Investigation of the pluripotency-specific response to LINE1 transposons in early development

Ferran Garcia Llagostera

Thesis submitted towards fulfilment of the degree of Doctor of Philosophy

MRC Laboratory of Medical Sciences
Institute of Clinical Sciences
Imperial College London

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Licence (CC BY NC-SA).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that; you credit the author, do not use it for commercial purposes and share any derivative works under the same licence.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Statement of originality

All experiments included in this thesis and writing of the present manuscript were performed by the author, unless stated otherwise. Embryo manipulation was performed by Dr. Bryony Leeke. RNA-seq data mapping was performed by Dr. Darren Taylor and analysis was performed by Dr. Michelle Percharde (Figures 4.2B,C, 4.3, 4.4). Experiments on figures 3.6, 4.1 and some replicates on figure 4.*6* were done by Alex Choromidis.

Acknowledgments

First, I would like to express my deepest gratitude to Dr. Michelle Percharde for her excellent guidance, supervision and patience for these last 4 years. I have learned a great deal from her about how to do science, specially always having the right controls! I also want to appreciate her encouragement and shared enthusiastic discussion of the last 4 years. I want to extend my thanks to my assessors Dr. Peter Sarkies and Dr. Veronique Azuara, and my co-supervisor Dr. Alexis Barr for their challenging and insightful comments. For their support, day to day teachings and shared lunches and events outside work, I am specially indebted to my lab colleagues Dr. Sheila Xie, Dr. Paul Chammas, Dr Bryony Leeke, Yi Xuan Low and Dr Darren Taylor for their help, advice, friendship, support and making the group a healthy environment to work in. From them, I would like to thank Alex Choromidis, for his very helpful work as a master's student and technician and a very personally enriching supervision experience, best of luck on your PhD.

Outside of my group, I would like to thank the people from the former 2nd floor of the CRB building and the current 5th floor for their support and encouragement. From them, my special appreciation for Lucas Morales, Keir Murison, Imanol Duran, Ilinca Patrascan, Dr. Teresa Buenaventura and Dr. Beatriz Fernández for their friendship and shared struggles.

My thanks to the ICL and LMS staff, past and present, that have kept the institute running, my professors at UAB and ICL for their teachings and knowledge, my first laboratory supervisor, now friend, Dr. Clara Mayayo and the public education system that has allowed me to come to this point.

Outside of work, I am forever thankful to my friends Javi, Nuria, Guille, Sergio, Blanca, Arek, Miquel, Anna, Jordi, Prarthna, and of course, Katya for their friendship and support all these years at home and from a distance, and the endless conversations, sharing experiences, joy and struggles.

No tinc paraules per agrair a la meva família, la immensa generositat i confiança dipositada en mi, així com l'esforç realitzat en proporcionar-me una bona educació i mantenir-me per la major part de la meva vida. Perdó per no haver estudiat medicina.

Finalment, estic per sempre agraït a les meves professores de biologia durant ESO i batxillerat, Maria Rosa Martorell i Cristina Matas per les seves excel·lents lliçons, així com per

compartir el seu inspirador entusiasme per la biologia. Tot plegat m'ha portat a descobrir un nou món i unes preguntes que estimo profundament, i a les quals m'omple dedicar-hi el meu temps.

Abbreviation glossary

2’3’cGAMP: 2’3’ cyclic GMP AMP.

BSA: Bovine serum albumin.

CMV: Citomegalovirus.

DAPI: 4′,6-diamidino-2-phenylindole

DNA: Deoxiribonucleic acid.

dNTPs: Deoxynucleotides triphosphate.

ERV: Endogenous retroviruses.

GOF: Gain of function.

HEK: Human embryonic kidney.

hESC: human embryonic stem cells.

ICM: Inner Cell Mass.

IF: Immunofluorescence.

IFN: Interferon.

IRF: Interferon regulating factor.

ISG: Interferon stimulated genes.

IP: immunoprecidpitation.

KD: Knock down.

KO: Knock out.

LINE: Long Interspaced Nuclear Element

LTR: Long terminal repeats

MEF: Mouse embryonic fibroblast.

mESC: Mouse embryonic stem cells.

NK: Natural Killer.

OE: overexpression.

ORF: Open reading frame.

PBS: Phosphate buffered saline.

PCR: Polymerase chain reaction.

PFA: Paraphormaldehyde.

PGC: Primordial germ cell.

PTM: Post-translational modifications.

RIPA: Radioimmunoprecipitation assay buffer.

RLT: RNeasy lysis buffer.

RNA: Ribonucleic acid.

RNP: Ribonucleoprotein.

Rosa: Reverse orientation splice acceptor.

STING: Stimulator of interferon genes.

SINE: short interspaced nuclear element

TBST: Tris buffered saline Tween 20.

TE: Transposable elements.

TPRT: Target prime reverse transcription.

UTR: Untranslated region.

WB: Western blot.

ZGA: zygotic genome activation.

Abstract

Long interspersed nuclear element (LINE1) retrotransposons comprise 17% of the human genome and some of them remain capable of retrotransposition. LINE1 is expressed in two different contexts: early development/pluripotency and pathogenic somatic contexts like senescence. In senescent cells, LINE1 expression consolidates the senescent phenotype through the retrotranscriptase and endonuclease enzymatic activities of its encoded protein, ORF2p. In a somatic context, cytosolic single-stranded DNA (ssDNA) intermediates generated by LINE1 trigger an innate immune response that establishes a state of chronic aseptic inflammation. On the other hand, LINE1 expression during pluripotency does not trigger any detrimental effect and new retrotransposition events are very rare. Moreover, LINE1 RNA is necessary for embryonic development progression. These different responses to LINE1 expression suggest an evolutionary pressure exerted by LINE1 to prevent its detrimental potential in pluripotency and allow the cell to tolerate its expression.

I have confirmed the deficiency of ssDNA sensing in mESCs via the cGAS/STING DNA sensing pathway. Expression analyses suggest that the lack of signalling seems to be caused by lack of expression in pluripotency of the DNA sensor cGAS and the innate immunity adapter STING, making mESCs incapable of producing IFN. Complementary to this, I also found that despite an attenuated response to type I IFN, activation of IRF3, the downstream effector of the cGAS/STING pathway, is enough to mount an antiviral response and drive defects in the self-renewal and pluripotency expression network of mESCs.

In parallel, I characterized the endogenous LINE1 ORF1p interactome in pluripotency. This has allowed for the description of new potential LINE1 interactors specific to this developmental stage. To identify potential repressors, I developed a screening platform, consisting of an homozygous LINE1 EGFP reporter insertion in the *Rosa26* locus of mESCs. This allowed me to identify a group of potential novel repressor candidates of LINE1 in pluripotency.

Table of contents

Chapter 1: Introduction 14

1.1 Introduction to transposable element biology 14
1.1.2. Introduction to LINE1 16
1.1.1.1. ORF1 and ORF2 proteins biology. 17
1.1.1.2. LINE1 replication cycle 19
1.1.2. TE repression mechanisms 20
1.1.2.1. Transcriptional repression of LINE1 21
1.1.2.2. Post-transcriptional LINE1 repression 22
1.1.2.3. Post-translational mechanisms suppressing LINE1 23
1.1.2.4. Post-integration repression of LINE1 integrations 23
1.1.3. Role of TEs in pathology and morbidity 25
1.1.4. TEs as an agent shaping homeostasis and genome evolution 27
1.1.5. Tools for the study of LINE1 29

1.2. The innate immune system and the cGAS/STING pathway in the detection of TE derived nucleic acids 30
1.2.1. Innate immunity is the first barrier of defence against challenges to host homeostasis 31
1.2.2. Innate immune sensors 32
1.2.2.1. TLR and NLR 33
1.2.2.2. RNA sensing in the cytoplasm 34
1.2.2.3. DNA sensing on the cytoplasm: the cGAS/STING pathway 34
1.2.2.4. Regulation of STING activation in the ER:Golgi interface 36
1.2.2.5. Other cytosolic DNA sensors 37
1.2.3. IRF3 and NFκB activation 38
1.2.4. Interferon signalling 40
1.2.5. The antiviral response 41
1.2.5.1. IFN response on cell proliferation and apoptosis 42
1.2.6. Impact of LINE1 activation cGAS/STING in somatic cells 43
1.2.6.1. Autoimmune diseases 43
1.2.6.2. Cancer and LINE1 induction of cGAS/STING 45
1.2.6.3. LINE1 and Senescence 45

1.3. LINE1 and Development 47
1.3.1. Embryonic stem cells as a model to recapitulate early development 48
1.3.2. LINE1 is necessary for early development progression 50

1.4. Current knowledge of innate immunity in development 52
1.4.1. Alternative innate immunity mechanisms in stem cells 53
1.4.2. Nucleic acid sensing in early development 55

1.5. Final considerations: Lessons from LINE1 control in the germline 57

1.6. Hypothesis 58

Chapter 2: Materials and methods 61

2.1. .Materials 61
2.1.1. Chemicals solvents and Reagents 61
2.1.2. . Kits 62
2.1.3. Antibiotics 63
2.1.4. Enzymes 63
2.1.5. Buffers, Solutions and Gels 64
2.1.6. Plasmids 67
2.1.7. Antibodies 68

2.1.8. Chemical inhibitors and activators 71
2.1.9. Tissue culture components 72
2.1.10. Media 72
2.1.10.1. Bacterial Media and Agar 72
2.1.10.2. Mammalian Tissue Culture Media 73
2.1.11. Strains/cells 73
2.1.11.1. Bacteria Strains 73
2.1.11.2. Mammalian Cell Lines 74
2.1.12. Resources and websites 74
2.2. Methods 75
2.2.1. Bacterial techniques 75
2.2.1.1. Transformation of bacteria 75
2.2.1.2. Plasmid amplification and purification 76
2.2.2. DNA Manipulation 76
2.2.2.1. Genomic DNA (gDNA) extraction 76
2.2.2.2. Polymerase Chain Reaction (PCR) 77
2.2.2.3. DNA Digestion 78
2.2.2.4. DNA Ligation 78
2.2.3. Generation of BFP pCW 57.1 constructs 79
2.2.4. Generating hSTING V155M and STING HAQ mESCs and MEF cell lines 79
2.2.5. Generation of mSTING V154M and IRF3-5D pcW 57.1 constructs 79
2.2.6. Generation of R26-LINE1-PFG construct 80
2.2.7. Generation of E14 R26-LINE1-PFG reporter cell lines 80
2.2.8. Cell culture methods 81
2.2.8.1. Maintenance of cell lines 81
2.2.8.2. Treatments of cell lines 82
2.2.8.3. ESC differentiation experiments 82
2.2.8.4. MEF isolation 82
2.2.8.5. DNA transfection 83
2.2.8.6. Colony formation essay 84
2.2.8.7. Flow cytometry 84
2.2.8.8. Virus work 85
2.2.9. Immunofluorescence (IF) 85
2.2.10. Embryo and mice work 86
2.2.11. Western blot 86
2.2.12. Immunoprecipitation 87
2.2.13. Mass spectrometry 87
2.2.14. RNA extraction 88
2.2.15. Quantitative PCR: 88
2.2.16. RNA-seq 90
2.2.16.1. RNA-seq analysis 90
Chapter 3: Results: Characterization of the sensitivity of mESCs to nucleic acids 91
3.1. Introduction 91
3.2. Results: DNA sensing in mESCs 92
3.2.1. mESCs do not react to G3-YSD and 2’3’cGAMP transfection 92
3.2.2. RNA analysis of cGAS/STING pathway components shows expression of all the components of the pathway. 93
3.2.3. WB analysis reveals no expression of cGAS and STING at protein level. 96
3.2.4. Human STING carrying gain of function mutations cannot rescue missing STING signalling in mESCS. 99
3.2.5. Mouse STING carrying gain of function mutations does not rescue STING signalling. IRF3-5D gain of function mutation does. 100
3.2.6. mSTING ER:Golgi dynamics in pluripotency 102

3.3. Discussion 105
Chapter 4: Sensitivity of mESCs and embryos to IFN signalling 106
4.1. Introduction 106
4.2. Results 107
4.2.1. mESCs are capable of upregulating intrinsically expressed ISGs but not canonical ISGs 107
4.2.2. mIRF3-5D OE in mESCs triggers an antiviral response 110
4.2.3. mIRF3-5D OE compromises mESCs self-renewal 115
4.3. Discussion 116
Chapter 5: : Results: Characterization of LINE1 dynamics in pluripotency and senescence 118
5.1. Introduction 118
5.2. 5.2. Results 119
5.2.1. LINE1 dynamics in pluripotency 119
5.2.2. LINE1 dynamics in senescence 121
5.2.2.1. Etoposide treatment induces senescence in MEFs 121
5.2.2.2. LINE1 RNA is gradually upregulated at later timepoints after senescence induction. 122
5.2.2.3. Endogenous Orf1p can be immunoprecipitated in senescent MEFs 124
5.2.3. ORF1p IP in ESCs successfully identifies ORF1p, previously known interactors and a set of potentially new LINE1 interactors. 125
5.2.4. ORF1p and MOV10 cannot be detected in senescent MEF ORF1p IP 128
5.2.5. No cytoplasmic DNA is detected in mESCs. 129
5.3. Discussion 131
Chapter 6: screening of novel LINE1 repressors in pluripotency. 135
6.1. Introduction. 135
6.2. Results 136
6.2.1. Generation of ROSA LINE1-PFG reporter cell lines 136
6.2.2. FACS screen of LINE1 interactors identifies potential repressors 139
6.2.3. γH2AX staining is a proxy for the identification of LINE1 repressors identification. 140
6.2.4. New LINE1 repressors in pluripotency are confirmed by FACS and genomic DNA qPCR. 143
6.3. Discussion 146
Chapter 7: Discussion 149
Chapter 8: Bibliography 164
Chapter 9: Annex I: Primers and siRNAs 212
Chapter 10: Annex II: list of chosen genes for siRNA screen with corresponding criteria 236
Annex III: Common LINE1 ORF1p interactors between mESCs and somatic datasets 239

Table of figures

Figure 1-1 TE phylogeny..... 15
Figure 1-2: LINE1 replication cycle. 20
Figure 1-3: LINE1-PFG retrotransposition reporter..... 30
Figure 1-4: cGAS/STING pathway schematics. 35
Figure 1-5: Consequences of the activation of the nucleic acid sensing pathway. 40
Figure 1-6: Temporal dynamics of LINE1 expression through early mouse development and correlation with global levels of DNA methylation and H3K9me3. 48
Figure 1-7: Pluripotent cells rely in alternative innate immunity mechanisms..... 53
Figure 1-8: A miRNA represses MAVS in pluripotency..... 56
Figure 1-9: Project hypothesis graphic summary..... 60
Figure 3-1: mESCs do not respond to DNA ligands and 2'3'cGAMP treatment. 93
Figure 3-2 Gene expression analysis in a pluripotency dataset. 94
Figure 3-3: mRNA expression levels of different components and regulators of the cGAS/STING pathway. 96
Figure 3-4: Protein expression of cGAS/STING components on mESCs and MEF. 98
Figure 3-5: STING gain of function mutation fails to activate an innate immune response on mESCs. 100
Figure 3-6: Mouse STING and IRF3 GOF constructs are functional 101
Figure 3-7: mSTING GOF variant cannot rescue IFN production in mESCs, but IRF3 GOF can. 102
Figure 3-8: STING GOF dynamics at the ER:Golgi interface in pluripotency. 104
Figure 4-1: mESCs cannot articulate a functional antiviral response. 109
Figure 4-2: IRF3-5D OE has broad transcriptional consequences in mESCs. 111
Figure 4-3: IRF3 activation in mESCs trigger the antiviral transcriptional programme. 113
Figure 4-4: IRF3-5D OE expression modifies the expression of pluripotency and cell-death associated genes. 114
Figure 4-5: IRF35D OE results in compromised self-renewal in mESCs..... 115
Figure 5-1: LINE1 is expressed in pluripotent mESCs..... 120
Figure 5-2: MEF undergo stable senescence upon etoposide treatment. 122
Figure 5-3: LINE1 and other markers are upregulated on later stages of senescence induction in MEF. 123
Figure 5-4: LINE1 Orf1p is detectable in senescent cells and it can be immunoprecipitated. 124
Figure 5-5: ORF1p endogenous interactome in pluripotency. 126
Figure 5-6: Cytoplasmic ssDNA identification and ORF1p IP in senescence mass spechtrometry analysis. 128
Figure 5-7: No ssDNA is detected in mESCs. 130
Figure 6-1. LINE1 activity seems inhibited in mESCs..... 137
Figure 6-2: Generation of LINE1 GFP reporter cell lines. 138
Figure 6-3: GFP LINE1 reporter and γH2AX screens were capable of identifying potential LINE1 repressors. 141
Figure 6-4: Selection of candidates to continue evaluation. 143
Figure 6-5: Detection of LINE1 GFP reporter retrotransposition events by genomic DNA qPCR. 145
Figure 7-1: Candidates Interaction network. 160

Table of tables

Table 1.1: Summary of proteins and complexes mentioned in this work repressing LINE1 activity. 25
Table 2.1: Chemicals 62
Table 2.2: Kits. 63
Table 2.3: Antibiotics 63
Table 2.4: Enzymes. 64
Table 2.5: Plasmids. 68
Table 2.6: Primary antibodies. 70
Table 2.7: Secondary antibodies. 71
Table 2.8: Chemical inhibitors and activators. 71
Table 2.9: Tissue culture components 72
Table 2.10: Transfection mix Optimem/Lipofectamine 2000 rates 83

Chapter 1: Introduction

1.1 Introduction to transposable element biology

The outbreak of the genomic era has permitted a more global overview of the genome. This has allowed us to appreciate that only a small part of mammalian DNA is represented by protein-coding genes, whereas most of the genome is composed of different types of repetitive sequences, including transposable elements (TEs). TEs are the most abundant type of repetitive sequences and represent around half of the mammalian genome. TEs have accumulated over millions of years using self-propagation mechanisms based on the cellular machinery and self-encoded proteins. For this reason, they have been considered since their discovery by Barbara McClintock, as genomic parasites (Bourque et al., 2018; McClintock, 1950).

TEs are classified into two classes based on their mechanism of transposition (Figure 1.1). Class I or retrotransposons follow a "copy-and-paste mechanism" involving a RNA intermediate that is retrocopied into the genome, creating a new element (Wells and Feschotte, 2020). They represent around 42% of the human genome and 38% of the mouse genome. Class II or DNA transposons move by a cut-and-paste mechanism. They represent 3% and 1% of the human and mouse genomes respectively. Each class is divided into subclasses based on their structure and replication mechanism (Wells and Feschotte, 2020). In the case of class I elements, which are the focus of this work, they are divided into different super-families: Long terminal repeat (LTR) elements, non-LTR elements and *Dictyostelium* repetitive sequences (DIRS). The LTR element structure is more complex and closely resembles that of retroviruses. They comprise endogenous retroviruses (ERVs) and other elements, with their characteristic replication mechanism involving an integrase. In the case of ERVs, the ERV-codified pol proteins would reverse transcribe the genomic RNA of the element in the cytoplasm within a viral-like capsid assembled from ERV-encoded gag proteins using a tRNA as a primer. Then, the integrase, which is also encoded by *pol,* binds the cDNA and translocates it into the nucleus, where it inserts the new element into the genome (Wells and Feschotte, 2020). Non-LTR elements are simpler and use an endonuclease. They comprise LINEs and the parasitic SINEs, whose replication depends upon the hijacking of LINE1 machinery. Finally, DIRS use a tyrosine-recombinase as an integration mechanism (Wells and Feschotte, 2020).

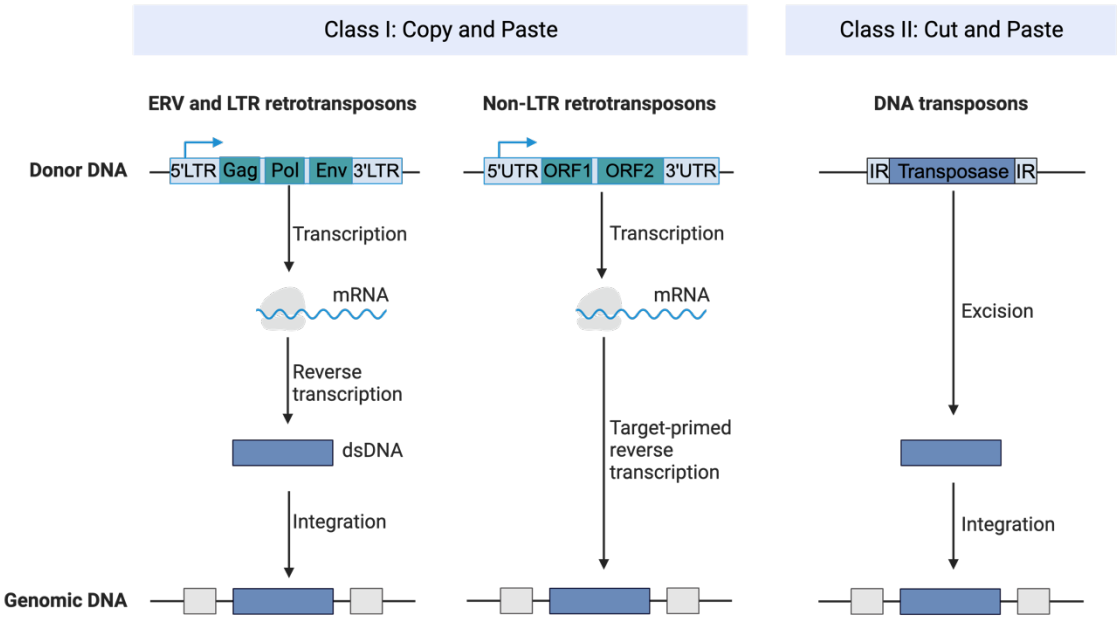


Figure 1-1 TE phylogeny There are two main classes of TEs. Class I TEs that comprise LINEs, SINEs and ERVs follow a copy-and-paste mechanism that generates a new copy. Class II TEs comprise DNA transposons and follow a cut-and-paste mechanism that does not generate a new copy. Further classification within Class I TEs depends on the replication mechanism.

As it can be seen, TEs replication cycle involves a large perturbation to homeostasis by mediating chromatin effects, triggering inflammation with their nucleic acid intermediates, DNA damage through their endonucleases and insertional mutagenesis (Cecco et al., 2019; Chen et al., 2005; Iskow et al., 2010; Jachowicz, 2023; Wallace et al., 2008). For this reason, TEs tend to be repressed by their hosts at multiple levels. However, to be maintained through evolution, TEs must also mobilize and be transmitted through the germline. Detrimental insertions are eliminated under purifying selection, co-opted elements are selected positively, but under no selection, TEs will accumulate mutations, eventually losing functionality. Hence, new insertions must generate new copies to avoid extinction (Sultana et al., 2017). At the same time TE activity cannot be too vigorous, in order to avoid compromising the hosts survival. This makes the coexistence of host and TE to be based in an equilibrium of opposing forces that repress TE activation, while maintaining windows for mobilization to occur (Philippe et al., 2016; Sultana et al., 2017). To be able to mobilize, a TE needs to be full length and possess a functional enzymatic machinery, bypass transcriptional control and host defence pathways. This greatly reduces the pool of TEs capable of creating new copies which will be later submitted to natural selection. These new copies, in turn will need to keep the aforementioned characteristics to keep propagating (Sultana et al., 2017). However, under

certain stressing contexts, like genotoxic stress or environmental changes, the mutation burden produced by TE activation, may provide a source of genetic variation that allows for rapid adaptation, especially in the context of unicellular organisms and tumour progression (Casacuberta and González, 2013; Chuong et al., 2017; Sultana et al., 2017).

TEs must therefore calibrate a delicate balance between expression and repression. This is a balance between enough retrotransposition to permit amplification, but not so vigorous as to perturbate homeostasis and lead to a fitness disadvantage for the host, which could further reduce TE propagation. This idea also aligns with the fact that TE-encoded enzymes are suboptimal for transposition (Bourque et al., 2018). Indeed, in this purely parsimonious scenario, TEs would prioritize mobilisation the germline (Allen and Nowacki, 2017; Haig, 2016). However, the development of germ cell-specific TE defence mechanism in the host, like piRNAs, makes TEs more likely to mobilise on the soma of the mammalian embryo than in the germ cell population (Di Giacomo et al., 2013; Kano et al., 2009; Pezic et al., 2014). Moreover, individual TEs can exhibit both locus and cell-type specific activation (Philippe et al., 2016).

Furthermore, the genomic distribution of TEs is not totally random, it is the result of two events: allocation of the insertion, then followed by selection (Sultana et al., 2017). TEs exhibit preference for insertion in certain genomic regions. For instance, in the case of LINE1 elements, they exhibit preference for AT-rich regions, while *Alu* sequences are enriched in GC-rich regions and the resurrected form of HERV-K exhibited preference for active transcriptional units (Sultana et al., 2019, 2017). Each element may have a distinct genomic distribution that results from the choice of integration site. However, the later selection that occurs will also be a factor in shaping the accumulation of TEs in specific genomic regions (Sultana et al., 2017).

1.1.2. Introduction to LINE1

LINE1 elements are class I, non-LTR TEs, which comprise ~17% of the human genome and ~19% of the mouse genome. In fact LINE1 is the only autonomously replicative-competent element in the human genome (Akagi et al., 2008; Mouse Genome Sequencing Consortium et al., 2002) (Gifford et al., 2013). LINE1 follows a “copy-and-paste” replication mechanism that involves the reverse transcription (3’ to 5’) of the transposon RNA to generate a new copy. A mouse retrotransposition-competent LINE1 element is around 7kb long and consists of a 5’UTR region with an RNA pol II sense and antisense promoter, two protein-coding genes, *ORF1* and *ORF2*, separated by a short intergenic region, and a 3’UTR with a polyadenylation

signal. Despite their huge representation in the genome, only a small percentage of LINE1 elements is intact and retrotransposition-competent (Beck et al., 2011; Jachowicz and Torres-Padilla, 2016; Kulpa and Moran, 2006; Sultana et al., 2017). In fact, due to their retro-copy insertion mechanisms, a lot of LINE1 copies are inactivated because of 5' truncations and accumulation of deletions (Beck et al., 2011; Jachowicz and Torres-Padilla, 2016; Kulpa and Moran, 2006; Sultana et al., 2017).

LINE1 insertion occurs normally in AT-rich regions such as intergenic spaces and introns. The presence of a LINE1 element at a particular locus can affect gene expression and even lead to the formation of alternative transcripts, which is in fact a significant contribution to the transcriptomic diversity and the different functionalities of the cells and tissues of an organism (Beck et al., 2011; Jachowicz and Torres-Padilla, 2016; Kulpa and Moran, 2006; Modzelewski et al., 2022; Sultana et al., 2017; Wells and Feschotte, 2020).

The features displayed above are generic features, LINE1 copies are not identical and some can display significant differences from their consensus sequence. LINE1 has colonized the human genome in different successive waves, which has allowed scientists to assign an evolutionary age to a LINE1 copy based on the mutational burden it has undergone. This has allowed the assignment of elements to chronologically ordered subfamilies. Old LINE1 elements have accumulated more mutations and are more divergent than young LINE1s (Beck et al., 2011; Jachowicz and Torres-Padilla, 2016; Kulpa and Moran, 2006; Sultana et al., 2017). Phylogenetic classification of mouse LINE1 has been based on its 5' UTR, which is composed of repetitive monomers of around 200bp organized into tandem repeats that follow single-copy, non-monomeric sequences (Goodier et al., 2001). These monomers differ in number and sequence among different LINE1 elements, in fact LINE1 promoter activity is proportional to the number of monomers (DeBerardinis and Kazazian, 1999). This divergence in the 5' UTR, has allowed to classify mouse LINE1s into 3 types of promoters: A, F, V and most recently Lx and N in the mouse genome (Sookdeo et al., 2013). Of these groups, A and F are the youngest and most abundant, while V and the others correspond to older elements. The promoter type has allowed scientists to classify the most active LINE1 elements into 3 families: Af (type A) and Gf and Tf (type F) (Jachowicz and Torres-Padilla, 2016; Sookdeo et al., 2013).

1.1.1.1. ORF1 and ORF2 proteins biology.

The LINE1 bicistronic transcript is translated into ORF1p and ORF2p non-stoichiometrically, with ORF1p being translated in excess, with some models suggesting one ORF2p being

translated per transcript (Kulpa and Moran, 2006; Taylor et al., 2013). These products then ensemble in cis preference to the source RNA transcript (Kulpa and Moran, 2006). Recently an additional product, named ORF0 has been described in the 5' UTR of primate LINE1, but its function is currently unknown, although it has been seen to influence LINE1 mobility, and by virtue of containing two splice donor sites, forming fusion proteins with proximal exons (Denli et al., 2015). ORF1 and ORF2 proteins can also complement in *trans* to promote the retrotransposition of retrotransposition-incompetent Short Interspersed Element (SINE) RNAs, a type of non-autonomous TE (Dewannieux et al., 2003; Raiz et al., 2012; Zhao et al., 2013), as well as cellular mRNAs, which lead to processed pseudogene formation, participating in genome expansion (Esnault et al., 2000; Wei et al., 2001).

Mouse ORF1p is a 41 kDa (40 kDa in human) nucleic acid-binding RNA chaperone that works as an homotrimer that coats LINE1 RNA and is necessary for LINE1 retrotransposition. Its coating of the LINE1 RNA has been proposed to protect it from degradation (Martin and Bushman, 2001). Besides its nucleic acid chaperone activity, capable of binding DNA, it has been proposed to facilitate the strand transfer required to anneal the DNA primer from the target side to the LINE1 RNA to prime reverse transcription (Martin and Bushman, 2001). It contains a coiled-coil trimerization domain, RNA recognition motifs and a basic charged patch in the N-terminal domain, which are all necessary for retrotransposition (Martin and Bushman, 2001). Despite being crucial for LINE1 activity, its role is still not clear. Some studies suggest that ORF1p, via its coiled coil domain, is necessary to form RNP condensates critical for retrotransposition and nuclear localization of LINE1 RNP (Freeman et al., 2019; Newton et al., 2021; Sil et al., 2023). Typically it has been observed colocalising with stress granules and processing bodies both in presence and absence of stress (Goodier et al., 2007; Hu et al., 2015; Jain et al., 2016).

ORF2p is a 150kDa protein with reverse-transcriptase and endonuclease activity (Feng et al., 1996; Luan et al., 1993). These properties are likely what necessitate its substantially lower stoichiometry in comparison to ORF1, which could allow for LINE1 retrotransposition without killing the cell. This is achieved through a still not fully understood unconventional termination reinitiation translation mechanism that discards any IRES-related mechanism that had been proposed previously (Alisch et al., 2006; Dmitriev et al., 2007). It is these two enzymatic activities which confer upon ORF2p expression a toxic activity, which is thought to be behind its very inefficient translation mechanism (Chen et al., 2005; Christian et al., 2016; Feng et al., 1996; Wallace et al., 2008). ORF2p overexpressing experiments have found it to cause a large quantity of double-strand breaks that may contribute to cell cycle arrest or apoptosis, as they greatly reduce the viability of the cell (Ardeljan et al., 2020; Gasior et al., 2006; Wallace et al.,

2008). The retrotranscriptase domain also has a significant impact on viability, as mutation of the ORF2p endonuclease domain only partially removes the negative influence on viability, as cDNA intermediates generated by ORF2p can compromise cell viability by activating the innate immune response (Cecco et al., 2019; Wallace et al., 2008). Only mutation of both domains remove all negative ORF2p consequences, and allow for its stable expression (Goodier et al., 2004; Wallace et al., 2008).

1.1.1.2. LINE1 replication cycle

LINE1 copies itself via a process known as target primed reverse transcription (TPRT) (Figure 1.2) (Cost and Boeke, 1998; Luan et al., 1993). LINE1 is transcribed in the nucleus from its 5' UTR, polyadenylated and exported to the cytoplasm. In the cytoplasm, LINE1 is translated into ORF1p and ORF2p that ensemble with cis- preference with their source RNA (Kulpa and Moran, 2006). There, other accessory proteins like PABPC1 assist in the formation of functional RNPs that are imported back into the nucleus (Dai et al., 2012). Import back into the nucleus is still a debated topic, as some authors argue that it is during mitosis, with nuclear membrane disappearance, that LINE1 can access the nucleus (Mita et al., 2018; Y. Xie et al., 2013). However, as LINE1 can retrotranspose in non-replicative cells, other mechanisms involving nuclear pore KPNA proteins, involved in the transport of molecules between the nucleus and the cytoplasm, have also been investigated (Freeman et al., 2019; Macia et al., 2017).

Upon entering the nucleus and reaching a genomic DNA locus, the ORF2p endonuclease domain recognizes a 5'-TTTT/AA-3' cleavage site and creates a single-stranded DNA break with a free 3'OH (Cost and Boeke, 1998; Flasch et al., 2019; Sultana et al., 2019). The transcript is attached via the polyA tail to the region, and ORF2p retrotranscribes the LINE1 RNA into DNA in a 3' to 5' direction. Subsequent steps of LINE1 integration like the second-strand DNA cleavage and the complementary strand cDNA synthesis, require elucidation. This mechanism involves the necessary participation of host proteins involved in DNA repair and replication, such as PARP1 and 2, PCNA and enzymes involved in the NHEJ pathway (Miyoshi et al., 2019; Taylor et al., 2018, 2013). After the integration of the cDNA, the remaining strand is synthesised using the host replication and repair machinery (Flasch et al., 2019; Miyoshi et al., 2019). TPRT ultimately leads to the integration of a new LINE1 copy at a new genomic location that is frequently 5' truncated, a 3' poly (A) tract, and flanking target-site duplications. Most retrotransposition events in dividing cells take place during S phase (Beck et al., 2011; Mita et al., 2018; Sultana et al., 2017).

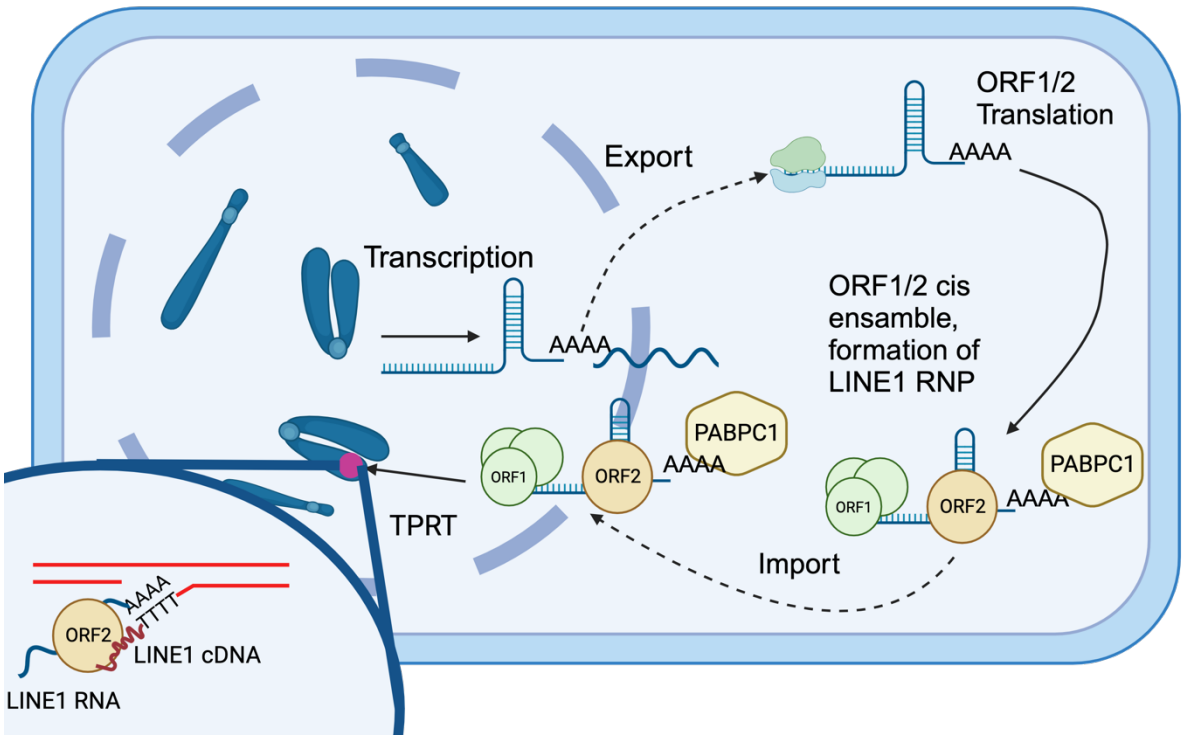


Figure 1-2: LINE1 replication cycle. Upon transcription, LINE1 RNA is exported into the cytoplasm where the two ORF proteins it encodes are translated in non-stoichiometric conditions in an unconventional termination reinitiation mechanism. The two proteins ensemble in cis from the transcript where they have been translated from. PABPC1 binds the LINE1 RNA polyA tail, other host proteins bind LINE1 or the ORFs and form the LINE1 RNP. LINE1 RNP is imported to the nucleus, there upon interaction with the chromatin via the polyA tail, ORF2 endonuclease activity nicks the host genome and uses its retrotranscriptase to retrocopy LINE1 RNA into cDNA: TPRT.

1.1.2. TE repression mechanisms

Nuclear DNA is not found in free linear strands, it is highly condensed and wrapped around nuclear proteins. This complex of DNA and proteins located in the nucleus is known as chromatin, and it forms the chromosomes of eukaryotic cells (Kornberg, 1977). Chromatin in cells exists either in an open, transcriptionally permissive state known as euchromatin, which has been characterised for having histone marks like H3K4 di- and tri- methylation and acetylation of H3K9, or in a densely compacted, more transcriptionally repressive state called heterochromatin, characterised by histone deacetylation and trimethylation of H3K9 and H3K27 (Jenuwein and Allis, 2001). Another feature of chromatin is the presence of DNA methylation, a DNA modification consisting of the addition of a methyl group to cytosine. Most CpG dinucleotides in vertebrates are methylated in somatic cells. DNA methylation is associated with repression when placed in gene promoters. This feature is used for

transposon control, in fact TE control is one of the oldest functions attributed to DNA methylation (Vanyushin et al., 1970; Yoder et al., 1997). In parallel, the gene body of highly transcribed genes is highly methylated, seemingly related to splicing (Lev Maor et al., 2015).

1.1.2.1. Transcriptional repression of LINE1

Chromatin is the first regulatory level for control of TE expression. To prevent homeostasis perturbation, TEs are typically silenced and under constitutive heterochromatin repression marks and DNA methylation (Goodier, 2016; Jachowicz and Torres-Padilla, 2016).

Going more specifically into LINE1, the potential destabilizing activity of LINE1 on host homeostasis has led to the development of a wide range of constraining mechanisms. The first mechanism involves preventing LINE1 transcription. LINE1 loci can be found within constitutive heterochromatin domains and have been found to be marked with trimethylation of H4K20, H3K9 and H3K27 in different cell types (Bulut-Karslioglu et al., 2014; Li et al., 2023; Pezic et al., 2014). H3K9me3 specifically decorates the 5'UTR of young LINE1 elements, a modification dependant on SUV39H1/H2 activity (Bulut-Karslioglu et al., 2014; Li et al., 2023; Pezic et al., 2014). SETDB1 is also involved in the deposition of H3K9me3, although its knock out leads to the upregulation of ERVs, it only weakly upregulates LINE1 (S. Liu et al., 2014). This may be a consequence of the redundancy of TE repression mechanisms, or the fact that different repression machineries target different TE families (S. Liu et al., 2014). In the absence of constitutive heterochromatin, facultative heterochromatin marked by H3K27me3 can take over in repressing LINE1 transcription (Martens et al., 2005; Peters et al., 2003; Walter et al., 2016).

Consistent with this repressive environment, the promoter of TEs is characterised for high levels of repressive DNA methylation, a mechanism though to have evolved precisely to silence TEs, established *de novo* by DNMT3a and DNMT3b enzymes (Meissner et al., 2008). Moreover, mouse male germ cells, express a fourth form of DNA methyltransferase, DNMT3c. This enzyme targets the promoters of evolutionary young TEs that burden the rodent genome, and could compromise germ cell viability, as the activity of this enzyme is necessary for mouse fertility (Barau et al., 2016). Overall, multiple layers of repression mechanisms operate over TEs in a redundant way that seemingly allows them to compensate each other in case of failure or compromise by the developmental momentum (Bulut-Karslioglu et al., 2014; Matsui et al., 2010).

Other host mechanisms that play a role in TE transcriptional repression are ncRNAs, which will be later reviewed (e.g piRNAs), and sequence-specific protein repressors like KRAB zinc-finger proteins, which have been evolved in the recognition and repression of specific families of TEs (Imbeault et al., 2017; Yang et al., 2017). These proteins contain different arrays of Zn^{2+} fingers that allow them to interact with specific DNA TE motifs and repress them thorough the KRAB domain, which acts as a scaffold for the recruitment, via the cofactor *Trim28,* of a repression complex that comprises SETDB1, DNA methyl transferases and NurD histone deacetylases, which will trigger the formation of local heterochromatin via deposition of H3K9me3 histone marks and methylation of CpG islands (Rowe et al., 2010; Tie et al., 2018; Turelli et al., 2014). In fact, Trim28 plays a critical role in the repression of a wide range of TE in mouse and human embryonic stem cells (ESC), as well as their embryonic counterparts. This activity persists in adult human cells. In the case of LINE1, this activity focuses in the 5' UTRs of old subfamilies (Rowe et al., 2010; Tie et al., 2018; Turelli et al., 2014). In some instances, young repeats evade the Zn^{2+}-finger silencing through mutations, gaining a selective advantage that results into an arms race between TEs and Zn^{2+}-fingers (Bruno et al., 2019).

1.1.2.2. Post-transcriptional LINE1 repression

In certain instances, LINE1 elements can escape the host's transcription repression mechanisms. In cases of LINE1 transcript expression, hosts have developed a network of proteins and ncRNAs that counter LINE1 retrotransposition post-transcriptionally through multiple mechanisms (Protasova et al., 2021).

Post-transcriptional mechanisms of LINE1 repression aim at reducing the level of LINE1 RNA through direct decay. These mechanisms rely on endo and exonucleases like *Dicer* and *Rnasel* (Bodak et al., 2017; Wu et al., 2019; A. Zhang et al., 2014; Z. Zhang et al., 2022), and RNA-editing enzymes like *Adar1* and *Apobec3a* deaminase activity. Editing can result in amino-acid recoding, or disrupt base pairing interaction in dsRNA secondary structures (Kinomoto et al., 2007; Orecchini et al., 2017; Wissing et al., 2011). LINE1 mobility can also be restricted by blocking translation of ORF1 or ORF2 by proteins (Tex19.1 and BRCA1), miRNAs, or obstructing RNP formation by Schlafen5 disassembling ORF1p from the RNA with its helicase activity (Ding et al., 2023; MacLennan et al., 2017; Mita et al., 2020; Tristán-Ramos et al., 2020). Mechanisms involving siRNAs, *Dicer* and piRNA are also involved in the control of LINE1 RNA levels, but due to their relevance at specific points of development relevant for this work, they will be covered later (Bodak et al., 2017; Di Giacomo et al., 2013).

1.1.2.3. Post-translational mechanisms suppressing LINE1

As a third layer, the host relies on mechanisms that directly target LINE1 at the site of potential LINE1 integrations. One mechanism for this is performed by the dNTPase activity of SAMHD1, which starves the retrotranscriptase of dNTPs, blocking retrotranscription of the LINE1 RNA (Coquel et al., 2018). Alternatively, MOV10/RNaseH2 and DNA repair pathways serve to eliminate of RNA:DNA replication intermediates (Choi et al., 2018; Xiaoyu Li et al., 2013). APOBEC3 is a well-studied protein that can deaminate LINE1 cDNA, leading to its degradation (Wissing et al., 2011). Other mechanisms involve obstruction of LINE1 copy integration by DNA repair mechanisms (Ariumi et al., 2018; Bona and Crossan, 2023; Flasch et al., 2019; Mita et al., 2020; Servant et al., 2017; Ardeljan et al., 2020). Repression of LINE1 also can involve the targeting for degradation by ubiquitin-ligases of proteins involved in TPRT, like ORF2p and PCNA (Ariumi et al., 2018; Zhen et al., 2023).

A lot of LINE1 RNPs accumulate in stress granules and processing bodies, but their role is still controversial (Briggs et al., 2021; Goodier et al., 2007). These are structures that play a big role in RNA metabolism and translation regulation. Stress granules form in response to stress, for example oxidative stress or heat shock, accumulate RNA and proteins that are stalled during translation initiation. Processing bodies are formed independently of stress, and contain effectors involved in RNA degradation, transcript silencing, RNA storage and translation repression (Ivanov et al., 2019). Several LINE1 repressors like MOV10 and SAMHD1 are involved in formation, sequestering and potential posterior degradation of stress granules, as they are assumed to be a mechanism to inhibit LINE1 (Hu et al., 2015; Xiaoyu Li et al., 2013).

Development is an interesting case regarding LINE1 repression, since the transcriptional repression mechanisms are released at certain developmental points. This will be discussed further down in this chapter.

1.1.2.4. Post-integration repression of LINE1 integrations

In case of a successful LINE1 copy integration, recent work has revealed the role of the HUSH complex in silencing LINE1 integrations. HUSH is a complex formed by TASOR, MPP8 and Periphilin. HUSH can discriminate newly inserted genomic sequences and silence them via H3K9me3 deposition, thereby inhibiting transgene expression. HUSH recognises sequences that are larger than 1.5kb, intronless and with high adenine content in their sense strand as exogenous. Therefore, HUSH particularly targets young full-length active LINE1 elements and transgenes (Robbez-Masson et al., 2018; Seczynska and Lehner, 2023; Tunbak et al., 2020).

Overall, this section highlights the diversity and redundancy of LINE1 repression mechanisms operating throughout the different layers of LINE1 life-cycle, forming a robust network that guards the genome. Key repressive factors/pathways that repress LINE1 are summarised below (Table 1.1).

	Activity	References
Adar	RNA deaminase activity catalyses the conversion of adenosine into inosine, deaminase independent mechanism	(Orecchini et al., 2017)
Apobec3	DNA and RNA deaminase activity catalyses the conversion of cytosine into uracil, deaminase independent activity	(Kinomoto et al., 2007; Wissing et al., 2011)
Ago2	RNA degradation	(Aporntewan et al., 2011)ra
BRCA1	ORF2 translation inhibition and prevention of LINE1 insertion into the replication fork	(Mita et al., 2020)
Dicer	LINE1 RNA degradation	(Bodak et al., 2017)
Dnmtc3a,b	*De novo* methylation of CpG dinucleotides on TE promoters.	(Meissner et al., 2008)
Dnmt3c	*De novo* CpG methylation of evolutionary young TEs promoters in the rodent male germline.	(Barau et al., 2016)
ERCC1	Cleavage of LINE1 replication intermediates	(Bona and Crossan, 2023)
FANCA/FANCD	Cleavage of LINE1 replication intermediates	(Bona and Crossan, 2023)
HELZ2	Destabilizes LINE1 RNA	(Luqman-Fatah et al., 2023)
HUSH complex	Deposits H3K9me3 in newly inserted sequences.	(Robbez-Masson et al., 2018; Seczynska and Lehner, 2023; Tunbak et al., 2020)
MOV10	Stress granule formation and RNA: DNA duplex degradation	(Choi et al., 2018; Xiaoyu Li et al., 2013)
Rad18	E3 ubiquitin-ligase of pro-TPRT PCNA	(Ariumi et al., 2018)
RNaseH2	RNA:DNA duplex degradation	(Benitez-Guijarro et al., 2018; Choi et al., 2018)
RNaseL	RNA degradation	(A. Zhang et al., 2014)
SAMHD1	Stress granule formation, dNTP ORF2 starvation	(Coquel et al., 2018; Hu et al., 2015)
Schlaffen5	Prevents formation of RNP	(Ding et al., 2023)
SETDB1	Deposition of H3K9me3	(S. Liu et al., 2014)
SUV39H1/H2	Deposition of H3K9me3	(Bulut-Karslioglu et al., 2014; Li et al., 2023; Pezic et al., 2014)
TDP-43	Unknown	(T. D. Li et al., 2022; E. Y. Liu et al., 2019)
Tex19.1	Blocks ORF1 translation	(MacLennan et al., 2017)
Trim28/ Kap1	Establishment of a repressive chromatin environment via DNA methylation and H3K9me3	(Rowe et al., 2010; Tie et al., 2018; Turelli et al., 2014)

Trim41	ORF2p degradation	(Zhen et al., 2023)
Zcchc3	RNA degradation	(Z. Zhang et al., 2022)
Zcchc8	RNA degradation	(Wu et al., 2019)

Table 1.1: Summary of proteins and complexes mentioned in this work repressing LINE1 activity.

1.1.3. Role of TEs in pathology and morbidity

There are multiple ways in which TEs can alter the biology of the cell, specifically LINE1, which is the only active TE in the human genome (Hancks and Kazazian, 2012). TEs can modify the genome through insertional mutagenesis and chromosomal rearrangements. They can also affect gene expression, acting as alternative promoters, creating or disrupting polyadenylation signals, splicing sites or generating chimeric transcripts (Deniz et al., 2020; Iskow et al., 2010; Tubio et al., 2014). Moreover, the endonuclease activity of some TE-encoded proteins such as LINE1 ORF2p can produce double-strand breaks (DSB) and through its retrotranscriptase activity, an aseptic inflammation. Overall these activities have a significant impact on the viability of the cell (Ardeljan et al., 2020; Belancio et al., 2009; Chen et al., 2005; Gasior et al., 2006). TE repressive mechanisms are reprogrammed in early development and germ cell development, opening a window for TE expression. TE mobilization in the germline will contribute to heritable changes, and in the developing embryo will produce somatic mosaicism. This activity has the potential to generate genetic variation between individuals such as has been seen in the brain and other organs (Beck et al., 2010; Ewing and Kazazian, 2010; Kano et al., 2009; Muotri et al., 2005; Newkirk and An, 2017).

The de-repression of TE activity, most significantly LINE1 activity in humans, has therefore been clinically related to many genetic diseases. This includes sporadic haemophilia cases, chloridaemia and type I neurofibromatosis (Ardeljan et al., 2020; Kazazian et al., 1988; van den Hurk et al., 2007; Wimmer et al., 2011). Recorded cases are reviewed in (Chen et al., 2005). Other TEs like SINEs, LINE1 parasites, have been linked to disease pathogenesis, with many SINE insertions having been mapped next to loci contributing to human diseases (Hancks and Kazazian, 2016, 2016; Payer et al., 2017). In recent years, the role of LINE1 in cancer has been gaining increasing prominence.

LINE1 expression has also been linked to many neurological disorders. Already in the early 2000s LINE1 activity was reported in neurons and its precursors (Muotri et al., 2005). LINE1 activity was linked to Rett syndrome when it was discovered that MECP2, a protein that promotes methylation-dependent repression of LINE1 by targeting the 5' UTR was found mutated in patients with Rett syndrome. This correlated with increased LINE1 expression and susceptibility to LINE1 retrotransposition. *In vitro* models derived from patients confirmed this

correlation, which was later confirmed by mouse studies and post-mortem analysis of patient samples (Jacob-Hirsch et al., 2018; Muotri et al., 2010; Boxun Zhao et al., 2019). However, despite multiple studies reporting a relation between LINE1 activity and neurological disorders, none of them has established a relationship between the increased LINE1 activity and the disease aetiology.

Early in the 1980s, cancer was described as a disease with a genetic basis caused from accumulated genetic mutations (Reddy et al., 1982). Due to its capacity for insertional mutagenesis, it was early on devised that particularly LINE1 could be a contributor to the accumulation of mutations.

General hypomethylation occurs in tumorigenesis and also involves LINE1 elements, as LINE1 expression is reactivated in multiple cancers and can contribute to their progression (Gualtieri et al., 2013; Rodić et al., 2014). LINE1 insertions have been reported to disturb tumour suppressor genes and drive proto-oncogenes by providing an alternative site of transcription or severing a negative regulation site (Miki et al., 1992; Morse et al., 1988; Shukla et al., 2013). Overall, the inactivation of tumour suppressor genes seems to be most common phenomenon (Lee et al., 2012). LINE1 activity differs widely among different types of cancer and stage of tumour progression (Lee et al., 2012). TE somatic insertions tend to occur in commonly mutated genes in cancer and cancer-specific DNA hypomethylated regions (Lee et al., 2012). Hence LINE1 hypomethylation was reported to be a marker of increased predisposition and poor prognosis in several cancers such as colorectal cancer, hepatocarcinoma, gastric cancer and lung cancer (Iskow et al., 2010; Kamiyama et al., 2012; Shigaki et al., 2013; Zhu et al., 2014). These mechanisms have also been observed in other mammals, such as in the infectious tumours found in Tasmanian devils and dogs (Hayes et al., 2023; Stammnitz et al., 2023).

Other TE-driven mutagenesis mechanisms are chromosome rearrangements triggered by somatic recombination between two LINE1 elements, or interstitial deletions caused by retrotransposition, leading to deletion of tumour suppressor genes, duplication of oncogenes or oncogenic fusion proteins (Gilbert et al., 2005; Han et al., 2005; Tubio et al., 2014). In addition, LINE1 may be transcribed downstream of its 3' UTR and retrotranspose into a target site in a process known as 3' transduction. In this process LINE1 transcript may incorporate downstream regulatory sequences or genes, which may have pro-oncogenic potential. At the same time, this process also contributes to the dissemination of genes and regulatory elements to new locations (Tubio et al., 2014). In addition, LINE1 can generate chimeric transcripts driven from sense or antisense promoters within or outside LINE1 sequences,

generate LINE1-derived regulatory RNAs or interfere in the host gene transcription by slowing down RNA pol II elongation, inducing premature termination and generate alternative splice sites or interfering in the splicing process (Cruickshanks et al., 2013; Xiao-Jie et al., 2016; Xu et al., 2023). Finally, TE-harboured regulatory regions can also have carcinogenic potential, as TEs such as ERVs can provide enhancers when dysregulated, altering the expression of adjacent genes involved in carcinogenesis (Deniz et al., 2020).

1.1.4. TEs as an agent shaping homeostasis and genome evolution

Even though most TEs no longer transpose, the continuous genomic infiltration, propagation and degradation of TEs has promoted genetic innovation at an evolutionary scale. TEs contain regulatory elements such as promoters, transcription factor binding sites and polyadenylation signals; and can encode proteins such as capsids, polymerases, transposases and envelopes (Miyawaki et al., 2020; Modzelewski et al., 2022). Because of this, TEs can also be a source of new host functions, a process known as co-option, and spread regulatory sequences in new areas of the genome. As mentioned before, once integrated, TEs are subject to natural selection (Bourque et al., 2018; Chuong et al., 2016; Frost et al., 2023; Mi et al., 2000; Modzelewski et al., 2022; Sakashita et al., 2020; Senft and Macfarlan, 2021). Most TE insertions are in intergenic regions and will be subjected to genetic drift and the few insertions in genes will be subjected to purifying selection, but in some cases TE integration can be selected positively for due to providing novel host functions as a source of new protein-coding genes, lncRNAs, cis-regulatory elements, etc that provide an adaptative advantage (Chuong et al., 2016; Cosby et al., 2021; Frost et al., 2023; Kruse et al., 2019; Modzelewski et al., 2022; Todd et al., 2019).

TE domestication has contributed to conserved and specific functions. In occasions the host has harnessed a TE encoded gene and conserved its function, while the rest of the element degenerated. An example of this is the DNA transposase protein, which has been co-opted multiple times into the emergence of new functions, like the *Rag* genes that catalyse V(D)J somatic recombination, key in the development of the vertebrate adaptative immune system (Huang et al., 2016; Jangam et al., 2017). Equally fascinating are the variety of functions derived from the co-option of ERV genes: *gag*, *pol* and *env*, as they have been identified to play a role in placentation and neuronal development. From *Env* gene co-option has been derived Syncitin, a protein necessary for the cell-cell fusion and formation of the syncitiotrophoblast (Mi et al., 2000). A *Gag* protein derived from a vertebrate lineage of retrotransposon: Ty3/*gypsy* retrotransposons, has been domesticated as Arc, a protein involved in synaptic plasticity (Ashley et al., 2018; Pastuzyn et al., 2018). This process is not

limited to protein-coding genes, as TEs can generate lncRNAs that act as additional layers of regulation into a cellular process (Bartonicek et al., 2022; Percharde et al., 2018). Other common mechanisms of TE contribution to gene diversity are exonization events of intronic TEs and 3' transductions (Schmitz and Brosius, 2011; Xing et al., 2006). Finally the retrotranscriptase activity of some TEs can generate processed retrogene copies that can be a source of additional functions (Carelli et al., 2016).

On average, around 20% of cell or tissue-specific, active chromatin elements in human, mouse and zebrafish are within TEs (Miao et al., 2020; M. Xie et al., 2013). Hence, TEs can be a source of alternative promoters and splicing sites and generate novel isoforms of pre-existing proteins that play important roles during development. For example, an isoform of *Cdk2ap1* derived form a TE alternative promoter, peaks at pre-implantation development and promotes proliferation, in contrast to the canonical form that peaks at mid gestation and represses proliferation (Modzelewski et al., 2021). Furthermore, LINE1 maintains naïve lymphocyte CD4 quiescence by being spliced into T CD4 cells activating gene transcripts, rendering them non-functional until activation of the cells promotes the splicing of the canonical variants (Marasca et al., 2022). Notably, these cases involve a strategy that attenuates the function of the canonical protein when needed in a specific developmental context, something perhaps inherent to the nature of the process.

Concerning other regulatory sequences, 37% mouse and 45% human enhancers are predicted to be TE-derived (Simonti et al., 2017). As mentioned before, TE-based spreading of regulatory elements has make them a rich resource for the modulation of eukaryotic gene expression (Bejerano et al., 2006; Chuong et al., 2016, 2013; Liang et al., 2023; Thompson et al., 2016; Trizzino et al., 2017). As they are spread throughout the genome, TEs can donate the same regulatory sequence to multiple genes dispersed throughout the genome, at the same time. This has therefore provided the foundation for the assembly of large genetic regulatory networks along evolution, a phenomenon anticipated many years ago (Davidson and Britten, 1979).

Since then, TE-dependent gene regulation has more recently been described in germ cells and pre-implantation embryos, which are characterized by epigenetic-reprogramming events that allow for TE induction (Durruthy-Durruthy et al., 2016; Fadloun et al., 2013; Hendrickson et al., 2017; Macfarlan et al., 2012; Ohno et al., 2013; Percharde et al., 2018; Sakashita et al., 2020; J. Wang et al., 2014). Nevertheless, TE-dependent gene regulation is also observed in other developmental contexts like neocortex development, the innate immune response and lymphocyte development (Chuong et al., 2016; Notwell et al., 2015; Ye et al., 2020). Moreover,

processes like placentation are also tightly regulated by TE-derived enhancers (Chuong et al., 2013; Frost et al., 2023).

1.1.5. Tools for the study of LINE1

A key requirement for studying LINE1 activity is being able to quantify retrotransposition and not just expression of LINE1. The most used tool for the studying of LINE1 biology through its retrotransposition capacity was developed by Moran et al: a retrotransposition reporter. It consists of an engineered LINE1 expression vector, containing a full-length codon optimised LINE1 expression cassette followed by an inverted reporter interrupted by an intron and driven by its own inverted promoter within the 3' UTR (Figure 1.3). When the construct is expressed, the intron is spliced and if the reporter is successfully retrotransposed by the encoded ORF1p/ORF2p, it will be integrated into the genome. Then, the reporter will be transcribed from its own promoter, making the cell discernible. This reporter can be of different natures such as an antibiotic resistance cassette or fluorescent protein (Figure 1.3) (Freeman et al., 1994; Moran et al., 1996; Ostertag and Kazazian Jr, 2001; Xie et al., 2011). The fact that the reporter is located in the 3'UTR of LINE1 makes it more likely that it will be inserted since integration is in a 3' to 5' direction. However it is notable that this readout is still an underestimate as most of LINE1 insertions are truncated below 500bp long (Garcia-Cañadas et al., 2023).

Antibiotic resistance cassettes typically require a 7- 14-day selection. For this reason EGFP-based reporters have been favoured as they allow a more immediate in real-time visualization (Ostertag et al., 2000). Native LINE1 elements have high adenosine content, which is thought to cause their inefficient expression because of a transcription elongation defect. The construction of a codon-optimized synthetic LINE1 (ORFeus) that reduced the adenosine content, lead to a larger transcript abundance of the construct and an increased retrotransposition rate (An et al., 2011, 2006; Han and Boeke, 2004).

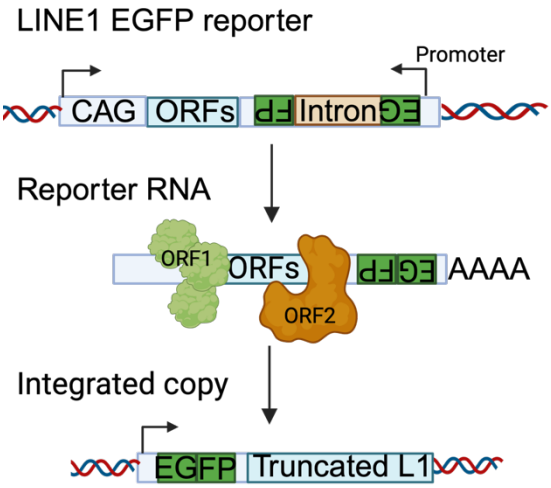


Figure 1-3: LINE1-PFG retrotransposition reporter. However, LINE1 insertion silencing by the HUSH complex and the epigenetic machinery, or insertion into an heterochromatic region, could also affect the reporter expression, leading to an underestimate of the real retrotransposition rate(Del Re et al., 2010; Garcia-Perez et al., 2010; Liu et al., 2018; Seczynska et al., 2022). In addition, the inclusion of an intron within the LINE1 3' UTR may alter LINE1 biology, as LINE1 lacks introns and is not processed by the splicing machinery. PFG stands for the reverse writing of GFP.

Normally transfected as a plasmid, the reporter consists of a synthetic (codon-optimised) LINE1 under the control of either the native 5'UTR promoter or a CAG promoter. The 3' UTR is interrupted by an EGFP inverted copy containing an intron and under the control of its own promoter, normally a CMV promoter. GFP therefore is not expressed. Transcription of the LINE1 reporter splices the intron out. In case of successful integration, EGFP is copied back in a sense direction and is intronless, allowing for its transcription and subsequent expression.

1.2. The innate immune system and the cGAS/STING pathway in the detection of TE derived nucleic acids

Having discussed the negative impacts to the cell of uncontrolled activity of TEs such as LINE1, it is important to introduce how LINE1 expression and/or activity is sensed in the cell. Immune activation is one consequence of this, which can serve to limit the effects of TE upregulation and as a source of detrimental consequences of LINE1 expression. How LINE1 nucleic acid intermediates are detected is centrally related to this thesis, since as described later, I set out to understand why LINE1 escapes immune surveillance in early development. In this section, I introduce how TEs such as LINE1 are sensed, what are the consequences of detecting increased LINE1 activity such as activation of the immune system, before finally describing the curious case of early development where LINE1 appears to be expressed without the activation of these pathways.

1.2.1. Innate immunity is the first barrier of defence against challenges to host homeostasis

One challenge that all organisms have to deal with is the elimination of microorganisms and abnormal or damaged cellular material to maintain homeostasis (M and Casey, 2022). Natural selection has selected those mechanisms that eliminate these challenges with the lowest possible homeostasis perturbation. This is the physiological role of the immune system, to protect the organism against microorganisms and substances identified as exogenous to the organism, and the cells and molecules responsible of immunity constitute the immune system. Immunity has been classified into two big categories: innate and adaptative (M and Casey, 2022).

The adaptative immunity is stimulated by the exposure to a specific antigen, and results in a specific response, such as the production of antibodies, that increases in magnitude and defensive capacity with every successive exposition to the antigen(M and Casey, 2022). This system develops during the life of an individual. For the purpose of this work, however, I will focus in the innate immune system as this is the most relevant for TE surveillance and response (M and Casey, 2022).

The innate immune system is an evolutionarily conserved system acting as a first-line of defence against invading microbial pathogens and other potential threats to the host. It consists of cellular and molecular mechanisms that are permanently active (M and Casey, 2022). These mechanisms are the anatomical barriers (epitheliums and mucosal surfaces), the soluble proteins of the complement system and the phagocytic cells (M and Casey, 2022).

The diversity of features recognized by the innate immune system is limited when compared to the adaptative immune system. Nevertheless, it can recognize any microorganism because of innate immune receptors called pattern recognition receptors (PRRs) that recognize conserved molecular features shared by different groups of microorganisms, which themselves are known as pathogen-associated molecular patterns (PAMPs) (M and Casey, 2022). These patterns are present in microbes such as viruses, bacteria, parasites and fungi, and can rarely mutate without a significant fitness cost to the microbe and are hence highly invariable. In contrast to adaptative receptors, PRRs are encoded and inherited through the germline, and therefore their specificities have been shaped throughout evolution in contrast to the random generation of antigen receptors of the adaptative immune system (M and Casey, 2022). Examples of PAMPs can be: bacterial lipopolysaccharide, flagellin, fungal

chitin, uncapped RNA, dsRNA and ssDNA. PAMPs concerning nucleic acids are specifically addressed later on this section (M and Casey, 2022).

As well as recognising PAMPs, PRRs are involved in sensing endogenous ‘danger’ signals by recognizing danger-associated molecular patterns (DAMPs) (M and Casey, 2022). These comprise host molecules released in case of damage, which can also activate the innate immune receptors. DAMPs can be a consequence of the infection; but they can also be induced by tissue damage or neoplastic transformation (M and Casey, 2022). Examples of DAMPs can be released mitochondrial DNA, released histones, heat-shock proteins, cytokines and reactive oxygen species (M and Casey, 2022).

The recognition of PAMPs or DAMPs by the PRRs triggers an inflammatory response which stimulates effectors while also helping to initiate an adaptive immune response (M and Casey, 2022). Inflammation has three essential roles in the context of infection: deliver additional effector molecules and cells from the blood, induce local blood clotting to avoid physically the spread of the infection, and to promote tissue repair after the infection has been cleared. Innate inflammatory responses include the secretion of cytokines/chemokines and lipid mediators, the induction of antimicrobial peptides, pyroptotic cell death and the recruitment of phagocytic cells that can also release more cytokines (M and Casey, 2022). However, phagocyte activation can also cause extensive tissue damage because of the hydrolytic enzymes, membrane-disrupting peptides and ROS that can be released into the extracellular environment (M and Casey, 2022).

The aforementioned changes are mediated by a variety of inflammatory mediators: cytokines and chemokines, which act on other immune cells (M and Casey, 2022). These cells are then recruited into target tissues to provide specific types of effector activities, such as killing or production of cytokines that have direct antiviral activity, all aimed to reduce or eliminate infection by pathogens (M and Casey, 2022). Chemokines are a subgroup of cytokines that work as chemoattractant for phagocytes and lymphoid cells that will attract immune cell types into the inflammation site. In the next section I will describe these PRRs and explain how they work in the context of TE detection (M and Casey, 2022).

1.2.2. Innate immune sensors

Because of their replication dynamics, retrotransposons can make immunogenic nucleic acids that activate the innate immune system of the cell. Their RNA can form secondary structures that activates the RNA sensing pathways, while their cDNA intermediates can activate the

DNA sensing pathways, mainly the cGAS/STING pathway, resulting in the activation of the innate immune system, generating an aseptic inflammation (Ardeljan et al., 2020, 2019; Cecco et al., 2019; Gázquez-Gutiérrez et al., 2021; Simon et al., 2019; Tunbak et al., 2020; Zhao et al., 2021). In the coming section, I will review the known innate immunity pathways involved in the detection of immunogenic nucleic acids derived from virus and TEs.

1.2.2.1. TLR and NLR

Toll like receptors (TLRs) were the first mechanism of PAMP recognition discovered. There are 10 receptors expressed in human and 12 in mice, each of them recognizing one or more distinct molecular pattern that is a component of a pathogenic or non-pathogenic microorganism, not generally found on vertebrate cells (Kawai and Akira, 2010). All these patterns are highly conserved, for example lipopolysaccharide (LPS) on the outer membrane of gram-negative bacteria (Kawai and Akira, 2010). TLRs are sensors of microbes present in the extracellular media, which is why they are localised on the surface of the plasma membrane or in the membranes of endosomes, where they detect pathogens or their components that have been taken into cells by any form of phagocytosis. In these internal compartments are located TLR7 and 8, which detect viral ssRNA, along with TLR3 that detects dsRNA and TLR9 that detects ssDNA sequences that contain unmethylated CpG dinucleotides characteristic of bacteria and viruses (Ewald et al., 2008; Heil et al., 2004; L. Liu et al., 2008, p. 3).

Activation of the TLR receptors initiates a signalling cascade, different for each TLR. In most cases, signalling will activate NFκB and interferon regulated factors (IRFs), which are transcription factors that also activate activator protein 1 (AP-1). NFκB and AP-1 primarily induce expression of pro-inflammatory cytokines and chemokines. IRFs factors primarily upregulate the expression of Type 1 Interferons (IFNs) (described in detail later).

Nucleotide-binding oligomerization domain- like (NOD) like receptors (NLRs) are a group of cytosolic innate sensors. NOD ligands may enter the cytoplasm as a result of intracellular infection or from materials captured by endocytosis. That said, not all of them are involved in pathogen detection (Chen et al., 2009).Some NLRs activate NFκB in the same way as TLRs, leading to the synthesis of pro-inflammatory mediators but also enzymes involved in the production of nitric oxide, which is toxic for intracellular parasites (Ting et al., 2010). Some NLRs can response to infection or cellular damage by inducing the formation of the inflammasome. These complexes are proteases that activate inflammatory caspases and some cytokines like IL-1β and IL18, and gasdermin D, a cytosolic protein that inserts into the

plasma membrane forming a pore. The function of the pore is to release the aforementioned cytokines and induce the cell death in a process known as pyropoptosis (He et al., 2015; Kayagaki et al., 2015; Shi et al., 2015).

It has also been suggested that TEs could act as amplifiers of TLRs signalling. It has been shown by a recent preprint that TLR4 activation by LPS can activate the transcription of LINE1 elements, that down the line amplify the immune response by generating nucleic acid intermediates that activate other PRRs (Rookhuizen et al., 2021).

1.2.2.2. RNA sensing in the cytoplasm

TLRs are primarily sensors of extracellular microbial products. Additional families of innate sensors have been identified to detect the products present in the cytosol. It is the case of RIG-I-like receptors, that serve as viral sensor for different types of viral dsRNA. Importantly here, they have also been shown to sense TE-derived RNA. RIG-I detects uncapped RNAs (most viral RNAs are not produced in the nucleus and are as a consequence not capped) and MDA5 senses long dsRNAs (Rehwinkel and Gack, 2020) LINE1 RNA is typically capped , therefore its RNA is mainly detected by MDA5 (Dmitriev et al., 2007; Tunbak et al., 2020). Once the pattern is detected, the sensors assemble into filaments around the RNA, promoting the aggregation of their signalling domains whose interactions are stabilized by polyubiquitin chains. The signalling domains interact later with MAVS and activate it. Activated MAVS also polymerises in the mitochondria, where is located, and propagates the signalling by recruiting TRAF E3 ubiquitin ligases. The cascade ends ultimately in the activation of NFκB and IRF3/7 (M and Casey, 2022).

1.2.2.3. DNA sensing on the cytoplasm: the cGAS/STING pathway

The cGAS/STING pathway is gaining increasing interest for its role in detecting TE-derived DNA in the cytoplasm of cells with high TE activity. cGAS is a cytosolic dsDNA sensor bound to the nuclear membrane and activated independently of sequence upon the recognition of B-type double-stranded DNA of at least 20bp *in vitro*, RNA:DNA hybrids, or ssDNA (Figure1.4)(Barnett et al., 2019; Civril et al., 2013; Herzner et al., 2015; Xin Li et al., 2013; Luecke et al., 2017; Mankan et al., 2014; Sun et al., 2013). Actually, cGAS has lower affinity for other DNA conformations like Z-DNA, a nucleic acid structure used as a strategy by several viruses to avoid detection (Zhao et al., 2023). This cytosolic DNA can have different sources: genomic DNA products of genotoxic stress, double strand brakes or micronuclei,

mitochondrial DNA, viral or bacterial DNA, and DNA intermediates TE reverse transcription (Figure 1.4).

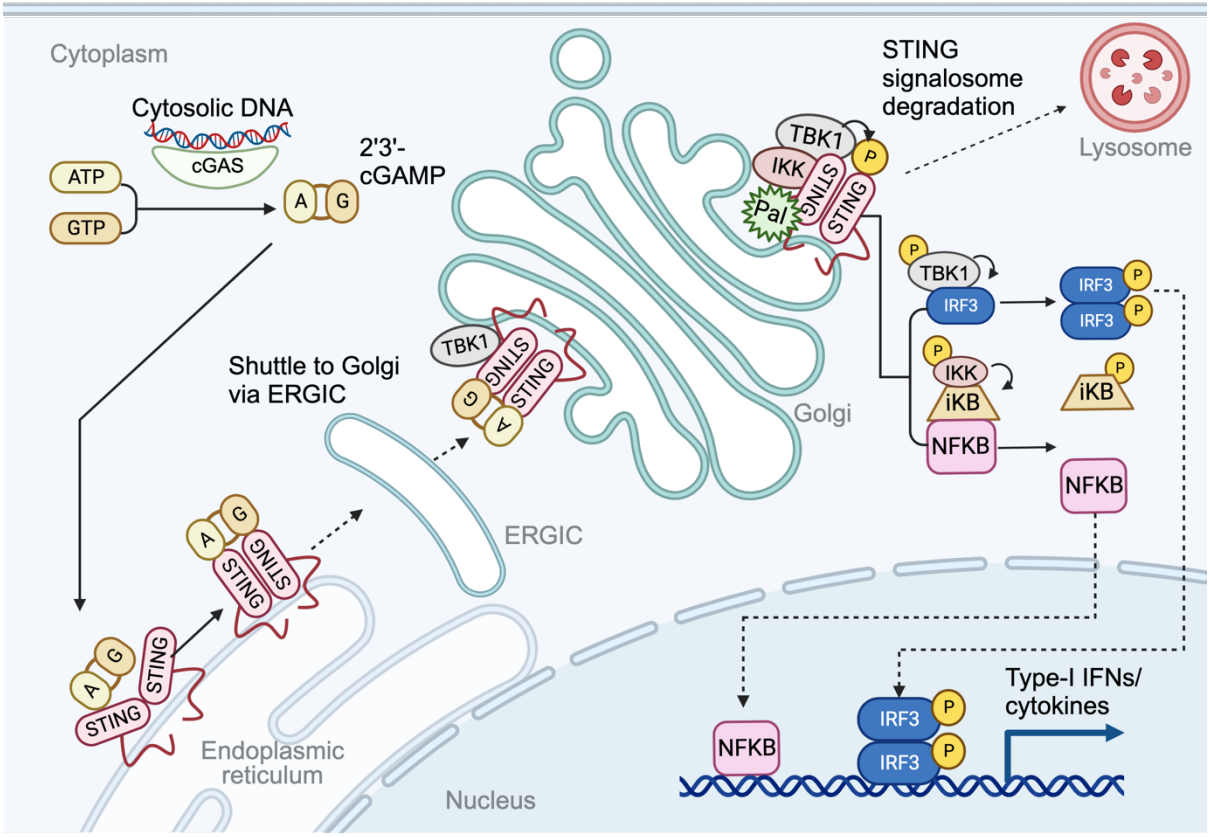


Figure 1-4: cGAS/STING pathway schematics. cGAS senses cytosolic DNA and catalyses the formation of 2'3'cGAMP that activates STING at the ER. STING undergoes a conformational change and is translocated to the Golgi via the ERGIC where it binds to Tbk1. At the Golgi. At the trans Golgi compartment STING also undergoes palmytoylation, which will enable the clustering of STING into signalosomes. STING is phosphorylated by Tbk1/IKK, which also get phosphorylated by contiguous STING dimers interacting with Tbk1/IKK. Phosphorylated Tbk1 and IKK will activate IRF3 and the Nfκb signalling pathway. IRF3 gets phosphorylated by Tbk1 and dimerizes. IKK phosphorylation of IκB releases Nfκb. Both transcription factors translocate into the nucleus and activate the expression of type I interferons and cytokines. Straight lines represent enzymatic reactions and conformational changes. Dotted lines represent shuttling or movement. Yellow circles marked with a "P" stand for phosphorylation and "Pal" stands for palmitoylation. Made with BioRender.

cGAS proteins oligomerize around DNA molecules and undergo liquid-phase condensation, forming a ladder structure that is thought to allow cells to recognize long DNA, of potential higher danger (Andreeva et al., 2017; X. Zhang et al., 2014; Zhou et al., 2018). The binding of cGAS to DNA catalyses the formation of cyclic 2'-3'GMP-AMP from ATP and GTP (2'3'cGAMP)(Diner et al., 2013; Gao et al., 2013). Activation of cGAS is highly regulated by post-translational modifications, ion concentrations like Zn^{2+} and Mn^{2+}, as well as by proteins that prevent its activation by genomic DNA, especially during replication, or that enhance its activation in case of viral infection (Barnett et al., 2019; Du and Chen, 2018; Guey et al., 2020;

Hertzog and Rehwinkel, 2020; Li et al., 2021; Pathare et al., 2020; Seo et al., 2018; Wang et al., 2018; Xia et al., 2016; Yu et al., 2021; Zhang et al., 2020). Intriguingly cGAS can also be inhibited by other DNA sensors like DNA-PK and AIM2, probably to prevent an overactivation of the innate immune response (Figure 1.4) (Banerjee et al., 2018; Burleigh et al., 2020; Wang et al., 2017).

2'-3'cGAMP from intracellular or extracellular origin binds the stimulator of interferon genes (STING), a transmembrane dimeric protein localized in the endoplasmic reticulum (ER) (Figure 1.4). Upon cGAMP binding, STING undergoes oligomerization fostered by disulfate bridges (Ergun et al., 2019; Shang et al., 2019; Zhou et al., 2020). In normal circumstances STING is tethered by STIM1, but upon activation, STING traffics using the COP II canonical mechanism to the ER:Golgi intermediate compartment (ERGIC), where according to some reports it recruits Tbk1 without yet activating it (Gui et al., 2019; Saitoh et al., 2009; Srikanth et al., 2019, p. 1). Then, STING and Tbk1 traffic to the Golgi (Decout et al., 2021; Taguchi et al., 2021)

At the Trans Golgi Network (TGN), STING undergoes palmytoilation at Cys-88 and Cys-91, allowing the clustering of STING into lipid rafts, forming peri-nuclear punctate structures (Hansen et al., 2019; Mukai et al., 2016; Saitoh et al., 2009; Taguchi et al., 2021). It is at this point that STING forms its activated complex. Trafficking STING binds to TBK1 and IKK, forming a signalosome at the TGN that is activated when TBK1 phosphorylates the contiguous STING (Ser-365 in mouse STING, 366 in human STING) and TBK1 molecules. The activated signalosome phosphorylates IRF3 (by TBK1) and inhibitor of NFκB (IκB) by IKK (also known as NEMO) (Decout et al., 2021; Liu et al., 2015; Baoyu Zhao et al., 2019). Oligomerization of STING is paramount, as one STING dimer cannot be at the same time the docking site of a Tbk1 and its phosphorylation target (Zhang et al., 2019). Activated, phosphorylated IRF3 undergoes dimerization, while phosphorylation of IKB marks it for degradation, which leads to the release of the canonical NFκB complex. Both are then translocated into the nucleus as transcription factors of type I interferons and other cytokines. Activated cGAS and STING oligomers later undergo polyubiquitination and are degraded in lysosomes (Figure 1.4) (Gonugunta et al., 2017; Gui et al., 2019).

1.2.2.4. Regulation of STING activation in the ER:Golgi interface

Due to the potential consequences that it can have on the cell, the activation of cGAS/STING signalling is tightly regulated also with the objective of preventing a persistent activation. Much of this regulation occurs in the regulation of the trafficking of STING from the ER to the Golgi

and the several post-translational modifications that are catalysed in the latter compartment that allow the activation of the pathway (Konno et al., 2013).

The transport of STING to the Golgi in COPII vesicles, involves multiple proteins (Jia et al., 2020; S. Li et al., 2022; Luo et al., 2016a; Ran et al., 2019; Sun et al., 2018). On the contrary, other proteins have been identified to inhibit this translocation, or promote the retrograde transport of STING from the Golgi to the ER to prevent its activation (Dewi Pamungkas Putri et al., 2019; Mukai et al., 2021). Post-translational modifications like Lys-224 polyubiquitination also seem important for STING export from the ER (Ni et al., 2017). Overall, STING can undergo multiple modifications of different effects on different sites, that involve the intervention of different enzymes and influence its dynamics. These modifications include phosphorylation, ubiquitination, sumoylation, palmitoylation, glycosylation, carbonylation and nitro-alkylation (Kang et al., 2022). Another way of regulation of STING activity and the innate immune response is through its degradation into lysosomes uncoupled of IFN activation. For this elimination, STING requires at least to be activated on the ERGIC (Chu et al., 2021; Gui et al., 2019; D. Liu et al., 2019; Prabakaran et al., 2018).

STING gain of function mutations have interestingly been described, which cause an autoinflammatory disease named STING-associated vasculopathy with onset in infancy (SAVI) (Bennion et al., 2020, 2019; Cerboni et al., 2017; Y. Liu et al., 2014). These mutations cause the constitutive activation of STING without 2'3'cGAMP stimulation and are permanently located in the TGN (Mukai et al., 2016). The structure of these variants is suggested to mimic that of 2'3'cGAMP-bound STING, which may allow its permanent traffic into COP-II vesicles. However, inhibition of palmitoylation or ER:Golgi trafficking prevents the activation of STING's downstream signalling cascade and the IFN response elicited by these mutations is inhibited. This highlights the paramount importance of STING clustering at the TGN for its activation (Mukai et al., 2016). STING also has loss of function alleles that are common in certain populations, STING HAQ, the most common of them, is an haplotype that contains 3 non-synonymous single nucleotide substitutions: R71H, G230A and R293Q in human (Jin et al., 2011)

1.2.2.5. Other cytosolic DNA sensors

Prior to the discovery of cGAS, other molecules were proposed or known to be DNA sensors: DAI, DDX42, DNA-PK and IFI16 (Paludan and Bowie, 2013). However, lack of any of these receptors still allows the cell to induce an IFN response upon detection of cytosolic DNA. In contrast, deletion of cGAS abolishes the induction of IFN (Almine et al., 2017; Ferguson et al.,

2012; Gray et al., 2016; Ishii et al., 2008). Therefore, the role of these proteins in the sensing of DNA is still unclear and poorly understood. Particularly interesting is the role of some these receptors like IFIT16, hnRNPA2B1 and DNA-PK as being primarily nuclear and capable of sensing ssDNA on this compartment in cooperation with cGAS and STING (Almine et al., 2017; Burleigh et al., 2020; Corrales et al., 2016; Dunphy et al., 2018; Ferguson et al., 2012; Karpova et al., 2002; Morchikh et al., 2017; Wu et al., 2019).

Notably, the inflammasome pyropoptotic pathway can also be activated by proteins other than NLRs, specifically the DNA sensor AIM2 can trigger the formation of the inflammasome. AIM2 is activated by binding to the B-form of dsDNA. However due to its potential, the threshold of activation, understood as the minimum length DNA needs to have to activate AIM2 is high, 80bp (Hornung et al., 2009; Sagulenko et al., 2013, p. 2).

1.2.3. IRF3 and NFκB activation

Activation of IKK/Nemo on the STING signalosome causes the phosphorylation of iKB, which releases the canonical NFκB complex. Once released from IκB, canonical NFκB formed by the p50 and p65 (RelA) subunits are released into the nucleus where they bind NFκB binding sites and activate the transcription of pro-inflammatory cytokines as well as their own inhibitor IKB (Nelson et al., 2004). Besides, activation of STING also activates the non-canonical NFκB pathway to limit the type I IFN response (Decout et al., 2021).

IRF3 is the downstream effector of the cGAS/STING pathway. Upon phosphorylation and activation, IRF3 binds to promoter elements within target genes termed as IFN-stimulated response elements (ISREs). This will induce type I IFNs. Type I interferons are a family of cytokines that mediate the innate immune response against virus that bind and signal through the heterodimeric IFNα receptor (IFNAR). They comprise in human 13 genes of *Ifn*a, one *Ifnb* gene and several poorly defined gene products (McNab et al., 2015). Secretion of IFNα/β will in turn stimulate the expression of several Interferon stimulated genes (ISGs) and IRF7, which can later heterodimerize with IRF3 and enhance its transcriptional activity (McNab et al., 2015).

Besides its role in stimulating type I IFN expression, IRF3 can directly contribute to the expression of other cytokines and chemokines, including, *Ccl5*, *Ifit1* and *Cxcl10* (Guinn and Petro, 2019; McWhirter et al., 2004; Petro, 2020). Moreover, IRF3 can induce apoptosis both independently and dependently of its role as a transcription factor, something not seen in its

homolog IRF7 (Eggenberger et al., 2019). Apoptosis can be induced by intrinsic and extrinsic signalling. The intrinsic pathway is initiated by different stimuli that influence the expression/activation of Bcl-2 family proteins that control the permeabilization of the outer mitochondria membrane that would release cytochrome c. The release of cytochrome c into the cytosol will lead to the formation of the apoptosome which will start a caspase-cleaving cascade that will lead to effectors activated caspases (3, 6 and 7) directly causing apoptosis(COHEN, 1997; Li et al., 1997; Peteranderl and Herold, 2017). The extrinsic apoptotic pathway relies on extracellular signal exerted by ligands like TNFα, TRAIL and FasL. They will interact with the respective cell surface-expressed death receptors, which leads to the activation of initiator caspases (Chawla-Sarkar and Lindner, n.d.; Dedoni et al., 2010; Locksley et al., 2001; McIlwain et al., 2013).

IRF3 can induce the pro-apoptotic factors *Noxa* and *Puma*, decreasing the activity of Bcl-2 proteins (Chattopadhyay et al., 2011; Goubau et al., 2009; Sharif-Askari et al., 2007; Solis et al., 2006). Alternatively, activated IRF3 can associate with Bax at the mitochondrial membrane, promoting the release of cytochrome c and triggering apoptosis (Chattopadhyay and Sen, 2017; Sharif-Askari et al., 2007; Chattopadhyay et al., 2010). Other studies have identified an IRF3-dependent and transcription- and Bax- independent necrosis pathway (Di Paolo et al., 2013). Cell-intrinsic responses to IRF3 such as ISG upregulation importantly synergise with paracrine and autocrine IFN signalling (below) in order to bring about a potent innate immune response (Figure 1.5).

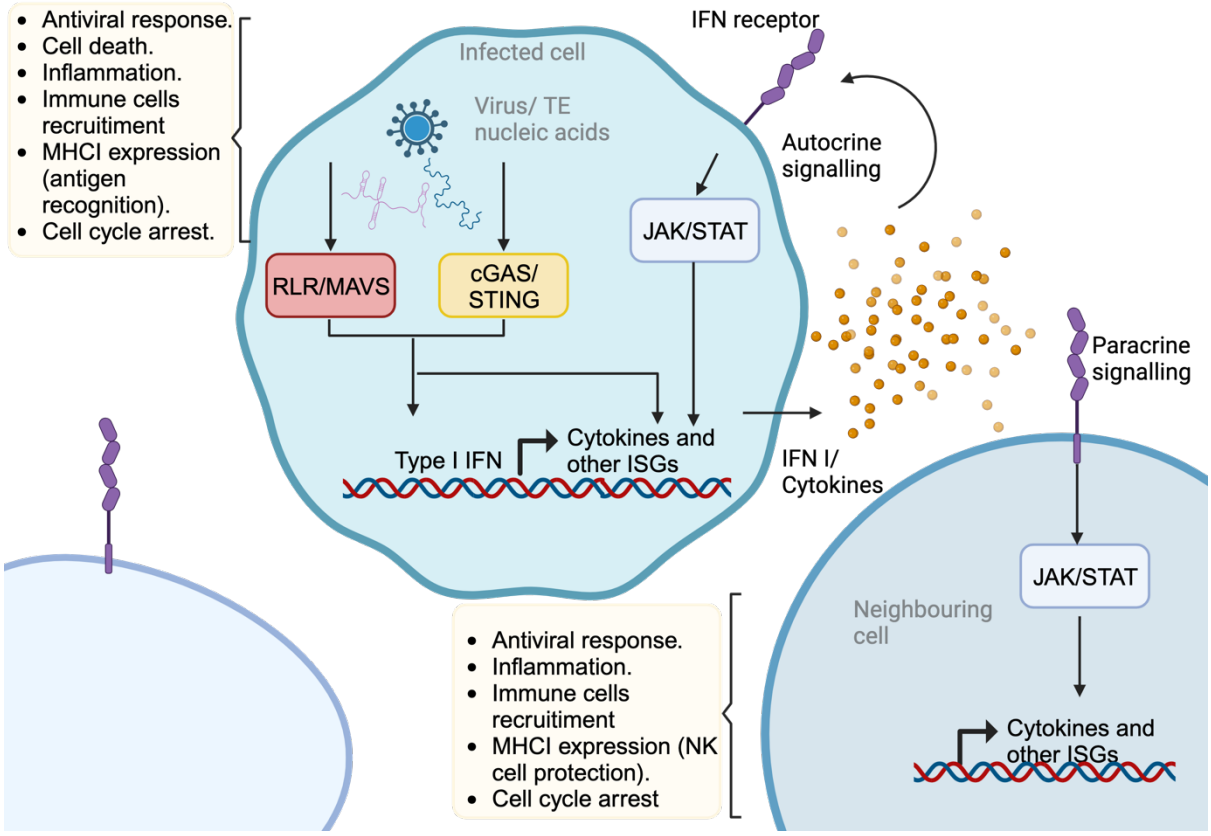


Figure 1-5: Consequences of the activation of the nucleic acid sensing pathway. The viral infection of a cell activates the nucleic acid sensing pathways, that in turn activate the transcription of type I IFNs and some cytokines and ISGs via NFκB and IRF3. IRF3 can also mediate the cell apoptosis dependent and independently of transcription. IFN will be secreted and activate autocrine the JAK/STAT pathway, which will activate the transcription of more ISGs. With this transcriptional programme the cell will establish an antiviral state to try to fight the infection, it will also arrest growth to prevent viral propagation, cause local inflammation and recruit immune cells through cytokines, express MHCI to facilitate the display of antigens and, if sustained, cell death in a cell intrinsic or extrinsic manner. Type I IFNs signalling can also be paracrine through the JAK/STAT1 pathway, to the neighbouring cells to induce an antiviral state and prevent their infection and secrete cytokines to promote inflammation and recruit the immune system. MHC I expression in this case will prevent the cell killing by NK cells, as it acts as an inhibitory receptor. Made with BioRender.

1.2.4. Interferon signalling

The IFNs generated by IRF3/ NFκB activation themselves activate downstream signalling pathways. Type I interferons can in theory be produced by any cell in the body, although dendritic cells seem to be capable of producing them at larger scale (Liu, 2005). IFN signalling works in an autocrine and paracrine manner. IFNα and IFNβ bind to a common heterodimeric receptor consisting of the IFNAR1 and IFNAR2 subunits (Peteranderl and Herold, 2017; Platanias, 2005). This receptor is present on the cell surface, and starts a signalling cascade through the Jak/STAT pathway. IFNAR uses the kinases Tyk2 and Jak1 to phosphorylate

STAT1 and STAT2 on conserved tyrosine residues (Peteranderl and Herold, 2017; Platanias, 2005). Consequently, they undergo a conformational change that enables heterodimerization and exposes a nuclear localisation signal. In the nucleus, they can interact with IRF9 and form a complex called ISGF3, which binds to specific nucleotide sequences known as IFN - stimulated response elements (ISREs) in the promoters of ISGs, inducing the transcription of these genes (Peteranderl and Herold, 2017; Platanias, 2005). There are thousands of ISGs, and their total pool is known as the interferome (Samarajiwa et al., 2009). An alternative signalling to the ISRE is through the formation of different homodimers and heterodimers of STAT genes, which bind to IFN-activated sites (GAS) elements in gene promoters. Actually, type I interferons can signal through ISRE and GAS, while type II IFN can only upregulate through GAS elements. Expression of ISGs will induce an antiviral state (Platanias, 2005). However, the expression of this programme can also have detrimental effects for the cell, especially if activated for a long time or exposed to a high IFN titter. To mitigate this, IFN signalling relies on a negative feedback loop consisting of specific phosphatases and expression of suppressors of cytokine signalling: SOCS1 and 3 that supress JAK/STAT signalling, as well as endocytosis of IFN receptors (Levin et al., 2011; Peteranderl and Herold, 2017; Platanias, 2005).

1.2.5. The antiviral response

The ultimate aim of the innate immune system is to prevent virus replication, entry and activity, as well as signalling the host immune system and neighbouring cells on the presence of a hazardous agent. This activity extends also to prevent TE activity and replication as they pose a similar thread to the host homeostasis. The induction of ISGs induces a state of resistance to viral replication in all cells. Many of the genes involved in nucleic acid sensing and signalling are ISGs, such that cells exposed to IFN have an increased ability to recognise nucleic acids (Honda et al., 2006; Kawai and Akira, 2006) (Figure 1.5).

ISGs exert their antiviral activities through multiple mechanisms. OAS proteins polymerize ATP in to 2',5' linked oligomers (in nucleic acids, nucleotides are normal 2',3' linked) that act as second messengers and activate endoribonucleases that degrade viral RNA (Hornung et al., 2014). Another activated ISG is PKR, which upon detection of dsRNA, supresses protein translation by phosphorylating eIF2 (Meurs et al., 1990; Williams, 2001). IFIT proteins will do the same by sequestering eIF3, impacting cellular proliferation (Diamond and Farzan, 2013).

Interferons also increase the expression of MHC class I molecules on all types of cells, which facilitates the recognition of viral infected cells by cytotoxic CD8 T lymphocytes, as MHCI

would display viral peptides on the cell surface in case the cell was infected (Gessani et al., 2014; Luster, 2002). Finally, they can also increase the production of cytokines and chemokines that recruit and activate lymphocytes, phagocytes, dendritic cells and NK cells to the site of the infection (Kawai and Akira, 2006; Luster, 2002; Sokol and Luster, 2015).

NK cells are recruited by chemokines released during the antiviral response and can recognise infected, stressed and transformed cells thanks to the expression of various combinations of activating and inhibitory receptors and (Lanier, 2008; Salazar-Mather et al., 2000). It is the overall balance of signalling by these receptors what determines if the NK cell kills the target. The receptors recognise ligands on the target cells membrane that are induced by infection, stress and other cellular perturbations (Brown et al., 2001; Vivier et al., 2011, 2008). On the other hand, inhibitory ligands, highly expressed in healthy cells, are downregulated in these kinds of stress/ anti-viral responses (Long, 2008; Parham, 2005). Once activated, NK cells will release proinflammatory cytokines like IFNγ and kill in two ways: release of granule contents (perforins and gramzines) that induce apoptosis in the target cell. Activation of the extrinsic apoptosis pathway signalling occurs via TRAIL and Fas ligands (Vivier et al., 2011, 2004).

1.2.5.1. IFN response on cell proliferation and apoptosis

In addition to antiviral and pro-inflammatory ISGs, type I interferons can also activate other pathways, such as the p38 MAP kinase, which mediates gene transcription, leading to antiviral and antiproliferative effects (C. Li et al., 2004; G. Li et al., 2004). These antiproliferative effects are mediated by modulation of retinoblastoma protein and cyclins and CDK expression.

Jun/STAT1 activation by type I IFNs has also been shown have proapoptotic effects through multiple mechanisms. In the context of a viral infection, these factors provide a mode to block viral spreading and reinfection by killing infected cells.

Some of the mechanisms are transcription-independent interactions with p53, TRADD, Protein kinase Cd and p65 (DeVries et al., 2004; Platanias, 2005). Moreover, STAT1 also promotes the upregulation of pro-apoptotic proteins like Bax, Bak, IFITM2, IP6K2, ISG12, ISG 54, NOXA, NUBI, IFI16, XAF-1, caspases 4 and 8 and AIM2 that promote cell death through the canonical apoptosis pathway (Kim and Lee, 2007; Kotredes and Gamero, 2013; Leaman et al., 2003; Panaretakis et al., 2008; Rosebeck and Leaman, 2008; Stang et al., 2006; Stawowczyk et al., 2011; Tsuno et al., 2012) as well as proteins from the extrinsic apoptotic pathway (Leaman et al., 2002; C. Li et al., 2004; Sato et al., 2001). STAT1 has also been

implicated in cell death by necrosis and autophagy (Kim and Lee, 2007). It has to be noted though, that the induction of apoptosis by IFN seems to be highly cell type-dependent. IFNs have also been shown to have an antiangiogenic effect (Dunn et al., 2005).

1.2.6. Impact of LINE1 activation cGAS/STING in somatic cells

Up until now we have seen how LINE1 activity can drive detrimental effects for the host through many ways. However, LINE1 pathogenic activity goes beyond insertional mutagenesis (Chen et al., 2005; Gilbert et al., 2002). LINE1 RNA can generate secondary structures that activate the innate immune system via MDA5 dsRNA sensor (Chen and Hur, 2022; Dmitriev et al., 2007; Tunbak et al., 2020). Moreover, ORF2p activity as an endonuclease and retrotranscriptase can mediate genotoxic effects and generate cDNA intermediates that activate the cGAS/STING pathway (Ardeljan et al., 2020; Cecco et al., 2019; Crow and Manel, 2015; Scopa et al., 2023; Simon et al., 2019). As it has been demonstrated, the persistent activation of this pathway can activate cell death and mediate multiple effects that are detrimental for the cell activity (Figure 1.5) (Petrasek et al., 2013; Zierhut et al., 2019). In the next section I will describe how these activities impact on the pathophysiology of some diseases.

1.2.6.1. Autoimmune diseases

TE-derived nucleic acids have been suggested in recent years to be molecules responsible for driving aseptic inflammation in the context of autoimmune diseases and other degenerative diseases. In this section I will showcase the diversity of autoimmune pathologies that can be caused by TEs, and more specifically LINE1.

Aicardi Goutières (AGS)

AGS is an inflammatory encephalopathy and genetic disease involving the elevation of type I IFN levels in patient serum and cerebrospinal fluids (Goutières et al., 1998). It is an early-onset disease characterized by neurological dysfunction, psychomotor retardation and skin inflammation (Goutières et al., 1998). The disease is caused by mutations of different genes involved in the metabolism of nucleic acids, all of them known to be LINE1 repressors: *Trex1, Samhd1, Adar1, RnaseH2* and *Apobec* genes (Crow et al., 2006a, 2006b; Rice et al., 2012; Zhao et al., 2013). These mutations lead to the abnormal accumulation and sensing of these nucleic acid intermediates. The origins of these nucleic acid intermediates are still unclear but

has been attributed to DNA damage or replication intermediates and retrotransposons (Crow and Manel, 2015; Volkman and Stetson, 2014).

In favour of the TE origin is the fact that *Trex1* mutant mice show accumulation of ssDNA of retrotransposon origin without the activation of the DNA damage response (Stetson et al., 2008). Moreover, a *Trex1* deficient human *in vitro* model showed that retroelements as the source of increased ssDNA accumulation in *Trex1* deficient neural cells and astrocytes which were rescued with nucleoside reverse transcriptase inhibitors (NRTIs). Interestingly, retrotranscriptase related toxicity occurred only in neurons but not in astrocytes and neural progenitor cells. In addition, astrocytes were still capable of producing IFN and ISGs, which then had a toxic effect on the neighbouring neurons (Thomas et al., 2017). However, the treatment of these mice with anti-retroviral drugs has not reduced the IFN activation, suggesting that a retrotransposon-independent mechanism may also contribute to the onset of the disease (Achleitner et al., 2017; Volkman and Stetson, 2014).

Systemic lupus erythemstosus (SLE)

SLE is an autoimmune disease characterized by the production of antibodies targeting nucleic acids and nuclear-associated proteins (Crow and Rehwinkel, 2009; Hooks et al., 1979). It has a major overlap with AGS as it can be caused by mutation of the same set of genes (Lee-Kirsch et al., 2007; Ravenscroft et al., 2011). Later on, LINE1 was found to be dysregulated in SLE concomitantly to the induction of the type I IFN response (Carter et al., 2020; Kelly et al., 2018; Mavragani et al., 2016). However, it has not been possible to attribute any causative role of LINE1 upregulation on the autoinflammation and onset of the disease. Furthermore, SLE patients have been found to produce auto antibodies against HERV elements proteins, suggesting aseptic inflammation can be also driven by HERVs (Perl et al., 1995).

Neurological disorders

In the case of Alzheimer disease (AD) increasing evidence has reported the aberrant de-repression and mobilization of LINE1 and LTR elements found in constitutive heterochromatin regions (Guo et al., 2018; Wanhe Li et al., 2013; Protasova et al., 2017; Tam et al., 2019). Recently, a study performed with patient-derived iPSCs dedifferentiated into hippocampal progenitors, CA3 neurons and cerebral organoids; demonstrated that c-JUN upregulation in AD, among other effects, causes the de-repression of TEs (Scopa et al., 2022). This leads to the formation of RNA:DNA hybrids that activate the cGAS/STING pathway, which ultimately leads to caspase-3 accumulation and apoptosis. This adds evidence on the role of

neuroinflammation on the pathogenesis of AD, in this case neuroinflammation would be primarily triggered by HERVs (Scopa et al., 2023).

1.2.6.2. Cancer and LINE1 induction of cGAS/STING

I have already reviewed the potential of TEs and LINE1 in particular to promote carcinogenesis. However, LINE1 has recently been found to contribute to tumorigenesis through its modulation of the innate immune system. TE activation in cancer can generate aberrant nucleic acid species like ssDNA or RNA:DNA hybrids that activate the innate immune response and signal for the host immune system to clear the transformed cell (Ardeljan et al., 2019; Belgnaoui et al., 2006; Zhao et al., 2021).

The capacity of TEs to generate aberrant nucleic acid species that activate RLRs and cGAS/STING pathway has proved instrumental for an organism to detect and eliminate neoplastic cells. The persisting activation of the type I IFN response through cGAS/STING or MDA5 caused by LINE1 and other TEs ssDNA and dsRNA in cancer, inhibits angiogenesis, promotes apoptosis, and enhances cell killing by NK cells and CD8 cytotoxic T cells and up (Belgnaoui et al., 2006; Cañadas et al., 2018; Chiappinelli et al., 2015; Haoudi et al., 2004; Roulois et al., 2015; Samson and Ablasser, 2022). Moreover, the 2'3'cGAMP produced by a cell can be exported intercellularly and act as a bystander (Ablasser et al., 2013; Carozza et al., 2020; Marcus et al., 2018). Consistently, cancer cells have developed strategies to inhibit the cGAS/STING pathway (Chen et al., 2023; Wu et al., 2021). For this reason the cGAS/STING pathway is inhibited in many cancers (Motwani et al., 2019).

However there are several instances where activation cGAS/STING signalling has been reported to accelerate tumour growth, metastasis or drug resistance in cases of persistent induction of pro-inflammatory cytokines, although none of them have been linked to TE expression (Budhwani et al., 2018; Hou et al., 2018; Novototskaya-Vlasova et al., 2022; Samson and Ablasser, 2022; Wu et al., 2020).

1.2.6.3. LINE1 and Senescence

The capacity of TEs to generate DNA cytosolic intermediates constitutes a major source of their detrimental potential. Aside from the clearly pathogenic context that I have discussed, TEs can also be active and detrimental in not strictly pathogenic contexts, such as senescence (Belancio et al., 2010). Senescence is perhaps the best example of how TE activity can interact with the innate immune system to influence cell fate. Cellular senescence is a stress

response characterized by a permanent cell cycle arrest (Hayflick, 1965; Hayflick and Moorhead, 1961). However, the senescent phenotype is not limited to growth arrest. Senescent cells undergo morphology changes, macromolecular damage, chromatin remodelling, metabolic reprogramming and a characteristic secretome known as the senescence associated secretory phenotype (SASP) (Gorgoulis et al., 2019; He and Sharpless, 2017; Herranz and Gil, 2018).

The senescent programme can be triggered by multiple different extrinsic and intrinsic stressful insults including telomere erosion, mitochondrial disfunction, oxidative damage, radiation exposure, chemotherapeutic drugs or hyperproliferation caused by oncogene activation (Herranz and Gil, 2018). The source of DNA damage can be diverse, in some cases it has been demonstrated that the endonuclease activity of TEs like LINE1 is capable of triggering senescence when derepressed (Cecco et al., 2019; Simon et al., 2019). These insults normally cause a persistent DNA damage response which leads to the activation of p53/p21 axis. Persistent stress activates p16, an inhibitor of CDK4/ 6 that contributes to the durable growth arrest (Krishnamurthy et al., 2004; Yamakoshi et al., 2009, p. 53).

Senescent cells can also influence the environment around them through SASP, involving the production of a combination of cytokines, chemokines, extracellular matrix proteases and other signalling molecules (Kuilman et al., 2008). The specific composition of this secretion will be variable depending on the tissue (Wiley et al., 2017). The SASP will reinforce within the same cell the senescent phenotype and via paracrine mechanisms will influence neighbouring cells to undergo senescence, aiming to stop the proliferation of cells that may be exposed to the same stressor and recruit the immune system to eliminate senescent cells (Acosta et al., 2013, 2008; Hubackova et al., 2012; Kuilman et al., 2008; Lujambio et al., 2013; Nelson et al., 2012; Wajapeyee et al., 2008). Finally, the aseptic chronic inflammation triggered by the SASP has also been strongly linked to the detrimental effects of aging and its related diseases, and its suppression linked to an improvement of the health-span (Baar et al., 2017; Baker et al., 2016, n.d.; Chang et al., 2016; Childs et al., 2016).

Senescence is curiously a very dynamic process. If the senescent cells persist for extended periods of time, they continue evolving and can be categorized as entering into late senescence, which can involve adaptation and diversification of the phenotype (Herranz and Gil, 2018). It is in this late senescence stage that studies report that LINE1 elements become derepressed and generate cytoplasmic cDNA intermediates that activate the cGAS/STING pathway and activate an IFN response that contributes to the maintenance of the SASP (Belancio et al., 2010; Simon et al., 2019; Cecco et al., 2019). This is supported by evidence

showing that the treatment with NRTIs that inhibit the LINE1 ORF2p reverse transcriptase (RT) activity led to a decay in LINE1 derived cDNA presence in the cytoplasm and ultimately the downregulation of interferons and other inflammatory SASP mediators. Additional evidence has been added by the specific inhibition of the cGAS/STING pathway and knock-out of interferon receptors, which has led to the same result (Simon et al., 2019; Cecco et al., 2019).

This phenomenon has been seen in organisms other than human, such as in aged primate and mouse tissue samples. The expression of LINE1 also correlated with the activation of STAT1 by type I IFN in the tissue microenvironment with other ISGs. It is relevant to point out that only a fraction of the tissue cells identified as senescent were also expressing LINE1, this percentage was seemingly increasing with age (Cecco et al., 2019). Treatment of aged mice with NRTIs contributed to a reduction in SASP and type I IFN responses, while maintaining LINE1 expression, producing a senostatic effect. In the long term, the treatment also contributed to ameliorating aging related phenotypes and increasing animals health-span (Cecco et al., 2019; Simon et al., 2019).

Cecco et al. explain the activation of LINE1 in senescence by the downregulation of the exonuclease *Trex1* and *Rb1*, a heterochromatin protein targeting repetitive elements (Ishak et al., 2016). The loss of *Rb1* is correlated with the loss of H3K9 and H3K27 trimethyl in repetitive regions. In parallel, the senescence upregulated factor FOXA1 binds LINE1 5'UTR and promotes its expression (Qian Li et al., 2013). Simon et al. also relies on the loss of *Sirt6* an heterochromatin proteins that typically represses LINE1 elements (Mostoslavsky et al., 2006), overexpression (OE) of *Sirt6* was seen to revert the inflammation phenotype (Simon et al., 2019; L. Zhang et al., 2022). Overall, LINE1 seems to be upregulated in senescence due to a broad chromatin remodulation that causes a genome-wide de-repression of repetitive elements.

1.3. LINE1 and Development

Up until now I have reviewed the basic biological aspects of TEs, more specifically LINE1 and their potential in in the pathogenesis of inflammatory diseases through the activation of the cell innate immune system when expressed in somatic contexts (Cecco et al., 2019). But central to this work is the fact that LINE1 is also expressed in early development, a stage where curiously LINE1 is activated without triggering any of the immune surveillance pathways (Fadloun et al., 2013; Jachowicz et al., 2017; Percharde et al., 2018). Therefore, in the coming section I am going to review the current knowledge of LINE1 in early development.

1.3.1. Embryonic stem cells as a model to recapitulate early development

One of the most fascinating events of nature, is the capacity of a zygote to give rise to an entire organism. This cell is therefore capable of generating any cell in the body, a property named totipotency (Casser et al., 2017; Tarkowski, 1959). In mice this property is only possessed by the zygote and the 2-cell embryo. This is also the stage when zygotic genome activation occurs in mice. After fertilization, the zygote undergoes a series of rapid cell divisions known as cleavage, the product of which is known as the morula. Within four days of fertilization, the morula undergoes blastulation and develops into the blastocyst that already contains 2 differentiated cell populations (Figure 1.6). The outer layer is made up of the trophectoderm, a population of cells that will later form the trophoblast involved in implantation, and further on, the embryonic component of the placenta in charge of supporting foetal development (Fleming, 1987; Yockey and Iwasaki, 2018). The other population, inserted within the trophectoderm is the inner cell mass (ICM). This population's differentiation potential is restricted to generating the embryo proper (Nagy et al., 1990). Therefore, the epiblast cells have the capacity to generate the three germ layers of the embryo, this property is known as pluripotency (Figure 1.6).

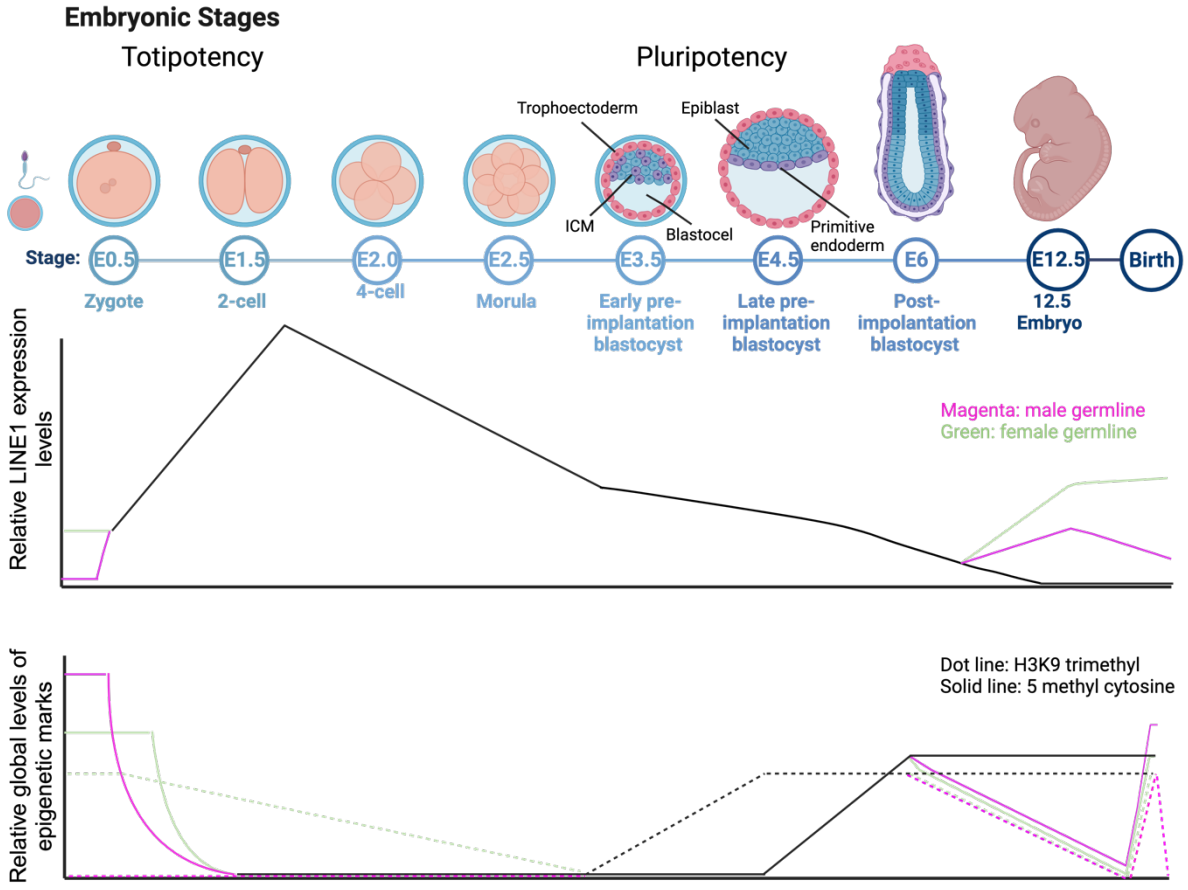


Figure 1-6: Temporal dynamics of LINE1 expression through early mouse development and correlation with global levels of DNA methylation and H3K9me3. The top graph is a schematic

of early development. The middle graph shows the levels of LINE1 expression through the stages on the top graph. The bottom graph shows the levels of DNA methylation (5 methyl cytosine) in solid lines and the levels of histone H3K9 trimethyl in dot lines. In the bottom two graphs male and female germ cells genome is discriminated in two colours. Made with Biorender. Adapted from (Jachowicz and Torres-Padilla, 2016)

The cells forming the distinct populations of the embryo can be isolated and stably cultured *in vitro* as stable cell lines. These are probably one of the most important tools for the study of mammalian development. In 1981, two groups were capable of deriving pluripotent cells from pre-implantation embryos that could recapitulate the development window of the ICM in the pre-implantation blastocyst (Evans and Kaufman, 1981; Martin, 1981). These cells were named mouse embryonic stem cells, mESCs and possess two defining characteristics. The first is indefinite self-renewal in culture as long as they keep their identity. The second is pluripotency, characterised experimentally by the ability of these cells to form teratocarcinomas containing cells of all three germ layers and to form mouse chimeras when injected into blastocyst (Robertson et al., 1986). These cells are also characterized by the expression of the core pluripotency transcription factors *Oct4*, *Sox2* and *Nanog*, key to maintain a transcriptional network that maintain their self-renewal and differentiated state (X. Liu et al., 2008).

Their post-implantation blastocyst counterparts, Epiblast stem cells (EpiSCs) share the self-renewal phenotype and the expression of the core pluripotency factors. However, owing to their more advance developmental state they show a decreased cellular plasticity, as they loss the capacity of forming efficient chimeras and expose features like X-chromosome inactivation, flattened morphology and dependency of FGF and Activin signalling for their propagation (Brons et al., 2007; Guo et al., 2009; Tesar et al., 2007). The other cell types forming the extra-embryonic tissues can also be cultured, having a differentiation potential restricted to their respective lineages (Ralston and Rossant, 2008; Rossant, 2008).

Pluripotency is by its own nature a transient state. To successfully keep mESCs in culture, ESC media must be supplemented with leukemic inhibitory factor (LIF) and serum, the active component of which is bone morphogenic protein 4 (BMP4) (Niwa et al., 1998; Qi et al., 2004). LIF activates the Jak-Stat3 pathway and the Pi(3)K pathway which maintain the expression of *Oct4, Sox2* and *Nanog* (Matsuda et al., 1999; Niwa et al., 2009). BMP4 signals via the TGF-B pathway and activates the expression of inhibitor of differentiation proteins, that block the expression of neuroectodermal differentiation.

A more *naïve* form of pluripotency can be achieved by substituting the serum with 2i, a combination of two inhibitors of pathways that negatively regulate the expression of pluripotency genes. These inhibitors are PD0325901, a MEK/ERK signalling, and CHIR99021, a GSK3B inhibitor (Batlle-Morera et al., 2008; Ying et al., 2008).

The ESC genome has a characteristic euchromatic environment that becomes increasingly heterochromatic with differentiation, and an hypomethylated genome (Figure 1.6). This pattern is generated after fertilization, when the genome undergoes a wide rapid reprogramming event that diminishes heterochromatin and DNA methylation levels. These levels are not restored until post-implantation. A similar type of event occurs a second time in development during the formation of the germline (Burton and Torres-Padilla, 2014; Greenberg and Bourc'his, 2019; Hemberger et al., 2009; Mayer et al., 2000; Smith et al., 2012). Moreover, totipotency and ZGA are characterised by high chromatin plasticity and histone mobility, a characteristic partially retained in pluripotency and gradually lost with differentiation (Bošković et al., 2014).

Overall, the ICM genome is rich in H3 and H4 acetylation, poor in H3K9 di- and tri- methylation and hypomethylated (Figure 1.6) (Arney and Fisher, 2004) . This is consistent with early passage ESCs, which exhibit global DNA hypomethylation, with a progressive increase in methylation with continued passaging. The hypomethylation has functional consequences at the level of particular loci, as the promoters of pluripotency factors are generally hypomethylated (Dean et al., 1998; Fouse et al., 2008; Nagy et al., 1993). Interestingly, culture conditions of mESCs can affect levels of epigenetic marks, as mESCs cultured in 2i+LIF conditions, show lower DNA methylation levels (Sim et al., 2017).

1.3.2. LINE1 is necessary for early development progression

The loss of heterochromatic features in ESCs has been normally associated with a loss of transcriptional control over TEs, which causes the upregulated expression of these elements in ESCs, primordial germ cells (PGCs) and their *in vivo* counterparts in a conserved manner (Figure 1.6) (Calarco and Brown, 1969; Fadloun et al., 2013; Friedli et al., 2014; Grow et al., 2015; Ohno et al., 2013; Percharde et al., 2018, 2017). Among such activated transposons are LINE1 elements.

LINE1 is also later expressed in the neuronal progenitor cells. In this cells, LINE1 has been proposed to be a driver of neurodiversity, as LINE1 retrotransposition bias towards open chromatin, would reshape the neural transcriptional network (Della Valle et al., 2020; Garza et al., 2023; Muotri, 2016).

In development, LINE1 expression is one of the earliest transcriptional events occurring with ZGA at the 2-cell stage, it is also during this stage that LINE1 expression peaks, although it is also expressed in pluripotency (Figure 1.6) (Jachowicz et al., 2017; Percharde et al., 2018). In addition, LINE1 RNA is stabilized by N^6-methyladenosine demethylation, catalysed by the enzyme FTO (Wei et al., 2022). Localization of LINE1 RNA during early development is mainly nuclear, while in contrast, ORF1p is mainly cytoplasmatic (Jachowicz et al., 2017; Percharde et al., 2018). Interestingly, it was found that in ESCs, LINE1 RNA forms a ribonucleocomplex with Kap1 and Nucleolin proteins and together has a lncRNA-like function. This complex interacts with the genomic locus for *Dux,* a master activator of the 2-cell transcriptional programme (De Iaco et al., 2017; Hendrickson et al., 2017; Macfarlan et al., 2012). This LINE1 complex is essential to silence *Dux*, in cooperation with a mechanism that moves the *Dux* loci to perinucleolar heterochromatin (Percharde et al., 2018; Xie et al., 2022). *Dux* silencing causes deactivation of the 2-cell transcriptional programme along the TE MERVL, which is normally expressed at the 2-cell stage under *Dux* transcriptional control. Independently of this process, LINE1 also interacts with the ribosomal DNA and activates rRNA transcription, which leads to a global increase in ribosome biogenesis and supports the rapid proliferation characteristic of the early embryo (Percharde et al., 2018). This effect of LINE1 in chromatin is consistent with data showing LINE1 RNA regulating global chromatin accessibility (Jachowicz et al., 2017). In fact, LINE1 expression reduction after the 2-cell stage and further development is accompanied by a progressive compaction of chromatin (Bošković et al., 2014; Jachowicz et al., 2017).

Together, these reports reveal that LINE1 RNA is necessary for maintaining an open chromatin state and the exit of the 2-cell stage into pluripotency (Jachowicz et al., 2017; Percharde et al., 2018). But contrary to what have been suggested until now, LINE1 does not seem to compromise the viability of mESCs. Moreover, the retrotransposition seems heavily diminished considering a similar expression in other somatic contexts (Arokium et al., 2014; Ewing and Kazazian, 2010; Garcia-Perez et al., 2007; Kano et al., 2009). On a new perspective contrary to their traditionally associated detrimental role, expression of TEs has been recently found to play an important role during multiple developmental contexts, challenging the old established simplistic view of TEs as pathogenic agents. The apparent lack of detrimental effects by LINE1 expression in early development is in stark contrast with its effects in somatic cells. This is therefore consistent with a conditional symbiotic situation where evolution must have developed mechanisms in the organism that allow tolerating LINE1 expression and contact with chromatin, while preventing its detrimental effects.

1.4. Current knowledge of innate immunity in development

One of the hypotheses that I set out to test in my PhD was that there are pluripotency specific changes to the IFN response in development that may serve to prevent detection and response to TEs, given the high expression of elements such as LINE1 (see above). Thus, in this section, I have summarised the existing knowledge about innate immunity in development.

It has been previously established that both human and mouse ESCs and embryos do not exhibit an IFN response to bacteria and viruses, which is instead a function acquired during differentiation (Burke et al., 1978; Chen et al., 2010; D'Angelo et al., 2018; Hertzog et al., 1994; Witteveldt et al., 2019; Wu et al., 2018; Zhao et al., 2013). It is particularly interesting that *in vitro* cultures of trophoblast stem cells can sense viral nucleic acids and seem capable of reacting to IFN (Aikawa et al., 2014; Fendereski et al., 2022), something consistent with the fact that during implantation the trophectoderm secretes several inflammatory cytokines, with implantation itself being an immune regulated process (Yockey and Iwasaki, 2018). Focusing in pluripotency, ESCs also exhibit a mild response to external IFN (R. Wang et al., 2014a). This phenotype has been suggested to be linked to the expression of pluripotency factors like *Klf4* and *Nanog*, which have been shown to attenuate the innate immune response in pluripotency by downregulating the NFκB pathway (Luo et al., 2016b; Torres and Watt, 2008). Moreover, the pluripotent phenotype seems to antagonize innate immunity; as Nanog represses NFκB by binding it and preventing its transcriptional activity. Nfκb is also inhibited by miR-290 cluster, which targets the NFκB subunit *RelA*/*p65* to prevent its translation. On the contrary, over-expression of NFκB results in differentiation (Lüningschrör et al., 2012; Torres and Watt, 2008; Yang et al., 2010). This antagonism is attributed to the already discussed anti-proliferative and pro-apoptotic properties of IFN pathway (Hertzog et al., 1994). Nevertheless, it has been established for decades that ESCs and other primitive cells resist viral infection better than their differentiated counterparts (Belzile et al., 2014; Swartzendruber and Lehman, 1975; Weichold et al., 1998). Therefore, stem cells have evolved alternative mechanisms to protect against virus. A key unanswered question is whether such mechanisms are also working against TEs or if there is TE-specific regulation operating in pluripotency. Furthermore, it is not known if these mechanisms could be a consequence of the evolutionary pressure exerted by the endogenous expression of TEs in early development.

1.4.1. Alternative innate immunity mechanisms in stem cells

To protect against virus, organisms are equipped with complex cell-intrinsic mechanisms to restrict viral replication and propagation (Figure 1.7). As we have seen, vertebrates rely on the protein-based IFN-driven immune system. In contrast, it is well established that in invertebrates and plants, RNA interference (RNAi) functions as an innate defence mechanism against virus. This system consists of Dicer cleaving virus-derived double-dsRNA, and converting it into siRNAs that join Argonaut proteins and accessory proteins to form the RISC, that cleaves complementary viral RNA through Ago2 (MacRae et al., 2008). In mammals, it was thought that this early mechanism had been supplanted by the non-sequence-specific IFN response. This was because although susceptible to synthetic siRNA and conserving the RNAi machinery, the existence of a natural RNAi mechanism in mammals was debated (Goubau et al., 2013; Haasnoot et al., 2007).

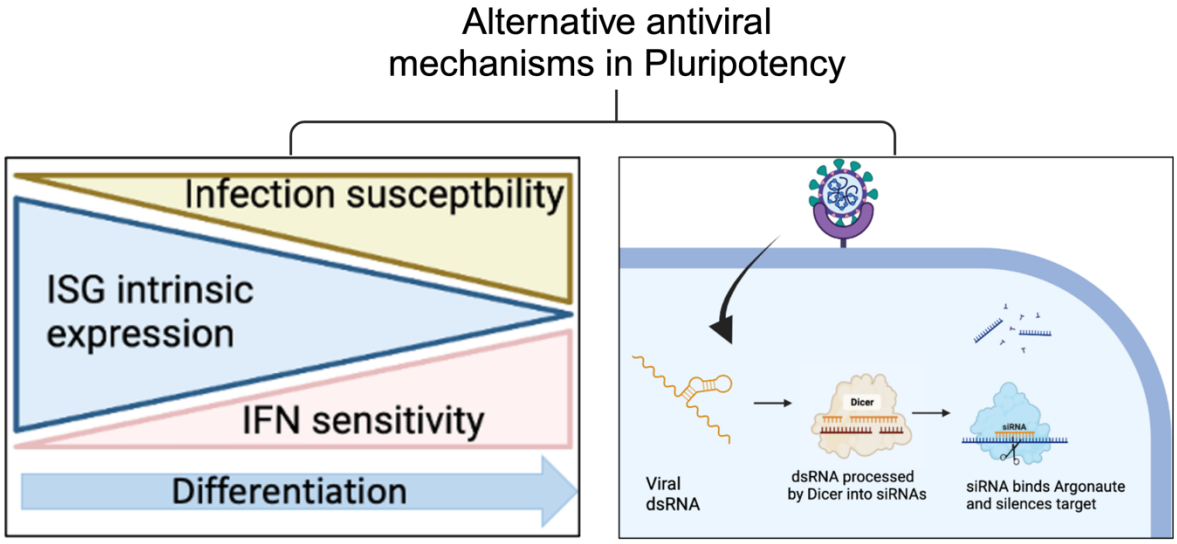


Figure 1-7: Pluripotent cells rely in alternative innate immunity mechanisms.Pluripotent cells have an attenuated innate immunity that prevents them from responding to IFN. However, they are surprisingly less susceptible to viral infection than non-IFN stimulated differentiated cells. This phenotype is achieved thanks to the intrinsic expression of a set of ISGs that confer the cell a basal state of anti-viral defence. The expression of this factors decays with differentiation in parallel to an increase in IFN sensitivity and infection susceptibility. Complimentary to this mechanism, pluripotent cells also rely on the only active RNAi pathway found in mammalian cells thanks the expression of a *Dicer* isoform. Made with Biorender. Based on (Maillard et al., 2013; Wu et al., 2018).

Nevertheless, in the particular case of early development, the apparent incompatibility with a functional IFN defence mechanism, has led to the recovery of alternative defence mechanisms against intracellular pathogens thorough evolution (Murchison et al., 2005). In 2013 Maillard et al. discovered the existence of a functional RNAi pathway in mESCs that depends on *Dicer*.

Cells infected with virus accumulated 22nt RNAs with all the characteristics of siRNAs. These siRNAs were later incorporated into Ago2 (Maillard et al., 2019, 2013). A subsequent study showed that the IFN system actively inhibits dsRNAi through the induction of LGP2 that binds Dicer and inhibits processing of long dsRNA into siRNAs (van der Veen et al., 2018).

Invertebrates and plants possess multiple copies of *Dicer* that are differentially involved in the function of gene regulation or antiviral defence. However, mammals only possess one copy of Dicer, with one canonical form that processes well pre-miRNA, but cleaves dsRNA very poorly (Ma et al., 2008). Further studies attributed this to a N-terminal helicase domain in canonical Dicer that inhibited the processing of long dsRNA (Kennedy et al., 2015; Provost et al., 2002). In ESCs and oocytes, two isoforms of *Dicer* (*aviD* and *Dicer0*) lacking the aforementioned domain, were found to be more efficient at processing dsRNA and to be more resistant to LGP2 binding. The expression of these isoforms is considerably lower compared to canonical Dicer and is preferentially expressed in ESCs and oocytes and decays with differentiation (Flemr et al., 2013; Poirier et al., 2021).

Gain and loss of function studies performed with Dicer in cells lacking IFN system like ESCs, have been found to have a small effect on the production of viral particles of certain viruses like *influenza A,* HIV, measles alphaviruses and flaviviruses (Bogerd et al., 2014; Kennedy et al., 2015; Maillard et al., 2016). The authors hypothesise that viral siRNA production in mammals is too weak to inhibit viral infection, as most viruses replication occurs within membranous structures, and generated RNA quickly loads into capsid proteins (Conzelmann, 1998; Romero-Brey and Bartenschlager, 2014).

Because of the potential of LINE1 RNA to form double-stranded secondary structures, Dicer has been associated with LINE1 repression by cleaving LINE1 RNA (Bodak et al., 2017). In mice, deletion of *Dicer* leads to early embryonic lethality and impaired pluripotency exit in culture, possibly caused by the upregulation of the IFN pathway upon *Dicer* removal (Gurung et al., 2021; Witteveldt et al., 2019). Bodak et al. added more evidence that Dicer has the potential to repress LINE1 with *Dicer* knockout mESCs, documenting a 6-fold increase in LINE1 RNA and 2 times-increase in accumulated retrotransposition events in mESCs 3 weeks after starting the assay, likely caused by the accumulation of LINE1 RNA (Bodak et al., 2017). Additionally, it was recently discovered that Dicer has RNA:DNA R-loop removal activity, which could potentially be influencing LINE1 repression through the elimination of TPRT DNA:RNA hybrids (Camino et al., 2023). However, ESCs have high levels of LINE1 RNA, which is also known to be translated into ORF1p, as high levels of the protein are expressed in the cytoplasm of ESCs (Fadloun et al., 2013; Percharde et al., 2018). Hence a mechanism,

like Dicer, relying on the elimination of LINE1 RNA, is not enough to explain the lack of detrimental effects of LINE1 activation in ESCs and early embryos.

As mentioned earlier, RNAi in vertebrates is not a very efficient mechanism (Bogerd et al., 2014; Kennedy et al., 2015; Maillard et al., 2016). Hence, ESCs are thought to have developed alternative antiviral mechanisms that involve a subset of ISGs which are intrinsically expressed in ESCs in absence of interferon (Figure 1.7). This confers to these cells an intrinsic IFN-independent antiviral state that makes them more resistant to viruses than their differentiated counterparts (Figure 1.6) (Wu et al., 2018). Moreover, this particular subset of ISGs have no attributed detrimental activity, allowing the cell to keep a basal level of anti-viral defence without facing any detrimental effect. This resistance capacity is however, lower than IFN stimulated differentiated cells carrying a full antiviral response capability (Figure 1.7) (Wu et al., 2018).

This feature is found in ESCs, germ layer cells, and tissue stem cells and it seems to be conserved across mammalian species (Wu et al., 2018). Of relevance, different types of stem cells display distinct intrinsic ISG signatures. As cells differentiate, the expression of these ISGs and their antiviral resistance decreases in parallel as cells become increasingly sensitive to IFN (Wu et al., 2018). Without the expression of these critical ISGs, stem cells are susceptible to infection and fail to regenerate tissues upon viral exposure.

What regulates this process and controls the differential expression of different ISGs is unknown. Moreover, what activates these ISGs without IFN is also still unknown. Romeinke et al. identified the ISG *Irf1* as a regulator of an enhancer that controls the expression of *Oct6*, a formative pluripotency factor, and a subset of ISGs differentially expressed between naïve and primed pluripotency. *Irf1* would be the factor in charge, and it would be itself under the control of a network of enhancers activated in the transition from naïve to formative pluripotency, and independent from the canonical activation by JAK/STAT signalling (Romeike et al., 2022). This suggests that the mechanisms responsible for the expression of these ISGs may be multifactorial and complex: different transcription factors may be responsible for the expression of different ISGs in different cell types. For example, OE of the HERVK product in hESCs is enough to increase the expression of *Iifitm1* (Grow et al., 2015).

1.4.2. Nucleic acid sensing in early development

As mentioned before, LINE1 can activate the immune response through both its nucleic acid intermediates: RNA, DNA and RNA: DNA hybrids. Hence, due to the potential negative impact

of a chronic inflammation, a lot of restriction mechanisms operate to limit LINE1 expression and life cycle in normal homeostatic conditions. However, early development raises a situation in which LINE1 expression has to be compatible with normal homeostasis. As seen earlier, the IFN response in what regards to bacteria and viruses seems to be attenuated in mESCs to prevent the negative consequences on development. This is a particularly relevant topic as many TEs are endogenously expressed in early development.

For the case of LINE1 dsRNA, it is a PAMP present in mESCs (Kano et al., 2009; Percharde et al., 2018), raising the question as to whether it is sensed as immunogenic during early development. Published studies examining the response to RNA viruses in ESCs have revealed that indeed there are mechanisms to prevent dsRNA-mediated IFN activation. In fact, during pluripotency, the miRNA miR-673 downregulates MAVS, the central adapter of dsRNA sensing (Figure 1.8). The expression of this miRNA is later silenced during differentiation. Eliminating Dicer or miR-673 reactivated MAVS protein expression and dsRNA sensing pathway. Therefore, *Dicer* deficient mESCs can acquire the capacity to produce IFN and the subsequent antiviral state in response to dsRNA that gives them higher resistance to viral infection (Gurung et al., 2021; Witteveldt et al., 2019).

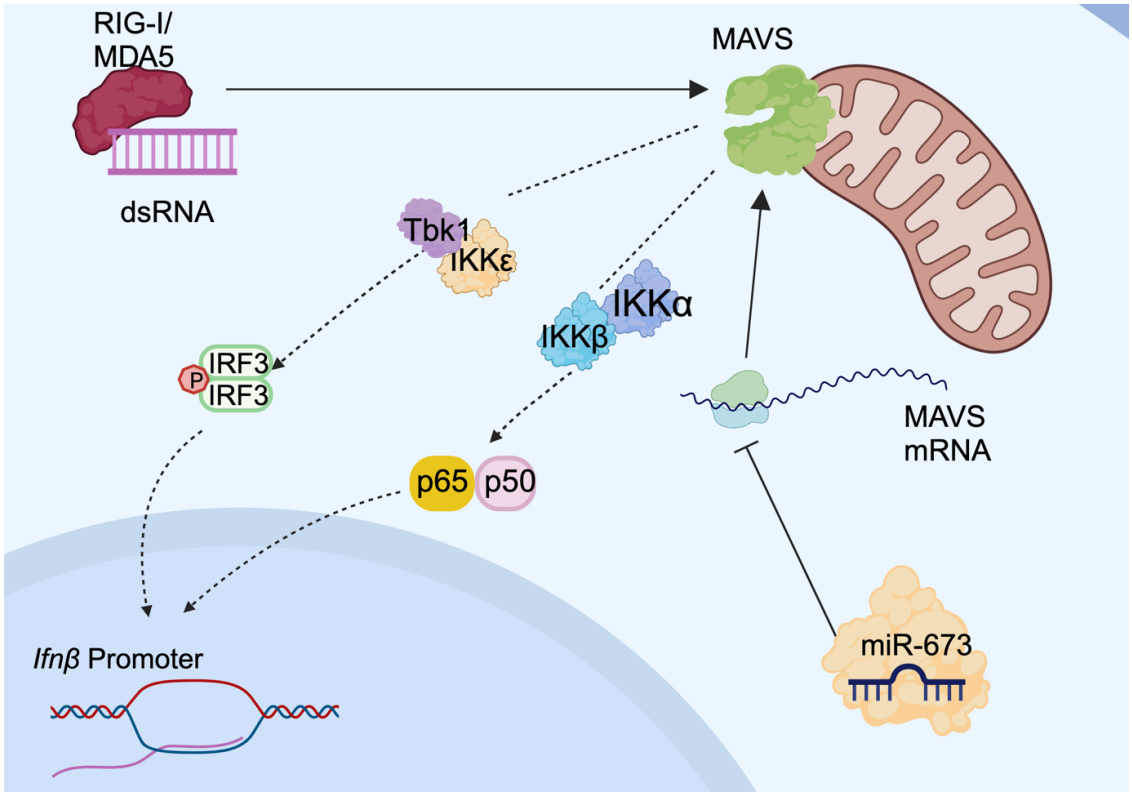


Figure 1-8: A miRNA represses MAVS in pluripotency. MAVS, a protein located in the mitochondrial is the adapter of dsRNA sensing, signalling through IRF3 and NFκB pathway. In pluripotency the miRNA miR-673 targets MAVS RNA and prevents its translation. Therefore, inhibiting dsRNA signalling.

Despite the discovery of the mechanism of RNA-sensing repression in pluripotency, at the beginning of this work, the mechanisms underlying TE DNA sensing in pluripotency were unknown. For this reason, and the role of TE-derived ssDNA in senescence, I set to explore the mechanisms underlying ESC sensing of ssDNA.

1.5. Final considerations: Lessons from LINE1 control in the germline

PGCs go through an equivalent epigenetic reprogramming process to the early embryo that upregulates the expression of TEs. Mammalian PGCs, like in other animals, rely on the piRNA pathway for silencing transposable elements in the developing gonad. piRNAs are a class of small-noncoding RNAs mostly derived from TEs (Di Giacomo et al., 2013; Wang, 2017), which target transposon RNA transcript for degradation by Piwi proteins in the cytoplasm, on parallel of targeting the TE genes in the nucleus for silencing. In mouse, mutation of any of the components of the pathway causes meiotic arrest and subsequent male infertility (Newkirk et al., 2017; Vourekas et al., 2015). Other mechanisms described in male and female germline mention specialized types of stress granules that sequester LINE1. Suppression of these granules leads to apoptotic death (Bao and Yan, 2012; Soper et al., 2008; Su et al., 2012; Tanaka et al., 2000). However, the importance of the piRNA pathway in the female germline seems to vary between species (Kohlrausch et al., 2022; Roovers et al., 2015).

LINE1 control is particularly relevant in the female germline, as genome remethylation does not occur until the time of birth (Kohlrausch et al., 2022). In fact, foetal oocyte attrition seems to correlate with LINE1 retrotransposition. This attrition does not seem to involve an antiviral response. Instead, it seems that LINE1 upregulation causes DNA damage, that would trigger meiotic arrest (Malki et al., 2014; Tharp et al., 2020).

LINE1 control in the germline is not only relevant to prevent infertility, but also to prevent the transfer of insertional mutations into the next generation. LINE1 expression in early development also creates a big challenge for embryo survival and prevention of somatic mosaicism (Kano et al., 2009). However, in contrast to germ cell development, another developmental stage with a similar LINE1 behaviour, no specific protection mechanism against LINE1 has been identified in ESCs (Percharde et al., 2018). Another contrast with germ cells, is that in early development ESCs need to tolerate LINE1 RNA, as it is necessary for developmental progression (Percharde et al., 2018), while a role for LINE1 RNA in germ

cells has not yet been described. The response of ESCs to potential endogenous sources of ssDNA coming from TEs that are active in early development, has not been investigated and is still unknown. Therefore, there must be a mechanism protecting mESCs from the detrimental activity of LINE1 that is compatible with LINE1 RNA integrity, as in a normal somatic context, the exposure of the cells to TE expression leads to an anti-proliferative state, which would have obvious detrimental consequences during early development if ESCs (Borden et al., 1982; Eggenberger et al., 2019; Hertzog et al., 1994). Moreover, the potential activation of RnaseL and the subsequent degradation of cellular RNAs, some of them maternally inherited, could severely compromise the viability of early cleavage embryos (Bisbal et al., 2001).

To sum up, LINE1 is expressed many detrimental contexts like neuroinflammation, cancer, and senescence. At this stage, LINE1 produces an aseptic chronic inflammation in response to its ssDNA intermediates. On the contrary, LINE1 expression in pluripotency does not seem to impair the ESCs and embryo viability, nor it seems to produce a huge mutation burden through retrotransposition (Arokium et al., 2014; Ewing and Kazazian, 2010; Garcia-Perez et al., 2007; Kano et al., 2009). In fact, LINE1 plays an important role, being necessary for development to progress. The nature of these differences remains to be understood.

1.6. Hypothesis

I hypothesise that LINE1 expression in pluripotency has exerted an evolutionary pressure in order to avoid irreversible detrimental effects caused by the activation of LINE1 ORF2p enzymatic endonuclease activity and retrotranscriptase activity that would activate the antiviral programme on these cells, causing cell cycle arrest and cell death as it happens in somatic cells (Ardeljan et al., 2020; Hertzog et al., 1994) (Figure 1.9). Beyond the intrinsic effects on the cell per se, a detrimental response to LINE1 expression could trigger via paracrine signalling, a local inflammation with the recruitment of immune cells in the context of an early pre-implantation embryo in the maternal endometrium, that could have an effect in the maternal/blastocyst interface and implantation (Hertzog et al., 1994; Yockey and Iwasaki, 2018). I therefore propose that there is a mechanism in pluripotency of protection against LINE1 detrimental activity by disruption of the innate immune sensing of TE-derived ssDNA.

Secondly, I hypothesise that there could be complementary, pluripotency-specific proteins to prevent retrotransposition. These could bind the LINE1 ribonucleoprotein, either preventing the activation of innate immunity or preventing the enzymatic activities of ORF2p; which would

be complimentary with the low transposition rate of LINE1 at this stage (Arokium et al., 2014; Ewing and Kazazian, 2010; Garcia-Perez et al., 2007; Kano et al., 2009)..

Therefore, in my PhD I aimed to unravel what enables pluripotent cells to escape the detrimental effects of LINE1 expression and its retrotransposition. I set out to test these hypotheses in the following objectives:

- Objective 1: investigate differences in the innate immune response and TE DNA sensing between pluripotent and somatic cells. With this objective I aimed to test the pluripotent cell reactivity to nucleic acids and IFN and investigate the status of the DNA sensing pathway in pluripotency, by checking the expression of its components. My ultimate goal was to induce an antiviral response in mESCs. I found that mESCs cannot react to DNA because *cGAS* and *STING* are not expressed in pluripotency. I also found that they cannot react to IFN either, although the IFN signalling pathway seems intact. Finally, I managed to prove the detrimental effects for pluripotency of the activation of the antiviral response by over expressing a gain-of-function (GOF) mutation of IRF3 in pluripotency.
- Objective 2: characterise pluripotency and senescence-specific LINE1 Orf1p interaction partners. I characterised endogenous LINE1 pluripotency interactors by immunoprecipitating ORF1p in ESCs. This allowed me to identify pluripotency-specific interactors and screen the potential of some of these to inhibit LINE1 retrotransposition. This has made it possible to narrow down a group of pluripotency-specific LINE1 repressors.

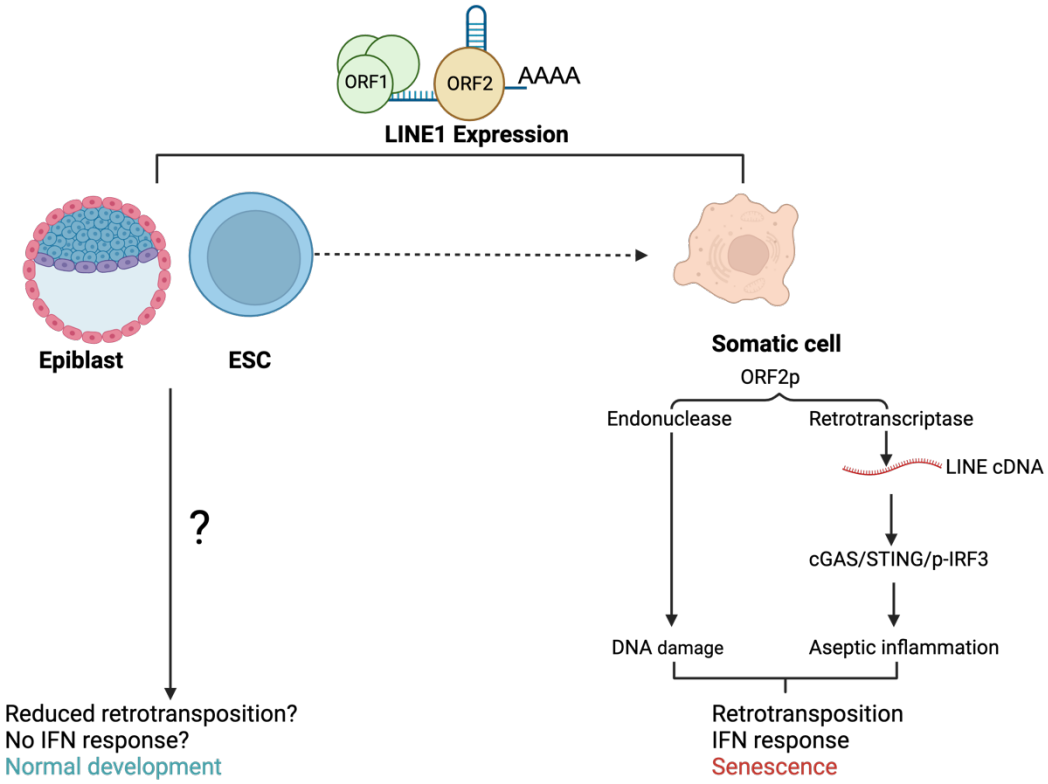


Figure 1-9: Project hypothesis graphic summary. My main aim is to understand why the expression of LINE1 in two different developmental contexts, can lead to two different response

Chapter 2: Materials and methods

2.1. .Materials

2.1.1. Chemicals solvents and Reagents

Chemicals and solvents were purchased from the suppliers listed below

Chemical	**Supplier**
Agarose	Sigma-Aldrich
β-Mercaptoethanol	Sigma-Aldrich
Bovine Serum Albumin	Sigma-Aldrich
BS3	Thermo-Fisher
Cut Smart buffer	New England Biolabs
Deoxynucleotide Triphosphate (dNTPs) 10mM	Invitrogen
DAPI	New England Biolabs
DNA loading dye	Qiagen
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich
Donkey serum	Abcam
EDTA	Invitrogen
Ethanol	Sigma-Aldrich
Ethidium Bromide	Thermo-Fisher
Foetal bovine serum	Gibco
Pierce™ 16% Formaldehyde (w/v), Methanol-free	Thermo-Fisher
Fugene HD Transfection Reagent	Sigma-Aldrich
HEPES	Thermo-Fisher
Hyalurionidase	Sigma-Aldrich
Human chorionic gonadotropin (hCG)	MSD Animal Health
Isopropanol	Sigma-Aldrich
Mouse IFNβ	Bio-Techne
KSOMaa	Sigma-Aldrich

LDS buffer 4x	Novex
Lipofectamine 2000 transfection reagent	Invitrogen
M-2 medium	Sigma-Aldrich
Methanol	ThermoFisher
Mineral oil	Sigma-Aldrich
MOPS running buffer (20X)	Invitrogen
OPTIMEM	Gibco
PageRuler Prestained Protein Ladder, 10 to 180kDa	ThermoFisher
Polyethylenimine	Merck
Phosphate-buffered saline (PBS)	Gibco
Polybrene	Bio-techne
Pregnant mare serum gonadotropin (PMSG)	MSD Animal Health
Protease inhibitors	Thermo-Fisher
Phosphatase inhibitors	Thermo-Fisher
siRNA buffer	Dharma
Skimmed milk Powder	Serva
Sodium Acetate	Invitrogen
Sodium Azide	Sigma-Aldrich
Sodium Chloride	Invitrogen
Sodium Deoxycholate	Sigma-Aldrich
Sodium Fluoride	Scientific laboratory supplies
Sodium Phosphate (dibasic)	G-Biosciences
Sodium Phosphate (monobasic)	Sigma-Aldrich
Sodium Orthovanadate	New England Biolabs
SYBR safe	New England Biolabs
Sytox Blue	Thermo Fisher
Triton X-100	Sigma-Aldrich
Tween-20	Santa Cruz
Vectashield with DAPI	Vector Laboratories

Table 2.1: Chemicals

2.1.2. . Kits

Kit Name	Supplier
Alkaline phosphatase staining kit	Vector
B-Galactosidase staining kit	Cell Signalling
Gateway cloning kit	Thermo Fisher
Pierce BCA Protein Assay Kit	Thermo Fisher
QIAquick PCR purification kit	QIAGEN
QIAquick PCR purification and Gel Extraction Kit	QIAGEN
QIAPrep Spin Miniprep Kit	QIAGEN
QIAPrep Spin Midiprep Kit	QIAGEN
Quick DNA 96-plus kit	Zymo
RNeasy Micro Kit	QIAGEN
Rneasy Mini Kit	QIAGEN

Table 2.2: Kits.

2.1.3. Antibiotics

Antibiotic	Dose	Supplier
Ampicillin	100µg/ml	Merck
Carbenicillin	100µg/ml	SLS
Chloramphenicol	20µg/ml	Sigma-Aldrich
Kanamycin	50µg/ml	Sigma-Aldrich
Puromycin dihydrochloride	1µg/ml	Stratech Scientific Ltd

Table 2.3: Antibiotics.

2.1.4. Enzymes

Enzyme	Supplier
DNase I (RNase-free)	QIAGEN
Protein A Dynabeads	Thermo-Fisher
Proteinase K	Invitrogen
Restriction Endonucleases (all)	New England Biolabs
RNase A (DNase-free)	Invitrogen
SYBR Green	Merck
Q5 high fidelity DNA polymerase	New England Biolabs
Quick calf intestinal alkaline phosphatase	New England Biolabs
Quick ligase kit	New England biolabs

FLASH Phusion high-fidelity DNA polymerase for cloning	New England Biolabs

Table 2.4: Enzymes.

2.1.5. Buffers, Solutions and Gels

TBE (Prepared by the Kitchen)

900mM Tris Borate
20mM EDTA pH8.0

TAE 50X

2M Tris base
1M acetic acid
50mM EDTA

TBS-T

130mM NaCl
20mM Tris pH 7.6
0.1%(v/v)

Blocking buffer

1/10 Donkey, goat or horse serum
25mg/ml BSA
Top up with PBS

FACS buffer

2% FBS

	0.5mM EDTA
Tris-EDTA (TE) Buffer	1mM EDTA 10mM Tris pH 8.0
Polyethylenimine Linear (PEI) 1mg/ml	Dissolve PEI in water heated at 80^{o}C Cool and adjust to pH 7.0 Sterile-filter (0.2um) (Store at -20^{o}C)
RIPA buffer	
	50mM Tris phH8 1mM EDTA 0,5mM EGTA 1% Triton CX100 0.1% deoxycholate de sodium 140mM NaCl Water (pure) 1:100 Protease inhibitor solution
BS3/sulfo-DSS 100mM	Volume/weigh 2mg BS3 34.93ul Conjugation buffer
0.1 M Na2HPO4 (Sodium Phosphate) buffer (1L)	
	20.214g Sodium Phosphate Dibasic 3.394g Sodium Phosphate Monobasic Water
Conjugation Buffer 50ml	

20mM Sodium Phosphate
0.15M NaCl
Water

Quenching Buffer

1M Tris HCl (pH 7.5)

TSE buffer

1% triton
2mM EDTA
20mM Tris pH8
150mM NaCl

Elution buffer

Desired concentration
50 mM HEPES, pH 8
1% (wt/vol) SDS
1% (vol/vol) Triton X-100
1% (vol/vol) NP-40
1% (vol/vol) Tween 20
1% (wt/vol) deoxycholate

50 mM NaCl

Blocking buffer

9ml PBS
1ml donkey serum
0.25g BSA

Bradley Lysis Buffer without Proteinase K

Tris-HCl (PH 7.5) 10 mM

EDTA 10 mM

SDS 0.5%

NaCl 10 mM

*Add Proteinase K just before use at final conc. 0.1 mg/ml

EtOH/NaCl mix

100% EtOH

5M NaCl

2.1.6. Plasmids

Plasmid Name	**Source**
4TO GFP	(Percharde et al., 2012)
mp783b	McManus lab (unpublished)
mp1052 (rosa-Multiple Cloning Site (MCS))	McManus/Percharde lab, unpublished
mp1052-rosa-smL1-Ef1a sense	Generated in this thesis
mp1072-rosa-dCas9-VPR	McManus/Percharde lab, unpublished
pENTR™ 1A Dual Selection Vector	Thermo Fisher Gateway cloning kit
pDONR 221	Thermo Fisher Gateway cloning kit
pCMV-3Tag-1a-BFP-T2A-STING HAQ V154M	Synthesized by GenScript
pCMV-3Tag-1a-BFP-T2A- STING V154M	Synthesized by GenScript

pCMV-3Tag-1a-BFP-T2A-IRF3 5D	Synthesized by GenScript
pCW 57.1	Addgene (ref: 41393)
pCW 57.1 hSTING V155M	Generated in this thesis
pCW 57.1 hSTING V155M HAQ	Generated in this thesis
pCW 57.1 BFP P2A	Generated in this thesis
pCW 57.1 BFP P2A mSTING V154M	Generated in this thesis
pCW 57.1 BFP P2A mSTING HAQ V154M	Generated in this thesis
pCW 57.1 BFP P2A IRF3 5D	Generated in this thesis
pFG04 (L1 UTR promoter)	(Percharde et al., 2018)
pJH447CAG_smL1_CMV	(An et al., 2006)
pJH447CAG_smL1_Ef1a	Generated in this thesis
pMDG2-vsvg	Gift from Louise Fets
pPax2	Gift from Louise Fets
pTRIP-SFFV-mtagBFP-2A	Addgene (ref: 102585)
pTRIP-SFFV-mtagBFP-2A STING V155M	Addgene (ref: 102591)
pTRIP-SFFV-mtagBFP-2A STINGHAQ V155M	Addgene (ref: 102592)
pX458-sgROSA-BFP	McManus/Percharde lab, unpublished
pWA125 (L1 CMV)	(Newkirk and An, 2017)

Table 2.5: Plasmids.

2.1.7. Antibodies

Antibodies used in the present work are listed below, including the dilutions used for immunofluorescence (IF) and western blot (WB).

Primary antibody	Host	Supplier	Catalogue number	Dilution (IF)	Dilution WB
αTubulin	Mouse	Sigma	T6074	NA	1:1000

cGAS	Rabbit	Cell Signalling Technologies	31659	NA	1:1000
Esrrb	Mouse	Novus Biological	H6705	NA	1:1000
Gapdh	Rabbit	Cell Signalling Technologies	2118	NA	1:2000
Gapdh	Mouse	Abcam	ab125247	NA	1:2000
IRF3	Rabbit	Cell Signalling Technologies	4302	NA	1:1000
Lin41/Trim71	Rabbit	Borrowed from Gregory Wulczyn	Present from Dr. Richard Gregory	NA	1:6000 for WB in PBS-T 0.2% Tween +3% BSA +Na azide 0.05%
Mov10	Rabbit	Abcam	ab80613	NA	1:1000
mORF1p	Rabbit	Abcam	ab216324	1:500	1:1000
Nfx1	Rabbit	Gift from Ann Roman	NA	NA	1:1000
Pdi	Mouse	Enzo	ADI-SPA-891-D	1:200	NA
Phospho γH2AX	Rabbit	Cell Signalling Technologies	2577	1:1000	NA
Phospho IRF3	Rabbit	Cell Signalling Technologies	4947	NA	1:1000
Phospho STAT1	Rabbit	Cell Signalling Technologies	9167	NA	1:1000

Phospho-STING	Rabbit	Cell Signalling Technologies	19781	1:500	NA
Phospho Tbk1	Rabbit	Cell Signalling Technologies	5483	NA	1:1000
p65	Rabbit	Cell Signalling Technologies	8242	NA	1:1000
Rb IgG	Rabbit	Cell Signalling Technologies	2729S	NA	NA
ssDNA	Mouse	Enzo	F7-26	1:250	NA
Sox2	Goat	R&Dsystems (Biotechne)	AF2018	1:500	1μg/ml
STAT1	Rabbit	Cell Signalling Technologies	14994	NA	1:1000
STING	Rabbit	Cell Signalling Technologies	13647	NA	1:1000
STING	Rabbit	Proteintech	19851-1-AP	1:500	NA
Tbk1	Rabbit	Abcam	ab40676	NA	1:1000
TGN38	Sheep	Biorad	AHP499G	1:500	NA
Vinculin	Rabbit	Cell Signalling Technologies	13901	NA	1:2000

Table 2.6: Primary antibodies.

Secondary antibody	Host	Supplier	Catalog number	Dilution (IF)	Dilution WB
Anti-rabbit IgG, HRP-linked Antibody	Goat	Cell Signaling Technology	7074S	NA	1:2000

Anti-mouse IgG, HRP-linked Antibody	Horse	Cell Signaling Technology	7076S	NA	1:2000
Anti-rat IgG, HRP-linked Antibody	Goat	Cell Signaling Technology	7077S	NA	1:2000
Anti-goat IgG, HRP-linked Antibody	Mouse	Santa Cruz	sc-2354	NA	1:2000
Anti-sheep 350	Donkey	Invitrogen	A-21097	1:1000	NA
Anti-Rabbit 488	Donkey	Invitrogen	A21206	1:1000	NA
Anti-mouse 555	Donkey	Invitrogen	A31570	1:500	NA
Anti-mouse 647	Donkey	Invitrogen	A32787	1:500	NA

Table 2.7: Secondary antibodies.

2.1.8. Chemical inhibitors and activators

Inhibitors were reconstituted in the appropriate solvent and aliquots stored at -80^{o}C.

Chemical inhibitor/activator	**Supplier**
2’3’cGAMP	Invivogen
5’-azacytidine	Selleckchem
Mg132	Selleckchem
Chloroquine	Merck
Doxycycline Hyclate	Merck
G3-YSD	Invivogen
PMA	Invivogen
Phosphatase inhibitors	Thermo-Fisher
Protease inhibitors cocktail	Thermo-Fisher
TSA	Selleckchem

Table 2.8: Chemical inhibitors and activators.

2.1.9. Tissue culture components

Foetal calf/bovine serum was stored long-term at -80°C, with working aliquots kept at -20°C. Aliquots of Penicillin/Streptomycin and Anti-Anti were stored at -20°C. All other components were stored at 4°C for up to two months before discarding.

Reagent	Supplier
ESGRO Anti-Anti	Gibco
Dublecco's MEM (DMEM), Modified Eagle Medium Glutamax	Gibco
Foetal Bovine Serum, Heat-inactivated (HI FBS)	Gibco
Gelatine solution 2%	Gibco
Laminin from Engelbreth-Holm-Swarm murine	Sigma-Aldrich
Leukaemia Inhibitory Factor (LIF)	Gibco
MEM Non-essential Amino Acids 10mM (NEAA) (100X)	Gibco
2-Mercaptoethanol 50mM (1000X)	Gibco
Neurobasal medium	Gibco
Penicillin/Streptomycin (100X)	Gibco
Triple Express	Gibco
Trypsin 0.25% 1X with phenol-red	Gibco

Table 2.9: Tissue culture components

2.1.10. Media

2.1.10.1. Bacterial Media and Agar

LB (prepared by the LMS kitchen)	10g Tryptone 10g NaCl 5g Yeast extract 950ml Water
LB Agar (prepared by the LMS kitchen)	10g Tryptone

	10g NaCl 5g Yeast extract 12g agar 938ml H_2O

2.1.10.2. Mammalian Tissue Culture Media

All media was stored at 4°C.

293T/MEF/3T3s	DMEM Modified Eagle Medium Glutamax supplemented with 10% HI-FBS
Senescent MEF	DMEM Modified Eagle Medium Glutamax supplemented with 10% HI-FBS 1X Anti-Anti
Embryonic Stem Cell (ESC) serum+ LIF Media	DMEM Modified Eagle Medium Glutamax supplemented with 10% HI-FBS 1X MEM NEAA 1X 2-Mercaptoethanol LIF 1:10,000
2X Freezing media	20% DMSO Serum

2.1.11. Strains/cells

2.1.11.1. Bacteria Strains

Stbl3. Chemically competent *Escherichi coli.* From ThermoFisher. Contains the *recA13* allele of the *recA* mutation which reduces nonspecific recombination occurrence in cloned DNA. The Stbl3 methylation-dependent restriction system is mutated (*mcrB, mrr* and *hsd*S20) allow construction of cloning genomic DNA. This strain has *end*A1+, a nonspecific endonuclease I that digests DNA.

2.1.11.2. Mammalian Cell Lines

293T. Human embryonic kidney cell line immortalized with adenovirus type 5 DNA (Belandia).

3T3. Immortalized fibroblast cell line that was isolated from a mouse embryo.

E14: mESCs isolated from a mouse blastocyst (Hooper et al., 1987).

E14 A1 cells: mESCs line with an EGFP under the control of a MuERVL LTR promoter reporter inserted in the Rosa26 locus. Generated in Percharde 2018 (Macfarlan et al., 2012; Percharde et al., 2018).

E14 hSTING-V155M: mESCs with a randomly inserted Dox-inducible *hSTING* V155M construct

E14 hSTING-V155M-HAQ: mESCs with a randomly inserted Dox-inducible *hSTING* V155M HAQ construct

E14 R26-LINE1-PFG: mESC line with a synthetic mouse LINE1 codon optimized EGFP reporter construct under the control of a CAG promoter. The EGFP is under the control of an *Ef1*α promoter. PFG stands for the reverse writing of GFP.

MEF. E13.5-derived mouse embryonic fibroblasts (MEFs). CD1 genetic background. Generated in lab.

2.1.12. Resources and websites

Software

- Cell Profiler.
- FlowJo.
- ImageJ.
- Leica LAS AF.
- Microsoft Office.

- Microsoft PowerPoint.
- Perseus.
- Prism.
- R studio.
- Zotero.

Online Tools

A single-cell molecular map of mouse gastrulation and early organogenesis	https://marionilab.cruk.cam.ac.uk/MouseGastrulation2018/
Benchlink.	https://www.benchling.com/
BioRender	https://biorender.com/
BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi
Clustal Omega	https://www.ebi.ac.uk/Tools/msa/clustalo/
EnrichR	https://maayanlab.cloud/Enrichr/
NEBcloner	https://nebcloner.neb.com/#!/
Pubmed	https://pubmed.ncbi.nlm.nih.gov/
Reprinnt Crapome	https://reprint-apms.org/
String	https://string-db.org/
UCSC Genome Browser.	https://genome.ucsc.edu/
VastDB	https://vastdb.crg.eu/wiki/Main_Page

2.2. Methods

2.2.1. Bacterial techniques

2.2.1.1. Transformation of bacteria

Stbl3 competent cells were used for this procedure. LB-agar plates (containing the relevant antibiotic) were removed from storage at 4ºC to room temperature. All manipulations were performed under a flame. 25~μL of competent cells were used per transformation (using always a non-DNA control) with 1-2μL of DNA in a microcentrifuge tube. The mix was incubated on ice for 25' and then heat-shocked at 42ºC for 45''. Tubes were put back on ice for 2'. Cells were then supplemented with SOC media without antibiotic, allowing bacteria to grow at 37ºC for 45' (1h for kanamycin) in order to develop the relevant antibiotic resistance

encoded in the backbone of the plasmid. The culture was then plated onto a 10cm LB agar plate containing the appropriate antibiotic. Plates were then incubated overnight with the agar base facing down to avoid condensation.

2.2.1.2. Plasmid amplification and purification

For small-scale plasmid DNA amplification, previously transformed colonies with the relevant construct were picked from plates and grown in 5mL cultures on LB-broth containing the right concentration of antibiotic. Cultures were grown overnight at 37ºC with agitation at 220rpm. The next day, bacteria from 4ml of the culture were pelleted by centrifugation at 4500xg, 15', and DNA extracted and purified utilizing the QIAprep Spin Miniprep Kit, according to manufacturer's guidelines. Purified DNA was eluted in 35µl of Elution buffer. The remaining 1mL was stored in the fridge in case the particular construct required a higher tier amplification.

For high-scale, high quality plasmid DNA suitable for transfection experiments, colonies grown during the day in 5mL cultures were inoculated in 100mL of LB-broth with the relevant antibiotic and incubated overnight in the same conditions stated before. Cultures were centrifuged at 6000xg for 15' and the pellets processed using the HiSpeed Plasmid Midi Kit. DNA was eluted in a final volume of 500-1000µL sterile TE buffer and stored at -20oC. In both cases the concentration of plasmid DNA was determined according to A260 readings using NanoDrop.

2.2.2. DNA Manipulation

2.2.2.1. Genomic DNA (gDNA) extraction

Two different protocols were used to extract gDNA depending on the posterior application.

<u>gDNA extraction for genotyping</u>

First, ES cells were grown in a 96-well plate until they were over-confluent. The plate was then quickly inverted to remove excess media and blot the cells on paper towels. The cells were washed once with 150µl of PBS and the buffer was removed by plate inversion.

Next, 50µl of Bradley Lysis Buffer was added to the wells, along with Proteinase K. The PCR plate sealing membrane was replaced and the plate was placed in a humidified chamber. The plate was then incubated at 60ºC for 5h to overnight.

After the incubation, the plates were cooled to room temperature. Then, 100µl of ice-cold EtOH/NaCl mix was added to the wells to precipitate DNA, and the mixture was mixed well. The plate was left to incubate at room temperature for 30'. The plate was then spun in a 96-well plate holder centrifuge at 3000 rpm for 20' to collect the DNA pellet. The liquid was decanted by inverting the plate. Next, 150µl of cold 70% EtOH was added to the wells, and the plate was spun at 3000 rpm for 10' to rinse the DNA pellet. The liquid was decanted by inverting the plate. This step was repeated, and the DNA pellet was air-dried for approximately 10-15'. After drying, 30µl of warm TE buffer (pH 8.0) was added to the wells. The PCR sealing membrane was replaced and the plate was incubated at 60ºC for 10-30' to fully resuspend the DNA pellet.

The plates were then ready for storage at -20ºC. Before use, they were incubated at 60ºC for 10-20' to fully resuspend the DNA.

gDNA extraction for qPCR

For gDNA qPCR, higher purity was needed. For this purpose, the Zymo Quick DNA 96-plus kit was used. Genomic DNA from pellets from up to a million cells was extracted from wild-type mESCs and F11s upon siRNA silencing. Washed pellets were then processed according to manufacturer's procedures. Concentrations will then be measured with nanodrop and the necessary dilutions performed.

2.2.2.2. Polymerase Chain Reaction (PCR)

PCR has been used for genotyping genetic insertions and for generating DNA fragments to be cloned into a target vector of interest. For the second case, oligonucleotide primers contained 20bp of the target sequence flanked by the relevant restriction endonuclease target site. Four bases were added to the 5' end of each primer to allow efficient digestion by endonucleases. For cloning, amplification was carried out utilizing the Q5 high fidelity DNA polymerase in a total volume of 50µL with the corresponding Q5 reaction buffer and Q5 High GC Enhancer, and 10mM dNTPs. Primer concentration was 10µM and the mass of template plasmid was 1-10ng.

For genotyping, amplification was carried out utilizing the Go-Taq polymerase reaction master mix in a total volume of 50µL, 25ng of gDNA, 500nM primer.

The following conditions were followed:

For generation of Ef1α promoter:

98ºC	30''	
98ºC	10''	X30 cycles
58ºC	30''	
72ºC	30''	
72ºC	2'	

For genotyping CAG promoter insertion in the *Rosa26* locus:

95ºC	3'	
95ºC	30''	X34 cycles
59ºC	20''	
72ºC	1'	
72ºC	5'	

In the case of a cloning experiment, to isolate the right size amplified fragment, each reaction was analysed by agarose gel electrophoresis 1% (w/v) containing SYBR safe for DNA visualization. Electrophoresis was carried out in 1X TBE buffer and run at 100V for 1h. DNA fragments were then excised and purified using the QIAquick PCR Gel Extraction Kit and eluted in 40µL H_2O, ready for digestion. Primers were obtained from published data or designed around 50% GC content and between 80 and 120bp PCR product length for genotyping. See primer sequences in Annex I.

2.2.2.3. DNA Digestion

Insert and linearized vector DNA were obtained by restriction endonuclease digestion of the relevant plasmid or purified PCR product according to New England Biolabs (NEB) guidelines. Reactions were performed in a total volume of 50µL containing around 1-2µL of each appropriate restriction endonuclease, in the specified NEB buffer. Mass of DNA digested corresponded to the whole PCR product, or 1-3µg of plasmid DNA, depending on the need. Digestions were performed at 37ºC for 2-5h**,** followed for the required plasmids, by an inactivation step of 82ºC for 20'. Digested PCR products were purified using the QIAquick PCR Purification kit. Fragments released form plasmid DNA digestion were subjected to gel electrophoresis, excised and purified as stated before.

2.2.2.4. DNA Ligation

Ligation reactions were performed using the NEB Quick ligation kit according to manufacturer's procedures. 50ng of linearized vector DNA was incubated with 3- or 5-fold

molar excess of insert DNA, 1X Quick ligation buffer and 1ul of Quick DNA Ligase in a reaction volume of 20µL. Reactions were performed for 5' at room temperature (RT) and then chilled on ice for 5'. 1-2µL of the product were then immediately transformed into Stbl3 *E.Coli* strains. All ligations had a vector only control that was included in transformations. Correct sequence was first checked by restriction digest followed by Sanger sequencing performed by the company TwistBiosciences.

2.2.3. Generation of BFP pCW 57.1 constructs

pTRIP-SFFV-mtagBFP-2A purchased from Dr. Nicolas Manel lab via Addgene and pcW57.1 purchases from Dr. David Root lab via Addgene, were digested with XhoI and NheI. Both plasmids were run in a 1% agarose gel and the bands corresponding to BFP and the pcW57.1 backbone were cut from the gel, extracted and purified, as reported previously. Then, ligation between both fragments was performed at 1:3 ratio and the product transformed into Stbl3 bacteria and amplified. Upon restriction map validation of the plasmid miniprep, sufficient stock of plasmid was built by midiprep.

2.2.4. Generating hSTING V155M and STING HAQ mESCs and MEF cell lines

pTRIP-SFFV-mtagBFP-2A STING V155M (Addgene: 102591) and pTRIP-SFFV-mtagBFP-2A STING HAQ were a gift from Nicolas Manel via Addgene. The STING construct was then cloned onto a pENTR^{TM} 1A Dual Selection vector (Thermo Fisher scientific, USA) using restriction digest with BamHI and XhoI. The constructs were then transferred into the destination vector pCW57.1 (Addgene: 41393), which contains a tetracycline/doxycycline inducible promoter, by gateway cloning (Thermo Fisher scientific, USA). Every step was checked by restriction digest and the final product was validated by sequencing.

Stable cell lines were later established by lentiviral infection and kept in posterior selection with puromycin (1µg/ml). Cells were then kept in puromycin constantly while in culture.

2.2.5. Generation of mSTING V154M and IRF3-5D pcW 57.1 constructs

pCMV-3T BFP T2A constructs containing mouse *STING* or mouse *Irf3* with the relevant mutation were synthesized by GeneScript. All constructs were flanked by attL sites which were used to transfer the constructs in to the pcW57.1 destination vectors by gateway cloning. Products were then transformed into bacteria and selected with ampicillin. Colonies were then grown and mini-preped as stipulated before and the final products validated by restriction digest followed by sequencing. Validated products were then amplified and midipreped.

Stable cell lines were later established by lentiviral infection and posterior selection with puromycin (1µg/ml). Cells were then kept in puromycin constantly while in culture.

2.2.6. Generation of R26-LINE1-PFG construct

The Ef1α promoter was amplified from the plasmid mp1072-Rosa-dCas9-VPR. Purified *Ef1α* fragments (1172bp) and pJH447CAG_smL1_CMV plasmid were then digested with NheI and BamHI, ligated and amplified. Miniprep products were checked by restriction digestion and checked by sequencing. Then, mp1052-ROSA_MCS and pJH447CAG_smL1_Ef1a were digested by NotI. Mp1052 was linearised and the pJH447CAG_smL_Ef1a fragment corresponding to the reporter was excised. Upon purification, ligation and transformation, mini-prepped plasmid clones were restriction digested. As NotI is a single cutter in both plasmids, the reporter insertion in the ROSA plasmid could be sense or anti-sense. The correct orientation was validated by further digestions with NheI and AhdI enzymes, then only the sense clones were amplified, checked by Sanger sequencing and one correct plasmid clone used for the following experiments. The reason to replace the CMV promoter for *Ef1α* promoter is owing to the fact that CMV is repressed in pluripotency (Cabrera et al., 2022; Herbst et al., 2012).

2.2.7. Generation of E14 R26-LINE1-PFG reporter cell lines

4.25µg of mp1052-Rosa-smL1-Ef1a sense was co-transfected with 0.5µg pX458-sgROSA-BFP into wt E14 mESCs. Excess reporter construct was used to ensure that once the BFP-positive cells were sorted, it would be more likely that a double transfection had occurred. BFP-positive cells were sorted 48h later and grown in dilutions in 10cm plates in order to ensure the emerging colonies were sparse enough to be discriminated from one another. A week later, colonies were picked and each one of them was transferred to an independent well of a 96-well plate. 24h later, the colonies in each well were trypsinised, the cells individualised, and transferred to 3 new plates. Some of the plates were frozen, others were used for DNA extraction in order to perform genotyping and the rest kept on culture to build up the stock of the previous two.

The genotyping strategy consisted of amplifying the CAG promoter inserted into the *Rosa26* locus and the amplification of the endogenous *Rosa26* locus upstream of the sgRNA cut site. Genomic DNA from a different cell line containing a CAG promoter insertion in the *Rosa26* locus was used as positive control. CAG insertion in the locus was genotyped with the primer pair RW174 and RW067 generating an amplicon of 650bp. The wild-type locus was amplified with RW174 and Rosa_endogenous_R.

A total of 11 clones containing an insertion were identified. Of these, 10 were heterozygotes while one (F11) contained a homozygous insertion. The homozygous clone was then later expanded and used for experiments. Heterozygous clones were also expanded and frozen stocks made.

2.2.8. Cell culture methods

2.2.8.1. Maintenance of cell lines

Mouse E14Tg2A (E14) ES cells (male) were cultured on 0.1% gelatine-coated plates in ES-FBS culture medium (high glucose DMEM GlutaMAX (Thermo Fisher Scientific, USA), 10% FBS (Thermofisher, USA) 0.1mM Mercaptoethanol (Millipore) and 1,000U/ml LIF supplement (ESGRO, Millipore). For spontaneous differentiation experiments, 200,000 cells were plated in non-LIF supplemented media for 5 days in 10cm plates, changing media every day. Media was routinely changed every day and cells passaged every other day using routine aseptic tissue culture practice. Media was removed, cells washed in PBS and kept in Gibco TrypLE Express for 5'. Then, TrypLE Express is neutralized with serum containing media and cells centrifuged at 400g for 3'. Cells are then resuspended in media and plated at the right dilution.

MEF cells were cultured in 10% serum DMEM media (Thermo Fisher Scientific). For senescence induction cells were grown until 70% confluence and treated with 20μM etoposide for 3 days, after which etoposide was removed and cells washed with PBS then cultured in normal media, changing media every 3 days. This procedure was performed with the guidance of Dr. Alexis Barr. For long-term senescence, senescent MEF were kept in culture for 72 days, renewing the media every 3 days. For maintenance in culture, Anti-Anti was used as indicated in manufacturer's procedures.

HEK 293T cells were cultured in 10% serum DMEM media (Thermo Fisher Scientific). Cells were passaged when 70% confluence was reached in routine tissue culture practice as previously explained.

Cells were frozen in serum 10%DMSO (freezing media) in total volume of 1mL in freezing vials (Nunc). For freezing clones in a 96-well plate, the media was removed completely and the cells were washed with 150μl of PBS per well. Then, 25 μl of trypsin was added and the cells were incubated for 5 minutes at 37°C. After that, 75 μl of 1X freezing medium was added. The plate was put on ice for about 2-3 minutes and then sealed with Parafilm. Two paper towels were taken and the plate was wrapped in them using sticky tape. The plate was then put at -20°C for about 15 to 30 minutes. Finally, the plate was stored at -80°C.

2.2.8.2. Treatments of cell lines

Dox treatment: stable and transfected cell lines containing a Dox inducible cassette were stimulated using 1-2µg/ml doses of doxycycline for 36-48h. This cell lines mainly being mESCs with randomly inserted Dox inducible hSTING-V155M, hSTING-HAQ-V155M and mESCs and HEK 293Ts and mESCs transfected with *mSTING*, *hSTING* and *mIrf3* with the corresponding mutation.

Interferon treatment: mESCs, MEFs and blastocyst were treated with 200 and 1000IU of mouse IFNβ at different timepoints. 100µg/ml IFNβ storage stocks (equivalent to $1.2*10^8$ IU/ml) are prepared in 0.1% PBS BSA (10mg of BSA in 10ml sterile PBS). From this, a 1000X working stock is made in 0.1% PBS BSA by diluting 1:120 the storage stock. The 1X working solution is then finally prepared in culture media.

2.2.8.3. ESC differentiation experiments

LIF withdrawal: ESC differentiation was accomplished via LIF withdrawal from ESC media. Media was replenished every 24h and cells were harvested 96h later.

Embryoid bodies (EBs): This protocol was performed by the MSc student, Alexandros Chorormidis. Cells were trypsinised and prepared in a falcon in a solution of half a million cells in 10mL of -LIF media, to achieve 1000 cells per 20µL (the size of a drop). Cells were then placed in a reservoir. Then, 20µl drops were placed with a multichannel on the inner part of a 15cm non treated plate lid, Separating drops enough to avoid fusion events. Following that, 20ml of PBS was added to the bottom of the plate to prevent the drops from drying. Then, the lid containing the droplets was turned upside down with a single, smooth motion. The plate was then left in the incubator for 48h.

Two days later, the EBs were collected from the lid with 10ml of -LIF media and transferred into a new plate with an additional 10mL of media. The plates were shaken regularly to prevent adhesion of the embryoid bodies to the surface. Half of the media was regularly changed every two days. At this point RNA and protein samples were taken at any time, up to 10 days of differentiation.

2.2.8.4. MEF isolation

MEF cells were isolated by Dr. Michelle Percharde from E13.5 CD1 embryos with the head and visceral tissues removed, then digested with 3ml of warmed trypsin plus 0.8mg/ml DNase

for 10'. Embryos were then dissociated by passing through 18G needles 5 times then incubating for a further 4' at 37°C. Trypsin was inactivated via addition of MEF media. Embryos were then pooled and centrifuged at 200xg, 5' and then resuspended in MEF media. Cells were then transferred to 15cm plates and cultured in MEF media plus Pen/Strep (Thermo Fisher Scientific) labelled as passage 0 and tested negative for Mycoplasma. These cells were used within 3 passages of their derivation.

2.2.8.5. DNA transfection

Depending on the size or the plasmids and the volume of culture (shown on table), plasmid DNA was transfected into cells using 3µl of Lipofectamine 2000 per µg of DNA in Optimem. Optimem plus Lipofectamine mix was first prepared and incubated for 5'. During this time, Optimem plus DNA mix was prepared. DNA and lipofectamine mix were then combined and incubated in the dark for 20' and then pipetted dropwise into the culture. Transfections were normally performed in suspension with newly plated cells. All transfections were performed in antibiotic free media. Media was renewed 8h later, or the next morning in case of overnight transfections.

Plate/Dish	Area (cm^2)	Culture Vol (µL)	Fold x 96wp	Tot vol transfection mix	Tot DNA for transfection
10cm Dish	60	10000	193.5483871	1mL	5-10µg
12-well Plate	3.66	1000	11.80645161	250µL	600ng
15cm Dish	147.8	15000	476.7741935		
24-well Plate	1.91	500	6.161290323		
35mm Dish	9.3	2000	30		
6-well Plate	9.3	2000	30	500µL	1.5µg
96-well Plate	0.31	100	1	20µL	20ng

Table 2.10: Transfection mix Optimem/Lipofectamine 2000 rates.

In the case of vectors containing a fluorescent marker, fluorescence was assessed every day.

<u>siRNA screen transfection:</u> the siRNA screen was performed in 96-well plate format transfecting 3pmol of siRNA per well in a 2000 cells suspension. The initial number of cells was determined after a series of optimizations. The siRNAs were reverse transfected in suspension the first day and again the next day, to the attached cells.

The 0.25nmols of siRNA were spined down and resuspended in siRNA buffer at a stock concentration of 10μM, and kept in shaking in the thermomixer for 30' at room temperature and spined down again. For the assay, a 0.5μM dilution in siRNA buffer was performed in v-bottom plates. Optimem plus siRNA mix was prepared by taking 13.2μl of the 0.5uM siRNA solution and adding 14.2μl of Optimem. In parallel an Optimem and Lipofectamine 2000 was prepared considering 0.17μl of Lipofectamine-2000 and 12.33μl of Optimem per well. The Optimem/ siRNA solution was then incubated in the dark for 20' with the Optimem/Lipofectamine-2000 mix. The resultant solution was then added to the cell suspension. Media was renewed next day. For analysis of GFP-positive retrotransposition events read by flow cytometry, cells were kept in culture for 96h before being analysed. In the case of assaying for gH2AX staining, only one transfection was performed and cells were prepared for IF 36h later.

cGAMP/G3-YSD transfection: 2'3'cGAMP and G3-YSD ligands powder was reconstituted in water at a stock concentration of 1mg/ml. 8 μg/ml of 2'3' cGAMP (Invivogen, USA) and 1μg/ml of G3-YSD ligands (Invivogen, USA) with the corresponding controls were transfected per well in a 12-well plate using 2.3μl of Lipofectamine 2000 per well (Thermo Fisher Scientific, USA), topping up with Optimem (Thermo Fisher Scientific, USA) to 250μl. After incubating in the dark for 20', DNA mix was added to each well dropwise. RNA and protein samples were harvested 6h after transfection in the case of DNA ligands and 8h post transfection in the case of 2'3'cGAMP, or at the indicated times noted in the figure legend.

2.2.8.6. Colony formation essay

Approximately 500-1000 mESCs per well were plated in a 6-well plate or 500 in a 12-well plate. Cells were kept in culture 6 days with daily renewals of media +/- 1μg/ml Dox. On the last day cells are fixed with 4% PFA and stained with an Alkaline phosphatase kit (Vector, USA) according to manufacturer's procedures. This kit stains undifferentiated cells red. Alkaline phosphatase staining is considered a stem cell marker. Alkaline phosphatase catalyses the hydrolysis of phosphate ester in an alkaline environment, resulting in the formation of an inorganic radical and an inorganic phosphate.

2.2.8.7. Flow cytometry

Cells were individualized, washed and transferred to filtered FACS tubes after resuspension in FACS buffer containing 1:8000 Sytox blue, unless sorted for BFP fluorescence. Sytox Blue was added into the FACS buffer as a viability dye. Samples were then run in a BD LSR II

analyser with the appropriate lasers. Data were then exported as fcs files and analysed using FlowJo. If the cells were not sorted and only analysed, 1:8000

2.2.8.8. Virus work

Confluent 6-well plates of 293Ts were transfected with 2.5μg of the expression vector (1 per well), 1.5μg of the lentiviral packaging plasmid, psPax2 and 0.6μg of the packaging plasmid pMDG2-vsvg. Polyethylenimine (PEI) (Polysciences) was then added to the mix at the ratio of 2.5μl per 1μg of DNA. The mix was then topped up to 250μl with OPTIMEM and incubated for 30', before adding dropwise to cells. Virus was harvested from the supernatant 48h post transfection, and filtered through a 0.45μm PES filter then concentrated using Amicon Ultra-4 Centrifugal Filter units (100,000NMWL) from Millipore and stored at -80°C or used immediately. For lentiviral transduction, 100μl of the supernatant containing virus was pipetted drop-wise onto cells (MEF/E14) in media containing 4-8μg/mL final concentration of polybrene and left overnight. Media was changed the next day and 1μg/ml puromycin selection started 48h post infection and kept on cells constantly. Manipulation was performed according to indicated safety procedure rules.

2.2.9. Immunofluorescence (IF)

mESCs and MEF were plated on matrigel-coated 8-chamber coverslips (Thermofisher, USA) for 2h and fixed for 15' with 4% paraformaldehyde (Thermofisher, USA) in Phosphate Buffered Saline (PBS) at room temperature (RT). To permeabilize, different concentrations of Triton were used depending on the staining, with the standard condition being 0.1% triton for 10'. Samples were then blocked for 30' in blocking buffer, then incubated overnight in a humidified compartment at 4^{o}C to prevent the samples from drying, with primary antibody (Ab) in blocking solution at the previously specified dilutions. Samples were washed 3 times the next day with Phosphate-buffer saline (PBS). Detection was performed by incubating samples with the appropriate secondary antibodies for 45' at RT protecting samples from light washed 3 times with PBS, and mounted in Vectashield solution (Thermofisher, USA) with 4′,6-diamidino-2-phenylindole (DAPI) as counterstaining. Stained samples were imaged using a confocal laser-scanning microscope (Leica TCS SP5) with 40x dry and 63x oil immersion objective lenses and processed using FIJI.

For the gH2AX screen, cells were transferred to a Perkin-Elmer 96-well black clear bottom plate coated with laminin. Then they were fixed, permeabilized and blocked as stated above and stained with anti gH2AX primary antibody in the condition found on table 2.6, following

the regular IF protocol. Plates were later imaged in an Operetta CLS High-Content Analysis System Microscope (PerkinElmer)

2.2.10. Embryo and mice work

All animal experiments were performed according to a UK Home Office Project License in a Home Office-designated facility using 4- to 6-week-old female and 2- to 6-months-old male CD1 mice purchases from Charles River. All mice and embryo experiments were performed by Dr. Bryony Leeke. Animals were maintained on a 12-h light/dark cycle and provided with ad libitum food and water in individually ventilated cages. Female mice were superovulated by intraperitoneal injection of 5 IU of pregnant mare serum gonadotropin (PMSG; Folligon, MSD Animal Health), followed by 5 IU of human chorionic gonadotropin (hCG; Chorulon, MSD Animal Health) 46–48h later, and then placed immediately with males. Zygotes were collected from oviducts ~22–24 h after hCG in M-2 medium (Sigma M7167), isolated from cumulus cells with 200 µg/mL hyalurionidase (Sigma H3506), washed through successive drops of M-2, and then cultured in preequilibrated KSOMaa (Sigma MR-106-D) in microdrops overlaid with mineral oil (Sigma M5310) or in four-well dishes. Zygotes were cultured in a humidified incubator at 37°C and 5% CO2 until early two-cell (31–33h after hCG), late two-cell (48– 49 h after hCG), morula (3 d after hCG), or blastocyst (4 d after hCG) stage.

2.2.11. Western blot

Whole cell extracts were prepared from ESCs in ice-cold RIPA buffer (50mM Tris phH8**,** 1mM EDTA, 0,5mM EGTA, 1% Triton X10, 0.1% sodium deoxycholate, 140mM NaCl) containing Halt protease inhibitors (Thermofisher, USA). To eliminate the debris, extracts were incubated in a rotating wheel at 4^{o}C for 15', followed by a 14,000rpm 15' centrifugation step and took the supernatant. For very DNA rich and viscous samples, the supernatant is run through a 1ml insulin needle multiple time. Samples are quantified using the Pierce BCA protein assay (Thermofisher, USA), according to manufacturer's recommendations using 96 well plates analysed at 467nm absorbance on a Fluostar omega plate reader (BMG Labtech, USA). 20-40µg of total protein was loaded per sample. Proteins were separated in a Bolt Bis-Tris Plus 4-12% gel, using Prestained protein plus ladder (Thermofisher, USA). After this step, a dry transfer was carried out onto a PVDF membrane using the iBlot2 Dry blotting system (Thermofisher, USA). The membrane was then blocked for an hour in 5% w/v powder milk in TBST. Membranes were then incubated overnight at 4^{o}C in primary antibody as depicted in table 2.6.

The next day membranes were washed in Tris-buffered saline TWEEN 0.2% (TBST) and incubated in the appropriate anti-rabbit/goat secondary antibodies conjugated to HRP in 1:2000 dilutions: anti goat-HRP (Santacruz, #sc- 2354), anti-rabbit-HRP (Cell signalling, #7074S) and anti-mouse-HRP (Cell signalling, #7076S) as depicted in table 2.7. Membranes were imaged in an Amersham imager 680 (Cytivia, UK). Quantification was performed using the software ImageJ and Prism for the statistical analysis.

2.2.12. Immunoprecipitation

Whole cell extracts were prepared from ESCs by adding to 10cm diameter confluent plates 1ml of ice-cold RIPA buffer containing protease inhibitors, prepared as for WB and quantified using Pierce BCA protein assay kit. Protein A magnetic beads were washed and prepared with 2.5µg the appropriate antibodies: ORF1(Abcam, #ab216324), anti-MOV10 (Abcam, #ab80613) and control Rabbit IgG (Cell Signalling Technology, #2729) as negative control. 30µl of protein A dynabeads were bound to antibodies by incubating them with 2.5µg of Ab in 200µl of RIPA buffer for 20 minutes at room temperature. The beads were then crosslinked to the antibody using 5mM BS3 in conjugation buffer (20mM NaH_2PO_4, 0.15M NaCl in H_2O) for 30 minutes at room temperature. This step was carried out in order to prevent antibody elution along with immunoprecipitated protein. Beads bound to the appropriate antibodies were used to immunoprecipitate the proteins of interest. 1mg of total protein extract plus inhibitors was incubated overnight with 25µl of magnetic beads in a rotating wheel. Next day immunocomplexes bound to the beads were washed first in RIPA and then twice for 1 minute in TSE buffer (1% triton 2mM EDTA, 20mM EDTA, 20mM Tris pH8 and150mM NaCl). In experiment one, samples were eluted by boiling 5' at 90^{o}C in LDS buffer + 2mM DTT. For subsequent experiments, 3 replicates of samples were eluted in elution buffer (50mM HEPES ph8, 1% (wt/vol) SDS, 1% (vol/vol) Triton X-100, 1% (vol/vol) NP-40, 1% (vol/vol) Tween 20, 1% (wt/vol) deoxycholate, 5 mM EDTA, 50 mM NaCl, 1% (vol/vol) glycerol) provided by Dr. Ilyana Kaneva, to allow submission to LMS mass-spectrometry (MS). The IP was validated by performing a Western blot for Orf1p and Mov10, using the previously mentioned protocol. For first experiment all the IP volume was loaded into the gel. For the second experiment (3 replicates), 5% of the total IP volume or unbound fraction was loaded into the gel for western blot validation, and the rest was submitted to mass spectrometry.

2.2.13. Mass spectrometry

IPs (ORF1 and Rb IgG control) with 3 replicates each, were submitted for MS analysis. Samples were processed with SP3 and digested with trypsin. Each sample was run in 2

technical replicates on the MS (HFX). Data was processed using the MaxQuant software, IPs were not normalized. Statistical analysis and visualizations were done in Perseus. Missing values were imputed assuming a normal distribution. A non-parametrical Welch t-test was used for the statistical analysis. Fisher's exact test was then applied for the gene ontology analysis. This procedure was performed by Dr. Alex Montoya and Dr Holger Kramer (LMS mass-spectrometry facility).

For the senescence IP second attempts, beads were handled by the LMS proteomics facility, and on-bead digestion was performed by the facility a posteriori.

2.2.14. RNA extraction

Total RNA was extracted from mESCs, 293T and MEF samples using the RNeasy Mini kit (Qiagen, USA) according to manufacturer's procedures, after washing with PBS and lysis in RLT buffer. To increase RNA preservation, 1:100 2-mercaptotenol was added to the RLT lysis buffer. Samples were digested with RNase-free DNase (Qiagen) during RNA extraction as recommended, eluted in 35μL of RNase-free water and quantified using nanodrop. Up to 1μg of RNA in a maximum volume of 9μl was used, when possible (MEF tended to give low concentrations of RNA) in the cDNA generation reaction using the High-capacity RNA to cDNA kit (Thermo Fisher, USA) according to the manufacturer's protocol. The effectiveness of the DNase digestion was measured using a minus RT enzyme control reaction. For experiments involving sorted cells and embryos the RNeasy Micro kit was used according to manufacturer's procedures. In this case the whole eluate was used for cDNA generation.

2.2.15. Quantitative PCR:

RT-qPCR of cDNA was performed using SYBR green kit (Thermo Fisher scientific, USA) on a QuantStudio5 (Thermo Fisher Scientific, USA) in 384 well plates (Thermo Fisher scientific, USA), using 4ng of cDNA per well. Each reaction per primer pair (see Annex I) was run in triplicates per each sample. Samples were amplified for 40 cycles. The arithmetic mean of *Rpl7* and *H2A* housekeeping genes was used to normalize the samples. Primers were used in 200nM final concentration.

gDNA qPCR: gDNA isolated from wt mESCs and F11 ESCs was used for this technique. To quantify the number of EGFP copies and genomes a series of standard curves for EGFP and intragenic region in ch.11 was first needed. For the EGFP standards, EGFP was amplified in serial dilutions of genomic DNA from the mESCs A1 cell line, which contains an EGFP

insertion. This product was obtained after overnight digestion with BamHI. For the genome standards, amplification of ch.11 in serial dilutions of wt mESCs genomic DNA was performed.

The reaction volume for 384-well plates was 10μl per well, with a total of 10ng of gDNA per well, divided in a volume of 4.6μl, being the other 5.4μl the primer and 2xSYBR mix. Primers were used at a final concentration of 200nM.

Considering I was using a ch. 11 specific intragenic-region, with 2 copies per mouse genome. A mouse haploid genome contains 2,728,222,451bp as per GRCm39 assembly. I considered that a bp is 1.096e-21g. Therefore, if a mouse haploid genome weighted 2.9901318e-12g, that would be 2.9pg. As ch. 11 had one copy per haploid genome, this means one copy per 2.99pg of gDNA. This allowed me to build standard curves by transforming the number of estimated copies into an estimated mass and then into an estimated concentration of material per well, which allowed me to build the standard curve using the formula C1*V1=C2*V2 to calculate the volumes needed to establish the serial dilutions. See example bellow:

Calculate calibration curve ch.11

N copies	pg/h genome	Mass gDNA	Vol sample	Concentration
4250	2.99	12707.5	4	3176.875
3000		8970		2242.5
2000		5980		1495
1000		2990		747.5
500		1495		373.75

	Initial (pg/μl)	volume of gDNA	Volume of diluent (μl)	Final volume (μl)	Final concentration
Stock	3200	19.85546875	0.14453125	20	3176.875
Dil 1	3176.875	14.11764706	5.882352941	20	2242.5
Dil 2	2242.5	13.33333333	6.666666667	20	1495
Dil 3	1495	10	10	20	747.5
Dil 4	747.5	10	10	20	373.75

For the EGFP calibration curve I used the E14 A1 cell-line, which harbours a GFP insertion in its genome. Therefore, I performed the same calculations assuming there is a copy of GFP per genome and half a copy per haploid genome. Each primer pair was run in triplicates per each sample. Samples were amplified for 40 cycles.

The technique allows for up to 10ng of DNA, considering the mouse haploid genome is 3pg (in this case a little bit less as we are using male cells and the Y chromosome is smaller), a qPCR reaction would include a maximum of 1,800 cells, which is a limitation if we consider the rarity of the event and the fact that in the flow cytometry experiment, each well contained ≅30,000- 50,000 cells.

2.2.16. RNA-seq

2.5 million mESCs were transfected in suspension overnight with 8µg of pCW57.1 BFP only vector, pCW57.1 BFP-T2A-STING-V154M-HAQ, pCW57.1 BFP-T2A-STING-V154M and, pCW57.1 BFP-T2A-IRF3-5D in the presence of 1.5ug/ml of Doxycycline (Dox). The next day, Dox media was renewed. 36-48h later, BFP positive cells were collected by FACS into mESC media. Cells were pelleted then lysed using RLT buffer and RNA extracted as described above. RNA quality and quantity of the samples was performed by Ivan Andrew from the LMS Genomics facility using a BioAnalyser. Samples were then sent for library preparation and sequencing by Novogene.

2.2.16.1. RNA-seq analysis

Stranded RNA-seq libraries were prepared and then sequenced by Novogene, to generate 9Gb of PE150bp raw data per sample. RNA-seq data were mapped by Dr Tarren Taylor. Library quality was evaluated using fastQC and multiQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ and https://github.com/ewels/MultiQC) . Reads were processed using TrimGalore (https://github.com/FelixKrueger/TrimGalore) and subsequently mapped to the mm10 genome and counted using SQuIRE on default settings (Yang et al., 2019).Subsequent analyses were performed by Dr Michelle Percharde, using statistical package DESeq2 for analysis and generation of graphs (Love et al., 2014).

Chapter 3: Results: Characterization of the sensitivity of mESCs to nucleic acids.

3.1. Introduction

Studies have established the incapability of embryos to react to viruses and other PAMPs and DAMPs (D'Angelo et al., 2018b; Witteveldt et al., 2019). When this project started, the inability of mESCs to sense dsRNA and cytosolic DNA was also reported. A mechanism was described in the case of dsRNA sensing, but the incapability of ESCs to react to cDNA remained elusive. Because immunogenic ssDNA derived from LINE1 is one of the main focuses of this project, I wanted to test the capability of mESCs to activate the cGAS/STING pathway. Moreover, I complemented this with a deep analysis of the expression of the different components of the cGAS/STING pathway in pluripotency. The intention was therefore to characterise the functionality of this DNA sensing pathway in pluripotency that could explain the lack of reactivity of mESCs to LINE1.

I assessed the expression levels of the components of the pathway and other innate immune system-related genes in mESC in different contexts: mESCs differentiation datasets, analysis of RNA expression levels by RT qPCR, analysis of protein levels by WB. Non-in *silico* experiments used MEF as a somatic cell positive control.

I first wanted to assess the capacity of the mESCs to react to DNA by testing the functionality of the cGAS/STING pathway. To do that, the pathway was first challenged at different points with stimuli that would activate it (Figure 1.4). G3-YSD, a 26-mer DNA ligand derived from the HIV genome, displays a Y-shaped palindromic structure that is a potent cGAS agonist (Herzner et al., 2015). Cells were also exposed to the second messenger 2'3'cGAMP to activate STING, then gain of function mutations of STING and IRF3 used to activate the pathway downstream of these proteins.

The objective of these different levels of targeting is to localize, in the case of defective functionality, the point at which the pathway would be dysfunctional. Finally, the functionality of the STING pathway was analysed by assessing the cellular localization of STING and IRF3 during the ER to TGN transition. This transition is necessary for STING to undergo several post-translational modifications (PTMs) that render it functional and capable of assembling a signalosome that activates IRF3 and NFκB in the TGN.

3.2. Results: DNA sensing in mESCs

3.2.1. mESCs do not react to G3-YSD and 2'3'cGAMP transfection

I wanted to test if mESCs can activate the cGAS/STING pathway, which is a key sensor of TE-derived DNA. To do so, E14 ESCs and mouse embryonic fibroblast (MEF) cells were transfected with G3-YSD DNA ligands and the RNA expression of *Ifnb* and *Ccl5* (an ISG) was assessed. As a control, a mock version of the ligands without Y-shaped structures was used. Due to the transient nature of the IFN response, the timing of the IFN response was first determined in a time-course (Figure 3.1A). The experiment was later repeated per triplicate 6h after treatment (Figure 3.1B). These results confirmed that mESCs are incapable of sensing cytosolic DNA, in contrast to somatic MEF cells.

To identify the source of this phenotype, the activation of the cGAS/STING pathway was challenged downstream of cGAS by transfecting 2'3'cGAMP. 2'3'cGAMP is the second messenger liberated by the DNA sensor cGAS once it binds to DNA. This dinucleotide plays the role of activating STING. A time-course was performed again, which revealed that the expression of IFN peaked in MEFs at 6h (Figure3.1C). The same experiment was repeated 8h after treatment to see *Ccl5* upregulation (Figure3.1D). These data revealed that despite producing a strong IFN response in MEF, 2'3'cGAMP was incapable of activating the cGAS/STING pathway in ESCs.

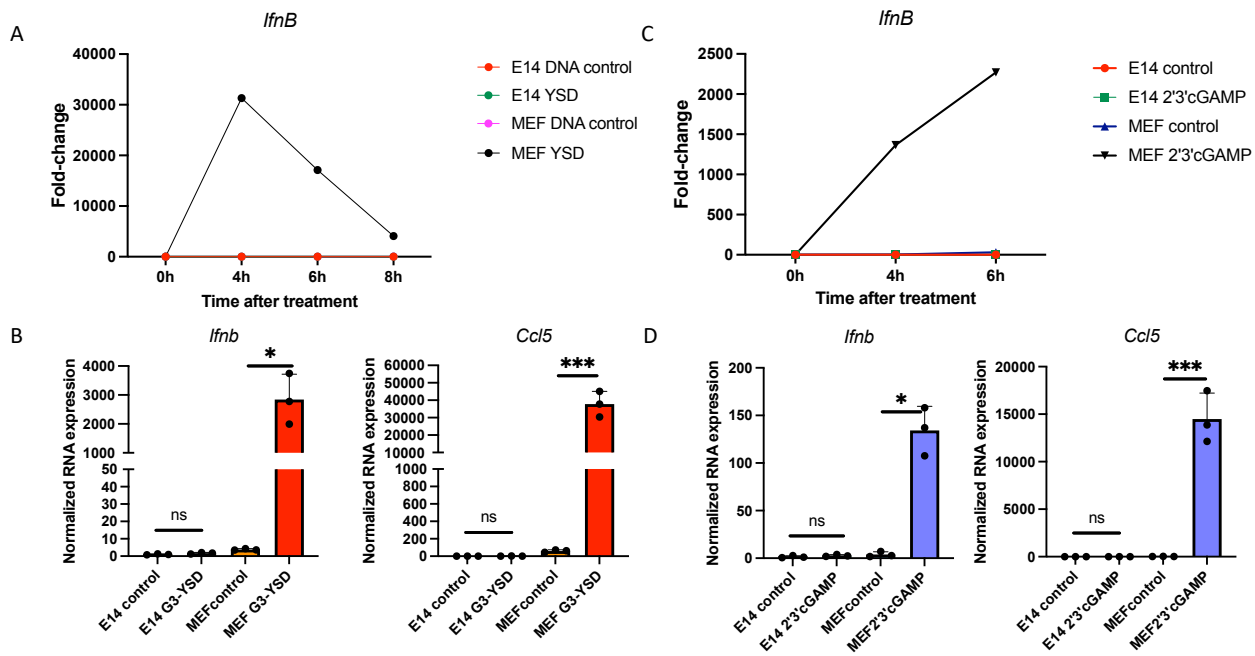


Figure 3-1: mESCs do not respond to DNA ligands and 2'3'cGAMP treatment. (A) Time-course of IFN response after DNA ligands (G3-YSD) and control DNA treatment of mESCs (E14) and MEF (N=1 independent experiments, no internal replicates). (B) Expression of *Ifnβ* and the ISG *Ccl5* 6h after treatment with DNA ligands and the corresponding controls(N=3). *Ifnb* expression t-test between MEF control and MEF G3-YSD (p-value= 0.03). Same comparison for *Ccl5* (p-value= 0.0008). (C) Time-course of IFN response 2'3'cGAMP treatment of mESCs (E14) and MEF (N=1 independent experiment, no internal replicates). (D) Expression of *Ifnb* and the ISG *Ccl5* on mESCs and MEF 8h after treatment with 2'3'cGAMP (N=3 biological replicates). Normalized mRNA expression, set to G3-YSD control E14 replicate where the expression of any gene is set to 1. *Ifnb* expression t-test between MEF control and MEF G3-YSD (p-value= 0.01). Same comparison for *Ccl5* (p-value= 0.007).Data in B, D corresponds to the mean of n=3 biological replicates and 3 independent experiments.

3.2.2. RNA analysis of cGAS/STING pathway components shows expression of all the components of the pathway.

To get an idea of the expression level of the innate immune system components in pluripotency, I used an RNA-seq dataset of mESCs undergoing differentiation at several time point. No obvious difference or detail in the expression of the assessed ISGs, components of the DNA sensing machinery and IFNs signalling genes was found (Figure 3.2).

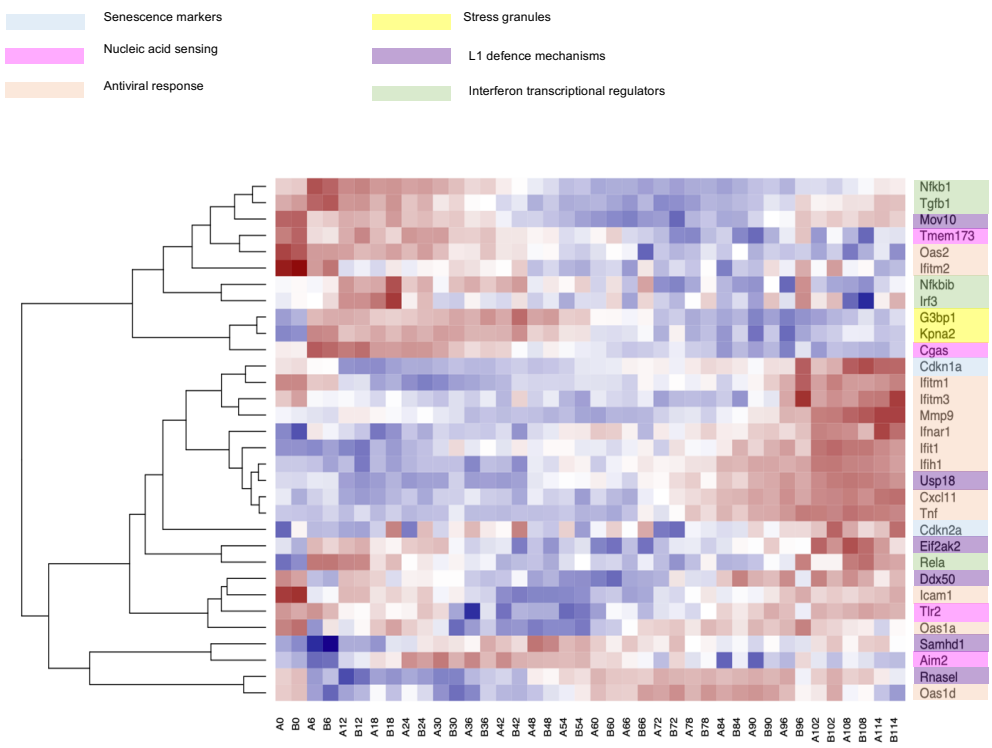


Figure 3-2 Gene expression analysis in a pluripotency dataset. The analysed genes are involved in nucleic acid sensing, antiviral response, interferon stimulated genes, senescence markers and L1 defence mechanism was created and its expression analysed (Y axis) in (A). An E14 mESCs spontaneous differentiation experiment where samples were extracted every 6 hours in two replicates (A and B), being A0 and B0 the initial conditions and A120 and B120 the samples harvested after 120h of LIF removal. STING is annotated with the name Tmem173.

The lack of responsiveness of mESCs to cytosolic DNA could be consistent with the deficiency of some of the components of the cytosolic DNA sensing pathway (cGAS/STING). However, the above transcriptomics analysis of pluripotent versus differentiating mESCs did not show any obvious cause for the insensitivity of mESCs to cytoplasmic DNA. Next, the mRNA expression of the cGAS/STING pathway components was assessed comparatively between mESCs and MEFs using qRT–PCR. Both cell types were treated with etoposide in order to simulate an innate immune response as this reagent generates DNA intermediaries in the nucleus and cytoplasm and has been shown to induce an IFN response(Dunphy et al., 2018). 24h later, RNA and protein samples were collected.

All components of the pathway showed reduced expression at the RNA level in mESCs compared with MEFs (Figure 3.3). These differences varied across components, being more similar for *Tbk1*, *Nfkb1* and *Irf3* and more marked for *cGAS*, *STING*, *Rela* and specially *Nemo*. Despite a reduction in expression, CTs values were not necessarily high (*STING* ~25, *cGAS* 30, *Irf3* 27, *Nfkb1* 28). As expected by reports from the literature, NFκB seems to be the most

inhibited arm of the pathway (Torres and Watt, 2008). Strikingly, the RNA levels of some of these genes seemed to be upregulated not only on the MEFs treated with etoposide but also on the mESCs, reaching in some cases levels comparable to non-treated MEFs. This is consistent with an IFN response occurring in both models, as some of these genes are interferon-stimulated genes (ISGs).

Dunphy et al. had previously shown that the IFN response to DNA damage follows a different pathway from the canonical response to cytosolic DNA (Dunphy et al., 2018). This pathway in mice requires the following DNA damage proteins: ATM, Parp-1, p53, Ifi16 and Traf6. However, the authors found that unlike in humans, in mice this response still requires *STING*, *Irf3*, *Tbk1* and *Nfkb1*. These results may suggest that mESCs may be capable of mounting an IFN response to DNA, only when the source of nucleic acids is DNA damage (Burleigh et al., 2020; Dunphy et al., 2018).

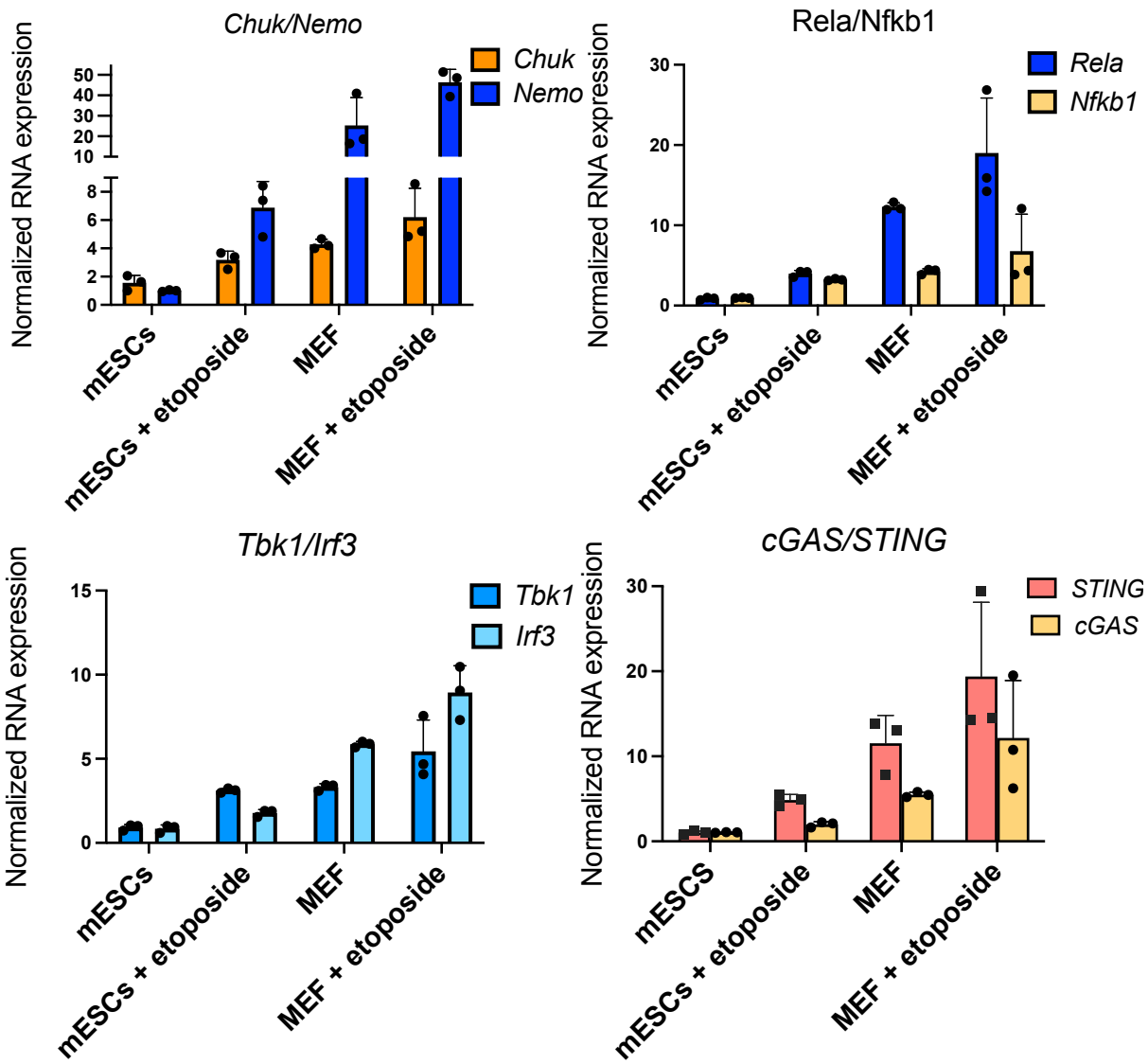


Figure 3-3: mRNA expression levels of different components and regulators of the cGAS/STING pathway. In mESCs (E14) and MEF treated and non-treated with etoposide Expression levels of *Chuk, Nemo, Rela, Nfkb1, Tbk1, Irf3, cGAS* and *STING*.). N=3 biological replicates. Normalized mRNA expression, set to control mESCs replicate =1. Data corresponds to +/- mean of 3 biological replicates.

3.2.3. WB analysis reveals no expression of cGAS and STING at protein level.

RNA expression analysis of the components of the cGAS/STING pathway did not show any conclusive result that explained the lack of response of the pathway in mESCs. The lack of expression at protein level due to mechanisms operating downstream of transcription next had to be considered.

Expression of cGAS/STING pathway components was analysed in E14 and MEF +/- etoposide protein samples. Nfkb1, Tbk1 and IRF3, the main cGAS/STING-downstream effectors, were found to be expressed in ESCs at a slightly reduced level when compared to

MEFs (Figure3.4A). The protein levels of these effectors increased in the etoposide-treated groups, consistent with an IFN response. In contrast, both cGAS and STING were not expressed at the protein level in mESCs (Figure 3.4B). Strikingly, the expression of cGAS was rescued in mESCs treated with etoposide, whereas STING was not. These results suggest that the absence of cGAS and STING expression at the protein level, but more specifically STING, may be the main cause of mESCs not responding to cytosolic DNA. The molecular mechanism of this phenomenon remains to be uncovered.

Considering that *STING* and *cGAS* are not expressed in pluripotency, the lack of functionality of this pathway was attributed to their absence. It was then assumed that rescuing their expression would restore functionality to the pathway. Rescue was attempted through different methods. Differentiation was first attempted to upregulate the expression of the components of the pathway (Witteveldt et al., 2019; D'Angelo et al., 2018b). mESCs were passively differentiated by removing LIF from the media for 6 days and collecting protein samples. Samples were quantified and analysed in immunoblots using MEFs as a positive control for *cGAS* and *STING* expression. Differentiation of mESCs could not rescue STING protein levels despite being able to upregulate Tbk1 (Figure 3.3C). Given the apparent expression of cGAS and STING at the RNA level, it was hypothesised that a post- or post-translational mechanism was responsible for the lack of protein. miRNA regulation was discarded as *Dicer*-/- cells are still insensitive to cytosolic DNA (Witteveldt et al., 2019) Therefore, the possibility that cGAS and STING were being degraded in mESCs was tested by treating the cells with proteasomal inhibitor, mg132 and lysosomal inhibitor, chloroquine at different time points and dosages. Nevertheless, the expression of *cGAS* and *STING* was not rescued upon treatment with these drugs (Figure3.4D).

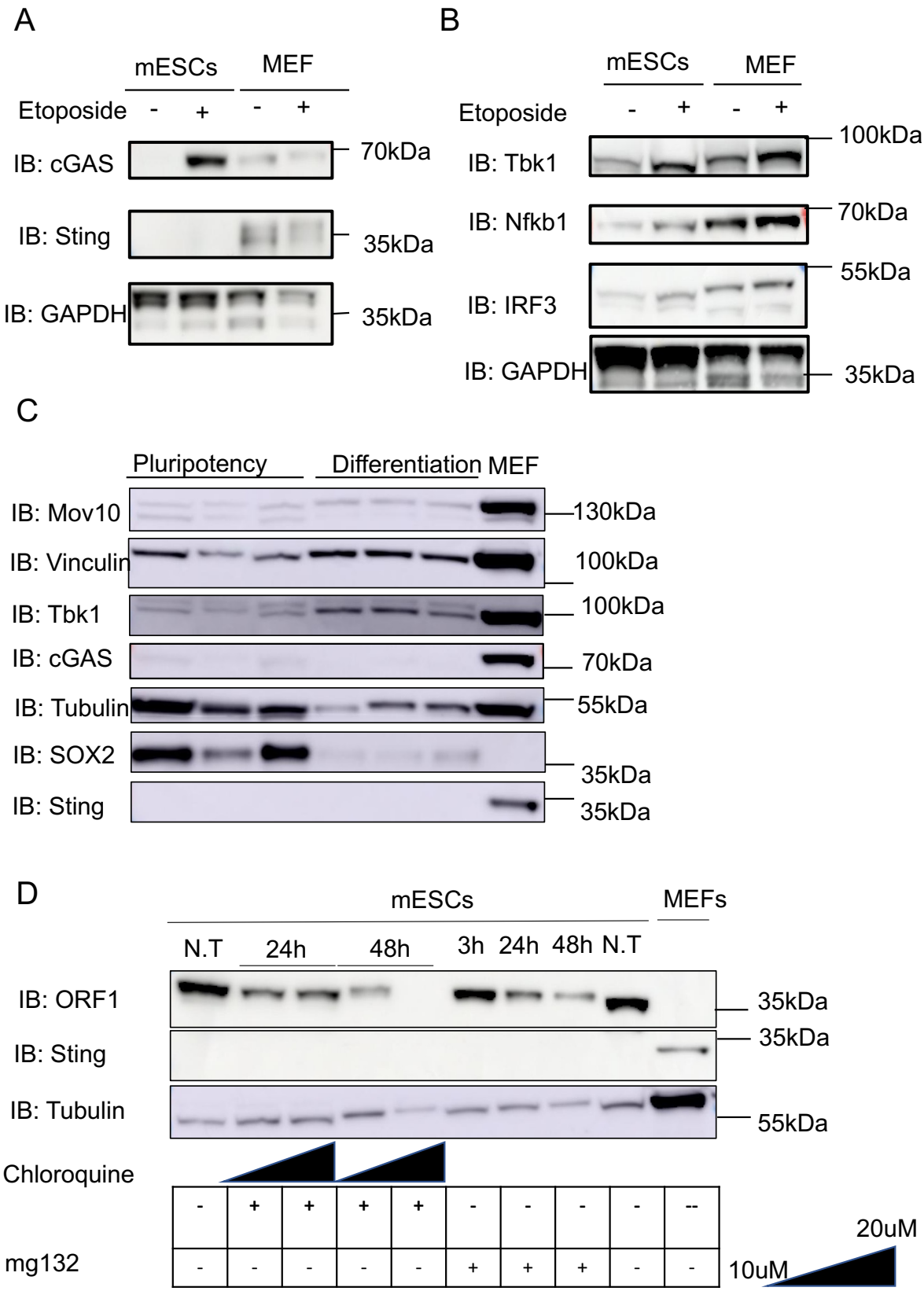


Figure 3-4: Protein expression of cGAS/STING components on mESCs and MEF. (A) Expression of Tbk1, Nfkb1 and IRF3 in E14 and MEF samples treated and non-treated with etoposide. This experiment corresponds to 2 technical replicates. (B) Expression of cGAS and STING proteins in mESCs and MEF samples treated and non-treated with etoposide (10μM). This experiment corresponds to two technical replicates. (C) Expression of innate immune system components in pluripotent mESCs and mESCs after 6 days of passive differentiation,

using MEFs as a positive control. This experiment corresponds to 3 independent experiments (D) Expression of STING and ORF1 upon treatment with different doses of Chloroquine and mg132 at different time-points, n=1 experiment. Chloroquine was used on doses of 10μM and 20μM. Mg132 3h treatment used a dose of 10μM 24 and 48h used 100nM.
MW: LINE1 ORF1p 40kDa, Sox2: 35kDa, STING: 35kDa, MOV10: 114kDa, cGAS: 62kDa, αTubulin: 50kDa, Tbk1: 84kDa, IRF3:45-55kDa, Nfkb1: 65kDa, GAPDH: 33kDa

3.2.4. Human STING carrying gain of function mutations cannot rescue missing STING signalling in mESCS.

I next tested whether the absence of STING protein causes the unresponsiveness of mESCs to cytosolic DNA by overexpressing STING in mESCs. Human *STING* (hSTING) V155M gain of function mutation was the chosen option as it was the only available *STING* expression plasmid containing the gain-of-function (GOF) mutation. This mutation is responsible for SAVI interferopathy and activates STING constantly, leading to a chronic aseptic inflammation (Bouis, 2019; Keskitalo et al., 2019; Y. Liu et al., 2014; Motwani et al., 2019). As a negative control, the hSTING-V155M variant also containing a "HAQ" mutation known to make STING non-functional, was used (Patel and Jin, 2019).

The *hSTING* cDNAs were sub-cloned from previous published vectors into lentiviral puromycin-resistant plasmids under the control of a doxycycline (Dox)-inducible promoter (Figure3.5A) (Cerboni et al., 2017; Jeremiah et al., 2014). Functionality of the construct was then tested by transfecting it into HEK 293T cells (Figure 3.5B). Stable inducible STING-expressing mESCs were generated and STING protein expression was validated (Figure 3.5C). Analysis of *Ifnb* and ISGs expression in STING GOF expressing mESCs showed how it was not capable of rescuing *Ifnb* production in pluripotency (Figure 3.5D). I found that MEFs could not generate long-term stable cell lines as from passage 3 onwards they were found to constituently upregulate IFN expression.

Activation of the pathway was also assessed through analysis of Tbk1 phosphorylation in hSTING GOF mESCs with the corresponding mutation expressing cell lines and MEF transfected with G3-YSD ligands as positive control. These results suggested that Tbk1 is not activated in STING V155M expressing cell lines, implying a defect in the pathway upstream of Tbk1 phosphorylation and consistent with the lack of *Ifnb* transcription in these cell lines (Figure3.5E).

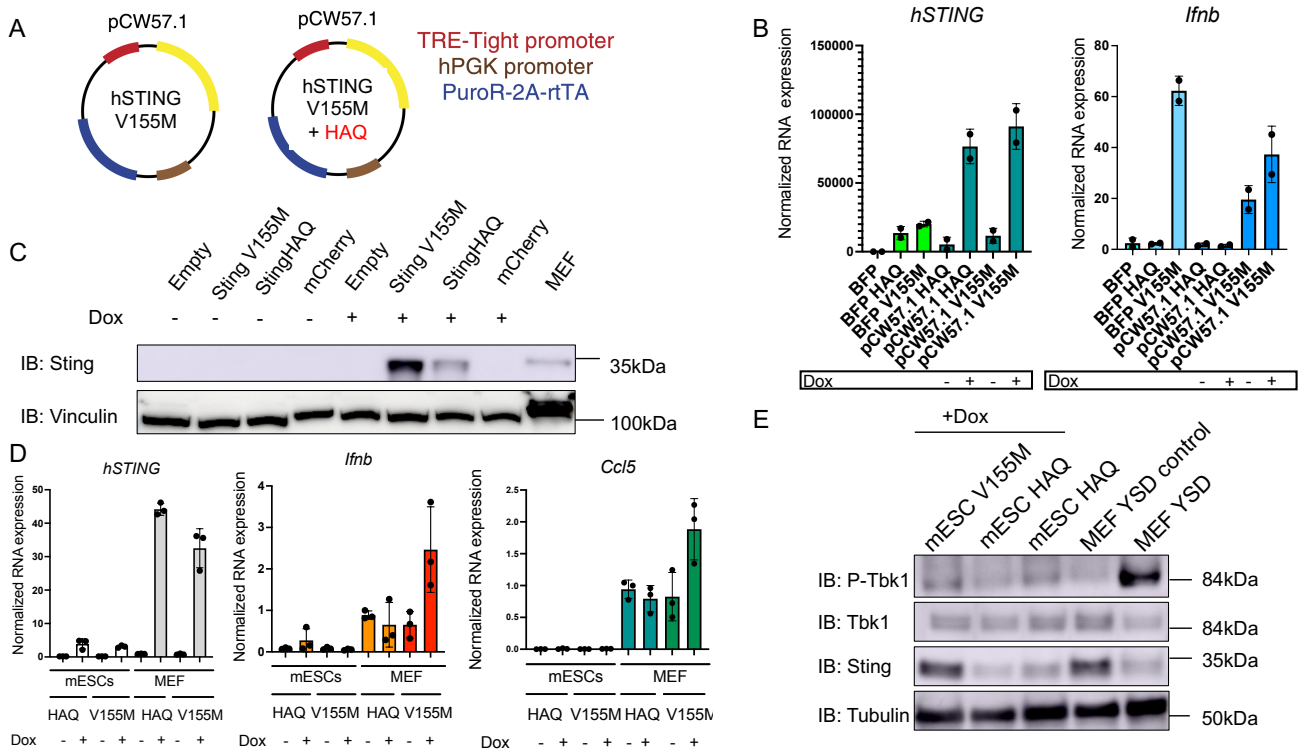


Figure 3-5: STING gain of function mutation fails to activate an innate immune response on mESCs. (A) Schematics of the *hSTING* V155M-HAQ and V155M expression vectors. (B) Quality control of plasmid functionality: Analysis of *hSTING* and *Ifnb* RNA expression in 293Ts transfected with the original *hSTING* expression vectors purchased from Addgene and the derived pCW57.1 STING expression vectors (C) WB analysis of STING protein expression in stable cell lines infected with lentivirus carrying the PCW57.1 STING expression constructs 24h post Dox addition. (D) *Ifn*b, *hSTING* and *Ccl5* response 12h upon Dox treatment in mESCs and MEF V155M and HAQ cell lines. N= 3 technical replicates. (E) Analysis of cGAS/STING pathway activation through phospho Tbk1 blotting in whole protein extracts of mESCs expressing STING gain and loss of function mutation and MEF transfected with G3-YSD and the corresponding controls. MW: STING: 35kDa Tbk1: 84kDa P-Tbk1: 84kDa αTubulin: 50kDa .

3.2.5. Mouse STING carrying gain of function mutations does not rescue STING signalling. IRF3-5D gain of function mutation does.

Unappreciated differences between human and mouse *STING* have previously led to misleading inferences of the effectiveness of certain drugs and STING biology (Shih et al., 2018). Hence, to confirm the inability of *STING* GOF forms to generate IFN signalling in ESCs, the same approach was repeated using mouse *STING* with its equivalent gain of function and LOF mutations: V154M and HAQ. In this iteration, a gain of function variant of *Irf3*, IRF3-5D was also generated as a potential positive control (Figure 6A) (Lin et al., 1999; J.-T. Wang et al., 2014). IRF3-5D is generated by replacing Serine/Threonin residues from the IRF3 transactivation domain by the phosphomimetic aspartic acid, which makes IRF3-5D constitutively active (Figure 3.6A). I also designed all the constructs with BFP-T2A followed by the protein of interest to validate expression and allow BFP+ cells to be sorted (Figure

3.6B). As a sorting control, I made a BFP only construct in the same vector than the transgenes. Constructs were cloned into the same lentiviral-puromycin resistant plasmid under the control of a doxycycline-inducible promoter (Figure 3.6B).

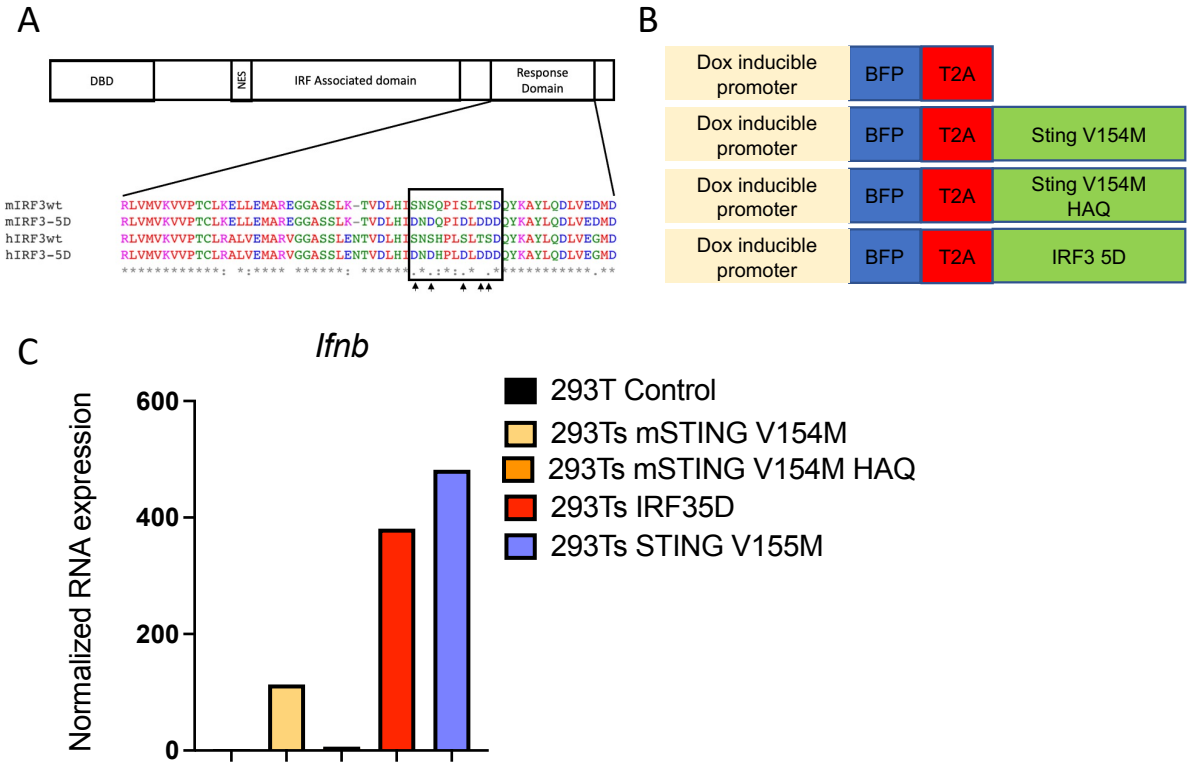


Figure 3-6: Mouse STING and IRF3 GOF constructs are functional (A) Schematic of the IRF3-5D gain of function (phosphomimetic variant). (B) Schematic of the lentiviral expressing vectors. (C) *mSTING*, *Irf3* and *hSTING* GOF constructs were transfected to 293T cells in the presence of doxycycline. Expression of *Ifnb* was checked 48h replicates. N=1 corresponds to 1 biological replicate.

The functionality of the constructs was confirmed by transfecting them into 293T cells (Figure 3.6C). Next, stable Dox-inducible cell lines for each construct were established by lentiviral-mediated transduction. Upon doxycycline stimulation, BFP was imperceptible in cultured cells and RT-qPCR analysis of stimulated cells carrying the construct showed only a low transgene upregulation. In addition, IRF3-5D OE showed little or no *Ccl5*/*Ifnb* upregulation (data not shown).

These results suggested that a higher construct expression was needed. Therefore, constructs were transiently transfected into mESCs and cells stimulated with doxycycline. 24h later, BFP was clearly visible in ~15% of cultured cells. 48h later, cells were harvested and BFP positive cells sorted. I assessed the RNA expression of the transgenic construct, as well as *Ifnb* and *ISG* transcripts by RT qPCR. The results confirmed a substantial increase in the construct expression and the inability of STING GOF to activate a functional immune response

(Figure 3.7). In contrast, *Irf3* GOF carrying samples induced a substantial upregulation of *Ifnb* and other ISGs directly regulated by IRF3 (Figure 3.7).

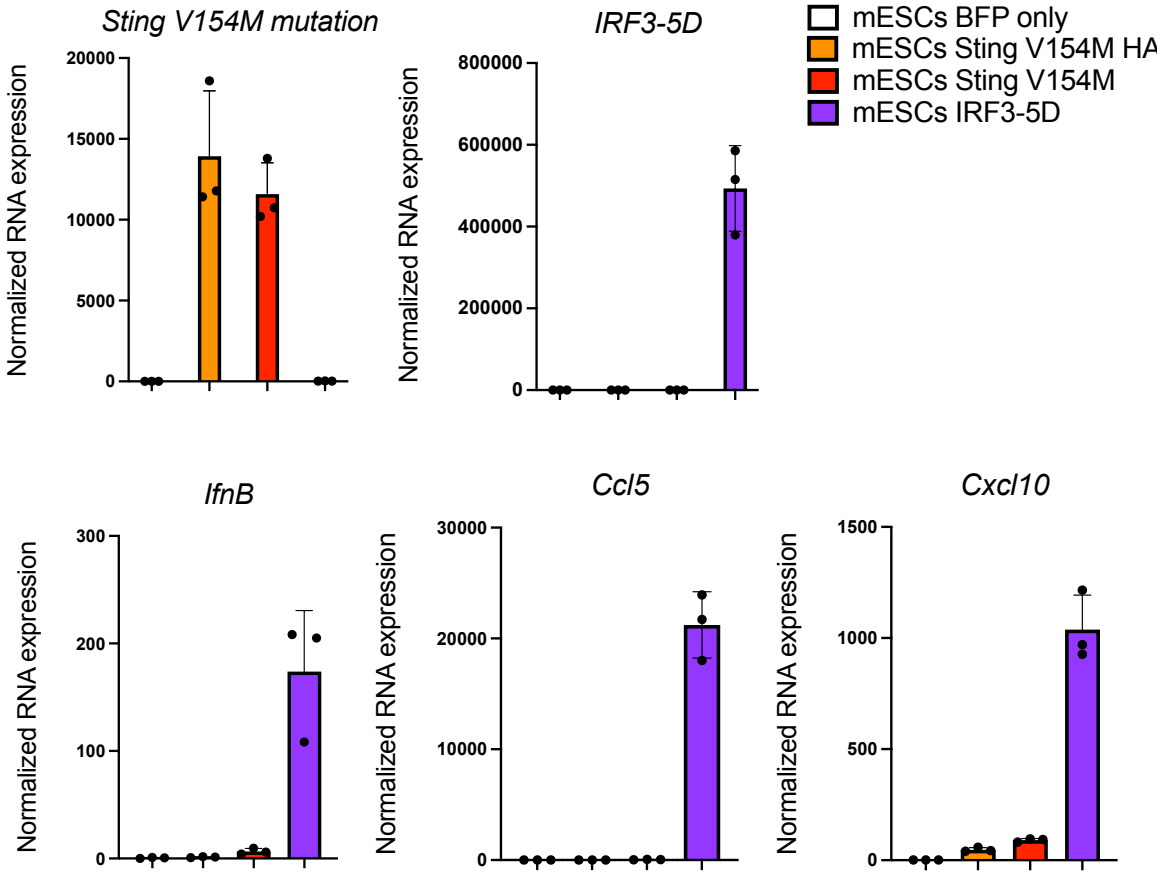


Figure 3-7: mSTING GOF variant cannot rescue IFN production in mESCs, but IRF3 GOF can. STING and IRF3-5D transgenic construct expression in mESCs 48h post-transfection. Expression of *Ifnb*, *Ccl5* and *Cxcl10* in mESCs 48h post doxycycline treatment. N=3 biological replicates.

These results suggest an additional level of cGAS/STING pathway dysfunction in pluripotency downstream of STING expression and upstream of IRF3 signalling. Additionally, I have demonstrated the capacity of mESCs to trigger an interferon response when exposed to constitutively active IRF3.

3.2.6. mSTING ER:Golgi dynamics in pluripotency

The inability of *mSTING* V154M OE to rescue IFN signalling in mESCs, in contrast to the acute response caused by IRF3-5D OE, prompted me to examine STING protein dynamics in the ER:Golgi interface in ESCs. It has previously been shown that post-translational modifications are needed for correct STING localisation and function (Mukai et al., 2016). Once activated in the ER, STING translocates into the Golgi and undergoes a set of post-translational modifications including palmitoylation, necessary for its functionality (Mukai et al., 2016).

STING V154M/HAQ location in mESCs was established by using PDI and TGN38 as ER and Golgi markers, respectively. As before, I transfected the *STING* and *Irf3* expressing constructs in Dox, followed by sorting of BFP positive cells 36h after transfection. BFP-only transfected controls showed no *STING* expression. *STING* V154M HAQ-expressing cells showed STING expansion through the cytoplasm and around the nucleus, overlapping with PDI (Figure 3.8A). This is consistent with a non-functional form of STING located on the ER or the ERGIC. In contrast, *mSTING* V154M expressing cells showed STING punctuate expression that overlapped with TGN38, consistent with Golgi localization of the activated form of STING (Figure 3.8A). This punctuate localization has been attributed by other to the activated form of STING (Mukai et al., 2016). Additional staining with an anti phospho-STING antibody, which marks the activated phosphorylated form of STING, revealed interestingly that both mSTING species GOF and loss of function, were phosphorylated while still showing the same localization and distribution pattern (Figure 3.8B). This result is inconsistent with the current literature on STING phosphorylation being a process catalysed by Tbk1 in the TGN (Zhang et al., 2019). Thus, STING seems to activate on the ER and shuttle to the Golgi, but without being yet functional.

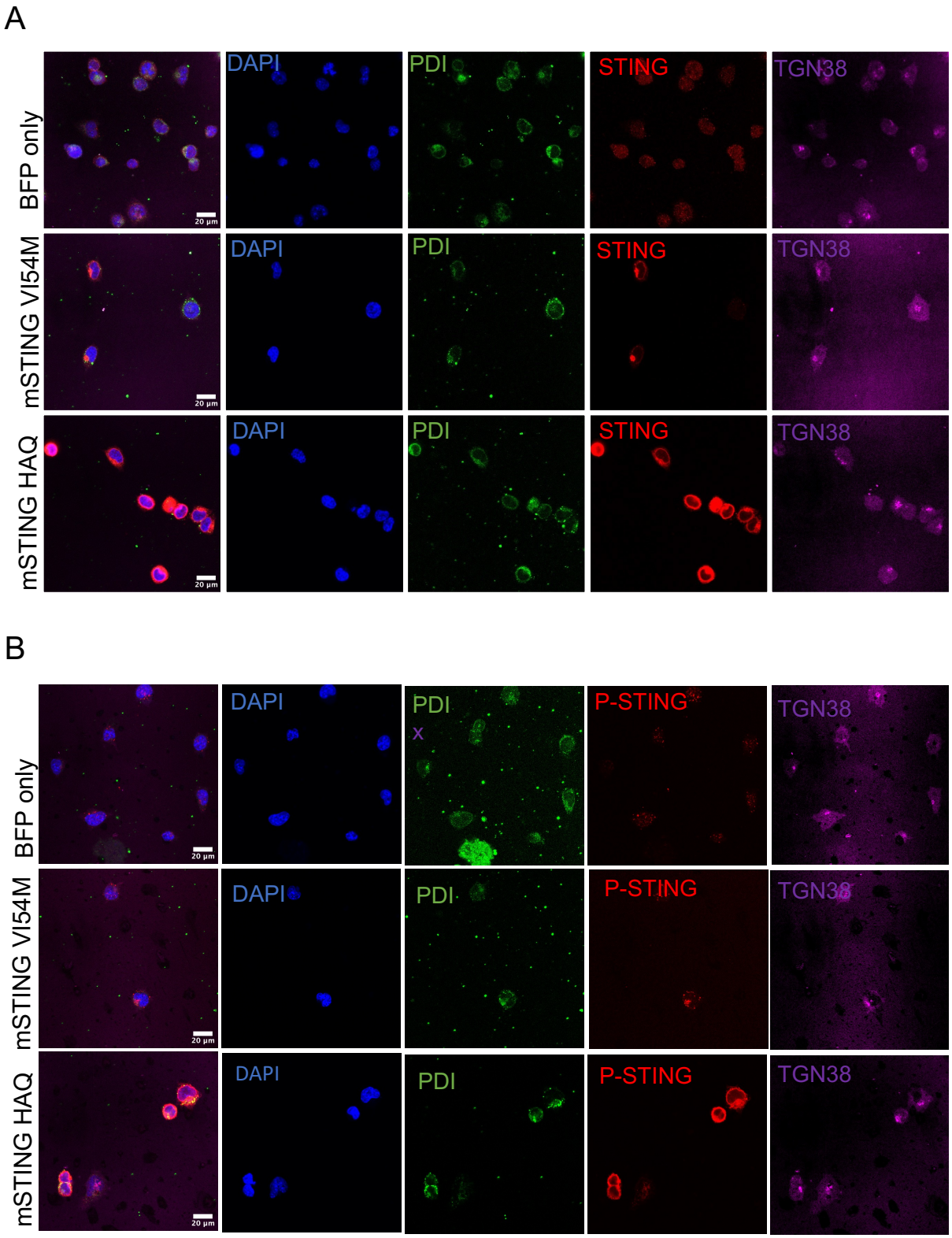


Figure 3-8: STING GOF dynamics at the ER:Golgi interface in pluripotency. (A) mESCs immunofluorescence with PDI (ER marker) in green, STING in red and TGN38 (Golgi marker) in magenta. Cells were previously transfected with BFP-T2A-STING V154M and HAQ constructs and a BFP only control. Construct expression was induced by Dox for 36h, upon which cells were plated in the coverslip and fixed. (B) In the previous set up, mESCs immunofluorescence with PDI (ER marker) in green, phospho-STING in red and TGN38 (Golgi marker) in magenta.

3.3. Discussion

In this chapter, I have demonstrated the inability of mESCs to respond to DNA and 2'3'cGAMP (Figure 3.1). This is consistent with the lack of protein expression of cGAS and STING (Figure 3.4A). The lack of protein expression of cGAS and STING contrasts with apparent expression of both genes at mRNA level, although notably reduced with the rest of the pathway, when compared to somatic cells (Figure 3.3). This reduction of cGAS/STING pathway components RNA expression is of different calibre depending on the factor. Data also shows that downstream components of the pathway such as *Tbk1*, *Irf3* and *p65* are expressed at the protein level in pluripotency, although at a lower level than their somatic counterparts, consistently with their reduced RNA expression (Figure 3.4B). The mechanism preventing cGAS and STING protein expression is still unknown. Interestingly, DNA damage seems to rescue cGAS protein expression (Figure 3.4A).

The expression of *hSTING* and *mSTING* gain of function mutations in mESCs was uncapable of rescuing the IFN signalling pathway in ESCs, with a seeming defect happening upstream Tbk1 activation (Figure 3.5D, E & 3.7). Meanwhile the IRF3-5D variant OE was capable of activating genes from the antiviral response and the production of *Ifnb* (Figure 3.7). In addition, *mSTING* GOF mutation was capable of translocating to the Golgi (Figure 3.8A). Therefore, these results suggested that an additional layer of negative regulation of the cGAS/STING pathway operates in pluripotency, between activation of STING in the ER and posterior translocation to the Golgi and activation of Tbk1 in the TGN, and preventing the downstream production of IFN. Moreover, STING phosphorylation was detected in STING V154M HAQ located at the ER and STING V154M located at the Golgi (Figure 3.8B). This challenges the current knowledge on STING activating phosphorylation, as it is in principle a modification carried out by Tbk1 at the TGN (Zhang et al., 2019). It is known that Tbk1 is recruited as early as the ERGIC, therefore it could be that the modification is occurring earlier, or that we are seeing an unspecific activity of the antibody used. Further experiments will need to determine the exact location of STING protein in the Golgi and the status of essential PTMs.

I am aware that this is an artificial system as we are over-expressing a construct, without considering the physiological levels of IRF3 and STING. However, it is of particular interest that IRF3-5D by itself seems to be capable, at least partially, of activating the antiviral response independently of IFN signalling in pluripotency. This sets up the scenario in which the detrimental effects of cGAS/STING activation are not fully dependent on the autocrine/ paracrine activation of the IFN signalling cascade, as activation of IRF3 by the pathway would potentially be enough to mediate detrimental effects.

Chapter 4: Sensitivity of mESCs and embryos to IFN signalling

4.1. Introduction

Up to this point, I have focused on the capacity of mESCs to produce interferon through their DNA sensing pathway, but not on the actual response to interferon that involves the upregulation of the antiviral transcriptional programme. This response is paramount to the assumption that restoring interferon production in response to LINE1-derived nucleic acids would result in the chronic activation of the antiviral transcription programme. The persistent activation of this programme could cause the detrimental effects that would compromise the viability of the cell and the embryo in an in vivo context.

It has been hypothesised for years that pluripotency may be incompatible with a functional antiviral response (Eggenberger et al., 2019; Torres and Watt, 2008). Several experiments have been performed exposing ESCs and their trophoblast counterparts to PAMPs and DAMPs, sometimes with contradictory results. While most of these studies report attenuated IFN signalling in pluripotency (Hertzog et al., 1994; Matfsson et al., 1992; D'Angelo et al., 2018b), some also report upregulation of particular ISGs (Fendereski et al., 2022; R. Wang et al., 2014b). Studies vary in the disparity of the ISGs used as markers, types of ISGs used, type of stimuli (LPS, polyI:C, virus…), IFN exposure, time points and model used. Therefore, it remained unclear for the interest of this project whether – and to what extent – mESCs could hypothetically respond to an IFNβ stimulus produced downstream of TE DNA sensing and IRF3 activation and how these could affect the functionality of the cell and the embryo. This was particularly interesting to explore given that the components of the IFN signalling pathway are expressed in pluripotency (de Weerd and Nguyen, 2012; Hertzog et al., 1994; R. Wang et al., 2014b). Moreover, the specific phenotypic and molecular consequences of the activation of the antiviral programme in pluripotent cells and embryos remain unclear.

For the experimental design of this study, it was also important to examine separately the response of "toxic" ISGs (such as *Tnf* and *Il1b*) versus the subset of ISGs previously found to be constitutively expressed in pluripotent cells that confer viral resistance (Wu et al., 2018). These genes, in contrast to the rest of the ISGs, have an apparent lower detrimental impact on the cell that seems to allow the embryo to tolerate their expression, like *Ifitm3* and *Eif2ak2* (PKR). Therefore, in the context of this work, I wanted to examine the expression of these "intrinsically expressed ISGs" as well as the canonical ISGs.

First, I wanted to address mESCs and embryos capacity to react to IFNβ. They were treated with IFNβ at different time points and later evaluated for the expression of ISGs downstream of IFNβ. Moreover, to elucidate if IFN signalling was active in mESCs, I checked for the activation of the JAK/STAT pathway by looking for phoshpho-STAT1 upon IFNβ treatment, the activated form of STAT1 (the transcription factor activator of the antiviral programme) (Platanias, 2005).

Then I tested the hypothesis that activated IRF3 can directly activate targets of the antiviral programme in a cell intrinsic fashion (R. Lin et al., 1998; Lin et al., 1999; Solis et al., 2006). IRF3 direct activation of the antiviral programme, may earn particular relevance in pluripotency (Solis et al., 2006). Hence, I tested the hypothesis, as it would be one of the direct consequences of STING activation by LINE1 ssDNA in mESCs. To investigate this, IRF3-5D expression plasmids were transfected into mESCs and I performed RNA-seq on the selected samples. This ties with the previous chapter results showing that IRF3-5D can indeed activate *Ifnb*, *Ccl5* and *Cxcl10.* With the RNA-seq I was able to have a wider scope and understanding of the transcriptional consequences for mESCs of IRF3 activation.

Overall, I found that despite IFN response being effectively blunted in mESCs, IRF3 activation has strong negative consequences for cell viability by itself. Thus, suggesting that TE sensing, if activated, would have profound detrimental consequences for development.

4.2. Results

4.2.1. mESCs are capable of upregulating intrinsically expressed ISGs but not canonical ISGs

To learn the appropriate window of time for IFNβ treatment, mESCs and MEFs were first cultured with 1000IU of IFNβ at different time points: 4h, 8h, 24h and 48h. For each time point, the RNA expression of different ISGs was evaluated. Some of these genes: *Ifitm1*, *Ifitm3*, *Oas1a* and *Eif2ak2*, were part of the pool of intrinsically expressed ISGs described by Wu et al (Wu et al., 2018). The others: *Il6*, *Ccl5*, *Tnf* and *cGAS* are “canonical” ISGs.

Analysis of MEFs treated with IFNβ showed that the antiviral response peaks at 8h (Figure 4.1A). At this time point, mESCs did not upregulate the canonical ISGs upregulated in MEFs. However, some intrinsically expressed ISGs were significantly yet mildly upregulated in

mESCs, indicating a low sensitivity to IFN treatment. Interestingly, at 24h and 48h, while the canonical ISGs in MEF are already turned down, the upregulation of intrinsic ISGs upon IFNβ treatment seems to persist in MEFs and mESCs (Figure 4.1B). In fact, they seem to peak at 24h. Altogether, these results suggest that mESCs can respond to exogenous IFNβ stimuli by upregulating a specific set of ISGs that are already expressed. Moreover, the window of expression of these genes seems to be longer than in the canonical genes.

Next, I wanted to assess the activation of the IFN signalling pathway by checking phospho STAT1 in whole cell lysates of samples treated with IFNβ. Phospho-STAT1 is the activated form of STAT1 (Platanias, 2005), the downstream transcription factor of the IFN type I sensing pathway. Considering the transcriptional response peaks at 8h, I decided to collect samples at an earlier timepoint to observe STAT1 activation. Whole cell lysates were collected 4h post IFNβ treatment. The results show the activation of pSTAT1 in MEF and mESCs. The detection of pSTAT1 implies that IFN signalling seems to be active in mESCs through the JAK/STAT pathway (Figure 4.1C). This may explain why a subset of ISGs can be mildly activated in mESCs.

Finally, culture of mouse pre-implantation embryos in IFN was performed to assess its consequences for embryo development. This did not seem to impact the viability or the development potential of the embryos *in vitro* (Figure 4.1D).

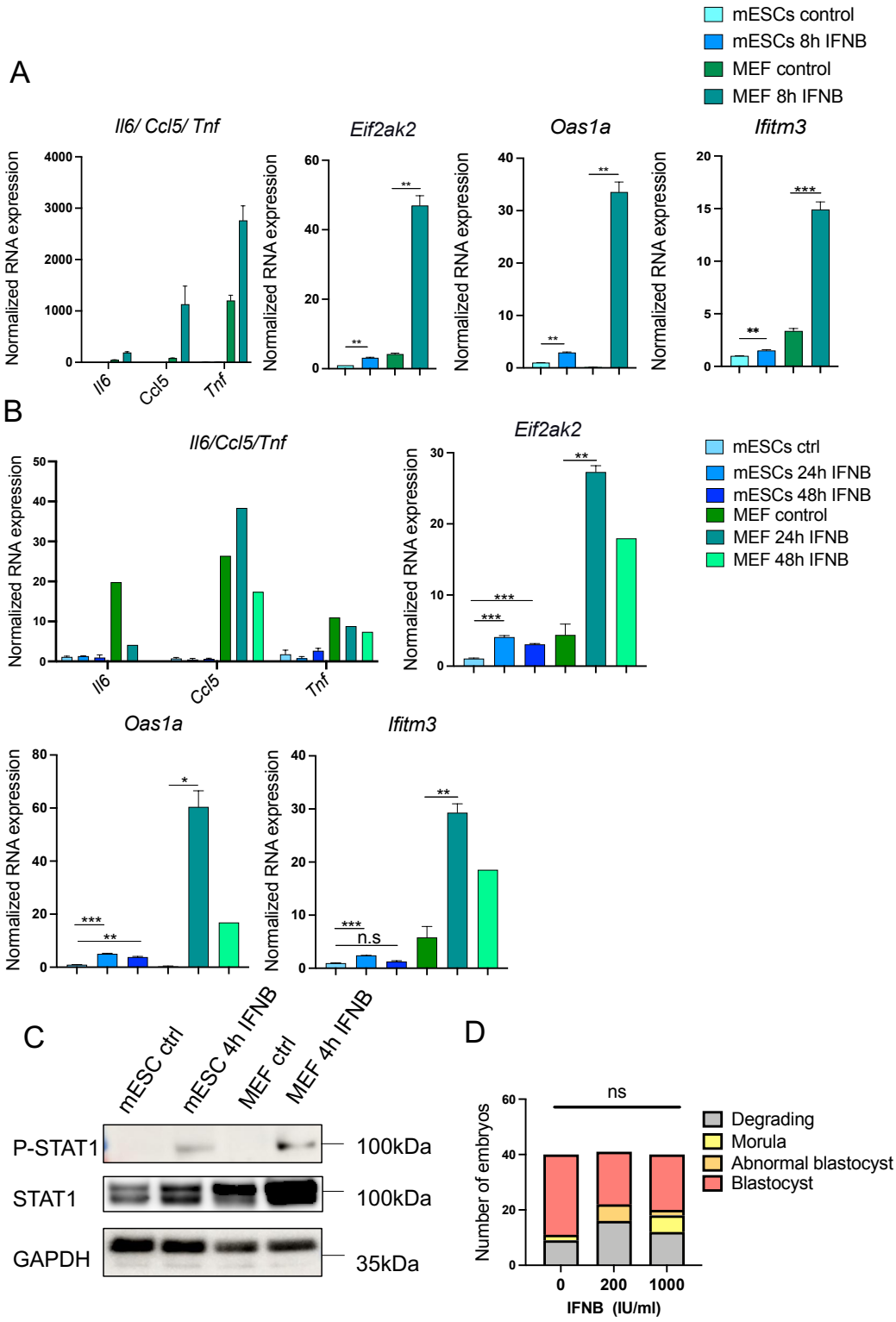


Figure 4-1: mESCs cannot articulate a functional antiviral response. (A) RNA expression analysis of several ISGs 24 and 48h post-IFN treatment. (B) RNA expression analysis of several ISGs 8h post IFN treatment. (C) Protein expression of STAT1 and its phosphorylated version in mESCs and MEF samples treated for 4h with IFNβ. This experiment is representative of to two independent experiments. MW: P-STAT1: 84, 91kDa, STAT1: 84,91kDa, GAPDH: 33kDa. Data for A and B are the mean +/- of 3 biological replicates. (D) Viability analysis of mouse pre-implantation embryos treated development potential upon treatment with different concentrations of IFNβ. Treatment was performed at E2.5.

4.2.2. mIRF3-5D OE in mESCs triggers an antiviral response

As previously reported, IRF3-5D overexpression (OE) in mESCs rescued the production of *Ifnb* mRNA and directly upregulated a subset of ISGs (see chapter 3). The upregulation of IRF3 targets could be particularly relevant in the context of pluripotency, where the downstream effects of TE DNA sensing are unknown. Therefore, to understand the full scope of this upregulation, I performed RNA-seq on sorted BFP+ mESCs transfected with plasmids bearing BFP-only, BFP-mSTING V154M HAQ, BFP-mSTING V154M and BFP-mIRF3-5D (Figure 4.2A). I also included samples from *mSTING* transfected cells to uncover whether there are any previously missed effects from the mSTING V154M gain of function mutant.

First, I examined whether ESCs transfected with the different constructs were transcriptionally distinct from one another, depicted in a PCA plot. This shows the IRF3-5D group is clearly different from other samples (Figure 4.2B). The PCA also suggested that some variation was introduced by STING OE; however, this was true for both the V154M and HAQ constructs. This could suggest that STING expression itself induces gene expression changes regardless of transcriptional activity. Differential expression analysis was next performed to identify statistically altered genes in each condition. A total of 1216 genes are differentially expressed in mIRF3-5D samples when compared to the wild-type (BFP only) considering a FDR of 5% and a log fold change >2 (Figure4.2C). The number of DEGs is 82 for mSTING V154M and 24 for mSTING V154M HAQ.

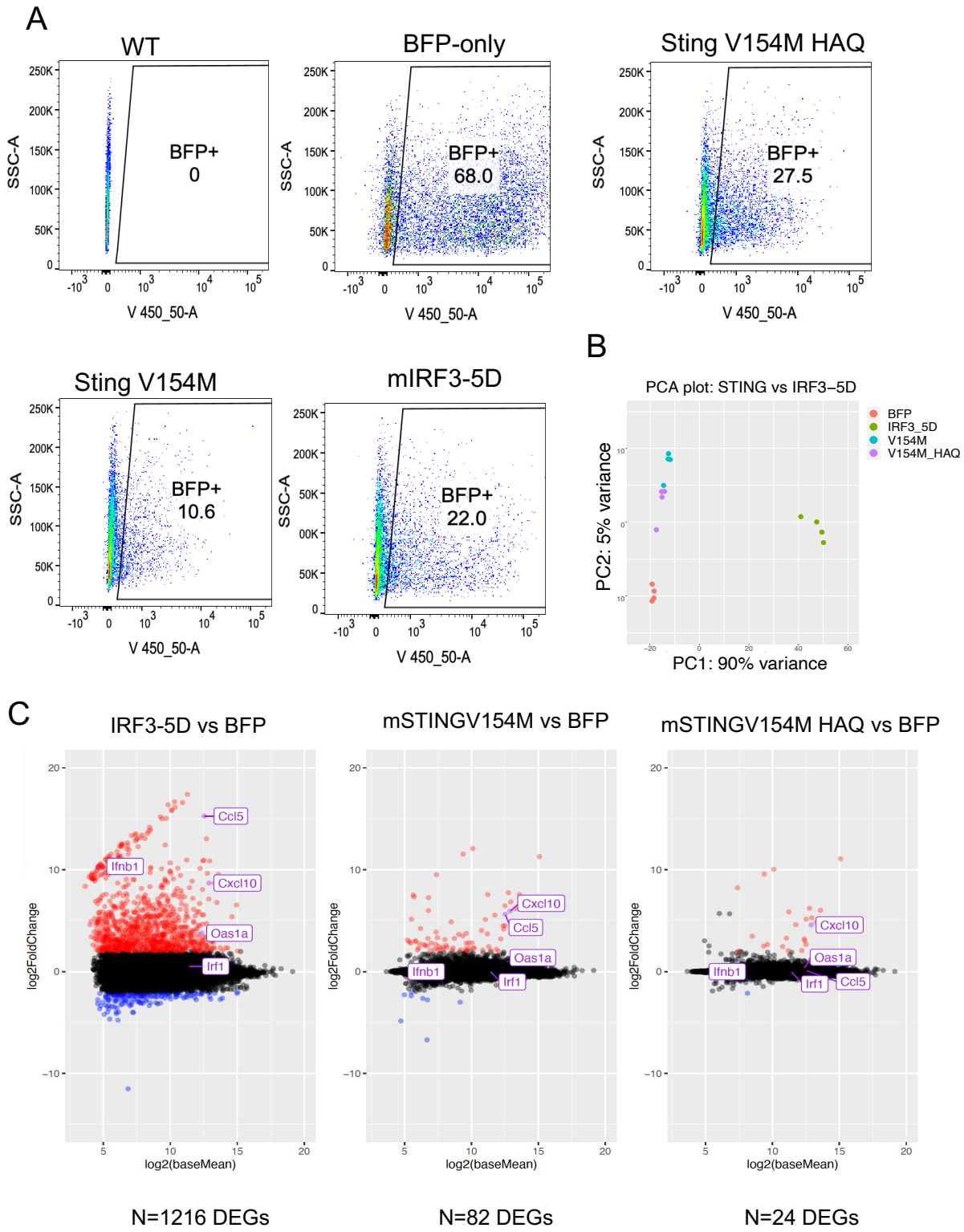


Figure 4-2: IRF3-5D OE has broad transcriptional consequences in mESCs. (A) mESCs transfected with Dox inducible plasmids carrying BFP + a protein of interest, were sorted 48h post transfection and Dox treatment. (B) PCA plot of sample similarity. (C) DEGs of IRF-5D, mSTING V154M HAQ and mSTING V154M samples when compared to BFP samples. Log2 fold-change >2 and FDR<0.05.

Next, gene ontology (GO) was performed in order to understand broadly the functional transcriptional consequences of activating the IRF3 in mESCs. Reactome and gene ontology analysis of mIRF3-5D gene expression show enrichment in terms mainly related to the innate immune response and defence against pathogens. Other relevant terms are interferon and cytokine signalling and T cell chemotaxis (Figure 4.3A). A general look at the top 1000 DEGs also shows mainly genes related to the antiviral programme: cytokines, chemokines, complement system components and other ISGs (data not shown). Overall, these data suggest that OE of IRF3-5D can trigger a functional antiviral response in mESCs independent of autocrine and paracrine downstream IFN signalling. As this response seems to rely solely on IRF3 activation, it is possible to speculate that restoration of STING signalling in mESCs that would allow for IRF3 phosphorylation could thus lead to a LINE1-triggered antiviral response in mESCs. Regarding DEGs induced by STING GOF OE, most of them seem to overlap with those upregulated by IRF3 (Figure 4.3 B). Nevertheless, IRF3-5D shows a substantially higher expression of the overlapping DEGs than STING V154M, most of them involved in the innate immune response (Figure 4.3C). It is therefore possible that overwhelmingly increasing activated STING stoichiometry in the cell, can mildly stimulate the pathway.

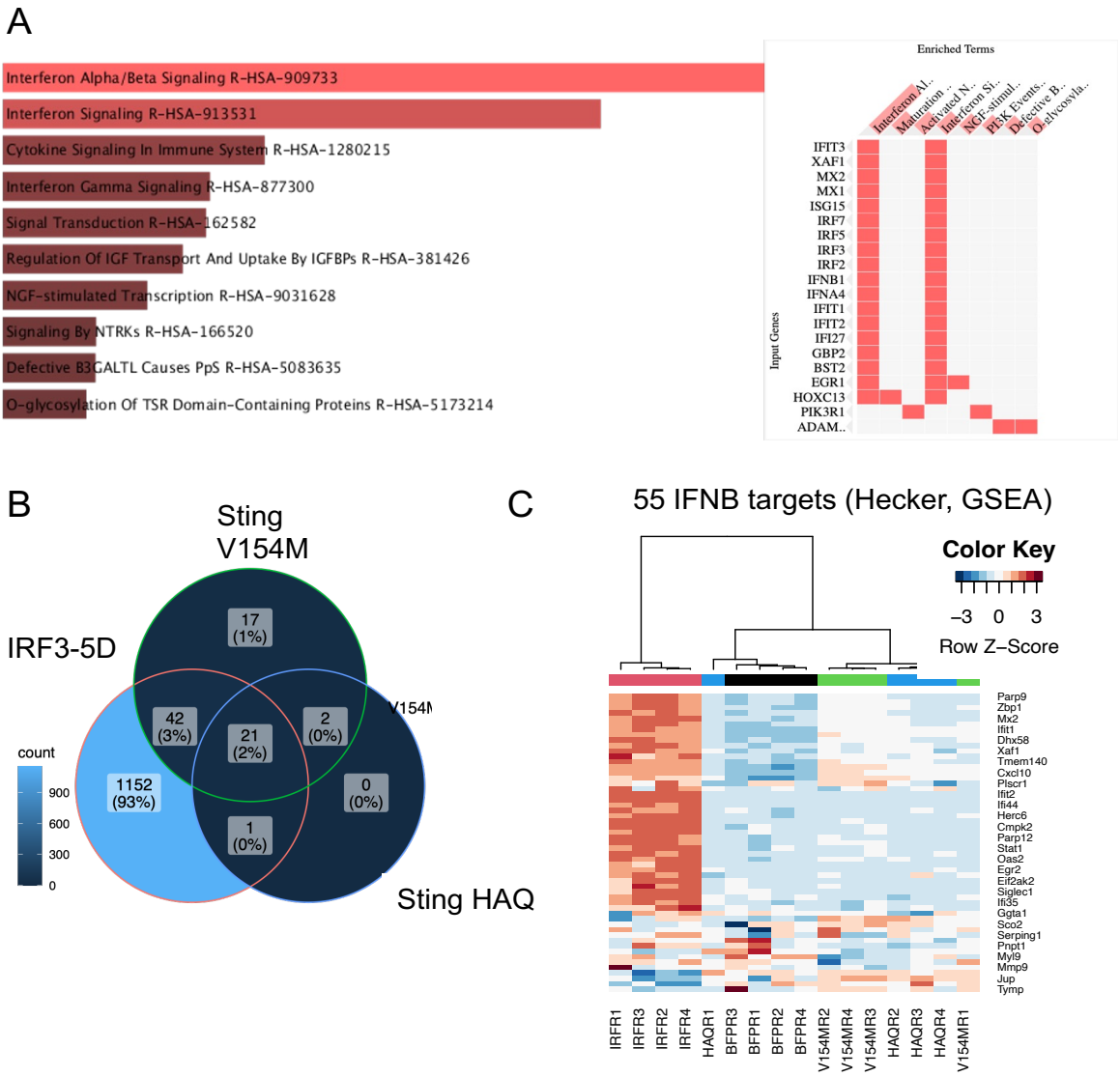


Figure 4-3: IRF3 activation in mESCs trigger the antiviral transcriptional programme.(A) Reactome 2022 analysis of enrichment in IRF3-5D samples. Log2 fold-change >2 and FDR<0.05. (B) Venn diagram depicting common DEGs between IRF-5D, mSTING V154M HAQ and mSTING V154M samples. (D) Heat-map depicting expression analysis comparison of the Hecker GSEA IFNβ target group among sample groups. A lot of these genes belong to the STING V154M OE DEGs.

Interestingly, I noticed that some of the downregulated genes were pluripotency genes. Thus, a set of pluripotency-related genes was plotted, revealing that there was an overall decrease in the expression of pluripotency-related genes upon IRF3-5D expression. Among them *Oct4* and *Nanog* (Figure 4.A).

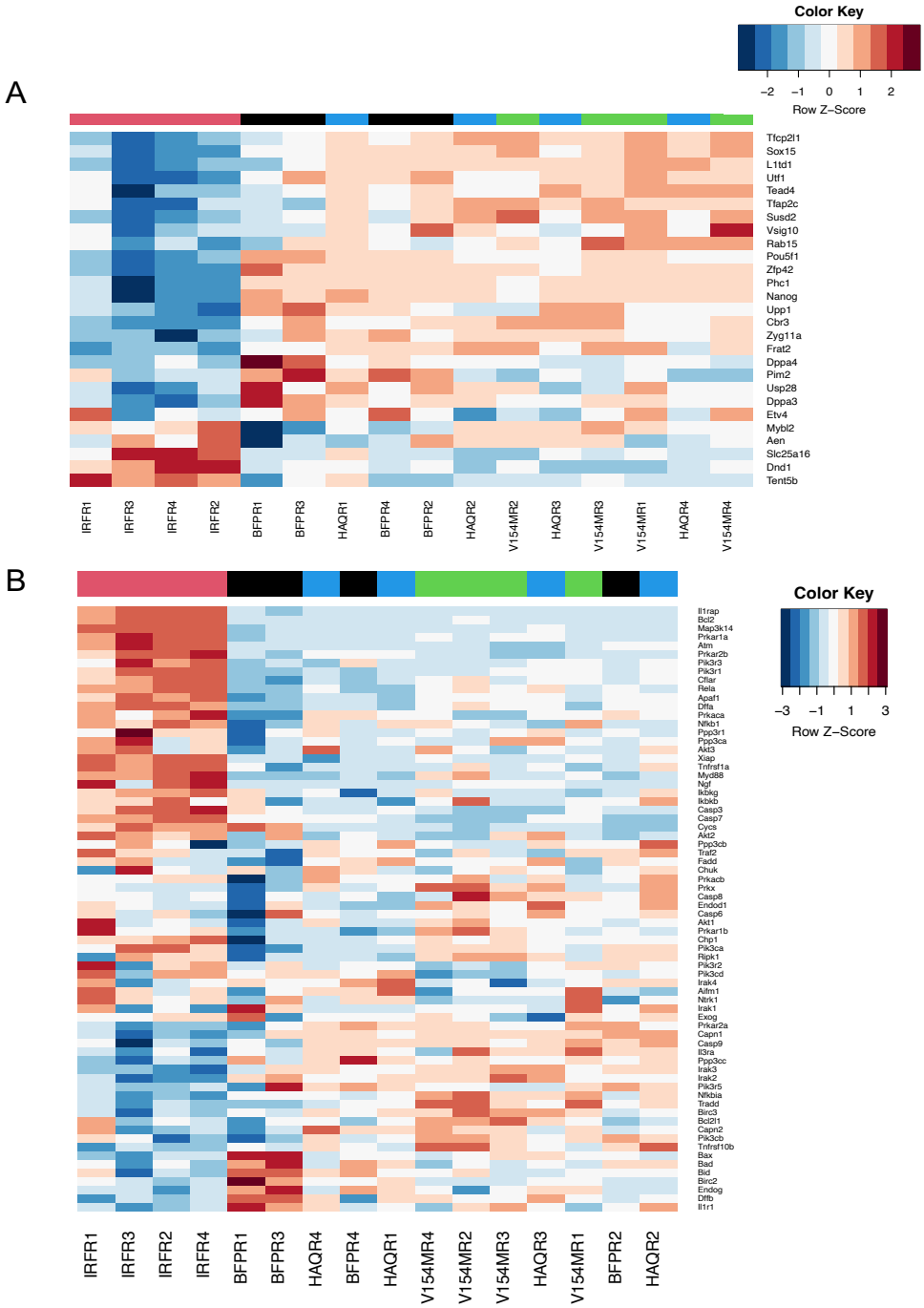


Figure 4-4: IRF3-5D OE expression modifies the expression of pluripotency and cell-death associated genes. A) Heat-map showing expression changes of pluripotency and stem cell associated genes across samples. (B) Heat-map showing expression changes of apoptosis associated genes across sample.

A sustained antiviral response in other cell contexts leads to cell death(Hertzog et al., 1994; Solis et al., 2006). Hence, I was particularly interested in the potential effect in ESCs of mIRF3-5D OE. GSEA analysis identified apoptosis as a significantly enriched term; heatmap analysis of an apoptotic gene set indeed revealed that a subset of pro-apoptotic genes are upregulated by mIRF3-5D. This suggests that activation of IRF3-5D in mESCs can lead to cell death

(Figure 4.4B). A lot of the genes are related to the NFκB pathway, something consistent with *Nanog* repression, being Nanog a repressor of the NFκB pathway (Torres and Watt, 2008).

4.2.3. mIRF3-5D OE compromises mESCs self-renewal

I was particularly interested in following up on the upregulation of apoptotic genes and downregulation of pluripotency related genes shown by the gene expression analysis. Next therefore, I wanted to determine the possible effect of these changes upon mESC self-renewal. To do this, I repeated the construct transfection, sorting BFP+ cells, then plating at lower cell density for 5 days for a colony-formation assay followed by AP staining. The self-renewal assays consistently showed a lower number of colonies in those populations transfected with mIRF3-5D OE plasmids (Figure 4.5A), while results were variable in those replicated expressing mSTING V154M. The same experiment was repeated with mESCs +/- IFNβ to compare any potential detrimental effects via the IFN signalling pathway. IFNβ was refreshed every day, with no differences being observed between both groups, suggesting therefore an IRF3 specific effect (Figure 4.5B).

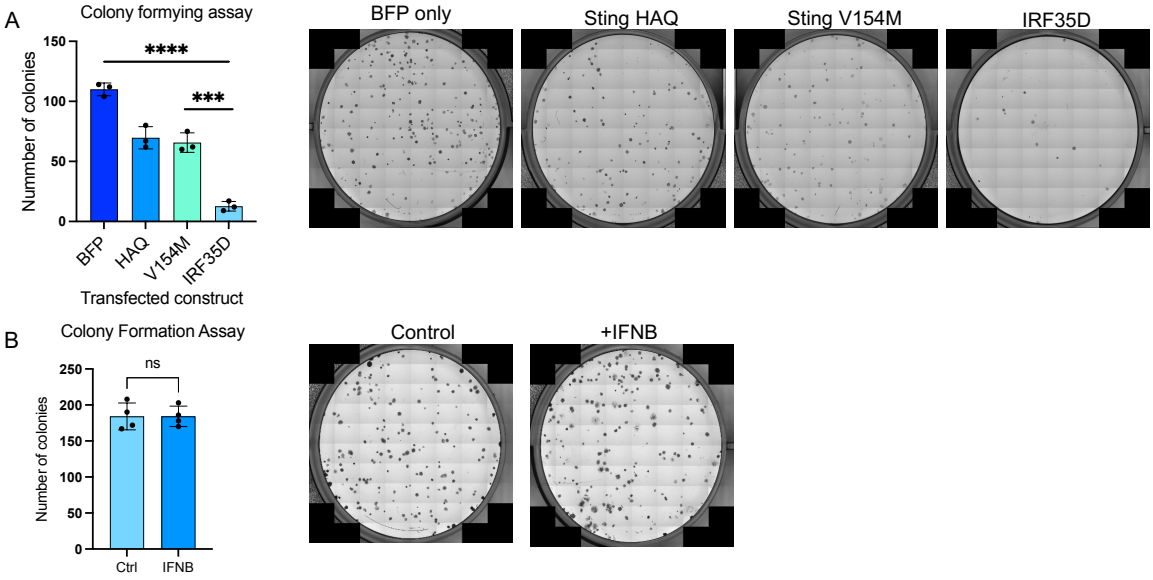


Figure 4-5: IRF35D OE results in compromised self-renewal in mESCs. A. Colony formation assay in mESCs transfected with plasmids OE BFP, BFP_mSTINGV154M HAQ, BFP_mSTINGV154M and BFP_IRF35D. T-test on the number of colonies between BFP and IRF35D groups (p-value= 0.0099), on the number of colonies between HAQ and IRF35D groups (p-value=0.017), on the number of colonies between V154M and IRF35D groups (p-value=0.03) Results are representative of 4 independent experiments. B. Colony forming assay on mESCs treated +/- IFNβ. Results are representative of 2 independent experiments. T-test on the number of colonies between both groups (p-value=0.657).

4.3. Discussion

In this chapter I tried to assess the sensitivity of mESCs to IFN and evaluate the effects that the activation of the anti-viral programme could have on mESCs. Moreover, I observed that upon IFNβ treatment, STAT1 is phosphorylated in mESCs, and therefore the JAK/STAT pathway is active in mESCs in response to IFNβ (Figure 4.1C). Despite that, my data suggest that mESCs are incapable of mounting a functional antiviral response when treated with IFNβ, as canonical ISGs expression remains unperturbed, while intrinsically expressed ISGs expression is mildly upregulated (Figure 4.1A, B). This agrees with an attenuated innate immune system in mESCs reported by previous authors (D'Angelo et al., 2018b; Hertzog et al., 1994). In contrast, intrinsic pluripotency-expressed ISGs seem to have a low sensitivity to IFNβ, probably caused by an open chromatin state that allows their expression. Besides this, the upregulation window seems to be larger for these ISGs than for the canonical ISGs, as shown by the fact that they stay upregulated 24h hours post IFNβ treatment, as their nature makes their expression more tolerable for extended periods. Nevertheless, this does not exclude potential detrimental effects during development caused by their OE, as IFITM proteins OE can inhibit placental syncitiotrophoblast formation and cause foetal demise (Buchrieser et al., 2019). This longer upregulation time also points to a mechanism allowing to discriminate between both sets of genes.

IFNβ treatment of E2.5 embryos showed that it did not alter their developmental potential (Figure 4.1D). However, since the treatment was performed before trophoblast specification, we are not accounting for possible effects that IFNβ could have on the trophoblast, as it is known to be a type I IFN-sensitive tissue (Aikawa et al., 2014; Fendereski et al., 2022).

It remains to be answered why mESCs show an attenuated IFN response, and at the same time can activate certain STAT1 target genes, but not others. The lower expression of STAT1 in pluripotency resulting in lower levels of p-STAT1 upon IFNβ exposure could be a reason. Another explanation can be a potentially restricted chromatin accessibility at canonical ISGs promoters caused by OSKM. This would be consistent with how Nanog interacts with Nfkb1 and represses its activation, while Klf4 has been observed to bind the promoter of *Ifnb* and repress it (Luo et al., 2016b; Torres and Watt, 2008). Epigenetic repression may explain the expression patterns for canonical ISG but it does not explain the minor upregulation of intrinsic ISGs in mESCs when compared to MEFs. This could be tested by motif analysis of the STAT1 binding sites in pluripotency, assessing chromatin accessibility on the promoters of these genes by analysing ATAQ-seq datasets available online.

Interestingly, I found that in contrast to IFNβ, the direct OE of IRF3-5D seems to at least partially rescue a substantial amount of the antiviral response (Figure 4.2 B,C and Figure 4.3). However, I do not have RNA-seq data of the antiviral response in a somatic context to have a clearer idea of the magnitude of the rescue. IRF3-transfected mESCs were capable of upregulating different interferon genes, as well as many genes associated with the canonical antiviral response (Figure 4.3 A, D). Other DEG genes included the downregulation of pluripotency-associated transcription factors, strengthening the picture of a distinct antagonism between canonical innate immunity and pluripotency (Figure 4.4A). Apoptosis-related genes were also upregulated (Figure 4.4B). Dysregulation of apoptosis and pluripotency related genes seems to translate into a self-renewal defect phenotype *in vitro* (Figure 4.5A). This phenotype depends on IRF3, as IFNβ treatment does not compromise mESCs self-renewal (Figure 4.5B). In this study, IRF3-5D OE was used to mimic a fully functional cGAS/STING signalling on mESCs, suggesting that a potential activation of the system by LINE1 or other TE-derived cDNA could indeed lead to a clear detrimental outcome, independent of IFN signalling. Nevertheless, a limitation is that these data come from an IRF3 OE scenario. Ideally, I would restore the upstream components to mimic a more physiological response with physiological protein levels.

Finally, the *in vivo* context of fertilization and implantation is very relevant to these findings. My results show a clear attenuated antiviral response in mESCs upon exposure to IFNβ that, indeed is consistent with the fact that IFNβ does not seem to affect the embryo development before implantation. However, ICM cells are surrounded by the trophectoderm and are in the environment of maternal endometrial cells. Thus, one hypothesis is that IFNβ production by cells in the ICM caused by LINE1 could trigger an antiviral response on the trophectoderm and the maternal endometrium, still needs to be tested, and it cannot be concluded categorically that IFNβ has no effects on development. This could potentially impair implantation and early post-implantation development, as they are tightly immune-regulated processes. Moreover, the pool of ISGs directly activated by IRF3 could compromise ICM viability directly. Some of these ISGs, like *Ccl5* and *Isg15* (NK cells activator), are chemokines and cytokines that activate the maternal cellular immune system. This could potentially have highly detrimental consequences for embryo development (Bhat et al., 2020; Perng and Lenschow, 2018).

Chapter 5: : Results: Characterization of LINE1 dynamics in pluripotency and senescence

5.1. Introduction

The different responses to LINE1 in pluripotency versus in other somatic contexts such as senescence could have multiple explanations. The first hypothesis was explored in chapters 3 and 4: that potential suppression of the innate immune system would dampen LINE1 cDNA sensing. In chapters 5 and 6, I explored a second hypothesis – that pluripotency-specific mechanisms may restrict LINE1 retrotransposition and thus cDNA/ssDNA generation.

The identification of protein-protein interactions is essential for understanding complex biological processes. Differences in the LINE1 interaction partners between pluripotency and a detrimental context like senescence may explain the different response to LINE1 expression in both contexts. For example, the expression of pluripotency-specific factors that inhibit LINE1 post-transcriptionally would allow LINE1 expression while preventing its detrimental functions. In contrast, identifying a lack of senescent/detrimental-specific interactions in pluripotency could also explain the uncoupling of LINE1 expression and pathology that is seen in ESCs and embryos.

Generating and analysing LINE1 interactomes has been a common way to identify modulators of retrotransposition. Multiple LINE1 ORF1p and ORF2p interactomes have been published to date(Goodier et al., 2013; Mita et al., 2018; Moldovan and Moran, 2015; Taylor et al., n.d.; Vuong et al., 2019). However, in nearly all these cases, the co-immunoprecipitated proteins have been identified from the over-expression of a LINE1 construct in a context when endogenous LINE1 is not expressed.

In this study, a LINE1 RNP interactome was obtained through the immunoprecipitation of endogenous ORF1p and analysis by mass-spectrometry. Ideally, the target would have been ORF2p due to its enzymatic activities being relevant to explain LINE1-related pathology. However, there is no suitable antibody for murine endogenous LINE1 ORF2p, and detection has not been achieved in mass spectrometry. In fact, ORF2p detection has only been possible with antibodies in ectopically expressed bicistronic constructs. Endogenous ORF2p expression has been to this day not possible even in tumour samples harbouring hundreds of somatic LINE1 insertions (Ardeljan et al., 2019). Using an overexpressed ORF2p was also

discarded because it for this study I wanted to obtain the endogenous LINE1 interactome, in the context of the assembled LINE1 RNP.

Therefore, immunoprecipitation of the endogenous ORF1p interactome was favoured over an ectopically expressed one. Endogenous immunoprecipitation better reflects a protein's expression in its native context, which may better reflect its normal cellular localization, PTMs, interactions with other proteins, and functional activities. Immunoprecipitating endogenous proteins also has the advantage of avoiding potential artefacts from OE caused by cellular stress or non-physiological protein levels that may activate cellular response pathways or cause potential off-target effects. Additionally, immunoprecipitating endogenous proteins allows for the preservation of native protein-protein interactions, which can be critical for studying the physiological relevance of these interactions in a cellular context.

Hence, in this chapter it was necessary to first establish models in which ORF1p could be immunoprecipitated endogenously. For pluripotency, mESCs express high levels of LINE1 RNA and ORF1p (Calarco and Brown, 1969; Fadloun et al., 2013). The expression of ORF1p needed to be verified in our model, to prove its pluripotency-specific expression. In addition, it was necessary to test whether endogenous ORF1p could be immunoprecipitated in pluripotency reliably.

In the case of an interactome for LINE1 where it is known to have a detrimental effect, first, I needed to establish a senescence model expressing LINE1. DNA damage-induced senescence was the model of choice due to practicality and its similarity to previous models in which ORF1p has been detected (Cecco et al., 2019); the protocol was obtained from Dr. Alexis Barr. I managed to establish a stressed-induced senescence model and describe LINE1 expression dynamics in it. This was of prime importance as LINE1 upregulation timepoint varies depending on the cell line and the type of senescence.

5.2. 5.2. Results

5.2.1. LINE1 dynamics in pluripotency

ESCs are the model of choice for pluripotency, and also express high LINE1 RNA and ORF1p levels (Calarco and Brown, 1969; Fadloun et al., 2013). However, ORF1p dynamics was not further investigated. I thus first validated ORF1p expression by IF in pluripotent ESCs, which was shown to decay sharply after passive differentiation. This was achieved by culturing ESCs

for 5 days in media lacking LIF (Figure 5.1A). The ORF1p staining is mainly cytoplasmic, as reported previously (Percharde et al., 2018). Hence, ORF1p expression can be described as a feature associated with pluripotency.

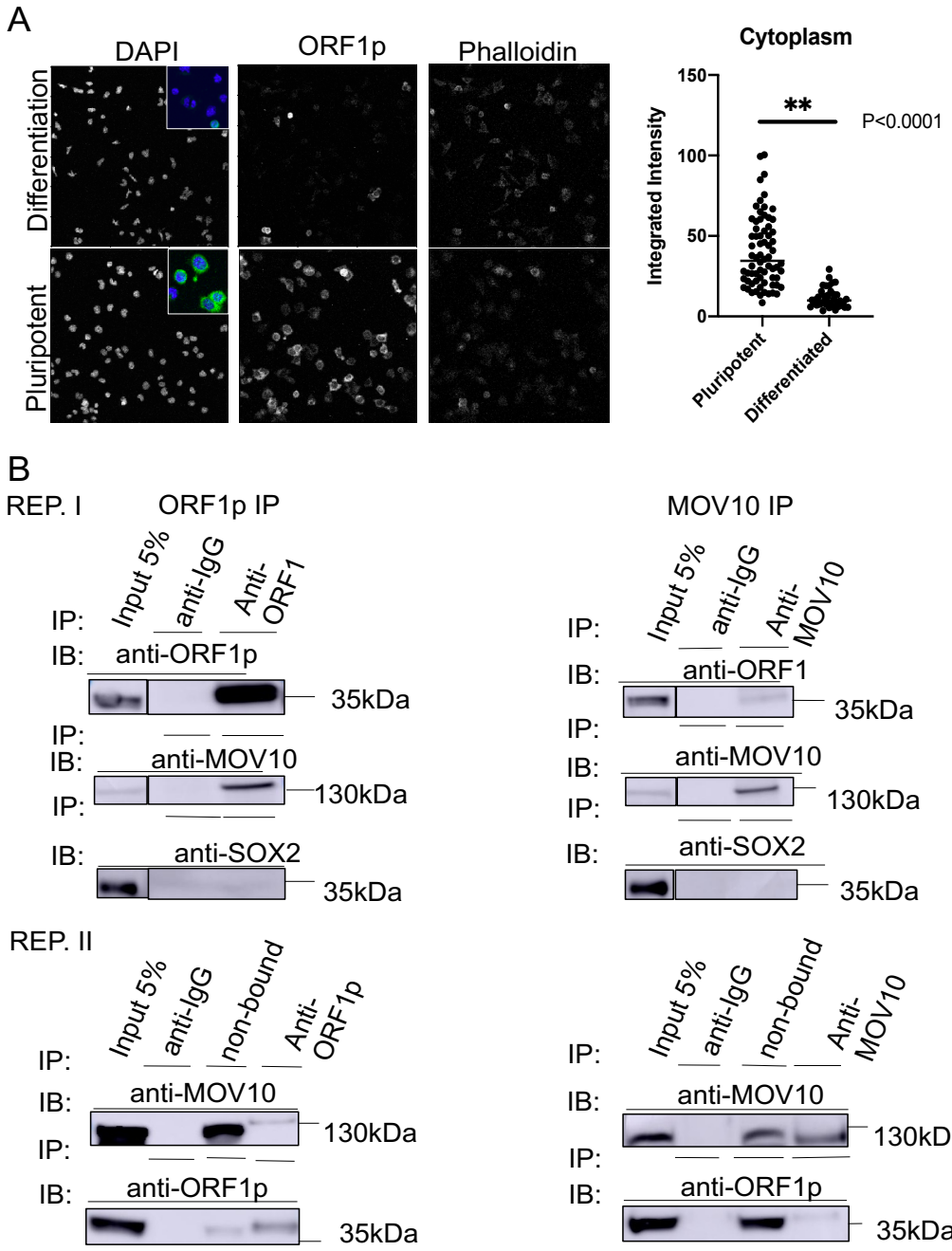


Figure 5-1: LINE1 is expressed in pluripotent mESCs. (A) ORF1P IF in pluripotent ESCs and after 5 days passive differentiation. Bellow, quantification of cytoplasmic ORF1p staining in mESCs. Unpaired two-tailed t-test was significant P<0.0001. (B) Immunoprecipitation of endogenous ORF1p and MOV10 in mESCs. Rabbit IgG was used as a negative control. Sox2 works also as a negative control on the IP, as it does not bind any of the proteins. Non-bound fraction refers to the supernatant of the incubation reaction between the antibody bound magnetic beads and the protein extract. In the second replicate, just 5% of the IP volume was loaded into the gel. MW: ORF1p: 40kDa, Sox2: 35kDa, MOV10: 114kDa.

Having confirmed the endogenous ORF1p expression in our model, two pilot IPs were performed as a proof of concept, using normal rabbit IgG as a negative control. ORF1p was successfully immunoprecipitated (Figure 5.1B). In addition, MOV10, a known ORF1p interactor, was co-immunoprecipitated, while Sox2 as a negative control, did not co-immunoprecipitate (Figure 5.1B) (Choi et al., 2018). A reciprocal IP using MOV10 antibodies was also performed, achieving the complementary result by co-immunoprecipitating ORF1p (Figure 5.1B**)**. However, a considerable fraction of protein remained in the non-bound fraction, suggesting a saturation or inefficiency of the MOV10 antibody.

5.2.2. LINE1 dynamics in senescence

5.2.2.1. Etoposide treatment induces senescence in MEFs

DNA damage-induced senescence was the model of choice, as it was already established by my collaborators. The first cell line used was the NIH 3T3 line, which was discarded as a model due to the inability of these cells to induce a senescence programme, as p16 is deleted from their genome (Klafter and Arbiser, 2000) (Figure 5.2A). In subsequent experiments, senescence was induced in mouse embryonic fibroblasts (MEFs) by treatment with 20μM of etoposide for 3 days followed by long-term culture. Etoposide is a topoisomerase inhibitor used in the treatment of several cancers due to its capacity to generate DNA damage. Specifically, it inhibits topoisomerase II, which is responsible for unwinding and resealing DNA stands during cell division, preventing the cell from properly replicating its DNA and leading to DNA damage (Hande, 1998).

Six days after etoposide treatment, MEFs were shown to be senescent by X-Gal staining, the standardised senescence marker (Figure 5.2A). Senescence induction was later confirmed by assessing the RNA and protein expression of known senescence markers: early upregulation of *Cdkn1a* (p21), later upregulation of *Cdkn2a* (p16), and downregulation of *Lmnb1* (Figure 5.2AB). This confirmed that our experimental conditions were capable of triggering senescence in MEFs. The maintenance of senescence was confirmed by assessing that the RNA expression levels of *Cdkn2a* and *Lmnb1* were maintained at later time-points (Figure 5.2C).

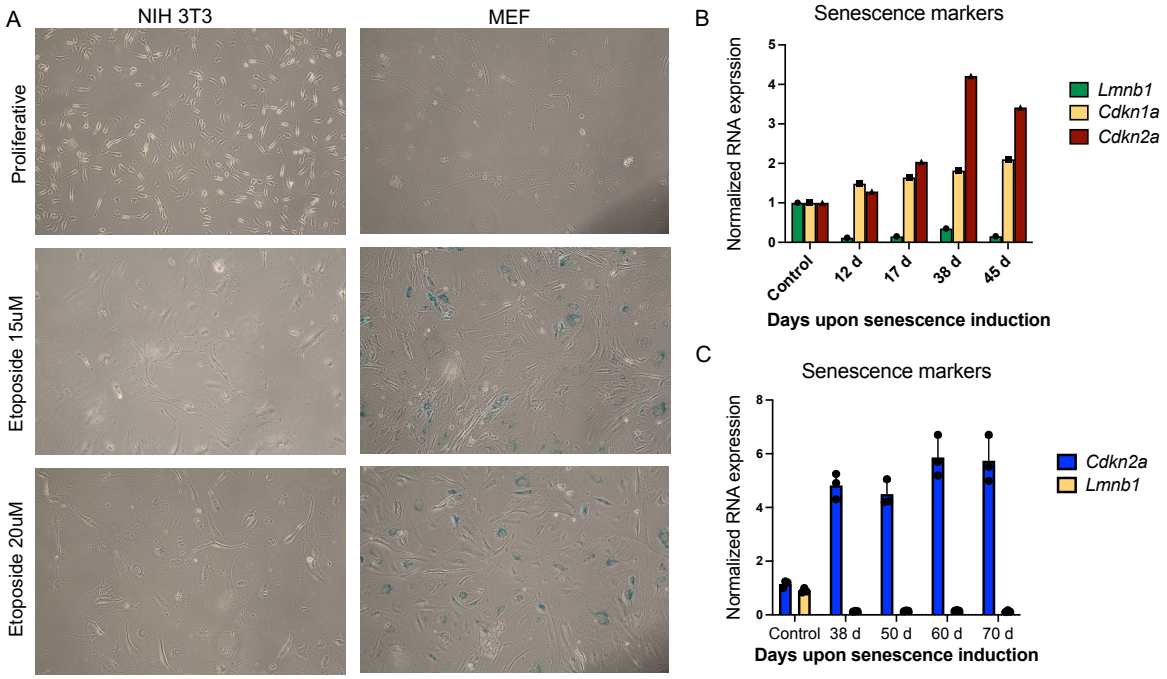


Figure 5-2: MEF undergo stable senescence upon etoposide treatment. **(**A) X-Gal staining of 3T3 and MEF cells treated with different concentrations of etoposide for 3 days and left an additional 3 days of rest. Images taken 20X. (B) RNA expression data of the senescence markers *p16*, *p21* and *Lmnb1* at different timepoints upon etoposide treatment. Data is a mean of 1 biological replicate1 (N=1) (C) Repeat of experiment in B in a longer time-course. Data is a mean of 3 biological replicates (N=3).

5.2.2.2. LINE1 RNA is gradually upregulated at later timepoints after senescence induction.

As described in Cecco 2019, LINE1 is upregulated in late-stage senescence (Cecco et al., 2019). The entrance into this state may vary depending on the cell line and the type of senescence. Therefore, it was necessary to first characterise LINE1 behaviour in our senescence model. Senescent MEF RNA samples were collected at different time points up to 45 days post-senescence (PS). Then, I assessed the expression of 3 of the youngest LINE1 families active in mice (L1Md_A, L1Md_Gf and L1Md_Tf) and *Ifnb* to link LINE1 upregulation with the chronic activation of an innate immune response (Figure 5.3A,B). The expression of other TEs like SINEs was also analysed in the same samples, as the same heterochromatin-loss mechanisms that lead to LINE1 expression can apply to them (Figure 5.3C).

During the 45 days of evaluation, I found that LINE1 is gradually upregulated in parallel to *Ifnb* and *Cdkn2a* but without seemingly reaching saturation (Figure 5.3A,B). Hence, a new time-course was set up, aiming at later time points. LINE1 RNA expression analysis showed a

clearly increased upregulation of LINE1 at later timepoints (Figure 5.3A,D). However, LINE1 upregulation did not seem to reach a plateau.

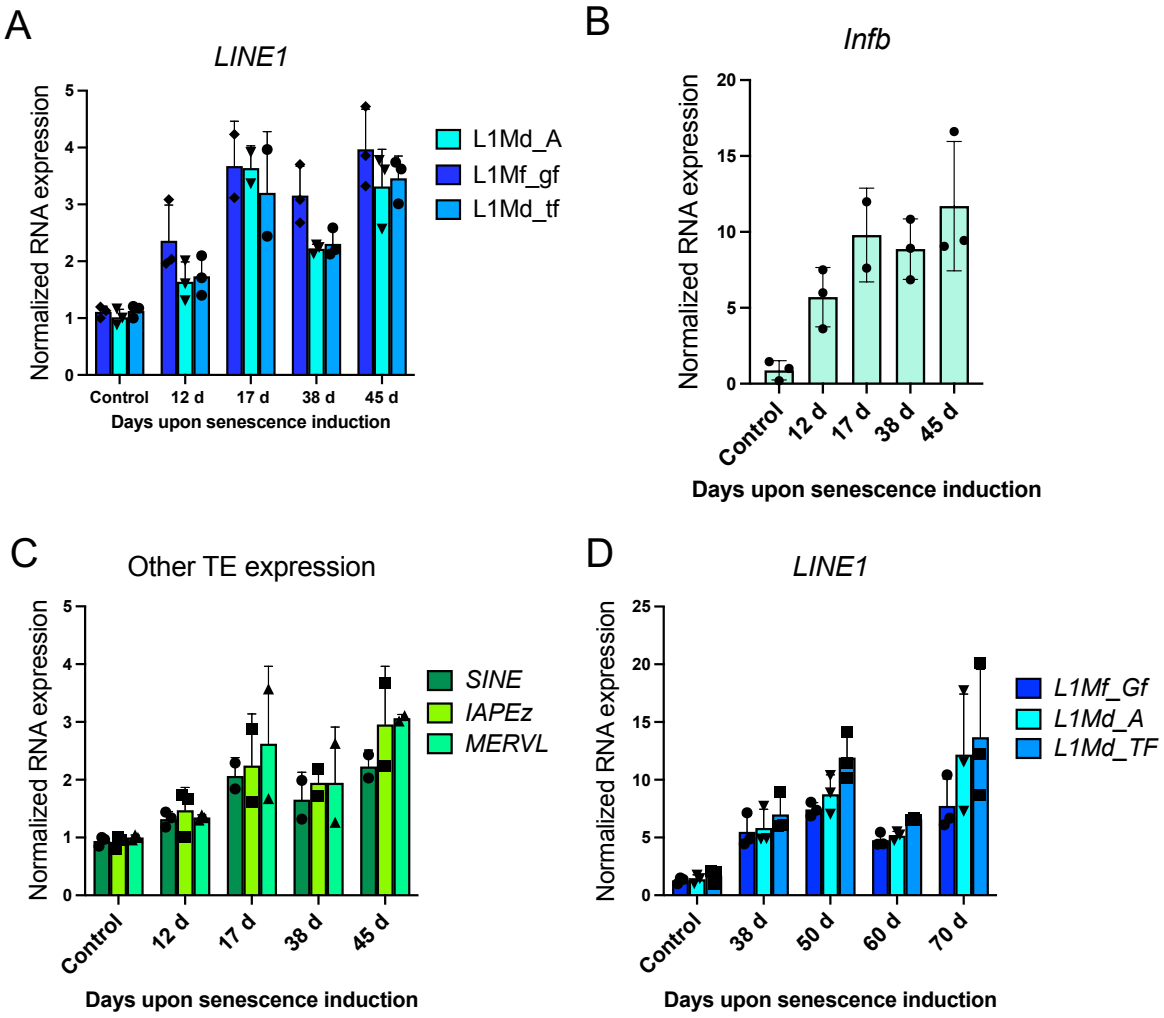


Figure 5-3: LINE1 and other markers are upregulated on later stages of senescence induction in MEF. (A) Time-course expression of *young* LINE1 families (L1Md_Gf, L1Md A, L1Md_Tf) across 45 days. (B) Time-course expression analysis of *Ifnb* (C) Time-course expression of different TEs (SINE, IAPEZ and MERVL) across 45 days. D. Time-course expression *of* LINE1 families active in mice (LINE1Md_Gf, LINE1Md_A, LINE1Md_Tf) across 72 days. Data in A, B, C and D are mean of 3 biological replicates (N=3) .

The experiment was replicated again per duplicate, this time checking for ORF1p protein expression. ORF1p was found to be stably upregulated by 72 days (Figure 5.4A, B, C). The expression of ORF1p in senescence was also compared to that in pluripotency, which highlighted the substantially higher expression of ORF1p in pluripotency (Figure 5.4B). For all samples, the total protein was quantified by BSA assay; therefore, differences in loading control between samples may be attributable to cell type.

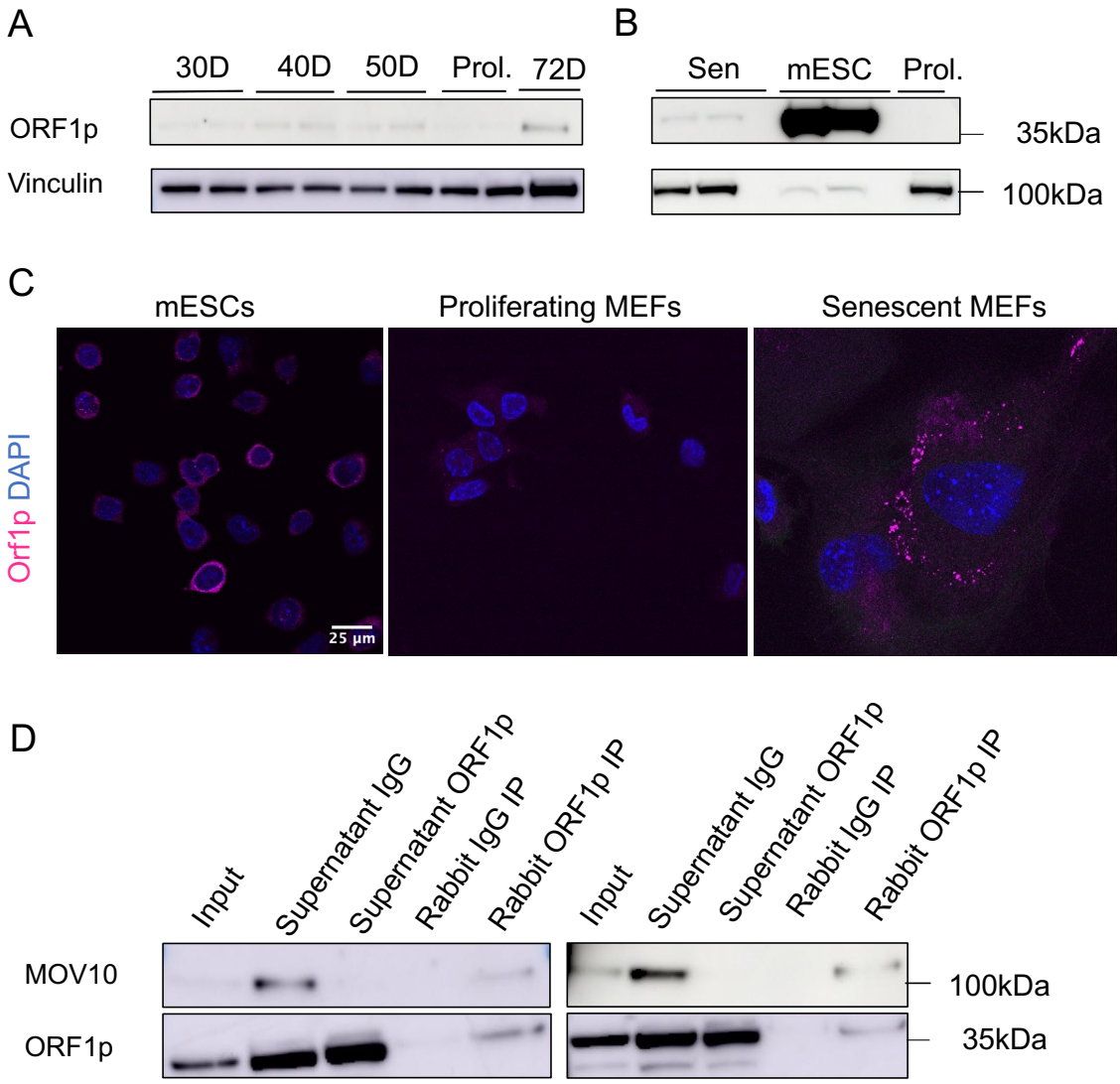


Figure 5-4: LINE1 Orf1p is detectable in senescent cells and it can be immunoprecipitated. (A) Immunoblot of ORF1p time-course expression in MEFs 30, 40, 50 and 72 days upon senescence induction compared with proliferative MEFs. (B) Immunoblot of ORF1p in MEFs 72 days after senescence induction, mESCs and Proliferative MEFs. N= 2 biological replicates. (C) ORF1p staining (magenta)) in mESCs, 72-day senescent MEFs and proliferative MEFs. MW: ORF1p (40kDa), Vinculin (124kDa). Mov10 (114kDa). Images are representative of (N=3 independent experiments). Scale bar: 25µm. (D) Immunoprecipitation of endogenous Orf1p in senescent cells (N=2). Rabbit IgG was used as negative control, while the ORF1p supernatant revealed the unbound protein fraction to beads upon overnight incubation in ORF1p antibody

5.2.2.3. Endogenous Orf1p can be immunoprecipitated in senescent MEFs

Previous experiments had revealed that many plates of MEFs were needed to achieve enough material for immunoprecipitations. Another concern was the lower expression of endogenous ORF1p compared to pluripotency. For this reason, a pilot IP experiment with ORF1p was next performed to visualise if it was possible to immunoprecipitate the endogenous protein at this stage. The IP was performed using 72-day senescent cells. The choosing of 72-days was decided to make sure that ORF1p was stably expressed, as the small RNA expression

decrease seen in day 60 (Figure 5.3D), gave me concerns that the expression of LINE1 could be very variable between samples in senescence. Therefore, choosing a later time-point aimed to achieve a higher confidence on upregulated LINE1 expression levels. For this experiment, the “non-bound” IP supernatant fraction was also loaded into the gel. ORF1p was successfully immunoprecipitated in senescent cells with no protein precipitated in the control and minimal presence of Orf1p in the channel corresponding to the non-bound fraction. ORF1p IP also co-immunoprecipitated MOV10 (Figure 5.4D). In conclusion, I successfully established a senescence LINE1 induction protocol for ORF1 IP harvesting.

5.2.3. ORF1p IP in ESCs successfully identifies ORF1p, previously known interactors and a set of potentially new LINE1 interactors.

Upon validation of endogenous LINE1 IP on mESCs and the submission of a pilot experiment (data not shown); 3 replicates of ORF1p IP in mESCs with the corresponding controls were submitted to mass spectrometry (Figure 5.5A). The raw data was analysed by the LMS proteomics facility and the appropriate quality controls were performed. Sample correlation analysis shows a ≅1 correlation in the pairwise comparison within controls and ORF IP interactomes. A total of 177 hits were only found on the ORF1p IP, while 336 were only binding to the control IgG. In total 1,120 hits were found to be significantly enriched in the ORF1p IP.

From the immunoprecipitation experiment, I first searched for the presence of already known LINE1 interactors identified in other interactomes, like MOV10, Larp1, Nucleolin and Pura (Figure 5.5A, Green) (Goodier et al., 2013; Mita et al., 2018; Moldovan and Moran, 2015; Taylor et al., n.d.; Vuong et al., 2019). This revealed that these factors were also found to interact with endogenous ORF1p in ESCs (Figure 5.5A,B). The aforementioned interactomes, except for Vuong 2019 which used hESCs, were obtained by over-expressing LINE1 in HeLa or HEK293T cells. Cross-comparison of these datasets allowed me to obtain a set of genes common in all ORF1p published interactomes (Figure 5.5A, green, and B). I also confirmed the presence of known somatic cell repressors of LINE1 such as Dicer, Ago proteins, Apobec proteins, Adar1, Mov10, Zcchc3, Tdp-43 (Tardbp), Helz and Trim5 (Figure 5.5A, Green and Magenta) (Adiliaghdam et al., 2020; Aporntewan et al., 2011; Bodak et al., 2017; Choi et al., 2018, 2018; Ding et al., 2023; T. D. Li et al., 2022; MacLennan et al., 2017; Orecchini et al., 2017; Volkmann et al., 2020; Wissing et al., 2011; Z. Zhang et al., 2022).

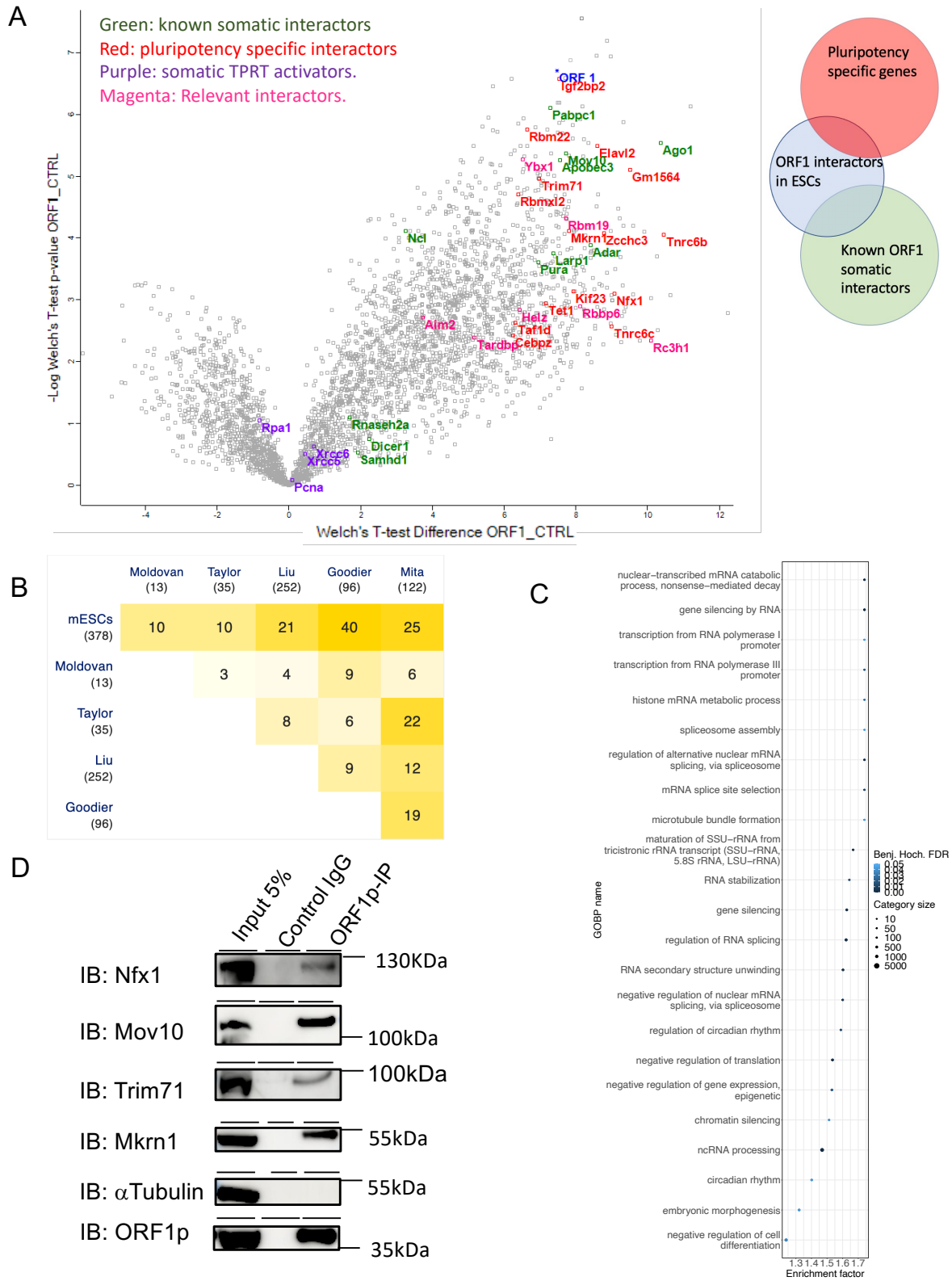


Figure 5-5: ORF1p endogenous interactome in pluripotency. (A) Volcano plot of ORF1 IP enriched hits. The Y axis registers the -log of the Welch t-test p-value depicting the significance of the hit. Therefore, significant hits are on the top of the graphs. The X axis depicts the Welch t-test difference of ORF1 versus control IPs, as a measure of enrichment. Therefore, hits enriched in the ORF1 IP compared to the control are on the right side of the graphs. Green shows previously known somatic LINE1 interactors, red depicts pluripotency specific LINE1

interactors, purple depicts factors involved in LINE1 cDNA genomic integration seemingly depleted from the interactome, magenta depicts interactors found relevant for the course of this work. (B) Table depicting pairwise comparisons between the ORF1p interactome in ESCs (this work) and other published LINE1 interactomes. (C) Plot of the gene ontology analysis depicting relevant categories in the GO biological processes. (D) Co-immunoprecipitation of Nfx1, Mkrn1, Trim71 and Mov10 with ORF1 IP. Normal rabbit IgG was used as a technical negative control, while αTubulin was used as a biological negative control. MW: ORF1p: 40kDa, Mov10: 114kDa, Trim71: 71kDa, Mkrn1: 51kDa, Nfx1: 123kDa, αTubulin: 50kDa.

My main interest was to identify LINE1 pluripotency-specific interactors. Therefore, the dataset was cleaned of already known somatic interactors and compared to a list of pluripotency-specific genes obtained from a mESCs transcriptomic metanalysis of 171 gene expression profiles measured in mESCs (De Cegli et al., 2013). This allowed me to identify a list of 20 pluripotency-specific potential new LINE1 interactors (Figure 5.5A red, Annex II green). Most of these interactors are interestingly related to translational regulation, RNA silencing, RNA binding and ubiquitination. Three of these interactors: Nfx1, Trim71 and Mkrn1 were next validated to be ORF1p interactors by ORF1p IP, using MOV10 as a positive control and αTubulin as a negative non-binding control. Therefore, I could assume that the IP was performed correctly and the detected interactions were probably real (Figure 5.5D).

Many of the pluripotency-specific interactors were present in the LINE1 ORF1p hESCs interactome published by Vuong et al. Comparison of this dataset with ours and other published datasets allowed me to identify a set of interactors only found in hESCs and mESCs datasets (Vuong et al., 2019). Other factors not labelled as pluripotency-specific but that appeared to be highly enriched only in ORF1p from ESCs were considered potentially interesting as LINE1 repressors and were also included (Annex II). This was the case notably for *Mff*, *Traf4*, *Rbbp6*, *Itch* and *Rc3h1*; the last three being also involved in innate immunity and protein degradation, according to GO analysis. These ORF1p interactors were compiled in a list of 74 potential LINE1 regulators, with the future aim of testing their LINE1 repression capacity (see next chapter) (Annex II).

GO ontology analysis of the enriched significant proteins in the dataset especially highlights spliceosome and splicing-related items, translation regulation, ribosome biogenesis, RNA-mediated gene silencing, helicase and retrovirus-related items like poly-purine tract binding (Figure 5.5C). Other enriched relevant terms not depicted in the graph are Zn^{2+} fingers, enhancers and DNA mediated regulation of gene expression, histone modifications, core promoter binding and negative regulation of differentiation.

Interestingly, a lot of the 20 pluripotency-specific interactors were ubiquitin-ligases. In the overall interactome “proteasomal ubiquitin-related protein catabolism” has a significant p-value, but its enrichment factor is <1, with 33 proteins significantly enriched out of 88 present in the category. Other relevant categories were RNA silencing, RNA binding and translational regulation.

5.2.4. ORF1p and MOV10 cannot be detected in senescent MEF ORF1p IP

Since LINE1 expression is detrimental in somatic contexts like senescence, I wanted to obtain an endogenous LINE1 ORF1p interactome to compare it to the pluripotent interactome. Three 72-day senescent cells ORF1p IP samples were submitted in the LMS proteomics facility with the appropriate controls, and the equivalent experiment in pluripotency repeated and resubmitted to allow comparison of datasets. ORF1p correct IP was checked by western blot in every submission.

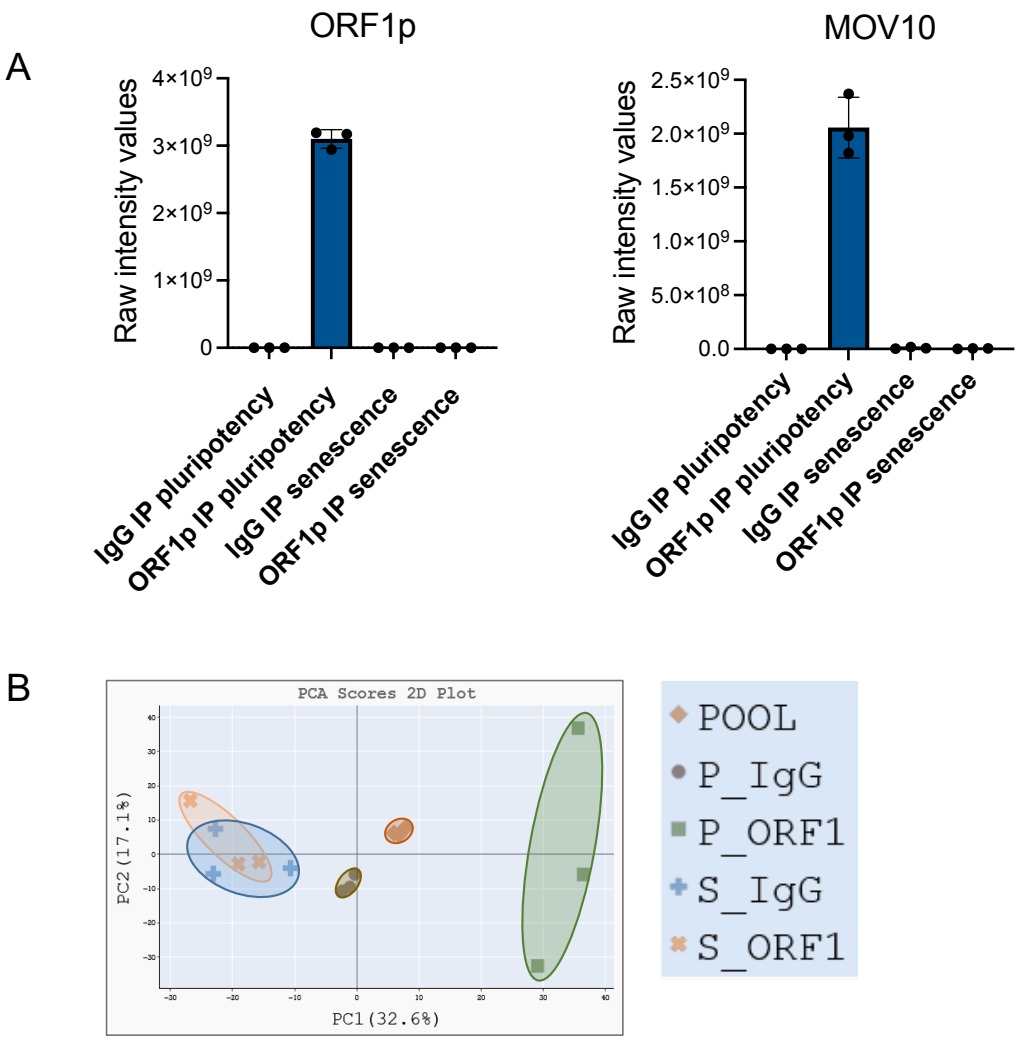


Figure 5-6: Cytoplasmic ssDNA identification and ORF1p IP in senescence mass spechtrometry analysis. (A) ORF1p raw intensity across tested conditions: ORF1p and IgG

IPs in pluripotency and senescence. (B) PCA plot depicting similarities among ORF1p IP in senescence (S_ORF1), normal rabbit IgG in senescence (S_IgG), ORF1p IP in pluripotency (P_ORF1) and normal rabbit IgG IP in pluripotency (P_IgG). Pool is referred as a combination of equal volumes of the two samples, representing a physical average of the protein signals in the samples. Works as an internal QC that is injected at regular intervals to gauge LC-MC performance over time. This would allow do detect any time-related batch effect.

The digestion efficiency was correct in all the samples, and samples correlated well with their same-group replicates. A total of 2299 proteins were detected in senescent samples, but in none of the three replicates of ORF1p IP, was ORF1p or any of its known interaction partners detected (Figure 5.6A). All three submitted ORF1p IPs in senescence failed to detect ORF1p protein or any of its known interactors. PCA analysis showed few differences between ORF1p IPs and negative controls in senescence (Figure 5.6A, B). Most of the proteins localized were cytoskeleton and fibrous proteins like actins. Therefore, it was not possible to characterise endogenous LINE1 ORF1p in senescence.

5.2.5. No cytoplasmic DNA is detected in mESCs.

Cytoplasmic ssDNA seems to be the immunogen triggering the innate immune response to LINE1 in senescence Hence, it was important to clarify its presence in mESCs (Benitez-Guijarro et al., 2018; Cecco et al., 2019; Simon et al., 2019). The absence of ssDNA would then point definitively to a mechanism preventing LINE1 retrotransposition activity.

To determine ssDNA presence, mESCs, proliferative MEFs and 72 days senescent MEFs were incubated with an anti-ssDNA DNA antibody, which has been previously used to image TE-derived ssDNA in senescence (Cecco et al., 2019). The same protocol was used: fixing with 4% PFA, followed by methanol and digestion with RNase P before blocking. The IF did not detect any quantity of ssDNA in mESCs and the proliferative MEFs but successfully detected cytoplasmic foci of ssDNA in senescent MEFs (Figure 5.7). These foci tend to localize around the nucleus and present in a subpopulation of cells (Cecco et al., 2019). This suggests an absence of ssDNA in the cytoplasm of mESCs. Nevertheless, other methods like cytoplasmic DNA extraction followed by qPCR would be necessary to confirm this result.

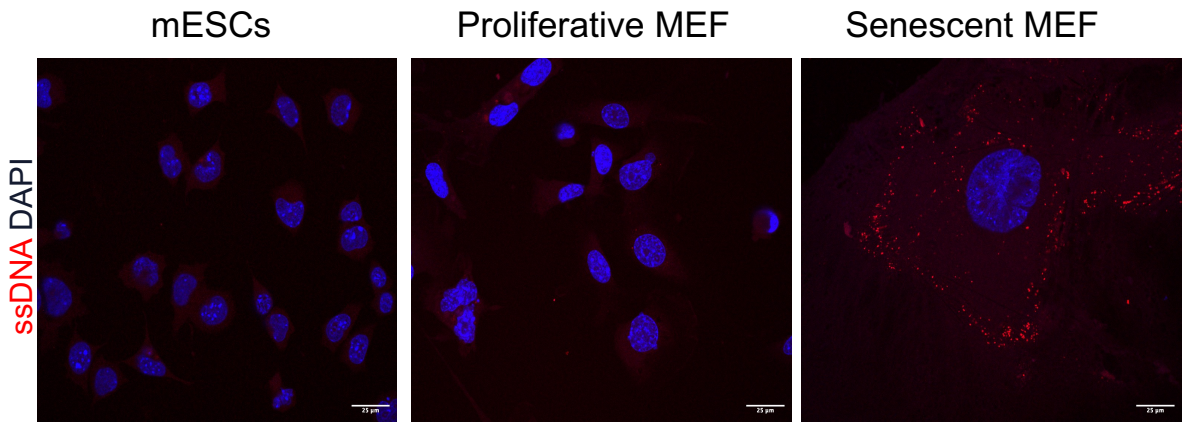


Figure 5-7: No ssDNA is detected in mESCs. No ssDNA staining (red) in mESCs, 72-day senescent MEFs and proliferative MEFs. Scale bar: 25μm.

Finally, I was also interested in whether known ORF1p interactors from somatic cells may be missing from ORF1p in pluripotency. It could be hypothesised that an absence of interaction in ESCs with a somatic LINE1 interactor could protect ESCs from LINE1 detrimental effects – for example, a necessary cofactor for retrotransposition. Following this rationale, a list of somatic-specific ORF1p interactors was obtained by comparing the list of common interactors present in at least two ORF1p somatic datasets with our ORF1p dataset. Those not appearing in mESCs were considered to be potentially somatic-specific. The GO analysis of this list highlighted terms such as: “ssDNA binding”, “activation of innate immune response” and other immune system-related categories (Annex III). From these factors, Rpa1 was significantly depleted in our mESCs dataset, while Xrcc5 and Xrcc6 were also not significant interactors (Figure 5.5A, purple).

Xrcc5 and Xrcc6 are involved in the NHEJ pathway required for DSB repair, and in the resolution of LINE1 integrations and are thus positive LINE1 regulators (Gasior et al., 2008; Goodier et al., 2013; Morrish et al., 2002; Suzuki et al., 2009). This complex has also recently been implicated in cytosolic DNA recognition (Morchikh et al., 2017; Roberts et al., 2010; Tuteja et al., 1994). Rpa1 is an ssDNA binding protein that plays an important role in DNA damage response, DNA repair, telomere maintenance and DNA replication (Grudic et al., 2007; Haahr et al., 2016; Y. L. Lin et al., 1998; Maréchal et al., 2014; Mita et al., 2020; Sleeth et al., 2007).

In contrast to ESCs, all these factors are significantly enriched in the somatic LINE1 interactomes supporting the scenario of LINE1 retrotransposition being a highly downregulated process in pluripotency.

5.3. Discussion

ESCs endogenously express ORF1p protein at high levels. ORF1p expression was used to model LINE1 expression in pluripotency and was observed to decay quickly. The level of expression of ORF1p in mESCs was optimal and the IP was performed satisfactorily in ESCs without further optimisation of the model (Figure 5.1A). The specificity of IP was confirmed by immunoprecipitating LINE1 interactors like MOV10 but no other proteins like Sox2. Co-immunoprecipitation of ORF1p in MOV10 IP added robustness to the experiment (Figure5.1B). This allowed me to obtain a LINE1 ORF1p endogenous interactome in mESCs (Figure5.5, 5.6A). Detection of previously known interactors verified our approach. I also identified a set of pluripotency-specific genes interacting with ORF1p in ESCs. This was important for my hypothesis: a pluripotency-specific interactor could prevent LINE1 detrimental action while allowing its necessary expression. I assumed that given the pluripotency-specific nature of the phenomenon, it would be highly likely that if such an interactor exists, it would be among those proteins specifically expressed in pluripotency. However, any potential LINE1 repressive activity of these interactors remained to be demonstrated.

A DNA damage-induced senescence model was also established in the laboratory. The model was shown to exhibit gradual upregulation of different LINE1 families alongside the upregulation of *Ifnb*, without proving to have reached a plateau (Figure 5.3). Other TE families were shown to be upregulated with LINE1, although their upregulation seemed to be milder. The upregulation of LINE1 was found to be clear and replicable by 72 days of senescence (Figure 5.3). At this point, endogenous ORF1p LINE1 protein was detected in increased levels compared to previous time points. It was not evident that the levels of ORF1p expression were peaking at this point, but for logistical reasons and time constraints, I decided to carry out the IP experiment at 72 days. Immunoprecipitation of ORF1p was finally achieved with 72-day senescent MEFs. Despite that, it was unfortunately not possible to obtain the interactome of ORF1p in this context due to unexplained technical reasons (Figure 5.6).

The reasons for the ORF1p IP failure in senescence can be multiple. A comparison of the input, the IP supernatant and the IP channels suggested that a substantial amount of protein may be lost in the washes in senescent samples (Figure 5.4C). This indicates that while most of the ORF1p present in the whole cell extract is captured, a lot of it is lost in the washes. ORF1p is perceptible in the IP WB membranes, since it is directly targeted by the antibody, but in the mass spectrometry, it falls below the detection threshold of the machine.

In a comparison of endogenous ORF1p in senescent and pluripotent cells, the latter showed a remarkably higher expression. Assuming ORF1p levels are a proxy of LINE1 expression levels, this makes the non-reactive phenotype of mESCs against LINE1 even more striking, considering that a much lower expression in senescence is capable of upregulating an IFN response and contributing to SASP (Cecco et al., 2019). The higher ORF1p expression in pluripotency, apart from implying a higher expression of LINE1 generally in pluripotency, also implies the preservation of its RNA. ORF1p detection implies that LINE1 RNA it is being translated and in parallel, said RNA forms a complex with Kap1 and Nucleolin that shuts down the 2C transcriptional programme (Percharde et al., 2018). Thus, a mechanism dependent on the regulation of LINE1 RNA levels like *Dicer* endonuclease is not enough to explain the LINE1 phenotype in mESCs as suggested by Ciaudo et al. (Bodak et al., 2017; Bodak and Ciaudo, 2016).

Considering the high amount of LINE1 RNA and ORF1p expression in mESCs. The fact that I did not manage to detect LINE1 ssDNA in mESCs as it would be expected, together with the low LINE1 retrotransposition rate in pluripotency compared to other somatic contexts, particularly transformed cells, when considering similar expression levels, as has been suggested by the work some authors and the present work (Figure 5.4A, 5.7. 6.1) (Arokium et al., 2014; Ewing and Kazazian, 2010; Garcia-Perez et al., 2007; Kano et al., 2009). This suggests that TPRT is highly repressed in this context (Figure 5.4A, 5.7. 6.1). This points to a mechanism acting at the level of ORF2p translation, degradation or activity: elimination of ORF2 enzymatic activities causes TPRT inhibition as well as inhibition of other detrimental effects caused by ORF2p endonuclease activity. Further evidence supporting this theory can be found in the ssDNA depletion in mESCs cytoplasm and the significant depletion of the ssDNA binding proteins Rpa1 and Xrcc5 (Goodier et al., 2013; Mita et al., 2020). These factors are well-known to be involved in the DNA repair mechanisms and the resolution of LINE1 integrations into the genome upon ORF2p endonuclease cutting DNA and retrocopying its RNA into DNA. It could be suggested that it is the lack of interaction with these factors that reduces LINE1's detrimental effects. However, considering the lack of cytoplasmic DNA in mESCs compared to senescent MEFs, it is more likely that the lack of binding of these factors to LINE1 is a symptom of the absence of ORF2p activity because LINE1 cDNA is never actually produced. Moreover, if we consider the GO analysis, we can observe a significant enrichment of translational repression that could align with the context of an ORF2-related mechanism.

The GO analysis also provides some evidence towards this theory, as the term “translation inhibition” appears as significantly enriched (Figure 5.6C). The term “proteasomal ubiquitin-

related protein catabolism" has a ~0.76 enrichment factor and a p-value<0.05, implying a depletion of this term in the interactome. While gene ontology is a powerful tool in detecting upregulated gene sets. For proteins that do not need to interact with others from the same functional grouping, like ubiquitin ligases, it is possible that they are still hold a functional relevance, despite not being enriched as a group. Assuming that a ubiquitin ligase is degrading ORF2p, only one E3-ligase would be needed. This suggests that enrichment factor may not always be a useful metric for GO or answering some specific biological questions. Moreover it would not be a new mechanism to control LINE1 retrotransposition, as it has already been described in the literature, but in a different cellular context and depending on cGAS (Zhen et al., 2023).

One of the most striking observations has been the magnitude of representation of spliceosome and splicing-related categories in our LINE1 ORF1p IP. Inhibition of splicing in ESCs has been associated with totipotent-like cells that recapitulate the 2-cell stage in the mouse embryo (Shen et al., 2021). This is also the stage when LINE1 expression peaks, forming complex with Nucleolin and Kap1 that will switch off the 2-cell transcriptional programme and advance into pluripotency (Jachowicz et al., 2017; Percharde et al., 2018). LINE1 is intronless, although it can be contained within introns, thus suggesting that its interaction with splicing machinery could be attributed to the generation of multiple LINE1 chimeric transcripts in early development, something that has been shown for other transposons at this stage (Modzelewski et al., 2021) and LINE1 in somatic cells (Belancio et al., 2006; Cervantes-Ayalc et al., 2020; Marasca et al., 2022). Other authors have proposed splicing as a defence mechanism against LINE1, that would prevent the formation of functional LINE1 transcripts by splicing out relevant sequences for retrotransposition (Belancio et al., 2008). In this scenario, I want to reiterate that LINE1 ORF1p IP is a mean to obtain the LINE1 interactome. This implies that a lot of the interactors do not necessarily have a direct interaction with LINE1 ORF1p. Some hits may be interacting with direct LINE1 ORF1p interactors or the LINE1 RNA and are pulled down with the rest of the complex with LINE1 ORF1p. Therefore, the functionality of a repression mechanism mediated by any hit of this interactome, does not necessarily require the mediation of LINE1 ORF1p.

Finally, the technical limitations of proteomics must be discussed. Most of the proteome is represented by highly abundant proteins, making the detection of lowly abundant proteins like ORF2p and any other potential interactor, extremely difficult. To be observable, peptides need to digested well by trypsin (contain acidic sequences), chromatograph well, ionize well, fragment well and be present in the sample at high enough abundance. Regarding that, PROSIT, a newly developed machine learning algorithm learning on sequences of identified

peptides in publicly available databases, predicts that only 40% of peptides are actually observable (Gessulat et al., 2019). In conclusion, this work does not have to discard the possibility that despite identifying some possible LINE1 repressors, a hypothetical LINE1 master repressor in pluripotency may not have been captured.

Chapter 6: screening of novel LINE1 repressors in pluripotency.

6.1. Introduction.

By characterising the LINE1 ORF1p interactome in mESCs, I identified in Chapter 5 a set of LINE1 interactors, among them pluripotency-specific proteins, some of which might act as LINE1 repressors. Next, it was necessary to assess the capacity of these factors to repress retrotransposition by knocking them down. This repression could happen at multiple levels: post-transcriptional transcript degradation, transcript editing, RNP formation, stress granules and sequestration at P-bodies, translation repression or inhibition or degradation of ORF1p/ORF2p. Overall, these repressors could impact LINE1 transcript availability (the most commonly described mechanism), and formation of a functional LINE1 RNP. This would decrease the likelihood for LINE1 retrotransposition events. LINE1 transcript integrity is required in pluripotency and for developmental progression (Percharde et al., 2018). Moreover, because of LINE1 ORF1p is detected in pluripotency, we know that LINE1 RNA is being translated. Therefore, a mechanism based on LINE1 RNA levels is not enough to explain the lack of detrimental phenotype in mESCs in response to LINE1 expression. Thus, I hypnotise the existence of a pluripotency-specific mechanism at the level of LINE1 functional RNP formation or ORF2p protein availability.

I hence set out to test the capacity of these factors to repress LINE1 using a knockdown (KD) screen of factor identified in the LINE1 ORF1p IP. This approach assumes that the knockdown of a LINE1 repressor would cause an increase in retrotransposition events. This increase in events is the final readout, quantified with a LINE1 EGFP retrotransposition reporter system (Chapter 1.1.4). Moreover, since the sample analysis of this screening method is done with flow cytometry, it allowed me to analyse multiple samples in a short time, making it an ideal screening platform.

Nevertheless, this approach has some limitations which I had to first overcome. First, the large size of the plasmid encoding the reporter system, and second, the low transfection rate of mESCs, together resulted in only a small proportion of the cell population expressing the reporter. Third, as reported by multiple authors, the cell epigenetic machinery represses newly inserted genetic material, making the GFP readout an under-estimate of the actual number of RT events (Del Re et al., 2010; Rosser and An, 2010; Seczynska et al., 2022). Third, most of LINE1 insertion are 5' truncated. To be detected, a LINE1 insertion must be long enough to span the reporter and its promoter. Fourth, the screening is limited to the detection of a single

LINE1 copy out of half a million existing in the genome. Lastly, measurement of GFP-positive/ negative cells is a qualitative variable, being not possible to distinguish single from multiple retrotransposition events in the same cell.

At this point, it is also important to consider that, due to their function and potential role as LINE1 repressors, some of the knockdowns may compromise the viability of mESCs or cause stress that could affect LINE1 retrotransposition.

Despite this, in this chapter, I managed to develop an optimised LINE1-PFG retrotransposition reporter and screen platform in mESCs (PFG stands for the reverse writing of GFP). This has allowed me to assess the repression capacity of a set of LINE1 repressors directly in mESCs, overcoming some of the caveats that these types of essays have in ESCs.

6.2. Results

6.2.1. Generation of ROSA LINE1-PFG reporter cell lines

Most of the studies published up to date make inferences of LINE1 repressors by using 293Ts, HeLas and other environments more permissive than mESCs to retrotransposition. However, it was important for the goal of this work to assess any potential repressor in a more physiologically relevant cell type. Before screening possible LINE1 repressors, I first therefore needed to set up the retrotransposition reporter assay in ESCs. I first attempted to optimize the reporter essay in mESCs compared to 293Ts by transfecting in the LINE1-EGFP reporter plasmid, but it was not possible to observe any retrotransposition (Figure 6.1A) Next, the problems of transfection efficiency and vector size were avoided by instead generating a stable cell line with the reporter randomly inserted into the genome of mESCs. Initial trials involved linearizing Pwa125 (mouse LINE1 construct under the control of a CMV promoter) and pFG04 (mouse LINE1 construct under the control of the LINE1 5'UTR) vectors and transfecting them into mESCs, then selecting for random integration by antibiotic selection (Newkirk and An, 2017). I then assayed for retrotransposition by using flow cytometry to analyse the number of GFP positive cells. This revealed that these cell lines also did not respond to 5-azacytidine treatment, traditionally known to result in in LINE1 upregulation (data not shown). Finally, I also tested whether the knockdown of previously known LINE1 repressors like *Dicer* and *Mov10* would upregulate retrotransposition. No positive events were seen (Figure 6.1B) (Bodak et al., 2014; Choi et al., 2018).

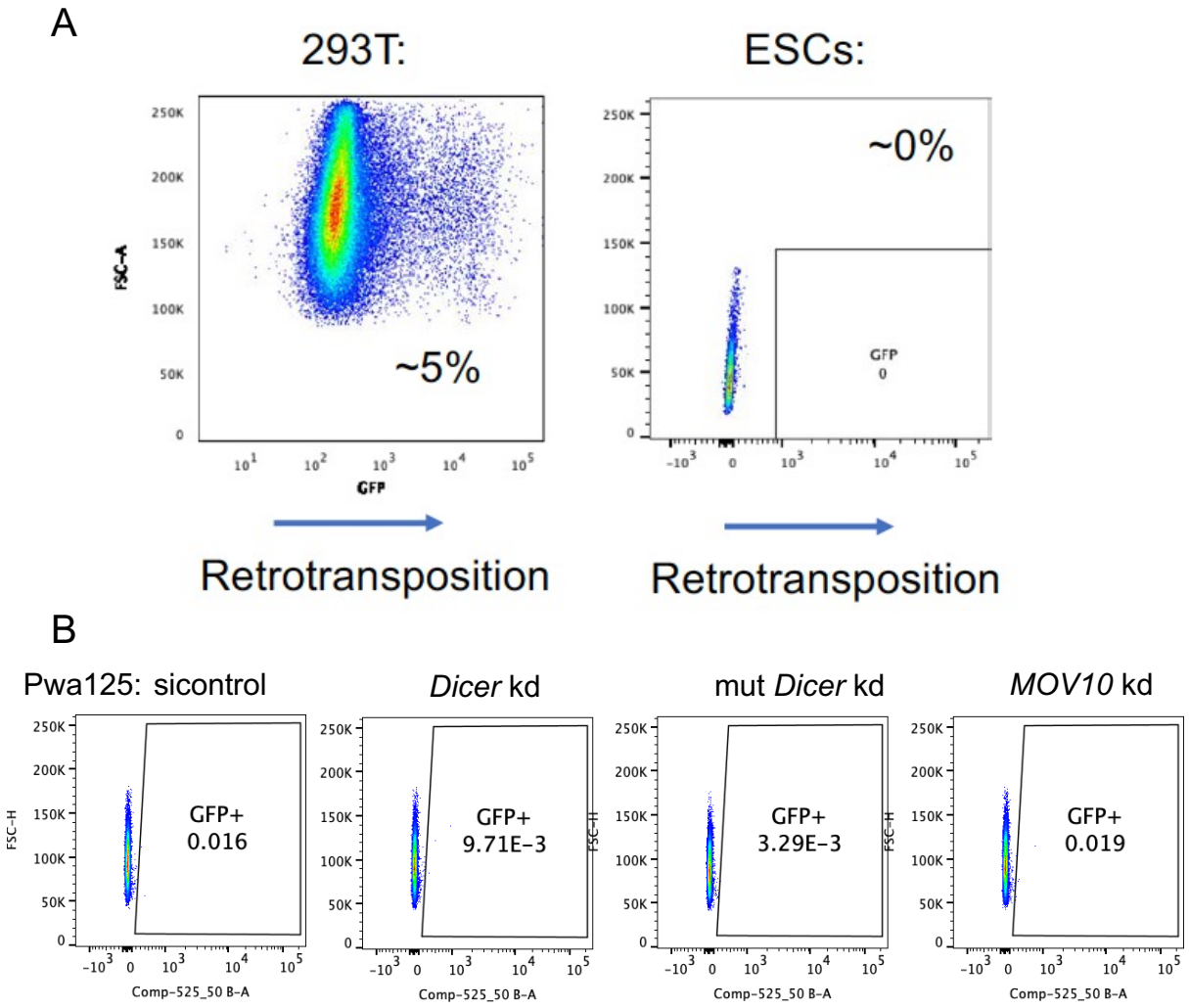


Figure 6-1. LINE1 activity seems inhibited in mESCs. (A) Percentage of GFP positive cells in mESCs and HEK 293T transfected with an smLINE1 reporter cassette. Both plots correspond to experiments done on different days. (B) Flow cytometry analysis of GFP positive cells on *Dicer* and *Mov10* KD in randomly integrated pwa125 mESCs line.

These results suggested that a randomly integrated reporter was not active enough for screening of retrotransposition in ESCs. I assumed that there may be several reasons to explain that: potential variability generated by the random integration of the construct, and silencing of the construct by the cell epigenetic machinery that overall would cause a bad expression of the construct (Seczynska et al., 2022). Moreover, the promoter driving the GFP construct expression was CMV (silenced in mESCs) and driving LINE1 – the endogenous LINE1 5' UTR (not very efficient). To circumvent the issues with randomly-integrated reporters, I designed a strategy involving the integration of smLINE1 EGFP reporter into the expression-permissive and safe-harbour locus, the *Rosa26* locus (Friedrich and Soriano, 1991). This construct was expected to allow homogeneous expression and be less susceptible to silencing by position effect and the cell epigenetic machinery, as it was driven by the CAG promoter and the ORFs were codon-optimized. In addition, the inverted CMV promoter driving the

expression of the reporter GFP was changed to an *Ef1α* promoter as CMV is inactive in mESCs (Figure 6.2A) (Cabrera et al., 2022; Herbst et al., 2012). The reporter construct was then cloned into a plasmid with *Rosa26* homology arms (Figure 6.2A) and GFP expression was confirmed by transient transfection in 293T cells (data not shown). Then, mESCs were co-transfected with the mp1052-rosa-smL1-Ef1a sense reporter plasmid, and a pX458-sgROSA-BFP plasmid that carries the *Rosa26* guide, the Cas9 and a BFP reporter (McManus, unpublished). Clones were then collected. For the detailed process, see Chapters 2.2.6 and 2.2.7.

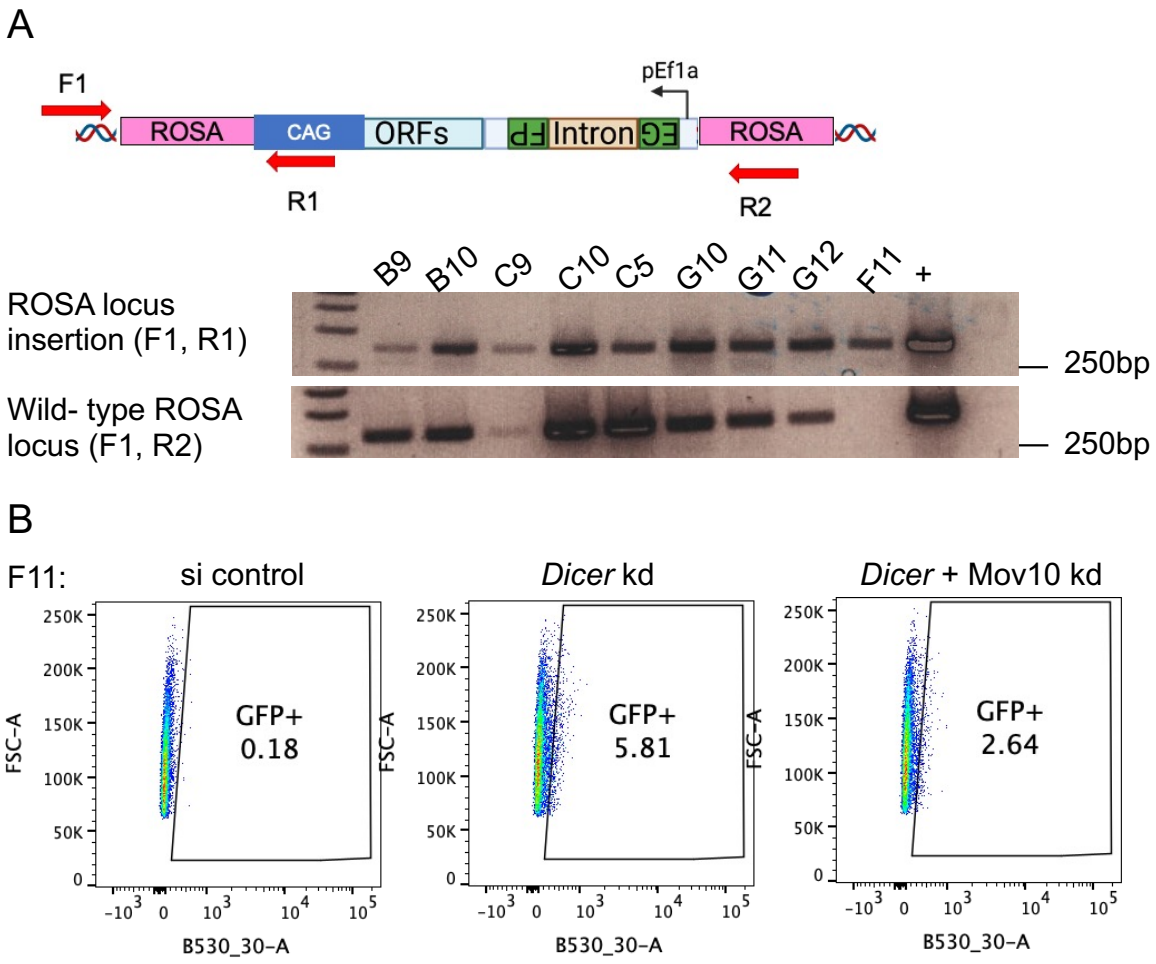


Figure 6-2: Generation of LINE1 GFP reporter cell lines. (A) Schematic of the genotyping strategy of smLINE1 GFP insertion on the *Rosa26* locus and gel electrophoresis of genotyping PCR products of some selected clones. F stands for forward and R stands for reverse. The forward primer used was RW174, while the reverse primers where RW067 for the CAG promoter and Endo-ROSA for the endogenous locus. (B) Flow cytometry analysis of GFP positive cells on *Dicer* and *Mov10* KD in ROSA smLINE1 PFG mESCs cell lines. Representative of N=3 independent experiments.

110 clones were identified, and screened for correct insertion by genomic PCR. 9 clones had the CAG promoter insertion in the *Rosa26* locus, of those 8 were heterozygous and 1 was homozygous (Figure 6.1B). Homozygosity was assessed by successful amplification in the

CAG promoter insertion PCR (F1 and R1 primers) and lack of amplification in the wild-type *Rosa26* locus PCR (F1 and R2 primers). The homozygous clone F11 was selected to continue the experiments, amplified and given the cell line name E14 R26-LINE1-PFG.

The functionality of the R26-LINE1-PFG cell line was next checked by assaying for GFP positive cells in normal culture conditions. Additionally, I performed siRNA mediated knockdown of *MOV10* and *Dicer* (Bodak et al., 2017; Choi et al., 2018). This indicated a low but detectable number of GFP-positive cells which increased by 4 days after KD (Figure 6.2B). These results together showed that I had generated a retrotransposition reporter ESC line, which could next be used to screen for potential repressors of LINE1 in mESCs.

6.2.2. FACS screen of LINE1 interactors identifies potential repressors

After successful validation of the reporter cell line, I carried out a siRNA-mediated KD screen against a set of LINE1 ORF1p interactors detected in my interactome. The candidate repressors for the KD screen were selected based on several lines of evidence. Firstly, the pluripotency-specific LINE1 interactors I identified in the ORF1p interactome in Chapter 5 (AnnexII) were included. Secondly, the specific interactors of my LINE1 ORF1p dataset in mESCs and the LINE1 ORF1p in hESCs interactome were added (Annex II) (Vuong et al., 2019). The rest of the candidates were other ORF1p interactors I selected from my dataset. They were chosen based on their high enrichment factor on the ORF1p IP versus the control IgG IP, their biological function potentially relevant for LINE1 repression (eg. Ubiquitin ligases, translation repressors, nucleases or embryonic lethal proteins), altogether resulting in a total of 74 candidate repressor proteins (Annex II).

The siRNA screen was carried out as detailed in Chapter 2, summarized in (Figure 6.3A). A total of two biological repeats were carried out, with 2 technical repeats in each of them. The Z-score was calculated using the average and the standard deviation of the siRNA controls. Results were consistent among experiments, suggesting that the KD of some LINE1 ORF1p interactors was capable of upregulating retrotransposition, notably more than *Dicer* KD (Figure 6.3B) (Bodak et al., 2014).

From the assessed genes, *Cebpz* seemed to be the most relevant candidate at this point, as it had the highest Z-score. *Cebpz* was part of the pluripotency-specific genes that were assessed to interact with LINE1 in our ORF1p interactome, like *Rbm22* and *Kif23*. Other potential repressors were ubiquitin-ligases like *Rbbp6* and *Itch*. For some KDs, a self-renewal defect was also observed, as I observed a lower number of cells in their corresponding wells

when compared to the rest of the plate. This could imply that these genes play an important role for cell viability, or that detrimental effects to self-renewal could be caused by the activation of LINE1.

6.2.3. γH2AX staining is a proxy for the identification of LINE1 repressors identification.

As mentioned before, the LINE1 reporter screen has multiple limitations, the main one being the possible silencing of most of the retrotransposition events reported in other studies (Del Re et al., 2010). Besides, due to the synthetic codon-optimized nature of our reporter (LINE1 5' UTR is substituted by the CAG promoter) it does not mimic the endogenous LINE1 elements from which the interactome was derived. The 5'UTR is an important site for LINE1 regulation, transcription factor binding and formation of secondary structures that are recognised by certain proteins (Bodak et al., 2014; Del Re et al., 2010; Heras et al., 2013). Removing the 5'UTR could be preventing some of the identified LINE1 interactors binding.

Furthermore, I also wanted to increase the capability to screen hits in a reporter-independent way. Therefore, as a complementary approach to help in the selection of targets for later confirmation, I decided to use DNA damage which might be caused by ORF2p activation as a proxy for LINE1 activity, by assaying γH2AX staining intensity following KD of repressor candidates. (Figure 6.3A). A caveat to this approach was that the half-life of histone γH2AX is 2-7h, and therefore this method detects only recent, and not cumulative, instances of retrotransposition (Bouquet et al., 2006).

I first confirmed the sensitivity of the experiment using *Rpa1* KD (a DNA repair enzyme) as a positive control, finding an increase in γH2AX in *Rpa1* KD samples (Figure 6.3C). In order to maximise the number of detectable retrotransposition events in the 7-hour window, I optimised the maximum cell number capacity of the assay. To do that, I transfected *Rpa1* siRNA into different cell number suspensions and evaluated the percentage of the population with a γH2AX fluorescent staining over 10,000AU (Figure 6.3C). 10,000 was a number derived when comparing the γH2AX staining intensity per cell distribution in the siRNA control versus *Rpa1* KD populations in the screen samples. These data showed ≅70% of the Rpa1 KD cell population being above a 10,000 intensity when the starting population was 2,000 cells. With 4,000 cells, this percentage dropped to ≅50%. I considered these values a proxy of transfection efficiency, and decided to proceed with transfection of 2,000 cells per target siRNA for the full-scale screen.

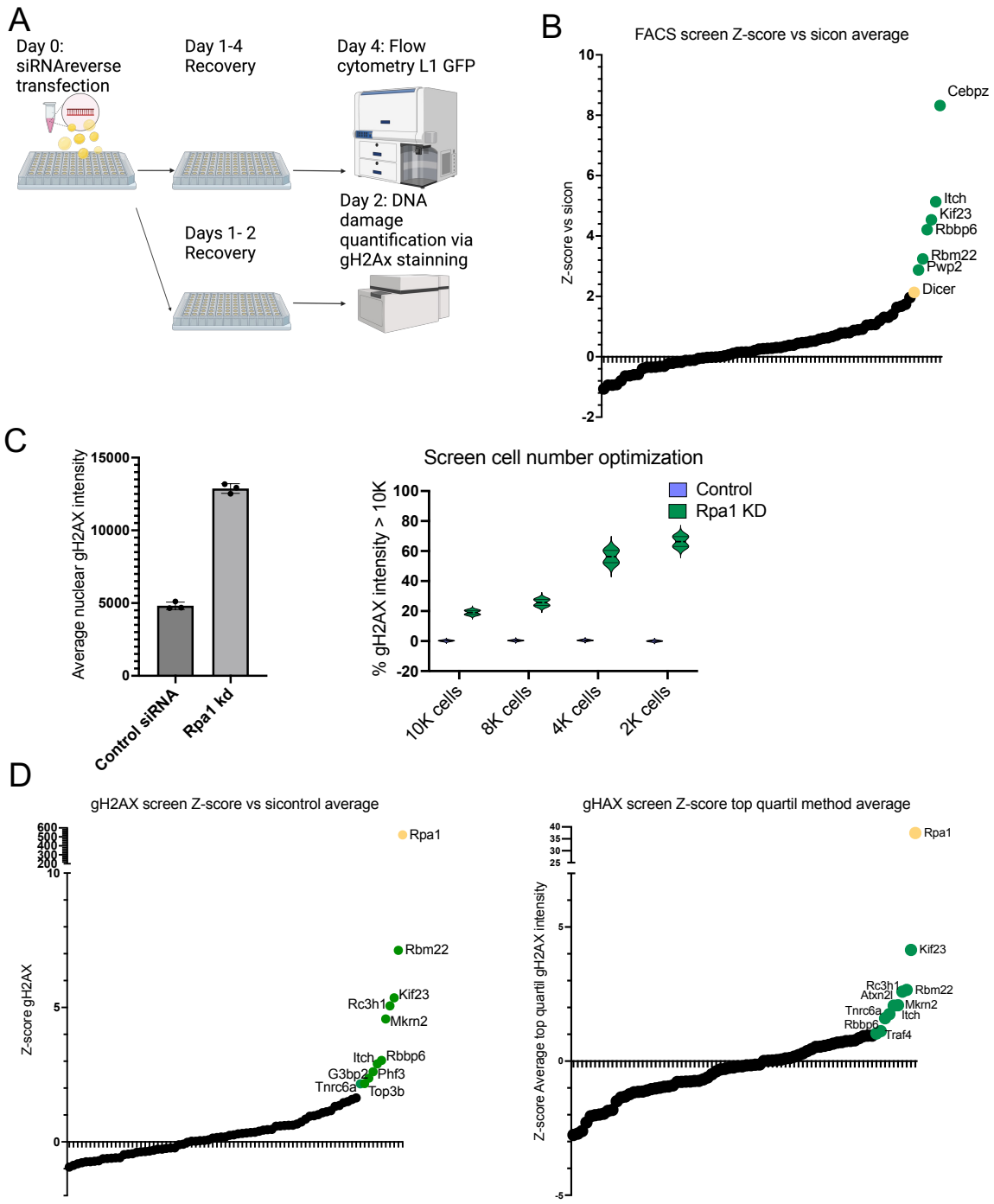


Figure 6-3: GFP LINE1 reporter and γH2AX screens were capable of identifying potential LINE1 repressors. (A) Schematic of screening strategy. (B) Z-score of smLINE1-PFG reporter flow cytometry screen. N=2 for two biological repeats, 2 technical replicates each. (C) Cell number optimization for transfection using γH2AX staining intensity in cells transfected with siRNA control and Rpa1 (DNA repair factor), n=2 for two biological repeats. (D) Z-score of γH2AX screen with two different analysis methods: % of the population above nuclear fluorescent intensity above 10,000AU and average of fluorescent intensity of the top quartile. N=2 for two biological repeats, 2 technical replicates.

Two biological replicates of the γH2AX screen were carried out with 2 technical replicates in each of them. Two types of analysis were done: the percentage of the population with a staining intensity above 10,000 AU and the average of the top quartile, which raised similar results (Figure 6.3D). As retrotransposition is a rare event, it was expected that increased retrotransposition would result in a minor shift in the upper end of the γH2AX distribution. Z-scores were calculated based on these values using the average of the siRNA control and standard deviation as reference (Figure 6.3D).

Finally, I correlated the Z-scores of the experiment averages of FACS and γH2AX screens, in order to see which hits increased both in LINE1-PFG and DNA damage. I observed that the analysis of the top candidates from both screens revealed a substantial overlap between both approaches, except for *Cebpz* which showed a drastically different behaviour in both screens (Figure 6.4A).

The top overlapping candidates of both experiments (*Rbm22, Rbbp6, Kif23, Itch, Rc3h1*) were selected for further investigation. Candidates like Cebpz, Mkrn1/2 and Tnrc6a were also considered after performing well in one of the two screens. Some of these genes have paralogous genes also present in the screen, which led to trying double KDs to explore potential redundancies. The LINE1-PFG reporter screen was repeated again with just these selected candidates including several technical replicates and an adjustment in the FACS buffer that allowed for better survival and less clumping. Results confirmed the previous two screens overlapping best hits (*Rbm22, Kif23, Rbbp6* and *Itch*) and *Cebpz* as the top hit (Figure 6.2F-G). For other included hits, no evidence of redundancy was found (Figure 6.4B). Out of these, the best performing candidates were *Cebpz* and *Rbbp6.* Summary graphs are presented with examples of the FACS plots for some candidates (Figure 6.4C).

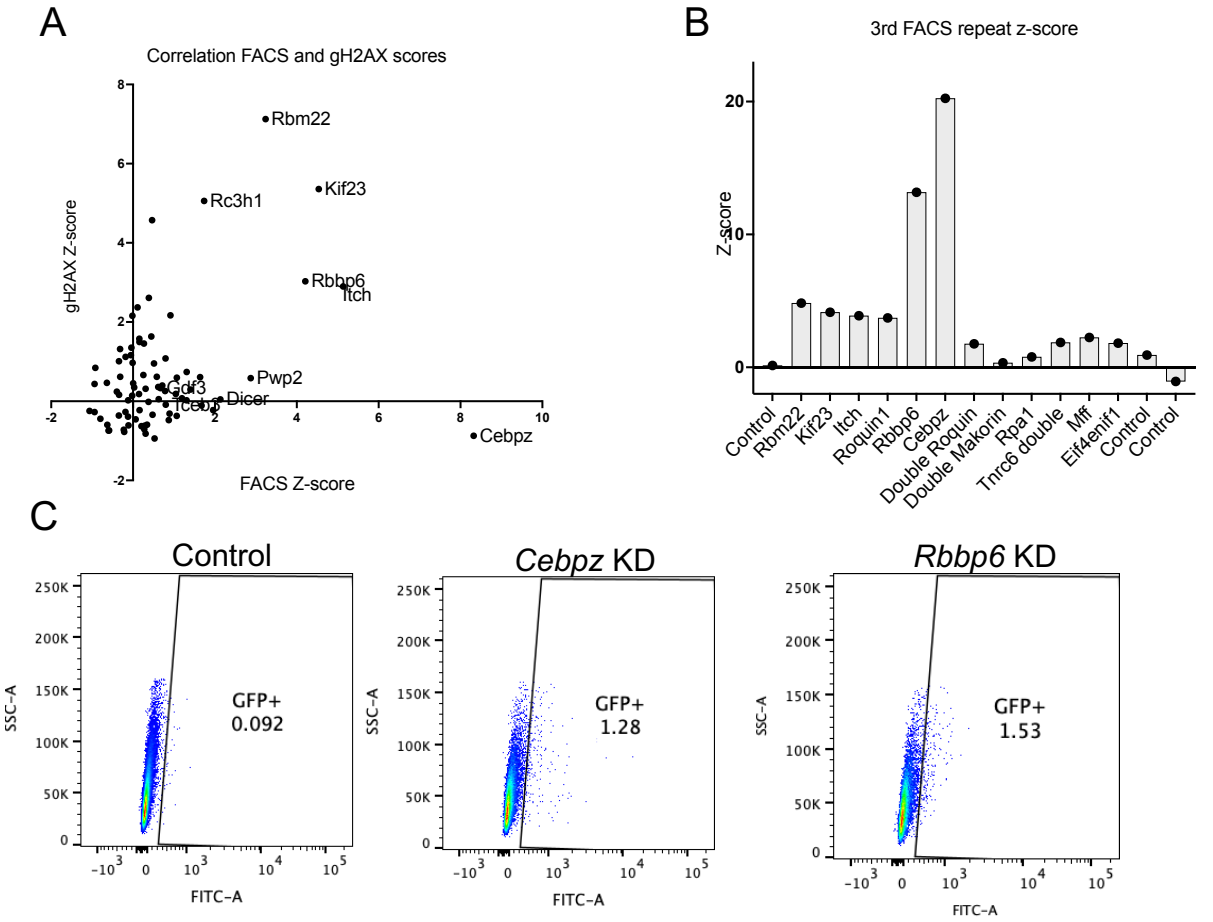


Figure 6-4: Selection of candidates to continue evaluation.(A) Correlation of Z-scores between both screens. (B) Z-score of LINE1-PFG cells upon KD of selected targets, comprising at least 3 technical replicates. (C) Flow cytometry analysis of GFP positive cells on siRNA control and *Cebpz* KD in R26-LINE1-PFG mESCs.

6.2.4. New LINE1 repressors in pluripotency are confirmed by FACS and genomic DNA qPCR.

Before proceeding to more specific functional essays that would involve the creation of stable knock-out cell lines, I still needed to reduce the list of candidates and get more robust evidence of the potential of my candidates to repress LINE1. To circumvent the possible repression of the reporter documented in previous assays (Del Re et al., 2010), I took the more direct approach to quantify the number of inserted sense copies (spliced) of the LINE1 GFP reporter by gDNA qPCR (Figure 6.5A) (Del Re et al., 2010; Newkirk et al., 2017; Rosser and An, 2010). This allowed me to circumvent any epigenetic event that could silence the reporter.

Before the optimization of the regression curves that allowed to quantify the exact number of LINE1 retrotransposition events, a pilot KD was attempted, analysing the data by normalizing

them to chromosome 11 and the siRNA control (Figure 6.5B, C). Preliminary results showed an increase of retrotransposition in *Rbm22*, *Rbbp6*, *Cebpz* and *Kif23* KDs. The same experiment was repeated for *Rbm22*, *Cebpz*, *Rc3h1* and *Rbbp6* with the optimized regression curves and using a new batch of siRNAs (Figure 6.5B, D). The control wells were stable and consistently showed a very low presence of GFP (equating to less than 1 GFP copy per 100 cells). However as also seen in the flow cytometry experiments, the KDs that led to the upregulation of retrotransposition tended to show a lot of variability.

Unfortunately, upon further repeats with new siRNAs and the same LINE1-PFG ESCs, I failed to show any increase in GFP copies upon KD of my candidate repressors, apart from for *Cebpz* (Figure 6.5F data not shown). A deeper examination of qPCR melting curves showed the appearance of a second peak in the melt curve of the EGFP qPCR for control wells and other target siRNA wells, indicating a secondary non-specific amplification (Figure 6.5F). This second product may therefore be obscuring any true increase in GFP copies upon candidate KD.

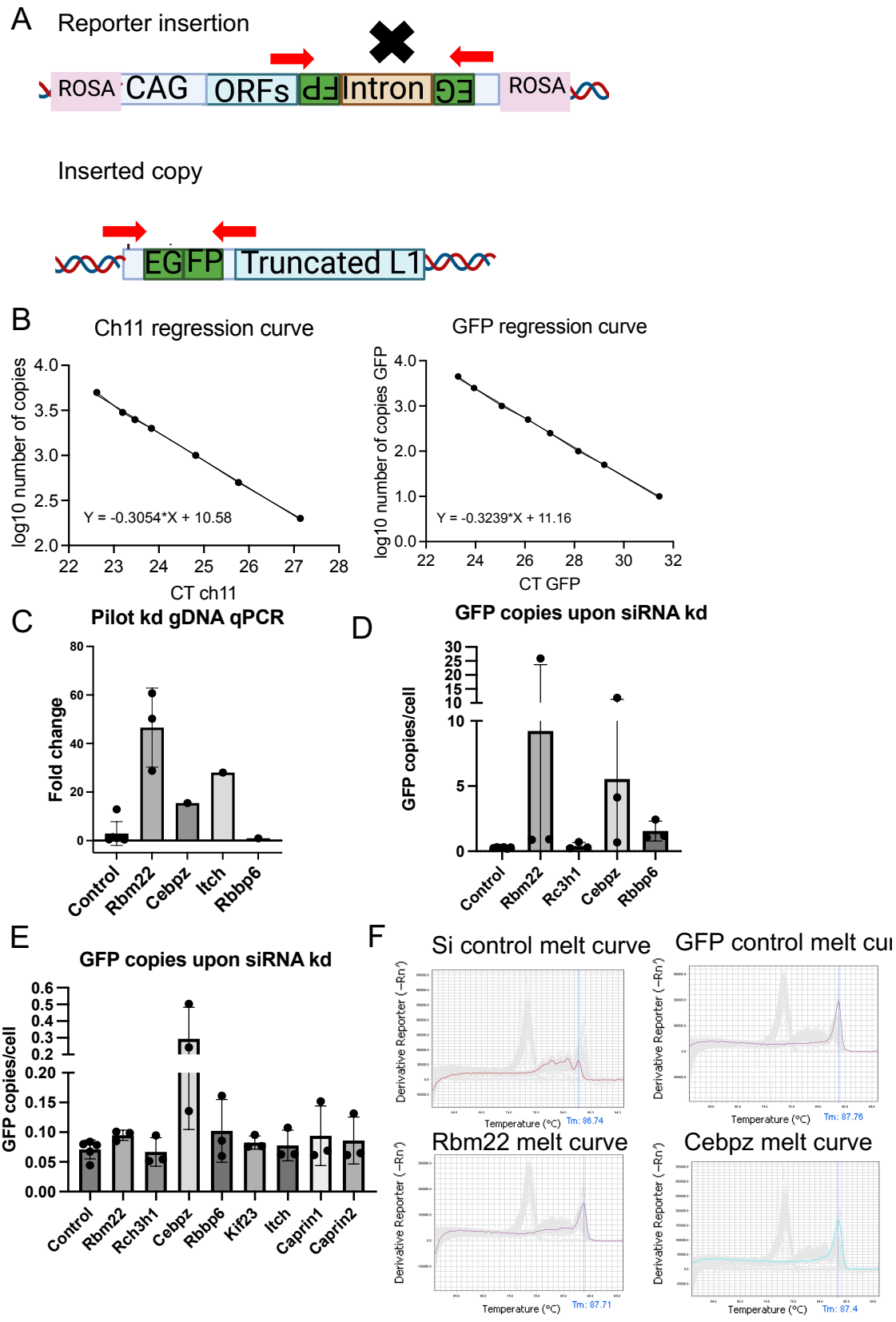


Figure 6-5: Detection of LINE1 GFP reporter retrotransposition events by genomic DNA qPCR.(A) Schematics of the GFP genomic DNA qPCR rationale. The intron does not allow amplification of the unspliced original GFP insertion. However, in case of *de novo* retrotransposition-derived insertion, the intron has been previously spliced and GFP can be

amplified. (B) Regression curves for chromosome 11 and GFP. (C) Pilot gDNA qPCR in KD reporter mESCs samples. Data analysed as 2^(-ddCT) normalizing with chromosome 11. (D) Quantification of the ratio of inserted copies of GFP per cell in KD reporter samples. Melt curves were showed only one peak for all samples. (E) Quantification of the ratio of inserted copies of GFP per cell in KD reporter samples. Melt curves showed unspecific amplification for siRNA control and some other samples. (F) Melt curves of some selected targets from repeats not showing any retrotransposition upregulation.

6.3. Discussion

In this chapter, I established a stable LINE1 retrotransposition reporter screening platform in mESCs. This platform has allowed the identification of potential LINE1 repressors among the ORF1p interactors previously described. This has been achieved through an R26-LINE1-PFG retrotransposition reporter screen and a DNA damage screen in mESCs that, with some exceptions, have been mutually consistent. The documented frequent silencing of most of the reporter insertions led me to suspect that the flow cytometry screen under-detected the real scope of retrotransposition (Del Re et al., 2010; Seczynska et al., 2022). This led me to explore different methods to validate retrotransposition – namely GFP qPCR. Preliminary results suggest that this method is able to accurately quantify retrotransposition, and seemed to highlight the importance of *Cebpz*, *Rbbp6* and *Rbm22* as potential LINE1 interactors. Nevertheless, the technical difficulties experienced in the end prevent me from making any reliable conclusions. Solving these technical problems would require extra optimization, which due to time constraints, was beyond the scope of this project. Furthermore, the disagreement between screening techniques for *Cebpz* may challenge the rationale of the γH2AX screen, as it may not be powerful enough to detect the increase of DNA damage caused by LINE1 activation.

Cebpz is a transcriptional activator, is upregulated in pluripotency and is associated with CCAAT elements binding, and involved in a stress response that involves heat shock factors (Lum et al., 1990). The way *Cebpz* may be directly repressing LINE1 is still intriguing given its known functionality. It could relate to an unknown function or depend on its interactors within LINE1 RNP.

Rbm22 is another pluripotency-upregulated Zn^{2+} finger protein. It has been related to pre-mRNA splicing by interacting with U6 snRNA and the pre-mRNA splice site, and also seems related to the stress response and gene regulation (Janowicz et al., 2011; Zhang et al., 2017). Regarding Rbm22's functionality in splicing, it is still unknown how splicing could relate to LINE1 repression. It could be assumed that by inhibiting alternative splicing (in which it is involved (Jiménez-Vacas et al., 2023; Zhang et al., 2017), more canonical LINE1 transcripts

would be formed, increasing retrotransposition. However, this mechanism does not make sense for the R26-LINE1-PFG reporter. As with *Cebpz,* a novel function of the protein could also explain LINE1 repression. In addition, its interactors within LINE1 RNP are worth exploring to have a larger idea of this potential repression mechanism.

Finally, Rbbp6 is also a Zn^{2+} finger protein and a E3 ubiquitin ligase (Chibi et al., 2008). It also seems to be involved in DNA replication and regulation of chromosomal fragile sites through its E3 ligase activity (Miotto et al., 2014). As a ubiquitin ligase, its repression mechanism may involve the degradation of LINE1 proteins (ORF1p or ORF2p) or other components of the LINE1 RNP that act as LINE1 activators.

However, it is also worth considering that the good performance of ubiquitin ligases like *Rbbp6*, *Rc3h1* and *Itch* in the retrotransposition reporter screen may be artefactual, as a result of EGFP stabilization, rather than an increase in retrotransposition. Moreover, these factors may play other roles related to LINE1 by targeting other proteins. Finally, the repressor potential of some candidates may have been overlooked, as cells are normally grown for longer to let events accumulate. For instance, Bodak et al, let their *Dicer* KO cell lines grow for 3 weeks before analysing, and other published assays relying on reporter transfection also involve longer time periods (Bodak et al., 2017; Garcia-Cañadas et al., 2023).

Further experiments will involve deeper investigation validation of the candidate repressors and their mechanism of action. This would especially relate to the capacity of knockdowns or knockouts of these factors to generate cytoplasmic DNA, which would crucially link both the DNA sensing and LINE1 repression sides of this project. Other validation strategies would include a LEAP assay: a LINE1 RNP purification assay designed to assess the retro-transcriptase capacity of LINE1 RNPs (Kopera et al., 2016). Other experiments may involve using HEK 293T cells, which are a much more permissive platform to retrotransposition. These experiments will involve co-expressing the LINE1-PFG reporter construct together with the potential repressor and assessing if it can repress retrotransposition.

To test the hypothesis that some repressor(s) may function by inhibiting ORF2p, I will use a LINE1-ORF2-cherry reporter construct. Co-expression of the ORF2p reporter with an ORF2p repressor should reduce the cherry signal. Conversely in mESCs, I would expect the ORF2 cherry signal to increase upon the KD of the repressor.

At this point it is important to acknowledge the limitations of siRNA as a system to assess the functionality of a LINE1 repressor. siRNA activity is transient, it may not be complete and it

takes some time to take effect. Moreover, siRNAs may have off targets and/or cause toxicity by interfering with the cells iRNA machinery, tittering out endogenous siRNAs and miRNAs. That is why, further down the line I aim to use a degradation TAG 13 (dTAG-13 or degron) system with a reduced pool of candidates, probably *Rbm22*, *Cebpz* and/or *Rbbp6*. The Auxin/degron system would allow for a finer-tuned, rapid, reversible and complete degradation of the target protein. In addition, it would allow suppression of the protein for longer periods of time (Nabet et al., 2018; Yesbolatova et al., 2020).

In addition, other downsides of my approach would include potential confounding effects in the γH2AX screen, if the putative repressor protein is directly or indirectly involved in DNA damage repair.

As a final consideration, it is also possible that the interaction of a protein with LINE1 has a function unrelated to repression that is relevant for stem cell biology (eg. Nucleolin and Kap1) (Percharde et al., 2018). Elimination of this interaction could compromise the viability of the cell. However, these types of functions fall beyond the scope of this project.

Chapter 7: Discussion

At the beginning of this project, I set out to investigate the lack of a detrimental response of ESCs and embryos to LINE1 expression. To achieve this, I developed two approaches that could explain the lack of responsiveness: on the one hand I investigated the status of the cellular immune system in pluripotency, particularly the DNA sensing machinery, and the cellular response to IFN. On the other hand, I wanted to test the hypothesis that a pluripotency-specific factor was inhibiting LINE1 activity in ESCs. To asses that I wanted to characterise LINE1 interactome in pluripotency and evaluate if any of its interactors could act as a LINE1 repressor.

In this work, I demonstrated the inactivation of the DNA sensing pathway in mESCs, the main pathway that is believed to detect TE-derived products of retrotransposition. This is consistent with the second discovery that two key components of the pathway: *cGAS* (the sensor) and *STING* (the adapter) are not expressed at the protein level in pluripotency. The expression of these proteins is still detected at mRNA level, although at a lower level than in their somatic counterparts. Meanwhile, the rest of the components of the cGAS/STING pathway seem to be expressed at protein level, although again downregulated somewhat when compared to their somatic counterparts. This downregulation is also observed at the RNA level. The mechanism for the lack of protein expression of *cGAS* and *STING* is still unknown and should be followed up in the future. It probably relies on a post-transcriptional mechanism, although chromatin effects that reduce mRNA expression should not be discarded. For what we know from the literature, this mechanism is independent of miRNA targeting, as knock out of the miRNA processing pathway does not seem to rescue a response to G3-YSD DNA ligands as it does for RNA ligand poly(I:C) sensing (Witteveldt et al., 2019). The inhibition of TE DNA sensing seems to be highly redundant in pluripotency, as I showed that rescue of STING with the expression of a gain of function (GOF) variant is not enough to rescue the pathway's functionality. However, I found that expression of GOF mutant of *Irf3* is capable of activating the antiviral response and the synthesis of IFN and other cytokines. This indicates that the pathway seems to be functional downstream of STING signalling.

STING activation is known to be a highly regulated process including by PTMs and compartment shuttling. I therefore decided to investigate more about the ER:Golgi dynamics of activated STING in pluripotency. My experiments allowed me to confirm that GOF STING is indeed capable of reaching the Golgi, while the loss of function HAQ mutant seems to be retained in the ER. Additionally, analysis of phospho-Tbk1, showed that both variants seemed

to be unable to activate Tbk1. Therefore, I suggest that an additional layer of regulation occurs at a point between STING arrival at the Golgi and the formation of the active STING signalosome in the TGN. Additional evidence in support of this is provided by the fact that GOF mutation of *Irf3* is capable of activating IFN synthesis and activating the antiviral response. Given this situation, future experiments aim to disclose the position of the STING GOF variant in the Golgi, as well as investigating if palmitoylation, an essential PTM for STING signalosome clustering at TGN and activation of downstream signalling is occurring in pluripotency (Mukai et al., 2016). Other approaches may involve the characterization of STING GOF variants interactomes in pluripotency and in a somatic context when STING is functional, by IP mass-spectrometry. Surprisingly, according to my data, both STING variants seem to be phosphorylated with the active phosphorylation mark at serine 357. This is contradictive with the current knowledge of this modification being dependent of Tbk1 and occurring in the TGN (Zhang et al., 2019). Phosphorylation in other residues of STING can have different effects rather than activation, and it may be possible that I am seeing an off-target effect of the antibody. However, this discovery does not fit with the current knowledge of STING activation dynamics.

STING target inactivation is understood in the context of preventing IRF3 activation and further downstream signalling that could compromise the viability of the embryo at a very early developmental stage. Being STING activation a highly regulated process, its post-transcriptional inactivation could occur at multiple levels: ER:Golgi transport, retrograde transport, STING protein degradation, etc. This can establish a parallel to how STING is inactivated in multiple cancers as a way as an immune evasion mechanism (Zhu et al., 2019). This inactivation does not necessitate the accumulation of mutations, as it has been seen in multiple instances that the repression of STING in cancer is epigenetic (Falahat et al., 2021; Konno et al., 2018). Moreover, some viral oncogenes can also interact with STING and inhibit it (Lau et al., 2015). For this reason, in the recent years, therapies aiming at the cGAS/STING pathway activation in cancer cells have become an increasing focus of research (Samson and Ablasser, 2022; Schreiber et al., 2011).

A key question raised by these findings is why cGAS/STING needs to be so strongly inhibited in pluripotency. I have hypothesised that this is because pluripotent cells of the epiblast need to prevent their highly expressed TEs from triggering the innate immune system and the production of IFNβ. It can however be argued that rescuing the capacity to produce IFN in the embryo is actually irrelevant, since as I and others have shown, pluripotent cells are insensitive to IFNβ. For example, my work shows that there is no upregulation of canonical/ non-

pluripotency ISGs upon exposing mESCs to IFNβ, except for a subset of intrinsically, presumable non-toxic, ISGs expressed in pluripotency such as those reported by Wu et al (Wu et al., 2018). The novelty of my analysis is that it discriminates between these genes and the canonical ISGs, such as *Tnf* and *Il6* (Peteranderl and Herold, 2017; Wu et al., 2018). Considering that STAT1 can be activated, the reason for this phenotype is uncertain. It could be attributed to the lower levels of STAT1 in pluripotency, but also to the potential chromatin state of the binding sites of STAT1 in the promoters of ISGs. The pluripotency programme TFs could be binding and repressing the STAT1 binding site of some, but not other, ISG promoters. This could also explain why pluripotency intrinsic ISGs are more sensitive to STAT1 activation, as since they need to be expressed, the chromatin conformation of their promoters remains more open.

The fact that ESCs are largely insensitive to IFNβ would suggest that there is no need to inactivate cGAS/STING and TE sensing. However, perhaps the mistake made until now has been to attribute the whole detrimental potential of the activation of the DNA sensing pathway to effects downstream of IFN signalling. My results reveal that activation of IRF3 as a direct consequence of STING activation, apart from activating the expression of IFNβ, can also activate at least a substantial subset of the antiviral response and cause a drastic self-renewal defect. Transcriptomic analysis of IRF3-5D expression in mESCs revealed a rescue of the innate immune response in pluripotency. These genes include upregulation of pro-apoptotic factors like *Xaf1* and *Caspase 3* and *7*, chemokines like *Ccl5* and *Isg15* together and genes inhibiting viral replication like *Mx2* and *Ifit2*. This suggests that the detrimental potential of the cGAS/STING pathway goes beyond the production of IFNβ and the subsequent autocrine IFN downstream signalling. Moreover, activation of the innate immune response seems incompatible with the pluripotency programme in mouse. This was suggested experiments in human by the OE of IRF7 in hESCs resulting in a mild dysregulation of some pluripotency markers that did not result in the loss of pluripotency, and a dysregulated differentiation potential (Eggenberger et al., 2019). Future, *in vitro* experiments will involve using a stable expressing system of STING and IRF3 GOF constructs in mESCs and hESCs, that allows to assess the RNA and protein expression of pluripotency markers, upon an antiviral response.

A completely new discovery that agrees with the data published in somatic cells, is that sustained IRF3 activation leads to the upregulation of apoptotic pathway-related genes. This activation is consistent with the self-renewal defect that I observed in the mESC population expressing IRF3-5D. However, I cannot still confirm the source of this phenotype. Indeed, activation of apoptosis by IRF3 can be transcription dependent or independent. I have seen

transcriptional upregulation of some apoptotic genes, but a transcription-independent mechanism cannot yet be discarded. Thus, which of the pathways may be activated in this context remains to be elucidated. Taken together, my data point to the importance of inhibiting the DNA sensing arm of the innate immune system in order to prevent the activation of IRF3 target genes, which would ultimately lead to loss of pluripotency or viability. This is therefore independent of downstream IFN-I signalling, which as others have shown and I have confirmed, is inhibited in development.

Overall, my results add to the evidence that TEs have created an evolutionary pressure in early development, which has shaped evolution towards the inhibition of the nucleic sensing pathways and IFN sensing. The main reason for this is that, in contrast with TE expression in the germline, the pre-implantation embryo needs to co-exist with expressed TEs which have been co-opted for essential roles (Jachowicz et al., 2017; Percharde et al., 2018), while at the same time avoiding the detrimental effects that the activation of the antiviral pathways could have at such an early stage. With the inhibition of the nucleic acid sensing pathways, evolution is inhibiting type I IFN production, although perhaps the main objective is preventing the activation of other IRF3 targets. This is consistent with the targeted NFκB repression in pluripotency by Nanog and miRNAs, and with the apparent incompatibility between innate immunity and pluripotency (Eggenberger et al., 2019; Lüningschrör et al., 2012; Torres and Watt, 2008).

One limitation of my work is that all the experiments were performed in an *in vitro* set up and only focused on a subpopulation of the embryo, the epiblast. In the *in vivo* context in which natural selection acts, an IFN-producing embryo in response to TE activation could generate high local concentrations of IFNβ that could activate the antiviral response in other cell types that may not have a blunted IFN-I response. This could include either the trophectoderm or also the maternal endometrium, two tissues that possess a functional IFN signalling system (Aikawa et al., 2014; Fendereski et al., 2022). In fact, the trophectoderm itself produces several cytokines like Il6, LIF and IL1 that are implicated in the implantation process with maternal uterine NK cells, macrophages and dendric cells (Yockey and Iwasaki, 2018). Actually, in ungulates the trophectoderm synthesises IFN-tau, a type-I IFN necessary for implantation that acts in an autocrine and paracrine manner (Roberts, 1996), with implantation actually a process highly regulated by the immune system. Therefore, embryonic IFN secretion could create a local inflammation that interferes in the implantation by disrupting the cytokine and immune cell type balance, mobilizing and recruiting the maternal immune system to target the embryo, or produce an aseptic autoinflammation that damages the embryo. In

fact, this could be achieved solely by the activation of IRF3, as it seems to trigger the expression of several chemotactic substances (*Cxcl10, Cxcl11, Isg15* and *Ccl5*) that activate NK and phagocytic cells and mobilize lymphocytes (Araujo et al., 2018; Gao et al., 2019; Iglesias-Guimarais et al., 2020; Seo et al., 2020). Therefore, future experiments will aim to assess the consequences of IRF3 activation *in vivo,* as a proxy of cGAS/STING activation, as well as the effects of IFN in the embryo implantation and pregnancy. For this reason, I am planning experiments involving IRF3-5D construct injection into zygotes and morula aggregation experiments with IRF3 expressing mESCs. Interestingly, it is in the *in vivo* context where the expression of cGAS earns relevance beyond a mere redundancy with STING inhibition. RNA sensors are expressed in pluripotency, however cGAS repression is necessary to prevent the generation of 2'3'cGAMP intermediates that could be exported to the trophectoderm in vesicles and activate cGAS/STING there, or even interact with the maternal somatic tissues at peri-implantation (Ablasser et al., 2013; Carozza et al., 2020; Desrochers et al., 2016)

The blastocyst attachment and invasion of the maternal endometrium is mediated by an evolutionary conserved inflammatory response that includes Il6, Il1, LIF and several immune cell types like NK cells, macrophages and dendritic cells (Yockey and Iwasaki, 2018). The issue of how the trophectoderm is capable of upregulating inflammatory mediators is itself highly intriguing. This at the same time has to be reconciled with the fact that LINE1 RNA is detected in the trophectoderm (Percharde et al., 2018). The origin of this RNA, as well as its implication is uncertain, as it could potentially activate IRF3 in the TE. For this reason, future experiments should test the sensitivity of TSCs to nucleic acid intermediates, as well as addressing the levels of expression of LINE1, and if applicable, its potential to generate ssDNA.

Considering the evidence presented above, it is possible to speculate that the inhibition of the DNA sensing pathway is something conserved in any organism that experiences a considerable transposition competent TE activation in its early phases of development. I have demonstrated how activation of IRF3, which would be the end result of the activation of the cGAS/STING pathway, is capable of compromising the self-renewal of pluripotent cells, just considering the own cell- intrinsic mechanism. This could be itself enough to consider TEs can exercise an evolutionary pressure to shut down DNA sensing in early development. However, it can be presumed that this pressure can have additional sources in eutherian animals like mouse and human. Considering these animals development involves a process of sustained implantation, invasion of the maternal endometrium and dependency on the mother's constant nutritional supply. It can be speculated that the immune signalling caused by the activation of

cGAS/STING, could recruit by paracrine signalling the maternal immune system into attacking the developing embryo from the maternal endometrium. Moreover, the trophectoderm is an IFN-sensitive tissue vital in the immune system-regulated implantation process (Fendereski et al., 2022; Yockey and Iwasaki, 2018). Hence it can be speculated that the production of IFN and other cytokines by the ICM could cause an immune response in the trophectoderm that disrupts the implantation process. This has to be reconciled with the fact that the trophectoderms of many eutherians are known to secrete inflammatory mediators necessary for implantation (Fendereski et al., 2022; Yockey and Iwasaki, 2018). In conclusion, apart from the cell-intrinsic consequences of cGAS/STING activation for the embryo, an additional source of pressure can be presumed in eutherian mammals. This source of pressure would be inexistent in animals in which fertilization takes place externally, but its hypothetical existence in non-placental animals with internal fertilization is beyond speculation.

Another related question is when DNA sensing becomes active again. This not only involves knowing when the embryo starts expressing cGAS and STING proteins. But also, at what timepoint the pathway is active to produce type IFN in response to aberrant cytoplasmatic DNA species generated by TEs. Moreover it is indispensable to comprehend at what timepoint of development this IFNs can activate the anti-viral programme, something that has been achieved in differentiation experiments of mESCs to fibroblast (D'Angelo et al., 2016). For my differentiation experiments, I know that the recovery of the innate immune sensing is not immediate after the loss of pluripotency, as mESCs passively differentiated for 6 days are still not expressing cGAS and STING. Elucidating this, will therefore require more established differentiation protocols like EBs and assessing the expression of cGAS and STING in embryos at different developmental points. Considering MEFs can sense nucleic acids, it can be speculated that this capacity is developed before E13.5.

As a technical note, it is important to consider that the methodology for this work has been highly reliant on transfection of *STING* and *Irf3* expression constructs, without being able to achieve inducible transgene expression of the proteins in stable mESC cell lines. The reasons for this are still unknown, however I suspect that it could be potentially be caused by the upstream expression of BFP-P2A in the constructs. This has increased the amount of time to perform experiments, and has reduced and limited the type of experiments we could carry out due to low numbers of transfected cells. For future experiments aiming at discerning STING dynamics and of the consequences of IRF3 activation, it will be necessary to develop stable mESC cell lines conditionally expressing *STING* variants and *Irf3*-5D.

One ideal aim of this project would be to restore cGAS/STING signalling in ESCs, and then examine the effects of LINE1 or other TEs expression on this pathway. However, my own data here also suggests that LINE1 TPRT is highly repressed in pluripotency, when compared to pathogenic contexts like senescence and cancer. Previous studies show that LINE1 retrotransposition levels are lower than what it would be expected in pluripotency and early development (Arokium et al., 2014; Ewing and Kazazian, 2010; Garcia-Perez et al., 2007; Kano et al., 2009). I also confirmed that in my reporter system and with novelty, I specifically showed that despite the high ORF1 expression, ESCs do not exhibit ssDNA intermediates in their nucleus or cytoplasm. This implies the existence of a second mechanism in ESCs to prevent TE-driven pathology, the inhibition of LINE1 activity and TPRT. This makes sense if we consider that inhibition of the cGAS/STING pathway would not mitigate the effects of ORF2p endonuclease activity. This must be a particularly strong mechanism that allows the cell to tolerate LINE1, since according to our evidence, LINE1 expression (using ORF1p as a proxy) is overwhelmingly higher in pluripotency than in senescence. Despite that, I could not find expression of ssDNA in pluripotent cells, while in senescence it was possible to locate ssDNA foci, which correlated with the upregulation of the IFN response. Moreover, proteins found in somatic LINE1 interactomes like the ssDNA binding protein Rpa1, involved in LINE1 retrocopy integration were significantly more enriched in the control IP, while PCNA, Xrcc5 and Xrcc6 are below level of significance (Goodier et al., 2013; Taylor et al., n.d.). Some mechanisms suggested by authors to reduce LINE1 retrotransposition rely on reducing the availability of LINE1 RNA in pluripotency (Bodak et al., 2017). Nevertheless, these mechanisms are not enough to explain this phenotype, as LINE1 basal expression in pluripotency is already higher than in pathogenic contexts. It could be argued that LINE1 RNA is trapped into the nucleus. However, the rapid cell division rate, which means a disassembly of the nuclear membrane, and the fact that ORF1p is being translated and in the cytoplasm, are indicative that LINE1 RNA can make it into the cytoplasm, and LINE1 replication intermediates could therefore also leak into the cytoplasm. Moreover, our interactome is enriched in several proteins involved in RNA export and RNA export is a category in the GO analysis. Moreover, we know that LINE1 RNA is being imported into the nucleus as several ORF1p interactors like NurD factors, Nucleolin or Cebpz are nuclear proteins and the LINE1 RNA FISH shows that LINE1 RNA location is nuclear (Percharde et al., 2018).

Evidence of TPRT inhibition in pluripotency lead me to investigate a mechanism acting at the level of TPRT/ORF2 inhibition, a mechanism that has been proven to exist in a somatic context being carried out by BRCA1 which inhibits ORF2p translation, Trim41 E3-ligase targeting ORF2p for degradation and, in human, the miRNA let-7 (Mita et al., 2020; Tristán-Ramos et al., 2020; Zhen et al., 2023). However, the lack of ssDNA in ESCs begs the question of how

LINE1 has displayed any evolutionary pressure to repress DNA sensing or vice-versa. We know that LINE1 is capable of retrotransposing in mESCs, but the basal levels are so low that is complicated to see it, proving that at least a reduced quantity of LINE1 ssDNA is generated, and could still be targeted by IFIT16 and DNA-PK, hnRNPA2b1 nuclear sensors (Almine et al., 2017; Burleigh et al., 2020; Dunphy et al., 2018; Wang et al., 2019). Moreover, dysregulated LINE1 ORF2p endonuclease activity would represent a genome danger for the cell. Finally, we cannot discern what evolved first, a TPRT mechanism or a DNA sensing repression mechanism. Alternatively, evolutionary pressure driving TE DNA sensing inhibition could be exerted by other TEs like ERVs, as IAPs and muERV-L are known active TEs in early mouse development (Gifford et al., 2013b).

At this point it is worth mentioning that the DNA sensor AIM2 was found to be interacting with LINE1 ORF1p in pluripotency (Figure 5.5A, Magenta). The enrichment factor of this interaction is not particularly high, although it is significant. Besides, the interaction has not been validated and the mechanism of said interaction remains unknown. However, the easiest explanation is that AIM2's interaction with LINE1 ORF1p is dependent on dsDNA, probably from LINE1 origin. It is clear that the AIM2 inflammasome does not seem to be active in mESCs and pyropoptosis does not seem to be occurring. Thus, it is possible that the AIM2 cell death pathway is inhibited downstream. Alternatively, it can be considered that AIM2 is expressed in pluripotency to activate pyropoptosis in response to an unresolvable threat.

With the goal of finding new LINE1 or ORF2p repressors in mind, I was able to characterise the endogenous LINE1 ORF1p interactome in mESCs cultured under serum/LIF conditions. At this point 2i/LIF conditions would have been preferred since they represent a more naïve state of pluripotency, they would be expected to have higher LINE1 expression and would have been a more accurate recapitulation of the ICM (Batlle-Morera et al., 2008; Garcia-Cañadas et al., 2023; Ying et al., 2008). Doing differential nucleus and a cytosolic-specific interactomes, could also have been very informative. My original intention was to compare the LINE1 ORF1p interactomes in pluripotency and senescence, a context where LINE1 clearly plays a pathogenic role (Cecco et al., 2019; Simon et al., 2019). This would have allowed a deeper understanding of LINE1 biology in senescence and facilitated the selection of targets to interrogate for their LINE1 repression potential. However, despite being able to immunoprecipitate ORF1p in senescence, failings in the mass spectrometry stage meant it was unfortunately not possible to obtain its interactome.

Despite this, my ORF1p interactome dataset confirmed the interaction in pluripotency of most of the already known LINE1 interactors found in LINE1 interactomes obtained with the

transgenic expression of LINE1 in somatic immortal cell lines like HeLa and HEK 293Ts. This included many repressors and activators that were known or have been recently discovered during the course of this work (T. D. Li et al., 2022; MacLennan et al., 2017; Nguyen et al., 2017). Remarkably, the DNA repair enzymes involved in LINE1 retrocopy integration such as Rpa1 were depleted from the ORF1p interactome. This provides further evidence for the repression of ORF2p expression or function, since they normally bind ssDNA. In contrast to ESCs, these proteins have been found in previous proteomic experiments in somatic cells.

Regarding the GO analysis of interacting proteins, it has highlighted categories related to splicing, ribosome biogenesis, translation regulation, helicase and chromatin modifications and gene expression. Ribosome biogenesis, the nucleolus and negative regulation of differentiation are consistent with the previous findings of LINE1 being involved in nucleolus formation, ribosome biogenesis activation and pluripotency (Lu et al., 2021, 2020; Percharde et al., 2018). Translation regulation makes sense in the context of LINE1 accumulation in stress granules, as most of the proteins associated with this category are also labelled as translational repressors (Coquel et al., 2018; Goodier et al., 2007). Translational repression of retrotransposons also occurs in meiotic cells (Laureau et al., 2021). The poly-purine tract is a characteristic feature of retroviruses that serves as the signal primer from which the reverse transcriptase starts the retrotranscription process. LINE1 RNAs also regulate early development chromatin dynamics acting as an scaffold to recruit epigenetic machinery like histone modifiers (Chen et al., 2021; Liu et al., 2021; Percharde et al., 2018). However, the enrichment of splicing and spliceosome-related categories remains to be understood.

At this point, considering the vast examples of TE co-option, it can be acknowledged that a lot of the interactors have a role as chromatin modifiers, transcriptional repressors and activators and the polycomb complex. These roles suggest and reinforce additional and current functions for LINE1 RNP beyond what is currently known, which may be relevant for chromatin dynamics in early development biology, as it has been suggested by some authors (Jachowicz et al., 2017). This activity can go beyond strictly chromatin effects. An example of this can be the E3 ubiquitin ligase *Trim71*, which is I found to be a LINE1 pluripotency-specific interactor. Trim71 maintains ESC self-renewal phenotype by targeting *Cdkn1a* RNA for elimination (Chang et al., 2012; Torres-Fernández et al., 2019). Considering that LINE1 KD reduces self-renewal, a phenotype partially explained by LINE1 promotion of ribogenesis, the interaction of Trim71 and LINE1 could also be relevant for ESC self-renewal, as its knockout is embryonic lethal before implantation.

As I hypothesised that pluripotent cells must have unique mechanisms of repressing LINE1 activity, I narrowed down my hits to 20 pluripotency-specific LINE1 interactors. One of them, *Zcchc3* was published as a repressor that degrades LINE1 RNA shortly after (Z. Zhang et al., 2022). Many of these 20 interactors are E3 ubiquitin ligases, which prompted me to consider that they might function in a targeted ORF2p degradation mechanism. Some of these proteins (Mkrn1, Trim71 and Nfx1), all of them ubiquitin ligases, were validated *in vitro* as interactors. To assess the retrotransposition repression activity of the candidates, I developed a screening platform that overcame the limitations of using a plasmid LINE1 reporter construct in mESCs. This platform consists of a stable mESC cell line with a codon-optimised LINE1-PFG reporter integrated into the *Rosa26* locus under the control of a CAG promoter, that replaced the 5'UTR (Han and Boeke, 2004). Using a construct that replaced the 5' UTR for the CAG promoter was potentially a mistake as some proteins may use the LINE1 5'UTR as a binding recognition site; it forms secondary structures recognised by certain proteins and it is important for LINE1 regulation (Bodak et al., 2014; Del Re et al., 2010; Heras et al., 2013). Despite its limitations, this cell line proved functional and allowed me to carry out a siRNA screening using the 20 pluripotency-specific genes, plus other cherry-picked interactors that had not been described in other contexts and had potentially relevant functions. The criteria for the rest of the genes were based on their high enrichment in LINE1 ORF1p IP and their reported function, prioritising E3 ubiquitin ligases or translational repressors. In retrospective, other approaches could have been taken, such as including more spliceosome related genes, giving the relevance of the category, or focusing on the lethality phenotypes of interactor gene knock outs on the embryo. 4 interactors cause failure morula compaction (Xab2, Rbm19, Npat and Dcaf13), and 26 cause embryo lethality before implantation. The ontologies of these genes involve primarily ribosome RNA processing, splicing and genomic reprogramming. Other categories that would have deserved more exploration are translational inhibition and endo and exonucleases like TREX1 which could degrade LINE1 cDNA and prevent retrotransposition (Thomas et al., 2017).

However, at this point I have to acknowledge the limitation of siRNAs as a system of screening: siRNA activity is transient, it may not be complete and it takes some time to take effect. Moreover, siRNAs may have off targets and or cause toxicity by interfering with the cell iRNA machinery, tittering out endogenous siRNAs and miRNAs. That is why, further down the line I aim to use the Degron system in a reduced pool of candidates. Nevertheless, the development of the E14 R26-LINE1-PFG cell line as a screening platform opens up avenues to use it for other approaches like CRISPR screens, genome-wide screens and chemical screens to test for proteins or drugs that could have an impact on LINE1 activation or repression.

Both the LINE1-PFG screen, and the γH2AX screen I used to readout DNA damage allowed me to identify a pool of potential LINE1 repressors. These were fairly consistent in the top candidates emerging from them, except for *Cebpz*. *Cebpz* has been the best performing candidate in successive LINE1-PFG reporter experiments. However, it failed to score highly in the γH2AX screen. This could be attributed to a reporter specific effect, a lack of resolution of the γH2AX screening. Moreover, giving its known function as transcriptional activator, known for stimulating transcription from the Hsp70 promoter (Lum et al., 1990), it is difficult to understand how could it be regulating LINE1 unless it is acting indirectly or has an additional not reported function.

Rbbp6 is an E3 ubiquitin ligase, and its knock out is embryonically lethal between implantation and somite formation, according to the mammalian phenotype ontology (Mungall et al., 2017; Smith et al., 2005). Due to its activity, it could potentially be stabilising EGFP in the first screen, but its good performance in the γH2AX makes it a more solid candidate. It could therefore potentially inhibit LINE1 by degrading ORF2p, preventing retrotransposition.

Finally, *Rbm22* is also a pluripotency specific interactor and part of the spliceosome, with its knock out also causing embryonic lethality according to “Mammalian phenotype ontology” (Mungall et al., 2017; Smith et al., 2005). The spliceosome seems to be a very relevant category in terms of LINE1 interactors in ESCs. However, its role in LINE1 repression is more difficult to comprehend, and may have relevance in LINE1 generating alternative splicing variants in pluripotency.

As none of the ORF1p interactors seem to have an obvious mechanism of repression, by submitting my ORF1p interactome to STRING, I can speculate on the function of my candidates based on the interactions between other components of the LINE1 interactome according to STRING. For instance, most of Cebpz’s interactors are surprisingly nucleolar proteins, proteins involved in the synthesis and processing of rRNA and ribosome biogenesis and helicases. The relationship of *Cebpz* to the nucleolus and ribosome biosynthesis seems new and considering that *Cebpz* is reported to be an activator, it may again be related to LINE1 role in activating ribosome biogenesis (Percharde et al., 2018). This would also make sense with the observation that its knockdown seems to substantially cause a self-renewal defect. Of its STRING interactors in the LINE1 ORF1p pluripotency interactome, *Rbm19* knockout is lethal during morula compaction and *Nop53* before implantation according to “Mammalian phenotype ontology” (Mungall et al., 2017; Smith et al., 2005). All of them are involved in ribosome biosynthesis (Baltz et al., 2012; Castello et al., 2012; Gaudet et al., 2011).

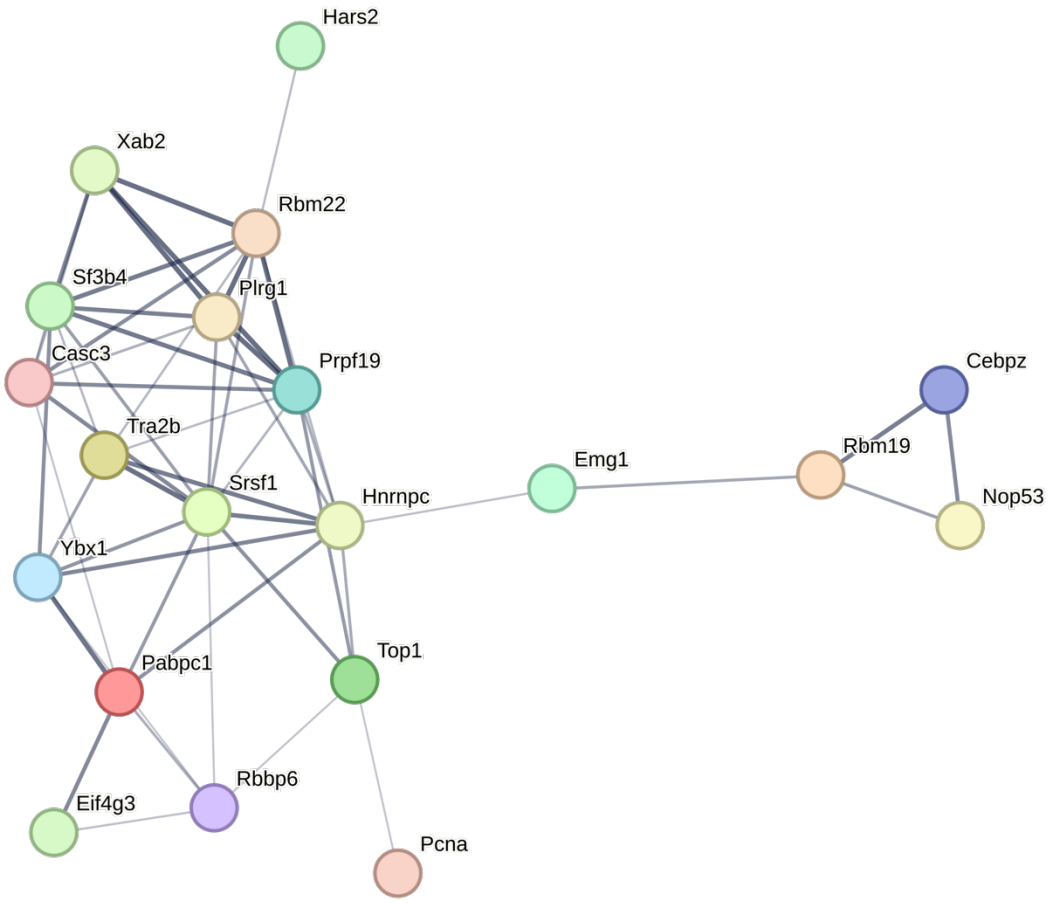


Figure 7-1: Candidates Interaction network. Interaction network of Rbm22, Pabpc1 and Cebpz with relevant interactors within the LINE1 ORF1p interactome in pluripotency. Based on STRING database.

Rbbp6 STRING interactors from LINE1 can have more potential interest as they are potential targets for degradation. One of them is *Pabpc1*, a PolyA binding protein necessary for LINE1 functional RNP formation (Dai et al., 2012). Another is Eif4g3, a translation initiation factor, which could also be assumed to be an activator of LINE1 RNA translation. In fact, its repression by a cell endogenous miRNA has been reported to lead to a decay in translation (Andreev et al., 2017; Sun et al., 2010). It can be speculated that the degradation of these proteins could block LINE1 retrotransposition. Another interactor is Ybx1, which has been detected in other somatic LINE1 interactomes, and is a documented target of degradation of Rbbp6. *Ybx1* is involved in multiple cellular processes like transcription, translation, splicing, DNA repair and mRNA packaging. It stabilizes the mRNA to prevent decay and facilitates its interaction with the translation machinery. As a DNA repair factor, it has endonuclease activity. Its knockout also results in embryonic lethality. Considering its activity as an E3 ubiquitin ligase, in future experiments I would like to test the capacity of Rbbp6 to target LINE1 ORF2p and other proteins relevant for LINE1 retrotransposition. Detection of ORF2p, specially endogenous protein is very complicated (Ardeljan et al., 2019). The best option in my hands

at this moment is the use of a LINE1-ORF2 cherry reporter construct. This can have multiple approaches like co-expression of the LINE1 ORF2p cherry reporter with a potential LINE1 repressor in a retrotransposition permissive environment like HEK 293T cells, which should reduce the cherry signal. Or on the contrary, co-expression of the LINE1 ORF2p cherry reporter, combined with inhibition of the potential LINE1 repressor in mESCs should lead to an increase in cherry signal.

Rbm22 is the final potential candidate repressor from my experiments. Rbm22 interacts mainly with helicases, spliceosome components and interestingly with Snip1, a protein involved on NFκB inhibition (Kim et al., 2001; Zhang et al., 2017). In my observation, its knock down seems to also cause a self-renewal defect, although notably this could also be a consequence of DNA damage from LINE1 activation. The knockout of one its interactors, *Xab2*, is lethal during morula compaction, and *Sf3b4*, *Prpf19* and *Plrg1* are lethal before implantation. Their roles involve primary splicing and in second term DNA repair. Considering it is a splicing factor, the most relevant topic in the gene ontology analysis, we can speculate in a role of the spliceosome splicing out ORF2p from LINE1 RNA. Studies have found LINE1 RNA isoforms where ORF1p has been spliced out (Belancio et al., 2006), these products have no capabilities of retrotransposition, and considering what we know about ORF2p translation, they are unlikely to translate ORF2p (Alisch et al., 2006). Considering mESCs express ORF1p, I speculate on a mechanism of LINE1 repression in pluripotency involving the exclusion by the spliceosome of ORF2p.

Potential repression of reporter insertions and potential confounding effects by ubiquitin-ligases lead me to develop an alternative approach to confirm LINE1 activity through genomic qPCR. This, however showed a very strong variability and background noise, which does not allow me to make any definitive statement on the repression potential of any of my candidates. A clearer result could come from using a Taqman approach, as the probe would reduce the unspecific amplification, but the persisting variability would still be a problem. Future experiments will involve a more efficient system to remove the activity of a potential repressor, like degrons. Degrons would allow for a finer-tuned, less toxic and complete degradation of the target protein and allow the suppression of protein expression for longer periods of time (Yesbolatova et al., 2020). New approaches should also not be limited to assess the reporter increase by qPCR or FACS, but also assessing the presence of ssDNA by IF or cytoplasmic DNA PCR and the use of the LINE1 ORF2p cherry reporter construct.

Regarding the reduced viability caused by the knockdown of these genes, there is not yet enough evidence to attribute this to LINE1 activity, as it could be attributed to siRNA toxicity.

Further experiments should assess and verify if this reduction is caused by de derepressing of ORF2p activity or any other pleiotropic effects.

It is also interesting to do some speculation on the similarities of LINE1 and innate immunity dynamics between the oocyte and early embryo. This can be useful as the embryo cytoplasm is inherited from the oocyte. While the piRNA pathways seem to be very conserved in the male germline as the established TE regulation mechanism, this conservation does not exist in the female germline, as mouse oocytes rely on a TE-derived *Dicer* isoform and human oocytes seem to rely in an additional Piwi protein (Flemr et al., 2013; Roovers et al., 2015). The lack of conservation could be the consequence of trying to prevent defence mechanisms from targeting certain TEs important for development when they are inherited via the cytoplasm of the oocyte. The result of this is very well illustrated by the fact that activation of LINE1 in the oocyte compromises its viability, but not through the activation of the cGAS/STING pathway, but presumably through ORF2p endonuclease activity (Malki et al., 2014; Tharp et al., 2020). Therefore, I can also speculate that *cGAS* and *STING* are also not expressed in the oocyte, preventing DNA sensing activation and inheritance of cGAS/STING to the early embryo.

We have to consider that early development is a highly redundant process. Redundancy is an expected characteristic in this type of system, as it adds robustness to gene networks and developmental processes. Some LINE1 repressors like *Dicer*, *Tdp-43* (no mechanism explained) and *Tex19.1* have been identified up until now in pluripotency, and several more will be identified (Bodak et al., 2017; T. D. Li et al., 2022; Luqman-Fatah et al., 2023). This work may provide some additional new names into the emerging network of factors repressing LINE1 in early development. In addition, it is important to take into account that the mechanism by which these potential repressors operate is still unclear and it might not necessary be what I predict. Still, the novelty and the relevance of this work also resides in the pluripotency-specific nature of the identified repressors, showing how LINE1 can be regulated by a developmental stage-specific network of factors and the establishment of a stable reliable LINE1 retrotransposition screening platform. As a final consideration, it is also possible that I have missed a repressor targeting ORF2p, either because it was overlooked despite being in the interactome, because the siRNAs targeting it were not effective enough, or because it was not detected by mass spectrometry. Future technical advances, especially those in ORF2p detection, will bring more progress to this field. Overall, future goals to prove my hypotheses should focus on the potential of repressor candidates to generate ssDNA intermediates that could trigger the cGAS/STING pathway. This would allow me ultimately to connect LINE1 TPRT repression with the repression of nucleic acid sensing.

Before concluding, giving the relevance of the translation inhibition category on the GO, and our hypothesis of ORF2p inhibition. The fact that some P-body markers are present in the interactome, like *Trim71*, *Mkrn1*, *Ago1* and *Ddx6* (Ayache et al., 2015; Miroci et al., 2012, p. 1; Torres-Fernández et al., 2019), could be mean that LINE1 is localizing in P-bodies. Considering the role of these structures in storage and translation inhibition, it can be speculated their role in the inhibition of LINE1 in pluripotency, particularly their role inhibiting ORF2p translation.

Overall, I have added new evidence supporting the idea that LINE1 has exercised evolutionary pressure that has shaped early development into developing a mechanism that permits the tolerance of LINE1 RNA in homeostasis but prevents its activity. In addition, my results support an additional redundant layer to inhibit the nucleic acid sensing and subsequent type I IFN generation. All in all, these mechanisms act in order to prevent the detrimental consequences for the fitness of the organism that activation of this pathway would have at this stage. Finding more about mechanisms of LINE1 evasion in mESCs may provide an insight in ways of inhibiting LINE1 detrimental activity in pathogenic somatic contexts, and shed new light on the aetiology of diseases like infertility and cancer.

Chapter 8: Bibliography

Ablasser, A., Schmid-Burgk, J.L., Hemmerling, I., Horvath, G.L., Schmidt, T., Latz, E., Hornung, V., 2013. Cell intrinsic immunity spreads to bystander cells via the intercellular transfer of cGAMP. Nature 503, 530–534. https://doi.org/10.1038/nature12640

Achleitner, M., Kleefisch, M., Hennig, A., Peschke, K., Polikarpova, A., Oertel, R., Gabriel, B., Schulze, L., Lindeman, D., Gerbaulet, A., Fiebig, U., Lee-Kirsch, M.A., Roers, A., Behrendt, R., 2017. Lack of Trex1 Causes Systemic Autoimmunity despite the Presence of Antiretroviral Drugs. J. Immunol. Baltim. Md 1950 199, 2261–2269. https://doi.org/10.4049/jimmunol.1700714

Acosta, J.C., Banito, A., Wuestefeld, T., Georgilis, A., Janich, P., Morton, J.P., Athineos, D., Kang, T.-W., Lasitschka, F., Andrulis, M., Pascual, G., Morris, K.J., Khan, S., Jin, H., Dharmalingam, G., Snijders, A.P., Carroll, T., Capper, D., Pritchard, C., Inman, G.J., Longerich, T., Sansom, O.J., Benitah, S.A., Zender, L., Gil, J., 2013. A complex secretory program orchestrated by the inflammasome controls paracrine senescence. Nat. Cell Biol. 15, 978–990. https://doi.org/10.1038/ncb2784

Acosta, J.C., O'Loghlen, A., Banito, A., Guijarro, M.V., Augert, A., Raguz, S., Fumagalli, M., Da Costa, M., Brown, C., Popov, N., Takatsu, Y., Melamed, J., d'Adda di Fagagna, F., Bernard, D., Hernando, E., Gil, J., 2008. Chemokine signaling via the CXCR2 receptor reinforces senescence. Cell 133, 1006–1018. https://doi.org/10.1016/j.cell.2008.03.038

Adiliaghdam, F., Basavappa, M., Saunders, T.L., Harjanto, D., Prior, J.T., Cronkite, D.A., Papavasiliou, N., Jeffrey, K.L., 2020. A Requirement for Argonaute 4 in Mammalian Antiviral Defense. Cell Rep. 30, 1690-1701.e4. https://doi.org/10.1016/j.celrep.2020.01.021

Aikawa, H., Tamai, M., Mitamura, K., Itmainati, F., Barber, G.N., Tagawa, Y., 2014. Innate immunity in an in vitro murine blastocyst model using embryonic and trophoblast stem cells. J. Biosci. Bioeng. 117, 358–365. https://doi.org/10.1016/j.jbiosc.2013.09.001

Akagi, K., Li, J., Stephens, R.M., Volfovsky, N., Symer, D.E., 2008. Extensive variation between inbred mouse strains due to endogenous L1 retrotransposition. Genome Res. 18, 869–880. https://doi.org/10.1101/gr.075770.107

Alisch, R.S., Garcia-Perez, J.L., Muotri, A.R., Gage, F.H., Moran, J.V., 2006. Unconventional translation of mammalian LINE-1 retrotransposons. Genes Dev. 20, 210–224. https://doi.org/10.1101/gad.1380406

Allen, S.E., Nowacki, M., 2017. Necessity Is the Mother of Invention: Ciliates, Transposons, and Transgenerational Inheritance. Trends Genet. TIG 33, 197–207. https://doi.org/10.1016/j.tig.2017.01.005

Almine, J.F., O'Hare, C.A.J., Dunphy, G., Haga, I.R., Naik, R.J., Atrih, A., Connolly, D.J., Taylor, J., Kelsall, I.R., Bowie, A.G., Beard, P.M., Unterholzner, L., 2017. IFI16 and cGAS cooperate in the activation of STING during DNA sensing in human keratinocytes. Nat. Commun. 8, 14392. https://doi.org/10.1038/ncomms14392

An, W., Dai, L., Niewiadomska, A.M., Yetil, A., O'Donnell, K.A., Han, J.S., Boeke, J.D., 2011. Characterization of a synthetic human LINE-1 retrotransposon ORFeus-Hs. Mob. DNA 2, 2. https://doi.org/10.1186/1759-8753-2-2

An, W., Han, J.S., Wheelan, S.J., Davis, E.S., Coombes, C.E., Ye, P., Triplett, C., Boeke, J.D., 2006. Active retrotransposition by a synthetic L1 element in mice. Proc. Natl. Acad. Sci. U. S. A. 103, 18662–18667. https://doi.org/10.1073/pnas.0605300103

Andreev, D.E., O'Connor, P.B.F., Loughran, G., Dmitriev, S.E., Baranov, P.V., Shatsky, I.N., 2017. Insights into the mechanisms of eukaryotic translation gained with ribosome profiling. Nucleic Acids Res. 45, 513–526. https://doi.org/10.1093/nar/gkw1190

Andreeva, L., Hiller, B., Kostrewa, D., Lässig, C., de Oliveira Mann, C.C., Jan Drexler, D., Maiser, A., Gaidt, M., Leonhardt, H., Hornung, V., Hopfner, K.-P., 2017. cGAS senses long and HMGB/TFAM-bound U-turn DNA by forming protein-DNA ladders. Nature 549, 394–398. https://doi.org/10.1038/nature23890

Aporntewan, C., Phokaew, C., Piriyapongsa, J., Ngamphiw, C., Ittiwut, C., Tongsima, S., Mutirangura, A., 2011. Hypomethylation of intragenic LINE-1 represses transcription in cancer cells through AGO2. PloS One 6, e17934. https://doi.org/10.1371/journal.pone.0017934

Araujo, J.M., Gomez, A.C., Aguilar, A., Salgado, R., Balko, J.M., Bravo, L., Doimi, F., Bretel, D., Morante, Z., Flores, C., Gomez, H.L., Pinto, J.A., 2018. Effect of CCL5 expression in the recruitment of immune cells in triple negative breast cancer. Sci. Rep. 8, 4899. https://doi.org/10.1038/s41598-018-23099-7

Ardeljan, D., Steranka, J.P., Liu, C., Li, Z., Taylor, M.S., Payer, L.M., Gorbounov, M., Sarnecki, J.S., Deshpande, V., Hruban, R.H., Boeke, J.D., Fenyö, D., Wu, P.-H., Smogorzewska, A., Holland, A.J., Burns, K.H., 2020. Cell fitness screens reveal a conflict between LINE-1 retrotransposition and DNA replication. Nat. Struct. Mol. Biol. 27, 168–178. https://doi.org/10.1038/s41594-020-0372-1

Ardeljan, D., Wang, X., Oghbaie, M., Taylor, M.S., Husband, D., Deshpande, V., Steranka, J.P., Gorbounov, M., Yang, W.R., Sie, B., Larman, H.B., Jiang, H., Molloy, K.R., Altukhov, I., Li, Z., McKerrow, W., Fenyö, D., Burns, K.H., LaCava, J., 2019. LINE-1 ORF2p expression is nearly imperceptible in human cancers. Mob. DNA 11, 1. https://doi.org/10.1186/s13100-019-0191-2

Ariumi, Y., Kawano, K., Yasuda-Inoue, M., Kuroki, M., Fukuda, H., Siddiqui, R., Turelli, P., Tateishi, S., 2018. DNA repair protein Rad18 restricts LINE-1 mobility. Sci. Rep. 8, 15894. https://doi.org/10.1038/s41598-018-34288-9

Arney, K.L., Fisher, A.G., 2004. Epigenetic aspects of differentiation. J. Cell Sci. 117, 4355–4363. https://doi.org/10.1242/jcs.01390

Arokium, H., Kamata, M., Kim, S., Kim, N., Liang, M., Presson, A.P., Chen, I.S., 2014. Deep sequencing reveals low incidence of endogenous LINE-1 retrotransposition in human induced pluripotent stem cells. PloS One 9, e108682. https://doi.org/10.1371/journal.pone.0108682

Ashley, J., Cordy, B., Lucia, D., Fradkin, L.G., Budnik, V., Thomson, T., 2018. Retrovirus-like Gag protein Arc1 binds RNA and traffics across synaptic boutons. Cell 172, 262-274.e11. https://doi.org/10.1016/j.cell.2017.12.022

Ayache, J., Bénard, M., Ernoult-Lange, M., Minshall, N., Standart, N., Kress, M., Weil, D., 2015. P-body assembly requires DDX6 repression complexes rather than decay or Ataxin2/2L complexes. Mol. Biol. Cell 26, 2579–2595. https://doi.org/10.1091/mbc.E15-03-0136

Baar, M.P., Brandt, R.M.C., Putavet, D.A., Klein, J.D.D., Derks, K.W.J., Bourgeois, B.R.M., Stryeck, S., Rijksen, Y., van Willigenburg, H., Feijtel, D.A., van der Pluijm, I., Essers, J., van Cappellen, W.A., van IJcken, W.F.J., Houtsmuller, A.B., Pothof, J., de Bruin, R.W.F., Madl, T., Hoeijmakers, J.H.J., Campisi, J., de Keizer, P.L.J., 2017. Targeted Apoptosis of Senescent Cells Restores Tissue Homeostasis in Response to Chemotoxicity and Aging. Cell 169, 132-147.e16. https://doi.org/10.1016/j.cell.2017.02.031

Baker, D.J., Childs, B.G., Durik, M., Wijers, M.E., Sieben, C.J., Zhong, J., Saltness, R., Jeganathan, K.B., Versoza, G.C., Pezeshki, A.-M., Khazaie, K., Miller, J.D., van Deursen, J.M., 2016. Naturally occurring p16Ink4a-positive cells shorten healthy lifespan. Nature 530, 184–189. https://doi.org/10.1038/nature16932

Baker, D.J., Wijshake, T., Tchkonia, T., LeBrasseur, N.K., Childs, B.G., Sluis, B. van de, Kirkland, J.L., Deursen, J.M. van, n.d. Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders. Nature 479, 232. https://doi.org/10.1038/nature10600

Baltz, A.G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., Wyler, E., Bonneau, R., Selbach, M., Dieterich, C., Landthaler, M., 2012. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol. Cell 46, 674–690. https://doi.org/10.1016/j.molcel.2012.05.021

Banerjee, I., Behl, B., Mendonca, M., Shrivastava, G., Russo, A.J., Menoret, A., Ghosh, A., Vella, A.T., Vanaja, S.K., Sarkar, S.N., Fitzgerald, K.A., Rathinam, V.A.K., 2018. Gasdermin D Restrains Type I Interferon Response to Cytosolic DNA by Disrupting Ionic Homeostasis. Immunity 49, 413-426.e5. https://doi.org/10.1016/j.immuni.2018.07.006

Bao, J., Yan, W., 2012. Male Germline Control of Transposable Elements1. Biol. Reprod. 86, 162, 1–14. https://doi.org/10.1095/biolreprod.111.095463

Barau, J., Teissandier, A., Zamudio, N., Roy, S., Nalesso, V., Hérault, Y., Guillou, F., Bourc'his, D., 2016. The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. Science 354, 909–912. https://doi.org/10.1126/science.aah5143

Barnett, K.C., Coronas-Serna, J.M., Zhou, W., Ernandes, M.J., Cao, A., Kranzusch, P.J., Kagan, J.C., 2019. Phosphoinositide Interactions Position cGAS at the Plasma Membrane to Ensure Efficient Distinction between Self- and Viral DNA. Cell 176, 1432-1446.e11. https://doi.org/10.1016/j.cell.2019.01.049

Bartonicek, N., Rouet, R., Warren, J., Loetsch, C., Rodriguez, G.S., Walters, S., Lin, F., Zahra, D., Blackburn, J., Hammond, J.M., Reis, A.L.M., Deveson, I.W., Zammit, N., Zeraati, M., Grey, S., Christ, D., Mattick, J.S., Chtanova, T., Brink, R., Dinger, M.E., Weatheritt, R.J., Sprent, J., King, C., 2022. The retroelement Lx9 puts a brake on the immune response to virus infection. Nature 608, 757–765. https://doi.org/10.1038/s41586-022-05054-9

Batlle-Morera, L., Smith, A., Nichols, J., 2008. Parameters influencing derivation of embryonic stem cells from murine embryos. Genes. N. Y. N 2000 46, 758–767. https://doi.org/10.1002/dvg.20442

Beck, C.R., Collier, P., Macfarlane, C., Malig, M., Kidd, J.M., Eichler, E.E., Badge, R.M., Moran, J.V., 2010. LINE-1 Retrotransposition Activity in Human Genomes. Cell 141, 1159–1170. https://doi.org/10.1016/j.cell.2010.05.021

Beck, C.R., Garcia-Perez, J.L., Badge, R.M., Moran, J.V., 2011. LINE-1 Elements in Structural Variation and Disease. Annu. Rev. Genomics Hum. Genet. 12, 187–215. https://doi.org/10.1146/annurev-genom-082509-141802

Bejerano, G., Lowe, C.B., Ahituv, N., King, B., Siepel, A., Salama, S.R., Rubin, E.M., Kent, W.J., Haussler, D., 2006. A distal enhancer and an ultraconserved exon are derived from a novel retroposon. Nature 441, 87–90. https://doi.org/10.1038/nature04696

Belancio, V.P., Deininger, P.L., Roy-Engel, A.M., 2009. LINE dancing in the human genome: transposable elements and disease. Genome Med. 1, 97. https://doi.org/10.1186/gm97

Belancio, V.P., Hedges, D.J., Deininger, P., 2006. LINE-1 RNA splicing and influences on mammalian gene expression. Nucleic Acids Res. 34, 1512–1521. https://doi.org/10.1093/nar/gkl027

Belancio, V.P., Roy-Engel, A.M., Deininger, P., 2008. The impact of multiple splice sites in human L1 elements. Gene 411, 38–45. https://doi.org/10.1016/j.gene.2007.12.022

Belancio, V.P., Roy-Engel, A.M., Pochampally, R.R., Deininger, P., 2010. Somatic expression of LINE-1 elements in human tissues. Nucleic Acids Res. 38, 3909–3922. https://doi.org/10.1093/nar/gkq132

Belgnaoui, S.M., Gosden, R.G., Semmes, O.J., Haoudi, A., 2006. Human LINE-1 retrotransposon induces DNA damage and apoptosis in cancer cells. Cancer Cell Int. 6, 13. https://doi.org/10.1186/1475-2867-6-13

Belzile, J.-P., Stark, T.J., Yeo, G.W., Spector, D.H., 2014. Human cytomegalovirus infection of human embryonic stem cell-derived primitive neural stem cells is restricted at several steps but leads to the persistence of viral DNA. J. Virol. 88, 4021–4039. https://doi.org/10.1128/JVI.03492-13

Benitez-Guijarro, M., Lopez-Ruiz, C., Tarnauskaitė, Ž., Murina, O., Mian Mohammad, M., Williams, T.C., Fluteau, A., Sanchez, L., Vilar-Astasio, R., Garcia-Canadas, M., Cano, D., Kempen, M.-J.H., Sanchez-Pozo, A., Heras, S.R., Jackson, A.P., Reijns, M.A., Garcia-Perez, J.L., 2018. RNase H2, mutated in Aicardi-Goutières syndrome, promotes LINE-1 retrotransposition. EMBO J. 37, e98506. https://doi.org/10.15252/embj.201798506

Bennion, B.G., Croft, C.A., Ai, T.L., Qian, W., Menos, A.M., Miner, C.A., Frémond, M.-L., Doisne, J.-M., Andhey, P.S., Platt, D.J., Bando, J.K., Wang, E.R., Luksch, H., Molina, T.J., Roberson, E.D.O., Artyomov, M.N., Rösen-Wolff, A., Colonna, M., Rieux-Laucat, F., Di Santo, J.P., Neven, B., Miner, J.J., 2020. STING Gain-of-Function Disrupts Lymph Node Organogenesis and Innate Lymphoid Cell Development in Mice. Cell Rep. 31, 107771. https://doi.org/10.1016/j.celrep.2020.107771

Bennion, B.G., Ingle, H., Ai, T.L., Miner, C.A., Platt, D.J., Smith, A.M., Baldridge, M.T., Miner, J.J., 2019. A Human Gain-of-Function STING Mutation Causes Immunodeficiency and Gammaherpesvirus-Induced Pulmonary Fibrosis in Mice. J. Virol. 93. https://doi.org/10.1128/JVI.01806-18

Bhat, H., Zaun, G., Hamdan, T.A., Lang, J., Adomati, T., Schmitz, R., Friedrich, S.-K., Bergerhausen, M., Cham, L.B., Li, F., Ali, M., Zhou, F., Khairnar, V., Duhan, V., Brandenburg, T., Machlah, Y.M., Schiller, M., Berry, A., Xu, H., Vollmer, J., Häussinger, D., Thier, B., Pandyra, A.A., Schadendorf, D., Paschen, A., Schuler, M., Lang, P.A., Lang, K.S., 2020. Arenavirus Induced CCL5 Expression Causes NK Cell-Mediated Melanoma Regression. Front. Immunol. 11.

Bisbal, C., Salehzada, T., Silhol, M., Martinand, C., Le Roy, F., Lebleu, B., 2001. The 2-5A/RNase L pathway and inhibition by RNase L inhibitor (RLI). Methods Mol. Biol. Clifton NJ 160, 183–198. https://doi.org/10.1385/1-59259-233-3:183

Bodak, M., Ciaudo, C., 2016. Monitoring Long Interspersed Nuclear Element 1 Expression During Mouse Embryonic Stem Cell Differentiation. Methods Mol. Biol. Clifton NJ 1400, 237–259. https://doi.org/10.1007/978-1-4939-3372-3_16

Bodak, M., Cirera-Salinas, D., Yu, J., Ngondo, R.P., Ciaudo, C., 2017. Dicer, a new regulator of pluripotency exit and LINE-1 elements in mouse embryonic stem cells. FEBS Open Bio 7, 204–220. https://doi.org/10.1002/2211-5463.12174

Bodak, M., Yu, J., Ciaudo, C., 2014. Regulation of LINE-1 in mammals. Biomol. Concepts 5, 409–428. https://doi.org/10.1515/bmc-2014-0018

Bogerd, H.P., Skalsky, R.L., Kennedy, E.M., Furuse, Y., Whisnant, A.W., Flores, O., Schultz, K.L.W., Putnam, N., Barrows, N.J., Sherry, B., Scholle, F., Garcia-Blanco, M.A., Griffin, D.E., Cullen, B.R., 2014. Replication of Many Human Viruses Is Refractory to Inhibition by Endogenous Cellular MicroRNAs. J. Virol. 88, 8065–8076. https://doi.org/10.1128/JVI.00985-14

Bona, N., Crossan, G.P., 2023. Fanconi anemia DNA crosslink repair factors protect against LINE-1 retrotransposition during mouse development. Nat. Struct. Mol. Biol. 30, 1434–1445. https://doi.org/10.1038/s41594-023-01067-8

Borden, E.C., Hogan, T.F., Voelkel, J.G., 1982. Comparative antiproliferative activity in vitro of natural interferons alpha and beta for diploid and transformed human cells. Cancer Res. 42, 4948–4953.

Bošković, A., Eid, A., Pontabry, J., Ishiuchi, T., Spiegelhalter, C., Raghu Ram, E.V.S., Meshorer, E., Torres-Padilla, M.-E., 2014. Higher chromatin mobility supports totipotency and precedes pluripotency in vivo. Genes Dev. 28, 1042–1047. https://doi.org/10.1101/gad.238881.114

Bouis, D., 2019. Severe combined immunodeficiency in stimulator of interferon genes (STING) V154M/wild-type mice. J ALLERGY CLIN IMMUNOL 143, 19.

Bouquet, F., Muller, C., Salles, B., 2006. The loss of gammaH2AX signal is a marker of DNA double strand breaks repair only at low levels of DNA damage. Cell Cycle Georget. Tex 5, 1116–1122. https://doi.org/10.4161/cc.5.10.2799

Bourque, G., Burns, K.H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin, H.L., Macfarlan, T.S., Mager, D.L., Feschotte, C., 2018. Ten things you should know about transposable elements. Genome Biol. 19, 1–12. https://doi.org/10.1186/s13059-018-1577-z

Briggs, E.M., McKerrow, W., Mita, P., Boeke, J.D., Logan, S.K., Fenyö, D., 2021. RIP-seq reveals LINE-1 ORF1p association with p-body enriched mRNAs. Mob. DNA 12, 5. https://doi.org/10.1186/s13100-021-00233-3

Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., Vallier, L., 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 448, 191–195. https://doi.org/10.1038/nature05950

Brown, M.G., Dokun, A.O., Heusel, J.W., Smith, H.R.C., Beckman, D.L., Blattenberger, E.A., Dubbelde, C.E., Stone, L.R., Scalzo, A.A., Yokoyama, W.M., 2001. Vital Involvement of a Natural Killer Cell Activation Receptor in Resistance to Viral Infection. Science 292, 934–937. https://doi.org/10.1126/science.1060042

Bruno, M., Mahgoub, M., Macfarlan, T.S., 2019. The Arms Race Between KRAB–Zinc Finger Proteins and Endogenous Retroelements and Its Impact on Mammals. Annu. Rev. Genet. 53, 393–416. https://doi.org/10.1146/annurev-genet-112618-043717

Buchrieser, J., Degrelle, S.A., Couderc, T., Nevers, Q., Disson, O., Manet, C., Donahue, D.A., Porrot, F., Hillion, K.-H., Perthame, E., Arroyo, M.V., Souquere, S., Ruigrok, K., Dupressoir, A., Heidmann, T., Montagutelli, X., Fournier, T., Lecuit, M., Schwartz, O., 2019. IFITM proteins inhibit placental syncytiotrophoblast formation and promote fetal demise. Science 365, 176–180. https://doi.org/10.1126/science.aaw7733

Budhwani, M., Mazzieri, R., Dolcetti, R., 2018. Plasticity of Type I Interferon-Mediated Responses in Cancer Therapy: From Anti-tumor Immunity to Resistance. Front. Oncol. 8.

Bulut-Karslioglu, A., De La Rosa-Velázquez, I.A., Ramirez, F., Barenboim, M., Onishi-Seebacher, M., Arand, J., Galán, C., Winter, G.E., Engist, B., Gerle, B., O'Sullivan, R.J., Martens, J.H.A., Walter, J., Manke, T., Lachner, M., Jenuwein, T., 2014. Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. Mol. Cell 55, 277–290. https://doi.org/10.1016/j.molcel.2014.05.029

Burke, D.C., Graham, C.F., Lehman, J.M., 1978. Appearance of interferon inducibility and sensitivity during differentiation of murine teratocarcinoma cells in vitro. Cell 13, 243–248. https://doi.org/10.1016/0092-8674(78)90193-9

Burleigh, K., Maltbaek, J.H., Cambier, S., Green, R., Gale, M., James, R.C., Stetson, D.B., 2020. Human DNA-PK activates a STING-independent DNA sensing pathway. Sci. Immunol. 5, eaba4219. https://doi.org/10.1126/sciimmunol.aba4219

Burton, A., Torres-Padilla, M.-E., 2014. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. Nat. Rev. Mol. Cell Biol. 15, 723–734. https://doi.org/10.1038/nrm3885

Cabrera, A., Edelstein, H.I., Glykofrydis, F., Love, K.S., Palacios, S., Tycko, J., Zhang, M., Lensch, S., Shields, C.E., Livingston, M., Weiss, R., Zhao, H., Haynes, K.A., Morsut, L., Chen, Y.Y., Khalil, A.S., Wong, W.W., Collins, J.J., Rosser, S.J., Polizzi, K., Elowitz, M.B., Fussenegger, M., Hilton, I.B., Leonard, J.N., Bintu, L., Galloway, K.E., Deans, T.L., 2022. The sound of silence: Transgene silencing in mammalian cell engineering. Cell Syst. 13, 950–973. https://doi.org/10.1016/j.cels.2022.11.005

Calarco, P.G., Brown, E.H., 1969. An ultrastructural and cytological study of preimplantation development of the mouse. J. Exp. Zool. 171, 253–283. https://doi.org/10.1002/jez.1401710303

Camino, L.P., Dutta, A., Barroso, S., Pérez-Calero, C., Katz, J.N., García-Rubio, M., Sung, P., Gómez-González, B., Aguilera, A., 2023. DICER ribonuclease removes harmful R-loops. Mol. Cell. https://doi.org/10.1016/j.molcel.2023.09.021

Cañadas, I., Thummalapalli, R., Kim, J.W., Kitajima, S., Jenkins, R.W., Christensen, C.L., Campisi, M., Kuang, Y., Zhang, Y., Gjini, E., Zhang, G., Tian, T., Sen, D.R., Miao, D.,

Imamura, Y., Thai, T., Piel, B., Terai, H., Aref, A.R., Hagan, T., Koyama, S., Watanabe, M., Baba, H., Adeni, A.E., Lydon, C.A., Tamayo, P., Wei, Z., Herlyn, M., Barbie, T.U., Uppaluri, R., Sholl, L.M., Sicinska, E., Sands, J., Rodig, S., Wong, K.K., Paweletz, C.P., Watanabe, H., Barbie, D.A., 2018. Tumor innate immunity primed by specific interferon-stimulated endogenous retroviruses. Nat. Med. 24, 1143–1150. https://doi.org/10.1038/s41591-018-0116-5

Carelli, F.N., Hayakawa, T., Go, Y., Imai, H., Warnefors, M., Kaessmann, H., 2016. The life history of retrocopies illuminates the evolution of new mammalian genes. Genome Res. 26, 301–314. https://doi.org/10.1101/gr.198473.115

Carozza, J.A., Böhnert, V., Nguyen, K.C., Skariah, G., Shaw, K.E., Brown, J.A., Rafat, M., von Eyben, R., Graves, E.E., Glenn, J.S., Smith, M., Li, L., 2020. Extracellular cGAMP is a cancer cell-produced immunotransmitter involved in radiation-induced anti-cancer immunity. Nat. Cancer 1, 184–196. https://doi.org/10.1038/s43018-020-0028-4

Carter, V., LaCava, J., Taylor, M.S., Liang, S.Y., Mustelin, C., Ukadike, K.C., Bengtsson, A., Lood, C., Mustelin, T., 2020. High Prevalence and Disease Correlation of Autoantibodies Against p40 Encoded by Long Interspersed Nuclear Elements in Systemic Lupus Erythematosus. Arthritis Rheumatol. Hoboken NJ 72, 89–99. https://doi.org/10.1002/art.41054

Casacuberta, E., González, J., 2013. The impact of transposable elements in environmental adaptation. Mol. Ecol. 22, 1503–1517. https://doi.org/10.1111/mec.12170

Casser, E., Israel, S., Witten, A., Schulte, K., Schlatt, S., Nordhoff, V., Boiani, M., 2017. Totipotency segregates between the sister blastomeres of two-cell stage mouse embryos. Sci. Rep. 7, 8299. https://doi.org/10.1038/s41598-017-08266-6

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., Krijgsveld, J., Hentze, M.W., 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149, 1393–1406. https://doi.org/10.1016/j.cell.2012.04.031

Cecco, M.D., Ito, T., Petrashen, A.P., Elias, A.E., Skvir, N.J., Criscione, S.W., Caligiana, A., Brocculi, G., Adney, E.M., Boeke, J.D., Le, O., Beauséjour, C., Ambati, J., Ambati, K., Simon, M., Seluanov, A., Gorbunova, V., Slagboom, P.E., Helfand, S.L., Neretti, N., Sedivy, J.M., 2019. LINE-1 derepression in senescent cells triggers interferon and inflammaging. Nature 566, 73–78. https://doi.org/10.1038/s41586-018-0784-9

Cerboni, S., Jeremiah, N., Gentili, M., Gehrmann, U., Conrad, C., Stolzenberg, M.-C., Picard, C., Neven, B., Fischer, A., Amigorena, S., Rieux-Laucat, F., Manel, N., 2017. Intrinsic antiproliferative activity of the innate sensor STING in T lymphocytes. J. Exp. Med. 214, 1769–1785. https://doi.org/10.1084/jem.20161674

Cervantes-Ayalc, A., Ruiz Esparza-Garrido, R., Velázquez-Flores, M.Á., 2020. Long Interspersed Nuclear Elements 1 (LINE1): The chimeric transcript L1-MET and its involvement in cancer. Cancer Genet. 241, 1–11. https://doi.org/10.1016/j.cancergen.2019.11.004

Chang, H.-M., Martinez, N.J., Thornton, J.E., Hagan, J.P., Nguyen, K.D., Gregory, R.I., 2012. Trim71 cooperates with microRNAs to repress Cdkn1a expression and promote embryonic stem cell proliferation. Nat. Commun. 3, 923. https://doi.org/10.1038/ncomms1909

Chang, J., Wang, Y., Shao, L., Laberge, R.-M., Demaria, M., Campisi, J., Janakiraman, K., Sharpless, N.E., Ding, S., Feng, W., Luo, Y., Wang, X., Aykin-Burns, N., Krager, K., Ponnappan, U., Hauer-Jensen, M., Meng, A., Zhou, D., 2016. Clearance of senescent cells by ABT263 rejuvenates aged hematopoietic stem cells in mice. Nat. Med. 22, 78–83. https://doi.org/10.1038/nm.4010

Chattopadhyay, S., Marques, J.T., Yamashita, M., Peters, K.L., Smith, K., Desai, A., Williams, B.R.G., Sen, G.C., 2010. Viral apoptosis is induced by IRF-3-mediated activation of Bax. EMBO J. 29, 1762–1773. https://doi.org/10.1038/emboj.2010.50

Chattopadhyay, S., Sen, G.C., 2017. RIG-I-like receptor-induced IRF3 mediated pathway of apoptosis (RIPA): a new antiviral pathway. Protein Cell 8, 165–168. https://doi.org/10.1007/s13238-016-0334-x

Chattopadhyay, S., Yamashita, M., Zhang, Y., Sen, G.C., 2011. The IRF-3/Bax-Mediated Apoptotic Pathway, Activated by Viral Cytoplasmic RNA and DNA, Inhibits Virus Replication. J. Virol. 85, 3708–3716. https://doi.org/10.1128/jvi.02133-10

Chawla-Sarkar, M., Lindner, D.J., n.d. Apoptosis and interferons: Role of interferon-stimulated genes as mediators of apoptosis.

Chen, C., Liu, W., Guo, J., Liu, Y., Liu, X., Liu, J., Dou, X., Le, R., Huang, Y., Li, C., Yang, L., Kou, X., Zhao, Y., Wu, Y., Chen, J., Wang, H., Shen, B., Gao, Y., Gao, S., 2021. Nuclear m6A reader YTHDC1 regulates the scaffold function of LINE1 RNA in mouse ESCs and early embryos. Protein Cell 12, 455–474. https://doi.org/10.1007/s13238-021-00837-8

Chen, G., Shaw, M.H., Kim, Y.-G., Nuñez, G., 2009. NOD-like receptors: role in innate immunity and inflammatory disease. Annu. Rev. Pathol. 4, 365–398. https://doi.org/10.1146/annurev.pathol.4.110807.092239

Chen, J.-M., Stenson, P.D., Cooper, D.N., Férec, C., 2005. A systematic analysis of LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease. Hum. Genet. 117, 411–427. https://doi.org/10.1007/s00439-005-1321-0

Chen, L.-L., Yang, L., Carmichael, G.G., 2010. Molecular basis for an attenuated cytoplasmic dsRNA response in human embryonic stem cells. Cell Cycle 9, 3552–3564. https://doi.org/10.4161/cc.9.17.12792

Chen, T., Xu, Z.-G., Luo, J., Manne, R.K., Wang, Z., Hsu, C.-C., Pan, B.-S., Cai, Z., Tsai, P.-J., Tsai, Y.-S., Chen, Z.-Z., Li, H., Lin, H.-K., 2023. NSUN2 is a glucose sensor suppressing cGAS/STING to maintain tumorigenesis and immunotherapy resistance. Cell Metab. 35, 1782-1798.e8. https://doi.org/10.1016/j.cmet.2023.07.009

Chen, Y.G., Hur, S., 2022. Cellular origins of dsRNA, their recognition and consequences. Nat. Rev. Mol. Cell Biol. 23, 286–301. https://doi.org/10.1038/s41580-021-00430-1

Chiappinelli, K.B., Strissel, P.L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N.S., Cope, L.M., Snyder, A., Makarov, V., Buhu, S., Slamon, D.J., Wolchok, J.D., Pardoll, D.M., Beckmann, M.W., Zahnow, C.A., Merghoub, T., Chan, T.A., Baylin, S.B., Strick, R., 2015. Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. Cell 162, 974–986. https://doi.org/10.1016/j.cell.2015.07.011

Chibi, M., Meyer, M., Skepu, A., G Rees, D.J., Moolman-Smook, J.C., Pugh, D.J.R., 2008. RBBP6 interacts with multifunctional protein YB-1 through its RING finger domain, leading to

ubiquitination and proteosomal degradation of YB-1. J. Mol. Biol. 384, 908–916. https://doi.org/10.1016/j.jmb.2008.09.060

Childs, B.G., Baker, D.J., Wijshake, T., Conover, C.A., Campisi, J., van Deursen, J.M., 2016. Senescent intimal foam cells are deleterious at all stages of atherosclerosis. Science 354, 472–477. https://doi.org/10.1126/science.aaf6659

Choi, J., Hwang, S.-Y., Ahn, K., 2018. Interplay between RNASEH2 and MOV10 controls LINE-1 retrotransposition. Nucleic Acids Res. 46, 1912–1926. https://doi.org/10.1093/nar/gkx1312

Christian, C.M., Kines, K.J., Belancio, V.P., 2016. The importance of L1 ORF2p cryptic sequence to ORF2p fragment-mediated cytotoxicity. Mob. Genet. Elem. 6. https://doi.org/10.1080/2159256X.2016.1198300

Chu, T.-T., Tu, X., Yang, K., Wu, J., Repa, J.J., Yan, N., 2021. Tonic prime-boost of STING signalling mediates Niemann–Pick disease type C. Nature 596, 570–575. https://doi.org/10.1038/s41586-021-03762-2

Chuong, E.B., Elde, N.C., Feschotte, C., 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science 351, 1083–1087. https://doi.org/10.1126/science.aad5497

Chuong, E.B., Elde, N.C., Feschotte, C., 2017. Regulatory activities of transposable elements: from conflicts to benefits. Nat. Rev. Genet. 18, 71–86. https://doi.org/10.1038/nrg.2016.139

Chuong, E.B., Rumi, M.A.K., Soares, M.J., Baker, J.C., 2013. Endogenous retroviruses function as species-specific enhancer elements in the placenta. Nat. Genet. 45, 325–329. https://doi.org/10.1038/ng.2553

Civril, F., Deimling, T., de Oliveira Mann, C.C., Ablasser, A., Moldt, M., Witte, G., Hornung, V., Hopfner, K.-P., 2013. Structural mechanism of cytosolic DNA sensing by cGAS. Nature 498, 332–337. https://doi.org/10.1038/nature12305

COHEN, G.M., 1997. Caspases: the executioners of apoptosis. Biochem. J. 326, 1–16. https://doi.org/10.1042/bj3260001

Conzelmann, K.K., 1998. Nonsegmented negative-strand RNA viruses: genetics and manipulation of viral genomes. Annu. Rev. Genet. 32, 123–162. https://doi.org/10.1146/annurev.genet.32.1.123

Coquel, F., Silva, M.-J., Técher, H., Zadorozhny, K., Sharma, S., Nieminuszczy, J., Mettling, C., Dardillac, E., Barthe, A., Schmitz, A.-L., Promonet, A., Cribier, A., Sarrazin, A., Niedzwiedz, W., Lopez, B., Costanzo, V., Krejci, L., Chabes, A., Benkirane, M., Lin, Y.-L., Pasero, P., 2018. SAMHD1 acts at stalled replication forks to prevent interferon induction. Nature 557, 57–61. https://doi.org/10.1038/s41586-018-0050-1

Corrales, L., Woo, S.-R., Williams, J., McWhirter, S.M., Dubensky, T.W., Gajewski, T.F., 2016. Antagonism of the STING pathway via activation of the AIM2 inflammasome by intracellular DNA. J. Immunol. Baltim. Md 1950 196, 3191–3198. https://doi.org/10.4049/jimmunol.1502538

Cosby, R.L., Judd, J., Zhang, R., Zhong, A., Garry, N., Pritham, E.J., Feschotte, C., 2021. Recurrent evolution of vertebrate transcription factors by transposase capture. Science 371, eabc6405. https://doi.org/10.1126/science.abc6405

Cost, G.J., Boeke, J.D., 1998. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. Biochemistry 37, 18081–18093. https://doi.org/10.1021/bi981858s

Crow, Y.J., Hayward, B.E., Parmar, R., Robins, P., Leitch, A., Ali, M., Black, D.N., van Bokhoven, H., Brunner, H.G., Hamel, B.C., Corry, P.C., Cowan, F.M., Frints, S.G., Klepper, J., Livingston, J.H., Lynch, S.A., Massey, R.F., Meritet, J.F., Michaud, J.L., Ponsot, G., Voit, T., Lebon, P., Bonthron, D.T., Jackson, A.P., Barnes, D.E., Lindahl, T., 2006a. Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 cause Aicardi-Goutières syndrome at the AGS1 locus. Nat. Genet. 38, 917–920. https://doi.org/10.1038/ng1845

Crow, Y.J., Leitch, A., Hayward, B.E., Garner, A., Parmar, R., Griffith, E., Ali, M., Semple, C., Aicardi, J., Babul-Hirji, R., Baumann, C., Baxter, P., Bertini, E., Chandler, K.E., Chitayat, D., Cau, D., Déry, C., Fazzi, E., Goizet, C., King, M.D., Klepper, J., Lacombe, D., Lanzi, G., Lyall, H., Martínez-Frías, M.L., Mathieu, M., McKeown, C., Monier, A., Oade, Y., Quarrell, O.W., Rittey, C.D., Rogers, R.C., Sanchis, A., Stephenson, J.B.P., Tacke, U., Till, M., Tolmie, J.L., Tomlin, P., Voit, T., Weschke, B., Woods, C.G., Lebon, P., Bonthron, D.T., Ponting, C.P., Jackson, A.P., 2006b. Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutières syndrome and mimic congenital viral brain infection. Nat. Genet. 38, 910–916. https://doi.org/10.1038/ng1842

Crow, Y.J., Manel, N., 2015. Aicardi–Goutières syndrome and the type I interferonopathies. Nat. Rev. Immunol. 15, 429–440. https://doi.org/10.1038/nri3850

Crow, Y.J., Rehwinkel, J., 2009. Aicardi-Goutieres syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity. Hum. Mol. Genet. 18, R130-136. https://doi.org/10.1093/hmg/ddp293

Cruickshanks, H.A., Vafadar-Isfahani, N., Dunican, D.S., Lee, A., Sproul, D., Lund, J.N., Meehan, R.R., Tufarelli, C., 2013. Expression of a large LINE-1-driven antisense RNA is linked to epigenetic silencing of the metastasis suppressor gene TFPI-2 in cancer. Nucleic Acids Res. 41, 6857–6869. https://doi.org/10.1093/nar/gkt438

Dai, L., Taylor, M.S., O'Donnell, K.A., Boeke, J.D., 2012. Poly(A) Binding Protein C1 Is Essential for Efficient L1 Retrotransposition and Affects L1 RNP Formation. Mol. Cell. Biol. 32, 4323. https://doi.org/10.1128/MCB.06785-11

D'Angelo, W., Acharya, D., Wang, R., Wang, J., Gurung, C., Chen, B., Bai, F., Guo, Y.-L., 2016. Development of Antiviral Innate Immunity During In Vitro Differentiation of Mouse Embryonic Stem Cells. Stem Cells Dev. 25, 648–659. https://doi.org/10.1089/scd.2015.0377

D'Angelo, W., Chen, B., Gurung, C., Guo, Y.-L., 2018a. Characterization of embryonic stem cell-differentiated fibroblasts as mesenchymal stem cells with robust expansion capacity and attenuated innate immunity. Stem Cell Res. Ther. 9, 278. https://doi.org/10.1186/s13287-018-1033-8

D'Angelo, W., Chen, B., Gurung, C., Guo, Y.-L., 2018b. Characterization of embryonic stem cell-differentiated fibroblasts as mesenchymal stem cells with robust expansion capacity and attenuated innate immunity. Stem Cell Res. Ther. 9, 278. https://doi.org/10.1186/s13287-018-1033-8

Davidson, E.H., Britten, R.J., 1979. Regulation of gene expression: possible role of repetitive sequences. Science 204, 1052–1059. https://doi.org/10.1126/science.451548

De Cegli, R., Iacobacci, S., Flore, G., Gambardella, G., Mao, L., Cutillo, L., Lauria, M., Klose, J., Illingworth, E., Banfi, S., di Bernardo, D., 2013. Reverse engineering a mouse embryonic stem cell-specific transcriptional network reveals a new modulator of neuronal differentiation. Nucleic Acids Res. 41, 711–726. https://doi.org/10.1093/nar/gks1136

De Iaco, A., Planet, E., Coluccio, A., Verp, S., Duc, J., Trono, D., 2017. DUX-family transcription factors regulate zygotic genome activation in placental mammals. Nat. Genet. 49, 941–945. https://doi.org/10.1038/ng.3858

de Weerd, N.A., Nguyen, T., 2012. The interferons and their receptors—distribution and regulation. Immunol. Cell Biol. 90, 483–491. https://doi.org/10.1038/icb.2012.9

Dean, W., Bowden, L., Aitchison, A., Klose, J., Moore, T., Meneses, J.J., Reik, W., Feil, R., 1998. Altered imprinted gene methylation and expression in completely ES cell-derived mouse fetuses: association with aberrant phenotypes. Development 125, 2273–2282. https://doi.org/10.1242/dev.125.12.2273

DeBerardinis, R.J., Kazazian, H.H., 1999. Analysis of the Promoter from an Expanding Mouse Retrotransposon Subfamily. Genomics 56, 317–323. https://doi.org/10.1006/geno.1998.5729

Decout, A., Katz, J.D., Venkatraman, S., Ablasser, A., 2021. The cGAS–STING pathway as a therapeutic target in inflammatory diseases. Nat. Rev. Immunol. 21, 548–569. https://doi.org/10.1038/s41577-021-00524-z

Dedoni, S., Olianas, M.C., Onali, P., 2010. Interferon-β induces apoptosis in human SH-SY5Y neuroblastoma cells through activation of JAK-STAT signaling and down-regulation of PI3K/Akt pathway. J. Neurochem. 115, 1421–1433. https://doi.org/10.1111/j.1471-4159.2010.07046.x

Del Re, B., Marcantonio, P., Capri, M., Giorgi, G., 2010. Evaluation of LINE-1 mobility in neuroblastoma cells by in vitro retrotransposition reporter assay: FACS analysis can detect only the tip of the iceberg of the inserted L1 elements. Exp. Cell Res. 316, 3358–3367. https://doi.org/10.1016/j.yexcr.2010.06.024

Della Valle, F., Thimma, M.P., Caiazzo, M., Pulcrano, S., Celii, M., Adroub, S.A., Liu, P., Alanis-Lobato, G., Broccoli, V., Orlando, V., 2020. Transdifferentiation of Mouse Embryonic Fibroblasts into Dopaminergic Neurons Reactivates LINE-1 Repetitive Elements. Stem Cell Rep. 14, 60–74. https://doi.org/10.1016/j.stemcr.2019.12.002

Deniz, Ö., Ahmed, M., Todd, C.D., Rio-Machin, A., Dawson, M.A., Branco, M.R., 2020. Endogenous retroviruses are a source of enhancers with oncogenic potential in acute myeloid leukaemia. Nat. Commun. 11, 3506. https://doi.org/10.1038/s41467-020-17206-4

Denli, A.M., Narvaiza, I., Kerman, B.E., Pena, M., Benner, C., Marchetto, M.C.N., Diedrich, J.K., Aslanian, A., Ma, J., Moresco, J.J., Moore, L., Hunter, T., Saghatelian, A., Gage, F.H., 2015. Primate-specific ORF0 contributes to retrotransposon-mediated diversity. Cell 163, 583–593. https://doi.org/10.1016/j.cell.2015.09.025

Desrochers, L.M., Bordeleau, F., Reinhart-King, C.A., Cerione, R.A., Antonyak, M.A., 2016. Microvesicles provide a mechanism for intercellular communication by embryonic stem cells during embryo implantation. Nat. Commun. 7, 11958. https://doi.org/10.1038/ncomms11958

DeVries, T.A., Kalkofen, R.L., Matassa, A.A., Reyland, M.E., 2004. Protein Kinase Cδ Regulates Apoptosis via Activation of STAT1*. J. Biol. Chem. 279, 45603–45612. https://doi.org/10.1074/jbc.M407448200

Dewannieux, M., Esnault, C., Heidmann, T., 2003. LINE-mediated retrotransposition of marked Alu sequences. Nat. Genet. 35, 41–48. https://doi.org/10.1038/ng1223

Dewi Pamungkas Putri, D., Kawasaki, T., Murase, M., Sueyoshi, T., Deguchi, T., Ori, D., Suetsugu, S., Kawai, T., 2019. PtdIns3P phosphatases MTMR3 and MTMR4 negatively regulate innate immune responses to DNA through modulating STING trafficking. J. Biol. Chem. 294, 8412–8423. https://doi.org/10.1074/jbc.RA118.005731

Di Giacomo, M., Comazzetto, S., Saini, H., De Fazio, S., Carrieri, C., Morgan, M., Vasiliauskaite, L., Benes, V., Enright, A.J., O'Carroll, D., 2013. Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. Mol. Cell 50, 601–608. https://doi.org/10.1016/j.molcel.2013.04.026

Di Paolo, N.C., Doronin, K., Baldwin, L.K., Papayannopoulou, T., Shayakhmetov, D.M., 2013. The Transcription Factor IRF3 Triggers "Defensive Suicide" Necrosis in Response to Viral and Bacterial Pathogens. Cell Rep. 3, 1840–1846. https://doi.org/10.1016/j.celrep.2013.05.025

Diamond, M.S., Farzan, M., 2013. The broad-spectrum antiviral functions of IFIT and IFITM proteins. Nat. Rev. Immunol. 13, 46–57. https://doi.org/10.1038/nri3344

Diner, E.J., Burdette, D.L., Wilson, S.C., Monroe, K.M., Kellenberger, C.A., Hyodo, M., Hayakawa, Y., Hammond, M.C., Vance, R.E., 2013. The innate immune DNA sensor cGAS produces a noncanonical cyclic dinucleotide that activates human STING. Cell Rep. 3, 1355–1361. https://doi.org/10.1016/j.celrep.2013.05.009

Ding, J., Wang, S., Liu, Qipeng, Duan, Y., Cheng, T., Ye, Z., Cui, Z., Zhang, A., Liu, Qiuyu, Zhang, Z., Zhang, N., Liu, Qian, An, N., Zhao, J., Yi, D., Li, Q., Wang, Jing, Zhang, Y., Ma, L., Guo, S., Wang, Jinhui, Liang, C., Zhou, J., Cen, S., Li, X., 2023. Schlafen-5 inhibits LINE-1 retrotransposition. iScience 26, 107968. https://doi.org/10.1016/j.isci.2023.107968

Dmitriev, S.E., Andreev, D.E., Terenin, I.M., Olovnikov, I.A., Prassolov, V.S., Merrick, W.C., Shatsky, I.N., 2007. Efficient translation initiation directed by the 900-nucleotide-long and GC-rich 5' untranslated region of the human retrotransposon LINE-1 mRNA is strictly cap dependent rather than internal ribosome entry site mediated. Mol. Cell. Biol. 27, 4685–4697. https://doi.org/10.1128/MCB.02138-06

Du, M., Chen, Z.J., 2018. DNA-induced liquid phase condensation of cGAS activates innate immune signaling. Science 361, 704–709. https://doi.org/10.1126/science.aat1022

Dunn, G.P., Bruce, A.T., Sheehan, K.C.F., Shankaran, V., Uppaluri, R., Bui, J.D., Diamond, M.S., Koebel, C.M., Arthur, C., White, J.M., Schreiber, R.D., 2005. A critical function for type I interferons in cancer immunoediting. Nat. Immunol. 6, 722–729. https://doi.org/10.1038/ni1213

Dunphy, G., Flannery, S.M., Almine, J.F., Connolly, D.J., Paulus, C., Jønsson, K.L., Jakobsen, M.R., Nevels, M.M., Bowie, A.G., Unterholzner, L., 2018. Non-canonical Activation of the DNA Sensing Adaptor STING by ATM and IFI16 Mediates NF-κB Signaling after Nuclear DNA Damage. Mol. Cell 71, 745-760.e5. https://doi.org/10.1016/j.molcel.2018.07.034

Durruthy-Durruthy, J., Sebastiano, V., Wossidlo, M., Cepeda, D., Cui, J., Grow, E.J., Davila, J., Mall, M., Wong, W.H., Wysocka, J., Au, K.F., Reijo Pera, R.A., 2016. The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming. Nat. Genet. 48, 44–52. https://doi.org/10.1038/ng.3449

Eggenberger, J., Blanco-Melo, D., Panis, M., Brennand, K.J., tenOever, B.R., 2019. Type I interferon response impairs differentiation potential of pluripotent stem cells. Proc. Natl. Acad. Sci. 116, 1384–1393.

Ergun, S.L., Fernandez, D., Weiss, T.M., Li, L., 2019. STING Polymer Structure Reveals Mechanisms for Activation, Hyperactivation, and Inhibition. Cell 178, 290-301.e10. https://doi.org/10.1016/j.cell.2019.05.036

Esnault, C., Maestre, J., Heidmann, T., 2000. Human LINE retrotransposons generate processed pseudogenes. Nat. Genet. 24, 363–367. https://doi.org/10.1038/74184

Evans, M.J., Kaufman, M.H., 1981. Establishment in culture of pluripotential cells from mouse embryos. Nature 292, 154–156. https://doi.org/10.1038/292154a0

Ewald, S.E., Lee, B.L., Lau, L., Wickliffe, K.E., Shi, G.-P., Chapman, H.A., Barton, G.M., 2008. The ectodomain of Toll-like receptor 9 is cleaved to generate a functional receptor. Nature 456, 658–662. https://doi.org/10.1038/nature07405

Ewing, A.D., Kazazian, H.H., 2010. High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. Genome Res. 20, 1262–1270. https://doi.org/10.1101/gr.106419.110

Fadloun, A., Le Gras, S., Jost, B., Ziegler-Birling, C., Takahashi, H., Gorab, E., Carninci, P., Torres-Padilla, M.-E., 2013. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. Nat. Struct. Mol. Biol. 20, 332–338. https://doi.org/10.1038/nsmb.2495

Fendereski, M., Neupane, B., Nazneen, F., Bai, F., Guo, Y.-L., 2022. Mouse Trophoblast Cells Can Provide IFN-Based Antiviral Protection to Embryonic Stem Cells via Paracrine Signaling. J. Immunol. 208, 2761–2770. https://doi.org/10.4049/jimmunol.2100679

Feng, Q., Moran, J.V., Kazazian, H.H., Boeke, J.D., 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87, 905–916. https://doi.org/10.1016/s0092-8674(00)81997-2

Ferguson, B.J., Mansur, D.S., Peters, N.E., Ren, H., Smith, G.L., 2012. DNA-PK is a DNA sensor for IRF-3-dependent innate immunity. eLife 1, e00047. https://doi.org/10.7554/eLife.00047

Flasch, D.A., Macia, Á., Sánchez, L., Ljungman, M., Heras, S.R., García-Pérez, J.L., Wilson, T.E., Moran, J.V., 2019. Genome-wide de novo L1 Retrotransposition Connects Endonuclease Activity with Replication. Cell 177, 837-851.e28. https://doi.org/10.1016/j.cell.2019.02.050

Fleming, T.P., 1987. A quantitative analysis of cell allocation to trophectoderm and inner cell mass in the mouse blastocyst. Dev. Biol. 119, 520–531. https://doi.org/10.1016/0012-1606(87)90055-8

Flemr, M., Malik, R., Franke, V., Nejepinska, J., Sedlacek, R., Vlahovicek, K., Svoboda, P., 2013. A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. Cell 155, 807–816. https://doi.org/10.1016/j.cell.2013.10.001

Fouse, S.D., Shen, Y., Pellegrini, M., Cole, S., Meissner, A., Van Neste, L., Jaenisch, R., Fan, G., 2008. Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. Cell Stem Cell 2, 160–169. https://doi.org/10.1016/j.stem.2007.12.011

Freeman, B.T., Sokolowski, M., Roy-Engel, A.M., Smither, M.E., Belancio, V.P., 2019. Identification of charged amino acids required for nuclear localization of human L1 ORF1 protein. Mob. DNA 10, 20. https://doi.org/10.1186/s13100-019-0159-2

Freeman, J.D., Goodchild, N.L., Mager, D.L., 1994. A modified indicator gene for selection of retrotransposition events in mammalian cells. BioTechniques 17, 46, 48–49, 52.

Friedli, M., Turelli, P., Kapopoulou, A., Rauwel, B., Castro-Díaz, N., Rowe, H.M., Ecco, G., Unzu, C., Planet, E., Lombardo, A., Mangeat, B., Wildhaber, B.E., Naldini, L., Trono, D., 2014. Loss of transcriptional control over endogenous retroelements during reprogramming to pluripotency. Genome Res. 24, 1251–1259. https://doi.org/10.1101/gr.172809.114

Friedrich, G., Soriano, P., 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Genes Dev. 5, 1513–1523. https://doi.org/10.1101/gad.5.9.1513

Frost, J.M., Amante, S.M., Okae, H., Jones, E.M., Ashley, B., Lewis, R.M., Cleal, J.K., Caley, M.P., Arima, T., Maffucci, T., Branco, M.R., 2023. Regulation of human trophoblast gene expression by endogenous retroviruses. Nat. Struct. Mol. Biol. 30, 527–538. https://doi.org/10.1038/s41594-023-00960-6

Gao, P., Ascano, M., Wu, Y., Barchet, W., Gaffney, B.L., Zillinger, T., Serganov, A.A., Liu, Y., Jones, R.A., Hartmann, G., Tuschl, T., Patel, D.J., 2013. Cyclic [G(2',5')pA(3',5')p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. Cell 153, 1094–1107. https://doi.org/10.1016/j.cell.2013.04.046

Gao, Q., Wang, S., Chen, X., Cheng, S., Zhang, Z., Li, F., Huang, L., Yang, Y., Zhou, B., Yue, D., Wang, D., Cao, L., Maimela, N.R., Zhang, B., Yu, J., Wang, L., Zhang, Y., 2019. Cancer-cell-secreted CXCL11 promoted CD8+ T cells infiltration through docetaxel-induced-release of HMGB1 in NSCLC. J. Immunother. Cancer 7, 42. https://doi.org/10.1186/s40425-019-0511-6

Garcia-Cañadas, M., Sanchez-Luque, F.J., Sanchez, L., Rojas, J., Garcia Perez, J.L., 2023. LINE-1 Retrotransposition Assays in Embryonic Stem Cells, in: Branco, M.R., de Mendoza Soler, A. (Eds.), Transposable Elements: Methods and Protocols, Methods in Molecular Biology. Springer US, New York, NY, pp. 257–309. https://doi.org/10.1007/978-1-0716-2883-6_13

Garcia-Perez, J.L., Marchetto, M.C.N., Muotri, A.R., Coufal, N.G., Gage, F.H., O'Shea, K.S., Moran, J.V., 2007. LINE-1 retrotransposition in human embryonic stem cells. Hum. Mol. Genet. 16, 1569–1577. https://doi.org/10.1093/hmg/ddm105

Garcia-Perez, J.L., Morell, M., Scheys, J.O., Kulpa, D.A., Morell, S., Carter, C.C., Hammer, G.D., Collins, K.L., O'Shea, K.S., Menendez, P., Moran, J.V., 2010. Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. Nature 466, 769–773. https://doi.org/10.1038/nature09209

Garza, R., Atacho, D.A.M., Adami, A., Gerdes, P., Vinod, M., Hsieh, P., Karlsson, O., Horvath, V., Johansson, P.A., Pandiloski, N., Matas-Fuentes, J., Quaegebeur, A., Kouli, A., Sharma, Y., Jönsson, M.E., Monni, E., Englund, E., Eichler, E.E., Gale Hammell, M., Barker, R.A., Kokaia, Z., Douse, C.H., Jakobsson, J., 2023. LINE-1 retrotransposons drive human neuronal transcriptome complexity and functional diversification. Sci. Adv. 9, eadh9543. https://doi.org/10.1126/sciadv.adh9543

Gasior, S.L., Roy-Engel, A.M., Deininger, P.L., 2008. ERCC1/XPF limits L1 retrotransposition. DNA Repair 7, 983–989. https://doi.org/10.1016/j.dnarep.2008.02.006

Gasior, S.L., Wakeman, T.P., Xu, B., Deininger, P.L., 2006. The human LINE-1 retrotransposon creates DNA double-strand breaks. J. Mol. Biol. 357, 1383–1393. https://doi.org/10.1016/j.jmb.2006.01.089

Gaudet, P., Livstone, M.S., Lewis, S.E., Thomas, P.D., 2011. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. Brief. Bioinform. 12, 449–462. https://doi.org/10.1093/bib/bbr042

Gázquez-Gutiérrez, A., Witteveldt, J., Heras, S.R., Macias, S., 2021. Sensing of transposable elements by the antiviral innate immune system. RNA 27, 735–752. https://doi.org/10.1261/rna.078721.121

Gessani, S., Conti, L., Cornò, M.D., Belardelli, F., 2014. Type I Interferons as Regulators of Human Antigen Presenting Cell Functions. Toxins 6, 1696–1723. https://doi.org/10.3390/toxins6061696

Gessulat, S., Schmidt, T., Zolg, D.P., Samaras, P., Schnatbaum, K., Zerweck, J., Knaute, T., Rechenberger, J., Delanghe, B., Huhmer, A., Reimer, U., Ehrlich, H.-C., Aiche, S., Kuster, B., Wilhelm, M., 2019. Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning. Nat. Methods 16, 509–518. https://doi.org/10.1038/s41592-019-0426-7

Gifford, W.D., Pfaff, S.L., Macfarlan, T.S., 2013a. Transposable elements as genetic regulatory substrates in early development. Trends Cell Biol. 23, 218–226. https://doi.org/10.1016/j.tcb.2013.01.001

Gifford, W.D., Pfaff, S.L., Macfarlan, T.S., 2013b. Transposable elements as genetic regulatory substrates in early development. Trends Cell Biol. 23, 218–226. https://doi.org/10.1016/j.tcb.2013.01.001

Gilbert, N., Lutz, S., Morrish, T.A., Moran, J.V., 2005. Multiple fates of L1 retrotransposition intermediates in cultured human cells. Mol. Cell. Biol. 25, 7780–7795. https://doi.org/10.1128/MCB.25.17.7780-7795.2005

Gilbert, N., Lutz-Prigge, S., Moran, J.V., 2002. Genomic Deletions Created upon LINE-1 Retrotransposition. Cell 110, 315–325. https://doi.org/10.1016/S0092-8674(02)00828-0

Gonugunta, V.K., Sakai, T., Pokatayev, V., Yang, K., Wu, J., Dobbs, N., Yan, N., 2017. Trafficking-Mediated STING Degradation Requires Sorting to Acidified Endolysosomes and Can Be Targeted to Enhance Anti-tumor Response. Cell Rep. 21, 3234–3242. https://doi.org/10.1016/j.celrep.2017.11.061

Goodier, J.L., 2016. Restricting retrotransposons: a review. Mob. DNA 7, 16. https://doi.org/10.1186/s13100-016-0070-z

Goodier, J.L., Cheung, L.E., Kazazian, H.H., 2013. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. Nucleic Acids Res. 41, 7401–7419. https://doi.org/10.1093/nar/gkt512

Goodier, J.L., Ostertag, E.M., Du, K., Kazazian, H.H., 2001. A novel active L1 retrotransposon subfamily in the mouse. Genome Res. 11, 1677–1685. https://doi.org/10.1101/gr.198301

Goodier, J.L., Ostertag, E.M., Engleka, K.A., Seleme, M.C., Kazazian Jr., H.H., 2004. A potential role for the nucleous in L1 retrotransposition. Hum. Mol. Genet. 13, 1041–1048. https://doi.org/10.1093/hmg/ddh118

Goodier, J.L., Zhang, L., Vetter, M.R., Kazazian, H.H., 2007. LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. Mol. Cell. Biol. 27, 6469–6483. https://doi.org/10.1128/MCB.00332-07

Gorgoulis, V., Adams, P.D., Alimonti, A., Bennett, D.C., Bischof, O., Bishop, C., Campisi, J., Collado, M., Evangelou, K., Ferbeyre, G., Gil, J., Hara, E., Krizhanovsky, V., Jurk, D., Maier, A.B., Narita, M., Niedernhofer, L., Passos, J.F., Robbins, P.D., Schmitt, C.A., Sedivy, J., Vougas, K., von Zglinicki, T., Zhou, D., Serrano, M., Demaria, M., 2019. Cellular Senescence: Defining a Path Forward. Cell 179, 813–827. https://doi.org/10.1016/j.cell.2019.10.005

Goubau, D., Deddouche, S., Reis e Sousa, C., 2013. Cytosolic sensing of viruses. Immunity 38, 855–869. https://doi.org/10.1016/j.immuni.2013.05.007

Goubau, D., Romieu-Mourez, R., Solis, M., Hernandez, E., Mesplède, T., Lin, R., Leaman, D., Hiscott, J., 2009. Transcriptional re-programming of primary macrophages reveals distinct apoptotic and anti-tumoral functions of IRF-3 and IRF-7. Eur. J. Immunol. 39, 527–540. https://doi.org/10.1002/eji.200838832

Goutières, F., Aicardi, J., Barth, P.G., Lebon, P., 1998. Aicardi-Goutières syndrome: an update and results of interferon-alpha studies. Ann. Neurol. 44, 900–907. https://doi.org/10.1002/ana.410440608

Gray, E.E., Winship, D., Snyder, J.M., Child, S.J., Geballe, A.P., Stetson, D.B., 2016. The AIM2-like Receptors Are Dispensable for the Interferon Response to Intracellular DNA. Immunity 45, 255–266. https://doi.org/10.1016/j.immuni.2016.06.015

Greenberg, M.V.C., Bourc'his, D., 2019. The diverse roles of DNA methylation in mammalian development and disease. Nat. Rev. Mol. Cell Biol. 20, 590–607. https://doi.org/10.1038/s41580-019-0159-6

Grow, E.J., Flynn, R.A., Chavez, S.L., Bayless, N.L., Wossidlo, M., Wesche, D.J., Martin, L., Ware, C.B., Blish, C.A., Chang, H.Y., Reijo Pera, R.A., Wysocka, J., 2015. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. Nature 522, 221–225. https://doi.org/10.1038/nature14308

Grudic, A., Jul-Larsen, A., Haring, S.J., Wold, M.S., Lønning, P.E., Bjerkvig, R., Bøe, S.O., 2007. Replication protein A prevents accumulation of single-stranded telomeric DNA in cells that use alternative lengthening of telomeres. Nucleic Acids Res. 35, 7267–7278. https://doi.org/10.1093/nar/gkm738

Gualtieri, A., Andreola, F., Sciamanna, I., Sinibaldi-Vallebona, P., Serafino, A., Spadafora, C., 2013. Increased expression and copy number amplification of LINE-1 and SINE B1 retrotransposable elements in murine mammary carcinoma progression. Oncotarget 4, 1882–1893. https://doi.org/10.18632/oncotarget.1188

Guey, B., Wischnewski, M., Decout, A., Makasheva, K., Kaynak, M., Sakar, M.S., Fierz, B., Ablasser, A., 2020. BAF restricts cGAS on nuclear DNA to prevent innate immune activation. Science 369, 823–828. https://doi.org/10.1126/science.aaw6421

Gui, X., Yang, H., Li, T., Tan, X., Shi, P., Li, M., Du, F., Chen, Z.J., 2019. Autophagy induction via STING trafficking is a primordial function of the cGAS pathway. Nature 567, 262–266. https://doi.org/10.1038/s41586-019-1006-9

Guinn, Z.P., Petro, T.M., 2019. Interferon regulatory factor 3 plays a role in macrophage responses to interferon-γ. Immunobiology 224, 565–574. https://doi.org/10.1016/j.imbio.2019.04.004

Guo, C., Jeong, H.-H., Hsieh, Y.-C., Klein, H.-U., Bennett, D.A., De Jager, P.L., Liu, Z., Shulman, J.M., 2018. Tau Activates Transposable Elements in Alzheimer's Disease. Cell Rep. 23, 2874–2880. https://doi.org/10.1016/j.celrep.2018.05.004

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., Smith, A., 2009. Klf4 reverts developmentally programmed restriction of ground state pluripotency. Dev. Camb. Engl. 136, 1063–1069. https://doi.org/10.1242/dev.030957

Gurung, C., Fendereski, M., Sapkota, K., Guo, J., Huang, F., Guo, Y.-L., 2021. Dicer represses the interferon response and the double-stranded RNA-activated protein kinase pathway in mouse embryonic stem cells. J. Biol. Chem. 296, 100264. https://doi.org/10.1016/j.jbc.2021.100264

Haahr, P., Hoffmann, S., Tollenaere, M.A.X., Ho, T., Toledo, L.I., Mann, M., Bekker-Jensen, S., Räschle, M., Mailand, N., 2016. Activation of the ATR kinase by the RPA-binding protein ETAA1. Nat. Cell Biol. 18, 1196–1207. https://doi.org/10.1038/ncb3422

Haasnoot, J., Westerhout, E.M., Berkhout, B., 2007. RNA interference against viruses: strike and counterstrike. Nat. Biotechnol. 25, 1435–1443. https://doi.org/10.1038/nbt1369

Haig, D., 2016. Transposable elements: Self-seekers of the germline, team-players of the soma. BioEssays News Rev. Mol. Cell. Dev. Biol. 38, 1158–1166. https://doi.org/10.1002/bies.201600125

Han, J.S., Boeke, J.D., 2004. A highly active synthetic mammalian retrotransposon. Nature 429, 314–318. https://doi.org/10.1038/nature02535

Han, K., Sen, S.K., Wang, J., Callinan, P.A., Lee, J., Cordaux, R., Liang, P., Batzer, M.A., 2005. Genomic rearrangements by LINE-1 insertion-mediated deletion in the human and chimpanzee lineages. Nucleic Acids Res. 33, 4040–4052. https://doi.org/10.1093/nar/gki718

Hancks, D.C., Kazazian, H.H., 2016. Roles for retrotransposon insertions in human disease. Mob. DNA 7, 9. https://doi.org/10.1186/s13100-016-0065-9

Hancks, D.C., Kazazian, H.H., 2012. Active human retrotransposons: variation and disease. Curr. Opin. Genet. Dev. 22, 191–203. https://doi.org/10.1016/j.gde.2012.02.006

Hande, K.R., 1998. Etoposide: four decades of development of a topoisomerase II inhibitor. Eur. J. Cancer 34, 1514–1521. https://doi.org/10.1016/S0959-8049(98)00228-7

Hansen, A.L., Mukai, K., Schopfer, F.J., Taguchi, T., Holm, C.K., 2019. STING palmitoylation as a therapeutic target. Cell. Mol. Immunol. 16, 236–241. https://doi.org/10.1038/s41423-019-0205-5

Haoudi, A., Semmes, O.J., Mason, J.M., Cannon, R.E., 2004. Retrotransposition-Competent Human LINE-1 Induces Apoptosis in Cancer Cells With Intact p53. J. Biomed. Biotechnol. 2004, 185–194. https://doi.org/10.1155/S1110724304403131

Hayes, A.M., Schiavo, L., Constantino-Casas, F., Desmas, I., Dobson, J., Draper, A., Elliot, J., Genain, M.-A., Wang, J., Murchison, E.P., 2023. Transmission of canine transmissible venereal tumour between two dogs in the UK. J. Small Anim. Pract. 64, 590–594. https://doi.org/10.1111/jsap.13607

Hayflick, L., 1965. THE LIMITED IN VITRO LIFETIME OF HUMAN DIPLOID CELL STRAINS. Exp. Cell Res. 37, 614–636. https://doi.org/10.1016/0014-4827(65)90211-9

Hayflick, L., Moorhead, P.S., 1961. The serial cultivation of human diploid cell strains. Exp. Cell Res. 25, 585–621. https://doi.org/10.1016/0014-4827(61)90192-6

He, S., Sharpless, N.E., 2017. Senescence in Health and Disease. Cell 169, 1000–1011. https://doi.org/10.1016/j.cell.2017.05.015

He, W., Wan, H., Hu, L., Chen, P., Wang, X., Huang, Z., Yang, Z.-H., Zhong, C.-Q., Han, J., 2015. Gasdermin D is an executor of pyroptosis and required for interleukin-1β secretion. Cell Res. 25, 1285–1298. https://doi.org/10.1038/cr.2015.139

Heil, F., Hemmi, H., Hochrein, H., Ampenberger, F., Kirschning, C., Akira, S., Lipford, G., Wagner, H., Bauer, S., 2004. Species-Specific Recognition of Single-Stranded RNA via Toll-like Receptor 7 and 8. Science 303, 1526–1529. https://doi.org/10.1126/science.1093620

Hemberger, M., Dean, W., Reik, W., 2009. Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. Nat. Rev. Mol. Cell Biol. 10, 526–537. https://doi.org/10.1038/nrm2727

Hendrickson, P.G., Doráis, J.A., Grow, E.J., Whiddon, J.L., Lim, J.-W., Wike, C.L., Weaver, B.D., Pflueger, C., Emery, B.R., Wilcox, A.L., Nix, D.A., Peterson, C.M., Tapscott, S.J., Carrell, D.T., Cairns, B.R., 2017. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. Nat. Genet. 49, 925–934. https://doi.org/10.1038/ng.3844

Heras, S.R., Macias, S., Plass, M., Fernandez, N., Cano, D., Eyras, E., Garcia-Perez, J.L., Cáceres, J.F., 2013. The Microprocessor controls the activity of mammalian retrotransposons. Nat. Struct. Mol. Biol. 20, 1173–1181. https://doi.org/10.1038/nsmb.2658

Herbst, F., Ball, C.R., Tuorto, F., Nowrouzi, A., Wang, W., Zavidij, O., Dieter, S.M., Fessler, S., Hoeven, F. van der, Kloz, U., Lyko, F., Schmidt, M., Kalle, C. von, Glimm, H., 2012. Extensive Methylation of Promoter Sequences Silences Lentiviral Transgene Expression During Stem Cell Differentiation In Vivo. Mol. Ther. 20, 1014–1021. https://doi.org/10.1038/mt.2012.46

Herranz, N., Gil, J., 2018. Mechanisms and functions of cellular senescence. J. Clin. Invest. 128, 1238–1246. https://doi.org/10.1172/JCI95148

Hertzog, J., Rehwinkel, J., 2020. Regulation and inhibition of the DNA sensor cGAS. EMBO Rep. 21, e51345. https://doi.org/10.15252/embr.202051345

Hertzog, P.J., Hwang, S.Y., Kola, I., 1994. Role of interferons in the regulation of cell proliferation, differentiation, and development. Mol. Reprod. Dev. 39, 226–232. https://doi.org/10.1002/mrd.1080390216

Herzner, A.-M., Hagmann, C.A., Goldeck, M., Wolter, S., Kübler, K., Wittmann, S., Gramberg, T., Andreeva, L., Hopfner, K.-P., Mertens, C., Zillinger, T., Jin, T., Xiao, T.S., Bartok, E., Coch, C., Ackermann, D., Hornung, V., Ludwig, J., Barchet, W., Hartmann, G., Schlee, M., 2015. Sequence-specific activation of the DNA sensor cGAS by Y-form DNA structures as found in primary HIV-1 cDNA. Nat. Immunol. 16, 1025–1033. https://doi.org/10.1038/ni.3267

Honda, K., Takaoka, A., Taniguchi, T., 2006. Type I Inteferon Gene Induction by the Interferon Regulatory Factor Family of Transcription Factors. Immunity 25, 349–360. https://doi.org/10.1016/j.immuni.2006.08.009

Hooks, J.J., Moutsopoulos, H.M., Geis, S.A., Stahl, N.I., Decker, J.L., Notkins, A.L., 1979. Immune interferon in the circulation of patients with autoimmune disease. N. Engl. J. Med. 301, 5–8. https://doi.org/10.1056/NEJM197907053010102

Hooper, M., Hardy, K., Handyside, A., Hunter, S., Monk, M., 1987. HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature 326, 292–295. https://doi.org/10.1038/326292a0

Hornung, V., Ablasser, A., Charrel-Dennis, M., Bauernfeind, F., Horvath, G., Caffrey, D.R., Latz, E., Fitzgerald, K.A., 2009. AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. Nature 458, 514–518. https://doi.org/10.1038/nature07725

Hornung, V., Hartmann, R., Ablasser, A., Hopfner, K.-P., 2014. OAS proteins and cGAS: unifying concepts in sensing and responding to cytosolic nucleic acids. Nat. Rev. Immunol. 14, 521–528. https://doi.org/10.1038/nri3719

Hou, Y., Liang, H., Rao, E., Zheng, W., Huang, X., Deng, L., Zhang, Y., Yu, X., Xu, M., Mauceri, H., Arina, A., Weichselbaum, R.R., Fu, Y.-X., 2018. Non-canonical NF-κB Antagonizes STING Sensor-Mediated DNA Sensing in Radiotherapy. Immunity 49, 490-503.e4. https://doi.org/10.1016/j.immuni.2018.07.008

Hu, S., Li, J., Xu, F., Mei, S., Le Duff, Y., Yin, L., Pang, X., Cen, S., Jin, Q., Liang, C., Guo, F., 2015. SAMHD1 Inhibits LINE-1 Retrotransposition by Promoting Stress Granule Formation. PLOS Genet. 11, e1005367. https://doi.org/10.1371/journal.pgen.1005367

Huang, S., Tao, X., Yuan, S., Zhang, Yuhang, Li, P., Beilinson, H.A., Zhang, Ya, Yu, W., Pontarotti, P., Escriva, H., Le Petillon, Y., Liu, X., Chen, S., Schatz, D.G., Xu, A., 2016. Discovery of an Active RAG Transposon Illuminates the Origins of V(D)J Recombination. Cell 166, 102–114. https://doi.org/10.1016/j.cell.2016.05.032

Hubackova, S., Krejcikova, K., Bartek, J., Hodny, Z., 2012. IL1- and TGFβ-Nox4 signaling, oxidative stress and DNA damage response are shared features of replicative, oncogene-induced, and drug-induced paracrine 'Bystander senescence.' Aging 4, 932–951.

Iglesias-Guimarais, V., Ahrends, T., de Vries, E., Knobeloch, K.-P., Volkov, A., Borst, J., 2020. IFN-Stimulated Gene 15 Is an Alarmin that Boosts the CTL Response via an Innate,

NK Cell–Dependent Route. J. Immunol. Author Choice 204, 2110–2121. https://doi.org/10.4049/jimmunol.1901410

Imbeault, M., Helleboid, P.-Y., Trono, D., 2017. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. Nature 543, 550–554. https://doi.org/10.1038/nature21683

Ishak, C.A., Marshall, A.E., Passos, D.T., White, C.R., Kim, S.J., Cecchini, M.J., Ferwati, S., MacDonald, W.A., Howlett, C.J., Welch, I.D., Rubin, S.M., Mann, M.R.W., Dick, F.A., 2016. An RB-EZH2 Complex Mediates Silencing of Repetitive DNA Sequences. Mol. Cell 64, 1074–1087. https://doi.org/10.1016/j.molcel.2016.10.021

Ishii, K.J., Kawagoe, T., Koyama, S., Matsui, K., Kumar, H., Kawai, T., Uematsu, S., Takeuchi, O., Takeshita, F., Coban, C., Akira, S., 2008. TANK-binding kinase-1 delineates innate and adaptive immune responses to DNA vaccines. Nature 451, 725–729. https://doi.org/10.1038/nature06537

Iskow, R.C., McCabe, M.T., Mills, R.E., Torene, S., Pittard, W.S., Neuwald, A.F., Van Meir, E.G., Vertino, P.M., Devine, S.E., 2010. Natural mutagenesis of human genomes by endogenous retrotransposons. Cell 141, 1253–1261. https://doi.org/10.1016/j.cell.2010.05.020

Ivanov, P., Kedersha, N., Anderson, P., 2019. Stress Granules and Processing Bodies in Translational Control. Cold Spring Harb. Perspect. Biol. 11, a032813. https://doi.org/10.1101/cshperspect.a032813

Jachowicz, J.W., 2023. Epigenetic Manipulation of Transposable and Repetitive Elements, in: Branco, M.R., de Mendoza Soler, A. (Eds.), Transposable Elements: Methods and Protocols, Methods in Molecular Biology. Springer US, New York, NY, pp. 355–368. https://doi.org/10.1007/978-1-0716-2883-6_16

Jachowicz, J.W., Bing, X., Pontabry, J., Bošković, A., Rando, O.J., Torres-Padilla, M.-E., 2017. LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. Nat. Genet. 49, 1502–1510. https://doi.org/10.1038/ng.3945

Jachowicz, J.W., Torres-Padilla, M.-E., 2016. LINEs in mice: features, families, and potential roles in early development. Chromosoma 125, 29–39. https://doi.org/10.1007/s00412-015-0520-2

Jacob-Hirsch, J., Eyal, E., Knisbacher, B.A., Roth, J., Cesarkas, K., Dor, C., Farage-Barhom, S., Kunik, V., Simon, A.J., Gal, M., Yalon, M., Moshitch-Moshkovitz, S., Tearle, R., Constantini, S., Levanon, E.Y., Amariglio, N., Rechavi, G., 2018. Whole-genome sequencing reveals principles of brain retrotransposition in neurodevelopmental disorders. Cell Res. 28, 187–203. https://doi.org/10.1038/cr.2018.8

Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., Parker, R., 2016. ATPase modulated stress granules contain a diverse proteome and substructure. Cell 164, 487–498. https://doi.org/10.1016/j.cell.2015.12.038

Jangam, D., Feschotte, C., Betrán, E., 2017. Transposable Element Domestication As an Adaptation to Evolutionary Conflicts. Trends Genet., Transposable Elements 33, 817–831. https://doi.org/10.1016/j.tig.2017.07.011

Janowicz, A., Michalak, M., Krebs, J., 2011. Stress induced subcellular distribution of ALG-2, RBM22 and hSlu7. Biochim. Biophys. Acta 1813, 1045–1049. https://doi.org/10.1016/j.bbamcr.2010.11.010

Jenuwein, T., Allis, C.D., 2001. Translating the histone code. Science 293, 1074–1080. https://doi.org/10.1126/science.1063127

Jeremiah, N., Neven, B., Gentili, M., Callebaut, I., Maschalidi, S., Stolzenberg, M.-C., Goudin, N., Frémond, M.-L., Nitschke, P., Molina, T.J., Blanche, S., Picard, C., Rice, G.I., Crow, Y.J., Manel, N., Fischer, A., Bader-Meunier, B., Rieux-Laucat, F., 2014. Inherited STING-activating mutation underlies a familial inflammatory syndrome with lupus-like manifestations. J. Clin. Invest. 124, 5516–5520. https://doi.org/10.1172/JCI79100

Jia, M., Qin, D., Zhao, C., Chai, L., Yu, Z., Wang, W., Tong, L., Lv, L., Wang, Y., Rehwinkel, J., Yu, J., Zhao, W., 2020. Redox homeostasis maintained by GPX4 facilitates STING activation. Nat. Immunol. 21, 727–735. https://doi.org/10.1038/s41590-020-0699-0

Jiménez-Vacas, J.M., Montero-Hidalgo, A.J., Gómez-Gómez, E., Sáez-Martínez, P., Fuentes-Fayos, A.C., Closa, A., González-Serrano, T., Martínez-López, A., Sánchez-Sánchez, R., López-Casas, P.P., Sarmento-Cabral, A., Olmos, D., Eyras, E., Castaño, J.P., Gahete, M.D., Luque, R.M., 2023. Tumor suppressor role of RBM22 in prostate cancer acting as a dual-factor regulating alternative splicing and transcription of key oncogenic genes. Transl. Res. 253, 68–79. https://doi.org/10.1016/j.trsl.2022.08.016

Jin, L., Xu, L.-G., Yang, I.V., Davidson, E.J., Schwartz, D.A., Wurfel, M.M., Cambier, J.C., 2011. Identification and characterization of a loss-of-function human MPYS variant. Genes Immun. 12, 263–269. https://doi.org/10.1038/gene.2010.75

Kamiyama, H., Suzuki, K., Maeda, T., Koizumi, K., Miyaki, Y., Okada, S., Kawamura, Y.J., Samuelsson, J.K., Alonso, S., Konishi, F., Perucho, M., 2012. DNA demethylation in normal colon tissue predicts predisposition to multiple cancers. Oncogene 31, 5029–5037. https://doi.org/10.1038/onc.2011.652

Kang, J., Wu, J., Liu, Q., Wu, X., Zhao, Y., Ren, J., 2022. Post-Translational Modifications of STING: A Potential Therapeutic Target. Front. Immunol. 13, 888147. https://doi.org/10.3389/fimmu.2022.888147

Kano, H., Godoy, I., Courtney, C., Vetter, M.R., Gerton, G.L., Ostertag, E.M., Kazazian, H.H., 2009. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. Genes Dev. 23, 1303–1312. https://doi.org/10.1101/gad.1803909

Karpova, A.Y., Trost, M., Murray, J.M., Cantley, L.C., Howley, P.M., 2002. Interferon regulatory factor-3 is an in vivo target of DNA-PK. Proc. Natl. Acad. Sci. 99, 2818–2823. https://doi.org/10.1073/pnas.052713899

Kawai, T., Akira, S., 2010. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nat. Immunol. 11, 373–384. https://doi.org/10.1038/ni.1863

Kawai, T., Akira, S., 2006. Innate immune recognition of viral infection. Nat. Immunol. 7, 131–137. https://doi.org/10.1038/ni1303

Kayagaki, N., Stowe, I.B., Lee, B.L., O'Rourke, K., Anderson, K., Warming, S., Cuellar, T., Haley, B., Roose-Girma, M., Phung, Q.T., Liu, P.S., Lill, J.R., Li, H., Wu, J., Kummerfeld, S., Zhang, J., Lee, W.P., Snipas, S.J., Salvesen, G.S., Morris, L.X., Fitzgerald, L., Zhang, Y., Bertram, E.M., Goodnow, C.C., Dixit, V.M., 2015. Caspase-11 cleaves gasdermin D for non-

canonical inflammasome signalling. Nature 526, 666–671. https://doi.org/10.1038/nature15541

Kazazian, H.H., Wong, C., Youssoufian, H., Scott, A.F., Phillips, D.G., Antonarakis, S.E., 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332, 164–166. https://doi.org/10.1038/332164a0

Kelly, M., Lihua, S., Zhe, Z., Li, S., Yoselin, P., Michelle, P., E, S.K., 2018. Transposable Element Dysregulation in Systemic Lupus Erythematosus and Regulation by Histone Conformation and Hsp90. Clin. Immunol. Orlando Fla 197, 6–18. https://doi.org/10.1016/j.clim.2018.08.011

Kennedy, E.M., Whisnant, A.W., Kornepati, A.V.R., Marshall, J.B., Bogerd, H.P., Cullen, B.R., 2015. Production of functional small interfering RNAs by an amino-terminal deletion mutant of human Dicer. Proc. Natl. Acad. Sci. U. S. A. 112, E6945–E6954. https://doi.org/10.1073/pnas.1513421112

Keskitalo, S., Haapaniemi, E., Einarsdottir, E., Rajamäki, K., Heikkilä, H., Ilander, M., Pöyhönen, M., Morgunova, E., Hokynar, K., Lagström, S., Kivirikko, S., Mustjoki, S., Eklund, K., Saarela, J., Kere, J., Seppänen, M.R.J., Ranki, A., Hannula-Jouppi, K., Varjosalo, M., 2019. Novel TMEM173 Mutation and the Role of Disease Modifying Alleles. Front. Immunol. 10, 2770. https://doi.org/10.3389/fimmu.2019.02770

Kim, H.S., Lee, M.-S., 2007. STAT1 as a key modulator of cell death. Cell. Signal. 19, 454–465. https://doi.org/10.1016/j.cellsig.2006.09.003

Kim, R.H., Flanders, K.C., Birkey Reffey, S., Anderson, L.A., Duckett, C.S., Perkins, N.D., Roberts, A.B., 2001. SNIP1 inhibits NF-kappa B signaling by competing for its binding to the C/H1 domain of CBP/p300 transcriptional co-activators. J. Biol. Chem. 276, 46297–46304. https://doi.org/10.1074/jbc.M103819200

Kinomoto, M., Kanno, T., Shimura, M., Ishizaka, Y., Kojima, A., Kurata, T., Sata, T., Tokunaga, K., 2007. All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. Nucleic Acids Res. 35, 2955–2964. https://doi.org/10.1093/nar/gkm181

Klafter, R., Arbiser, J.L., 2000. Regulation of angiogenesis and tumorigenesis by signal transduction cascades: lessons from benign and malignant endothelial tumors. J. Investig. Dermatol. Symp. Proc. 5, 79–82. https://doi.org/10.1046/j.1087-0024.2000.00007.x

Kohlrausch, F.B., Berteli, T.S., Wang, F., Navarro, P.A., Keefe, D.L., 2022. Control of LINE-1 Expression Maintains Genome Integrity in Germline and Early Embryo Development. Reprod. Sci. 29, 328–340. https://doi.org/10.1007/s43032-021-00461-1

Konno, H., Konno, K., Barber, G.N., 2013. Cyclic dinucleotides trigger ULK1 (ATG1) phosphorylation of STING to prevent sustained innate immune signaling. Cell 155, 688–698. https://doi.org/10.1016/j.cell.2013.09.049

Kopera, H.C., Flasch, D.A., Nakamura, M., Miyoshi, T., Doucet, A.J., Moran, J.V., 2016. LEAP: L1 Element Amplification Protocol. Methods Mol. Biol. Clifton NJ 1400, 339–355. https://doi.org/10.1007/978-1-4939-3372-3_21

Kornberg, R.D., 1977. Structure of Chromatin. Annu. Rev. Biochem. 46, 931–954. https://doi.org/10.1146/annurev.bi.46.070177.004435

Kotredes, K.P., Gamero, A.M., 2013. Interferons as Inducers of Apoptosis in Malignant Cells. J. Interferon Cytokine Res. 33, 162–170. https://doi.org/10.1089/jir.2012.0110

Krishnamurthy, J., Torrice, C., Ramsey, M.R., Kovalev, G.I., Al-Regaiey, K., Su, L., Sharpless, N.E., 2004. Ink4a/Arf expression is a biomarker of aging. J. Clin. Invest. 114, 1299–1307. https://doi.org/10.1172/JCI22475

Kruse, K., Díaz, N., Enriquez-Gasca, R., Gaume, X., Torres-Padilla, M.-E., Vaquerizas, J.M., 2019. Transposable elements drive reorganisation of 3D chromatin during early embryogenesis. https://doi.org/10.1101/523712

Kuilman, T., Michaloglou, C., Vredeveld, L.C.W., Douma, S., van Doorn, R., Desmet, C.J., Aarden, L.A., Mooi, W.J., Peeper, D.S., 2008. Oncogene-Induced Senescence Relayed by an Interleukin-Dependent Inflammatory Network. Cell 133, 1019–1031. https://doi.org/10.1016/j.cell.2008.03.039

Kulpa, D.A., Moran, J.V., 2006. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. Nat. Struct. Mol. Biol. 13, 655–660. https://doi.org/10.1038/nsmb1107

Lanier, L.L., 2008. Evolutionary struggles between NK cells and viruses. Nat. Rev. Immunol. 8, 259–268. https://doi.org/10.1038/nri2276

Lau, L., Gray, E.E., Brunette, R.L., Stetson, D.B., 2015. DNA tumor virus oncogenes antagonize the cGAS-STING DNA-sensing pathway. Science 350, 568–571. https://doi.org/10.1126/science.aab3291

Laureau, R., Dyatel, A., Dursuk, G., Brown, S., Adeoye, H., Yue, J.-X., Chiara, M.D., Harris, A., Ünal, E., Liti, G., Adams, I.R., Berchowitz, L.E., 2021. Meiotic Cells Counteract Programmed Retrotransposon Activation via RNA-Binding Translational Repressor Assemblies. Dev. Cell. https://doi.org/10.1016/j.devcel.2020.11.008

Leaman, D.W., Chawla-Sarkar, M., Jacobs, B., Vyas, K., Sun, Y., Ozdemir, A., Yi, T., Williams, B.R., Borden, E.C., 2003. Novel growth and death related interferon-stimulated genes (ISGs) in melanoma: greater potency of IFN-beta compared with IFN-alpha2. J. Interferon Cytokine Res. Off. J. Int. Soc. Interferon Cytokine Res. 23, 745–756. https://doi.org/10.1089/107999003772084860

Leaman, D.W., Chawla-Sarkar, M., Vyas, K., Reheman, M., Tamai, K., Toji, S., Borden, E.C., 2002. Identification of X-linked inhibitor of apoptosis-associated factor-1 as an interferon-stimulated gene that augments TRAIL Apo2L-induced apoptosis. J. Biol. Chem. 277, 28504–28511. https://doi.org/10.1074/jbc.M204851200

Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L.J., Lohr, J.G., Harris, C.C., Ding, L., Wilson, R.K., Wheeler, D.A., Gibbs, R.A., Kucherlapati, R., Lee, C., Kharchenko, P.V., Park, P.J., Cancer Genome Atlas Research Network, 2012. Landscape of somatic retrotransposition in human cancers. Science 337, 967–971. https://doi.org/10.1126/science.1222077

Lee-Kirsch, M.A., Gong, M., Chowdhury, D., Senenko, L., Engel, K., Lee, Y.-A., de Silva, U., Bailey, S.L., Witte, T., Vyse, T.J., Kere, J., Pfeiffer, C., Harvey, S., Wong, A., Koskenmies, S., Hummel, O., Rohde, K., Schmidt, R.E., Dominiczak, A.F., Gahr, M., Hollis, T., Perrino, F.W., Lieberman, J., Hübner, N., 2007. Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus. Nat. Genet. 39, 1065–1067. https://doi.org/10.1038/ng2091

Lev Maor, G., Yearim, A., Ast, G., 2015. The alternative role of DNA methylation in splicing regulation. Trends Genet. 31, 274–280. https://doi.org/10.1016/j.tig.2015.03.002

Levin, D., Harari, D., Schreiber, G., 2011. Stochastic Receptor Expression Determines Cell Fate upon Interferon Treatment. Mol. Cell. Biol. 31, 3252–3266. https://doi.org/10.1128/MCB.05251-11

Li, C., Chi, S., He, N., Zhang, X., Guicherit, O., Wagner, R., Tyring, S., Xie, J., 2004. IFNα induces Fas expression and apoptosis in hedgehog pathway activated BCC cells through inhibiting Ras-Erk signaling. Oncogene 23, 1608–1617. https://doi.org/10.1038/sj.onc.1207273

Li, G., Xiang, Y., Sabapathy, K., Silverman, R.H., 2004. An Apoptotic Signaling Pathway in the Interferon Antiviral Response Mediated by RNase L and c-Jun NH2-terminal Kinase*. J. Biol. Chem. 279, 1123–1131. https://doi.org/10.1074/jbc.M305893200

Li, P., Nijhawan, D., Budihardjo, I., Srinivasula, S.M., Ahmad, M., Alnemri, E.S., Wang, X., 1997. Cytochrome c and dATP-Dependent Formation of Apaf-1/Caspase-9 Complex Initiates an Apoptotic Protease Cascade. Cell 91, 479–489. https://doi.org/10.1016/S0092-8674(00)80434-1

Li, Qian, Zhang, Y., Fu, J., Han, L., Xue, L., Lv, C., Wang, P., Li, G., Tong, T., 2013. FOXA1 mediates p16INK4a activation during cellular senescence. EMBO J. 32, 858–873. https://doi.org/10.1038/emboj.2013.35

Li, S., Qian, N., Jiang, C., Zu, W., Liang, A., Li, M., Elledge, S.J., Tan, X., 2022. Gain-of-function genetic screening identifies the antiviral function of TMEM120A via STING activation. Nat. Commun. 13, 105. https://doi.org/10.1038/s41467-021-27670-1

Li, T., Huang, T., Du, M., Chen, X., Du, F., Ren, J., Chen, Z.J., 2021. Phosphorylation and chromatin tethering prevent cGAS activation during mitosis. Science 371, eabc5386. https://doi.org/10.1126/science.abc5386

Li, T.D., Murano, K., Kitano, T., Guo, Y., Negishi, L., Siomi, H., 2022. TDP-43 safeguards the embryo genome from L1 retrotransposition. Sci. Adv. 8, eabq3806. https://doi.org/10.1126/sciadv.abq3806

Li, Wanhe, Prazak, L., Chatterjee, N., Grüninger, S., Krug, L., Theodorou, D., Dubnau, J., 2013. Activation of transposable elements during aging and neuronal decline in Drosophila. Nat. Neurosci. 16, 529–531. https://doi.org/10.1038/nn.3368

Li, Xin, Shu, C., Yi, G., Chaton, C.T., Shelton, C.L., Diao, J., Zuo, X., Kao, C.C., Herr, A.B., Li, P., 2013. Cyclic GMP-AMP Synthase Is Activated by Double-Stranded DNA-Induced Oligomerization. Immunity 39, 1019–1031. https://doi.org/10.1016/j.immuni.2013.10.019

Li, Xiaoyu, Zhang, J., Jia, R., Cheng, V., Xu, X., Qiao, W., Guo, F., Liang, C., Cen, S., 2013. The MOV10 Helicase Inhibits LINE-1 Mobility. J. Biol. Chem. 288, 21148–21160. https://doi.org/10.1074/jbc.M113.465856

Li, Z., Duan, S., Hua, X., Xu, X., Li, Y., Menolfi, D., Zhou, H., Lu, C., Zha, S., Goff, S.P., Zhang, Z., 2023. Asymmetric distribution of parental H3K9me3 in S phase silences L1 elements. Nature 1–9. https://doi.org/10.1038/s41586-023-06711-3

Liang, L., Cao, C., Ji, L., Cai, Z., Wang, D., Ye, R., Chen, J., Yu, X., Zhou, J., Bai, Z., Wang, R., Yang, X., Zhu, P., Xue, Y., 2023. Complementary Alu sequences mediate enhancer–promoter selectivity. Nature 619, 868–875. https://doi.org/10.1038/s41586-023-06323-x

Lin, R., Heylbroeck, C., Genin, P., Pitha, P.M., Hiscott, J., 1999. Essential role of interferon regulatory factor 3 in direct activation of RANTES chemokine transcription. Mol. Cell. Biol. 19, 959–966. https://doi.org/10.1128/MCB.19.2.959

Lin, R., Heylbroeck, C., Pitha, P.M., Hiscott, J., 1998. Virus-Dependent Phosphorylation of the IRF-3 Transcription Factor Regulates Nuclear Translocation, Transactivation Potential, and Proteasome-Mediated Degradation. Mol. Cell. Biol. 18, 2986–2996. https://doi.org/10.1128/MCB.18.5.2986

Lin, Y.L., Shivji, M.K., Chen, C., Kolodner, R., Wood, R.D., Dutta, A., 1998. The evolutionarily conserved zinc finger motif in the largest subunit of human replication protein A is required for DNA replication and mismatch repair but not for nucleotide excision repair. J. Biol. Chem. 273, 1453–1461. https://doi.org/10.1074/jbc.273.3.1453

Liu, D., Wu, H., Wang, C., Li, Y., Tian, H., Siraj, S., Sehgal, S.A., Wang, X., Wang, J., Shang, Y., Jiang, Z., Liu, L., Chen, Q., 2019. STING directly activates autophagy to tune the innate immune response. Cell Death Differ. 26, 1735–1749. https://doi.org/10.1038/s41418-018-0251-z

Liu, E.Y., Russ, J., Cali, C.P., Phan, J.M., Amlie-Wolf, A., Lee, E.B., 2019. Loss of Nuclear TDP-43 Is Associated with Decondensation of LINE Retrotransposons. Cell Rep. 27, 1409-1421.e6. https://doi.org/10.1016/j.celrep.2019.04.003

Liu, J., Gao, M., He, J., Wu, K., Lin, S., Jin, L., Chen, Y., Liu, H., Shi, J., Wang, X., Chang, L., Lin, Y., Zhao, Y.-L., Zhang, X., Zhang, M., Luo, G.-Z., Wu, G., Pei, D., Wang, J., Bao, X., Chen, J., 2021. The RNA m6A reader YTHDC1 silences retrotransposons and guards ES cell identity. Nature 591, 322–326. https://doi.org/10.1038/s41586-021-03313-9

Liu, L., Botos, I., Wang, Y., Leonard, J.N., Shiloach, J., Segal, D.M., Davies, D.R., 2008. Structural basis of toll-like receptor 3 signaling with double-stranded RNA. Science 320, 379–381. https://doi.org/10.1126/science.1155406

Liu, N., Lee, C.H., Swigut, T., Grow, E., Gu, B., Bassik, M.C., Wysocka, J., 2018. Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. Nature 553, 228–232. https://doi.org/10.1038/nature25179

Liu, S., Brind'Amour, J., Karimi, M.M., Shirane, K., Bogutz, A., Lefebvre, L., Sasaki, H., Shinkai, Y., Lorincz, M.C., 2014. Setdb1 is required for germline development and silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells. Genes Dev. 28, 2041–2055. https://doi.org/10.1101/gad.244848.114

Liu, S., Cai, X., Wu, J., Cong, Q., Chen, X., Li, T., Du, F., Ren, J., Wu, Y.-T., Grishin, N.V., Chen, Z.J., 2015. Phosphorylation of innate immune adaptor proteins MAVS, STING, and TRIF induces IRF3 activation. Science 347, aaa2630. https://doi.org/10.1126/science.aaa2630

Liu, X., Huang, J., Chen, T., Wang, Y., Xin, S., Li, J., Pei, G., Kang, J., 2008. Yamanaka factors critically regulate the developmental signaling network in mouse embryonic stem cells. Cell Res. 18, 1177–1189. https://doi.org/10.1038/cr.2008.309

Liu, Y., Jesus, A.A., Marrero, B., Yang, D., Ramsey, S.E., Montealegre Sanchez, G.A., Tenbrock, K., Wittkowski, H., Jones, O.Y., Kuehn, H.S., Lee, C.-C.R., DiMattia, M.A., Cowen, E.W., Gonzalez, B., Palmer, I., DiGiovanna, J.J., Biancotto, A., Kim, H., Tsai, W.L., Trier, A.M., Huang, Y., Stone, D.L., Hill, S., Kim, H.J., St. Hilaire, C., Gurprasad, S., Plass, N., Chapelle, D., Horkayne-Szakaly, I., Foell, D., Barysenka, A., Candotti, F., Holland, S.M., Hughes, J.D., Mehmet, H., Issekutz, A.C., Raffeld, M., McElwee, J., Fontana, J.R., Minniti, C.P., Moir, S., Kastner, D.L., Gadina, M., Steven, A.C., Wingfield, P.T., Brooks, S.R., Rosenzweig, S.D., Fleisher, T.A., Deng, Z., Boehm, M., Paller, A.S., Goldbach-Mansky, R., 2014. Activated STING in a Vascular and Pulmonary Syndrome. N. Engl. J. Med. 371, 507–518. https://doi.org/10.1056/NEJMoa1312625

Liu, Y.-J., 2005. IPC: professional type 1 interferon-producing cells and plasmacytoid dendritic cell precursors. Annu. Rev. Immunol. 23, 275–306. https://doi.org/10.1146/annurev.immunol.23.021704.115633

Locksley, R.M., Killeen, N., Lenardo, M.J., 2001. The TNF and TNF Receptor Superfamilies: Integrating Mammalian Biology. Cell 104, 487–501. https://doi.org/10.1016/S0092-8674(01)00237-9

Long, E.O., 2008. Negative signaling by inhibitory receptors: the NK cell paradigm. Immunol. Rev. 224, 70–84. https://doi.org/10.1111/j.1600-065X.2008.00660.x

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Lu, J.Y., Chang, L., Li, T., Wang, T., Yin, Y., Zhan, G., Han, X., Zhang, K., Tao, Y., Percharde, M., Wang, L., Peng, Q., Yan, P., Zhang, H., Bi, X., Shao, W., Hong, Y., Wu, Z., Ma, R., Wang, P., Li, W., Zhang, J., Chang, Z., Hou, Y., Zhu, B., Ramalho-Santos, M., Li, P., Xie, W., Na, J., Sun, Y., Shen, X., 2021. Homotypic clustering of L1 and B1/Alu repeats compartmentalizes the 3D genome. Cell Res. 31, 613–630. https://doi.org/10.1038/s41422-020-00466-6

Lu, J.Y., Shao, W., Chang, L., Yin, Y., Li, T., Zhang, H., Hong, Y., Percharde, M., Guo, L., Wu, Z., Liu, L., Liu, W., Yan, P., Ramalho-Santos, M., Sun, Y., Shen, X., 2020. Genomic Repeats Categorize Genes with Distinct Functions for Orchestrated Regulation. Cell Rep. 30, 3296-3311.e5. https://doi.org/10.1016/j.celrep.2020.02.048

Luan, D.D., Korman, M.H., Jakubczak, J.L., Eickbush, T.H., 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72, 595–605. https://doi.org/10.1016/0092-8674(93)90078-5

Luecke, S., Holleufer, A., Christensen, M.H., Jønsson, K.L., Boni, G.A., Sørensen, L.K., Johannsen, M., Jakobsen, M.R., Hartmann, R., Paludan, S.R., 2017. cGAS is activated by DNA in a length-dependent manner. EMBO Rep. 18, 1707–1715. https://doi.org/10.15252/embr.201744017

Lujambio, A., Akkari, L., Simon, J., Grace, D., Tschaharganeh, D.F., Bolden, J.E., Zhao, Z., Thapar, V., Joyce, J.A., Krizhanovsky, V., Lowe, S.W., 2013. Non-cell-autonomous tumor suppression by p53. Cell 153, 449–460. https://doi.org/10.1016/j.cell.2013.03.020

Lum, L.S., Sultzman, L.A., Kaufman, R.J., Linzer, D.I., Wu, B.J., 1990. A cloned human CCAAT-box-binding factor stimulates transcription from the human hsp70 promoter. Mol. Cell. Biol. 10, 6709–6717. https://doi.org/10.1128/mcb.10.12.6709-6717.1990

Lüningschrör, P., Stöcker, B., Kaltschmidt, B., Kaltschmidt, C., 2012. miR-290 Cluster Modulates Pluripotency by Repressing Canonical NF-κB Signaling. STEM CELLS 30, 655–664. https://doi.org/10.1002/stem.1033

Luo, W.-W., Li, S., Li, C., Lian, H., Yang, Q., Zhong, B., Shu, H.-B., 2016a. iRhom2 is essential for innate immunity to DNA viruses by mediating trafficking and stability of the adaptor STING. Nat. Immunol. 17, 1057–1066. https://doi.org/10.1038/ni.3510

Luo, W.-W., Lian, H., Zhong, B., Shu, H.-B., Li, S., 2016b. Krüppel-like factor 4 negatively regulates cellular antiviral immune response. Cell. Mol. Immunol. 13, 65–72. https://doi.org/10.1038/cmi.2014.125

Luqman-Fatah, A., Watanabe, Y., Uno, K., Ishikawa, F., Moran, J.V., Miyoshi, T., 2023. The interferon stimulated gene-encoded protein HELZ2 inhibits human LINE-1 retrotransposition and LINE-1 RNA-mediated type I interferon induction. Nat. Commun. 14, 203. https://doi.org/10.1038/s41467-022-35757-6

Luster, A.D., 2002. The role of chemokines in linking innate and adaptive immunity. Curr. Opin. Immunol. 14, 129–135. https://doi.org/10.1016/s0952-7915(01)00308-9

M, M., Kenneth, Casey, W., 2022. Janeway's Immunobiology: Tenth International Student Edition with Registration Card. W.W. Norton & Company.

Ma, E., MacRae, I.J., Kirsch, J.F., Doudna, J.A., 2008. Autoinhibition of human dicer by its internal helicase domain. J. Mol. Biol. 380, 237–243. https://doi.org/10.1016/j.jmb.2008.05.005

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., Pfaff, S.L., 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature 487, 57–63. https://doi.org/10.1038/nature11244

Macia, A., Widmann, T.J., Heras, S.R., Ayllon, V., Sanchez, L., Benkaddour-Boumzaouad, M., Muñoz-Lopez, M., Rubio, A., Amador-Cubero, S., Blanco-Jimenez, E., Garcia-Castro, J., Menendez, P., Ng, P., Muotri, A.R., Goodier, J.L., Garcia-Perez, J.L., 2017. Engineered LINE-1 retrotransposition in nondividing human neurons. Genome Res. 27, 335–348. https://doi.org/10.1101/gr.206805.116

MacLennan, M., García-Cañadas, M., Reichmann, J., Khazina, E., Wagner, G., Playfoot, C.J., Salvador-Palomeque, C., Mann, A.R., Peressini, P., Sanchez, L., Dobie, K., Read, D., Hung, C.-C., Eskeland, R., Meehan, R.R., Weichenrieder, O., García-Pérez, J.L., Adams, I.R., 2017. Mobilization of LINE-1 retrotransposons is restricted by Tex19.1 in mouse embryonic stem cells. eLife 6, e26152. https://doi.org/10.7554/eLife.26152

MacRae, I.J., Ma, E., Zhou, M., Robinson, C.V., Doudna, J.A., 2008. In vitro reconstitution of the human RISC-loading complex. Proc. Natl. Acad. Sci. U. S. A. 105, 512–517. https://doi.org/10.1073/pnas.0710869105

Maillard, P.V., Ciaudo, C., Marchais, A., Li, Y., Jay, F., Ding, S.W., Voinnet, O., 2013. Antiviral RNA Interference in Mammalian Cells. Science 342, 235–238. https://doi.org/10.1126/science.1241930

Maillard, P.V., Van der Veen, A.G., Deddouche-Grass, S., Rogers, N.C., Merits, A., Reis e Sousa, C., 2016. Inactivation of the type I interferon pathway reveals long double-stranded RNA-mediated RNA interference in mammalian cells. EMBO J. 35, 2505–2518. https://doi.org/10.15252/embj.201695086

Maillard, P.V., van der Veen, A.G., Poirier, E.Z., Reis e Sousa, C., 2019. Slicing and dicing viruses: antiviral RNA interference in mammals. EMBO J. 38, e100941. https://doi.org/10.15252/embj.2018100941

Malki, S., van der Heijden, G.W., O'Donnell, K.A., Martin, S.L., Bortvin, A., 2014. A Role for Retrotransposon LINE-1 in Fetal Oocyte Attrition in Mice. Dev. Cell 29, 521–533. https://doi.org/10.1016/j.devcel.2014.04.027

Mankan, A.K., Schmidt, T., Chauhan, D., Goldeck, M., Höning, K., Gaidt, M., Kubarenko, A.V., Andreeva, L., Hopfner, K.-P., Hornung, V., 2014. Cytosolic RNA:DNA hybrids activate the cGAS-STING axis. EMBO J. 33, 2937–2946. https://doi.org/10.15252/embj.201488726

Marasca, F., Sinha, S., Vadalà, R., Polimeni, B., Ranzani, V., Paraboschi, E.M., Burattin, F.V., Ghilotti, M., Crosti, M., Negri, M.L., Campagnoli, S., Notarbartolo, S., Sartore-Bianchi, A., Siena, S., Prati, D., Montini, G., Viale, G., Torre, O., Harari, S., Grifantini, R., Soldà, G., Biffo, S., Abrignani, S., Bodega, B., 2022. LINE1 are spliced in non-canonical transcript variants to regulate T cell quiescence and exhaustion. Nat. Genet. 54, 180–193. https://doi.org/10.1038/s41588-021-00989-7

Marcus, A., Mao, A.J., Lensink-Vasan, M., Wang, L., Vance, R.E., Raulet, D.H., 2018. Tumor-Derived cGAMP Triggers a STING-Mediated Interferon Response in Non-tumor Cells to Activate the NK Cell Response. Immunity 49, 754-763.e4. https://doi.org/10.1016/j.immuni.2018.09.016

Maréchal, A., Li, J.-M., Ji, X.Y., Wu, C.-S., Yazinski, S.A., Nguyen, H.D., Liu, S., Jiménez, A.E., Jin, J., Zou, L., 2014. PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATR activation via a ubiquitin-mediated circuitry. Mol. Cell 53, 235–246. https://doi.org/10.1016/j.molcel.2013.11.002

Martens, J.H.A., O'Sullivan, R.J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., Jenuwein, T., 2005. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. EMBO J. 24, 800–812. https://doi.org/10.1038/sj.emboj.7600545

Martin, G.R., 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. U. S. A. 78, 7634–7638. https://doi.org/10.1073/pnas.78.12.7634

Martin, S.L., Bushman, F.D., 2001. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. Mol. Cell. Biol. 21, 467–475. https://doi.org/10.1128/MCB.21.2.467-475.2001

Matfsson, R., Matfsson, A., Holmdahl, R., Scheynius, A., Van Der Meide, P.H., 1992. In Vivo Treatment with Interferon-Gamma during Early Pregnancy in Mice Induces Strong Expression of Major Histocompatibility Complex Class I and II Molecules in Uterus and Decidua But not in Extra-Embryonic Tissues1. Biol. Reprod. 46, 1176–1186. https://doi.org/10.1095/biolreprod46.6.1176

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., Yokota, T., 1999. STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. EMBO J. 18, 4261–4269. https://doi.org/10.1093/emboj/18.15.4261

Matsui, T., Leung, D., Miyashita, H., Maksakova, I.A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M.C., Shinkai, Y., 2010. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature 464, 927–931. https://doi.org/10.1038/nature08858

Mavragani, C.P., Sagalovskiy, I., Guo, Q., Nezos, A., Kapsogeorgou, E.K., Lu, P., Liang Zhou, J., Kirou, K.A., Seshan, S.V., Moutsopoulos, H.M., Crow, M.K., 2016. Expression of Long Interspersed Nuclear Element 1 Retroelements and Induction of Type I Interferon in Patients With Systemic Autoimmune Disease. Arthritis Rheumatol. Hoboken NJ 68, 2686–2696. https://doi.org/10.1002/art.39795

Mayer, W., Niveleau, A., Walter, J., Fundele, R., Haaf, T., 2000. Demethylation of the zygotic paternal genome. Nature 403, 501–502. https://doi.org/10.1038/35000656

McCLINTOCK, B., 1950. The origin and behavior of mutable loci in maize. Proc. Natl. Acad. Sci. U. S. A. 36, 344–355. https://doi.org/10.1073/pnas.36.6.344

McIlwain, D.R., Berger, T., Mak, T.W., 2013. Caspase Functions in Cell Death and Disease. Cold Spring Harb. Perspect. Biol. 5, a008656. https://doi.org/10.1101/cshperspect.a008656

McNab, F., Mayer-Barber, K., Sher, A., Wack, A., O'Garra, A., 2015. Type I interferons in infectious disease. Nat. Rev. Immunol. 15, 87–103. https://doi.org/10.1038/nri3787

McWhirter, S.M., Fitzgerald, K.A., Rosains, J., Rowe, D.C., Golenbock, D.T., Maniatis, T., 2004. IFN-regulatory factor 3-dependent gene expression is defective in Tbk1-deficient mouse embryonic fibroblasts. Proc. Natl. Acad. Sci. 101, 233–238. https://doi.org/10.1073/pnas.2237236100

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., Gnirke, A., Jaenisch, R., Lander, E.S., 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454, 766–770. https://doi.org/10.1038/nature07107

Meurs, E., Chong, K., Galabru, J., Thomas, N.S.B., Kerr, I.M., Williams, B.R.G., Hovanessian, A.G., 1990. Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon. Cell 62, 379–390. https://doi.org/10.1016/0092-8674(90)90374-N

Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.-Y., Edouard, P., Howes, S., Keith, J.C., McCoy, J.M., 2000. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature 403, 785–789. https://doi.org/10.1038/35001608

Miao, B., Fu, S., Lyu, C., Gontarz, P., Wang, T., Zhang, B., 2020. Tissue-specific usage of transposable element-derived promoters in mouse development. Genome Biol. 21, 255. https://doi.org/10.1186/s13059-020-02164-3

Miki, Y., Nishisho, I., Horii, A., Miyoshi, Y., Utsunomiya, J., Kinzler, K.W., Vogelstein, B., Nakamura, Y., 1992. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. Cancer Res. 52, 643–645.

Miotto, B., Chibi, M., Xie, P., Koundrioukoff, S., Moolman-Smook, H., Pugh, D., Debatisse, M., He, F., Zhang, L., Defossez, P.-A., 2014. The RBBP6/ZBTB38/MCM10 axis regulates DNA replication and common fragile site stability. Cell Rep. 7, 575–587. https://doi.org/10.1016/j.celrep.2014.03.030

Miroci, H., Schob, C., Kindler, S., Ölschläger-Schütt, J., Fehr, S., Jungenitz, T., Schwarzacher, S.W., Bagni, C., Mohr, E., 2012. Makorin Ring Zinc Finger Protein 1 (MKRN1), a Novel Poly(A)-binding Protein-interacting Protein, Stimulates Translation in Nerve Cells. J. Biol. Chem. 287, 1322–1334. https://doi.org/10.1074/jbc.M111.315291

Mita, P., Sun, X., Fenyö, D., Kahler, D.J., Li, D., Agmon, N., Wudzinska, A., Keegan, S., Bader, J.S., Yun, C., Boeke, J.D., 2020. BRCA1 and S phase DNA repair pathways restrict LINE-1 retrotransposition in human cells. Nat. Struct. Mol. Biol. 27, 179–191. https://doi.org/10.1038/s41594-020-0374-z

Mita, P., Wudzinska, A., Sun, X., Andrade, J., Nayak, S., Kahler, D.J., Badri, S., LaCava, J., Ueberheide, B., Yun, C.Y., Fenyö, D., Boeke, J.D., 2018. LINE-1 protein localization and functional dynamics during the cell cycle. eLife 7, e30058. https://doi.org/10.7554/eLife.30058

Miyawaki, S., Kuroki, S., Maeda, R., Okashita, N., Koopman, P., Tachibana, M., 2020. The mouse Sry locus harbors a cryptic exon that is essential for male sex determination. Science 370, 121–124. https://doi.org/10.1126/science.abb6430

Miyoshi, T., Makino, T., Moran, J.V., 2019. Poly(ADP-Ribose) Polymerase 2 Recruits Replication Protein A to Sites of LINE-1 Integration to Facilitate Retrotransposition. Mol. Cell 75, 1286-1298.e12. https://doi.org/10.1016/j.molcel.2019.07.018

Modzelewski, A.J., Gan Chong, J., Wang, T., He, L., 2022. Mammalian genome innovation through transposon domestication. Nat. Cell Biol. 24, 1332–1340. https://doi.org/10.1038/s41556-022-00970-4

Modzelewski, A.J., Shao, W., Chen, J., Lee, A., Qi, X., Noon, M., Tjokro, K., Sales, G., Biton, A., Anand, A., Speed, T.P., Xuan, Z., Wang, T., Risso, D., He, L., 2021. A mouse-specific retrotransposon drives a conserved Cdk2ap1 isoform essential for development. Cell 184, 5541-5558.e22. https://doi.org/10.1016/j.cell.2021.09.021

Moldovan, J.B., Moran, J.V., 2015. The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition. PLOS Genet. 11, e1005121. https://doi.org/10.1371/journal.pgen.1005121

Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D., Kazazian, H.H., 1996. High frequency retrotransposition in cultured mammalian cells. Cell 87, 917–927. https://doi.org/10.1016/s0092-8674(00)81998-4

Morchikh, M., Cribier, A., Raffel, R., Amraoui, S., Cau, J., Severac, D., Dubois, E., Schwartz, O., Bennasser, Y., Benkirane, M., 2017. HEXIM1 and NEAT1 Long Non-coding RNA Form a Multi-subunit Complex that Regulates DNA-Mediated Innate Immune Response. Mol. Cell 67, 387-399.e5. https://doi.org/10.1016/j.molcel.2017.06.020

Morrish, T.A., Gilbert, N., Myers, J.S., Vincent, B.J., Stamato, T.D., Taccioli, G.E., Batzer, M.A., Moran, J.V., 2002. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nat. Genet. 31, 159–165. https://doi.org/10.1038/ng898

Morse, B., Rotherg, P.G., South, V.J., Spandorfer, J.M., Astrin, S.M., 1988. Insertional mutagenesis of the myc locus by a LINE-1 sequence in a human breast carcinoma. Nature 333, 87–90. https://doi.org/10.1038/333087a0

Mostoslavsky, R., Chua, K.F., Lombard, D.B., Pang, W.W., Fischer, M.R., Gellon, L., Liu, P., Mostoslavsky, G., Franco, S., Murphy, M.M., Mills, K.D., Patel, P., Hsu, J.T., Hong, A.L., Ford, E., Cheng, H.-L., Kennedy, C., Nunez, N., Bronson, R., Frendewey, D., Auerbach, W., Valenzuela, D., Karow, M., Hottiger, M.O., Hursting, S., Barrett, J.C., Guarente, L., Mulligan, R., Demple, B., Yancopoulos, G.D., Alt, F.W., 2006. Genomic Instability and Aging-like Phenotype in the Absence of Mammalian SIRT6. Cell 124, 315–329. https://doi.org/10.1016/j.cell.2005.11.044

Motwani, M., Pesiridis, S., Fitzgerald, K.A., 2019. DNA sensing by the cGAS–STING pathway in health and disease. Nat. Rev. Genet. 20, 657–674. https://doi.org/10.1038/s41576-019-0151-1

Mouse Genome Sequencing Consortium, Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S.E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M.R., Brown, D.G., Brown, S.D., Bult, C., Burton, J., Butler, J., Campbell, R.D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A.T., Church, D.M., Clamp, M., Clee, C., Collins, F.S., Cook, L.L., Copley, R.R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K.D., Deri, J., Dermitzakis, E.T., Dewey, C., Dickens, N.J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D.M., Eddy, S.R., Elnitski, L., Emes, R.D., Eswara, P., Eyras, E., Felsenfeld, A., Fewell, G.A., Flicek, P., Foley, K., Frankel, W.N., Fulton, L.A., Fulton, R.S., Furey, T.S., Gage, D., Gibbs, R.A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Grafham, D., Graves, T.A., Green, E.D., Gregory, S., Guigó, R., Guyer, M., Hardison, R.C., Haussler, D., Hayashizaki, Y., Hillier, L.W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D.B., Johnson, L.S., Jones, M., Jones, T.A., Joy, A., Kamal, M., Karlsson, E.K., Karolchik, D., Kasprzyk, A., Kawai, J., Keibler, E., Kells, C., Kent, W.J., Kirby, A., Kolbe, D.L., Korf, I., Kucherlapati, R.S., Kulbokas, E.J., Kulp, D., Landers, T., Leger, J.P., Leonard, S., Letunic, I., Levine, R., Li, J., Li, M., Lloyd, C., Lucas, S., Ma, B., Maglott, D.R., Mardis, E.R., Matthews, L., Mauceli, E., Mayer, J.H., McCarthy, M., McCombie, W.R., McLaren, S., McLay, K., McPherson, J.D., Meldrim, J., Meredith, B., Mesirov, J.P., Miller, W., Miner, T.L., Mongin, E., Montgomery, K.T., Morgan, M., Mott, R., Mullikin, J.C., Muzny, D.M., Nash, W.E., Nelson, J.O., Nhan, M.N., Nicol, R., Ning, Z., Nusbaum, C., O'Connor, M.J., Okazaki, Y., Oliver, K., Overton-Larty, E., Pachter, L., Parra, G., Pepin, K.H., Peterson, J., Pevzner, P., Plumb, R., Pohl, C.S., Poliakov, A., Ponce, T.C., Ponting, C.P., Potter, S., Quail, M., Reymond, A., Roe, B.A., Roskin, K.M., Rubin, E.M., Rust, A.G., Santos, R., Sapojnikov, V., Schultz, B., Schultz, J., Schwartz, M.S., Schwartz, S., Scott, C., Seaman, S., Searle, S., Sharpe, T., Sheridan, A., Shownkeen, R., Sims, S., Singer, J.B., Slater, G., Smit, A., Smith, D.R., Spencer, B., Stabenau, A., Stange-Thomann, N., Sugnet, C., Suyama, M., Tesler, G., Thompson, J., Torrents, D., Trevaskis, E., Tromp, J., Ucla, C., Ureta-Vidal, A., Vinson, J.P., Von Niederhausern, A.C., Wade, C.M., Wall, M., Weber, R.J., Weiss, R.B., Wendl, M.C., West, A.P., Wetterstrand, K., Wheeler, R., Whelan, S., Wierzbowski, J., Willey, D., Williams, S., Wilson, R.K., Winter, E., Worley, K.C., Wyman, D., Yang, S., Yang, S.-P., Zdobnov, E.M., Zody, M.C., Lander, E.S., 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520–562. https://doi.org/10.1038/nature01262

Mukai, K., Konno, H., Akiba, T., Uemura, T., Waguri, S., Kobayashi, T., Barber, G.N., Arai, H., Taguchi, T., 2016. Activation of STING requires palmitoylation at the Golgi. Nat. Commun. 7, 11932. https://doi.org/10.1038/ncomms11932

Mukai, K., Ogawa, E., Uematsu, R., Kuchitsu, Y., Kiku, F., Uemura, T., Waguri, S., Suzuki, T., Dohmae, N., Arai, H., Shum, A.K., Taguchi, T., 2021. Homeostatic regulation of STING by retrograde membrane traffic to the ER. Nat. Commun. 12, 61. https://doi.org/10.1038/s41467-020-20234-9

Mungall, C.J., McMurry, J.A., Köhler, S., Balhoff, J.P., Borromeo, C., Brush, M., Carbon, S., Conlin, T., Dunn, N., Engelstad, M., Foster, E., Gourdine, J.P., Jacobsen, J.O.B., Keith, D., Laraway, B., Lewis, S.E., NguyenXuan, J., Shefchek, K., Vasilevsky, N., Yuan, Z., Washington, N., Hochheiser, H., Groza, T., Smedley, D., Robinson, P.N., Haendel, M.A., 2017. The Monarch Initiative: an integrative data and analytic platform connecting phenotypes to genotypes across species. Nucleic Acids Res. 45, D712–D722. https://doi.org/10.1093/nar/gkw1128

Muotri, A.R., 2016. L1 Retrotransposition in Neural Progenitor Cells, in: Garcia-Pérez, J.L. (Ed.), Transposons and Retrotransposons: Methods and Protocols, Methods in Molecular Biology. Springer, New York, NY, pp. 157–163. https://doi.org/10.1007/978-1-4939-3372-3_11

Muotri, A.R., Chu, V.T., Marchetto, M.C.N., Deng, W., Moran, J.V., Gage, F.H., 2005. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature 435, 903–910. https://doi.org/10.1038/nature03663

Muotri, A.R., Marchetto, M.C.N., Coufal, N.G., Oefner, R., Yeo, G., Nakashima, K., Gage, F.H., 2010. L1 retrotransposition in neurons is modulated by MeCP2. Nature 468, 443–446. https://doi.org/10.1038/nature09544

Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., Hannon, G.J., 2005. Characterization of Dicer-deficient murine embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. 102, 12135–12140. https://doi.org/10.1073/pnas.0505479102

Nabet, B., Roberts, J.M., Buckley, D.L., Paulk, J., Dastjerdi, S., Yang, A., Leggett, A.L., Erb, M.A., Lawlor, M.A., Souza, A., Scott, T.G., Vittori, S., Perry, J.A., Qi, J., Winter, G.E., Wong, K.-K., Gray, N.S., Bradner, J.E., 2018. The dTAG system for immediate and target-specific protein degradation. Nat. Chem. Biol. 14, 431–441. https://doi.org/10.1038/s41589-018-0021-8

Nagy, A., Gócza, E., Diaz, E.M., Prideaux, V.R., Iványi, E., Markkula, M., Rossant, J., 1990. Embryonic stem cells alone are able to support fetal development in the mouse. Dev. Camb. Engl. 110, 815–821. https://doi.org/10.1242/dev.110.3.815

Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W., Roder, J.C., 1993. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. 90, 8424–8428. https://doi.org/10.1073/pnas.90.18.8424

Nelson, D.E., Ihekwaba, A.E.C., Elliott, M., Johnson, J.R., Gibney, C.A., Foreman, B.E., Nelson, G., See, V., Horton, C.A., Spiller, D.G., Edwards, S.W., McDowell, H.P., Unitt, J.F., Sullivan, E., Grimley, R., Benson, N., Broomhead, D., Kell, D.B., White, M.R.H., 2004. Oscillations in NF-κB Signaling Control the Dynamics of Gene Expression. Science 306, 704–708. https://doi.org/10.1126/science.1099962

Nelson, G., Wordsworth, J., Wang, C., Jurk, D., Lawless, C., Martin-Ruiz, C., von Zglinicki, T., 2012. A senescent cell bystander effect: senescence-induced senescence. Aging Cell 11, 345–349. https://doi.org/10.1111/j.1474-9726.2012.00795.x

Newkirk, S.J., An, W., 2017. L1 Regulation in Mouse and Human Germ Cells, in: Cristofari, G. (Ed.), Human Retrotransposons in Health and Disease. Springer International Publishing, Cham, pp. 29–61. https://doi.org/10.1007/978-3-319-48344-3_2

Newkirk, S.J., Lee, S., Grandi, F.C., Gaysinskaya, V., Rosser, J.M., Vanden Berg, N., Hogarth, C.A., Marchetto, M.C.N., Muotri, A.R., Griswold, M.D., Ye, P., Bortvin, A., Gage, F.H., Boeke, J.D., An, W., 2017. Intact piRNA pathway prevents L1 mobilization in male meiosis. Proc. Natl. Acad. Sci. U. S. A. 114, E5635–E5644. https://doi.org/10.1073/pnas.1701069114

Newton, J.C., Naik, M.T., Li, G.Y., Murphy, E.L., Fawzi, N.L., Sedivy, J.M., Jogl, G., 2021. Phase separation of the LINE-1 ORF1 protein is mediated by the N-terminus and coiled-coil domain. Biophys. J. 120, 2181–2191. https://doi.org/10.1016/j.bpj.2021.03.028

Nguyen, D.T.T., Richter, D., Michel, G., Mitschka, S., Kolanus, W., Cuevas, E., Wulczyn, F.G., 2017. The ubiquitin ligase LIN41/TRIM71 targets p53 to antagonize cell death and differentiation pathways during stem cell differentiation. Cell Death Differ. 24, 1063–1078. https://doi.org/10.1038/cdd.2017.54

Ni, G., Konno, H., Barber, G.N., 2017. Ubiquitination of STING at lysine 224 controls IRF3 activation. Sci. Immunol. 2, eaah7119. https://doi.org/10.1126/sciimmunol.aah7119

Niwa, H., Burdon, T., Chambers, I., Smith, A., 1998. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev. 12, 2048–2060.

Niwa, H., Ogawa, K., Shimosato, D., Adachi, K., 2009. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature 460, 118–122. https://doi.org/10.1038/nature08113

Notwell, J.H., Chung, T., Heavner, W., Bejerano, G., 2015. A family of transposable elements co-opted into developmental enhancers in the mouse neocortex. Nat. Commun. 6, 6644. https://doi.org/10.1038/ncomms7644

Novototskaya-Vlasova, K.A., Neznanov, N.S., Molodtsov, I., Hall, B.M., Commane, M., Gleiberman, A.S., Murray, J., Haber, M., Norris, M.D., Leonova, K.I., Gudkov, A.V., 2022. Inflammatory response to retrotransposons drives tumor drug resistance that can be prevented by reverse transcriptase inhibitors. Proc. Natl. Acad. Sci. U. S. A. 119, e2213146119. https://doi.org/10.1073/pnas.2213146119

Ohno, R., Nakayama, M., Naruse, C., Okashita, N., Takano, O., Tachibana, M., Asano, M., Saitou, M., Seki, Y., 2013. A replication-dependent passive mechanism modulates DNA demethylation in mouse primordial germ cells. Dev. Camb. Engl. 140, 2892–2903. https://doi.org/10.1242/dev.093229

Orecchini, E., Doria, M., Antonioni, A., Galardi, S., Ciafrè, S.A., Frassinelli, L., Mancone, C., Montaldo, C., Tripodi, M., Michienzi, A., 2017. ADAR1 restricts LINE-1 retrotransposition. Nucleic Acids Res. 45, 155–168. https://doi.org/10.1093/nar/gkw834

Ostertag, E.M., Kazazian Jr, H.H., 2001. Biology of Mammalian L1 Retrotransposons. Annu. Rev. Genet. 35, 501–538. https://doi.org/10.1146/annurev.genet.35.102401.091032

Ostertag, E.M., Luning Prak, E.T., DeBerardinis, R.J., Moran, J.V., Kazazian, H.H., 2000. Determination of L1 retrotransposition kinetics in cultured cells. Nucleic Acids Res. 28, 1418–1423.

Paludan, S.R., Bowie, A.G., 2013. Immune Sensing of DNA. Immunity 38, 870–880. https://doi.org/10.1016/j.immuni.2013.05.004

Panaretakis, T., Hjortsberg, L., Tamm, K.P., Björklund, A.-C., Joseph, B., Grandér, D., 2008. Interferon α Induces Nucleus-independent Apoptosis by Activating Extracellular Signal-regulated Kinase 1/2 and c-Jun NH2-Terminal Kinase Downstream of Phosphatidylinositol 3-Kinase and Mammalian Target of Rapamycin. Mol. Biol. Cell 19, 41–50. https://doi.org/10.1091/mbc.E07-04-0358

Parham, P., 2005. MHC class I molecules and KIRs in human history, health and survival. Nat. Rev. Immunol. 5, 201–214. https://doi.org/10.1038/nri1570

Pastuzyn, E.D., Day, C.E., Kearns, R.B., Kyrke-Smith, M., Taibi, A.V., McCormick, J., Yoder, N., Belnap, D.M., Erlendsson, S., Morado, D.R., Briggs, J.A.G., Feschotte, C., Shepherd, J.D., 2018. The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. Cell 172, 275-288.e18. https://doi.org/10.1016/j.cell.2017.12.024

Patel, S., Jin, L., 2019. TMEM173 variants and potential importance to human biology and disease. Genes Immun. 20, 82–89. https://doi.org/10.1038/s41435-018-0029-9

Pathare, G.R., Decout, A., Glück, S., Cavadini, S., Makasheva, K., Hovius, R., Kempf, G., Weiss, J., Kozicka, Z., Guey, B., Melenec, P., Fierz, B., Thomä, N.H., Ablasser, A., 2020. Structural mechanism of cGAS inhibition by the nucleosome. Nature 587, 668–672. https://doi.org/10.1038/s41586-020-2750-6

Payer, L.M., Steranka, J.P., Yang, W.R., Kryatova, M., Medabalimi, S., Ardeljan, D., Liu, C., Boeke, J.D., Avramopoulos, D., Burns, K.H., 2017. Structural variants caused by Alu insertions are associated with risks for many human diseases. Proc. Natl. Acad. Sci. U. S. A. 114, E3984–E3992. https://doi.org/10.1073/pnas.1704117114

Percharde, M., Bulut-Karslioglu, A., Ramalho-Santos, M., 2017. Hypertranscription in Development, Stem Cells, and Regeneration. Dev. Cell 40, 9–21. https://doi.org/10.1016/j.devcel.2016.11.010

Percharde, M., Lavial, F., Ng, J.-H., Kumar, V., Tomaz, R.A., Martin, N., Yeo, J.-C., Gil, J., Prabhakar, S., Ng, H.-H., Parker, M.G., Azuara, V., 2012. Ncoa3 functions as an essential Esrrb coactivator to sustain embryonic stem cell self-renewal and reprogramming. Genes Dev. 26, 2286–2298. https://doi.org/10.1101/gad.195545.112

Percharde, M., Lin, C.-J., Yin, Y., Guan, J., Peixoto, G.A., Bulut-Karslioglu, A., Biechele, S., Huang, B., Shen, X., Ramalho-Santos, M., 2018. A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. Cell 174, 391-405.e19. https://doi.org/10.1016/j.cell.2018.05.043

Perl, A., Colombo, E., Dai, H., Agarwal, R., Mark, K.A., Banki, K., Poiesz, B.J., Phillips, P.E., Hoch, S.O., Reveille, J.D., 1995. Antibody reactivity to the HRES-1 endogenous retroviral element identifies a subset of patients with systemic lupus erythematosus and overlap syndromes. Correlation with antinuclear antibodies and HLA class II alleles. Arthritis Rheum. 38, 1660–1671. https://doi.org/10.1002/art.1780381119

Perng, Y.-C., Lenschow, D.J., 2018. ISG15 in antiviral immunity and beyond. Nat. Rev. Microbiol. 16, 423–439. https://doi.org/10.1038/s41579-018-0020-5

Peteranderl, C., Herold, S., 2017. The Impact of the Interferon/TNF-Related Apoptosis-Inducing Ligand Signaling Axis on Disease Progression in Respiratory Viral Infection and Beyond. Front. Immunol. 8.

Peters, A.H.F.M., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A.H.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., Martens, J.H.A., Jenuwein, T., 2003. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol. Cell 12, 1577–1589. https://doi.org/10.1016/s1097-2765(03)00477-5

Petrasek, J., Iracheta-Vellve, A., Csak, T., Satishchandran, A., Kodys, K., Kurt-Jones, E.A., Fitzgerald, K.A., Szabo, G., 2013. STING-IRF3 pathway links endoplasmic reticulum stress

with hepatocyte apoptosis in early alcoholic liver disease. Proc. Natl. Acad. Sci. U. S. A. 110, 16544–16549. https://doi.org/10.1073/pnas.1308331110

Petro, T.M., 2020. IFN Regulatory Factor 3 in Health and Disease. J. Immunol. 205, 1981–1989. https://doi.org/10.4049/jimmunol.2000462

Pezic, D., Manakov, S.A., Sachidanandam, R., Aravin, A.A., 2014. piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. Genes Dev. 28, 1410–1428. https://doi.org/10.1101/gad.240895.114

Philippe, C., Vargas-Landin, D.B., Doucet, A.J., van Essen, D., Vera-Otarola, J., Kuciak, M., Corbin, A., Nigumann, P., Cristofari, G., 2016. Activation of individual L1 retrotransposon instances is restricted to cell-type dependent permissive loci. eLife 5, e13926. https://doi.org/10.7554/eLife.13926

Platanias, L.C., 2005. Mechanisms of type-I- and type-II-interferon-mediated signalling. Nat. Rev. Immunol. 5, 375–386. https://doi.org/10.1038/nri1604

Poirier, E.Z., Buck, M.D., Chakravarty, P., Carvalho, J., Frederico, B., Cardoso, A., Healy, L., Ulferts, R., Beale, R., Reis e Sousa, C., 2021. An isoform of Dicer protects mammalian stem cells against multiple RNA viruses. Science 373, 231–236. https://doi.org/10.1126/science.abg2264

Prabakaran, T., Bodda, C., Krapp, C., Zhang, B., Christensen, M.H., Sun, C., Reinert, L., Cai, Y., Jensen, S.B., Skouboe, M.K., Nyengaard, J.R., Thompson, C.B., Lebbink, R.J., Sen, G.C., van Loo, G., Nielsen, R., Komatsu, M., Nejsum, L.N., Jakobsen, M.R., Gyrd-Hansen, M., Paludan, S.R., 2018. Attenuation of cGAS-STING signaling is mediated by a p62/SQSTM1-dependent autophagy pathway activated by TBK1. EMBO J. 37, e97858. https://doi.org/10.15252/embj.201797858

Protasova, M.S., Andreeva, T.V., Rogaev, E.I., 2021. Factors Regulating the Activity of LINE1 Retrotransposons. Genes 12, 1562. https://doi.org/10.3390/genes12101562

Protasova, M.S., Gusev, F.E., Grigorenko, A.P., Kuznetsova, I.L., Rogaev, E.I., Andreeva, T.V., 2017. Quantitative analysis of L1-retrotransposons in Alzheimer's disease and aging. Biochem. Mosc. 82, 962–971. https://doi.org/10.1134/S0006297917080120

Provost, P., Dishart, D., Doucet, J., Frendewey, D., Samuelsson, B., Rådmark, O., 2002. Ribonuclease activity and RNA binding of recombinant human Dicer. EMBO J. 21, 5864–5874. https://doi.org/10.1093/emboj/cdf578

Qi, X., Li, T.-G., Hao, J., Hu, J., Wang, J., Simmons, H., Miura, S., Mishina, Y., Zhao, G.-Q., 2004. BMP4 supports self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways. Proc. Natl. Acad. Sci. U. S. A. 101, 6027–6032. https://doi.org/10.1073/pnas.0401367101

Raiz, J., Damert, A., Chira, S., Held, U., Klawitter, S., Hamdorf, M., Löwer, J., Strätling, W.H., Löwer, R., Schumann, G.G., 2012. The non-autonomous retrotransposon SVA is trans-mobilized by the human LINE-1 protein machinery. Nucleic Acids Res. 40, 1666–1683. https://doi.org/10.1093/nar/gkr863

Ralston, A., Rossant, J., 2008. Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. Dev. Biol. 313, 614–629. https://doi.org/10.1016/j.ydbio.2007.10.054

Ran, Y., Xiong, M., Xu, Z., Luo, W., Wang, S., Wang, Y.-Y., 2019. YIPF5 Is Essential for Innate Immunity to DNA Virus and Facilitates COPII-Dependent STING Trafficking. J. Immunol. 203, 1560–1570. https://doi.org/10.4049/jimmunol.1900387

Ravenscroft, J.C., Suri, M., Rice, G.I., Szynkiewicz, M., Crow, Y.J., 2011. Autosomal dominant inheritance of a heterozygous mutation in SAMHD1 causing familial chilblain lupus. Am. J. Med. Genet. A. 155A, 235–237. https://doi.org/10.1002/ajmg.a.33778

Reddy, E.P., Reynolds, R.K., Santos, E., Barbacid, M., 1982. A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature 300, 149–152. https://doi.org/10.1038/300149a0

Rehwinkel, J., Gack, M.U., 2020. RIG-I-like receptors: their regulation and roles in RNA sensing. Nat. Rev. Immunol. 20, 537–551. https://doi.org/10.1038/s41577-020-0288-3

Rice, G.I., Kasher, P.R., Forte, G.M.A., Mannion, N.M., Greenwood, S.M., Szynkiewicz, M., Dickerson, J.E., Bhaskar, S.S., Zampini, M., Briggs, T.A., Jenkinson, E.M., Bacino, C.A., Battini, R., Bertini, E., Brogan, P.A., Brueton, L.A., Carpanelli, M., De Laet, C., de Lonlay, P., del Toro, M., Desguerre, I., Fazzi, E., Garcia-Cazorla, A., Heiberg, A., Kawaguchi, M., Kumar, R., Lin, J.-P.S.-M., Lourenco, C.M., Male, A.M., Marques, W., Mignot, C., Olivieri, I., Orcesi, S., Prabhakar, P., Rasmussen, M., Robinson, R.A., Rozenberg, F., Schmidt, J.L., Steindl, K., Tan, T.Y., van der Merwe, W.G., Vanderver, A., Vassallo, G., Wakeling, E.L., Wassmer, E., Whittaker, E., Livingston, J.H., Lebon, P., Suzuki, T., McLaughlin, P.J., Keegan, L.P., O'Connell, M.A., Lovell, S.C., Crow, Y.J., 2012. Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type I interferon signature. Nat. Genet. 44, 1243–1248. https://doi.org/10.1038/ng.2414

Robbez-Masson, L., Tie, C.H.C., Conde, L., Tunbak, H., Husovsky, C., Tchasovnikarova, I.A., Timms, R.T., Herrero, J., Lehner, P.J., Rowe, H.M., 2018. The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes. Genome Res. 28, 836–845. https://doi.org/10.1101/gr.228171.117

Roberts, R.M., 1996. Interferon-tau and pregnancy. J. Interferon Cytokine Res. Off. J. Int. Soc. Interferon Cytokine Res. 16, 271–273. https://doi.org/10.1089/jir.1996.16.271

Roberts, S.A., Strande, N., Burkhalter, M.D., Strom, C., Havener, J.M., Hasty, P., Ramsden, D.A., 2010. Ku is a 5'-dRP/AP lyase that excises nucleotide damage near broken ends. Nature 464, 1214–1217. https://doi.org/10.1038/nature08926

Robertson, E., Bradley, A., Kuehn, M., Evans, M., 1986. Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. Nature 323, 445–448. https://doi.org/10.1038/323445a0

Rodić, N., Sharma, Reema, Sharma, Rajni, Zampella, J., Dai, L., Taylor, M.S., Hruban, R.H., Iacobuzio-Donahue, C.A., Maitra, A., Torbenson, M.S., Goggins, M., Shih, I.-M., Duffield, A.S., Montgomery, E.A., Gabrielson, E., Netto, G.J., Lotan, T.L., De Marzo, A.M., Westra, W., Binder, Z.A., Orr, B.A., Gallia, G.L., Eberhart, C.G., Boeke, J.D., Harris, C.R., Burns, K.H., 2014. Long interspersed element-1 protein expression is a hallmark of many human cancers. Am. J. Pathol. 184, 1280–1286. https://doi.org/10.1016/j.ajpath.2014.01.007

Romeike, M., Spach, S., Huber, M., Feng, S., Vainorius, G., Elling, U., Versteeg, G.A., Buecker, C., 2022. Transient upregulation of IRF1 during exit from naive pluripotency confers viral protection. EMBO Rep. 23, e55375. https://doi.org/10.15252/embr.202255375

Romero-Brey, I., Bartenschlager, R., 2014. Membranous replication factories induced by plus-strand RNA viruses. Viruses 6, 2826–2857. https://doi.org/10.3390/v6072826

Rookhuizen, D.C., Bonte, P.-E., Ye, M., Hoyler, T., Gentili, M., Burgdorf, N., Durand, S., Aprahamian, F., Kroemer, G., Manel, N., Waterfall, J.J., Milne, R., Goudot, C., Towers, G.J., Amigorena, S., 2021. Induction of transposable element expression is central to innate sensing. https://doi.org/10.1101/2021.09.10.457789

Roovers, E.F., Rosenkranz, D., Mahdipour, M., Han, C.-T., He, N., Chuva de Sousa Lopes, S.M., van der Westerlaken, L.A.J., Zischler, H., Butter, F., Roelen, B.A.J., Ketting, R.F., 2015. Piwi proteins and piRNAs in mammalian oocytes and early embryos. Cell Rep. 10, 2069–2082. https://doi.org/10.1016/j.celrep.2015.02.062

Rosebeck, S., Leaman, D.W., 2008. Mitochondrial localization and pro-apoptotic effects of the interferon-inducible protein ISG12a. Apoptosis Int. J. Program. Cell Death 13, 562–572. https://doi.org/10.1007/s10495-008-0190-0

Rossant, J., 2008. Stem cells and early lineage development. Cell 132, 527–531. https://doi.org/10.1016/j.cell.2008.01.039

Rosser, J.M., An, W., 2010. Repeat-induced gene silencing of L1 transgenes is correlated with differential promoter methylation. Gene 456, 15–23. https://doi.org/10.1016/j.gene.2010.02.005

Roulois, D., Loo Yau, H., Singhania, R., Wang, Y., Danesh, A., Shen, S.Y., Han, H., Liang, G., Jones, P.A., Pugh, T.J., O'Brien, C., De Carvalho, D.D., 2015. DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. Cell 162, 961–973. https://doi.org/10.1016/j.cell.2015.07.056

Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P.V., Layard-Liesching, H., Verp, S., Marquis, J., Spitz, F., Constam, D.B., Trono, D., 2010. KAP1 controls endogenous retroviruses in embryonic stem cells. Nature 463, 237–240. https://doi.org/10.1038/nature08674

Sagulenko, V., Thygesen, S.J., Sester, D.P., Idris, A., Cridland, J.A., Vajjhala, P.R., Roberts, T.L., Schroder, K., Vince, J.E., Hill, J.M., Silke, J., Stacey, K.J., 2013. AIM2 and NLRP3 inflammasomes activate both apoptotic and pyroptotic death pathways via ASC. Cell Death Differ. 20, 1149–1160. https://doi.org/10.1038/cdd.2013.37

Saitoh, T., Fujita, N., Hayashi, T., Takahara, K., Satoh, T., Lee, H., Matsunaga, K., Kageyama, S., Omori, H., Noda, T., Yamamoto, N., Kawai, T., Ishii, K., Takeuchi, O., Yoshimori, T., Akira, S., 2009. Atg9a controls dsDNA-driven dynamic translocation of STING and the innate immune response. Proc. Natl. Acad. Sci. U. S. A. 106, 20842–20846. https://doi.org/10.1073/pnas.0911267106

Sakashita, A., Maezawa, S., Takahashi, K., Alavattam, K.G., Yukawa, M., Hu, Y.-C., Kojima, S., Parrish, N.F., Barski, A., Pavlicev, M., Namekawa, S.H., 2020. Endogenous retroviruses drive species-specific germline transcriptomes in mammals. Nat. Struct. Mol. Biol. 27, 967–977. https://doi.org/10.1038/s41594-020-0487-4

Salazar-Mather, T.P., Hamilton, T.A., Biron, C.A., 2000. A chemokine-to-cytokine-to-chemokine cascade critical in antiviral defense. J. Clin. Invest. 105, 985–993. https://doi.org/10.1172/JCI9232

Samarajiwa, S.A., Forster, S., Auchettl, K., Hertzog, P.J., 2009. INTERFEROME: the database of interferon regulated genes. Nucleic Acids Res. 37, D852–D857. https://doi.org/10.1093/nar/gkn732

Samson, N., Ablasser, A., 2022. The cGAS–STING pathway and cancer. Nat. Cancer 3, 1452–1463. https://doi.org/10.1038/s43018-022-00468-w

Sato, K., Hida, S., Takayanagi, H., Yokochi, T., Kayagaki, N., Takeda, K., Yagita, H., Okumura, K., Tanaka, N., Taniguchi, T., Ogasawara, K., 2001. Antiviral response by natural killer cells through TRAIL gene induction by IFN-alpha/beta. Eur. J. Immunol. 31, 3138–3146. https://doi.org/10.1002/1521-4141(200111)31:11<3138::aid-immu3138>3.0.co;2-b

Schmitz, J., Brosius, J., 2011. Exonization of transposed elements: A challenge and opportunity for evolution. Biochimie 93, 1928–1934. https://doi.org/10.1016/j.biochi.2011.07.014

Schreiber, R.D., Old, L.J., Smyth, M.J., 2011. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. Science 331, 1565–1570. https://doi.org/10.1126/science.1203486

Scopa, C., Barnada, S.M., Cicardi, M.E., Singer, M., Trotti, D., Trizzino, M., 2023. JUN upregulation drives aberrant transposable element mobilization, associated innate immune response, and impaired neurogenesis in Alzheimer's disease. Nat. Commun. 14, 1–16. https://doi.org/10.1038/s41467-023-43728-8

Scopa, C., Barnada, S.M., Cicardi, M.E., Singer, M., Trotti, D., Trizzino, M., 2022. JUN upregulation drives aberrant transposable element mobilization, associated innate immune response, and impaired neurogenesis in Alzheimer's disease. https://doi.org/10.1101/2022.11.24.517794

Seczynska, M., Bloor, S., Cuesta, S.M., Lehner, P.J., 2022. Genome surveillance by HUSH-mediated silencing of intronless mobile elements. Nature 601, 440–445. https://doi.org/10.1038/s41586-021-04228-1

Seczynska, M., Lehner, P.J., 2023. The sound of silence: mechanisms and implications of HUSH complex function. Trends Genet. 39, 251–267. https://doi.org/10.1016/j.tig.2022.12.005

Senft, A.D., Macfarlan, T.S., 2021. Transposable elements shape the evolution of mammalian development. Nat. Rev. Genet. 22, 691–711. https://doi.org/10.1038/s41576-021-00385-1

Seo, G.J., Kim, C., Shin, W.-J., Sklan, E.H., Eoh, H., Jung, J.U., 2018. TRIM56-mediated monoubiquitination of cGAS for cytosolic DNA sensing. Nat. Commun. 9, 613. https://doi.org/10.1038/s41467-018-02936-3

Seo, W., Shimizu, K., Kojo, S., Okeke, A., Kohwi-Shigematsu, T., Fujii, S., Taniuchi, I., 2020. Runx-mediated regulation of CCL5 via antagonizing two enhancers influences immune cell function and anti-tumor immunity. Nat. Commun. 11, 1562. https://doi.org/10.1038/s41467-020-15375-w

Servant, G., Streva, V.A., Derbes, R.S., Wijetunge, M.I., Neeland, M., White, T.B., Belancio, V.P., Roy-Engel, A.M., Deininger, P.L., 2017. The Nucleotide Excision Repair Pathway Limits L1 Retrotransposition. Genetics 205, 139–153. https://doi.org/10.1534/genetics.116.188680

Shang, G., Zhang, C., Chen, Z.J., Bai, X., Zhang, X., 2019. Cryo-EM structures of STING reveal its mechanism of activation by cyclic GMP–AMP. Nature 567, 389–393. https://doi.org/10.1038/s41586-019-0998-5

Sharif-Askari, E., Nakhaei, P., Oliere, S., Tumilasci, V., Hernandez, E., Wilkinson, P., Lin, R., Bell, J., Hiscott, J., 2007. Bax-dependent mitochondrial membrane permeabilization enhances IRF3-mediated innate immune response during VSV infection. Virology 365, 20–33. https://doi.org/10.1016/j.virol.2007.03.011

Shen, H., Yang, M., Li, S., Zhang, J., Peng, B., Wang, C., Chang, Z., Ong, J., Du, P., 2021. Mouse totipotent stem cells captured and maintained through spliceosomal repression. Cell 184, 2843-2859.e20. https://doi.org/10.1016/j.cell.2021.04.020

Shi, J., Zhao, Y., Wang, K., Shi, X., Wang, Y., Huang, H., Zhuang, Y., Cai, T., Wang, F., Shao, F., 2015. Cleavage of GSDMD by inflammatory caspases determines pyroptotic cell death. Nature 526, 660–665. https://doi.org/10.1038/nature15514

Shigaki, H., Baba, Y., Watanabe, M., Murata, A., Iwagami, S., Miyake, K., Ishimoto, T., Iwatsuki, M., Baba, H., 2013. LINE-1 hypomethylation in gastric cancer, detected by bisulfite pyrosequencing, is associated with poor prognosis. Gastric Cancer Off. J. Int. Gastric Cancer Assoc. Jpn. Gastric Cancer Assoc. 16, 480–487. https://doi.org/10.1007/s10120-012-0209-7

Shih, A.Y., Damm-Ganamet, K.L., Mirzadegan, T., 2018. Dynamic Structural Differences between Human and Mouse STING Lead to Differing Sensitivity to DMXAA. Biophys. J. 114, 32–39. https://doi.org/10.1016/j.bpj.2017.10.027

Shukla, R., Upton, K.R., Muñoz-Lopez, M., Gerhardt, D.J., Fisher, M.E., Nguyen, T., Brennan, P.M., Baillie, J.K., Collino, A., Ghisletti, S., Sinha, S., Iannelli, F., Radaelli, E., Dos Santos, A., Rapoud, D., Guettier, C., Samuel, D., Natoli, G., Carninci, P., Ciccarelli, F.D., Garcia-Perez, J.L., Faivre, J., Faulkner, G.J., 2013. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. Cell 153, 101–111. https://doi.org/10.1016/j.cell.2013.02.032

Sil, S., Keegan, S., Ettefa, F., Denes, L.T., Boeke, J.D., Holt, L.J., 2023. Condensation of LINE-1 is critical for retrotransposition. eLife 12, e82991. https://doi.org/10.7554/eLife.82991

Sim, Y.-J., Kim, M.-S., Nayfeh, A., Yun, Y.-J., Kim, S.-J., Park, K.-T., Kim, C.-H., Kim, K.-S., 2017. 2i Maintains a Naive Ground State in ESCs through Two Distinct Epigenetic Mechanisms. Stem Cell Rep. 8, 1312–1328. https://doi.org/10.1016/j.stemcr.2017.04.001

Simon, M., Van Meter, M., Ablaeva, J., Ke, Z., Gonzalez, R.S., Taguchi, T., De Cecco, M., Leonova, K.I., Kogan, V., Helfand, S.L., Neretti, N., Roichman, A., Cohen, H.Y., Meer, M.V., Gladyshev, V.N., Antoch, M.P., Gudkov, A.V., Sedivy, J.M., Seluanov, A., Gorbunova, V., 2019. LINE1 Derepression in Aged Wild-Type and SIRT6-Deficient Mice Drives Inflammation. Cell Metab. 29, 871-885.e5. https://doi.org/10.1016/j.cmet.2019.02.014

Simonti, C.N., Pavlicev, M., Capra, J.A., 2017. Transposable Element Exaptation into Regulatory Regions Is Rare, Influenced by Evolutionary Age, and Subject to Pleiotropic Constraints. Mol. Biol. Evol. 34, 2856–2869. https://doi.org/10.1093/molbev/msx219

Sleeth, K.M., Sørensen, C.S., Issaeva, N., Dziegielewski, J., Bartek, J., Helleday, T., 2007. RPA mediates recombination repair during replication stress and is displaced from DNA by checkpoint signalling in human cells. J. Mol. Biol. 373, 38–47. https://doi.org/10.1016/j.jmb.2007.07.068

Smith, C.L., Goldsmith, C.-A.W., Eppig, J.T., 2005. The Mammalian Phenotype Ontology as a tool for annotating, analyzing and comparing phenotypic information. Genome Biol. 6, R7. https://doi.org/10.1186/gb-2004-6-1-r7

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., Meissner, A., 2012. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature 484, 339–344. https://doi.org/10.1038/nature10960

Sokol, C.L., Luster, A.D., 2015. The Chemokine System in Innate Immunity. Cold Spring Harb. Perspect. Biol. 7, a016303. https://doi.org/10.1101/cshperspect.a016303

Solis, M., Goubau, D., Romieu-Mourez, R., Genin, P., Civas, A., Hiscott, J., 2006. Distinct functions of IRF-3 and IRF-7 in IFN-alpha gene regulation and control of anti-tumor activity in primary macrophages. Biochem. Pharmacol. 72, 1469–1476. https://doi.org/10.1016/j.bcp.2006.06.002

Sookdeo, A., Hepp, C.M., McClure, M.A., Boissinot, S., 2013. Revisiting the evolution of mouse LINE-1 in the genomic era. Mob. DNA 4, 3. https://doi.org/10.1186/1759-8753-4-3

Soper, S.F.C., van der Heijden, G.W., Hardiman, T.C., Goodheart, M., Martin, S.L., de Boer, P., Bortvin, A., 2008. Mouse Maelstrom, a Component of Nuage, Is Essential for Spermatogenesis and Transposon Repression in Meiosis. Dev. Cell 15, 285–297. https://doi.org/10.1016/j.devcel.2008.05.015

Srikanth, S., Woo, J.S., Wu, B., El-Sherbiny, Y.M., Leung, J., Chupradit, K., Rice, L., Seo, G.J., Calmettes, G., Ramakrishna, C., Cantin, E., An, D.S., Sun, R., Wu, T.-T., Jung, J.U., Savic, S., Gwack, Y., 2019. The Ca2+ sensor STIM1 regulates the type I interferon response by retaining the signaling adaptor STING at the endoplasmic reticulum. Nat. Immunol. 20, 152–162. https://doi.org/10.1038/s41590-018-0287-8

Stammnitz, M.R., Gori, K., Kwon, Y.M., Harry, E., Martin, F.J., Billis, K., Cheng, Y., Baez-Ortega, A., Chow, W., Comte, S., Eggertsson, H., Fox, S., Hamede, R., Jones, M., Lazenby, B., Peck, S., Pye, R., Quail, M.A., Swift, K., Wang, J., Wood, J., Howe, K., Stratton, M.R., Ning, Z., Murchison, E.P., 2023. The evolution of two transmissible cancers in Tasmanian devils. Science. https://doi.org/10.1126/science.abq6453

Stang, M.T., Armstrong, M., Watson, G., Liu, Y., Sung, K.-Y., Yim, J., 2006. IRF-1 mediated apoptosis is FADD and caspase 8 dependent. J. Am. Coll. Surg. 203, S82. https://doi.org/10.1016/j.jamcollsurg.2006.05.214

Stawowczyk, M., Van Scoy, S., Kumar, K.P., Reich, N.C., 2011. The interferon stimulated gene 54 promotes apoptosis. J. Biol. Chem. 286, 7257–7266. https://doi.org/10.1074/jbc.M110.207068

Stetson, D.B., Ko, J.S., Heidmann, T., Medzhitov, R., 2008. Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell 134, 587–598. https://doi.org/10.1016/j.cell.2008.06.032

Su, Y.-Q., Sun, F., Handel, M.A., Schimenti, J.C., Eppig, J.J., 2012. Meiosis arrest female 1 (MARF1) has nuage-like function in mammalian oocytes. Proc. Natl. Acad. Sci. U. S. A. 109, 18653–18660. https://doi.org/10.1073/pnas.1216904109

Sultana, T., van Essen, D., Siol, O., Bailly-Bechet, M., Philippe, C., Zine El Aabidine, A., Pioger, L., Nigumann, P., Saccani, S., Andrau, J.-C., Gilbert, N., Cristofari, G., 2019. The Landscape of L1 Retrotransposons in the Human Genome Is Shaped by Pre-insertion

Sequence Biases and Post-insertion Selection. Mol. Cell 74, 555-570.e7. https://doi.org/10.1016/j.molcel.2019.02.036

Sultana, T., Zamborlini, A., Cristofari, G., Lesage, P., 2017. Integration site selection by retroviruses and transposable elements in eukaryotes. Nat. Rev. Genet. 18, 292–308. https://doi.org/10.1038/nrg.2017.7

Sun, F., Palmer, K., Handel, M.A., 2010. Mutation of Eif4g3, encoding a eukaryotic translation initiation factor, causes male infertility and meiotic arrest of mouse spermatocytes. Dev. Camb. Engl. 137, 1699–1707. https://doi.org/10.1242/dev.043125

Sun, L., Wu, J., Du, F., Chen, X., Chen, Z.J., 2013. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. Science 339, 786–791. https://doi.org/10.1126/science.1232458

Sun, M.-S., Zhang, J., Jiang, L.-Q., Pan, Y.-X., Tan, J.-Y., Yu, F., Guo, L., Yin, L., Shen, C., Shu, H.-B., Liu, Y., 2018. TMED2 Potentiates Cellular IFN Responses to DNA Viruses by Reinforcing MITA Dimerization and Facilitating Its Trafficking. Cell Rep. 25, 3086-3098.e3. https://doi.org/10.1016/j.celrep.2018.11.048

Suzuki, J., Yamaguchi, K., Kajikawa, M., Ichiyanagi, K., Adachi, N., Koyama, H., Takeda, S., Okada, N., 2009. Genetic Evidence That the Non-Homologous End-Joining Repair Pathway Is Involved in LINE Retrotransposition. PLOS Genet. 5, e1000461. https://doi.org/10.1371/journal.pgen.1000461

Swartzendruber, D.E., Lehman, J.M., 1975. Neoplastic differentiation: Interaction of simian virus 40 and polyoma virus with murine teratocarcinoma cells in vitro. J. Cell. Physiol. 85, 179–187. https://doi.org/10.1002/jcp.1040850204

Taguchi, T., Mukai, K., Takaya, E., Shindo, R., 2021. STING Operation at the ER/Golgi Interface. Front. Immunol. 12. https://doi.org/10.3389/fimmu.2021.646304

Tam, O.H., Ostrow, L.W., Gale Hammell, M., 2019. Diseases of the nERVous system: retrotransposon activity in neurodegenerative disease. Mob. DNA 10, 32. https://doi.org/10.1186/s13100-019-0176-1

Tanaka, S.S., Toyooka, Y., Akasu, R., Katoh-Fukui, Y., Nakahara, Y., Suzuki, R., Yokoyama, M., Noce, T., 2000. The mouse homolog of Drosophila Vasa is required for the development of male germ cells. Genes Dev. 14, 841–853.

Tarkowski, A.K., 1959. Experiments on the development of isolated blastomers of mouse eggs. Nature 184, 1286–1287. https://doi.org/10.1038/1841286a0

Taylor, M.S., Altukhov, I., Molloy, K.R., Mita, P., Jiang, H., Adney, E.M., Wudzinska, A., Badri, S., Ischenko, D., Eng, G., Burns, K.H., Fenyö, D., Chait, B.T., Alexeev, D., Rout, M.P., Boeke, J.D., LaCava, J., 2018. Dissection of affinity captured LINE-1 macromolecular complexes. eLife 7, e30094. https://doi.org/10.7554/eLife.30094

Taylor, M.S., Altukhov, I., Molloy, K.R., Mita, P., Jiang, H., Adney, E.M., Wudzinska, A., Badri, S., Ischenko, D., Eng, G., Burns, K.H., Fenyö, D., Chait, B.T., Alexeev, D., Rout, M.P., Boeke, J.D., LaCava, J., n.d. Dissection of affinity captured LINE-1 macromolecular complexes. eLife 7. https://doi.org/10.7554/eLife.30094

Taylor, M.S., LaCava, J., Mita, P., Molloy, K.R., Huang, C.R.L., Li, D., Adney, E.M., Jiang, H., Burns, K.H., Chait, B.T., Rout, M.P., Boeke, J.D., Dai, L., 2013. Affinity proteomics

reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. Cell 155, 1034–1048. https://doi.org/10.1016/j.cell.2013.10.021

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., McKay, R.D.G., 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature 448, 196–199. https://doi.org/10.1038/nature05972

Tharp, M.E., Malki, S., Bortvin, A., 2020. Maximizing the ovarian reserve in mice by evading LINE-1 genotoxicity. Nat. Commun. 11, 330. https://doi.org/10.1038/s41467-019-14055-8

Thomas, C.A., Tejwani, L., Trujillo, C.A., Negraes, P.D., Herai, R.H., Mesci, P., Macia, A., Crow, Y.J., Muotri, A.R., 2017. Modeling of TREX1-Dependent Autoimmune Disease using Human Stem Cells Highlights L1 Accumulation as a Source of Neuroinflammation. Cell Stem Cell 21, 319-331.e8. https://doi.org/10.1016/j.stem.2017.07.009

Thompson, P.J., Macfarlan, T.S., Lorincz, M.C., 2016. Long Terminal Repeats: From Parasitic Elements to Building Blocks of the Transcriptional Regulatory Repertoire. Mol. Cell 62, 766–776. https://doi.org/10.1016/j.molcel.2016.03.029

Tie, C.H., Fernandes, L., Conde, L., Robbez-Masson, L., Sumner, R.P., Peacock, T., Rodriguez-Plata, M.T., Mickute, G., Gifford, R., Towers, G.J., Herrero, J., Rowe, H.M., 2018. KAP1 regulates endogenous retroviruses in adult human cells and contributes to innate immune control. EMBO Rep. 19, e45000. https://doi.org/10.15252/embr.201745000

Todd, C.D., Deniz, Ö., Taylor, D., Branco, M.R., 2019. Functional evaluation of transposable elements as enhancers in mouse embryonic and trophoblast stem cells. eLife 8, e44344. https://doi.org/10.7554/eLife.44344

Torres, J., Watt, F.M., 2008. Nanog maintains pluripotency of mouse embryonic stem cells by inhibiting NFkappaB and cooperating with Stat3. Nat. Cell Biol. 10, 194–201. https://doi.org/10.1038/ncb1680

Torres-Fernández, L.A., Jux, B., Bille, M., Port, Y., Schneider, K., Geyer, M., Mayer, G., Kolanus, W., 2019. The mRNA repressor TRIM71 cooperates with Nonsense-Mediated Decay factors to destabilize the mRNA of CDKN1A/p21. Nucleic Acids Res. 47, 11861–11879. https://doi.org/10.1093/nar/gkz1057

Tristán-Ramos, P., Rubio-Roldan, A., Peris, G., Sánchez, L., Amador-Cubero, S., Viollet, S., Cristofari, G., Heras, S.R., 2020. The tumor suppressor microRNA let-7 inhibits human LINE-1 retrotransposition. Nat. Commun. 11, 5712. https://doi.org/10.1038/s41467-020-19430-4

Trizzino, M., Park, Y., Holsbach-Beltrame, M., Aracena, K., Mika, K., Caliskan, M., Perry, G.H., Lynch, V.J., Brown, C.D., 2017. Transposable elements are the primary source of novelty in primate gene regulation. Genome Res. 27, 1623–1633. https://doi.org/10.1101/gr.218149.116

Tsuno, T., Mejido, J., Zhao, T., Phillips, T., Myers, T.G., Bekisz, J., Zoon, K.C., 2012. BID is a critical factor controlling cell viability regulated by IFN-α. J. Immunother. Hagerstown Md 1997 35, 23–31. https://doi.org/10.1097/CJI.0b013e3182372dcf

Tubio, J.M.C., Li, Y., Ju, Y.S., Martincorena, I., Cooke, S.L., Tojo, M., Gundem, G., Pipinikas, C.P., Zamora, J., Raine, K., Menzies, A., Roman-Garcia, P., Fullam, A., Gerstung, M., Shlien, A., Tarpey, P.S., Papaemmanuil, E., Knappskog, S., Van Loo, P., Ramakrishna, M., Davies, H.R., Marshall, J., Wedge, D.C., Teague, J.W., Butler, A.P., Nik-Zainal, S.,

Alexandrov, L., Behjati, S., Yates, L.R., Bolli, N., Mudie, L., Hardy, C., Martin, S., McLaren, S., O'Meara, S., Anderson, E., Maddison, M., Gamble, S., Foster, C., Warren, A.Y., Whitaker, H., Brewer, D., Eeles, R., Cooper, C., Neal, D., Lynch, A.G., Visakorpi, T., Isaacs, W.B., Veer, L.V., Caldas, C., Desmedt, C., Sotiriou, C., Aparicio, S., Foekens, J.A., Eyfjörd, J.E., Lakhani, S.R., Thomas, G., Myklebost, O., Span, P.N., Børresen-Dale, A.-L., Richardson, A.L., Van de Vijver, M., Vincent-Salomon, A., Van den Eynden, G.G., Flanagan, A.M., Futreal, P.A., Janes, S.M., Bova, G.S., Stratton, M.R., McDermott, U., Campbell, P.J., ICGC Breast Cancer Group, ICGC Bone Cancer Group, ICGC Prostate Cancer Group, 2014. Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. Science 345, 1251343. https://doi.org/10.1126/science.1251343

Tunbak, H., Enriquez-Gasca, R., Tie, C.H.C., Gould, P.A., Mlcochova, P., Gupta, R.K., Fernandes, L., Holt, J., van der Veen, A.G., Giampazolias, E., Burns, K.H., Maillard, P.V., Rowe, H.M., 2020. The HUSH complex is a gatekeeper of type I interferon through epigenetic regulation of LINE-1s. Nat. Commun. 11, 5387. https://doi.org/10.1038/s41467-020-19170-5

Turelli, P., Castro-Diaz, N., Marzetta, F., Kapopoulou, A., Raclot, C., Duc, J., Tieng, V., Quenneville, S., Trono, D., 2014. Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements. Genome Res. 24, 1260–1270. https://doi.org/10.1101/gr.172833.114

Tuteja, N., Tuteja, R., Ochem, A., Taneja, P., Huang, N.W., Simoncsits, A., Susic, S., Rahman, K., Marusic, L., Chen, J., 1994. Human DNA helicase II: a novel DNA unwinding enzyme identified as the Ku autoantigen. EMBO J. 13, 4991–5001.

van den Hurk, J.A.J.M., Meij, I.C., del Carmen Seleme, M., Kano, H., Nikopoulos, K., Hoefsloot, L.H., Sistermans, E.A., de Wijs, I.J., Mukhopadhyay, A., Plomp, A.S., de Jong, P.T.V.M., Kazazian, H.H., Cremers, F.P.M., 2007. L1 retrotransposition can occur early in human embryonic development. Hum. Mol. Genet. 16, 1587–1592. https://doi.org/10.1093/hmg/ddm108

van der Veen, A.G., Maillard, P.V., Schmidt, J.M., Lee, S.A., Deddouche-Grass, S., Borg, A., Kjær, S., Snijders, A.P., Reis e Sousa, C., 2018. The RIG-I-like receptor LGP2 inhibits Dicer-dependent processing of long double-stranded RNA and blocks RNA interference in mammalian cells. EMBO J. 37, e97479. https://doi.org/10.15252/embj.201797479

Vanyushin, B.F., Tkacheva, S.G., Belozersky, A.N., 1970. Rare Bases in Animal DNA. Nature 225, 948–949. https://doi.org/10.1038/225948a0

Vivier, E., Nunès, J.A., Vély, F., 2004. Natural Killer Cell Signaling Pathways. Science 306, 1517–1519. https://doi.org/10.1126/science.1103478

Vivier, E., Raulet, D.H., Moretta, A., Caligiuri, M.A., Zitvogel, L., Lanier, L.L., Yokoyama, W.M., Ugolini, S., 2011. Innate or adaptive immunity? The example of natural killer cells. Science 331, 44–49. https://doi.org/10.1126/science.1198687

Vivier, E., Tomasello, E., Baratin, M., Walzer, T., Ugolini, S., 2008. Functions of natural killer cells. Nat. Immunol. 9, 503–510. https://doi.org/10.1038/ni1582

Volkman, H.E., Stetson, D.B., 2014. The enemy within: endogenous retroelements and autoimmune disease. Nat. Immunol. 15, 415–422. https://doi.org/10.1038/ni.2872

Volkmann, B., Wittmann, S., Lagisquet, J., Deutschmann, J., Eissmann, K., Ross, J.J., Biesinger, B., Gramberg, T., 2020. Human TRIM5α senses and restricts LINE-1 elements. Proc. Natl. Acad. Sci. 117, 17965–17976. https://doi.org/10.1073/pnas.1922366117

Vourekas, A., Zheng, K., Fu, Q., Maragkakis, M., Alexiou, P., Ma, J., Pillai, R.S., Mourelatos, Z., Wang, P.J., 2015. The RNA helicase MOV10L1 binds piRNA precursors to initiate piRNA processing. Genes Dev. 29, 617–629. https://doi.org/10.1101/gad.254631.114

Vuong, L.M., Pan, S., Donovan, P.J., 2019. Proteome Profile of Endogenous Retrotransposon-Associated Complexes in Human Embryonic Stem Cells. Proteomics 19, e1900169. https://doi.org/10.1002/pmic.201900169

Wajapeyee, N., Serra, R.W., Zhu, X., Mahalingam, M., Green, M.R., 2008. Oncogenic BRAF Induces Senescence and Apoptosis through Pathways Mediated by the Secreted Protein IGFBP7. Cell 132, 363–374. https://doi.org/10.1016/j.cell.2007.12.032

Wallace, N.A., Belancio, V.P., Deininger, P.L., 2008. L1 mobile element expression causes multiple types of toxicity. Gene 419, 75–81. https://doi.org/10.1016/j.gene.2008.04.013

Walter, M., Teissandier, A., Pérez-Palacios, R., Bourc'his, D., 2016. An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. eLife 5, e11418. https://doi.org/10.7554/eLife.11418

Wang, C., Guan, Y., Lv, M., Zhang, R., Guo, Z., Wei, X., Du, X., Yang, J., Li, T., Wan, Y., Su, X., Huang, X., Jiang, Z., 2018. Manganese Increases the Sensitivity of the cGAS-STING Pathway for Double-Stranded DNA and Is Required for the Host Defense against DNA Viruses. Immunity 48, 675-687.e7. https://doi.org/10.1016/j.immuni.2018.03.017

Wang, J., Xie, G., Singh, M., Ghanbarian, A.T., Raskó, T., Szvetnik, A., Cai, H., Besser, D., Prigione, A., Fuchs, N.V., Schumann, G.G., Chen, W., Lorincz, M.C., Ivics, Z., Hurst, L.D., Izsvák, Z., 2014. Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. Nature 516, 405–409. https://doi.org/10.1038/nature13804

Wang, J.-T., Chang, L.-S., Chen, C.-J., Doong, S.-L., Chang, C.-W., Chen, M.-R., 2014. Glycogen synthase kinase 3 negatively regulates IFN regulatory factor 3 transactivation through phosphorylation at its linker region. Innate Immun. 20, 78–87. https://doi.org/10.1177/1753425913485307

Wang, L., Wen, M., Cao, X., 2019. Nuclear hnRNPA2B1 initiates and amplifies the innate immune response to DNA viruses. Science 365, eaav0758. https://doi.org/10.1126/science.aav0758

Wang, P.J., 2017. Tracking LINE1 retrotransposition in the germline. Proc. Natl. Acad. Sci. 114, 7194–7196. https://doi.org/10.1073/pnas.1709067114

Wang, R., Teng, C., Spangler, J., Wang, J., Huang, F., Guo, Y.-L., 2014a. Mouse Embryonic Stem Cells Have Underdeveloped Antiviral Mechanisms That Can Be Exploited for the Development of mRNA-Mediated Gene Expression Strategy. Stem Cells Dev. 23, 594–604. https://doi.org/10.1089/scd.2013.0417

Wang, R., Wang, J., Acharya, D., Paul, A.M., Bai, F., Huang, F., Guo, Y.-L., 2014b. Antiviral Responses in Mouse Embryonic Stem Cells: DIFFERENTIAL DEVELOPMENT OF CELLULAR MECHANISMS IN TYPE I INTERFERON PRODUCTION AND RESPONSE. J. Biol. Chem. 289, 25186–25198. https://doi.org/10.1074/jbc.M113.537746

Wang, Y., Ning, X., Gao, P., Wu, S., Sha, M., Lv, M., Zhou, X., Gao, J., Fang, R., Meng, G., Su, X., Jiang, Z., 2017. Inflammasome Activation Triggers Caspase-1-Mediated Cleavage of cGAS to Regulate Responses to DNA Virus Infection. Immunity 46, 393–404. https://doi.org/10.1016/j.immuni.2017.02.011

Wei, J., Yu, X., Yang, L., Liu, X., Gao, B., Huang, B., Dou, X., Liu, J., Zou, Z., Cui, X.-L., Zhang, L.-S., Zhao, X., Liu, Q., He, P.C., Sepich-Poore, C., Zhong, N., Liu, W., Li, Y., Kou, X., Zhao, Y., Wu, Y., Cheng, X., Chen, C., An, Y., Dong, X., Wang, H., Shu, Q., Hao, Z., Duan, T., He, Y.-Y., Li, X., Gao, S., Gao, Y., He, C., 2022. FTO mediates LINE1 m6A demethylation and chromatin regulation in mESCs and mouse development. Science 376, 968–973. https://doi.org/10.1126/science.abe9582

Wei, W., Gilbert, N., Ooi, S.L., Lawler, J.F., Ostertag, E.M., Kazazian, H.H., Boeke, J.D., Moran, J.V., 2001. Human L1 retrotransposition: cis preference versus trans complementation. Mol. Cell. Biol. 21, 1429–1439. https://doi.org/10.1128/MCB.21.4.1429-1439.2001

Weichold, F.F., Zella, D., Barabitskaja, O., Maciejewski, J.P., Dunn, D.E., Sloand, E.M., Young, N.S., 1998. Neither Human Immunodeficiency Virus-1 (HIV-1) nor HIV-2 Infects Most-Primitive Human Hematopoietic Stem Cells as Assessed in Long-Term Bone Marrow Cultures. Blood 91, 907–915. https://doi.org/10.1182/blood.V91.3.907

Wells, J.N., Feschotte, C., 2020. A Field Guide to Eukaryotic Transposable Elements. Annu. Rev. Genet. 54, 539–561. https://doi.org/10.1146/annurev-genet-040620-022145

Wiley, C.D., Flynn, J.M., Morrissey, C., Lebofsky, R., Shuga, J., Dong, X., Unger, M.A., Vijg, J., Melov, S., Campisi, J., 2017. Analysis of individual cells identifies cell-to-cell variability following induction of cellular senescence. Aging Cell 16, 1043–1050. https://doi.org/10.1111/acel.12632

Williams, B.R., 2001. Signal integration via PKR. Sci. STKE Signal Transduct. Knowl. Environ. 2001. https://doi.org/10.1126/stke.2001.89.re2

Wimmer, K., Callens, T., Wernstedt, A., Messiaen, L., 2011. The NF1 gene contains hotspots for L1 endonuclease-dependent de novo insertion. PLoS Genet. 7, e1002371. https://doi.org/10.1371/journal.pgen.1002371

Wissing, S., Montano, M., Garcia-Perez, J.L., Moran, J.V., Greene, W.C., 2011. Endogenous APOBEC3B Restricts LINE-1 Retrotransposition in Transformed Cells and Human Embryonic Stem Cells*. J. Biol. Chem. 286, 36427–36437. https://doi.org/10.1074/jbc.M111.251058

Witteveldt, J., Knol, L.I., Macias, S., 2019. MicroRNA-deficient mouse embryonic stem cells acquire a functional interferon response. eLife 8, e44171. https://doi.org/10.7554/eLife.44171

Wu, J., Dobbs, N., Yang, K., Yan, N., 2020. Interferon-Independent Activities of Mammalian STING Mediate Antiviral Response and Tumor Immune Evasion. Immunity 53, 115-126.e5. https://doi.org/10.1016/j.immuni.2020.06.009

Wu, S.-Y., Xiao, Y., Wei, J.-L., Xu, X.-E., Jin, X., Hu, X., Li, D.-Q., Jiang, Y.-Z., Shao, Z.-M., 2021. MYC suppresses STING-dependent innate immunity by transcriptionally upregulating DNMT1 in triple-negative breast cancer. J. Immunother. Cancer 9, e002528. https://doi.org/10.1136/jitc-2021-002528

Wu, X., Dao Thi, V.L., Huang, Y., Billerbeck, E., Saha, D., Hoffmann, H.-H., Wang, Y., Silva, L.A.V., Sarbanes, S., Sun, T., Andrus, L., Yu, Y., Quirk, C., Li, M., MacDonald, M.R., Schneider, W.M., An, X., Rosenberg, B.R., Rice, C.M., 2018. Intrinsic Immunity Shapes Viral Resistance of Stem Cells. Cell 172, 423-438.e25. https://doi.org/10.1016/j.cell.2017.11.018

Wu, Y., Liu, W., Chen, J., Liu, S., Wang, M., Yang, L., Chen, C., Qi, M., Xu, Y., Qiao, Z., Yan, R., Kou, X., Zhao, Y., Shen, B., Yin, J., Wang, H., Gao, Y., Gao, S., 2019. Nuclear Exosome Targeting Complex Core Factor Zcchc8 Regulates the Degradation of LINE1 RNA in Early Embryos and Embryonic Stem Cells. Cell Rep. 29, 2461-2472.e6. https://doi.org/10.1016/j.celrep.2019.10.055

Xia, P., Ye, B., Wang, S., Zhu, X., Du, Y., Xiong, Z., Tian, Y., Fan, Z., 2016. Glutamylation of the DNA sensor cGAS regulates its binding and synthase activity in antiviral immunity. Nat. Immunol. 17, 369–378. https://doi.org/10.1038/ni.3356

Xiao-Jie, L., Hui-Ying, X., Qi, X., Jiang, X., Shi-Jie, M., 2016. LINE-1 in cancer: multifaceted functions and potential clinical implications. Genet. Med. 18, 431–439. https://doi.org/10.1038/gim.2015.119

Xie, M., Hong, C., Zhang, B., Lowdon, R.F., Xing, X., Li, D., Zhou, X., Lee, H.J., Maire, C.L., Ligon, K.L., Gascard, P., Sigaroudinia, M., Tlsty, T.D., Kadlecek, T., Weiss, A., O'Geen, H., Farnham, P.J., Madden, P.A.F., Mungall, A.J., Tam, A., Kamoh, B., Cho, S., Moore, R., Hirst, M., Marra, M.A., Costello, J.F., Wang, T., 2013. DNA hypomethylation within specific transposable element families associates with tissue-specific enhancer landscape. Nat. Genet. 45, 836–841. https://doi.org/10.1038/ng.2649

Xie, S.Q., Leeke, B.J., Whilding, C., Wagner, R.T., Garcia-Llagostera, F., Low, Y., Chammas, P., Cheung, N.T.-F., Dormann, D., McManus, M.T., Percharde, M., 2022. Nucleolar-based Dux repression is essential for embryonic two-cell stage exit. Genes Dev. https://doi.org/10.1101/gad.349172.121

Xie, Y., Mates, L., Ivics, Z., Izsvák, Z., Martin, S.L., An, W., 2013. Cell division promotes efficient retrotransposition in a stable L1 reporter cell line. Mob. DNA 4, 10. https://doi.org/10.1186/1759-8753-4-10

Xie, Y., Rosser, J.M., Thompson, T.L., Boeke, J.D., An, W., 2011. Characterization of L1 retrotransposition with high-throughput dual-luciferase assays. Nucleic Acids Res. 39, e16. https://doi.org/10.1093/nar/gkq1076

Xing, J., Wang, H., Belancio, V.P., Cordaux, R., Deininger, P.L., Batzer, M.A., 2006. Emergence of primate genes by retrotransposon-mediated sequence transduction. Proc. Natl. Acad. Sci. U. S. A. 103, 17608–17613. https://doi.org/10.1073/pnas.0603224103

Xu, B., Li, X., Zhang, S., Lian, M., Huang, W., Zhang, Y., Wang, Y., Huang, Z., 2023. Pan cancer characterization of genes whose expression has been associated with LINE-1 antisense promoter activity. Mob. DNA 14, 13. https://doi.org/10.1186/s13100-023-00300-x

Yamakoshi, K., Takahashi, A., Hirota, F., Nakayama, R., Ishimaru, N., Kubo, Y., Mann, D.J., Ohmura, M., Hirao, A., Saya, H., Arase, S., Hayashi, Y., Nakao, K., Matsumoto, M., Ohtani, N., Hara, E., 2009. Real-time in vivo imaging of p16Ink4a reveals cross talk with p53. J. Cell Biol. 186, 393–407. https://doi.org/10.1083/jcb.200904105

Yang, C., Atkinson, S.P., Vilella, F., Lloret, M., Armstrong, L., Mann, D.A., Lako, M., 2010. Opposing Putative Roles for Canonical and Noncanonical NFκB Signaling on the Survival,

Proliferation, and Differentiation Potential of Human Embryonic Stem Cells. STEM CELLS 28, 1970–1980. https://doi.org/10.1002/stem.528

Yang, P., Wang, Y., Macfarlan, T.S., 2017. The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. Trends Genet. TIG 33, 871–881. https://doi.org/10.1016/j.tig.2017.08.006

Yang, W.R., Ardeljan, D., Pacyna, C.N., Payer, L.M., Burns, K.H., 2019. SQuIRE reveals locus-specific regulation of interspersed repeat expression. Nucleic Acids Res. 47, e27. https://doi.org/10.1093/nar/gky1301

Ye, M., Goudot, C., Hoyler, T., Lemoine, B., Amigorena, S., Zueva, E., 2020. Specific subfamilies of transposable elements contribute to different domains of T lymphocyte enhancers. Proc. Natl. Acad. Sci. U. S. A. 117, 7905–7916. https://doi.org/10.1073/pnas.1912008117

Yesbolatova, A., Saito, Y., Kitamoto, N., Makino-Itou, H., Ajima, R., Nakano, R., Nakaoka, H., Fukui, K., Gamo, K., Tominari, Y., Takeuchi, H., Saga, Y., Hayashi, K., Kanemaki, M.T., 2020. The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. Nat. Commun. 11, 5701. https://doi.org/10.1038/s41467-020-19532-z

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., Smith, A., 2008. The ground state of embryonic stem cell self-renewal. Nature 453, 519–523. https://doi.org/10.1038/nature06968

Yockey, L.J., Iwasaki, A., 2018. Interferons and Proinflammatory Cytokines in Pregnancy and Fetal Development. Immunity 49, 397–412. https://doi.org/10.1016/j.immuni.2018.07.017

Yoder, J.A., Walsh, C.P., Bestor, T.H., 1997. Cytosine methylation and the ecology of intragenomic parasites. Trends Genet., Epigenetics 13, 335–340. https://doi.org/10.1016/S0168-9525(97)01181-5

Yu, Q., Chu, L., Li, Y., Wang, Q., Zhu, J., Wang, C., Cui, S., 2021. miR-23a/b suppress cGAS-mediated innate and autoimmunity. Cell. Mol. Immunol. 18, 1235–1248. https://doi.org/10.1038/s41423-021-00668-x

Zhang, A., Dong, B., Doucet, A.J., Moldovan, J.B., Moran, J.V., Silverman, R.H., 2014. RNase L restricts the mobility of engineered retrotransposons in cultured human cells. Nucleic Acids Res. 42, 3803–3820. https://doi.org/10.1093/nar/gkt1308

Zhang, C., Shang, G., Gui, X., Zhang, X., Bai, X.-C., Chen, Z.J., 2019. Structural basis of STING binding with and phosphorylation by TBK1. Nature 567, 394–398. https://doi.org/10.1038/s41586-019-1000-2

Zhang, L., Hughes, B., Tombline, G., Angelini, L., Yousefzadeh, M., Gorbunova, V., Niedernhofer, L., Robbins, P., 2022. FUCOIDANS ARE NOVEL SENOTHERAPEUTICS THAT ENHANCE SIRT6 AND DNA REPAIR ACTIVITY. Innov. Aging 6, 130. https://doi.org/10.1093/geroni/igac059.522

Zhang, X., Yan, C., Hang, J., Finci, L.I., Lei, J., Shi, Y., 2017. An Atomic Structure of the Human Spliceosome. Cell 169, 918-929.e14. https://doi.org/10.1016/j.cell.2017.04.033

Zhang, X., Wu, J., Du, F., Xu, H., Sun, L., Chen, Z., Brautigam, C.A., Zhang, Xuewu, Chen, Z.J., 2014. The cytosolic DNA sensor cGAS forms an oligomeric complex with DNA and undergoes switch-like conformational changes in the activation loop. Cell Rep. 6, 421–430. https://doi.org/10.1016/j.celrep.2014.01.003

Zhang, Y., Li, M., Li, L., Qian, G., Wang, Y., Chen, Z., Liu, J., Fang, C., Huang, F., Guo, D., Zou, Q., Chu, Y., Yan, D., 2020. β-arrestin 2 as an activator of cGAS-STING signaling and target of viral immune evasion. Nat. Commun. 11, 6000. https://doi.org/10.1038/s41467-020-19849-9

Zhang, Z., Zhang, N., Guo, S., Liu, Q., Wang, S., Zhang, A., Yi, D., Zhao, J., Li, Q., Wang, J., Zhang, Y., Ma, L., Ding, J., Cen, S., Li, X., 2022. The Zinc-Finger protein ZCCHC3 inhibits LINE-1 retrotransposition. Front. Microbiol. 13, 891852. https://doi.org/10.3389/fmicb.2022.891852

Zhao, Baoyu, Du, F., Xu, P., Shu, C., Sankaran, B., Bell, S.L., Liu, M., Lei, Y., Gao, X., Fu, X., Zhu, F., Liu, Y., Laganowsky, A., Zheng, X., Ji, J.-Y., West, A.P., Watson, R.O., Li, P., 2019. A conserved PLPLRT/SD motif of STING mediates the recruitment and activation of TBK1. Nature 569, 718–722. https://doi.org/10.1038/s41586-019-1228-x

Zhao, Boxun, Wu, Q., Ye, A.Y., Guo, J., Zheng, X., Yang, X., Yan, L., Liu, Q.-R., Hyde, T.M., Wei, L., Huang, A.Y., 2019. Somatic LINE-1 retrotransposition in cortical neurons and non-brain tissues of Rett patients and healthy individuals. PLOS Genet. 15, e1008043. https://doi.org/10.1371/journal.pgen.1008043

Zhao, C., Ma, Y., Zhang, M., Gao, X., Liang, W., Qin, Y., Fu, Y., Jia, M., Song, H., Gao, C., Zhao, W., 2023. Polyamine metabolism controls B-to-Z DNA transition to orchestrate DNA sensor cGAS activity. Immunity. https://doi.org/10.1016/j.immuni.2023.09.012

Zhao, K., Du, J., Han, X., Goodier, J.L., Li, P., Zhou, X., Wei, W., Evans, S.L., Li, L., Zhang, W., Cheung, L.E., Wang, G., Kazazian, H.H., Yu, X.-F., 2013. Modulation of LINE-1 and Alu/SVA Retrotransposition by Aicardi-Goutières Syndrome-Related SAMHD1. Cell Rep. 4, 1108–1115. https://doi.org/10.1016/j.celrep.2013.08.019

Zhao, Y., Oreskovic, E., Zhang, Q., Lu, Q., Gilman, A., Lin, Y.S., He, J., Zheng, Z., Lu, J.Y., Lee, J., Ke, Z., Ablaeva, J., Sweet, M.J., Horvath, S., Zhang, Z., Nevo, E., Seluanov, A., Gorbunova, V., 2021. Transposon-triggered innate immune response confers cancer resistance to the blind mole rat. Nat. Immunol. 22, 1219–1230. https://doi.org/10.1038/s41590-021-01027-8

Zhen, Z., Chen, Y., Wang, H., Tang, H., Zhang, H., Liu, H., Jiang, Y., Mao, Z., 2023. Nuclear cGAS restricts L1 retrotransposition by promoting TRIM41-mediated ORF2p ubiquitination and degradation. Nat. Commun. 14, 8217. https://doi.org/10.1038/s41467-023-43001-y

Zhou, C., Chen, X., Planells-Cases, R., Chu, J., Wang, L., Cao, L., Li, Z., López-Cayuqueo, K.I., Xie, Y., Ye, S., Wang, X., Ullrich, F., Ma, S., Fang, Y., Zhang, X., Qian, Z., Liang, X., Cai, S.-Q., Jiang, Z., Zhou, D., Leng, Q., Xiao, T.S., Lan, K., Yang, J., Li, H., Peng, C., Qiu, Z., Jentsch, T.J., Xiao, H., 2020. Transfer of cGAMP into Bystander Cells via LRRC8 Volume-Regulated Anion Channels Augments STING-Mediated Interferon Responses and Anti-viral Immunity. Immunity 52, 767-781.e6. https://doi.org/10.1016/j.immuni.2020.03.016

Zhou, W., Whiteley, A.T., de Oliveira Mann, C.C., Morehouse, B.R., Nowak, R.P., Fischer, E.S., Gray, N.S., Mekalanos, J.J., Kranzusch, P.J., 2018. Structure of the Human cGAS-DNA Complex Reveals Enhanced Control of Immune Surveillance. Cell 174, 300-311.e11. https://doi.org/10.1016/j.cell.2018.06.026

Zhu, C., Utsunomiya, T., Ikemoto, T., Yamada, S., Morine, Y., Imura, S., Arakawa, Y., Takasu, C., Ishikawa, D., Imoto, I., Shimada, M., 2014. Hypomethylation of long interspersed nuclear element-1 (LINE-1) is associated with poor prognosis via activation of c-MET in hepatocellular carcinoma. Ann. Surg. Oncol. 21 Suppl 4, S729-735. https://doi.org/10.1245/s10434-014-3874-4

Zhu, Y., An, X., Zhang, X., Qiao, Y., Zheng, T., Li, X., 2019. STING: a master regulator in the cancer-immunity cycle. Mol. Cancer 18, 152. https://doi.org/10.1186/s12943-019-1087-y

Zierhut, C., Yamaguchi, N., Paredes, M., Luo, J.-D., Carroll, T., Funabiki, H., 2019. The Cytoplasmic DNA Sensor cGAS Promotes Mitotic Cell Death. Cell 178, 302-315.e23. https://doi.org/10.1016/j.cell.2019.05.035

Chapter 9: Annex I: Primers and siRNAs

Primers

AIM2_F	GATTCAAAGTGCAGGTGCGG	L1_A-F	GGATTCCACACGTGATCCTAA
AIM2_R	TCTGAGGGTTAGCTTGAGGAC	L1_A-R	TCCTCTATGAGCAGACCTGGA
AMPEF1A_F	AATTGGATCCAAAGGAGTGGGAATTGGCTCC	L1_GF-F	CTCCTTGGCTCCGGGACT
AMPEF1A_R	AATTGCTAGCTACGTCACGACACCTGAAATG	L1_GF-R	CAGGAAGGTGGCCGGTTGT
APOBEC3B_2_-F	GCTACATCTCGGTCCCTTCC	L1_ORF2_ETP_F	AGTGCAGAGTTCTATCAGACCTTC
APOBEC3B_2_-R	TCACTTAGCGGGTCCATTCG	L1_ORF2_ETP_R	AACCTACTTGGTCAGGATGGATG
CCL5-F	CTCACCATATGGCTCGGACA	L1_TF-F	CAGCGGTCGCCATCTTG
CCL5-R	ACAAACACGACTGCAAGATTGG	L1_TF-R	CACCCTCTCACCTGTTCAGACTAA
CDKN1A_1-R	TCCTCTTGGAGAAGATCAGCCG	LMNB1-F	CAGATCGAGCTGGGCAAGTT
CDKN1A_2-F	AGGTGGACCTGGAGACTCTCAG	LMNB1-R	CCGCCTCATACTCTCGAAGC
CDKN2A_2-F	TGTTGAGGCTAGAGAGGATCTTG	LRRFIP1_F	GCAGTTCCAGTTTGCCG
CDKN2A_2-R	CGAATCTGCACCGTAGTTGAGC	LRRFIP1_R	GTGTCTGAAGTCTCTCCG
CGAS_F	GCCTGCGCATTCAAAACT	MERVL-LTR-F	CTTCCATTCACAGCTGCGACTG
CGAS_F2	5 -GCCGAGACGGTGAATAAAGT-3	MERVL-LTR-R	CTAGAACCACTCCTGGTACCAAC
CGAS_R	TGTGAGAGAAGGATAGCCGC	MMERVK10C-F	CAAATAGCCCTACCATATGTCAG
CGAS_R2	- CATTAGGAGCAGAAATCTTCACA	MMERVK10C-R	GTATACTTTCTTCTTCAGGTCCAC
CHUK-F	AACATCCTCTGACATGTGTGGT	MOV10_2_-F	TGCGGACGATTTACAACCAC

CHUK-R	TGGTCCTCATTTGCTTCACGA
CXCL10 FORWARD	GGTCTGAGTGGGACTCAAGG
CXCL10 REVERSE	GTGGCAATGATCTCAACACG
EGFP2(SML1)_F	GGTCACGAACTCCAGCAG
EGFP2(SML1)_R	CAGAAGAACGGCATCAAGG
ENDO-ROSA_R	ATCTGTGGGAAGTCTTGTC
FGL_169_ROSA_CAG_F	CTGCAACTCCAGTCTTTCTAG
GHPRT_F	CTTCCAGCACCTCGTAATTTGAC
GHPRT_R	AGCCTGGCCAAGTTCCAGTT
HGAPDH_FW	AGCACATCGCTCAGACAC
HGAPDH_REV	GCCCAATACGACCAAATCC
HNRNP_A2B1_F	CCATAACCAGGGCTACCTCCAA
HNRNP_A2B1_R	CGGTGGCAATTTTGGACCAGGA
IAPEZ-F	CTCCATGTGCTCTGCCTTCC
IAPEZ-R	CCCCGTCCCTTTTTTAGGAGA
IBKKG-F	TCCAAGACTACGACAGCCACA
IFIT1_F	5'-CCAAGTGTTCCAATGCTCCT-3'
IFIT1_R	5'-GGATGGAATTGCCTGCTAGA-3'
IFITM1_F	ATCTCCCTGGGACTCTAGCA
IFITM3_F	CCGTGAAGTCTAGGGATCGG
IFITM3_R	ACAATGGTGATAACAACCATCAGG
IFNA_F	CTGCTGGCTGTGAGGACATA
IFNA_R	CTGCTGGCTGTGAGGACATA
MOV10_2_-R	CTGGTAACTGCACATCAGCC
MSB2-ETP-F	GAGCACCTGACTGCTCTTCC
MSB2-ETP-R	ACACACCAGAAGAGGGCATC
NFKB1_F	TCTGGCACAGAAGTTGGGTC
NFKB1_R	CCCGGAGTTCATCTCATAGTTGT
NFKBIA-F	CACCAACTACAATGGCCACAC
NFKBIB_F	ACAACAGAGATGAGGGCGATG
NFKBIB_R	CTCTGGAGGGTCGGGACTTA
NFKBIB-R	GAGCGAAACCAGGTCAGGAT
OAS1A_F	GCCTGATCCCAGAATCTATGC
OAS1A_F	TGTCCTGGGTCATGTTAATAC
OAS1A_R	GAGCAACTCTAGGGCGTACTG
OAS1A_R	CCGTGAAGCAGGTAGAGA
P204-F	GGTCCCAAACAAGTGATGGTGC-
P204-R	TCAGTTTCAGTAGCCACGGTAGCA
PKR_F	GGAAAATCCCGAACAAGGAG
PKR_R	CCCAAAGCAAAGATGTCCAC
PQBP1_F	AGCGGGCATTTGACAAGGCAGA
PQBP1_R	GGTGCATCTGAGTATGAGCTGG
RELA_F	AGGCTTCTGGGCCTTATGTG
RELA_R	TGCTTCTCTCGCCAGGAATAC
RNASE_L_F	TGGACTTGGGAGAACCGCTA
RNASE_L_R	GGCTCAAGTGTCTGATGCTGC

IFNB_F	CCCTATGGAGATGACGGAGA
IFNB_R	ACCCAGTGCTGGAGAAATTG
IFNB2-F	GCTCCAAGAAAGGACGAACA
IFNB2-R	CATCTCTTGGATGGCAAAGG
IFNB3-F	AAGAGTTACACTGCCTTTGCCATC
IFNB3-R	CACTGTCTGCTGGTGGAGTTCATC
IFNG_F	GAAAGACAATCAGGCCATCA
IFNG_R	TTGCTGTTGCTGAAGAAGGT
IFTIM1_R	TTTGAAGAGGTGTGGTCGAGG
IKBKB-F	CACCCAATGATTTGCCACTGC
IKBKB-R	TTTCTGGCTTCAGGTCTCGATG
IKBKE-F	TTCTGAGCCACTTGGAGTGC
IKBKE-R	TTTACAGCAACCACCTCCCC
IKBKG-R	AGATATCCGCCTGGGCCTTC
IL15F	GTAGGTCTCCCTAAAACAGAGGC
IL15R	TCCAGGAGAAAGCAGTTCATTGC
IL6_F	CAAGAAAGACAAAGCCAGAGTC
IL6_R	GAAATTGGGGTAGGAAGGAC
IRF3_F	CAATTCCTCCCCTGGCTAGA
IRF3_R	GGGATCCTGAACCTCGTTCG
IRF35D F	
IRF35D R	
IRF7_F	ACCGTGTTTACGAGGAACCC
IRF7_R	ATACCCATGGCTCCAGCTTC
ISG15 FORWARD	AGGTCTTTCTGACGCAGACTG
RPL7-F	AGCGGATTGCCTTGACAGAT
RPL7-R	AACTTGAAGGGCCACAGGAA
RW067-ROSA5-R	GAACTCCATATATGGGCTATGAAC
RW174-ROSA5'SHORT-F1	TCTGGGGGAGTCGTTTTACC
SEQ_BFP	TAGCAAACTGGGGCACAAG
SEQ-SFVV	ATTGATTGACTGCCCACCTC
SEQEF1AINSERTED	GTCCCCATAATTTTTGGCAGAGGG
STING_F	TAGAGAGCTTTGGGGCCTCT
STING_R	CTGAATGGGGAGACAGCAGC
STINGMUT_F	CCTGCCTGGGATGCCCCATC
STINMUT_R1	AGGCCAGCCCGTGGGCCAT
STINMUT_R2	AGGCCAGCCCGTGGGCCTT
TBK1_F	TGTACGCGGCGGACACTCGCG
TBK1_R	TCCTCTTCAATAGCAAATAAC-
TLR3_F	GCCTAGTTGTTTCTGGGAGT
TLR3_R	AGTGAAACTCGCTTTATCTTCTTTG
TLR9_F	CAAGAACCTGGTGTCACTGC
TLR9_R	TGCGATTGTCTGACAAGTCC
TNF-F	TCCCCAAAGGGATGAGAAGTT
TNF-R	GTTTGCTACGACGTGGGCTAC
TRAF6-F	CCTGACGGTAAAGTGCCCAA
TRAF6-R	ACCTGGCACTTCTGGAAAGG
TREX1-F	TTGAGATGGCCCCCATGGTTC
TREX1-R	CACAGCTCTGTGACTTCGGG
ZBP1-F	GCTATGACGGACAGACGTGG

ISG15 REVERSE	GGGGCTTTAGGCCATACTC C	ZBP1-R	GTTGGCTCCTTGTTG GCAGA

siRNAs

All siRNAs were reconstituted in Dharmacon siRNA buffer to a final concentration of 20μM and aliquots stored at -20°C.

Gene Symbol	GENE ID	Gene Accession	GI Number	Sequence
Acer2	230379	NM_139306	21314857	UCAUUUGAUUGGCGAUUUA
Acer2	230379	NM_139306	21314857	CCACAAGCGAAGCCCGGUA
Acer2	230379	NM_139306	21314857	GGGUGAAAGGAGCCGAGUA
Acer2	230379	NM_139306	21314857	GCUCUGUUCUGUUCCGAAU
Adar	56417	NM_019655	31981008	GCUAUUUGCUGCCGCGUGA
Adar	56417	NM_019655	31981008	CAGUGAAAGCCAUGGCGAU
Adar	56417	NM_019655	31981008	CCAAAAGACAGUCCGGAAA
Adar	56417	NM_019655	31981008	CUGAAGGGAGAGACGGUCA
Ago2	239528	NM_153178	33859782	GUUGUAUUGUUUAGCGAUU
Ago2	239528	NM_153178	33859782	UGAGAAACGCCCUCGGAGA
Ago2	239528	NM_153178	33859782	CGUGUUUACAAGUCGGACA
Ago2	239528	NM_153178	33859782	GUGUGUGGGACAUGCGAAA

Atxn2l	233871	NM_183020	33942088	GGGUUCUACUUGUGACGUA
Atxn2l	233871	NM_183020	33942088	GUUUAAACCAAGCGACGUU
Atxn2l	233871	NM_183020	33942088	GCUUCAAAGUAAUCCACGA
Atxn2l	233871	NM_183020	33942088	GGGGACAGACCACCGGAAA
Cebpz	12607	NM_001024806	67188875	CCAGAUAGGUGACGACAAA
Cebpz	12607	NM_001024806	67188875	GAGUAGUGACGGCGAACUU
Cebpz	12607	NM_001024806	67188875	GAACAGAAUUGCUACGAAA
Cebpz	12607	NM_001024806	67188875	GCAACAAGGACUCGAGAGA
Cpsf4	54188	NM_178576	30466246	GGACAUCAGUGCCUCGUUU
Cpsf4	54188	NM_178576	30466246	CCAAGUUCGGUUAUAAGAA
Cpsf4	54188	NM_178576	30466246	CUACAAGUGUGGUGAGAAA
Cpsf4	54188	NM_178576	30466246	CUUUCUAUGAGGUUGAUUA
Dazl	13164	NM_010021	31542548	GAUGUUAGGAUGGAUGAAA
Dazl	13164	NM_010021	31542548	GAGCAGAGGAGUUAUGUUA
Dazl	13164	NM_010021	31542548	AGAAUGUGCUAAAGUGAAU
Dazl	13164	NM_010021	31542548	CUAAAUUUCACUAGGUGUA

Ddx6	13209	NM_181324	31342194	CAUAUGUAACGGAGCGCCA
Ddx6	13209	NM_181324	31342194	CAUCAAGAAAGGCGUGGCA
Ddx6	13209	NM_181324	31342194	ACUAACUGCUGGUCUUUAA
Ddx6	13209	NM_181324	31342194	GGAUAUUAUUCUCACACUA
Dicer1	192119	NM_148948	22507358	GGUAGACUGUGGACCGUUU
Dicer1	192119	NM_148948	22507358	GGAAAUACCUGUACAACCA
Dicer1	192119	NM_148948	22507358	GCAAUUUGGUGGUUCGUUU
Dicer1	192119	NM_148948	22507358	ACAGGAAUCAGGAUAAUUA
Dnd1	213236	NM_173383	27734199	CUUAAACCGCCGAGCGUUA
Dnd1	213236	NM_173383	27734199	CGUCAGGGUCCGAGGUGUA
Dnd1	213236	NM_173383	27734199	CCUGCUAUCCUUAGUCAAU
Dnd1	213236	NM_173383	27734199	CUGAGUGAACCUAGGACCA
Eif4enif1	74203	NM_023743	12963810	GAUAAGGACCUGCGAGACU
Eif4enif1	74203	NM_023743	12963810	ACUACAAGGACAAACGUUU
Eif4enif1	74203	NM_023743	12963810	GAACAAGAUUAUCGACCUA
Eif4enif1	74203	NM_023743	12963810	GAGCGUAGAAGAAACGAUU

Eif4g3	230861	NM_001256198	371876056	GGACAAGGCUUCACGACGA
Eif4g3	230861	NM_001256198	371876056	GGCCAAGAACAGUCGUGUA
Eif4g3	230861	NM_001256198	371876056	CAAGAAACCCUGUAGCGUA
Eif4g3	230861	NM_001256198	371876056	GAGAGAUGUUGGAGACCGU
Elavl2	15569	NM_207686	46592825	GCGACAUUAUAAACGCUUA
Elavl2	15569	NM_207686	46592825	CAGGAAAUUUGCAACGGUU
Elavl2	15569	NM_207686	46592825	CUAAAGUGAACAAGCGAGU
Elavl2	15569	NM_207686	46592825	GAGCUUGGGUUACGGAUUU
Etl4	208618	XM_001003433	94367904	CGCCAGAGGCAGUCGUACU
Etl4	208618	XM_001003433	94367904	GGAUCCGGUUGAACAUUUA
Etl4	208618	XM_001003433	94367904	UGGUGAGGUCGGCCAAUCA
Etl4	208618	XM_001003433	94367904	GUGACUCUUUGGAAGCCAU
Fmr1	14265	NM_008031	51036612	GAUGAGCAGUUGCGACAAA
Fmr1	14265	NM_008031	51036612	CCAGAAGAUUUACGACAAA
Fmr1	14265	NM_008031	51036612	GCACCAAGUUGUCUCUUAU
Fmr1	14265	NM_008031	51036612	GAUCAUUCCCGAACAGAUA

Fxr1	14359	NM_008053	17998693	CUAAUUACACCUCCGGUUA
Fxr1	14359	NM_008053	17998693	CGGACGAAGUUGAUGCUUA
Fxr1	14359	NM_008053	17998693	UAGAGUAUCACAUCGCUUA
Fxr1	14359	NM_008053	17998693	UGAAGACACCGGAACGUUU
G3bp2	23881	NM_011816	31982756	UCGAGGAGAAGUACGUUUA
G3bp2	23881	NM_011816	31982756	GAAUGAUGCGUGAUCGUGA
G3bp2	23881	NM_011816	31982756	GCGCUUACUAUGACGCACA
G3bp2	23881	NM_011816	31982756	GUACCUAACUCAUCGUGUU
Gdf3	14562	NM_008108	75677586	GGGCCAGACAUCCGAGUUU
Gdf3	14562	NM_008108	75677586	GGACUUAUGCUACGUGAAG
Gdf3	14562	NM_008108	75677586	CCACCAAGCUCUCGCCCAU
Gdf3	14562	NM_008108	75677586	UCUCAAUGACCACGUAUUU
Hnrnpdl	50926	NM_016690	148664249	CGAAAUCAAAGUUGCGCAA
Hnrnpdl	50926	NM_016690	148664249	AAACCAUCUUACAGGACGA
Hnrnpdl	50926	NM_016690	148664249	ACCAGCAGGAUGACGGUAA
Hnrnpdl	50926	NM_016690	148664249	ACAUAGAGGAGUUCGCAGA

Ilf3	16201	NM_010561	6754339	CAUCAUCACUUCAUCACAA
Ilf3	16201	NM_010561	6754339	GAAGAGACGUGGCCUCAAA
Ilf3	16201	NM_010561	6754339	GCAAGAACCCUGUUAUGGA
Ilf3	16201	NM_010561	6754339	GGAGAAACGCUAUCAGUCA
Itch	16396	XM_902524	82887085	CAUCUCAGCCAAACUUAAA
Itch	16396	XM_902524	82887085	CAGCUUGAUUCCAUGGGUA
Itch	16396	XM_902524	82887085	GAAACGUCAUGCUCCGAGA
Itch	16396	XM_902524	82887085	AUUGGGAACUGCUGGAUUA
Kif22	110033	NM_145588	21704181	GGUCAAGGUAGAUCAGCGU
Kif22	110033	NM_145588	21704181	CGGCUUAACCAGCGCUCUU
Kif22	110033	NM_145588	21704181	CCCCAAAGCGAGAACGAAU
Kif22	110033	NM_145588	21704181	GGAAGUCUAUGUAGGUUCA
Kif23	71819	NM_024245	256985167	ACAUAUGGCGUGACGGGAA
Kif23	71819	NM_024245	256985167	UGAGAGAGAACCAGACGUA
Kif23	71819	NM_024245	256985167	GAAAGAACCAACCGUACUA
Kif23	71819	NM_024245	256985167	ACUGGUAGUCGGAAACGAA

Larp1	73158	XM_985836	94391540	GAGUUGAUGUUGCGGAGUA
Larp1	73158	XM_985836	94391540	GCAAGAAUACCUCGGCAAA
Larp1	73158	XM_985836	94391540	ACACACUGACUUUGCGAAA
Larp1	73158	XM_985836	94391540	UGGACAAUUUAGAGCGAGA
Lsm14b	241846	NM_177727	31343196	CGGUCAGGCAAUCGGAGAA
Lsm14b	241846	NM_177727	31343196	GGUCUAGAGAAGUUGGUAA
Lsm14b	241846	NM_177727	31343196	AGUUCUAGGCGGACAACAU
Lsm14b	241846	NM_177727	31343196	CUACUUAGUUUGGACUUAA
Meioc	268491	XM_921161	82936686	GCUAAUAUCUCUACUGCUC
Meioc	268491	XM_921161	82936686	GAAAGUCCAUGAAAGCAUA
Meioc	268491	XM_921161	82936686	GGUGAAACAAGUAAACAUU
Meioc	268491	XM_921161	82936686	GAUAGACACUUGGAGUCAA
Mex3a	72640	NM_001029890	71274133	GACCAUCCGUGUACGAGUA
Mex3a	72640	NM_001029890	71274133	CCACACAAGCCAUCCGAAU
Mex3a	72640	NM_001029890	71274133	GAUAUCUGUUUCCGGGCUA
Mex3a	72640	NM_001029890	71274133	GAUGGAGGAUGUCGAUAAA

Mff	75734	NM_029409	40254162	CGAAGGUAUUAGUCAGCGA
Mff	75734	NM_029409	40254162	GAGAAUAAAGAGCGUGCUA
Mff	75734	NM_029409	40254162	GAAAAUGCUGUUCGCCAAA
Mff	75734	NM_029409	40254162	CAAAUUAAAUAGACGUCUC
Mkrn1	54484	NM_018810	9055273	CGAUGGACUUGGUGGGCGA
Mkrn1	54484	NM_018810	9055273	GCGAAGCUGAGUCACGAAA
Mkrn1	54484	NM_018810	9055273	CCUUUGACAACGACGAAGA
Mkrn1	54484	NM_018810	9055273	GUCAGUAGUUGUUAACGUA
Mkrn2	67027	NM_023290	88853569	GCAAGUGCCUUUACCGUCA
Mkrn2	67027	NM_023290	88853569	GACCCGAGGUGACGUGGUA
Mkrn2	67027	NM_023290	88853569	UAAUAGAGUUGUCGAGCCU
Mkrn2	67027	NM_023290	88853569	GGGAAGAAGGCAUGCAAAU
Mov10	17454	NM_001163441	254540182	CGAAUAAAGCCAGGGUCGA
Mov10	17454	NM_001163441	254540182	CUACAAGGGUUUCGUGCAC
Mov10	17454	NM_001163441	254540182	GGUACUAACCCUACGGCUU
Mov10	17454	NM_001163441	254540182	UCACCAAACUGCUCCGCAA

Msi1	17690	NM_008629	6678939	AGGAAGGGCUGCGCGAAUA
Msi1	17690	NM_008629	6678939	GGUCAUGCCCUACGGAAUG
Msi1	17690	NM_008629	6678939	CCAAUGGGUACCACUGAAA
Msi1	17690	NM_008629	6678939	GUAUUGGGAUGCUGGGUUA
N4bp2	333789	NM_001024917	67906815	GGAAUUGGUAUCCGGAACA
N4bp2	333789	NM_001024917	67906815	GAUUUAUAUGAGCGAUGUA
N4bp2	333789	NM_001024917	67906815	CCUUAUAUCCAGCACACGA
N4bp2	333789	NM_001024917	67906815	AAUCUGUAAUGGAGCGACA
Nfx1	74164	NM_023739	24475690	CGGCAUAAAUGUAACGAGA
Nfx1	74164	NM_023739	24475690	UGAUAUCUCUUGCGGAUUA
Nfx1	74164	NM_023739	24475690	GGAAGUAAUGCUCGGAAAA
Nfx1	74164	NM_023739	24475690	GCCCAAAGUCUGAGGACGA
Nsd1	18193	NM_008739	76443702	GGGCAUAUCCCACGAGCUA
Nsd1	18193	NM_008739	76443702	ACAAAGGAGCAGCGGUUAA
Nsd1	18193	NM_008739	76443702	CCUUAAAGACUGAGCGGAA
Nsd1	18193	NM_008739	76443702	CCUAGAGUGUCUUGGGAAU

Nsun5	100609	NM_145414	21703825	GGUCUCUUGCUGCGAGUUA
Nsun5	100609	NM_145414	21703825	AGACAGAUCUGCACGAGCA
Nsun5	100609	NM_145414	21703825	CUGGAAGACUUGCGAGCUC
Nsun5	100609	NM_145414	21703825	GGUGCUAGUGUAUGAAUUA
Nufip1	27275	NM_013745	31543354	GAAUGAUCCUGUUUCGUUU
Nufip1	27275	NM_013745	31543354	ACUGAUAGAAAGCCGUAAA
Nufip1	27275	NM_013745	31543354	UGGCCUUAGUGGACGAGUA
Nufip1	27275	NM_013745	31543354	GUAUACAAAUCAAACGUCA
Nufip2	68564	NM_001024205	66730552	AGGAAAGCUAGGCGCAAUA
Nufip2	68564	NM_001024205	66730552	AAACACAGGUAGUAGCGAA
Nufip2	68564	NM_001024205	66730552	AGGCAAACCUCAAUUGAUA
Nufip2	68564	NM_001024205	66730552	CUGCGAGAAUGCUGGAUAA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUCGACUAA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUGUGUGA

ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUUUCUGA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUUUCCUA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUCGACUAA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUGUGUGA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUUUCUGA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUUUCCUA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUCGACUAA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUGUGUGA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUUUCUGA

ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUUUCCUA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUCGACUAA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUCGACUAA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUGUGUGA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUUUCUGA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUUUCCUA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUCGACUAA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUGUGUGA
ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUUUCUGA

ON-TARGETplus Non-targeting Control	0		0	UGGUUUACAUGUUUUCCUA
Pabpc4	230721	NM_148917	22507390	UCAAUGACCGUAAAGUGUU
Pabpc4	230721	NM_148917	22507390	GUUGUGACCCAGAUGAUAA
Pabpc4	230721	NM_148917	22507390	GUACCCAAGACUUAAGUUU
Pabpc4	230721	NM_148917	22507390	AGAAAUGAAUGGAAAGGAA
Paip1	218693	NM_145457	31981785	GGGAUGAAGUGACUCGAAA
Paip1	218693	NM_145457	31981785	UCAGGUUGGUCUUCGGGAA
Paip1	218693	NM_145457	31981785	UGUCUAAGCUGUCUGCGAA
Paip1	218693	NM_145457	31981785	CUUCAAAGGUGCAGGACGG
Pan3	72587	NM_028291	61744462	GCAAGCGAAUUUAUACCUA
Pan3	72587	NM_028291	61744462	GAUCCAACCUGGUCAGAGA
Pan3	72587	NM_028291	61744462	CGACUUACUUCUAUACAGA
Pan3	72587	NM_028291	61744462	GUAACGAUCGGAACAGUAA
Patl1	225929	NM_172635	27369915	GAUCAUGUGUAUCCGAAAA
Patl1	225929	NM_172635	27369915	GUGAAUAAUCCUCGGAAGA

Patl1	225929	NM_172635	27369915	AAAUAAACCAGCCGAUAUA
Patl1	225929	NM_172635	27369915	GCUAUAGAAUCAACGCAAA
Phf3	213109	NM_001081080	124487308	GUCUAAUGAUGCUCGGUUA
Phf3	213109	NM_001081080	124487308	AAUCAGAAUUUCCGUCUAA
Phf3	213109	NM_001081080	124487308	AGUAAGAAGUUUGCGACAA
Phf3	213109	NM_001081080	124487308	CAAAAUAUGCAUAGGUCGA
Pum1	80912	NM_030722	13507675	GAAACAACCGGUACCCUAA
Pum1	80912	NM_030722	13507675	GAACAAGGGUACGUUUGAU
Pum1	80912	NM_030722	13507675	GAAUGGUUCUGAGGGCUUA
Pum1	80912	NM_030722	13507675	UGGAUCCUCUUCAAGCUUA
Pwp2	110816	NM_029546	30102929	GCGAGAGUUCAAUGCGUUU
Pwp2	110816	NM_029546	30102929	CUUCACAGGUACCGGAAUU
Pwp2	110816	NM_029546	30102929	GAGUGGUGCUUUACCGAUA
Pwp2	110816	NM_029546	30102929	CGUCAGUGGUCUUUGCUUU
Rbbp6	19647	NM_175023	163937859	CCUCUAAACUCAACUACGA
Rbbp6	19647	NM_175023	163937859	UCACCAACGCACAGACGAA

Rbbp6	19647	NM_175023	163937859	UGUUAUAAGUCGAACUGAA
Rbbp6	19647	NM_175023	163937859	CAGAUUAUGGGGAGAGAAA
Rbm22	66810	XM_994591	94405411	UAACAAGCUGAUCGUGAAU
Rbm22	66810	XM_994591	94405411	UAUCAAAGACCGGUACUAU
Rbm22	66810	XM_994591	94405411	CCUGUUAGACCUUGAAUAU
Rbm22	66810	XM_994591	94405411	CUUAUAUACGAAUGACCAA
Rbpms2	71973	NM_028030	31980693	GGAUGGAAGUACCGCCAGU
Rbpms2	71973	NM_028030	31980693	GUUUAUGAAUUGUGACACU
Rbpms2	71973	NM_028030	31980693	GAUCUGAGGCCUUGGAUAC
Rbpms2	71973	NM_028030	31980693	UUGUUCAGCCUGUCAGAUA
Rc3h1	381305	NM_001024952	68131568	ACUUGAAGGUAGCCGGAAA
Rc3h1	381305	NM_001024952	68131568	GCGAAUUUGAACCGACUAA
Rc3h1	381305	NM_001024952	68131568	AAGAGAACAAUUACGGCUA
Rc3h1	381305	NM_001024952	68131568	GAUCAGGAGUUGUUUCCAA
Rc3h2	319817	NM_001100591	155030208	ACUAGUAUGUGCCGAGAUU
Rc3h2	319817	NM_001100591	155030208	GAGUGUGAGUGGCGCAAAA

Rc3h2	319817	NM_001100591	155030208	GGAUCAUGUUAUUGGUCGU
Rc3h2	319817	NM_001100591	155030208	GAUUAUGUCCCUUACGUGA
Rpa1	68275	NM_026653	18390320	CGCAUGAUCUUAUCGGCAA
Rpa1	68275	NM_026653	18390320	AGGGAGACAAGGCGGACUA
Rpa1	68275	NM_026653	18390320	AAUCAAAGGCCUAUCGUUA
Rpa1	68275	NM_026653	18390320	GAGAAGUGUGACCGGGAAU
Safb2	224902	NM_001029979	71533180	GGAGGAAGCGAGUGCGUUU
Safb2	224902	NM_001029979	71533180	ACACAAACAUGACGGGCCA
Safb2	224902	NM_001029979	71533180	AGGCAGAGGGAGCGUGAAA
Safb2	224902	NM_001029979	71533180	GGACAGAUACCGAGACUUU
Secisbp2	75420	NM_029279	116812874	ACAAGCGACCCGACCGCAA
Secisbp2	75420	NM_029279	116812874	UCCUAGGAGACGUCGGCAA
Secisbp2	75420	NM_029279	116812874	GGUUAUAUUGUUCGGCAGA
Secisbp2	75420	NM_029279	116812874	CUGUCAGUUGCGUCGGAAA
Spop	20747	NM_025287	13384633	AAAUGGUGUUUGCGAGUAA
Spop	20747	NM_025287	13384633	AAGAUUACCUGUCGCUUUA

Spop	20747	NM_025287	13384633	CGAAAGGGCUAGAUGAAGA
Spop	20747	NM_025287	13384633	GCCGAAAGCUGGUGCUAUA
Taf1d	75316	NM_029248	21312925	CCUACUGUGAAACGUGUUA
Taf1d	75316	NM_029248	21312925	UCACUAGAAGGUCGCAAAU
Taf1d	75316	NM_029248	21312925	GCCAUUAACUGAAGAUGAA
Taf1d	75316	NM_029248	21312925	GUUCAGAAUUCUCAAGGAA
Tceb3	27224	NM_013736	31543846	CAGAAUAGCCCUUUGCGAU
Tceb3	27224	NM_013736	31543846	GGAUAAAGCCAGCGCCAUA
Tceb3	27224	NM_013736	31543846	UCUCUUAGCUGACGUUAAA
Tceb3	27224	NM_013736	31543846	ACACAAGGACUCCGAGAAA
Tdrd3	219249	NM_172605	31982086	UCGUUAAGUUCACGGACUA
Tdrd3	219249	NM_172605	31982086	AGGAUGGAAUGCCGAGCAA
Tdrd3	219249	NM_172605	31982086	GGAAAUACUGAGUUCGAUA
Tdrd3	219249	NM_172605	31982086	GUGGAAACCUGGCGAUGAA
Tet1	52463	XM_920196	94388307	AGGAAGAGGCGACUACGUU
Tet1	52463	XM_920196	94388307	GGUUGGAAGCUGUUCGCUA

Tet1	52463	XM_920196	94388307	AUUUCAACUCCGACGUAAA
Tet1	52463	XM_920196	94388307	CCACAUCAACACACGGAGU
Tnrc6a	233833	XM_901687	94380990	AAUUAAUACUGUGCGGGAA
Tnrc6a	233833	XM_901687	94380990	AAGUUAAGAUACUGACGUA
Tnrc6a	233833	XM_901687	94380990	ACUAAAGAAUUUCGCUCUA
Tnrc6a	233833	XM_901687	94380990	GAGUGAUGUUCUCCGUCUA
Tnrc6b	213988	NM_177124	67782333	CCCAGAAGGUCACGGAACA
Tnrc6b	213988	NM_177124	67782333	ACAGUCAAUUAGUCGGAAA
Tnrc6b	213988	NM_177124	67782333	ACAAAGCACUGGAGAUCGA
Tnrc6b	213988	NM_177124	67782333	ACUGAAUCUUCUUCCGAAA
Tnrc6c	217351	XM_982911	94391253	CAGCAGGCCAAUACCGAAA
Tnrc6c	217351	XM_982911	94391253	CUGCCUACGUGGUCGGGAA
Tnrc6c	217351	XM_982911	94391253	UGUUACAGGCCCAGCGGAA
Tnrc6c	217351	XM_982911	94391253	GCACAGAGCUGAGGCGUGU
Top3b	21976	NM_011624	31981463	CAUCAACCGCCCUCGGAAA
Top3b	21976	NM_011624	31981463	GUGCGUGGAAUGCGAGAAU

Top3b	21976	NM_011624	31981463	ACACCUAUGAAGUCGGCUA
Top3b	21976	NM_011624	31981463	AGACGAGCCCACCGGACUA
Traf4	22032	NM_009423	31543888	CCUCUCUAGUGGACGAUAU
Traf4	22032	NM_009423	31543888	CAGCUAUGGGCGUCGUCUA
Traf4	22032	NM_009423	31543888	GGAACUACGUGCGGGAUGA
Traf4	22032	NM_009423	31543888	GGGUAAGGGUAGUUAGAAG
Trim25	217069	NM_009546	46849790	CCGUACAACCUUCGAAGAA
Trim25	217069	NM_009546	46849790	AGGAAUAAACUCACUAUCA
Trim25	217069	NM_009546	46849790	GGAAAUAGAAACAGACGCA
Trim25	217069	NM_009546	46849790	CUGCAAAAGAACACGGUAA
Trim71	636931	XM_980887	94387742	GAUCAUAGUGGCCGACAAA
Trim71	636931	XM_980887	94387742	CAUCAUAGUGGCUGACCGA
Trim71	636931	XM_980887	94387742	CAAGAAGAUGACCGCAUUA
Trim71	636931	XM_980887	94387742	GGGCAAGAUCCUCGUUUCA
Ubap2l	74383	NM_153489	23943917	CAAUAGCAGUGGCAAUACA
Ubap2l	74383	NM_153489	23943917	CAACAGAACCAGACGCAGA

Ubap2l	74383	NM_153489	23943917	AGGUAGUUCUGGUAGGCGA
Ubap2l	74383	NM_153489	23943917	GAUAUGGGUCUCAUGGAUA
Uhrf1	18140	NM_010931	31981595	CCACACACUCUUCGAUUAU
Uhrf1	18140	NM_010931	31981595	CACGUGAGCUAUACGGCAA
Uhrf1	18140	NM_010931	31981595	GGAAGUGUUUCACGUGGAA
Uhrf1	18140	NM_010931	31981595	CUGCAGAGGGCAACCGCUA
Usp10	22224	NM_009462	6678492	GAGGGAAACCGCUGACGUA
Usp10	22224	NM_009462	6678492	CAGUAUAUCUUUGGCGAUU
Usp10	22224	NM_009462	6678492	CACCAUGGCAACAGCGCUA
Usp10	22224	NM_009462	6678492	GCACCAGAGGAUCGAGUUU
Wwp2	66894	NM_025830	31543965	CCUCAGGACUCCCGACAUA
Wwp2	66894	NM_025830	31543965	GAAGAACAAUGGCGGCAAA
Wwp2	66894	NM_025830	31543965	CAUGUGAAGAUCAGCGUUU
Wwp2	66894	NM_025830	31543965	GAACAAAGCCCUGGCGCUA
Xrn1	24127	NM_011916	6756024	GAAGAUGGGCGUCGAUGUA
Xrn1	24127	NM_011916	6756024	GAGGUGUUGUUUCGAAUUA

Xrn1	24127	NM_011916	6756024	AAACAUGACUCGCAACGAA
Xrn1	24127	NM_011916	6756024	GCAGAGUAGUCGUGGUGAA
Ybx2	53422	NM_016875	8393790	CGACGUAGGUUCCGGAGGU
Ybx2	53422	NM_016875	8393790	CAUUGGAGGGUGAUCAACA
Ybx2	53422	NM_016875	8393790	GGACAGAACCUGGCAGCGA
Ybx2	53422	NM_016875	8393790	GGUUACGGAUUCAUCAACA
Zc3hav1	78781	NM_028864	142383007	CCAUGAUCUUGCUGCGUCA
Zc3hav1	78781	NM_028864	142383007	CCAGUGAGCCCUCGAGGAU
Zc3hav1	78781	NM_028864	142383007	GCAAAUAUUCUCACGAUGU
Zc3hav1	78781	NM_028864	142383007	AUGCACAGUCCCAGCGGAA
Zcchc3	67917	NM_175126	34328413	GGAACGAGACCGUGGACGU
Zcchc3	67917	NM_175126	34328413	CGACGAAACUGGCAGCGGU
Zcchc3	67917	NM_175126	34328413	GCAUGGACCCUGACGACAU
Zcchc3	67917	NM_175126	34328413	CGAAGAUGCCACCGAGCGA
Zfr	22763	NM_011767	7106452	GAGAUGUAACCUCGGGUAU
Zfr	22763	NM_011767	7106452	CAAAUGAAUCAGCGCUUUA

Zfr	22763	NM_011767	7106452	GUUAUGAAGCAGCCGUAUA
Zfr	22763	NM_011767	7106452	UGUCUGAGGCAGCGAUAAU

Chapter 10: Annex II: list of chosen genes for siRNA screen with corresponding criteria

List of selected proteins for the LINE1 repressor screening in pluripotency. Highlighted in green, there are the proteins that were identified in the LINE1 ORF1p interactome as pluripotency-specific. In blue, those proteins overlapping between our dataset and the ORF1p interactome in hESCs (Vuong et al., 2019). In beaux are highlighted the rest of included proteins in the screen for the relevance their biological function could have for LINE1 repression or their level of enrichment.

Symbol	**Reasoning**
NFX1	Stem cells specific according to mESCs specific database. Ubiquitin ligase
MKRN1	Stem cells specific according to mESCs specific database. Ubiquitin ligase
TNRC6C	Stem cells specific according to mESCs specific database. Translation control
Meioc	Stem cells specific according to mESCs specific database
DND1	Stem cells specific according to mESCs specific database
ELAVL2	Stem cells specific according to mESCs specific database
TOP3B	Stem cells specific according to mESCs specific database
G3BP2	Stem cells specific according to mESCs specific database
KIF23	Stem cells specific according to mESCs specific database
CEBPZ	Stem cells specific according to mESCs specific database
TRIM71	Stem cells specific according to mESCs specific database. Ubiquitin ligase
NUFIP1	Stem cells specific according to mESCs specific database
KIF22	Stem cells specific according to mESCs specific database
NSUN5	Stem cells specific according to mESCs specific database
Zcchc3	ESC specific but palready detected in other datasets. Some role in immune system, cgas activator
PWP2	Stem cells specific according to mESCs specific database
GDF3	Stem cells specific according to mESCs specific database

TAF1D	Stem cells specific according to mESCs specific database
PAIP1	Stem cells specific according to mESCs specific database
DAZL1	Stem cells specific according to mESCs specific database
TET1	Stem cells specific according to mESCs specific database
Tnrc6b	hESC and mESC dataset concordance, not in other datasets
Mex3a	hESC and mESC dataset concordance, not in other datasets
Pum1	hESC and mESC dataset concordance, not in other datasets/ translational regulator
Msi1	hESC and mESC dataset concordance, not in other datasets
Lsm14b	hESC and mESC dataset concordance, not in other datasets
Eif4enif1	hESC and mESC dataset concordance, not in other datasets,Translational repression through P-bodies
Atxn2l	hESC and mESC dataset concordance, not in other datasets
Nufip2	hESC and mESC dataset concordance, not in other datasets
Ubap2l	hESC and mESC dataset concordance, not in other datasets
Zfr	hESC and mESC dataset concordance, not in other datasets
Phf3	hESC and mESC dataset concordance, not in other datasets
Hnrnpdl	hESC and mESC dataset concordance, not in other datasets
Rbpms2	hESC and mESC dataset concordance, not in other datasets
Rbm22	hESC and mESC dataset concordance, not in other datasets
Fmr1	The protein encoded by this gene binds RNA and is associated with polysomes. The encoded protein may be involved in mRNA trafficking from the nucleus to the cytoplasm. VERY MAINSTREAM PROTEIN PROBABLY VERY STUDIED
Cpsf4	hESC and mESC dataset concordance, not in other datasets
Mff	2nd hit
Traf4	Very enriched, very significative, top 5, seems to have relation with innate immune system.
Rc3h1	Very enriched, top 5, seems to have relation with innate immune system.
Tnrc6a	Very enriched, very significative, top 8 Seems to have a role in postranscriptional regulation
Nsd1	Histone methyltransferase, quite enriched
Wwp2	Ubiquitin ligase, Innate immune response negative regulation https://pubmed.ncbi.nlm.nih.gov/23479606/

Spop	Ubiquitin ligase, Innate immune response negative regulation https://www.nature.com/articles/s41590-019-0454-6
Acer2	Top hit, regulates cell proliferation, golgi
Ybx2	RNA transport and storage, maternal RNA retention, expressed in germ cells
Eif4g3	Translation negative regulation
Tdrd3	Transcriptional activator recognizes H3R17me2a and H4R3me2a, regulation of assembly and disassembly of stress granules, regulation of target RNAs translation
Etl4	Very enriched, not known function at all
Secisbp2	involved in selenocysteine incorporation. Among its related pathways are Viral mRNA Translation and Metabolism. Essential for embryonic development. Lethal before gastrulation. Regulation of mRNA decay, regulation of translation elongation and readthrough
Safb2	involved in selenocysteine incorporation. Among its related pathways are Viral mRNA Translation and Metabolism. Essential for embryonic development. Lethal before gastrulation
N4bp2	Role in DNA repair, not very studied, very enriched
Adar	RNA editing, known L1 repressor
Itch	ubiquitin ligase
Pabpc4	mRNA stability
Trim25	ubiquitin ligase, involved in RIG-I signalling
Zc3hav1	Degradation of viral RNAs
Larp1	Known interactor, translational inhibitor blocks eif4F
Rbbp6	ubiquitin ligase
Uhrf1	E3 ubiquitin-protein ligase UHRF1
Mkrn2	Putative E3 ubiquitin-protein ligase UHRF1
Rc3h2	Protein ligase activity
Eloa	Protein ligase activity
Usp10	Ubiquitin carboxyl-terminal hydrolase 10
Ago2	Translation inhibition
Pan3	RNA degradation
Patl1	RNA degradation
Xrn1	RNA degradation
Ddx6	P-bodie formation, translation control, RNA decay
Ilf3	Upon viral infection, ILF3 accumulates in the cytoplasm and participates in the innate antiviral response (PubMed:21123651). Mechanistically, ILF3 becomes phosphorylated and activated by the double-stranded RNA-activated protein kinase/PKR which releases ILF3 from cellular mature circRNAs. In turn, unbound ILF3 molecules are able to interact with and thus inhibit viral mRNAs
Fxr1	anatomical structure. Highly enriched
Dicer1	Known interactor

Rpa1	Positive control gamma H2AX
Mov10	Known interactor

Annex III: Common LINE1 **ORF1p interactors** between mESCs and somatic datasets

Item	Occurences	Present In
Mov10	6	mESCs, Moldovan, Taylor, Liu, Goodier, Mita
Pura	6	mESCs, Moldovan, Taylor, Liu, Goodier, Mita
Pabpc1	5	mESCs, Taylor, Liu, Goodier, Mita
Pabpc4	5	mESCs, Taylor, Liu, Goodier, Mita
Upf1	5	mESCs, Moldovan, Taylor, Liu, Mita
Dhx9	4	mESCs, Moldovan, Liu, Goodier
Larp1	4	mESCs, Moldovan, Goodier, Mita
Matr3	4	mESCs, Moldovan, Goodier, Mita
Top1	4	mESCs, Taylor, Goodier, Mita
Ddx17	3	mESCs, Goodier, Mita
Hnrnpc	3	mESCs, Goodier, Mita
Hnrnpk	3	mESCs, Goodier, Mita
Hnrnpl	3	mESCs, Moldovan, Goodier
Igf2bp1	3	mESCs, Goodier, Mita
Igf2bp3	3	mESCs, Moldovan, Mita
Ilf3	3	mESCs, Moldovan, Goodier
PARP1	3	Taylor, Liu, Mita
PCNA	3	Taylor, Liu, Mita
PURB	3	Taylor, Liu, Mita
Raly	3	mESCs, Goodier, Mita
Srsf1	3	mESCs, Liu, Goodier
Syncrip	3	mESCs, Liu, Goodier
Tra2a	3	mESCs, Goodier, Mita
Tra2b	3	mESCs, Goodier, Mita
XRCC5	3	Liu, Goodier, Mita
Ybx1	3	mESCs, Goodier, Mita
Ybx3	3	mESCs, Liu, Mita
Zcchc3	3	mESCs, Taylor, Mita
Ago2	2	mESCs, Liu
Cpsf7	2	mESCs, Liu
Ddx5	2	mESCs, Goodier
Ddx6	2	mESCs, Taylor
Elavl1	2	mESCs, Goodier
Ezh2	2	mESCs, Liu

Fam120a	2	Moldovan, Goodier
Fam98a	2	mESCs, Goodier
FKBP4	2	Taylor, Mita
HAX1	2	Taylor, Mita
Hist1h1b	2	mESCs, Goodier
Hist1h1e	2	mESCs, Goodier
HMCES	2	Taylor, Mita
Hnrnpa1	2	mESCs, Goodier
Hnrnpa2b1	2	mESCs, Goodier
Hnrnpf	2	mESCs, Mita
Hnrnph1	2	mESCs, Mita
Hnrnph3	2	mESCs, Goodier
Hnrnpm	2	mESCs, Mita
Hnrnpr	2	mESCs, Goodier
HNRNPU	2	Goodier, Mita
HSP90AA1	2	Taylor, Mita
HSP90AB1	2	Taylor, Mita
HSPA1A	2	Taylor, Mita
HSPA8	2	Taylor, Mita
Igf2bp2	2	mESCs, Goodier
Ilf2	2	mESCs, Goodier
IPO7	2	Taylor, Mita
Larp7	2	mESCs, Taylor
Luc7l3	2	mESCs, Goodier
MARS	2	Liu, Goodier
Mcm10	2	mESCs, Liu
Morc1	2	mESCs, Liu
Mphosph8	2	mESCs, Liu
NAP1L1	2	Taylor, Mita
Nat10	2	mESCs, Goodier
Ncbp1	2	mESCs, Mita
Ncl	2	mESCs, Moldovan, Goodier
Nop56	2	mESCs, Goodier
Pabpc4l	2	mESCs, Taylor
PCBP2	2	Goodier, Mita
Ppan	2	mESCs, Goodier
PRKDC	2	Liu, Mita
Prrc2a	2	mESCs, Liu
Ptbp1	2	mESCs, Goodier
Rbm10	2	mESCs, Liu
RPS27A	2	Taylor, Mita
Safb	2	mESCs, Liu
Sltm	2	mESCs, Liu

Srsf10	2	mESCs, Goodier
Srsf6	2	mESCs, Goodier
Srsf7	2	mESCs, Mita
Stau1	2	mESCs, Goodier
STK38	2	Goodier, Mita
Thrap3	2	mESCs, Mita
TROVE2	2	Taylor, Goodier
TUBB	2	Taylor, Mita
TUBB4B	2	Taylor, Mita
Utp15	2	mESCs, Liu
XRCC6	2	Liu, Mita
Zfr	2	mESCs, Liu