Screening for transposon regulators in mouse development

Dissertation

Zur Erlangung des Grades

Doktor der Naturwissenschaften

Am Fachbereich Biologie

Der Johannes Gutenberg-Universität Mainz

Styliani-Eirini Kanta

geboren am 08.03.1994 in Kavala

Mainz, 2025

Dekan: Prof. Dr. Eckhard Thines

1. Berichterstatter: Prof. Dr. Rene Ketting

2. Berichterstatter: Dr. Jan Padeken

Tag der mündlichen Prüfung: 15. Juli 2025

In loving memory of my beloved grandma,
and to my parents, sister and friends.
Thank you for always being there for me.

Table of Contents

Acknowledgments

Abstract

Transposable elements (TEs) are mobile genetic elements that contribute to genome evolution but pose potential threats to genomic stability. Their activity is tightly regulated by host defense mechanisms, including DNA methylation, histone modifications, and transcription factor (TF)-mediated control. This thesis investigates the regulation of TEs in two model systems: mouse embryonic stem cells (mESCs) and the mouse male germline.

In the first part of this thesis, we aimed to establish a CRISPR/Cas9-based screening using an innovative endogenous readout to identify TE regulators in mESCs. By generating KAP1-degron and *Suv39h*-dKO cell lines we aimed to monitor transcriptional and translational activation of endogenous ERV and LINE-1 elements. Despite significant efforts, technical challenges prevented the successful implementation of the screening approach. However, this work highlights the complexities of studying endogenous TE regulation and provides a foundation for future optimization.

The second part of this thesis investigates TE reactivation during DNA methylation loss in mouse spermatogenesis. Using *Dnmt3C*$^{KO/KO}$ mice, we found that young ERVs remain active throughout postnatal germ cell stages in the absence of DNA methylation, and in contrast to other studies, we observed LINE-1 reactivation before meiosis, which further increased during meiosis. DNA pulldown assays identified NRF1, a DNA methylation-sensitive TF, as a potential regulator of unmethylated TEs in the male germline and chromatin profiling analysis confirmed NRF1 binding to unmethylated TEs. Conditional knockout of *Nrf1* in *Dnmt3C*$^{KO/KO}$ mice resulted in significant downregulation of IAP protein expression prior to meiosis, demonstrating NRF1 as a trans-activator of these elements. Additionally, H3K27me3 was identified as a potential barrier to NRF1 binding, suggesting a complex interplay between histone modifications and TF-mediated regulation.

Together, these studies enhance our understanding of the multilayered regulatory networks controlling TE activity in the mouse genome. While challenges remain in developing high-throughput approaches for studying TE regulation, our findings emphasize the importance of both epigenetic modifications and TFs in balancing TE repression and activation during development.

Zusammenfassung

Transposable Elemente (TEs) sind weit verbreitete genetische Elemente, die zur Evolution des Genoms beitragen, aber auch potenzielle Gefahren für die genomische Stabilität darstellen. Ihre Aktivität wird durch Abwehrmechanismen des Wirts streng reguliert, darunter DNA-Methylierung, Histon-Modifikationen und Transkriptionsfaktor(TF)-vermittelte Kontrolle. Diese Dissertation untersucht die Regulation von TEs in zwei Modellsystemen: Maus-embryonalen Stammzellen (mESCs) und der männlichen Keimbahn der Maus.

Im ersten Teil dieser Arbeit versuchten wir, ein CRISPR/Cas9-basiertes Screening mit einem innovativen endogenen Readout zu etablieren, um TE-Regulatoren in mESCs zu identifizieren. Durch die Generierung von KAP1-Degron und *Suv39h*-dKO Zelllinien wollten wir die transkriptionelle und translationale Aktivierung endogener ERV- und LINE-1-Elemente überwachen. Trotz erheblicher Bemühungen verhinderten technische Herausforderungen die Umsetzung des Screening-Ansatzes. Dennoch hebt diese Arbeit die Komplexität der Untersuchung der Regulation endogener TEs hervor und bietet eine Grundlage für zukünftige Optimierungen.

Der zweite Teil dieser Dissertation untersucht die Reaktivierung von TEs beiVerlust von DNA-Methylierung in der Maus-Spermatogenese. Mithilfe von *Dnmt3C*$^{KO/KO}$ Mäusen stellten wir fest, dass junge ERVs während der postnatalen Keimzell-Stadien aktiv bleiben, wenn sie unmethyliert bleiben. Im Gegensatz zu anderen Studien beobachteten wir zudem eine Reaktivierung von LINE-1 vor der Meiose, die sich während der Meiose weiter verstärkte. DNA-Pulldown-Assays identifizierten NRF1, einen DNA-Methylierungs-sensitiven TF, als potenziellen Regulator unmethylierter TEs in der männlichen Keimbahn, und weitere Chromatin-Profilierung Analysen bestätigten die Bindung von NRF1 an unmethylierte TEs. Der konditionelle Knockout von *Nrf1* in *Dnmt3C*$^{KO/KO}$ Mäusen führte zu einer signifikanten Herunterregulierung der IAP-Proteinexpression vor der Meiose und belegte, NRF1 als Transaktivator dieser Elemente. Darüber hinaus wurde H3K27me3 als potenzielle Barriere für die NRF1-Bindung identifiziert, was auf ein komplexes Zusammenspiel zwischen Histon-Modifikationen und TF-vermittelter Regulation hinweist.

Insgesamt liefern diese Studien neue Einblicke in die vielschichtigen regulatorischen Netzwerke, die die Aktivität von TEs im Mausgenom steuern. Obwohl weiterhin Herausforderungen bei der Entwicklung hochdurchsatz-fähiger Ansätze zur Untersuchung der TE-Regulation bestehen, betonen unsere Ergebnisse die Bedeutung sowohl epigenetischer Modifikationen als auch TFs beim Ausgleich von TE-Repression und -Aktivierung während der Entwicklung.

List of Figures

List of Tables

Supplementary material

List of Abbreviations

5mC	5-Methylcytosine
ADD	ATRX-DNMT3-DNMT3L domain
C	Control
CPM	Counts per million
CUT&Tag	Cleavage Under Targets & Tagmentation
DBD	DNA binding domain
dKO	double knock-out: *Nrf1*$^{cKO/KO}$*; Dnmt3C*$^{KO/KO}$
DNMTs	DNA methyltransferases
DNMT3L	DNMT3-like enzyme
Dox	Doxycycline
E	Embryonic day
ERVs	Endogenous retroviruses
ERVKs	ERV-K promoter
ERVLs	ERV-L promoter
F	Offspring Generation
FACS	Fluorescence-activated cell sorting
FBS	Fetal Bovine Serum
FC	Fold Change
FISH	Fluorescent In Situ Hybridization
CGI	CpG Island
H3K4	Histone H3 Lysine 4
H3K9	Histone H3 Lysine 9
H3K27	Histone H3 Lysine 27
HP1	Heterochromatin Protein 1

HUSH	Human Silencing Complex
IAP	Intracisternal A-type Particle
IF	Immunofluorescence
IP	Immunoprecipitation
IP-MS	Immunoprecipitated Pulldown followed by Mass-Spectrometry
KRAB-ZFPs	Krueppel-associated box domain containing zinc finger proteins
KAP1	KRAB-Associated Protein 1
Knockin	KI
Knockout	KO
L1s	LINE1s
LINEs	Long Interspersed Nuclear particle
LTRs	Long terminal repeats
MaLR	Mammalian-apparent LTR retrotransposons
MBD	Methyl binding domain proteins
mESCs	mouse Embryonic Stem Cells
me2	dimethylation
me3	trimethylation
MLV	Mouse Leukemia Virus
NGS	Next Generation Sequencing
NRF1	Nuclear Respiratory Factor 1
ORFs	Open Reading Frames
ORF1p	Open Reading Frame 1 protein
P	Postnatal day
PCNA	Proliferating cell nuclear antigen
PFA	Paraformaldehyde
PGCs	Primordial Germ Cells
PIC	Protease Inhibitor Cocktail

piRNAs	PIWI interacting RNAs
PIWI	PIWI-clade Argonaute
Pol II	RNA polymerase II
PRC2	Polycomb Repressive Complex 2
PWWP	Pro-Trp-Trp-Pro
qPCR	quantitative PCR
RNP	Ribonucleoprotein
ROI	Region of Interest
RT	Room Temperature
RT-qPCR	Reverse transcription quantitative real-time PCR
SAM	S-Adenosylmethionine
SETDB1	SET Domain Bifurcated Histone Lysine Methyltransferase 1 (ESET)
SINEs	Short Interspersed Nuclear Elements
Spc	Spermatocytes
Spg	Spermatogonia
SSCs	Spermatogonial Stem Cells
SUV39h	Suppressor of Variegation (3-9) Homologue
SVA	SINE-VNTR-Alu
TEs	Transposable Elements
Tet-On	Tetracycline-controlled transcriptional activation
TFs	Transcription Factors
TKO	*Dnmt3A*$^{KO/KO}$*; Dnmt3B*$^{KO/KO}$*; Dnmt1*$^{KO/KO}$ mESCs
TIRs	Terminal Inverted Repeats
TPM	Transcripts per Million
TSDs	Tandem Site Duplications
TSS	Transcription Start Site

TPRT	Target-Primed Reverse Transcription
UHRF1	Ubiquitin-like containing PHD and RING finger domains, 1
UTR	Untranslated Region
WB	Western Blot
WT	Wild-type
x	Crossed

1. Introduction

1.1 Transposable elements

1.1.1 Discovery of transposable elements

Transposable elements (TEs), also known as “jumping genes”, are highly abundant and repetitive genetic sequences capable of changing their location within the genome. During 1940 to 1950, Barbara McClintock first discovered genetic mobility in maize chromosomes correlating TEs to gene expression and phenotypic variation (McCLINTOCK 1950). Her revolutionary findings predated the DNA’s double-helix structure discovery, and challenged the prevailing concept of genetic stability in a male-dominated scientific community at the time (Wu 2024; Watson and Crick 1953; Maddox 2003). Despite facing gender discrimination, when TEs were found across different species, her work was validated with the Nobel Prize in Physiology or Medicine in 1983. Today, TEs are known to pervade all kingdoms of life, comprising nearly half of the genome in humans and mice (Wells and Feschotte 2020); (Kidwell 2002). TEs are now recognized as significant contributors in genome regulation and evolution while also serving as valuable research tools for genetic modification.

1.1.2 Classes of transposable elements

TEs are classified in two major groups based on their transposition mechanism: Class I or retrotransposons, and class II or DNA transposons (Wicker et al. 2007).

Class I or retrotransposons move via a “copy-and-paste” mechanism that involves an RNA intermediate. The integration process requires the transcription of the TE sequence into mRNA, which in turn is reverse transcribed into DNA. The resulting cDNA complement is then inserted into a new position in the genome (Boeke et al. 1985). Retrotransposons are broadly categorised into two main subclasses: the long terminal repeats (LTRs) and the non-LTRs. The non-LTR retrotransposons include the superfamilies of long and short interspersed nuclear elements (LINEs and SINEs), while the endogenous retrovirus (ERV) superfamily belongs to the LTR subclass (Bourque et al. 2018).

Class II or DNA transposons move via a “cut-and-paste” mechanism without an RNA intermediate. DNA transposons are flanked by terminal inverted repeats (TIRs), recognized by their transposase, an enzyme encoded by autonomous DNA transposons, to enable

mobilization into a new location in the genome (Muñoz-López and García-Pérez 2010) (Figure 1.1A). During the transposition short Target Site Duplications (TSDs) are created flanking the insert. DNA transposons represent a small portion (~3%) of the mice and human genomes and have been inactive in both genomes (Lander et al. 2001; Mouse Genome Sequencing Consortium et al. 2002). In humans, DNA transposons have been inactive for over 40 million years (Myrs) (Pace and Feschotte 2007). Although inactive, DNA transposons have been engineered into valuable molecular biology tools (e.g PiggyBac system) having significant applications on genetic manipulation and genome engineering regulation (Wilson, Coates, and George 2007).

Since Class I or retrotransposons represent the only active elements in humans and mice, the following chapter will focus on their biology and impact on genome regulation and evolution (Figure 1.1B).

1.1.2.1 Long Terminal Repeat (LTR) retrotransposons

Long Terminal Repeat (LTR) retrotransposons are mobile genetic elements occupying approximately 8-10% of the human and mouse genome (Cordaux and Batzer 2009; Mouse Genome Sequencing Consortium et al. 2002). While LTR elements exhibit substantial transcriptional activity in mice, they remain predominantly inactive in humans (Mills et al. 2007; Cordaux and Batzer 2009; Crichton et al. 2014). LTR elements possess specific structural features important for mobilization. A typical autonomous LTR element contains two identical LTR sequences around 300 to 1,000 bp long flanking a coding region ranging in size from 6,000 to 9,000 bp (Sandmeyer and Craig 2015). This internal coding region harbors at least two essential genes: group-specific antigen (*Gag*), polymerase (*Pol*). *Gag* encodes proteins assembling virus-like particles (VLPs) in the cytoplasm, while *Pol* encodes four proteins: a protease (PR), a reverse-transcriptase (RT), RNase H (RNH) and an integrase (IN) (Wicker et al. 2007) (Figure 1.1C).

The retrotransposition cycle of LTRs starts with the mRNA transcription initiated from the 5' LTR promoter by host RNA polymerase II. The primary transcript undergoes standard cellular processing including 5' capping and 3' polyadenylation, prior to nuclear export (Hughes 2015). Once exported to the cytoplasm, the *Gag-Pol* bicistronic transcript functions both as a template for protein translation and as genomic RNA for novel insertions. The GAG and POL proteins are produced through two different mechanisms: either by the cleavage of a single ORF fusion or by a ribosomal frameshift between the two ORFs (Gao et al. 2003). The reverse transcription occurs within the VLPs, where host tRNAs bind to the

primer binding site (PBS). PBSs serve as primers and the RT domain of POL synthesizes complementary DNA using the packaged RNA as a template (Lin and Levin 1998). This process yields a double-stranded DNA copy of the element and the resulting cDNA forms a complex with integrase proteins for nuclear transport (Hughes 2015). Once in the nucleus, the integrase mediates the insertion of the DNA into a new genomic location (Curcio, Lutz, and Lesage 2015).

LTR retrotransposons can be classified into three main subfamilies based on their structure and evolutionary origin: Endogenous retroviruses (ERVs), *Gypsy* and *Copia (Wicker et al. 2007)*. While ERVs are just one superfamily of LTR retrotransposons, they represent the predominant type found in the mouse genome. Thus, the terms “ERVs” and “LTRs” are often used interchangeably. ERVs originated from ancient exogenous retroviral infections of germline cells (X. Lu 2024; Johnson 2019; Vogt 1997). The primary distinction between ERVs and *bona fide* retroviruses resides in their transmission pattern. While exogenous retroviruses (XRVs) can form infectious particles and transmit horizontally within organisms and cells, ERVs are (with a few exceptions (Tarlinton, Meers, and Young 2006)) primarily non-infectious: they restrict their mobility to intracellular transposition and are inherited vertically (Magiorkinis et al. 2012). This difference stems from the presence or absence of a functional envelope (env) gene (Magiorkinis et al. 2012). Most ERVs either lack the env gene entirely or they retain an env-like gene, which carries mutations that result in a typically dysfunctional and non-infectious gene (Magiorkinis et al. 2012).

In mammalian genomes, the majority of ERVs exist as solo-LTRs, which have formed via homologous recombination between the 5' and 3' LTRs of full-length elements (M. Chen et al. 2024). Besides solo-LTRs, ERVs are further categorized into three classes (Figure 1.1D): class I or ERV-1, class II or ERV-K, and class III or ERV-L, based on the family of XRVs they are related to (Magiorkinis, Belshaw, and Katzourakis 2013; Stocking and Kozak 2008). ERV-1 elements comprise the smallest fraction (~1%) of ERVs in the mouse genome and a characteristic example includes Murine Leukemia Virus (MuLV). ERV-K elements constitute 5% of the mouse genome and includes the Intracisternal A Particles (IAPs), RLTR10, MMERVK10C. IAPs are young, highly mobile and mutagenic elements, comprising 1,000 to 2,000 copies in the mouse genome, hundreds of which remain active for retrotransposition (Magiorkinis et al. 2012; C. Lu, Contreras, and Peterlin 2011). Class III elements represent the most abundant ERVs in mice (5-6%) and include the Murine Endogenous Retrovirus-L (MERVL) and the non-autonomous MalRs (Smit 1993; Stocking and Kozak 2008). Active ERVs, such as IAPs, exemplify the persistent activity of ERVs to continue to generate new insertions and influence genome stability.

1.1.2.2 non-LTR retrotransposons

1.1.2.2.1 Long Interspersed Nuclear Element 1 (LINE-1)

Long interspersed nuclear elements (LINEs) are autonomous non-LTR retrotransposons and can be classified into five main subfamilies in eukaryotes: RTE, R2, I, Jockey and LINE-1 (L1) elements (Wicker et al. 2007). Among all LINEs, LINE-1 elements are particularly abundant occupying 20% and 17% of the mice and human genome, respectively (Mouse Genome Sequencing Consortium et al. 2002) (Lander et al. 2001). Although LINEs exist in ~500,000 to 600,000 copies in both genomes, their activity differs significantly: only 3,000 copies are considered potentially active for retrotransposition in mice and just ~80 to 100 copies remain active in humans (DeBerardinis et al. 1998) (Brouha et al. 2003).

Typical full-length mammalian LINE-1 elements span 6,000 to 7,000bp long and contain distinct structural components essential for mobilization: a 5' untranslated region (UTR) region, two open reading frames (ORFs) and a 3' UTR (Bodak, Yu, and Ciaudo 2014) (Figure 1.1E). The 5' UTR functions as an internal RNA polymerase II promoter, exhibiting species-specific characteristics: mouse elements contain GC-rich tandemly repeated monomers, whereas human elements comprise approximately 900 base pairs without repeated motifs (DeBerardinis and Kazazian 1999) (Khan, Smit, and Boissinot 2006) (Becker et al. 1993; Alexandrova et al. 2012). In both humans and mice the 5' UTRs harbor transcription factor binding motifs (e.g YY1) important for their retrotransposition (Becker et al. 1993; Athanikar, Badge, and Moran 2004). In mice, active LINE-1 subfamilies (e.g L1MdF, L1MdA, L1MdTf, and L1MdGf) are mainly classified based on their unique monomer sequences and arrangements within these promoter regions. This diversity of LINE-1 promoters in mice, can contribute to faster adaptation compared to humans by enabling frequent retrotransposition events and driving genomic and evolutionary innovation (Kong et al. 2022; Sookdeo et al. 2013) (Figure 1.1F).

LINE-1s encode two proteins crucial for retrotransposition: ORF1p and ORF2p. ORF1p, a 40kDa protein, forms homotrimeric complexes with high-affinity RNA-binding and nucleic acid chaperone activities (Martin and Garfinkel 2003). ORF2p, a 150kDa protein, contains both reverse transcriptase (RT) and endonuclease (EN) domains. The 3' UTR region consists of a polyadenylation signal and a poly(A) tail, which contribute to LINE-1 mRNA stability and facilitate retrotransposition (Belancio, Whelton, and Deininger 2007; Beck et al. 2011; Doucet et al. 2015) (Figure 1.1E).

LINE-1 retrotransposition occurs through a mechanism called Target-Primed Reverse Transcription (TPRT) (Luan et al. 1993; Cost et al. 2002). The process initiates with RNA

polymerase II-mediated transcription of LINE-1 mRNA, followed by nuclear export and the translation of ORF1p and ORF2p. These proteins preferentially bind in cis to their L1 encoding mRNA, and not to other cellular RNAs (Kulpa and Moran 2006) forming ribonucleoprotein (RNP) complexes located in cytoplasmic stress granules (Kazazian 2004; Goodier et al. 2007; Sil et al. 2023) (Figure 1.1E). Once the RNPs are imported back into the nucleus, nicks in the genomic DNA are generated by the endonuclease activity of ORF2p at the target site consensus sequence (5'-TTTT/AA-3') (Mita et al. 2018; Feng et al. 1996; X. Zhang et al. 2024). Subsequently, RNA-DNA hybrids are formed at the target site cleavage enabling reverse transcription via the RT activity of ORF2p (Mathias et al. 1991). The synthesized cDNA is subsequently integrated at the insertion site (Cost et al. 2002). Understanding these molecular mechanisms has revealed how LINE-1 elements contribute to both genome instability and evolution, contributing to both physiological and pathological processes.

1.1.2.2.2 Short Interspersed Nuclear Elements (SINEs)

Short interspersed nuclear elements (SINEs) are non-autonomous retrotransposons spanning 100 to 700bp that occupy approximately 13% of the human genome and 8% of the mouse genome (Schmitz 2012; Mandal and Kazazian 2008). In humans, Alu elements represent the most abundant SINE family (over 1 million copies), while a smaller number of SVA elements (~2,700 copies) are among the youngest and most active in primates (Häsler and Strub 2006; H. Wang et al. 2005). In mice, B1 and B2 are the most abundant elements of the SINE family (Kass and Jamison 2007). SINEs lack protein-coding regions (Figure 1.1G) and hijack LINE-1 retrotransposition machinery to mobilize, representing an example of molecular nested parasitism. Specifically, SINE transcripts form complexes with LINE ORF2 proteins and integrate into the genome through target-primed reverse transcription (TPRT), creating short target site duplications upon insertion (Dewannieux, Esnault, and Heidmann 2003). SINEs possess an internal RNA polymerase III promoter, and a 3' region essential for mobility, mimicking the L1 mRNA (Schmitz 2012; Kramerov and Vassetzky 2011). The promoter region determines their evolutionary origins: Alu and B1 elements harbor a 7SL promoter, while B2 elements are tRNA-related (Vassetzky and Kramerov 2013). Recent research has revealed that SINEs, although considered purely parasitic, have evolved to serve various regulatory functions in mammalian genomes, including roles in gene regulation as they can act as promoters, enhancer or insulator elements (X.-O. Zhang, Pratt, and Weng 2021) and could regulate pluripotency by hijacking transcription factors and act as enhancers (Ge 2017; Kravchenko and Tachibana 2025).

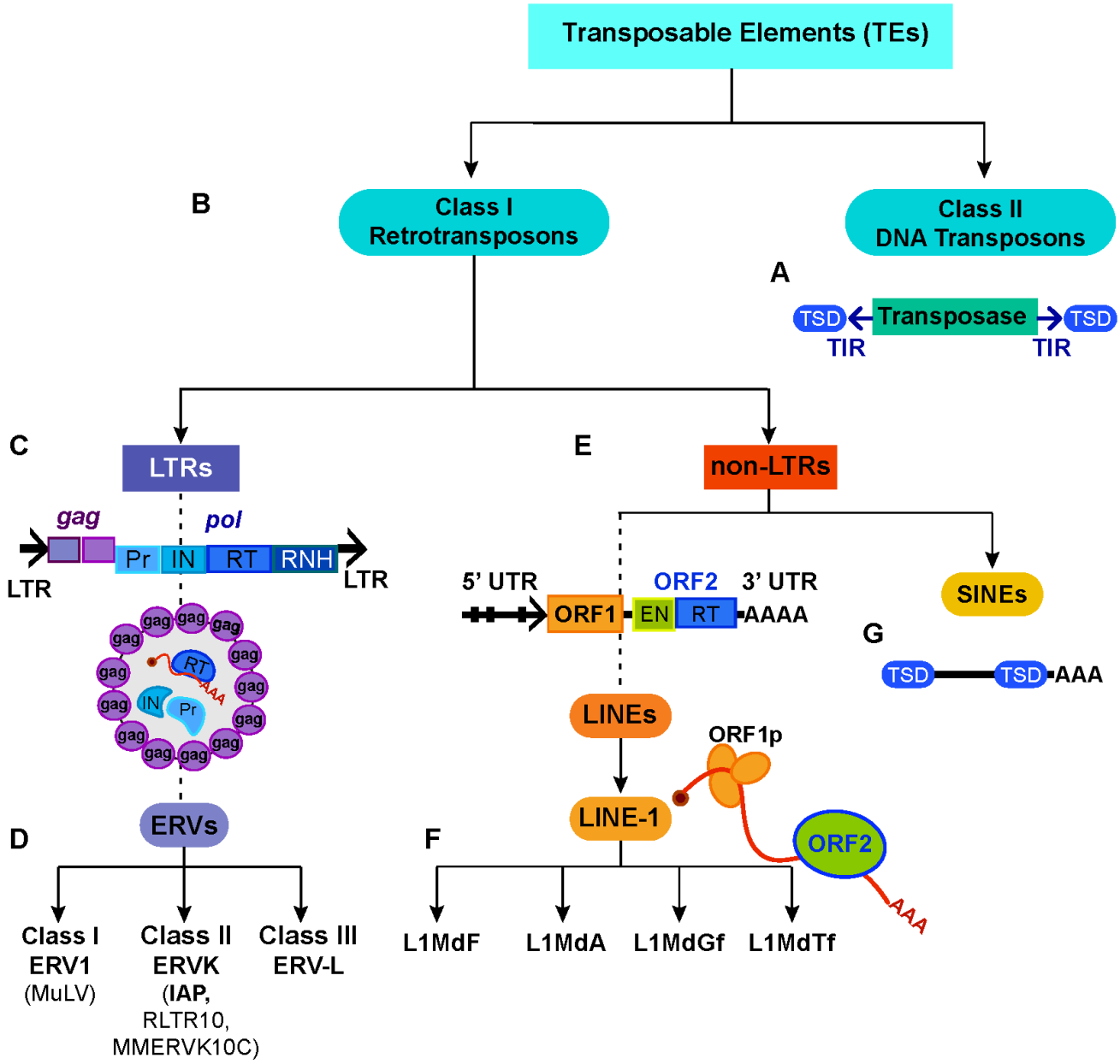


Figure 1.1. Classes of Transposable Elements (TEs). This classification is focused mostly on ERVs and LINE-1s, as they are the most abundant and active elements in mice (Dewannieux et al. 2004) and the focus of this thesis (therefore the other subfamilies of LTRs and LINEs are not depicted). **A.** Structure of DNA transposon. Terminal Inverted Repeats (TIRs) are flanking the transposase gene. Two target site duplications (TSD) are flanking the insert. **B.** Class I or retrotransposons: the only active TEs in mice. **C.** Structure of LTR elements. Two LTRs are flanking the polyprotein-coding genes (*gag*, *pol*); Protease (Pr), Integrase (IN), Reverse transcriptase (RT), RNAse H (RNH) domains. Virus-like particles formed in the cytoplasm; Adapted from the PhD thesis of Marius Walter. **D.** Subgroups of ERVs: ERV1, ERVK, ERVL. **E.** non-LTRs subdivided in LINEs and SINEs. Structure of LINEs; Promoter with tandem repeats (5' UTR); Endonuclease (EN) and Reverse transcriptase (RT) domains, poly-A tail (3' UTR); structure of LINE-1 RNP complex. **F.** Subfamilies of LINE-1. **G.** Structure of SINEs; lack of coding sequence.

1.1.3 TEs: From “junk DNA” to key contributors in evolution and development

Transposable elements have exhibited a positive role in evolutionary innovation, gene regulation and development, exhibiting both deleterious and beneficial effects on their host organisms (Platt, Vandewege, and Ray 2018).

A striking example of evolutionary innovation is the insertion of a SINE Alu element in the *TBXT* gene, resulting in tail-loss in humans and hominoid ancestors (Xia et al. 2024) (Figure 1.2A). When inserted into the genome, the hominoid-specific Alu element paired with the neighboring ancestral Alu element encoded in the reverse genomic orientation, leading to the formation of a secondary RNA structure and subsequently to an alternative splicing event (Xia et al. 2024). Another iconic example of a phenotypic change in response to adaptation, is the industrial melanism in the peppered moth caused by a DNA transposon insertion into the first intron of the *cortex* gene (Van’t Hof et al. 2016) (Figure 1.2A). This TE-mediated adaptive change enhanced moth survival through effective camouflage in industrial environments and reducing predator detection. TEs have also been evolutionarily co-opted to be involved in post-transcriptional regulatory networks. For instance, LINE-2 elements have given rise to microRNAs (miRNAs) that regulate gene expression by targeting L2-derived transcripts in their 3’ UTR. These miRNAs positively contribute to the maintenance of housekeeping genes and their dysregulation could lead to brain disease (Petri et al. 2019)

While TEs have contributed to evolutionary adaptation, they also have a profound impact in normal developmental processes. Their developmental importance initiates during Zygotic Genome Activation (ZGA), the process in which embryonic gene expression and maternal-to-zygotic transition begins. In mice, ZGA occurs at the two-cell stage, where MERVLs are among the first transcripts to be expressed and demonstrate a remarkable upregulation in two-cell embryos compared to oocytes (Peaston et al. 2004; Svoboda et al. 2004; Macfarlan et al. 2012). MERVLs can act as alternative promoters for ZGA genes (Macfarlan et al. 2012; Peaston et al. 2004) and are crucial for preimplantation, since the absence of MERVL transcripts results in embryonic lethality (Sakashita et al. 2023). Alongside MERVLs, LINE1s have crucial roles in early development. During ZGA, LINE-1 elements are also active and regulate chromatin accessibility and influence global chromatin organization (Jachowicz et al. 2017) and have been reported to act as enhancers (Xiufeng Li et al. 2024) (Figure 1.2B). In mESCs, LINE-1 transcripts were found to be important for self-renewal (Percharde et al. 2018). The developmental significance of TEs extends

beyond early embryogenesis, as co-opted retroviral genes (syncytins) have been shown to be essential for placental formation in both humans and mice (Keighley et al. 2023; Dupressoir et al. 2009).

Although TEs have proven to be beneficial, their selfish activity can be highly deleterious. TEs, by virtue of their mobility, can be mutagenic when disrupting coding regions causing transcriptional interference, genome instability and double-strand breaks (Hancks and Kazazian 2016). Subsequently, TEs have been associated with a repertoire of multiple cancer tissues and diseases. For instance, oncogenic Alu insertions in BRCA1/2 genes have been reported to increase the probability of breast cancer development (Teugels et al. 2005; Bouras et al. 2021; Qian et al. 2017) (Figure 1.2C). Additionally, LINE-1 ORF1p expression has emerged as a biomarker in many epithelial cancers, including ovarian carcinoma (Ardeljan et al. 2020). Beyond cancer, TEs have been implicated in various age-related conditions and neurodegenerative disorders. Their expression increases with age and their dysregulation has been linked to Alzeiheimer's disease via the activation of inflammatory pathways (Guo et al. 2018; Pabis et al. 2024). TE dysregulation can also have severe consequences during development. During early development, at specific developmental windows, TEs escape repression and threaten genome integrity. In the mouse germline, aberrant TE activity can cause meiotic failure leading to male infertility (Barau et al. 2016) (Figure 1.2D).

Throughout evolution, although TEs have been co-opted for both beneficial functions, their uncontrolled mobilization has been proven detrimental for the genome. However, this raises a fundamental question: How does the genome regulate TE expression to maintain beneficial functions while preventing uncontrolled mobilization? This dual nature of TEs has developed a complicated interplay between TEs and their host genomes, necessitating regulatory mechanisms.

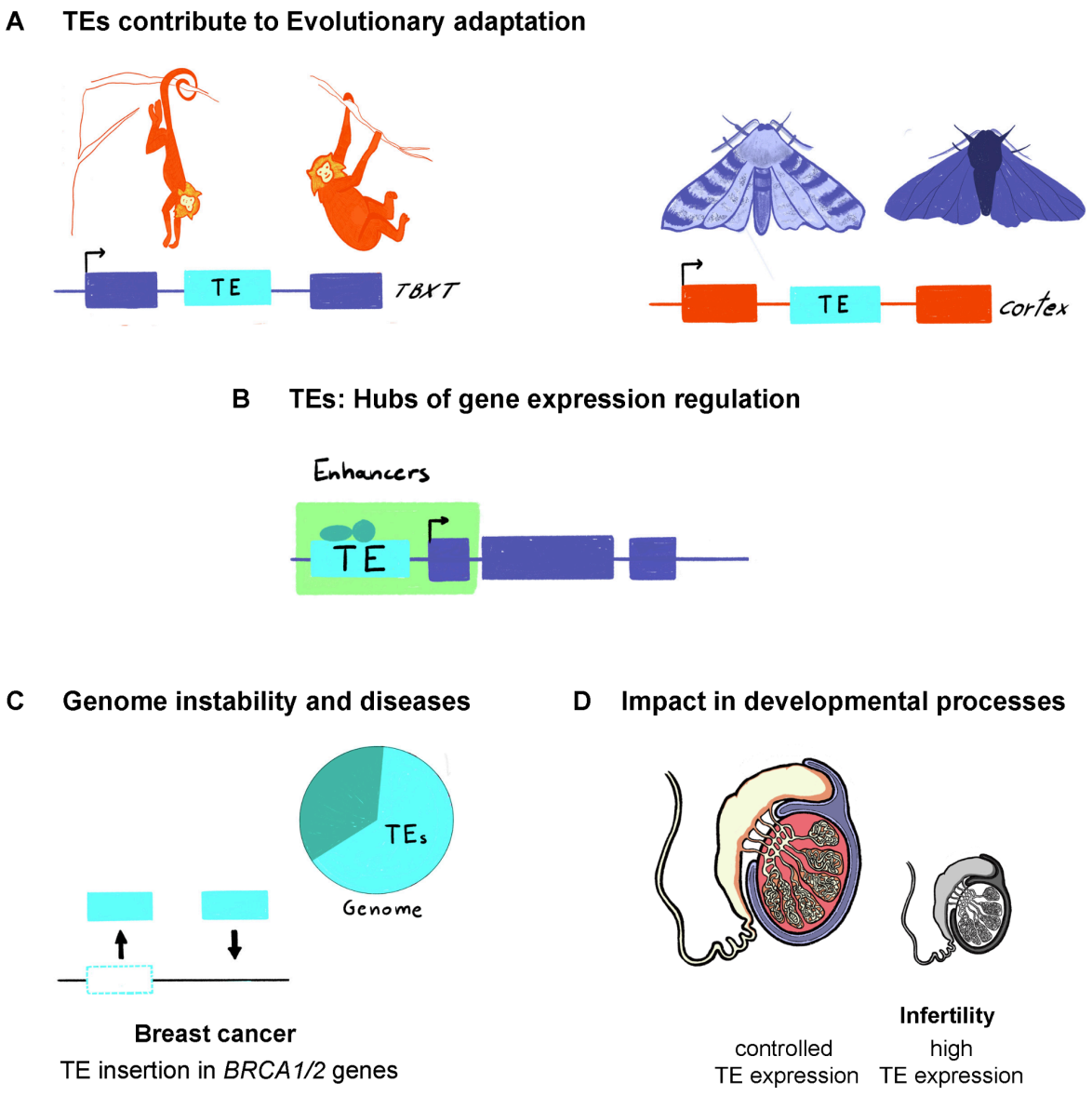


Figure 1.2 TEs: Two sides of the same coin. A. Two illustrations exemplifying TE contribution to evolutionary adaptation. An Alu insertion disrupted the *TBXT* gene expression resulting in tail-loss in hominids (left). A DNA transposon disruption into the *cortex* gene of the peppered moth led to a darker phenotype (right). **B.** Illustration of TEs as enhancers of neighbouring genes. **C.** TEs are a source of genome instability and diseases. Example of an Alu insertions into *BRCA1/2* genes, resulting in breast cancer. Image adapted from (Bourque et al. 2018; Ecco, Imbeault, and Trono 2017). **D.** Illustration of the impact of TEs in developmental processes. High TE expression can cause infertility in the mouse male germline (small grey testis). Illustrations were created by Dr. Dimitra Kanta.

1.2 Regulation of Transposable elements

The intricate relationship between TEs and their host genomes as well as their harmful effects when uncontrolled, has generated a selective pressure for the genome to evolve control mechanisms acting upon TEs at all levels of their life cycle. TE regulation operates in an interconnected regulatory network through DNA methylation, the piRNA pathway, histone modifications as well as sequence-specific protein-mediated regulators, which I will discuss in the following chapters. This balance between repressing harmful TE activity and permitting beneficial TE expression at specific developmental stages may reflect a dynamic interplay: host regulatory mechanisms adapt to constrain TE threats, while TEs themselves evolve strategies to evade silencing during permissive time windows. Such equilibrium especially during early embryonic and germline development, when regulatory controls are relaxed and TEs seize the opportunity to propagate to the next generation through the germline. Like TEs themselves, such regulatory mechanisms have often been co-opted for the physiological genome regulation. Thus, understanding how TEs are regulated is a powerful proxy to uncover novel ways of regulating the genome itself.

1.2.1 DNA methylation

DNA methylation is a chemical modification where a methyl group is covalently deposited to the 5' carbon of cytosine residues forming a 5-methylcytosine (5mC). In mammals, DNA methylation primarily occurs at CG, also known as CpG sites. CpGs-rich genomic regions are found in gene bodies, transposable elements and CpG islands (CGIs) (Ehrlich et al. 1982; Lister et al. 2009). CGIs are genomic regions consisting of more than 50% CG content and typically located near transcription sites, thus often having a regulatory function (Saxonov, Berg, and Brutlag 2006). While 70-80% of CpGs are methylated, non-CpG methylation exists in some cell types like embryonic stem cells (ESC) (Lister et al. 2009; En Li and Zhang 2014). DNA methylation plays crucial genome regulatory roles, predominantly associated with transcriptional repression (Ben-Hattar and Jiricny 1988). Promoters of transcriptionally active genes are typically unmethylated, whereas their bodies remain methylated (Lister et al. 2009) (Figure 1.3A). DNA methylation is essential for normal development and is associated with several key processes including genomic imprinting (E. Li, Beard, and Jaenisch 1993) X-chromosome inactivation (XCI) (Mohandas, Sparkes, and Shapiro 1981), and repression of TEs ((E. Li, Beard, and Jaenisch 1993; Mohandas, Sparkes, and Shapiro 1981; Walsh, Chaillet, and Bestor 1998).

1.2.1.1 DNA methyltransferases (DNMTs)

Mammalian DNA methylation is maintained and established by DNA methyltransferases (DNMTs): DNMT1 and DNMT3 enzymes, respectively. During DNA replication, DNMT1 is responsible for the maintenance of DNA methylation (T. Bestor et al. 1988), in coordination with UHRF1, by binding to the hemimethylated CpGs and copying from parent to daughter strands, ensuring epigenetic inheritance (Bostick et al. 2007). DNMT1 is essential in development and its depletion leads to embryonic lethality (E. Li, Bestor, and Jaenisch 1992). Historically, DNMT1 maintains existing patterns, but studies have challenged this idea by revealing *de novo* activity of DNMT1 targeting IAPs (Haggerty et al. 2021). While DNMT1 is mostly responsible for 5mC maintenance, DNMT3 enzymes establish novel 5mC sites, with each enzyme having specialized roles. DNMT3A and DNMT3B have similar structures. They both share a C' terminal catalytic methyltransferase (MTase) domain, an ATRX-DNMT3-DNMT3L (ADD) domain and a Pro-Trp-Trp-Trp-Pro PWWP domain (Okano et al. 1999) (Figure 1.3B).

DNMT3A and DNMT3B play a crucial role in development. DNMT3A has a pivotal role in germline methylation and has been reported to be necessary for spermatogonial stem cell (SSC) differentiation (Dura et al. 2022). In synergy with DNMT3L, DNMT3A establishes maternal genomic imprints by methylating bodies of oocyte-specific genes (D. Bourc'his et al. 2001). DNMT3L, a catalytically inactive cofactor of *de novo* DNA methyltransferases, has crucial roles in global germline methylation. *Dnmt3L*$^{KO/KO}$ mice have been reported to develop dysfunctional gametes and demonstrate TE upregulation resulting in male sterility (D. Bourc'his et al. 2001; Déborah Bourc'his and Bestor 2004; Zamudio et al. 2015). DNMT3B is important for early embryonic *de novo* methylation and has been reported to be crucial for placental formation (Andrews et al. 2023). The absence of either *Dnmt3A* or *Dnmt3B* leads to developmental dysregulation and embryonic lethality in mice (Okano et al. 1999).

The final member of the DNMT3 family, DNMT3C, has evolved from a gene duplication event of *Dnmt3B* and is muroid-specific (Barau et al. 2016; Jain et al. 2017). In contrast to DNMT3A/B, DNMT3C only carries a MTase and ADD domain but not a PWWP domain and it specifically methylates promoters of evolutionary young TEs in the male germline (Barau et al. 2016). Furthermore, *Dnmt3C*$^{KO/KO}$ mice mutants demonstrate high TE expression, resulting in meiotic failure and male mice sterility (Barau et al. 2016).

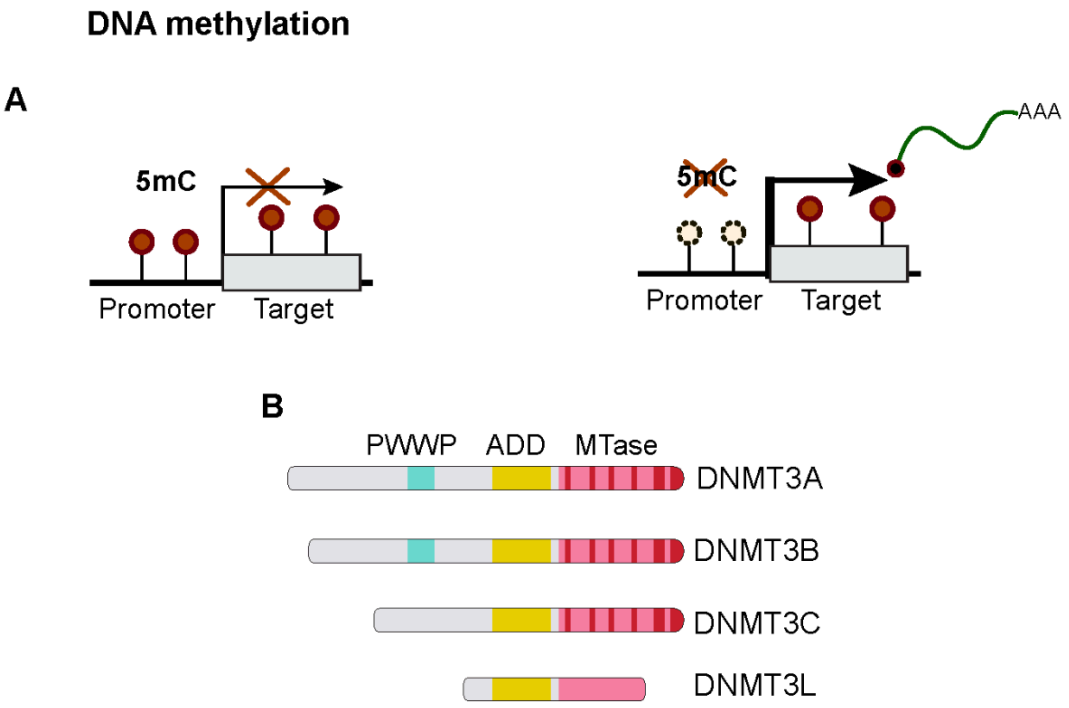


Figure 1.3 DNA methylation. A. Schematic depiction of DNA methylation repressing a gene target promoter. The presence of 5mC on gene promoters results in repression, whereas the absence leads to expression. **B.** Domains of *de novo* DNA methyltransferases (DNMTs): DNMT3A, DNMT3B, DNMT3C, DNMT3L; Figure adapted from (Barau et al. 2016).

1.2.1.2 DNA methylation Reprogramming

In mouse embryonic development, DNA methylation is very dynamic and undergoes reprogramming. Cells exhibit two waves of DNA methylation reprogramming in order to facilitate normal development. The first wave, also known as embryonic wave, happens after zygote formation, where DNA methylation is erased from normal levels to approximately 20% at the blastocyst stage at embryonic day (E)3.5, allowing developmental genes to be expressed (L. Wang et al. 2014) (Figure 1.4A). During this window, some TEs get expressed at the 2-cell (2C) stage, and are beneficial for normal development, as discussed in a previous chapter (1.1.3) while others retain DNA methylation levels. DNA methylation levels are restored by DNMT3A/B-mediated *de novo* methylation (Auclair et al. 2014). The second wave of DNA methylation reprogramming occurs in primordial germ cells (PGCs) starting at E9.5 and continues until E13.5, where DNA methylation levels drop to a lower level than the first wave, reaching as low as 5-7% (L. Wang et al. 2014) (Figure 1.4B). This marks the initiation of physiological germline gene expression, while some IAP promoters retain methylation (Seisenberger et al. 2012). Between E13.5 to E16.5 DNA methylation is restored

at the majority of the genome except of the evolutionarily young TEs, which are methylated by a second germline wave by DNMT3C/DNMT3L (Barau et al. 2016). Although DNA methylation reprogramming is imperative for development, hypomethylation could create an opportunity for TE upregulation, posing a germline genomic threat (Seisenberger et al. 2012). To overcome this vulnerability, the hosts have evolved genome defense mechanisms to repress TEs during this critical period.

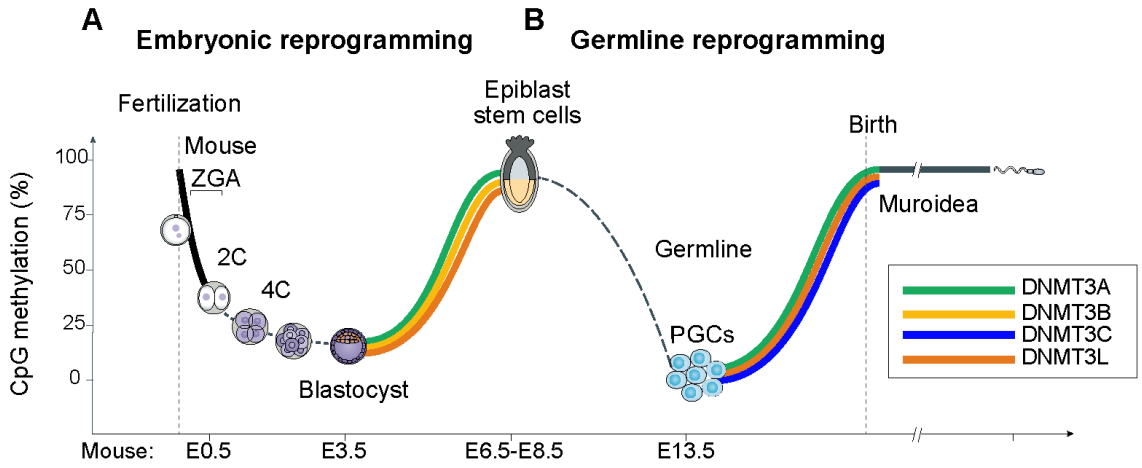


Figure 1.4. DNA methylation reprogramming. A. First wave: Embryonic reprogramming; 2-cell stage (2C), 4-cell stage (4C). After fertilization the zygote has normal DNA methylation levels, which drop gradually until the blastocyst stage. The methylation levels are re-established by DNMT3A and DNMT3B until the Epiblast stage (E.5-E8.5). **B.** Second wave: germline reprogramming; The DNA methylation levels get erased until the Primordial Germ Cells (PGCs). *De novo* methylation is restored to the male gametes by DNMT3A, DNMT3C and their cofactor DNMT3L; The image was extracted and adapted from (Greenberg and Bourc'his 2019).

1.2.2 piRNA pathway

One such specialized mechanism is the piRNA pathway. PIWI-interacting RNAs (piRNAs) are 24-35nt-long small RNAs that protect genome integrity by TE regulation in the germline (Vagin et al. 2006). These small RNAs are associated with PIWI proteins, a subgroup of Argonaute proteins. Mice express three PIWI proteins: MIWI, MILI and MIWI2 (Chuma and Nakano 2013). During male mice spermatogenesis, MIWI2 is expressed in early embryonic germ cells from E14.5 to postnatal day (P)3, MILI functions in both embryonic and postnatal meiotic cells (Figure 1.5A), and MIWI is present only in meiotic cells from P14 onwards (Chuma and Nakano 2013). Mutations of PIWI proteins lead to the

loss of DNA methylation from TE promoters and subsequently to TE upregulation and male sterility (Aravin et al. 2007; Zamudio et al. 2015; Barau et al. 2016; Vasiliauskaitė et al. 2018) (Figure 1.5A).

piRNAs protect the genome from TE activity through two distinct mechanisms. In the cytoplasm, TEs are repressed post-transcriptionally by the “ping-pong” cycle (Figure 1.5B). MILI cleaves TE transcripts guided by piRNAs, producing antisense secondary piRNAs, which target more TE transcripts for degradation creating an amplification loop (Aravin et al. 2008, 2007). The piRNAs are then loaded onto MIWI2, which translocates to the nucleus at E16.5 and binds to nascent TE transcripts recruiting the DNA methylation machinery and several proteins essential to achieve transcriptional silencing (Figure 1.5C) (Aravin et al. 2008; Zoch et al. 2024; Barau et al. 2016; Déborah Bourc’his and Bestor 2004; Zoch et al. 2020; Dias Mirandela et al. 2024). Among these proteins, DNMT3C is the most downstream enzyme guided by the nuclear piRNA-pathway to methylate evolutionary young TEs promoters (Barau et al. 2016; Zoch et al. 2024, 2020). Deficiency of the piRNA pathway components (such as *Dnmt3C* (Barau et al. 2016), *Dnmt3L* (Zamudio et al. 2015), *Miwi2* (Carmell et al. 2007) mutants), during the germline reprogramming leads to upregulation of TEs, deregulation of the germline transcriptome (Vasiliauskaitė et al. 2018), and subsequently infertility.

While the piRNA pathway guides DNA methylation to silence TEs, additional layers of transcriptional regulation, such as histone modification, are needed to ensure robust TE repression. Histone modifications can silence TEs both independently and by a DNA methylation interplay that I will discuss in the next chapters.

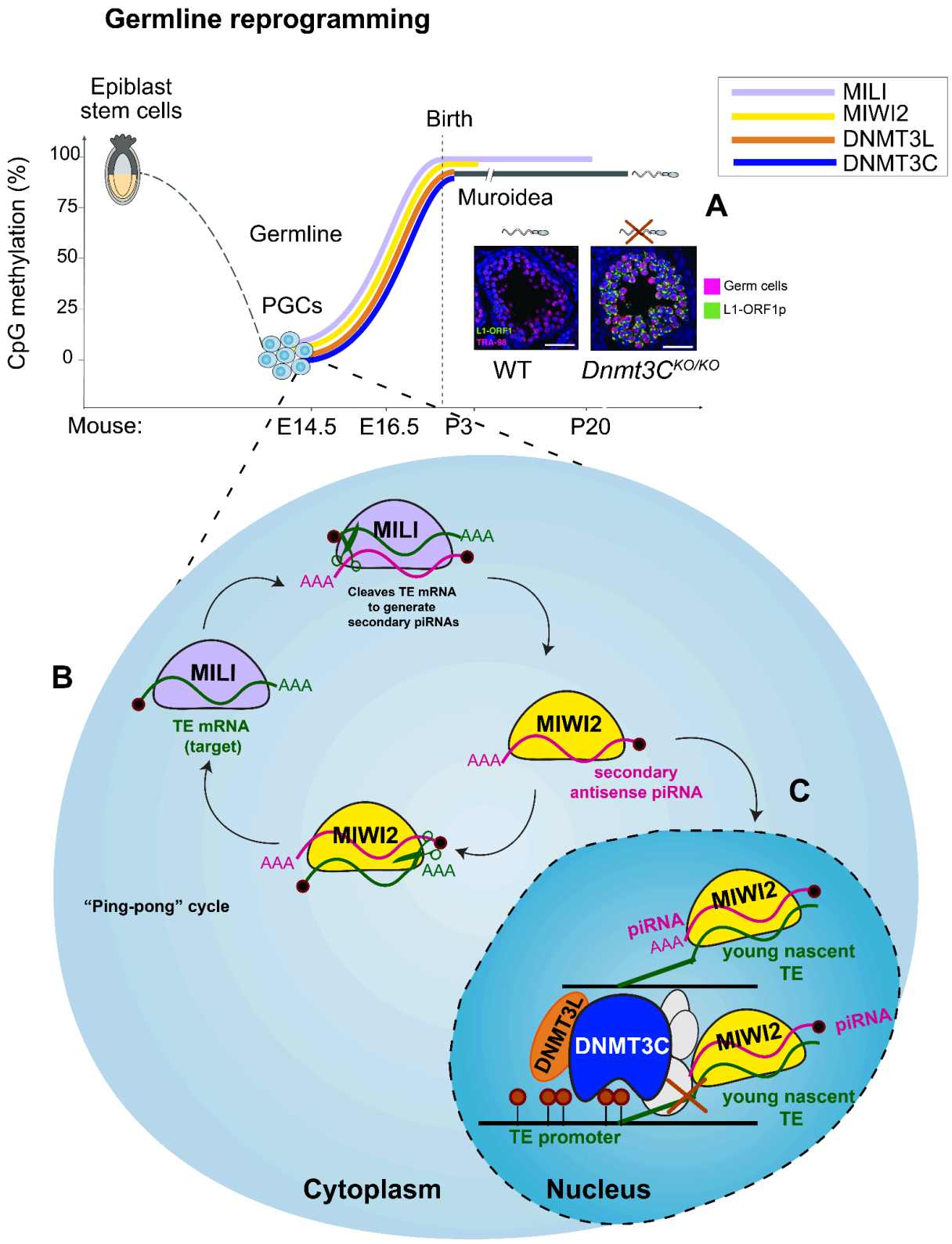


Figure 1.5. Transcriptional and post-transcriptional piRNA pathway. A. Example of a pi-RNA deficient mutant testis, *Dnmt3C*$^{KO/KO}$. The mutant germ cells undergo L1-ORF1 expression (green) and infertility. Image extracted and adapted from (Barau et al. 2016); MILI (lilac) is expressed in embryonic and postnatal meiotic stages; MIWI2 (yellow) is expressed from embryonic stages until P3; Image of germline reprogramming was adapted from (Greenberg and Bourc'his 2019). **B.** Illustration of the post-transcriptional piRNA pathway (ping-pong cycle) in the cytoplasm; **C.** Transcriptional piRNA

pathway; the secondary piRNAs are loaded onto MIWI2 (yellow) and target TEs in the nucleus. DNMT3C (blue), its cofactor DNMT3L (orange) and other enzymes (grey) are guided to repress TEs.

1.2.3 Histone modifications

Histone modifications play a fundamental role not only in chromatin structure regulation but also in gene transcriptional regulation (Bannister and Kouzarides 2011). These modifications can establish a distinct chromatin state: euchromatin, characterized by active marks and chromatin accessibility, and heterochromatin, which can have a constitutive and facultative form. Constitutive heterochromatin remains permanently condensed and is often found in repetitive sequences, like TEs (Grewal and Jia 2007), whereas facultative heterochromatin is transient and often found in developmental genes, allowing their dynamic regulation (Trojer and Reinberg 2007). While activating marks, such as tri-methylation at the lysine 4 residue of histone H3 (H3K4me3) facilitate an open chromatin configuration and gene activation, repressive marks like trimethylation at lysine 9 or 27 on histone H3 (H3K9me3 or H3K27me3, respectively) promote chromatin compaction and gene silencing. In the following sections, I will focus on two major repressing modifications, H3K27me3 and H3K9me3.

1.2.3.2 H3K27me3

H3K27me3 is a repressing histone modification that contributes to gene silencing through dynamic facultative heterochromatin formation. The Polycomb repressive Complex (PRC2) catalyzes H3K27me3 through the SET domain of its catalytic subunit Enhancer of Zeste Homolog 1 or 2 (EZH1 or EZH2) (Czermin et al. 2002). PRC2 has another two subunits that are important for the function of EZH1/2: Embryonic Ectoderm Development (EED) and Suppressor of Zeste 12 (SUZ12). Depletion of EZH2 and EED abolishes global H3K27me3 levels, whereas deletion of EZH1 does not affect global H3K27me3 levels, due to its lower activity (Raphael Margueron et al. 2008, 2009; Lee et al. 2018).

H3K27me3 is an abundant modification that marks 15% of H3 tails in mESCs (Peters et al. 2003). It is found in both small domains, like promoters of developmental genes that are not required in pluripotency (Boyer et al. 2006) and very large domains, such as Hox clusters or at the inactive X-chromosome of mammalian females (Pauler et al. 2009; Inoue et al. 2017; Boyer et al. 2006; Plath et al. 2003).

The relationship between H3K27me3 and DNA methylation is mostly antagonistic. In mESCs, H3K27me3 together with H3K9me3 are essential for repressing TEs in the absence

of DNA methylation (Walter et al. 2016). Similarly, in PGCs, TEs were detected to be enriched with H3K27me3, during decreased levels of DNA methylation (S. Liu et al. 2014). During development, non-canonical imprinting genes, which are initially silenced by H3K27me3, exhibit a switch to DNA methylation-mediated imprinting during their transition from preimplantation to postimplantation development (Z. Chen et al. 2019). H3K27me3 is typically found in unmethylated CGIs, as methylated CpGs sequences have been reported to counteract H3K27me3 (Jermann et al. 2014). However, a recent study reported that when cells exit from naive pluripotency to EpiLCs, H3K27me3 restriction can occur in a DNA methylation-independent manner, mediated by *Ezhip*, a PRC2 antagonist (Richard Albert et al. 2024).

PRC2 is required throughout early embryonic development from ZGA to gastrulation, and its disruption can cause embryonic lethality in mice (Raphaël Margueron and Reinberg 2011; Condemi et al. 2024). During development, H3K27me3 often coexists with the transcriptional activation mark H3K4me3, forming bivalent domains (Bernstein et al. 2006). This signature called “bivalency” is crucial in development by maintaining developmental genes in a poised state for either future activation or repression (Mohn et al. 2008). For example, bivalency has been identified in promoters of lineage-specific or germline genes in mESCs (Mohn et al. 2008) or PGCs, respectively, to facilitate activation later in differentiation (Sachs et al. 2013).

1.2.3.1 H3K9me3

While H3K27me3 associates with facultative heterochromatin and often antagonizes DNA methylation, H3K9me3 establishes transcriptional silencing through heterochromatin formation. This modification is often found at telomeres, centromeres and TEs (Richards and Elgin 2002; Karimi et al. 2011; Matsui et al. 2010). Histone methyltransferases (HMTs) are responsible for catalyzing H3K9me3 (Figure 1.6). One well characterized HMT is SET domain Bifurcated 1 (SETDB1), also known as ERG-associated protein with SET domain (ESET). SETDB1 transfers methyl groups via its SET domain to H3K9 using S-adenosyl methionine (SAM) as a donor. SETDB1 is often recruited to genomic targets through its interaction with KRAB-associated protein 1 (KAP1), also known as TRIM28 (Schultz et al. 2002). In mESCs, around 40% of SETDB1 binding sites were found in close proximity to ERVs, possibly explained by its interaction with KAP1 (Karimi et al. 2011). Additionally, depletion of SETDB1 leads to strong upregulation of ERVs and modest LINE-1 reactivation (Matsui et al. 2010; N. Liu et al. 2018). SETDB1 is essential for development since its absence causes lethality in pre-implantation and in mESCs (Dodge et al. 2004).

Suppressor of Variegation (3-9) homologue (SUV39H1) and its homolog SUV39H2 were the first HMT to be discovered (Figure 1.6). These enzymes were found to catalyze H3K9me3 via their SET domain *in vitro* (Aagaard et al. 1999; O'Carroll et al. 2000). However, *in vivo* SUV39H1/2 do not directly deposit H3K9me3 but rather primarily bind to existing H3K9me3 through their chromodomain and further spread H3K9me3 along the chromatin region (Lachner et al. 2001). SUV39H enzymes interact with HP1, which also binds H3K9me3, creating a positive feedback loop where HP1 binds to H3K9me3, recruits more SUV39H, leading to more heterochromatin spreading. Consequently, *Suv39h-double knockout (dKO)* leads to loss of HP1 in pericentric regions (Lachner et al. 2001), and results in LINE-1 upregulation and but not ERV reactivation in mESCs (Bulut-Karslioglu et al. 2014).

H3K9me3 has been found to typically co-occur with the presence of DNA methylation. In mouse E13.5 PGCs, during global hypomethylation, a subset of methylated IAPs was found to retain both DNA methylation and H3K9me3 enrichment (S. Liu et al. 2014). Depletion of *Setdb1* in PGCs, results in the loss of both H3K9me3 and DNA methylation from the LTR promoters of these IAPs (S. Liu et al. 2014). Similarly, when mESCs are cultured in hypomethylating condition, some IAPs and other genomic regions that maintain residual DNA methylation are additionally marked by H3K9me3 (Ficz et al. 2013; Habibi et al. 2013).

1.2.4 KRAB-ZFPs

The establishment of repressive histone modifications often requires targeting mechanisms. Krüppel-Associated box domain of zinc finger proteins (KRAB-ZFPs) bring another layer of TE control via sequence-specific binding to TEs serving as a molecular bridge to recruit both DNA methylation and histone modifications. KRAB-ZFPs recognize specific DNA sequences and bind close to the LTR of ERVs or the 5' UTR of LINEs. The silencing role of KRAB-ZFPs stems from the recruitment of KAP1, which in turn recruits HP1, SETDB1 and DNMTs creating a repressive chromatin environment established by the deposition of H3K9me3 and secondarily DNA methylation, to assure TE repression (Figure 1.6) (Ryan et al. 1999; Quenneville et al. 2012; Groner et al. 2010).

In early embryogenesis, during the first wave of reprogramming, some ERVs remain silent through a KRAB-ZFP-dependent mechanism. Specifically, depletion of the KRAB-ZFP, ZFP809, led to ERV reactivation and an epigenetic switch from repressive to active histone marks (Wolf et al. 2015). Similarly, KAP1 depletion in mESCs and early mice embryos, results in dramatic ERV upregulation and embryonic lethality highlighting its essential role in TE control during development (Rowe et al. 2010). KAP1 also regulates LINE1 elements in

an age-dependent manner. In both humans and mice, KAP1 primarily binds to LINE-1 elements of moderate age (~8-31 Myrs and ~4-6 Myrs for humans and mice, respectively), with these elements showing H3K9me3 enrichment (Castro-Diaz et al. 2014). While ancient L1 elements have degraded into genomic fossils, KRAB-ZFP/KAP1-mediating silencing mostly targets L1s of moderate age and the youngest elements often escape this control through mutations (Castro-Diaz et al. 2014).

Although KRAB-ZFPs were believed to have evolved to silence TEs, studies have also revealed the involvement of KRAB-ZFPs in the evolution of beneficial regulatory networks (Michaël Imbeault, Helleboid, and Trono 2017; Kosuge, Ito, and Hamada 2024). However, their contribution to TE repression has been well characterized with the majority of KRAB-ZFPs being associated with at least one TE retrotransposon subfamily (Michaël Imbeault, Helleboid, and Trono 2017), illustrating an ongoing evolutionary arms race with TEs. Multiple KRAB-ZFPs have continuously evolved to target the promoters of new L1 elements. While new L1 elements emerge, in order to escape, they accumulate mutations to lose the KRAB-ZFP binding sites on their sequences, and the hosts respond by adapting new KRAB-ZFPs that can recognize these mutations. In some cases young L1 elements (e.g. human L1PA1) have escaped by evolving a 129 bp deletion, excluding the binding site of a KRAB-ZFP, whereas in other cases the mutations are very subtle (Michaël Imbeault, Helleboid, and Trono 2017; Michael Imbeault and Trono 2014; Jacobs et al. 2014). Interestingly, Ecco et al. proposed a domestication model where KRAB-ZFPs facilitate TE integration into beneficial gene regulatory functions, instead of just repressing them (Ecco, Imbeault, and Trono 2017). This ongoing molecular Tom-and-Jerry game exemplifies how genomic conflict has led to both genome defense and regulatory innovation.

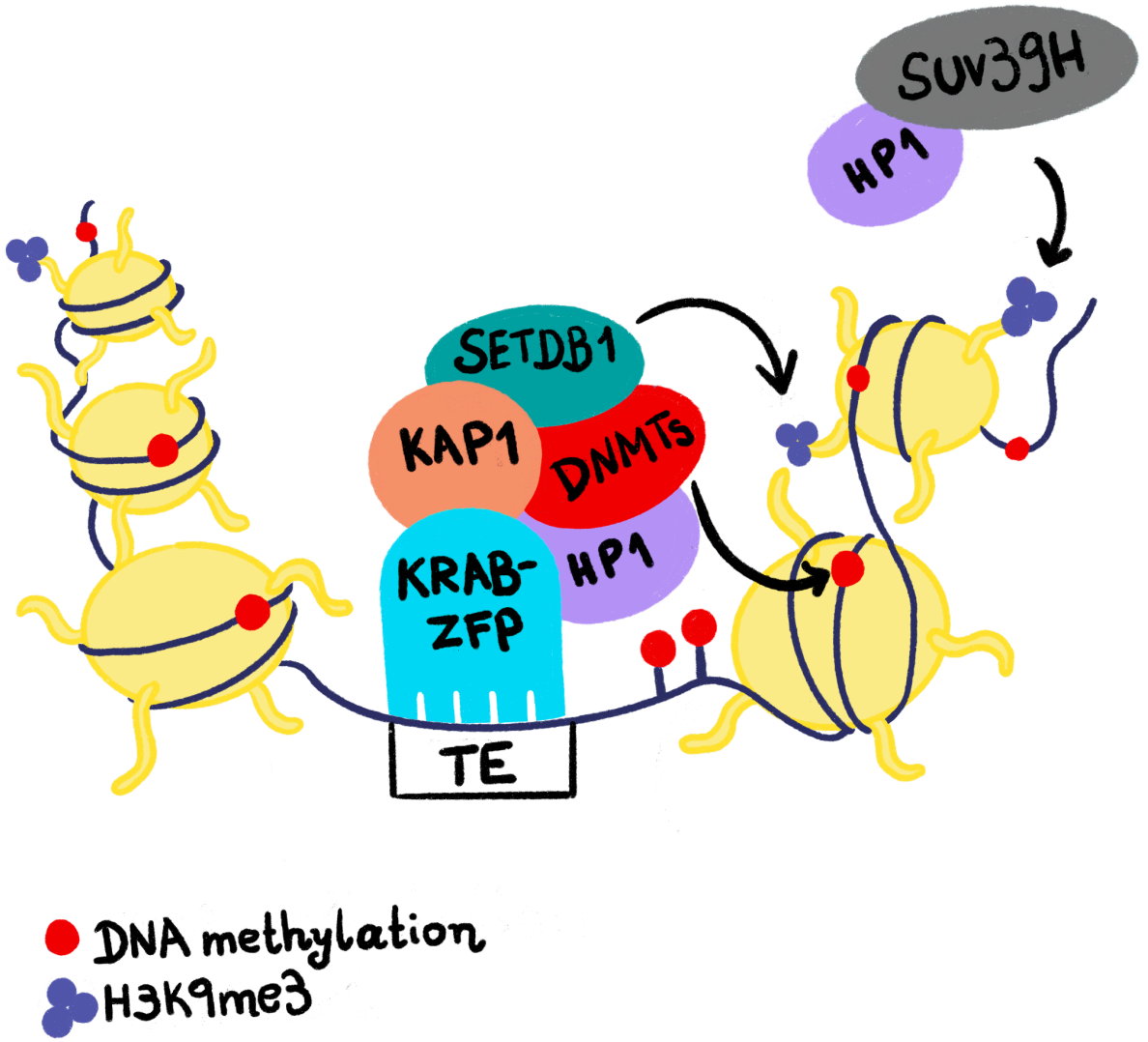


Figure 1.6 KRAB-ZFPs: targeted TE repressive complex. Illustration depicting KRAB-ZFPs (blue) binding to sequence-specific TE targets and recruit repressors to silence TEs; KAP1 (orange), SETDB1 (dark green) catalyses H3K9me3 (purple); HP1 (lilac); DNMTs (red) catalyze 5mC (red); SUV39H enzymes (grey) catalyze H3K9me3; histones (yellow); DNA string wrapped around histones (purple). Illustration was created by Dr. Dimitra Kanta.

1.2.5 Transcription Factors (TFs)

Besides the epigenetic landscape that shapes genome regulation and TE repression, transcription factors (TFs) have a determining role in the cellular identity. TFs are DNA-binding regulators that recognize and bind to specific short sequences in regulatory regions and initiate expression or repression of cell-type specific genes (Jacob and Monod 1961; Wunderlich and Mirny 2009; Maston, Evans, and Green 2006). Although TF binding motifs are abundant in the genome, TFs only bind to a limited number of their potential binding sites in a cell-type specific manner (Biggin 2011).

TF binding to DNA is largely dependent on chromatin accessibility. The nucleosome positioning, post-translational histone modifications, and DNA methylation could prohibit or promote TF binding. Nucleosome occupancy is associated with low TF-binding affinity (Yuan et al. 2005; Dai et al. 2009), and TF-nucleosome binding is affected by several factors including positional or orientation preferences (Zhu et al. 2018). Pioneer TFs represent an exception since they are capable of binding to nucleosomes and initiating chromatin remodeling (Cirillo and Zaret 1999; Iwafuchi-Doi and Zaret 2014). The pluripotency factors (OCT4, SOX2, KLF4, and c-MYC) are typical pioneer factors that bind to nucleosomes and initiate chromatin remodeling by inducing H3K4 methylation and subsequently gene activation (Soufi, Donahue, and Zaret 2012; Koche et al. 2011). However, even these factors can be restricted by certain chromatin states. In early development, their binding has been reported to be blocked from H3K9me3 regions and only bind upon removal of this repressive mark (Soufi, Donahue, and Zaret 2012). TFs often act cooperatively to access their binding can be restricted from other layers.

DNA methylation also plays a crucial role in TF-binding (Moore, Le, and Fan 2013). The presence of DNA methylation is often associated with TF binding inhibition (Héberlé and Bardet 2019). The binding could be affected either directly by the the presence of 5mC altering the binding site, so the TF can't recognize the motif anymore (Dantas Machado et al. 2015), or indirectly by methyl-CpG-binding domain (MBDs) proteins that recruit deacetylases to form chromatin compaction and block TF binding sites (Nan et al. 1996, 1998). *In vitro* large-scale assays have revealed TFs that bind to both unmethylated and methylated DNA templates (Mann et al. 2013). For example, CTCF can bind methylated sequences both *in vitro* and *in vivo.* However, *in vivo* CTCF binding is context-dependent: CTCF binds to CpG-poor regions regardless of the methylation status, while its binding is inhibited at methylated CpG-rich regions (Holwerda and de Laat 2013; Stadler et al. 2011). Similarly, REST can bind methylated DNA and stimulate demethylation (Feldmann et al. 2013; Stadler et al. 2011). *In vivo*, TF binding is a more complex process due to the diverse biological environment. TFs often act cooperatively to access their motifs and their binding could be restricted by several layers including chromatin structure, co-factors, or other regulatory proteins.

TFs play an important role in TE regulation, though no TE-specific TF has been identified as a master regulator. Instead, different TFs regulate TEs in a cell-type specific manner, with TEs acting as cell-specific enhancers (Karttunen et al. 2023; Hermant and Torres-Padilla 2021). For example, p53 was found to prevent LINE-1 and SINE expression in human cancer cells (Tiwari et al. 2020; Lopez et al. 2023). LINE-1 elements include TF binding site (TFBS) of YY1 in their 5' UTR promoters, crucial for L1 transcription in humans and mice (Athanikar, Badge, and Moran 2004; Saha et al. 2024). In differentiated neurons,

CREB1 binds and activates methylated IAPs (Kaluscha et al. 2022). In humans, pluripotency transcription factors such as OCT4 and KLF4 facilitate young ERV expression under hypomethylation conditions (Grow et al. 2015), potentially driving transcription preimplantation (Pontis et al. 2019). Similarly, in humans and mice DUX, a TF important for zygotic genome transcription (Sugie et al. 2020) activates MERVLs, which are essential for ZGA and embryo development in mice preimplantation (Sakashita et al. 2023). In mouse male germ cells, A-MYB, also known as MYBL1, is an important meiotic regulator (Bolcun-Filas et al. 2011), which binds to an enhancer LTR element, RLTR10, to activate germline genes (Sakashita et al. 2020). However, TF-dependent regulation of TEs in germ cells and their contribution to TE activity or repression remains largely unexplored. Understanding the interplay between TFs and TEs requires appropriate model systems, which we will discuss in the following chapter.

1.3. Model systems to study TE regulation

1.3.1 Mouse male germline

Germ cells represent a unique system to study TEs, as they are the central point, where both life itself and TEs can be passed onto future generations. As previously discussed (1.2.1.2), germ cells undergo genome-wide DNA methylation reprogramming creating a suitable substrate for TE activation and propagation. The piRNA pathway initiates TE silencing by MILI expression at E13.5, followed by nuclear MIWI2 expression at E16.5 (Molaro et al. 2014) (2.2). Subsequently, the piRNA pathway recruits DNMT3C, the expression of which peaks at E16.5-E18.5, to methylate young TE promoters (Barau et al. 2016; Zoch et al. 2024). Of note, TEs get exclusively upregulated in male DNMT3C-deficient mice, while females maintain TE silencing independent of DNMT3C (Barau et al. 2016). Importantly, in male germ cells, *Dnmt3C*$^{KO/KO}$ specifically depletes DNA methylation at young TE promoters while leaving global CpG methylation levels stable (Barau et al. 2016; Dura et al. 2022). This unique feature allows us to study TE regulation without the confounding side effects of deregulated gene expression caused by global DNA methylation loss (Barau et al. 2016; Dura et al. 2022). These characteristics make male DNMT3C mutant germ cells one of the best mammalian systems to study regulation of unmethylated young TE promoters.

1.3.1.1 Spermatogenesis

In mice, PGCs are the precursors of germ cells. PGCs start forming around E6.5 at the epiblast and migrate at E10.5 to the genital ridges, where they rapidly proliferate from a

few tens to several thousands of cells until E13.5, when they transition into prospermatogonia (Pro-Spg) in males (Saitou and Yamaji 2012; Tam and Snow 1981). Pro-Spg are characterized by large spherical nuclei with apparent nucleoli. During this germline developmental stage, specifically at E16.5, the piRNA pathway is important as its failure results in TE upregulation and subsequently infertility (Schöpp et al. 2020). At this stage, they undergo mitotic arrest in the G0/G1 phase and stop proliferating until birth (Vergouwen et al. 1991).

Shortly after birth around P3, the mitotic phase begins, where prospermatogonia cells start dividing again and differentiate into spermatogonial stem cells (SSCs) to initiate spermatozoa production, a process known as spermatogenesis. Although SSCs comprise a rare population of total germ cells in the testis (Fayomi and Orwig 2018), they are fundamental for spermatogenesis, since they possess the ability to self-renew and give rise to differentiated cells (de Rooij and Russell 2000). SSCs are located at the basal lamina of the seminiferous epithelium, and the differentiation of germ cells progresses towards the lumen of the seminiferous tubule, and continues along the tubule (Figure 1.7A). A subset of somatic cells, sertoli cells, induce self-renewal and differentiation of SSCs, into their progeny, spermatogonia cells, which undergo further proliferation and differentiation. Interestingly, during the spermatogonia germline developmental stage, undifferentiated spermatogonia show IAP upregulation in *Miwi2* and *Dnmt3L* mutants (Vasiliauskaitė et al. 2018), while in differentiated spermatogonia (Spg) LINE-1 elements remain repressed in a DNA-methylation independent manner (Zamudio et al. 2015). Differentiated spermatogonia cells undergo several mitotic divisions before transitioning into spermatocytes (Spc) at around P10, when the meiotic phase starts (Bellvé et al. 1977).

Spermatocytes are meiotic cells dedicated to producing haploid cells through meiosis I, characterized by homologous chromosome segregation followed by meiosis II, characterized by sister chromatids segregation, and subsequently leading to haploid daughter cells (Handel and Schimenti 2010). The first step of meiosis I is prophase I, which is very crucial as most piRNA pathway-deficient mutants exhibit meiotic arrest resulting in spermatogenesis failure and infertility (Y. Li, Zhang, and Liu 2021). Prophase I consists of four different stages (Figure 1.7). Leptotene is the first stage, which occurs after DNA replication and chromosomes become more compacted. The second stage is zygotene, where chromosomes pair with their homologous paternal and maternal chromosomes. Pachytene is the third stage that occurs around P14 to P18 (Bellvé et al. 1977), characterized by complete homologous recombination of the paired chromosomes, a process known as synapsis. Pachytene is an important stage in which germ cells of a lot of piRNA-deficient mice mutants, including *Dnmt3C*$^{KO/KO}$, arrest and cause TE upregulation (Barau et al. 2016; Zamudio et al. 2015; F. Yang et al. 2008). Diplotene is the final stage of

prophase I, in which the chromatids of the homologous chromosome start segregating, generating haploid round spermatids.

The last stage of spermatogenesis is spermiogenesis, during which the haploid cells experience dramatic morphological and cytological changes without further cell division (Jan et al. 2012). These changes include an acrosome development in the sperm head, formation of a midpiece containing mitochondria, and development of a flagellum (sperm tail). The cells also undergo cytoplasmic reduction and chromatin remodelling, where histones are replaced from protamines. Spermatozoa maturation is finalized after the cells are released into the lumen of the seminiferous tubule and further travel in the epididymis where they gain motility (Oakberg 1956).

In adults, spermatogenesis is a continuous and asynchronous process of germ cell differentiation, essential for male reproduction. Of note, this means that a testis is composed of a mixed population of germ cells depending on the developmental stage (Figure 1.7B) (Hess and Renato de Franca 2008). For example, fetal mouse testes consist only of prospermatogonia cells, while at P5 both SSCs and spermatogonia are present, and at P15 SSC, spermatogonia and spermatocytes can be found. At post-meiosis, as the germline development progresses, cells differentiate into round and elongated spermatids, eventually forming spermatozoa (Figure 1.7A).

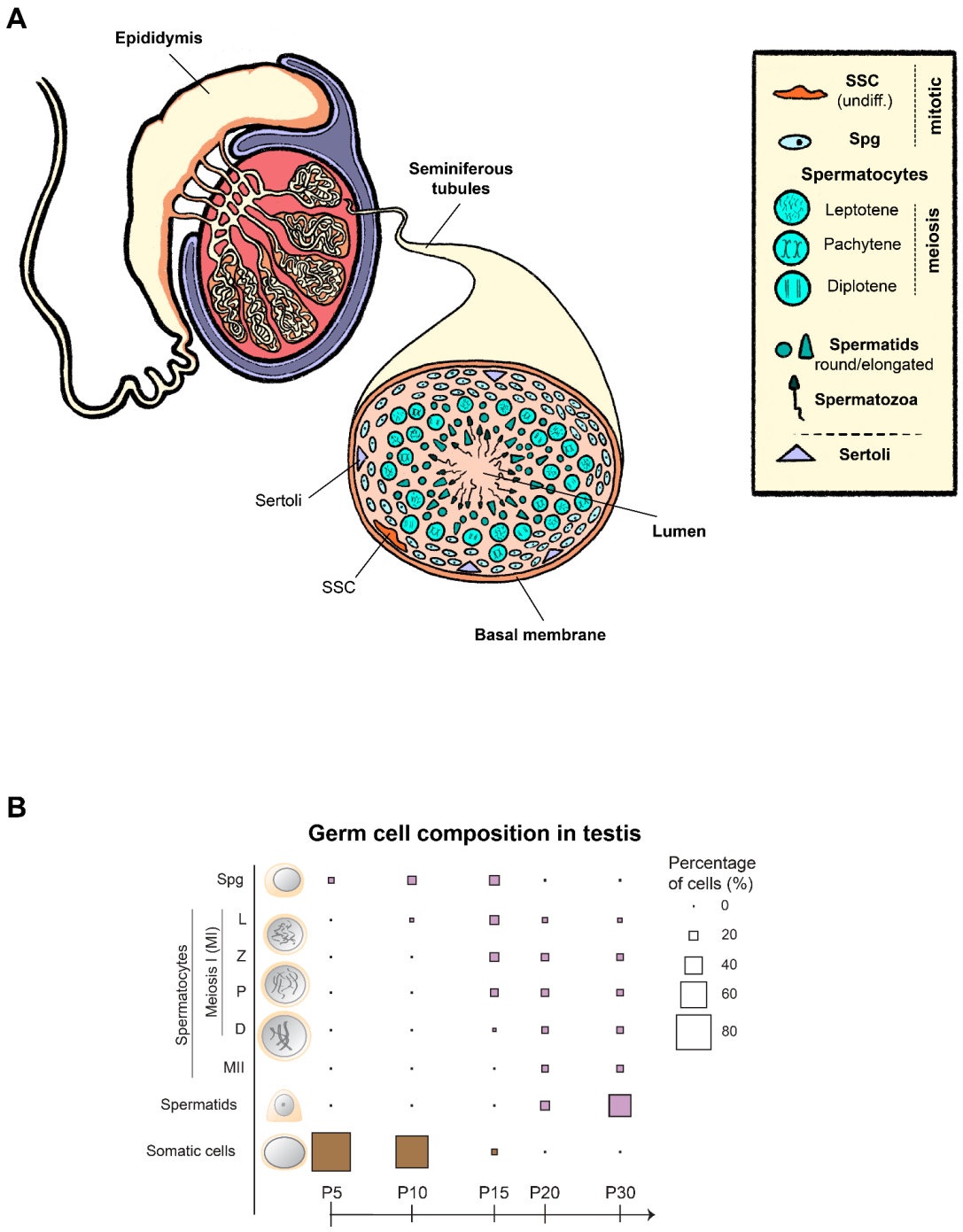


Figure 1.7 Male mice spermatogenesis. A. Illustration of male mouse testis anatomy. Spermatogenesis occurs from the basal lamina of the seminiferous epithelium towards the lumen. Three main phases of spermatogenesis: mitotic phase, meiotic phase and spermiogenesis; mitotic: undifferentiated SSC, Spermatogonia (Spg); meiotic: Spermatocytes (Spc), leptotene, pachytene, diplotene; Spermiogenesis: round or elongated spermatids, mature spermatozoa. Illustration was created by Dr. Dimitra Kanta. **B.** Schematic representation showing percentage of cells from mice testes with different age classified into developmental stages of Spg, Spc, spermatids and somatic

cells. Meiosis I with Leptotene (L), Zygotene (Z), Pachytene (P) and Diplotene stages as well as Meiosis II are depicted in Spc. This image is modified from (Ernst et al. 2019).

1.3.2 Mouse embryonic stem cells (mESCs)

Mouse embryonic stem cells (mESCs) are among the few mammalian cells that remain viable in the absence of DNA methylation (Tsumura et al. 2006). Although DNMT-depleted mESCs exhibit TE silencing by compensatory defense mechanisms, such as H3K9me3 and H3K27me3 (Walter et al. 2016), different TEs showed varied chromatin silencing response patterns. Specifically, LINE-1 elements switch from H3K9me3 to H3K27me3, MERVLs showed enrichment in H3K27me3, while IAPs remained enriched with H3K9me3 (Walter et al. 2016), indicating that different TE sequences depend on distinct silencing control mechanisms in the same cell type. The exact reasons behind these differential responses remain largely elusive. mESCs provide an amenable system to study these questions due to faster genetic manipulation, an aspect that is more challenging and time-consuming to apply in mice. Previous studies have successfully depleted key factors of TE genome defense mechanism including components of KRAB/ZFPs-mediated (Rowe et al. 2010) or H3K9me3-mediated TE silencing pathways (Bulut-Karslioglu et al. 2014). The ease of genetic manipulation in mESCs, compared to the complexity of mouse genetic crosses, makes them an ideal experimental model to explore TE regulation. One could consider mESCs the “yeast of mammalian cells”, as they even enable genome-wide forward genetics screens to discover novel TE regulatory pathways, which we will discuss in the following chapter.

1.4 Screening for TE regulators

Although several TE repressors have been discovered so far, many transposon silencing mechanisms remain elusive. Over the past years, several attempts have been made to identify new TE regulators by genome-wide loss-of-function screens. These screens have utilized gene-editing tools such as small interfering RNA (siRNA) and CRISPR/Cas9, which enable targeted gene silencing or disruption. siRNA screens operate by degrading mRNA transcripts to reduce gene expression, while CRISPR/Cas9 directly edits the genome by introducing double-strand breaks at specific loci. Both tools have been instrumental in uncovering key transposon regulators.

For instance, siRNA loss-of-function screens were initially performed to identify transposon silencing factors in mouse Embryonic Carcinoma cells (ECCs) (B. X. Yang et al.

2015). In this screen, they used a Moloney murine leukemia virus (MMLV) sequence driving GFP expression. GFP fluorescence served as a screen readout and led to the identification of new regulators such as small ubiquitin related modifier 2 (SUMO2) involved in the deposition of SUMO on KAP1 and subsequent recruitment of KZFP binding on DNA. However, knockdown screens only transiently degrade transcripts without complete gene abolishment.

CRISPR/Cas9-based knockout forward genetic screens have emerged as a powerful alternative approach to discover novel TE regulators. Fukuda et al. performed a genome-wide CRISPR-based knockout screen in mESCs (Fukuda et al. 2018). As a screen readout, they used a reporter retrovirus expressing GFP under the control of a MMLV LTR promoter, assuming it behaves as a synthetic, *bona fide* TE. They sorted cells based on high GFP expression to identify putative regulators through sgRNA sequences. In CRISPR/Cas9-based screens, sgRNA sequences serve as unique identifiers for the targeted genes, allowing researchers to map GFP expression changes back to specific gene knockouts. The analysis of this screen specifically revealed a catalogue of genes and the top cellular factors were picked as candidate transposon repressors. Despite being successful, their approach was somewhat limited as it led to the identification of components, such as MORC2a, which are part of previously known repressive pathways (Tchasovnikarova et al. 2017). Chelmicky et al. conducted a genome-scale CRISPR/Cas9 knockout screen in mESCs to study IAP regulation (Chelmicki et al. 2021). They used an IAPez reporter driving inducible blasticidin-resistance-GFP cassette, enabling both antibiotic selection and fluorescence-based sorting at different timepoints with enhanced sensitivity compared to GFP-only selection. Interestingly, among known top candidates they identified N6-methyladenosine (m6A) regulators and confirmed m6A RNA methylation as a potentially novel ERV repressor. Groh et al. performed a CRISPR/Cas9 screen in mESCs using a small heterochromatin inducing sequence (SHIN) as a reporter (Groh et al. 2021). SHIN is a small sequence within the *gag* region common to many IAPez elements, that was previously shown to induce heterochromatin formation and recruit repressing factors (Sadic et al. 2015). Both these two screens, although using very similar readout approaches, found different candidates, indicating that there might be different silencing pathways for different IAPez regions.

CRISPR/Cas9 screens have also been performed to discover LINE-1 regulators. Liu et al. developed an innovative neomycin-based retrotransposition cassette to detect both repressors and activators only upon retrotransposition events. This readout enabled them to screen for post-transcriptional regulators in K562 cells (N. Liu et al. 2018). At a recent CRISPR/Cas9 screening for LINE-1 elements, they used a LINE-1 5’ UTR as a reporter and

they identified both regulators and activators (Xiufeng Li et al. 2024). While this study largely confirmed known L1 regulators, it revealed a novel function of L1s as long-range enhancers.

Although all these screenings have provided valuable insights, they share an ostensible limitation: the assumption that the artificial reporter constructs fully reflect endogenous TE behavior. The truncated TE promoters, absent of the whole TE coding sequences, and the artificial reporters have primarily led to the identification of transcriptional and in some cases post-transcriptional regulators. These approaches may have overlooked *bona fide* transposon silencing factors or even entirely new pathways. Furthermore, while transcriptional and post-transcriptional TE regulation has been studied, the regulation and fate of TE-encoded proteins remains largely unexplored.

In mouse models, high-throughput and CRISPR/Cas9 screens for TE regulators remain technically very challenging. Instead, insights into TE expression regulation in the male germline have primarily come from more targeted approaches. These approaches, such as immunoprecipitation of whole testes (De Luca, Gupta, and Bortvin 2023) or RNA-seq analysis of sorted germ cells (Sakashita et al. 2020), could be viewed as strategies to “screen” for TE regulation, but with a narrower scope than high-throughput methods. Technical limitations and challenges in both mammalian cells and *in vivo* approaches highlight the need for complementary and alternative strategies to understand TE regulation in physiological contexts.

Hypothesis and Aim of the Thesis

TE activity across development is most pronounced in time windows where retrotransposition can generate heritable copies, such as in preimplantation embryos and the germline. While TE activity can be highly deleterious, silencing pathways have evolved to repress their activity and safeguard genome integrity throughout development. In preimplantation stages, when DNA methylation is erased, KAP1-mediated TE silencing plays a crucial role in maintaining genome stability (Coluccio et al. 2018). In the mouse male germline, DNMT3C silences TEs through *de novo* DNA methylation (Barau et al. 2016). However, TE activity can also be an integral and important part of certain developmental stages.

An unexplored question in the field is how cells restrict retrotransposition while still allowing beneficial TE expression. We hypothesize that post-transcriptional and post-translational restrictive pathways may be critical for safeguarding genome stability during beneficial TE expression. Additionally, we propose that transcriptional activators drive TE expression under harmful conditions when key repressors, such as DNA methylation are absent. In my two PhD projects, I envisioned addressing these hypotheses by discovering novel TE regulators in mESCs and in the male germline, respectively.

In my first project, I aimed to establish a whole-genome wide CRISPR/Cas9-based forward genetic screen for repressors of endogenous TEs to gain a deeper understanding of transcriptional and translational repression mechanisms in fairly unexplored parts of the TE life-cycle. Unlike previous forward genetic screening approaches for TE silencing, I aimed to optimize a pioneering TE screen using direct readouts of endogenous TE activity rather than relying on reporter systems by:

1. Detecting endogenous TE mRNA abundance by fluorescent in situ hybridization of TE mRNA (RNA FISH), and
2. Monitoring the translation products of endogenous TE mRNA by immunofluorescence (IF).

Fluorescence-activated cell sorting (FACS) would allow phenotypical sorting of cells highly expressing TEs and/or accumulating their translated products, followed by sequencing of the gRNAs to discover putative genes involved (Figure 2.1).

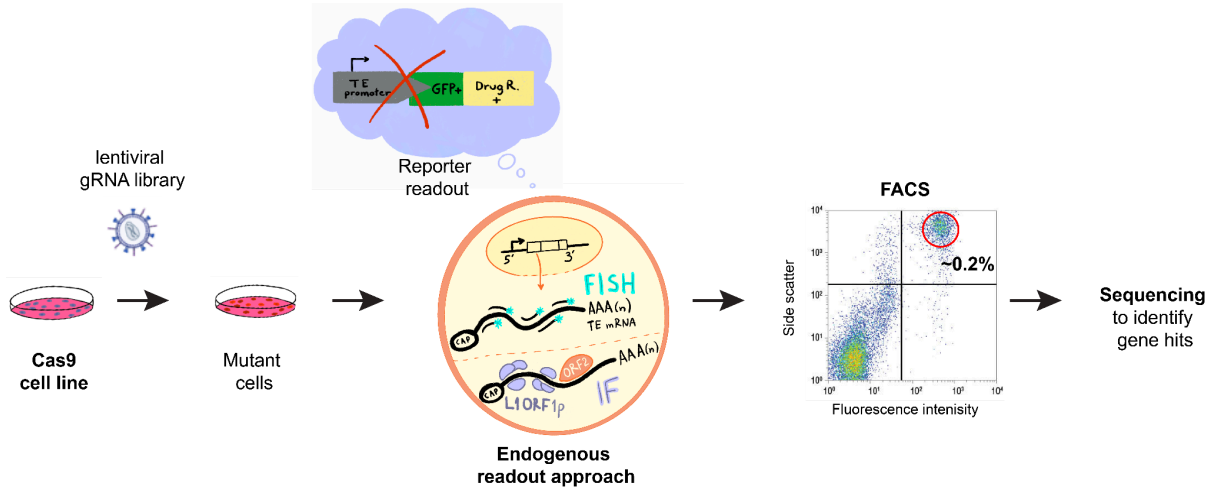


Figure 2.1 Scheme of CRISPR/Cas9-based forward genetic screen of novel transposon regulators by lentiviral gRNA library. mESCs expressing an inducible- degron Cas9 system, are infected with a lentiviral gRNA library. Further, TE-encoded proteins (such as L1-ORF1p) are labeled by IF and TE RNA by RNA FISH. The positive cells are selected by FACS and gRNA sequences are PCR-amplified and sequenced. Illustrations were generated by Dr. Dimitra Kanta; The scatter plot was adapted from Cell signaling.

In my second PhD project, which was a collaboration with my colleague, Jessica Leismann, we investigated TE reactivation and TE regulators in male germ cells. Utilizing *Dnmt3C*$^{KO/KO}$ mice we addressed two main questions:

1. What are the dynamics of TE expression when DNMT3C-mediated DNA methylation is lost throughout mouse spermatogenesis?
2. Is NRF1 a potential trans-activator of unmethylated TEs in the mouse male germline?

This study provides comprehensive and pioneer *in vivo* insights into TE regulation across spermatogenesis.

2. Results

2.1 An innovative CRISPR/Cas9 screen readout based on endogenous TE activity

Chapter 2.1 represents my primary PhD project, which I pursued for nearly 3 years. Despite significant challenges and time constraints ultimately leading to the suspension of this project (2.1.3), identifying endogenous TE regulators remains an intriguing and relevant topic in the field. Cell lines generated in this work facilitated collaborations with the Bourc'his lab, the Imbeault lab and the Dikic lab. All experiments and figures presented in the following subchapters were conceptualized, performed and generated by me.

2.1.1 Generation of an inducible degron-Cas9 cell line

To enable tightly controlled genome editing for discovering TE repressors, we sought to develop a system for conditional Cas9 expression in mESCs. While constitutive Cas9 activity is commonly used as a baseline in CRISPR/Cas9 screenings, it can restrict genome editing when specific temporal control is required. Therefore, we aimed for a system with precise control over the timing of Cas9 activity, to minimize potential off-target effects and to allow for dynamic investigation of TE regulatory mechanisms. We achieved this by combining doxycycline (Dox)-inducible transcription with dTAG-mediated protein degradation (Figure 2.1.1A).

To generate the cell line, we utilized a previously generated mESC line stably expressing the Tet-On 3G activator under the EF1α promoter in the *Rosa26* locus (data not shown), a safe harbor locus known for stable and ubiquitous transgene expression (Xiaoping Li et al. 2014). Although Tet-On 3G is constitutively expressed from the EF1α promoter, it could only induce activation in the presence of Dox (Das, Tenenbaum, and Berkhout 2016). This system was chosen for its precise and inducible control of gene expression. Specifically, upon Dox treatment, Tet-On 3G binds to pTRE3G and subsequently drives downstream Cas9 transcription (Figure 2.1.1B). Dox-inducible systems can exhibit leaky expression, often due to intrinsic properties of the system such as weak binding of the reverse tetracycline transactivator (rtTA) to the promoter even in the absence of Dox (Das et al. 2016; Costello et al. 2019). To prevent any potential leakiness, we further engineered the

system to incorporate a dTAG-degron system (Nabet et al. 2018) enabling a double-locked control of Cas9 expression. In contrast to other degron systems, the dTAG-13 system does not require the expression of an additional proteasome component. This system uses a mutated F-box protein, FKBP-V, as an in-frame tag of the protein of interest. Upon application of a heterobifunctional synthetic molecule, FKBP-V engages to an E3 ubiquitin ligase (CRBN) inducing degradation of the tagged protein (Nabet et al. 2018) (Figure 2.1.1C). The pTRE3G-FKBPV-3xFlAG-Cas9 cassette was inserted by homologous recombination in the *Tigre* locus, another safe harbor site optimized for tightly regulated and inducible transgene expression without interference from endogenous genes (Ahmadzadeh et al. 2020; Zeng et al. 2008). We generated the cell line by co-transfecting the donor plasmid carrying the cassette and the editor plasmid carrying SpCas9 and sgRNA targeting the *Tigre* locus.

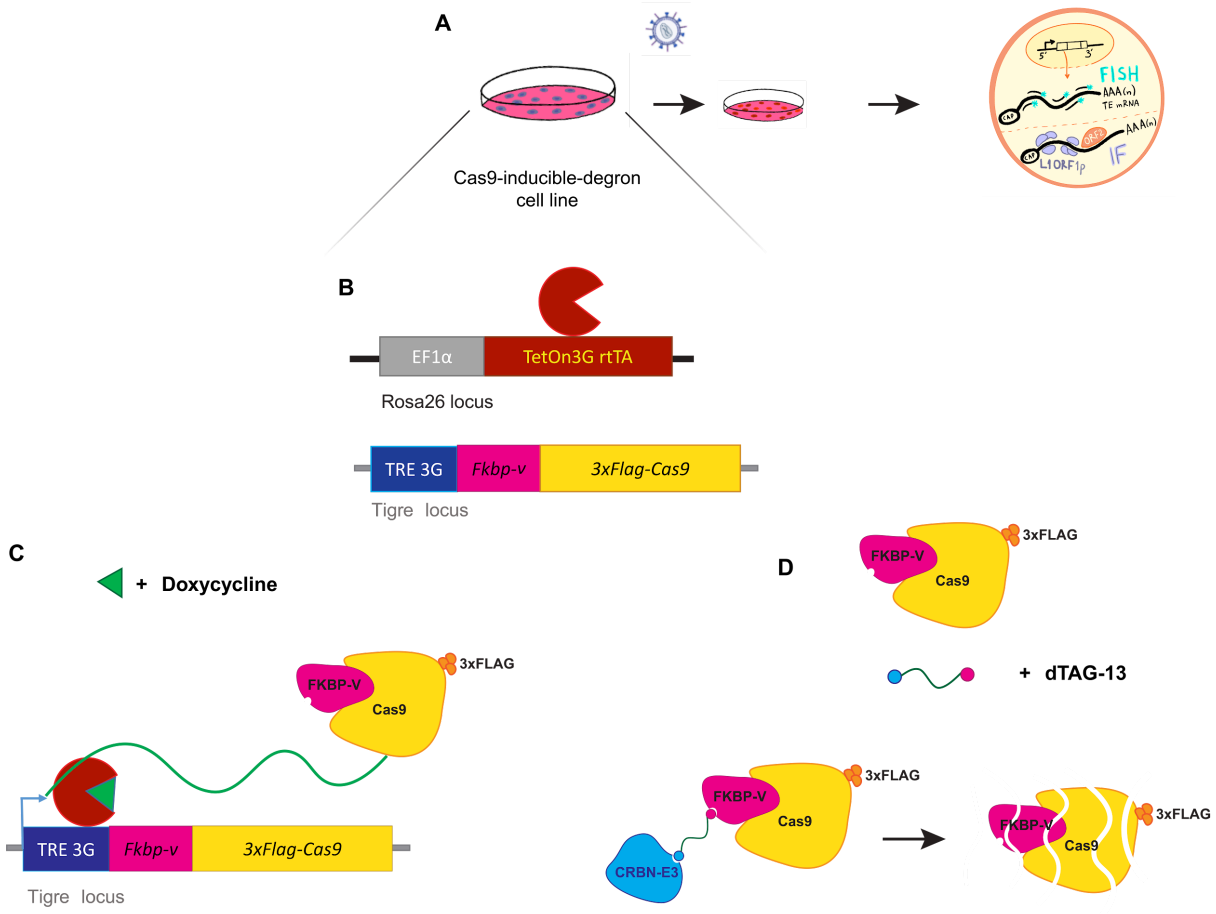


Figure 2.1.1 Schematic depiction of dox-inducible dTAG-mediated Cas9 cell line strategy. A. Scheme of CRISPR/Cas9 screen upstream of the readout utilizing Cas9-inducible-degron cell line. **B.** Schematic depiction of the Tet-On 3G (red) integration at the mouse *Rosa26* locus under EF1α promoter (grey) and pTRE3G-FKBPV-3xFlAG-Cas9 cassette at the mouse *Tigre* locus. **C.** Tet-On 3G binds to the pTRE3G promoter (dark blue) only in the presence of doxycycline (Dox) (green); FKBPV-Cas9 (magenta-yellow) transcription is induced. **C.** FKBPV-Cas9 protein degradation upon dTAG-13 treatment; dTAG-13 binds to FKBPV (magenta) and CRBN-E3 (blue).

To characterize the function of the inducible Cas9 system, we assessed both Dox-induced Cas9 expression and dTAG-mediated Cas9 degradation using Western blot analysis. To validate Dox-induced Cas9 expression, we treated the cells with serial dilutions of Dox (0.1, 1, and 4 µg) and harvested the cells at various time points (1h, 6h, 24h) (Figure 2.1.2A,B). Western blot analysis revealed that no Cas9 expression was detectable 1 hour post-Dox treatment compared to the untreated sample (Figure 2.1.2A). Pronounced Cas9 expression, however, was observed at 6 and 24 hours post-treatment (Figure 2.1.2A,B). Notably, we detected some background Cas9 expression (leakiness) in the absence of Dox, indicating a low level of basal transcription from the pTRE3G promoter (Figure 2.1.2A,B).

Similarly, to assess dTAG-mediated Cas9 degradation, we treated the cells with serial dilutions of dTAG-13 (100, 250, and 500 nM). Cas9 expression was analyzed at various time points (1h, 6h, 24h) by Western blot (Figure 2.1.2C,D). We observed a striking degradation of Cas9 after only 1 hour of exposure to dTAG-13, even at concentrations as low as 100 nM (Figure 2.1.2C). Moreover, no Cas9 expression was detectable at 6 and 24 hours post-treatment, demonstrating the potent and rapid degradation effect of the dTAG system (Figure 2.1.2D).

Finally, to demonstrate the reversibility of the system, cells were pre-treated with 100 nM dTAG-13 for 48 hours to deplete Cas9 (Figure 2.1.2E). Dox was then added to induce Cas9 expression, and we analyzed the expression at 6, 16, and 30 hours post-Dox treatment. We confirmed Cas9 protein recovery upon Dox treatment following dTAG-mediated degradation by Western blot (Figure 2.1.2E). This demonstrates that the system is not only inducible and degradable but also reversible, allowing for dynamic control of Cas9 activity. Overall, these data demonstrate the successful establishment of a tightly controlled, Dox-inducible, and dTAG-degradable Cas9 system in mESCs. This system provides a valuable and dynamic tool to perform the CRISPR/Cas9 screening with optimal conditions and dissect novel TE regulators involved in new biological processes. In the next sections, we focused on optimizing readouts to ensure robust detection of transcriptional and translational activation upon TE repressor depletion.

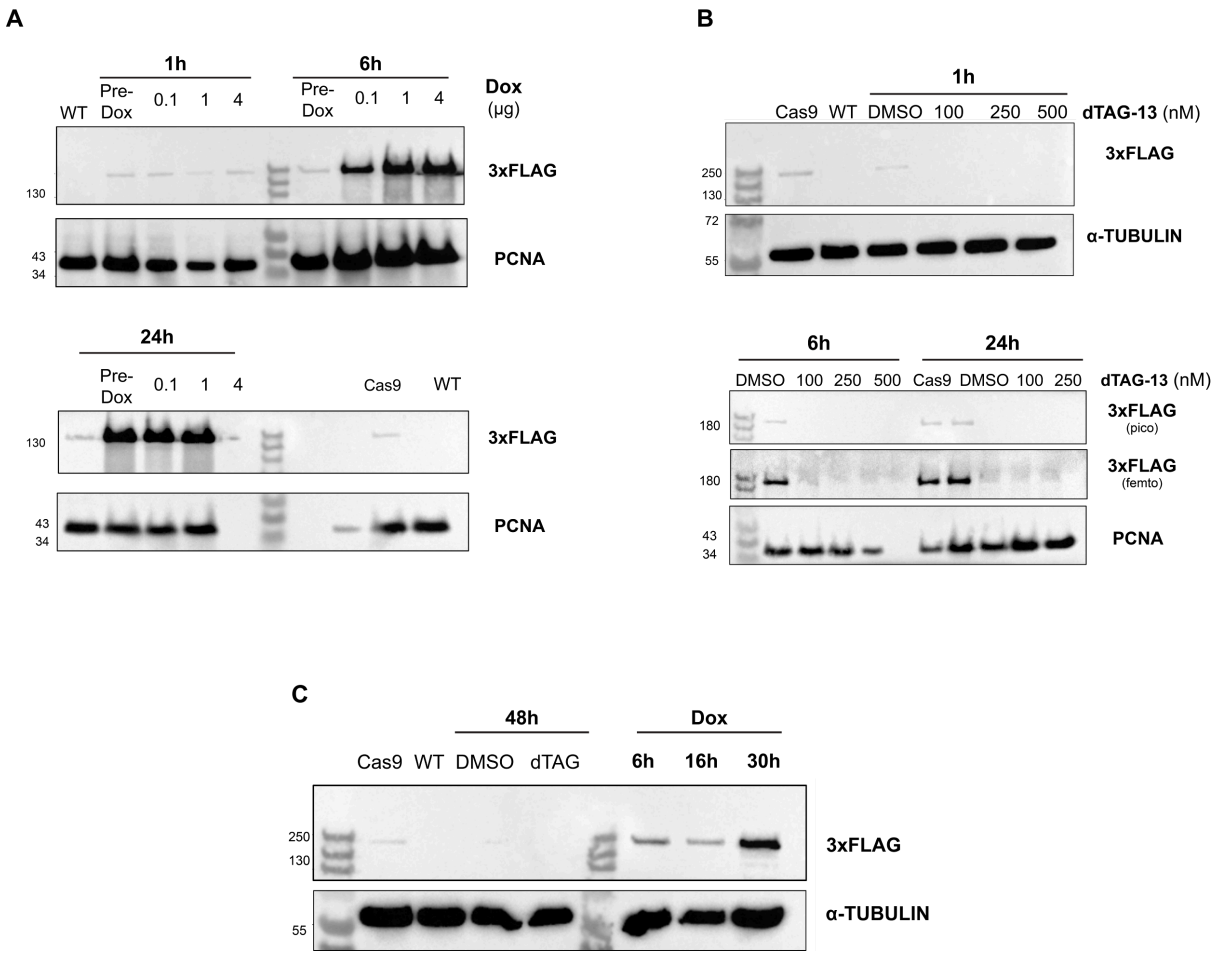


Figure 2.1.2 Cas9 kinetics upon doxycycline (Dox) induction and post-dTAG treatment. A. Western blot analysis of Cas9-3xFLAG reactivation upon serial dilutions of Dox and without Dox induction (Pre-Dox) at the indicated time-course; WT mESCs were used as a negative control; PCNA was used as a loading control. **B.** Cas9-3xFLAG degradation upon serial dilutions of dTAG-13 treatment at the given time-course; Cas9 cell line (without DMSO) and WT mESCs were used as a controls; The blot of 6 hours and 24 hours was probed with both pico and femto chemistry; α-TUBULIN (1hour) and PCNA (6 and 24 hours) were used as loading controls. **C.** Reversibility of the inducible system demonstrated by Western blot analysis at the indicated time points; In all blots Cas9 expression was detected using an anti-3xFLAG antibody.

2.1.2 Readout optimization of CRISPR/Cas9-forward genetic screening for identifying TE regulators

Building upon the foundational Cas9-expressing cell line described in the previous chapter, we next turned our attention to establishing robust and reliable readouts for our screening approach. As a proof-of-concept for our screening approach, we developed control systems targeting well-characterized TE repressors to serve as reliable readouts for identifying novel TE regulators. For ERVs, we selected KAP1 (TRIM28) for degradation, to induce robust transcriptional activation of ERVs (Rowe et al. 2010), for optimizing downstream assays such as RNA-FISH and immunofluorescence (IF). For LINE-1 elements, we chose to delete *Suv39h1/2*, genes encoding SUV39H1/2 methyltransferases, which regulate LINE-1 activity via H3K9me3-mediated silencing (Bulut-Karslioglu et al. 2014). These systems aimed to allow for controlled, rapid upregulation of both TE families, providing a foundation for establishing screening protocols with proper controls.

2.1.2.1 Generation of KAP1-degron cell line for ERV reactivation

To establish a robust system for conducting the screening, we required a model in which the depletion of a *bona fide* TE repressor would result in strong transcriptional activation of ERVs. As a proof-of-concept, we degraded KAP1 (TRIM28) protein, a well-characterized TE repressor in mESCs (Rowe et al. 2010), serving as a readout for our experimental setup. KAP1 mediates transcriptional silencing of ERVs by recruiting chromatin modifiers to their loci and its depletion causes significant ERV reactivation (Schultz et al. 2002; Rowe et al. 2010), providing a reliable control system for optimizing the RNA-FISH and IF (Hypothesis and Aim).

Considering that *Kap1* is an essential gene (Rowe et al. 2010), we chose the dTAG-13 system to cause rapid and inducible KAP1 degradation (Figure 2.1.3A). To achieve precise temporal control over KAP1 expression, we tagged its C-terminus with the FKBP-V degron using CRISPR-Cas9-mediated homologous recombination (Figure 2.1.3B). The cell line was generated by co-transfecting the donor plasmid carrying the FKBP-V-2xHA-P2A-puromycin cassette with the editor carrying the sgRNA in wild-type (WT) mESCs (Figure 2.1.3B). The in-frame homozygous knockin (KI) integration was further confirmed by PCR (Figure 2.1.3C).

After generating the homozygous KAP1 degron KI cell line, we evaluated the kinetics of KAP1 degradation upon serial dilution of dTAG-13 (250nM, 500nM, 1μM) treatment and different time points (Figure 2.1.3D). Western blot analysis revealed partial KAP1 degradation already at 2 hours post-treatment, a remarkable degradation after 4 hours

exposure to dTAG-13 and complete degradation after 24 and 48 hours with as little as 250 nM of the drug (Figure 2.1.3D). These results demonstrate the dTAG system enables rapid and efficient KAP1 degradation.

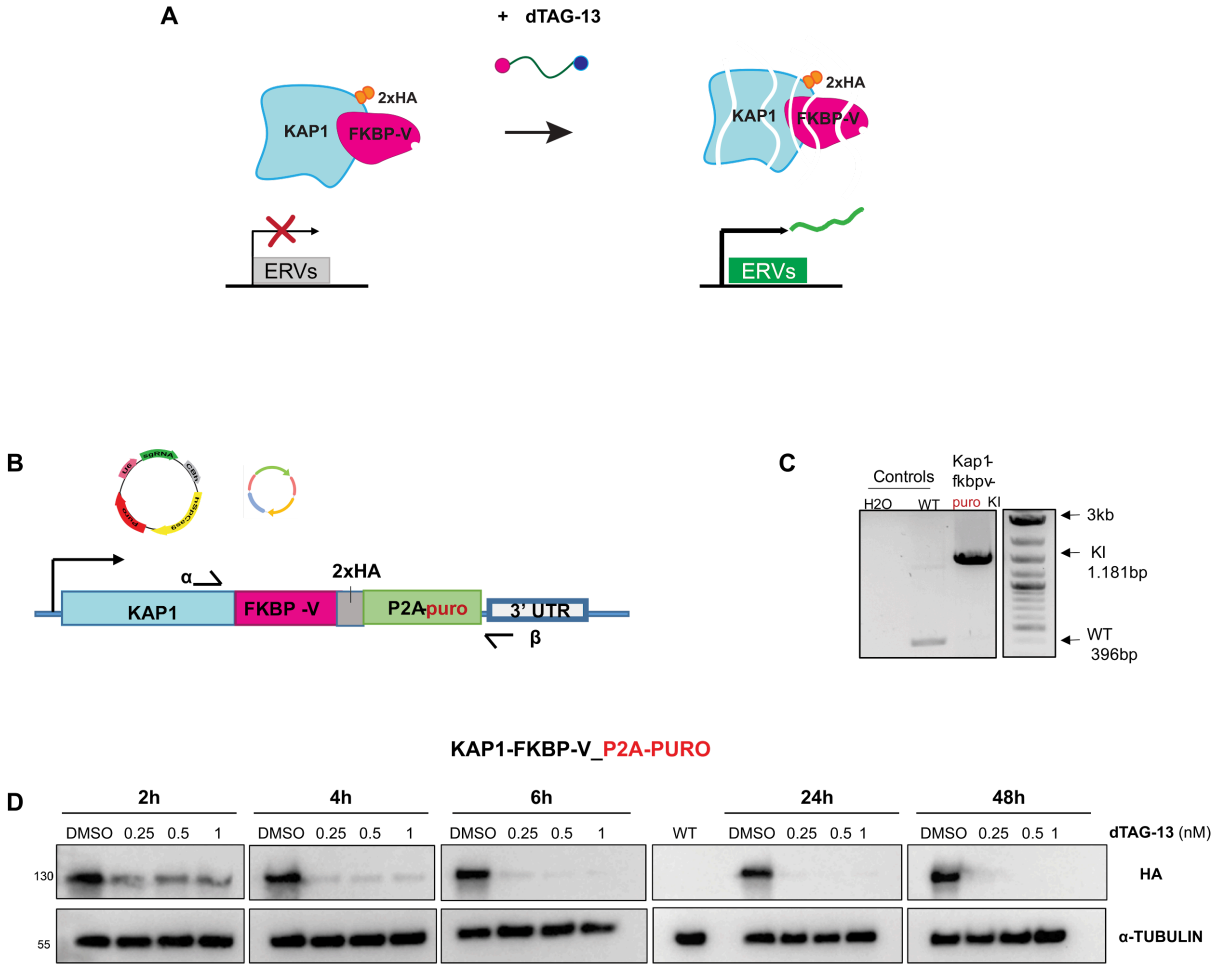


Figure 2.1.3 Generation of KAP1-degron cell line and KAP1 kinetics upon dTAG-13 treatment. **A.** Schematic representation of the dTAG degron system. In the absence of dTAG-13 KAP1 represses ERVs. Upon dTAG-13 treatment KAP1 is degraded leading to ERV transcriptional reactivation. **B.** Schematic depiction of FKBP-V KI stable puromycin selection strategy. Co-transfection of the FKBP-V donor plasmid and the editor plasmid in E14 WT ESCs; Primers used for the PCR screening are depicted by arrows (α,β) **C.** PCR screening for genotyping. The cell line has a 1181bp KI fragment (α,β); WT mESCs were used as a positive control and H2O as a negative control. **D.** Western blot analysis expressing the KAP1-FKBP-V fusion protein in the KI cell line. Protein extracts were probed with an anti-HA antibody to detect the FKBP-V tag; α-TUBULIN was used as a loading control.

2.1.2.2 Transcriptional ERV reactivation upon KAP1 depletion

To monitor the TE transcriptional expression levels upon acute KAP1-degradation, we isolated total RNA from the KAP1-FKBP-V-puromycin cell line. Similar to the kinetics experiment (Figure 2.1.3D), the cells were treated with the same dTAG-13 concentrations for the same time-course (Figure 2.1.4A,B). We additionally included 72-hour and 96-hour dTAG-13 exposure time points and performed quantitative reverse transcription PCR (RT-qPCR) for all samples. Although KAP1 degradation was observed four hours post-treatment (Figure 2.1.3D), LTR-element transcripts showed reactivation after 48 hours (Figure 2.1.4A). More specifically, when targeting IAPez, IAP-Δ1, MuERV-L and MMERK10C transcripts we observed a 12- to 15-fold increase in expression between 48 hours and 96 hours post-treatment, whereas IAP-gag showed more than 20-fold upregulation at 96 hours post-treatment (Figure 2.1.4A,B). When targeting the 5' UTR of LTR elements, we observed a remarkable 40-fold increase in expression (Figure 2.1.4B), confirming the results of Rowe et al. (Rowe et al. 2010). It is not surprising that IAP 5' UTR exhibited the highest upregulation in KAP1 depletion for two reasons. First, KAP1 has previously been found to be enriched at the 5′ UTR of IAP genomes (Rowe et al. 2010) and second, in mammalian genomes recombination deletion between LTRs is a conserved mechanism across mammals, leading to a much higher abundance of solo LTRs compared to full-length ERVs (Stoye 2001; Thompson, Macfarlan, and Lorincz 2016; M. Chen et al. 2024).

Having confirmed the upregulation of IAPs in KAP1-depleted cells, we optimized fluorescent in-situ hybridization (RNA-FISH), to monitor the mRNA abundance of LTR-elements in a potential screening-set-up. We generated homemade nick translation RNA-FISH probes targeting full-length LTR elements (Figure 2.1.4C,D). First, we tested their efficacy on KAP1-FKBPV cells at 48 hours post-dTAG treatment, the time point at which IAP reactivation is initiated (Figure 2.1.4A), while including DMSO-treated KAP1-degron cells at the same time point as a negative control. We detected stronger IAP RNA-FISH foci on the KAP1-depleted cells compared to the negative control (Figure 2.1.4C,D), confirming results from previous studies (Asimi et al. 2022). Next, we performed the RNA-FISH in-suspension to optimize the transcriptional readout of the endogenous LTR-elements. Although we detected positive IAP mRNA expression in the absence of KAP1, we encountered limitations in producing stronger signals compared to the negative control. As a result, this limitation would restrict efficient sorting of positive populations in future genome-wide screening experiments (Figure 2.1.4E,F).

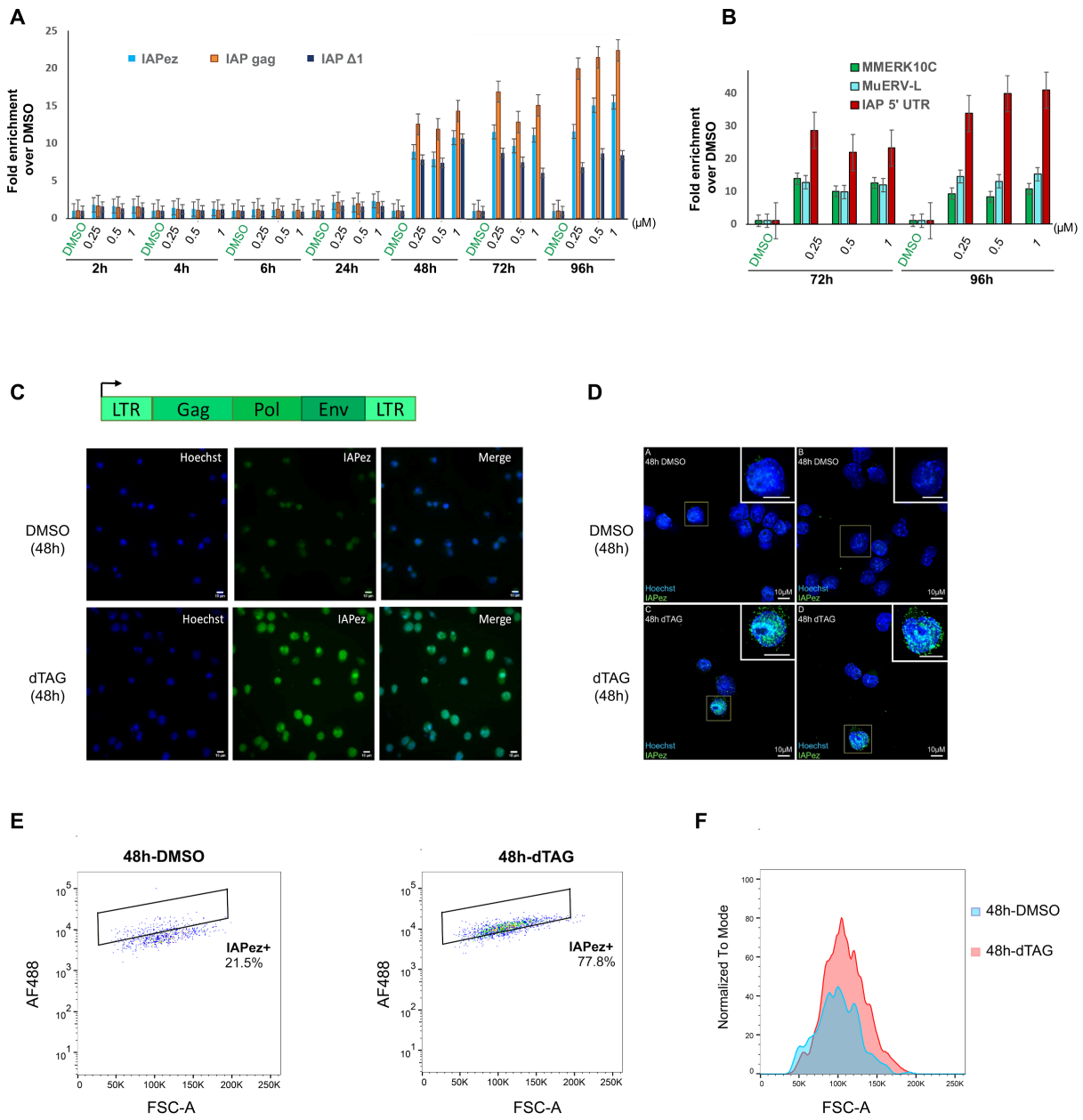


Figure 2.1.4 ERV reactivation in KAP1-depleted cells upon dTAG-13 treatment by RT-qPCR and RNA-FISH. A. Gene expression levels of IAPez (light blue), IAP-gag (orange), IAP-Δ1 (dark blue) of KAP1-FKBP-V-puro cell line treated with DMSO or serial dilutions of dTAG-13 for the indicated time-course. **B.** MMERK10C (green), MuERV-L (turquoise), IAP 5' UTR (red) upregulation at 72 hours and 96 hours post-treatment. **C.** Schematic depiction of the targeted full-length LTR element. Representative images of RNA-FISH signals on KAP1-depleted cells, and an image of the merged channels. The nuclear staining is determined by Hoechst (blue); Alexa-488-labeled probes detect IAPez transcripts (green); scale bars, 10μm. **D.** As C. Representative deconvoluted image of RNA-FISH; The image was deconvoluted with the assistance of our Postdoc. **E.** Scatter plots of KAP1-depleted cells and DMSO-treated cells (negative control) after hybridization with probes against IAPez mRNA. AF488 fluorescence was recorded using a 488-nm laser and 530/30-A BL filter. A total of 1,000 events per sample were analyzed. **F.** Histogram showing fluorescence intensity distribution for IAPez-positive populations from flow cytometry analysis of E., normalized to mode; Normalization ensures that the peaks of both distributions are scaled to the same height, for direct comparison of fluorescence intensity shifts between DMSO- (blue) and dTAG-treated (red) samples.

2.1.2.3 Translational regulation of IAPs

2.1.2.3.1 Validation of in-suspension IF for L1-ORF1p in mouse testes

As a readout for mRNA translation, we aimed to perform an in-suspension IF for TE-encoded proteins in mESCs and test their isolation by FACS. To validate the efficiency of the technique, we used juvenile and adult from *Dnmt3C*$^{KO/KO}$ testes known to highly translate L1-ORF1p (Barau et al. 2016) (Figure 2.1.5). Single cell suspensions were prepared from WT and *Dnmt3C*$^{KO/KO}$ juvenile and adult testes and incubated with L1-ORF1 antibody following a standard IF protocol. Flow cytometry analysis detected L1-ORF1 negative and positive cell populations, with up to 45.3% positive cells in the *Dnmt3C*$^{KO/KO}$ adult testes (Figure 2.1.5A) and 35.8% in juvenile testes (Figure 2.1.5B), compared to only 3.35% in WT testes (Figure 2.1.5C). These percentages of positive populations are consistent with the proportion of germ cells typically found in mouse testes (Figure 1.7B). A no-primary antibody control confirmed the specificity of the staining, showing only 1.34% background (Alexa546-positive cells) (Figure 2.1.5D). These results ensured that, technically, the in-suspension IF protocol works effectively in a system known to express L1-ORF1p. However, while we did not test the isolation of positive cells by FACS, these results strongly indicate that this approach could be feasible for future experiments.

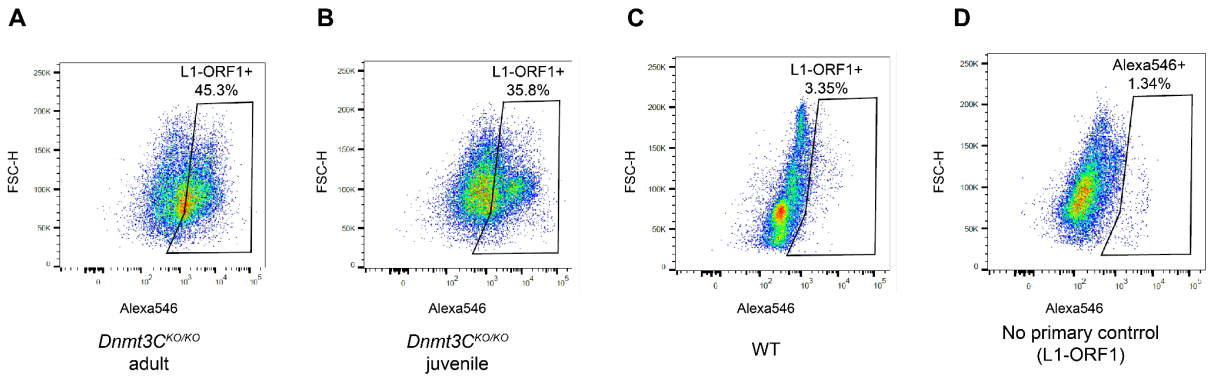


Figure 2.1.5 Flow cytometric analysis of L1-ORF1 protein expression in *Dnmt3C*$^{KO/KO}$ germ cells. A. Scatter plot showing 45.3% frequency of germ cells, from adult *Dnmt3C*$^{KO/KO}$ testis, positive for L1-ORF1 protein after incubation with Alexa 546 secondary antibody. **B.** Scatter plot showing 35.8% frequency of germ cells, from 17dpp *Dnmt3C*$^{KO/KO}$ testis, positive for L1-ORF1 protein after incubation with the same antibodies dilution, shown in A. **C.** Scatter plot showing 3.35% frequency of germ cells, from a WT mice, after incubation with the same antibodies dilution, described in A and B. **D.** Scatter plot showing AF546 emission above the set threshold after incubation with secondary antibody only (negative control). Alexa 564 fluorescence was recorded using a 561-nm laser and a 610/20 bandpass (BP) filter. A total of 10,000 events per sample were analyzed.

2.1.2.3.2 Validation of custom-made IAP-GAG antibodies

After validating the in-suspension IF for L1-ORF1 detection *in vivo* (Figure 2.1.5), we sought to establish the technique in our KAP1-degron cell line as a readout for IAP mRNA translation. Ultimately, we aimed to perform IF against IAP-GAG in KAP1-depleted cells and test their sorting by FACS. Given that KAP1 knockout mice have been shown to cause IAP-GAG reactivation in E3.5 blastocysts (Asimi et al. 2022), we aimed to see whether this reactivation also occurs in our degron system. Since IAP-GAG antibodies are not commercially available, we collaborated with the IMB Protein Production Core Facility to generate custom-made antibodies targeting the epitope of IAP-GAG. Two polyclonal antibodies (IAP-GAG 1294 and IAP-GAG 1295) were purified from two individual rabbits immunized with an MBP-GAG fusion protein.

To evaluate the efficiency of both custom-made IAP-GAG antibodies, we conducted an IF staining on KAP1-degron cells and *Dnmt3C*$^{KO/KO}$ testes, a system known for TE reactivation in meiosis (Barau et al. 2016) (Figure 2.1.6). KAP1-degron cells were treated with DMSO or dTAG-13 for 48 hours. HA-tag staining confirmed successful KAP1 degradation in dTAG-treated cells, as KAP1 was detected in DMSO-treated cells but absent after 48 hours of dTAG treatment (Figure 2.1.6A,B). However, neither IAP-GAG antibody detected specific foci in either condition (Figure 2.1.6A,B), despite previous evidence of ERV mRNA reactivation at 48 hours post-dTAG treatment (2.1.2.2) (Asimi et al. 2022). Next, we leveraged *Dnmt3C*$^{KO/KO}$ testes, and performed an IF at postnatal day (P10) (Figure 2.1.6C), when germ cells enter meiosis (1.3.1 Mouse male germline) (Figure 1.7B). TRA98 staining was used as a germ cell marker, while L1-ORF1 served as a positive control for TE-encoded proteins. Although we detected L1-ORF1, we were not able to detect IAP-GAG using either antibodies (1294 and 1295) by microscopy (Figure 2.1.6C). Taken together, these results indicate that IAP-GAG protein is not detectable under these experimental conditions, either due to low expression levels or insufficient antibody sensitivity.

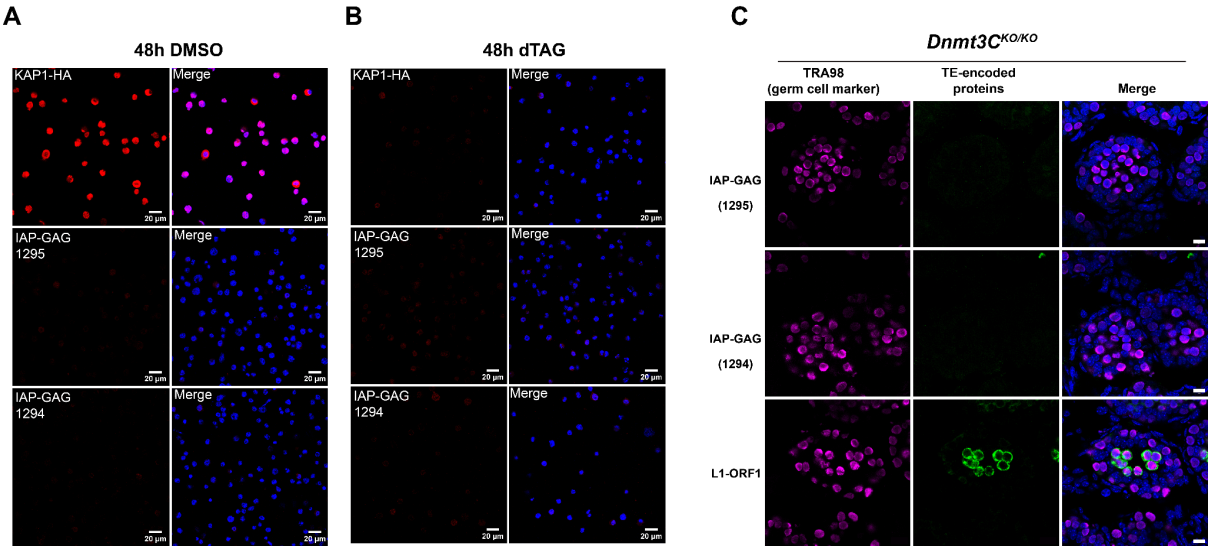


Figure 2.1.6 Custom IAP-GAG antibodies did not detect protein in KAP1-degron and *Dnmt3C*$^{KO/KO}$ testes by microscopy. A. Immunofluorescence staining of KAP1-HA protein (red) and IAP-GAG (1295 and 1294) in KAP1-degron cells treated with DMSO for 48 hours; scale bars, 20μm **B.** as A. for 48 hours post-dTAG-13 treatment in KAP1-degron cells. **C.** Immunofluorescence of L1-ORF1 protein, IAP-GAG (1295 and 1294) (green), TRA98 (germ cell marker, magenta) and DAPI (nucleus staining, blue) on cryosections of WT and *Dnmt3C*$^{KO/KO}$ testes at P10; scale bar, 10μm.

To further validate the IAP-GAG antibody we performed IF on single cell suspension at P25 and WT and *Dnmt3C*$^{KO/KO}$ mice testes (Figure 2.1.7). As a positive control we included L1-ORF1 staining for *Dnmt3C*$^{KO/KO}$ testes (Figure 2.1.7A), which confirmed that the technique worked as expected (2.1.2.3.1). However, when testing both IAP-GAG antibodies, no difference in signal was observed between WT and *Dnmt3C*$^{KO/KO}$ (Figure 2.1.7B,C). We next tested whether in-suspension IF would work for KAP1-degron cells treated with DMSO or dTAG-13, as this is ultimately our goal for screening applications (Figure 2.1.7D). Surprisingly, flow cytometry revealed extremely high percentages (~99%) of L1-ORF1-positive cells in both DMSO- and dTAG-treated populations (Figure 2.1.7D). This result was unexpected because DMSO-treated cells serve as a negative control where no TE reactivation should occur, raising concerns about the antibody specificity. While no-primary antibody control showed minimal background (~0%), non-specific primary antibody interactions could still contribute to this signal. The high L1-ORF1-positive signal suggests potential non-specific binding of the antibody.

To further assess antibody performance, we conducted Western blot analysis using protein extracts from *Dnmt3C*$^{KO/KO}$ and WT testes (Figure 2.1.7E). Although the expected full-length IAP-GAG is around 55kDa, we observed bands at around 95kDa that do not correspond to either the full-length IAP-GAG or the epitope. Given that these bands were observed in both WT and *Dnmt3C*$^{KO/KO}$ samples, this suggests that they are likely background signals. To verify whether the antibody can recognize its antigen, we loaded a

purified MBP (Maltose-Binding Protein)-GAG fusion protein. Additionally, to confirm that the antibody doesn't nonspecifically bind to any MBP, which is a commonly used fusion tag for protein expression and purification (Reuten et al. 2016), we loaded another MBP fusion protein (MBP-X), as a control. The antibody successfully detected both MBP-GAG (~70 kDa) and cleaved GAG (~13 kDa) following incubation with 3C protease and it did not strongly bind to MBP-X (Figure 2.1.7E). These results confirmed that the antibody can bind specifically to its target epitope under ideal conditions but fails to detect endogenous GAG protein. Overall, these data demonstrate that while the KAP1-degron system has strong potential for investigating novel ERV regulators, further optimization of this readout approach is required, which was beyond the scope of this study due to limitations in resources and time.

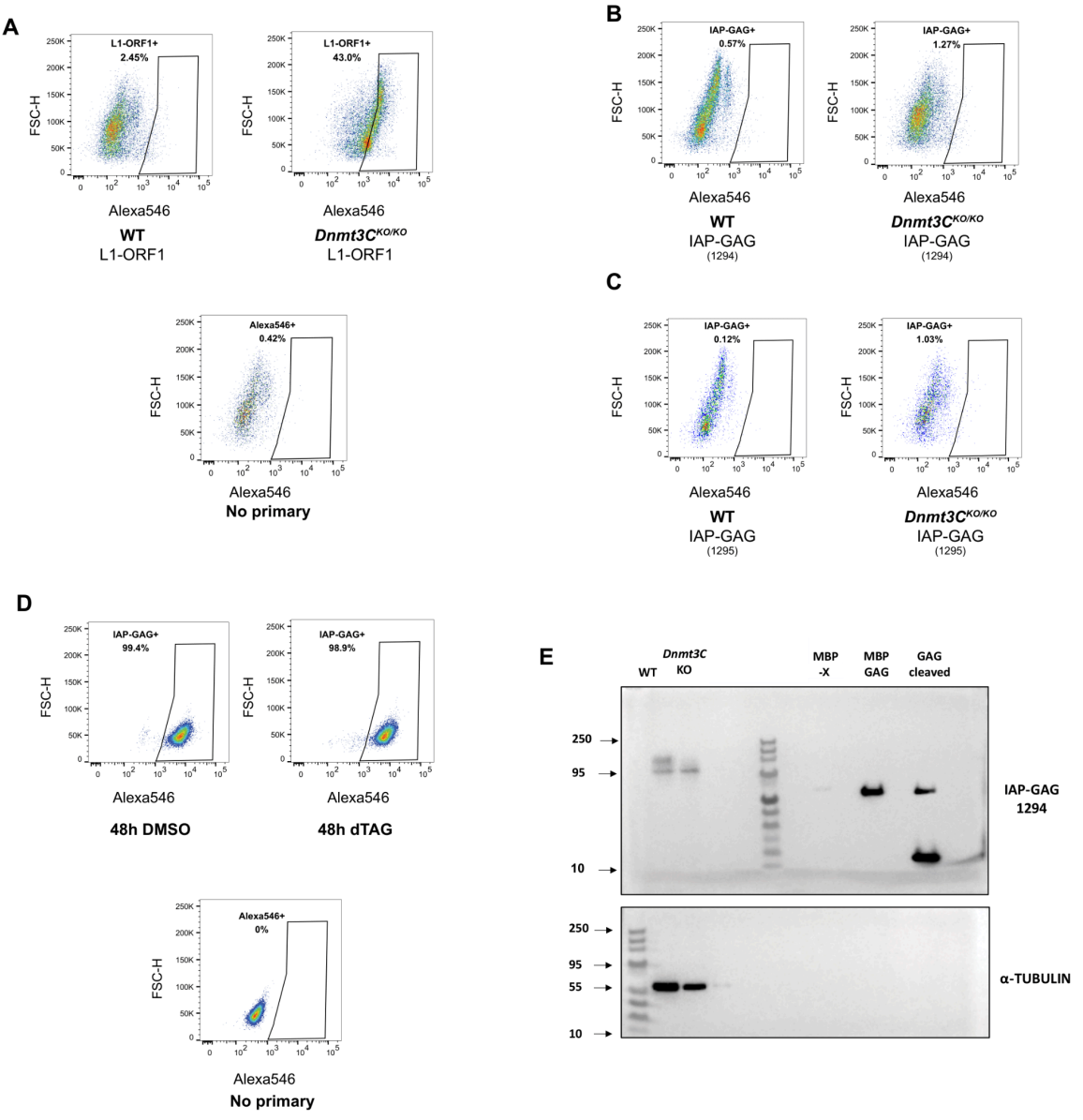


Figure 2.1.7 Custom IAP-GAG antibodies bind to the epitope *in vitro* but not *in vivo*. A. Scatter plot showing 43.0% frequency of germ cells from *Dnmt3C*$^{KO/KO}$ testes positive for L1-ORF1 protein after incubation with Alexa 546 secondary antibody; WT testes showed 2.45% L1-ORF1-positive cells; No primary (negative control) showing AF546 emission above the set threshold after incubation with secondary antibody only **B.** as A. Scatter plot showing IAP-GAG positive cells (1294 antibody) in WT (0.57%) and *Dnmt3C*$^{KO/KO}$ (1.27%) testes. frequency of germ cells from Dnmt3CKO/KO testes **C.** as A. Scatter plot showing IAP-GAG positive cells (1295 antibody) in WT (0.12%) and *Dnmt3C*$^{KO/KO}$ (1.03%) testes. **D.** as A. Scatter plot showing IAP-GAG-positive cells in KAP1-degron cells treated with DMSO or dTAG-13 for 48 hours (~99% Alexa546-positive in both conditions); no-primary antibody control showed minimal background fluorescence (~0%). **E.** Western blot analysis of IAP-GAG detection using the 1294 antibody. Protein extracts from WT and *Dnmt3C*$^{KO/KO}$ whole testes; MBP-X control protein, MBP-GAG fusion protein (~70 kDa), and cleaved GAG antigen (~13 kDa) were analyzed; α-TUBULIN was used as a loading control.

2.1.2.4 Depletion of *Suv39h1/2* and KAP1 as a system for endogenous LINE-1 and IAP regulation

In the previous chapter, we explored the KAP1-degron system as a readout for ERV reactivation. However, we faced challenges optimizing this approach due to limitations in the efficiency of IAP-GAG custom-made antibodies. Given these restrictions, we shifted our focus to optimizing the readout for LINE-1 upregulation, leveraging an available antibody that we previously validated to effectively detect L1-ORF1 protein in mouse testes by IF (2.1.2.3.1). Although KAP1 is an efficient control for ERV upregulation, it is ineligible for monitoring young LINE-1 upregulation (Castro-Diaz et al. 2014). Hence, we directed our attention to SUV39H1/2 methyltransferases, that have been previously reported to regulate LINE-1s in mESCs by H3K9me3-mediated silencing (Bulut-Karslioglu et al. 2014) (Figure 2.1.8A). We generated a homozygous *Suv39h1/2* double-KO (*Suv39h*-dKO) ES cell line by a deletion in the catalytic domain of *Suv39h1* (Figure 2.1.8B) and *Suv39h2* (Figure 2.1.8C), respectively, leading to a frameshift mutation confirmed by PCR (Figure 2.1.8D,E). We additionally used the previously generated KAP1-FKBP-V cell line to obtain a compound *Suv39h*-dKO-KAP1-degron-based mutant cell line, to simultaneously induce ERV and LINE-1 reactivation.

We performed a RT-qPCR to validate the LINE-1 and LTR mRNA expression levels upon *Suv39h1/2* and KAP1 deletion in the generated mutant cell lines (Figure 2.1.8F,G). First, we validated the two *Suv39h1/2* cell lines generated in both WT mESCs and *Kap1-fkbpv* (no dTAG-13) background by comparing the mRNA transcription of previously characterized targets, L1MdA_I (L1-promoter) and L1-ORF1 (Bulut-Karslioglu et al. 2014). We observed up to 2-fold increase in expression levels between both cell lines compared to WT mESCs (Figure 2.1.8F). Next, we evaluated the LINE-1 expression levels of evolutionary young LINE-1 subfamilies in the absence of *Suv39h1/2* and KAP1 (Figure 2.1.8G). We

observed a mild upregulation in L1MdA_I (L1-promoter), L1-ORF1, L1-ORF2, and L1-Tf in the absence of *Suv39h1/2* compared to WT cells. This reactivation did not become stronger in the compound *Suv39h*-dKO-KAP1-degron-based mutant cell line upon 48 hours post-dTAG treatment (Figure 2.1.8G), confirming the minimal effect of KAP1 in LINE-1 regulation (Castro-Diaz et al. 2014). Interestingly, in *Suv39h-dKO* we observed a 5-fold increase of expression when targeting L1-Gf, which mildly increased further in the absence of *Suv39h1/2* and KAP1 (Figure 2.1.8G). Moreover, IAP-GAG transcripts showed around 10-fold upregulation in *Suv39h*-dKO-KAP1-degron-based cells, confirming our previous results (Figure 2.1.4A). These data indicate that the *Suv39h*-dKO-KAP1-degron-based mutant cell could be utilized to monitor the regulation of both LINE-1 and ERV elements.

To further validate L1-ORF1 expression on a translational level in the absence of *Suv39h1/2*, we performed an IF staining against L1-ORF1p and visualized the results by microscopy and flow cytometry (Figure 2.1.8H-K). Interestingly, while we observed L1-ORF1p expression in *Suv39h*-dKO-*Kap1-fkbpv* (without dTAG-13 treatment), we also detected a baseline L1-ORF1p expression in WT mESCs (Figure 2.1.8H). We then validated the microscopy results by a single-cell suspension IF against L1-ORF1p (Figure 2.1.8I-K). We detected high levels of L1-ORF1p in *Suv39h*-dKO cells (43.4%) (Figure 2.1.8I) and *Suv39h*-dKO-*Kap1-fkbpv* (no dTAG-13 treatment) (46.3%) (Figure 2.1.8J), along with slightly lower L1-ORF1p expression in WT mESCs (36.1%) (Figure 2.1.8K). Overall, while we observed L1-ORF1p expression in *Suv39h-dKO* cells, the modest upregulation and high baseline expression levels in WT mESCs limit their utility as a reliable readout for genome-wide screening experiments. Moreover, to mimic a genome-wide screening, with an infection rate of one sgRNA per cell, additional validation would be required for single-gene knockouts (*Suv39h1* or *Suv39h2*), which is beyond the scope of this study.

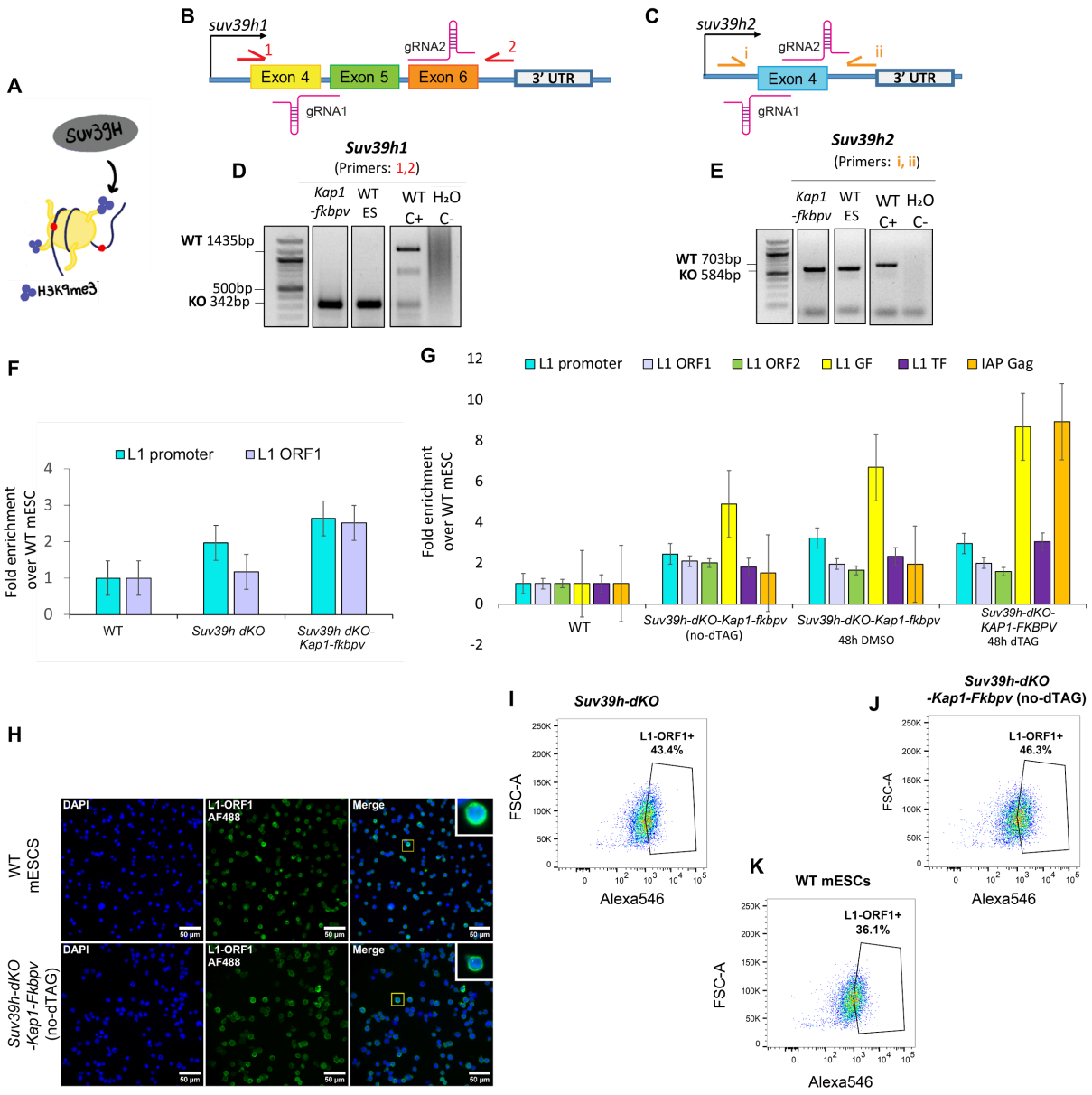


Figure 2.1.8 Analysis of *Suv39h-dKO* and compound *Suv39h-dKO*-KAP1-depleted cell lines and their impact on LINE-1 and IAP upregulation. A. Schematic representation of SUV39H-mediated H3K9me3 deposition at retrotransposon loci; Illustration generated by Dimitra Kanta. **B. and C.** CRISPR/Cas9 strategy for generating *Suv39h1* (B) and *Suv39h2* (C) knockouts; guide RNAs (gRNAs) target exon 5 and exon 6 for *Suv39h1* and exon 4 for *Suv39h2*; Primer pairs used for genotyping are indicated (1, 2 for Suv39h1; i, ii for Suv39h2). **D. and E.** Genotyping PCR results confirming *Suv39h1* (D) and *Suv39h2* (E) knockouts. WT bands correspond to the expected product sizes (1435bp for *Suv39h1*; 703bp for *Suv39h2*), while knockout (KO) bands show shorter amplicons due to CRISPR-induced deletions (500bp for *Suv39h1*; 584bp for *Suv39h2*). WT mESCs were used as Positive (+) control (C) for the WT band and H_2O was used as negative (-) control. **F.** RT-qPCR analysis showing fold enrichment of L1 promoter and L1 ORF1 transcripts in WT mESCs, *Suv39h-dKO* cells, and *Suv39h-dKO-Kap1-fkbpv* cells relative to WT mESCs. **G.** RT-qPCR analysis showing fold enrichment of various LINE-1 subfamilies (L1 promoter, L1 ORF1, L1 ORF2, L1-Gf, L1-Tf) and IAP-Gag transcripts in WT mESCs, *Suv39h-dKO* cells, and *Suv39h-dKO-Kap1-fkbpv* cells under different conditions: no dTAG treatment, 48 hours DMSO treatment, and 48 hours dTAG treatment. **H.** Immunofluorescence staining of L1-ORF1p (green) in WT mESCs and *Suv39h-dKO-Kap1-fkbpv* cells without dTAG treatment; nuclei were stained by DAPI (blue); Scale

bars, 50 µm. **I–K.** Flow cytometry analysis of L1-ORF1p expression in **Suv39h-dKO** cells (I), *Suv39h-dKO-Kap1-fkbpv* cells without dTAG treatment (J), and WT mESCs (K); Percentages of Alexa546-positive cells are indicated: 43.4% for Suv39h-dKO cells, 46.3% for Suv39h-dKO-Kap1-fkbpv cells, and 36.1% for WT mESCs

2.1.3 Schematic overview of the project

In light of the multiple challenges encountered, and given the limited timeframe, we made a strategic decision to suspend this project. Considering the ongoing research priorities in our laboratory, I directed my efforts in assisting to complete the project of Jessica Leismann, which at the time was in risk of remaining unfinished due to time constraints. The subsequent chapters will present the results of this collaborative project, to which I contributed significantly and led to a co-first author publication (in revision). The following chart summarizes the outcomes and challenges of our initial investigation (Figure 2.1.9).

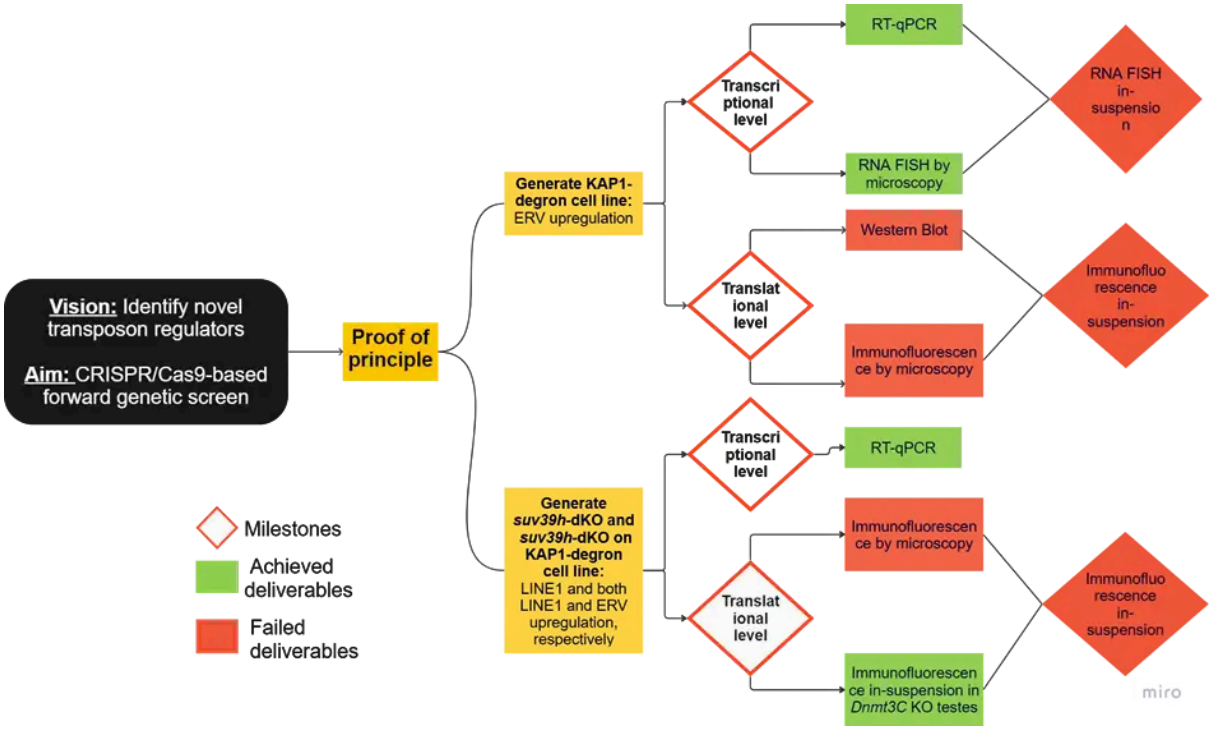


Figure 2.1.9 Flow chart diagram of my previous project. Milestones are depicted in a diamond shape. Achieved deliverables (experiments) are illustrated in green. Failed deliverables (experiments) are depicted in red. The chart was created using Miro.

2.2 NRF1 as a trans-activator of young unmethylated IAPs in the male germline

Chapter 2.2 represents a collaborative effort between my colleague Jessica Leismann and myself that resulted in a co-first authored manuscript submitted to EMBO Reports (currently under revision). This project constituted Jessica's primary research focus for over 5.5 years and benefited from my contributions over a 1.5-year period. Jessica conducted all mice crosses/breadings and testes collection at specific developmental time points, as well as germ cell-specific sorting by FACS (2.2.1), (2.2.2), the DNA pulldown experiment (2.2.2), and NRF1 chromatin profiling throughout spermatogenesis (2.2.2). The bioinformatic analysis was performed by IMB Bioinformatics Core Facility and Jessica Leismann.

RNA isolation was performed collaboratively by Jessica Leismann and myself (2.2.1.1). I conducted all immunofluorescence/microscopy experiments and Western blots shown in the following sections. Additionally, I performed chromatin profiling of NRF1 and H3K27me3 in mESCs (2.2.4). The scripts for the quantification of microscopy images were written by the IMB Microscopy and Histology Core Facility.

2.2.1 TE expression dynamics throughout *Dnmt3C*$^{KO/KO}$ spermatogenesis

To understand the role of DNA methylation specifically on the promoters of evolutionary young and active retrotransposons throughout spermatogenesis, we comprehensively explored their expression on both transcriptional (Figure 2.2.1) and translational levels (Figure 2.2.2). DNMT3C specifically methylates evolutionarily young retrotransposon promoters in mice (Barau et al. 2016). Despite defects in its function, mice retain overall stable global DNA methylation levels, while young retrotransposons lack DNA methylation at their promoters (Barau et al. 2016). We used *Dnmt3C*$^{KO/KO}$ mice as our experimental model to address TE activity in spermatogenesis, and compared the expression levels to WT mice.

2.2.1.1 Transcriptional TE reactivation in the absence of DNA methylation

To assess the transcriptional expression of young retrotransposon copies with respect to the DNA methylation status of their promoters, we first performed RNA-seq on FACS-sorted germ cells of *Dnmt3C*$^{KO/KO}$ and WT mice. Our analysis focused on two key postnatal spermatogenesis stages: mitotic stage or pre-meiotic cells/spermatogonia (Spg) and meiotic stage or meiotic cells/spermatocytes (Spc).

In *Dnmt3C*$^{KO/KO}$ spermatogonia (pre-meiotic cells), the most significantly upregulated TEs belonged to the the ERV-K subclass (class II ERVs), the most active and mutagenic TEs in mice (Dewannieux et al. 2004) (Figure 2.2.1A). The predominantly affected ERV-K elements included IAPs (with LTR1 and LTR2 variants), MMERK10C, and RLTR10C. This is consistent with a previous study, in which IAP upregulation was reported due to piRNA pathway deficiency and subsequent loss of DNA methylation (Vasiliauskaitė et al. 2018). Interestingly, we detected upregulation of young and active LINE-1 elements (L1MdGf, L1MdT, L1MdA) in *Dnmt3C*$^{KO/KO}$ spermatogonia. This phenotype was previously observed in *Dnmt3L*$^{KO/KO}$ and *Miwi2*$^{KO/KO}$ sorted spermatogonia (Vasiliauskaitė et al. 2018), but there was no translational L1-ORF1 activity detected at this stage (Vasiliauskaitė et al. 2018; Zamudio et al. 2015).

In *Dnmt3C*$^{KO/KO}$ spermatocytes (meiotic cells), both ERVs and LINE-1 elements remained transcriptionally upregulated compared to WT, while several ancient LTRs and DNA transposons showed downregulation (Figure 2.2.1B). These changes likely reflect the altered cellular composition of testes from *Dnmt3C*$^{KO/KO}$ compared to WT mice in meiosis, as

spermatocytes arrest at the pachytene stage in the mutants (Barau et al. 2016). To visualize how the loss of *Dnmt3C* affects ERVs and LINE-1s expression patterns across spermatogonia and spermatocytes, we generated a bubble chart combining Z-scores and logarithmic transformed (Log) Transcripts per Million (TPM) expression values obtained from the RNA-seq (Figure 2.2.1C). LINE-1 elements (L1Md_T, L1Md_G, L1Md_A, L1_Mm) showed progressive upregulation from spermatogonia to spermatocytes in *Dnmt3C*$^{KO/KO}$ mice, while ERVs maintained consistent upregulation in both *Dnmt3C*$^{KO/KO}$ stages.

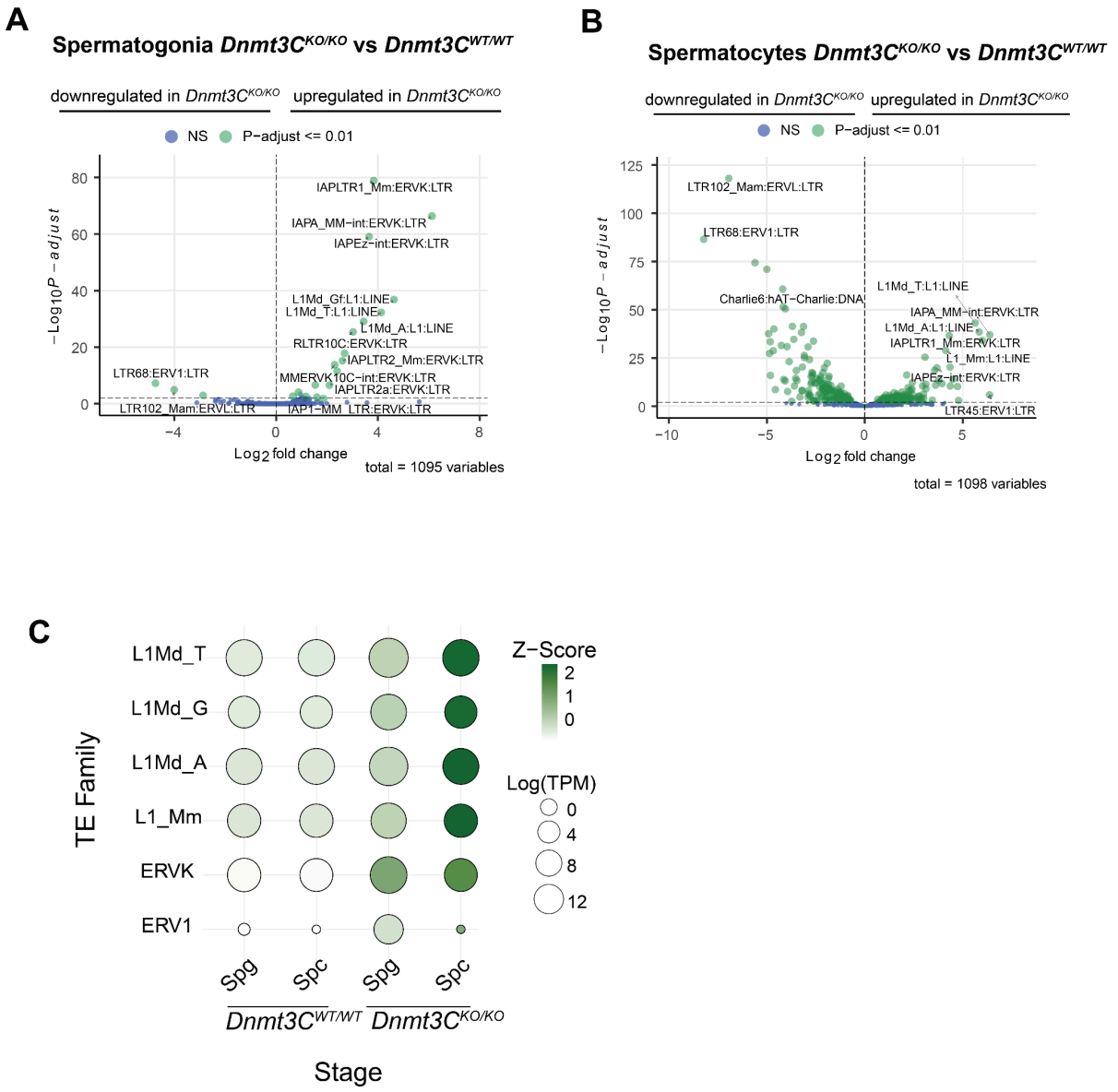


Figure 2.2.1 Young retrotransposons are transcriptionally upregulated prior to meiosis in *Dnmt3C*$^{KO/KO}$ germ cells. A. Volcano plot showing log10P-adjusted values and log2 fold changes of differentially expressed TE-transcripts between wild type and *Dnmt3C*$^{KO/KO}$ sorted germ cells at spermatogonia stages. **B.** as **A.** at spermatocyte stages. **C.** Bubble chart displaying Z-score and log(TPM) values in postnatal wild type (spermatogonia and spermatocytes) and *Dnmt3C*$^{KO/KO}$ (spermatogonia and spermatocytes) for TE families that were differentially expressed in

spermatogonia and spermatocytes. Due to the high similarity between wild type and *Dnmt3C*$^{KO/KO}$ in postnatal wild type spermatogonia and spermatocytes,these data points were combined.

2.2.1.2 Translational TE reactivation in the absence of DNA methylation

To validate the transcriptional changes and gain more developmental resolution of the protein expression dynamics, we next analyzed retrotransposon activity at specific developmental time points. We focused on the detection of L1-encoded ORF1 proteins (L1-ORF1p) across WT and *Dnmt3C*$^{KO/KO}$ spermatogenesis. Taking advantage of the synchronized nature of the first wave of spermatogenesis, we collected testes at key embryonic and postnatal stages to obtain a comprehensive view of the linear developmental progression. Specifically, we collected testes at E15.5 and E18.5, to examine L1-ORF1p dynamics in prospermatogonia at two critical time points: before MIWI2 nuclear localization (E15.5) and after MIWI2 enters the nucleus (E18.5), when the piRNA pathway guides DNMT3C to silence TEs (Aravin et al. 2008)(Zoch et al. 2020, 2024). We then collected testes at P5, a stage containing exclusively spermatogonia (undifferentiated and differentiated) germ cells, to represent a time point when DNA methylation is restored in WT testes (Drumond, Meistrich, and Chiarini-Garcia 2011). At P10, the first wave of meiosis begins with most cells being in spermatogonia and a few entering meiosis (Bellvé et al. 1977). By P15, most germ cells are meiotic and P20 refers to a post-pachytene-arrest in *Dnmt3C*$^{KO/KO}$ mutants. By P30 the first wave of meiosis is complete and round and elongated spermatids start to form (Bellvé et al. 1977) (Figure 2.2.2A).

We first examined L1-ORF1p expression by IF in testes cryosections across germline development (Figure 2.2.2A). In prospermatogonia E15.5, L1-ORF1p was expressed in both *Dnmt3C*$^{KO/KO}$ and WT mice, consistent with the time period before the initiation of DNMT3C-mediated *de novo* DNA methylation (Barau et al. 2016). After E18.5, when the establishment of *de novo* DNA methylation had already been initiated (Molaro et al. 2014), as anticipated, we did not detect L1-ORF1p expression in WT germ cells (Figure 2.2.1A,C). However, L1ORF1p persists in *Dnmt3C*$^{KO/KO}$ germ cells at E18.5, indicating the importance of *Dnmt3C* in L1 silencing, a phenotype also reported in *Dnmt3L* mutants (Zamudio et al. 2015). Of note, we detected L1-ORF1p expression in *Dnmt3C*$^{KO/KO}$ germ cells at P5 confirming the RNA-seq results in spermatogonia (Figure 2.2.1A,C). This expression persisted at comparable levels as in prospermatogonia (Figure 2.2.2B,C). We report L1-ORF1p translational activity in spermatogonia, in contrast to previous observations of L1-ORF1p immunostainings in *Dnmt3L*$^{KO/KO}$ (Zamudio et al. 2015; Vasiliauskaitė et al. 2018), *Mili*$^{KO/KO}$ (Di Giacomo et al. 2013) or *Miwi2*$^{KO/KO}$ (Vasiliauskaitė et al. 2018). This distinct

phenotype likely reflects the consequences of completely eliminating DNMT3C, which specifically methylates young TE promoters. While DNMT3L, MILI, and MIWI2 are upstream components of the piRNA-directed DNA methylation pathway, DNMT3C is the ultimate effector that deposits methyl marks at TEs (Barau et al. 2016). At the onset of meiosis, consistent with previous observations (Barau et al. 2016), L1-ORF1p intensified and remained elevated through P30 (Figure 2.2.2A,B,C).

Next we confirmed the IF findings by Western blot analysis in whole testes protein lysates at key developmental stages (Figure 2.2.2D). At E15.5, L1-ORF1p expression levels remained similar between WT and *Dnmt3C*$^{KO/KO}$ and the differences emerged in E18.5. At P5, L1-ORF1p showed reduced signal in the mutant compared to E18.5 (Figure 2.2.2D), confirming the quantification obtained from the IF (Figure 2.2.2C). L1-ORF1p expression increased markedly during meiosis (P15) (Figure 2.2.2D). Additionally, we tested the IAP expression by Western blot at the same stages using an IAP-POL antibody (obtained by external collaborators) and we observed a similar expression pattern as for L1-ORF1p (Figure 2.2.2D). At E15.5, IAP-POL showed similar band intensity between WT and *Dnmt3C*$^{KO/KO}$ with overall stronger signal compared to L1-ORF1p. IAP-POL levels decreased at E18.5 in both genotypes and the clear difference between WT and *Dnmt3C*$^{KO/KO}$ seen in L1-ORF1p was not observed, suggesting a compensatory IAP regulation in the absence of DNMT3C (Figure 2.2.2D). Differences between WT and *Dnmt3C*$^{KO/KO}$ in IAP-POL expression became apparent at P5 and increased in meiosis (Figure 2.2.2D), indicating that IAP silencing could occur developmentally later than LINE-1 elements. These results showing L1-ORF1p and IAP-POL upregulation in *Dnmt3C*$^{KO/KO}$ spermatogonia, while L1-ORF1p increases at meiosis, further validate our RNA-seq findings, confirming LINE-1 and ERV upregulation in a cell-and-stage-specific manner. Overall, our data suggest additional factors regulating unmethylated retrotransposon expression patterns during spermatogenesis.

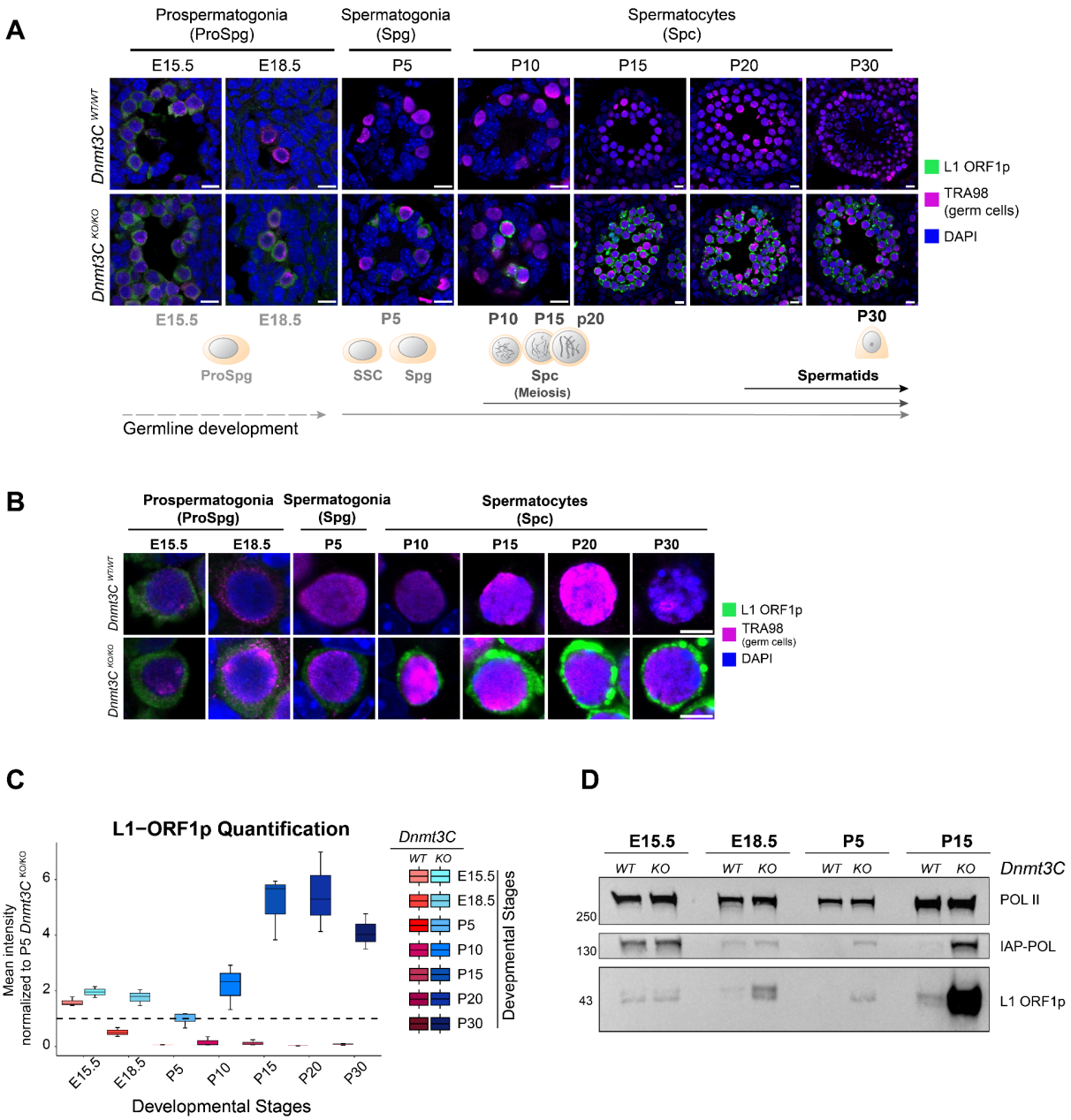


Figure 2.2.2 Young retrotransposons are translationally upregulated prior to meiosis in *Dnmt3C*$^{KO/KO}$ germ cells. A. Immunofluorescence panel of L1-ORF1 protein (green), TRA98 (germ cell marker, magenta) and DAPI (nucleus staining, blue) on cryosections of wild-type (WT) and *Dnmt3C*$^{KO/KO}$ testes at specific developmental stages. Illustration of germline development in the bottom. Scale bars, 10µm. **B.** centred on a single germ cell nucleus, representing one cell used for the quantification in C; zoom10, scale bars, 5µm. **C.** Box plot displaying mean intensity of L1-ORF1p quantification from B. **D.** Western blot analysis of LINE1-ORF1p, IAP-POL and RNA polymerase 2 (POL II) as a loading control from WT and *Dnmt3C*$^{KO/KO}$ testes at the indicated developmental time points. One testis was used per genotype and time point. The blot was developed by Pico chemistry.

2.2.2 Previous work in the lab revealed NRF1 as potential activating TF of unmethylated TEs

To understand the mechanisms driving TE expression in the absence of DNA methylation, previous work performed by Jessica Leismann sought to investigate potential transcription factors (TFs) that could activate these elements. Using a proteomic approach with P17 whole testes nuclear extracts, she performed biotin/streptavidin-based DNA pulldowns using TE sequences as baits that are highly activated in *Dnmt3C*$^{KO/KO}$. Specifically, an unmethylated and methylated IAP LTR1a promoter and L1MdT monomers as well as scrambled sequences, as controls, were utilized. Liquid chromatography-mass spectrometry (LC-MS) analysis identified several significantly enriched candidates. Among these candidates, NRF1 was enriched on unmethylated versus methylated of both IAP (Figure 2.2.3A) and LINE-1 (Figure 2.2.3B) promoters. NRF1 is an activating TF, initially reported to regulate mitochondrial biogenesis (Scarpulla 2002). Moreover, NRF1 plays an essential role in germline gene regulation, as demonstrated by *Nrf1* conditional knockout (cKO) mice showing spermatogenesis dysregulation and male infertility (J. Wang et al. 2017). Recent studies have also revealed that NRF1 promotes the development of post-migrating primordial germ cells (PGCs) (P. Wang et al. 2024). Importantly, NRF1 was previously identified as a DNA-methylation sensitive TF in mESCs (Domcke et al. 2015), consistent with our findings that NRF1 preferentially binds to unmethylated TE promoters.

We first confirmed NRF1 expression by IF in embryonic germ cells (prospermatogonia) at E15.5, postnatal pre-meiotic (spermatogonia) at P5 and meiotic germ cells (spermatocytes) at P15 (Figure 2.2.3C). Next, to validate NRF1 binding *in vivo*, CUT&Tag experiments were performed to profile NRF1 across the three key stages of spermatogenesis in both WT and *Dnmt3C*$^{KO/KO}$ sorted germ cells (Figure 2.2.3D). In E15.5 sorted prospermatogonia, prior to MIWI2 nuclear localization (Aravin et al. 2008), both WT and *Dnmt3C*$^{KO/KO}$ were considered equivalent, as the transposon genome defense mechanism has not yet been initiated (Molaro et al. 2014). At this developmental stage, NRF1 profiling revealed low occupancy at DNMT3C targets (Figure 2.2.3D). However, in sorted spermatogonia and spermatocytes a profound difference in the occupancy at DNMT3C targets was detected between WT and *Dnmt3C*$^{KO/KO}$ (Figure 2.2.3D). Examining the tracks of specific unmethylated LINE-1 (Figure 2.2.3E) and LTR (Figure 2.2.3F) copies, higher NRF1 enrichment was observed in *Dnmt3C*$^{KO/KO}$ compared to WT. For instance, one LINE-1 copy showed a higher NRF1 peak in both spermatogonia and spermatocytes *Dnmt3C*$^{KO/KO}$ (Figure 2.2.3E), while an example LTR copy revealed higher NRF1 enrichment in *Dnmt3C*$^{KO/KO}$ spermatogonia but not in *Dnmt3C*$^{KO/KO}$ spermatocytes (Figure 2.2.3F). This

work identified NRF1 as a candidate activator of evolutionarily young unmethylated TE promoters in the male germline.

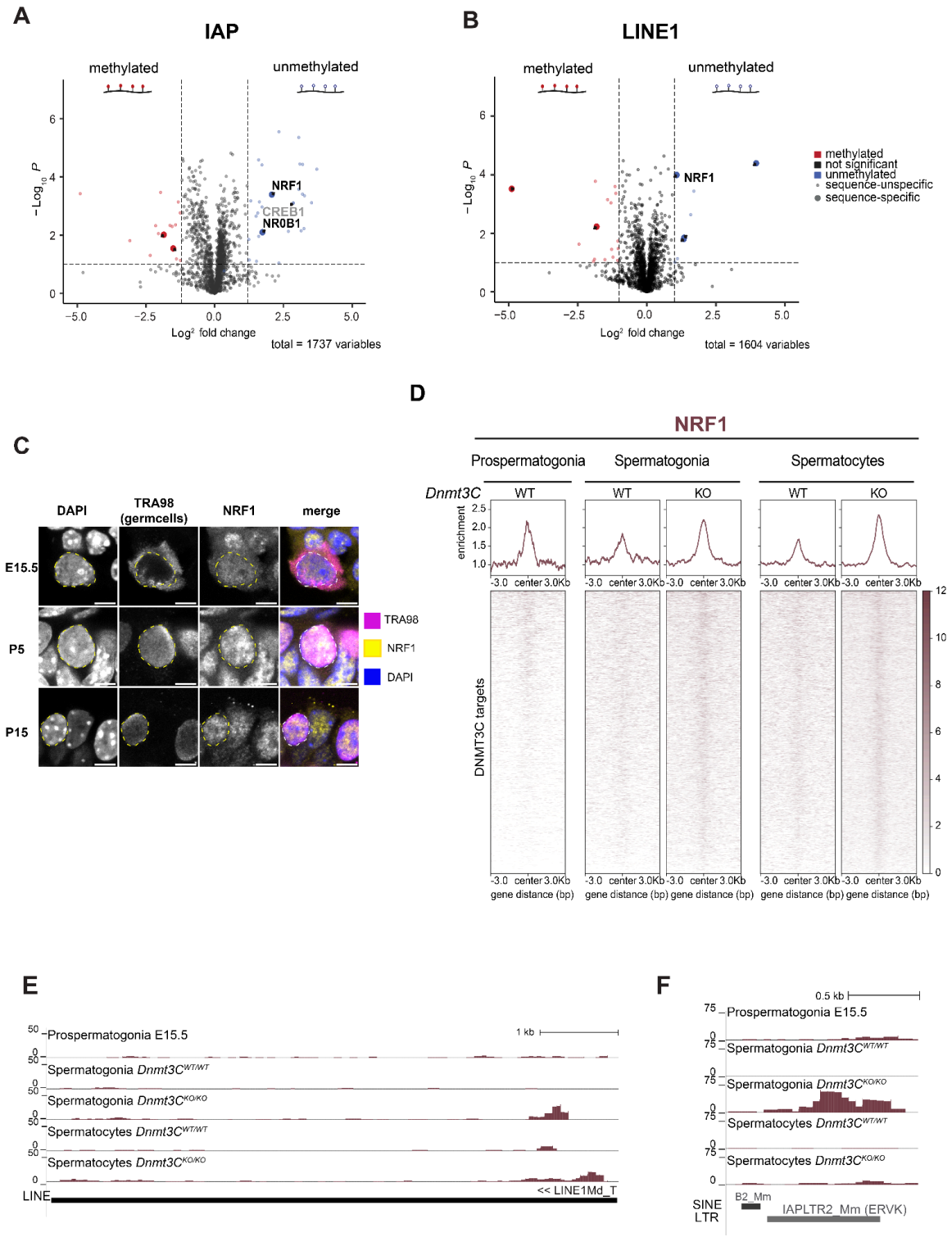


Figure 2.2.3 Methylation-sensitive TF NRF1 is a potential activator of unmethylated young TEs. **A.** Enhanced volcano plot showing -$\log_{10}$P-adjusted values and $\log_2$ fold changes of proteins enriched on methylated and unmethylated IAP LTR1a from biological triplicates detected by mass-spectrometry in DNA-pulldown experiments. Sequence-unspecific proteins that were not detected as enriched on

the scrambled controls are represented as small points, while sequence-specific proteins are highlighted by larger, filled points. **B.** as in A for methylated and unmethylated LINE1MdT. **C.** Immunofluorescence staining indicates NRF1 expression (yellow) in germ cells (TRA98, magenta). Scale bar, 5µm. **D.** Heatmaps displaying normalized coverage and metaplots showing mean enrichment of NRF1 CUT&Tag, centered on DNMT3C targets from biological duplicates of sorted wild-type embryonic germ cells at E15.5, wild-type and *Dnmt3C*$^{KO/KO}$ sorted Spg and Spc. **E.** Representative NRF1 CUT&Tag tracks in sorted germ cells at ProSpg, wild-type, and *Dnmt3C*$^{KO/KO}$ Spg and Spc at LINE1Md_T. **F.** Representative track example of NRF1 CUT&Tag as in an IAPLTR2 (ERVK).

2.2.3 The absence of NRF1 represses unmethylated IAPs

To further investigate potential effects of NRF1 on unmethylated young retrotransposon promoters specifically in mice spermatogenesis, we studied TE expression in the absence of NRF1 and *Dnmt3C*.

2.2.3.1 Generation of a mouse model with conditional *Nrf1* loss in the absence of DNA methylation

Nrf1 is an essential gene, and its disruption leads to embryonic lethality between E3.5 and E6.5 (Huo and Scarpulla 2001). Therefore, we generated a germ-cell specific conditional *Nrf1* knockout (*Nrf1*$^{cKO/KO}$) mouse model and bred them into the *Dnmt3C*$^{KO/KO}$ background. Specifically, we crossed female *Dnmt3C*$^{KO/KO}$ (Barau et al. 2016) mice carrying loxP sites flanking *Nrf1 (Nrf1*$^{fl/fl}$*) (J. Wang et al. 2017)* with male mice heterozygous for the *Nrf1* allele (*Nrf1*$^{fl/WT}$) carrying *Cre* under the *Vasa* (*Ddx4*) promoter to ensure *Cre* expression only in the germ cells (Gallardo et al. 2007) (Figure 5.1). The genotypes were confirmed by PCR reactions, by screening every animal for the floxed allele, the knockout *Nrf1* allele, the presence of *Vasa-Cre* allele, and the *Dnmt3C* knockout allele (Figure 2.2.4A). CRE-mediated recombination and subsequent allele conversion were further validated by IF. Nuclear staining for NRF1 and the germ cell marker TRA98 confirmed NRF1 expression in WT and *Dnmt3C*$^{KO/KO}$ germ cells, while showing absence of NRF1 specifically in the germ cells of *Nrf1*$^{cKO/KO}$ and double knockout (dKO, *Nrf1*$^{cKO/KO}$;*Dnmt3C*$^{KO/KO}$) mutant testes (Figure 2.2.4B). To determine the most optimal stage to study *Nrf1* loss, we analyzed NRF1 expression in different developmental stages (P5, P8 and P10). We observed extreme germ cell loss at P8 that increased at P10, consistent with Wang et al. (J. Wang et al. 2017) (Figure 2.2.4C). Quantifying the percentage of germ cells in whole testes, we detected around 50% loss at P5 and severe reduction at P8 and P10 (Figure 2.2.4D). Thus, we focused on P5, where spermatogonia and spermatogonial stem cells are the sole germ cell

populations (Figure 1.7B), and before the massive germ cell loss occurs, to characterize the effect of *Nrf1* ablation in spermatogonia. We also extended the developmental stage to 5 weeks and observed a stronger hypogonadism phenotype in the dKO compared to *Dnmt3C*$^{KO/KO}$ testes, suggesting an earlier arrest in germline development (Figure 2.2.4E).

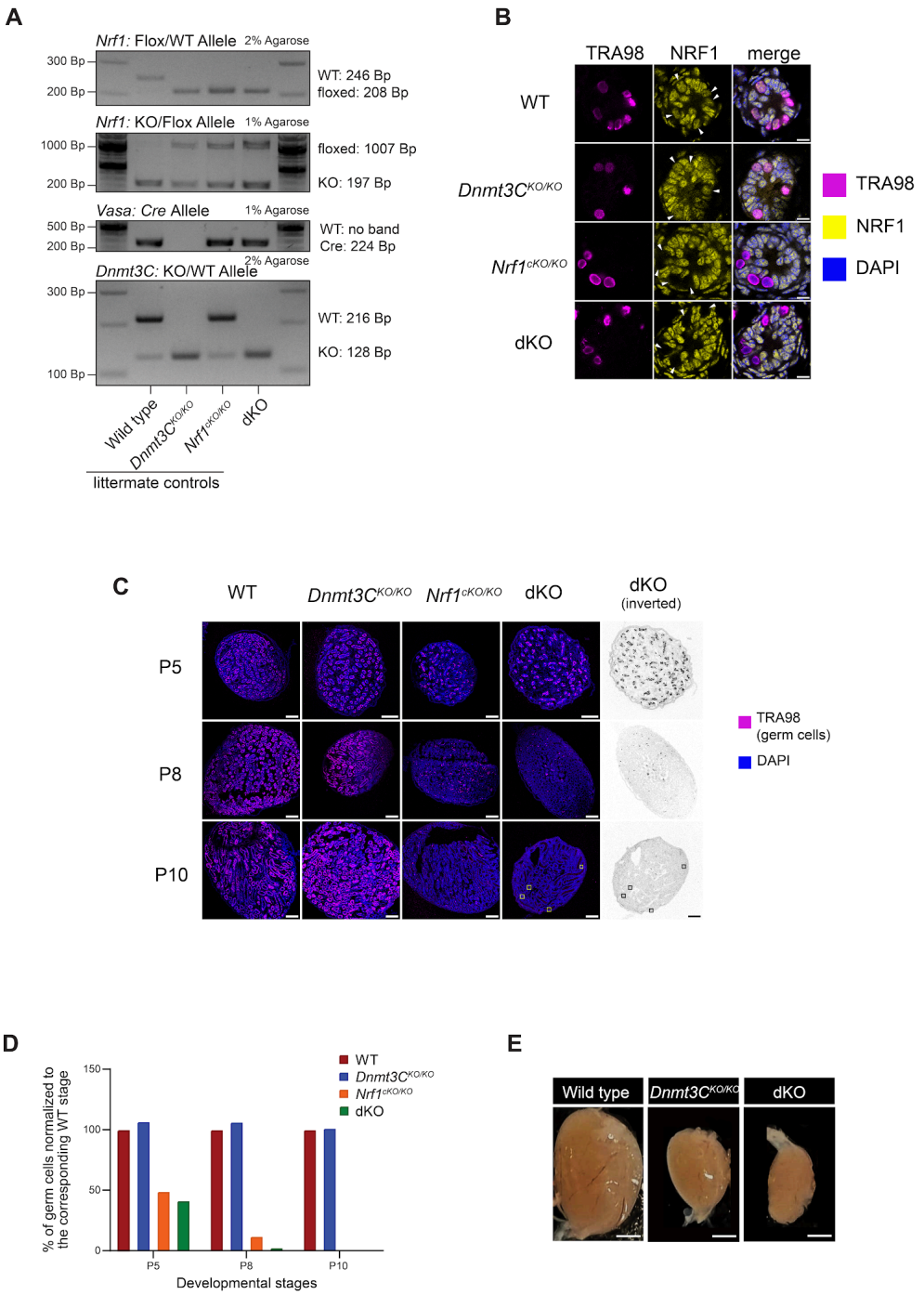


Figure 2.2.4 Generation of germline conditional *Nrf1* knockout. A. Representative agarose gel depicting genotyping PCR results for WT, *Dnmt3C*$^{KO/KO}$, *Nrf1*$^{cKO/KO}$, dKO mutants. **B.** Representative immunostaining of NRF1 (yellow), TRA98 (magenta), DAPI (blue) on P5 testes cryosections of WT, *Dnmt3C*$^{KO/KO}$, *Nrf1*$^{cKO/KO}$, dKO; arrows pointing to germ cells; scale bars, 10µm. **C.** Representative

immunostaining of whole testes cryosections at P5, P8, P10, of WT, *Dnmt3C*$^{KO/KO}$, *Nrf1*$^{cKO/KO}$, dKO mutants; TRA98 (magenta) showing germ cells; inverted images showing the contrast of germ cells (black) to the background (DAPI); scale bar, 200µm. **D.** Bar diagram showing germ cell percentage obtained from the mutants in C. TRA98/DAPI counts generated by Fiji ImageJ automated counting from one stage were normalized to TRA98/DAPI counts in the corresponding WT stage. **E.** Representative image showing hypogonadism of dKO compared to *Dnmt3C*$^{KO/KO}$ and WT testes.

2.2.3.2 NRF1 regulates IAP expression in the absence of *Dnmt3C*

To examine NRF1's role as an activator of young retrotransposons, we collected biological triplicates of WT, *Dnmt3C*$^{KO/KO}$, *Nrf1*$^{cKO/KO}$ and dKO testes at P5 and extracted RNA and protein from whole testes. All genotypes were confirmed by genotyping PCR followed by IF. To assess the transcriptional expression of young TEs we performed poly-A-enriched RNA-seq on the four different genotypes. Initial comparison of *Nrf1*$^{cKO/KO}$ versus WT revealed minimal transcriptional changes of TEs, with only IAP1-MM_I-int showing significant downregulation (Figure 2.2.5A). This limited impact on retrotransposon expression validates choosing the P5 time point for subsequent analyses despite partial germ cell loss.

Analysis of *Dnmt3C*$^{KO/KO}$ versus WT identified significant differential upregulation of young ERVK elements, particularly of IAPA_MM-int, IAPLTR1_Mm, MMERVK10C, IAPEz-int and their associated LTR elements (Figure 2.2.5B), consistent with our previous RNA-seq data in sorted spermatogonia (Figure 2.2.1A). However, LINE-1 upregulation was not detected, likely due to background RNA coming from somatic cells in whole testes, similar to previous whole-testes RNA-seq analysis of *Dnmt3L* mutants at P10, in which LINE-1 transcripts where not detected (Zamudio et al. 2015).

To investigate whether young retrotransposons are downregulated in the dKO mutant testes, we made systematic pairwise comparisons of dKO with different genotypes using *Dnmt3C*$^{KO/KO}$ versus WT as a reference. Comparison of dKO versus *Nrf1*$^{cKO/KO}$, revealed LTR elements significantly downregulated compared to *Dnmt3C*$^{KO/KO}$ versus WT (Figure 2.2.5C). Downregulation of IAP elements, such as IAPA_MM-int, IAPLTR2_Mm, IAPEz-int, and their associated LTRs, suggested the importance of *Nrf1* in activating IAP elements (Figure 2.2.5C). Interestingly, although LINE-1 upregulation was not observed in *Dnmt3C*$^{KO/KO}$ testes (Figure 2.2.5B), we detected differential upregulation of L1Md_A in the dKO testes compared to *Nrf1*$^{cKO/KO}$ (Figure 2.2.5C) and WT testes (Figure 2.2.5D). This suggests a combined effect between the absence of both *Nrf1* and *Dnmt3C*. Comparison of dKO versus WT (Figure 2.2.5D) or dKO versus *Dnmt3C*$^{KO/KO}$ (Figure 2.2.5E) revealed a further decrease in IAP element expression, as shown by the downward shifts in the volcano plots. However, these changes should be interpreted considering the potential impact of varying germ cell

numbers across genotypes. To visualize how ablation of *Nrf1* impacts the expression patterns of ERVs and L1Md_A elements, we generated a bubble chart illustrating the $\log_{10}$ p-value combined with the Log_2 Fold Change (FC) expression values obtained from the RNA-seq (Figure 2.2.5F). L1Md_A and MMERVK10C, showed indeed upregulation between *Dnmt3C*$^{KO/KO}$ versus WT and dKO versus *Nrf1*$^{cKO/KO}$ comparison, whereas RLTR10C, RLTR10B2C and other elements appeared unaffected, indicating NRF1-independent regulation of these elements (Figure 2.2.5F). Notably, we confirmed upregulation of IAP elements obtained from the differential analysis (e.g IAPA_MM-int, IAPLTR1_Mm, IAPLTR2_Mm, IAPEz-int) by visualization of the Z-score in a heatmap (Figure 2.2.5G).

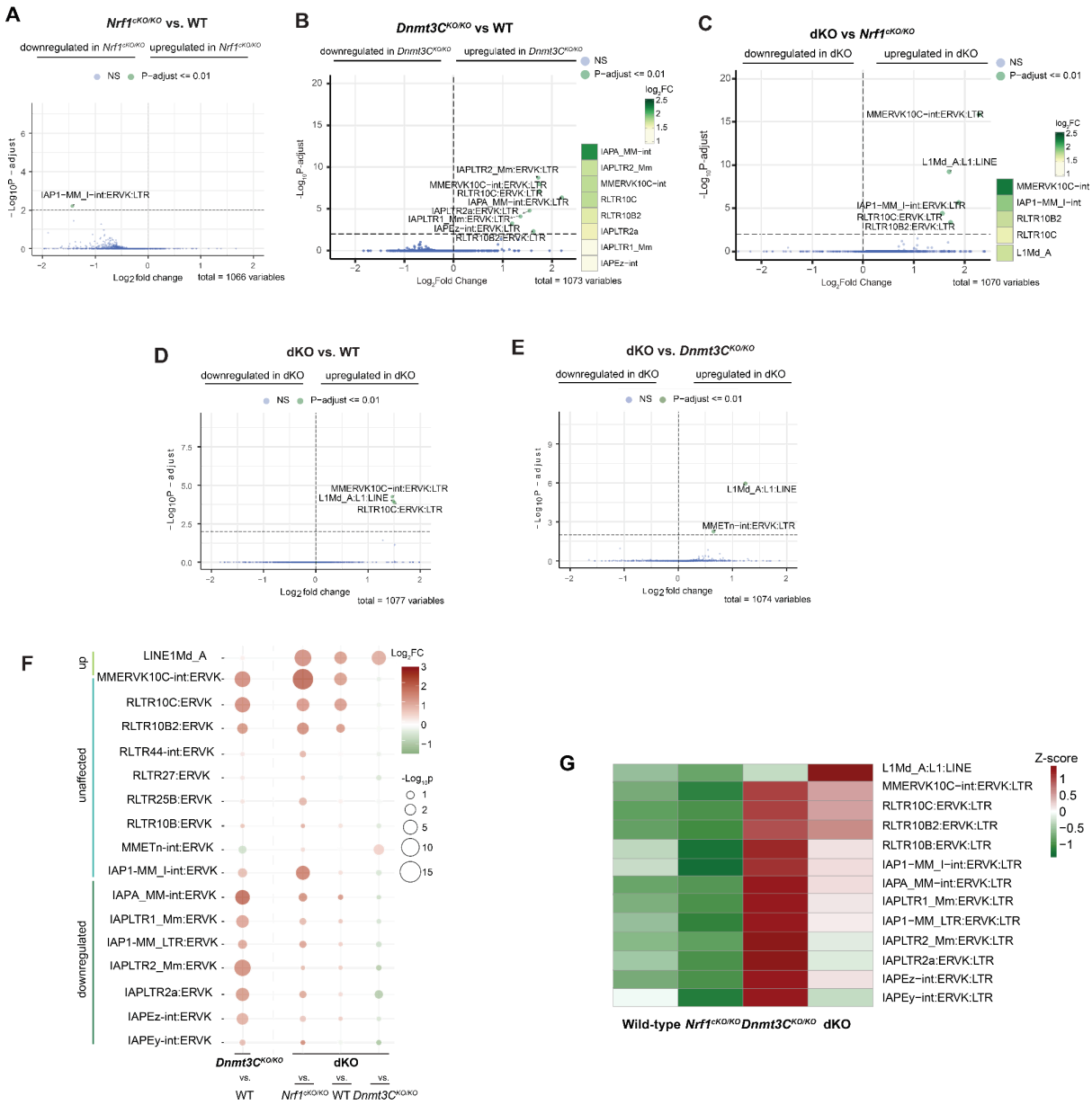


Figure 2.2.5 Transcriptional IAP downregulation in the loss of *Nrf1* and *Dnmt3C* at P5 testes. A. Enhanced Volcano plot showing -log10P-adjusted and log2 fold change values for differentially expressed TEs from biological triplicates in *Nrf1*$^{cKO/KO}$ vs. WT displayed as TEname:family:TEclass. Non-significant (NS) data points depicted in blue, significant values in green. The heatmap displays log2 fold change values for the same comparison. **B.** as A for *Dnmt3C*$^{KO/KO}$ vs WT. **C.** as A for dKO (*Nrf1*$^{cKO/KO}$; *Dnmt3C*$^{KO/KO}$) vs *Nrf1*$^{cKO/KO}$. **D.** as A for dKO vs WT. **E.** as A dKO vs *Dnmt3C*$^{KO/KO}$. **F.** Bubble chart showing log_2FC as a heatmap, with -log_{10}P-adjusted values represented as circle size, for TE subfamilies (TEname:family:TEclass) differentially enriched in pairwise mutant comparisons from biological triplicates. The pairwise comparisons were analyzed for *Dnmt3C*$^{KO/KO}$ vs WT and dKO vs *Nrf1*$^{cKO/KO}$, WT and *Dnmt3C*$^{KO/KO}$. **G.** Heatmaps showing Z-score of expression values from differentially expressed TE families in all mutants.

To monitor the translational expression of LINE-1 and IAP elements, we performed IF at P5 testes cryosections. L1-ORF1 protein (L1-ORF1p) was detected in *Dnmt3C*$^{KO/KO}$ germ cells but not in WT and *Nrf1*$^{cKO/KO}$ testes (Figure 2.2.6A), aligning with the RNA-seq data (Figure 2.2.5A,B). In the dKO mutant, similar to the transcriptome data (Figure 2.2.5) we did not observe a reduction of L1-ORF1p signal (Figure 2.2.6A). Quantification of the L1-ORF1p intensity, revealed a non-significant difference in the expression between *Dnmt3C*$^{KO/KO}$ and dKO (Figure 2.2.6B). Similar to LINE-ORF1p, IAP-POL was expressed in *Dnmt3C*$^{KO/KO}$ but undetectable in WT and *Nrf1*$^{cKO/KO}$ germ cells (Figure 2.2.6C). Notably, stainings in the dKO revealed the absence of IAP-POL expression confirming the RNA-seq results and suggesting that NRF1 is crucial for IAP expression in the absence of DNA methylation (Figure 2.2.6C).

These observations were further confirmed by whole testes protein extraction followed by Western blot analysis. At P5, L1-ORF1p levels remained comparable between dKO and *Dnmt3C*$^{KO/KO}$ samples (Figure 2.2.6D), although levels decreased in dKO by P8 (Figure 2.2.6E), likely due to progressive germ cell depletion. Strikingly, the IAP-POL protein, which was enhanced in *Dnmt3C*$^{KO/KO}$ testes, showed marked reduction in dKO samples, returning to levels similar to WT and *Nrf1*$^{cKO/KO}$ (Figure 2.2.6F,G).

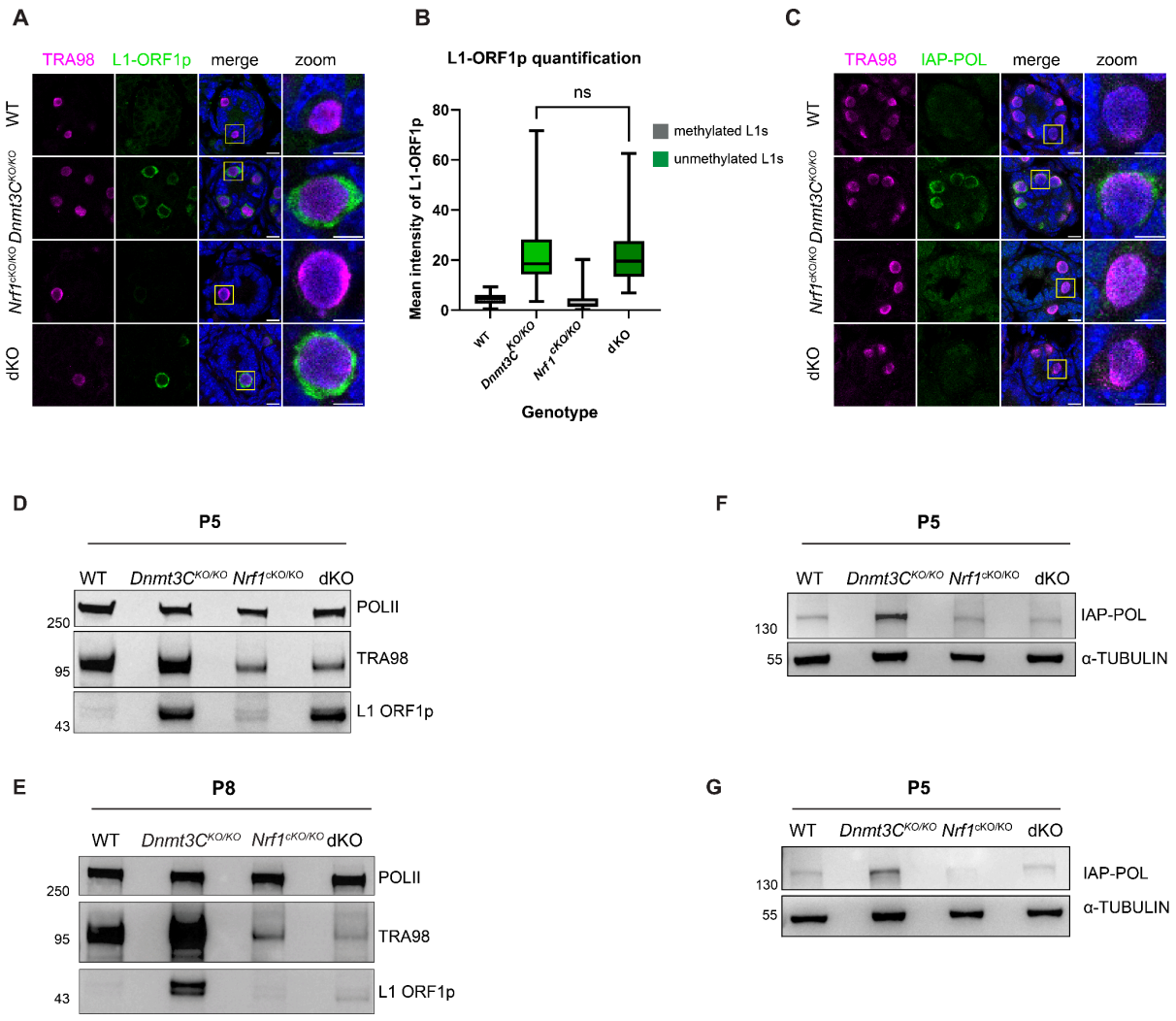


Figure 2.2.6 Translational IAP downregulation in the loss of *Nrf1* and *Dnmt3C* at P5 and P8 testes. A. Representative immunofluorescence of L1-ORF1p on WT, *Dnmt3C*$^{KO/KO}$, *Nrf1*$^{cKO/KO}$, dKO at P5 testes cryosections; TRA98 (germ cell maker); Scale bars, 10μm; zoom10, scale bar 5μm. **B.** Quantification of L1-ORF1p signal from A; WT, n=89; *Dnmt3C*$^{KO/KO}$, n= 79; *Nrf1*$^{cKO/KO}$, n=85; dKO, n= 118; t-test and graph was performed and generated in GraphPad Prism v10. **C.** as A immunofluorescence for IAP-POL. **D.** Western blot analysis of L1-ORF1p on WT, *Dnmt3C*$^{KO/KO}$, *Nrf1*$^{cKO/KO}$, dKO at P5 whole testes; TRA98 depicting germ cell loss; POL II serving as loading control; 15μg protein per sample was loaded. **E.** as C Western blot at P8 testes. **F.** as C Western blot for IAP-POL at P5 testes; α-TUBULIN serving as loading control. **G.** as F but different biological replicates; 25μg protein per sample was loaded; harsher washes were applied (5.8 Western Blot).

2.2.4 H3K27me3 restricts NRF1 binding to its target sites

Having established NRF1's importance in activating unmethylated IAPs, we next explored the potential interplay between its binding and histone modifications to elucidate the NRF1 recruitment mechanism. Previous work performed by Jessica Leismann identified a likely H3K4me3-H3K27me3 bivalent chromatin mark at unmethylated TEs in spermatogonia (Figure 3.2A). These TEs were shown to become prone for reactivation later in meiosis (data not shown). We showed that LINE-1 reactivation was increased in meiosis when DNA methylation was absent at its promoters (Figure 2.2.1C) and that NRF1 binding to DNMT3C targets was more prominent in spermatocytes compared to spermatogonia (Figure 2.2.3D). Additionally, analysis of different TE categories (Figure 2.2.5F) revealed that downregulated TEs in dKO exhibited increased NRF1 binding in *Dnmt3C*$^{KO/KO}$ spermatogonia, coinciding with reduced H3K27me3 (Figure 2.2.7A). In contrast, TEs maintaining H3K4me3-H3K27me3 bivalency showed limited NRF1 binding (Figure 2.2.7A). This pattern shifted in spermatocytes, where upregulated TEs lost H3K27me3 and gained stronger NRF1 binding, suggesting H3K27me3 might inhibit NRF1 binding (Figure 2.2.7A). Given that histone modifications can inhibit TF binding (2.5 Transcription Factors), we hypothesized that H3K27me3 could prevent NRF1 binding to its target sites.

To gain mechanistic insights into this hypothesis, we explored NRF1 binding in mESCs, which provide a more amenable system for faster genetic manipulation compared to mice models. Given that NRF1 is a DNA-methylation sensitive TF, we profiled NRF1 in DNA methylation-deficient TKO (*Dnmt1*$^{KO/KO}$, *Dnmt3a*$^{KO/KO}$, *Dnmt3b*$^{KO/KO}$) cells. NRF1 CUT&Tag-qPCR analysis of previously reported NRF1 targets (Domcke et al. 2015; J. Wang et al. 2017) confirmed stronger binding to unmethylated targets (Figure 2.2.7B).

While NRF1 binding was significantly enriched at positive control loci (e.g. *Asz1*) in TKO cells, it was minimal at TE loci despite DNA hypomethylation (Figure 2.2.7B). This suggests that DNA hypomethylation alone is insufficient to enable NRF1 binding at these specific TE targets. The lack of NRF1 enrichment may reflect either the absence of NRF1 motifs at these loci or restricted accessibility due to repressive histone marks such as H3K9me3 and H3K27me3, known to repress TEs in the absence of DNA methylation (Walter et al. 2016). To test whether H3K27me3 inhibits NRF1 binding, we treated the cells with EZH2 inhibitor (EZH2i). CUT&Tag-qPCR revealed further increased NRF1 binding at its unmethylated canonical targets but not at TE targets upon H3K27me3 depletion (Figure 2.2.7B). Genome browser tracks at representative loci of NRF1 targets (*Asz1* and *Spata22*) (J. Wang et al. 2017) demonstrated progressive NRF1 binding

enhancement, with strongest enrichment observed in TKO cells treated with EZH2i (Figure 2.2.7C).

To analyze this relationship genome-wide, we performed peak calling on NRF1 CUT&Tag data from EZH2i-treated cells (Figure 2.2.7D). For each identified NRF1 binding site, we quantified both NRF1 binding and H3K27me3 levels across all conditions (DMSO and EZH2i). We generated a heatmap where these genomic regions were arranged based on their H3K27me3 levels after EZH2i treatment, with regions showing the strongest H3K27me3 depletion placed at the top. This analysis revealed that genomic regions where H3K27me3 was effectively removed showed the highest NRF1 binding intensity. Importantly, while these regions showed some NRF1 binding in DMSO conditions, the signal intensity increased substantially following H3K27me3 removal (Figure 2.2.7D). Overall, this data suggests that the presence of H3K27me3 restricts the extent of NRF1 binding in addition to DNA methylation rather than completely preventing it in a binary way.

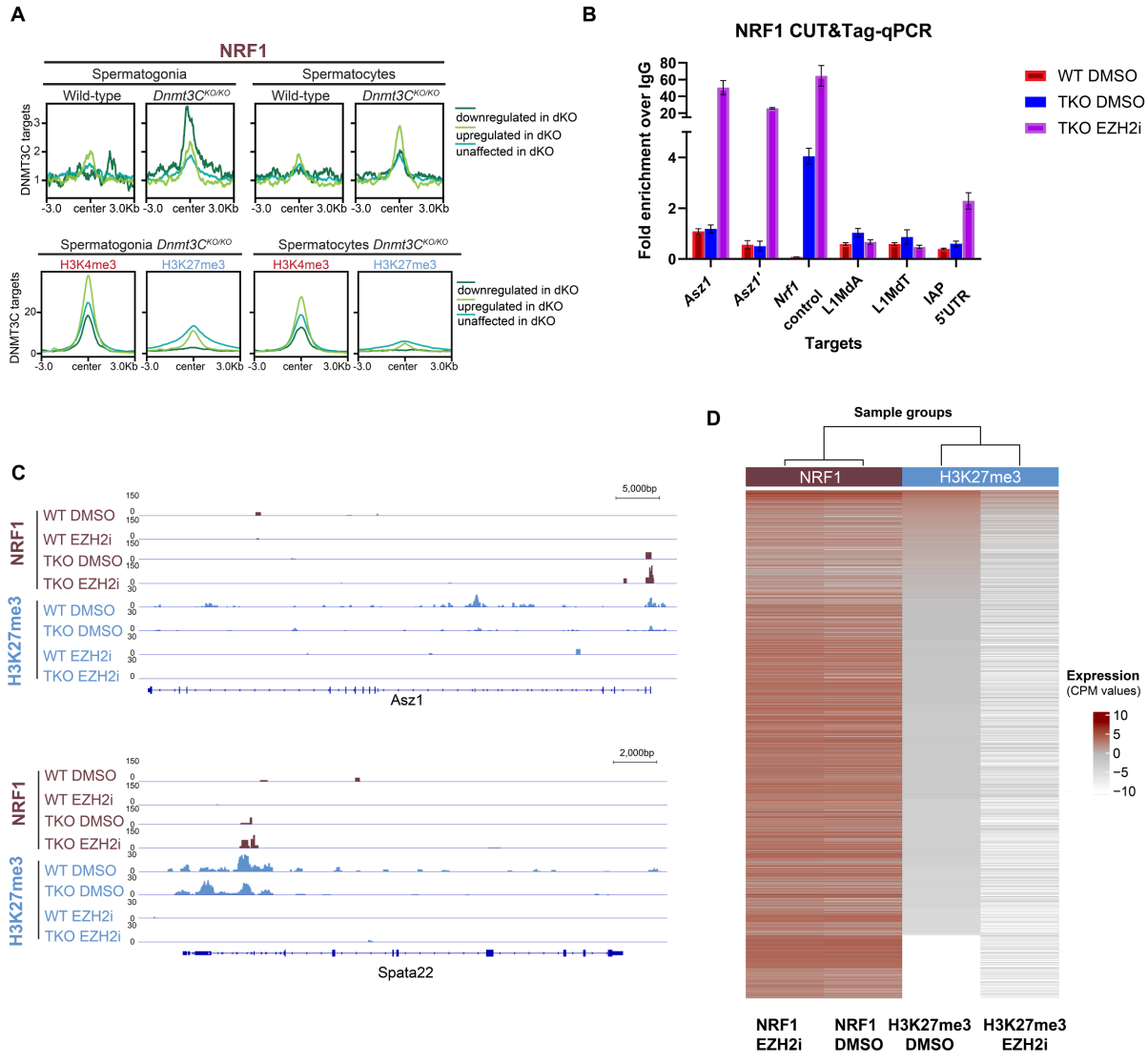


Figure 2.2.7 H3K27me3 loss enhances NRF1 binding. A. Metaplots showing enrichment of NRF1, H3K4me3, H3K27me3 CUT&Tag in at DNMT3C targets WT and *Dnmt3C*$^{KO/KO}$ in Spg and Spc. The lines represent different TE categories: downregulated (dark green), upregulated (light green), and unaffected (green) in dKO. **B.** NRF1 CUT&Tag-qPCR analysis in WT (red) and TKO mESCs treated with DMSO (blue) or EZH2i (magenta). Data shows fold enrichment over IgG control at known NRF1 targets; two different primer sets were used to target *Asz1*; The *Nrf1* control target represents a known unmethylated target that NRF1 bind to (Domcke et al. 2015) and transposable elements. Error bars represent standard deviation from three technical replicates. **C.** Representative track example of NRF1 (dark red) and H3K27me3 (blue) CUT&Tag at *Asz1* (top) (chr6:18,049,933-18,112,030) and *Spata22* (bottom) (chr11:73,324,537-73,348,434) genes across different conditions; WT=normal DNA methylation, TKO= no DNA methylation, EZH2i= no H3K27me3. **D.** Heatmap showing NRF1 binding and H3K27me3 (blue) levels at NRF1 (dark red) peak regions in TKO DMSO and EZH2i. Genomic regions are sorted by H3K27me3 levels in EZH2i condition. Color intensity represents normalized signal intensity (CPM values).

3. Discussion

3.1 The importance of an innovative endogenous readout approach for identifying TE regulators in a genetic screening

TEs have been co-opted by the genome for essential physiological processes, but they can also pose a threat to genome integrity. Despite extensive efforts to elucidate transcriptional and post-transcriptional control mechanisms, certain aspects of the TE-life cycle remain poorly understood. In particular, the translational regulation of TEs has been largely unexplored. CRISPR/Cas9-based knockout forward genetic screenings have emerged as powerful tools for discovering novel TE regulators. Previous studies have utilized reporter systems to identify TE repressors. Although successful, these approaches have an ostensible limitation: the assumption that the artificial construct used for the screen readout fully behaves as an endogenous TE. The use of artificial and truncated TE coding promoter sequences as reporters, may lead to the overlooking of *bona fide* transposon silencing factors and potentially entire novel regulatory pathways.

In this context, the objective of my PhD was to discover novel transcriptional and translational regulators of endogenous TEs. To achieve this, I sought to establish a CRISPR/Cas9-based forward genetic screen for repressors of endogenous TEs. In order to explore novel pathways, I attempted to develop a pioneering TE screen utilizing innovative readout approaches for high TE activity. This involved performing RNA-FISH to monitor mRNA abundance of endogenous TEs, and IF to assess the endogenous TE-encoded proteins. This approach aimed to provide more accurate representation of *bona fide* TEs and subsequently provide deeper insights into the mechanisms of TE silencing and their participation into broader regulatory implications. Additionally, utilizing an endogenous readout approach by performing RNA-FISH or IF provides the flexibility of identifying both transcriptional and translational regulators. Utilizing methods like PrimeFlow, which involves co-staining by fluorescent labeled antibodies and RNA FISH to simultaneously detect protein and mRNA transcripts, offers a comprehensive view of genome defense mechanisms against TEs at both transcriptional and translational levels (Fulco et al. 2019).

3.1.1 Technical challenges and future directions

3.1.1.1 The usage of CRISPR/Cas9 and other inducible editing systems in genetic screenings

As a baseline for the screen we generated a drug-inducible degron-based Cas9 cell line (Figure 2.1.1), (Figure 2.1.2). This system allows for controlled and flexible gene editing with reduced background, which is crucial for identifying essential regulators without causing unintended off-target effects. A common concern in knockout screens is the potential overlook of vital regulators. Essential genes targeted by sgRNAs, may lead to cell death, resulting in their exclusion as potential candidates. Drug-inducible CRISPR/Cas9 systems have emerged to surpass this limitation, enabling gene knockouts to be induced only when needed, reducing the risk of cell death associated with continuous disruption of essential genes. A recent study, developed a drug-inducible degron Cas9 system, achieving minimal background and tight control suitable for studying essential genes. This reinforces our approach as a potential powerful tool for similar applications (Srinivasan et al. 2024).

Inducible systems are versatile and beneficial even when not used for Cas9 induction. Groh et al. developed an alternative inducible readout approach to study the dynamics of ERV silencing under controlled conditions (Groh et al. 2021). They successfully identified *Kap1* among the top candidates, despite it being an essential gene, indicating the versatility of inducible systems even without utilizing an inducible Cas9 system.

Alternatively, CRISPRi is often a preferred sensitive method for gene knockdown. CRISPRi utilizes catalytically dead Cas9 (dCas9), to specifically bind DNA sequences without generating double-strand breaks, but instead by blocking RNA polymerase and thereby preventing gene transcription. This approach allows the discovery of essential genes without complete gene abolishment. Additionally, drug-inducible degron-based CRISPRi systems can also be utilized to overcome any potential limitations (Srinivasan et al. 2024).

Although our system appears promising, we still need to validate the activity of Cas9 within the system. A straightforward and cost-effective method to test this is the T7 assay (Sentmanat et al. 2018). After delivering Cas9 and a gRNA to an endogenous sequence, the edited locus is amplified and mixed with the same amplicon from unedited cells. The DNA is denatured and hybridized, forming heteroduplexes between an edited strand and an unedited strand. Heteroduplexes are formed when mutations, such as insertions and deletions (indels), are present in the amplicon pool. T7 Endonuclease recognizes the mismatches and cleaves the DNA, resulting in the generation of smaller fragments, thereby indicating successful editing. Upon confirmation of Cas9 activity in our system, more powerful methods such as NGS-based assays will be used to confirm the editing activity and

investigate the off-target landscape. We anticipate that our tunable system will have a higher specificity than constitutive expression of Cas9 and gRNAs, as long lingering of CRISPR in cells has been shown to decrease specificity (Shin et al. 2017). While our Cas9 approach is potentially powerful, it introduces additional layers of optimization that must be carefully managed to ensure precision without compromising project timelines.

3.1.1.2 Endogenous readout for identifying TE regulators: Limitations and troubleshooting

Our study aimed to develop an innovative readout approach to target endogenous ERVs and LINE-1 elements. We selected KAP1 as a well-characterized ERV regulator (Rowe et al. 2010) and SUV39H1/2 to target LINE-1 elements (Bulut-Karslioglu et al. 2014). Although we detected transcriptional upregulation of ERVs in KAP1-depleted cells (Figure 2.1.4), we encountered challenges when optimizing in-suspension RNA-FISH using probes generated by nick translation.

Traditional RNA-FISH methods, such as nick translation probes, have been largely superseded in high-throughput screenings by newer techniques, such as single molecule (sm)RNA-FISH (Arrigucci et al. 2017) or RNAscope (F. Wang et al. 2012; Hanley et al. 2013), which offer improved sensitivity and specificity. smRNA-FISH targets RNA single molecules with multiple fluorescent probes, increasing specificity and signal intensity. Similarly, RNAscope uses "Z" probes that require dual hybridization for signal amplification, minimizing non-specific binding and enhancing sensitivity. Alternatively, hybridization chain reaction (HCR)- FlowFISH (Reilly et al. 2021) is a recent approach for targeting mRNA abundance in high-throughput screenings. HCR-FlowFISH employs hairpin probes that amplify signals via a chain reaction, minimizing background noise. These modern methods are more costly than traditional RNA-FISH approaches but offer improved sensitivity and reduced background noise, making them more suitable for detecting mRNA transcripts, especially in high-throughput applications analyzing cells in suspension.

To target endogenous IAP-encoded proteins, we generated a custom IAP-GAG antibody. However, we encountered difficulties during validation and despite the negative results, we cannot conclusively exclude that the antibody is inefficient. The antibody efficiently detects its epitope *in vitro* (Figure 2.1.7E), but its performance in detecting endogenous IAP-GAG protein remains uncertain. Validation of custom antibodies can often be time-consuming and plagued by technical challenges. Different antibodies may require distinct permeabilization or washing conditions, and their performance can vary across different cell types, tissues, or even detection methods (e.g. Western blot vs IF). Notably,

given that ERV mRNA transcription is present in KAP1-depleted cells, we anticipated that the corresponding mRNA translation product should be expressed, too. However, to validate this assumption, we would first need to perform LC-MS to confirm the detection of IAP-derived peptides in our KAP1-degron system. Additionally, after acquiring the IAP-POL antibody from collaborators, used in the 2.2 NRF1 project (Figure 2.2.6 C,F,G), it would be interesting to revisit its efficiency in detecting IAP-encoded proteins in the KAP1-degron system. The IAP-POL antibody could also serve as a positive control to further validate the effectiveness of our custom IAP-GAG antibody.

Targeting endogenous LINE-1 elements is more challenging than ERV elements, due to their weaker promoters compared to the strong LTR promoters of ERVs. This is reflected in the RT-qPCR, in which we detected lower fold-enrichment of LINE-1 targets in the mutant cells (Figure 2.1.8 F,G) compared to ERVs targets in KAP1-depleted cells (Figure 2.1.4 A,B). Furthermore, the high baseline L1-ORF1p expression in WT mESCs presented a limitation when compared to the *Suv39h-dKO* cells (Figure 2.1.8 H-K). Besides SUV39H enzymes, LINE-1 elements are also regulated by the HUSH complex (Tchasovnikarova et al. 2015; N. Liu et al. 2018) as well as the ChAHP2 and ChAHP complexes in mESCs, which were demonstrated by a recent study, to be regulators of a broader TE repertoire including ERVs, LINEs and SINEs (Ahel et al. 2024). However, targeting these regulatory components, would require a simultaneous depletion strategy involving multiple components of the complex. This approach would not be the most efficacious, since the infectious rate of the lentiviral sgRNA library is typically one sgRNA per cell. To circumvent the labor-intensive and time-consuming process of generating multiple cell lines, we could overexpress L1-ORF1p in WT mESCs. This would determine if higher L1-ORF1p expression enables efficient detection and sorting of L1-ORF1p-positive cells compared to WT mESCs. While overexpressing L1-ORF1p could facilitate setting up a robust readout for L1-ORF1p accumulation, allowing for a more robust comparison with WT mESCs, it should be interpreted cautiously. This is necessary to avoid masking subtle changes in L1-ORF1p expression that are biologically relevant.

In conclusion, while our concept and some preliminary results were promising, multiple technical challenges hindered the progress of the project. Future genome-wide screening projects would benefit from collaborative efforts with experienced researchers and laboratories to overcome these challenges and optimize both the readout and the downstream processes. Such collaborations would be particularly valuable for starting-laboratories without prior expertise in complex screening approaches.

3.1.2 Proof-of-principle experiment

As a proof-of principle for the endogenous readout approach, it would be valuable to simulate the screening process, where only a subpopulation of the cells might exhibit altered TE expression. To achieve this, we propose spiking ~0.2% KAP1-depleted or *Suv39h*-dKO cells into a population of ~99.8% WT mESCs to mimic the conditions of a genome-wide screen (Figure 3.1). Assuming that the RNA-FISH and IF have been successfully optimized, we could then perform RNA-FISH to detect TE mRNA transcripts, and IF to target TE-encoded proteins within the mixed population. Next, FACS could be employed to sort mutant cells based on their fluorescence intensity. Then, we would extract genomic DNA (gDNA) from the isolated cells to determine whether the original genotype can be recovered. This experiment would aim to validate the efficacy of the readout approach in reliably targeting and isolating mutant cells within a heterogeneous population.

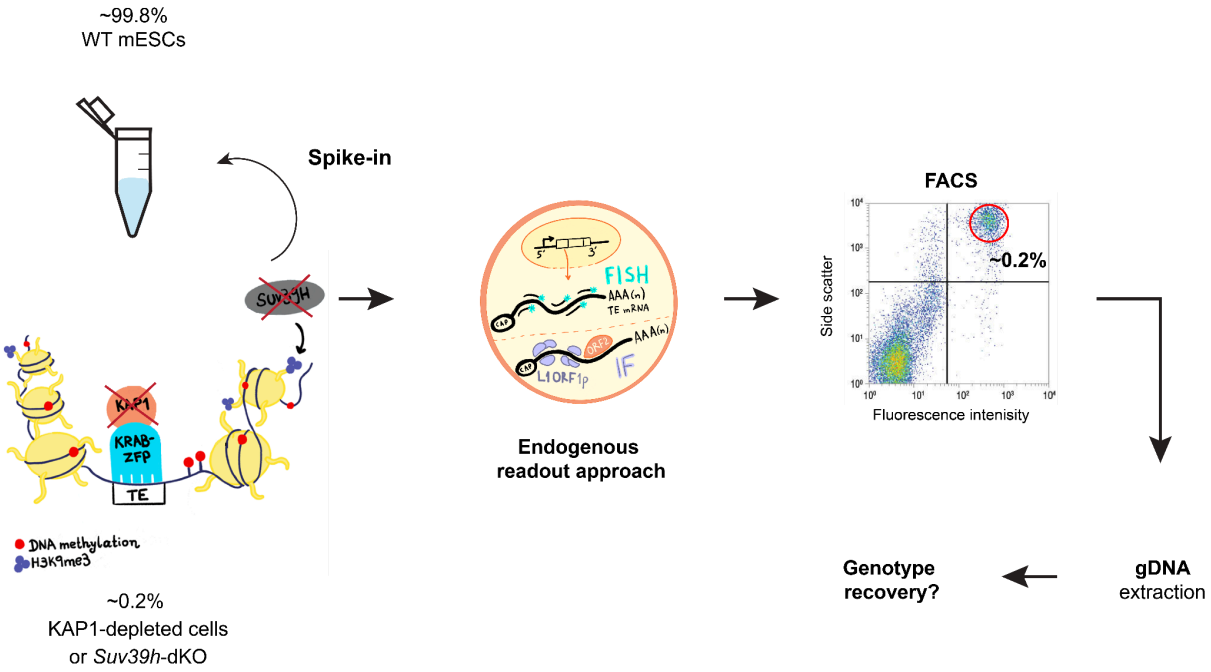


Figure 3.1 Scheme of proof-of-principal for endogenous TE readout. ~0.2% of KAP1-depleted and/or Suv39h-dKO cells spiked-in ~99.8% WT mESCs; TE mRNA and proteins targeting via RNA-FISH and IF, respectively; Isolation of positive populations FACS; gDNA extraction and genotype recovery by PCR amplification.

3.1.3 Alternative approaches to identify TE translational regulators

How TE expression can integrate as part of normal development, in an independent-damaging manner remains an active area of current research. Moreover, TE translational controlling mechanisms have remained an enigmatic part of their life cycle and the aim of my PhD project was to uncover such pathways. But how can we identify novel TE translational regulators? We chose the challenging path of optimizing an endogenous readout approach for a genome-wide CRISPR/Cas9 forward genetic screening. However, there are alternative screening approaches to address this question.

Despite the challenge to enrich TE ribonucleoparticles biochemically, immunoprecipitation (IPs) has served as a primary method to identify TE-encoded protein interactions by mass spectrometry (De Luca, Gupta, and Bortvin 2023; Ardeljan et al. 2020). Recently, proximity biotinylation has arisen as robust interaction proteomics assay that can *in situ* monitor direct and indirect interactions in the proximal proteome of a target protein of interest (Santos-Barriopedro, van Mierlo, and Vermeulen 2021). Proximity biotinylation has never been previously attempted in this context, and the *in situ* format of the assay can prove valuable in overcoming the problems associated with the biochemical enrichment of TE protein particles. This methods could be applied to various biological samples known to highly express L1-ORF1p, including *Dnmt3C*$^{KO/KO}$ testis (Barau et al. 2016), neural progenitor cells (NPCs) (Coufal et al. 2009) and a big repertoire of cancer cell lines, such as human breast adenocarcinoma (Belgnaoui et al. 2006), ovarian and testis embryonal carcinoma (Garcia-Perez et al. 2010). This approach could contribute significantly to our understanding of TE-protein biology and potentially lead to the development of novel cancer biomarkers.

Additionally, an alternative way to shed light on the mechanisms regulating the dynamics of TE-encoded proteins, could be through the use of a tandem fluorescent protein timer (tFT) designed to quantitatively assess protein stability (Khmelinskii et al. 2012). By applying this timer approach on TE-encoded proteins as a readout of a genetic screen, we could potentially identify regulators of protein stability. These regulators could act as host factors either hijacked by TEs to enhance folding and stability or as repressors that destabilize TE proteins. Although this method still utilizes an artificial readout, it offers quantitative insights into protein stability. This nuanced view of protein dynamics, provides an improved readout over traditional reporter systems. By employing these diverse screening approaches, we anticipate unveiling novel regulatory pathways and gaining deeper insights into the mechanisms of TE regulation and their participation into genome regulation.

3.2 NRF1 as a potential TE regulator in spermatogenesis

In contrast to the CRISPR/Cas9 genetic screening approach described in Section 3.1, which aimed to identify novel TE regulators in mESCs, this section focuses on a complementary strategy: the investigation and "screening" of specific factors influencing TE expression in mice. While the endogenous readout approach faced technical challenges in high-throughput applications, it highlighted the importance of understanding TE regulation at both the transcriptional and translational levels. To further investigate the complex regulation of TEs, particularly in the context of DNA methylation, we turned to the *Dnmt3C*$^{KO/KO}$ mouse model. This model allowed us to examine the *in vivo* dynamics of unmethylated TEs and to identify factors involved in their regulation during spermatogenesis.

Through RNA-seq on sorted pre-meiotic and meiotic cells (Figure 2.2.1), immunostainings and Western blot analysis (Figure 2.2.2), we found that although LINE-1s and ERVs remain active throughout spermatogenesis in the absence of DNA methylation, LINE-1s increase in meiosis. DNA pulldown assays and CUT&Tag experiments identified NRF1 as a potential activator of unmethylated TEs in germ cells. Notably, IAP silencing in pre-meiotic cells was rescued when depleting *Nrf1* in the *Dnmt3C*$^{KO/KO}$ background. In the following sections I will discuss the role of NRF1 in regulating IAPs and potentially LINE-1s and factors mediating or hindering NRF1's binding to its targets and subsequently to TEs.

3.2.1 L1s and ERVs exhibit distinct expression patterns across spermatogenesis

DNA methylation is a key regulatory mechanism for TEs in mammals (Yoder, Walsh, and Bestor 1997) and therefore utilizing hypomethylated systems is essential for understanding TE regulation. While mESCs are viable in the absence of DNA methylation (Tsumura et al. 2006), loss of *Dnmts* are lethal in mice (E. Li, Bestor, and Jaenisch 1992; Okano et al. 1999). *Dnmt3C* mouse mutants represent the perfect model to study TE upregulation upon hypomethylation *in vivo*, as DNMT3C targets only the promoters of young TEs to silence them (Barau et al. 2016). We investigated TE expression dynamics in the *Dnmt3C*$^{KO/KO}$ mouse model and observed distinct patterns of IAP and LINE-1 elements in embryonic and postnatal stages. At embryonic stages, our western blot analysis indicated a potential delay of IAP silencing by *de novo* methylation. Specifically, IAPs showed upregulation for the first time at P5 in *Dnmt3C*$^{KO/KO}$ compared to WT testes, implicating a

potential compensatory epigenetic switch of IAP regulation in the absence of DNMT3C (Figure 2.2.2D). Conversely, LINE-1s responded earlier to the loss of *Dnmt3C*, as their upregulation was observed at E18.5 (Figure 2.2.2D). At postnatal stages, ERVs remain expressed throughout both pre-meiotic and meiotic stages, whereas LINE-1s exhibited a significant increase from pre-meiotic to meiotic cells (Figure 2.2.1),(Figure 2.2.2). These results revealed a novel role of DNA methylation on LINE-1 regulation in pre-meiotic cells, that has not previously been reported in similar mutants (Zamudio et al. 2015; Vasiliauskaitė et al. 2018).

This varied TE expression between IAPs and LINE-1s during hypomethylation may correlate with their distinct chromatin landscape. In embryonic male germ cells, LINE-1s obtain a H3K4me9-K9me3 bivalent mark, facilitating piRNA-directed DNA methylation on LINE-1 elements (Dias Mirandela et al. 2024), while IAPs rely on SETDB1-dependent H3K9me3 (Dias Mirandela et al. 2024; S. Liu et al. 2014). In hypomethylated embryonic oocytes, TEs gain H3K27me3, while some LINE-1 copies (L1-Gf) retain repression through dual H3K9me3/H3K27me3 enrichment (Huang et al. 2021). In postnatal stages, ChIP-qPCR in *Dnmt3L*$^{KO/KO}$ demonstrated that IAPs and LINE-1s are marked by both H3K9me2 and H3K9me3 before meiosis, whereas at meiosis they lose both marks and gain H3K4me3. Strikingly, L1-Gf resists activation in *Dnmt3L* mutants despite losing H3K9me2/H3K9me3 during meiosis. (Zamudio et al. 2015). Interestingly, L1-Gf responded to the loss of SUV39H1/2-dependent H3K9me3 in our *Suv39h*-dKO cell line during normal DNA methylation levels (Figure 2.1.8G). Hypomethylated mESCs further underscore this divergence: IAPs remain exclusively silenced by H3K9me3, whereas LINE-1s recruit compensatory H3K9me3 and H3K27me3 (Walter et al. 2016).

Why different TE families exhibit distinct epigenetic landscapes upon DNA methylation loss remains an open question. Future studies could perform CUT&Tag with tagmentation-based bisulfite sequencing (CUT&Tag-BS) to simultaneously monitor the epigenetic dynamics upon loss of DNA methylation at key developmental time windows [(e.g from early embryonic (E13.5) to early postnatal stages (P1)] (R. Li, Grimm, and Wade 2021). Additionally, long-read NGS-based methods can be particularly beneficial for TE sequencing (Gerdes et al. 2023). Comprehensive epigenetic profiling by the recently developed scNanoseq-CUT&Tag, which combines CUT&Tag with Nanopore sequencing (Q. Li et al. 2024), could elucidate the nuanced regulation of unmethylated TE families or subfamilies by histone modifications.

3.2.2 NRF1 could regulate IAPs before meiosis but LINE-1s at meiosis

Another approach to explore and screen for TE regulators affecting the varied TE expression in germ cells, is by performing DNA pulldown assays. Our pulldown experiment identified NRF1 as a DNA methylation-sensitive TF that could regulate unmethylated TEs in postnatal male germ cells (Figure 2.2.3 A,B). Combining CUT&Tag data of NRF1 profiling, we confirmed its binding to DNMT3C targets in sorted pre-meiotic cells and reported an enhanced binding in meiotic cells (Figure 2.2.3 C). Ablation of *Nrf1* in *Dnmt3C*$^{KO/KO}$ (dKO) germ cells showed that IAPs were transcriptionally downregulated (Figure 2.2.5). Moreover, IF analysis indicated a rescue of the IAP-POL protein silencing in the dKO (Figure 2.2.6C). Additionally, we found that the downregulated IAPs in dKO exhibited a stronger NRF1 binding in pre-meiotic compared to meiotic cells (Figure 2.2.7A). This data indicates that NRF1 activates IAPs before meiosis.

Interestingly, while IAPs were downregulated in dKO testes, we observed increased L1Md_A transcriptional levels (Figure 2.2.5E) but no reduction in L1-ORF1 protein levels (Figure 2.2.6B). NRF1 is an ubiquitous TF and its disruption may lead to dysregulation of genes beyond germline genes. For instance, NRF1 is known to regulate EHMT1/GLP (Palmer et al. 2019), a histone methyltransferase that together with EHMT2/G9a catalyze H3K9me2, known to silence LINE-1 elements in spermatogonia (Di Giacomo et al. 2013; Zamudio et al. 2015). We cannot exclude that L1Md_A transcriptional reactivation is an indirect effect caused by EHMT1 dysregulation. To provide insights into this hypothesis, we could measure the levels of H3K9me2 and further confirm NRF1 binding to unmethylated LINE-1 promoters *in vitro*, by a fluorescent polarization assay.

The stronger NRF1 binding that we observed in meiosis (Figure 2.2.3D) coinciding with the high expression of LINE-1s at meiosis (Figure 2.2.1C) (Figure 2.2.2), suggests that LINE-1 transcriptional activation might depend on NRF1 at the onset of meiosis. However, testing this effect in the conditional mutants was complicated due to the extensive spermatogenic arrest and germ cell loss occurring before meiosis when using *Vasa-Cre* driver for conditional *Nrf1* depletion (Figure 2.2.4 C,D,E). Alternatively, similarly to the *Vasa-Cre* crosses (Figure 5.1) we generated dKO mutants in meiotic cells using a *Stra8-Cre* driver (data not shown), which created knockout allele conversion at the onset of meiosis (Sadate-Ngatchou et al. 2008). Phenotypic validation of these mutants would need to carefully determine the stage at which meiotic arrest occurs. This involves performing meiotic spreads using a meiotic-specific marker like SYCP3, followed by counting the percentage of cells in each meiotic stage (leptotene, zygotene, pachytene, diplotene). Next, co-staining for NRF1 and L1-ORF1p on meiotic spreads would be necessary to confirm

knockout allele conversion and assess L1-ORF1p expression at the same meiotic stage. Furthermore, any potential downregulation of L1-ORF1p should be quantified and compared to the levels in *Dnmt3C*$^{KO/KO}$ at the same meiotic stage. Finally, given that testes at meiotic stages consist of both meiotic and pre-meiotic cells (Figure 1.7B), monitoring whether LINE-1 silencing could be rescued transcriptionally would necessitate sorting meiotic cells. However, this approach would require significant time and resources, involving numerous mice, which was beyond the scope of this project. While these results provide valuable insights into NRF1's role in TE regulation, further studies are needed to fully elucidate the mechanisms underlying its stage-specific effects on different TE families.

3.2.3 How does NRF1 regulate TEs?

In our study we identified NRF1 as a potential regulator of TEs, particularly IAPs, before meiosis. NRF1 regulation involves structural features of NRF1 itself, complex interactions between DNA methylation or an interplay with histone modifications. In the following sections, I discuss how these factors influence NRF1's binding dynamics and how we could further explore its regulatory mechanisms. Understanding these multifaceted regulatory mechanisms will be crucial for developing a comprehensive model of TE regulation during spermatogenesis.

3.2.3.1 The role of structural features and binding patterns of NRF1

NRF1's ability to regulate TEs might be closely linked to its structural domains. A crystal structural study revealed that NRF1's DNA-binding domain (DBD) recognizes TGCGC motifs, but its dimerization domain (DD) stabilizes the interaction by binding the full consensus sequence (GCGCATGCGC) (K. Liu et al. 2024). While the DBD can bind DNA as a monomer, transcriptional activation requires the intact homodimer to stabilize binding and recruit transcriptional machinery. This is important to note, as NRF1 binding to TEs per se does not necessarily mean regulation. We could leverage our chromating profiling data to further analyze NRF1 binding motifs and determine whether NRF1 preferentially binds full consensus sequences or half-sites at TE promoters.

Moreover, a recent study revealed NRF1 can act either as an activator or a repressor depending on its binding position relative to the transcriptional start site (TSS) (Duttke et al. 2024). Specifically, NRF1 binding upstream of the TSS promotes transcription activation, while downstream binding promotes repression by blocking RNA polymerase II activity. Interestingly, we observed positional shifts in NRF1 binding at some representative TE loci (Figure 2.2.3E). However, further motif analysis is required to precisely determine where

NRF1 binds and at which TE promoters. While NRF1's structural domains and positional binding patterns are critical for its regulatory function, these mechanisms do not operate in isolation but they are further modulated by epigenetic factors, such as DNA methylation, which play a pivotal role in determining NRF1's binding ability.

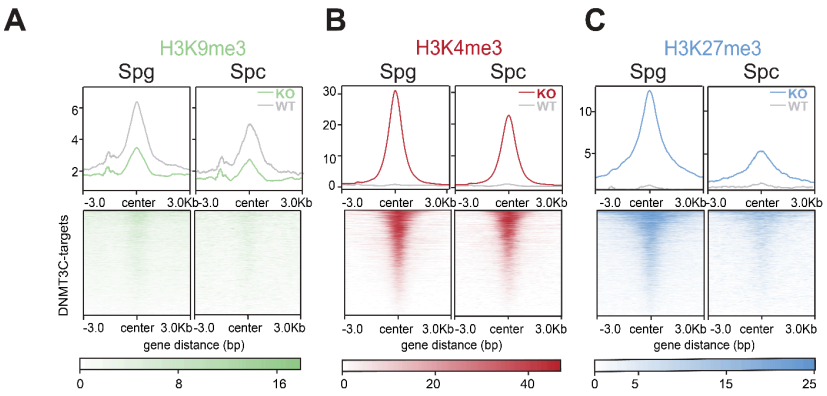
3.2.3.1 The role of DNA methylation in NRF1 binding

DNA methylation is a well-established mechanism for silencing TEs, and it directly influences the binding capacity of TFs like NRF1 (Domcke et al. 2015), adding another layer of complexity to its regulation of TEs. An open question in the field is to understand how DNA methylation silences TEs. As previously discussed (1.2.5), DNA methylation-sensitive TFs have been implicated in TE silencing. But how do TFs regulate TEs? TFs' binding could be directly inhibited by the presence of DNA methylation itself, or indirectly by methyl-binding proteins (MBD) (Dantas Machado et al. 2015; Nan et al. 1998). A recent study demonstrated that DNA methylation directly inhibits the binding of the methylation-sensitive TF, CREB1, to unmethylated IAPs in neuronal cell culture models (Kaluscha et al. 2022). Specifically, they reported that the absence of the four functional MBDs (MBD1, MBD2, MBD4 and MeCP2) was not sufficient for TE upregulation either alone or in combination, excluding that these MBDs indirectly mediate DNA methylation silencing at TEs. Interestingly, we also identified CREB1 in our *in vitro* unmethylated IAP pulldown, but it was not significantly enriched (Figure 2.2.3A). In the same study they also demonstrated that in mESCs loss of MBDs did not lead to the same level of chromatin accessibility changes compared to the loss of DNA methylation, particularly in regions containing NRF1 motif. This suggests that MBDs are not the primary factors maintaining the inaccessibility of these regions (Kaluscha et al. 2022). In agreement with this study, we did not identify significantly enriched MBD proteins in our pulldown experiments using *in vitro* methylated IAP promoter, indicating a direct inhibition of DNA methylation (data not shown). However, it is worth noting that the mechanism by which methylated TEs are silenced may vary between germ cells and somatic cells, with the latter relying on the scaffolding protein KAP1 for repression (Rowe et al. 2010; Boulard et al. 2020). While DNA methylation directly inhibits the binding of TFs like NRF1, it might often interplay with histone modifications to further shape the epigenetic landscape and accessibility of TE promoters.

3.2.3.2 Interplay between DNA methylation and histone modifications

The relationship between DNA methylation and histone modifications is a critical factor in TE regulation, as these marks often co-occur or influence one another to establish repressive or permissive chromatin states (1.2.3). Jessica Leismann had previously

conducted a chromatin profiling in pre-meiotic (Spg) and meiotic (Spc) cells to explore the role of histone modifications in TE expression throughout spermatogenesis (Leismann and Kanta et al., Embo Reports - paper in revision) (Figure 3.2A,B,C). Specifically, Jessica showed that H3K9me3 was present in methylated TEs, but lost in unmethylated TEs (Figure 3.2A). This aligns with previous studies that report co-occurrence of DNA methylation and H3K9me3 on IAP elements in PGCs (S. Liu et al. 2014) and mESCs (S. Liu et al. 2014; Ficz et al. 2013; Habibi et al. 2013). Additionally, Jessica found that the activating H3K4me3 mark was present in unmethylated TEs (Figure 3.2B), which aligns with their high transcriptional (Figure 2.2.1) and translational expression (Figure 2.2.2). We suggest that NRF1 binds to unmethylated IAPs in pre-meiotic cells in the presence of H3K4me3 but in the absence of DNA methylation and H3K9me3 (Figure 3.2D). Thus, NRF1 may regulate IAPs through a combination of DNA methylation and histone modifications.



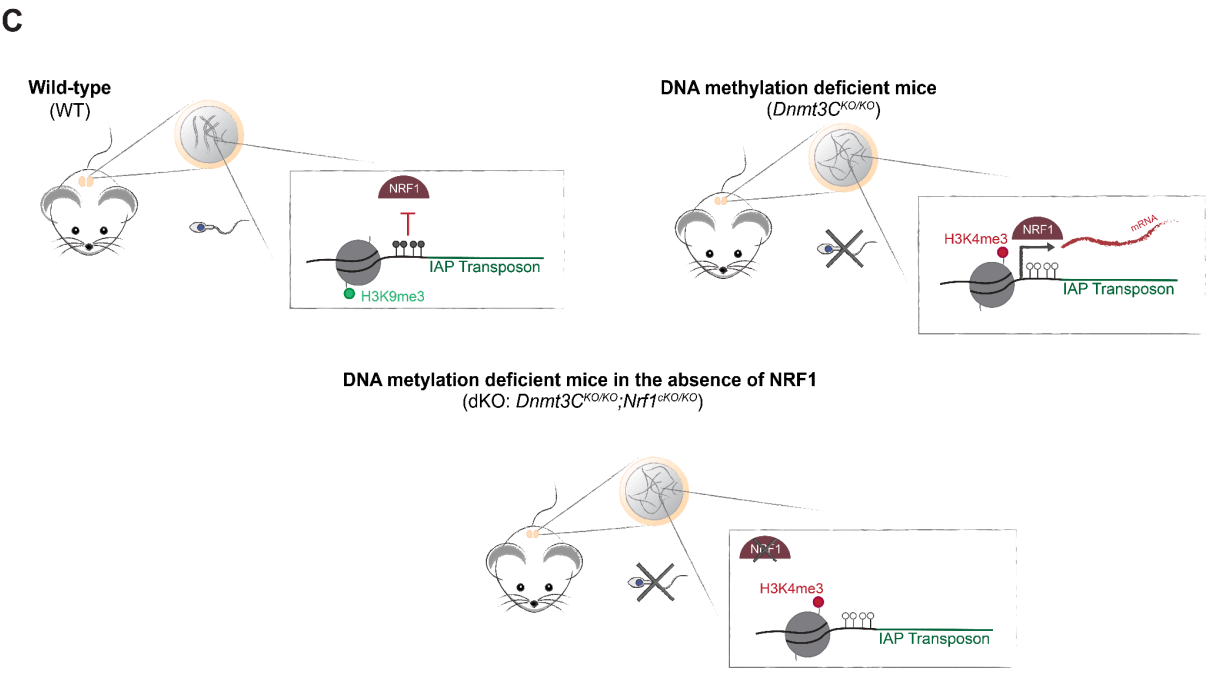


Figure 3.2. The role of DNA methylation and histone marks in silencing TEs. A,B,C. Chromatin landscape of TEs in pre-meiotic (Spg) and meiotic cells (Spc); heatmaps displaying normalized coverage and metaplots showing mean enrichment of H3K9me3 **(A.)**, H3K4me3 **(B.)** and H3K27me3

(C.) CUT&Tag from biological duplicates, centered on DNMT3C targets in sorted *Dnmt3C*$^{KO/KO}$ pre-meiotic/spermatogonia (Spg) and meiotic/spermatocytes (Spc); WT is illustrated as a gray line in the same metaplot. **D.** Proposed model depicted how NRF1 binds to IAPs in WT, *Dnmt3C*$^{KO/KO}$ and dKO spermatogonia cells.

Jessica also reported that besides H3K4me3, unmethylated TEs gained the repressive H3K27me3 mark in pre-meiotic cells, which was reduced in meiosis (Figure 3.2C). This data suggests a likely bivalent H3K27me3-H3K4me3 mark on TEs, which is commonly observed in germline genes (Mohn et al. 2008; Sachs et al. 2013). Interestingly, downregulated IAPs in dKO testes gained H3K27me3 in pre-meiotic cells, which was subsequently reduced during meiosis, coinciding with stronger NRF1 binding (Figure 2.2.7A). CUT&Tag-qPCR and representative track examples further showed increased NRF1 binding to some targets in the absence of DNA methylation and H3K27me3 in mESCs (Figure 2.2.7B,C). Further CUT&Tag analysis revealed that some genomic regions exhibited stronger NRF1 binding after losing H3K27me3 (Figure 2.2.7D). While these preliminary observations are promising, further validation and detailed analysis will be needed to fully understand the relationship between H3K27me3 removal and NRF1 binding patterns.

CUT&Tag-qPCR did not show substantial enrichment of NRF1 on some TE targets tested (L1MdA, L1MdT), although some enrichment was detected on the IAP 5'UTR (LTR) (Figure 2.2.7B). This pattern could potentially be explained by the presence of H3K9me3 on these regions in hypomethylated conditions. As detailed by Walter et al., H3K9me3 primarily marks IAPez and the 5' UTRs of LINE-1s, while the 3' UTR of LINE-1 is marked by H3K27me3 in hypomethylated mESCs (Walter et al. 2016). The general lack of strong NRF1 enrichment detected by CUT&Tag-qPCR on these targets might suggest that H3K9me3 further hinders efficient NRF1 binding to unmethylated TEs in TKO ES cells. To further explore this hypothesis, we could leverage the *Suv39h*-dKO-KAP1-degron cell lines previously generated (2.1.2.1) (2.1.2.4). In these cell lines, where loss of SUV39H1/2 and KAP1 disrupts H3K9me3 occupancy on TEs (Bulut-Karslioglu et al. 2014; Stoll et al. 2022), we could further reduce H3K27me3 levels (via EZH2i treatment) and induce hypomethylation conditions. This would allow us to address whether NRF1 binding increases on unmethylated TEs when both major repressive marks (H3K9me3 and H3K27me3) are reduced, providing a clearer picture of NRF1's potential binding capabilities in a permissive chromatin environment.

Moreover, to directly explore how specific chromatin states affect NRF1 binding, we could use epigenome editing to deposit or remove specific histone marks at canonical NRF1 binding sites and measure NRF1 occupancy by CUT&Tag (Policarpi et al. 2022). Alternatively, the dynamics of NRF1 binding could be investigated *in vitro* by assembling

nucleosomes with specific histone modifications and performing antibody-based pulldown assays. Elucidating this direct NRF1-chromatin interplay will be crucial for a comprehensive understanding of TE regulation during spermatogenesis.

3.2.4 NRF1: from ancestral functions to TE regulation

The evolution of NRF1 as a methyl-sensitive TF and its role in TE regulation reflect a fascinating interplay between host defense mechanisms and TE survival strategies. Interestingly, NRF1 is conserved across diverse species, including those with hypermethylated genomes, such as the marine sponge *Amphimedon queenslandica*, where NRF1 preferentially binds to unmethylated CpG-rich promoters (de Mendoza et al. 2019). *Nrf1* is also conserved in species lacking DNA methylation, such as *Drosophila melanogaster*, where the *Nrf1* homolog, erect wing (*ewg*), regulates muscle development independently of methylation (Schaefer, Engman, and Miller 2000). This conservation pattern suggests that NRF1's ancestral function might not have been inherently tied to DNA methylation. De Mendoza et al. propose that NRF1 could have likely convergently evolved in organisms with robust DNA methylation systems (de Mendoza et al. 2019).

Crucially, TEs require host transcriptional activators in the germline for vertical transmission (Bao and Yan 2012; Maupetit-Mehouas and Vaury 2020). While hosts cannot eliminate TEs embedded in their genomes, they evolve mechanisms to suppress harmful activation. However, essential factors like NRF1 cannot be discarded entirely, as their loss disrupts essential processes and could lead to sterility (J. Wang et al. 2017). Instead, DNMT3C's presence and NRF1's methylation sensitivity might act as a safeguard: by avoiding methylated DNA, NRF1 could regulate active genes while ignoring methylated TEs, balancing host survival and TE repression (T. H. Bestor 2003).

NRF1's interaction with IAPs might be lineage-specific, as analogous elements are absent in species like *Drosophila melanogaster* (Keegan et al. 2021). The temporal mismatch between NRF1's ancient conservation (de Mendoza et al. 2019) and IAPs' recent integration (Stocking and Kozak 2008) supports that NRF1's ancestral function was probably unrelated to TE regulation and that its role in controlling IAPs might represent a derived adaptation specific to rodents (Chuong, Elde, and Feschotte 2017). Future research could explore how TFs, like NRF1, evolved methylation sensitivity across species with varying levels of DNA methylation and could provide valuable insights into the evolutionary dynamics of TF-TE interactions and their implications for genome regulation.

3.2.5 The importance of identifying TFs

Our study on NRF1 and its role in regulating TEs opens up new avenues for understanding complex regulatory networks governing TE expression in mammalian genomes. While our study focused on NRF1, it's important to note that TEs likely employ various other TFs for their regulation. Our pulldown experiments identified other TFs among NRF1, including NR0B1, and CREB1 (Figure 2.2.3A), suggesting a complex regulatory network governing TE expression. A-MYB, a TF that was previously reported to bind RLTR10 during spermatogenesis, was not identified in our pulldown experiments (Sakashita et al. 2020). Consistent with this observation, RLTR10 remained among the unaffected DNMT3C targets in dKO testes (Figure 2.2.5E,G), indicating that its regulation involves NRF1-independent factors. This finding reinforces the specificity of our *in vitro* IAP promoter pulldown experiment and highlights the distinct regulatory mechanisms governing different TE families. Our results provide a robust platform for further investigation into the intricate mechanisms of TE regulation. Understanding how different TFs interact with TEs could further provide insights into evolutionary adaptations and the development of species-specific regulatory networks.

Looking beyond the fundamental understanding of TE regulation, the study of TFs involved in this process has significant potential for future applications. TFs regulating TEs could serve as valuable biomarkers in cancer diagnostics and prognostics. Aberrant expression of these TFs might indicate dysregulation of TEs, which is often associated with various cancers (Islam et al. 2021). Abnormal levels of specific TFs could lead to increased TE activity, potentially serving as an early indicator of malignant tumors. Furthermore, inhibition of specific TFs that activate oncogenic TEs (Bouras et al. 2021; Garcia-Perez et al. 2010), could potentially repress TE expression and mitigate their detrimental effects in cancer progression. This approach could lead to the development of more targeted and less toxic cancer therapies. As we continue to unravel the intricate relationships between TFs and TEs, we anticipate exciting discoveries that could revolutionize our approach to treating diseases and manipulating cellular processes.

4.Conclusions

In my PhD thesis, I explored two distinct approaches to uncover novel regulators of TEs in two developmental time windows that have high risk for TE activity due to low DNA methylation levels. The first project aimed to develop an innovative endogenous readout for a genome-wide CRISPR/Cas9-based screen in mESCs, while the second project utilized the *Dnmt3C*$^{KO/KO}$ mouse model to investigate TE regulation in spermatogenesis. Both projects highlighted the complexity of TE regulation and the challenges in studying these elements. The endogenous readout approach, although promising, faced technical barriers in optimizing RNA-FISH and IF techniques for a high-throughput screening. The *Dnmt3C*$^{KO/KO}$ study successfully identified NRF1 as a potential regulator of unmethylated IAPs, demonstrating the power of combining genetic models with biochemical approaches. Of note, this study brings value to the field of transposon biology as, to our knowledge, it provides the first comprehensive characterization of a TF regulating TEs in an *in vivo* mouse model. Studies like this also have the potential to bridge understanding in various additional fields like development, evolution, transposon biology and epigenetics.

Both projects underscore that screening for TE regulators can be accomplished through various methods, including genome-wide CRISPR/Cas9 screens, genetic models, and biochemical assays such as DNA pulldowns. However, each approach presents unique challenges, from technical optimizations for high-throughput screens to the complexity of genetic mouse crosses and the application of techniques in low-input samples. Overall, this work contributes to our understanding of TE regulation and provides insights into potential screening strategies, while also highlighting the need for continued methodological innovations in this field.

5. Materials and methods

5.1 Mouse Embryonic Stem Cell (mESC) culture

E14 mESCs were seeded in gelatin-coated plates and were culture with the following media conditions: Dulbecco's Modified Eagle Medium (Thermo Fisher Scientific 21969035), 15% Fetal calf serum (ES approved), 1mM Sodium Pyruvate (Thermo Fisher Scientific 11360039), 0.1mM MEM non-essential amino acids (Thermo fisher Scientific 11140050), 2mM L-Glutamine (Thermo Fisher Scientific 25030024), 100 U/mL Pen Strep (Thermo Fisher Scientific 15140122), 0.1mM 2-Mercaptoethanol, 1000U/mL Leukemia inhibitory factor (LIF, in-house) and in titrated 2i (normal DNA methylation levels): 200 nM MEK inhibitor PD0325901 (Axon Medchem 1408), 3 mM Gsk3 inhibitor CHIR99021 (Axon Medchem 1386). The media was added fresh every day, and the cells were passed every other day at ~70% confluence using TrypLE Express (Gibco). All lines were cultured at 37^{o}C with 5% CO_2 and 80% humidity.

5.2 Generation of ESC lines

5.2.1 Generation KAP1-degron (puromycin) ESC line

To knock in in the degron fused to the C-terminal region of KAP1, we designed a sgRNA targeting the last coding exon of KAP1 and cloned it in a CRISPR-Cas9 genome editing plasmid pX459 (Ran et al. 2013). Degron donor sequences consisting of a linker and FKBP-V-2xHA-P2A-puromycin coding sequences were amplified from pCRIS-PITChv2-C-dTAG-Puro (Nabet et al. 2018) containing 600bp of flanking DNA sequences homologous to the desired integration site in the *Kap1* gene. These PCR products were cloned into a pBluescript backbone using Gibson assembly. We co-transfected this plasmid together with *Kap1*-sgRNA-Cas9 vector into WT E14 mESCs using Lipofectamine2000 (Thermo Fisher, Catalog number: 11668019) to select for the in-frame knock-in integration of the FKBP-V-containing cassette using stable puromycin (1µg/ml) selection.

5.2.2 Generation Cas9-degron inducible ESC line

To generate the Dox-inducible-3xFlag-Fkbpv-Cas9 cell line, we cloned the pTRE3G-FKBPV-Cas9 cassette in a donor plasmid carrying 672 homology arms of the mouse *Tigre* locus. Degron donor plasmid consisted of a ATG, 3xFlag, NLS of SV40, FKBP-V, a linker, and the SpCas9 sequence. The FKBP-V was PCR amplified from the pCRIS-PITChv2-N-dTAG-Puro plasmid (Nabet et al. 2018). We co-transfected the donor plasmid together with a CRISPR/Cas9 genome editing plasmid pX459 carrying a sgRNA targeting the *Tiger* locus of the previously generated cell line stably expressing a Tet-On 3G activator under a EF1α promoter on mouse Rosa26 locus. We first transiently selected the cells by adding 1µg/ml puromycin (expressed from the editing plasmid) followed by 48 hours of 8µm/ml blasticidin selection. After 48 hours blasticidin concentration was reduced to 6µg/ml and kept until the colony picking selection.

5.2.3 Generation of *Suv39h1/2* KO ESC line

To knock out *Suv39h1* we designed two sgRNAs targeting exon 4 and cloned them in a CRISPR-Cas9 genome editing plasmid pDual Guide pX459. The cell line was generated by creating a frame-shift in *Suv39h1* exon 4. Similarly, we designed two sgRNAs to target *Suv39h2* and the cell line was generated by a 119bp deletion on exon 4. E14 mESCs were co-transfected with the two editor plasmids (one for targeting *Suv39h1* and one for *Suv39h2*) and transiently selected the clones 24 hours post-transfection, for 48 hours by adding 1µg/ml puromycin.

5.2.4 Generation of *Suv39h*-dKO-KAP1-degron (no puro) ESC line

To generate this cell line we first had to generate a KAP1-degron cell line with no puromycin (puro) stably expressed under *Kap1* (5.2.1). The plasmids generated at (5.2.3), carried a puromycin resistance cassette, so we would not be able to select cells by using the cell line generated before as a background (5.2.1). Similar to (5.2.1), degron donor sequences consisting of a linker and FKBP-V-2xHA were amplified from CRIS-PITChv2-C-dTAG-Puro (Nabet et al. 2018) containing 45bp of flanking DNA sequences homologous to the desired integration site in the Kap1 gene. These microhomology arms were included as primer extensions and designed to selectively knock-in the dTAG in frame with the last exon of the *Kap1* gene, producing a tagged KAP1-FKBP-V-2XHA protein. Co-transfection of the purified PCR product and the *kap1*-sgRNA-Cas9 vector into WT E14 mESCs was performed using lipofectamine under optimized conditions, and selection for Cas9 expression was applied transiently for 48 h, 24

h after transfection, by adding 1µg/ml puromycin. *Suv39h*-dKO-KAP1-degron cell line was then generated by co-transfecting the editor plasmids targeting *Suv39h1/2* genes (5.2.3) into the KAP1-degron cell line (no puro). The clones were selected transiently by puromycin.

5.3 Colony picking and 96-well plate genotyping

Transfected cells were always seeded in 10cm^2 dishes, in three dilutions, to avoid over-confluency and cross-contamination within different clones. The media was changed fresh every day followed by 1-2 1x DPBS washes to remove dead cells or floating clones, to avoid cross-contamination. Individual clones were picked in 96-well U-bottom plates approximately 10 days post-transfection. The media was aspirated and 20 ml of DPBS were added in the dish. The clones were picked carefully using a P20 pipette (adjusted in 10µl). 50µl of TrypLE were added into each well and quenched with 100µl media after 7 minutes of incubation at 37ºC. The cells were then seeded into gelatin-coated 96-well flat plates. The remaining clones were pooled and the pellets were used for DNA extraction and further used for the PCR genotyping optimization. The primers were always designed spanning the sequence of interest. For knockout cell lines, one extra primer was designed within the anticipated deleted sequence.

For genotyping a 96-well plate, the clones were picked and maintained for approximately four days, with daily media changes. A replica plate was generated when the 96-well plate reached 70-80% confluency. This involved aspirating media, washing with PBS, trypsinizing with 50 µl of trypsin for 5 minutes, and transferring cells to gelatinized plates with media to inhibit trypLE. One plate (replica) was kept in low confluency for maintenance and the genotyping plate was used to extract DNA. The genotyping plate was prepared by aspirating media, washing with PBS, freezing at -80°C for 15–20 minutes, and then adding a lysis buffer containing 100 mM Tris-HCl, 1 mM EDTA, 200 mM NaCl, and 0.5% Tween-20, along with Proteinase K. The genotyping plates were incubated at 55°C for 2–3 hours, then heated at 99°C for 15 minutes, and placed on ice. PCR reactions used 1 µl of extracted DNA. HIgh viscosity at this step is expected (the higher the confluence per well the more viscous it will be). Therefore, careful handling is crucial to avoid cross-contamination within the cells. After confirming the genotype, only positive clones were transferred into individual wells of a 12-well plate for further maintenance. The genotyping of all generated cell lines was further confirmed by sanger sequencing.

5.4 RNA analysis

The RNA was extracted from sorted germ cells (spermatogonia and spermatocytes) (Figure 2.2.1), whole testes (Figure 2.2.5), and mESCs (Figure 2.1.4 A,B), (Figure 2.1.8 F,G).

5.4.1 RNA extraction and library preparation

Sorted germ cells (Figure 2.2.1) or whole testes (Figure 2.2.5) of littermates were collected by Jessica Leismann with the assistance of IMB's Flow Cytometry core facility and frozen in Trizol until triplicates of each genotype were collected. RNA extraction was performed together with Jessica Leismann using the DirectZol RNA MicroPrep kit (Zymo, R2061). NGS library prep was performed by IMB's Genomic core facility with Illumina's Stranded mRNA Prep Ligation Kit following Stranded mRNA Prep Ligation ReferenceGuide (April 2021) (Document 1000000124518 v02). Libraries were prepared with a starting amount of 10 ng for RNA extracted from sorted Spg and Spc and 29 ng for RNA extracted from whole testes mutants and amplified in 13 PCR cycles and 12 PCR cycles, respectively. Two post PCR purification steps were performed to exclude residual primer and adapter dimers. Libraries were profiled in a DNA High Sensitivity chip on a 2100 Bioanalyzer (Agilent technologies) and quantified using the Qubit 1x dsDNA HS Assay Kit, in a Qubit 4.0 Fluorometer (Invitrogen by Thermo Fisher Scientific). 16 Samples from sorted Spg and Spc and 17 samples from whole testes mutants were pooled in equimolar ratio aiming for at least 50 million reads per sample and sequenced on an Illumina NextSeq 2000 sequencing device with 1 dark cycle upfront R1 and R2 as 2x108-nt paired-end reads, yielding 50-60 million read pairs per sample.

5.4.2 RNA sequencing analysis

The RNA sequencing analysis was performed by IMB's Bioinformatics core facility. Sequencing reads were mapped on the *Mus musculus* genome assembly GRCm38 (GENCODE release M25) utilizing STAR (v.2.7.3a, parameters: --outMultimapperOrder Random --outSAMattributes NH HI AS nM MD --outSJfilterReads All --outSAMunmapped Within --outFilterMismatchNoverReadLmax 0.04 --outFilterMismatchNmax 999 --winAnchorMultimapNmax 100 --outFilterMultimapNmax 100) (Dobin et al. 2013). The TEtranscripts program (v.2.2.3, parameters: --mode multi --sortByPos --stranded reverse) was implemented to annotate reads to both genes and transposable elements (Jin et al. 2015). Differentially expressed analysis was performed using DESeq2 (v.1.44.0) with corrected p-value (Benjamini and Hochberg, FDR) < 0.01 (Love, Huber, and Anders 2014).

TElocal (v.1.1.1, parameters: --mode multi --sortByPos --stranded reverse) was employed to quantify transposable element expression at the locus level, and the counts were normalized to TPM (Transcripts Per Million) values.

5.4.3 mESCs: RNA extraction cDNA synthesis

Whole mESCs pellets were harvested and resuspended in 500 µl of TRIZOL and incubated for 15 min on ice. 100µl of chloroform were added, then the samples were vortexed and centrifuged at full speed for 15min at 4ºC. The upper aqueous phase containing the RNA was then transferred to a fresh low-binding tube, 250 µl of isopropanol were added and samples were incubated for 10 min on ice to allow precipitation. RNA was collected by centrifugation at full speed for 30 min at 4ºC and the pellet was washed twice with 500µl of 70% ethanol to remove remaining salt residues. The pellet was resuspended in 50µl of free-nuclease H_2O after being air-dried under the hood for 10-15 min. The RNA concentration was measured with NanoDrop. After RNA extraction, DNase treatment was performed adding DNase I enzyme and its buffer (NEB) to 3µg of RNA per condition (final volume 11.9µl). Samples were incubated for 30 min at 37 °C and the reaction was inhibited by adding EDTA (1µl of 50 mM EDTA into 11.9µl reaction), incubated at 65 °C for 10 min, followed by 5 min incubation on ice. cDNA synthesis was performed by M-MLV reverse transcriptase (Promega).

5.5 RNA FISH

5.5.1 RNA FISH homemade probes

We generated home-made RNA FISH nick translation probes, using the Nick translation kit (Abbott Molecular Inc, REF: 07J00-001, Lot#: 517634). IAP probes consisted of: the full length IAP element IAPez (7156bp) (Supplementary material) cloned into pBluescript labeled with homemade Alexa488. The reaction mixture was prepared with 1 µg of plasmid carrying IAPez DNA sequence, 5 µl dTTP (0.1 mM), 10 µl dNTP mix (0.3 mM dATP, 0.3 mM dCTP, and 0.3 mM dGTP) , 5 µl 10x Buffer, 10 µl Enzyme, 2.5 µl 0.2mM dUTP Fluorophore (Alexa488), and nuclease-free water to a final volume of 50 µl. The mixture was incubated at 15°C for 10 minutes. To determine the probe size, different incubation times (10min, 15min, 45min, 2h, 4h, 6h, 8h, 16h) were tested by tapestation or agarose gel, to generate fragments of probes from a range of 200 bp to 500 bp. The reaction was heat inactivated at 70ºC for 10 minutes and then placed on ice or stored in -20ºC. Around 100 ng (5µl) of probe was ethanol precipitated by adding 1.5µl of 20 µg tRNA, 1:10 sodium acetate, 2 volumes of 100%

ethanol, incubated at -80ºC for 30 minutes and further centrifuged for 30 min at 4ºC at maximum speed. The pellet was resuspended thoroughly in 7µl per sample of 2X hybridization buffer (4X SSC, 20% dextran sulfate, 2mg/ml BSA, 20mM Vanadyl Ribonucleoside Complex) and kept on ice.

5.5.2 RNA FISH (microscopy)

KAP1-degron cells were cultured with 250 nM dTAG and DMSO (control) for 48 hours. Coverslips were treated with poly-D-lysine overnight and washed with DPBS. The cells were harvested and 150,000 cells were seeded on the coverslips for 3-4 hours at 37ºC. The cells were washed in RNase-free PBS, permeabilized in 1X DPBS with 0.5% Triton X-100 for 10 min at room temperature (RT), and fixed in 3% paraformaldehyde/PBS for 10 minutes at room temperature. Cells were dehydrated through an ethanol series (80%, 95%, 100%) for 3 min each. The denatured probe was added to a slide with the coverslip placed cell-side down. Hybridization was performed overnight at 37°C in a humid chamber. Subsequent washes were carried out in 50% formamide/2X SSC and 2X SSC, followed by DAPI or Hoechst counterstaining. Coverslips were mounted using a glycerol-based medium for microscopic analysis. Images were collected on a SP5 confocal microscope (Leica) (Figure 2.1.4C) and THUNDER widefield (Figure 2.1.4D) with a 63x oil objective. Image analysis was performed with the software ImageJ (1.53o version) (Schneider, Rasband, and Eliceiri 2012) and Huygens was used for deconvolution (Figure 2.1.4D).

5.5.3 RNA FISH in-suspension (Flow cytometry)

The single-cell suspension RNA FISH was optimized by adapting the protocols of Reilly et al. and Arrigucci et al. (Reilly et al. 2021; Arrigucci et al. 2017). The probes were labeled via nick translation using 1 µg of DNA per 50 µl reaction. Approximately 0.1 µg of probe was ethanol-precipitated and resuspended in 50 µl hybridization buffer (HB 10% dextran sulfate, 0.1 ng/µl probe). The probes were denatured at 85°C for 5 minutes. KAP1-dTAG and KAP1-DMSO cells were harvested after 48 hours post-treatment and 2 million cells per sample (we recommend to start with 8-10 million cells) were fixed in 4% paraformaldehyde/PBS for 10–15 minutes, washed with DPBS, and permeabilized in 0.2% Tween20 in DPBS for 10 minutes. Cells were then incubated in cold 70% ethanol for 10 minutes at 4°C, washed with PBST, and resuspended in 50 µl denatured probe solution for overnight hybridization at 37°C in the dark, rotating in a hot air oven. Post-hybridization, cells were washed three times in 50% formamide/2xSSC (pH 7.2) and twice in 2xSSC (5 minutes each at 42°C), with centrifugation between washes. DNA was counterstained with 1 µg/ml

DAPI in 400 µl 2xSSC for 10 minutes. Samples were filtered into FACS tubes and analyzed using a Fortessa flow cytometer. All washes were performed by centrifugation in 600 x g (unfixed cells) and 1000xg (after fixation) for 5 min. The samples were analyzed using FlowJo software.

5.6 Immunofluorescence

5.6.1 Testes cryosections

Cryosections of testes at different developmental stages were fixed overnight at 4°C with paraformaldehyde. Next day, the testes were washed twice with 1X PBS for 5-10 min at RT. Two sequentially overnight incubations at 4°C were performed in 15% and 30% sucrose solutions (in 1X PBS), respectively. Testes were embedded in O.C.T compound (FSC 22 Frozen Section Compound, Leica) for 1 hour at RT. After dissecting the epididymis, the testes were embedded in O.C.T in a cryoblock and frozen on dry ice. 10µm sections were cut and mounted onto Superfrost Plus slides (Epredia) and stored at -80°C.

5.6.2 Immunofluorescence staining (microscopy)

For immunofluorescence stainings on testes cryosection, slides were calibrated at RT, marked with a hydrophobic pen (ImmEdge Pen, VectorLabs) around the O.C.T, washed in 1X PBS for 10 minutes to remove the O.C.T), and the blocked and permeabilized (10% donkey serum (Sigma), 3% BSA, 0.3% TritonX-100, 1X PIC in PBS) for 1 hour at RT. The sections were incubated with primary antibodies overnight at 4°C in blocking buffer. Next day they were washed with 1X PBST (0.1% Tween-20 in 1X PBS) washes. The sections were incubated for 1h at RT with donkey anti-rabbit Alexa Fluor 647, donkey anti-rat Alexa Fluor 488 secondary antibodies (ThermoFischer) at 1:800 dilution and DAPI (2 µg/µl), to stain nuclei. The following primary antibodies were used in the indicated dilutions: rabbit monoclonal anti-mouse L1-ORF1p at 1:400 (Abcam, 216324), rat monoclonal anti-mouse TRA98 at 1:500 (Abcam, 82527) and rabbit monoclonal NRF1 at 1:600 (Abcam, 175932), rabbit monoclonal IAP-POL at 1:400 (gift from Boulard lab); For the IAP-POL antibody due to high background, all washes were conducted in 0.3% Tween-20 in 1x PBS shaking for 15 minutes per wash. Three 1X PBST washes were followed and the sections were mounted with ProLong Gold antifade mount (Life Technologies). All images were acquired with a Stellaris 8 confocal microscope (Leica Microsystems, Wetzlar, Germany) applying the same setting and all images were processed with ImageJ (1.53o version) (Schneider, Rasband, and Eliceiri 2012).

For staining mESCs, the protocol was the same except for the first step. 100,000 cells were spotted for 1 hour on a poly-D-lysine coated to attach and fixed with 4% PFA for 15 min. The PFA was washed twice with 1x PBS and then we proceeded to blocking and permeabilization.

5.6.3 Immunofluorescence analysis

Quantification of TRA98 and L1ORF1 signals in the nucleus and cytoplasm was performed by Microscopy Core Facility using 256 x 256-pixel images of manually cropped cells from a larger field of view. Nuclei were segmented using the DAPI channel. A Gaussian filter (sigma=6) was applied, and thresholds were set using the OTSU algorithm (Otsu, 1979). Segmentation was refined by filling holes and using a watershed operation to separate touching nuclei, generating the region of interest (ROI) for the nucleus. The cytoplasm ROI was defined by creating a ring around the nucleus ROI, dilating it 10 times, and subtracting the nucleus ROI. Mean TRA98 and L1ORF1 intensities were measured within the nucleus and cytoplasm ROIs. The total number of nuclei in whole sections was calculated using the Analyze Particles tool in ImageJ after using a threshold. GraphPad Prism v10 was used to generate the bar plot.

5.6.4 Immunofluorescence staining (flow cytometry)

For single-cell suspension preparation and immunostaining of *Dnmt3C*$^{KO/KO}$ and WT testicular cells, testes were dissected, decapsulated in PBS, and transferred to a 2 mL tube containing 600 µl collagenase solution (HBSS + 0.5 mM $CaCl_2$, 1 U/ml collagenase). Tubules were dissociated at 37°C for 6-10 minutes with periodic flicking, followed by incubation with 600 µl TrypLE at 37°C for 4-8 minutes, with gentle pipetting using truncated P1000 tips (to avoid cell breakage). The reaction was neutralized with 300 µl FBS, filtered to remove clumps, and centrifuged at 600xg for 4 minutes. Cells were washed sequentially with PBS/10% FBS and PBS, then fixed in 4% PFA in PBS on ice for 15 minutes. After PBS washes, cells were permeabilized in 200 µl ice-cold permeabilization buffer (1× PBS, 0.1% Triton X-100, 1% BSA, 10% goat serum, 1xPIC) for 10 minutes. Primary antibody staining was conducted using anti-L1ORF1p or anti-IAP-GAG (1:400 dilution in permeabilization buffer) at 4°C for 30 minutes, followed by two staining buffer (SB) (1× PBS, 1% BSA) washes. Secondary antibody (AlexaFluor546, 1:800) incubation proceeded for 30 minutes at room temperature, with subsequent SB washes. Cells were resuspended in DAPI (1 µg/ml in SB) for nuclear counterstaining. All solutions were prepared fresh, including collagenase solution (HBSS + 0.5 mM $CaCl_2$ + 91 µl 11 U/ml collagenase), SB (1× PBS, 1% BSA), and

permeabilization buffer (1× PBS, 0.1% Triton X-100, 1% BSA, 10% goat serum, 1× protease inhibitor). Centrifugation steps were consistently performed at 600-1,000 xg for 4-5 minutes (after fixation the speed can be increased up to 1000 xg to avoid cell loss).

Similarly, for single-cell suspension preparation and immunostaining mESCs, the same protocol was followed skipping the testes dissociation and instead we harvested 4 million cells per sample, washed the pellet once with SB and proceeded to 4% PFA fixation. The samples were analyzed using FlowJo software.

5.7 IAP-GAG antibody generation

The production of the MmLTR1-GAG antibody involved cloning, expression, and purification steps, ultimately utilizing the His-MBP-tagged expression vector for antigen production.We first cloned the amplified and cloned the GAG epitope (Supplementary material) in the backbone of the expression vector by Gibson assembly. The generated plasmid was confirmed by sanger sequencing.

The IMB Protein Production Core Facility expressed the antigen using E. coli BL21(DE3) cells under IPTG induction at 18°C overnight. Following lysis by high-pressure homogenization, His-MBP-tagged GAG was purified using immobilized metal affinity chromatography (IMAC) with high-salt washes. The His-MBP tag was cleaved with His-3C protease during dialysis, and untagged GAG was separated from His-MBP by reverse IMAC. Untagged GAG was further purified via gel filtration using a Superdex 75 column, yielding monomeric protein (~12 kDa). The final yield of untagged GAG was 2.7 mg from 4 L of culture, stored as aliquots at 1 mg/ml in PBS.

For antibody production, 8 vials of purified antigen (200 µg/vial) were sent to Eurogentec for rabbit immunization. Serum from two rabbits (#SY1294 and #SY1295) was used for affinity purification. Antigen coupling to SulfoLink resin was performed by reacting 1 mg of GAG protein (0.5 mg/ml in coupling buffer with 2 mM TCEP) with 1.5 ml SulfoLink resin at RT for 2 hours. After washing and quenching nonspecific binding sites with L-cysteine, the resin was applied for affinity purification of antibodies from rabbit serum. Serum was filtered and applied to the resin overnight at 4°C, followed by washing with PBS + 0.1% Triton X-100 and PBS alone. Antibodies were eluted with glycine buffer (pH 2.5), neutralized with Tris buffer (pH 8.5), and rebuffered into PBS containing 10% glycerol and 0.05% NaN_3 using PD-10 columns. The antibodies were concentrated to 1 mg/ml using Amicon spin concentrators, aliquoted, snap-frozen in liquid nitrogen, and stored at -80°C. Purified antibodies were diluted and stored into PBS containing 10% glycerol (buffer) and concentrated to 1 mg/ml. Final yields were 1.35 mg for rabbit #SY1294 (27 aliquots of 50 µl

each) and 1 mg for rabbit #SY1295 (20 aliquots of 50 µl each). Working aliquots were stored at 4°C (in 40-50% glycerol) to minimize freeze-thaw cycles.

5.8 Western blot

Protein samples were extracted by homogenizing testis or cell pellets in standard RIPA buffer supplemented with cOmplete, Mini, EDTA-free Protease-inhibitor-cocktail (PIC) (Roche, 11836170001); Every total protein lysate was extracted from one testis not a pool of testes. The samples were sonicated for 5 minutes (30 sec ON, 30 sec OFF) and centrifuged for 30 min at 16,000g at 4°C. The total protein concentration was quantified by Bradford assay. Protein samples were diluted in 1X LDS and boiled at 95oC for 5 minutes. The proteins were separated on SDS-PAGE (GeneScript, M00653, M00654) and transferred to nitrocellulose membranes (ThermoFischer, 88018). Membranes were blocked for 40 minutes with 5% BSA in TBST and incubated with primary antibodies overnight at 4°C. All antibodies with the corresponding concentration are listed in Table 2. Antibodies. The membranes were washed twice with TBST for 15 min; 15µg were used from all protein lysates. In Figure 2.2.6G the total protein load was increased to 25µg and after the primary antibody the washes were performed in 0.5 M NaCl, 0.5% Triton X-110, PBS, to reduce the Δ between *Dnmt3C*$^{KO/KO}$ and WT or any potential background. The blots were incubated with HRP-conjugated secondary antibodies (1:5000) for 1h at RT. The protein bands were visualized using chemiluminescence reagents (SuperSignal West Pico, Femto, ThermoFischer, 1863096).

5.9 Mouse housing and breedings

The mouse housing and all breedings were handled by Jessica Leismann. All the animal procedures complied with institutional and EU animal welfare regulations and received formal approval from the Animal Ethics Committee of Rhineland-Palatinate, Germany (Approval ID: G 23-5-049). Mice aged 7–24 weeks were maintained in ventilated cages enriched with environmental stimuli under controlled conditions (humidity: 40–70%, temperature: 22 ± 2°C, 12-hour light/dark cycle) with unrestricted access to food and water. Euthanasia was performed via cervical dislocation under approved protocols. Timed pregnancies were established to obtain prenatal testes, with embryonic day 0.5 (E0.5) designated as the first day postcoitum, while postnatal samples were collected starting at birth (designated P1). All mouse strains had been previously described and were bred on a C57Bl6/J genetic background (*Dnmt3C*$^{KO/KO}$, *Nrf1*$^{f/f}$, *Ddx4*^{Cre} (also known as *Vasa*^{Cre}) and *Oct4-eGFP* (Jackson Laboratories, stock 008214).

To generate *Nrf1* conditional knockout in spermatogonia germ cells, $Nrf1^{fl/fl}$ mice were crossed with *Ddx4-Cre* mice. To do so, $Nrf1^{fl/KO}$ females were bred with $Nrf1^{fl/WT}$;$Ddx4^{Cre}$ males (Figure 5.1). For generate $Nrf1^{fl/KO}$;$Ddx4^{Cre}$; $Dnmt3C^{KO/KO}$ male offsprings, $Nrf1^{fl/KO}$;$Dnmt3C^{KO/KO}$ females were crossed with $Nrf1^{fl/WT}$;$Ddx4^{Cre}$;$Dnmt3C^{KO/WT}$. Prenatal testes were collected from timed pregnancies (E0.5 designated as the first day postcoitum), and postnatal samples were obtained starting at birth (P1). Genotyping was performed via PCR and further IF to confirm allelic recombination and *Cre* activity ($Dnmt3C^{KO/KO}$;$Nrf1^{cKO/KO}$).

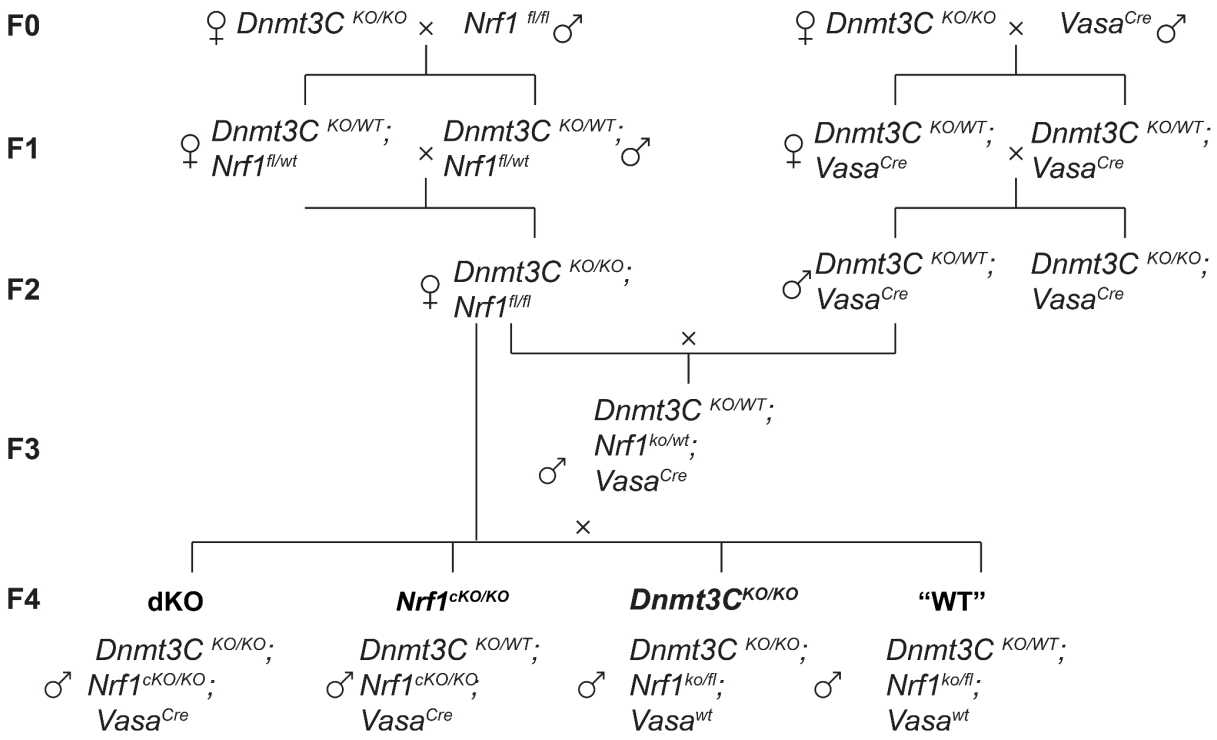


Figure 5.1 Mice breeding scheme.

5.10 FACS

Prospermatogonia were isolated from E15.5 Oct4-eGFP male mice testes and collected in 1X PBS. Testes were decapsulated and transferred to tubes with 100 µl collagenase solution in HBSS (Collagenase A, 2x AAs, 2x Na-pyruvate, and 25 mM HEPES-KOH, pH 7.5) and dissociated at 37 °C for 8 min. After adding 200 µl TrypLE, the cell suspension was incubated for 5 minutes at 37 °C and pipetted carefully to obtain a single-cell suspension. TrypLE was quenched by adding 70 µl pre-warmed FBS, and cells were washed twice with FACS buffer (PBS, 2 mM EDTA, 25 mM HEPES-KOH, 1.5% BSA, 10% FBS) after centrifugation at 600xg for 4 min. Cells were resuspended in

DAPI-supplemented (2 µg/mL) FACS buffer, filtered through 40µm cell strainers, and sorted using an Invitrogen Bigfoot with a 100-µm nozzle. DAPI-negative and OCT4-eGFP-positive cells were sorted into low-binding tubes containing 100 µl PBS.

Spermatogonia cells were isolated from P14 mice testes and processed similarly. Single-cell suspensions were stained with 8.25 mg per 5 million cells of CD326 (Ep-CAM)-Alexa Fluor647 (BioLegened, 118212) and β-2-Microglobulin-PE (Santa Cruz Biotechnology, SC-32241) for 20 min on ice, washed with FACS buffer twice and resuspended in DAPI-supplemented (1 µg/mL) FACS buffer, then sorted using the same conditions. DAPI-negative, β-2-Microglobulin-negative, and Ep-CAM-positive cells were collected in 100 µl PBS.

Spermatocytes were isolated from mice testes older than P20. Testes were decapsulated and dissociated in 2 mL collagenase solution (40U ofCollagenase A, 1.5x AAs, 1.5x Na-pyruvate, and 25 mM HEPES-KOH, pH 7.5) at 37 °C. The cell suspension was sedimented for 2 min, and the supernatant was aspirated. After adding 250 µl TrypLE and filtering, 10 million cells were stained with 0.5 µL Hoechst 3342 (Thermo Scientific, 62249) for 8 min. Tryple was quenched by FBS followed by a centrifugation, cells were resuspended in buffer supplemented with 1µL Hoechst 3342 per 10 million cells at 35 °C for another 8 minutes, filtered, and stained with Propidium Iodide (Invitrogen, P3566). Cells were sorted using the same conditions by Propidium Iodide exclusion and Hoechst 3342 staining, indicating 2N DNA content.

5.11 DNA pulldown

The DNA pulldown was performed by Jessica Leismann. Nuclear proteins were isolated from P15 mouse testes (n=20 males per replicate, three replicates). The tissue was first processed in a cytoplasmic extraction buffer containing 10 mM Hepes-KOH (pH 7.5), 10 mM KCl, 1.5 mM MgCl2, 1 mM DTT, and protease inhibitor cocktail. Following a 30-minute ice incubation and centrifugation (1100 g, 4°C, 10 minutes), nuclei were lysed in a high-salt buffer (20 mM Hepes-KOH pH 7.9, 420 mM NaCl, 2 mM MgCl2, 20% glycerol, 1 mM DTT, 0.2% IGEPAL CA-630, protease inhibitors). The insoluble nuclear fraction underwent sonication (10 cycles of 10s on/50s off) in the same buffer. After centrifugation (14,000 g, 4°C, 20 minutes), the soluble fractions were pooled and protein concentration was determined using Bradford assay. Following the protocol adapted from (Harris et al. 2018), streptavidin-conjugated magnetic beads were incubated with biotinylated bait DNA (60 µg) in coupling buffer (20 mM Tris-HCl pH 8.0, 2M NaCl, 1 mM EDTA, 0.1% Tween-20) for 1 hour at RT. The DNA-bound beads were then equilibrated in the pulldown buffer (50 mM Tris-HCl

pH 8.0, 150 mM NaCl, 5 mM MgCl2, 0.5% IGEPAL CA-630, 1 mM DTT, protease inhibitors) before incubation with nuclear protein extracts (250 µg) for 3 hours at 4°C. Following three wash steps, protein complexes were eluted using LDS buffer supplemented with DTT, heat-denatured (80°C, 10 minutes), and processed for mass spectrometry analysis.

5.12 LC-MS preparation and analysis

The LC-MS preparation and analysis by the Butter lab. Proteins were resolved using 4%–12% NuPAGE Novex Bis-Tris precast gels with 1x NuPAGE MES running buffer at 180 V for 8 minutes. Gels were fixed in a solution of 7% acetic acid and 40% methanol, stained with Coomassie G250, and destained with water. Individual gel lanes were excised and fragmented, followed by destaining in a buffer containing 50% ethanol and 25 mM ammonium bicarbonate (pH 8.0). After dehydration with acetonitrile (ACN), proteins underwent reduction with 10 mM DTT (56°C, 60 minutes) and alkylation with 50 mM iodoacetamide (room temperature, 45 minutes, dark conditions). The dehydrated gel fragments were subjected to overnight tryptic digestion at 37°C using MS-grade trypsin. Peptides were extracted using 30% ACN, followed by complete dehydration. After ACN removal, peptides underwent desalting on StageTips. Tryptic peptides were separated on a Reprosil C18-packed nanocapillary using an Easy nLC 1000 system. The separation employed a 90-minute gradient (2-60% ACN with 0.1% formic acid) at 225 nL/min flow rate. Analysis was performed on a Q Exactive Plus mass spectrometer using HCD fragmentation with Top10 MS/MS data-dependent acquisition. MaxQuant software (version 1.6.5.0) was used for data analysis against a mouse uniprot database, employing match-between-runs and LFQ quantitation. Protein groups were filtered to require two peptides minimum (one unique), excluding contaminants and reverse hits. Further analysis and visualization were performed using R.

5.13 CUT&Tag

The CUT&Tag in germ cells was conducted by Jessica Leismann. CUT&Tag experiments were performed on freshly harvested mESCs in technical triplicates or isolated germ cells obtained by FACS in biological duplicates. CUT&Tag was conducted following the established protocol with minor modifications (Kaya-Okur et al. 2019). For germ cells 10,000-20,000 cells and for mESCs 100,000 were bound to 10 µL of Concanavalin A-coated beads (Polysciences, 86057-10) and incubated with primary antibodies targeting the protein of interest, as detailed in (Table 2). After binding with secondary antibodies, protein A-Tn5 transposase (Diagenode, C01070001) was added at a dilution of 1:400 to insert sequencing

adapters at target sites. The Wash Digest Buffer was adjusted with 0.05% Digitonin for prospermatogonia, spermatogonia and mESCs, whereas 0.025% Digitonin for spermatocytes. DNA extraction was done using the DNA Clean & Concentrator kit (Zymo, D4014), and libraries were amplified via PCR using Q5 High-Fidelity DNA Polymerase (M0491S) with 15 cycles of amplification for all germ cells and mESCs samples. AMPure beads clean-up was performed to remove primer-dimers with a double-sided size selection: fragments bigger than 10,000 bp were removed by adding 0.3X beads and fragments smaller than 200 bp were removed by adding 1.3X ratio. The resulting libraries were pooled and sequenced on an Illumina NextSeq 2000 platform using paired-end reads (2x155 nt), generating 7-15 million read pairs per sample.

5.13.1 CUT&Tag-qPCR

CUT&Tag-qPCR was performed to evaluate the profiling method of NRF1 in mESCs. The primers used for testing the NRF1 enrichment at known positive control targets and TEs are listed in Table 1. The qPCR was performed using Power SYBR Green PCR Master Mix. 5 µL of each sample was used per reaction.

5.13.2 Bioinformatic analysis

Paired-end reads were quality-trimmed using Trim Galore (v.0.6.10) with a minimum read length of 10 bp, and quality reports were generated using FastQC. Trimmed reads were aligned to the mm10 reference genome using Bowtie2 (v.2.5.3) in local alignment mode allowing a minimum fragment length of 10 bp and a maximum of 1,000 bp, while excluding mixed and discordant alignments. Multimapping reads were assigned randomly as previously described (Teissandier et al. 2019). Duplicates were retained, as CUT&Tag fragments can share identical starting and ending positions because the Tn5 integration sites are affected by DNA accessibility. BAM files were first indexed using Samtools (v.1.10). Then, BigWig files were generated from the BAM files using bamCoverage (v. 2.27.1) , with normalization set to RPGC. DNMT3C-targets were identified using published WGBS data (Barau et al. 2016). Downstream visualization of reads as heatmaps or metaplots were generated using deepTools (v.3.5.5) and was performed by Jessica Leismann and IMB's Bioinformatics facility. BAM files of replicates were merged using Samtools (v.1.10). Peak calling was performed using MACS2 (v.2.1.2) on aligned BAM files, with no control BAM files specified. The analysis was conducted using broad peak calling mode, with genome size set to 'mm' (for mouse) and otherwise default settings.

5.14 EZH2i treatment

WT and TKO mESCs were seeded on 0.2% gelatin-coated plates before treatment. For chemical EZH2 inhibition, 5μM of GSK16 (MedChem Express, HY-13470) inhibitor was mixed with mESCs culture media (supplemented with titrated 2i and LIF) and cultured for 5 days. In parallel, WT and TKO mESCs were treated with DMSO and used as a negative control treatment; media was changed daily adding fresh GSK16.

Table 1. Primers and Oligonucleotides

Primers and Oligonucleotides				
Name	**Application**	**Full Sequence (5' to 3')**	**Annealing (5' to 3')**	**5' Extension**
sgRNA1 ss mouse Rosa26	Cloning	ACCGTGTATGTAAC TAATCTGTCGT	ACCGTGTATGTAACT AATCTGTCGT	
sgRNA1 as mouse Rosa26	Cloning	TAAAACGTGTATGTA ACTAATCTGT	TAAAACGTGTATGTA ACTAATCTGT	
sgRNA2 ss mouse Rosa26	Cloning	ACCGTTACATACAC CACAAATCGGT	ACCGTTACATACAC CACAAATCGGT	
sgRNA2 as mouse Rosa26	Cloning	TAAAACCGATTTGT GGTGTATGTAA	TAAAACCGATTTGTG GTGTATGTAA	
sgRNA1 ss mouse TIGRE	Cloning	CACCACGTAAATTA GAGTTCTAAG	CACCACGTAAATTA GAGTTCTAAG	
sgRNA1 as mouse *Tigre*	Cloning	AAACACCTTAGAAC TCTAATTTACGT	AAACACCTTAGAAC TCTAATTTACGT	
sgRNA2 ss mouse *Tigre*	Cloning	ACCGCTAATTTACGT ACCCCCTCCGT	ACCGCTAATTTACGT ACCCCCTCCGT	
sgRNA2 as mouse *Tigre*	Cloning	TAAAACGGAGGGG GTACGTAAATTAG	TAAAACGGAGGGGG TACGTAAATTAG	
sgRNA1 ss mouse KAP1/Trim28	Cloning	CACCGCACAAGAG CCCCAGCTTCAG	CACCGCACAAGAGC CCCAGCTTCAG	

Primers and Oligonucleotides				
Name	**Application**	**Full Sequence (5' to 3')**	**Annealing (5' to 3')**	**5' Extension**
sgRNA1 as mouse KAP1/Trim28	Cloning	AAACCTGAAGCTGGGGCTCTTGTGC	AAACCTGAAGCTGGGGCTCTTGTGC	
dTAG-2HA-P2A-puro-NO microHR-KAP1 FWD	Generating PCR product (microhomology arms) for -FKBPV-xHA-NOpuro for KnockIn in *Kap1*)	GCTGGCCTAAGTTCTCAGGAGCTCTCTGGCCCTGGTGATGGCCCCGGTGGCGGTGGCTCG	GGTGGCGGTGGCTCG	GCTGGCCTAAGTTCTCAGGAGCTCTCTGGCCCTGGTGATGGCCCC
dTAG-2HA-P2A-puro-NO microHR-KAP1 REV	Generating PCR product for -FKBPV-xHA-NOpuro for KnockIn in *Kap1*)	CCAGAGCTGGACTGGGCTGACCACAAGAGCCCCAGCTTCATGCATAGTCCGGGACATCAT	TGCATAGTCCGGGACATCAT	CCAGAGCTGGACTGGGCTGACCACAAGAGCCCCAGCTTCA
mouse dTAG-2HA-P2A-puro-microHR-KAP1 REV	Cloning (PCR amplification) -FKBPV-xHA-puro and clone into a plasmid donor for KnockIn in *Kap1*)	GAGACCAGAGCTGGACTGGGCTGACCACAAGAGCCCCAGCTTCATCAGGCACCGGGCTTG	TCAGGCACCGGGCTTG	GAGACCAGAGCTGGACTGGGCTGACCACAAGAGCCCCAGCTTCA
dTAG-2HA-P2A-puro-microHR-KAP1 FWD	Cloning (PCR amplification) -FKBPV-xHA-puro and clone into a plasmid donor for KnockIn in *Kap1*)	GCTGGCCTAAGTTCTCAGGAGCTCTCTGGCCCTGGTGATGGCCCCGGTGGCGGTGGCTCG	GGTGGCGGTGGCTCG	GCTGGCCTAAGTTCTCAGGAGCTCTCTGGCCCTGGTGATGGCCCC
dTAG KI/WT *Kap1* FWD	For amplifying RHA of KAP1 locus and clone into a plasmid donor to transfect for KnockIn in *Kap1*)	TGCAGGTGGAAACCATCTCC	TGCAGGTGGAAACCATCTCC	
dTAG KI/WT *Kap1* REV	PCR genotyping for KI/WT allele *Kap1*	TTTGTGAAGAGGGACTGAGGG	TTTGTGAAGAGGGACTGAGGG	
dTAG KI *Kap1* FWD	PCR genotyping for KI specific allele *Kap1*	TGGTGCATGACCCGCAAG	TGGTGCATGACCCGCAAG	

Primers and Oligonucleotides				
Name	**Application**	**Full Sequence (5' to 3')**	**Annealing (5' to 3')**	**5' Extension**
dTAG-Cas9 KI in mouse *Tigre* FWD	PCR genotyping for Cas9 KI	ACTAGTTAACTCTCAGGTCCCC	ACTAGTTAACTCTCAGGTCCCC	
dTAG-Cas9 KI in mouse *Tigre* REV	PCR genotyping for Cas9 KI	TCAGTTTGGCTCTCTGACCC	TCAGTTTGGCTCTCTGACCC	
dTAG-Cas9 KI in mouse *Tigre* REV	PCR genotyping for Cas9 KI	ACCGAATACGGACTCGTTTCTTGGGT	ACCGAATACGGACTCGTTTCTTGGGT	
sgRNA1 fwd *Suv39h1* KO	Cloning/ Generating *Suv39h*-dKO cell line	CACCGGCCAGATCTACGACCGCCA	CACCGGCCAGATCTACGACCGCCA	
sgRNA1 rev *Suv39h1* KO	Cloning/ Generating *Suv39h*-dKO cell line	AAACTGGCGGTCGTAGATCTGGCC	AAACTGGCGGTCGTAGATCTGGCC	
sgRNA2 rev *Suv39h1* KO	Cloning/ Generating *Suv39h*-dKO cell line	TAAAACCCAAGAAACGAGTCCGTATTC	TAAAACCCAAGAAACGAGTCCGTATTC	
sgRNA2 fwd *Suv39h1* KO	Cloning/ Generating *Suv39h*-dKO cell line	ACCGAATACGGACTCGTTTCTTGGGT	ACCGAATACGGACTCGTTTCTTGGGT	
sgRNA1 fwd *Suv39h2* KO	Cloning/ Generating *Suv39h*-dKO cell line	CACCGTCTTCACTTGTGATCACCTA	CACCGTCTTCACTTGTGATCACCTA	
sgRNA1 rev *Suv39h2* KO	Cloning/ Generating *Suv39h*-dKO cell line	TAAAACTAGGTGATCACAAGTGAAGAC	TAAAACTAGGTGATCACAAGTGAAGAC	
Suv39h1 ko fwd	genotyping KO/WT	CATCGCTCCTGTTCACTGTTC	CATCGCTCCTGTTCACTGTTC	
Suv39h1 ko fwd	Genotyping WT (anneals only in WT)	TCCTTGTTCTGAGCCTCTGG	TCCTTGTTCTGAGCCTCTGG	

Primers and Oligonucleotides				
Name	**Application**	**Full Sequence (5’ to 3’)**	**Annealing (5’ to 3’)**	**5’ Extension**
Suv39h1 ko rev	genotyping ko/wt	GTAGGTGAGGACAAGTGGAG	GTAGGTGAGGACAAGTGGAG	
Suv39h2 ko fwd	Genotyping ko/wt	GTCACCTGTGGTCCTATGTG	GTCACCTGTGGTCCTATGTG	
Suv39h2 ko fwd	genotyping ko/wt	GACCTGGACTACGAGTCTGA	GACCTGGACTACGAGTCTGA	
Suv39h2 ko rev	genotyping ko/wt	GGACCAGAGCTAAGTAGGCA	GGACCAGAGCTAAGTAGGCA	
IAP Δ1 FWD (Zamudio et al. 2015)	RT-qPCR	AACGCTGCTGCTTTAACTCC	AACGCTGCTGCTTTAACTCC	
IAP Δ1 REV (Zamudio et al. 2015)	RT-qPCR	TGCACATAAAGCTGGCACA	TGCACATAAAGCTGGCACA	
Gapdh FWD (Walter et al. 2016)	RT-qPCR	TCCATGACAACTTTGGCATTG	TCCATGACAACTTTGGCATTG	
Gapdh REV (Walter et al. 2016)	RT-qPCR	CAGTCTTCTGGGTGGCAGTGA	CAGTCTTCTGGGTGGCAGTGA	
IAPez_FWD (Deniz et al. 2018)	RT-qPCR	AAGCAGCAATCACCCACTTTGG	AAGCAGCAATCACCCACTTTGG	
IAPez_REV (Deniz et al. 2018)	RT-qPCR	CAATCATTAGATGTGGCTGCCAAG	CAATCATTAGATGTGGCTGCCAAG	
IAP_Gag_FWD (Deniz et al. 2018)	RT-qPCR	AATCTCAGAACCGCTCCATGA	AATCTCAGAACCGCTCCATGA	
IAP_Gag_REV (Deniz et al. 2018)	RT-qPCR	TTTCTTAAAATGCCCAGGCTTT	TTTCTTAAAATGCCCAGGCTTT	
MMERKV10C_FWD (Deniz et al. 2018)	RT-qPCR	CAAATAGCCCTACCATATGTCAG	CAAATAGCCCTACCATATGTCAG	

Primers and Oligonucleotides				
Name	**Application**	**Full Sequence (5’ to 3’)**	**Annealing (5’ to 3’)**	**5’ Extension**
MMERKV10C_REV (Deniz et al. 2018)	RT-qPCR	GTATACTTTCTTCTTCAGGTCCAC	GTATACTTTCTTCTTCAGGTCCAC	
MuERV-L FWD (Zamudio et al. 2015)	RT-qPCR	CAATGGGAAGGTCCAGAAGA	CAATGGGAAGGTCCAGAAGA	
MuERV-L REV (Zamudio et al. 2015)	RT-qPCR	CCTTGTTACCTCGGAATCCA	CCTTGTTACCTCGGAATCCA	
IAP 5' UTR FWD (Rowe et al. 2010)	RT-qPCR, CUT&Tag-qPCR	CGGGTCGCGGTAATAAAGGT	CGGGTCGCGGTAATAAAGGT	
IAP 5' UTR REV (Rowe et al. 2010)	RT-qPCR, CUT&Tag-qPCR	ACTCTCGTTCCCCAGCTGAA	ACTCTCGTTCCCCAGCTGAA	
L1_promoter - FWD (Bulut-Karslioglu et al. 2014)	RT-qPCR	ACTGCGGTACATAGGGAAGC	ACTGCGGTACATAGGGAAGC	
L1_promoter - REV (Bulut-Karslioglu et al. 2014)	RT-qPCR	TGTGATCCACTCACCAGAGG	TGTGATCCACTCACCAGAGG	
L1_ORF1 - FWD (Bulut-Karslioglu et al. 2014)	RT-qPCR	CACTCCCACCCCACCTAGT	CACTCCCACCCCACCTAGT	
L1_ORF1 REV (Bulut-Karslioglu et al. 2014)	RT-qPCR	TAACTCTTTAGCAGTGCTCTCCTGT	TAACTCTTTAGCAGTGCTCTCCTGT	

Primers and Oligonucleotides				
Name	**Application**	**Full Sequence (5' to 3')**	**Annealing (5' to 3')**	**5' Extension**
L1_ORF2 - FWD (Bulut-Karslioglu et al. 2014)	RT-qPCR	ACCTGGACGAAATGGACAAA	ACCTGGACGAAATGGACAAA	
L1_ORF2 - REV (Bulut-Karslioglu et al. 2014)	RT-qPCR	CATCTGGTCCTGGGCTTTT	CATCTGGTCCTGGGCTTTT	
L1 Gf FWD (Zamudio et al. 2015)	RT-qPCR	CTCCTTGGCTCCGGGACT	CTCCTTGGCTCCGGGACT	
L1 Gf REV (Zamudio et al. 2015)	RT-qPCR	CAGGAAGGTGGCCGGTTGT	CAGGAAGGTGGCCGGTTGT	
L1 Tf FWD (Zamudio et al. 2015)	RT-qPCR	CAGCGGTCGCCATCTTG	CAGCGGTCGCCATCTTG	
L1 Tf REV (Zamudio et al. 2015)	RT-qPCR	CACCCTCTCACCTGTTCAGACTAA	CACCCTCTCACCTGTTCAGACTAA	
Asz1 FWD	CUT&Tag-qPCR Nrf1 binding site 1 *Asz1*	TCACTATCGCTGCTCTCGC	TCACTATCGCTGCTCTCGC	
Asz1 REV	CUT&Tag-qPCR Nrf1 binding site 1 *Asz1*	AGCTCTGATAGTGCGCAGG	AGCTCTGATAGTGCGCAGG	
Asz1' FWD	CUT&Tag-qPCR Nrf1 binding site 2 *Asz1*	CGCACTATCAGAGCTTGAGC	CGCACTATCAGAGCTTGAGC	
Asz1' REV	CUT&Tag-qPCR *Nrf1* binding site 2 *Asz1*	AGGTTCCGGTGTGAAAGGCC	AGGTTCCGGTGTGAAAGGCC	
Nrf1 target (control)	CUT&Tag-qPCR *Nrf1* methyl-sensitive binding	AGCCATGCCAGTTGGAGA	AGCCATGCCAGTTGGAGA	

Primers and Oligonucleotides				
Name	**Application**	**Full Sequence (5' to 3')**	**Annealing (5' to 3')**	**5' Extension**
Nrf1 target (control)	CUT&Tag-qPCR *Nrf1* methyl-sensitive binding	CAGTTCCCAAAGGTGCAGAA	CAGTTCCCAAAGGTGCAGAA	
L1MdA FWD	CUT&Tag-qPCR	CTGCCCCAATCCAATCGCG	CTGCCCCAATCCAATCGCG	
L1MdA REV	CUT&Tag-qPCR	GAGCGGAAGGGACTTGTGC	GAGCGGAAGGGACTTGTGC	
L1MdT FWD	CUT&Tag-qPCR	AGAGGGTGCGCCAGAGAA	AGAGGGTGCGCCAGAGAA	
L1MdT REV	CUT&Tag-qPCR	CCTGTCCTCTGGTCCGGA	CCTGTCCTCTGGTCCGGA	
OPP150 T7 term FWD	Cloning backbone expression vector (IAP-GAG Ab)	CCTTGGGGCCTCTAAACGGGTCTTG	CCTTGGGGCCTCTAAACGGGTCTTG	
OPP1 REV	Cloning backbone expression vector (IAP-GAG Ab)	GGGCCCCTGGAACAGAACTTCCAG	GGGCCCCTGGAACAGAACTTCCAG	
GAG epitope	for cloning the Gag epitope-	CTGTTCCAGGGGCCCTTAACGGGTCAGGGAGCTTA	TTAACGGGTCAGGGAGCTTA	CTGTTCCAGGGGCCCC
GAG epitope	for cloning the Gag epitope- rev	TTAGAGGCCCCAAGGTCATTATAAACCTGCATTGCTAAGGG	TCATTATAAACCTGCATTGCTAAGGG	TTAGAGGCCCCAAGG

Table 2. Antibodies

Antibodies					
Antibody target	**Species**	**Source**	**Cat. No.**	**Application**	**Dilution**
HA-tag	Rabbit	Cell Signaling	3724	IF/WB	1:1000/1:5000
α-TUBULIN	Mouse	Sigma	T5168	WB	1:10000

Antibodies					
FLAG-tag	Mouse	Sigma	F1804-200UG	WB	1:5000
PCNA	Mouse	DAKO (Agilent)	M0879	WB	1:2000
L1ORF1p	Rabbit	Abcam	ab216324	IF/WB	1:400/1:1000
TRA98	Rat	Abcam	ab82527	IF/WB	1:500
POL II (phospho S5)	Rabbit	Abcam	ab5131	WB	1:100
IAP-POL	Rabbit	produced by Boulard lab		IF/WB	1:400/ 1:1000
IAP-GAG (1294/1295)	Rabbit	Produced by Eurogentec and IMB Core Facility		IF/WB	1:400/ 1:1000
NRF1	Rabbit	Abcam	ab34682	CUT&Tag	1:50
NRF1	Rabbit	Abcam	ab175932	IF	1:600
IgG control	Rabbit	Abcam	ab37415	CUT&Tag	1:50
Rabbit IgG	Guinea pig	Antibodies	ABIN101961	CUT&Tag	1:100

Supplementary material

Sequences (5’ to 3’)

1. **Full length ERVK (IAPez) LTR_Mm (used for 5.5.1 RNA FISH homemade probes):**

TGTGGGAAGCCGCCCCCACATTCGCCGTCACAAGATGGCGCTGACATCCTGTGTTCTAAGTTGGTAAA
CAAATAATCTGCGCATGAGCCAAGGGTATTTACGACTACTTGTACTCTGTTTTTCCCGTGAACGTCAGCT
CGGCCATGGGCTGCAGCCAATCAGGGAGTGATGCGCCCTAGGCAATGGTTGTTCTCTTTAAAATAGAA
GGGGTTTCGTTTTTCTCGCTCTCTTGCTTCTCTCTCTTGCTTCTCTCTCTTGCTTCTTACACTCTGGCCC
CATAAGATGTAAGCAATAAAGCTTTGCCGTAGAAGATTCTGGTTGTTGTGTTCTTCCTGGCCGGTCGTG
AGAACGCGACGAATAACAATTGGTGCCGAATTCCGGGACGAGAAAATCCGGGACGAGAAAAAACTCCG
GACTGGCGCAGGAGGGATACTTCATTTCAGAACCAGAACTACGGATCACGTTTATAAAGGTTCCCGTAA
CACAGACTGTTGAGAAGGATTCAACTACCGAATTCAGAACTCATCAGCTGGGGAACGACGGTGATAAA
GGTTCCCGTAAAGCAGACTGTTAAGAAGGATTCAACTGTATGAATTCAGAACTTTTCAGCTGGGGAACG
AGAGTACCAGTGAGTATGTTTGGCCTTGAATTTTTTCTGGTGTTAGGAGCCCTTTTGTTCCTTTTCACAT
GTTATATAGTGGTTAAGGCAGGGCTGAAAATTCTGGATGAAATTCAGGGCAGTCTATCAGAAGTAAAGC
GGGGAGAGAGAGTAGGAGCAAGGAGAAACGGTAAGTATACAGGCCTTTCCAAGGGTCTTGAACCCGA
GGAAAAGTTAAGGTTAGGTAGGAATACCTGGAGAGAGATTAGAAGAAAAAGAGGAAAAAGGGAAAAGA

AAAAAGATCGATTAGCGGAGGTCTCTAGGAGATACTCGTCACTAGATGAGCTCAGGAAGCCAGCTCTTA
GTAGCTCTGAAGCAAGTGAAGAATCCTCCTCTGAGGAAACAGACTGGGAGGAAGAAGCAGCCCATTAC
CAGCCAGCTAATTGGTCAAGAAAAAAGCCAAAAGCGGCTGGCGAAAGTCAGCGTACTGTTCAACCTCC
CGGCAGTCGGTTTCAAGGTCCGCCCTATGCGGAGCCCCCGCCCTGCGTAGTGCGTCAGCAATGCGCA
GAGAGGCAATGCGCAGAGAGGTGCGCAGAGAGGCAGTGCGCAGACAGGTGCGCAGAGAGGCAGTG
CGCAGAGAGGCAGTGCGCAGACTCATTCATTCCCCGAGAGGAACAAAAGAAAATAGAACAGGCATTTC
CAGTCTTTGAAGGAGCCGAGGGTGGGCGTGTCCACGCTCCGGTAGAATACGTACAGATTAAGGAAATT
GCCGAGTCGGTTCGTAAATACGGAACCAATGCTAATTTCACCTTGGTGCAGTTAGACAGGCTCGCTGGT
ATGGCACTAACGCCTGCTGATTGGCAGACGGTTGTAAAAGCCGCTCTTCCTAGTATGGGCAAATATATG
GAATGGAAAGCGCTTTGGCACGAAGCTGCACAGGCGCAGGCCCGAGCAAACGCAGCTGCTTTGACTC
CAGAGCAGAGAGATTGGACTTTTGACTTGTTAACGGGTCAGGGAGCTTATTCTGCTGATCAGACAAACT
ACCATTGGGGAGCTTATGCCCAGATTTCTTCCACGGCTATTAGGGCCTGGAAGGCGCTCTCTCGAGCA
GGTGAAACCACTGGTCAGTTAACAAAAATAATCCAGGGACCTCAGGAATCTTTCTCAGATTTTGCGGCC
AGAATGACAGAGGCAGCAGAGCGTATTTTTGGAGAGCCAGAGCAAGCTGCGCCTCTCATAGAACAGCT
AATCTACGAGCAAGCCACAAAGGAGTGCCGAGCGGCCATAGCCCCAAGAAAGAACAAAGGCTTACAAG
ACTGGCTCAGGGTCTGTCGAGAGCTTGGGGGACCCCTTAGCAATGCAGGTTTAGCGGCTGCCATCCTT
CAATCCCAAAACCGCTCCATGGGCAGAAATGATCAGAGGACATGTTTTAACTGCGGAAAGCCTGGGCA
TTTTAAGAAAGATTGCAGAGCTCCAGATAAACAGGGAGGGACTCTCACTCTTTGCTCTAAGTGTGGCAA
GGGTTATCATAGAGCTGACCAGTGTCGCTCTGTGAGGGATATAAAGGGCAGAATTCTTCCCCCACCTGA
TAGTCAATCAGCTGATGTGCCAAAAAACGGGTCACCGGGCCCTCGGTCCCAGGGCCCTCAAAGATATG
GGAACCGGTTTGTCAGGACCCAGGAAGCAGTCAGAGAGACGACCCAGGAAGACCCACAAGGGTGGA
CCTGCGTGCCGCCTCCGACTTCCTATTAATGCCTCAAATGAGTATTCAGCCGGTGCCGGTGGAGCCTAT
ACCATCCTTGCCCCCGGGAACCATGGGCCTTATTCTCGGCCGAGGTTCACTCACCTTGCAGGGCTTAG
TAGTCCACCCTGGAATTATGGATTGTCAACATTCCCCTGAAATACAGGTTCTGTGCTCAAGCCCTAAAGG
CGTTCTTTCTATTAGTAAAGGAGATAGGATAGCTCAGCTGCTGCTCCTCCCTGATAATACCAGGGAGAAA
TCTGCAGGACCTGAGATAAAGAAAATGGGCTCCTCAGGAAATGATTCTGCCTATTTGGTTGTATCTTTAA
ATGATAGACCTAAGCTCCGCCTTAAGATTAATGGAAAAGAGTTTGAAGGCATCCTTGATACCGGAGCAG
ATAAAAGTATAATTTCTACACATTGGTGGCCCAAAGCATGGCCCACCACAGAGTCACCTCATTCATTACA
GGGCCTAGGATATCAATCATGTCCCACTATAAGCTCCGTTGCCTTGACGTGGGAATCCTCTGAAGGGCA
GCAAGGGAAATTCATACCTTATGTGCTCCCACTCCCGGTTAACCTCTGGGGAAGGGATATTATGCAGCA
TTTGGGCCTTATTTTGTCCAATGAAAACGCCCCATCAGGAGGGTATTCAGCTAAAGCAAAAAATGTCATG
GCAAAGATGGGTTATAAAGAAGGAAGAGGGTTAGGACATCAAGAACAGGGAAGGATAGAGCCCATCTC
ACCTAATGGAAACCAAGACAGACAGGGTCTGGGTTTTCCATAGCGGCCATTGGGGCAGCACGACCCAT
ACCATGGAAAACAGGGGACCCAGTGTGGGTTCCTCAATGGCACCTATCCTCTGAAAAACTAGAAGCTG
TGATTCAACTGGTAGAGGAACAATTAAAATTAGGCCATATTGAACCCTCTACCTCACCTTGGAATACTCC
AATTTTTGTAATTAAGAAAAAGTCAGGAAAGTGGAGACTGCTCCATGACCTCAGAGCCATTAATGAGCAA
ATGAACTTATTTGGCCCAGTACAGAGGGGTCTCCCTGTATTTTCCGCCTTACCACGTGGCTGGAATTTAA
TCATTATAGATATTAAAGATTGTTTCTTTTCTATACCTTTGTGTCCAAGGGATAGGCCCAGATTTGCCTTTA
CCATCCCCTCTATTAATCACATGGAACCTGATAAGAGGTATCAATGGAAGGTCTTACCACAGGGAATGTC
CAATAGTCCTACTATGTGTCAACTTTATGTACAAGAAGCTCTTTTGCCAGTGAGGGAACAATTCCCCTCT
TTAATTTTGCTCCTTTACATGGATGACATCCTCCTGTGCCATAAAGACCTTACCATGCTACAAAAGGCATA
TCCTTTTCTACTTAAAACTTTAAGTCAGTGGGGTCTACAGATAGCCACAGAAAAGGTCCAAATTTCTGAT
ACAGGACAATTCTTGGGCTCTGTGGTGTCCCCAGATAAGATTGTGCCCCAAAAGGTAGAGATAAGAAGA
GATCACCTCCATACCTTAAATGATTTTCAAAAGCTGTTGGGAGATATTAATTGGCTCAGACCCTTTTTAAA
GATTCCTTCTGCTGAATTAAGGCCTTTGTTTAGTATTTTAGAAGGAGATCCTCATATCTCCTCCCCTAGGA
CTCTTACTCTAGCTGCTAACCAGGCCTTACAAAAAGTGGAAAAAGCCTTACAGAATGCACAATTACAACG
TATTGAGGATTCGCAGCCTTTCAGTTTGTGTGTCTTTAGGACAGCACAATTGCCAACTGCAGTTTTGTG
GCAAAGTGGGCCATTGTTGTGGATCCATCCAAACGTATCCCCAGCTAAAATAATAGATTGGTATCCTGAT
GCAATTGCACAGCTTGCCCTTAAAGGCCTAAAAGCAGCAATCACCCACTTTGGGCAAAGTCCATATCTT
TTAATTGTACCTTATACCGCTGCACAGGTTCAAACCTTGGCAGCCACATCTAATGATTGGGCAGTTTTAG
TTACCTCCTTTTCAGGAAAAATAGGTAACCATTATCCAAAACATCCAATCTTACAGTTTGCCCAAAATCAA
TCTGTTGTGTTTCCACAAATAACAGTAAGAAACCCACTTAAAAATGGGATTGTGGTATATACTGATGGATC
AAAAACTGGCATAGGTGCCTATGTGGCTAATGGTAAAGTGGTATCCAAACAATATAATGAAAATTCACCTC
AAGTGGTAGAATGTTTAGTGGTCTTAGAAGTTTTAAAAACCTTTTTAGAACCCCTTAATATTGTGTCAGAT
TCCTGTTATGTGGTAAATGCAGTAAATCTTTTAGAAGTGGCTGGAGTGATTAAGCCTTCCAGTAGAGTTG

CCAATATTTTTCAGCAGATACAATTAGTTTTGTTATCTAGAAGATCTCCTGTTTATATTACTCATGTTAGAGC
CCATTCAGGCCTACCTGGCCCCATGGCTCTGGGAAATGATTTGGCAGATAAGGCCACTAAAGTGGTGG
CTGCTGCCCTATCATCCCCGGTAGAGGCTGCAAGAAATTTTCATAACAATTTTCATGTGACGGCTGAAAC
ATTACGCAGTCGTTTCTCCTTGACAAGAAAAGAAGCCCGTGACATTGTTACTCAATGTCAAAGCTGCTG
TGAGTTCTTGCCAGTTCCTCATGTGGGAATTAACCCACGCGGTATTCGACCTCTACAGGTCTGGCAAAT
GGATGTTACACATGTTTCTTCCTTTGGAAAACTTCAATATCTCCATGTGTCCATTGACACATGTTCTGGCA
TCATGTTTGCCTCTCCGTTAACCGGAGAAAAAGCCTCACATGTGATTCAACATTGTCTTGAGGCATGGA
GTGCTTGGGGGAAACCCAGATTCCTTAAGACTGATAATGGACCAGCTTATACATCCCAAAAATTCCAACA
GTTCTGCCGTCAGATGGACGTGACCCACCTGACTGGACTTCCATACAACCCTCAAGGACAGGGTATTG
TTGAGCGTGCGCATCGCACCCTCAAAGCCTATCTTTTAAAACAGAAGAGGGGAACTTTTGAGGAGACT
GTACCCCGAGCACCAAGAGTGTCGGTGTCTTTGGCACTCTTTACACTCAATTTTTTAAATATTGATGCTC
ATGGCCATACTGCGGCTGAACGTCATTGTTCAGAGCCAGATAGGCCCAATGAGATGGTTAAATGGAAAA
ATGTCCTTGATAATAAATGGTATGGCCCGGATCCTATCTTGATAAGATCCAGGGGAGCTATCTGTGTTTTC
CCACAGAATGAAGACAACCCATTTTGGGTACCAGAAAGACTCACCCGAAAAATCCAGACTGACCAAGG
GAATACTAATGTCCCTCGTCTTGGTGATGTCCAGGGCGTCAATAATAAAGAGAGAGCAGCGTTGGGGG
ATAATGTCGACATTTCCACTCCCAATGACGGTGATGTATAATGCTCAAGTATTCTCCTGCTTTTTTACCAC
TAACTAGGAACTGGGTTTGGCCTTAATTCAGACAGCCTTGGCTCTGTCCGGACAGGTCCAGACAACTG
ACACCATTAACACTTTGTCAGCCTCAGTGACTACAGTCATAGATGAACAGGCCTCAGCTAATGTCAAGAT
ACAGAGAGGTCTCATGCTGGTTAATCAACTCATAGATCTTGTCCAGAAACAACTAGATGTATTATGACAA
ATAACTCAGCAGGGATGTGAACAAAAGTTTCCGGGATTGTGTGTTATTTCCATTCAGTATGTTAAATTTAC
TAGGACAGCTAATTTGTCAAAAAGTCTTTTTCAGTATATGTTACAGAATTGGATGGCTGAATTTGAATAGA
TCCTTCGAGAATTGAGACTTCAGGTCAACTCCACGCGCTTGGACCTGTCGCTGACCAAAGGATTACCC
AATTGGATCTCCTCAGCATTTTCTTTCTTTAAAAAATTGGGTGGGATTAATATTATTTGGAGATACACTTTG
CTGTGGATTAGTGTTGCTTCTTTGATTGGTCTGTAAGCTTAAGGCCCAAACTAAGAGAGACAAGGTGGT
TATTGCCCAGGCGCTTGCAGGACTAGAACATGGAGCTTCCCCTGATATATCTATGCTTAAGCAATAGGTC
GCTGGCCACTCAGCTCTTATATCTCACGAGGCTAGTCTCATTGCACGAGATAGAGTGAGTGTGCTTCAG
CAGCCCGAGAGAGTTGCAGGGCTAAGCACTGCAGTAGAAGGGCTCTGCGGCACATATGAGCCTATTCT
AGGGAGACATGTCATCTTTCATGAAGGTTCAGTGTCCTAGTTCCCTTCCCCCAGGCAAAACGACACGG
GAGCAGGTCAGGGTTGCTCTGGGTAAAAGCCTGTAAGCCTAAGAGCTAATCCTGTACATGGCTCCTTTA
CCTACACACTGGGGATTTGACCTCTATCTCCACTCTCATCAATATGGGTGGCCTATTTGCTCCTATTAAAA
GAAAAAGGGGGAACTGTGGGAAGCCGCCCCCACATTCGCCGTCACAAGATGGCGCTGACATCCTGTG
TTCTAAGTTGGTAAACAAATAATCTGCGCATGAGCCAAGGGTATTTACGACTACTTGTACTCTGTTTTTCC
CGTGAACGTCAGCTCGGCCATGGGCTGCAGCCAATCAGGGAGTGATGCGCCCTAGGCAATGGTTGTT
CTCTTTAAAATAGAAGGGGTTTCGTTTTTCTCGCTCTCTTGCTTCTCTCTCTTGCTTCTCTCTCTTGCTTC
TTACACTCTGGCCCCATAAGATGTAAGCAATAAAGCTTTGCCGTAGAAGATTCTGGTTGTTGTGTTCTTC
CTGGCCGGTCGTGAGAACGCGACGAATAA

2. **GAG epitope (used for [5.7 IAP-GAG antibody generation]):**

TTAACGGGTCAGGGAGCTTATTCTGCTGATCAGACAAACTACCATTGGGGAGCTTATGCCCAGATTTCT
TCCACGGCTATTAGGGCCTGGAAGGCGCTCTCTCGAGCAGGTGAAACCACTGGTCAGTTAACAAAAAT
AATCCAGGGACCTCAGGAATCTTTCTCAGATTTTGCGGCCAGAATGACAGAGGCAGCAGAGCGTATTTT
TGGAGAGCCAGAGCAAGCTGCGCCTCTCATAGAACAGCTAATCTACGAGCAAGCCACAAAGGAGTGCC
GAGCGGCCATAGCCCCAAGAAAGAACAAAGGCTTACAAGACTGGCTCAGGGTCTGTCGAGAGCTTGG
GGGACCCCTTAGCAATGCAGGTTTATAATGA

References

Aagaard, L., G. Laible, P. Selenko, M. Schmid, R. Dorn, G. Schotta, S. Kuhfittig, et al. 1999. “Functional Mammalian Homologues of the Drosophila PEV-Modifier Su(var)3-9 Encode Centromere-Associated Proteins Which Complex with the Heterochromatin Component M31.” *The EMBO Journal* 18 (7): 1923–38.

Ahel, Josip, Aparna Pandey, Michaela Schwaiger, Fabio Mohn, Anja Basters, Georg Kempf, Aude Andriollo, Lucas Kaaij, Daniel Hess, and Marc Bühler. 2024. “ChAHP2 and ChAHP Control Diverse Retrotransposons by Complementary Activities.” *Genes & Development* 38 (11-12): 554–68.

Ahmadzadeh, Elham, N. Sumru Bayin, Xinli Qu, Aditi Singh, Linda Madisen, Daniel Stephen, Hongkui Zeng, Alexandra L. Joyner, and Alberto Rosello-Diez. 2020. “A Collection of Genetic Mouse Lines and Related Tools for Inducible and Reversible Intersectional Mis-Expression.” *Development (Cambridge, England)* 147 (10): dev186650.

Alexandrova, Elena A., Ivan A. Olovnikov, Galina V. Malakhova, Anastasia A. Zabolotneva, Maria V. Suntsova, Sergey E. Dmitriev, and Anton A. Buzdin. 2012. “Sense Transcripts Originated from an Internal Part of the Human Retrotransposon LINE-1 5’ UTR.” *Gene* 511 (1): 46–53.

Andrews, Simon, Christel Krueger, Maravillas Mellado-Lopez, Myriam Hemberger, Wendy Dean, Vicente Perez-Garcia, and Courtney W. Hanna. 2023. “Mechanisms and Function of de Novo DNA Methylation in Placental Development Reveals an Essential Role for DNMT3B.” *Nature Communications* 14 (1): 371.

Aravin, Alexei A., Ravi Sachidanandam, Deborah Bourc’his, Christopher Schaefer, Dubravka Pezic, Katalin Fejes Toth, Timothy Bestor, and Gregory J. Hannon. 2008. “A piRNA Pathway Primed by Individual Transposons Is Linked to de Novo DNA Methylation in Mice.” *Molecular Cell* 31 (6): 785–99.

Aravin, Alexei A., Ravi Sachidanandam, Angelique Girard, Katalin Fejes-Toth, and Gregory J. Hannon. 2007. “Developmentally Regulated piRNA Clusters Implicate MILI in Transposon Control.” *Science (New York, N.Y.)* 316 (5825): 744–47.

Ardeljan, Daniel, Xuya Wang, Mehrnoosh Oghbaie, Martin S. Taylor, David Husband, Vikram Deshpande, Jared P. Steranka, et al. 2020. “LINE-1 ORF2p Expression Is Nearly Imperceptible in Human Cancers.” *Mobile DNA* 11 (1): 1.

Arrigucci, Riccardo, Yuri Bushkin, Felix Radford, Karim Lakehal, Pooja Vir, Richard Pine, December Martin, et al. 2017. “FISH-Flow, a Protocol for the Concurrent Detection of mRNA and Protein in Single Cells Using Fluorescence in Situ Hybridization and Flow Cytometry.” *Nature Protocols* 12 (6): 1245–60.

Asimi, Vahid, Abhishek Sampath Kumar, Henri Niskanen, Christina Riemenschneider, Sara Hetzel, Julian Naderi, Nina Fasching, et al. 2022. “Hijacking of Transcriptional Condensates by Endogenous Retroviruses.” *Nature Genetics* 54 (8): 1238–47.

Athanikar, Jyoti N., Richard M. Badge, and John V. Moran. 2004. “A YY1-Binding Site Is Required for Accurate Human LINE-1 Transcription Initiation.” *Nucleic Acids Research* 32 (13): 3846–55.

Auclair, Ghislain, Sylvain Guibert, Ambre Bender, and Michael Weber. 2014. “Ontogeny of CpG Island Methylation and Specificity of DNMT3 Methyltransferases during Embryonic Development in the Mouse.” *Genome Biology* 15 (12): 545.

Bannister, Andrew J., and Tony Kouzarides. 2011. “Regulation of Chromatin by Histone Modifications.” *Cell Research* 21 (3): 381–95.

Bao, Jianqiang, and Wei Yan. 2012. “Male Germline Control of Transposable Elements.” *Biology of Reproduction* 86 (5): 162, 1–14.

Barau, Joan, Aurélie Teissandier, Natasha Zamudio, Stéphanie Roy, Valérie Nalesso, Yann Hérault, Florian Guillou, and Déborah Bourc’his. 2016. “The DNA Methyltransferase DNMT3C Protects Male Germ Cells from Transposon Activity.” *Science (New York,*

N.Y.) 354 (6314): 909–12.
Beck, Christine R., José Luis Garcia-Perez, Richard M. Badge, and John V. Moran. 2011. "LINE-1 Elements in Structural Variation and Disease." *Annual Review of Genomics and Human Genetics* 12 (1): 187–215.
Becker, K. G., G. D. Swergold, K. Ozato, and R. E. Thayer. 1993. "Binding of the Ubiquitous Nuclear Transcription Factor YY1 to a Cis Regulatory Sequence in the Human LINE-1 Transposable Element." *Human Molecular Genetics* 2 (10): 1697–1702.
Belancio, Victoria P., Megan Whelton, and Prescott Deininger. 2007. "Requirements for Polyadenylation at the 3' End of LINE-1 Elements." *Gene* 390 (1-2): 98–107.
Belgnaoui, S. Mehdi, Roger G. Gosden, O. John Semmes, and Abdelali Haoudi. 2006. "Human LINE-1 Retrotransposon Induces DNA Damage and Apoptosis in Cancer Cells." *Cancer Cell International* 6 (1): 13.
Bellvé, A. R., J. C. Cavicchia, C. F. Millette, D. A. O'Brien, Y. M. Bhatnagar, and M. Dym. 1977. "Spermatogenic Cells of the Prepuberal Mouse. Isolation and Morphological Characterization." *The Journal of Cell Biology* 74 (1): 68–85.
Ben-Hattar, J., and J. Jiricny. 1988. "Methylation of Single CpG Dinucleotides within a Promoter Element of the Herpes Simplex Virus Tk Gene Reduces Its Transcription in Vivo." *Gene* 65 (2): 219–27.
Bernstein, Bradley E., Tarjei S. Mikkelsen, Xiaohui Xie, Michael Kamal, Dana J. Huebert, James Cuff, Ben Fry, et al. 2006. "A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells." *Cell* 125 (2): 315–26.
Bestor, Timothy H. 2003. "Cytosine Methylation Mediates Sexual Conflict." *Trends in Genetics: TIG* 19 (4): 185–90.
Bestor, T., A. Laudano, R. Mattaliano, and V. Ingram. 1988. "Cloning and Sequencing of a cDNA Encoding DNA Methyltransferase of Mouse Cells. The Carboxyl-Terminal Domain of the Mammalian Enzymes Is Related to Bacterial Restriction Methyltransferases." *Journal of Molecular Biology* 203 (4): 971–83.
Biggin, Mark D. 2011. "Animal Transcription Networks as Highly Connected, Quantitative Continua." *Developmental Cell* 21 (4): 611–26.
Bodak, Maxime, Jian Yu, and Constance Ciaudo. 2014. "Regulation of LINE-1 in Mammals." *Biomolecular Concepts* 5 (5): 409–28.
Boeke, J. D., D. J. Garfinkel, C. A. Styles, and G. R. Fink. 1985. "Ty Elements Transpose through an RNA Intermediate." *Cell* 40 (3): 491–500.
Bolcun-Filas, Ewelina, Laura A. Bannister, Alex Barash, Kerry J. Schimenti, Suzanne A. Hartford, John J. Eppig, Mary Ann Handel, Lishuang Shen, and John C. Schimenti. 2011. "A-MYB (MYBL1) Transcription Factor Is a Master Regulator of Male Meiosis." *Development (Cambridge, England)* 138 (15): 3319–30.
Bostick, Magnolia, Jong Kyong Kim, Pierre-Olivier Estève, Amander Clark, Sriharsa Pradhan, and Steven E. Jacobsen. 2007. "UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells." *Science (New York, N.Y.)* 317 (5845): 1760–64.
Boulard, Mathieu, Sofia Rucli, John R. Edwards, and Timothy H. Bestor. 2020. "Methylation-Directed Glycosylation of Chromatin Factors Represses Retrotransposon Promoters." *Proceedings of the National Academy of Sciences of the United States of America* 117 (25): 14292–98.
Bouras, Ahmed, Melanie Leone, Valerie Bonadona, Marine Lebrun, Alain Calender, and Nadia Boutry-Kryza. 2021. "Identification and Characterization of New Alu Element Insertion in the BRCA1 Exon 14 Associated with Hereditary Breast and Ovarian Cancer." *Genes* 12 (11): 1736.
Bourc'his, Déborah, and Timothy H. Bestor. 2004. "Meiotic Catastrophe and Retrotransposon Reactivation in Male Germ Cells Lacking Dnmt3L." *Nature* 431 (7004): 96–99.
Bourc'his, D., G. L. Xu, C. S. Lin, B. Bollman, and T. H. Bestor. 2001. "Dnmt3L and the Establishment of Maternal Genomic Imprints." *Science (New York, N.Y.)* 294 (5551): 2536–39.
Bourque, Guillaume, Kathleen H. Burns, Mary Gehring, Vera Gorbunova, Andrei Seluanov,

Molly Hammell, Michaël Imbeault, et al. 2018. “Ten Things You Should Know about Transposable Elements.” *Genome Biology* 19 (1): 199.
Boyer, Laurie A., Kathrin Plath, Julia Zeitlinger, Tobias Brambrink, Lea A. Medeiros, Tong Ihn Lee, Stuart S. Levine, et al. 2006. “Polycomb Complexes Repress Developmental Regulators in Murine Embryonic Stem Cells.” *Nature* 441 (7091): 349–53.
Brouha, Brook, Joshua Schustak, Richard M. Badge, Sheila Lutz-Prigge, Alexander H. Farley, John V. Moran, and Haig H. Kazazian Jr. 2003. “Hot L1s Account for the Bulk of Retrotransposition in the Human Population.” *Proceedings of the National Academy of Sciences of the United States of America* 100 (9): 5280–85.
Bulut-Karslioglu, Aydan, Inti A. De La Rosa-Velázquez, Fidel Ramirez, Maxim Barenboim, Megumi Onishi-Seebacher, Julia Arand, Carmen Galán, et al. 2014. “Suv39h-Dependent H3K9me3 Marks Intact Retrotransposons and Silences LINE Elements in Mouse Embryonic Stem Cells.” *Molecular Cell* 55 (2): 277–90.
Carmell, Michelle A., Angélique Girard, Henk J. G. van de Kant, Deborah Bourc’his, Timothy H. Bestor, Dirk G. de Rooij, and Gregory J. Hannon. 2007. “MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline.” *Developmental Cell* 12 (4): 503–14.
Castro-Diaz, Nathaly, Gabriela Ecco, Andrea Coluccio, Adamandia Kapopoulou, Benyamin Yazdanpanah, Marc Friedli, Julien Duc, Suk Min Jang, Priscilla Turelli, and Didier Trono. 2014. “Evolutionally Dynamic L1 Regulation in Embryonic Stem Cells.” *Genes & Development* 28 (13): 1397–1409.
Chelmicki, Tomasz, Emeline Roger, Aurélie Teissandier, Mathilde Dura, Lorraine Bonneville, Sofia Rucli, François Dossin, Camille Fouassier, Sonia Lameiras, and Deborah Bourc’his. 2021. “m6A RNA Methylation Regulates the Fate of Endogenous Retroviruses.” *Nature* 591 (7849): 312–16.
Chen, Mingyue, Xiaolong Huang, Chunlei Wang, Shibo Wang, Lei Jia, and Lin Li. 2024. “Endogenous Retroviral Solo-LTRs in Human Genome.” *Frontiers in Genetics* 15 (March):1358078.
Chen, Zhiyuan, Qiangzong Yin, Azusa Inoue, Chunxia Zhang, and Yi Zhang. 2019. “Allelic H3K27me3 to Allelic DNA Methylation Switch Maintains Noncanonical Imprinting in Extraembryonic Cells.” *Science Advances* 5 (12): eaay7246.
Chuma, Shinichiro, and Toru Nakano. 2013. “piRNA and Spermatogenesis in Mice.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 368 (1609): 20110338.
Chuong, Edward B., Nels C. Elde, and Cédric Feschotte. 2017. “Regulatory Activities of Transposable Elements: From Conflicts to Benefits.” *Nature Reviews. Genetics* 18 (2): 71–86.
Cirillo, L. A., and K. S. Zaret. 1999. “An Early Developmental Transcription Factor Complex That Is More Stable on Nucleosome Core Particles than on Free DNA.” *Molecular Cell* 4 (6): 961–69.
Coluccio, Andrea, Gabriela Ecco, Julien Duc, Sandra Offner, Priscilla Turelli, and Didier Trono. 2018. “Individual Retrotransposon Integrants Are Differentially Controlled by KZFP/KAP1-Dependent Histone Methylation, DNA Methylation and TET-Mediated Hydroxymethylation in Naïve Embryonic Stem Cells.” *Epigenetics & Chromatin* 11 (1): 7.
Condemi, Livia, Ivano Mocavini, Sergi Aranda, and Luciano Di Croce. 2024. “Polycomb Function in Early Mouse Development.” *Cell Death and Differentiation*, July, 1–10.
Cordaux, Richard, and Mark A. Batzer. 2009. “The Impact of Retrotransposons on Human Genome Evolution.” *Nature Reviews. Genetics* 10 (10): 691–703.
Costello, Alan, Nga T. Lao, Clair Gallagher, Berta Capella Roca, Lourdes A. N. Julius, Srinivas Suda, Jens Ducrée, et al. 2019. “Leaky Expression of the TET-on System Hinders Control of Endogenous miRNA Abundance.” *Biotechnology Journal* 14 (3): e1800219.
Cost, Gregory J., Qinghua Feng, Alain Jacquier, and Jef D. Boeke. 2002. “Human L1 Element Target-Primed Reverse Transcription in Vitro.” *The EMBO Journal* 21 (21): 5899–5910.

Coufal, Nicole G., José L. Garcia-Perez, Grace E. Peng, Gene W. Yeo, Yangling Mu, Michael T. Lovci, Maria Morell, K. Sue O'Shea, John V. Moran, and Fred H. Gage. 2009. "L1 Retrotransposition in Human Neural Progenitor Cells." *Nature* 460 (7259): 1127–31.
Crichton, James H., Donncha S. Dunican, Marie Maclennan, Richard R. Meehan, and Ian R. Adams. 2014. "Defending the Genome from the Enemy within: Mechanisms of Retrotransposon Suppression in the Mouse Germline." *Cellular and Molecular Life Sciences: CMLS* 71 (9): 1581–1605.
Curcio, M. Joan, Sheila Lutz, and Pascale Lesage. 2015. "The Ty1 LTR-Retrotransposon of Budding Yeast, Saccharomyces Cerevisiae." *Microbiology Spectrum* 3 (2): MDNA3–0053 – 2014.
Czermin, Birgit, Raffaella Melfi, Donna McCabe, Volker Seitz, Axel Imhof, and Vincenzo Pirrotta. 2002. "Drosophila Enhancer of Zeste/ESC Complexes Have a Histone H3 Methyltransferase Activity That Marks Chromosomal Polycomb Sites." *Cell* 111 (2): 185–96.
Dai, Zhiming, Xianhua Dai, Qian Xiang, and Jihua Feng. 2009. "Nucleosomal Context of Binding Sites Influences Transcription Factor Binding Affinity and Gene Regulation." *Genomics, Proteomics & Bioinformatics* 7 (4): 155–62.
Dantas Machado, Ana Carolina, Tianyin Zhou, Satyanarayan Rao, Pragya Goel, Chaitanya Rastogi, Allan Lazarovici, Harmen J. Bussemaker, and Remo Rohs. 2015. "Evolving Insights on How Cytosine Methylation Affects Protein-DNA Binding." *Briefings in Functional Genomics* 14 (1): 61–73.
Das, Atze T., Liliane Tenenbaum, and Ben Berkhout. 2016. "Tet-on Systems for Doxycycline-Inducible Gene Expression." *Current Gene Therapy* 16 (3): 156–67.
Das, Atze T., Xue Zhou, Stefan W. Metz, Monique A. Vink, and Ben Berkhout. 2016. "Selecting the Optimal Tet-On System for Doxycycline-Inducible Gene Expression in Transiently Transfected and Stably Transduced Mammalian Cells." *Biotechnology Journal* 11 (1): 71–79.
DeBerardinis, R. J., J. L. Goodier, E. M. Ostertag, and H. H. Kazazian Jr. 1998. "Rapid Amplification of a Retrotransposon Subfamily Is Evolving the Mouse Genome." *Nature Genetics* 20 (3): 288–90.
DeBerardinis, R. J., and H. H. Kazazian Jr. 1999. "Analysis of the Promoter from an Expanding Mouse Retrotransposon Subfamily." *Genomics* 56 (3): 317–23.
De Luca, Chiara, Anuj Gupta, and Alex Bortvin. 2023. "Retrotransposon LINE-1 Bodies in the Cytoplasm of piRNA-Deficient Mouse Spermatocytes: Ribonucleoproteins Overcoming the Integrated Stress Response." *PLoS Genetics* 19 (6): e1010797.
Deniz, Özgen, Lorenzo de la Rica, Kevin C. L. Cheng, Dominik Spensberger, and Miguel R. Branco. 2018. "SETDB1 Prevents TET2-Dependent Activation of IAP Retroelements in Naïve Embryonic Stem Cells." *Genome Biology* 19 (1): 6.
Dewannieux, Marie, Anne Dupressoir, Francis Harper, Gérard Pierron, and Thierry Heidmann. 2004. "Identification of Autonomous IAP LTR Retrotransposons Mobile in Mammalian Cells." *Nature Genetics* 36 (5): 534–39.
Dewannieux, Marie, Cécile Esnault, and Thierry Heidmann. 2003. "LINE-Mediated Retrotransposition of Marked Alu Sequences." *Nature Genetics* 35 (1): 41–48.
Dias Mirandela, Madeleine, Ansgar Zoch, Jessica Leismann, Shaun Webb, Rebecca V. Berrens, Devisree Valsakumar, Yuka Kabayama, et al. 2024. "Two-Factor Authentication Underpins the Precision of the piRNA Pathway." *Nature*, September, 1–7.
Di Giacomo, Monica, Stefano Comazzetto, Harpreet Saini, Serena De Fazio, Claudia Carrieri, Marcos Morgan, Lina Vasiliauskaite, Vladimir Benes, Anton J. Enright, and Dónal O'Carroll. 2013. "Multiple Epigenetic Mechanisms and the piRNA Pathway Enforce LINE1 Silencing during Adult Spermatogenesis." *Molecular Cell* 50 (4): 601–8.
Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics (Oxford, England)* 29 (1): 15–21.
Dodge, Jonathan E., Yong-Kook Kang, Hideyuki Beppu, Hong Lei, and En Li. 2004. "Histone H3-K9 Methyltransferase ESET Is Essential for Early Development." *Molecular and*

Cellular Biology 24 (6): 2478–86.
Domcke, Silvia, Anaïs Flore Bardet, Paul Adrian Ginno, Dominik Hartl, Lukas Burger, and Dirk Schübeler. 2015. “Competition between DNA Methylation and Transcription Factors Determines Binding of NRF1.” *Nature* 528 (7583): 575–79.
Doucet, Aurélien J., Jeremy E. Wilusz, Tomoichiro Miyoshi, Ying Liu, and John V. Moran. 2015. “A 3’ poly(A) Tract Is Required for LINE-1 Retrotransposition.” *Molecular Cell* 60 (5): 728–41.
Drumond, Ana Luiza, Marvin L. Meistrich, and Hélio Chiarini-Garcia. 2011. “Spermatogonial Morphology and Kinetics during Testis Development in Mice: A High-Resolution Light Microscopy Approach.” *Reproduction* 142 (1): 145–55.
Dupressoir, Anne, Cécile Vernochet, Olivia Bawa, Francis Harper, Gérard Pierron, Paule Opolon, and Thierry Heidmann. 2009. “Syncytin-A Knockout Mice Demonstrate the Critical Role in Placentation of a Fusogenic, Endogenous Retrovirus-Derived, Envelope Gene.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (29): 12127–32.
Dura, Mathilde, Aurélie Teissandier, Mélanie Armand, Joan Barau, Clémentine Lapoujade, Pierre Fouchet, Lorraine Bonneville, et al. 2022. “DNMT3A-Dependent DNA Methylation Is Required for Spermatogonial Stem Cells to Commit to Spermatogenesis.” *Nature Genetics* 54 (4): 469–80.
Duttke, Sascha H., Carlos Guzman, Max Chang, Nathaniel P. Delos Santos, Bayley R. McDonald, Jialei Xie, Aaron F. Carlin, Sven Heinz, and Christopher Benner. 2024. “Position-Dependent Function of Human Sequence-Specific Transcription Factors.” *Nature* 631 (8022): 891–98.
Ecco, Gabriela, Michael Imbeault, and Didier Trono. 2017. “KRAB Zinc Finger Proteins.” *Development (Cambridge, England)* 144 (15): 2719–29.
Ehrlich, M., M. A. Gama-Sosa, L. H. Huang, R. M. Midgett, K. C. Kuo, R. A. McCune, and C. Gehrke. 1982. “Amount and Distribution of 5-Methylcytosine in Human DNA from Different Types of Tissues of Cells.” *Nucleic Acids Research* 10 (8): 2709–21.
Ernst, Christina, Nils Eling, Celia P. Martinez-Jimenez, John C. Marioni, and Duncan T. Odom. 2019. “Staged Developmental Mapping and X Chromosome Transcriptional Dynamics during Mouse Spermatogenesis.” *Nature Communications* 10 (1): 1251.
Fayomi, Adetunji P., and Kyle E. Orwig. 2018. “Spermatogonial Stem Cells and Spermatogenesis in Mice, Monkeys and Men.” *Stem Cell Research* 29 (May):207–14.
Feldmann, Angelika, Robert Ivanek, Rabih Murr, Dimos Gaidatzis, Lukas Burger, and Dirk Schübeler. 2013. “Transcription Factor Occupancy Can Mediate Active Turnover of DNA Methylation at Regulatory Regions.” *PLoS Genetics* 9 (12): e1003994.
Feng, Q., J. V. Moran, H. H. Kazazian Jr, and J. D. Boeke. 1996. “Human L1 Retrotransposon Encodes a Conserved Endonuclease Required for Retrotransposition.” *Cell* 87 (5): 905–16.
Ficz, Gabriella, Timothy A. Hore, Fátima Santos, Heather J. Lee, Wendy Dean, Julia Arand, Felix Krueger, et al. 2013. “FGF Signaling Inhibition in ESCs Drives Rapid Genome-Wide Demethylation to the Epigenetic Ground State of Pluripotency.” *Cell Stem Cell* 13 (3): 351–59.
Fukuda, Kei, Akihiko Okuda, Kosuke Yusa, and Yoichi Shinkai. 2018. “A CRISPR Knockout Screen Identifies SETDB1-Target Retroelement Silencing Factors in Embryonic Stem Cells.” *Genome Research* 28 (6): 846–58.
Fulco, Charles P., Joseph Nasser, Thouis R. Jones, Glen Munson, Drew T. Bergman, Vidya Subramanian, Sharon R. Grossman, et al. 2019. “Activity-by-Contact Model of Enhancer-Promoter Regulation from Thousands of CRISPR Perturbations.” *Nature Genetics* 51 (12): 1664–69.
Gallardo, Teresa, Lane Shirley, George B. John, and Diego H. Castrillon. 2007. “Generation of a Germ Cell-Specific Mouse Transgenic Cre Line, Vasa-Cre.” *Genesis (New York, N.Y.: 2000)* 45 (6): 413–17.
Gao, Xiang, Ericka R. Havecker, Pavel V. Baranov, John F. Atkins, and Daniel F. Voytas. 2003. “Translational Recoding Signals between Gag and Pol in Diverse LTR

Retrotransposons.” *RNA (New York, N.Y.)* 9 (12): 1422–30.
Garcia-Perez, Jose L., Maria Morell, Joshua O. Scheys, Deanna A. Kulpa, Santiago Morell, Christoph C. Carter, Gary D. Hammer, et al. 2010. “Epigenetic Silencing of Engineered L1 Retrotransposition Events in Human Embryonic Carcinoma Cells.” *Nature* 466 (7307): 769–73.
Gerdes, Patricia, Dorothy Chan, Mischa Lundberg, Francisco J. Sanchez-Luque, Gabriela O. Bodea, Adam D. Ewing, Geoffrey J. Faulkner, and Sandra R. Richardson. 2023. “Locus-Resolution Analysis of L1 Regulation and Retrotransposition Potential in Mouse Embryonic Development.” *Genome Research* 33 (9): 1465–81.
Ge, Steven Xijin. 2017. “Exploratory Bioinformatics Investigation Reveals Importance of ‘Junk’ DNA in Early Embryo Development.” *BMC Genomics* 18 (1): 200.
Goodier, John L., Lili Zhang, Melissa R. Vetter, and Haig H. Kazazian Jr. 2007. “LINE-1 ORF1 Protein Localizes in Stress Granules with Other RNA-Binding Proteins, Including Components of RNA Interference RNA-Induced Silencing Complex.” *Molecular and Cellular Biology* 27 (18): 6469–83.
Greenberg, Maxim V. C., and Deborah Bourc’his. 2019. “The Diverse Roles of DNA Methylation in Mammalian Development and Disease.” *Nature Reviews. Molecular Cell Biology* 20 (10): 590–607.
Grewal, Shiv I. S., and Songtao Jia. 2007. “Heterochromatin Revisited.” *Nature Reviews. Genetics* 8 (1): 35–46.
Groh, Sophia, Anna Viktoria Milton, Lisa Katherina Marinelli, Cara V. Sickinger, Angela Russo, Heike Bollig, Gustavo Pereira de Almeida, et al. 2021. “Morc3 Silences Endogenous Retroviruses by Enabling Daxx-Mediated Histone H3.3 Incorporation.” *Nature Communications* 12 (1): 5996.
Groner, Anna C., Sylvain Meylan, Angela Ciuffi, Nadine Zangger, Giovanna Ambrosini, Nicolas Dénervaud, Philipp Bucher, and Didier Trono. 2010. “KRAB-Zinc Finger Proteins and KAP1 Can Mediate Long-Range Transcriptional Repression through Heterochromatin Spreading.” *PLoS Genetics* 6 (3): e1000869.
Grow, Edward J., Ryan A. Flynn, Shawn L. Chavez, Nicholas L. Bayless, Mark Wossidlo, Daniel J. Wesche, Lance Martin, et al. 2015. “Intrinsic Retroviral Reactivation in Human Preimplantation Embryos and Pluripotent Cells.” *Nature* 522 (7555): 221–25.
Guo, Caiwei, Hyun-Hwan Jeong, Yi-Chen Hsieh, Hans-Ulrich Klein, David A. Bennett, Philip L. De Jager, Zhandong Liu, and Joshua M. Shulman. 2018. “Tau Activates Transposable Elements in Alzheimer’s Disease.” *Cell Reports* 23 (10): 2874–80.
Habibi, Ehsan, Arie B. Brinkman, Julia Arand, Leonie I. Kroeze, Hindrik H. D. Kerstens, Filomena Matarese, Konstantin Lepikhov, et al. 2013. “Whole-Genome Bisulfite Sequencing of Two Distinct Interconvertible DNA Methylomes of Mouse Embryonic Stem Cells.” *Cell Stem Cell* 13 (3): 360–69.
Haggerty, Chuck, Helene Kretzmer, Christina Riemenschneider, Abhishek Sampath Kumar, Alexandra L. Mattei, Nina Bailly, Judith Gottfreund, et al. 2021. “Dnmt1 Has de Novo Activity Targeted to Transposable Elements.” *Nature Structural & Molecular Biology* 28 (7): 594–603.
Hancks, Dustin C., and Haig H. Kazazian Jr. 2016. “Roles for Retrotransposon Insertions in Human Disease.” *Mobile DNA* 7 (1): 9.
Handel, Mary Ann, and John C. Schimenti. 2010. “Genetics of Mammalian Meiosis: Regulation, Dynamics and Impact on Fertility.” *Nature Reviews. Genetics* 11 (2): 124–36.
Hanley, Mary Beth, Woodrow Lomas, Dev Mittar, Vernon Maino, and Emily Park. 2013. “Detection of Low Abundance RNA Molecules in Individual Cells by Flow Cytometry.” *PloS One* 8 (2): e57002.
Harris, C. Jake, Marion Scheibe, Somsakul Pop Wongpalee, Wanlu Liu, Evan M. Cornett, Robert M. Vaughan, Xueqin Li, et al. 2018. “A DNA Methylation Reader Complex That Enhances Gene Transcription.” *Science (New York, N.Y.)* 362 (6419): 1182–86.
Häsler, Julien, and Katharina Strub. 2006. “Alu Elements as Regulators of Gene Expression.” *Nucleic Acids Research* 34 (19): 5491–97.

Héberlé, Éléa, and Anaïs Flore Bardet. 2019. “Sensitivity of Transcription Factors to DNA Methylation.” *Essays in Biochemistry* 63 (6): 727–41.
Hermant, Clara, and Maria-Elena Torres-Padilla. 2021. “TFs for TEs: The Transcription Factor Repertoire of Mammalian Transposable Elements.” *Genes & Development* 35 (1-2): 22–39.
Hess, Rex A., and Luiz Renato de Franca. 2008. “Spermatogenesis and Cycle of the Seminiferous Epithelium.” *Advances in Experimental Medicine and Biology* 636:1–15.
Holwerda, Sjoerd Johannes Bastiaan, and Wouter de Laat. 2013. “CTCF: The Protein, the Binding Partners, the Binding Sites and Their Chromatin Loops.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 368 (1620): 20120369.
Huang, Tien-Chi, Yi-Fang Wang, Eric Vazquez-Ferrer, Ina Theofel, Cristina E. Requena, Courtney W. Hanna, Gavin Kelsey, and Petra Hajkova. 2021. “Sex-Specific Chromatin Remodelling Safeguards Transcription in Germ Cells.” *Nature* 600 (7890): 737–42.
Hughes, Stephen H. 2015. “Reverse Transcription of Retroviruses and LTR Retrotransposons.” In *Mobile DNA III*, 3:1051–77. American Society of Microbiology.
Huo, L., and R. C. Scarpulla. 2001. “Mitochondrial DNA Instability and Peri-Implantation Lethality Associated with Targeted Disruption of Nuclear Respiratory Factor 1 in Mice.” *Molecular and Cellular Biology* 21 (2): 644–54.
Imbeault, Michaël, Pierre-Yves Helleboid, and Didier Trono. 2017. “KRAB Zinc-Finger Proteins Contribute to the Evolution of Gene Regulatory Networks.” *Nature* 543 (7646): 550–54.
Imbeault, Michael, and Didier Trono. 2014. “As Time Goes by: KRABs Evolve to KAP Endogenous Retroelements.” *Developmental Cell* 31 (3): 257–58.
Inoue, Azusa, Lan Jiang, Falong Lu, and Yi Zhang. 2017. “Genomic Imprinting of Xist by Maternal H3K27me3.” *Genes & Development* 31 (19): 1927–32.
Islam, Zeyaul, Ameena Mohamed Ali, Adviti Naik, Mohamed Eldaw, Julie Decock, and Prasanna R. Kolatkar. 2021. “Transcription Factors: The Fulcrum between Cell Development and Carcinogenesis.” *Frontiers in Oncology* 11 (June):681377.
Iwafuchi-Doi, Makiko, and Kenneth S. Zaret. 2014. “Pioneer Transcription Factors in Cell Reprogramming.” *Genes & Development* 28 (24): 2679–92.
Jachowicz, Joanna W., Xinyang Bing, Julien Pontabry, Ana Bošković, Oliver J. Rando, and Maria-Elena Torres-Padilla. 2017. “LINE-1 Activation after Fertilization Regulates Global Chromatin Accessibility in the Early Mouse Embryo.” *Nature Genetics* 49 (10): 1502–10.
Jacob, F., and J. Monod. 1961. “Genetic Regulatory Mechanisms in the Synthesis of Proteins.” *Journal of Molecular Biology* 3 (3): 318–56.
Jacobs, Frank M. J., David Greenberg, Ngan Nguyen, Maximilian Haeussler, Adam D. Ewing, Sol Katzman, Benedict Paten, Sofie R. Salama, and David Haussler. 2014. “An Evolutionary Arms Race between KRAB Zinc-Finger Genes ZNF91/93 and SVA/L1 Retrotransposons.” *Nature* 516 (7530): 242–45.
Jain, Devanshi, Cem Meydan, Julian Lange, Corentin Claeys Bouuaert, Nathalie Lailler, Christopher E. Mason, Kathryn V. Anderson, and Scott Keeney. 2017. “Rahu Is a Mutant Allele of Dnmt3c, Encoding a DNA Methyltransferase Homolog Required for Meiosis and Transposon Repression in the Mouse Male Germline.” *PLoS Genetics* 13 (8): e1006964.
Jan, Sabrina Z., Geert Hamer, Sjoerd Repping, Dirk G. de Rooij, Ans M. M. van Pelt, and Tinke L. Vormer. 2012. “Molecular Control of Rodent Spermatogenesis.” *Biochimica et Biophysica Acta* 1822 (12): 1838–50.
Jermann, Philip, Leslie Hoerner, Lukas Burger, and Dirk Schübeler. 2014. “Short Sequences Can Efficiently Recruit Histone H3 Lysine 27 Trimethylation in the Absence of Enhancer Activity and DNA Methylation.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (33): E3415–21.
Jin, Ying, Oliver H. Tam, Eric Paniagua, and Molly Hammell. 2015. “TEtranscripts: A Package for Including Transposable Elements in Differential Expression Analysis of RNA-Seq Datasets.” *Bioinformatics (Oxford, England)* 31 (22): 3593–99.

Johnson, Welkin E. 2019. “Origins and Evolutionary Consequences of Ancient Endogenous Retroviruses.” *Nature Reviews. Microbiology* 17 (6): 355–70.
Kaluscha, Sebastian, Silvia Domcke, Christiane Wirbelauer, Michael B. Stadler, Sevi Durdu, Lukas Burger, and Dirk Schübeler. 2022. “Evidence That Direct Inhibition of Transcription Factor Binding Is the Prevailing Mode of Gene and Repeat Repression by DNA Methylation.” *Nature Genetics* 54 (12): 1895–1906.
Karimi, Mohammad M., Preeti Goyal, Irina A. Maksakova, Misha Bilenky, Danny Leung, Jie Xin Tang, Yoichi Shinkai, et al. 2011. “DNA Methylation and SETDB1/H3K9me3 Regulate Predominantly Distinct Sets of Genes, Retroelements, and Chimeric Transcripts in mESCs.” *Cell Stem Cell* 8 (6): 676–87.
Karttunen, Konsta, Divyesh Patel, Jihan Xia, Liangru Fei, Kimmo Palin, Lauri Aaltonen, and Biswajyoti Sahu. 2023. “Transposable Elements as Tissue-Specific Enhancers in Cancers of Endodermal Lineage.” *Nature Communications* 14 (1): 5313.
Kass, David H., and Nicole Jamison. 2007. “Identification of an Active ID-like Group of SINEs in the Mouse.” *Genomics* 90 (3): 416–20.
Kaya-Okur, Hatice S., Steven J. Wu, Christine A. Codomo, Erica S. Pledger, Terri D. Bryson, Jorja G. Henikoff, Kami Ahmad, and Steven Henikoff. 2019. “CUT&Tag for Efficient Epigenomic Profiling of Small Samples and Single Cells.” *Nature Communications* 10 (1): 1930.
Kazazian, Haig H., Jr. 2004. “Mobile Elements: Drivers of Genome Evolution.” *Science (New York, N.Y.)* 303 (5664): 1626–32.
Keegan, Richard M., Lillian R. Talbot, Yung-Heng Chang, Michael J. Metzger, and Josh Dubnau. 2021. “Intercellular Viral Spread and Intracellular Transposition of Drosophila Gypsy.” *PLoS Genetics* 17 (4): e1009535.
Keighley, Laura M., Chiemi F. Lynch-Sutherland, Suzan N. Almomani, Michael R. Eccles, and Erin C. Macaulay. 2023. “Unveiling the Hidden Players: The Crucial Role of Transposable Elements in the Placenta and Their Potential Contribution to Pre-Eclampsia.” *Placenta* 141 (September):57–64.
Khan, Hameed, Arian Smit, and Stéphane Boissinot. 2006. “Molecular Evolution and Tempo of Amplification of Human LINE-1 Retrotransposons since the Origin of Primates.” *Genome Research* 16 (1): 78–87.
Khmelinskii, Anton, Philipp J. Keller, Anna Bartosik, Matthias Meurer, Joseph D. Barry, Balca R. Mardin, Andreas Kaufmann, et al. 2012. “Tandem Fluorescent Protein Timers for in Vivo Analysis of Protein Dynamics.” *Nature Biotechnology* 30 (7): 708–14.
Kidwell, Margaret G. 2002. “Transposable Elements and the Evolution of Genome Size in Eukaryotes.” *Genetica* 115 (1): 49–63.
Koche, Richard P., Zachary D. Smith, Mazhar Adli, Hongcang Gu, Manching Ku, Andreas Gnirke, Bradley E. Bernstein, and Alexander Meissner. 2011. “Reprogramming Factor Expression Initiates Widespread Targeted Chromatin Remodeling.” *Cell Stem Cell* 8 (1): 96–105.
Kong, Lingqi, Karabi Saha, Yuchi Hu, Jada N. Tschetter, Chase E. Habben, Leanne S. Whitmore, Changfeng Yao, et al. 2022. “Subfamily-Specific Differential Contribution of Individual Monomers and the Tether Sequence to Mouse L1 Promoter Activity.” *Mobile DNA* 13 (1): 13.
Kosuge, Masato, Jumpei Ito, and Michiaki Hamada. 2024. “Landscape of Evolutionary Arms Races between Transposable Elements and KRAB-ZFP Family.” *Scientific Reports* 14 (1): 23358.
Kramerov, D. A., and N. S. Vassetzky. 2011. “Origin and Evolution of SINEs in Eukaryotic Genomes.” *Heredity* 107 (6): 487–95.
Kravchenko, Pavel, and Kikuë Tachibana. 2025. “Rise and SINE: Roles of Transcription Factors and Retrotransposons in Zygotic Genome Activation.” *Nature Reviews. Molecular Cell Biology* 26 (1): 68–79.
Kulpa, Deanna A., and John V. Moran. 2006. “Cis-Preferential LINE-1 Reverse Transcriptase Activity in Ribonucleoprotein Particles.” *Nature Structural & Molecular Biology* 13 (7): 655–60.

Lachner, M., D. O'Carroll, S. Rea, K. Mechtler, and T. Jenuwein. 2001. "Methylation of Histone H3 Lysine 9 Creates a Binding Site for HP1 Proteins." *Nature* 410 (6824): 116–20.
Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, et al. 2001. "Initial Sequencing and Analysis of the Human Genome." *Nature* 409 (6822): 860–921.
Lee, Chul-Hwan, Marlene Holder, Daniel Grau, Ricardo Saldaña-Meyer, Jia-Ray Yu, Rais Ahmad Ganai, Jenny Zhang, et al. 2018. "Distinct Stimulatory Mechanisms Regulate the Catalytic Activity of Polycomb Repressive Complex 2." *Molecular Cell* 70 (3): 435–48.e5.
Li, E., C. Beard, and R. Jaenisch. 1993. "Role for DNA Methylation in Genomic Imprinting." *Nature* 366 (6453): 362–65.
Li, E., T. H. Bestor, and R. Jaenisch. 1992. "Targeted Mutation of the DNA Methyltransferase Gene Results in Embryonic Lethality." *Cell* 69 (6): 915–26.
Li, En, and Yi Zhang. 2014. "DNA Methylation in Mammals." *Cold Spring Harbor Perspectives in Biology* 6 (5): a019133.
Lin, Jia-Hwei, and Henry L. Levin. 1998. "Reverse Transcription of a Self-Primed Retrotransposon Requires an RNA Structure Similar to the U5-IR Stem-Loop of Retroviruses." *Molecular and Cellular Biology* 18 (11): 6859–69.
Li, Qingqing, Yuqing Guo, Zixin Wu, Xueqiang Xu, Zhenhuan Jiang, Shuyue Qi, Zhenyu Liu, Lu Wen, and Fuchou Tang. 2024. "scNanoSeq-CUT&Tag: A Single-Cell Long-Read CUT&Tag Sequencing Method for Efficient Chromatin Modification Profiling within Individual Cells." *Nature Methods* 21 (11): 2044–57.
Li, Ruifang, Sara A. Grimm, and Paul A. Wade. 2021. "CUT&Tag-BS for Simultaneous Profiling of Histone Modification and DNA Methylation with High Efficiency and Low Cost." *Cell Reports Methods* 1 (8): 100118.
Lister, Ryan, Mattia Pelizzola, Robert H. Dowen, R. David Hawkins, Gary Hon, Julian Tonti-Filippini, Joseph R. Nery, et al. 2009. "Human DNA Methylomes at Base Resolution Show Widespread Epigenomic Differences." *Nature* 462 (7271): 315–22.
Liu, Ke, Weifang Li, Yuqing Xiao, Ming Lei, Ming Zhang, and Jinrong Min. 2024. "Molecular Mechanism of Specific DNA Sequence Recognition by NRF1." *Nucleic Acids Research* 52 (2): 953–66.
Liu, Nian, Cameron H. Lee, Tomek Swigut, Edward Grow, Bo Gu, Michael C. Bassik, and Joanna Wysocka. 2018. "Selective Silencing of Euchromatic L1s Revealed by Genome-Wide Screens for L1 Regulators." *Nature* 553 (7687): 228–32.
Liu, Sheng, Julie Brind'Amour, Mohammad M. Karimi, Kenjiro Shirane, Aaron Bogutz, Louis Lefebvre, Hiroyuki Sasaki, Yoichi Shinkai, and Matthew C. Lorincz. 2014. "Setdb1 Is Required for Germline Development and Silencing of H3K9me3-Marked Endogenous Retroviruses in Primordial Germ Cells." *Genes & Development* 28 (18): 2041–55.
Li, Xiaoping, Yi Yang, Lei Bu, Xiaogang Guo, Chengcheng Tang, Jun Song, Nana Fan, et al. 2014. "Rosa26-Targeted Swine Models for Stable Gene over-Expression and Cre-Mediated Lineage Tracing." *Cell Research* 24 (4): 501–4.
Li, Xiufeng, Luyao Bie, Yang Wang, Yaqiang Hong, Ziqiang Zhou, Yiming Fan, Xiaohan Yan, et al. 2024. "LINE-1 Transcription Activates Long-Range Gene Expression." *Nature Genetics* 56 (7): 1494–1502.
Li, Yinuo, Yue Zhang, and Mingxi Liu. 2021. "Knockout Gene-Based Evidence for PIWI-Interacting RNA Pathway in Mammals." *Frontiers in Cell and Developmental Biology* 9 (July):681188.
Lopez, Omar M., Michelle Zhang, Peter G. Hendrickson, Jianguo Huang, Andrea R. Daniels, Jane Blackmer, Lixia Luo, Laura D. Attardi, David Corcoran, and David G. Kirsch. 2023. "P53 and RB Cooperate to Suppress Transposable Elements." *bioRxiv*. https://doi.org/10.1101/2023.02.06.527304.
Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2." *Genome Biology* 15 (12): 550.
Luan, D. D., M. H. Korman, J. L. Jakubczak, and T. H. Eickbush. 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.
Lu, Chunye, Xavier Contreras, and B. Matija Peterlin. 2011. "P Bodies Inhibit Retrotransposition of Endogenous Intracisternal a Particles." *Journal of Virology* 85 (13): 6244–51.
Lu, Xinyi. 2024. "Regulation of Endogenous Retroviruses in Murine Embryonic Stem Cells and Early Embryos." *Journal of Molecular Cell Biology* 15 (8): mjad052.
Macfarlan, Todd S., Wesley D. Gifford, Shawn Driscoll, Karen Lettieri, Helen M. Rowe, Dario Bonanomi, Amy Firth, Oded Singer, Didier Trono, and Samuel L. Pfaff. 2012. "Embryonic Stem Cell Potency Fluctuates with Endogenous Retrovirus Activity." *Nature* 487 (7405): 57–63.
Maddox, Brenda. 2003. "The Double Helix and the 'Wronged Heroine.'" *Nature* 421 (6921): 407–8.
Magiorkinis, Gkikas, Robert Belshaw, and Aris Katzourakis. 2013. "'There and Back Again': Revisiting the Pathophysiological Roles of Human Endogenous Retroviruses in the Post-Genomic Era." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 368 (1626): 20120504.
Magiorkinis, Gkikas, Robert J. Gifford, Aris Katzourakis, Joris De Ranter, and Robert Belshaw. 2012. "Env-Less Endogenous Retroviruses Are Genomic Superspreaders." *Proceedings of the National Academy of Sciences of the United States of America* 109 (19): 7385–90.
Mandal, Prabhat K., and Haig H. Kazazian Jr. 2008. "SnapShot: Vertebrate Transposons." *Cell* 135 (1): 192–192.e1.
Mann, Ishminder K., Raghunath Chatterjee, Jianfei Zhao, Ximiao He, Matthew T. Weirauch, Timothy R. Hughes, and Charles Vinson. 2013. "CG Methylated Microarrays Identify a Novel Methylated Sequence Bound by the CEBPB|ATF4 Heterodimer That Is Active in Vivo." *Genome Research* 23 (6): 988–97.
Margueron, Raphael, Neil Justin, Katsuhito Ohno, Miriam L. Sharpe, Jinsook Son, William J. Drury 3rd, Philipp Voigt, et al. 2009. "Role of the Polycomb Protein EED in the Propagation of Repressive Histone Marks." *Nature* 461 (7265): 762–67.
Margueron, Raphael, Guohong Li, Kavitha Sarma, Alexandre Blais, Jiri Zavadil, Christopher L. Woodcock, Brian D. Dynlacht, and Danny Reinberg. 2008. "Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms." *Molecular Cell* 32 (4): 503–18.
Margueron, Raphaël, and Danny Reinberg. 2011. "The Polycomb Complex PRC2 and Its Mark in Life." *Nature* 469 (7330): 343–49.
Martin, Sandra L., and David J. Garfinkel. 2003. "Survival Strategies for Transposons and Genomes." *Genome Biology* 4 (4): 313.
Maston, Glenn A., Sara K. Evans, and Michael R. Green. 2006. "Transcriptional Regulatory Elements in the Human Genome." *Annual Review of Genomics and Human Genetics* 7 (1): 29–59.
Mathias, S. L., A. F. Scott, H. H. Kazazian Jr, J. D. Boeke, and A. Gabriel. 1991. "Reverse Transcriptase Encoded by a Human Transposable Element." *Science (New York, N.Y.)* 254 (5039): 1808–10.
Matsui, Toshiyuki, Danny Leung, Hiroki Miyashita, Irina A. Maksakova, Hitoshi Miyachi, Hiroshi Kimura, Makoto Tachibana, Matthew C. Lorincz, and Yoichi Shinkai. 2010. "Proviral Silencing in Embryonic Stem Cells Requires the Histone Methyltransferase ESET." *Nature* 464 (7290): 927–31.
Maupetit-Mehouas, Stéphanie, and Chantal Vaury. 2020. "Transposon Reactivation in the Germline May Be Useful for Both Transposons and Their Host Genomes." *Cells (Basel, Switzerland)* 9 (5): 1172.
McCLINTOCK, B. 1950. "The Origin and Behavior of Mutable Loci in Maize." *Proceedings of the National Academy of Sciences of the United States of America* 36 (6): 344–55.
Mendoza, Alex de, William L. Hatleberg, Kevin Pang, Sven Leininger, Ozren Bogdanovic, Jahnvi Pflueger, Sam Buckberry, et al. 2019. "Convergent Evolution of a Vertebrate-like Methylome in a Marine Sponge." *Nature Ecology & Evolution* 3 (10): 1464–73.

Mills, Ryan E., E. Andrew Bennett, Rebecca C. Iskow, and Scott E. Devine. 2007. “Which Transposable Elements Are Active in the Human Genome?” *Trends in Genetics: TIG* 23 (4): 183–91.
Mita, Paolo, Aleksandra Wudzinska, Xiaoji Sun, Joshua Andrade, Shruti Nayak, David J. Kahler, Sana Badri, et al. 2018. “LINE-1 Protein Localization and Functional Dynamics during the Cell Cycle.” *eLife* 7 (January). https://doi.org/10.7554/eLife.30058.
Mohandas, T., R. S. Sparkes, and L. J. Shapiro. 1981. “Reactivation of an Inactive Human X Chromosome: Evidence for X Inactivation by DNA Methylation.” *Science (New York, N.Y.)* 211 (4480): 393–96.
Mohn, Fabio, Michael Weber, Michael Rebhan, Tim C. Roloff, Jens Richter, Michael B. Stadler, Miriam Bibel, and Dirk Schübeler. 2008. “Lineage-Specific Polycomb Targets and de Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors.” *Molecular Cell* 30 (6): 755–66.
Molaro, Antoine, Ilaria Falciatori, Emily Hodges, Alexei A. Aravin, Krista Marran, Shahin Rafii, W. Richard McCombie, Andrew D. Smith, and Gregory J. Hannon. 2014. “Two Waves of de Novo Methylation during Mouse Germ Cell Development.” *Genes & Development* 28 (14): 1544–49.
Moore, Lisa D., Thuc Le, and Guoping Fan. 2013. “DNA Methylation and Its Basic Function.” *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology* 38 (1): 23–38.
Mouse Genome Sequencing Consortium, Robert H. Waterston, Kerstin Lindblad-Toh, Ewan Birney, Jane Rogers, Josep F. Abril, Pankaj Agarwal, et al. 2002. “Initial Sequencing and Comparative Analysis of the Mouse Genome.” *Nature* 420 (6915): 520–62.
Muñoz-López, Martín, and José L. García-Pérez. 2010. “DNA Transposons: Nature and Applications in Genomics.” *Current Genomics* 11 (2): 115–28.
Nabet, Behnam, Justin M. Roberts, Dennis L. Buckley, Joshiawa Paulk, Shiva Dastjerdi, Annan Yang, Alan L. Leggett, et al. 2018. “The dTAG System for Immediate and Target-Specific Protein Degradation.” *Nature Chemical Biology* 14 (5): 431–41.
Nan, X., H. H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eisenman, and A. Bird. 1998. “Transcriptional Repression by the Methyl-CpG-Binding Protein MeCP2 Involves a Histone Deacetylase Complex.” *Nature* 393 (6683): 386–89.
Nan, X., P. Tate, E. Li, and A. Bird. 1996. “DNA Methylation Specifies Chromosomal Localization of MeCP2.” *Molecular and Cellular Biology* 16 (1): 414–21.
Oakberg, E. F. 1956. “A Description of Spermiogenesis in the Mouse and Its Use in Analysis of the Cycle of the Seminiferous Epithelium and Germ Cell Renewal.” *The American Journal of Anatomy* 99 (3): 391–413.
O’Carroll, Dónal, Harry Scherthan, Antoine H. F. M. Peters, Susanne Opravil, Andrew R. Haynes, Götz Laible, Stephen Rea, et al. 2000. “Isolation and Characterization of *Suv39h2*, a Second Histone H3 Methyltransferase Gene That Displays Testis-Specific Expression.” *Molecular and Cellular Biology* 20 (24): 9423–33.
Okano, M., D. W. Bell, D. A. Haber, and E. Li. 1999. “DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for de Novo Methylation and Mammalian Development.” *Cell* 99 (3): 247–57.
Pabis, Kamil, Diogo Barardo, Olga Sirbu, Kumar Selvarajoo, Jan Gruber, and Brian K. Kennedy. 2024. “A Concerted Increase in Readthrough and Intron Retention Drives Transposon Expression during Aging and Senescence.” *eLife* 12 (April). https://doi.org/10.7554/eLife.87811.
Pace, John K., 2nd, and Cédric Feschotte. 2007. “The Evolutionary History of Human DNA Transposons: Evidence for Intense Activity in the Primate Lineage.” *Genome Research* 17 (4): 422–32.
Palmer, Nathan, S. Zakiah A. Talib, Chandrahas Koumar Ratnacaram, Diana Low, Xavier Bisteau, Joanna Hui Si Lee, Elisabeth Pfeiffenberger, et al. 2019. “CDK2 Regulates the NRF1/Ehmt1 Axis during Meiotic Prophase I.” *The Journal of Cell Biology* 218 (9): 2896–2918.
Pauler, Florian M., Mathew A. Sloane, Ru Huang, Kakkad Regha, Martha V. Koerner, Ido

Tamir, Andreas Sommer, Andras Aszodi, Thomas Jenuwein, and Denise P. Barlow. 2009. “H3K27me3 Forms BLOCs over Silent Genes and Intergenic Regions and Specifies a Histone Banding Pattern on a Mouse Autosomal Chromosome.” *Genome Research* 19 (2): 221–33.
Peaston, Anne E., Alexei V. Evsikov, Joel H. Graber, Wilhelmine N. de Vries, Andrea E. Holbrook, Davor Solter, and Barbara B. Knowles. 2004. “Retrotransposons Regulate Host Genes in Mouse Oocytes and Preimplantation Embryos.” *Developmental Cell* 7 (4): 597–606.
Percharde, Michelle, Chih-Jen Lin, Yafei Yin, Juan Guan, Gabriel A. Peixoto, Aydan Bulut-Karslioglu, Steffen Biechele, Bo Huang, Xiaohua Shen, and Miguel Ramalho-Santos. 2018. “A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity.” *Cell* 174 (2): 391–405.e19.
Peters, Antoine H. F. M., Stefan Kubicek, Karl Mechtler, Roderick J. O’Sullivan, Alwin A. H. A. Derijck, Laura Perez-Burgos, Alexander Kohlmaier, et al. 2003. “Partitioning and Plasticity of Repressive Histone Methylation States in Mammalian Chromatin.” *Molecular Cell* 12 (6): 1577–89.
Petri, Rebecca, Per Ludvik Brattås, Yogita Sharma, Marie E. Jönsson, Karolina Pircs, Johan Bengzon, and Johan Jakobsson. 2019. “LINE-2 Transposable Elements Are a Source of Functional Human microRNAs and Target Sites.” *PLoS Genetics* 15 (3): e1008036.
Plath, Kathrin, Jia Fang, Susanna K. Mlynarczyk-Evans, Ru Cao, Kathleen A. Worringer, Hengbin Wang, Cecile C. de la Cruz, Arie P. Otte, Barbara Panning, and Yi Zhang. 2003. “Role of Histone H3 Lysine 27 Methylation in X Inactivation.” *Science (New York, N.Y.)* 300 (5616): 131–35.
Platt, Roy N., 2nd, Michael W. Vandewege, and David A. Ray. 2018. “Mammalian Transposable Elements and Their Impacts on Genome Evolution.” *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* 26 (1-2): 25–43.
Pontis, Julien, Evarist Planet, Sandra Offner, Priscilla Turelli, Julien Duc, Alexandre Coudray, Thorold W. Theunissen, Rudolf Jaenisch, and Didier Trono. 2019. “Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic Genome Activation and Control Transcription in Naive Human ESCs.” *Cell Stem Cell* 24 (5): 724–35.e5.
Qian, Yaping, Debora Mancini-DiNardo, Thaddeus Judkins, Hannah C. Cox, Krystal Brown, Maria Elias, Nanda Singh, et al. 2017. “Identification of Pathogenic Retrotransposon Insertions in Cancer Predisposition Genes.” *Cancer Genetics* 216-217 (October):159–69.
Quenneville, Simon, Priscilla Turelli, Karolina Bojkowska, Charlène Raclot, Sandra Offner, Adamandia Kapopoulou, and Didier Trono. 2012. “The KRAB-ZFP/KAP1 System Contributes to the Early Embryonic Establishment of Site-Specific DNA Methylation Patterns Maintained during Development.” *Cell Reports* 2 (4): 766–73.
Ran, F. Ann, Patrick D. Hsu, Jason Wright, Vineeta Agarwala, David A. Scott, and Feng Zhang. 2013. “Genome Engineering Using the CRISPR-Cas9 System.” *Nature Protocols* 8 (11): 2281–2308.
Reilly, Steven K., Sager J. Gosai, Alan Gutierrez, Ava Mackay-Smith, Jacob C. Ulirsch, Masahiro Kanai, Kousuke Mouri, et al. 2021. “Direct Characterization of Cis-Regulatory Elements and Functional Dissection of Complex Genetic Associations Using HCR-FlowFISH.” *Nature Genetics* 53 (8): 1166–76.
Reuten, Raphael, Denise Nikodemus, Maria B. Oliveira, Trushar R. Patel, Bent Brachvogel, Isabelle Breloy, Jörg Stetefeld, and Manuel Koch. 2016. “Maltose-Binding Protein (MBP), a Secretion-Enhancing Tag for Mammalian Protein Expression Systems.” *PloS One* 11 (3): e0152386.
Richard Albert, Julien, Teresa Urli, Ana Monteagudo-Sánchez, Anna Le Breton, Amina Sultanova, Angélique David, Margherita Scarpa, Mathieu Schulz, and Maxim V. C. Greenberg. 2024. “DNA Methylation Shapes the Polycomb Landscape during the Exit from Naive Pluripotency.” *Nature Structural & Molecular Biology*, October, 1–12.
Richards, Eric J., and Sarah C. R. Elgin. 2002. “Epigenetic Codes for Heterochromatin

Formation and Silencing: Rounding up the Usual Suspects." *Cell* 108 (4): 489–500.
Rooij, D. G. de, and L. D. Russell. 2000. "All You Wanted to Know about Spermatogonia but Were Afraid to Ask." *Journal of Andrology* 21 (6): 776–98.
Rowe, Helen M., Johan Jakobsson, Daniel Mesnard, Jacques Rougemont, Séverine Reynard, Tugce Aktas, Pierre V. Maillard, et al. 2010. "KAP1 Controls Endogenous Retroviruses in Embryonic Stem Cells." *Nature* 463 (7278): 237–40.
Ryan, Robert F., David C. Schultz, Kasirajan Ayyanathan, Prim B. Singh, Josh R. Friedman, William J. Fredericks, and Frank J. Rauscher III. 1999. "KAP-1 Corepressor Protein Interacts and Colocalizes with Heterochromatic and Euchromatic HP1 Proteins: A Potential Role for Krüppel-Associated Box–zinc Finger Proteins in Heterochromatin-Mediated Gene Silencing." *Molecular and Cellular Biology* 19 (6): 4366–78.
Sachs, Michael, Courtney Onodera, Kathryn Blaschke, Kevin T. Ebata, Jun S. Song, and Miguel Ramalho-Santos. 2013. "Bivalent Chromatin Marks Developmental Regulatory Genes in the Mouse Embryonic Germline in Vivo." *Cell Reports* 3 (6): 1777–84.
Sadate-Ngatchou, Patricia I., Christopher J. Payne, Andrea T. Dearth, and Robert E. Braun. 2008. "Cre Recombinase Activity Specific to Postnatal, Premeiotic Male Germ Cells in Transgenic Mice." *Genesis (New York, N.Y.: 2000)* 46 (12): 738–42.
Sadic, Dennis, Katharina Schmidt, Sophia Groh, Ivan Kondofersky, Joachim Ellwart, Christiane Fuchs, Fabian J. Theis, and Gunnar Schotta. 2015. "Atrx Promotes Heterochromatin Formation at Retrotransposons." *EMBO Reports* 16 (7): 836–50.
Saha, Karabi, Grace I. Nielsen, Raj Nandani, Lingqi Kong, Ping Ye, and Wenfeng An. 2024. "YY1 Is a Transcriptional Activator of Mouse LINE-1 Tf Subfamily." *bioRxivorg*. https://doi.org/10.1101/2024.01.03.573552.
Saitou, Mitinori, and Masashi Yamaji. 2012. "Primordial Germ Cells in Mice." *Cold Spring Harbor Perspectives in Biology* 4 (11): a008375–a008375.
Sakashita, Akihiko, Tomohiro Kitano, Hirotsugu Ishizu, Youjia Guo, Harumi Masuda, Masaru Ariura, Kensaku Murano, and Haruhiko Siomi. 2023. "Transcription of MERVL Retrotransposons Is Required for Preimplantation Embryo Development." *Nature Genetics* 55 (3): 484–95.
Sakashita, Akihiko, So Maezawa, Kazuki Takahashi, Kris G. Alavattam, Masashi Yukawa, Yueh-Chiang Hu, Shohei Kojima, et al. 2020. "Endogenous Retroviruses Drive Species-Specific Germline Transcriptomes in Mammals." *Nature Structural & Molecular Biology* 27 (10): 967–77.
Sandmeyer, Suzanne, and Nancy Craig. 2015. "Mammalian Endogenous Retroviruses." In *Mobile DNA III*, 1079–1100. American Society of Microbiology.
Santos-Barriopedro, Irene, Guido van Mierlo, and Michiel Vermeulen. 2021. "Off-the-Shelf Proximity Biotinylation for Interaction Proteomics." *Nature Communications* 12 (1): 5015.
Saxonov, Serge, Paul Berg, and Douglas L. Brutlag. 2006. "A Genome-Wide Analysis of CpG Dinucleotides in the Human Genome Distinguishes Two Distinct Classes of Promoters." *Proceedings of the National Academy of Sciences of the United States of America* 103 (5): 1412–17.
Scarpulla, Richard C. 2002. "Nuclear Activators and Coactivators in Mammalian Mitochondrial Biogenesis." *Biochimica et Biophysica Acta* 1576 (1-2): 1–14.
Schaefer, L., H. Engman, and J. B. Miller. 2000. "Coding Sequence, Chromosomal Localization, and Expression Pattern of Nrf1: The Mouse Homolog of Drosophila Erect Wing." *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 11 (2): 104–10.
Schmitz, J. 2012. "SINEs as Driving Forces in Genome Evolution." *Genome Dynamics* 7 (June):92–107.
Schneider, Caroline A., Wayne S. Rasband, and Kevin W. Eliceiri. 2012. "NIH Image to ImageJ: 25 Years of Image Analysis." *Nature Methods* 9 (7): 671–75.
Schöpp, Theresa, Ansgar Zoch, Rebecca V. Berrens, Tania Auchynnikava, Yuka Kabayama, Lina Vasiliauskaitė, Juri Rappsilber, Robin C. Allshire, and Dónal O'Carroll. 2020. "TEX15 Is an Essential Executor of MIWI2-Directed Transposon DNA Methylation and

Silencing." *Nature Communications* 11 (1): 3739.
Schultz, David C., Kasirajan Ayyanathan, Dmitri Negorev, Gerd G. Maul, and Frank J. Rauscher 3rd. 2002. "SETDB1: A Novel KAP-1-Associated Histone H3, Lysine 9-Specific Methyltransferase That Contributes to HP1-Mediated Silencing of Euchromatic Genes by KRAB Zinc-Finger Proteins." *Genes & Development* 16 (8): 919–32.
Seisenberger, Stefanie, Simon Andrews, Felix Krueger, Julia Arand, Jörn Walter, Fátima Santos, Christian Popp, Bernard Thienpont, Wendy Dean, and Wolf Reik. 2012. "The Dynamics of Genome-Wide DNA Methylation Reprogramming in Mouse Primordial Germ Cells." *Molecular Cell* 48 (6): 849–62.
Sentmanat, Monica F., Samuel T. Peters, Colin P. Florian, Jon P. Connelly, and Shondra M. Pruett-Miller. 2018. "A Survey of Validation Strategies for CRISPR-Cas9 Editing." *Scientific Reports* 8 (1): 888.
Shin, Jiyung, Fuguo Jiang, Jun-Jie Liu, Nicolas L. Bray, Benjamin J. Rauch, Seung Hyun Baik, Eva Nogales, Joseph Bondy-Denomy, Jacob E. Corn, and Jennifer A. Doudna. 2017. "Disabling Cas9 by an Anti-CRISPR DNA Mimic." *Science Advances* 3 (7): e1701620.
Sil, Srinjoy, Sarah Keegan, Farida Ettefa, Lance T. Denes, Jef D. Boeke, and Liam J. Holt. 2023. "Condensation of LINE-1 Is Critical for Retrotransposition." *eLife* 12 (April). https://doi.org/10.7554/eLife.82991.
Smit, A. F. 1993. "Identification of a New, Abundant Superfamily of Mammalian LTR-Transposons." *Nucleic Acids Research* 21 (8): 1863–72.
Sookdeo, Akash, Crystal M. Hepp, Marcella A. McClure, and Stéphane Boissinot. 2013. "Revisiting the Evolution of Mouse LINE-1 in the Genomic Era." *Mobile DNA* 4 (1): 3.
Soufi, Abdenour, Greg Donahue, and Kenneth S. Zaret. 2012. "Facilitators and Impediments of the Pluripotency Reprogramming Factors' Initial Engagement with the Genome." *Cell* 151 (5): 994–1004.
Srinivasan, Rajini, Tao Sun, Azalia Sandles, Diana Wu, Liang Wang, Amy Heidersbach, Clark Ho, Shiqi Xie, Andrew Ng, and Benjamin Haley. 2024. "Chemically-Inducible CRISPR/Cas9 Circuits for Ultra-High Dynamic Range Gene Perturbation." *bioRxiv*. https://doi.org/10.1101/2024.12.27.630546.
Stadler, Michael B., Rabih Murr, Lukas Burger, Robert Ivanek, Florian Lienert, Anne Schöler, Erik van Nimwegen, et al. 2011. "DNA-Binding Factors Shape the Mouse Methylome at Distal Regulatory Regions." *Nature* 480 (7378): 490–95.
Stocking, C., and C. A. Kozak. 2008. "Murine Endogenous Retroviruses." *Cellular and Molecular Life Sciences: CMLS* 65 (21): 3383–98.
Stoll, Guido A., Ninoslav Pandiloski, Christopher H. Douse, and Yorgo Modis. 2022. "Structure and Functional Mapping of the KRAB-KAP1 Repressor Complex." *The EMBO Journal* 41 (24): e111179.
Stoye, J. P. 2001. "Endogenous Retroviruses: Still Active after All These Years?" *Current Biology: CB* 11 (22): R914–16.
Sugie, Kenta, Satoshi Funaya, Machika Kawamura, Toshinobu Nakamura, Masataka G. Suzuki, and Fugaku Aoki. 2020. "Expression of Dux Family Genes in Early Preimplantation Embryos." *Scientific Reports* 10 (1): 19396.
Svoboda, Petr, Paula Stein, Martin Anger, Emily Bernstein, Gregory J. Hannon, and Richard M. Schultz. 2004. "RNAi and Expression of Retrotransposons MuERV-L and IAP in Preimplantation Mouse Embryos." *Developmental Biology* 269 (1): 276–85.
Tam, P. P., and M. H. Snow. 1981. "Proliferation and Migration of Primordial Germ Cells during Compensatory Growth in Mouse Embryos." *Journal of Embryology and Experimental Morphology* 64 (August):133–47.
Tarlinton, Rachael E., Joanne Meers, and Paul R. Young. 2006. "Retroviral Invasion of the Koala Genome." *Nature* 442 (7098): 79–81.
Tchasovnikarova, Iva A., Richard T. Timms, Christopher H. Douse, Rhys C. Roberts, Gordon Dougan, Robert E. Kingston, Yorgo Modis, and Paul J. Lehner. 2017. "Hyperactivation of HUSH Complex Function by Charcot-Marie-Tooth Disease Mutation in MORC2."

Nature Genetics 49 (7): 1035–44.
Tchasovnikarova, Iva A., Richard T. Timms, Nicholas J. Matheson, Kim Wals, Robin Antrobus, Berthold Göttgens, Gordon Dougan, Mark A. Dawson, and Paul J. Lehner. 2015. “GENE SILENCING. Epigenetic Silencing by the HUSH Complex Mediates Position-Effect Variegation in Human Cells.” *Science (New York, N.Y.)* 348 (6242): 1481–85.
Teissandier, Aurélie, Nicolas Servant, Emmanuel Barillot, and Deborah Bourc’his. 2019. “Tools and Best Practices for Retrotransposon Analysis Using High-Throughput Sequencing Data.” *Mobile DNA* 10 (1): 52.
Teugels, Erik, Sylvia De Brakeleer, Guido Goelen, Willy Lissens, Erica Sermijn, and Jacques De Grève. 2005. “De Novo Alu Element Insertions Targeted to a Sequence Common to the BRCA1 and BRCA2 Genes.” *Human Mutation* 26 (3): 284.
Thompson, Peter J., Todd S. Macfarlan, and Matthew C. Lorincz. 2016. “Long Terminal Repeats: From Parasitic Elements to Building Blocks of the Transcriptional Regulatory Repertoire.” *Molecular Cell* 62 (5): 766–76.
Tiwari, Bhavana, Amanda E. Jones, Candace J. Caillet, Simanti Das, Stephanie K. Royer, and John M. Abrams. 2020. “p53 Directly Represses Human LINE1 Transposons.” *Genes & Development* 34 (21-22): 1439–51.
Trojer, Patrick, and Danny Reinberg. 2007. “Facultative Heterochromatin: Is There a Distinctive Molecular Signature?” *Molecular Cell* 28 (1): 1–13.
Tsumura, Akiko, Tomohiro Hayakawa, Yuichi Kumaki, Shin-Ichiro Takebayashi, Morito Sakaue, Chisa Matsuoka, Kunitada Shimotohno, et al. 2006. “Maintenance of Self-Renewal Ability of Mouse Embryonic Stem Cells in the Absence of DNA Methyltransferases Dnmt1, Dnmt3a and Dnmt3b.” *Genes to Cells: Devoted to Molecular & Cellular Mechanisms* 11 (7): 805–14.
Vagin, Vasily V., Alla Sigova, Chengjian Li, Hervé Seitz, Vladimir Gvozdev, and Phillip D. Zamore. 2006. “A Distinct Small RNA Pathway Silences Selfish Genetic Elements in the Germline.” *Science (New York, N.Y.)* 313 (5785): 320–24.
Van’t Hof, Arjen E., Pascal Campagne, Daniel J. Rigden, Carl J. Yung, Jessica Lingley, Michael A. Quail, Neil Hall, Alistair C. Darby, and Ilik J. Saccheri. 2016. “The Industrial Melanism Mutation in British Peppered Moths Is a Transposable Element.” *Nature* 534 (7605): 102–5.
Vasiliauskaitė, Lina, Rebecca V. Berrens, Ivayla Ivanova, Claudia Carrieri, Wolf Reik, Anton J. Enright, and Dónal O’Carroll. 2018. “Defective Germline Reprogramming Rewires the Spermatogonial Transcriptome.” *Nature Structural & Molecular Biology* 25 (5): 394–404.
Vassetzky, Nikita S., and Dmitri A. Kramerov. 2013. “SINEBase: A Database and Tool for SINE Analysis.” *Nucleic Acids Research* 41 (Database issue): D83–89.
Vergouwen, R. P., S. G. Jacobs, R. Huiskamp, J. A. Davids, and D. G. de Rooij. 1991. “Proliferative Activity of Gonocytes, Sertoli Cells and Interstitial Cells during Testicular Development in Mice.” *Journal of Reproduction and Fertility* 93 (1): 233–43.
Vogt, P. K. 1997. “Historical Introduction to the General Properties of Retroviruses.” In *Retroviruses*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
Walsh, C. P., J. R. Chaillet, and T. H. Bestor. 1998. “Transcription of IAP Endogenous Retroviruses Is Constrained by Cytosine Methylation.” *Nature Genetics* 20 (2): 116–17.
Walter, Marius, Aurélie Teissandier, Raquel Pérez-Palacios, and Déborah Bourc’his. 2016. “An Epigenetic Switch Ensures Transposon Repression upon Dynamic Loss of DNA Methylation in Embryonic Stem Cells.” *eLife* 5 (January):e11418.
Wang, Fay, John Flanagan, Nan Su, Li-Chong Wang, Son Bui, Allissa Nielson, Xingyong Wu, Hong-Thuy Vo, Xiao-Jun Ma, and Yuling Luo. 2012. “RNAscope: A Novel in Situ RNA Analysis Platform for Formalin-Fixed, Paraffin-Embedded Tissues.” *The Journal of Molecular Diagnostics: JMD* 14 (1): 22–29.
Wang, Hui, Jinchuan Xing, Deepak Grover, Dale J. Hedges, Kyudong Han, Jerilyn A. Walker, and Mark A. Batzer. 2005. “SVA Elements: A Hominid-Specific Retroposon Family.” *Journal of Molecular Biology* 354 (4): 994–1007.
Wang, Junpeng, Chao Tang, Qian Wang, Jun Su, Ting Ni, Wenjing Yang, Yongsheng Wang,

et al. 2017. “NRF1 Coordinates with DNA Methylation to Regulate Spermatogenesis.” *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 31 (11): 4959–70.
Wang, Lu, Jun Zhang, Jialei Duan, Xinxing Gao, Wei Zhu, Xingyu Lu, Lu Yang, et al. 2014. “Programming and Inheritance of Parental DNA Methylomes in Mammals.” *Cell* 157 (4): 979–91.
Wang, Pengxiang, Jun Su, Junpeng Wang, Yilin Xie, Wei Chen, Jinhai Zhong, and Yuan Wang. 2024. “NRF1 Promotes Primordial Germ Cell Development, Proliferation and Survival.” *Cell Proliferation* 57 (1): e13533.
Watson, J. D., and F. H. Crick. 1953. “Molecular Structure of Nucleic Acids; a Structure for Deoxyribose Nucleic Acid.” *Nature* 171 (4356): 737–38.
Wells, Jonathan N., and Cédric Feschotte. 2020. “A Field Guide to Eukaryotic Transposable Elements.” *Annual Review of Genetics* 54 (1): 539–61.
Wicker, Thomas, François Sabot, Aurélie Hua-Van, Jeffrey L. Bennetzen, Pierre Capy, Boulos Chalhoub, Andrew Flavell, et al. 2007. “A Unified Classification System for Eukaryotic Transposable Elements.” *Nature Reviews. Genetics* 8 (12): 973–82.
Wilson, Matthew H., Craig J. Coates, and Alfred L. George Jr. 2007. “PiggyBac Transposon-Mediated Gene Transfer in Human Cells.” *Molecular Therapy: The Journal of the American Society of Gene Therapy* 15 (1): 139–45.
Wolf, Gernot, Peng Yang, Annette C. Füchtbauer, Ernst-Martin Füchtbauer, Andreia M. Silva, Chungoo Park, Warren Wu, Anders L. Nielsen, Finn S. Pedersen, and Todd S. Macfarlan. 2015. “The KRAB Zinc Finger Protein ZFP809 Is Required to Initiate Epigenetic Silencing of Endogenous Retroviruses.” *Genes & Development* 29 (5): 538–54.
Wunderlich, Zeba, and Leonid A. Mirny. 2009. “Different Gene Regulation Strategies Revealed by Analysis of Binding Motifs.” *Trends in Genetics: TIG* 25 (10): 434–40.
Wu, Qinyan. 2024. “Controlling Systems and Controlling Legacies: Barbara McClintock’s 1961 Conversation with Two Bacterial Geneticists.” *History and Philosophy of the Life Sciences* 46 (3): 31.
Xia, Bo, Weimin Zhang, Guisheng Zhao, Xinru Zhang, Jiangshan Bai, Ran Brosh, Aleksandra Wudzinska, et al. 2024. “On the Genetic Basis of Tail-Loss Evolution in Humans and Apes.” *Nature* 626 (8001): 1042–48.
Yang, Bin Xia, Chadi A. El Farran, Hong Chao Guo, Tao Yu, Hai Tong Fang, Hao Fei Wang, Sharon Schlesinger, et al. 2015. “Systematic Identification of Factors for Provirus Silencing in Embryonic Stem Cells.” *Cell* 163 (1): 230–45.
Yang, Fang, Sigrid Eckardt, N. Adrian Leu, K. John McLaughlin, and Peijing Jeremy Wang. 2008. “Mouse TEX15 Is Essential for DNA Double-Strand Break Repair and Chromosomal Synapsis during Male Meiosis.” *The Journal of Cell Biology* 180 (4): 673–79.
Yoder, J. A., C. P. Walsh, and T. H. Bestor. 1997. “Cytosine Methylation and the Ecology of Intragenomic Parasites.” *Trends in Genetics: TIG* 13 (8): 335–40.
Yuan, Guo-Cheng, Yuen-Jong Liu, Michael F. Dion, Michael D. Slack, Lani F. Wu, Steven J. Altschuler, and Oliver J. Rando. 2005. “Genome-Scale Identification of Nucleosome Positions in S. Cerevisiae.” *Science (New York, N.Y.)* 309 (5734): 626–30.
Zamudio, Natasha, Joan Barau, Aurélie Teissandier, Marius Walter, Maté Borsos, Nicolas Servant, and Déborah Bourc’his. 2015. “DNA Methylation Restrains Transposons from Adopting a Chromatin Signature Permissive for Meiotic Recombination.” *Genes & Development* 29 (12): 1256–70.
Zeng, Hongkui, Kyoji Horie, Linda Madisen, Maria N. Pavlova, Galina Gragerova, Alex D. Rohde, Brian A. Schimpf, et al. 2008. “An Inducible and Reversible Mouse Genetic Rescue System.” *PLoS Genetics* 4 (5): e1000069.
Zhang, Xiao-Ou, Henry Pratt, and Zhiping Weng. 2021. “Investigating the Potential Roles of SINEs in the Human Genome.” *Annual Review of Genomics and Human Genetics* 22 (1): 199–218.
Zhang, Xuanming, Ivana Celic, Hannah Mitchell, Sam Stuckert, Lalitha Vedula, and Jeffrey

S. Han. 2024. “Comprehensive Profiling of L1 Retrotransposons in Mouse.” *Nucleic Acids Research* 52 (9): 5166–78.
Zhu, Fangjie, Lucas Farnung, Eevi Kaasinen, Biswajyoti Sahu, Yimeng Yin, Bei Wei, Svetlana O. Dodonova, et al. 2018. “The Interaction Landscape between Transcription Factors and the Nucleosome.” *Nature* 562 (7725): 76–81.
Zoch, Ansgar, Tania Auchynnikava, Rebecca V. Berrens, Yuka Kabayama, Theresa Schöpp, Madeleine Heep, Lina Vasiliauskaitė, et al. 2020. “SPOCD1 Is an Essential Executor of piRNA-Directed de Novo DNA Methylation.” *Nature* 584 (7822): 635–39.
Zoch, Ansgar, Gabriela Konieczny, Tania Auchynnikava, Birgit Stallmeyer, Nadja Rotte, Madeleine Heep, Rebecca V. Berrens, et al. 2024. “C19ORF84 Connects piRNA and DNA Methylation Machineries to Defend the Mammalian Germ Line.” *Molecular Cell* 84 (6): 1021–35.e11.

Styliani (Stella) Kanta

she/her

Molecular Biologist

08 March 1994

+49 15227974093

stella.eirini.kanta@gmail.com

Mainz

Tiny Cover Letter

I am a PhD candidate in molecular biology with experience in mammalian cell culture, CRISPR/Cas9 genome editing, immunofluoresence and chromatin profiling. I am a highly motivated and observant scientist and I enjoy working with and helping other. With my educational background and my desire for lifelong learning, I aspire to positively impact every professional environment I work at.

Education

2020-2025 **PhD Candidate** JGU Mainz, Institute of Molecular Biology
"Regulation of transposon transcription and translation in mice"

2018-2020 **Master in Biomedical and Molecular Sciences in Diagnosis and Treatment of Diseases** Medicine School of Alexandroupoli

2013-2017 **Bachelor in Biomedical Science** International Hellenic University
(grade 8,79/10 Excellent)

Experience

2020-2025 **PhD candidate** Institute of Molecular Biology,Dr. Joan Barau Lab

In my PhD I focused on characterizing the expression of a transcription factor in mice spermatogenesis. I gained expertise in immunostaining, microscopy, flow cytometry, RNA in-situ hybridization and NGS (chromatin profiling). I have expertise in mammalian cell culture including cancer cells and I have generated cell lines by CRISPR/Cas9 genome editing techniques and piggyBac transposon vector in mESCs and HEK239T cells.

2/2019-9/2019 **Master Thesis** IMB, Institute of Molecular Biology

My thesis focused on the "Characterization of a novel helicase involved in the m6A pathway". I gained experience in several techniques such as Western Blot, RT-qPCR, protein Pull-down, cloning, RNA isolation, PCR and generating CRISPR/Cas9 edited Drosophila melanogaster strains.

12/2018-1/2019 **Trainee Student** Papageorgiou Hospital, Dept Molecular Biology

I volunteered to work in the Molecular Biology Department, where I got familiar with fundamental diagnostic techniques, including DNA extraction, PCR, restriction enzyme digestion.

6/2018-11/2018 **Trainee Student**
424 Military Hospital of Thessaloniki, Papageorgiou Hospital

I volunteered to get trained in venipuncture in adults, young chil-dren and infants. I was in daily contact with patients,doctors and medical staff.

4/2017-10/2017 **Internship as Biomedical Laboratory Scientist**
Oncology Hospital ("Theageneio") in Thessaloniki

I did a rotation in all medical departments of the hospital, such as: Microbiolocal, Biochemical, Blood Donation, Cytological, Histological, etc. I was in close daily contact with doctors and patients.

Styliani (Stella) Kanta

she/her

Molecular Biologist

08 March 1994

+49 15227974093

stella.eirini.kanta@gmail.com

Mainz

Soft Skills

Team Work
Time Management
Leadership
Reliability
Networking
Empathy

Languages

Greek (native)
English (full professional)
German (intermediate)

Volunteer Work

2022-2024 **Events Manager of Pint of Science** Mainz

In collaboration with fellow passionate scientists, we launched the first Pint of Science (PoS) event in Mainz in 2022. I contributed to venue selection, coordinated event setups and promoted the events.

2023-2024 **IPP PhD representative** IMB, JGU, UMC Mainz

As a PhD representative, I enhanced communication skills through managing multiple social media accounts. In collaboration with another colleague, we initiated the creation of the first social media platform on Instagram for our PhD program. I participated in events organization and advocated for equity in academic intercultural contexts.

7/2014-9/2014 **AIESEC Volunteer Program** Novi Sad, Serbia

Teaching greek culture/mythology and entertaining hospitalized kids at the Decija bolnica children's hospital and I co-organize a fundraising.

Further qualifications

2023-2024 **ADA Lovelace Mentoring Program** JGU,Mainz

As an ADA Lovelace mentee, I had the opportunity to find a mentor and receive a 16-month training. During this period, I acquired leadership skills, teamwork, and negotiations.