md-Tf expression is maintained after 3 days of random mES cells differentiation, in the presence of selection drugs

The impact of the selection drugs was further studied during differentiation, in female mES cells containing the transcriptional repression system, KRAB-ZFP, in one mES clone (A6 clone). For that purpose, mES cells underwent a 2-day and 3-day random differentiation process by LIF withdrawal in ES differentiation medium supplemented with 1 μg/ml of puromycin. Doxycycline-treated (1 μg/ml of doxycycline for 48 hours) and untreated samples were harvested and the expression levels of KRAB-ZFP transgene and L1Md-Tf elements during the differentiation process were assessed by qPCR (**Figure 5A**). Although a decrease in KRAB-ZFP transgene expression is observed between days 2 and 3 of differentiation, the levels of transgene expression are still considerably high after 3 days of differentiation (~ 300-fold change). At the level of L1Md-Tf expression, a genome-wide L1Md-Tf de-repression is observed between days 2 and 3 of differentiation, but again, a considerable level of repression is still observed after 3 days of differentiation (~ 0.3-fold change). The decrease in expression of Nanog and Klf4 pluripotency markers, as well as the increase in expression of Xist and Fgf5 differentiation markers, during differentiation (**Figure 5B**) show that the differentiation process was successful. Altogether, these results led us to conclude that, after 3 days of mES cells random differentiation by LIF withdrawal, in the presence of 1 μg/ml of puromycin, the L1Md-Tf transcriptional repression system seems to be functional and thus, being able to induce a considerable genome-wide repression of L1Md-Tf expression. Furthermore, the expression levels presented in **Figure 5** are similar to the ones that were previously obtained for undifferentiated mES cells without selection drugs (**Figure 2A**), which seems to be a good indication that indeed the use of selection drugs during mES cells differentiation may increase the performance of the L1Md-Tf transcriptional repression system.

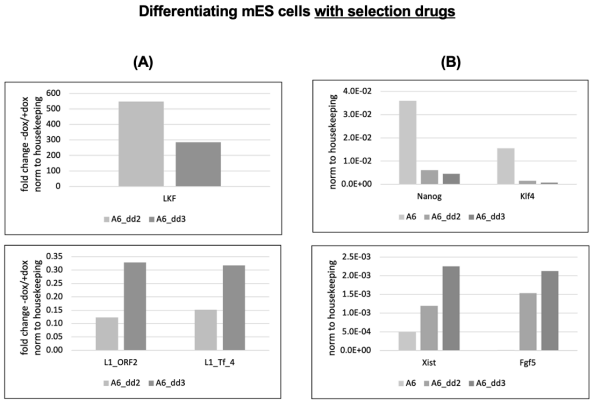

Figure 5: KRAB-ZFP-mediated repression of L1Md-Tf expression is maintained after 3 days of random mES cells differentiation, in the presence of selection drugs. (A) Relative expression levels of KRAB-ZFP transgene and L1Md-Tf elements during mES cells random differentiation by LIF withdrawal in ES differentiation medium supplemented with1 μg/ml of puromycin. **(B)** Relative expression levels of pluripotency markers (Nanog and Klf4) and differentiation markers (Xist and Fgf5), before and during mES cells random differentiation by LIF withdrawal in ES differentiation medium supplemented with1 μg/ml of puromycin. Expression levels were obtained by qPCR and analyzed with the $2^{-\Delta\Delta Ct}$ method, using Gapdh as housekeeping gene. Data is normalized to the housekeeping gene and, in (A), represented as fold change over the control (doxycycline-untreated cells). Doxycycline treatment was performed with 1 μg/ml of doxycycline for 48 hours. LKF, L1_ORF2 and L1_Tf correspond to primer pairs that amplify a genomic region of the KRAB-ZFP transgene, ORF2 coding region of L1 elements and L1Md-Tf elements, respectively. A6 corresponds to undifferentiated mES cell clone. A6_dd2 and A6_dd3 correspond to A6 clone after 2 and 3 days of mESC differentiation, respectively.

3.5. CRISPRa-mediated activation levels of L1Md-Tf expression increase after treatment with selection drugs, in differentiated NPCs

Finally, the impact of the selection drugs was studied in differentiated female NPCs containing the transcriptional activation CRISPRa system, obtained after differentiation of the mES clone B4 (NPC_B4 clone). For that purpose, NPCs were grown in NPC medium supplemented with 1 μg/ml of puromycin and 10 μg/ml of blasticidin. Doxycycline-treated (1 μg/ml of doxycycline for 48 hours) and untreated samples were harvested and the expression levels of dCas9-VPR transgene and L1Md-Tf elements were assessed by qPCR (**Figure 6**). A considerable increase in expression of dCas9-VPR transgene (~ 17.5-fold change) and a modest increase in expression of L1Md-Tf elements (~ 2.5-fold change) were observed after growing the cells with selection drugs. Although these expression levels continue to be lower than the expression levels of undifferentiated mES cells without selection drugs (**Figure 1A**), they are higher than the previously obtained ones for differentiated NPCs without selection drugs (**Figure 1C**).

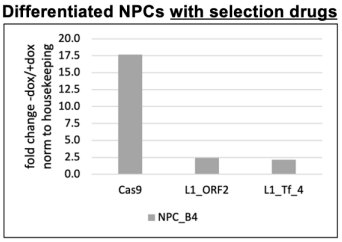

Figure 6: CRISPRa-mediated activation levels of L1Md-Tf expression increase after treatment with selection drugs, in differentiated NPCs. Relative expression levels of dCas9-VPR transgene and L1Md-Tf elements after 48h treatment with 1 μg/ml of doxycycline in differentiated female NPCs that were grown in NPC medium supplemented with 1 μg/ml of puromycin and 10 μg/ml of blasticidin. Expression levels were obtained by qPCR and analyzed with the $2^{-\Delta\Delta Ct}$ method, using Gapdh as housekeeping gene. Data is normalized to the housekeeping gene and represented as

fold change over the control (doxycycline-untreated cells). Cas9, L1_ORF2 and L1_Tf correspond to primer pairs that amplify a genomic region of the dCas9-VPR transgene, ORF2 coding region of L1 elements and L1Md-Tf elements, respectively.

3.6. Establishing a CRISPRa system for L1Md-Tf elements in fully differentiated, somatic NIH-3T3 cells

3.6.1. CRISPR-mediated homologous recombination of dCas9-VPR transgene

Next, we sought to implement the L1Md-Tf transcriptional activation system in fully differentiated, somatic cells. For that purpose, female NIH-3T3 cell line, which is a standard cell line of mouse embryonic fibroblasts, was used. The dCas9-VPR transgene was integrated into the TIGRE safe harbor locus by CRISPR-mediated homologous recombination, i.e., by co-transfection of pX330-sgTIGRE-Cas9 and pEN366-dCas9-VPR plasmids. The cells where the dCas9-VPR transgene was successfully integrated (somewhere in the genome, not necessarily in the desired location) were selected with 2.5 μg/ml of puromycin for one week. Afterwards, 48 puromycin-resistant clonal colonies were picked into a 96-well plate and, after performing quick cell lysis and PCR genotyping (using primers that amplify the left transgene/TIGRE junction after CRISPR-mediated homologous recombination) on the 48 cell clones, only 4 of them seemed to have the dCas9-VPR transgene integrated at the desired location. Next, these 4 clones were further expanded and cells were harvested for proper phenol/chloroform DNA extraction. PCR genotyping was performed on these DNA samples and only two (G4 and H6) out of the four clones were confirmed to have the dCas9-VPR transgene integrated at the TIGRE safe harbor locus.

3.6.2. Induction of dCas9-VPR expression with doxycycline

Next, the confirmed positive clones, G4 and H6, were cultured in 3T3 medium and expression of dCas9-VPR transgene was inducted by treating the cells with doxycycline for 48 hours. Two different concentrations of doxycycline were tested, 1 μg/ml and 2 μg/ml. Then, doxycycline-treated and untreated cells were harvested and the expression levels of dCas9-VPR transgene were assessed by qPCR (**Figure 7A**). The expression levels of L1-ORF2 coding region, L1Md-A, L1Md-Gf and L1Md-Tf elements were also assessed in both clones by qPCR (**Figure 7B**). For H6 clone, no activation of L1 elements was observed after treatment with doxycycline, which was the expected behavior for cells lacking guide RNAs complementary with 5'UTR monomers of L1 elements. G4 clone, on the other hand, showed an unexpected behavior and, thus, it was not considered for further studies. Although we were able to obtain one positive clone, i.e., with the dCas9-VPR transgene integrated at the TIGRE safe harbor locus and a functional CRISPRa system in general, the efficiency of CRISPR-mediated homologous recombination in NIH-3T3 cells seems to be sub-optimal.

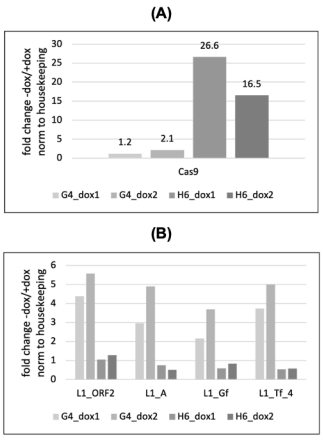

Figure 7: dCas9-VPR transgene expression levels increase, unlike L1 expression levels, in H6 cell clone after treatment with 1 μg/ml of doxycycline. Relative expression levels of **(A)** dCas9-VPR transgene and **(B)** L1 elements, after 48h treatment with 1 μg/ml (dox1) or 2 μg/ml (dox2) of doxycycline in puromycin-resistant NIH-3T3 cell clones. Expression levels were obtained by qPCR and analyzed with the $2^{-\Delta\Delta C_t}$ method, using Gapdh as housekeeping gene. Data is normalized to the housekeeping gene and represented as fold change over the control (doxycycline-untreated cells). Cas9, L1_ORF2, L1_A, L1_Gf and L1_Tf correspond to primer pairs that amplify a genomic region of the dCas9-VPR transgene, ORF2 coding region of L1 elements, L1Md-A elements, L1Md-Gf elements and L1Md-Tf elements, respectively.

At the protein level, immunofluorescence was performed with rabbit anti-Cas9 primary antibody to check the dCas9 nuclease levels of H6 clone, before and after treatment with 1 μg/ml of doxycycline for 48h (**Figure 8**). Doxycycline-untreated population of cells showed no Cas9-positive cells, which indicates that the system is not leaky, and the population of doxycycline-treated cells presented approximately 15% of Cas9-positive cells. Although this is not a high percentage of Cas9-positive cells, as a preliminary result, it does indicate that dCas9 nuclease is only being translated in the presence of doxycycline and that the implemented system, so far, is functional.

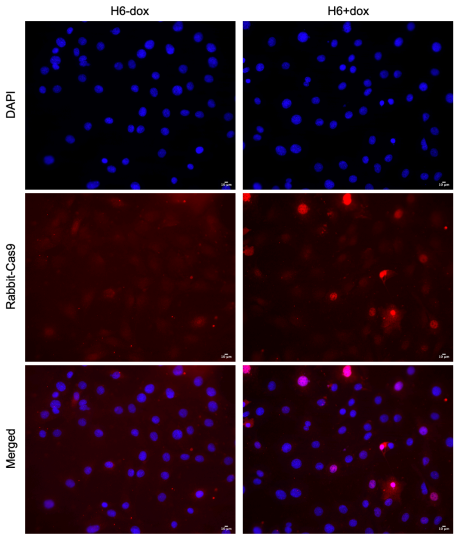

Figure 8: dCas9 nuclease translation is activated in NIH-3T3 cell clone H6 after induction with doxycycline. Representative immunofluorescence images of doxycycline-untreated and treated (1 μg/ml for 48h) NIH-3T3 cells of H6 clone. DAPI is presented here as a nuclear marker. Quantification was performed on both conditions, resulting in 0% and 15% of Cas9-positive cells for untreated and treated cell populations, respectively. Scale bars: 10 μm.

3.6.3. Random integration of sgRNAs for 5'UTR of L1Md-Tf elements

To achieve a complete CRISPRa system, the next step was to transfect sgRNAs that were complementary to 5'UTR monomers of L1Md-Tf elements. For this purpose, H6 clone was cultured in 3T3 medium and transfected with pLK01-sgTfmono2-3 plasmid. The cells where the transgene was successfully integrated, by random integration, were selected with 10 μg/ml of blasticidin for one week. Then, a 48-hour treatment with 1 μg/ml of doxycycline was performed in part of the blasticidin-resistant cells to induce the expression of the dCas9-VPR transgene and study the impact of the implemented CRISPRa system on L1Md-Tf elements, at the population level. For that purpose, doxycycline-treated and untreated cells were harvested and the expression levels of dCas9-VPR transgene and L1Md-Tf elements were assessed by qPCR (**Figure 9**). At this point, a great increase in expression of dCas9-VPR transgene (~ 55-fold change) and only a modest activation of L1Md-Tf elements (~ 1.6-fold change) was observed at the population of cells after transfection of sgRNAs. Although at sub-optimal levels, the preliminary results presented in **Figure 9** indicate that the transcriptional activation system is, at least, functional and leading to some L1Md-Tf activation, when compared with the control clone (clone H6 without sgRNAs for L1Md-Tf elements). Thus, a clonal and stable CRISPRa cell line could be established at this point, by picking individual colonies of cells transfected with both dCas9-

VPR and sgRNAs transgenes. Then, further validation and characterization should be performed by qPCR, western blot and immunofluorescence, in accordance with what was previously done in mES cells.

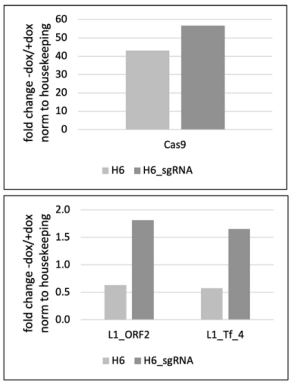

Figure 9: CRISPRa-mediated activation of L1Md-Tf expression was observed by qPCR, in fully differentiated and somatic NIH-3T3 cells. Relative expression levels of dCas9-VPR transgene and L1Md-Tf elements after 48h treatment with 1 μg/ml of doxycycline in female somatic NIH-3T3 cells. Expression levels were obtained by qPCR and analyzed with the $2^{-\Delta\Delta C_t}$ method, using Gapdh as housekeeping gene. Data is normalized to the housekeeping gene and represented as fold change over the control (doxycycline-untreated cells). Cas9, L1_ORF2 and L1_Tf correspond to primer pairs that amplify a genomic region of the dCas9-VPR transgene, ORF2 coding region of L1 elements and L1Md-Tf elements, respectively. H6 is a NIH-3T3 cell clone without sgRNAs for L1Md-Tf elements. H6_sgRNA represents the population of cells after transfection of sgRNAs for L1Md-Tf elements into H6 clone.

4. Discussion

After treating differentiated cells – such as NPCs – with selection drugs, although a great increase in expression of dCas9-VPR transgene was observed (**Figure 6**) in comparison to the results without selection drugs (**Figure 1C**), it only translated into a small increase in L1Md-Tf expression. This suggests that the selection drugs may, indeed, be increasing the accessibility of the transgenes (dCas9-VPR and sgRNAs) to the transcriptional machinery by creating an open chromatin configuration near each transgene of interest. However, the increased efficiency of the CRISPRa system does not translate into a proportional increase in the expression levels of L1Md-Tf elements, which indicates that the level of silencing occurring locally, at the 5'UTR of L1Md-Tf elements, is hampering the binding between the activation complex and the promoter region of these elements. Following induction of dCas9-VPR with doxycycline, somatic NIH-3T3 cells showed L1Md-Tf expression levels (**Figure 9**) similar to the ones observed in differentiated NPCs without selection drugs (**Figure 1C**), which supports the previously-mentioned hypothesis.

The gradual decrease in expression of dCas9-VPR transgene that is observed during the mESC differentiation process (**Figure 1B**) suggests that, indeed, some structural changes might occur during differentiation, possibly at the chromatin level, that may lead to a closed chromatin configuration. The same trend was observed for the transcriptional repression system, in which the expression levels of KRAB-ZFP transgene decrease during differentiation (**Figure 2B**). Nevertheless, the expression levels of KRAB-ZFP transgene after 4 days of differentiation should still be enough to induce, at least, some repression of L1Md-Tf elements, which is not happening since the repression system after 4 days of differentiation is no longer able to interfere with L1Md-Tf expression levels. Only after increasing greatly the expression levels of KRAB-ZFP transgene, by using the selection drugs during differentiation, we are able to observe that some L1Md-Tf repression is maintained after 3 days of differentiation (**Figure 5A**). These results suggest, once again, that the implemented interference system is functional and becomes more efficient when selection drugs are used, but not being able to proportionally interfere with L1Md-Tf expression levels, while the cells become more differentiated.

Undifferentiated mES cells, on the other hand, show a completely different behavior, with high expression levels of activation/repression transgenes and a considerably higher level of interference on L1Md-Tf expression (**Figure 1** and **Figure 2**). Furthermore, no significant changes are observed when selection drugs are used, neither at the expression level of the transgene nor of L1Md-Tf elements (**Figure 3**). These results suggest that the implemented translational activation/repression systems in undifferentiated mES cells are functional and efficient, even when selection drugs are not used. Thus, we may hypothesize that, at the undifferentiated state, a closed chromatin configuration is not present near the transgenes of interest and that the promoter region of L1Md-Tf elements is more accessible to the activation/repression complexes, than it is in further differentiated cells, which would explain the greater levels of L1 interference that one can achieve in undifferentiated mES cells.

After analyzing all the obtained results, we conclude that, although the implemented interference systems show a great potential in interfering with L1Md-Tf expression in cellular models with different stages of differentiation, their efficiency during differentiation and in differentiated cells is still sub-optimal. Here, two problems may arise that decrease the efficiency of these systems: i) low transcription levels of the activation/repression transgenes due to closed chromatin configuration, that is possibly established during the differentiation process, and ii) low accessibility of L1Md-Tf 5'UTR to the activation/repression complexes, due to epigenetic silencing marks that are deposited on these regions. As previously mentioned, growing the cells in the presence of selection drugs may impose an open chromatin configuration in the drugs resistance genes, that are located adjacent to the activation/repression transgenes, favoring the accessibility of the transgenes by the transcriptional machinery of the cells. Thus, the use of selection drugs seems to solve problem (i). Regarding problem (ii), (Wiznerowicz et al., 2007) have shown that epigenetic silencing of L1 elements occur once mESCs undergo differentiation. Thus, to be able to considerably interfere with L1Md-Tf expression levels during differentiation and in differentiated cells, besides increasing the expression levels of activation/repression transgenes, one must also remove the "epigenetic armor" from the 5'UTR of L1Md-Tf elements to increase the accessibility of these regions. Only then, one would be able to establish a stronger L1 interference system that could be used to study the impacts of L1 expression levels in X-chromosome inactivation and genome stability.

4.1. Future work

In general, more biological replicates of the experiments should be obtained to increase the robustness of the data, not only in NIH-3T3 cells but also in the previously mentioned experiments with mES cells and NPCs.

To correlate the translation of dCas9 and ORF1p, at the protein level, immunofluorescence images of doxycycline-treated and untreated B4 and D4 mES cell clones could be acquired with anti-dCas9 and anti-ORF1p primary antibodies, simultaneously. Then, intensity quantification could be performed on ImageJ to i) determine the level of ORF1p signal enhancement that is induced by the doxycycline treatment, ii) evaluate if the cells that show a higher enhancement of ORF1p signal are, simultaneously, dCas9-positive cells and also iii) determine the percentage of dCas9-positive cells that show a strong ORF1p signal enhancement.

Finally, a higher level of perturbation could be achieved if we were able to interfere simultaneously with all currently active subfamilies of L1 elements (L1Md-A, L1Md-Tf and L1Md-Gf) instead of interfering with L1Md-Tf elements only. Briefly, this could be achieved by cloning two additional pairs of sgRNAs, one against L1Md-A elements and another against L1Md-Gf elements, into the pLK01-sgTfmono2-3 plasmid. This new construct could then be used to establish a stable and clonal CRISPRa cell line.

4.2. Concluding remarks

Transposable elements are DNA sequences with the ability to change their position in the genome, accounting for more than 50% of the mouse genome. Among these, L1 retrotransposons have remained highly active in the mouse lineage. Although L1 elements are usually silenced by epigenetic marks, there are developmental time windows, like the early developmental stages, during which L1 elements are not silenced or fully silenced. This time windows represent a great window of opportunity for the activation of L1 elements, which has been hypothesized to play an important role in development. Here, one appealing hypotheses is that the activation of L1 elements during early development may facilitate the process of X-chromosome inactivation. Thus, the development of functional cellular models in which we are able to interfere (by activation or repression) with L1 expression genome-wide is essential to further investigate the role of L1 elements in XCI and in genome stability.

In the present study, an attempt to optimize L1 perturbation systems during differentiation of female mES cells and in female differentiated NPCs was tested, by culturing the cells with selection drugs. The obtained preliminary results suggest that the use of selection drugs leads to a considerable improvement in the expression levels of the activation/repression transgenes, but not in the levels of L1Md-Tf perturbation. Nevertheless, both transcriptional activation and repression systems are apparently functional and lead to L1Md-Tf expression interference, in comparison to previous results without the use of selection drugs. Furthermore, a CRISPRa system for L1Md-Tf activation was also implemented in fully differentiated mouse embryonic fibroblasts (female NIH-3T3 cells). Preliminary results suggest that the activation system is functional in NIH-3T3 cells, although only a modest level of L1Md-Tf perturbation has been achieved.

We hypothesize that L1Md-Tf perturbation levels may be considerably improved if one is able to specifically remove the “epigenetic armor” that is deposited on the promoter region of these elements, once differentiation is induced. Thus, cellular models with optimized levels of genome-wide L1Md-Tf perturbation could possibly be obtained, allowing for the investigation of the role of L1 elements in XCI and in genome stability, during the several stages of development – from undifferentiated to fully differentiated states.

5. References

Adey, N. B., Schichman, S. A., Graham, D. K., Peterson, S. N., Edgell, M. H., & Hutchison, C. A. (1994). Rodent L1 evolution has been driven by a single dominant lineage that has repeatedly acquired new transcriptional regulatory sequences. *Molecular Biology and Evolution*, *11*(5), 778–789.

Bourque, G., Burns, K. H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin, H. L., Macfarlan, T. S., Mager, D. L., & Feschotte, C. (2018). Ten things you should know about transposable elements. *Genome Biology*, *19*(1).

Burwinkel, B., & Kilimann, M. (1998). Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. *Journal of Molecular Biology*, *277*(3), 513–517.

Buzin, C. H., Mann, J. R., & Singer-Sam, J. (1994). Quantitative RT-PCR assays show Xist RNA levels are low in mouse female adult tissue, embryos and embryoid bodies. *Development*, *120*(12), 3529–3536.

Chaumeil, J., Okamoto, I., & Heard, E. (2004). X-chromosome inactivation in mouse embryonic stem cells: analysis of histone modifications and transcriptional activity using immunofluorescence and FISH. *Methods in Enzymology*, *376*, 405–419.

Chow, J. C., Ciaudo, C., Fazzari, M. J., Mise, N., Servant, N., Glass, J. L., Attreed, M., Avner, P., Wutz, A., Barillot, E., Greally, J. M., Voinnet, O., & Heard, E. (2010). LINE-1 Activity in Facultative Heterochromatin Formation during X Chromosome Inactivation. *Cell*, *141*(6), 956–969.

DeBerardinis, R. J., & Kazazian Jr., H. H. (1999). Analysis of the promoter from an expanding mouse retrotransposon subfamily. *Genomics*, *56*(3), 317–323.

Feng, Q., Moran, J. v, Kazazian Jr., H. H., & Boeke, J. D. (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell*, *87*(5), 905–916.

Fuentes, D. R., Swigut, T., & Wysocka, J. (2018). Systematic perturbation of retroviral LTRs reveals widespread long-range effects on human gene regulation. *ELife*, *7*.

Gilbert, N., Lutz-Prigge, S., & Moran, J. v. (2002). Genomic Deletions Created upon LINE-1 Retrotransposition. *Cell*, *110*(3), 315–325.

Goodier, J. L., Ostertag, E. M., Du, K., Kazazian, H. H., & Jr. (2001). A Novel Active L1 Retrotransposon Subfamily in the Mouse. *Genome Research*, *11*(10), 1677.

Han, J. S., Szak, S. T., & Boeke, J. D. (2004). Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature*, *429*(6989), 268–274.

Honda, T., Nishikawa, Y., Nishimura, K., Teng, D., Takemoto, K., & Ueda, K. (2020). Effects of activation of the LINE-1 antisense promoter on the growth of cultured cells. *Scientific Reports*, *10*(1), 1–8.

Jachowicz, J. W., Bing, X., Pontabry, J., Bošković, A., Rando, O. J., & Torres-Padilla, M.-E. (2017). LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nature Genetics*, *49*(10), 1502–1510.

Liu, N., Lee, C. H., Swigut, T., Grow, E., Gu, B., Bassik, M. C., & Wysocka, J. (2017). Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature*, *553*(7687), 228–232.

Luan, D. D., Korman, M. H., Jakubczak, J. L., & Eickbush, T. H. (1993). Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell*, *72*(4), 595–605.

Lyon, M. F. (1961). Gene Action in the X-chromosome of the Mouse (Mus musculus L.). *Nature*, *190*(4773), 372–373.

Lyon, M. F. (1962). Sex Chromatin and Gene Action in the Mammalian X-Chromosome. *American Journal of Human Genetics*, *14*(2), 135.

Lyon, M. F. (1998). X-chromosome inactivation: a repeat hypothesis. *Cytogenetics and Cell Genetics*, *80*(1–4), 133–137.

Mathias, S., Scott, A., Kazazian, H., Boeke, J., & Gabriel, A. (1991). Reverse transcriptase encoded by a human transposable element. *Science, 254*(5039), 1808–1810.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., & Haaf, T. (2000). Demethylation of the zygotic paternal genome. *Nature*, *403*(6769), 501–502.

Meissner, A., Mikkelsen, T. S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B. E., Nusbaum, C., Jaffe, D. B., Gnirke, A., Jaenisch, R., & Lander, E. S. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, *454*(7205), 766.

Moran, J. v, Holmes, S. E., Naas, T. P., DeBerardinis, R. J., & Boeke, J. D. (1996). High frequency retrotransposition in cultured mammalian cells. *Cell*, *87*(5), 917–927.

Percharde, M., Lin, C.-J., Yin, Y., Guan, J., Peixoto, G. A., Bulut-Karslioglu, A., Biechele, S., Huang, B., Shen, X., & Ramalho-Santos, M. (2018). A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. *Cell*, *174*(2), 391-405.

Smith, Z. D., Chan, M. M., Mikkelsen, T. S., Gu, H., Gnirke, A., Regev, A., & Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature*, *484*(7394), 339–344.

Sookdeo, A., Hepp, C. M., McClure, M. A., & Boissinot, S. (2013). Revisiting the evolution of mouse LINE-1 in the genomic era. *Mobile DNA*, *4*(1), 1–15.

Symer, D. E., Connelly, C., Szak, S. T., Caputo, E. M., Cost, G. J., Parmigiani, G., & Boeke, J. D. (2002). Human L1 Retrotransposition Is Associated with Genetic Instability In Vivo. *Cell*, *110*(3), 327–338.

Thomas, C. A., Paquola, A. C. M., & Muotri, A. R. (2012). LINE-1 Retrotransposition in the Nervous System. *Annual Review of Cell and Developmental Biology*, *28*, 555–573.

Wiznerowicz, M., Jakobsson, J., Szulc, J., Liao, S., Quazzola, A., Beermann, F., Aebischer, P., & Trono, D. (2007). The Krüppel-associated box repressor domain can trigger de novo promoter methylation during mouse early embryogenesis. *Journal of Biological Chemistry*, *282*(47), 34535–34541.

Wutz, A., & Jaenisch, R. (2000). A Shift from Reversible to Irreversible X Inactivation Is Triggered during ES Cell Differentiation. *Molecular Cell*, *5*(4), 695–705.

Xiao-Jie, L., Hui-Ying, X., Qi, X., Jiang, X., & Shi-Jie, M. (2015). LINE-1 in cancer: multifaceted functions and potential clinical implications. *Genetics in Medicine*, *18*(5), 431–439.