LUND UNIVERSITY

# Transposable Elements in Neural Progenitor Cells

Fasching, Liana

2015

Link to publication

*Citation for published version (APA):*
Fasching, L. (2015). *Transposable Elements in Neural Progenitor Cells*. [Doctoral Thesis (compilation), Neurobiology]. Molecular Neurogenetics, Faculty of Medicine, Lund University.

*Total number of authors:*
1

**General rights**
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

**Take down policy**
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

*Academic Dissertation*

# Transposable Elements in Neural Progenitor Cells

**Liana Fasching**

2015

With approval of the Faculty of Medicine of Lund University,
This thesis will be defended
at 09:00 on October 23rd, 2015 in Segerfalksalen,
Wallenberg Neuroscience Center, Lund, Sweden.

*Faculty Opponent:*

**Dr. Joshua Dubnau**
Cold Spring Harbor Laboratory,
New York, USA.

LUND
UNIVERSITY

Organization
LUND UNIVERSITY
Molecular Neurogenetics,
Department of Experimental Medical Sciences,
Faculty of Medicine, Lund University, Sweden

Document name
DOCTORAL DISSERTATION

Date of issue
2015-10-23

Sponsoring organization

Author(s)
Liana Fasching

Title and subtitle
Transposable Elements in Neural Progenitor Cells

Abstract More than 90% of DNA does not code for proteins and for a long time these sequences were referred to as "junk DNA" due to their unknown purpose. With the advent of new technologies it is now known, that the non-coding part of the genome is of great importance for regulating gene expression and is therefore indispensable.

Transposable elements comprise about 50% of the genome and co-exist as symbionts regulated by epigenetic mechanisms - a highly defined machinery that controls gene expression and is mandatory for a proper development and maintenance of an organism. Although transposable elements are associated with diseases, their role in fine-tuning the host gene expression becomes more and more evident, which seems to justify the positive selection during evolution.

A transposable element called *Line-1* was found to be active in neural progenitor cells and in the brain. Several studies report *Line-1* transcription and frequent retrotransposition during normal brain development, with further evidence that *Line-1* induced retrotransposition can influence neuronal gene expression. Today, there is few published data focusing on epigenetic regulation of transposable elements in neural progenitor cells.

In this thesis, I identify TRIM28 as key regulator of certain groups of transposable elements in mouse and human neural progenitor cells. This feature is unique compared to other somatic tissues, where DNA-methylation is prevalent.

Here I demonstrate, that transposable elements *MMERVK10C* and *IAP1* in mouse neural progenitor cells are repressed by the establishment of H3K9me3-associated heterochromatin. De-repressed *MMERVK10C* and *IAP1* furthermore activate nearby genes and generate long non-coding RNAs. Homozygous TRIM28 knockout is lethal, while mice with mono-allelic TRIM28 expression are characterised with a distinct behavioural phenotype.

Moreover we are also able to show that TRIM28 is regulating a fraction of young *Alu*-elements in human neural progenitor cells, which is not the case in human embryonic stem cells. Furthermore, we report that transcribed *Alu*-elements influence gene expression of close-by genes.

Studying pluripotent cells revealed that TRIM28 modulates transposable elements in mouse embryonic stem cells. Activation of transposable elements upon TRIM28 depletion induces changes in gene expression of close-by genes and causes alteration of the repressive chromatin mark H3K9me3 at transposable element loci. Upregulated genes were shown to have bivalent promoters, characterised by H3K4me3 and H3K27me3 and lay close to H3K9me3 regulated transposable elements. These findings in mouse embryonic stem cells are highly relevant for the interpretation of my studies in neural progenitor cells.

Taken together, this thesis demonstrates that the regulation of transposable elemenets in mouse and human neural progenitor cells is distinct compared to previous reports regarding somatic tissues. These results provide novel insights into why the brain has developed into such a complex organ with so many different cell types.

Key words:
Transposable Elements, TRIM28, Epigenetic Regulation, Gene Expression, Neural Progenitor Cells

Classification system and/or index termes (if any):

Supplementary bibliographical information:

Language
English

ISSN and key title:
1652-8220

ISBN
978-91-7619-186-6

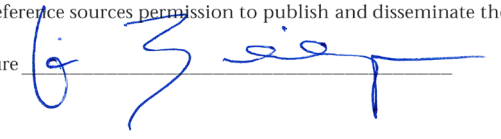
Recipient's notes

Number of pages
116

Price

Security classification

Distribution by (name and address)
I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature 

Date 2015-09-10

# Transposable Elements in Neural Progenitor Cells

**Liana Fasching**

2015

Laboratory of Molecular Neurogenetics, Wallenberg Neuroscience Center,
Department of Experimental Medical Sciences, Faculty of Medicine,
Lund University, Lund Sweden.

LUND
UNIVERSITY

Cover: “When the scientist is away, the mice will play”.
Beautifully interpreted and painted by Dieter Fasching.

ISSN 1652-8220
ISBN 978-91-7619-186-6
Lund University, Faculty of Medicine, Doctoral Dissertation Series 2015:107

© 2015 Liana Fasching and the respective publishers

Printed by Exactaprinting AB, Malmö, Sweden

*To my parents*

In any given moment we have two options:
to step forward into growth or to step back into safety.

**Abraham Maslow**

And those who were seen dancing were thought
to be insane by those who could not hear the music.

**Friedrich Nietzsche**

# Table of Contents:


**Original Papers and Manuscripts** 9
**Abstract** 11
**Lay Summary** 13
**Populärvetenskaplig Sammanfattning** 15
**Abbreviations** 17
**Introduction** 19
Transposable Elements 19
*Endogenous Retroviruses -ERVs* 20
*Intracisternal A Particles - IAPs* 22
*Long INterspersed repeated Elements – LINEs* 22
*SVA-elements* 23
Epigenetic Regulation of Transposable Elements 23
The role of TRIM28 in Regulation of Transposable Elements 25
Transposable Elements – Drivers of gene expression 25
Transposable Elements – Contribution to Genome Evolution 26
Transposable Elements in the Brain 27
Transposable Elements - Implications in other Diseases 28
Neural Progenitor Cells 29
**Aims** 31
**Results and Comments** 33
Establishment of a conditional TRIM28 Knockout in mouse Neural Progenitor Cells (Paper I) 33
TRIM28 deficiency in mouse Neural Progenitor Cells leads to Activation of Transposable Elements (Paper I) 35
TRIM28 deficient mouse Neural Progenitor Cells are able to differentiate into Neurons and Astrocytes (Paper I) 36
TRIM28 regulates *MMERVK10C* in mouse Neural Progenitor Cells (Paper I) 37
*MMERVK10C* is associated with H3K9me3 in mouse Neural Progenitor Cells (Paper I) 38
Analysis of *IAP1* Expression in mouse Neural Progenitor Cells (Paper I) 39
Activation of *MMERVK10C* and *IAP1* leads to Transcription of Nearby Genes (Paper I) 40
TRIM28 knockout-dependent Activation of Transposable Elements results in Expression of Long Non-coding RNAs (Paper I) 41
TRIM28 homozygous knockout is lethal (Paper I) 42
Mono-allelic TRIM28 Expression results in Behavioural Phenotype (Paper I) 43

Expression of Transposable Elements in human Neural Progenitor Cells (Paper II) 44
TRIM28 knockdown in human Neural Progenitor Cells leads to Upregulation of *Alu*-elements (Paper II) 46
TRIM28 controls Young *Alu*-elements in human Neural Progenitor Cells (Paper II) 47
TRIM28-controlled *Alu*-elements affect Expression of Nearby Genes (Paper II) 49
TRIM28 repression of Retrotransposon-based Enhancers is necessary to preserve Transcriptional Dynamics in Embryonic Stem Cells (Paper III) 49
**Concluding Remarks and Future Prospects** 51
**Materials and Methods** 53
Transgenic Mice 53
*Purification of total DNA from adult mouse-tail Biopsies* 53
*Purification of total DNA from embryonic mouse-tail Biopsies* 53
*Purification of total DNA from Cultured Cells* 54
*Polymerase Chain Reaction to determine the Nestin Cre TRIM28 Genotype* 54
*PCR-product Amplification via agarose-gel electrophoresis* 55
*Dissection of mouse embryos* 55
Cell Culture 56
*Culturing Neurospheres of dissected mouse embryonic forebrains* 56
*Generation of Neural Progenitor Cell Cultures* 57
*Differentiation of Neural Progenitor Cell Cultures* 58
*hNES Cell Cultures* 58
*TRIM28 Knockdown in human Neural Epithelial Stem Cells using Lenti Viral Vectors* 58
Chromatin Immuoprecipitation – ChIP 58
*Crosslinking* 58
*Cell lysis and Chromatin Shearing* 59
*Evaluation of Sonication* 59
*Immunoprecipitation* 59
*Analysis of Gene expression – RNA extraction, cDNA-synthesis and quantitative Real Time PCR* 60
Immunochemistry 61
*Immunocytochemistry* 61
*Immunohistochemistry on mouse embryos* 62
Behavioural Tests 62
*Open Field* 63
*Elevated Plus Maze* 64
**Acknowledgements** 65
**References** 69
**Paper I** 77
**Paper II** 91
**Paper III** 107

# Original Papers and Manuscripts

**Paper I**

TRIM28 represses transcription of endogenous retroviruses in neural progenitor cells.

**Fasching L.**, Kapopoulou A., Sachdeva R., Petri R., Jönsson M.E., Männe C., Turelli P., Jern P., Cammas F., Trono D. and Jakobsson J.

*Cell Reports 2015 Jan 6;10(1):20-8*

**Paper II**

TRIM28-controlled *Alu*-elements compose a gene regulatory network during human neuronal differentiation.

**Fasching L.**, Brattås P.L., Jönsson M.E., Pircs K., Shahsavani M., Pekkanen-Mattila M., Falk R., Falk A. and Jakobsson J.
*Manuscript 2015.*

**Paper III**

TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells.

Rowe H.M., Kapopoulou A., Corsinotti A., **Fasching L.**, Macfarlan T.S., Tarabay Y., Viville S., Jakobsson J., Pfaff S.L. and Trono D.

*Genome Research 2013 Mar;23(3):452-61*

# Abstract

More than 90% of DNA does not code for proteins and for a long time these sequences were referred to as "junk DNA" due to their unknown purpose. With the advent of new technologies it is now known, that the non-coding part of the genome is of great importance for regulating gene expression and is therefore indispensable.

Transposable elements comprise about 50% of the genome and co-exist as symbionts regulated by epigenetic mechanisms - a highly defined machinery that controls gene expression and is mandatory for a proper development and maintenance of an organism. Although transposable elements are associated with diseases, their role in fine-tuning the host gene expression becomes more and more evident, which seems to justify the positive selection during evolution.

A transposable element called *Line-1* was found to be active in neural progenitor cells and in the brain. Several studies report *Line-1* transcription and frequent retrotransposition during normal brain development, with further evidence that *Line-1* induced retrotransposition can influence neuronal gene expression. Today, there is few published data focusing on epigenetic regulation of transposable elements in neural progenitor cells.

In this thesis, I identify TRIM28 as key regulator of certain groups of transposable elements in mouse and human neural progenitor cells. This feature is unique compared to other somatic tissues, where DNA-methylation is prevalent.

Here I demonstrate, that transposable elements *MMERVK10C* and *IAP1* in mouse neural progenitor cells are repressed by the establishment of H3K9me3-associated heterochromatin. De-repressed *MMERVK10C* and *IAP1* furthermore activate nearby genes and generate long non-coding RNAs. Homozygous TRIM28 knockout is lethal, while mice with mono-allelic TRIM28 expression are characterised with a distinct behavioural phenotype.

Moreover we are also able to show that TRIM28 is regulating a fraction of young *Alu*-elements in human neural progenitor cells, which is not the case in human embryonic stem cells. Furthermore, we report that transcribed *Alu*-elements influence gene expression of close-by genes.

Studying pluripotent cells revealed that TRIM28 modulates transposable elements in mouse embryonic stem cells. Activation of transposable elements upon TRIM28 depletion induces changes in gene expression of close-by genes and causes alteration of the repressive chromatin mark H3K9me3 at transposable element loci. Upregulated genes were shown to have bivalent promoters, characterised by H3K4me3 and H3K27me3 and lay close to H3K9me3 regulated transposable elements. These findings in mouse embryonic stem cells are highly relevant for the interpretation of my studies in neural progenitor cells.

Taken together, this thesis demonstrates that the regulation of transposable elemenets in mouse and human neural progenitor cells is distinct compared to previous reports regarding somatic tissues. These results provide novel insights into why the brain has developed into such a complex organ with so many different cell types.

# Lay Summary

In an organism the DNA contains the entire biological information that is needed to be able to exist and function. Those segments, also known as DNA sequences are a genetic code, which is further packaged into units called genes. The entity of all genes is called genome. Today we know, that only a small proportion of genes encode for proteins, which are responsible for major biological functions in an organism, thus more than 90% of the DNA is referred to as non-coding DNA. For decades scientists were debating the purpose of the large non-coding proportion, and called it "junk DNA".

Currently we know that the non-coding part of the genome actually plays an important role regarding regulation of gene expression. Gene expression converts the information that is saved in DNA sequences into cellular components with a specific function. Since the entire genetic code is stored in each cell of an organism, gene expression is a highly regulated process. Not all genes can be active in all cells of the body at the same time. Therefore it has to be assured that only those genes, which are important for that specific cell type are activated. Different cell types are the basis for generating specific tissues, which are then further organised into organs. What happens to the majority of the genome that is non-coding DNA? Non-coding DNA has a distinct function in regulating gene expression. Processes that regulate gene expression are called epigenetic mechanisms. Those mechanisms can be seen like a light switch having two functions: switching "on" or "off". Genes that are needed for the cell to function are switched "on" while genes that are not needed are shut "off".

By winding DNA around histones, which is a certain type of proteins, DNA gets condensed and less accessible to be activated. DNA that is wound tightly around histones is called heterochromatin and keeps the DNA silenced and therefore inactive.

Since a few decades it is known, that about 50% of our genome are transposable elements, which are mobile genetic elements inherited over generations. Evolution is a continuous process characterised by optimal adaptation of an organism over millions of years to a changing environment. Transposable elements, if correctly regulated by epigenetic mechanisms, seem to have a positive effect on the host organism and are debated to drive evolution. If these transposable elements are not correctly regulated, they can cause many different diseases, like for example cancer. By now we know, that transposable elements can be active in the brain.

In my thesis, I investigate the regulation of transposable elements in mouse and human neural progenitor cells, which is a characterised cell type that is able to develop into several brain-specific cells. Therefore I have activated transposable elements in neural progenitor cells by removing their regulatory mechanism. I looked for resulting changes and found that these mobile elements are able

to switch "on" genes. I show that transposable elements seem to be important for the brain. The studies included in my thesis demonstrate that the regulation of transposable elements is different compared to what has been previously reported for other organs like heart or skin. These gained results provide novel insights into why the brain has developed into such a complex organ with so many different cell types.

# Populärvetenskaplig Sammanfattning

DNA innehåller hela den biologiska information som behövs för en levande organism att kunna existera och fungera. Denna information består av segment, även kända som DNA-sekvenser, som bildar en genetisk kod, som i sin tur bildar enheter som kallas gener. Helheten av alla gener kallas för arvsmassan. Numera vet vi att endast en liten andel av gener kodar för proteiner som ansvarar för grundläggande biologiska funktioner i en organism. Av denna anledning beskrivs mer än 90% av DNA som icke-kodande DNA. I årtionden har forskarna diskuterat syftet med den proportionerligt stora andelen av icke-kodande DNA och kallat den för "skräp-DNA".

Numera vet vi att den icke-kodande delen av genomet spelar en viktig roll när det gäller reglering av genuttryck. Genuttryck omvandlar den informationen som sparas i DNA-sekvenser i cellulära komponenter med specifika funktioner. Eftersom organismens hela genetiska kod finns lagrad i dess varje cell, är genuttryck en mycket strikt reglerad process. Alla gener kan inte vara aktiva i alla celler i kroppen samtidigt. Av denna anledning säkerställs att endast de gener som är viktiga för en specifik celltyp är aktiverade. Olika celltyper är grunden för att specifika vävnader skapas, vävnader som sedan grupperas i olika organtyper. Men vad är det som händer med den icke-kodande delen av DNA? Jo, icke-kodande DNA har en distinkt funktion i regleringen av genuttryck. De regulatoriska processerna i genuttryck kallas för epigenetiska mekanismer. Dessa mekanismer kan jämföras med en strömbrytare som har två funktioner - "på" eller "av". Gener som behövs för att cellen ska fungera slås "på" medan gener som inte behövs stängs "av". Genom att linda sig runt histoner, som är en viss typ av proteiner, kondenseras DNA och blir mindre tillgängligt för att aktiveras. DNA som är tätt packat runt histoner kallas för heterokromatin och håller det tystat och därför inaktivt.

Det har varit känt i ett par decennier nu, att cirka 50% av vår arvsmassa består av transposabla element som är rörliga genetiska element. Dessa element förs vidare från generation till generation. Evolution är en pågående process därigenom en organism adapteras optimalt till en föränderlig miljö under miljontals år. Transposabla element, i fall de regleras på rätt sätt genom epigenetiska mekanismer, tycks ha positiva effekter på värdorganismen och det debatteras i fall de för själva evolutionen framåt. Om dessa transposabla element inte regleras på ett korrekt sätt, kan de orsaka många olika sjukdomar, som till exempel cancer. Numera vet vi även att transposabla element kan vara aktiva i hjärnan. I min avhandling undersöker jag regleringen av transposabla element i mus och mänskliga neurala stamceller, som är celler som kan utvecklas till olika typer av hjärnspecifika celler. Jag aktiverade därför transposabla element i neurala stamceller genom att ta bort deras regleringsmekanism. Därefter letade jag efter förändringar och fann att dessa mobila element kan slå på gener. Jag påvisar att transposabla element verkar vara viktiga för hjärnan. De studier som ingår i min avhandling visar att regleringen av transposabla element är annorlunda i hjärnan jämfört med vad som tidigare rapporterats om andra organ, t.ex. hjärta eller hud. Dessa resultat ger nya insikter om varför hjärnan har utvecklats till ett sådant oerhört komplext organ med så många olika celltyper.

# Abbreviations

| | |
|---|---|
| ALS | amyotrophic lateral sclerosis |
| Ago3 | Argonaute 3 |
| Aub | Aubergine |
| bFGF | basic fibroblast growth factor |
| bp | base pairs |
| ChIP | chromatin immunoprecipitation |
| ChIP-seq | chromatin immunoprecipitation sequencing |
| Cre | Cre recombinase |
| CSF | cerebrospinal fluid |
| DG | dentate gyrus |
| DNA | deoxyribonucleic acid |
| DNMT 1 | DNA methyltransferase 1 |
| E 13.5 | embryonic day 13.5 |
| EGF | epidermal growth factor |
| EGFP | enhanced greenfFluorescent protein |
| EGFR | epidermal growth factor receptor |
| env | gene encoding envelope protein |
| EPM | elevated plus maze |
| ERV | endogenous retrovirus |
| fl | flox |
| gag | gene encoding for group antigens |
| H3K27ac | histone 3-lysine 27-acetylation |
| H3K27me3 | histone 3-lysine 27-tri-methylation |
| H3K4me1 | histone 3-lysine 4-mono-methylation |
| H3K4me3 | histone 3-lysine 4-tri-methylation |
| H3K9me3 | histone 3-lysine 9-tri-methylation |
| hERV | human endogenous retrovirus |
| hES cells | human embryonic stem cells |
| hNES cells | human neuroepithelial-like stem cells |
| hNP cells | human neural progenitor cells |
| HP1 | heterochromatin protein 1 |
| IAP | intracisternal A particles |
| IAP1 | intracisternal A-particles class 1 |
| Kap1 | KRAB-associated protein-1 |
| kb | kilo base |
| KD | knockdown |

| | |
|---|---|
| KO | knockout |
| KRAB-ZFP | Kruppel-associated box zinc finger protein |
| lincRNA | long non-coding RNA |
| LINE | long interspersed repeated element |
| Line-1 | long interspersed repeated element-1 |
| LTR | long terminal repeat |
| MeCP2 | methyl-CpG binding protein 2 |
| mES cells | mouse embryonic stem cells |
| Mili | Piwi-like protein 2 - mus musculus |
| Miwi2 | Piwi-like protein 4 - mus musculus |
| MMERVK10C | mus musculus ERV using $tRNA^{Lys}$ type 10C |
| mNP cells | mouse neural progenitor cells |
| mRNA | messenger RNA |
| OF | open field |
| ORF | open reading frame |
| PHD | plant homeo domain |
| pi-RNA | Piwi-interacting RNA |
| pol | gene encoding revere transcriptase |
| POL | polymerase |
| RC-seq | Retrotransposon capture sequencing |
| RING finger | Really Interesting New Gene finger |
| RNA-seq | RNA-sequencing |
| RT | reverse transcriptase |
| RV | retrovirus |
| SETDB1 | set-domain protein 1 |
| SINE | short interspersed repeated element |
| SOX1 | sex determining region Y-box 1 |
| SOX2 | sex determining region Y-box 2 |
| SUMO | small ubiquitin-like modifier |
| SVA-family | Sines, VNRT and Alu-elements |
| SVZ | subventricular zone |
| t-RNA | transfer RNA |
| TDB-43 | TAR DNA-binding protein 43 |
| TE | transposable element |
| TF | transcription factor |
| Tif1β | transcriptional intermediary factor 1β |
| TRIM28 | tripartite motif-containing 28 |
| UTR | untranslated region |
| WT | wild type |

# Introduction

## Transposable Elements

Transposable elements (TEs) are mobile genetic elements, which comprise almost 50% of the genome (Baillie et al., 2011; Bannert and Kurth, 2004; Castro-Diaz et al., 2014; Cordaux and Batzer, 2009; Cowley and Oakey, 2013; Feschotte, 2008; Hancks and Kazazian, 2012; Hua-Van et al., 2011; Mills et al., 2007; Muotri et al., 2007; Pi et al., 2010; Reilly et al., 2013; Rowe and Trono, 2011; Sundaram et al., 2014; Xing et al., 2009). Barbara McClintock pioneered the field of TEs already in the mid-20$^{th}$ century, although initially criticised for her hypothesis, she was finally awarded with the Nobel Prize in 1983 for her discovery of TEs (Bannert and Kurth, 2004; Cowley and Oakey, 2013; Hua-Van et al., 2011; Reilly et al., 2013). For a long time, the non-coding part of the genome was referred to as "junk" because those parts are known to be non-protein-coding regions (Hua-Van et al., 2011; Muotri et al., 2007; Pi et al., 2010; Reilly et al., 2013).

Today we know, that the life of a cell depends on the interplay of genetics, epigenetics and environment, which are all essential factors for gene regulation (Hua-Van et al., 2011; Muotri et al., 2007). The previous assumptions that TEs are only parasitic genome invaders seem to be out of date. Nowadays we rather pursue the question of benefits for landscaping the host genome due to the positive selection during evolution (Hua-Van et al., 2011; Reilly et al., 2013).

Recent research has recognised the impact of TEs on host regulatory networks and therefore acknowledged their role in shaping the genome (Feschotte, 2008; Hua-Van et al., 2011; Kunarso et al., 2010; Muotri et al., 2007). Mobility of these elements can result in nucleotide changes and chromosome rearrangements, which are furthermore passed on to following generations (Hua-Van et al., 2011; Muotri et al., 2007). Certain copies of TEs can be very specific to individuals, while several families of TEs are unique for certain species (Hua-Van et al., 2011). Interestingly, no alleles are found to be spared of integration (Bannert and Kurth, 2004). Transposons are characterised by the ability to move within the same genome, without infecting other cells (Hua-Van et al., 2011). They are divided into two major classes: Class I and Class II transposons (Feschotte, 2008). Class I transposons have the ability to transpose via "copy-and-paste" mechanism. Therefore DNA is transcribed into an RNA intermediate and later reverse transcribed by the help of *reverse transcriptase (RT)*. The eukaryotic genome is mainly composed of few *Long Terminal Repeat (LTR)* transposons and a large group of non-*LTR* transposons consisting of *Long INterspersed repeated Elements (LINEs), Short INterspersed repeated Elements (SINEs)* and *Alu*-elements (Bannert and Kurth, 2004; Cordaux and Batzer, 2009; Hancks and Kazazian, 2012; Hua-Van et al., 2011; Kannan et al., 2015; Reilly et al., 2013). *SINEs* and *LINEs* are the most abundant and active elements in the human genome (Cordaux and Batzer, 2009) and characterised by a poly-A tail at their 3´end (Dewannieux and Heidmann, 2013; Hancks and Kazazian,

2012; Muotri et al., 2007). Activated *SINEs* and *LINEs* are associated with diseases (Bannert and Kurth, 2004).

The major difference between *LTR*-transposons (or also called *Endogenous RetroViruses – ERVs*) to *SINEs* and *LINEs* is the existence of the *LTR* (Hua-Van et al., 2011). *LINEs* are autonomous transposons, while *SINEs* are non-autonomous and need *LINEs* for their retrotransposition (Kannan et al., 2015; Muotri et al., 2007; Reilly et al., 2013). Class II transposons are DNA-transposons (Cordaux and Batzer, 2009; Feschotte and Pritham, 2007) and transpose to another genomic location via "cut and paste" mechanism; avoiding the RNA intermediate (Bannert and Kurth, 2004; Dewannieux and Heidmann, 2013; Hua-Van et al., 2011; Kannan et al., 2015; Reilly et al., 2013).

This introduction focuses only on Class I transposons, because DNA-transposons are known to be inactive or extinct since approximately 37 million years (Cordaux and Batzer, 2009; Reilly et al., 2013). In general, eukaryotic genomes seem to have a higher copy number of TEs when compared to prokaryotic genomes (Hua-Van et al., 2011). Taking a closer look into the human genome, the majority of TEs are comprised of *LINEs* and *SVA*-elements (*SINEs* and *Alu*-elements), while *ERVs* appear in a lower copy number. The major impact of TEs is caused by their mobility within the genome, because they can "jump" close to or even insert themselves into genes, which can drastically influence gene expression (Hua-Van et al., 2011). The consequences reach from total inactivation of the "invaded" gene up to alternative splicing events (Hua-Van et al., 2011). Alternative splicing events eventually lead to genetic variations and enhance the mosaicism of gene expression (Hua-Van et al., 2011). TEs in the genome are found to be highly abundant in heterochromatin, especially in centromers and telomers (Dewannieux and Heidmann, 2013; Hua-Van et al., 2011).

*Endogenous Retroviruses - ERVs*

*ERVs* comprise about 7-8 % of the human and mouse genome and are Class I transposons that were discovered in the early 1970s (Ryan, 2004). The origin of *ERVs* is not resolved yet. There are different theories regarding their emergence. The most pursued ideas are: either *ERVs* descent from retroviral infections of the germ line or retrotransposons in the host cell developed a mechanism to escape from cells (Bannert and Kurth, 2004; Dewannieux and Heidmann, 2013; Rowe et al., 2013; Rowe and Trono, 2011). Through vertical transmission the infection gets passed on to the following generations and the virus becomes hereditable. When looking at the phylogenetic tree of different *ERV* families it is evident that each family descends from a single retroviral infection (Dewannieux and Heidmann, 2013). Only *retroviruses* (*RVs*) with simple genomes became endogenous with the exception of *spumavirus* (Weiss, 2006). Simple *RVs* are classified into *alpha*-, *beta*-, *gamma*-, *delta*- and *epsilon*-*RVs* (Dewannieux and Heidmann, 2013; Weiss, 2006). So far no *ERVs* were found to be associated with *delta*-retroviral infection (Dewannieux and Heidmann, 2013).

The *ERV* provirus is characterised by open reading frames (ORFs) flanked by one *LTR* on each side (Dewannieux and Heidmann, 2013). These *LTRs* vary between 100 bp to 5000 bp in size (Kannan et al., 2015) and contain regulatory elements like promoters, enhancers, silencers and poly-A signals (Ryan, 2004). The *ERV*-ORFs can either correspond to retroviral ORFs containing genes that encode for the three major proteins: *gag*, *pol* and *env*, or in some cases *ERVs* have lost the viral env-gene due to homologous recombination (Dewannieux and Heidmann, 2013). Some *ERVs* possess

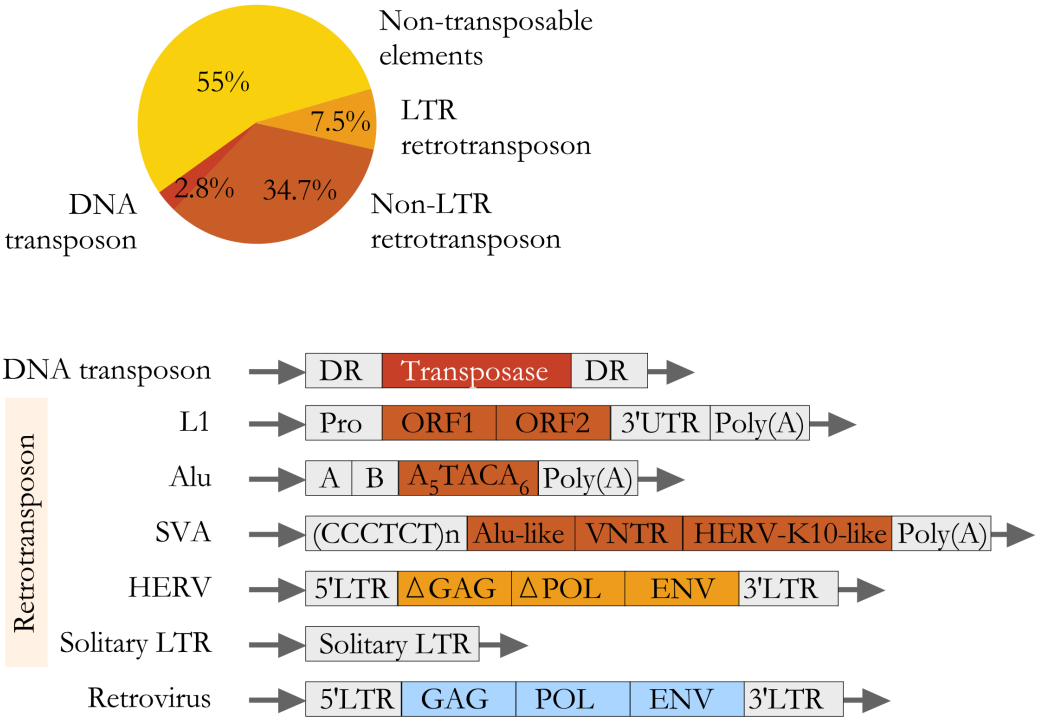



**Figure 1**. *Transposable elements*

ORFs that are malfunctioning due to frame-shifts or premature stop-codons (Bannert and Kurth, 2004; Dewannieux and Heidmann, 2013) and are therefore inactive (Cordaux and Batzer, 2009).

Currently more than 200 *LTR*-retrotransposons are described in Repbase, a database of consensus sequences of repetitive elements (Bannert and Kurth, 2004; Bao et al., 2015). The nomenclature of h*ERVs* is based on the amino acid of the t-RNA that hybridises to the primer-binding site (Bannert and Kurth, 2004).

*Intracisternal A Particles - IAPs*

*Intracisternal A Particles* (*IAPs*) are a rodent-specific group of Class I *LTR*-retrotransposons, which are related to *type D simian retroviruses, type B mouse mammary tumour virus* and *type C avian sarcoma virus. IAPs* are randomly distributed over all chromosomes (Qin et al., 2010). This is evident through thousands of copies with a few hundred characterised as being autonomous (Ribet et al., 2008). Particularly in the mouse genome, these elements appear to be highly polymorphic (Qin et al., 2010). *IAPs* are associated with most mutagenic insertions in mouse characterised by promoter activity (Huang et al., 2012; Qin et al., 2010). The 7 kb long *IAP* provirus contains a protein-coding region (with functional *gag, pro* and *pol* genes (Ribet et al., 2008)) surrounded by *LTRs* and are divided into 4 family members: *IAP1-4*. More than 60% of the *LTR* insertions are found in anti-sense orientation. *IAP1* and *IAP1a*, which are classified as the youngest and most active members, are able to generate viral proteins (Qin et al., 2010). Actively transcribed *IAP* sequences are only found in pre-implantation embryos and are inactive in somatic tissues where they are hyper-methylated (Hutnick et al., 2010). *IAP*-elements in general are described as rodent *ERVs* that have "lost" their *env*-gene. However a single active *env*-gene harbouring *IAPE-D* provirus with infectious properties was characterised. This finding leads to the theory that a single retroviral progenitor became endogenous by germ line infection (Ribet et al., 2008).

*Long INterspersed repeated Elements - LINEs*

About 20% of the mammalian genome consists of *Line-1* sequences (Boeke, 1997; Garcia-Perez et al., 2007; Hancks and Kazazian, 2012; Kano et al., 2009; Muotri et al., 2007; Reilly et al., 2013) and 3.2% of *Line-2* elements (Bannert and Kurth, 2004). *Line-1* is known to be autonomous and can therefore independently transpose within the genome (Cordaux and Batzer, 2009; Muotri et al., 2007; Reilly et al., 2013). *Line-1* elements, that are transcribed via polymerase II (Pol II) are approximately 6-7 kb long (Reilly et al., 2013). Their elements contain a 5´-untranslated region (UTR) with internal promoter, two ORFs and a 3´-UTR characterised by a poly-A tail (Boeke, 1997; Cordaux and Batzer, 2009; Hancks and Kazazian, 2012; Kannan et al., 2015). ORF1 encodes an RNA-binding protein and ORF2 encodes for a protein with *reverse-transcriptase* and *endonuclease* activities (Boeke, 1997; Cordaux and Batzer, 2009; Feng et al., 1996; Kannan et al., 2015; Mills et al., 2007). The stop-codons are mostly found within those ORFs (Mills et al., 2007) Both ORFs are necessary for retrotransposition (Hancks and Kazazian, 2012). *LINEs* consisting of only a solo-ORF2 are debated to be potential drivers for retrotransposition of *Alu*-elements (Mills et al., 2007) Transcribed *Line-1* sequences are stable for more than 24 hours and can be transmitted to the next generation (Kuramochi-Miyagawa et al., 2008). *Line-1* elements become inactive due to mutations, truncations and internal rearrangements (Cordaux and Batzer, 2009). On average there are 80-100 transposable competent *Line-1* elements in every individual (Brouha et al., 2003). Transposition mainly occurs during embryogenesis, but has also been detected at a lower level in male and female germ cells (Kano et al., 2009; Kuramochi-Miyagawa et al., 2008) and in the brain (Baillie et al., 2011; Muotri et al., 2005; Muotri et al., 2010; Upton et al., 2015).

*SVA-elements*

The *SVA*-family of retrotransposons consists of *SINEs*, *VNRT* and *Alu*-elements (Hancks and Kazazian, 2010; Wang et al., 2005b). These 2 kb long elements (Cordaux and Batzer, 2009) are known to be the most active elements in human (Baillie et al., 2011). Around 3000 copies of *SVA*-elements are found in the human genome (Wang et al., 2005b; Xing et al., 2006).

*SINEs* are transcribed by Pol III, do not contain any ORFs and are therefore non-autonomous transposons (Cordaux and Batzer, 2009; Kannan et al., 2015). These elements are completely dependent on *RT* derived from other transposable elements e.g. *LINEs*, which act *in trans* (Bannert and Kurth, 2004; Hancks and Kazazian, 2012; Kannan et al., 2015; Mills et al., 2007). The most abundant groups of *SVAs* in human are members of the *Alu*-family (Muotri et al., 2007; Wang et al., 2005b). The 300 bp long *Alu*-elements are represented by more than 200 subfamilies in approximately 1 million copies per genome. *Alu*-elements are active since about 65 million years (Cordaux and Batzer, 2009). Today, 22 *AluY* and six *AluS* subfamilies were identified as most active among *Alu*-elements in humans. *AluY* members can also stand minor changes in their sequence and still retain their ability to transpose (Mills et al., 2007)

## Epigenetic Regulation of Transposable Elements

Since some TE are retrotransposition competent (Bannert and Kurth, 2004) and thereby alter gene expression (Douville et al., 2011; Herquel et al., 2013), it is essential that these mobile elements are highly regulated.

The Miwi2 protein is involved in silencing *Line-1* and *IAP*-elements in the male germ line. Miwi2 deficiency was shown to result in loss of DNA-methylation at *Line-1* loci (Carmell et al., 2007; Kuramochi-Miyagawa et al., 2008). Further studies indicate that interaction of Mili and Miwi2 proteins with pi-RNAs, are essential for de-novo DNA-methylation of retrotransposons in male fetal germ cells (Aravin et al., 2007; Hancks and Kazazian, 2012; Kuramochi-Miyagawa et al., 2008). *IAP* expression is not only repressed through DNA- and histone-methylation, but also through small RNAs (Qin et al., 2010). In mouse embryonic stem (mES) cells *IAP* regulation is not dependent on DNA-methylation, which was confirmed by DNA methyltransferase 1 (DNMT1) deletion. Interestingly, differentiation of DNMT1 null mES cells leads to enrichment of *IAP* mRNA and proteins (Hutnick et al., 2010). These data suggest that *IAP*-elements in mES cells are regulated in an alternative manner (Hutnick et al., 2010), which was later described to be via TRIM28 dependent histone (H3) lysine 9 (K9) tri-methylation (H3K9me3) (Rowe et al., 2010). In most human tissues (ES cells and adult tissues) *SVA* promoters are hypermethylated with the exception of germ line cells (Hancks and Kazazian, 2012). *SINEs*, which are transcribed via their own Pol III are primarily suppressed by H3K9me3 (Varshney et al., 2015).

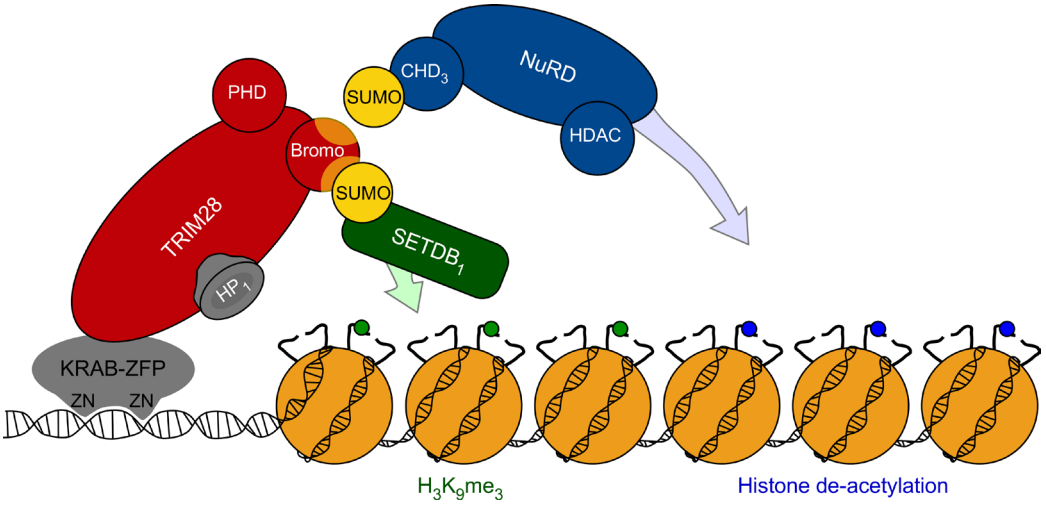



**Figure 2**. *TRIM28-mediated regulation of transposable elements.*

## The role of TRIM28 in Regulation of Transposable Elements

TRIM28 (Kap1 or TIF1β) is known as epigenetic co-repressor of transcription (Cammas et al., 2000; Friedman et al., 1996; Sripathy et al., 2006; Wiznerowicz et al., 2007). Kruppel-associated box zinc finger proteins (KRAB-ZFPs) are transcriptional repressors that interact with the primer-binding site (PBS) of the TE and initiate epigenetic repression in ES cells by recruitment of TRIM28 to the locus of interest (Ellis et al., 2007; Urrutia, 2003).

TRIM28 was described to consist of an alanine-rich amino-terminus, a RING finger, B1 and B2 boxes of conserved cysteine and histidine spacing, a leucine-zipper and an α-helical coiled-coil structure. The carboxyl-terminus is characterised by Plant Homeo Domain (PHD)-finger and a bromo domain, which are involved in chromatin-mediated gene regulation (Cammas et al., 2000; Friedman et al., 1996; Peng et al., 2000; Wang et al., 2005a; Wiznerowicz et al., 2007). The bromo domain plays an important role in chromatin targeting by interacting with lysine-acetylated peptides from the histones H3 and H4 (Cammas et al., 2000; Friedman et al., 1996; Sripathy et al., 2006). Apart from this, the PHD-finger possesses an intermolecular E3 ligase activity, which triggers SUMOylation on several sites, especially nearby the bromo domain known as K779 and K804 (Ivanov et al., 2007) and thus

recruits the set-domain protein 1 (SETDB1) – a histone methyl-transferase (Schultz et al., 2002). SETDB1 methylates lysine 9 of histone 3 and induces H3K9me3-mediated repression (Bilodeau et al., 2009). SETDB1 is essential for the recruitment of heterochromatin protein 1 (HP1) (Matsui et al., 2010). Proper HP1/TRIM28 interaction is absolutely necessary for transcriptional repression (Schultz et al., 2002; Sripathy et al., 2006).

Only during embryonic development does the interaction of KRAB-ZFP/TRIM28 induce epigenetic repression by the establishment of cytosine methylation (Quenneville et al., 2012). Several studies highlight the importance of TRIM28 regarding repression of TEs in mouse and human ES cells, as well as during early embryonic development (Rowe et al., 2013; Rowe et al., 2010; Turelli et al., 2014). Deletion of SETDB1 and TRIM28 leads to activation of transposable elements, which is in contrast to mouse embryonic fibroblasts (MEFs) (Matsui et al., 2010). Recently, it was proposed that ZFP809 plays an important role for the silencing complex during the establishment of heterochromatin in mES cells at *ERV* loci (Wolf et al., 2015). These results further indicate that DNA-methylation does not regulate TEs in mES cells, which was already previously described (Hutnick et al., 2010). TRIM28/SETDB1-associated H3K9me3 is silencing transposable elements in mES cells, especially during de-novo DNA methylation in early embryonic development (Karimi et al., 2011; Matsui et al., 2010). Recently, a study showed that the regulation of *Line-1* elements in mouse and human ES cells varies among their subtypes, with older *Line-1* members regulated by KRAB-ZFP/TRIM28, while silencing of the newer emerged members was clearly DNA-methylation dependent (Castro-Diaz et al., 2014). In another study it was presented that the regulatory protein SIRT6 ribosylates TRIM28 and together they establish heterochromatin at *Line-1* loci. Loss of SIRT6 at those loci is associated with aging (Van Meter et al., 2014).

## Transposable Elements – Drivers of gene expression

About 25% of the human promoters comprise of transposon-derived sequences (Bannert and Kurth, 2004). TEs in ES cells play a role during development and are beneficial for the host are dynamically regulated *e.g.* through recruitment of transcription factors (TFs) or the production of non-coding RNAs. Transposons that are associated with diseases are kept silent (Robbez-Masson and Rowe, 2015). Very recently it was shown that h*ERV-K* is not only actively transcribed during embryonic development but also produces *gag*-proteins and virus-like particles (Grow et al., 2015). The link between TEs and transcription factors was recently reported by finding transcription factor-binding sides at transposons. Chromatin immunoprecipitation-sequencing (ChIP-seq) data reveals that TF-binding peaks are concomitant with "active" chromatin marks and a reduction in DNA-methylation (Sundaram et al., 2014). h*ERVs* are also shown to have distinct roles during gene expression: h*ERV-LTRs* possess promoter and enhancer traits that can influence expression of close-by genes (Douville et al., 2011; Herquel et al., 2013). Activated h*ERVs* were also described during early development. h*ERV-H* is expressed in hES cells and has a role in maintaining pluripotency by interacting with cell-specific transcription factors (Lu et al., 2014). h*ERV-R* for example was found to be actively expressed in fetal tissue (Ryan, 2004). Recently it was shown that TEs have an enhancer activity in the rodent placenta and are therefore important contributors to the gene regulatory network (Chuong et al., 2013)

*Alu*-elements can render the host transcriptome by altering their own methylation state (Ryan, 2004). In hES cells, *Line-1* retrotransposition into genes is associated with genomic DNA deletions (Garcia-Perez et al., 2007). Moreover, human *LINEs* are shown to generate processed pseudogenes but only if both ORFs are present (Esnault et al., 2000). Even from a distance, activated TEs not only influence transcriptional termination in *cis* (Li et al., 2012a) but also reduce elongation and furthermore induce changes in epigenetic marks of introns (Isbel and Whitelaw, 2012).

Besides exon-shuffling, retrotransposon-mediated 3´ transduction also leads to gene duplications and generation of novel gene families with *SVA* derived promoters, with a huge impact on the host genome (Wang et al., 2005b; Xing et al., 2006).

*Alu*- and *B*-elements, which are both members of the *SVA*-family, are significantly enriched in upstream and introgenic loci of genes with known functions. These numerous findings imply that TEs seem to contribute to host gene regulation as a consequence of positive selection during evolution (Tsirigos and Rigoutsos, 2009).

## Transposable Elements – Contribution to Genome Evolution

From an evolutionary perspective transposable elements are of great interest, since they comprise such a large part of the genome and have furthermore the potential to actively contribute by adding new coding sequences or regulatory units (Feschotte, 2008; Kannan et al., 2015; Reilly et al., 2013). *Line-1* and *SVA*-members independently evolved from single lineages (Cordaux and Batzer, 2009). Therefore, the high copy number of those non-*LTR* transposons as well as the fact of being active over millions of years gives a hint on their success of replication during evolution (Cordaux and Batzer, 2009). Following TE-associated events are known to be involved in genome shaping during evolution: insertional mutagenesis; creating and repairing of DNA double-strand breaks; micro satellite seeding; gene conversion; insertion-mediated deletions; ectopic recombination and transduction (Cordaux and Batzer, 2009).

Younger members of *Alu* and *Line-1* are still found to actively retrotranspose in humans (Bannert and Kurth, 2004; Mills et al., 2007; Ryan, 2004; Wang et al., 2005b). The large proportion of transposable elements which was found in the genome obviously has an impact in shaping the genome by establishing genetic variation (Bannert and Kurth, 2004). Retrotransposition events occurring upstream of a coding region can result in exon-shuffling (Cowley and Oakey, 2013; Hancks and Kazazian, 2012).
Insertions of TEs were also found in 69% of human and 51% of mouse long non-coding RNAs (lincRNAs) and occur more frequently in introns as compared to exons, with most insertions occurring in promoter regions of lincRNAs. These results indicate that TEs play a role in evolutionary development of lincRNAs (Kannan et al., 2015).

*Line-1* activation in general is associated with insertions, deletions, generation of novel splice sites, and fine-tuning of gene expression (Muotri et al., 2010). *Line-1* insertions predominantly occur in adult brains, but are found to be absent in other somatic tissues (Muotri et al., 2010) with exceptions in various cancers (Carreira et al., 2014; Helman et al., 2014; Lee et al., 2012; Solyom et al., 2012; Tubio et al., 2014). *Line-1* is therefore proposed to be a major contributor to genetic mosaicism in the adult brain (Kuramochi-Miyagawa et al., 2008; Muotri et al., 2010) but also in ES cells (Garcia-Perez et al., 2007) and during early stages of development (Garcia-Perez et al., 2007; Reilly et al., 2013). Several studies report *Line-1* transcription and retrotransposition in neural progenitor cells (Coufal et al., 2009; Muotri et al., 2005; Muotri et al., 2010), with further evidence that *Line-1* induced retrotransposition can influence neuronal gene expression (Muotri et al., 2005). Frequent retrotransposition events were reported during normal brain development (Li et al., 2012b). Designing a human *Line-1* element reveals retrotransposition events in adult rat neural progenitor (NP) cells *in vitro* as well as *in vivo* transposition in mouse brains (Muotri et al., 2005). NP cells, where active *Line-1* retrotransposition was monitored, are still capable of differentiation but preferably into neurons. Interestingly, decreased SOX2 expression upon differentiation seems to correlate with increased *Line-1* transposition. These findings propose SOX2 as regulator of *Line-1* associated retrotransposition in hippocampal neural progenitor cells (Muotri et al., 2005). Furthermore retrotransposition can be triggered by environmental stimuli. This was shown by an engineered *Line-1* EGFP reporter, which detected enhanced retrotransposition in hippocampal neurons *in vivo* of mice upon voluntarily performed exercise (Muotri et al., 2009). Similar results were observed in human NP cells, where a *Line-1* reporter assay showed 19 retrotransposition events in progenitors but not primary fibroblasts or astrocytes. Moreover significantly more *Line-1* copies were detected in cells derived from the hippocampus compared to heart or liver (Coufal et al., 2009). Retrotransposon capture sequencing (RC-seq) leads to further evidence of somatic *Line-1* retrotransposition in human post mortem hippocampus and caudate nucleus. The vast majority of insertions (more than 80%) are strongly associated with younger *Line-1* members (Baillie et al., 2011). A recent study proposes improved RC-seq (Baillie et al., 2011) and reveals 13.7 *Line-1* insertions in hippocampal neurons (Upton et al., 2015), which is in strong contrast to an earlier publication that reported transposition at a very low frequency (less than 0.6 insertions per neuron) in human cerebral cortex and caudate nucleus (Evrony et al., 2012). In contrast to the retrotransposition frequency in neurons, only 6.5 Line-1 insertions were found in glia. Genes that were found to be upregulated in the hippocampus were prevalent for *Line-1* insertions in neurons and glia. Somatic inserations occur in both, neurons and glia and have a strong preference for protein-coding regions (Upton et al., 2015).

The methyl-CpG binding protein 2 (MeCP2) is involved in the establishment of DNA-methylation. A study in mouse shows, that neurons lacking MeCP2 demonstrate increased *Line-1* retrotranspostion (Muotri et al., 2010). Disturbance of DNA methylation in neural progenitor cells leads to *Line-1* activation (about 80 new insertions) and is associated with neurodevelopmental diseases like the RETT syndrome (Hancks and Kazazian, 2012; Muotri et al., 2010). TAR DNA-binding protein 43 (TDB-43) was described as regulator of TEs in rodents and humans. TDB-43 dysfunction is associated with upregulation of TEs and furthermore results in neurodegeneration (Li et al., 2012b). Interestingly in mice, TRIM28 deletion in the adult hippocampus causes significant anxious-like behaviour and elevated movement and a shift from repressive chromatin (H3K9me3) to "open"

chromatin marks (histone-3 and –4 acetylation) (Jakobsson et al., 2008). Another study shows that in the hippocampus, environmental stress leads to an increase of H3K9me3 at loci of transposable elements as a response to acute stress (Hunter et al., 2012). Single cell sequencing of human neuronal nuclei reveals, that *Line-1* retrotransposition rarely occurs in the cerebral cortex and caudate nucleus in human brains, but may play a higher role in other brain areas *e.g.* the hippocampus with a potential role in neuronal diversity (Evrony et al., 2012). Mutations that cause neurofibromatosis type 1 are associated with 18 insertions mediated by retrotransposition events (Hancks and Kazazian, 2012). h*ERVs* have many implementations in neurological disorders, h*ERV-H* for example is associated with multiple sclerosis (Ryan, 2004). h*ERV-K* which belongs to the younger *ERV*-members (Grow et al., 2015), is highly expressed in neurons of the frontal lobe of patients suffering from amyotrophic lateral sclerosis (ALS). *RT* expression was not only found in serum of patients but also in serum of first grade relatives. These observations indicate the possibility of inheriting an activated h*ERV*, and furthermore propose h*ERV-K* as suitable marker for ALS (Douville et al., 2011). h*ERV-W* expression was found in the cerebrospinal fluid (CSF) of recent-onset but also chronic schizophrenia patients as well as in post mortem samples of frontal cortex (Karlsson et al., 2001) Besides *exogenous viruses, endogenous retroviruses* also show implications in the development of prion diseases: RNA transcripts of several h*ERV* members are enriched in the CSF of patients diagnosed with sporadic Creutzfeldt-Jakob disease; especially h*ERV-W*, h*ERV-K* and h*ERV-T* (Jeong et al., 2010)

Furthermore retrotransposition competent elements were described in *Drosophila melanogaster* (Perrat et al., 2013; Waddell et al., 2014). The *Drosophila* genome harbours around 1500 mobile elements categorised into approximately 100 transposon classes. TEs in the fruit fly show a lot of similarities to transposons found in vertebrates. A very well characterised *Drosophila LTR*-transposon for example is *gypsy*, which was described to possess three ORFs (Perrat et al., 2013; Waddell et al., 2014). Active and therefore mobile *LTR*- and non-*LTR*-like transposons are found in the brains of fruit flies, which is similar to *Line-1* retrotransposition in rodents and human. Silencing of transpositions in the *Drosophila* brain is associated with the establishment of H3K9me3. Although there is no evidence for adult neurogenesis in the fly brain, deep sequencing suggests that active retrotransposition events contribute to somatic mosaicism specifically in the population of αβ neurons in mushroom bodies (MB) of *Drosophila* brains (Waddell et al., 2014). Furthermore pi-RNA mediated repression was found to be decreased in αβ neurons which was concomitant with low levels of RNA-binding proteins Argonaute 3 (Ago3) and Aubergine (Aub) in αβ neurons and correlates with increased retrotransposition (Perrat et al., 2013; Waddell et al., 2014).

## Transposable Elements Implications in other Diseases

About 10% of spontaneous mutations are caused by transposable elements (Matsui et al., 2010). *LINEs* for example are associated with mutagenesis in humans (Esnault et al., 2000). Already in the late 1980s it was shown that *Line-1* insertions have the potential to cause haemophilia A (Kazazian et al., 1988) *SVA*-elements are highly active in the human genome and the consequence of their insertions are implicated in diseases (Mills et al., 2007; Wang et al., 2005b) About 0.4% of genetic diseases are associated with *Alu*-elements that cause mutations, which influence gene expression and regulation as well as ectopic recombination. *Alu*-associated diseases include: for example neurofibromatosis, breast cancer and type 2 diabetes (Ryan, 2004).

h*ERVs* on the other hand are associated with several types of cancer as well as autoimmune disorders (Ryan, 2004). For example h*ERV-K* derived viral particles are found in cancer cells (Belshaw et al., 2005) *IAP* insertions in viable yellow agouti mice not only result in a change of coat colour (Cowley and Oakey, 2013; Isbel and Whitelaw, 2012) but these mice are furthermore prone to cancer and obesity (Qin et al., 2010). Although there was a strong belief that retrotransposition mainly happens in germ cells, it became more and more evident that most insertions happen in somatic cells and therefore play a major role in tumorigenesis (Hancks and Kazazian, 2012).

## Neural Progenitor Cells

Neural progenitor (NP) cells are multipotent cells of the central nervous system that can self-renew, proliferate (Graham et al., 2003) and furthermore give rise to neurons, astrocytes and oligodendrocytes (Carpenter et al., 1999; Taupin and Gage, 2002; Torrado et al., 2014). The progenitor defining properties are regulated by the interplay of Notch signaling and the EGFR growth factor (Aguirre et al., 2010; Graham et al., 2003). For long time it was thought that these cells are restricted to embryonic development but now we know that neurogenesis occurs in two specific locations of the adult brain: the subventricular zone (SVZ) of the lateral ventricle and the dentate gyrus (DG) of the hippocampus (Aguirre et al., 2010). These findings were confirmed by bromodeoxyuridine labeling of dividing cells (Gage, 2000; Taupin and Gage, 2002). Nestin (Taupin and Gage, 2002), SOX1 and SOX2 (Graham et al., 2003) are validated markers of neural progenitors. NP cells can be isolated from embryonic and adult brain and furthermore cultured for many different *in vitro* studies (Torrado et al., 2014). These cells can be expanded as neurospheres *in vitro* (Ahlenius and Kokaia, 2010; Torrado et al., 2014). Considering their properties of proliferation, NP cells are suitable models for studying cellular processes *in vitro* and qualify for studies of transposable elements.

# Aims

The overall goal of my thesis was to study transposable elements in neural progenitor cells. In particular I was interested in what mechanisms are essential for their regulation and what impact their de-regulation might have on a genome-wide scale. Here, I would like to list the major aims of this thesis on which we based our studies:

I. To show that TRIM28 plays a major role in regulating transposable elements in mouse neural progenitor cells by the establishment of the heterochromatin mark H3K9me3, which is unique for somatic tissues. Furthermore, I wanted to investigate the impact of transposon activation on the whole transcriptome.

II. To investigate if our findings regarding TRIM28 dependent regulation of transposable elements in mouse neural progenitor cells are translatable to human neural progenitor cells. Additionally it was of great interest to study the fraction of transposable elements, which is controlled by TRIM28 mediated repression and furthermore to investigate the effect of their activation.

# Results and Comments

The transcriptional co-repressor TRIM28 is known to regulate transposable elements in embryonic stem cells (Rowe et al., 2013; Rowe et al., 2010; Turelli et al., 2014) but not in somatic tissues (Matsui et al., 2010). Several studies show that certain families of transposable elements are highly active in mouse and human brain (Coufal et al., 2009; Kuramochi-Miyagawa et al., 2008; Muotri et al., 2005; Muotri et al., 2010).

In this thesis, I demonstrate that regulation of transposons in mouse and human neural progenitor cells is distinct compared to reported studies regarding other somatic tissues (Matsui et al., 2010). In this section, I am summarising the key findings of Paper I and II and I will relate those to findings from Paper III, a study about TRIM28 as regulator of transposable elements in mouse embryonic stem cells.

## Establishment of a conditional TRIM28 Knockout in mouse Neural Progenitor Cells (Paper I)

In the initial step of this study, I have set up a mouse breeding of male mice expressing Cre recombinase under a Nestin promoter with females carrying homozygous floxed TRIM28 alleles as described previously (Tronche et al., 1999; Weber et al., 2002). Embryos were collected at E 13.5 and mouse neural progenitor (mNP) cell cultures of embryonic forebrain were established (Figure 3A). The knockout was verified by genotyping all collected embryos (Figure 3B) as well as by immunocytochemistry demonstrating the lack of TRIM28 protein (Figure 3C). Furthermore, I wanted to investigate whether TRIM28 deficiency in the embryonic forebrain has an impact on brain development, in particular in regards to brain morphology. Thus, embryos were collected at two different time points E 13.5 and E 15.5 (data for E 15.5 not shown) and histological analysis was performed. TRIM28 deficiency does not result in malformation of the developing forebrain or in cell loss, which was shown by Nestin and DAPI expression (Figure 3D).

Interestingly, homozygous TRIM28 knockout embryos were underrepresented in number (Table 1) comprising only 11% instead of the expected 25% of the total offspring.

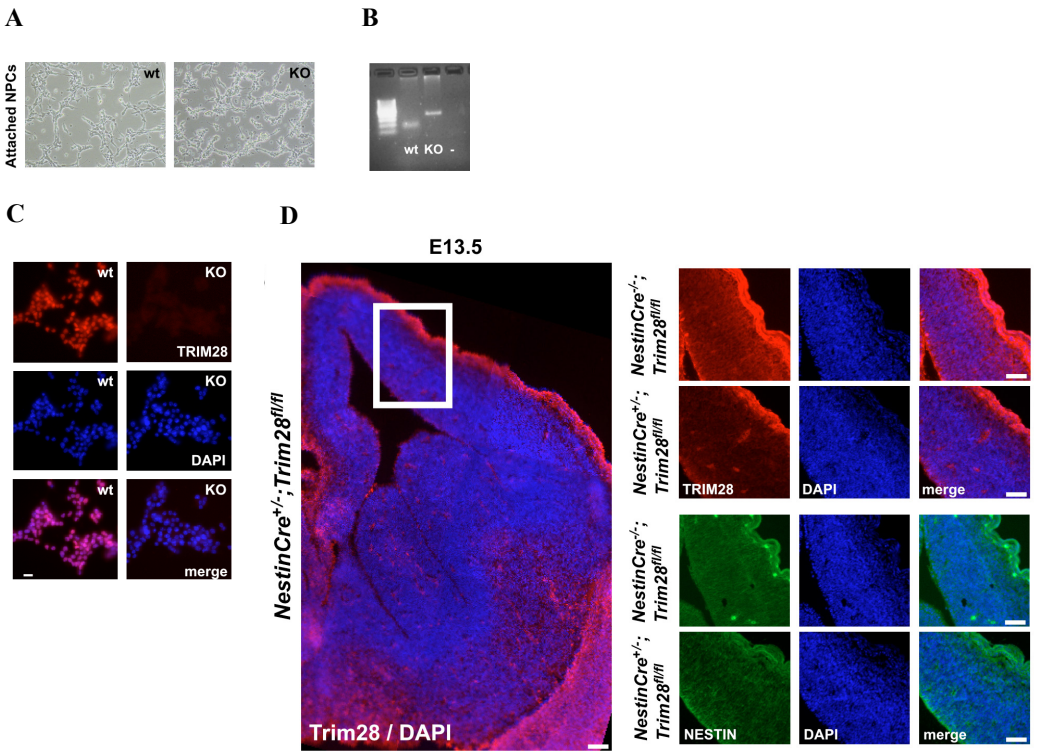



**Figure 3.** *Generation of TRIM28 knockout mouse embryos (E 13.5) and culturing of neural progenitor cells.*

(A) TRIM28 deficient mNP cell cultures. (B) Genotyping of TRIM28 knockout mNP cells. (C) Immunofluorescent analysis of TRIM28 protein expression in mNP cells. (D) Immunohistochemical analysis revealing the normal morphology of TRIM28 knockout embryonic forebrain and verifying the lack of TRIM28 protein as well as the normal expression of the neural progenitor marker Nestin.

**Table 1**. *Genotypic analysis of offspring from mating NestinCre$^{(+/-)}$;TRIM28$^{(fl/wt)}$ males with TRIM28$^{(fl/fl)}$ females at E 13.5.*

| Genotypes | # of Embryos | Percentage | exp. Percentage |
|---|---|---|---|
| NestinCre $^{+/-}$ Trim28 $^{fl/fl}$ | 8 | 11 % | 25 % |
| NestinCre $^{+/-}$ Trim28 $^{fl/wt}$ | 20 | 41 % | 25 % |
| NestinCre $^{-/-}$ Trim28 $^{fl/fl}$ | 13 | 19 % | 25 % |
| NestinCre $^{-/-}$ Trim28 $^{fl/wt}$ | 29 | 29 % | 25 % |
| **Total** | **70** | **100 %** | **100 %** |

# TRIM28 deficiency in mouse Neural Progenitor Cells leads to Activation of Transposable Elements (Paper I)

We performed RNA-sequencing (RNA-seq) analysis of mNP cells and further mapped those results against Repbase (Jurka et al., 2005), a database containing consensus sequences of repetitive elements. We discovered that in TRIM28 deficient mNP cells a certain fraction of transposable elements was highly upregulated compared to wild type controls. Two groups or transposable elements attracted our attention: *Mus musculus ERV* using $tRNA^{Lys}$ *type 10C (MMERVK10C)* and *IAP1*, which we found to be distinctively upregulated in mNP cells lacking TRIM28 protein. We also looked at other transposon families but either detected only a slight elevation in expression, when looking for example at *MusD* and *Line-1* (Figure 4A), or did not see any difference when compared to controls. To confirm these RNA-seq findings, we performed q-RT PCR using specific primer pairs detecting the expression level of *MMERVK10C, IAP1, IAP-consensus, MusD* and *Line-1*. *IAP-consensus* primers were designed to detect the expression level of the entire *IAP*-family. These results show a less significant upregulation of the transcript compared to the specific *IAP1* expression (Figure 4B). From these results I conclude, that TRIM28 regulates a certain fraction of transposable elements.

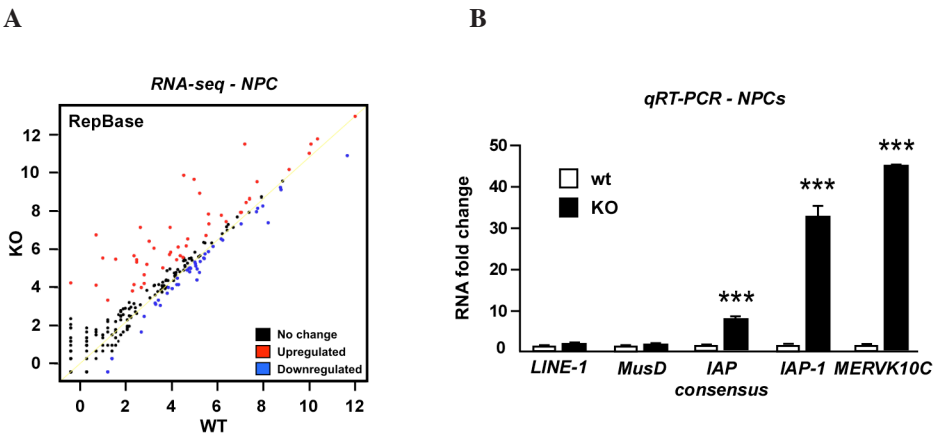



**Figure 4**. *Analysis of upregulated transposable elements upon TRIM28 knockdown in mNP cells.*

(A) RNA-seq analysis: knockout samples were plotted against wild type controls; dots represent individual Repbase sequences. (B) q-RT PCR of knockout mNP cells show upregulation of a certain fraction of transposable elements.

## TRIM28 deficient mouse Neural Progenitor Cells are able to differentiate into Neurons and Astrocytes (Paper I)

mNP cells from TRIM28 depleted embryos show all expected properties of a neural progenitor cell. The cells continued to proliferate and could be expanded for more than 60 passages and expressed the neural progenitor marker Nestin (Figure 5A). However, when comparing TRIM28 knockout mNP cell cultures to wild type control cells, I observed that TRIM28 deficient cells had a tendency to grow in clusters and were less prone to attach to the culturing flask surface. In order to verify that TRIM28 deletion had no impact on the capability of differentiation, TRIM28 knockout mNP cells were subjected to a differentiation assay revealing that differentiated cells expressed β-III-tubulin and GFAP and thus demonstrated that these cells were able to give rise to neurons and astrocytes (Figure 5B). Hence, I conclude that neither TRIM28 deficiency, nor significant upregulation of transposable elements influence the potential for differentiation.

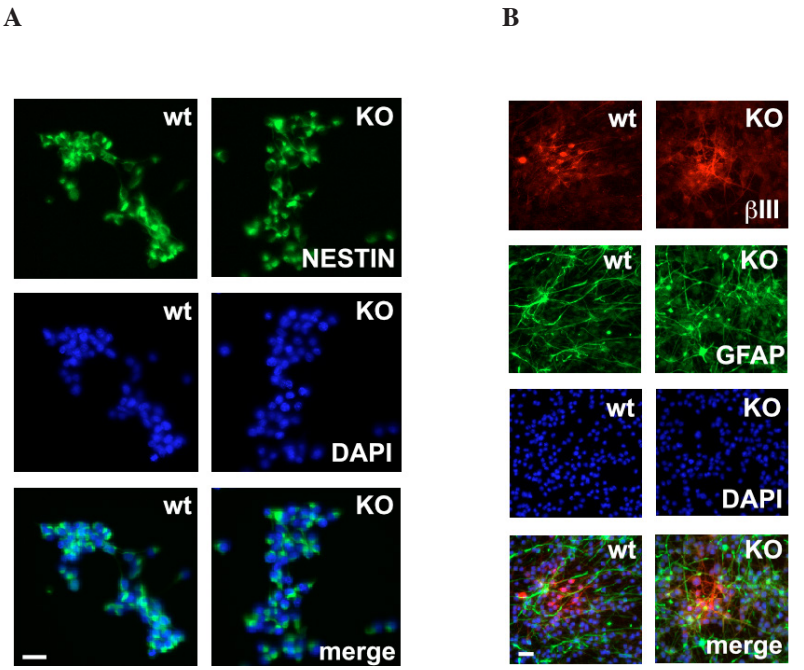



**Figure 5**. *Immunocytochemistry of undifferentiated and differentiated mNP cells.*

(A) TRIM28 knockout derived mNP cells express Nestin and are able to differentiate which is indicated by β-III-tubulin and GFAP expression (B).

# TRIM28 regulates *MMERVK10C* in mouse Neural Progenitor Cells (Paper I)

We found *MMERVK10C* to be one of the two transposable elements in TRIM28 deficient mNP cells, which showed the strongest upregulation. *MMERVK10C* was characterised as *beta*-like *endogenous retrovirus*, which recently invaded the mouse genome and was described to be present as provirus in about 20 loci but also in over 1000 incomplete loci (Reichmann et al., 2012). We analysed the *MMERVK10C* provirus using RetroTractor software (Sperber et al., 2007) To investigate the expression level of the *MMERVK10C* provirus in TRIM28 knockout mNP cells, we designed primers recognising *LTR, gag, pol* and *env* and found a 170-fold upregulation compared to wild type controls (Figure 6A). By looking at the transcription of *MMERVK10C* provirus in TRIM28 deficient embryonic forebrain tissue, we discovered elevated levels of *LTR, gag, pol* and *env* (Figure 6B).

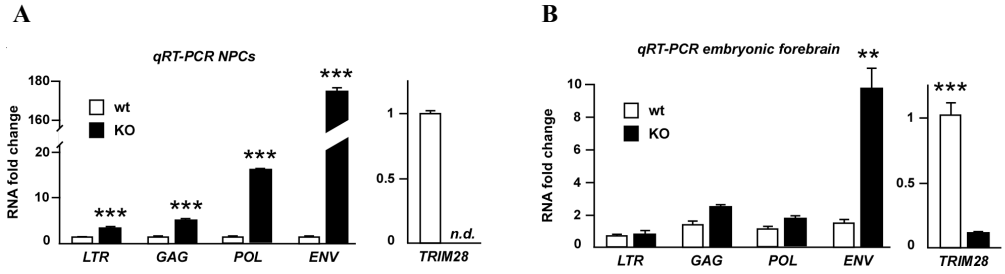



**Figure 6**. *q-RT PCR analysis of MMERVK10C provirus expression.*
(A) in mNP cells and (B) in mouse embryonic forebrain (E 13.5).

## *MMERVK10C* is associated with H3K9me3 in mouse Neural Progenitor Cells (Paper I)

Previously it was shown, that TRIM28 is involved in the establishment of H3K9me3 associated heterochromatin (Rowe et al., 2010). Since we discovered that a certain proportion of transposable elements in mNP cells are regulated by TRIM28, this raised the important question, if TRIM28-mediated repression follows the same mechanisms in mNP cells as described for mES cells (Rowe et al., 2010). Therefore, I performed chromatin immunoprecipitation (ChIP) experiments on TRIM28 knockout mNP cells, as well as on wild type controls, using a specific antibody against H3K9me3. For this analysis the same primers were used as designed for detecting the *MMERVK10C* provirus (*LTR, gag, pol* and *env*). The results revealed that *MMERVK10C* sequences are enriched for H3K9me3 in mNP cells by showing a substantial loss of H3K9me3 at *MMERVK10C* proviral loci upon TRIM28 depletion (Figure 7). These findings suggest, that *MMERVK10C* is repressed by TRIM28-mediated heterochromatin.

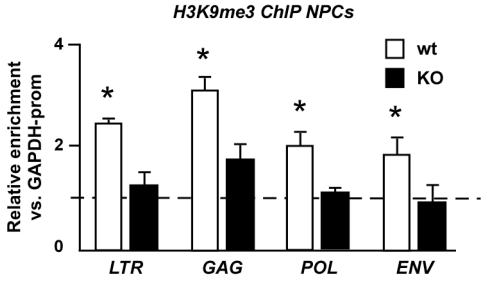



**Figure 7**. *Chromatin immunoprecipitation (ChIP) for H3K9me3 investigating the heterochromatin state of MMERVK10C provirus in mNP cells.*

## Analysis of *IAP1* Expression in mouse Neural Progenitor Cells (Paper I)

According to our RNA-seq data, the second group of transposable elements that we discovered to be highly expressed upon TRIM28 depletion were *IAP1*-elements. These transposons are known for their ability to retrotranspose and are specifically found in the mouse genome (Qin et al., 2010). The RNA-seq results were confirmed by q-RT PCR, using primer pairs that detect *IAP1-LTR* and *IAP-pol* (Figure 8A). When investigating the chromatin state of *IAP1* loci upon TRIM28 knockdown (using the above mentioned primers *IAP1-LTR* and *IAP-pol*), we were able to determine that these loci are less enriched for H3K9me3 compared to wild type mNP cells (Figure 8B). These data are in line with our findings regarding TRIM28 mediated H3K9me3 of *MMERVK10C* proviral loci. By performing immunocytochemistry using an *IAP* antibody, we were able to detect IAP-gag protein expression (Figure 8C). Taken together these data not only demonstrate active *IAP1-gag* transcription in TRIM28 deficient mNP cells, but furthermore shows that in fact the transcribed *IAP1-gag* actually results in active translation of a transposon derived protein.

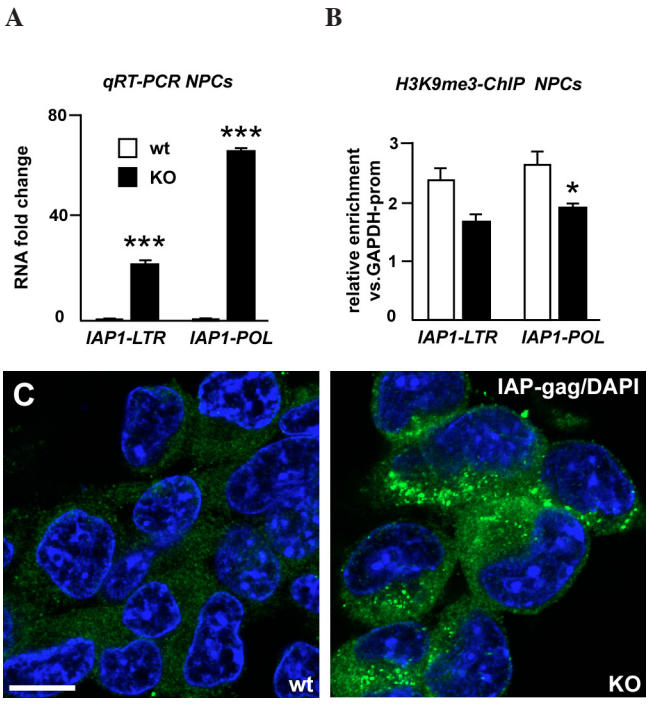



**Figure 8**. *Analysis of IAP1 expression in mNP cells. (A) q-RT PCR analysis of different regions of the IAP1 provirus.* (B) Chromatin immunoprecipitation (ChIP) for H3K9me3. (C) Confocal analysis - investigating immunofluorescence of *IAP-gag* staining of TRIM28 deficient cells.

## Activation of *MMERVK10C* and *IAP1* leads to Transcription of Nearby Genes (Paper I)

RNA-seq analysis of TRIM28 depleted mNP cells revealed upregulation of 26 *MMERVK10C* and 361 *IAP1* proviruses, which were mapped to exact genomic locations. Taken all proviruses together, 90 were found to be located nearby genes (<50 kb). In 25 of those genes the expression level was highly enhanced, while a reduced gene expression was only detected in 6 genes (Figure 9A). Genes with increased gene expression (due to the upregulation of the provirus), were on average 3-fold upregulated (Figure 9B). Proviruses, that were not upregulated upon TRIM28 depletion, did not alter the gene expression of close by genes (Figure 9B). Five genes that we found to be upregulated were validated by q-RT PCR: Fbxw19, Klrb1a, 240018L13Rik, Olfl1350 and ZFP932. These results generated by q-RT PCR confirmed the significantly increased expression of those genes, which is consistent with our findings from RNA-seq data (Figure 9C).

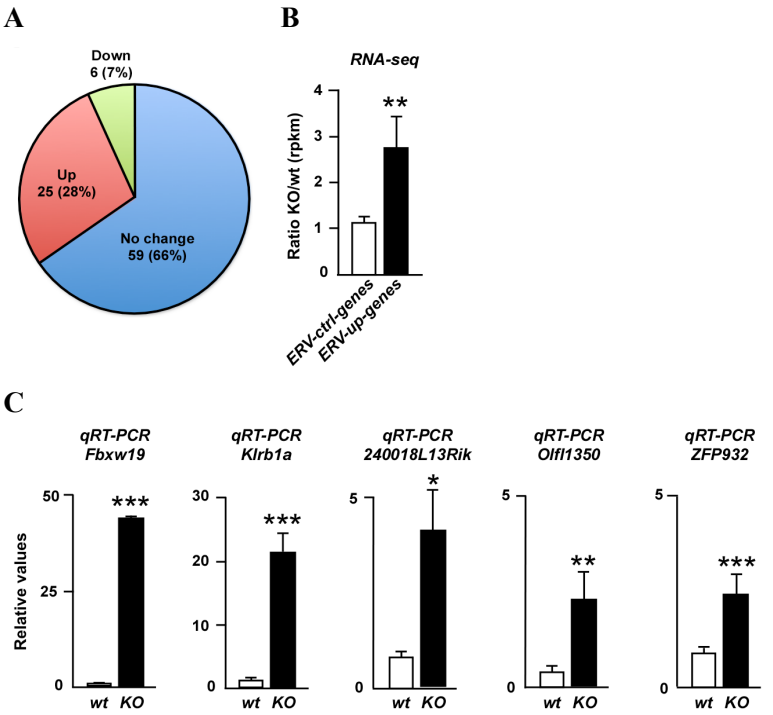



**Figure 9**. *Activation of transposable elements results in transcription of nearby genes.*

(A) Change of gene expression upon TRIM28-dependent upregulation of transposable elements. (B) Mean change of host genes transcription upon activation of transposable elements. (C) q-RT PCR analysis of five upregulated genes.

# TRIM28 knockout-dependent Activation of Transposable Elements results in Expression of Long Non-coding RNAs (Paper I)

We further investigated a highly activated *IAP*-element, which was only found expressed upon TRIM28 knockdown. 5 kb downstream of that particular *IAP*-element we found BC048671, a protein-coding transcript that is not expressed in wild type mNP cells. RNA-seq data revealed that transcriptional initiation at that *IAP*-element results in formation of a >10 kb long transcript, which extends into the coding sequence of BC048671 (Figure 10A and B). These findings demonstrate that activated transposable elements have the potential to act as transcriptional start sites.
Looking closer into non-coding regions, we discovered that both *MMERVK10C* and *IAP1* have the potential to activate expression of long non-coding RNAs (Figure 10 C-F).

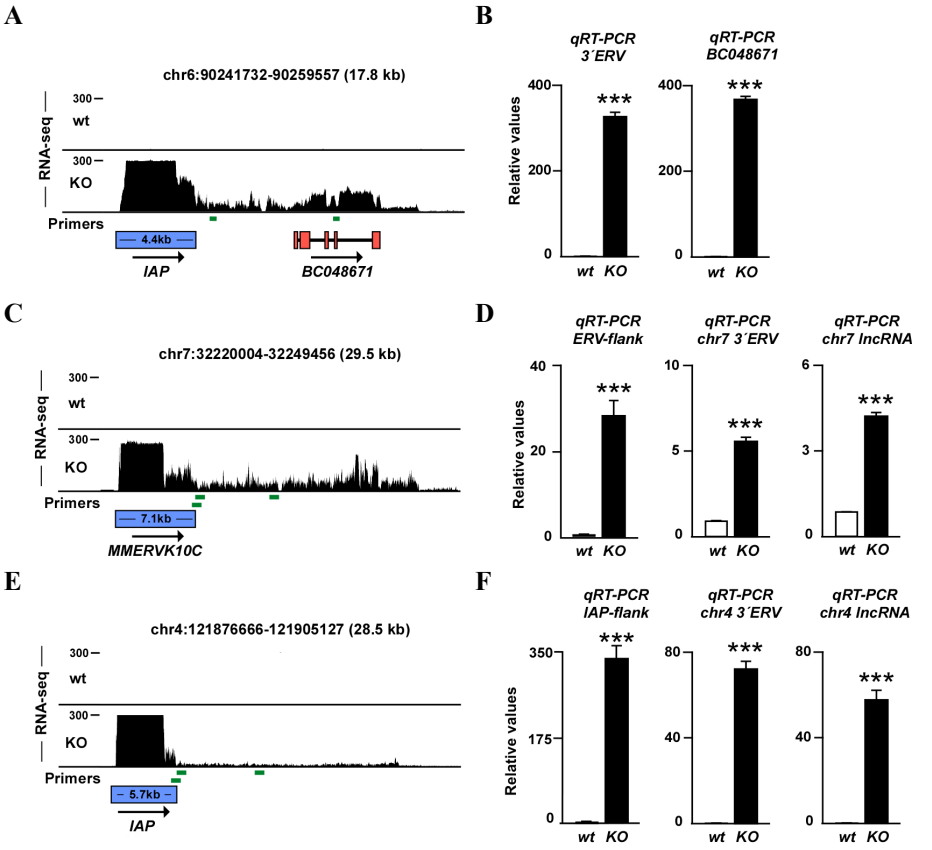



**Figure 10**. *Activation of transposable elements in TRIM28 deficient mNP cells results in transcription of long non-coding RNAs.*

(A) Screen shot from UCSC genome browser shows induced transcription of BC048671 in TRIM28 depleted mNP cells. (B) q-RT PCR analysis of extended *IAP-1* and BC048671 transcripts. (C and E) *MMERVK10C* and *IAP-1* activate expression of long non-coding RNAs (D and F) q- RT PCR validation of non-coding RNAs.

## TRIM28 homozygous knockout is lethal (Paper I)

Our *in vitro* studies of mNP cells derived from TRIM28 knockout embryos show a distinct impact on activation of transposable elements, which leads to the question, if transposon activation alters the transcriptome of the adult mouse brain. Therefore, I applied the same breeding strategy as used for culturing TRIM28 knockout mNP cells. Interestingly, when the 21 days (P21) old offspring was genotyped, none of the homozygous mice survived (Table 2). Here, I would like to add that one week after one litter was born, I discovered two mice smaller in size showing apathetic-like behaviour. Very soon after, these mice died for unknown reason and genotyping revealed that both mice had been homozygous TRIM28 knockouts (data not shown).

**Table 2**. *Genotypic analysis of offspring (P21) from mating NestinCre$^{(+/-)}$;TRIM28$^{(fl/wt)}$ males with TRIM28$^{(fl/fl)}$ females.*

| Genotypes | # of Mice | Percentage | exp. Percentage |
|---|---|---|---|
| NestinCre $^{+/-}$ Trim28 $^{fl/fl}$ | 0 | 0 % | 25 % |
| NestinCre $^{+/-}$ Trim28 $^{fl/wt}$ | 33 | 26 % | 25 % |
| NestinCre $^{-/-}$ Trim28 $^{fl/fl}$ | 42 | 32 % | 25 % |
| NestinCre $^{-/-}$ Trim28 $^{fl/wt}$ | 54 | 42 % | 25 % |
| **Total** | **129** | **100 %** | **100 %** |

# Mono-allelic TRIM28 Expression results in Behavioural Phenotype (Paper I)

Observing the born offspring consisting of TRIM28 heterozygous mice and their wild type littermates, I could see that a subset of mice was behaving in a more active manner. Therefore I decided to perform behavioural experiments. I have chosen the Open Field (OF) to study activity and the Elevated Plus Maze (EPM) to investigate anxiety-like behaviour. Interestingly, mice with a mono-allelic TRIM28 expression were significantly more active than their wild type littermates when performing the OF test (Figure 11A). Furthermore, I measured that heterozygous TRIM28 mice were spending more time in the open arms of the EPM indicating decreased anxiety-like behaviour compared to control mice (Figure 11B).

.

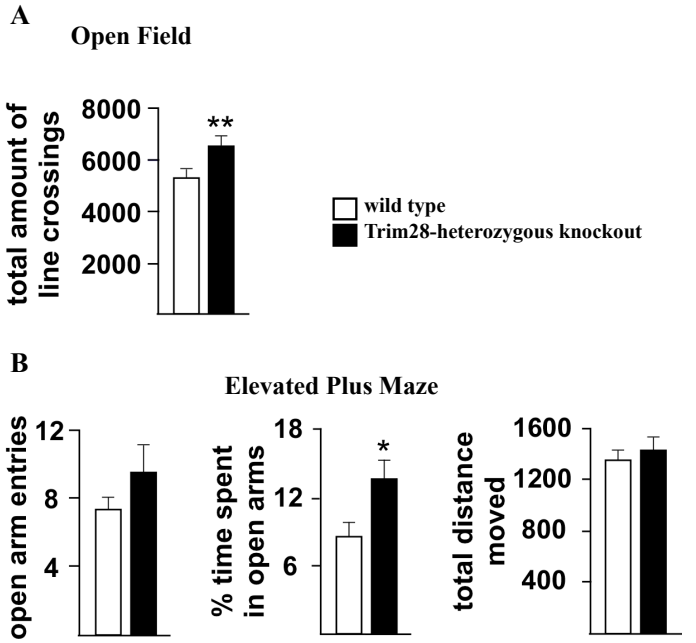



**Figure 11**. *Heterozygous TRIM28 deletion leads to behavioural phenotype in adult mice.*

(A) Examination of total amount of line crossing in the Open Field behavioural test. (B) Investigation of open arm entries, % time spent in open arms and total distance moved in the Elevated Plus Maze.

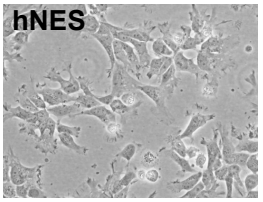
# Expression of Transposable Elements in human Neural Progenitor Cells (Paper II)

Our findings in Paper I lead to the question if our results in mNP cells are applicable to the human neural progenitor (hNP) cells. For studying transposable elements in hNP cells, we obtained human neuroepithelial-like stem (hNES) cells (Figure 12A), which have been previously described (Falk et al., 2012). hNES cells were characterised as such by immunocytochemistry expressing the progenitor markers Nestin, SOX1 and SOX2 (Figure 12B). Subsequently, it was interesting to investigate the baseline expression of transposable elements in hNES cells and compare the results to human hES cells. We performed RNA-seq of hNES cells and compared all reads of transposable elements to RNA-seq data, which was previously published about human embryonic stem (hES) cells (Turelli et al., 2014). We revealed higher transcriptional expression of *Alu*-elements and a similar expression of *Line-1* when comparing hNES cells to hES cells. Interestingly we noticed *ERVs* are predominantly silenced in hNES cells (Figure 12C).

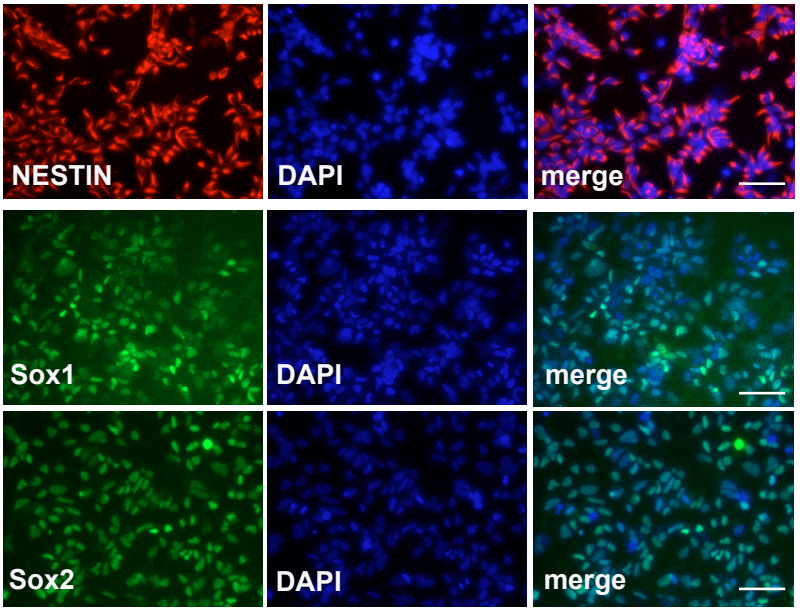
**Figure 12**. *Characterisation of hNES cells.*

(A) Brightfield image of hNES cells. (B) Immunohistochemical analysis of neural progenitor markers: Nestin, SOX1 and SOX2. (C) RNA-seq data of hNES and hES cells comparing mean expression of *Line-1, Alu* and *ERV*.

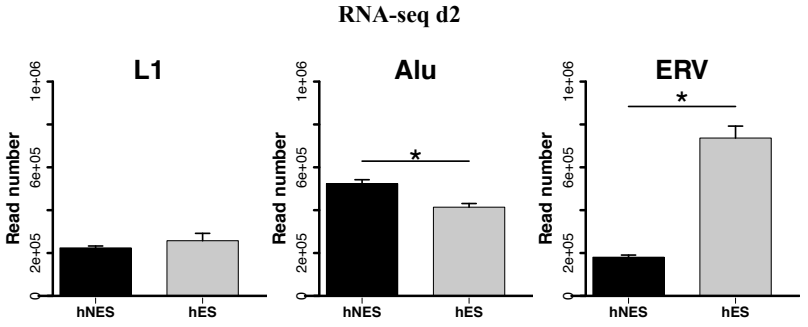
A


hNES

B


NESTIN
DAPI
merge
Sox1
DAPI
merge
Sox2
DAPI
merge

C


RNA-seq d2
L1
Alu
ERV
Read number
0
2e+05
6e+05
1e+06
*
hNES
hES

# TRIM28 knockdown in human Neural Progenitor Cells leads to Upregulation of *Alu*-elements (Paper II)

To achieve an efficient shRNA-mediated knockdown of TRIM28, I performed lenti viral transduction of hNES cells. hNES cells were transduced and collected 48 hours post transduction (Figure 13A). At first the knockdown was verified by q-RT PCR analysis (Figure 13B) and subsequently RNA-seq was performed.

RNA-seq analysis revealed an increased expression of human-specific *Alu*-elements upon knockdown, while the expression level of *Line-1* and *ERVs* was not altered (Figure 13C). These results were compared to a published data set (Turelli et al., 2014), where the same shRNA-mediated knockdown was performed on hES cells. Interestingly, in hES cells *Alu*-elements were not upregulated upon TRIM28 knockdown (Figure 13D). We found that TRIM28 regulates different transposable elements in hNES cells compared to hES cells.

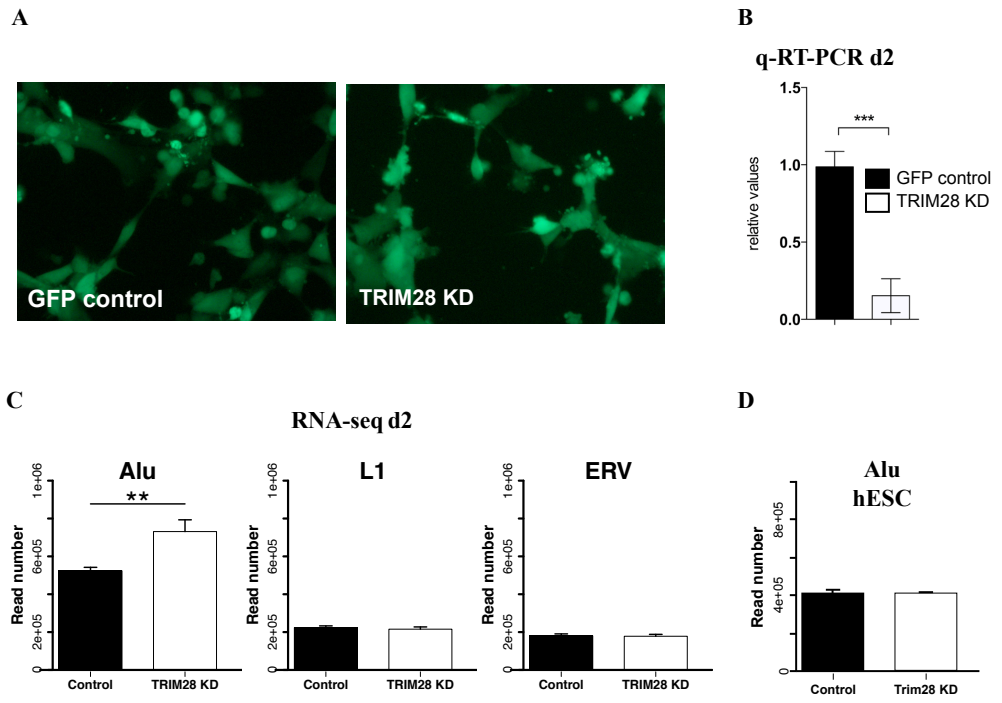



**Figure 13**. *TRIM28 knockdown and validation of Alu, Line-1 and ERV expression in hNES and hES cells.*

(A) GFP expression of hNES cells transduced with TRIM28 knockdown and GFP-control vector. (B) q-RT PCR validation of TRIM28 knockdown in hNES cells. (C) RNA-seq analysis of *Alu, Line-1* and *ERV* expression upon TRIM28 knockdown (day 2). (D) RNA-seq data of *Alu* expression in hES cells after TRIM28 knockdown.

## TRIM28 controls Young *Alu*-elements in human Neural Progenitor Cells (Paper II)

The *Alu*-family (more than 1.1 million elements) is classified according to the time point they invaded the human genome. *AluJ* was described to be an older *Alu* subpopulation compared to the younger members *AluS* and *AluY* (Tsirigos and Rigoutsos, 2009)>. When comparing different members of the *Alu*-family in our RNA-seq data, we found that *AluY* appears to be activated upon TRIM28 knockdown (Figure 14A). *Alu*-elements, which are usually controlled by Pol III driven promoters, have the tendency to be transcribed with help of the host Pol II when located within a gene.

Looking at the genome of TRIM28-deficient hNES cells, we were especially interested in which *Alu*-elements are highly expressed and furthermore their exact genomic location. Upon TRIM28 knockdown we found *AluYm1* and *AluYa5* to be independently expressed, suggesting that the activation of these elements relies on their own Pol III-driven promoter, indicating that their transcription is not a downstream effect of host gene transcription (Figure 14B).

It was recently described (Varshney et al., 2015) that histone modifications play a distinct role in regulation of *Alu*-elements. In line with these data we found that DNA-methyltransferase DNMT1 knockdown in hNES cells does not result in activation of *AluYa5* transcription (Figure 14C), revealing that TRIM28 has a significant role in silencing a subset of *Alu*-elements *e.g. AluYa5* in hNES cells (Figure 14D).

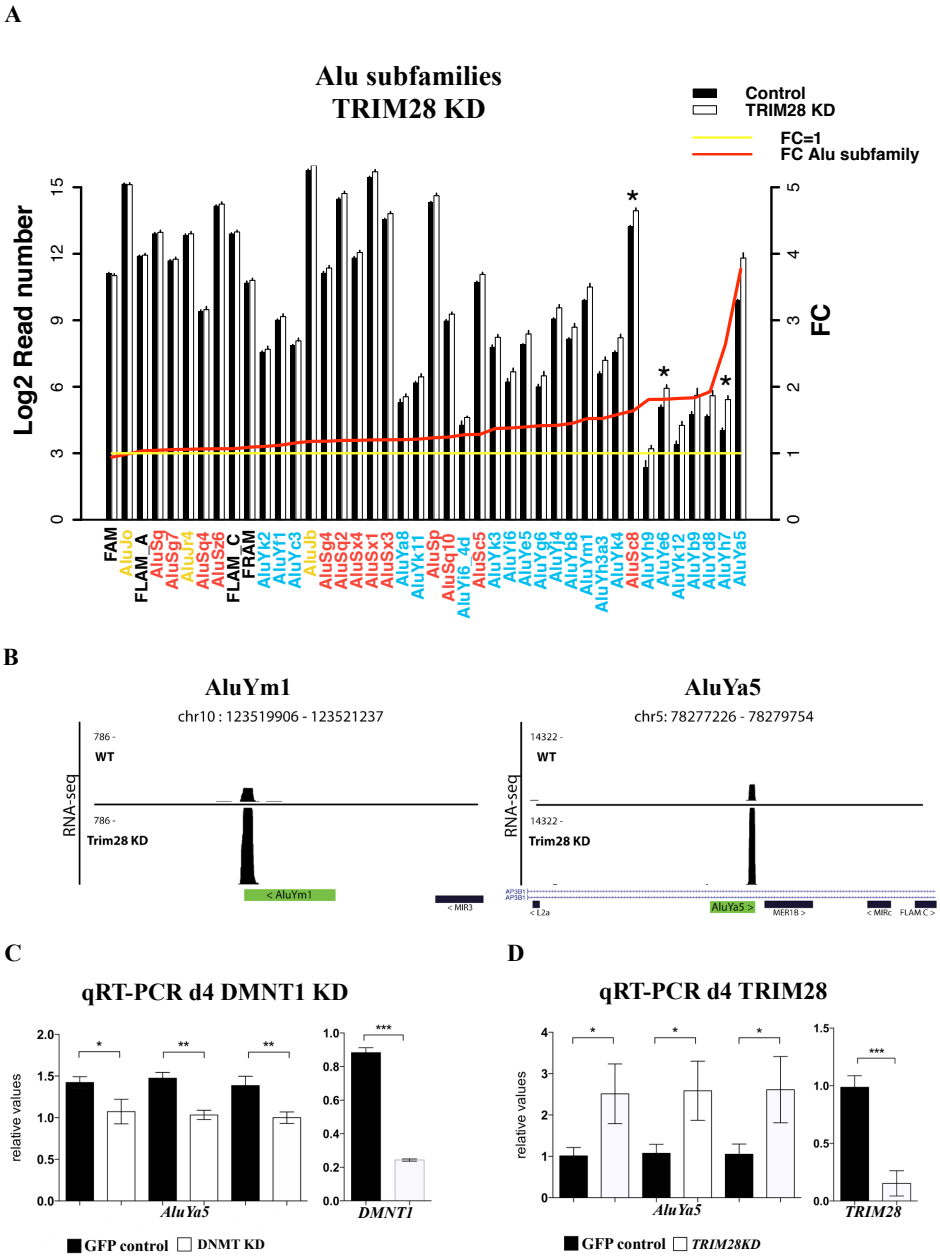



**Figure 14**. *Analysis of Alu-elements in TRIM28 and DNMT1 knockdown.*

(A) Mean global expression of *Alu*-subfamilies in TRIM28 knockdown and control. (B) Screen shots from UCSC genome browser showing *AluYm1* and *AluYa5* expression upon TRIM28 knockdown. (C) q-RT PCR analysis of *AluYa5* expression upon DNMT1 knockdown in hNES cells. (D) *AluYa5* expression upon TRIM28 knockdown in hNES cells.

## TRIM28-controlled *Alu*-elements affect Expression of Nearby Genes (Paper II)

We found 392 genes within a distance of 50 kb of 154 significantly upregulated *Alu*-elements (Benjamini-Hochberg corrected, $p<0.05$); when looking at RNA-seq data of TRIM28 knockout hNES cells. The majority of these genes were upregulated. Therefore we suggest that host genes close to *Alu*-elements are repressed as consequence of TRIM28 mediated silencing of these transposons (Figure 15).

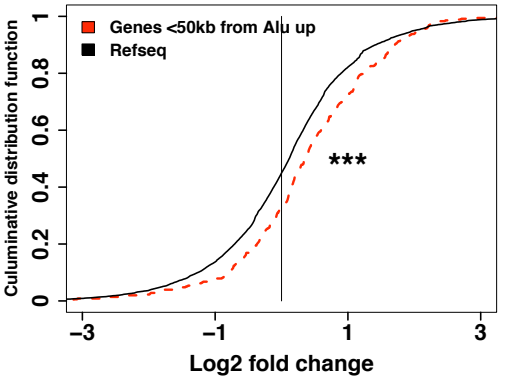



**Figure 15**. *Alu upregulation effects nearby gene expression.*

## TRIM28 repression of Retrotransposon-based Enhancers is necessary to preserve Transcriptional Dynamics in Embryonic Stem Cells (Paper III)

In brief, I would like to summarise the most relevant findings regarding TRIM28-mediated repression of transposable elements in mES cells and relate those to our results in mouse and human NP cells (Paper I and II).

TRIM28 deletion in mES cells was accomplished using a tamoxifen inducible Cre-lox-P system. As previously shown, TRIM28 silences endogenous retroviruses in mES cells with a distinct enrichment of H3K9me3 on transposable sequences (Rowe et al., 2010). Paper III reveals that TRIM28 depletion leads to activation of transposable elements in mES cells, inducing a change in gene expression of nearby genes. These findings are consistent with results we obtained in studying mouse and human

NP cells (Paper I and II). In mES cells, TRIM28-mediated regulation of transposable elements is associated with establishment of H3K9me3. TRIM28 deleted mNS cells show remarkable reduction of H3K9me3 associated heterochromatin (Paper I).

In the mES cell study (Paper III) we show that TRIM28 deletion causes a change of chromatin state regarding H3K9me3 at loci of transposable elements. Interestingly those genes, which were found to be upregulated have bivalent promoters characterised by H3K4me3 and H3K27me3 marks and are located close to H3K9me3 regulated transposable elements. Transposon activation was associated with a change from the repressive H3K9me3 state to active chromatin, characterised through H3K4me1 and H3K27ac histone marks.

These results from Paper III are very relevant for our studies in mouse and human NP cells, since similar experiments might give us a better insight in the consequence of activation of transposable elements. More research is required to improve the knowledge about the impact of transposable elements on the transcriptome.

# Concluding Remarks and Future Prospects

More than half a century has passed since transposable elements were discovered to comprise a large fraction of our genome. For a long time the general idea was that TEs were parasitic invaders with no benefit to the host. If that was the case, why did evolution select in favour of these elements. Transposons are useful for phylogenetic studies, since the comparison of different families and subgroups of TEs can give important insights into the evolution of different species and time points of separation from one another.

Prior studies have noted that TEs are highly regulated by epigenetic mechanisms (Kuramochi-Miyagawa et al., 2008; Rowe et al., 2013; Rowe et al., 2010; Rowe and Trono, 2011). Interestingly, TEs, which are well integrated into the host machinery, perform different tasks that include fine-tuning of gene regulation (Muotri et al., 2005)and are involved in exon-shuffling to create a larger genetic diversity (Cowley and Oakey, 2013; Hancks and Kazazian, 2012). Therefore the epigenetic machinery has the important function to balance between the benefits and drawbacks for the host. TEs that are found to be associated with diseases are kept silent (Robbez-Masson and Rowe, 2015). Deregulation and resulting activation of those elements was reported in many studies of cancer, autoimmune disorders or diseases of the central nervous system (Belshaw et al., 2005; Carreira et al., 2014; Helman et al., 2014; Lee et al., 2012; Qin et al., 2010; Reilly et al., 2013; Ryan, 2004; Solyom et al., 2012; Tubio et al., 2014). It has been demonstrated, that the regulatory mechanism of these mobile elements is dependent on the cell type, the transposon family and the phylogenetic age. In pluripotent ES cells TE regulation is dependent on TRIM28, while DNA-methylation induces stable silencing in somatic tissues (Matsui et al., 2010; Rowe et al., 2013; Rowe et al., 2010; Turelli et al., 2014).

This thesis presents an exception to this rule. Here I have identified TRIM28 as key regulator of a certain fraction of transposons in mouse and human NP cells, that is unique for somatic tissues. Although the types of TEs that underlie TRIM28-regulated repression vary between mouse and human, TRIM28 seems to regulate younger and more active transposons. Thus TRIM28-mediated control of specific TEs in mouse and human NP cells suggest that this repression mechanism is evolutionary conserved.

Here we discovered that TRIM28 represses different families of TE in hNP cells compared to the published data in hES cells (Turelli et al., 2014).

Another major finding of this thesis reveals that de-repression of TEs in NP cells activates transcriptional networks, which leads to transcription of nearby genes and long non-coding RNAs.

Taken together, these data demonstrate the potential of transposons on activation of transcription in multipotent neural progenitor cells. These findings indicate that transcriptional repression is a dynamic and highly adjustable mechanism and not a unified process.

Studies over the past decade revealed that certain families of transposable elements are highly active in mouse and human brain (Coufal et al., 2009; Kuramochi-Miyagawa et al., 2008; Muotri et al., 2005; Muotri et al., 2010; Muotri et al., 2009).

TEs are more active in NP cells or certain areas of brain (Coufal et al., 2009; Muotri et al., 2005; Muotri et al., 2010). These findings suggest that the cell diversity in the brain is affiliated with the activity of TEs, especially during neuronal differentiation (Muotri et al., 2005). Since retrotransposition was discovered in the hippocampus, the role of TEs in learning and memory is debated (Evrony et al., 2012).

Interestingly, TE-derived transcripts were found in patients suffering from psychiatric disorders (Karlsson et al., 2001). Nevertheless no evidence was found that those diseases are actually caused by activated TEs. Hypothetically, if disorders like schizophrenia are caused by mobile elements could these conditions be seen as a step during evolution to select for or against transposons.

In our mouse study for example we observed that mice with mono-allelic TRIM28 depletion show higher activity and less anxious-like behaviour. However, we can not conclude that these behavioural phenotypes are caused by transposable elements.

At present, several questions remain unanswered and research of the next few decades will hopefully help us to better understand, where transposable elements originate from and if they are the main drivers of evolution. More studies are required, to investigate how the host genome distinguishes between transposons that are beneficial and elements that interfere with its transcriptome. Furthermore we need to understand how the epigenetic machinery adapts the repression of TEs.

These and other findings will hopefully provide more insight and unravel the capacity of transposable elements as well as their mechanisms in disorders. Further development of existing techniques for example: retrotransposition assays, whole genome and single cell sequencing, as well as inventions of novel techniques will be highly valuable to study transoposons and their role in the genome.

# Materials and Methods

In this section I would like describe in more detail the most important techniques that I used for the studies in my thesis:

## Transgenic Mice

All procedures were approved by and conducted in accordance with the committee for use of laboratory animals at Lund University. The generation as well as genotyping of mice with a floxed TRIM28 allele was previously described by (Weber et al., 2002). The *Nestin-Cre* mouse was generated and described as model to achieve *Cre* recombination in the developing forebrain (Tronche et al., 1999). Transgenic mice were backcrossed to a C57/Bl6-background for at least 8 generations.

*Purification of total DNA from adult mouse-tail Biopsies*

For DNA-extraction of a small tail biopsy (~0.2 cm) was lysed in buffer containing Tris [pH 8.5], 5 mM EDTA [pH 8.0], 100 mN NaCl, 0,5% SDS and proteinase K (20.2mg/ml – Fermentas). The samples were incubated either for at least 4 hours or optionally over night at 56°C. Saturated NaCl was added to each sample, subsequently vortexed and incubated on ice for 20 minutes; followed by a centrifugation step of 19 600g. The supernatant was washed with 99.5% EtOH followed by centrifugation of 5 minutes at 15 000g. Next the pellet was washed by adding 70% EtOH, followed by another centrifugation step at the same conditions.

The pellet was air-dried for at least 1 hour. The pellet was dissolved in 100ml TE-buffer (optional MilliQ-$H_2O$) at 55°C for 10-20 minutes.

*Purification of total DNA from embryonic mouse-tail Biopsies*

Although the protocol for adult mouse-tail biopsies is applicable to embryonic tissues, I have observed that using DNeasy Blood and Tissue Kit (Quiagen) leads to a clearer result when comparing DNA band after electrophoresis. Therefore, I decided to use the kit for all DNA extractions of embryonic tissues.

Embryonic mouse-tails were lysed in ATL buffer after adding proteinase K. The samples were vortexed und incubated until the tissue was lysed (approximately for 40 minutes) at 56°C. The samples were vortexed before and after AL buffer was added. EtOH was added and the biopsies were vortexed again before being loaded onto a DNeasy Mini spin column placed in a 2 ml collection tube. The

flow-through was discharged after the samples were centrifuged at > 6000 g for 1 minute at room temperature.

The column membranes were washed with AW1 at > 6000 g for 1 minute followed by a washing step with AW2 buffer for 3 minutes at 20 000 g. DNA was eluted in 200 µl AE buffer.

*Purification of total DNA from Cultured Cells*

DNA extraction from cells was performed using DNeasy Blood and Tissue Kit (Quiagen). Cells were harvested, the cell pellet was resuspended in PBS and Proteinase K and incubated for 10 minutes at 56°C in AL buffer. Hereafter the extraction was preceded like described in the paragraph above.

*Polymerase Chain Reaction (PCR) to determine the Nestin Cre TRIM28 Mouse Genotype*

The PCR reaction mix for TRIM28 and Nestin Cre was prepared as shown in Table 3 using following PCR primer (Sigma) summarised in Table 4.

**Table 3**: *PCR Conditions for Genotyping*

| PCR - Reaction MIX | Trim28 | Nestin-Cre |
|---|---|---|
| MilliQ - $H_2O$ | 13.4 µl | 14.4 µl |
| 10xPCR buffer | 2.0 µl | 2.0 µl |
| dNTPs (each 2,5mM) | 0.5 µl | 0.5 µl |
| per Primer (10 mM) | 1.0 µl | 1.0 µl |
| Taq polymerase (DreamTaq™ Green DNA Polymerase - 500u) | 0.1 µl | 0.1 µl |
| DNA sample | 1.0 µl | 1.0 µl |
| **Total volume** | **20.0** µl | **20.0** µl |

PCR conditions for TRIM28:

The PCR was performed under following conditions: 3 minutes at 95°C, 30 seconds at 95°C, 30 seconds at 55°C, 30 seconds at 72°C, 7 minutes at 72°C, forever on 4°C - 35 cycles. Genotyping was preceded as previously described by (Weber et al., 2002).

Amplified sizes of DNA fragments - wild type:152 bp, mutant:180 bp.

PCR conditions for Nestin Cre:

The PCR was performed under following conditions: 3 minutes at 95°C, 30 seconds at 95°C, 30 seconds at 55°C, 30 seconds at 72°C, 7 minutes at 72°C, forever on 4°C - 30 cycles. Genotyping was preceded as earlier described by (Tronche et al., 1999).
Amplified size of DNA fragment: 100 bp.

**Table 4**: PCR *Primer for Genotyping*

| Primer | Sequence |
|---|---|
| *TRIM28 - 1* | GGAATGGTTGTTCATTGGTG |
| *TRIM28 - 2* | ACCTTGGCCCATTTATTGATAAAG |
| *TRIM28 - 3* | GCGAGCAGAATCAAGGTCAG |
| *Nestin Cre - forward* | GCCACCAGCTTGCATGATC |
| *Nestin Cre - reverse* | GGAGCCGCGCGAGAAT |

*PCR-product Amplification via agarose-gel electrophoresis*

2 % agarose gel was made using 1X TEA buffer and either SYBR.Safe DNA gel stain (5 µl / 50 ml gel - Invitrogen) or Advanced DNA-stain (3-4 µl / 50 ml gel – Nippon Genetics) was added.

The electrophoresis was performed using following systems: Easy-Cast™ Electrophoresis System with power supply PS 250-2 or VWR Power Source. The amplified DNA bands were compared to a 100 bp DNA Ladder (Fermentas).

*Dissection of mouse embryos*

Mouse embryos were obtained after mating *TRIM28*$^{(fl/fl)}$ homozygous females with *Nestin Cre*$^{(+/-)}$ *TRIM28 flox*$^{(fl/wt)}$ heterozygous males. Vaginal plugs were examined carefully to determine the plugging date, which was designated as E 0.5.

Mouse uteri were excised on embryonic day 13.5 (E 13.5) after abdominal incision and kept on ice in KPBS. Embryos were collected from the uteri and embryonic forebrains were dissected by using the dissection microscope (LEICA MZ APO).

*Culturing Neurospheres of dissected mouse embryonic forebrains*

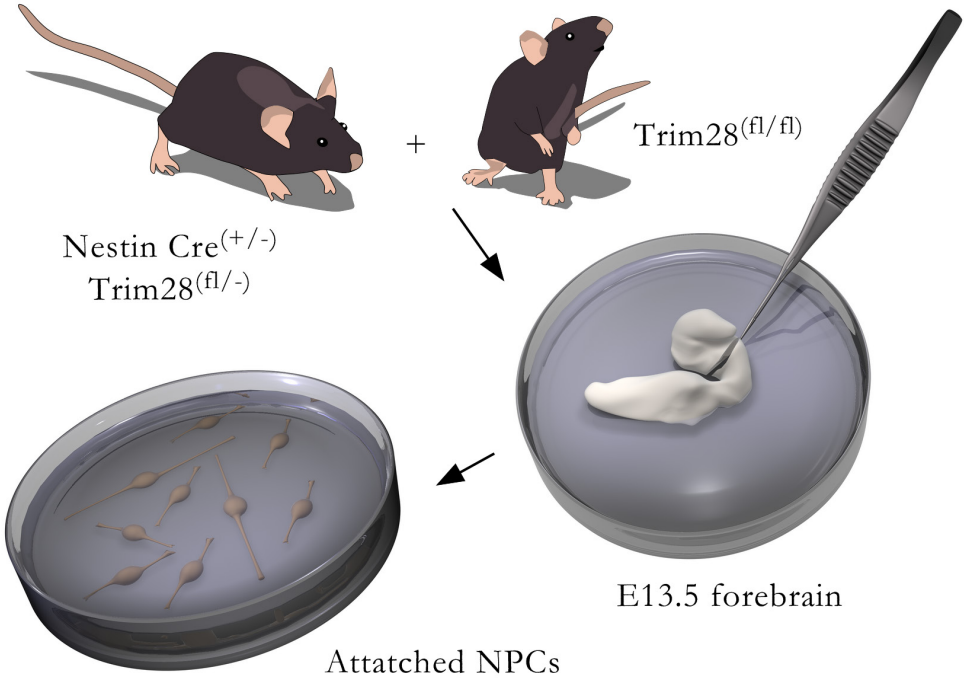



**Figure 16**. *Embryo dissection of forebrain and cell culture*.

For cell culture purposes embryonic forebrains were dissected in Leibovitz`s L15 medium 1x ([+] L-Glutamine, [+] L-Amino Acids - Invitrogen) or DMEM/F12 (Gibco); both of these media are equally fine to use. The embryonic forebrains were dissociated by adding trypsin and DNAse and incubated for 30 minutes at 37°C. By observation these cells had a lower survival rate compared to mechanical dissociation. Therefore, the forebrains were dissociated mechanically via pipette and cultured in mouse neurosphere complete medium (Table 5) containing DEMEM/F12, Penicillin/ Streptomycin, L-Glutamine, Glucose stock 30%, 100X N2 supplement adding growth factors EGF and bFGF. Neurosphere formation was visible 2-3 days after generating primary cell cultures. Every other day 1/5 of the medium was added to the free-floating cell cultures. Importantly, the cell cultures should be prevented from attaching to the surface, to guarantee sphere-formation and therefore the flasks should be gently tapped on a daily basis. To prevent infections of the cultures, the cells should preferably be cultures in flasks compared to plates.

The cells were split every 7-10 days by dissociating the spheres and plated in a density of 500 000 cells / ml; counted by using trypan blue (Sigma) and a Bürker-chamber (0.1 mm / 0.0025 mm / 0,04 mm). The protocol was adapted from (Ahlenius and Kokaia, 2010).

**Table 5**: *Cell Culture Media*

| Neurospheres Basic | 250ml |
|---|---|
| DEMEM/F12 (Gibco) | 237.5 ml |
| Penicillin / Streptomycin | 2.5 ml |
| Glucose (30% Sigma) | 5.0 ml |
| L-Glutamin (Sigma) | 2.5 ml |
| 100X N2 supplement (Invitrogen) | 2.5 ml |

| Neurospheres Complete | 50ml |
|---|---|
| Neurospheres Basic | 50 ml |
| EGF (R&D System) | 10 µl |
| bFGF recombinant Human FGF (R&D System) | 25 µl |

| NSC Basic | 250ml |
|---|---|
| Euromed-N (Euroclone) | 242.5 ml |
| Penicillin / Streptomycin | 2.5 ml |
| L-Glutamin (Sigma) | 2.5 ml |
| 100X N2 supplement (Invitrogen) | 2.5 ml |

| NSC Complete | 50ml |
|---|---|
| NSC Basic | 50 ml |
| EGF (R&D System) | 10 µl |
| bFGF recombinant Human FGF (R&D System) | 50 µl |

| hNES Basic | 250ml |
|---|---|
| Euromed-N (Euroclone) | 242.5 ml |
| Penicillin / Streptomycin | 2.5 ml |
| L-Glutamin (Sigma) | 2.5 ml |
| 100X N2 supplement (Invitrogen) | 2.5 ml |

| hNES Complete | 50ml |
|---|---|
| NSC Basic | 50 ml |
| EGF (R&D System) | 5 µl |
| bFGF recombinant Human FGF (R&D System) | 25 µl |
| B27 (Invitrogen) | 50 µl |

*Generation of Neural Progenitor Cell Cultures*

Neurosphere cultures are a heterogenous cell population. To be able to work with cells that are expanding faster and are more homogenous, I decided to generate neural progenitor cells from the established neurospheres. The expansion of these cells turned out to be more efficient.

Neurospheres were dissociated as performed during normal splitting procedure and cultured as previously described (Conti et al., 2005) on 0.1 % gelatine coated T-25 flasks (Nunc) in neural stem

cell (NSC) complete medium (Table 5) containing Euromed-N medium, Penicillin/Streptomycin, L-Glutamine, 100X N2 supplement adding growth factors EGF and bFGF recombinant Human FGF. The cells were split every 2-3 days depending on their state of confluence using Accutase (PAA).

*Differentiation of Neural Progenitor Cell Cultures*

For differentiation, NP cells were plated in a density of 30 000 cells / $cm^2$ on laminin (Invitrogen) coated 4-well plates using mouse neural stem cell basic medium. Medium was changed every other day for a period of 7 days.

*hNES Cell Cultures*

hNES cells were cultured like described (Falk et al., 2012) on 0.01 % Poly-L-Ornithine and 0.5 % Laminin (in PBS – both Sigma) and using following medium conditions (Table 5) DMEM/F12, Penicillin/Streptomycin, L-Glutamine, 100X N2 supplement; adding growth factors EGF, bFGF recombinant Human FGF and B27. The cells were split every 2-3 days depending on their state of confluence; half of the culturing medium was changed on a daily basis. For the splitting procedure: medium was aspirated from the adherent cells and incubated in TrypLE Express for 3-4 minutes at 4°C until the cells detach. The reaction was stopped by adding trypsin inhibitor (both Invitrogen). The cells were spun down in preheated medium (without B27, EGF and bFGF) at 300 g for 4 minutes at room temperature. The supernatant needs to be aspired completely before the cells are resupended in medium for plating. Cells are plated in a density of 1:2 or 1:3 dependent on the confluency of the harvested cells. To be safe 1:2 split is recommended.

*TRIM28 Knockdown in human Neural Epithelial Stem Cells (hNES) using Lenti Viral Vectors*

Cells were plated at a density of 35 000 cells / $cm^2$. For the experiment hNES cells were immediately transduced with two different sh-TRIM28 vectors (plko.1_shKap1B_GFP and plko.1_shKap1D_GFP) as well as the control vector sh- plko.1_shLucA_GFP. The cells were either collected day 2 or day 4 post-transduction. RNA extraction and qPCR analysis were performed using specific TRIM28 primers (Table 6). RNA from both knockdown samples was sent for sequencing and the transcriptome was compared.

## Chromatin Immuoprecipitation – ChIP

*Crosslinking*

The crosslinking of proteins to DNA was performed using the following protocol. Cells were harvested as usual and counted. Per crosslinking reaction 1 x $10^7$ cells were used. Pelleted cells were resuspended in 10 ml medium. For the crosslinking 275 µl 37 % formaldehyde (Sigma) was added to the cells suspension and incubated for exactly 10 minutes at 4°C under slow rotation. The reaction was stopped using 625 µl 2.5 M Glycine (Sigma) and incubated for 5 minutes at 4°C under slow rotation.

The cells were pelleted at a centrifugation step at 1700 g at 4°C. The cell pellet was resuspended in 12 ml ice-cold PBS and spun at 1700 g at 4°C. That step was repeated twice. In the final step of crosslinking the cells were resuspended in 2 ml PBS and pelleted at 1700 g and 4°C. The cell pellets were frozen and stored at -80°C until used for further sonication. The crosslinking process is a very delicate procedure. The cells should be kept ice-cold during the whole procedure.

*Cell lysis and Chromatin Shearing*

Cell lysis and preparation of DNA for chromatin shearing was preformed using iDeal ChIP-seq kit (Diagenode) according to supplier's recommendations. Shearing chromatin was performed by sonication using Bioruptor® (Diagenode). During the entire process of sonication the samples were kept ice-cold. The following protocol is approved for a volume of 200 µl of lysed cells. For a minimal variation in sonication, the procedures should be precisely performed under the same conditions. The Bioruptor® was cooled down with ice for about 30 minutes. The ice was removed and cold water was filled until 1 cm below the indication mark for maximal volume to be added and filled up with floating ice until reaching the mark. The sonication was performed for 40 cycles (30 seconds "ON" and 30 seconds "OFF") at "HIGH" power setting. After an interval of 5 sonication cycles, the melting ice was replace with fresh ice to maintain the same sonication conditions.

*Evaluation of Sonication*

There are many different options how to validate sonication. Therefore I decided to take aliquots of each sonicated sample to determine the sample quality for performing ChIP. 80 µl of sheared chromatin, 3 µl NaCl (Invitrogen), 5 µl Proteinase K, 1 µl RNAse A (both Thermo Fisher Scientific) was incubated at 65°C over night. 4 µl Glycine and 200 µl Phenol Chloroform (Invitrogen) was added, vortexed and spun in Phase Lock tubes (heavy gel – EppendORF) for 10 minutes at 4°C. The aqueous phase was collected and 3M KAC (pH 5.5 – Invitrogen) and >2.5 volume cold 100% EtOH were added and incubated for 30 minutes at -80°C to precipitate the DNA. The samples were centrifuged 15 minutes at 4°C and 16 000 g. The pellet was washed with 70% EtOH and centrifuged for 5 minutes at 4°C and 7 600 g. Then the pellets were dried and resuspended in MilliQ-$H_2O$ and incubated at 55°C. Loading dye (Thermo Fisher Scientific) was added the samples and the DNA fragments were amplified by agarose gelelectrophoresis.

*Immunoprecipitation*

ChIP was preformed using iDeal ChIP-seq kit (Diagenode) according to supplier's recommendations. The H3K9me3 antibody (Diagenode, pAb-056-050) was used at 2 µg / reaction. Primer sequences are summarised in Table 7. PCR SYBR green quantitative real-time PCR reactions were run in triplicates with Roche SYBR Green PCR Master Mix (Roche) using standard procedures.

To quantify the relative enrichment of each sequence a ΔCt for each sample was determined (CtInput - CtSample). The relative enrichment was then calculated by raising 2 to the ΔCt power. Relative quantification between KO and WT samples was performed by calculating a ΔΔCt-value for each pair of samples that were run in parallel (KOΔCt – WTΔCt). The fold difference was then

determined by raising 2 to the ΔΔCt power. The fold difference amongst pairs was then normalized to the Gapdh. All data are expressed as mean +/- S.E.M, based on the results of 3 complete independent experiments.

*Analysis of Gene expression – RNA extraction, cDNA-synthesis and quantitative Real Time PCR (q-RT PCR)*

RNA extractions were performed using RNAeasy mini / micro kit (Quiagen) followed by cDNA-synthesis (Fermentas); both were performed according to supplier´s recommendations. For cDNA synthesis the RNA concentrations of samples was measured. The RNA concentration that was used was in a range between 400ng – 2µg. For each experiment the same amount of RNA was used in all samples. Furthermore 2 negative controls were included; one of them excluded RNA, while the other one excluded Enzyme Mix.

For maintaining accuracy and reproducibility, the pipetting was performed using either the VarispanArm robot (Perkin Elmer) or the Bravo robot (Agilent Technologies).

q-RT PCR was performed using DNA-dye SYBR Green Mastermix and SYBR Light Cycler 480® (both Roche). Primer sequences are summarised in Tables 6 and 7.

**Table 6:** *q-RT PCR Primers - human*

| Primer | Sequence | Primer | Sequence |
|---|---|---|---|
| *b-Actin* | CCTTGCACATGCCGGAG<br>GCACAGAGCCTCGCCTT | *AluY5 P1* | CAGGAGATCGAGACCATCCC<br>CACTACGCCCGGCTAATTTT |
| *Gapdh* | TTGAGGTCAATGAAGGGGTC<br>GAAGGTGAAGGTCGGAGTCA | *AluY5 P2* | GCTCACGCCTGTAATCCCA<br>GGGATGGTCTCGATCTCCTG |
| *Hprt-1* | ACCCTTTCCAAATCCTCAGC<br>GTTATGGCGACCCGCAG | *AluY5 P3* | GCAGGAGAATGGCGTGAAC<br>AGTCTCGCTCTGTCGCCC |
| *TRIM28* | GTCAATGATGCCCAGAAGGT<br>GTCACTCTCCAGAGCCCAAG | *DNMT-1* | GATCGAGACCACGGTTCCTC<br>CGGCCTCGTCATAACTCTCC |

**Table 7:** *q-RT PCR Primers - mouse*

| Primer | Sequence | Primer | Sequence |
|---|---|---|---|
| *b-Actin* | TAG GCA CCA GGG TGT GAT GG<br>CAT GGC TGG GGT GTT GAA GG | *IAP1-POL* | TGG CCA TAC CCC AAA GAT AA<br>CCA GTT TAC TGG GGC TGG TA |
| *Gapdh* | TCC ATG ACA ACT TTG GCA TTG<br>CAG TCT TCT GGG TGG CAG TGA | *2410018L13Rik* | CCC ACT GCC TCT AGC TTC AC<br>TTC TTC CAG GGA CAT TTT GC |
| *Gapdh (ChIP)* | CCC ACT CCG CGA TTT TCA<br>CCT ACT CCG CGA TTT TCA | *Fbxw19* | TGT GTA CGT GTG GGA GGA GA<br>AGA AAG CAG GGA ATG GGA CT |
| *TRIM28* | GCC TCT GAC TGA AGG TCC TG<br>TCC AAG CCT GAG CTG GTA CT | *Olfr1350* | AGA TAT CCC TCC CAG CCT GT<br>GGG CAA GGA GAA AGT GTT GA |
| *IAP* | CGG GTC GCG GTA ATA AAG GT<br>ACT CTC GTT CCC CAG CTG AA | *Klrb1a* | ACC ATG AAA CCC TGA GCA AC<br>TGA GAG GCA GAC AGC AGA GA |
| *MerkLZ10* | CAA ATA GCC CTA CCA TAT GTC AG<br>GTA TAC TTT CTT CTT CAG GTC CAC | *Zfp932* | CAG GCT TGA ATG GTC CCT TA<br>TCA GCA AAG CCC ATT CTT CT |
| *IAP-LTR* | TGT GCC AGG CAG TAA ACA AG<br>ACC AAT CAC CAC AGG TCA CA | *MERVlnc-flank* | GCA GTC AAT GCT CTC CCA AT<br>CCC ATT CTT GAG GTT TTC TCT TT |
| *Line1* | TTT GGG ACA CAA TGA AAG CA<br>CTG CCG TCT ACT CCT CTT GG | *IAPlnc-flank* | GGA TCT GGT TGT CCG AGT GT<br>TCT GTT CCT GGC AAT CCT TC |
| *MusD* | GAT TGG TGG AAG TTT AGC TAG CAT<br>TAG CAT TCT CAT AAG CCA ATT GCA T | *MERV3-UTR* | AAC TAC AAA ACA ACA AAC AAA TAA GCA<br>AAA CTT GAC TTC TTA AAC CCA TTC TT |
| *ERVK10C-LTR* | GTG TGA GAC ACG CCT CTC CT<br>GGG AGA GCT TGA TTG CAG AG | *MERVlncRNA* | GTT TTG GAA GTT GCC TTG GA<br>CCC AAG AAC AGA AGC AGA GC |
| *ERVK10C-GAG* | TCA GGA TCA TGC TCA ACA GC<br>TGG CAT TGT GAG CCA ATC TA | *IAP3-UTR* | TTG AAG CCA GGT GCA GTA AC<br>TTC TGT TCC TGG CAA TCC TT |
| *ERVK10C-POL* | GCC ACC AGA GAC ATG GTT TT<br>CGG GCT TCT TTT CTT GTG AG | *IAPlncRNA* | TCA GGA TGT TGA GCC TGT TC<br>GGG TTT CCT AGG TGC TGA CA |
| *ERVK10C-ENV* | TAT CGC CTC AGG GTT AAT GC<br>TGG ATG CCA CAC AAC TCA TT | *IAP-Gene3UTR* | GGT GAA CTG CCT GGA AGA GA<br>TGG GGT CCT AGT CAC CTT TG |
| *IAP1-LTR* | TGT GCC AGG CAG TAA ACA AG<br>ACC AAT CAC CAC AGG TCA CA | *IAP-Gene3Inside* | TCA TGC CCA CCA TCT TGT AA<br>CGT GTT GGC ACC AGA TTC TT |

# Immunochemistry

*Immunocytochemistry*

Medium was rinsed off the cells with potassium-buffered PBS (kPBS) and fixed in 4% PFA (Sigma) for 10 minutes as previously described (Sachdeva et al., 2010; Thompson et al., 2005). After the protein fixation the cells were washed again with KPBS and incubated for 30 minutes in blocking solution containing (5 % normal serum, 0.25 % Triton X-100 and KPBS). In the following step the samples were incubated in blocking solution containing primary antibody at 4°C either over night or for 48 hours. The primary antibody incubation gives a better result with the prolonged incubation time. The primary antibody concentrations are found in Table 8.

The cells were rinsed in KPBS and incubated in fluorophore-conjugated secondary antibody (1:200 - Jackson Laboratories, 1:500 - Molecular Probes, 1:200 biotinylated antibody – Vector Laboratories) and DAPI (Sigma) for 30 minutes at room temperature in the dark. Samples that were incubated in biotinylated antibody were followed by an incubation of fluorphore-conjugated streptavidin for one hour at room temperature in the dark. The samples were rinsed and further kept in KPBS. The fluorescent images were acquired either using an inverted fluorescence microscope (Leica DFC360TX), or using the confocal microscope (Leica TCS SP8).

*Immunohistochemistry on mouse embryos*

Embryos were incubated at 4°C over night in 4 % formaldehyde solution (Sigma, Stock: 36.5 %) followed by incubation in 25 % sucrose-solution for 36 - 48 hours. The embryos were fixed and frozen in Tissue-tek (Sakura O.C.T™ COMPOUND) and sectioned on the cryostat (MICROM HM500M) in 14 μm coronal sections. The immunochemistry procedure was performed as described in the paragraph above (Immunocytochemistry).

**Table 6:** *Primary Antibody List*

| Antibody | Host | Dilution | Company |
|---|---|---|---|
| β–III-tubulin | Mouse | 1:1000 | Promega |
| GFAP | Rabbit | 1:1000 | DAKO |
| IAP-Gag | Mouse | 1:2000 | gift from Dr. Cullen |
| NESTIN | Mouse | 1:200 | BD 556309 |
| SOX 1 | rabbit | 1:100 | Cellsignalling |
| SOX 2 | mouse | 1:50 | RND systems |
| TRIM 28 | mouse | 1:1000 | Millipore |

## Behavioural Tests

For the behavioral tests *NestinCre*$^{(+/-)}$ *TRIM28*$^{(fl/wt)}$ males were mated with *TRIM28*$^{(fl/fl)}$ females. A total number of age-matched male and female offspring were used (16 heterozygous and 44 wild type mice were tested).

All animals, starting at 3 months of age, were exposed to the open field test and elevated plus maze as previously described in (Jakobsson et al., 2008), with an interval of one week in between the tests. All behavioral testing took place during day light cycle. To eliminate odor cues, each apparatus was thoroughly cleaned with 70 % ethanol and dried after each animal.

*Open Field – OF*

**Figure 17:** *Open field test*

Locomotion and reactivity to an open field was assessed in a white box (50 x 50 x 37 cm) under dimed and dispersed light conditions. The OF test was reported previously as standard test that evaluates locomotor activity consisting of a simple squared boxes where two adjacent sides of this square contain rows of beams. These beams form a coordinate system connected to a data processing computer. Each cube was examined for functionality prior starting the test; 8 tests were performed simultaneously. In this paradigm, each mouse was placed into the center of the field and allowed free locomotion during the 60 minutes test period. The total amount of line crossings of all four paws was captured throughout beam breaks (PASdata). Measures of total amount of line crossings are used as an index of activity.

*Elevated Plus Maze – EPM*

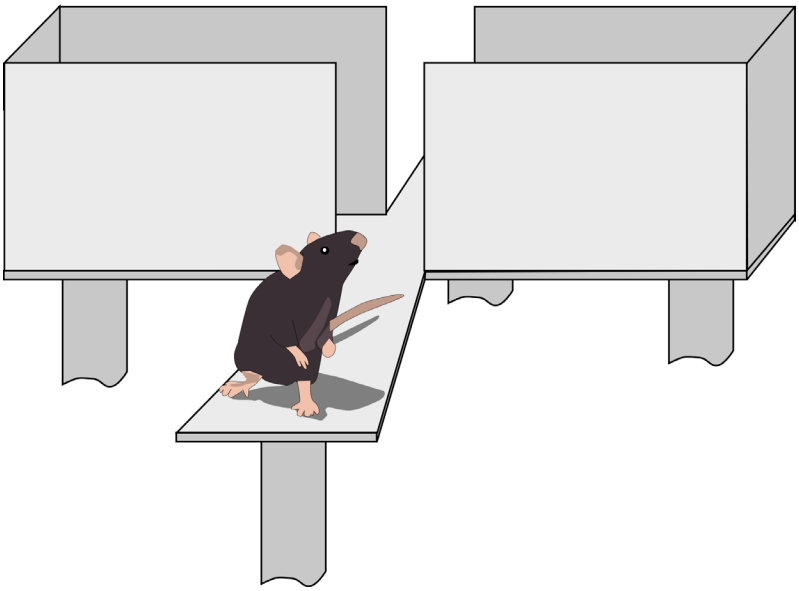


**Figure 18:** *EPM test*

To investigate the level of anxiety the EPM paradigm was performed, which was previously reported as a standardized test. The maze consists of two opposite open arms and two opposite closed arms (66 x 6 x 14 cm) arranged at 90° angles. The four arms are connected by a central platform (6 x 6 cm). The maze itself was elevated on a 70 cm translucent and stable platform under dimed and dispersed light conditions. In the beginning of the test mice were placed in the center of the maze being allowed to freely explore for 5 minutes. During that time, a video tracking software (Ethovision 3.1.16, Noldus) was recording the total distance moved, time spent in the center, open and closed arms, number and latency of entries to the open and closed arms. The total distance moved served as indicator of spontaneous locomotor activity, while the differences in times spent in the open compared to the closed arms are a measurement of anxiety.

For the statistical analysis, data from males and females as well as the two wild type genotypes *NestinCre*$^{(-/-)}$ *TRIM28*$^{(fl/wt)}$ and *NestinCre*$^{(-/-)}$ *TRIM28*$^{(fl/fl)}$ were pooled since we never found behavioral differences when comparing these groups.

# Acknowledgements

Doing a PhD is like a roller coaster ride – with all the ups and downs, turns and loops, thinking you would never come out in one piece and suddenly it´s over and you think " wow that was it ?? ". This section I would like to dedicate to those, without who I would have never been able to accomplish this ride. I would like to thank the ones who helped me to evolve scientifically; the ones who became extraordinary good friends and made me feel home during my stay in Sweden but also the precious people back in Austria. I would like to express my sincere gratitude to all of you!

First and foremost, I would like thank you **Johan**, for taking me into your team and for giving me the chance and freedom to evolve and develop over those past years. Your door was always open and you took time to discuss data und to help out at moments I got stuck. I´m very grateful for the trust you´ve put in me. But I don´t only want to take the chance to thank you for supervising me during the years of my PhD and for guiding me through this "funny project" but also for sharing the opinion about "Cow Girl Inn" being THE place to be in Santa Fe and convincing me that PBR should be my favorite beer. I will keep those trips in great memory. I wish you all the best for your future! **Malin**, thank you for being my co-supervisor over the past years. I´m not just grateful for your scientific input to my projects but I also appreciated our non-scientific talks and discussions and for the opportunity to talk to you when I needed advice. I also wish you all the best for your group.

**Anders**, thank you for all the scientific input to my projects. **Cilla**, thank you for the encouragements during thesis writing and the nice chats at the coffee table. **Eva**, thank you for all your administrative support during my first year of the PhD. **Paulina**, you are the biggest admin-talent I´ve ever seen. It feels like you have taken care of things before they even happened. Chapeau!! You´ve always been extremely helpful and patient - every time, any time and especially in times where I came running by your office panicking about last minute stuff. Over those last years you also became a very valuable friend to me! I would not only like to thank you for the hundreds of nonscientific talks and discussions we had over drinks on the terrace, which kept my mind cleared, but also for the fun trips to Lausanne/ Geneva and Paris.

None of this work could have been carried out in the same way without the help of our brilliant technical stuff. I am very grateful for all the support during the last years:

**Elsy**, you took me on a Masters trainee with basically no lab experience at all and you taught me all aspects of PCR with remarkable patience. It was a great pleasure to work side by side with you. Hardly anyone else can understand the excitement of having bands! I´m very grateful for your support regarding all mouse breedings and genotyping and the ability to keep an overview, especially when I lost it. **AnnaKarin**, for keeping the A10 cell lab organised and tidy and for being extremely helpful any time. I will always keep our cell lab chats in great memory, especially in times when I really

needed someone to talk to. **Ingar**, for all your help in the molecular lab and for refreshing laughs when sharing office. **Christina** "min kaffe kompis", it was a pleasure to share coffee weeks with you! Thank you for your support and help during my last experiments, I will truly never forget that! **Jenny J.** the "one-woman-AAV factory", for your help with the virus production and especially for rescuing me when I was fighting with the robot. **Ulla**, thank you for your help, for the nice chats in the histology lab and for the lovely decoration of the lab and lunch area. **Bengt**, I´m very grateful for your help formatting this thesis and making those beautiful illustrations. Your art is very inspiring to me. **Marie VP**, for your support when I needed to order stuff and for the nice chats in the lab. **Sol**, for organising the molecular lab. **Anna H.**, for all the help with FACS. **Micke** and **Anneli** for all the help in the animal house.

In the following paragraph I would like to thank the Jakobsson/Parmar/Björklund and Lundberg crowd:

**Rohit**, I learned a lot from you - scientifically and non-scientifically! First of all I will always be grateful for preventing that I got homeless. Thank you for all the great times at conferences, parties, road trips and for all the good memories. **Malin**, without you I´ve probably never have become "Swedish". Thanks for your trust by co-signing several contracts so I actually could use internet and phone and live a life in civilisation. I´m very thankful for joining me on the "cockroach hunt" and for being roomie at most of the conferences we´ve been too and the trips to NY. **Josephine**, thank you for introducing me into the world of Morris Water Maze and virus surgeries as well as for the nice chats, interesting discussions at the coffee table and the trip to Mexico. **Rebecca**, thanks for all the laughs and fun moments, especially when we seriously thought the song "My Pony is over the Ocean" is about a horse that died and for being brave enough to dine with me in public. **Per**, I´m very grateful for your help and support regarding bioinformatics; especially for the Alu manuscript. Thanks for all the good times at parties and for joining me drinking Palinka No.4 from the bottle. **Marie** - "it´s just another Margarita Monday", well we luckily only had one! Thanks for always being in a good mood, for encouragements during writing and for the fun trips to Santa Fe. Socializing at conferences has never been so much fun! Bazdmeg, **Karolina**!! I´m so happy you joined the lab. Your vivid personality was exactly what was missing here! You are the living prove of what´s needed to be a great scientist – a highly competent researcher paired with total madness! I´m very happy that I met **Peti** and you; both became very important to me! Thank you for the memories we have created starting with our first night out in Bishops/Arimans, which was absolutely "magical". I would like to thank my students **Emily**, **Christian** and **Anja** for their help in the lab and for challenging me.

**Ulrich** - " piiitschen, ja was macht denn da Uli?? ", thank you for joining me on that trip - *per aspera ad astra*. This journey would not be the same without you; I would have gotten lost somewhere on the way. I´m very grateful for: the countless talks & discussions, exchanges of music, singing of "Leck´s mi", office pranks, fun at the ISSCR meetings – especially the Maid of the Mist at the Niagara Falls and many more memorable moments – ESSEN!! In all these years you became like a brother to me, the voice of reason for many times - I owe you a big one! **Shane** – "Schnau Schnau", if I would have to sum up "the adventures of Shiana" this book would first of all extend the page limit and second - which is actually more problematic - I doubt I would have found a company who would agree on printing that content! When I first met you I was afraid of you, but that changed quite quickly – I think very much to the disadvantage of several people who actually wanted to enjoy their coffee breaks. Over all those years you became family for me, which included the invention of a new

language, regular after work shopping events at City Gross, food intakes like "Amazing A..`s Pachenga Curry" or "Mc Drive-Thru – Frappés & Bananas" and insulting car rides "Halloo!!". I still have not lost the hope that we will get a TV-show at some point, pfui deifel!! If there is one person who would drop everything to help in an emergency and drive you to the airport in Copenhagen, that´s **Andi**! I will truly never forget that. Thank you for the " let´s have 1!! beer " Friday after works, which usually ended only cause the bar was closing, the road trip to Paris, the deep conversations and the "hiding incognito in the bushes" with a coffee to go and Shane to make sure I was safe. **Olof**, I had the pleasure to sit between you and Ulrich most of the time during my PhD and I´ve truly enjoyed every minute! Your positive attitude was really contagious and your creative songs were always cheering me up! I truly miss the game "If you have to choose, who..". **Jenny N.W.**, thank you for all the nice chats we had while we were office mates and for the friendship that evolved. **Agnete**, I´m very greatful for your scientific input but also for all the fun and good conversations while having coffee breaks or drinks. All the best for your group! **Martin L.**, thank you for not only sharing the same sense of humor but also for having the same opinion about driving and parking skills. **Jenny S.**, Thank you for all the cat-sitting!

**Lu** - "ma chère amie", you literally walked with me - in all weathers - through the last parts of my PhD and with each step our friendship developed further. Thank you for all the encouraging chats and advices during the past two years, the countless parties and celebrations, the food and flower deliveries while writing this thesis and for basically everything! I will miss you a lot!

Further I would like to thank following people for creating a great working environment combined with nice chats and good laughs at coffee breaks or over drinks: **Daniela**, **Poonam**, **Maria**, **Shong**, **Janelle, Tiago, Daniella, Erika, Winnie, Luis, Thies, Monika, Sofia, Pina, Dan, Giuseppe, Sara, Marcus, Yolanda** and **Ylva**.

I´m also very grateful that I had the chance to meet so many lovely people at the BMC: **Natalie** – "Miss Jäger", meeting up with you is always a guarantee for fun either at parties or road trips. Thank you for joining me at the anti-stress painting sessions before writing this thesis and for always being there for me! **Zisis** – "der Kommissar", for all the good times at parties, **Tomas** – "Oh là là", for the convulsions of laughter and inappropriate jokes. **Ilknur**, for the good chats during the hours we worked side by side in the cell lab. **Martin W.**, for the advices regarding ChIP and nice conversations. I will never forget the fun at parties, midsommar celebrations, retreats or smoke/coffee breaks and therefore I would like to thank following people: **Sajedeh, Irene, Francesco, Barbara, Tim, Sanaz, Sofia, Rana, Krzysztof, Erik, Katja, Hélène, Gaby, Thea, Nina** and **Micke, Trine, Madde, Roshanak, Carolina, Hanna I., Hanna L., Elisabet, Itzia, Mehdi, Wen, Meike, Andy, Staffan** and **Marcus**.

I´m furthermore grateful that I met two lovely German ladies who joined A11: **Anna**-"Speidelman", thank you for all the good times we had during the stay for your Masters project, for making the effort to come over just for my birthday and for joining me driving like a crazy person on the German Autobahn. **Dani** – thanks for the fun at dinners, BBQs, nights out and joining me for crazy shopping.

I would also like to express my sincere gratitude to all the collaborators: **Didier Trono, Pilla, Mado, Gabriella, Florance Cammas** and **Patric Jern**.

I would like to thank all members of the **Yuan** and **Bellodi** - **groups** for the good discussions during journal clubs and also all the members of the **Kokaia** and **Ekdahl Clementson – groups** for all the great floor meetings and chats during lunch breaks.

A big thank you to my beloved friends from University in Vienna **Valerie** and **Peter D**., without those two I would not have ended up where I am. Valerie, thank you for being such a good friend and for checking on me while studying "herst geh bitte lern was – du stresst miiich!". I will always keep the great times at the University but also the ones we had here in Lund in great memory! I´m very lucky to have you as friends!

**Peter P.**, I know you will never forgive me, that I´ve chosen Lund over Hawaii! Thank you for being such a true friend, for all the hours we have spent on the phone chatting and for all the dinners you make when I´m back in Austria.

**Sirvart** - my childhood friend, I´m incredibly thankful for our true and long lasting friendship. Thank you for all the Skype sessions we made happen to keep each other updated, your encouragements during the past years, for always believing in me, for sending a care package while writing, for being there when I needed someone to talk to and for becoming a precious part of my life.

**Marion** – my biology teacher from high school, thank you for really challenging me during the biology lessons! Without you, I would have never discovered my passion for science. I´m very grateful, that we are still in touch over all those years.

Finally, I would like to give all my love to my parents for all their love and unlimited support. **Natascha** – thank you for being a good friend rather than a classical mum. I admire you for always seeing things in a positive way, for being permanently in a good mood, for happily joining for an adventurous trip to Las Vegas and for deep conversations while having a glass of Single Malt and a smoke. **Dieter** – thank you for painting this beautiful thesis cover, for making me laugh with your dark sense of humour, for the times we painted together, for helping me moving here within Lund including the crazy trip to IKEA followed by loads of fun during furniture assembling in the middle of the night and for reminding me that we stick together in rough times. Your career and commitment to work were always very inspiring to me, I wish one day I would just achieve half as much.

# References


Aguirre, A., Rubio, M.E., and Gallo, V. (2010). Notch and EGFR pathway interaction regulates neural stem cell number and self-renewal. Nature 467, 323-327.

Ahlenius, H., and Kokaia, Z. (2010). Isolation and generation of neurosphere cultures from embryonic and adult mouse brain. Methods Mol Biol 633, 241-252.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science 316, 744-747.

Baillie, J.K., Barnett, M.W., Upton, K.R., Gerhardt, D.J., Richmond, T.A., De Sapio, F., Brennan, P.M., Rizzu, P., Smith, S., Fell, M., et al. (2011). Somatic retrotransposition alters the genetic landscape of the human brain. Nature 479, 534-537.

Bannert, N., and Kurth, R. (2004). Retroelements and the human genome: new perspectives on an old relation. Proc Natl Acad Sci U S A 101 Suppl 2, 14572-14579.

Bao, W., Kojima, K.K., and Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob DNA 6, 11.

Belshaw, R., Dawson, A.L., Woolven-Allen, J., Redding, J., Burt, A., and Tristem, M. (2005). Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K(HML2): implications for present-day activity. J Virol 79, 12507-12514.

Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R.A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes Dev 23, 2484-2489.

Boeke, J.D. (1997). LINEs and Alus--the polyA connection. Nat Genet 16, 6-7.

Brouha, B., Schustak, J., Badge, R.M., Lutz-Prigge, S., Farley, A.H., Moran, J.V., and Kazazian, H.H., Jr. (2003). Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci U S A 100, 5280-5285.

Cammas, F., Mark, M., Dolle, P., Dierich, A., Chambon, P., and Losson, R. (2000). Mice lacking the transcriptional corepressor TIF1beta are defective in early postimplantation development. Development 127, 2955-2963.

Carmell, M.A., Girard, A., van de Kant, H.J., Bourc'his, D., Bestor, T.H., de Rooij, D.G., and Hannon, G.J. (2007). MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell 12, 503-514.

Carpenter, M.K., Cui, X., Hu, Z.Y., Jackson, J., Sherman, S., Seiger, A., and Wahlberg, L.U. (1999). In vitro expansion of a multipotent population of human neural progenitor cells. Exp Neurol 158, 265-278.

Carreira, P.E., Richardson, S.R., and Faulkner, G.J. (2014). L1 retrotransposons, cancer stem cells and oncogenesis. FEBS J 281, 63-73.

Castro-Diaz, N., Ecco, G., Coluccio, A., Kapopoulou, A., Yazdanpanah, B., Friedli, M., Duc, J., Jang, S.M., Turelli, P., and Trono, D. (2014). Evolutionally dynamic L1 regulation in embryonic stem cells. Genes Dev 28, 1397-1409.

Chuong, E.B., Rumi, M.A., Soares, M.J., and Baker, J.C. (2013). Endogenous retroviruses function as species-specific enhancer elements in the placenta. Nat Genet 45, 325-329.

Conti, L., Pollard, S.M., Gorba, T., Reitano, E., Toselli, M., Biella, G., Sun, Y., Sanzone, S., Ying, Q.L., Cattaneo, E., et al. (2005). Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. PLoS Biol 3, e283.

Cordaux, R., and Batzer, M.A. (2009). The impact of retrotransposons on human genome evolution. Nat Rev Genet 10, 691-703.

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Yeo, G.W., Mu, Y., Lovci, M.T., Morell, M., O'Shea, K.S., Moran, J.V., and Gage, F.H. (2009). L1 retrotransposition in human neural progenitor cells. Nature 460, 1127-1131.

Cowley, M., and Oakey, R.J. (2013). Transposable elements re-wire and fine-tune the transcriptome. PLoS Genet 9, e1003234.

Dewannieux, M., and Heidmann, T. (2013). Endogenous retroviruses: acquisition, amplification and taming of genome invaders. Curr Opin Virol 3, 646-656.

Douville, R., Liu, J., Rothstein, J., and Nath, A. (2011). Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis. Ann Neurol 69, 141-151.

Ellis, J., Hotta, A., and Rastegar, M. (2007). Retrovirus silencing by an epigenetic TRIM. Cell 131, 13-14.

Esnault, C., Maestre, J., and Heidmann, T. (2000). Human LINE retrotransposons generate processed pseudogenes. Nat Genet 24, 363-367.

Evrony, G.D., Cai, X., Lee, E., Hills, L.B., Elhosary, P.C., Lehmann, H.S., Parker, J.J., Atabay, K.D., Gilmore, E.C., Poduri, A., et al. (2012). Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. Cell 151, 483-496.

Falk, A., Koch, P., Kesavan, J., Takashima, Y., Ladewig, J., Alexander, M., Wiskow, O., Tailor, J., Trotter, M., Pollard, S., et al. (2012). Capture of neuroepithelial-like stem cells from pluripotent stem cells provides a versatile system for in vitro production of human neurons. PloS one 7, e29597.

Feng, Q., Moran, J.V., Kazazian, H.H., Jr., and Boeke, J.D. (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87, 905-916.

Feschotte, C. (2008). Transposable elements and the evolution of regulatory networks. Nat Rev Genet 9, 397-405.

Feschotte, C., and Pritham, E.J. (2007). DNA transposons and the evolution of eukaryotic genomes. Annu Rev Genet 41, 331-368.

Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., and Rauscher, F.J., 3rd (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev 10, 2067-2078.

Gage, F.H. (2000). Mammalian neural stem cells. Science 287, 1433-1438.

Garcia-Perez, J.L., Marchetto, M.C., Muotri, A.R., Coufal, N.G., Gage, F.H., O'Shea, K.S., and Moran, J.V. (2007). LINE-1 retrotransposition in human embryonic stem cells. Hum Mol Genet 16, 1569-1577.

Graham, V., Khudyakov, J., Ellis, P., and Pevny, L. (2003). SOX2 functions to maintain neural progenitor identity. Neuron 39, 749-765.

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., et al. (2015). Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. Nature 522, 221-225.

Hancks, D.C., and Kazazian, H.H., Jr. (2010). SVA retrotransposons: Evolution and genetic instability. Semin Cancer Biol 20, 234-245.

Hancks, D.C., and Kazazian, H.H., Jr. (2012). Active human retrotransposons: variation and disease. Curr Opin Genet Dev 22, 191-203.

Helman, E., Lawrence, M.S., Stewart, C., Sougnez, C., Getz, G., and Meyerson, M. (2014). Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. Genome Res 24, 1053-1063.

Herquel, B., Ouararhni, K., Martianov, I., Le Gras, S., Ye, T., Keime, C., Lerouge, T., Jost, B., Cammas, F., Losson, R., et al. (2013). Trim24-repressed VL30 retrotransposons regulate gene expression by producing noncoding RNA. Nat Struct Mol Biol 20, 339-346.

Hua-Van, A., Le Rouzic, A., Boutin, T.S., Filee, J., and Capy, P. (2011). The struggle for life of the genome's selfish architects. Biol Direct 6, 19.

Huang, C.R., Burns, K.H., and Boeke, J.D. (2012). Active transposition in genomes. Annu Rev Genet 46, 651-675.

Hunter, R.G., Murakami, G., Dewell, S., Seligsohn, M., Baker, M.E., Datson, N.A., McEwen, B.S., and Pfaff, D.W. (2012). Acute stress and hippocampal histone H3 lysine 9 trimethylation, a retrotransposon silencing response. Proc Natl Acad Sci U S A 109, 17657-17662.

Hutnick, L.K., Huang, X., Loo, T.C., Ma, Z., and Fan, G. (2010). Repression of retrotransposal elements in mouse embryonic stem cells is primarily mediated by a DNA methylation-independent mechanism. J Biol Chem 285, 21082-21091.

Isbel, L., and Whitelaw, E. (2012). Endogenous retroviruses in mammals: an emerging picture of how ERVs modify expression of adjacent genes. Bioessays 34, 734-738.

Ivanov, A.V., Peng, H., Yurchenko, V., Yap, K.L., Negorev, D.G., Schultz, D.C., Psulkowski, E., Fredericks, W.J., White, D.E., Maul, G.G., et al. (2007). PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. Mol Cell 28, 823-837.

Jakobsson, J., Cordero, M.I., Bisaz, R., Groner, A.C., Busskamp, V., Bensadoun, J.C., Cammas, F., Losson, R., Mansuy, I.M., Sandi, C., et al. (2008). KAP1-mediated epigenetic repression in the forebrain modulates behavioral vulnerability to stress. Neuron 60, 818-831.

Jeong, B.H., Lee, Y.J., Carp, R.I., and Kim, Y.S. (2010). The prevalence of human endogenous retroviruses in cerebrospinal fluids from patients with sporadic Creutzfeldt-Jakob disease. J Clin Virol 47, 136-142.

Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 110, 462-467.

Kannan, S., Chernikova, D., Rogozin, I.B., Poliakov, E., Managadze, D., Koonin, E.V., and Milanesi, L. (2015). Transposable Element Insertions in Long Intergenic Non-Coding RNA Genes. Front Bioeng Biotechnol 3, 71.

Kano, H., Godoy, I., Courtney, C., Vetter, M.R., Gerton, G.L., Ostertag, E.M., and Kazazian, H.H., Jr. (2009). L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. Genes Dev 23, 1303-1312.

Karimi, M.M., Goyal, P., Maksakova, I.A., Bilenky, M., Leung, D., Tang, J.X., Shinkai, Y., Mager, D.L., Jones, S., Hirst, M., et al. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. Cell Stem Cell 8, 676-687.

Karlsson, H., Bachmann, S., Schroder, J., McArthur, J., Torrey, E.F., and Yolken, R.H. (2001). Retroviral RNA identified in the cerebrospinal fluids and brains of individuals with schizophrenia. Proc Natl Acad Sci U S A 98, 4634-4639.

Kazazian, H.H., Jr., Wong, C., Youssoufian, H., Scott, A.F., Phillips, D.G., and 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.

Kunarso, G., Chia, N.Y., Jeyakani, J., Hwang, C., Lu, X., Chan, Y.S., Ng, H.H., and Bourque, G. (2010). Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nat Genet 42, 631-634.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T.W., et al. (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev 22, 908-917.

Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L.J., 3rd, Lohr, J.G., Harris, C.C., Ding, L., Wilson, R.K., et al. (2012). Landscape of somatic retrotransposition in human cancers. Science 337, 967-971.

Li, J., Akagi, K., Hu, Y., Trivett, A.L., Hlynialuk, C.J., Swing, D.A., Volfovsky, N., Morgan, T.C., Golubeva, Y., Stephens, R.M., et al. (2012a). Mouse endogenous retroviruses can trigger premature transcriptional termination at a distance. Genome Res 22, 870-884.

Li, W., Jin, Y., Prazak, L., Hammell, M., and Dubnau, J. (2012b). Transposable elements in TDP-43-mediated neurodegenerative disorders. PLoS One 7, e44099.

Lu, X., Sachs, F., Ramsay, L., Jacques, P.E., Goke, J., Bourque, G., and Ng, H.H. (2014). The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. Nat Struct Mol Biol 21, 423-425.

Matsui, T., Leung, D., Miyashita, H., Maksakova, I.A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M.C., and Shinkai, Y. (2010). Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature 464, 927-931.

Mills, R.E., Bennett, E.A., Iskow, R.C., and Devine, S.E. (2007). Which transposable elements are active in the human genome? Trends Genet 23, 183-191.

Muotri, A.R., Chu, V.T., Marchetto, M.C., Deng, W., Moran, J.V., and Gage, F.H. (2005). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature 435, 903-910.

Muotri, A.R., Marchetto, M.C., Coufal, N.G., and Gage, F.H. (2007). The necessary junk: new functions for transposable elements. Hum Mol Genet 16 Spec No. 2, R159-167.

Muotri, A.R., Marchetto, M.C., Coufal, N.G., Oefner, R., Yeo, G., Nakashima, K., and Gage, F.H. (2010). L1 retrotransposition in neurons is modulated by MeCP2. Nature 468, 443-446.

Muotri, A.R., Zhao, C., Marchetto, M.C., and Gage, F.H. (2009). Environmental influence on L1 retrotransposons in the adult hippocampus. Hippocampus 19, 1002-1007.

Peng, H., Begg, G.E., Harper, S.L., Friedman, J.R., Speicher, D.W., and Rauscher, F.J., 3rd (2000). Biochemical analysis of the Kruppel-associated box (KRAB) transcriptional repression domain. J Biol Chem 275, 18000-18010.

Perrat, P.N., DasGupta, S., Wang, J., Theurkauf, W., Weng, Z., Rosbash, M., and Waddell, S. (2013). Transposition-driven genomic heterogeneity in the Drosophila brain. Science 340, 91-95.

Pi, W., Zhu, X., Wu, M., Wang, Y., Fulzele, S., Eroglu, A., Ling, J., and Tuan, D. (2010). Long-range function of an intergenic retrotransposon. Proc Natl Acad Sci U S A 107, 12992-12997.

Qin, C., Wang, Z., Shang, J., Bekkari, K., Liu, R., Pacchione, S., McNulty, K.A., Ng, A., Barnum, J.E., and Storer, R.D. (2010). Intracisternal A particle genes: Distribution in the mouse genome, active subtypes, and potential roles as species-specific mediators of susceptibility to cancer. Mol Carcinog 49, 54-67.

Quenneville, S., Turelli, P., Bojkowska, K., Raclot, C., Offner, S., Kapopoulou, A., and Trono, D. (2012). The KRAB-ZFP/KAP1 system contributes to the early embryonic establishment of site-specific DNA methylation patterns maintained during development. Cell Rep 2, 766-773.

Reichmann, J., Crichton, J.H., Madej, M.J., Taggart, M., Gautier, P., Garcia-Perez, J.L., Meehan, R.R., and Adams, I.R. (2012). Microarray analysis of LTR retrotransposon silencing identifies Hdac1 as a regulator of retrotransposon expression in mouse embryonic stem cells. PLoS computational biology 8, e1002486.

Reilly, M.T., Faulkner, G.J., Dubnau, J., Ponomarev, I., and Gage, F.H. (2013). The role of transposable elements in health and diseases of the central nervous system. J Neurosci 33, 17577-17586.

Ribet, D., Harper, F., Dupressoir, A., Dewannieux, M., Pierron, G., and Heidmann, T. (2008). An infectious progenitor for the murine IAP retrotransposon: emergence of an intracellular genetic parasite from an ancient retrovirus. Genome Res 18, 597-609.

Robbez-Masson, L., and Rowe, H.M. (2015). Retrotransposons shape species-specific embryonic stem cell gene expression. Retrovirology 12, 45.

Rowe, H.M., Friedli, M., Offner, S., Verp, S., Mesnard, D., Marquis, J., Aktas, T., and Trono, D. (2013). De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET. Development 140, 519-529.

Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P.V., Layard-Liesching, H., Verp, S., Marquis, J., et al. (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. Nature 463, 237-240.

Rowe, H.M., and Trono, D. (2011). Dynamic control of endogenous retroviruses during development. Virology 411, 273-287.

Ryan, F.P. (2004). Human endogenous retroviruses in health and disease: a symbiotic perspective. J R Soc Med 97, 560-565.

Sachdeva, R., Jonsson, M.E., Nelander, J., Kirkeby, A., Guibentif, C., Gentner, B., Naldini, L., Bjorklund, A., Parmar, M., and Jakobsson, J. (2010). Tracking differentiating neural progenitors in pluripotent cultures using microRNA-regulated lentiviral vectors. Proc Natl Acad Sci U S A 107, 11602-11607.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev 16, 919-932.

Solyom, S., Ewing, A.D., Rahrmann, E.P., Doucet, T., Nelson, H.H., Burns, M.B., Harris, R.S., Sigmon, D.F., Casella, A., Erlanger, B., et al. (2012). Extensive somatic L1 retrotransposition in colorectal tumors. Genome Res 22, 2328-2338.

Sperber, G.O., Airola, T., Jern, P., and Blomberg, J. (2007). Automated recognition of retroviral sequences in genomic data--RetroTector. Nucleic acids research 35, 4964-4976.

Sripathy, S.P., Stevens, J., and Schultz, D.C. (2006). The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. Mol Cell Biol 26, 8623-8638.

Sundaram, V., Cheng, Y., Ma, Z., Li, D., Xing, X., Edge, P., Snyder, M.P., and Wang, T. (2014). Widespread contribution of transposable elements to the innovation of gene regulatory networks. Genome Res 24, 1963-1976.

Taupin, P., and Gage, F.H. (2002). Adult neurogenesis and neural stem cells of the central nervous system in mammals. J Neurosci Res 69, 745-749.

Thompson, L., Barraud, P., Andersson, E., Kirik, D., and Bjorklund, A. (2005). Identification of dopaminergic neurons of nigral and ventral tegmental area subtypes in grafts of fetal ventral mesencephalon based on cell morphology, protein expression, and efferent projections. J Neurosci 25, 6467-6477.

Torrado, E.F., Gomes, C., Santos, G., Fernandes, A., Brites, D., and Falcao, A.S. (2014). Directing mouse embryonic neurosphere differentiation toward an enriched neuronal population. Int J Dev Neurosci 37, 94-99.

Tronche, F., Kellendonk, C., Kretz, O., Gass, P., Anlag, K., Orban, P.C., Bock, R., Klein, R., and Schutz, G. (1999). Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety. Nature genetics 23, 99-103.

Tsirigos, A., and Rigoutsos, I. (2009). Alu and b1 repeats have been selectively retained in the upstream and intronic regions of genes of specific functional classes. PLoS Comput Biol 5, e1000610.

Tubio, J.M., Li, Y., Ju, Y.S., Martincorena, I., Cooke, S.L., Tojo, M., Gundem, G., Pipinikas, C.P., Zamora, J., Raine, K., et al. (2014). Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. Science 345, 1251343.

Turelli, P., Castro-Diaz, N., Marzetta, F., Kapopoulou, A., Raclot, C., Duc, J., Tieng, V., Quenneville, S., and Trono, D. (2014). Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements. Genome Res 24, 1260-1270.

Upton, K.R., Gerhardt, D.J., Jesuadian, J.S., Richardson, S.R., Sanchez-Luque, F.J., Bodea, G.O., Ewing, A.D., Salvador-Palomeque, C., van der Knaap, M.S., Brennan, P.M., et al. (2015). Ubiquitous L1 mosaicism in hippocampal neurons. Cell 161, 228-239.

Urrutia, R. (2003). KRAB-containing zinc-finger repressor proteins. Genome Biol 4, 231.

Van Meter, M., Kashyap, M., Rezazadeh, S., Geneva, A.J., Morello, T.D., Seluanov, A., and Gorbunova, V. (2014). SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. Nat Commun 5, 5011.

Varshney, D., Vavrova-Anderson, J., Oler, A.J., Cowling, V.H., Cairns, B.R., and White, R.J. (2015). SINE transcription by RNA polymerase III is suppressed by histone methylation but not by DNA methylation. Nat Commun 6, 6569.

Waddell, S., Barnstedt, O., and Treiber, C. (2014). Neural transposition in the Drosophila brain: is it all bad news? Adv Genet 86, 65-92.

Wang, C., Ivanov, A., Chen, L., Fredericks, W.J., Seto, E., Rauscher, F.J., 3rd, and Chen, J. (2005a). MDM2 interaction with nuclear corepressor KAP1 contributes to p53 inactivation. EMBO J 24, 3279-3290.

Wang, H., Xing, J., Grover, D., Hedges, D.J., Han, K., Walker, J.A., and Batzer, M.A. (2005b). SVA elements: a hominid-specific retroposon family. J Mol Biol 354, 994-1007.

Weber, P., Cammas, F., Gerard, C., Metzger, D., Chambon, P., Losson, R., and Mark, M. (2002). Germ cell expression of the transcriptional co-repressor TIF1beta is required for the maintenance of spermatogenesis in the mouse. Development 129, 2329-2337.

Weiss, R.A. (2006). The discovery of endogenous retroviruses. Retrovirology 3, 67.

Wiznerowicz, M., Jakobsson, J., Szulc, J., Liao, S., Quazzola, A., Beermann, F., Aebischer, P., and Trono, D. (2007). The Kruppel-associated box repressor domain can trigger de novo promoter methylation during mouse early embryogenesis. J Biol Chem 282, 34535-34541.

Wolf, G., Yang, P., Fuchtbauer, A.C., Fuchtbauer, E.M., Silva, A.M., Park, C., Wu, W., Nielsen, A.L., Pedersen, F.S., and Macfarlan, T.S. (2015). The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses. Genes Dev 29, 538-554.

Xing, J., Wang, H., Belancio, V.P., Cordaux, R., Deininger, P.L., and Batzer, M.A. (2006). Emergence of primate genes by retrotransposon-mediated sequence transduction. Proc Natl Acad Sci U S A 103, 17608-17613.

Xing, J., Zhang, Y., Han, K., Salem, A.H., Sen, S.K., Huff, C.D., Zhou, Q., Kirkness, E.F., Levy, S., Batzer, M.A., et al. (2009). Mobile elements create structural variation: analysis of a complete human genome. Genome Res 19, 1516-1526.

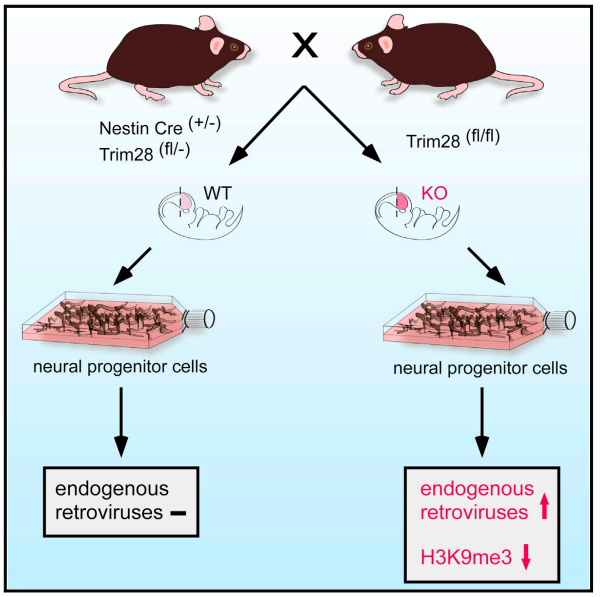
# Paper I

Report

Cell Reports

# TRIM28 Represses Transcription of Endogenous Retroviruses in Neural Progenitor Cells

## Graphical Abstract




## Highlights

- Deletion of TRIM28 in NPCs results in transcriptional activation of ERVs
- ERVs are marked by H3K9me3 in NPCs, which is lost upon TRIM28 deletion
- Activation of ERVs in NPCs influences expression levels of nearby genes
- Activation of ERVs in NPCs results in the production of long noncoding RNAs

## Authors

Liana Fasching, Adamandia Kapopoulou, ..., Didier Trono, Johan Jakobsson

## Correspondence

johan.jakobsson@med.lu.se

## In Brief

Endogenous retroviruses, which make up almost 10% of the mouse genome, are thought to be transcriptionally silenced by DNA methylation in adult tissues. Fasching et al. now show that endogenous retroviruses are controlled by TRIM28-mediated histone modifications in neural progenitor cells, suggesting a role for these elements in the control of transcriptional dynamics in the brain.

## Accession Numbers

GSE45930

Fasching et al., 2015, Cell Reports *10*, 20–28
January 6, 2015 ©2015 The Authors
http://dx.doi.org/10.1016/j.celrep.2014.12.004

# TRIM28 Represses Transcription of Endogenous Retroviruses in Neural Progenitor Cells

Liana Fasching,[1] Adamandia Kapopoulou,[2] Rohit Sachdeva,[1] Rebecca Petri,[1] Marie E. Jönsson,[1] Christian Männe,[1] Priscilla Turelli,[2] Patric Jern,[3] Florence Cammas,[4] Didier Trono,[2] and Johan Jakobsson[1,*]
[1]Laboratory of Molecular Neurogenetics, Department of Experimental Medical Science, Wallenberg Neuroscience Center and Lund Stem Cell Center, Lund University, 221 84 Lund, Sweden
[2]School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland
[3]Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, 751 23 Uppsala, Sweden
[4]IRCM, Institut de Recherche en Cancérologie de Montpellier, INSERM, U896, Université Montpellier; Institut Régional du Cancer Montpellier, Montpellier 34298, France
*Correspondence: johan.jakobsson@med.lu.se
http://dx.doi.org/10.1016/j.celrep.2014.12.004
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

## SUMMARY

**TRIM28 is a corepressor that mediates transcriptional silencing by establishing local heterochromatin. Here, we show that deletion of TRIM28 in neural progenitor cells (NPCs) results in high-level expression of two groups of endogenous retroviruses (ERVs): *IAP1* and *MMERVK10C*. We find that NPCs use TRIM28-mediated histone modifications to dynamically regulate transcription and silencing of ERVs, which is in contrast to other somatic cell types using DNA methylation. We also show that derepression of ERVs influences transcriptional dynamics in NPCs through the activation of nearby genes and the expression of long noncoding RNAs. These findings demonstrate a unique dynamic transcriptional regulation of ERVs in NPCs. Our results warrant future studies on the role of ERVs in the healthy and diseased brain.**

## INTRODUCTION

The mammalian brain is an extremely complex organ harboring more than a thousand different types of neurons that serve a wide variety of functions. How this complexity is achieved remains largely unknown. However, epigenetic mechanisms such as DNA methylation, histone modification, and noncoding RNAs are thought to be important in establishing a high diversity of gene expression from the same template, leading to a spatial pattern of transcription. How distinct transcriptional programs are established in different neuronal populations remains poorly understood, but one interesting recently proposed hypothesis suggests transposable elements (TEs) to be involved in this process (Muotri et al., 2007; Reilly et al., 2013). TEs are repetitive mobile genetic elements that were originally considered to be parasitic DNA without any function, popularly termed "junk DNA." Today, it is becoming increasingly clear that TEs can act as gene regulatory elements by serving as hubs for chromatin modifications and by acting as transcriptional start sites for noncoding RNAs. Consequently, TEs are very well suited to influence gene expression and may play an important role in controlling and fine-tuning gene networks in the brain (Jern and Coffin, 2008; Cowley and Oakey, 2013).

Retroviruses are found in most vertebrates and can transform their genetic material and integrate into the host genome as proviruses to produce new viruses. Occasionally, retroviruses infect germline cells allowing the integrated proviruses to be passed on to the offspring as an endogenous retrovirus (ERV). Around 8%–10% of the human and mouse genome are composed of this type of TE, and, despite up to millions of years since their integration in host germline, many ERVs contain sequences that can serve as transcriptional start sites or as *cis*-acting regulatory elements in the host genomes (Jern and Coffin, 2008). The large amount of ERVs in mammalian genomes suggest that they play important roles in the host organisms, for instance, by influencing gene regulatory networks (Kunarso et al., 2010; Feschotte, 2008), but ERVs have also been linked to diseases. In humans, aberrant expression of ERVs has been found in both cancer and brain disorders, although causality remains to be established (Jern and Coffin, 2008; Douville et al., 2011). Thus, ERVs may contribute both beneficial and detrimental effects, which have been balanced throughout evolution, to the host organism.

ERVs are silenced during the first few days of embryogenesis by TRIM28 (tripartite motif-containing protein 28, also known as KAP1 or TIF1beta), a transcriptional corepressor essential for early mouse development (Cammas et al., 2000; Rowe et al., 2010). During the extensive genome reprogramming that takes place at this period, TRIM28 is recruited to ERVs via sequence-specific Krüppel-associated box zinc-finger proteins (KRAB-ZFPs), a family of transcription factors that has undergone a rapid expansion in mammalian genomes in parallel with the expansion of ERVs (Wolf and Goff, 2009; Thomas and Schneider, 2011). TRIM28 then induces repressive histone modifications by recruiting multiprotein complexes including the H3K9 methyltransferase SETDB1 (also known as ESET), the histone deacetylase-containing NuRD complex, and heterochromatin protein 1 (HP1) (Schultz et al., 2002; Sripathy et al.,

2006). Deletion of *Trim28* or *Setdb1* in ESCs leads to loss of the H3K9me3-mark at ERVs, resulting in transcriptional activation of these elements (Matsui et al., 2010; Rowe et al., 2010).

However, KRAB-ZFP/TRIM28 histone-based repression of ERVs rapidly gives place to a more permanent silencing mechanism, as the TRIM28-mediated recruitment of de novo DNA methyltransferases leads to cytosine methylation at CpG dinucleotides (Ellis et al., 2007; Wiznerowicz et al., 2007; Rowe and Trono, 2011). The maintenance DNA methyltransferase complex then ensures that DNA methylation is maintained, alleviating the need for sequence-specific KRAB-ZFPs and TRIM28. In mouse embryonic fibroblasts as well as in all adult tissues examined so far, TRIM28 depletion has no impact on ERV silencing, which is instead released by drugs such as 5-azacytidine or by deletion of DNA methyltransferases (Jackson-Grusby et al., 2001; Hutnick et al., 2010).

DNA methylation has long been considered as a stable epigenetic mark resulting in maintenance of DNA-methylation patterns throughout the lifespan of an organism. However, several recent studies demonstrate a unique dynamic regulation of DNA-methylation patterns in the brain (Sweatt, 2013). There is also evidence that retroelements and transposons are highly active during brain development and in neural progenitor cells (NPCs) (Muotri et al., 2005, 2010; Baillie et al., 2011; Evrony et al., 2012; Li et al., 2013; Perrat et al., 2013). For example, *LINE-1* elements have been found to be transcriptionally active and to retrotranspose in NPCs (Muotri et al., 2005, 2010; Coufal et al., 2009). In addition, we have previously found that deletion of TRIM28 in postmitotic forebrain neurons results in complex behavioral alterations, including vulnerability to stress (Jakobsson et al., 2008). In the present work, we demonstrate that NPCs use TRIM28-mediated histone modifications to dynamically regulate the transcription and silencing of ERVs, rather than the DNA methylation at play in other somatic tissues. We also unveil that derepression of ERVs influences transcriptional dynamics in NPCs, by activating nearby genes and the expression of long noncoding RNAs (lncRNAs).

## RESULTS

### TRIM28-Deficient NPCs Express High Levels of ERVs

To investigate if TRIM28 contributes to ERV silencing in NPCs we established *Trim28*-deficient NPC cultures. We crossed transgenic *NestinCre* mice (Tronche et al., 1999) with mice carrying floxed *Trim28*-alleles ($Trim28^{fl/fl}$) (Weber et al., 2002), resulting in excision of *Trim28* in neural progenitors at the time when Nestin-expression is initiated, starting around embryonic day 10 (E10). At E13.5, we collected embryos, dissected the forebrain, and established NPC cultures from individual embryos (Figures 1A and 1B). We confirmed the deletion of *Trim28* by genotyping for the excised allele and by verifying the absence of TRIM28 protein (Figures 1C and 1D). We collected RNA from $Trim28^{-/-}$ NPCs and wild-type controls and performed RNA extraction followed by deep sequencing (RNA-seq). The resulting reads were mapped against reference sequences from Repbase, a database containing consensus sequences for known repetitive elements (Jurka et al., 2005). We found that several ERVs were highly upregulated in $Trim28^{-/-}$ NPCs, including, e.g., *Mus musculus ERV using* $tRNA^{Lys}$ *type 10C* (*MMERVK10C*) and *intracisternal A-particles class 1* (*IAP1*) (Figure 1E; Tables S1 and S2). Other retroelements such as *MusD* and *LINE-1* were modestly upregulated, whereas several other types of common repetitive elements were unaffected (Figure 1E; Tables S1 and S2).

We confirmed increased transcription of *MMERVK10C* and *IAP1* elements using quantitative RT-PCR (qRT-PCR) (Figure 1F). In contrast, when we used primer pairs designed to recognize the consensus sequence of the entire *IAP*-family, including more ancient *IAP* elements, we detected only a modest upregulation (Figure 1F). This finding is in line with the results of the RNA-seq, which indicated that only certain types of IAP elements were upregulated in $Trim28^{-/-}$ NPCs. Also in agreement with the RNA-seq, qRT-PCR analyses indicated that deletion of *Trim28* in NPCs only modestly increased the expression of other retroelements such as *LINE-1* or *MusD* (Figure 1F). We confirmed these results in cultures derived from two separate embryos (data not shown).

$Trim28^{-/-}$ NPCs proliferated at a similar rate compared to cells generated from wild-type and heterozygous siblings and could be expanded for more than 60 passages. However, we observed that $Trim28^{-/-}$ NPCs were growing in dense cluster-like formations, which seemed to attach less to the flask surface compared to the wild-type control. $Trim28^{-/-}$ NPCs could also be differentiated to both neurons and astrocytes suggesting that TRIM28 has no major influence on the self-renewal and differentiation of NPCs (Figures 1G and 1H).

### *MMERVK10C* Elements Are Controlled by TRIM28

The RNA-seq analysis indicated that *MMERVK10C* elements were among the most upregulated ERVs following *Trim28*-deletion in NPCs. *MMERVK10C* is a beta-like ERV similar to *HERVK* (*HML2*), one of the most recent ERVs to invade the human genome (Reichmann et al., 2012) (Belshaw et al., 2005). *MMERVK10C* sequences flanked by *RLTR10C* make up putative proviral sequences of around 8.4 kb. In the mouse genome, *MMERVK10C* is present in a few complete provirus loci (~20) and more than 1,000 incomplete loci (Reichmann et al., 2012). We performed sequence analysis of the *MMERVK10C* provirus for the presence of retroviral features using the RetroTector software (Sperber et al., 2007). Based on this analysis, we designed primers recognizing the *LTRs*, *gag*, *pol*, and *env* of the *MMERVK10C* provirus and investigated expression levels in $Trim28^{-/-}$ NPCs (schematics in Figure 2A). We found that transcripts over the entire region of the provirus were increased, including a massive expression of *env* sequences when compared to wild-type controls (170-fold; Figure 2B).

Ascertaining that the ERV induction observed in NPCs isolated from $Trim28^{-/-}$ animals was not secondary to more general developmental anomalies, knocking down TRIM28 in wild-type NPCs by lentivector-mediated RNA interference led to a marked upregulation of these retroelements (Figure 2C). Furthermore, increased ERV expression was detected in forebrain tissue from $Trim28^{-/-}$ embryos (Figure 2D).

In ESCs, TRIM28 controls ERV expression via histone modifications including H3K9 trimethylation (Rowe et al., 2010), whereas it is DNA methylation that instead prevails in somatic

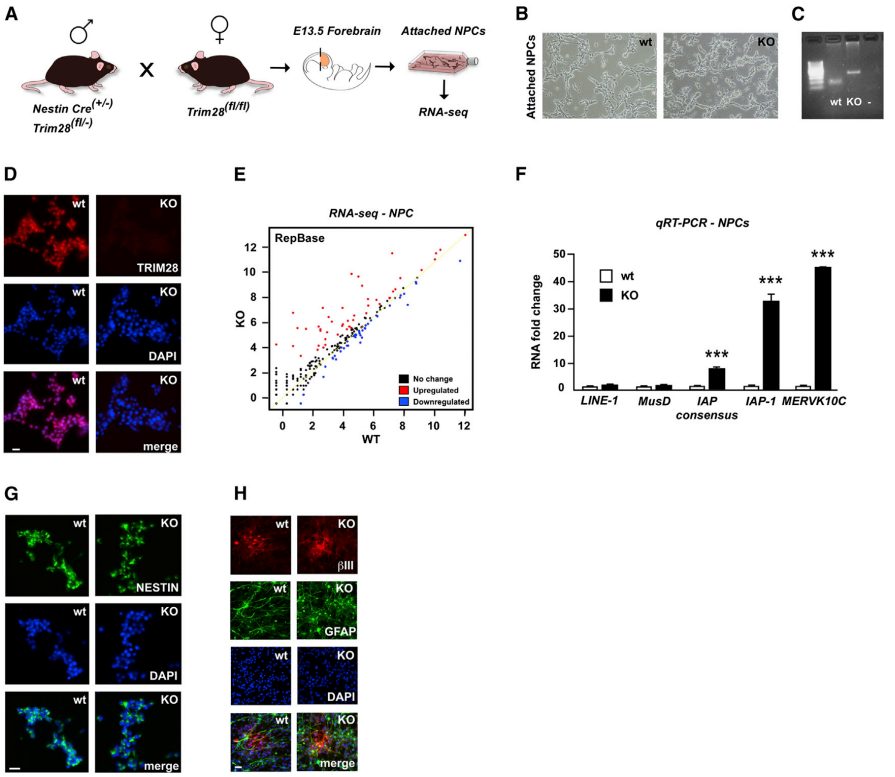



**Figure 1. Establishment of *Trim28*-Deficient Neural Progenitor Cultures**

(A) Illustration of the experimental approach.

(B) Representative images of early passage *Trim28*$^{-/-}$ NPCs.

(C) PCR analysis of genomic DNA from wild-type and *Trim28*$^{-/-}$ NPCs demonstrates the presence of the 152 and 290 bp products corresponding to *loxP*-flanked or excised *Trim28*, respectively.

(D) Verification of a complete lack of TRIM28 protein via immunocytochemistry.

(E) RNA-seq analysis. The graph shows KO samples plotted versus wild-type samples, where each dot represents a Repbase sequence.

(F) qRT-PCR of RNA isolated from wild-type and *Trim28*$^{-/-}$ NPCs.

(G) *Trim28*-deficient NPCs display a homogenous expression of NESTIN.

(H) Immunofluorescent analysis of differentiated NPCs.

Data are presented as mean of relative values ± SEM. **p < 0.01, ***p < 0.001, Student's t test. Scale bars, 200 (A) and 50 (B) μm. See also Tables S1 and S2.

tissues. In NPCs, we found that the *MMERVK10C* provirus was enriched in H3K9me3, and that this repressive mark was markedly reduced in *Trim28*$^{-/-}$ NPCs (Figure 2E).

Because *MMERVK10C* appeared to be under TRIM28 control in NPCs, we hypothesized that at least a proportion of these retroelements escaped DNA methylation in these cells. To probe this issue, we examined the DNA methylation status of full-length *MMERVK10C*, which were among the most highly upregulated retroelements in *Trim28*$^{-/-}$ NPCs. Bisulfite sequencing of a CpG-island located in the 3′ region of *MMERVK10C* revealed several clones with some unmethylated CpGs (17% unmethylated CpGs, Figure 2F) in NPCs, whereas this region was almost fully methylated in DNA extracted from mouse tail (7% unmethylated CpGs, Figure 2F, Fisher's exact test one-sided $p < 0.05$). Moreover, we found no difference in the level of CpG methylation between wild-type and *Trim28*$^{-/-}$ NPCs. In summary, these data suggest that a proportion of the *MMERVK10C* elements are spared from undergoing DNA methylation specifically in NPCs during early development.

## Increased Expression of *IAP1* Results in ERV-Derived Protein Expression

*IAP1* elements, which lose H3K9me3 marks and were also highly upregulated in *Trim28*$^{-/-}$ NPCs (Figures 3A and 3B), are

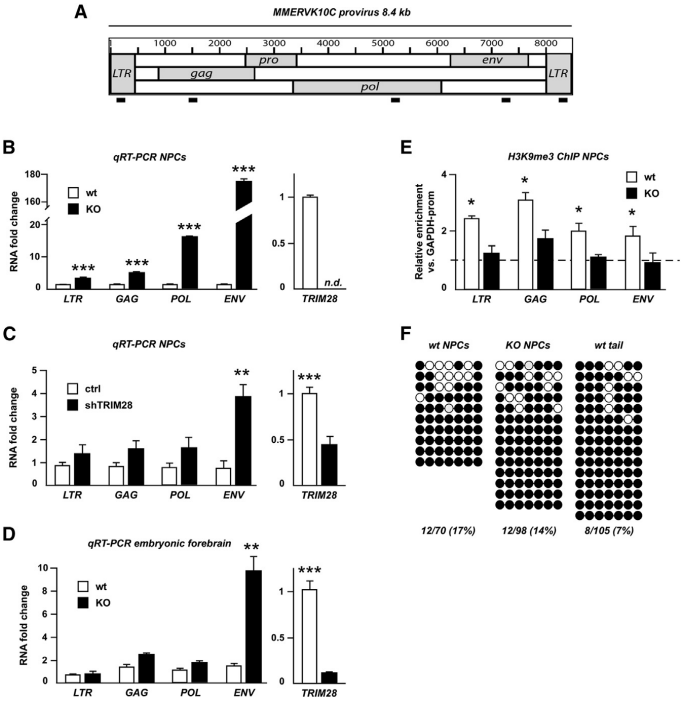



**Figure 2. Analysis of the Putative *MMERVK10C* Provirus**

(A) Schematic drawing of the *MMERVK10C* provirus and approximate primer positions.

(B) Quantitative analysis of transcript levels of different regions of the *MMERVK10C* provirus in *Trim28*$^{-/-}$ and wild-type NPCs.

(C) qRT-PCR analysis of *MMERVK10C* following TRIM28-shRNA knockdown.

(D) qRT-PCR analysis of E13.5 forebrain dissected from intercrosses of *NestinCre Trim28*$^{floxed}$ mice.

(E) ChIP for H3K9me3 in *Trim28*$^{-/-}$ and wild-type NPCs.

(F) Bisulfite sequencing analysis of the 3′ end region of *MMERVK10C*. Empty and full circles represent unmethylated and methylated CpGs, respectively.

Data are presented as mean of relative values ± SEM. *p < 0.05, Student's t test.

internalized *env*-lacking mouse ERVs that demonstrate a large degree of polymorphism among different mouse strains and maintain the capacity to retrotranspose. Using immunocytochemistry with an IAP-specific antibody, we found a uniform, high-level IAP-gag expression located to the cytoplasm in *Trim28*$^{-/-}$ NPCs (Figure 3C).

Taken together, these data demonstrate that deletion of TRIM28 in NPCs results in a massive transcriptional increase of ERVs, including *MMERVK10C* and *IAP1*. NPCs thus appear to constitute a cellular environment distinct from that of other somatic cells studied so far, with the TRIM28-induced histone-based repressive mechanism playing a role in ERV control.

## Activation of ERVs Correlates with Increased Transcription of Nearby Genes

The ability of ERVs to attract transcription factors and silencing complexes has led to a reassessment of their role in the host genome. ERVs are now considered to be important transcriptional regulatory elements that shape and influence gene expression during early development (Isbel and Whitelaw, 2012). For example, we have recently found that TRIM28 controls the expression of developmental genes by repressing ERV-associated enhancers in pluripotent cells (Rowe et al., 2013). Twenty-six *MMERVK10C* proviruses and 361 *IAP* proviruses that were upregulated in *Trim28*$^{-/-}$ NPCs were mapped to precise genomic locations (Figure S1). Out of these 387 proviruses, 90 were situated close to genes (<50 kb). We found that 25 of those genes (28%) demonstrated significantly increased expression, whereas expression of only six of them was decreased (7%) (Figure 4A). We also found that those 90 genes located close to upregulated ERVs (ERV-up genes) were on average 3-fold upregulated in *Trim28*$^{-/-}$ cells (Figure 4B). In contrast, a random selection of ERVs that was not upregulated in *Trim28*$^{-/-}$ cells (n = 129, *MMERVK10C* and *IAP1* elements) did not affect nearby genes (ERVs-ctrl genes, n = 50, Figure 4B). Interestingly, we also found that ERV-up genes were expressed at low levels in wild-type cells (Figure 4C), which is in agreement with a model where ERVs mediate repressive regulation of nearby genes caused by the attraction of the TRIM28 silencing complex to ERV sequences. We validated the increased expression of five ERV-up genes in *Trim28*$^{-/-}$ cells using qRT-PCR (Figure 4D).

## ERVs Produce Long Noncoding RNAs

We looked in detail at *BC048671*, which is a protein-coding transcript that is induced in *Trim28*$^{-/-}$ NPCs but completely absent in wild-type NPCs. *BC048671* is located 5 kb downstream of an *IAP* element, which is also highly upregulated in *Trim28*$^{-/-}$ NPCs. The RNA-seq data show that transcriptional initiation at the *IAP* element results in the formation of a long transcript (>10 kb) that extends into the coding sequence of *BC048671* (Figure 4E). The presence of high levels of this long transcript was verified using qRT-PCR primers located both upstream and within the coding sequence of *BC048671* (Figure 4F). Thus, readthrough of an ERV-derived transcript into another locus is likely to be one of several mechanisms by which nearby gene expression can be affected (see also Figure S2). This finding supports the notion that a general feature of ERVs might

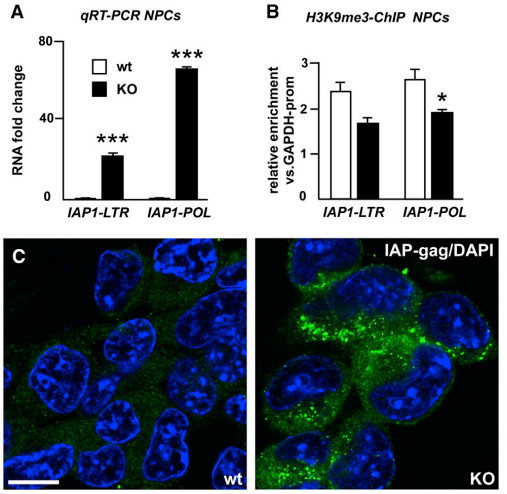



**Figure 3. Analysis of *IAP1* Expression**

(A) Quantitative analysis of transcript levels of different regions of *IAP1* provirus in *Trim28*$^{-/-}$ and wild-type NPCs.

(B) ChIP for H3K9me3 in *Trim28*$^{-/-}$ and wild-type NPCs.

(C) Confocal analysis of immunofluorescence staining for IAP-gag on *Trim28*$^{-/-}$ and wild-type NPCs. Scale bar, 10 μm.

Data are presented as mean of relative values ± SEM. *p < 0.05, Student's t test.

be to act as transcriptional start sites for long noncoding RNAs (lncRNAs). Indeed, when we scrutinized ERV elements located in gene free regions, we found that both *IAP* and *MMERVK10C* elements serve as start sites for lncRNAs (Figures 4G and 4I). Using qRT-PCR, we confirmed high-level expression of two ERV-derived lncRNAs in *Trim28*$^{-/-}$ NPCs (Figures 4H and 4J). The length of the ERV-derived lncRNAs did in many cases extend 25 kb (Figure 4K). These data demonstrate that derepression of ERVs in NPCs can result in the expression of multiple lncRNAs. The functional role of lncRNAs in NPCs remains largely unexplored, but they are thought to play important regulatory roles and have been implicated as scaffolds for nuclear protein complexes and as antisense transcripts in the control of epigenetic pathways (Guttman and Rinn, 2012).

## DISCUSSION

In pluripotent stem cells, TRIM28 is a master corepressor of retroelements including ERVs (Matsui et al., 2010; Rowe et al., 2010). When these cells differentiate into various somatic cell types, DNA methylation is instated on ERV sequences, which ultimately results in stable silencing that is no longer dependent on TRIM28 (Wiznerowicz et al., 2007; Rowe et al., 2013). Thus, when TRIM28 is deleted from various somatic cell types such as fibroblasts, hepatocytes, and white blood cells, no increased ERV expression is detected (Rowe et al., 2010; Bojkowska et al., 2012; Santoni de Sio et al., 2012a, 2012b). Here, we describe an exception to this rule. When TRIM28 is deleted in NPCs, several ERVs become highly expressed. This finding unravels a unique transcriptional regulation of ERVs in NPCs.

ERVs regulated by TRIM28 in NPCs are recent invaders of the mouse genome. *IAP1* is the most recent member of the well-studied IAP ERVs (Qin et al., 2010). IAPs are ERVs that have lost the *env* gene and adopted an intracellular life cycle (Ribet et al., 2008). *IAP1* has been shown to retrotranspose and has distinct integration patterns in different strains of laboratory mice (Li et al., 2012). *MMERVK10C*, another ERV massively upregulated in *Trim28*$^{-/-}$ NPCs, is poorly characterized, and it is unclear if it is still endowed with retrotransposition potential, whether on its own or with the support of factors provided in *trans*. A previous study that analyzed the structure of *MMERVK10C* elements in the mouse genome found that the majority of these elements have 3′ deletions removing the start of the *gag* open reading frame as well as the major part of *env* (Reichmann et al., 2012). Our data demonstrate that, in NPCs, TRIM28 controls the rare copies of *env*-containing *MMERVK10C* elements, which are most likely to be the youngest ones, raising the possibility that these recent invaders of the mouse genome contain *cis*-acting genomic elements that allow them to escape DNA methylation in NPCs.

The classic view of repetitive mobile genetic elements as parasitic DNA without beneficial function to the host is challenged in many ways. There are a number of recent studies indicating that transposable elements (TEs) play important roles in establishing and rewiring gene networks (Kunarso et al., 2010; Chuong et al., 2013). TEs have been shown to act as enhancers, repressors, and alternative promoters. In addition, TEs can affect splicing patterns and produce peptides with important functional roles (Jern and Coffin, 2008). In this study, we demonstrate that activated ERVs can influence gene expression of nearby genes, such as *BC048671*, and serve as start sites for lncRNAs. Taken together, our findings indicate that ERVs participate in the control of gene networks in the brain.

We have previously demonstrated that deletion of *Trim28* in postmitotic forebrain neurons results in complex behavioral changes (Jakobsson et al., 2008). In addition, heterozygous germline deletion of *Trim28* has been described to result in abnormal behavioral phenotypes (Whitelaw et al., 2010). In this study, we found that deletion of *Trim28* during brain development is lethal (Figure S3). In addition, we also noted that heterozygous deletion of *Trim28* during brain development resulted in behavioral changes characterized by hyperactivity (Figure S3). Together, these findings demonstrate that disruption of TRIM28 levels in the mouse brain results in behavioral changes that are similar to impairments found in humans with certain psychiatric disorders. With this in mind, it is noteworthy that increased levels of ERV transcripts have been detected in patients with several neurological and psychiatric disorders (Jeong et al., 2010; Douville et al., 2011; Li et al., 2012; Karlsson et al., 2001). The significance of these findings has been questioned because the human genome does not appear to harbor ERVs with known retrotransposing capacity (Jern and Coffin, 2008). However, the increasing evidence that derepression of

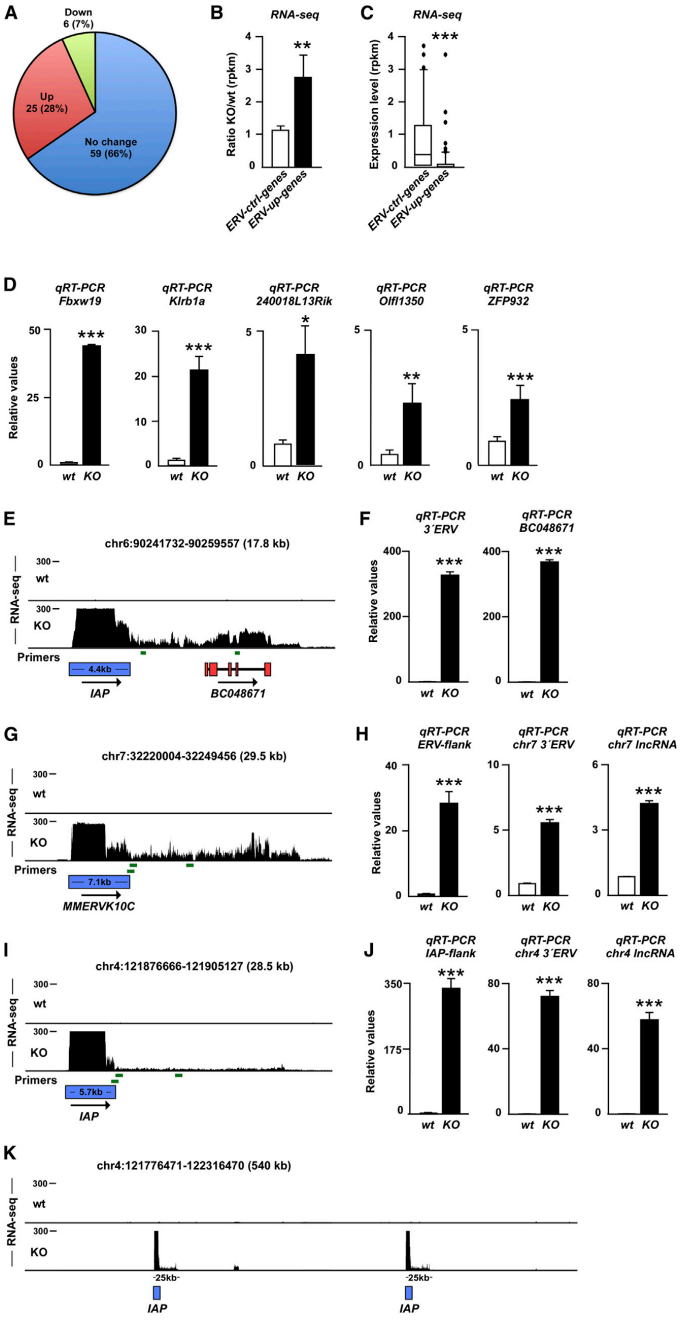



**Figure 4. Activation of ERVs Influences Expression of Nearby Genes and Results in the Expression of lncRNAs**

(A) Transcriptional change of genes located close (<50 kb) to ERVs in *Trim28*$^{-/-}$ NPCs.

(B) Mean transcriptional change of genes located to ERVs with increased transcription (ERV-up genes) and genes located close to unchanged ERVs (ERV-ctrl genes) in *Trim28*$^{-/-}$ NPCs.

(C) Absolute expression level of ERV-up genes and ERV-ctrl genes in wild-type NPCs.

(D) qRT-PCR of RNA isolated from wild-type and *Trim28*$^{-/-}$ NPCs.

(E) Screen shot from the USCS genome browser (mm9) showing induced transcription of *BC048671 in Trim28*$^{-/-}$ NPCs.

(G, I, and K) Activation of ERVs results in the expression of lncRNAs. Screen shot from the USCS genome browser (mm9).

(F, H, and J) qRT-PCR of RNA isolated from wild-type and *Trim28*$^{-/-}$ NPCs. Primers are indicated as green bars and include primers over the ERV junction as well as close and more distant from the 3′ end of the ERVs.

Data are presented as mean of relative values ± SEM. *p < 0.05, Student's t test. See also Figures S1 and S2.

ERVs influence gene networks, including the findings presented here, provides a potential mechanistic explanation for these observations.

In summary, our data suggest that ERVs may be involved in the regulation of gene expression in NPCs and may hereby offer a link between ERVs and brain disorders. It seems unlikely that behavioral phenotypes would arise from the derepression of a single ERV-induced gene. Instead, the presence of ERVs in multiple copies scattered throughout the genome allows for a powerful network-like control of gene expression, where dysregulation could result in widespread consequences. However, due to the large numbers of ERVs present in the mouse and human genome and their sequence variation, it is currently unfeasible to demonstrate a causal role for ERVs in controlling complex behavior or brain disorders using loss-of-function approaches, such as gene targeting and small hairpin RNA (shRNA) knockdown. Instead, improving our knowledge of critical host factors and networks controlling ERVs is essential to appreciate their impact on the genome and pathologies that may stem from their dysregulation. The demonstration that there is an ongoing dynamic TRIM28-mediated regulation of ERVs in NPCs is a step in this direction and warrants future studies of epigenetic and posttranscriptional regulation of ERVs in the healthy and diseased brain.

## EXPERIMENTAL PROCEDURES

Detailed experimental procedures can be found in the Supplemental Experimental Procedures.

### Procedures

#### *Transgenic Animals*

All animal-related procedures were approved by and conducted in accordance with the committee for use of laboratory animals at Lund University. NestinCre and floxed *Trim28* mice have been described previously (Weber et al., 2002; Tronche et al., 1999).

#### *Cell Culture*

NPC was established from embryonic day 13.5 (E13.5) forebrain and cultured as previously described (Conti et al., 2005).

#### *Immunofluorescence*

Immunofluorescence was performed as previously described (Thompson et al., 2005; Sachdeva et al., 2010).

#### *RNA Studies*

RNA-seq and qRT-PCR was performed as previously described (Rowe et al., 2010). The 50-base-paired end reads were mapped onto the RepBase version 16.08 (Jurka et al., 2005) and to the mouse genome (mm9) assembly. Mapping was done using the bowtie short read aligner (Langmead et al., 2009).

#### *Chromatin Immunoprecipitation*

Chromatin immunoprecipitation was performed with iDeal chromatin immunoprecipitation sequencing (ChIP-seq) kit (Diagenode) according to supplier's recommendations.

#### *DNA-Methylation Analysis*

Bisulfite sequencing was performed with the EpiTect bisulfite kit (QIAGEN) according to the supplier's recommendations. Sequence data were analyzed with the QUantification tool for Methylation Analysis (Kumaki et al., 2008).

#### *Statistical Analysis*

An unpaired t test was performed in order to test for statistical significance. Data are presented as mean ± SEM.

## ACCESSION NUMBERS

The RNA-seq data were deposited in the NCBI Gene Expression Omnibus and are available under accession number GSE45930.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.12.004.

## AUTHOR CONTRIBUTIONS

L.F., A.K., R.S., R.P., M.E.J., and C.M. designed and performed research and analyzed data. P.J., P.T., and D.T. designed research and analyzed data. F.C. contributed reagents. J.J. designed and coordinated the project and analyzed data. L.F. and J.J. wrote the paper, and all authors reviewed the manuscript.

## ACKNOWLEDGMENTS

We are grateful to A. Björklund, S. Quenneville, and all members of the J.J. and Parmar laboratories for stimulating discussions. We thank U. Jarl, A. Josefsson, C. Isaksson, I. Nilsson, A.-K. Oldén, E. Ling, S. Smiljanic, M. Sparrenius, and E. Tjon for technical assistance. This study was supported by grants from Swedish Research Council (J.J.), Formas (P.J.), the Swedish Cancer Foundation (J.J.), the Lundqvist, Jeansson, and Crafoord foundations (J.J.), the Swedish Government Initiative for Strategic Research Areas *MultiPark* (J.J.), the French government, CNRS and INSERM (F.C.), and the French Agence Nationale pour la Recherche (F.C.).

Received: June 10, 2014
Revised: October 28, 2014
Accepted: December 1, 2014
Published: December 24, 2014

## REFERENCES

Baillie, J.K., Barnett, M.W., Upton, K.R., Gerhardt, D.J., Richmond, T.A., De Sapio, F., Brennan, P.M., Rizzu, P., Smith, S., Fell, M., et al. (2011). Somatic retrotransposition alters the genetic landscape of the human brain. Nature *479*, 534–537.

Belshaw, R., Dawson, A.L., Woolven-Allen, J., Redding, J., Burt, A., and Tristem, M. (2005). Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K(HML2): implications for present-day activity. J. Virol. *79*, 12507–12514.

Bojkowska, K., Aloisio, F., Cassano, M., Kapopoulou, A., Santoni de Sio, F., Zangger, N., Offner, S., Cartoni, C., Thomas, C., Quenneville, S., et al. (2012). Liver-specific ablation of Krüppel-associated box-associated protein 1 in mice leads to male-predominant hepatosteatosis and development of liver adenoma. Hepatology *56*, 1279–1290.

Cammas, F., Mark, M., Dollé, P., Dierich, A., Chambon, P., and Losson, R. (2000). Mice lacking the transcriptional corepressor TIF1beta are defective in early postimplantation development. Development *127*, 2955–2963.

Chuong, E.B., Rumi, M.A., Soares, M.J., and Baker, J.C. (2013). Endogenous retroviruses function as species-specific enhancer elements in the placenta. Nat. Genet. *45*, 325–329.

Conti, L., Pollard, S.M., Gorba, T., Reitano, E., Toselli, M., Biella, G., Sun, Y., Sanzone, S., Ying, Q.L., Cattaneo, E., and Smith, A. (2005). Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. PLoS Biol. *3*, e283.

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Yeo, G.W., Mu, Y., Lovci, M.T., Morell, M., O'Shea, K.S., Moran, J.V., and Gage, F.H. (2009). L1 retrotransposition in human neural progenitor cells. Nature *460*, 1127–1131.

Cowley, M., and Oakey, R.J. (2013). Transposable elements re-wire and fine-tune the transcriptome. PLoS Genet. *9*, e1003234.

Douville, R., Liu, J., Rothstein, J., and Nath, A. (2011). Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis. Ann. Neurol. *69*, 141–151.

Ellis, J., Hotta, A., and Rastegar, M. (2007). Retrovirus silencing by an epigenetic TRIM. Cell *131*, 13–14.

Evrony, G.D., Cai, X., Lee, E., Hills, L.B., Elhosary, P.C., Lehmann, H.S., Parker, J.J., Atabay, K.D., Gilmore, E.C., Poduri, A., et al. (2012). Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. Cell *151*, 483–496.

Feschotte, C. (2008). Transposable elements and the evolution of regulatory networks. Nat. Rev. Genet. *9*, 397–405.

Guttman, M., and Rinn, J.L. (2012). Modular regulatory principles of large non-coding RNAs. Nature *482*, 339–346.

Hutnick, L.K., Huang, X., Loo, T.C., Ma, Z., and Fan, G. (2010). Repression of retrotransposal elements in mouse embryonic stem cells is primarily mediated by a DNA methylation-independent mechanism. J. Biol. Chem. *285*, 21082–21091.

Isbel, L., and Whitelaw, E. (2012). Endogenous retroviruses in mammals: an emerging picture of how ERVs modify expression of adjacent genes. Bioessays *34*, 734–738.

Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M., Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E., and Jaenisch, R. (2001). Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. Nat. Genet. *27*, 31–39.

Jakobsson, J., Cordero, M.I., Bisaz, R., Groner, A.C., Busskamp, V., Bensadoun, J.C., Cammas, F., Losson, R., Mansuy, I.M., Sandi, C., and Trono, D. (2008). KAP1-mediated epigenetic repression in the forebrain modulates behavioral vulnerability to stress. Neuron *60*, 818–831.

Jeong, B.H., Lee, Y.J., Carp, R.I., and Kim, Y.S. (2010). The prevalence of human endogenous retroviruses in cerebrospinal fluids from patients with sporadic Creutzfeldt-Jakob disease. J. Clin. Virol. *47*, 136–142.

Jern, P., and Coffin, J.M. (2008). Effects of retroviruses on host genome function. Annu. Rev. Genet. *42*, 709–732.

Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. *110*, 462–467.

Karlsson, H., Bachmann, S., Schröder, J., McArthur, J., Torrey, E.F., and Yolken, R.H. (2001). Retroviral RNA identified in the cerebrospinal fluids and brains of individuals with schizophrenia. Proc. Natl. Acad. Sci. USA *98*, 4634–4639.

Kumaki, Y., Oda, M., and Okano, M. (2008). QUMA: quantification tool for methylation analysis. Nucleic Acids Res. *36* (Web Server issue), W170-5.

Kunarso, G., Chia, N.Y., Jeyakani, J., Hwang, C., Lu, X., Chan, Y.S., Ng, H.H., and Bourque, G. (2010). Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nat. Genet. *42*, 631–634.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Li, J., Akagi, K., Hu, Y., Trivett, A.L., Hlynialuk, C.J., Swing, D.A., Volfovsky, N., Morgan, T.C., Golubeva, Y., Stephens, R.M., et al. (2012). Mouse endogenous retroviruses can trigger premature transcriptional termination at a distance. Genome Res. *22*, 870–884.

Li, W., Prazak, L., Chatterjee, N., Grüninger, S., Krug, L., Theodorou, D., and Dubnau, J. (2013). Activation of transposable elements during aging and neuronal decline in Drosophila. Nat. Neurosci. *16*, 529–531.

Matsui, T., Leung, D., Miyashita, H., Maksakova, I.A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M.C., and Shinkai, Y. (2010). Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature *464*, 927–931.

Muotri, A.R., Chu, V.T., Marchetto, M.C., Deng, W., Moran, J.V., and Gage, F.H. (2005). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature *435*, 903–910.

Muotri, A.R., Marchetto, M.C., Coufal, N.G., and Gage, F.H. (2007). The necessary junk: new functions for transposable elements. Hum. Mol. Genet. *16 Spec No. 2*, R159–R167.

Muotri, A.R., Marchetto, M.C., Coufal, N.G., Oefner, R., Yeo, G., Nakashima, K., and Gage, F.H. (2010). L1 retrotransposition in neurons is modulated by MeCP2. Nature *468*, 443–446.

Perrat, P.N., DasGupta, S., Wang, J., Theurkauf, W., Weng, Z., Rosbash, M., and Waddell, S. (2013). Transposition-driven genomic heterogeneity in the Drosophila brain. Science *340*, 91–95.

Qin, C., Wang, Z., Shang, J., Bekkari, K., Liu, R., Pacchione, S., McNulty, K.A., Ng, A., Barnum, J.E., and Storer, R.D. (2010). Intracisternal A particle genes: Distribution in the mouse genome, active subtypes, and potential roles as species-specific mediators of susceptibility to cancer. Mol. Carcinog. *49*, 54–67.

Reichmann, J., Crichton, J.H., Madej, M.J., Taggart, M., Gautier, P., Garcia-Perez, J.L., Meehan, R.R., and Adams, I.R. (2012). Microarray analysis of LTR retrotransposon silencing identifies Hdac1 as a regulator of retrotransposon expression in mouse embryonic stem cells. PLoS Comput. Biol. *8*, e1002486.

Reilly, M.T., Faulkner, G.J., Dubnau, J., Ponomarev, I., and Gage, F.H. (2013). The role of transposable elements in health and diseases of the central nervous system. J. Neurosci. *33*, 17577–17586.

Ribet, D., Harper, F., Dupressoir, A., Dewannieux, M., Pierron, G., and Heidmann, T. (2008). An infectious progenitor for the murine IAP retrotransposon: emergence of an intracellular genetic parasite from an ancient retrovirus. Genome Res. *18*, 597–609.

Rowe, H.M., and Trono, D. (2011). Dynamic control of endogenous retroviruses during development. Virology *411*, 273–287.

Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P.V., Layard-Liesching, H., Verp, S., Marquis, J., et al. (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. Nature *463*, 237–240.

Rowe, H.M., Kapopoulou, A., Corsinotti, A., Fasching, L., Macfarlan, T.S., Tarabay, Y., Viville, S., Jakobsson, J., Pfaff, S.L., and Trono, D. (2013). TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells. Genome Res. *23*, 452–461.

Sachdeva, R., Jönsson, M.E., Nelander, J., Kirkeby, A., Guibentif, C., Gentner, B., Naldini, L., Björklund, A., Parmar, M., and Jakobsson, J. (2010). Tracking differentiating neural progenitors in pluripotent cultures using microRNA-regulated lentiviral vectors. Proc. Natl. Acad. Sci. USA *107*, 11602–11607.

Santoni de Sio, F.R., Barde, I., Offner, S., Kapopoulou, A., Corsinotti, A., Bojkowska, K., Genolet, R., Thomas, J.H., Luescher, I.F., Pinschewer, D., et al. (2012a). KAP1 regulates gene networks controlling T-cell development and responsiveness. FASEB J. *26*, 4561–4575.

Santoni de Sio, F.R., Massacand, J., Barde, I., Offner, S., Corsinotti, A., Kapopoulou, A., Bojkowska, K., Dagklis, A., Fernandez, M., Ghia, P., et al. (2012b). KAP1 regulates gene networks controlling mouse B-lymphoid cell differentiation and function. Blood *119*, 4675–4685.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd. (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev. *16*, 919–932.

Sperber, G.O., Airola, T., Jern, P., and Blomberg, J. (2007). Automated recognition of retroviral sequences in genomic data—RetroTector. Nucleic Acids Res. *35*, 4964–4976.

Sripathy, S.P., Stevens, J., and Schultz, D.C. (2006). The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. Mol. Cell. Biol. *26*, 8623–8638.

Sweatt, J.D. (2013). The emerging field of neuroepigenetics. Neuron *80*, 624–632.

Thomas, J.H., and Schneider, S. (2011). Coevolution of retroelements and tandem zinc finger genes. Genome Res. *21*, 1800–1812.

Thompson, L., Barraud, P., Andersson, E., Kirik, D., and Björklund, A. (2005). Identification of dopaminergic neurons of nigral and ventral tegmental area

subtypes in grafts of fetal ventral mesencephalon based on cell morphology, protein expression, and efferent projections. J. Neurosci. *25*, 6467–6477.

Tronche, F., Kellendonk, C., Kretz, O., Gass, P., Anlag, K., Orban, P.C., Bock, R., Klein, R., and Schütz, G. (1999). Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety. Nat. Genet. *23*, 99–103.

Weber, P., Cammas, F., Gerard, C., Metzger, D., Chambon, P., Losson, R., and Mark, M. (2002). Germ cell expression of the transcriptional co-repressor TIF1beta is required for the maintenance of spermatogenesis in the mouse. Development *129*, 2329–2337.

Whitelaw, N.C., Chong, S., Morgan, D.K., Nestor, C., Bruxner, T.J., Ashe, A., Lambley, E., Meehan, R., and Whitelaw, E. (2010). Reduced levels of two modifiers of epigenetic gene silencing, Dnmt3a and Trim28, cause increased phenotypic noise. Genome Biol. *11*, R111.

Wiznerowicz, M., Jakobsson, J., Szulc, J., Liao, S., Quazzola, A., Beermann, F., Aebischer, P., and Trono, D. (2007). The Kruppel-associated box repressor domain can trigger de novo promoter methylation during mouse early embryogenesis. J. Biol. Chem. *282*, 34535–34541.

Wolf, D., and Goff, S.P. (2009). Embryonic stem cells use ZFP809 to silence retroviral DNAs. Nature *458*, 1201–1204.

# Paper II

# Paper III

# TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells

Helen M. Rowe,[1] Adamandia Kapopoulou,[1,2] Andrea Corsinotti,[1] Liana Fasching,[3] Todd S. Macfarlan,[4] Yara Tarabay,[5] Stéphane Viville,[5] Johan Jakobsson,[3] Samuel L. Pfaff,[4] and Didier Trono[1,6]

[1]*School of Life Sciences and Frontiers in Genetics Program, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland;* [2]*Swiss Bioinformatics Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland;* [3]*Wallenberg Neuroscience Center, Lund University, BMC A11, 221 84 Lund, Sweden;* [4]*Gene Expression Laboratory and the Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, California 92037, USA;* [5]*Institute of Genetics and Molecular and Cellular Biology (IGBMC), University of Strasbourg, BP10142, Illkirch Cedex, France*

**TRIM28 is critical for the silencing of endogenous retroviruses (ERVs) in embryonic stem (ES) cells. Here, we reveal that an essential impact of this process is the protection of cellular gene expression in early embryos from perturbation by *cis*-acting activators contained within these retroelements. In TRIM28-depleted ES cells, repressive chromatin marks at ERVs are replaced by histone modifications typical of active enhancers, stimulating transcription of nearby cellular genes, notably those harboring bivalent promoters. Correspondingly, ERV-derived sequences can repress or enhance expression from an adjacent promoter in transgenic embryos depending on their TRIM28 sensitivity in ES cells. TRIM28-mediated control of ERVs is therefore crucial not just to prevent retrotransposition, but more broadly to safeguard the transcriptional dynamics of early embryos.**

**[Supplemental material is available for this article.]**

TRIM28 (tripartite motif-containing protein 28, also known as KAP1, KRAB-associated protein 1, or TIF1β) is a co-repressor that is highly expressed in embryonic stem (ES) cells and is crucial to early mouse development, because homozygous *Trim28* knock-out (KO) embryos arrest shortly after implantation and fail to gastrulate (Cammas et al. 2000). TRIM28 is tethered to DNA by sequence-specific Krüppel-associated box zinc finger proteins (KRAB-ZFPs) (Friedman et al. 1996; Emerson and Thomas 2009; Thomas and Schneider 2011) and induces local heterochromatin formation through the histone methyltransferase SETDB1 (or ESET), responsible for trimethylating histone 3 at lysine 9 (Schultz et al. 2002; Ivanov et al. 2007; Frietze et al. 2010), the NuRD (nucleosome remodeling and deacetylation) complex (Schultz et al. 2001), which contains the histone deacetylases HDAC1 and HDAC2 (for review, see McDonel et al. 2009), and heterochromatin protein 1 (HP-1) (Lechner et al. 2000; Sripathy et al. 2006). TRIM28 is required for proper oocyte-to-embryo transition (Messerschmidt et al. 2012), for the maintenance of imprinting marks immediately after fertilization (Li et al. 2008; Quenneville et al. 2011; Zuo et al. 2012), and for the self-renewal of ES cells, which rapidly die or undergo differentiation upon its removal (Wolf and Goff 2007; Fazzio et al. 2008; Hu et al. 2009; Rowe et al. 2010; Seki et al. 2010). However, which specific genes are controlled by TRIM28 during this early embryonic period remains largely unknown.

**[6]Corresponding author**
**E-mail Didier.trono@epfl.ch**
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.147678.112.

In contrast, it has now been firmly established that TRIM28, in part through SETDB1, is responsible for maintaining endogenous retroviruses (ERVs) in a silent state in ES cells and early embryos (Matsui et al. 2010; Rowe et al. 2010). TRIM28-mediated repression acts on multiple subsets of ERVs including intracisternal A-type particles (IAPs) and early transposon (Etn)/MusD elements, as well as on MERVL and ERVK families (for review, see Rowe and Trono 2011), and also partakes in blocking the replication of murine leukemia virus (MLV) in murine embryonic cells (Wolf and Goff 2007, 2009). Preventing the genomic spread of these retroelements may intuitively appear as the primary role of this process, yet the vast majority of ERVs carry mutations that inactivate their retrotransposition potential. Accordingly, it is noteworthy that the long terminal repeats (*LTRs*) of ERVs harbor binding sites for numerous transcription factors, as expected from the needs of their own replication. Furthermore, rare ERV-contained sequences have been found to function as *cis*-acting regulatory elements during mouse, human, and chick development through their recruitment of proteins such as POU5F1 (also called OCT4), GATA4, and CTCF (Bourque et al. 2008; Kunarso et al. 2010; Mey et al. 2012; Schmidt et al. 2012). ERVs and cellular genes can additionally be coordinately controlled in ES cells (Karimi et al. 2011; Macfarlan et al. 2011, 2012). Based on this premise, we asked here whether a component of the TRIM28-mediated maintenance of ES cell homeostasis might be the control of cryptic ERV-associated transcriptional activators. Our results indicate that ERVs are, indeed, transcriptional landmines, the TRIM28-mediated control of which is essential to preserve the transcriptional dynamics of ES cells. Regulation of retrotransposons by a TRIM28 pathway is thus critical not just to


23:452–461 © 2013, Published by Cold Spring Harbor Laboratory Press; ISSN 1088-9051/13; www.genome.org

prevent retrotransposition, but more broadly to safeguard the timely activation of genes during early development.

## Results

### Transcriptional deregulation in *Trim28* knock-out ES cells

Using a previously described tamoxifen-inducible *Cre/lox* system (Rowe et al. 2010), we first compared mRNA-sequencing (mRNA-seq) data from control and *Trim28*-deleted murine ES cells (Fig. 1A,B). Transcripts from ~20,000 genes were detected in control cells. Four days after *Cre* induction, based on a twofold cutoff and a significant difference of $P \leq 0.05$, around 5700 of them were up-regulated (29%, including 1850 transcripts that were more than fivefold up-regulated), while around 720 were down-regulated (4%) and 13,600 unchanged (67%). From now on, we refer to these gene groups as "Up," "Down," and "Stable," respectively. In contrast, in mouse embryonic fibroblasts (MEFs), transcriptional deregulation was only modest upon *Trim28* deletion (Fig. 1A). This correlates the difference between the dramatic phenotype of *Trim28*-deleted ES cells, which die or differentiate after a few days and overexpress ERVs, and MEFs, which can be stably maintained and do not up-regulate ERVs (Rowe et al. 2010). Of note, genes affected by *Trim28* deletion (both Up and Down) in ES cells were lowly expressed at baseline compared with genes unaffected by removal of this regulator (according to a Wilcoxon rank-sum test that was used to calculate significance here and for all boxplots) (Supplemental Fig. S1A). We decided to focus on up-regulated genes since they represented the larger category and Gene Ontology analysis indicated these genes to be involved in developmental pathways (see Supplemental Fig. S1B; Supplemental Table 1), including through expression at the embryonic two-cell stage as recently described (Macfarlan et al. 2012).

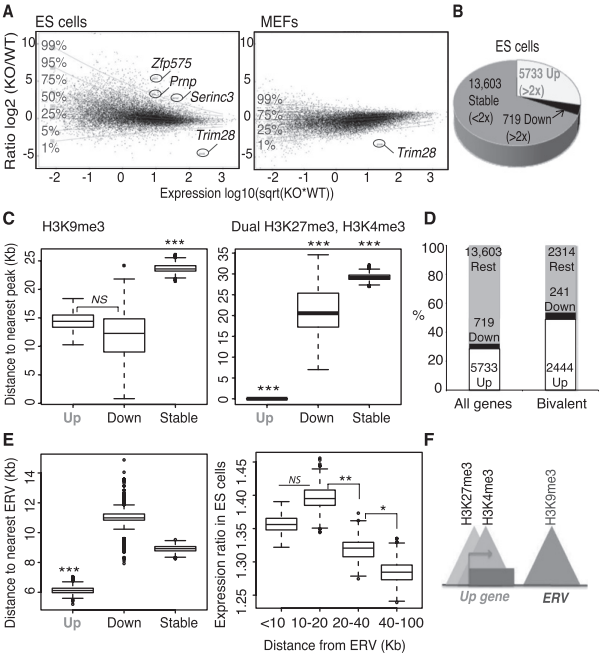



**Figure 1.** *Trim28* deletion in ES cells leads to up-regulation of genes close to ERVs, including many bivalent genes. (*A*) mRNA-seq in *Trim28* wild-type (WT) and knock-out (KO) embryonic stem (ES) cells (*left* panel) or *Trim28* WT and KO MEFs (*right* panel). Transcripts (assembly mm9) are plotted in black with the ratio on the *y*-axis and expression level on the *x*-axis. (Sqrt) Square root. (Horizontal lines) Levels of gene deregulation (e.g., only 1% of genes lie above the 99% line). The genes *Zfp575*, *Prnp*, and *Serinc3* (referred to later) are highlighted, as well as *Trim28*. (*B*) Data from ES cells in *A* were used to group transcripts depending on whether they were greater than twofold up-regulated (Up), greater than twofold down-regulated (Down), or less than twofold affected (Stable). Up and Down genes were significantly changed based on a DESeq test (Anders and Huber 2010) (adjusted *P*-values ≤0.05). (*C*) The distance to the nearest peak (of either H3K9me3 on the *left* panel, 19,128 peaks, or dual H3K27me3, H3K4me3 peaks on the *right* panel, 12,390 peaks) from Up, Down, and Stable gene groups. (*Left P*-values) Up versus Down, not significant (NS), $P = 0.48$; Up versus Stable, $P = 7.7 \times 10^{-10}$; Down versus Stable, $P = 0.0010$. (*Right P*-values) Up versus Down, $P = 9.9 \times 10^{-11}$; Up versus Stable, $P \leq 2.2 \times 10^{-16}$; Down versus Stable, $P = 4.1 \times 10^{-4}$. (*D*) Bivalent genes (as defined above by the presence of dual H3K27me3, H3K4me3 peaks) are enriched for up-regulated genes compared with all genes. (*E*) ERV locations ($N = 82,382$) were downloaded from the UCSC Genome Browser to include the categories ERV, ERV1, ERVK, and ERVL as defined by Repbase with a size cutoff of 500-bp minimum and used to plot the distance to the nearest ERV from Up, Down, and Stable gene groups (*left*). A Mann-Whitney Wilcoxon test was used to calculate significance: Up genes were significantly closer than the other two gene groups; (***) $P \leq 0.001$. (*Right*) All genes were divided into groups based on their distance to the nearest ERV and their ratio between *Trim28* WT and KO ES cells plotted on the *y*-axis. (*P*-values) The groups 10–20 versus 20–40 and 20–40 versus 40–100 are different: $P = 0.0048$ and $P = 0.01$, respectively. (*F*) Model showing that Up genes are close to H3K9me3 marks and ERVs and are often bivalent.

### Chromatin state at genes affected by *Trim28* deletion

Surprisingly, confrontation of these results with TRIM28 ChIP-seq data performed in the same cells revealed that <1% of up-regulated gene promoters were direct targets of TRIM28 (Supplemental Table 2). This suggested that Up genes could be indirectly affected by *Trim28* deletion and/or were normally subjected to TRIM28-controlled nearby *cis*-acting influences. We thus compared the chromatin status of Up, Down, and Stable genes more broadly using available ChIP-seq data (Mikkelsen et al. 2007). We focused on H3K4me3, a Trithorax group– or TrxG-deposited mark typically associated with active transcription, H3K9me3, frequently a signature of TRIM28/SETDB1 recruitment (Matsui et al. 2010; Rowe et al. 2010), and H3K27me3, another repressive histone modification induced by the Polycomb repressive complex 2 (PRC2) (Bernstein et al. 2006; Gan et al. 2007; Guenther and Young 2010). As previously observed (Mikkelsen et al. 2007), H3K4me3 and H3K27me3 were significantly enriched at gene promoters, while H3K9me3 was generally depleted from these functional domains (Supplemental Fig. S1C). Genes deregulated upon TRIM28 depletion, whether up or down, were significantly closer to H3K9me3-enriched regions than

unaffected genes (Fig. 1C, left). More revealingly, Up genes almost completely coincided with H3K27me3 peaks (Supplemental Fig. S1D). In ES cells, the H3K27me3 repressive mark is found together with its activating counterpart H3K4me3 at so-called bivalent promoters, which are rapidly induced upon differentiation (Bernstein et al. 2006). We thus compared the relative distribution of these two marks over the three gene groups. Genes unaffected by TRIM28 removal were the closest to H3K4me3-alone peaks and the farthest away from H3K27me3-alone peaks (Supplemental Fig. S1E), consistent with their average higher levels of expression than Up or Down genes. In contrast and most strikingly, Up genes almost completely overlapped bivalent H3K4me3/H3K27me3 peaks (Fig. 1C, right), indicating that the promoters of many of the genes induced upon *Trim28* deletion are poised for transcription. Reciprocally, up-regulated genes (2444) were enriched among bivalent genes (4999) (Mikkelsen et al. 2007), compared with all genes (Fig. 1D, Fisher's exact test: *P*-value $\leq 1 \times 10^{-16}$).

### Genes up-regulated upon *Trim28* deletion are located close to ERVs

Since few gene promoters were direct targets of TRIM28 (see above), we hypothesized that up-regulation of many genes could reflect the deregulation of TRIM28-controlled *cis*-acting elements situated in their nearby vicinity. In that respect, TRIM28, together with H3K9me3, is found enriched at ERV sequences in ES cells but not MEFs (Matsui et al. 2010; Rowe et al. 2010). Because ERVs are known to contain transcription-regulating sequences, we asked whether they were spatially associated with genes induced upon *Trim28* deletion. Indeed, matching the genomic locations of ERVs (82,382 sites) with the three gene groups differentially affected by TRIM28 removal revealed that Up genes were on average significantly closer to these elements than Down or Stable genes (Fig. 1E, left). We also verified that it is not the case that all bivalent genes are enriched in ERVs but rather that bivalent Up genes (2444) are on average closer to ERVs than bivalent stable genes (2314, $P = 0.001470$) (Supplemental Fig. S2A). Interestingly, Up genes also clustered with long interspersed nuclear elements (*LINE1s*) but lay further from short interspersed nuclear elements (*SINEs*) than Down and Stable genes (Supplemental Fig. S2B–D), consistent with the previous observation that *LINEs* but not *SINEs* are modestly up-regulated in *Trim28*-deleted ES cells (Rowe et al. 2010). Reciprocally, the closer genes were to an ERV or particularly to an ERV of the subclass IAPs, the higher their average up-regulation upon TRIM28 removal, with genes also affected (although to a lesser extent) at distances of 100 kb (Fig. 1E, right; data not shown). Of note, this phenomenon of nearby *cis*-acting regulation is consistent with the previously documented modulation of the *Agouti* gene by an IAP located some 100 kb away, leading to variable coat colors in mice (Duhl et al. 1994; Michaud et al. 1994). In sum, these data indicate that many Up genes harbor bivalent promoters and lie close to H3K9me3 and ERVs (Fig. 1F).

### *Trim28* deletion triggers a switch from repressive to active chromatin marks at ERVs

Mapping the genomic location of specific TRIM28-regulated ERVs based on a TRIM28 ChIP-seq is problematic because of the sharpness of the corresponding peaks, which only rarely extend beyond the borders of these multicopy elements. We thus turned to a comparison of H3K9me3 peaks in wild-type and *Trim28*-deleted ES cells, since this histone modification can spread a few kilobases into the junction of ERV proviruses with their flanking regions (Karimi et al. 2011; Rebollo et al. 2011). We found around 19,000 H3K9me3 peaks, that is, about half of those detected in control ES cells, to be TRIM28 dependent as indicated by their absence in knock-out cells (Fig. 2A, left). In agreement with their noted proximity to ERVs (see Fig. 1E), Up genes lay closer to TRIM28-dependent H3K9me3 peaks than Down and Stable genes (Fig. 2A, right). Likewise, in an element-centric analysis, we used the TRIM28-dependent H3K9me3 peaks to determine the nearest gene, generating a list significantly enriched for up-regulated genes (giving 2220 Up genes, Fisher's exact test, $P \leq 2.2 \times 10^{-16}$) (Supplemental Fig. S3A; Supplemental Table 3), in line with the gene-centric analysis above. Of note, upon further examination of the high number of H3K9me3 peaks "newly present" in *Trim28* knock-out cells, we found them to be in the same locations as the WT peaks but just slightly displaced and smaller in height and diameter rather than representing new peaks (Fig. 2A). These peaks thus most likely represent remnants of TRIM28-specific peaks, which is not surprising considering that our analyses

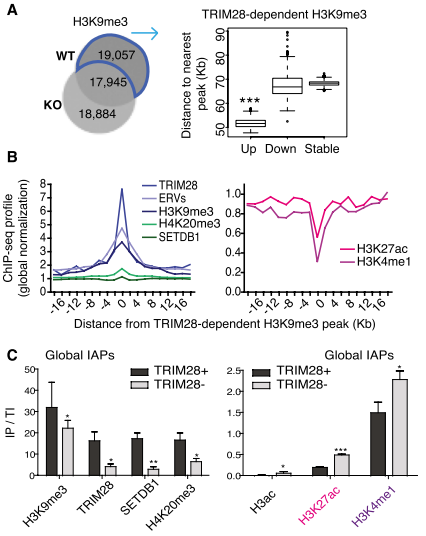



**Figure 2.** *Trim28* deletion triggers a switch from repressive to active chromatin marks at ERVs. (*A*) Venn diagram of H3K9me3 ChIP-seq peaks in WT versus KO ES cells (*left*). 19,057 peaks are present in WT but lost in KO cells and so are defined as TRIM28-dependent peaks, which cluster closer to Up genes than Down ($P = 0.001418$) and Stable ($P \leq 2.2 \times 10^{-16}$) genes (*right*). (*B*) TRIM28-dependent H3K9me3 peaks (see above) were assessed for correlation with ChIP-seq data sets. Positive correlations are shown on the *left* graph and anti-correlations on the *right*. All data displayed after global normalization of ChIP-seq counts. (*C*) ChIP results for repressive (*left* panel) and active (*right* panel) marks present at global IAPs (using IAP 5′-UTR primers). Bars show the mean and SD of three to four ChIPs per antibody with immunoprecipitate values normalized to total inputs (IP/TI) relative to *Gapdh*. Negative controls of no antibody were used in all experiments giving no enrichments, while the *Pou5f1* enhancer served as a positive control with high enrichments for both H3K27ac and H3K4me1 of 1.1 and 7.5, respectively. Results were also reproduced in an independent ES cell line (*Rex1*). Paired *t*-tests were used to compare WT and TRIM28-depleted samples for each antibody: H3K9me3, $P = 0.014$; TRIM28, $P = 0.027$; SETDB1, $P = 0.0036$; H4K20me3, $P = 0.0308$; H3ac, $P = 0.0337$; H3K27ac, $P \leq 0.0001$; H3K4me1, $P = 0.011$.

were performed only 4 d after inducing *Trim28* excision to avoid lethality.

Interestingly, we observed that the TRIM28-dependent H3K9me3 peaks not only correlated with repressive histone marks, TRIM28, SETDB1 peaks (the latter data set obtained from Bilodeau et al. 2009), and with ERVs, but anti-correlated with H3K4me1 and H3K27ac, marks typically found together on active enhancers (Creyghton et al. 2010; Rada-Iglesias et al. 2010; Shen et al. 2012), while displaying no particular association with H3K4me3 or H3K27me3 (Fig. 2B; data not shown). In line with this, Up genes themselves also lay far from enhancer marks (Supplemental Fig. S3B). We therefore hypothesized that ERVs may gain these marks upon *Trim28* deletion, thereby enhancing expression of neighboring genes. To test this idea, we focused on IAPs since we identified a motif highly represented in our H3K9me3 ChIP-seq peaks (in 64% of peaks) normally present in IAP consensus sequences (Supplemental Fig. S3C,D). Supporting this model, ChIP-qPCR with primers designed to amplify the majority of IAPs revealed that, indeed, in *Trim28* knock-out ES cells, these elements not only lost TRIM28, SETDB1, and the repressive marks H3K9me3 and H4K20me3, but also gained active marks, including H3K27ac and H3K4me1 (Fig. 2C). This observation fits with the recent detection of H3K9me3 at poised enhancers (Zentner et al. 2011), and indicates that loss of this mark upon TRIM28 depletion may be sufficient to activate such regulatory elements, notably those located within IAPs and likely other ERVs. The derepression of cryptic enhancers within ERVs thus appears to be one prominent mechanism in the transcriptional deregulation triggered by *Trim28* deletion in ES cells.

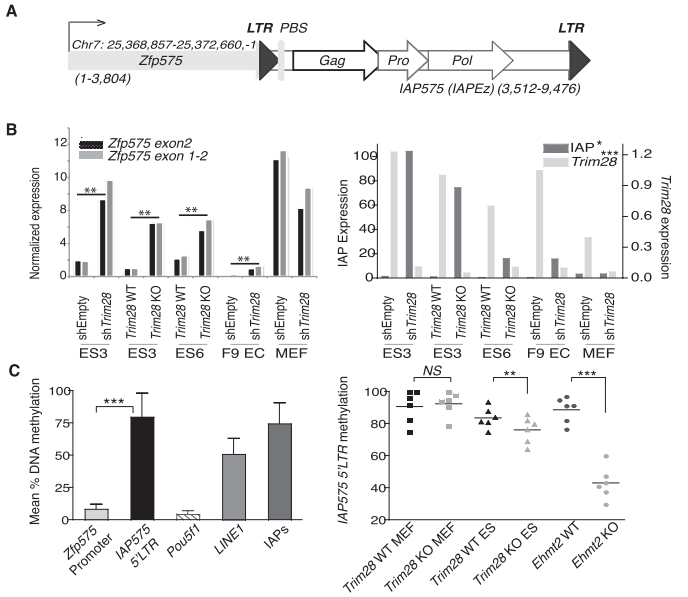



**Figure 3.** Expression and cytosine methylation at the *Zfp575* gene and adjacent IAP. (*A*) Map (drawn to scale) of the *Zfp575* gene that just overlaps a full-length IAP (named *IAP575* and of the *IAPEz* type) with both gene and IAP in the same orientation (sizes of each are stated). (LTR) Long terminal repeat; (PBS) primer binding site; (*Gag*) group-specific antigen; (*Pro*) protease; (*Pol*) polymerase. (*B*) TRIM28 knock-out and knockdown (comparing control, shEmpty and KD, shTRIM28) cell lines were assessed for their expression of *Zfp575* (*left* panel) using two different primer sets, or TRIM28 or IAPs as controls (*right* panel). Unpaired *t*-tests were used to compare controls with TRIM28-depleted samples for all ES and EC cell lines: *Zfp575*, $P = 0.0015$; IAP, $P = 0.0344$; TRIM28, $P = 0.0008$. Since *Zfp575* is normally expressed specifically in brain, we also verified it to be expressed in primary neurospheres and brain (data not shown). (*C*) Quantitative pyrosequencing was used to measure DNA methylation levels at the *Zfp575* promoter versus the flanking 5′-LTR IAP575 promoter (*left* panel). Control primers were specific for the *Pou5f1* promoter or global *LINE1s* or global IAP LTRs (IAPs). Bars represent means over multiple CpG positions with error bars showing the SD across all CpGs. (*Right* panel) Samples were compared (across six CpG positions) for their methylation levels at the IAP575 promoter. Primordial germ cells were also used to show that IAP575 is demethylated in germ cells to a level not much lower than in *Trim28*-deleted ES cells (e.g., to an average of 69% instead of 76%) (data not shown). Two-tailed paired *t*-tests display all significant differences: *Trim28* WT versus KO ES, $P = 0.0088$; *Ehmt2* WT versus KO, $P = 0.0001$.

## Activation of specific ERV-based enhancers upon loss of TRIM28 leads to activation of nearby genes

To explore the molecular mechanism of this process further, we examined transcription and chromatin state at specific ERV–Up gene pairs. We first focused on an element that was 90% identical to IAP sequences previously found to be TRIM28 regulated (Rowe et al. 2010) and named this ERV *IAP575* because of its position 3′ to the bivalent gene *Zfp575* (Mikkelsen et al. 2007; Bilodeau et al. 2009) in the sense orientation (Fig. 3A). *Zfp575* was markedly upregulated in TRIM28-depleted ES cells but not MEFs, consistent with our mRNA-seq data, paralleling the modulation of IAPs in these targets (Figs. 3B, 1A). Similar to its *Pou5f1* counterpart, the *Zfp575* promoter was unmethylated in ES cells. In contrast, the *IAP575 LTR* displayed high rates of CpG methylation, as did the IAP family as a whole, and to a lesser extent *LINEs* (Fig. 3C, left). The failure of DNA methylation to extend from the *IAP575 LTR* to the promoter of the adjacent *Zfp575* gene fits with recent observations that (1) DNA methylation only spreads a few kilobases from TRIM28 binding sites (Quenneville et al. 2012; Rowe et al. 2013), and (2) ERV methylation rarely affects flanking regions (Rebollo et al. 2011). Interestingly, while methylation of the *IAP575 LTR* was unaltered by *Trim28* deletion in MEFs, it significantly decreased in their ES cell counterparts, albeit not as dramatically as in ES cells deleted for *Ehmt2* (*G9a*), a histone methyltransferase involved in the maintenance of DNA methylation (Fig. 3C, right; Dong et al. 2008; Tachibana et al. 2008). Perhaps explaining this latter difference, TRIM28 loss is lethal after a few days in ES cells (Rowe et al. 2010), while EHMT2-depleted cells can be stably maintained for many passages, allowing for extensive loss of cytosine methylation through multiple rounds of DNA replication. However, since this only modest decrease in DNA methylation was observed in parallel to the striking up-regulation of genes, it is possible that it contributes to this phenotype.

We then mapped histone marks across the *Zfp575*/*IAP575* locus (Fig. 4). TRIM28, SETDB1, H3K9me3, and H4K20me3 were

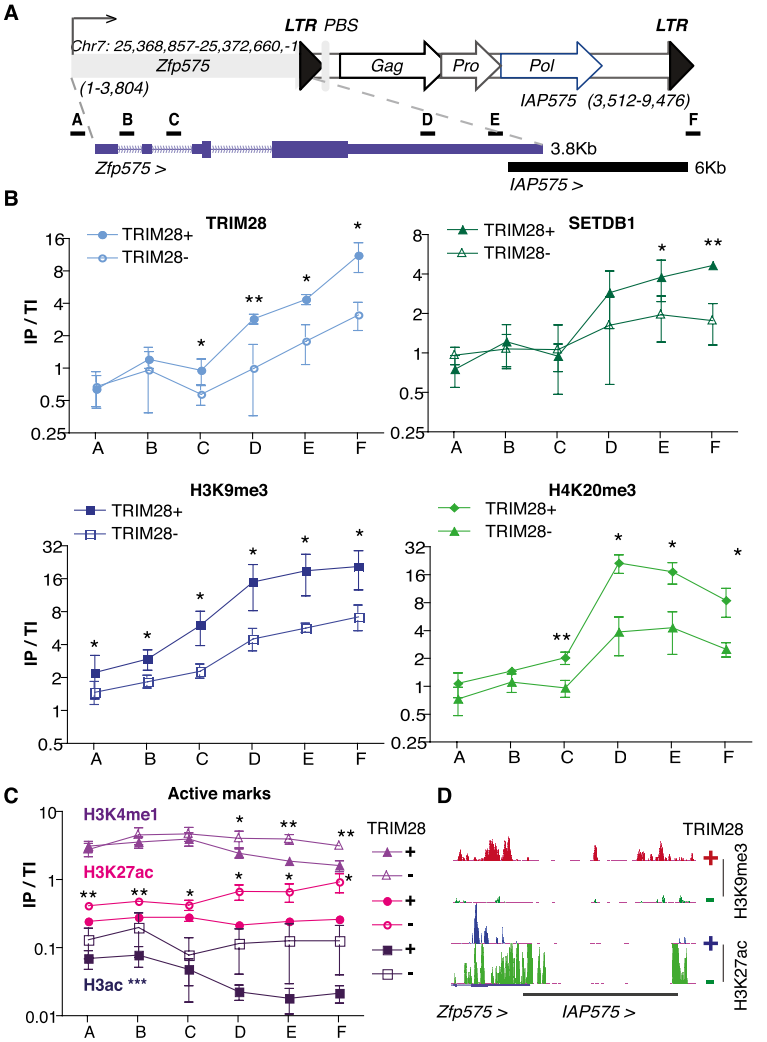



**Figure 4.** *Zfp575* is regulated by a gain of active chromatin marks at its adjacent IAP575. (*A*) Map of *Zfp575* and its adjacent IAP575 (for details, see Fig. 3A) with an enlargement shown underneath to show where primer pairs for ChIP are located. (*B*) ChIP results of repressive marks. (IP/TI) Immunoprecipitate values were normalized to their respective total inputs and to *Gapdh*. Bars represent the mean and SD of three to four ChIPs per antibody, and experiments were also reproduced in another ES cell line (*Rex1*) (data not shown). In each experiment, controls of no antibody were included giving no enrichments. Differences between WT and TRIM28-depleted samples were assessed for each primer set using paired *t*-tests with all significant differences given; (*) $P \leq 0.05$, (**) $P \leq 0.01$. (*C*) ChIPs this time on active marks were performed as described in *B* with data representing three to four ChIPs per antibody. Additionally, here the *Pou5f1* enhancer was used as a positive control (data not shown) showing high enrichment for both H3K4me1 and H3K27ac but not for TRIM28 or H3K9me3. For H3K4me1 and H3K27ac, all significant differences are shown for each primer set, while for H3ac, WT samples were significantly different from TRIM28-depleted ones, not for individual points but over all primer sets; (***) $P \leq 0.001$. (*D*) ChIP-seq maps of H3K9me3 and H3K27ac in TRIM28 WT and depleted ES cells (set to the same vertical scale) at the *Zfp575-IAP575* locus. Note that reads within ERVs, especially conserved ones (in black), are usually missing due to the inability to map reads within highly repeated sequences. However, reads are present at the borders of these elements.

markedly enriched at *IAP575*, yet did not spread back to the *zfp575* promoter. Upon *Trim28* deletion, these repressive histone modifications collectively decreased, to be replaced by the active marks H3K4me1, H3K27ac, and H3Ac over the whole locus, albeit in the most pronounced fashion over its *IAP575* part (Fig. 4B–D). We then further validated the up-regulation of several other ERV–Up gene

pairs and verified that at these loci, TRIM28-dependent H3K9me3 is substituted by the active mark H3K27ac, as documented by ChIP-seq (Supplemental Figs. S4–S6), in support of our model.

### ERV sequences that escape TRIM28-mediated repression can act as activators during embryogenesis

These results indicate that some ERVs carry intrinsic enhancer sequences that are silenced at the ES cell stage via TRIM28-induced repression. To probe this model further, we tested previously identified TRIM28-sensitive and TRIM28-resistant IAP sequences (Rowe et al. 2010) for their ability to modulate a nearby cellular promoter during embryonic development. To this end, we placed these elements in the antisense direction upstream of a phosphoglycerate kinase (PGK) promoter because at baseline this promoter drives only weak expression of GFP in embryos. We then used these lentiviral vectors for transgenesis via transduction of fertilized murine oocytes. Examination of the resulting embryos at E13 revealed that, while a TRIM28-sensitive IAP-derived sequence (IAP4) was able to limit expression from the PGK promoter contained in the lentiviral provirus, its TRIM28-resistant counterpart (IAP1, ~87% identical) (see Rowe et al. 2010), in contrast, enhanced GFP expression (Fig. 5). Thus, TRIM28 susceptibility can condition the *cis*-acting transcriptional impact of specific ERV sequences in vivo during embryonic development.

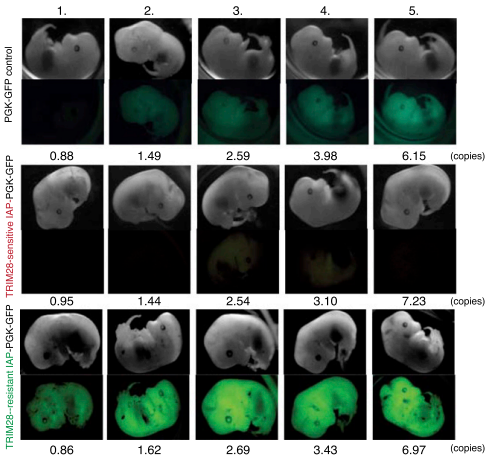



**Figure 5.** ERV sequences that escape TRIM28-mediated repression can act as activators during embryogenesis. Lentiviral transgenesis was performed with an empty PGK-GFP vector (PGK-GFP control, *upper* panels), or with the same vector including either an IAP4 (TRIM28-sensitive IAP-PGK-GFP, *middle* panels) or an IAP1 (TRIM28-resistant IAP-PGK-GFP, *lower* panels) sequence cloned antisense upstream of the PGK promoter. At E13, embryos were scored for GFP expression and vector copy numbers. For the PGK-GFP control, 13/29 embryos were green. For the TRIM28-sensitive IAP-PGK-GFP, 4/19 embryos were green (all with copy numbers above 16), and 4/19 pale green (including numbers 3 and 4 in this figure). For the TRIM28-resistant IAP-PGK-GFP, 12/17 embryos were green (including one with a copy number above 10), and 2/17 pale green (with copy numbers of 0.95 and 0.89). Embryos with similar copy numbers per vector group are shown in each column with increasing copy numbers by row. Vectors were injected twice with similar results. In one experiment, MEFs were derived from embryos to verify that microscopy differences were reproduced by flow cytometry (data not shown).

## Discussion

The present work unveils a fundamental aspect of transcriptional regulation during the early embryogenesis of higher vertebrates. At the heart of this system lies, on one side, retroelements that have colonized eukaryotic genomes from the earliest times, and on the other side, the tetrapod-specific KRAB-ZFP gene family (Urrutia 2003; Huntley et al. 2006; Emerson and Thomas 2009; Wolf and Goff 2009; Thomas and Schneider 2011), which acts as the targeting machinery for TRIM28. We previously demonstrated that TRIM28 is responsible for the silencing of ERVs in ES cells and early embryos (Rowe et al. 2010). Here, we reveal that an important role of this process is to protect the transcriptional dynamics of early embryos from perturbation by *cis*-acting activators contained in these mobile elements.

For this, we deleted *Trim28* in ES cells and monitored chromatin signatures at deregulated genes and ERVs. We found that half of the ~5700 transcriptional units up-regulated upon *Trim28* deletion in ES cells bore, at baseline, the bivalent histone marks H3K4me3 and H3K27me3 characteristic of genes poised for transcription (Bernstein et al. 2006). Moreover, we noted that, remarkably, these genes were on average located closer to ERVs than genes down-regulated or unaffected following TRIM28 removal. We then further observed that, while in wild-type ES cells, ERVs bound TRIM28 and SETDB1 and accordingly were enriched in H3K9me3 and H4K20me3, they lost these repressive marks upon *Trim28* deletion and instead acquired chromatin modifications typically associated with active enhancers such as H3K4me1 and H3K27ac, a phenomenon that was documented both at global IAPs and at the level of specific ERV-up-regulated gene loci. Finally, we could demonstrate that ERV-derived sequences could either repress or activate an adjacent cellular promoter in transgenic mouse embryos, depending on whether they were recognized or not by a TRIM28-containing complex in ES cells.

The model emerging from our study (Fig. 6) is one whereby, in ES cells, the recruitment of TRIM28 and its partners, including SETDB1, at ERV-contained enhancers leads to the maintenance of H3K9me3, H4K20me3, and DNA methylation, which prevents the untimely activation of nearby genes, in particular, those harboring bivalent promoters. Indeed, DNA methylation is known to anti-correlate with active marks (Okitsu and Hsieh 2007; Ooi et al. 2007; Weber et al. 2007; Stadler et al. 2011), and SETDB1 has previously been shown to maintain H3K9 trimethylation and, secondarily, the Suv420H1/2-mediated mark H4K20me3 at ERVs (Matsui et al. 2010). Inactivation of this machinery leads not only to the loss of silent histone marks and to a mild decrease in cytosine methylation but also to the acquisition of active enhancer marks at these loci, which tilts nearby genes, notably those poised for transcription, toward expression. Noteworthy, the NuRD complex, also recruited by TRIM28, is known to mediate deacetylation of H3K27 through its HDAC1 and HDAC2 subunits (Reynolds et al. 2011), which would explain the genome-wide anti-correlation observed between H3K27ac and TRIM28 target sites at baseline. Likewise, LSD1, which shares at least some targets with TRIM28 and NuRD (Macfarlan et al. 2011, 2012), is able to demethylate and therefore decommission the active mark H3K4me1 (Whyte et al. 2012). Accordingly, disruption of either SETDB1 or LSD1 leads to effects on cellular transcripts (Bilodeau et al. 2009; Yuan et al. 2009; Karimi et al. 2011; Macfarlan et al. 2011, 2012). In the case of SETDB1 deletion, this includes the induction of chimeric transcripts initiating from derepressed ERVs, which we also see evidence for here, since some of the same transcripts are induced (Karimi et al. 2011;

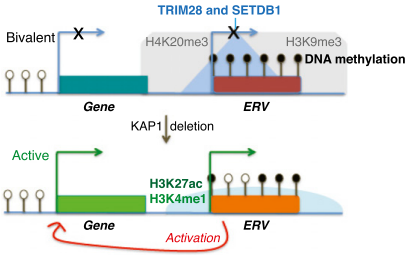



**Figure 6.** Summary model: Substitution of TRIM28-dependent repressive chromatin by the active marks H3K4me1 and H3K27ac at specific ERV-Up gene pairs parallels activation of gene expression.

this study). Here we demonstrate that in the absence of TRIM28, retrotransposon-based enhancers become active.

The heterogeneity of the TRIM28-recruiting ERV loci uncovered here, with sequences intrinsic to IAP, MERVL, and ERVK families, suggests that a large number of different KRAB-ZFPs engage in directing TRIM28 to ERVs in ES cells. Additionally, TRIM28 can also interact with KRAB-O proteins that lack zinc fingers but bridge DNA through other factors such as SRY (Peng et al. 2009). Remarkably, TRIM28 and some KRAB-ZFPs are also detected in adult tissues, albeit along exquisitely cell- and stage-specific fashions, where they have become coopted to influence tissue-specific gene regulation (Jakobsson et al. 2008; Bojkowska et al. 2012; Chikuma et al. 2012; Krebs et al. 2012; Santoni de Sio et al. 2012a,b). Whether some ERV-derived enhancers serve as docking sites for this repressor system in these adult tissues warrants exploration. There is evidence that some ERV sequences function as authentic regulators, including enhancers, in certain cells, not only during development but also in adult tissues (Pi et al. 2004; Bourque et al. 2008; Kunarso et al. 2010; Teng et al. 2011; Mey et al. 2012; Schmidt et al. 2012). Our data indicate that these rare coopted elements represent only exceptions within a large group, most members of which are repressed through TRIM28. This may explain why most KRAB-ZFP genes are expressed in both mouse and human ES cells, while at least in this latter species, most if not all endogenous retroviruses have accumulated mutations that would anyway preclude their retrotransposition. The need to preserve the transcription dynamics of ES cells, rather than to protect the genome from further spread of these elements, is likely what constitutes the strongest selective pressure on the KRAB/TRIM28 system in higher species.

## Methods

### Lentiviral vectors

For in vivo experiments, the transfer vector pRRLSIN.cPPT.PGK-GFP.WPRE (available from Addgene) was used with either IAP1 or IAP4 sequences (Rowe et al. 2010) included upstream of the PGK (phosphoglycerate kinase-1) promoter in the antisense orientation (Rowe et al. 2013). For TRIM28 knockdown experiments, shRNA lentiviral plasmids (against mouse *Trim28* or the empty vector control) were ordered from Sigma-Aldrich (pLKO.1-puro). All vectors were produced by transient transfection of 293T cells with the transfer vector, packaging, and VSVG envelope plasmids (Barde et al. 2010) and titrated on 3T3 fibroblasts.

### Cell culture

ES cells were cultured in standard conditions as described (Rowe et al. 2013). The ES cell lines used were two $Trim28_{loxP/loxP}$ lines called ES3 and ES6 and their derived *Trim28*-conditional knock-out cell lines that are transduced with a tamoxifen (4-0HT)–inducible *Cre* vector (Rowe et al. 2010). For analysis of expression and chromatin marks, knock-out cells were collected 4 d after treatment with 4-0HT (used overnight at 1 μM, Sigma-Aldrich: H7904) due to the lethality of *Trim28* knock-out for longer time periods. *Rex1GFP* ES cells (Wray et al. 2011) were additionally used where stated (kind gift from A.G. Smith, University of Cambridge, UK) or *Ehmt2* parental or stable knock-out ES cells (Dong et al. 2008; Tachibana et al. 2008) (a kind gift from Yoichi Shinkai, RIKEN Institute, Japan). TRIM28-knockdown was induced with shRNA vectors (see above), and cells selected with puromycin 2 d post-transduction and collected 4 d post-puromycin selection, a time point giving similar expression changes to 4 d post-knock-out. Knockdown efficiency was verified by qRT-PCR. $TRIM28_{loxP/loxP}$ 4-0HT-inducible MEFs were used to delete *Trim28*, while TRIM28 knockdowns were also performed in MEFs and F9 EC cells where stated.

### Flow cytometry

Vector titers and GFP repression were measured by FACS, as well as the differentiation status of ES cells as monitored by staining with an SSEA-1 PE- conjugated antibody or isotype control (BD Pharmingen: 560142 and 555584).

### RNA extraction and quantification

Total RNA was extracted with TRIzol (Invitrogen: 15596-018), purified using a PureLink RNA kit (Ambion: 12183018A), treated with DNase (Ambion: AM1907) and 500 ng reverse-transcribed using random primers and SuperScript II (Invitrogen: 18064-022). Primers (see Supplemental Table 4) were designed for an Applied Biosystems 7900HT machine using Primer Express (Applied Biosystems) and used for SYBR Green qPCR. Primer specificity was confirmed by dissociation curves and samples were normalized to *Gapdh*, although *Actin* gave similar results.

### mRNA sequencing

Total RNA (10 μg) from TRIM28 WT and KO ES cells and MEFs was subject to mRNA selection, fragmentation, cDNA synthesis, and library preparation for Illumina high-throughput sequencing, after checking RNA quality on a Bioanalyzer. Single read sequencing was performed on a Genome Analyzer IIx machine with 40 cycles generating ~33 million reads per sample. Additionally, mRNA sequencing was performed on *Trim28* control (shEmpty) and knockdown (shTRIM28) *Rex1* ES cells with 50 cycles on an Illumina HiSeq 2000 machine generating around 200 million reads per sample and confirming our knock-out ES cell results.

### Chromatin immunoprecipitation (ChIP)

ES cell samples were washed twice (in PBS + 2% FCS), counted to normalize by cell number, cross-linked (10 min rotation in 1% formaldehyde), quenched with glycine (at 125 mM on ice), washed three times (PBS), and pelleted at $10^7$ cells per Eppendorf. Pellets were lysed, resuspended in 1 mL of sonication buffer on ice (10 mM Tris at pH 8, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% NaDOC, 0.25% NLS, and protease inhibitors), transferred to glass 12 × 24-mm tubes (Covaris: 520056), and sonicated (Covaris settings: 20% duty cycle, intensity 5, 200 cyles/ burst, 30 min).

Sonication was then assessed by reverse cross-linking overnight in the presence of proteinase K and RNase, followed by DNA extraction and quantification on a Bioanalyzer (Agilent 2100 machine). Fragment sizes were equivalent between wild-type and knock-out samples, which were done in parallel (with mean fragment sizes of ~200 bp for Experiment 1 and ~400 bp for Experiments 2 and 3). Samples were also checked for the absence of single-stranded DNA by Exonuclease I treatment. Immunoprecipitations were performed in duplicates or triplicates with Dynabeads (100.03D) using $1 \times 10^6$ to $2 \times 10^6$ cells, 80 μL of pre-blocked beads, and 5 μg of antibody (or no antibody as a control) per sample in IP buffer (167 mM NaCl, 16.7 mM Tris at pH 8.1, 1.2 mM EDTA, 0.5 mM EGTA, 1.1% Triton X-100, and protease inhibitors) overnight. After washing and reverse cross-linking (also overnight) and DNA extraction, results were quantified by SYBR Green qPCR (for primers, see Supplemental Table 4). The antibodies used were TRIM28 (Tronolab, rabbit polyclonal SY 3267-68, 30–50 μL per sample), H3K9me3 (Abcam: ab8898), SETDB1 (Santa Cruz, 50 μL per sample), H4K20me3 (Millipore: 07-463), H3ac (Millipore: 06-599), H3K27ac (Abcam: ab4729), and H3K4me1 (Abcam: ab8895).

### ChIP sequencing

Total input (TI) and corresponding immunoprecipitated (IP) ChIP libraries were prepared using 10 ng of material with gel selection of 200-bp- to 300-bp-sized fragments. Libraries were ligated with Illumina adaptors and paired-end sequenced (or single-end for H3K27ac) on an Illumina HiSeq 2000 machine with 50–100 cycles and two samples multiplexed in one lane, generating ~100 million sequences per sample. TI samples gave background enrichment patterns distinct from IPs.

### Quantitative bisulfite pyrosequencing

Genomic DNA was converted (200 ng/sample) and used for PCR and pyrosequencing as previously described (Rowe et al. 2013). We thank A. Reymond (CIG, UNIL, Lausanne) for kind use of the pyrosequencer. Results were analyzed using Pyro Q-CpG Software.

### Lentiviral transgenesis

Lentiviral vectors for transgenesis were prepared using Episerf medium (Invitrogen: 10732022), the particle concentration obtained by p24 ELISA (PerkinElmer: NEK050B001KT), and the infectious titer determined on HCT116 cells by GFP flow cytometry. Ratios for the three vectors were between 1/319 and 1/428 of infectious to physical particles with titers between 2 and $2.4 \times 10^9$ infectious units/mL. Transgenesis was performed by perivitelline injection of vectors into fertilized oocytes that were transferred to foster mothers (strain B6D2F1/J) and then recovered at embryonic day 13 (E13). Photographs were taken using the same saturation, gain, and exposure settings and image settings for all embryos.

### Bioinformatics analyses and statistics

#### *mRNA–seq analysis*

Reads were mapped to the mouse genome mm9 using the short read aligner program Bowtie (Langmead et al. 2009) with reads (three mismatches allowed) excluded that mapped more than five times. The SAMtools and bedtools suites (Li et al. 2009; Quinlan and Hall 2010) were used to generate files to be visualized on the UCSC Genome Browser (http://genome.ucsc.edu/) (Kent et al. 2002).

#### *MA plots*

MA plots were generated from rpkm values (number of reads normalized by gene length and total reads) using the maplot Python package (https://github.com/delafont/maplot).

#### *Boxplots*

Boxplots showing bootstrapped values (generated using R: http://www.R-project.org) were used in gene-centric analyses to determine if up-regulated (Up) genes were closer to the indicated histone marks/ERVs compared with two control gene groups (down-regulated, "Down" or unaffected, "Stable" genes). Statistical significance was calculated using the Wilcoxon rank-sum test.

#### *H3K9me3 ChIP–seq analysis*

Paired-end reads were mapped to the mouse genome (three mismatches allowed) mm9 using the short read aligner program Bowtie (Langmead et al. 2009). Several analyses were performed, showing the same global results where reads were either excluded if mapping more than one time, five times, or 20 times to the genome. Peaks were called from the data where reads were mapped with a cutoff of 20 to allow more coverage of repeats, although individual peaks of interest were validated using the analysis where a cutoff of one was used (in this case, only exact matches were allowed). Enriched regions were defined using the ChIP-Part analysis module from the ChIP-seq analysis suite (http://ccg.vital-it.ch/chipseq/). H3K27ac ChIP-seq data were confirmed to correlate (by 53%) with previous H3K27ac ChIP-seq in ES cells (Creyghton et al. 2010) and verified to be normally present at active genes and gained at specific ERV loci (see Supplemental Figs. S5, S6). TRIM28 ChIP-seq peaks were defined using MACS (default threshold $P$-value $<1 \times 10^{-5}$) and normalized to the total input generating 3099 peaks. Direct binding sites to promoters of up-regulated genes were identified using a cutoff of ±2 kb from the TSS giving 49 genes, 13 of which were excluded due to the binding being through an ERV.

#### *Public ChIP–seq data*

Raw or already mapped reads were downloaded from publicly available ChIP-seq data (GEO IDs: GSE12241, GSE18371, and GSE24165) and peaks called using MACS. ChIP-correlation analyses were performed with bed files, using the online tool ChIP-Cor (http://ccg.vital-it.ch/chipseq/chip_cor.php). Histograms were analyzed using raw counts and count densities, and those showing a correlation were displayed after global normalization, where ChIP-seq counts are normalized by the total number of counts and the window width to allow visualization of multiple data sets on the same plot.

#### *Motif identification*

The MotifRegressor and motifsComparator softwares were used to identify DNA sequence binding motifs (Conlon et al. 2003; Carat et al. 2010).

### Other statistical analyses

GraphPad Prism version 4.00 (http://www.graphpad.com) was used for other statistical analyses, where control and knock-out groups were compared with paired or unpaired $t$-tests (as noted) that were one-tailed except where stated as two-tailed.

## Data access

All next-generation sequencing data have been submitted to the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) and are accessible with the accession no. GSE41903.

## Acknowledgments

We thank P.V. Maillard and J. Marquis for advice, S. Verp and S. Offner for technical assistance through the EPFL Lentiviral Transgenesis platform, B. Khubieh and J. Rougemont for bioinformatics help (EPFL Biostatistics and Bioinformatics core facility), Y. Shinkai and D. Schübeler for the *Ehmt2* knock-out and parental ES cells, and K. Harshman and A. Reymond (Center for Integrative Genomics, University of Lausanne) for high-throughput sequencing or use of a pyrosequencer, respectively. All computing for high-throughput sequencing was done on the Vital-it cluster. This work was supported through grants from the Swiss National Science Foundation and the European Research Council to D.T. and an NIH grant to S.L.P.

*Author contributions:* H.M.R. conceived the study, designed and performed the experiments, analyzed the data, and wrote the manuscript. A.K. performed bioinformatics analyses. A.C., L.F., T.S.M., and Y.T. performed experiments. J.J., S.V., and S.L.P. designed experiments. D.T. conceived the study, designed experiments, and wrote the manuscript.

## References

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* **11:** R106.

Barde I, Salmon P, Trono D. 2010. Production and titration of lentiviral vectors. *Curr Protoc Neurosci* **53:** 4.21.1–4.21.23.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125:** 315–326.

Bilodeau S, Kagey MH, Frampton GM, Rahl PB, Young RA. 2009. SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev* **23:** 2484–2489.

Bojkowska K, Aloisio F, Cassano M, Kapopoulou A, de Sio FS, Zangger N, Offner S, Cartoni C, Thomas C, Quenneville S, et al. 2012. Liver-specific ablation of KRAB associated protein 1 in mice leads to male-predominant hepatosteatosis and development of liver adenoma. *Hepatology* **56:** 1279–1290.

Bourque G, Leong B, Vega VB, Chen X, Lee YL, Srinivasan KG, Chew JL, Ruan Y, Wei CL, Ng HH, et al. 2008. Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome Res* **18:** 1752–1762.

Cammas F, Mark M, Dolle P, Dierich A, Chambon P, Losson R. 2000. Mice lacking the transcriptional corepressor TIF1β are defective in early postimplantation development. *Development* **127:** 2955–2963.

Carat S, Houlgatte R, Bourdon J. 2010. A parallel scheme for comparing transcription factor binding sites matrices. *J Bioinform Comput Biol* **8:** 485–502.

Chikuma S, Suita N, Okazaki IM, Shibayama S, Honjo T. 2012. TRIM28 prevents autoinflammatory T cell development in vivo. *Nat Immunol* **13:** 596–603.

Conlon EM, Liu XS, Lieb JD, Liu JS. 2003. Integrating regulatory motif discovery and genome-wide expression analysis. *Proc Natl Acad Sci* **100:** 3339–3344.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, et al. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci* **107:** 21931–21936.

Dong KB, Maksakova IA, Mohn F, Leung D, Appanah R, Lee S, Yang HW, Lam LL, Mager DL, Schubeler D, et al. 2008. DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J* **27:** 2691–2701.

Duhl DM, Vrieling H, Miller KA, Wolff GL, Barsh GS. 1994. Neomorphic agouti mutations in obese yellow mice. *Nat Genet* **8:** 59–65.

Emerson RO, Thomas JH. 2009. Adaptive evolution in zinc finger transcription factors. *PLoS Genet* **5:** e1000325.

Fazzio TG, Huff JT, Panning B. 2008. An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* **134:** 162–174.

Friedman JR, Fredericks WJ, Jensen DE, Speicher DW, Huang XP, Neilson EG, Rauscher FJ III. 1996. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev* **10:** 2067–2078.

Frietze S, O'Geen H, Blahnik KR, Jin VX, Farnham PJ. 2010. ZNF274 recruits the histone methyltransferase SETDB1 to the 3′ ends of ZNF genes. *PLoS ONE* **5:** e15082.

Gan Q, Yoshida T, McDonald OG, Owens GK. 2007. Concise review: Epigenetic mechanisms contribute to pluripotency and cell lineage determination of embryonic stem cells. *Stem Cells* **25:** 2–9.

Guenther MG, Young RA. 2010. Transcription. Repressive transcription. *Science* **329:** 150–151.

Hu G, Kim J, Xu Q, Leng Y, Orkin SH, Elledge SJ. 2009. A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev* **23:** 837–848.

Huntley S, Baggott DM, Hamilton AT, Tran-Gyamfi M, Yang S, Kim J, Gordon L, Branscomb E, Stubbs L. 2006. A comprehensive catalog of human KRAB-associated zinc finger genes: Insights into the evolutionary history of a large family of transcriptional repressors. *Genome Res* **16:** 669–677.

Ivanov AV, Peng H, Yurchenko V, Yap KL, Negorev DG, Schultz DC, Psulkowski E, Fredericks WJ, White DE, Maul GG, et al. 2007. PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Mol Cell* **28:** 823–837.

Jakobsson J, Cordero MI, Bisaz R, Groner AC, Busskamp V, Bensadoun JC, Cammas F, Losson R, Mansuy IM, Sandi C, et al. 2008. KAP1-mediated epigenetic repression in the forebrain modulates behavioral vulnerability to stress. *Neuron* **60:** 818–831.

Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, Tang JX, Shinkai Y, Mager DL, Jones S, Hirst M, et al. 2011. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* **8:** 676–687.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The Human Genome Browser at UCSC. *Genome Res* **12:** 996–1006.

Krebs CJ, Schultz DC, Robins DM. 2012. The KRAB zinc finger protein RSL1 regulates sex- and tissue-specific promoter methylation and dynamic hormone-responsive chromatin configuration. *Mol Cell Biol* **32:** 3732–3742.

Kunarso G, Chia NY, Jeyakani J, Hwang C, Lu X, Chan YS, Ng HH, Bourque G. 2010. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat Genet* **42:** 631–634.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25.

Lechner MS, Begg GE, Speicher DW, Rauscher FJ III. 2000. Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: Direct chromoshadow domain–KAP-1 corepressor interaction is essential. *Mol Cell Biol* **20:** 6449–6465.

Li X, Ito M, Zhou F, Youngson N, Zuo X, Leder P, Ferguson-Smith AC. 2008. A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev Cell* **15:** 547–557.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25:** 2078–2079.

Macfarlan TS, Gifford WD, Agarwal S, Driscoll S, Lettieri K, Wang J, Andrews SE, Franco L, Rosenfeld MG, Ren B, et al. 2011. Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. *Genes Dev* **25:** 594–607.

Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL. 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487:** 57–63.

Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H, Kimura H, Tachibana M, Lorincz MC, Shinkai Y. 2010. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **464:** 927–931.

McDonel P, Costello I, Hendrich B. 2009. Keeping things quiet: Roles of NuRD and Sin3 co-repressor complexes during mammalian development. *Int J Biochem Cell Biol* **41:** 108–116.

Messerschmidt DM, de Vries W, Ito M, Solter D, Ferguson-Smith A, Knowles BB. 2012. Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science* **335:** 1499–1502.

Mey A, Acloque H, Lerat E, Gounel S, Tribollet V, Blanc S, Curton D, Birot AM, Nieto MA, Samarut J. 2012. The endogenous retrovirus ENS-1 provides active binding sites for transcription factors in embryonic stem cells that specify extra embryonic tissue. *Retrovirology* **9:** 21.

Michaud EJ, van Vugt MJ, Bultman SJ, Sweet HO, Davisson MT, Woychik RP. 1994. Differential expression of a new dominant agouti allele ($A^{iapy}$) is correlated with methylation state and is influenced by parental lineage. *Genes Dev* **8:** 1463–1472.

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448:** 553–560.

Okitsu CY, Hsieh CL. 2007. DNA methylation dictates histone H3K4 methylation. *Mol Cell Biol* **27:** 2746–2757.

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, et al. 2007. DNMT3L connects unmethylated

lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448:** 714–717.
Peng H, Ivanov AV, Oh HJ, Lau YF, Rauscher FJ III. 2009. Epigenetic gene silencing by the SRY protein is mediated by a KRAB-O protein that recruits the KAP1 co-repressor machinery. *J Biol Chem* **284:** 35670–35680.
Pi W, Yang Z, Wang J, Ruan L, Yu X, Ling J, Krantz S, Isales C, Conway SJ, Lin S, et al. 2004. The LTR enhancer of ERV-9 human endogenous retrovirus is active in oocytes and progenitor cells in transgenic zebrafish and humans. *Proc Natl Acad Sci* **101:** 805–810.
Quenneville S, Verde G, Corsinotti A, Kapopoulou A, Jakobsson J, Offner S, Baglivo I, Pedone PV, Grimaldi G, Riccio A, et al. 2011. In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol Cell* **44:** 361–372.
Quenneville S, Turelli P, Bojkowska K, Raclot C, Offner S, Kapopoulou A, Trono D. 2012. The KRAB-ZFP/KAP1 system contributes to the early embryonic establishment of site-specific DNA methylation patterns maintained during development. *Cell Rep* **2:** 766–773.
Quinlan AR, Hall IM. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26:** 841–842.
Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. 2010. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470:** 279–283.
Rebollo R, Karimi MM, Bilenky M, Gagnier L, Miceli-Royer K, Zhang Y, Goyal P, Keane TM, Jones S, Hirst M, et al. 2011. Retrotransposon-induced heterochromatin spreading in the mouse revealed by insertional polymorphisms. *PLoS Genet* **7:** e1002301.
Reynolds N, Salmon-Divon M, Dvinge H, Hynes-Allen A, Balasooriya G, Leaford D, Behrens A, Bertone P, Hendrich B. 2011. NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. *EMBO J* **31:** 593–605.
Rowe HM, Trono D. 2011. Dynamic control of endogenous retroviruses during development. *Virology* **411:** 273–287.
Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, Aktas T, Maillard PV, Layard-Liesching H, Verp S, Marquis J, et al. 2010. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463:** 237–240.
Rowe HM, Friedli M, Offner S, Verp S, Mesnard D, Marquis J, Aktas T, Trono D. 2013. De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET. *Development* **140:** 519–529.
Santoni de Sio F, Barde I, Offner S, Kapopoulou A, Corsinotti A, Bojkowska K, Genolet R, Thomas JH, Luescher IF, Pinschewer D, et al. 2012a. KAP1 regulates gene networks controlling T-cell development and responsiveness. *FASEB J* **26:** 4561–4575.
Santoni de Sio FR, Massacand J, Barde I, Offner S, Corsinotti A, Kapopoulou A, Bojkowska K, Dagklis A, Fernandez M, Ghia P, et al. 2012b. KAP1 regulates gene networks controlling mouse B lymphoid cell differentiation and function. *Blood* **119:** 4675–4685.
Schmidt D, Schwalie PC, Wilson MD, Ballester B, Goncalves A, Kutter C, Brown GD, Marshall A, Flicek P, Odom DT. 2012. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* **148:** 335–348.
Schultz DC, Friedman JR, Rauscher FJ III. 2001. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: The PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2α subunit of NuRD. *Genes Dev* **15:** 428–443.
Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ III. 2002. SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev* **16:** 919–932.
Seki Y, Kurisaki A, Watanabe-Susaki K, Nakajima Y, Nakanishi M, Arai Y, Shiota K, Sugino H, Asashima M. 2010. TIF1β regulates the pluripotency of embryonic stem cells in a phosphorylation-dependent manner. *Proc Natl Acad Sci* **107:** 10926–10931.
Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV, et al. 2012. A map of the *cis*-regulatory sequences in the mouse genome. *Nature* **488:** 116–120.
Sripathy SP, Stevens J, Schultz DC. 2006. The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. *Mol Cell Biol* **26:** 8623–8638.
Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Scholer A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, et al. 2011. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480:** 490–495.
Tachibana M, Matsumura Y, Fukuda M, Kimura H, Shinkai Y. 2008. G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO J* **27:** 2681–2690.
Teng L, Firpi HA, Tan K. 2011. Enhancers in embryonic stem cells are enriched for transposable elements and genetic variations associated with cancers. *Nucleic Acids Res* **39:** 7371–7379.
Thomas JH, Schneider S. 2011. Coevolution of retroelements and tandem zinc finger genes. *Genome Res* **21:** 1800–1812.
Urrutia R. 2003. KRAB-containing zinc-finger repressor proteins. *Genome Biol* **4:** 231.
Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. 2007. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* **39:** 457–466.
Whyte WA, Bilodeau S, Orlando DA, Hoke HA, Frampton GM, Foster CT, Cowley SM, Young RA. 2012. Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature* **482:** 221–225.
Wolf D, Goff SP. 2007. TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. *Cell* **131:** 46–57.
Wolf D, Goff SP. 2009. Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* **458:** 1201–1204.
Wray J, Kalkan T, Gomez-Lopez S, Eckardt D, Cook A, Kemler R, Smith A. 2011. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat Cell Biol* **13:** 838–845.
Yuan P, Han J, Guo G, Orlov YL, Huss M, Loh YH, Yaw LP, Robson P, Lim B, Ng HH. 2009. Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. *Genes Dev* **23:** 2507–2520.
Zentner GE, Tesar PJ, Scacheri PC. 2011. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res* **21:** 1273–1283.
Zuo X, Sheng J, Lau HT, McDonald CM, Andrade M, Cullen DE, Bell FT, Iacovino M, Kyba M, Xu G, et al. 2012. Zinc finger protein ZFP57 requires its co-factor to recruit DNA methyltransferases and maintains DNA methylation imprint in embryonic stem cells via its transcriptional repression domain. *J Biol Chem* **287:** 2107–2118.

*Received August 10, 2012; accepted in revised form December 6, 2012.*